

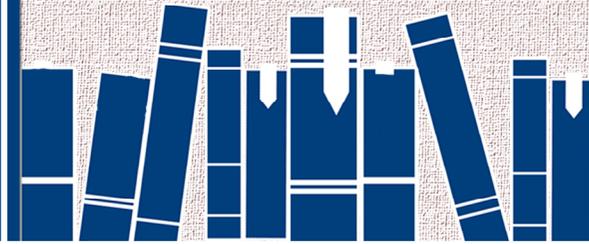
А. А. Мицель, А. А. Шелестов, В. В. Романенко \*\*\* МЕТОДЫ ОПТИМИЗАЦИИ 2



Учебное издание ТУСУРа, 2022



СЕРИЯ



А. А. Мицель  
А. А. Шелестов  
В. В. Романенко

# МЕТОДЫ ОПТИМИЗАЦИИ

## Часть 2

Министерство науки и высшего образования Российской Федерации  
Томский государственный университет систем управления  
и радиоэлектроники

**А.А. Мицель, А.А. Шелестов, В.В. Романенко**

# **МЕТОДЫ ОПТИМИЗАЦИИ**

**Учебное пособие**

**В двух частях**

**Часть 2**

Томск  
Издательство ТУСУРа  
2022

УДК 519.852(075.8)

ББК 22.183я73

М701

Рецензенты:

**Кошкин Г. М.**, д-р физ.-мат. наук, проф. Ин-та прикладной мат.  
и компьютерных наук Нац. исслед. Томск. гос. ун-та

**Кочегуров А. И.**, канд. техн. наук, доц. отд-ния информ. технологий  
Инженерная шк. информ. технологий и робототехники  
Нац. исслед. Томск. политехн. ун-та

**Мицель, Артур Александрович**

М701 Методы оптимизации : учеб. пособие. В 2 ч. Ч. 2 / А.А. Мицель,  
А.А. Шелестов, В.В. Романенко. – Томск : Изд-во Томск. гос. ун-та  
систем упр. и радиоэлектроники, 2022. – 360 с.

ISBN 978-5-86889-894-5 (Ч. 2)

ISBN 978-5-86889-892-1

Вторая часть учебного пособия посвящена методам условной оптимизации для решения задач нелинейного и квадратичного программирования, а также решению оптимизационных задач особого вида методами динамического программирования, вариационного исчисления и эвристическими методами. Каждый раздел сопровождается контрольными вопросами для самопроверки.

Для студентов технических специальностей, аспирантов, преподавателей и инженеров, сталкивающихся с необходимостью численного или аналитического решения оптимизационных задач.

ISBN 978-5-86889-894-5 (Ч. 2)

ISBN 978-5-86889-892-1

© Мицель А.А., Шелестов А.А.,  
Романенко В.В., 2022

© Томск. гос. ун-т систем упр.  
и радиоэлектроники, 2022

## Оглавление

Введение.....	6
<b>Глава 6. НЕЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ</b>	
6.1. Задача определения наилучшего положения приемной антенны .....	8
6.2. Классификация методов решения задач нелинейного программирования .....	12
6.3. Задачи с ограничениями в виде равенств.....	16
6.3.1. Метод замены переменных .....	16
6.3.2. Метод множителей Лагранжа.....	21
6.3.3. Экономическая интерпретация множителей Лагранжа .....	25
6.4. Необходимые и достаточные условия оптимальности.....	27
6.4.1. Влияние ограничений на решение задач оптимизации .....	27
6.4.2. Необходимые и достаточные условия оптимальности задач с ограничениями общего вида.....	28
6.4.3. Необходимые и достаточные условия оптимальности второго порядка .....	34
6.5. Методы штрафов.....	38
6.5.1. Понятие штрафных функций.....	38
6.5.2. Квадратичный штраф .....	41
6.5.3. Логарифмический штраф .....	48
6.5.4. Штраф типа обратной функции .....	51
6.5.5. Штраф типа квадрата срезки .....	54
6.5.6. Выбор штрафного параметра .....	59
6.5.7. Обобщенный алгоритм .....	60
6.6. Методы, основанные на линеаризации.....	62
6.6.1. Базовый метод линеаризации .....	62
6.6.2. Алгоритм Франка – Вульфа .....	69
6.6.3. Метод допустимых направлений Зойтендейка.....	77
6.6.4. Метод условного градиента .....	91
6.7. Метод проекции градиента .....	104
6.7.1. Задачи с линейными ограничениями.....	104
6.7.2. Задачи с нелинейными ограничениями.....	118
Вопросы для самопроверки.....	125
<b>Глава 7. КВАДРАТИЧНОЕ ПРОГРАММИРОВАНИЕ</b>	
7.1. Задача выбора портфеля ценных бумаг.....	128
7.2. Классификация методов.....	133
7.3. Условие Куна – Таккера для задачи квадратичного программирования .....	134
7.4. Метод решения ЗКП с помощью искусственного базиса.....	136

7.5. Решение с помощью симплексного преобразования таблицы коэффициентов уравнений .....	140
7.6. Метод решения задач о дополнителности .....	143
7.7. Метод Мицеля – Хващевского .....	150
Вопросы для самопроверки .....	157

## **Глава 8. ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ**

8.1. Общая постановка задачи динамического программирования.....	158
8.2. Принцип оптимальности и уравнения Беллмана .....	162
8.3. Задача о распределении средств между предприятиями .....	167
8.4. Задача об оптимальном распределении ресурсов между отраслями .....	172
8.5. Задача о замене оборудования .....	177
8.6. Задача о кратчайшем пути.....	185
8.7. Редукция линейной задачи динамического программирования к задаче линейного программирования .....	186
Вопросы для самопроверки .....	190

## **Глава 9. ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ**

9.1. Функционалы. Основные понятия.....	191
9.2. Необходимое и достаточное условие существования экстремума функционалов .....	194
9.3. Вариационные задачи с закрепленными концами .....	198
9.4. Уравнение Эйлера для вариационных задач с закрепленными концами .....	199
9.5. Многомерный случай.....	206
9.6. Уравнения Эйлера – Пуассона .....	208
Вопросы для самопроверки .....	209

## **Глава 10. ЭВРИСТИЧЕСКИЕ МЕТОДЫ ПОИСКА ЭКСТРЕМУМОВ**

10.1. Понятие и особенности эвристических методов .....	210
10.2. Классификация эвристических методов .....	211
10.3. Эволюционные методы оптимизации .....	222
10.3.1. Генетические алгоритмы .....	222
10.3.2. Методы иммунных систем .....	226
10.3.3. Метод рассеивания.....	229
10.3.4. Эволюционная стратегия преобразования ковариационной матрицы.....	232
10.3.5. Метод динамических сеток .....	234
10.4. Генетические алгоритмы оптимизации .....	237
10.4.1. Основные понятия генетических алгоритмов.....	237
10.4.2. Кодирование параметров задачи.....	242
10.4.3. Оператор селекции .....	247

10.4.4. Скрещивание .....	252
10.4.5. Оператор мутации.....	253
10.4.6. Операторы отбора особей в новую популяцию .....	254
10.4.7. Основные отличия генетических алгоритмов от традиционных методов поиска решений.....	254
10.5. Поиск минимума функции одной переменной .....	255
10.6. Многомерная безусловная оптимизация при помощи генетических алгоритмов .....	260
10.6.1. Постановка задачи .....	260
10.6.2. Селекция в задаче многомерной оптимизации .....	261
10.6.3. Кроссинговер в задаче многомерной оптимизации .....	262
10.6.4. Мутация в задаче многомерной оптимизации.....	265
Вопросы для самопроверки .....	267
<b>Глава 11. ПРИМЕРЫ ЗАДАЧ, РЕШАЕМЫХ С ИСПОЛЬЗОВАНИЕМ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ</b>	
11.1. Задача коммивояжера .....	269
11.2. Использование генетических алгоритмов для решения социально-экономических задач .....	277
11.2.1. Условия эффективного применения генетических алгоритмов .....	277
11.2.2. Применение генетических алгоритмов для решения задачи прогнозирования развития города.....	278
11.3. Применение генетических алгоритмов для решения задачи оценки эффективности региональной промышленной политики .....	281
11.4. Формирование системы прогнозирующих правил в деятельности страховых компаний.....	285
Вопросы для самопроверки.....	287
Заключение .....	288
Список сокращений и условных обозначений.....	289
Список литературы.....	292
<b>Приложение А. Краткий математический справочник</b>	
П.А.1. Приближенные числа.....	296
П.А.2. Уравнения и неравенства.....	299
П.А.3. Векторы и матрицы .....	320
П.А.4. Дифференцирование функций .....	341

## Введение

В наиболее общем смысле теория оптимизации представляет собой совокупность фундаментальных математических результатов и численных методов, ориентированных на нахождение и идентификацию наилучших вариантов из множества альтернатив и позволяющих избежать полного перебора и оценивания возможных вариантов.

Одним из классов задач на экстремум, традиционно рассматриваемых в математике, являются задачи вариационного исчисления. Следы интереса к ним можно найти в античной математике (разного рода изопериметрические проблемы), однако подлинное рождение вариационного исчисления произошло в конце XVIII века, когда И. Бернулли сформулировал знаменитую задачу о брахистохроне (греч. «брахисто» – кратчайший, «хронос» – время).

Условия экстремума 1-го порядка в вариационном исчислении были получены Эйлером (уравнения Эйлера), а 2-го порядка – Лежандром и Якоби. Важный вопрос о существовании решения в вариационном исчислении был впервые поставлен Вейерштрассом во второй половине XIX века.

Значительный прогресс в развитии теории оптимизации произошел при изучении так называемых задач оптимального управления. Необходимые условия оптимальности для этих задач были сформулированы и доказаны Понтрягиным, Болтянским и Гамкрелидзе в 1956–1958 гг. в форме «принципа максимума». В иной форме условия оптимальности для подобных задач были получены Беллманом на основе идей динамического программирования.

С необходимостью решения задач оптимизации (ЗО) столкнулись и специалисты по теории автоматического управления (ТАУ). Труды В.В. Казакевича, А.А. Фельдбаума, А.А. Первозванского в 50-х годах XX в. была создана теория экстремального регулирования и предложены специальные математические модели объектов, действующих в реальном масштабе времени.

К середине 1960-х годов в рамках вычислительной математики сложилось самостоятельное направление, связанное с численными методами оптимизации (МО). С тех пор непрерывно шло интенсивное развитие этого направления как вширь (разработка новых методов,

исследование новых классов задач и т.д.), так и вглубь (создание единого аппарата для анализа сходимости и скорости сходимости, классификации методов). В настоящее время эта область вычислительной математики может считаться окончательно сформировавшейся, хотя существуют проблемы построения эффективных методов для некоторых специальных типов задач и отдельных специфических ситуаций. Так, например, для составления маршрута доставки тяжелой баллистической ракеты-носителя с завода в г. Сиэтле на северо-западе США до космодрома на мысе Канаверал (полуостров Флорида) был специально разработан новый оптимизационный алгоритм, учитывающий массу, габариты груза, загруженность автомагистралей, грузоподъемность мостов, размеры туннелей и т.д.

Теория оптимизации, успешному применению которой способствует бурное развитие средств вычислительной техники, вносит заметный вклад в ускорение научно-технического прогресса. Трудно назвать такую отрасль инженерной деятельности, где бы ни возникали задачи оптимизационного характера.

Приведем список классов алгоритмов оптимизации и научной литературы, рекомендуемой для более глубокого изучения материала учебника.

1. Методы нелинейного программирования с ограничениями:

- методы штрафных и барьерных функций [1, 2, 5, 10, 11];
- методы линеаризации [1, 2, 4, 5, 7];
- сепарабельное программирование [1];
- методы условного градиента и проекции градиента [1, 2, 11];
- квадратичное программирование [1, 2, 8, 9].

2. Методы решения оптимизационных задач особого вида:

- динамическое программирование [1, 2, 5, 10–16];
- вариационное исчисление [2, 5, 11, 17, 18];
- теория оптимального управления [1–6, 19];
- эвристические методы оптимизации [19–21].

Для успешного овладения теоретическими материалами по методам оптимизации и решения практических оптимизационных задач в приложении А представлены некоторые сведения из школьного курса алгебры, а также вузовского курса высшей математики.

## Глава 6. НЕЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

### 6.1. Задача определения наилучшего положения приемной антенны

В первой части учебного пособия мы рассматривали задачи условной оптимизации с линейными целевыми функциями (ЦФ) и ограничениями. Теперь рассмотрим нелинейные задачи. Такие задачи называют *задачами нелинейного программирования* (ЗНП) и под ними понимают класс задач, в которых ЦФ или ограничения (или и то и другое вместе) нелинейные. Как уже указывалось ранее, ограничения существенно уменьшают размеры области, в которой проводится поиск экстремума. Изложим методы решения таких задач на примере определения наилучшего положения приемной антенны.

Пусть в южной оконечности озера, форма которого напоминает параболу  $y = x^2$  (как в рассмотренной ранее задаче проектирования канала наименьшей длины), т.е. в точке  $O$ , находится населенный пункт (рисунок 6.1). В точке  $A$ , расположенной на 6 км восточнее и 8 км севернее, расположена телевизионная вышка мощностью 1 кВт. Интенсивность излучения вышки уменьшается пропорционально квадрату расстояния по закону

$$E(x, y) = \frac{E_0}{1 + (x - a)^2 + (y - b)^2},$$

где  $E_0 = 1000$  Вт – начальная мощность сигнала;  $x, y$  – координаты точки приема сигнала;  $a, b$  – координаты вышки ( $a = 6$  км,  $b = 8$  км). Таким образом, в населенном пункте интенсивность приходящего от вышки излучения составляет

$$E(0, 0) = \frac{1000}{1 + 6^2 + 8^2} \approx 9,9 \text{ Вт/м}^2,$$

чего слишком мало для устойчивого приема телевизионного сигнала.

Поэтому было решено установить ближе к вышке (в некоторой точке  $B$ ) приемную антенну и передавать от нее телевизионный сигнал по кабелю. Очевидно, что антенна может быть установлена только

на суше, а кабель можно протянуть как по суше, так и по дну озера. Имеющиеся запасы кабеля составляют  $L = 2$  км. В каком месте следует расположить приемную антенну для получения сигнала максимальной интенсивности?

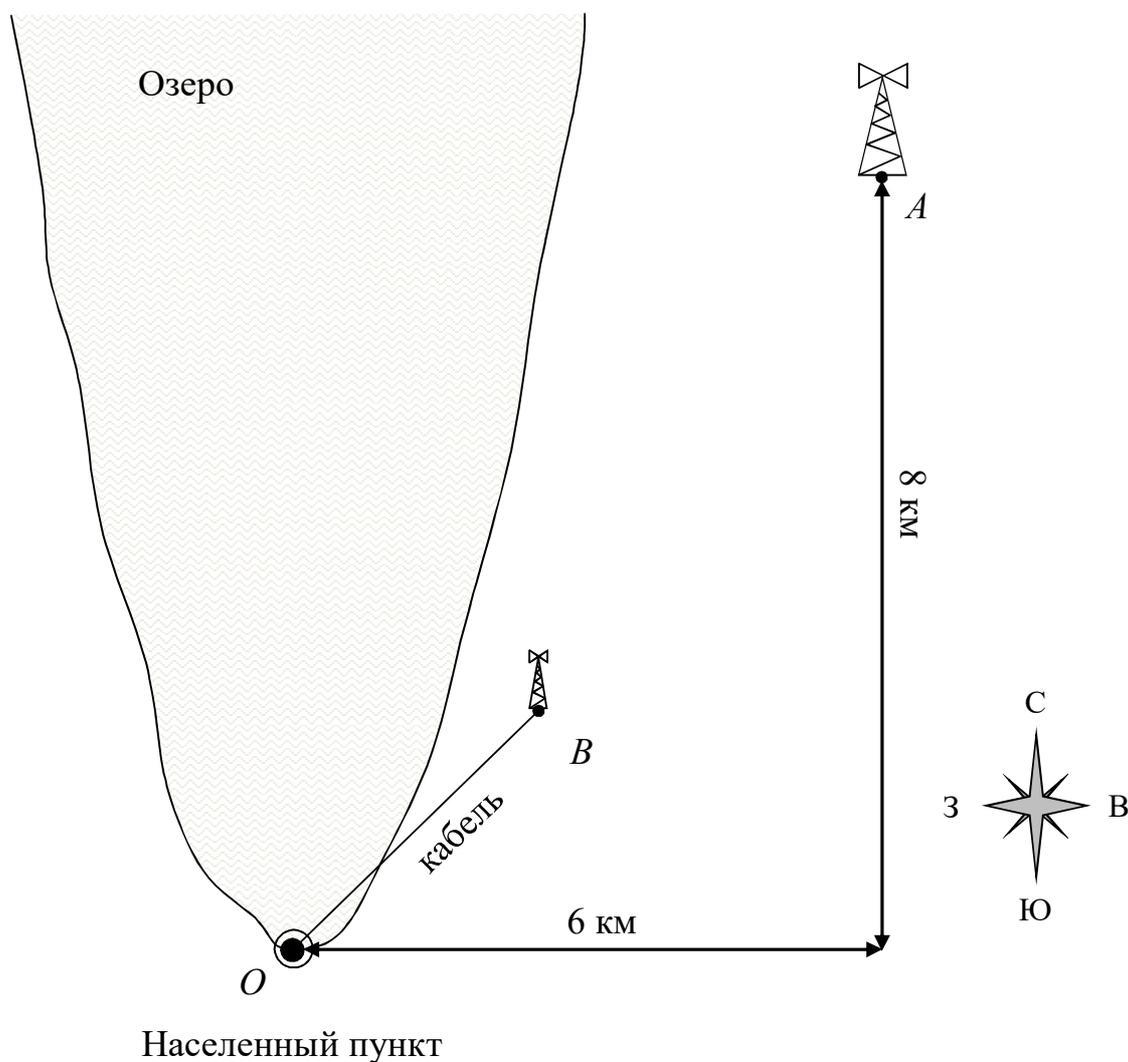


Рисунок 6.1 – К задаче определения оптимального положения приемной антенны

Выполним анализ задачи. Очевидно, что целевой функцией будет зависимость  $E(B)$ , определяющая интенсивность сигнала в точке  $B$ :

$$E(B) \rightarrow \max.$$

Чтобы перейти от максимизации к минимизации, необходимо изменить знак ЦФ:

$$E(B) = -\frac{E_0}{1 + (B_x - a)^2 + (B_y - b)^2} \rightarrow \min.$$

Также очевидно, что точка  $B$  будет располагаться в первом квадранте, между вышкой и населенным пунктом, поэтому добавлять в задачу ограничения  $B_x \geq 0$  и  $B_y \geq 0$  нет необходимости. Чем она ближе к вышке, тем лучше, но мы ограничены длиной кабеля. Отсюда вытекает первое ограничение задачи – расстояние между населенным пунктом и приемной антенной не может превышать  $L$ , т.е.

$$\sqrt{(O_x - B_x)^2 + (O_y - B_y)^2} \leq L.$$

Получается, что точка  $B$  должна лежать в пределах окружности с центром в точке  $O$  и радиусом 2 км.

Второе ограничение задачи вытекает из того факта, что вышка может располагаться только на суше. Береговая линия озера описывается функциональной зависимостью  $y = (x - O_x)^2 + O_y$ , поэтому

$$B_y \leq (B_x - O_x)^2 + O_y.$$

Если принять местоположение населенного пункта за начало координат, эти соотношения можно переписать в виде

$$B_x^2 + B_y^2 \leq L^2,$$

$$B_y \leq B_x^2.$$

В качестве неизвестных задачи удобно принять координаты точки  $B$ :  $x_1 = B_x$ ,  $x_2 = B_y$ .

Область допустимых решений (ОДР) задачи изображена на рисунке 6.2 (отмечена штриховкой).

Окончательно получаем следующую ЗНП: минимизировать ЦФ

$$f(x) = -\frac{1000}{1 + (x_1 - 6)^2 + (x_2 - 8)^2}$$

при ограничениях

$$x_1^2 + x_2^2 \leq 4,$$

$$x_2 \leq x_1^2.$$

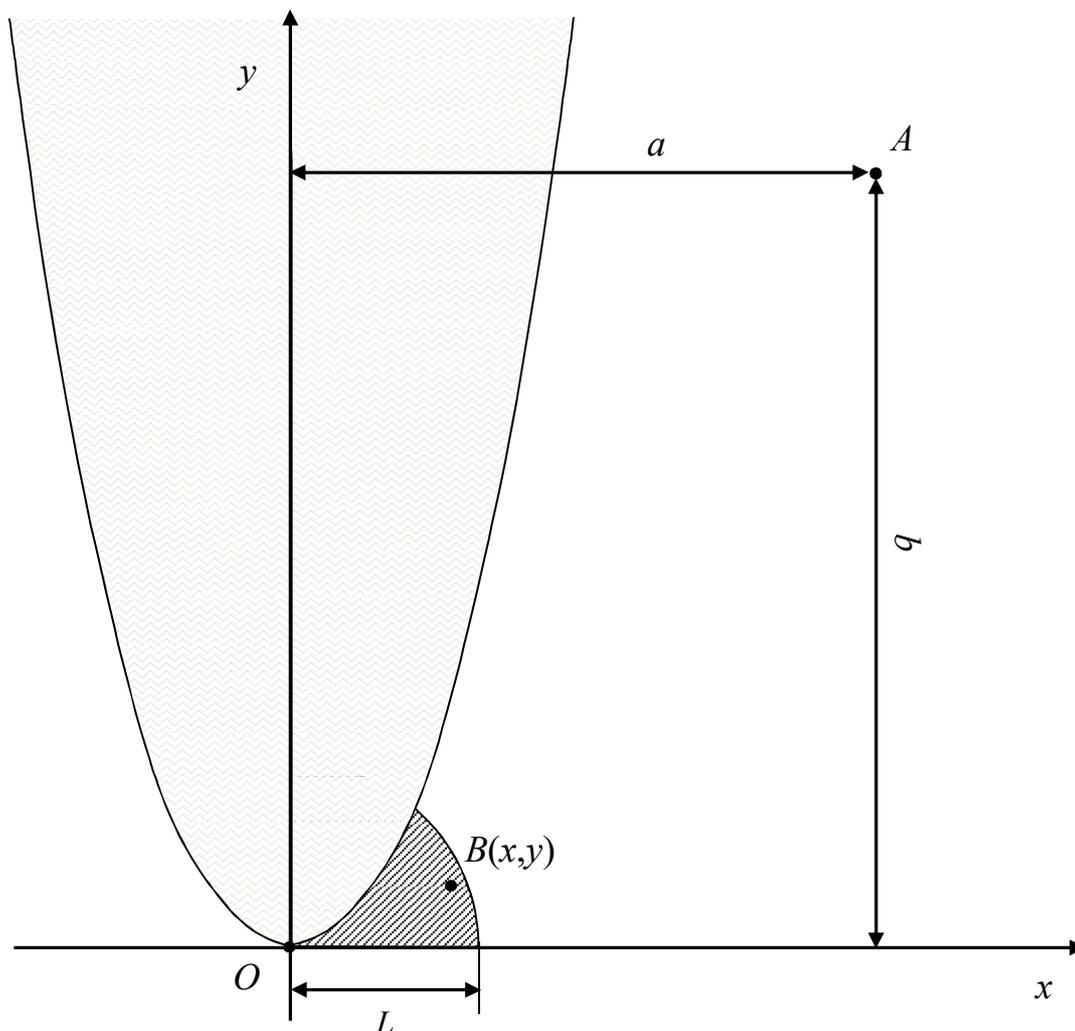


Рисунок 6.2 – Область допустимых решений задачи

В данной задаче можно не максимизировать интенсивность принимаемого излучения, а минимизировать расстояние от приемной антенны до вышки, т.е. принять

$$f(x) = \sqrt{1 + (x_1 - 6)^2 + (x_2 - 8)^2} \rightarrow \min.$$

Или, как уже было показано выше, минимизировать не расстояние, а его квадрат, что упростит ЦФ:

$$f(x) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min. \quad (6.1)$$

Обычно ограничения-неравенства принято записывать в виде  $g(x) \geq 0$ , таким образом,

$$g_1(x) = 4 - x_1^2 - x_2^2 \geq 0, \quad (6.2)$$

$$g_2(x) = x_1^2 - x_2 \geq 0. \quad (6.3)$$

Не все методы решения ЗНП предназначены для решения задач с ограничениями такого вида. Поэтому ниже будут предложены другие варианты ограничений.

Сразу определим векторы-градиенты ЦФ и ограничений:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - 12 \\ 2x_2 - 16 \end{pmatrix};$$

$$\nabla g_1(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix}; \quad \nabla g_2(x) = \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}.$$

## 6.2. Классификация методов решения задач нелинейного программирования

Выполним классификацию методов решения ЗНП в зависимости от типов ЗО, для решения которых они предназначены.

1. ЗНП с ограничениями-неравенствами:

$$f(x) \rightarrow \min, \quad x \in R^n, \quad (6.4)$$

$$g_j(x) \geq 0, \quad j = 1, 2, \dots, m, \quad (6.5)$$

где  $f(x)$  и  $g_j(x)$  – нелинейные функции. К этому виду ЗНП относится рассмотренная выше задача. Для решения таких задач применяются следующие методы:

- линеаризации;
- допустимых направлений Зойтендейка;
- проекции градиента для случая нелинейных ограничений;
- один из вариантов метода условного градиента;
- ряд штрафных методов.

*Замечание.* В методах условного градиента и проекции градиента требуется находить проекцию точки на ОДР задачи. Поиск проекции точки даже на такие фигуры, как парабола, эллипс и т.п., является достаточно сложной самостоятельной нелинейной задачей. В этих и некоторых других методах будем использовать только одно ограничение вида (6.2), как если бы озера не было. Найти проекцию точки на окружность достаточно просто.

2. ЗНП с ограничениями-равенствами:

$$f(x) \rightarrow \min, \quad x \in R^n, \quad (6.6)$$

$$h_k(x) = 0, \quad k = 1, 2, \dots, m, \quad (6.7)$$

где  $f(x)$  и  $h_k(x)$  – нелинейные функции. Для решения таких задач применяются следующие методы:

- замены переменных;
- множителей Лагранжа;
- линеаризации;
- ряд штрафных методов.

Как будет показано далее, решение задачи (6.1)–(6.3) не изменится, если ее ограничения заменить равенствами. Поэтому в данных методах будем использовать ограничение

$$h_1(x) = 4 - x_1^2 - x_2^2 = 0 \quad (6.8)$$

и/или

$$h_2(x) = x_1^2 - x_2 = 0. \quad (6.9)$$

3. ЗНП общего вида:

$$f(x) \rightarrow \min, \quad x \in R^n, \quad (6.10)$$

$$g_j(x) \geq 0, \quad j = 1, 2, \dots, m, \quad (6.11)$$

$$h_k(x) = 0, \quad k = 1, 2, \dots, p. \quad (6.12)$$

Это самый сложный тип ЗНП. Задачи такого вида можно решать методом линеаризации, но он очень неустойчивый. Также можно использовать штрафные функции с комбинацией двух видов штрафов –

первого для ограничений-равенств и второго для ограничений-неравенств. В некоторых случаях удается свести неравенства к равенствам (или наоборот), тогда для решения можно использовать другие методы, перечисленные в первых двух пунктах.

4. ЗНП с линейными ограничениями:

$$f(x) \rightarrow \min, \quad x \in R^n, \quad (6.13)$$

$$g_j(x) = a^T x \geq 0, \quad j = 1, 2, \dots, m. \quad (6.14)$$

Для решения таких задач, в принципе, можно использовать методы, перечисленные в первом пункте, так как это частный случай ЗНП с ограничениями-неравенствами. Но при этом может ухудшиться сходимость. Дополнительно можно использовать:

- метод Франка – Вульфа;
- один из вариантов метода условного градиента;
- метод проекции градиента для случая линейных ограничений.

В этом случае будем использовать ограничение

$$g(x) = 4 - x_1 - x_2 \geq 0, \quad (6.15)$$

а если в методе требуются условия неотрицательности переменных, то дополнительно еще ограничения

$$x_1 \geq 0, \quad (6.16)$$

$$x_2 \geq 0, \quad (6.17)$$

т.е. из задачи убираем озеро, а приемная антенна будет располагаться не в радиусе 2 км от населенного пункта, а в пределах треугольной области (рисунок 6.3). Например, можно положить, что это берег реки или канала и техническая возможность перенести антенну на другой берег отсутствует.

5. ЗНП вида

$$f(x) \rightarrow \min, \quad x \in R^n, \quad (6.18)$$

$$a_i \leq x_i \leq b_i, \quad i = 1, 2, \dots, n. \quad (6.19)$$

Здесь ОДР представляет собой прямоугольник (при  $n > 2$  – прямоугольный параллелепипед). Задачи такого типа решаются одним из вариантов метода условного градиента, хотя могут быть решены и другими методами, например перечисленными в пункте 4.

Для использования данного метода зададим ОДР соответствующей формы (рисунок 6.4):

$$0 \leq x_1 \leq 10, 0 \leq x_2 \leq 2. \quad (6.20)$$

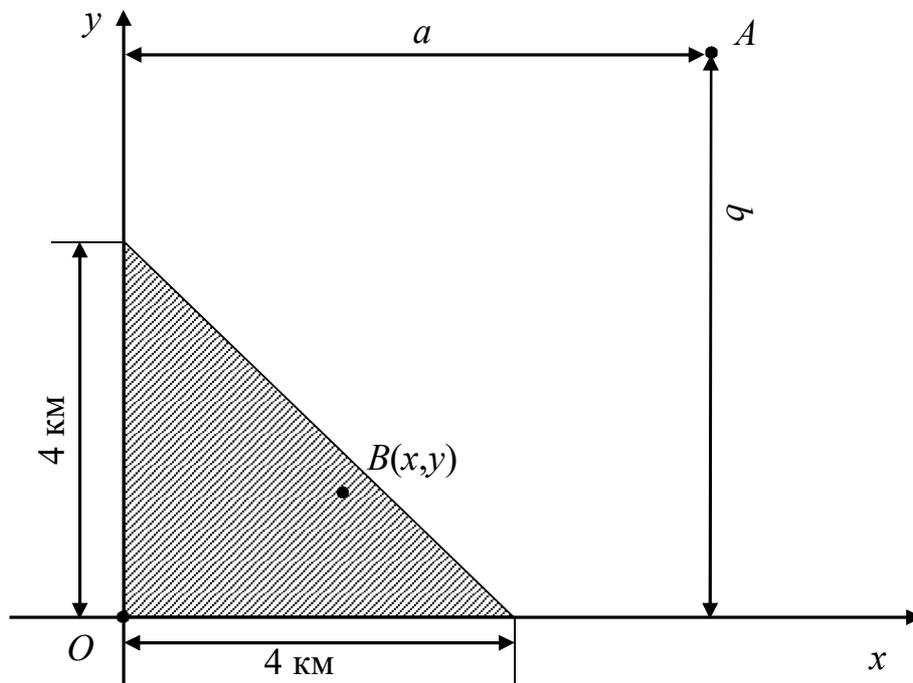


Рисунок 6.3 – Область допустимых решений задачи с линейными ограничениями

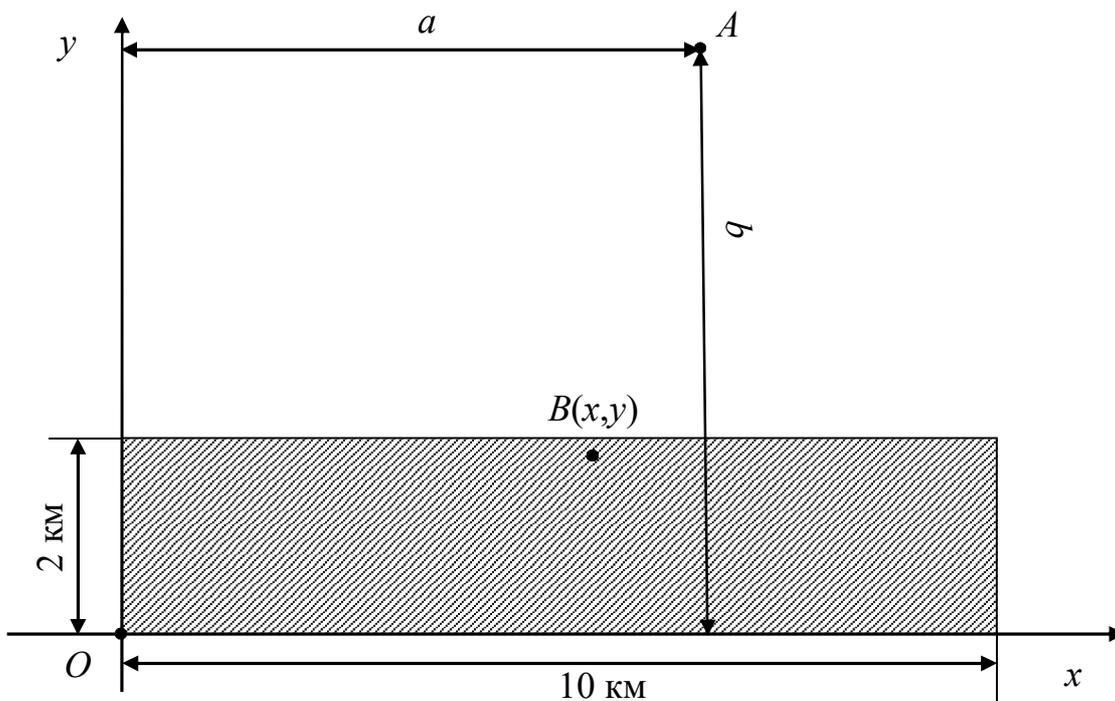


Рисунок 6.4 – Задача с ОДР в форме прямоугольника

Указанные ограничения также можно трактовать как невозможность перенести антенну через водоем, железную дорогу и т.п.

## 6.3. Задачи с ограничениями в виде равенств

### 6.3.1. Метод замены переменных

Рассмотрим общую задачу оптимизации с ограничениями-равенствами (6.6), (6.7). Функции  $f(x)$  и  $h_k(x)$  предполагаются непрерывными и дифференцируемыми.

Задача нелинейного программирования, в принципе, может быть решена как задача безусловной оптимизации путем исключения из целевой функции (6.6) независимых переменных с помощью заданных равенств (6.7). Наличие ограничений в виде равенств фактически позволяет уменьшить размерность исходной задачи с  $n$  до  $n - m$ . Для решения задачи безусловной минимизации можно использовать методы, изложенные ранее.

Пусть имеется  $m$  ограничений равенств, причем  $m < n$ . Если  $m = n$ , то задача оптимизации сводится к решению системы уравнений (6.7) и оптимизация не нужна.

Таким образом, экстремальная точка функции  $f(x)$  ищется лишь по тем значениям  $x$ , которые удовлетворяют независимой системе (6.7) при  $m < n$ , т.е. в некоторой части  $D$  пространства  $R^n$  ( $D \subset R^n$ ).

Данный метод называется *методом замены переменных* (МЗП) и является аналитическим. Запишем его алгоритм.

1. Представить первые  $m$  переменных  $x_1, \dots, x_m$ , относительно которых мы хотим разрешить систему (6.7), в виде вектора  $y$ , а остальные переменные в виде вектора  $u$ :

$$x = \begin{pmatrix} y \\ u \end{pmatrix}; \quad y = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \equiv \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}; \quad u = \begin{pmatrix} x_{m+1} \\ \vdots \\ x_n \end{pmatrix} \equiv \begin{pmatrix} u_1 \\ \vdots \\ u_{n-m} \end{pmatrix}. \quad (6.21)$$

*Замечание.* В общем случае можно выбирать не первые переменные, а любые  $m$  переменных.

2. Решить систему (6.7):

$$y = \Psi(u), \quad u \in R^{n-m}. \quad (6.22)$$

Если  $m = n$ , то перейти на шаг 4. В противном случае необходимо подставить  $y$  в целевую функцию  $F(u) = f(y, u) = f[\Psi(u), u]$ . Получим функцию  $n - m$  переменных  $u_1, \dots, u_{n-m}$ , на которую не наложено никаких ограничений, т.е. получим задачу оптимизации без ограничений.

3. Использовать один из известных методов безусловной оптимизации для идентификации стационарных точек целевой функции.

4. Выбрать из полученных точек оптимальное решение. Для этого следует использовать рассмотренные далее необходимые и достаточные условия оптимальности решения ЗНП.

*Пример 1.* Решим задачу определения наилучшего положения приемной антенны с ограничениями-равенствами:

$$f(x) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min,$$

$$h_1(x) = 4 - x_1^2 - x_2^2 = 0,$$

$$h_2(x) = x_1^2 - x_2 = 0.$$

Шаг 1. Так как в данном случае  $n = m = 2$ , то вектор  $y = x$ , а вектор  $u$  остается пустым.

Шаг 2. Получаем систему уравнений

$$\begin{cases} 4 - y_1^2 - y_2^2 = 0, \\ y_1^2 - y_2 = 0 \end{cases}$$

и решаем ее. Из второго уравнения  $y_1^2 = y_2$ , подставляем в первое уравнение:

$$4 - y_2 - y_2^2 = 0 \Rightarrow y_2^{(1)} \approx 1,562, \quad y_2^{(2)} \approx 2,562.$$

Тогда  $y_1^{(1)} \approx \pm 1,25$ ,  $y_1^{(2)} \approx \pm 1,6$ . Таким образом, получили четыре решения системы:

$$y^{(1)} \approx \begin{pmatrix} 1,25 \\ 1,562 \end{pmatrix}, \quad y^{(2)} \approx \begin{pmatrix} -1,25 \\ 1,562 \end{pmatrix}, \quad y^{(3)} \approx \begin{pmatrix} 1,6 \\ 2,562 \end{pmatrix}, \quad y^{(4)} \approx \begin{pmatrix} -1,6 \\ 2,562 \end{pmatrix}.$$

Так как  $n = m$ , переходим на шаг 4.

Шаг 4. Очевидно, что точки  $y^{(2)}$  и  $y^{(4)}$  не могут быть оптимальными, потому что выходят за пределы первого квадранта. Но это нужно показать математически. Необходимые и достаточные условия оптимальности решения ЗНП будут изучены ниже, а пока можно проверить, какие из точек не удовлетворяют ограничениям, т.е. они точно не будут являться решением задачи:

$$\begin{aligned} h_1(y^{(1)}) &= 0, & h_2(y^{(1)}) &= 0, \\ h_1(y^{(2)}) &= 0, & h_2(y^{(2)}) &= 0, \\ h_1(y^{(3)}) &= 0, & h_2(y^{(3)}) &\approx -5,123 \neq 0, \\ h_1(y^{(4)}) &= 0, & h_2(y^{(4)}) &\approx -5,123 \neq 0. \end{aligned}$$

Следовательно, точки  $y^{(3)}$  и  $y^{(4)}$  не являются решениями задачи. Вычислим значение ЦФ в двух оставшихся точках:

$$f(y^{(1)}) \approx 65,02, \quad f(y^{(2)}) \approx 95,011.$$

Минимальный квадрат расстояния до вышки в точке  $y^{(1)}$ , таким образом, составит

$$x^* \approx \begin{pmatrix} 1,25 \\ 1,562 \end{pmatrix}, \quad f^* \approx 65,02.$$

Данный критерий не является надежным, так как в точке локального максимума или в седловой точке значение ЦФ может оказаться меньше, чем в точке локального минимума. Поэтому на практике следует использовать специальные критерии оптимальности, которые мы рассмотрим далее.

Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 8,063$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 15,38 \text{ Вт/м}^2.$$

*Пример 2.* Решим задачу определения наилучшего положения приемной антенны с одним ограничением:

$$\begin{aligned} f(x) &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ h(x) &= 4 - x_1^2 - x_2^2 = 0. \end{aligned}$$

Шаг 1. В данном случае  $n = 2$ ,  $m = 1$ ,  $y = x_1$ ,  $u = x_2$ ,  $x = \begin{pmatrix} y \\ u \end{pmatrix}$ .

Шаг 2. В системе будет только одно уравнение:

$$4 - y^2 - u^2 = 0.$$

Решаем его относительно независимой переменной  $u$ :

$$y = \Psi(u) = \pm\sqrt{4 - u^2}, \quad u \in [-2, 2].$$

Шаг 3. Получили две задачи безусловной оптимизации:

$$\begin{aligned} F_1(u) &= 1 + \left(\sqrt{4 - u^2} - 6\right)^2 + (u - 8)^2 \rightarrow \min; \\ F_2(u) &= 1 + \left(-\sqrt{4 - u^2} - 6\right)^2 + (u - 8)^2 \rightarrow \min. \end{aligned}$$

Минимизируем только функцию  $F_1(x)$ , так как выше было установлено, что точки с отрицательными координатами не являются решением задачи:

$$\begin{aligned} \frac{dF_1(u)}{du} &= -\frac{2u(\sqrt{4 - u^2} - 6)}{\sqrt{4 - u^2}} + 2(u - 8) = \frac{-2u(\sqrt{4 - u^2} - 6) + 2(u - 8)\sqrt{4 - u^2}}{\sqrt{4 - u^2}} = \\ &= \frac{12u - 16\sqrt{4 - u^2}}{\sqrt{4 - u^2}} = 0, \quad u \in [-2, 2]. \end{aligned}$$

Соответственно  $\sqrt{4 - u^2} \neq 0 \Rightarrow u \neq \pm 2 \Rightarrow u \in (-2, 2)$ , и далее в уравнении можно оставить только числитель:

$$12u - 16\sqrt{4 - u^2} = 0 \Rightarrow 12u = 16\sqrt{4 - u^2}, \quad u \in (-2, 2).$$

Возведем обе части уравнения в квадрат, при этом получим дополнительное ограничение  $u \geq 0$ , так как корень не может быть отрицательным:

$$144u^2 = 256(4 - u^2), \quad u \in [0, 2) \Rightarrow$$

$$400u^2 = 1024 \Rightarrow u^2 = 2,56.$$

В диапазоне  $u \in [0, 2)$  существует только одно решение:  $u = 1,6$ .

Так как  $y = \sqrt{4 - u^2}$ , то получаем  $y = 1,2$ .

Шаг 4. Получена только одна точка, поэтому ее и принимаем за оптимальную:

$$x^* = \begin{pmatrix} 1,2 \\ 1,6 \end{pmatrix}, \quad f^* = 65.$$

Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 8,062$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 15,39 \text{ Вт/м}^2.$$

МЗП применим лишь в случаях, когда уравнения ограничения можно разрешить относительно некоторого набора независимых переменных. При наличии большого числа ограничений МЗП становится весьма трудоемкой процедурой. Кроме того, возможны ситуации, когда уравнение не удастся разрешить относительно переменной, например

$$h(x) = x_1^2 \cdot x_3 + x_2 \cdot x_3^2 + x_2^{-1} \cdot x_1.$$

В этом случае целесообразно использовать метод множителей Лагранжа.

### 6.3.2. Метод множителей Лагранжа

С помощью *метода множителей Лагранжа* (ММЛ) устанавливают необходимое условие, позволяющее идентифицировать точки экстремума в задаче оптимизации с ограничениями-равенствами. При этом задача с ограничениями преобразуется в эквивалентную задачу безусловной оптимизации, в которой фигурируют некоторые неизвестные параметры – множители Лагранжа.

Рассмотрим задачу оптимизации:

$$f(x) \rightarrow \min, \quad (6.23)$$

$$h(x) = 0. \quad (6.24)$$

В соответствии с ММЛ эта задача преобразуется в задачу оптимизации без ограничений

$$L(x, \lambda) = f(x) + \lambda h(x) \rightarrow \min, \quad (6.25)$$

где  $L(x, \lambda)$  – функция Лагранжа,  $\lambda = \text{const}$  – множитель Лагранжа, на знак которого никаких требований не накладывается.

ММЛ можно распространить на случай, когда задача оптимизации имеет несколько ограничений-равенств. Рассмотрим общую задачу (6.6)–(6.7). Функция Лагранжа принимает следующий вид:

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i h_i(x), \quad (6.26)$$

где  $\lambda_1, \dots, \lambda_m$  – множители Лагранжа, т.е. неизвестные параметры, значения которых необходимо определить.

Запишем алгоритм ММЛ.

1. Составить функцию Лагранжа.
2. Найти частные производные

$$\frac{\partial L}{\partial x_j}, \quad \frac{\partial L}{\partial \lambda_i}, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m,$$

и, приравнивая их нулю, составить расширенную систему  $n + m$  уравнений с  $n + m$  неизвестными  $x_j, \lambda_i$ :

$$\begin{cases} \nabla_x L = 0 \\ \nabla_\lambda L = 0 \end{cases} \Rightarrow \begin{cases} \frac{\partial L}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^m \lambda_i \frac{\partial h_i(x)}{\partial x_j} = 0, \quad j = 1, 2, \dots, n, \\ \frac{\partial L}{\partial \lambda_i} = h_i(x) = 0, \quad i = 1, 2, \dots, m. \end{cases} \quad (6.27)$$

Это и есть *необходимое условие экстремума функции Лагранжа*.

3. Решить полученную систему и определить стационарные точки функции  $L$ .

4. Классифицировать полученные стационарные точки на основе вычисления элементов матрицы Гессе функции  $L$ , рассматриваемой как функции только аргумента  $x$ :

$$H_L(x, \lambda) = \begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2} & \dots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial^2 L}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 L}{\partial x_n^2} \end{pmatrix}. \quad (6.28)$$

Выбрать из полученных точек оптимальное решение  $x^*$ .

Точка  $(x^*, \lambda^*)$  является седловой точкой функции Лагранжа  $L(x, \lambda)$ , если функцию Лагранжа рассматривать как функцию переменных  $x$  и  $\lambda$ .

Проиллюстрируем ММЛ на примере.

*Пример.* Решим задачу определения наилучшего положения приемной антенны с одним ограничением:

$$\begin{aligned} f(x) &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ h(x) &= 4 - x_1^2 - x_2^2 = 0. \end{aligned}$$

Шаг 1. Соответствующая задача оптимизации без ограничений записывается в следующем виде:

$$L(x, \lambda) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 + \lambda(4 - x_1^2 - x_2^2) \rightarrow \min.$$

Шаг 2. Находим частные производные и составляем расширенную систему уравнений:

$$\begin{cases} \frac{\partial L}{\partial x_1} = 2(x_1 - 6) - 2\lambda x_1 = 0; \\ \frac{\partial L}{\partial x_2} = 2(x_2 - 8) - 2\lambda x_2 = 0; \\ \frac{\partial L}{\partial \lambda} = 4 - x_1^2 - x_2^2 = 0. \end{cases}$$

Шаг 3. Решаем полученную систему. Из первого уравнения

$$2x_1 - 2\lambda x_1 - 12 = 0 \Rightarrow x_1 = \frac{12}{2 - 2\lambda} = \frac{6}{1 - \lambda},$$

из второго уравнения

$$2x_2 - 2\lambda x_2 - 16 = 0 \Rightarrow x_2 = \frac{16}{2 - 2\lambda} = \frac{8}{1 - \lambda}.$$

Подставляем в третье уравнение и решаем его относительно  $\lambda$ :

$$4 - \left(\frac{6}{1 - \lambda}\right)^2 - \left(\frac{8}{1 - \lambda}\right)^2 = \frac{4(1 - \lambda)^2 - 36 - 64}{(1 - \lambda)^2} = 0.$$

Следовательно,  $\lambda \neq 1$ , поэтому знаменатель отбрасываем:

$$4(1 - \lambda)^2 - 36 - 64 = 0 \Rightarrow 4(1 - \lambda)^2 = 100 \Rightarrow \lambda^{(1)} = -4, \lambda^{(2)} = 6.$$

Соответственно

$$x^{(1)} = \begin{pmatrix} \frac{6}{1 - \lambda^{(1)}} \\ 8 \\ \frac{8}{1 - \lambda^{(1)}} \end{pmatrix} = \begin{pmatrix} 1, 2 \\ 1, 6 \end{pmatrix}, \quad x^{(2)} = \begin{pmatrix} \frac{6}{1 - \lambda^{(2)}} \\ 8 \\ \frac{8}{1 - \lambda^{(2)}} \end{pmatrix} = \begin{pmatrix} -1, 2 \\ -1, 6 \end{pmatrix}.$$

Шаг 4. Для того чтобы проверить, соответствует ли стационарная точка  $x^*$  минимуму, вычислим матрицу Гессе функции  $L(x, \lambda)$ , рассматриваемой как функции только от аргумента  $x$ :

$$H_L(x, \lambda) = \begin{pmatrix} 2 - 2\lambda & 0 \\ 0 & 2 - 2\lambda \end{pmatrix}.$$

В точке  $(x^{(1)}, \lambda^{(1)})$  имеем

$$H_L(x^{(1)}, \lambda^{(1)}) = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}.$$

т.е. гессиан положительно определен. Это означает, что в окрестности точки  $(x^{(1)}, \lambda^{(1)})$  функция Лагранжа  $L(x, \lambda)$  является выпуклой. Следовательно, координаты  $x^* = (1, 2; 1, 6)^T$  определяют точку минимума.

В точке  $(x^{(2)}, \lambda^{(2)})$  имеем

$$H_L(x^{(2)}, \lambda^{(2)}) = \begin{pmatrix} -10 & 0 \\ 0 & -10 \end{pmatrix}.$$

т.е. гессиан отрицательно определен. Это означает, что в окрестности точки  $(x^{(1)}, \lambda^{(1)})$  функция Лагранжа  $L(x, \lambda)$  является вогнутой. Следовательно, координаты  $x^* = (-1, 2; -1, 6)^T$  определяют точку максимума.

Таким образом, глобальный минимум достигается в точке с координатами  $x_1^* = 1, 2$  и  $x_2^* = 1, 6$ . Значение ЦФ  $f^* = 65$ . Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 15,39 \text{ Вт/м}^2.$$

Для некоторых задач расширенная система  $n + m$  уравнений с  $n + m$  неизвестными может не иметь решений и ММЛ окажется неприемлемым. Однако на практике такие задачи встречаются достаточно редко.

### 6.3.3. Экономическая интерпретация множителей Лагранжа

Множители Лагранжа рассматривались выше как параметры, значения которых выбираются таким образом, чтобы выполнялись ограничения задачи. С экономической точки зрения множители Лагранжа интерпретируются как неявные (теневые) цены ресурсов, определяемых ограничениями; оптимальные значения множителей Лагранжа играют важную роль в анализе чувствительности решений. Для того чтобы пояснить эту интерпретационную схему, рассмотрим следующую задачу оптимизации с двумя переменными и одним ограничением в виде равенства:

$$\begin{aligned} f(x_1, x_2) &\rightarrow \min, \\ h(x_1, x_2) &= b, \end{aligned} \quad (6.29)$$

где константа  $b$  характеризует наличие некоторого ресурса. Составим функцию Лагранжа:

$$L(x, \lambda) = f(x) + \lambda [h(x) - b].$$

Предположим, что стационарная точка функции  $L$  соответствует глобальному минимуму:

$$\begin{cases} \frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} + \lambda \frac{\partial h}{\partial x_1} = 0, \\ \frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} + \lambda \frac{\partial h}{\partial x_2} = 0; \end{cases} \quad (6.30)$$

$$\frac{\partial L}{\partial \lambda} = h(x) - b = 0. \quad (6.31)$$

Пусть  $\lambda^*$  – оптимальное значение множителя Лагранжа, а  $x^*$  – оптимальное решение задачи.

Далее, пусть минимум функции  $L(x, \lambda)$  при  $\lambda = \lambda^*$  достигается в точке  $x = x^*$ , причем

$$h(x^*) = b,$$

$$f(x^*) = L(x^*, \lambda^*) = f^*.$$

Очевидно, что оптимальные значения  $(x^*, \lambda^*)$  связаны функциональной зависимостью с величиной  $b$ , задающей границу наличия дефицитного ресурса.

Изменения оптимального значения целевой функции  $f^*$ , обусловленные изменениями  $b$ , описываются частной производной  $\partial f^* / \partial b$ . По правилу дифференцирования сложной функции имеем

$$\frac{\partial f^*}{\partial b} = \frac{\partial f^*}{\partial x_1^*} \cdot \frac{\partial x_1^*}{\partial b} + \frac{\partial f^*}{\partial x_2^*} \cdot \frac{\partial x_2^*}{\partial b}. \quad (6.32)$$

Дифференцируя обе части ограничения  $h(x) - b = 0$ , получаем

$$\frac{dh}{db} = \frac{\partial h}{\partial x_1^*} \cdot \frac{\partial x_1^*}{\partial b} + \frac{\partial h}{\partial x_2^*} \cdot \frac{\partial x_2^*}{\partial b} - 1 = 0. \quad (6.33)$$

Умножим обе части равенства (6.33) на  $\lambda^*$  и суммируем с выражением (6.32), получим

$$\frac{\partial f^*}{\partial b} = -\lambda^* + \sum_{j=1}^2 \left[ \frac{\partial f^*}{\partial x_j^*} + \lambda^* \cdot \frac{\partial h}{\partial x_j^*} \right] \cdot \frac{\partial x_j^*}{\partial b}. \quad (6.34)$$

Так как  $x^*$  и  $\lambda^*$  удовлетворяют уравнениям (6.31) и (6.32), то равенство (6.34) приводится к виду

$$\frac{\partial f^*}{\partial b} = -\lambda^*. \quad (6.35)$$

Таким образом, из выражения (6.35) следует, что скорость изменения оптимального значения  $f^*$ , вызываемого изменением  $b$ , определяется оптимальным значением  $\lambda^*$ . Другими словами, величина изменения оптимального значения целевой функции, обусловленного единичным увеличением правой части ограничения, задается множителем Лагранжа. В зависимости от знака  $\lambda^*$  значения  $f^*$  при

изменении  $b$  могут увеличиваться или уменьшаться. При  $m$  ограничениях имеем

$$\frac{\partial f^*}{\partial b_i} = -\lambda_i^*, \quad i = 1, 2, \dots, m.$$

## 6.4. Необходимые и достаточные условия оптимальности

### 6.4.1. Влияние ограничений на решение задач оптимизации

Ранее были рассмотрены необходимые и достаточные условия оптимальности решения задач оптимизации без ограничений. Ограничения на управляемые переменные существенно уменьшают размеры области, в которой проводится поиск экстремума (ОДР). На первый взгляд может показаться, что уменьшение размеров ОДР должно упростить процедуру поиска экстремума. Но напротив, процесс оптимизации становится более сложным, так как установленные выше критерии оптимальности нельзя использовать при наличии ограничений. Может нарушаться даже основное условие – равенство нулю первой производной в стационарной точке.

Например, безусловный минимум функции  $f(x) = (x - 2)^2$  имеет место при  $x^* = 2$  (т.е.  $\nabla f(x^*) = 0$ ). Но если задача решается с учетом ограничения  $x \geq 4$ , то условный минимум достигается в точке  $x^* = 4$ , где  $\nabla f(x^*) = 4 \neq 0$ . Может показаться, что условный минимум достигается на границе ОДР, но и это не так. Например, если решать задачу с учетом ограничения  $x \leq 4$ , то минимумом останется точка  $x^* = 2$ .

Если ограничений в задаче несколько, то процесс ее решения усложняется еще сильнее, так как условный минимум:

– может совпадать с безусловным минимумом, а именно являться внутренней точкой ОДР, т.е. не лежать на поверхности ни одного из ограничений (в этом случае  $\nabla f(x^*) = 0$ );

– может лежать на поверхности некоторых ограничений, но не лежать на поверхности других, т.е. являться граничной точкой ОДР (в этом случае градиент  $\nabla f(x^*)$  может быть нулевым, а может не быть);

– может лежать на пересечении поверхностей всех ограничений (в этом случае обычно  $\nabla f(x^*) \neq 0$ , но не обязательно).

Будем называть ограничение-неравенство  $g(x)$  *активным* в некоторой точке  $x$ , если данная точка лежит на его поверхности (т.е. обращает его в равенство), и *неактивным*, если данная точка не лежит на его поверхности.

Например, ограничение  $g(x) = 4 - x_1^2 - x_2^2 \geq 0$  в точке  $x = (0, 2)^T$  является активным, так как  $4 - 0 - 4 = 0$ , а в точке  $x = (1, 1)^T$  – неактивным, так как  $4 - 1 - 1 = 2 > 0$ .

Если в точке  $x$  ограничение  $g(x)$  выполняется, то такая точка называется *внутренней* (т.е. она лежит внутри ОДР, определяемой данным ограничением), точка называется *внешней*, если ограничение  $g(x)$  для нее не выполняется (т.е. она лежит вне ОДР).

Например, для ограничения  $g(x) = 4 - x_1^2 - x_2^2 \geq 0$  точка  $x = (1, 1)^T$  является внутренней, так как  $4 - 1 - 1 = 2 \geq 0$ , а точка  $x = (2, 2)^T$  – внешней, так как  $4 - 4 - 4 = -4 < 0$ .

### **6.4.2. Необходимые и достаточные условия оптимальности задач с ограничениями общего вида**

Мы установили, что множители Лагранжа можно использовать при построении критериев оптимальности для задач оптимизации с ограничениями в виде равенств. Кун и Таккер обобщили этот подход на случай общей задачи нелинейного программирования (НП) с ограничениями как в виде равенств, так и в виде неравенств.

Рассмотрим общую задачу нелинейного программирования

$$\begin{aligned}
f(x) &\rightarrow \min, \\
g_i(x) &\begin{matrix} \geq \\ \leq \end{matrix} 0, \quad i = 1, 2, \dots, m, \\
h_l(x) &= 0, \quad l = 1, 2, \dots, p,
\end{aligned} \tag{6.36}$$

где  $f(x)$ ,  $g_i(x)$ ,  $h_l(x)$  – непрерывные и дифференцируемые функции.

Необходимые и достаточные условия первого порядка выражаются теоремой.

*Теорема.* Для того чтобы точка  $x^*$  была глобальным минимумом задачи (6.36), необходимо и достаточно, чтобы выполнялись следующие условия:

- а) ЦФ  $f(x)$  была выпукла;
- б) функция-ограничение  $g_i(x)$  была вогнута, если  $g_i(x) \geq 0$ , и выпукла, если  $g_i(x) \leq 0$ ;
- в) ограничения-равенства  $h_l(x)$  являлись линейными функциями;
- г) существовали такие числа  $\mu_i^* \geq 0$  и  $\lambda_l^*$  (произвольного знака), что

$$\nabla f(x^*) \mp \sum_{i=1}^m \mu_i^* \nabla g_i(x^*) + \sum_{l=1}^p \lambda_l^* \nabla h_l(x^*) = 0, \tag{6.37}$$

$$\mu_i^* g_i(x^*) = 0, \quad i = 1, 2, \dots, m, \tag{6.38}$$

$$h_l(x^*) = 0, \quad l = 1, 2, \dots, p. \tag{6.39}$$

Знак «минус» перед вторым слагаемым ЦФ ставится в том случае, если  $g_i(x) \geq 0$ , «плюс» – если  $g_i(x) \leq 0$ . Для выполнения условий Куна – Таккера необходимо:

- 1) чтобы существовала такая точка  $\bar{x}$ , что  $g_i(\bar{x}) > 0$  (либо  $g_i(\bar{x}) < 0$ , если  $g_i(x) \leq 0$ ) и  $h_l(\bar{x}) = 0$ ;

2) чтобы в точке  $x^*$  градиенты  $\nabla g_i(x^*)$  и  $\nabla h_l(x^*)$  были линейно независимы для всех  $i \in I = \{i \mid g_i(x^*) = 0\}$  и  $l = 1, 2, \dots, p$ , т.е. для тех  $i$ , для которых ограничения  $g_i$  активны в точке  $x^*$ .

В задачах с ограничениями только типа равенств

$$\begin{aligned} f(x) &\rightarrow \min, \\ h_l(x) &= 0, \quad l = 1, 2, \dots, p, \end{aligned}$$

функция Лагранжа имеет вид

$$L(x, \lambda) = f(x) + \sum_{l=1}^p \lambda_l h_l(x).$$

Для этой функции условия оптимальности первого порядка выражаются следующими уравнениями:

$$\begin{aligned} \nabla_x L(x, \mu) &= \nabla f(x) + \sum_{l=1}^p \lambda_l \nabla h_l(x); \\ \nabla_\lambda L(x, \mu) &= h_l(x) = 0, \quad l = 1, 2, \dots, p. \end{aligned}$$

То есть условия Куна – Таккера (6.37) и (6.39) совпадают с условиями оптимальности для задачи Лагранжа.

*Пример.* Рассмотрим задачу определения наилучшего положения приемной антенны с ограничениями-неравенствами:

$$\begin{aligned} f(x) &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ g_1(x) &= 4 - x_1^2 - x_2^2 \geq 0, \\ g_2(x) &= x_1^2 - x_2 \geq 0. \end{aligned}$$

Проверим выполнение условий Куна – Таккера. Вычислим матрицы вторых производных целевой функции и функций ограничений:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - 12 \\ 2x_2 - 16 \end{pmatrix}, \quad H_f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Матрица Гессе  $H_f$  положительно определена при всех значениях  $x$ , следовательно,  $f(x)$  – выпуклая функция.

Матрица Гессе для функции  $g_1(x)$  имеет вид

$$\nabla g_1(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix}, \quad H_{g_1}(x) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Матрица  $H_{g_1}$  отрицательно определена, следовательно, функция  $g_1(x)$  является вогнутой.

Матрица Гессе для функции  $g_2(x)$  имеет вид

$$\nabla g_2(x) = \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}, \quad H_{g_2}(x) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Матрица  $H_{g_2}$  положительно полуопределена, следовательно, функция  $g_2(x)$  является выпуклой.

Ограничений-равенств в задаче нет. Так как условия (б) теоремы не выполнены, то необходимые и достаточные условия оптимальности для данной задачи не выполняются.

Тем не менее попробуем найти точку оптимума  $x^*$ . Уравнение (6.37) принимает следующий вид:

$$\frac{\partial f}{\partial x_j} - \mu_1 \frac{\partial g_1}{\partial x_j} - \mu_2 \frac{\partial g_2}{\partial x_j} = 0, \quad j = 1, 2.$$

При  $j = 1$  и  $j = 2$  соответственно можно получить выражения

$$2x_1 - 12 + 2x_1\mu_1 - 2x_1\mu_2 = 0;$$

$$2x_2 - 16 + 2x_2\mu_1 + \mu_2 = 0.$$

Таким образом, условия Куна – Таккера (6.37)–(6.39) для нашего примера имеют следующий вид:

$$2x_1 - 12 + 2x_1\mu_1 - 2x_1\mu_2 = 0;$$

$$2x_2 - 16 + 2x_2\mu_1 + \mu_2 = 0;$$

$$\mu_1(4 - x_1^2 - x_2^2) = 0;$$

$$\mu_2(x_1^2 - x_2) = 0;$$

$$\mu_{1,2} \geq 0.$$

Из четвертого уравнения находим  $\mu_2 = 0$  (т.е. второе ограничение не является активным) или  $x_1^2 - x_2 = 0$  (является активным). Рассмотрим оба случая.

1. Пусть  $\mu_2 = 0$ . Тогда

$$2x_1(1 + \mu_1) - 12 = 0;$$

$$2x_2(1 + \mu_1) - 16 = 0;$$

$$\mu_1(4 - x_1^2 - x_2^2) = 0;$$

$$\mu_1 \geq 0.$$

Из первых двух уравнений находим

$$x_1 = \frac{6}{1 + \mu_1}, \quad x_2 = \frac{8}{1 + \mu_1},$$

подставляем в третье уравнение:

$$\mu_1 \left( 4 - \left( \frac{6}{1 + \mu_1} \right)^2 - \left( \frac{8}{1 + \mu_1} \right)^2 \right) = 0,$$

получаем еще два варианта решения.

1.1. Полагаем  $\mu_1 = 0$  (т.е. первое ограничение также не является активным). В этом случае мы решаем задачу так, как будто у нее нет ограничений. Очевидно, что решением будет точка, в которой расположена вышка, так как в ней уровень сигнала максимальный. Проверим это:

$$\mu_1^{(1)} = 0, \quad \mu_2^{(1)} = 0, \quad x_1^{(1)} = \frac{6}{1+0} = 6, \quad x_2^{(1)} = \frac{8}{1+0} = 8.$$

Также очевидно, что эта точка не является решением, так как не удовлетворяет ограничениям задачи.

1.2. Полагаем 
$$4 - \left( \frac{6}{1 + \mu_1} \right)^2 - \left( \frac{8}{1 + \mu_1} \right)^2 = 0, \quad \text{откуда}$$

$$\frac{4(1 + \mu_1)^2 - 36 - 48}{(1 + \mu_1)^2} = 0. \quad \text{По условию } \mu_1 \geq 0, \text{ знаменатель не может}$$

быть равен нулю и его можно сократить:

$$4(1 + \mu_1)^2 - 36 - 48 = 0 \Rightarrow 1 + \mu_1 = \pm 5.$$

Так как  $\mu_1 \geq 0$ , выбираем решение  $\mu_1 = 4$ , тогда

$$\mu_1^{(2)} = 4, \mu_2^{(2)} = 0, x_1^{(2)} = \frac{6}{1+4} = 1,2, x_2^{(2)} = \frac{8}{1+4} = 1,6.$$

Данная точка также не удовлетворяет ограничениям задачи. Отсюда можно сделать вывод, что второе ограничение в точке оптимума должно быть активным.

2. Пусть  $x_1^2 - x_2 = 0$ . Тогда  $x_2 = x_1^2$ , получаем систему уравнений

$$2x_1 - 12 + 2x_1\mu_1 - 2x_1\mu_2 = 0;$$

$$2x_1^2 - 16 + 2x_1^2\mu_1 + \mu_2 = 0;$$

$$\mu_1(4 - x_1^2 - x_1^4) = 0;$$

$$\mu_{1,2} \geq 0.$$

Из третьего уравнения следуют два варианта решения: либо  $\mu_1 = 0$  (т.е. первое ограничение не является активным, но при этом, в отличие от предыдущего случая, второе ограничение активно), либо  $4 - x_1^2 - x_1^4 = 0$ .

2.1. Полагаем  $\mu_1 = 0$ , тогда

$$2x_1(1 - \mu_2) - 12 = 0;$$

$$2x_1^2 - 16 + \mu_2 = 0;$$

$$\mu_2 \geq 0.$$

Из второго уравнения  $\mu_2 = -2x_1^2 + 16$ , подставляем в первое уравнение:  $2x_1(1 + 2x_1^2 - 16) - 12 = 4x_1^3 - 30x_1 - 12 = 0$ . Решая данное уравнение численно, находим лишь один положительный корень  $x_1^{(3)} \approx 2,92$ . Тогда  $x_2^{(3)} \approx 2,92^2 \approx 8,53$ . При этом  $\mu_2^{(3)} \approx -2 \cdot 8,53 + 16 \approx -1,05$ . Полученное решение

$$\mu_1^{(3)} = 0, \mu_2^{(3)} \approx -1,05, x_1^{(3)} \approx 2,92, x_2^{(3)} \approx 8,53$$

является недопустимым, так как  $\mu_2 < 0$ . Отсюда можно сделать вывод, что первое ограничение в точке оптимума также должно быть активным.

2.2. Полагаем  $4 - x_1^2 - x_1^4 = 0$ , т.е. получили биквадратное уравнение  $4 - z - z^2 = 0$ , где  $z = x_1^2$ . Соответственно  $z \geq 0$ . У данного уравнения только один положительный корень, который мы находили ранее, —  $z \approx 1,562$ . Соответственно  $x_1^{(4)} = \sqrt{z} \approx 1,25$ ,  $x_2^{(4)} = z \approx 1,562$ .

Найдем  $\mu_1$  и  $\mu_2$ :

$$2 \cdot 1,25 - 12 + 2\mu_1 \cdot 1,25 - 2\mu_2 \cdot 1,25 = 0;$$

$$2 \cdot 1,562 - 16 + 2\mu_1 \cdot 1,562 + \mu_2 = 0;$$

$$\mu_{1,2} \geq 0.$$

Решением данной СЛАУ является  $\mu_1^{(4)} \approx 4,045$ ,  $\mu_2^{(4)} \approx 0,244$ . Так как оба параметра положительны, а точка  $x^{(4)}$  удовлетворяет ограничениям задачи, то решение  $(x^{(4)}, \mu^{(4)})$  потенциально является оптимальным. Но условия оптимальности не выполняются, из-за чего утверждать это пока нельзя — это может быть и точка максимума, и седловая точка. В таких случаях следует использовать условия оптимальности второго порядка.

### 6.4.3. Необходимые и достаточные условия оптимальности второго порядка

Мы рассмотрели необходимые и достаточные условия первого порядка — это условия Куна — Таккера. Они применимы к задачам условной оптимизации и формулируются через градиенты ЦФ и функций, входящих в ограничения. При этом на функции ограничения  $g_i(x)$  накладываются жесткие условия выпуклости (или вогнутости), а ограничения-равенства  $h_l(x)$  должны быть линейными, что выполняется не для всех ЗНП. Необходимые и достаточные условия оптимальности второго порядка для задач с дважды дифференцируемыми функциями построены Мак-Кормиком.

Рассмотрим следующую задачу нелинейного программирования:

$$\begin{aligned} f(x) &\rightarrow \min, \quad x \in R^n; \\ g_i(x) &\geq 0, \quad i = 1, 2, \dots, m; \\ h_l(x) &= 0, \quad l = 1, 2, \dots, p. \end{aligned} \quad (6.40)$$

Пусть  $f(x)$ ,  $g_i(x)$ ,  $h_l(x)$  – дважды дифференцируемые функции,  $x^*$  – допустимое решение (6.40).

Множество активных в точке  $x^*$  ограничений определяется множеством  $I = \{i \mid g_i(x^*) = 0\}$ .

Предположим, что  $\nabla g_i(x^*)$  при  $i \in I$  и  $\nabla h_l(x^*)$  при  $l = 1, 2, \dots, p$  линейно независимы.

*Теорема.* Для того чтобы точка  $x^*$  была точкой строгого локального минимума, необходимо и достаточно выполнение следующих условий:

1) должен существовать такой вектор  $(\lambda^*, \mu^*)$ , что вектор  $(x^*, \mu^*, \lambda^*)$  будет определять точку Куна – Таккера;

2) для любого ненулевого вектора  $y \in R^n$ , удовлетворяющего системе уравнений и неравенств

$$\nabla g_i(x^*)^T y = 0, \quad i \in I_1 = \{i \mid g_i(x^*) = 0, \mu_i^* > 0\}, \quad (6.41)$$

$$\nabla g_i(x^*)^T y \geq 0, \quad i \in I_2 = \{i \mid g_i(x^*) = 0, \mu_i^* = 0\}, \quad (6.42)$$

$$\nabla h_l(x^*)^T y = 0, \quad l = 1, 2, \dots, p, \quad (6.43)$$

должно выполняться неравенство

$$y^T H_L(x^*, \lambda^*, \mu^*) y > 0, \quad (6.44)$$

где  $H_L$  – матрица Гессе функции  $L(x, \lambda, \mu)$ :

$$H_L(x, \lambda, \mu) = \left\| \frac{\partial^2 L}{\partial x_i \partial x_j} \right\|_{\substack{x=x^* \\ \lambda=\lambda^* \\ \mu=\mu^*}},$$

$$L(x, \lambda, \mu) = f(x) - \sum_{i=1}^m \mu_i g_i(x) + \sum_{l=1}^p \lambda_l h_l(x).$$

*Примечание.* Множество  $I = I_1 \cup I_2$  содержит номера всех активных ограничений в точке  $x^*$ .

*Пример.* Рассмотрим задачу определения наилучшего положения приемной антенны с ограничениями-неравенствами:

$$f(x) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min,$$

$$g_1(x) = 4 - x_1^2 - x_2^2 \geq 0,$$

$$g_2(x) = x_1^2 - x_2 \geq 0.$$

Выше из условий Куна – Таккера была получена подозрительная на экстремум точка  $\bar{x}_1 \approx 1,25$ ,  $\bar{x}_2 \approx 1,562$ ,  $\bar{\mu}_1 \approx 4,045$ ,  $\bar{\mu}_2 \approx 0,244$ .

Оба ограничения являются активными, т.е.  $g_1(\bar{x}) = 0$  и  $g_2(\bar{x}) = 0$ , поэтому  $I = \{1, 2\}$ . Найдем градиенты ограничений в точке  $\bar{x}$ :

$$\nabla g_1(\bar{x}) = \begin{pmatrix} -2\bar{x}_1 \\ -2\bar{x}_2 \end{pmatrix} \approx \begin{pmatrix} -2,499 \\ -3,123 \end{pmatrix}, \quad \nabla g_2(\bar{x}) = \begin{pmatrix} 2\bar{x}_1 \\ -1 \end{pmatrix} \approx \begin{pmatrix} 2,499 \\ -1 \end{pmatrix}.$$

Ограничений-равенств в задаче нет, поэтому проверять линейную независимость векторов  $\nabla g_i(\bar{x})$  и  $\nabla h_l(\bar{x})$  не требуется. Равенство  $\nabla g_1(\bar{x})^T y = 0$  будет выполняться для любого вектора

$$y = \begin{pmatrix} y_1 \\ -\frac{2,499}{3,192} y_1 \end{pmatrix},$$

а равенство  $\nabla g_2(\bar{x})^T y = 0$  – для любого вектора

$$y = \begin{pmatrix} y_1 \\ 2,499y_1 \end{pmatrix}.$$

Запишем функцию  $L(x, \mu)$  (параметр  $\lambda$  не требуется, так как ограничения-равенства отсутствуют) и ее гессиан:

$$\begin{aligned} L(x, \mu) &= f(x) - \mu_1 g_1(x) - \mu_2 g_2(x) = \\ &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 - \mu_1 (4 - x_1^2 - x_2^2) - \mu_2 (x_1^2 - x_2), \end{aligned}$$

$$H_L(x, \mu) = \begin{pmatrix} 2 + 2\mu_1 - 2\mu_2 & 0 \\ 0 & 2 + 2\mu_1 \end{pmatrix}.$$

Найдем гессиан в точке  $(\bar{x}, \bar{\mu})$ :

$$H_L(\bar{x}, \bar{\mu}) \approx \begin{pmatrix} 2 + 2 \cdot 4,045 - 2 \cdot 0,244 & 0 \\ 0 & 2 + 2 \cdot 4,045 \end{pmatrix} \approx \begin{pmatrix} 9,603 & 0 \\ 0 & 10,09 \end{pmatrix}.$$

Матрица  $H_L$  положительно определена в любой точке  $(\bar{x}, \bar{\mu})$ , поэтому для любого вектора  $y \neq 0$  имеем  $y^T H_L y > 0$ , в том числе и для найденных выше векторов. Таким образом, точка  $(x^*, \mu^*) = (1,25; 1,562; 4,045; 0,244)^T$  является точкой строгого локального минимума.

Ответ:

$$x^* = \begin{pmatrix} 1,25 \\ 1,562 \end{pmatrix}, \quad f^* \approx 65,02.$$

Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 8,063$  км.  
Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 15,38 \text{ Вт/м}^2.$$

Еще раз обратим внимание, что в точке  $x^*$  оба ограничения являются активными. Это означает, что точка оптимума лежит на поверхности обоих ограничений, т.е. находится на пересечении береговой

линии озера и линии окружности радиусом  $L = 2$  км с центром в населенном пункте. Следовательно, если в данной задаче ограничения заменить равенствами, ее решение не изменится. Но это не означает, что решение любой задачи с ограничениями-равенствами будет эквивалентно ее решению с ограничениями-неравенствами. Например, если переместить телевизионную вышку в точку  $(6,0)$ , то решение задачи с ограничениями-равенствами не изменится, так как оптимальная точка обязана будет находиться на пересечении ограничений. А в задаче с ограничениями-неравенствами оптимальной станет точка  $x^* = (2,0)^T$ , так как именно она в ОДР будет ближе всего к вышке.

## 6.5. Методы штрафов

### 6.5.1. Понятие штрафных функций

Продолжим изучение ЗО с ограничениями. Нас будут интересовать методы, существенно отличающиеся от рассмотренных выше. Их общий принцип заключается в замене исходной задачи на решение последовательности экстремальных задач без ограничений.

Рассмотрим задачу

$$f(x) \rightarrow \min, \quad x \in R^n, \quad (6.45)$$

$$\begin{aligned} g_i(x) &\geq 0, \quad i = 1, 2, \dots, m; \\ h_l(x) &= 0, \quad l = 1, 2, \dots, p. \end{aligned} \quad (6.46)$$

Предполагается, что точка  $x^*$  является решением этой задачи, известно некоторое начальное приближение  $x^0$ , возможно, недопустимое, т.е. не удовлетворяющее отношениям (6.46). С помощью рассматриваемых далее алгоритмов в пространстве  $R^n$  строится конечная последовательность точек  $x^t, t = 0, 1, \dots, k$ , которая начинается с заданной точки  $x^0$  и заканчивается точкой  $x^k$ , дающей наилучшее приближение к точке экстремума  $x^*$  среди всех точек построенной последовательности. В качестве точек последовательности  $\{x^t\}$  берутся

стационарные точки *штрафной функции* (ШФ) – ЦФ вспомогательной задачи безусловной минимизации.

Конкретные методы, основанные на указанной общей схеме, определяются видом штрафной функции, а также правилами, по которым производится пересчет штрафных параметров по окончании очередного цикла безусловной минимизации. Идея преобразования задачи с ограничениями в последовательность задач без ограничений, построенную надлежащим образом, представляется заманчивой главным образом в связи с наличием эффективных и надежных методов безусловной минимизации. При этом будем искать условный минимум с приемлемой точностью путем решения относительно небольшого числа не слишком сложных задач.

В зависимости от того, являются ли элементы последовательности  $\{x^t\}$  допустимыми или недопустимыми точками, говорят соответственно о методах *внутренней* (рисунок 6.5) или *внешней* (рисунок 6.6) *точки*. Иногда их называют *методами внутреннего* или *внешнего штрафа*. Методы внутренних штрафных функций называют также *методами барьерных функций*. На рисунках 6.5 и 6.6 точка безусловного минимума обозначена как  $x'$ .

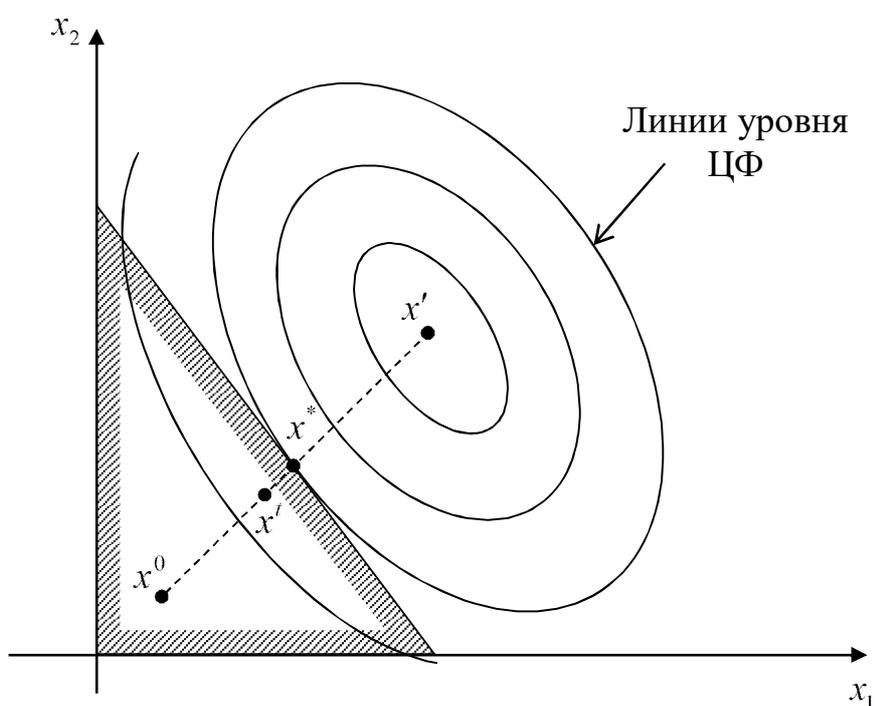


Рисунок 6.5 – Метод внутренних штрафов

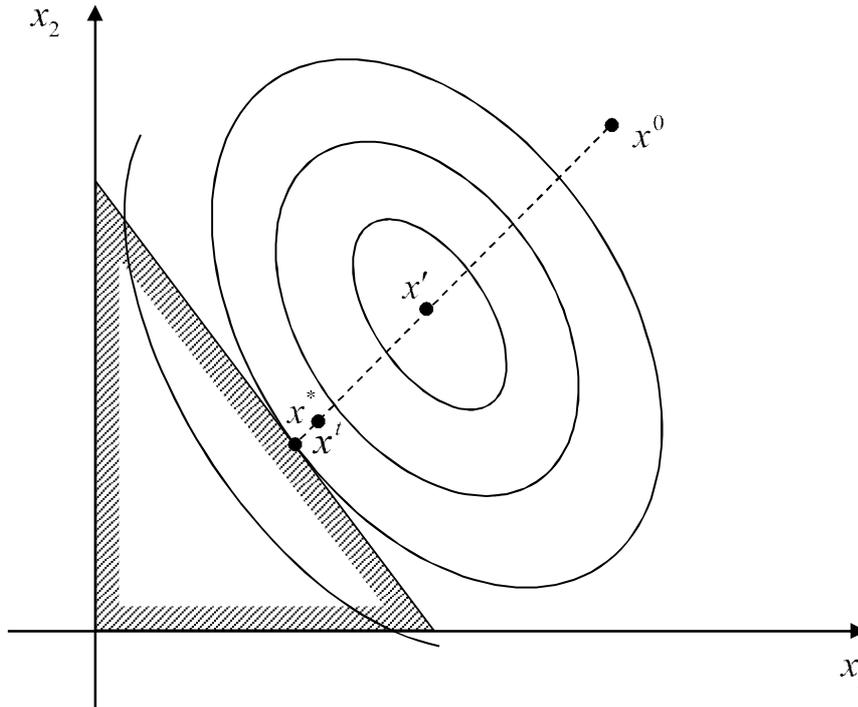


Рисунок 6.6 – Метод внешних штрафов

Штрафная функция определяется выражением

$$P(x, R) = f(x) + H[R, g(x), h(x)], \quad (6.47)$$

где  $R$  – набор штрафных параметров;  $H$  – штраф-функция  $R$  и ограничений,

$$H[R, g(x), h(x)] = R\varphi[g(x), h(x)].$$

Выражение для штрафа  $H$  определяется таким образом, чтобы допустимые точки задачи имели преимущество перед недопустимыми в отношении безусловной минимизации ШФ.

Методы внутренней точки связаны с такими функциями  $H$ , при которых стационарные точки функции  $P(x, R)$  оказываются заведомо допустимыми. Эти методы называют также методами барьеров, поскольку здесь штраф как бы создает вдоль границы допустимой области  $S$  барьер из бесконечно больших значений функции  $P$ .

*Методы штрафных функций* (МШФ) представляют большой интерес лишь при выполнении следующих требований:

1) решения подзадач должны стремиться к решению исходной задачи нелинейного программирования вида

$$\lim x^t \rightarrow x^*, \quad t \rightarrow k < \infty;$$

2) сложность минимизации  $P(x, R)$  должна быть того же порядка, что и для функции  $f(x)$ ;

3) правило пересчета  $R^{t+1} = F(R^t)$  должно быть простым.

Рассмотрим широко используемые типы штрафов и различные процедуры учета ограничений при переходе к задачам безусловной минимизации.

### 6.5.2. Квадратичный штраф

*Квадратичный штраф* (рисунок 6.7) используется для учета ограничений-равенств и имеет вид

$$H = R \cdot [h(x)]^2. \quad (6.48)$$

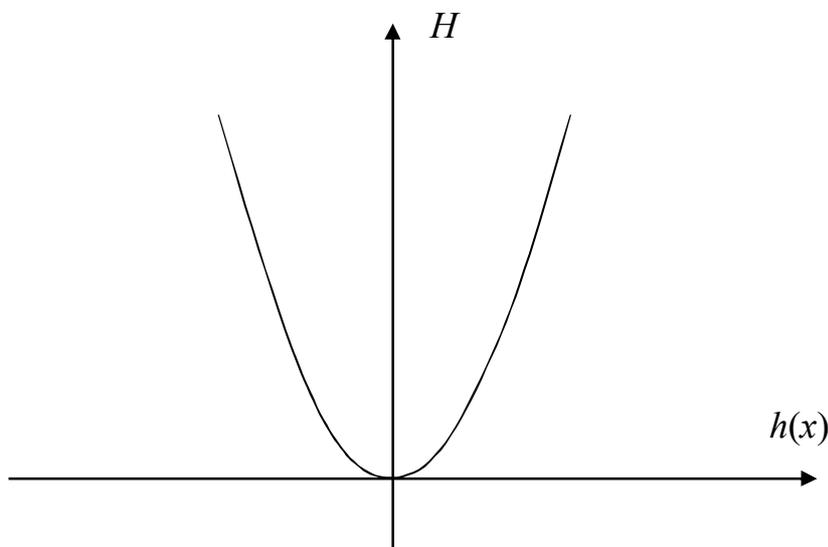


Рисунок 6.7 – Квадратичный штраф

При минимизации этот штраф препятствует отклонению величины  $h(x)$  от нуля (как вправо, так и влево). При увеличении  $R$

стационарная точка соответствующей ШФ  $P(x, R)$  приближается к точке  $x^*$ , так как в пределе  $h(x^k) = 0$ . Поскольку допустимая область  $S$  определяется ограничением  $h(x) = 0$ , то в допустимой области значение штрафа  $H = 0$ , а вне допустимой области  $H > 0$  и тем больше, чем дальше точка  $x^t$  выйдет за пределы допустимой области и чем больше  $R$ .

*Пример.* Решим задачу определения наилучшего положения приемной антенны с одним ограничением (6.15), преобразовав его в равенство:

$$\begin{aligned} f(x) &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ h(x) &= 4 - x_1 - x_2 = 0. \end{aligned}$$

Используем квадратичный штраф. Введем ШФ вида

$$P(x, R) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 + R(4 - x_1 - x_2)^2.$$

Запишем уравнения, определяющие стационарную точку функции  $P(x, R)$ :

$$\begin{aligned} \frac{\partial P}{\partial x_1} &= 2(x_1 - 6) - 2R(4 - x_1 - x_2) = 0; \\ \frac{\partial P}{\partial x_2} &= 2(x_2 - 8) - 2R(4 - x_1 - x_2) = 0. \end{aligned}$$

Выразим из первого уравнения

$$2R(4 - x_1 - x_2) = 2(x_1 - 6)$$

и подставим во второе уравнение:

$$2(x_2 - 8) - 2(x_1 - 6) = 2x_2 - 2x_1 - 4 = 0 \Rightarrow x_2 = x_1 + 2.$$

Тогда

$$\begin{aligned} 2R(4 - x_1 - (x_1 + 2)) &= 2(x_1 - 6) \Rightarrow R(2 - 2x_1) = x_1 - 6 \Rightarrow \\ 2Rx_1 + x_1 - 2R - 6 &= 0 \Rightarrow x_1 = \frac{2R + 6}{2R + 1}. \end{aligned}$$

Переходя к пределу, получим

$$x_1 = \lim_{R \rightarrow \infty} \frac{2R + 6}{2R + 1} = 1.$$

Соответственно  $x_2 = x_1 + 2 = 3$ . Таким образом, метод сходится к точке  $x^* = (1; 3)^T$  и  $f^* = 51$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 7,141$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 19,61 \text{ Вт/м}^2.$$

Мы нашли решение задачи аналитически, как и во всех предыдущих случаях. Однако методы штрафов позволяют решать ЗНП численно. Рассмотрим, как это происходит. В таблице 6.1 приведены координаты стационарных точек ШФ  $P(x, R)$  для различных значений параметра  $R$ .

На рисунке 6.8 изображена ЦФ и ограничение задачи. В точке  $x'$  находится безусловный минимум задачи (без учета ограничения), в точке  $x^*$  – минимум ШФ и оптимальное решение задачи.

Таблица 6.1 – Стационарные точки квадратичного штрафа

$R$	$x_1^*; x_2^*$	$P(x, R)$	$f(x)$	$h(x)$
0	6; 8	1	1	-10
0,1	5,167; 7,167	9,333	2,389	-8,333
1	2,667; 4,667	34,33	23,22	-3,333
10	1,238; 3,238	48,62	46,35	-0,476
100	1,025; 3,025	50,75	50,50	-0,050
$\infty$	1; 3	51	51	0

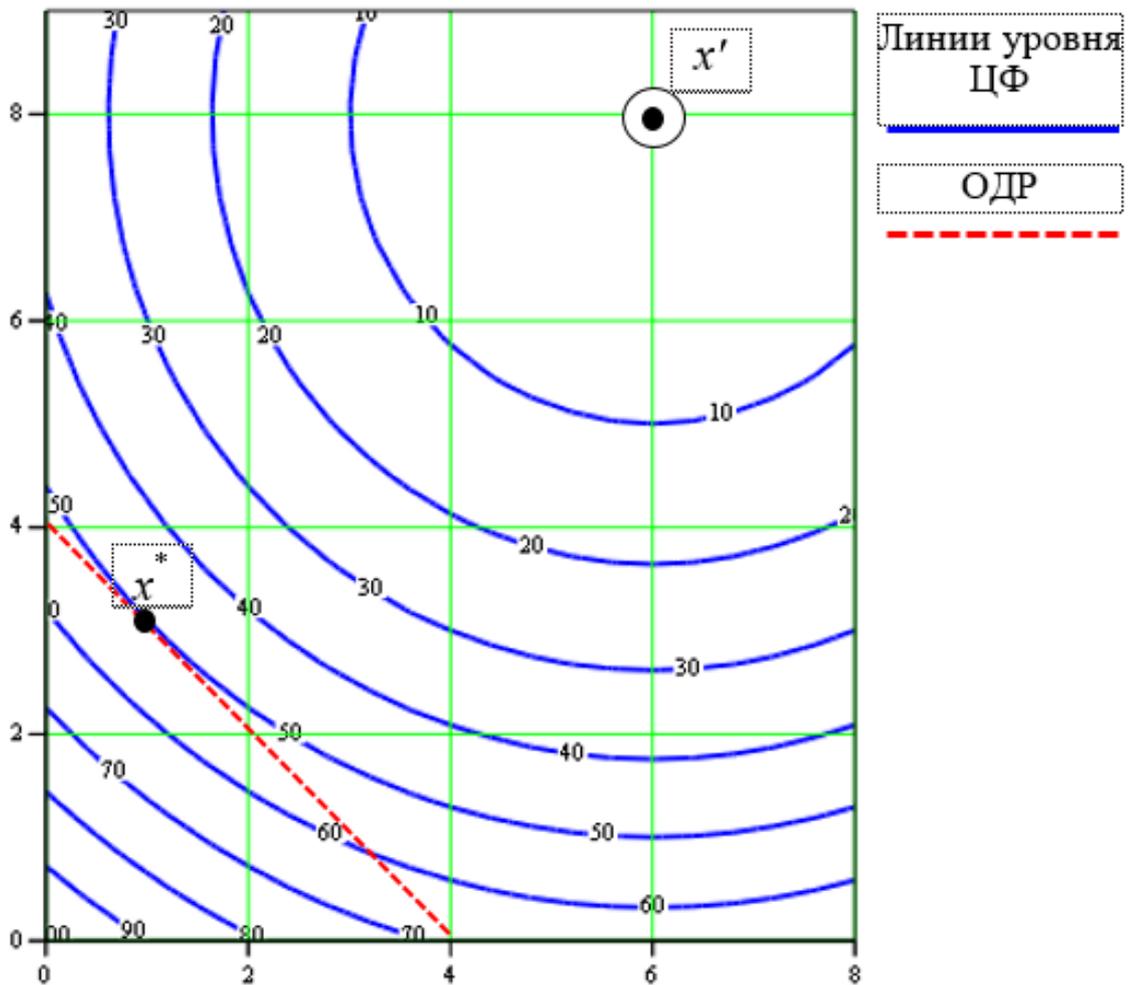


Рисунок 6.8 – Линии уровня ЦФ

Линии уровня функции  $P(x, R)$  при  $R \rightarrow 0$  сближаются с линиями уровня функции  $f(x)$ , а при  $R = 0$  совпадают с ними. При  $R = 1$  штрафная функция имеет вид (рисунок 6.9)

$$P(x, R) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 + (4 - x_1 - x_2)^2.$$

Минимум ШФ находится в точке  $(2,667; 4,667)^T$ .

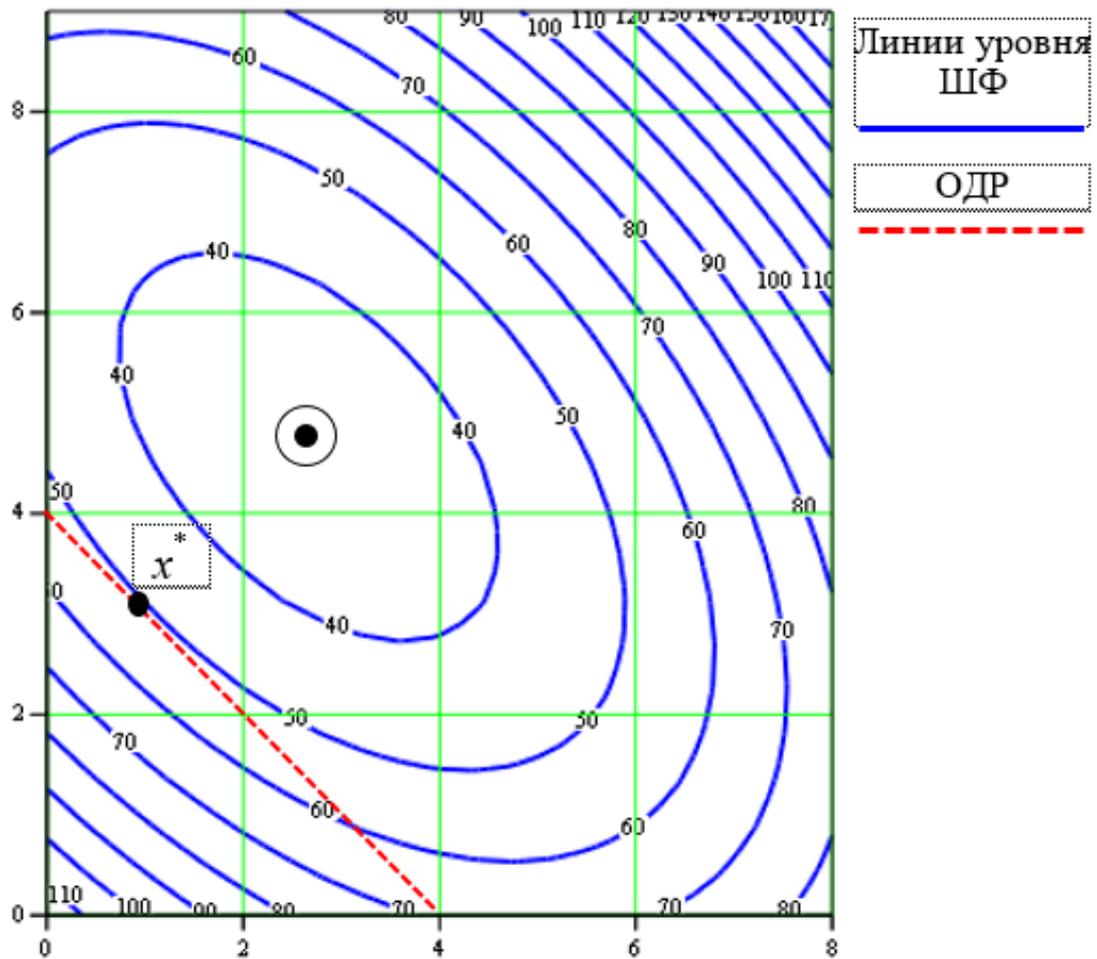


Рисунок 6.9 – Линии уровня штрафной функции для  $R = 1$

При  $R = 10$  штрафная функция имеет вид (рисунок 6.10)

$$P(x, R) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 + 10(4 - x_1 - x_2)^2.$$

Минимум ШФ находится в точке  $(1,238; 3,238)^T$ .

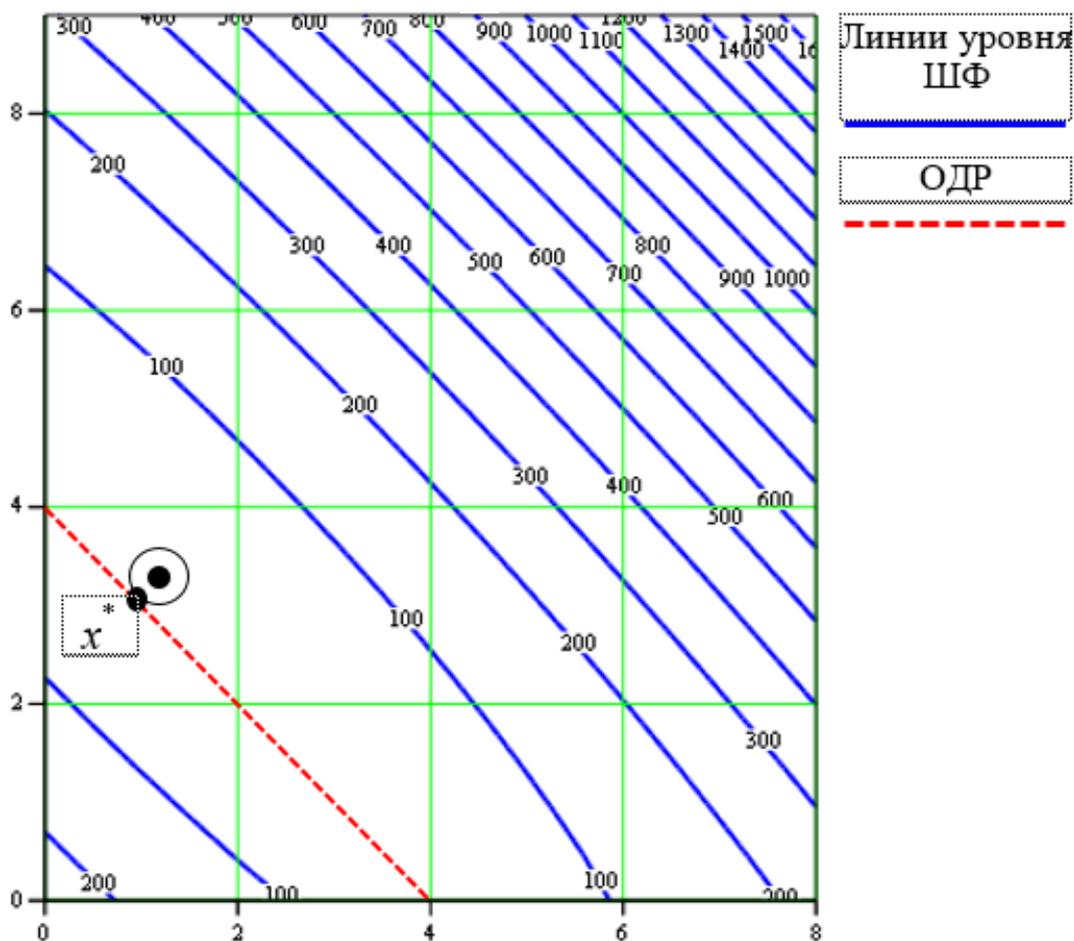


Рисунок 6.10 – Линии уровня штрафной функции для  $R = 10$

При  $R = 100$  штрафная функция имеет вид (рисунок 6.11)

$$P(x, R) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 + 100(4 - x_1 - x_2)^2.$$

Минимум ШФ находится в точке  $(1,025; 3,025)^T$ , что уже очень близко к оптимальному значению  $x^* = (1; 3)^T$ .

Обсудим данный пример подробнее. Рассмотрим стационарную точку функции  $P(x, 10)$ . Эта точка уже довольно близка к  $x^*$ . Во многих прикладных задачах такое приближенное решение может считаться вполне удовлетворительным, однако в общем случае величина невязки  $h(x)$ , как правило, неудовлетворительна. Возникает вопрос, как изменить ШФ, чтобы ее стационарная точка приблизилась к  $x^*$ ?

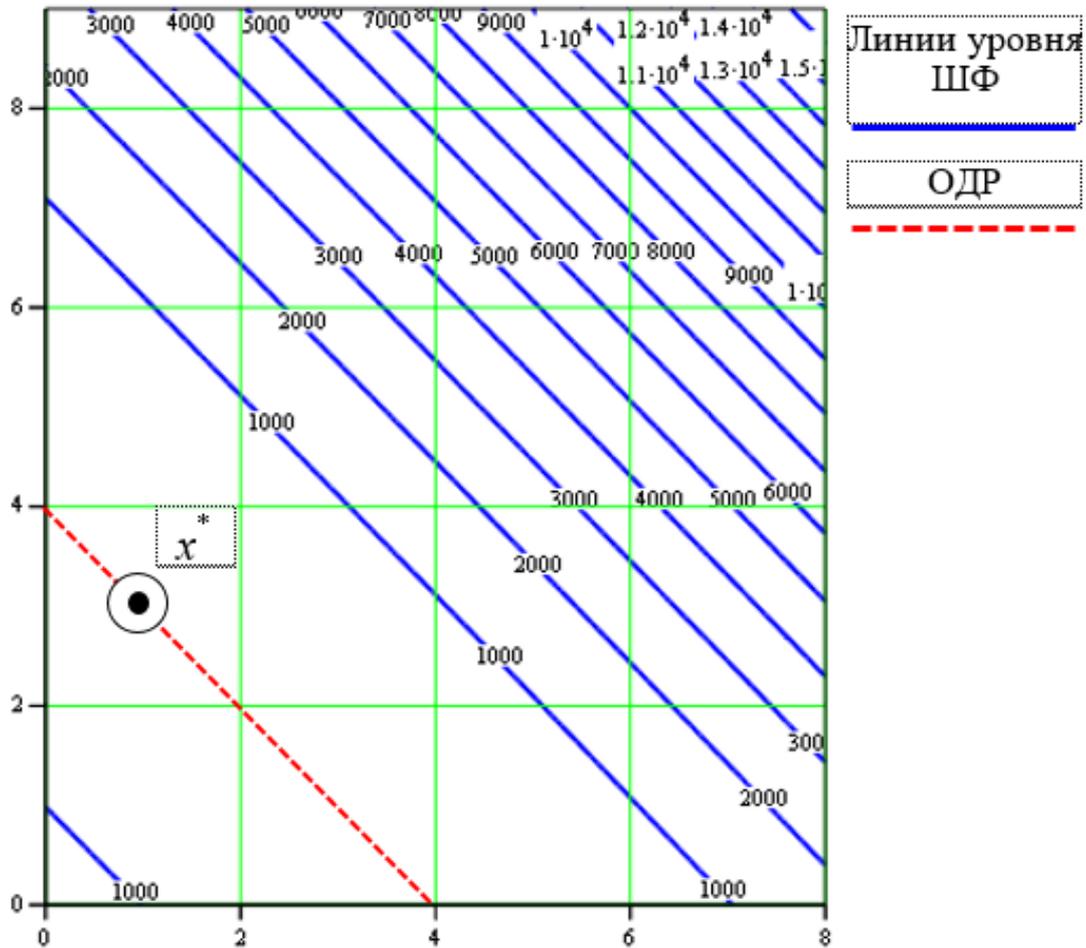


Рисунок 6.11 – Линии уровня штрафной функции для  $R = 100$

Заметим, что в точно вычисленной стационарной точке функции  $P(x, 10)$  значение ЦФ  $f(x^t) < P(x^t, 10)$  (оптимального) из-за нарушения условия  $h(x^t) = 0$ . При этом соотношение между значениями  $f(x)$  и  $P(x, R)$  связано с величиной  $R$ , в данном случае  $R = 10$ . По существу,  $R$  играет роль весового коэффициента, определяющего относительную значимость ограничения и ЦФ. Ясно, что при увеличении  $R$  стационарная точка функции  $P(x, R)$  будет смещаться в сторону более точного выполнения ограничения  $h(x) = 0$  и тем самым приближаться к  $x^*$ .

Если бы можно было положить  $R = \infty$ , то точное решение соответствующей подзадачи на безусловный минимум совпало бы с решением исходной задачи. Но при этом линии уровня ШФ становятся слишком вытянутыми – получаем *овражную функцию* и сложность безусловной минимизации резко возрастает.

Предпочтительнее стратегия многоэтапного последовательного приближения к решению:

1) процесс начинают с относительно малого  $R$ , например  $R = 0,01$ , и находят точку минимума функции  $P(x, 0,01)$ ;

2) увеличивают  $R$ , например на порядок ( $R = 0,1$ ), и находят минимум  $P(x, 0,1)$ , используя ранее полученную точку в качестве начальной;

3) параметр  $R$  увеличивают до тех пор, пока элементы итерационных последовательностей  $x^t$ ,  $f(x^t)$ ,  $P(x^t, R)$  не станут изменяться от шага к шагу достаточно мало. Однако при этом увеличивается число итераций и время решения задачи. В этом основная трудность МШФ.

### 6.5.3. Логарифмический штраф

Рассмотрим несколько типов штрафов для ограничений-неравенств. *Логарифмический штраф* (рисунок 6.12) имеет вид

$$H = -R \cdot \ln[g(x)].$$

Штраф положителен при всех  $x$  таких, что  $0 < g(x) < 1$ , и отрицателен при всех  $x$  таких, что  $g(x) > 1$ . В данном случае вводится как бы искусственная дискриминация точек допустимой области: внутренним точкам отдается предпочтение перед граничными точками. Отрицательных значений штрафа можно избежать, положив  $H = 0$  для таких  $x$ , что  $g(x) > 1$ . Однако при этом у градиента  $\nabla P$  появляются разрывы.

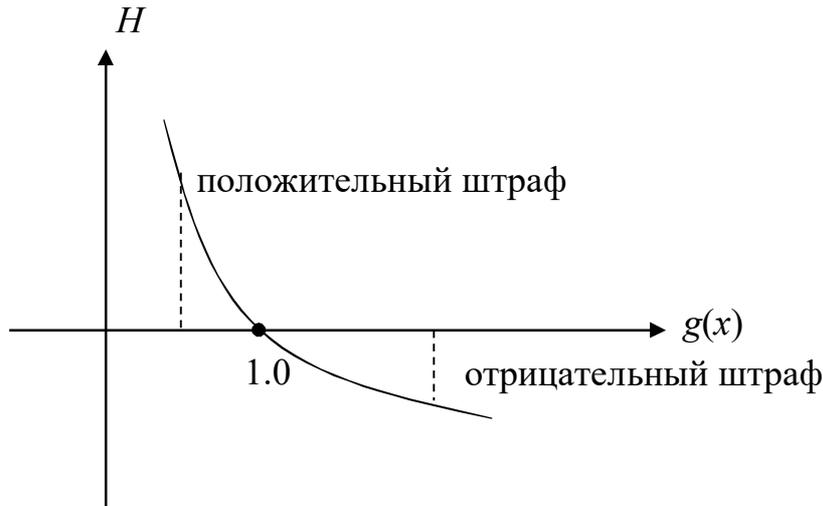


Рисунок 6.12 – Логарифмический штраф

Логарифмический штраф – это барьерная функция, неопределенная в недопустимых точках (т.е. для таких точек  $x$ , в которых  $g(x) < 0$ ). Поэтому в тех случаях, когда приходится иметь дело с недопустимыми точками (например, когда точка  $x^0 \notin S$ ), требуется процедура специального типа, обеспечивающая попадание в допустимую область  $S$ .

Итерационный процесс желательно начинать из допустимой области точки  $x^0$  при положительном начальном значении  $R$  (например,  $R = 10$  или  $R = 100$ ). После решения каждой подзадачи безусловной минимизации параметр  $R$  уменьшается и в пределе стремится к нулю.

*Пример.* Решим задачу определения наилучшего положения приемной антенны с одним ограничением (6.15):

$$f(x) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min,$$

$$g(x) = 4 - x_1 - x_2 \geq 0,$$

т.е. ограничение-равенство в предыдущем примере заменим неравенством. При этом решение  $x^*$  задачи должно остаться прежним, так как оптимальное решение лежит на границе ОДР. В данном случае

$$P(x, R) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 - R \cdot \ln(4 - x_1 - x_2).$$

Условия стационарности:

$$\begin{cases} \frac{\partial P}{\partial x_1} = 2(x_1 - 6) + R \cdot \left[ \frac{1}{4 - x_1 - x_2} \right] = 0, \\ \frac{\partial P}{\partial x_2} = 2(x_2 - 8) + R \cdot \left[ \frac{1}{4 - x_1 - x_2} \right] = 0. \end{cases}$$

Выразим из первого уравнения

$$R \cdot \left[ \frac{1}{4 - x_1 - x_2} \right] = -2(x_1 - 6)$$

и подставим во второе уравнение:

$$2(x_2 - 8) - 2(x_1 - 6) = 2x_2 - 2x_1 - 4 = 0 \Rightarrow x_2 = x_1 + 2.$$

Тогда

$$R \cdot \left[ \frac{1}{4 - x_1 - (x_1 + 2)} \right] = -2(x_1 - 6) \Rightarrow \frac{R}{2 - 2x_1} = -2(x_1 - 6) \Rightarrow$$

$$R = -2(x_1 - 6)(2 - 2x_1) \Rightarrow \frac{R}{4} = (x_1 - 6)(x_1 - 1) = 0 \Rightarrow$$

$$x_1^2 - 7x_1 + 6 - \frac{R}{4} = 0 \Rightarrow x_1^{(1,2)} = \frac{7 \pm \sqrt{25 + R}}{2}.$$

Переходя к пределу, получим

$$x_1^{(1,2)} = \lim_{R \rightarrow 0} \frac{7 \pm \sqrt{25 + R}}{2} = \frac{7 \pm 5}{2} \Rightarrow x_1^{(1)} = 1, x_1^{(2)} = 6.$$

Соответственно  $x_2^{(1)} = x_1^{(1)} + 2 = 3$ ,  $x_2^{(2)} = x_1^{(2)} + 2 = 8$ . В точке  $x^{(2)}$  ограничение задачи нарушается, поэтому оптимальной является точка  $x^{(1)}$ . Таким образом, метод сходится к точке  $x^* = (1; 3)^T$  и  $f^* = 51$ .

Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 7,141$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 19,61 \text{ Вт/м}^2.$$

Метод оказывается сходящимся в предположении, что безусловная минимизация проводится точно.

В таблице 6.2 приведены координаты стационарных точек логарифмического штрафа.

Таблица 6.2 – Стационарные точки логарифмического штрафа

$R$	$x_1^*; x_2^*$	$P(x, R)$	$f(x)$	$g(x)$
100	-2,090; -0,090	-50,24	131,9	6,180
10	0,542; 2,542	61,46	60,58	0,916
1	0,950; 2,950	54,31	52,00	0,099
0,1	0,995; 2,995	51,56	51,10	$9,990 \times 10^{-3}$
0,01	0,9995; 2,9995	51,08	51,01	$9,999 \times 10^{-4}$
0,001	1,000; 3,000	51,01	51,001	$9,999 \times 10^{-5}$
0	1; 3	–	51	0

#### 6.5.4. Штраф типа обратной функции

*Штраф типа обратной функции* имеет вид

$$H = R \frac{1}{g(x)}.$$

Здесь  $H$ , в отличие от логарифмического штрафа, не имеет отрицательных значений в допустимой области (рисунок 6.13). Данный штраф называют также барьерной функцией.

Очевидно, что необходимо предотвращение появления недопустимых точек, в которых  $H < 0$ . В точках  $x \in S$  вблизи границы значения штрафа положительны и быстро убывают при удалении от границы внутрь допустимой области. На самой границе  $P(x, R)$  и  $\nabla P$  не определены.

Итерации начинаются с точки  $x^0$  – начальной допустимой точки при  $R > 0$ , причем в процессе вычисления  $R \rightarrow 0$ .

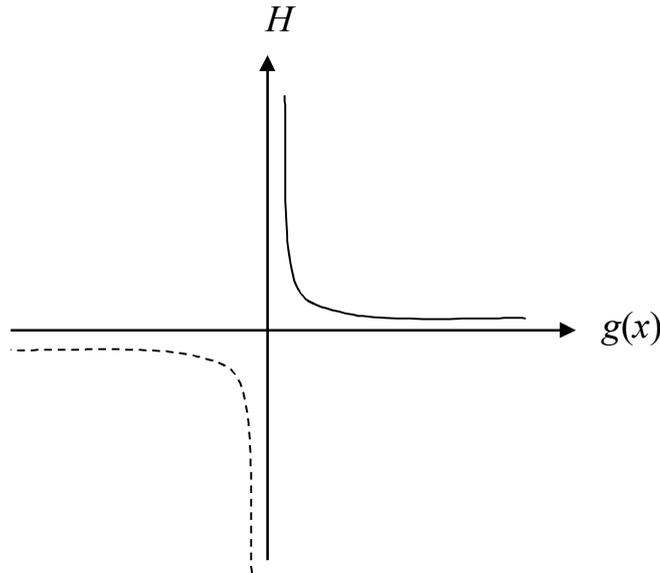


Рисунок 6.13 – Штраф типа обратной функции

*Пример.* Решим задачу определения наилучшего положения приемной антенны с одним ограничением (6.15):

$$f(x) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min,$$

$$g(x) = 4 - x_1 - x_2 \geq 0.$$

Для нашей задачи имеем

$$P(x, R) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 + \frac{R}{4 - x_1 - x_2}.$$

Условия стационарности:

$$\begin{cases} \frac{\partial P}{\partial x_1} = 2(x_1 - 6) + \frac{R}{(4 - x_1 - x_2)^2}, \\ \frac{\partial P}{\partial x_2} = 2(x_2 - 8) + \frac{R}{(4 - x_1 - x_2)^2}. \end{cases}$$

Выразим из первого уравнения

$$\frac{R}{(4 - x_1 - x_2)^2} = -2(x_1 - 6)$$

и подставим во второе уравнение:

$$2(x_2 - 8) - 2(x_1 - 6) = 2x_2 - 2x_1 - 4 = 0 \Rightarrow x_2 = x_1 + 2.$$

Тогда

$$\frac{R}{(4 - x_1 - (x_1 + 2))^2} = -2(x_1 - 6) \Rightarrow R = -2(x_1 - 6)(2 - 2x_1)^2 \Rightarrow$$

$$x_1^3 - 8x_1^2 + 13x_1 - 6 + \frac{R}{8} = 0.$$

Получили кубическое уравнение. Аналитически его решать затруднительно, особенно при наличии неизвестного параметра  $R$ , поэтому приведем только численное решение задачи.

Исследуем стационарные точки функции  $P(x, R)$  при различных  $R$  (таблица 6.3). Кубическое уравнение имеет три корня, будем выбирать один из них, принадлежащий ОДР.

Таблица 6.3 – Стационарные точки штрафа типа обратной функции

$R$	$x_1^*; x_2^*$	$P(x, R)$	$f(x)$	$g(x)$
100	-0,398; 1,602	118,6	82,86	2,796
10	0,522; 2,522	165,7	61,01	0,955
1	0,844; 2,844	375,3	54,16	0,311
0,1	0,950; 2,950	1057	52,00	0,100
0,01	0,984; 2,984	3127	51,32	0,032
0,001	0,995; 2,995	10055	51,10	$9,996 \times 10^{-3}$
0	1; 3	$\infty$	51	0

Согласно таблице метод сходится к точке  $x^* = (1; 3)^T$  и  $f^* = 51$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 7,141$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 19,61 \text{ Вт/м}^2.$$

### 6.5.5. Штраф типа квадрата срезки

Штраф типа квадрата срезки имеет вид (рисунок 6.14)

$$H = R \cdot \langle g(x) \rangle^2,$$
$$\langle g(x) \rangle = \begin{cases} g(x), & \text{если } g(x) \leq 0, \\ 0, & \text{если } g(x) > 0. \end{cases}$$

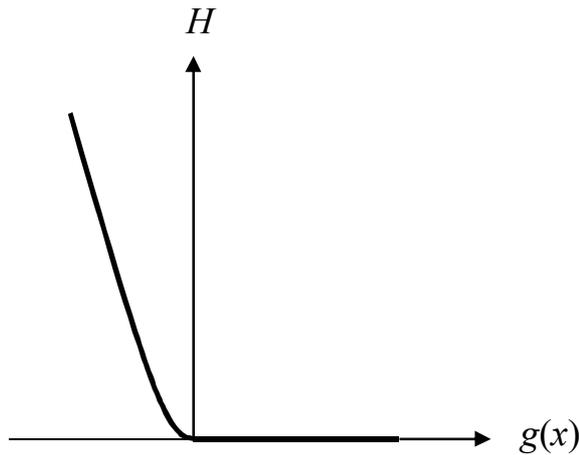


Рисунок 6.14 – Штраф типа квадрата срезки

Отметим, что  $H$  – внешний штраф, поэтому стационарные точки функции  $P(x, R)$  могут оказаться недопустимыми. С другой стороны, недопустимые точки не создают в данном случае дополнительных сложностей по сравнению с допустимыми точками. Различие между ними состоит лишь в том, что в допустимых и граничных точках  $H = 0$ .

Достоинством этого штрафа является то, что функция  $P(x, R)$  определена и непрерывна всюду. Вычисления проводятся с положительными  $R$ ; после решения очередной задачи безусловной минимизации  $R$  увеличивается.

*Пример.* Решим задачу определения наилучшего положения приемной антенны с одним ограничением (6.15):

$$f(x) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min,$$
$$g(x) = 4 - x_1 - x_2 \geq 0.$$

Используем штраф типа квадрата срезки. Для нашего случая

$$P(x, R) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 + R \cdot \langle 4 - x_1 - x_2 \rangle^2.$$

Уравнения, определяющие стационарную точку функции, имеют вид

$$\begin{cases} \frac{\partial P}{\partial x_1} = 2(x_1 - 6) - 2R \cdot \langle 4 - x_1 - x_2 \rangle = 0, \\ \frac{\partial P}{\partial x_2} = 2(x_2 - 8) - 2R \cdot \langle 4 - x_1 - x_2 \rangle = 0. \end{cases}$$

Выразим из первого уравнения

$$2R \cdot \langle 4 - x_1 - x_2 \rangle = 2(x_1 - 6)$$

и подставим во второе уравнение:

$$2(x_2 - 8) - 2(x_1 - 6) = 2x_2 - 2x_1 - 4 = 0 \Rightarrow x_2 = x_1 + 2.$$

С учетом этого получим

$$(x_1 - 6) - R \cdot \underbrace{\langle 2 - 2x_1 \rangle}_{\alpha} = 0.$$

Проанализируем результаты, когда аргумент оператора срезки  $\alpha \leq 0$  и  $\alpha > 0$ . Пусть  $\alpha > 0$ , т.е.  $\langle g(x) \rangle = 0$ . Это случай, когда ограничение не является активным, а оптимальная точка находится внутри ОДР (безусловный минимум). Тогда имеем

$$x_1 = 6, \quad x_2 = 8.$$

Однако данное решение не удовлетворяет ограничению задачи, т.е. является недопустимым ( $x \notin S$ ). Пусть  $\alpha \leq 0$ , т.е.  $\langle g(x) \rangle = 2 - 2x_1$ . Следовательно, ограничение полагается активным (лежит на границе ОДР). Тогда штраф становится эквивалентен квадратичному штрафу:

$$\begin{aligned} (x_1 - 4) - R(2 - 2x_1) &= 0 \Rightarrow \\ 2Rx_1 + x_1 - 2R - 6 &= 0 \Rightarrow x_1 = \frac{2R + 6}{2R + 1}. \end{aligned}$$

Переходя к пределу, получим

$$x_1 = \lim_{R \rightarrow \infty} \frac{2R + 6}{2R + 1} = 1.$$

Соответственно  $x_2 = x_1 + 2 = 3$ . Таким образом,  $x^* = (1; 3)^T$ ,  $f^* = 51$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 7,141$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 19,61 \text{ Вт/м}^2.$$

Получается, что если в оптимальной точке все ограничения задачи активны, то штраф типа квадрата срезки аналогичен квадратичному штрафу. Но в отличие от квадратичного штрафа, штраф типа квадрата срезки (ШКС) работает с ограничениями-неравенствами, поэтому может находить оптимальные точки не только на границе ОДР, но и внутри ОДР. В общем случае, если задача имеет  $m$  ограничений, при решении необходимо рассмотреть  $2^m$  различных вариантов, когда оптимальная точка лежит или не лежит на поверхности каждого из ограничений. То есть при  $m = 1$  получим два варианта решения (как в данном случае), при  $m = 2$  – уже четыре и т.д. Обратите внимание на пример решения задачи с помощью условий оптимальности Куна – Таккера – там как раз получены четыре варианта решения для задачи с двумя ограничениями, в каждом из которых набор активных ограничений меняется при помощи параметров  $\mu_1$  и  $\mu_2$ .

В таблице 6.4 исследованы стационарные точки функции  $P(x, R)$  для различных значений параметра  $R$ .

Таблица 6.4 – Стационарные точки штрафа типа квадрата срезки

$R$	$x_1^*; x_2^*$	$P(x, R)$	$f(x)$	$g(x)$
0	6; 8	1	1	-10
0,1	5,167; 7,167	9,333	2,389	-8,333
1	2,667; 4,667	34,33	23,22	-3,333
10	1,238; 3,238	48,62	46,35	-0,476
100	1,025; 3,025	50,75	50,50	-0,050
$\infty$	1; 3	51	51	0

Мы получили результаты, аналогичные квадратичному штрафу (КШ). Сравним вид ШФ для этих двух штрафов. При  $R = 1$  штрафная функция имеет вид

$$P(x,1) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 + \langle 4 - x_1 - x_2 \rangle^2.$$

На рисунке 6.15 изображены линии уровня штрафной функции для ШКС (толстые сплошные линии) и КШ (тонкие сплошные линии). Видим, что за пределами ОДР линии уровня для обоих штрафов совпадают и только в пределах ОДР наблюдаются различия. Они связаны с тем, что КШ считает эти точки недопустимыми (ограничение-равенство описывает только точки, лежащие на границе ОДР, отмеченной красной линией), а ШКС – допустимыми.

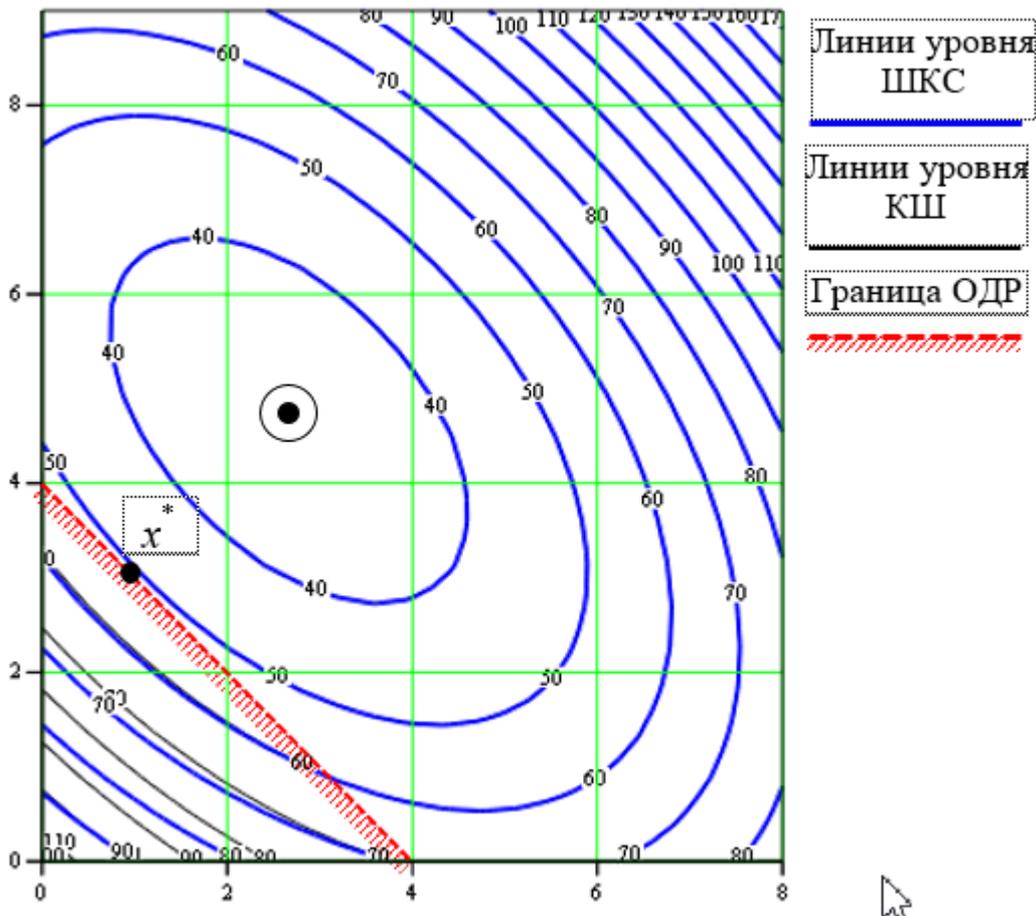


Рисунок 6.15 – Линии уровня штрафной функции для  $R = 1$

При  $R = 10$  штрафная функция имеет вид

$$P(x,10) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 + 10 \langle 4 - x_1 - x_2 \rangle^2.$$

На рисунке 6.16 также изображены линии уровня штрафной функции для ШКС (толстые сплошные линии) и КШ (тонкие сплошные линии). Видим, что за пределами ОДР эти штрафы вновь совпадают, но внутри ОДР штрафная функция для ШКС имеет гораздо меньшее значение, чем для КШ. Поэтому мы видим линии уровня только для КШ, а линии уровня ШКС переместились глубже в ОДР. Чем больше величина параметра  $R$ , тем большую «овражность» приобретают линии уровня ШФ за пределами ОДР, но внутри ОДР сохраняют свой начальный вид, как бы «заталкивая» решение в пределы ОДР (рисунок 6.17).

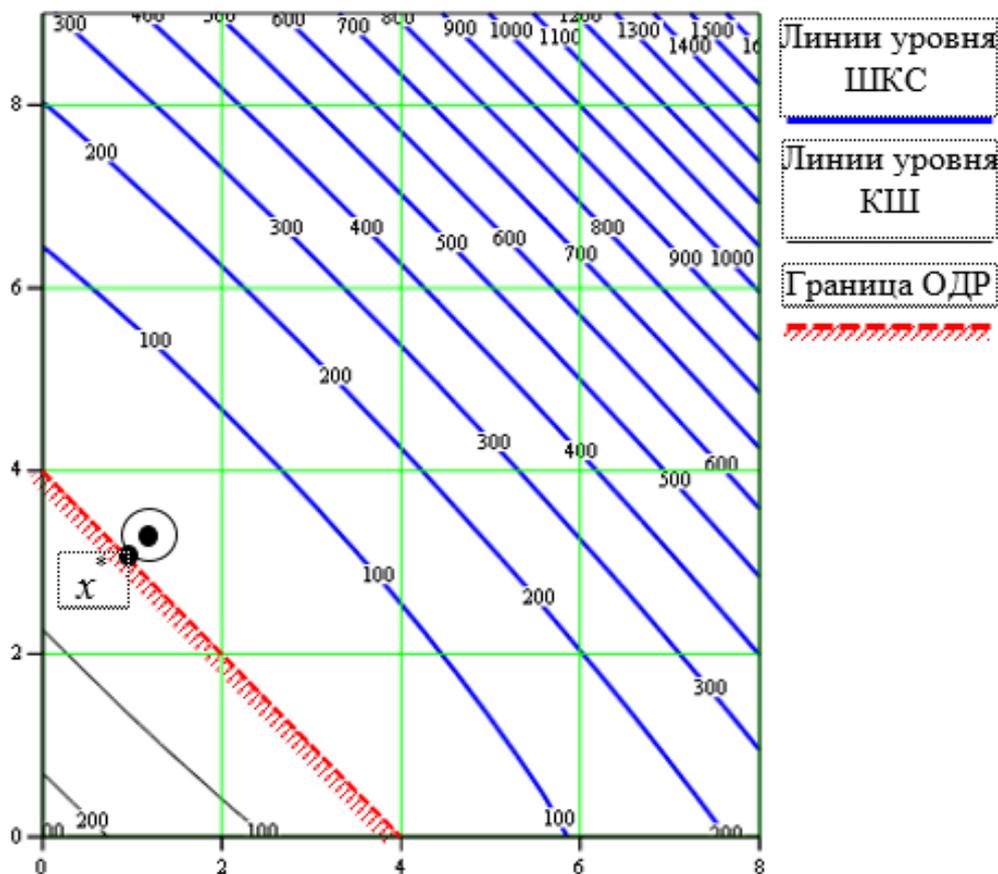


Рисунок 6.16 – Линии уровня штрафной функции для  $R = 10$

Таким образом, стационарная точка функции  $P(x, 0)$  есть точка  $x = (6; 8)^T$ . При изменении  $R$  от 0 до  $\infty$  стационарная точка функции  $P(x, R)$  перемещается вдоль прямой, соединяющей точку  $(6; 8)^T$  –

безусловный минимум  $f(x)$ , с точкой  $(1;3)^T$  ее условного минимума. При любом (конечном)  $R$  соответствующая стационарная точка недопустима, т.е. мы имеем дело с методом внешней точки.

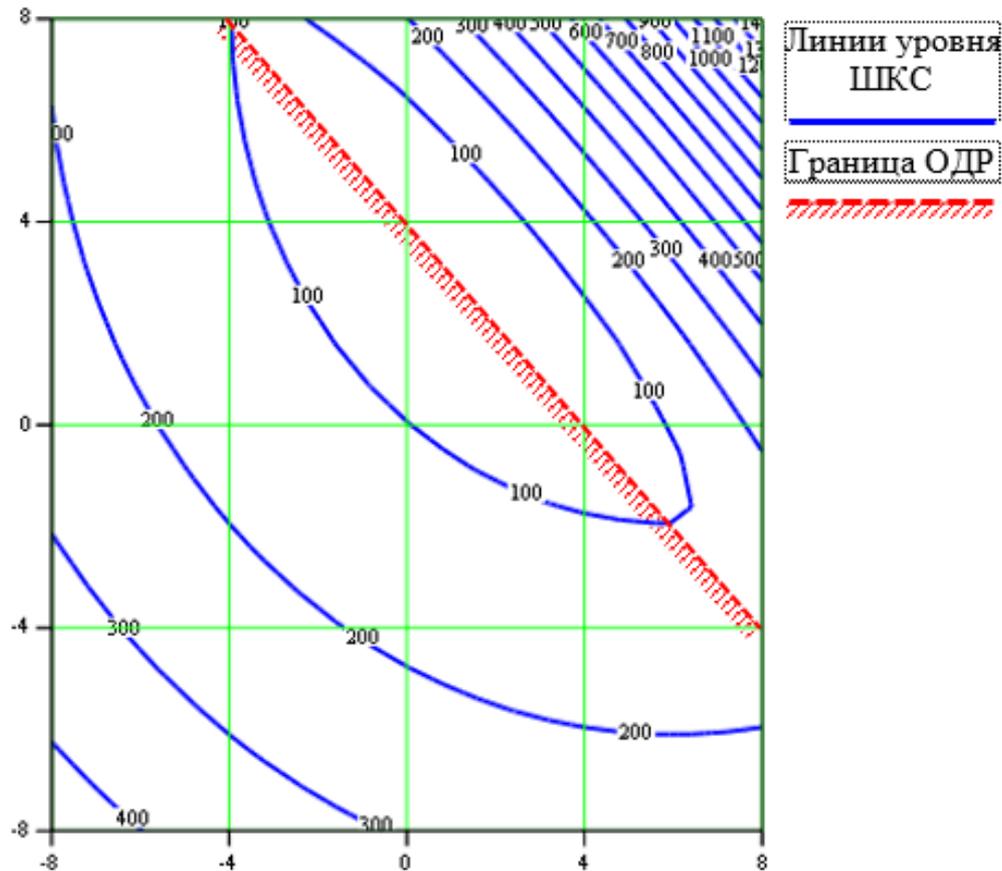


Рисунок 6.17 – Деформация линии уровня штрафной функции для  $R = 10$

### 6.5.6. Выбор штрафного параметра

Рассмотренные примеры демонстрируют, с одной стороны, принципиальную возможность использования методов ШФ, а с другой стороны, их общие недостатки, а именно сходимость метода связана с ярко выраженной нелинейностью («овражностью») ШФ, и здесь традиционные методы безусловной минимизации могут оказаться неэффективными.

При практическом использовании любой из представленных выше штрафных функций требуется выбрать начальное значение  $R$

и изменять его после решения каждой подзадачи безусловной минимизации, чтобы обеспечить сходимость последовательности стационарных точек  $x^t$ .

Значение параметра  $R$  необходимо выбирать следующим образом: при использовании внешних штрафов, переходя от одной подзадачи к другой, значение  $R$  надо увеличивать, а при использовании внутренних (барьерных) штрафов – уменьшать.

При этом необходимо установить, как именно будет изменяться  $R$  при переходе от одной подзадачи к другой. Например, для обычного квадратичного штрафа, учитывающего ограничения-равенства, целесообразно начинать с  $R = 0$ , т.е. безусловной минимизации функции, а затем последовательно увеличивать  $R$  на некоторое значение  $\Delta R$ . Вычислительные эксперименты показывают, что точно найденная стационарная точка  $x^t(R)$  перемещается вдоль некоторой гладкой кривой к точке  $x^*$ .

С другой стороны, пересчет параметра  $R$ , имеющий целью обеспечить сходимость метода, автоматически приводит к ухудшению обусловленности вспомогательных задач (появление «овражной» структуры штрафной функции). Во многих методах безусловной оптимизации используется матрица Гессе  $\nabla^2 P(x, R)$ . Часть собственных значений этой матрицы зависит от величины  $R$ . При этом обусловленность матрицы Гессе  $\nabla^2 P(x, R)$  неограниченно ухудшается, когда  $R \rightarrow \infty$ . Тем самым усложняется процесс решения вспомогательных задач безусловной оптимизации.

### 6.5.7. Обобщенный алгоритм

Построение эффективных алгоритмов на основе метода штрафных функций не представляет особых трудностей. Приведем простейший алгоритм, в котором используется штрафная функция.

1. Задать значения:

$n$  – размерность вектора  $x$ ;

$m$  – число ограничений-неравенств;

$p$  – число ограничений-равенств;

$\varepsilon_1$  – параметр окончания одномерного поиска;

$\varepsilon_2$  – параметр окончания процедуры безусловной минимизации;

$\varepsilon_3$  – параметр окончания работы алгоритма;

$x^0$  – начальное приближение для  $x^*$ ;

$R^0$  – начальный вектор штрафных параметров.

2. Построить  $P(x, R) = f(x) + H[R, g(x), h(x)]$ .

3. Найти значение  $x^{t+1}$ , доставляющее минимум функции  $P(x^{t+1}, R^t)$  при фиксированном  $R^t$ . В качестве начальной точки использовать  $x^t$ , в качестве параметра окончания шага – константу  $\varepsilon_2$ .

4. Проверить выполнение условий окончания поиска:

$$\left| P(x^{t+1}, R^t) - P(x^t, R^{t-1}) \right| \leq \varepsilon_3, \quad \|x^{t+1} - x^t\| \leq \varepsilon_3.$$

Да: положить  $x^* = x^{t+1}$  и закончить работу алгоритма.

Нет: перейти на шаг 5.

5. Положить  $R^{t+1} = R^t + \Delta R^t$  в соответствии с используемым правилом пересчета штрафных параметров и перейти на шаг 2.

При наличии ограничений в форме неравенств и равенств используют следующие штрафные функции:

$$P(x, R) = f(x) + R \cdot \sum_{j=1}^m \frac{1}{g_j(x)} + \frac{1}{R} \cdot \sum_{l=1}^p [h_l(x)]^2, \quad R^0 \gg 1, \quad R^t \rightarrow 0;$$

$$P(x, R) = f(x) - R \cdot \sum_{j=1}^m \ln[g_j(x)] + \frac{1}{R} \cdot \sum_{l=1}^p [h_l(x)]^2, \quad R^0 \gg 1, \quad R^t \rightarrow 0;$$

$$P(x, R) = f(x) + R \cdot \sum_{j=1}^m \langle g_j(x) \rangle^2 + R \cdot \sum_{l=1}^p [h_l(x)]^2, \quad R^0 \ll 1, \quad R^t \rightarrow \infty$$

и другие.

## 6.6. Методы, основанные на линеаризации

### 6.6.1. Базовый метод линеаризации

Наиболее изученными в теории оптимизации являются два класса задач: задачи без ограничений и задачи с линейными ограничениями. Поэтому в большинстве случаев для решения задач с нелинейными ограничениями используются методы, разработанные для упомянутых классов задач оптимизации. В предыдущих подразделах применялся первый подход – попытка свести ЗНП к задачам без ограничений.

Теперь рассмотрим второй подход, основанный на линеаризации, который позволяет свести общую задачу к задаче с линейными ограничениями (или даже линейной ЦФ). Использование линеаризации дает возможность применять методы линейного программирования (ЛП).

Все рассматриваемые ниже методы основываются на разложении нелинейной функции общего вида  $f(x)$  в ряд Тейлора до членов первого порядка в окрестности некоторой точки  $x^0$ :

$$f(x) = f(x^0) + \nabla f^T(x^0) \cdot (x - x^0) + O \cdot (\|x - x^0\|)^2,$$

где  $O \cdot (\|x - x^0\|)^2$  – отбрасываемый член второго порядка малости.

Таким образом, функция  $f(x)$  аппроксимируется в точке  $x^0$  линейной функцией:

$$\bar{f}(x, x^0) \cong f(x^0) + \nabla f(x^0) \cdot (x - x^0), \quad (6.49)$$

где  $x^0$  – точка линеаризации. Аналогично линеаризуются ограничения задачи, если они представлены нелинейными функциями.

Рассмотрим общую задачу нелинейного программирования:

$$\begin{aligned} f(x) &\rightarrow \min, \\ g_i(x) &\geq 0, \quad i = 1, \dots, m, \\ h_l(x) &= 0, \quad l = 1, \dots, p, \\ a_j &\leq x_j \leq b_j. \end{aligned}$$

Применим к ней описанный выше подход и сведем ее к ЗЛП.

*Алгоритм линеаризации.* Исходные данные: ЦФ, ограничения, начальная точка  $x^0 \in S$ , точность решения по аргументу и по функции – абсолютная ( $\varepsilon_x$  и  $\varepsilon_y$ ) или относительная ( $\delta_x$  и  $\delta_y$ ).

1. Задать номер итерации  $k = 0$ . Вычислить  $f(x^0)$ .

2. Вычислить  $g_i(x^k)$ ,  $h_l(x^k)$ ,  $\nabla f(x^k)$ ,  $\nabla g_i(x^k)$ ,  $\nabla h_l(x^k)$ . Если безусловное оптимальное решение лежит внутри ОДР (т.е. условный минимум совпадает с безусловным минимумом), можно проверить условие  $\|\nabla f(x^k)\| \leq \varepsilon_y$ . Если оно выполняется, то положить  $x^* \approx x^k$  и закончить вычисления.

3. Иначе линеаризовать задачу в точке  $x^k$  – текущей оценке решения. Использование непосредственной линеаризации приводит к следующей задаче:

$$\begin{aligned} f(x^k) + \nabla f(x^k) \cdot (x - x^k) &\rightarrow \min, \\ g_i(x^k) + \nabla g_i(x^k) \cdot (x - x^k) &\geq 0, \\ h_l(x^k) + \nabla h_l(x^k) \cdot (x - x^k) &= 0, \\ a_j \leq x_j \leq b_j. \end{aligned}$$

Эта задача представляет собой ЗЛП.

4. Решить ЗЛП и найти новое приближение  $x^{k+1}$ , которое может и не принадлежать допустимой области решений  $S$ . Если  $x^{k+1} \notin S$ , то оптимальное значение линеаризованной целевой функции, удовлетворяющее неравенству

$$\bar{f}(x^{k+1}, x^k) < \bar{f}(x^k, x^k),$$

может не быть точной оценкой истинного значения оптимума, поэтому следует вернуться на шаг 2. Иначе перейти к шагу 5.

5. Если  $\|x^{k+1} - x^k\| \leq \varepsilon_x$  и  $|f(x^{k+1}) - f(x^k)| \leq \varepsilon_y$  (если заданы относительные погрешности, то  $\|x^{k+1} - x^k\| / \|x^{k+1}\| \leq \delta_x$ ,  $|f(x^{k+1}) - f(x^k)| / |f(x^{k+1})| \leq \delta_y$ ), положить  $x^* \approx x^{k+1}$  и закончить вычисления. Иначе выполнить присваивание  $k = k + 1$  и вернуться на шаг 2.

Для сходимости к экстремуму достаточно, чтобы для последовательности точек  $\{x^k\}$ , полученных в результате решения последовательности подзадач ЛП, значение целевой функции и невязки по ограничениям в точке  $x^{k+1}$  было меньше их значений в точке  $x^k$ .

*Пример.* Решим задачу определения наилучшего положения приемной антенны. Так как метод линеаризации позволяет решать задачи со смешанными ограничениями, используем это и заменим одно из условий задачи на равенство:

$$\begin{aligned} f(x) &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ g(x) &= x_1^2 - x_2 \geq 0, \\ h(x) &= 4 - x_1^2 - x_2^2 = 0. \end{aligned}$$

Как мы уже убедились, в оптимальной точке оба ограничения являются активными и это не должно повлиять на результат решения.

Так как приемная антенна расположена между населенным пунктом и телевизионной вышкой, то в качестве ограничений вида  $a_j \leq x_j \leq b_j$  примем

$$0 \leq x_1 \leq 6, \quad 0 \leq x_2 \leq 8.$$

Построим допустимую область  $S$  (рисунок 6.18). Допустимая область  $S$  состоит из точек кривой  $h(x) = 0$ , лежащих между точкой  $(2; 0)$ , определяемой ограничением  $x_2 \geq 0$ , и точкой  $(1,25; 1,562)$ , определяемой ограничением  $g(x) \geq 0$ .

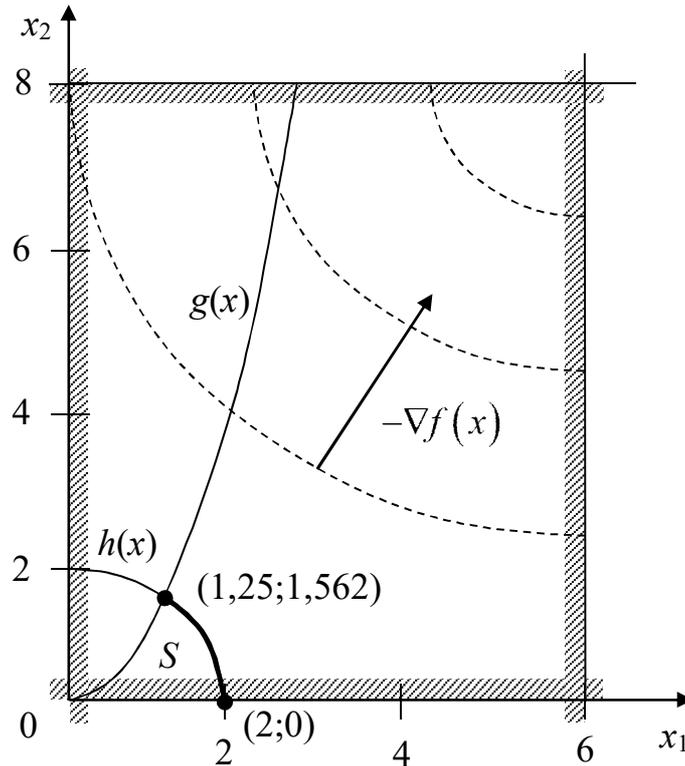


Рисунок 6.18 – Область допустимых решений задачи и линии уровня целевой функции

Зададим  $x^0 = (2;0)^T \in S$ ,  $\varepsilon_x = \varepsilon_y = 0,03$ .

Шаг 1. Выполняем присваивание  $k = 0$ . Вычисляем  $f(x^0) = 81$ .

Шаг 2. Вычисляем  $g(x^0) = 4$ ,  $h(x^0) = 0$ ,

$$\nabla f(x^0) = \begin{pmatrix} 2 \cdot 2 - 12 \\ 2 \cdot 0 - 16 \end{pmatrix} = \begin{pmatrix} -8 \\ -16 \end{pmatrix},$$

$$\nabla g(x^0) = \begin{pmatrix} 2 \cdot 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}, \quad \nabla h(x^0) = \begin{pmatrix} -2 \cdot 2 \\ -2 \cdot 0 \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \end{pmatrix}.$$

Так как  $\|\nabla f(x^0)\| \approx 17,889 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3. Составим ЗЛП:

$$\bar{f}(x, x^0) = 81 - 8 \cdot (x_1 - 2) - 16 \cdot (x_2 - 0) \rightarrow \min,$$

$$\bar{g}(x, x^0) = 4 + 4 \cdot (x_1 - 2) - 1 \cdot (x_2 - 0) \geq 0,$$

$$\bar{h}(x, x^0) = 0 - 4 \cdot (x_1 - 2) - 0 \cdot (x_2 - 0) = 0,$$

$$0 \leq x_1 \leq 6, \quad 0 \leq x_2 \leq 8.$$

Здесь  $\bar{S}$  представляет собой отрезок прямой  $\bar{h}(x) = 0$  ( $-4x_1 + 8 = 0$ ), ограниченный точками  $(2;0)$  и  $(2;4)$  – точкой пересечения с прямой  $\bar{g}(x) = 0$  ( $4x_1 - x_2 - 4 = 0$ ) (рисунок 6.19). Линии уровня линеаризованной целевой функции  $\bar{f}(x)$  представляют собой прямые  $97 - 8x_1 - 16x_2 = \text{const}$ , тогда как линии уровня исходной целевой функции – окружности с центром в точке  $(6;8)$ .

Шаг 4. Ясно, что решением линеаризованной задачи является точка  $x^1 = (2;4)^T$ . В этой точке имеем:

$$g(x^1) = 0,$$

$$h(x^1) = -16 \neq 0,$$

так что ограничение-равенство нарушается. Возвращаемся на шаг 2.

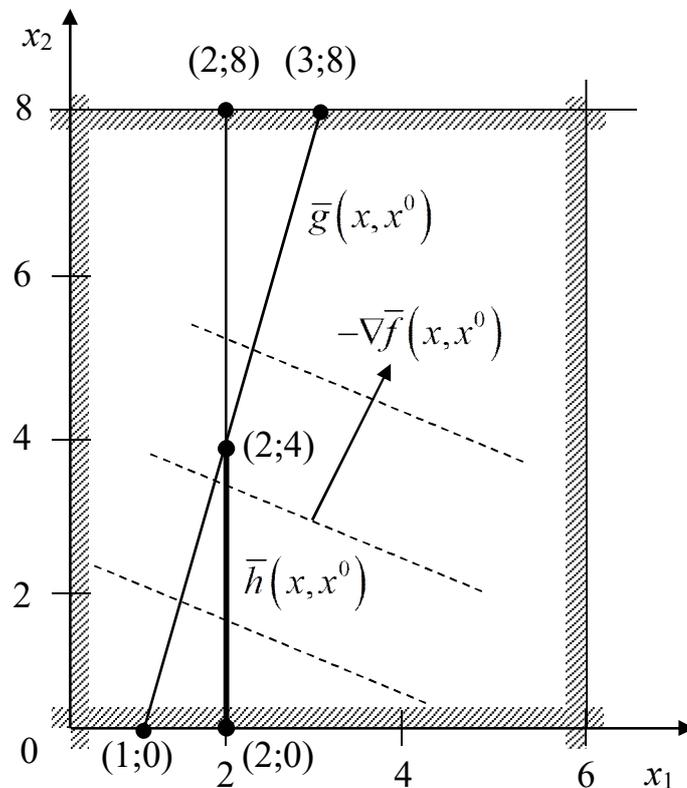


Рисунок 6.19 – Линии уровня целевой функции

Шаг 2<sup>(1)</sup>. Вычисляем  $f(x^1) = 33$ ,  $g(x^1) = 0$ ,  $h(x^1) = -16$ ,

$$\nabla f(x^1) = \begin{pmatrix} 2 \cdot 2 - 12 \\ 2 \cdot 4 - 16 \end{pmatrix} = \begin{pmatrix} -8 \\ -8 \end{pmatrix},$$

$$\nabla g(x^1) = \begin{pmatrix} 2 \cdot 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}, \quad \nabla h(x^1) = \begin{pmatrix} -2 \cdot 2 \\ -2 \cdot 4 \end{pmatrix} = \begin{pmatrix} -4 \\ -8 \end{pmatrix}.$$

Так как  $\|\nabla f(x^1)\| \approx 11,314 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3<sup>(1)</sup>. Составим ЗЛП:

$$\bar{f}(x, x^1) = 33 - 8 \cdot (x_1 - 2) - 8 \cdot (x_2 - 4) \rightarrow \min,$$

$$\bar{g}(x, x^1) = 0 + 4 \cdot (x_1 - 2) - 1 \cdot (x_2 - 4) \geq 0,$$

$$\bar{h}(x, x^1) = -16 - 4 \cdot (x_1 - 2) - 8 \cdot (x_2 - 4) = 0,$$

$$0 \leq x_1 \leq 6, \quad 0 \leq x_2 \leq 8.$$

И так далее, остальные итерации решения сведены в таблицу 6.5.

Таблица 6.5 – Решение задачи методом линейризации

$k$	$x^k$	$f(x^k)$	$\nabla f(x^k)$			$\ \nabla f(x^k)\ $
		$g(x^k)$	$\nabla g(x^k)$	$\nabla h(x^k)$	$\ x^{k+1} - x^k\ $	$ f(x^{k+1}) - f(x^k) $
1	2	3	4	5	6	7
0	2	81	-8	4	-4	17,89
	0	4	-16	-1	0	4
		0				48
1	2	33	-8	4	-4	11,31
	4	0	-8	-1	-8	5,657
		-16				32
2	6	65	0	12	-12	16
	0	36	-16	-1	0	4,807
		-32				40,89

Окончание таблицы 6.5

1	2	3	4	5	6	7
3	3,333 4	24,11 7,111 -23,11	-5,333 -8	6,667 -1	- 6,667 -8	9,615 2,222 26,27
4	2 2,222	50,38 1,778 -4,938	-8 -11,56	4 -1	-4 - 4,444	14,06 0,827 12,11
5	1,410 1,642	62,49 0,348 -0,685	-9,179 -12,72	2,821 -1	- 2,821 - 3,283	15,68 0,170 2,416
6	1,259 1,563	64,91 0,0228 -0,0290	-9,481 -12,87	2,519 -1	- 2,519 - 3,126	15,99 $9,777 \times 10^{-3}$ 0,111
7	1,250 1,562	65,02 $9,022 \times 10^{-5}$ $-1,053 \times 10^{-4}$	-9,501 -12,88	2,499 -1	- 2,499 - 3,121	16,00 $3,123 \times 10^{-6}$ $4,022 \times 10^{-5}$
8	1,250 1,562	65,02 $8,709 \times 10^{-5}$ $-1,151 \times 10^{-4}$	-9,501 -12,88	2,499 -1	- 2,499 - 3,121	16,00

С учетом того что  $\varepsilon_x = \varepsilon_y = 0,03$ , округляем значения аргумента и функции до двух цифр после запятой:  $x^* \approx x^8 \approx (1,25; 1,56)^T$ ,  $f^* = f(x^8) \approx 65,02$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 8,063$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 15,38 \text{ Вт/м}^2.$$

*Замечание.* Линеаризацию следует использовать с большой осторожностью, поскольку иногда она дает весьма грубое приближение.

Из таблицы (6.5) видно, что значения  $f(x)$ ,  $g(x)$  и  $h(x)$  монотонно улучшаются. Однако такая монотонность характерна для задач, функции которых являются «умеренно» нелинейными. В случае функций с ярко выраженной нелинейностью монотонность улучшения нарушается и алгоритм перестает сходиться.

Существуют три способа усовершенствования методов непосредственной линеаризации:

- 1) использование линейного приближения для отыскания направления спуска;
- 2) глобальная аппроксимация нелинейной функции задачи при помощи кусочно-линейчатой функции;
- 3) применение последовательных линеаризаций на каждой итерации для уточнения допустимой области  $S$ .

Далее рассмотрим эти три способа.

### 6.6.2. Алгоритм Франка – Вульфа

Наиболее простым способом линеаризации является замена общей нелинейной задачи на задачу, получаемую при помощи линеаризации всех функций исходной задачи. Таким образом, получается задача линейного программирования, которую можно решать при помощи методов линейного программирования.

Рассмотрим задачу нелинейного программирования с линейными ограничениями:

$$f(x) \rightarrow \min, \quad (6.50)$$

$$Ax \leq b, \quad x \geq 0, \quad (6.51)$$

где  $f(x)$  – нелинейная целевая функция.

Допустимая область  $S$  является выпуклым многогранником, образованным пересечением полуплоскостей. Так как функция  $f(x)$  нелинейная, то оптимальное решение может не совпадать с вершиной (или угловой точкой) области  $S$ . Кроме того, если  $f(x)$  не выпуклая, то задача нелинейного программирования может иметь несколько локальных минимумов.

Рассмотрим задачу линейного программирования, полученную линеаризацией исходной задачи нелинейного программирования в некоторой допустимой точке  $x^0 \in S$ :

$$\begin{aligned} \bar{f}(x, x^0) &\rightarrow \min, \\ Ax &\leq b, \quad x \geq 0. \end{aligned}$$

Данная задача линейного программирования имеет оптимальное решение в угловой точке  $\bar{x}^* \in S$ .

Важен вопрос о соотношении  $\bar{x}^*$  и  $x^*$  – решением исходной задачи, причем исходная задача может иметь несколько локальных минимумов. Выясним, будет ли  $\bar{x}^* \rightarrow x^*$ . В силу допустимости точек  $x^0$  и  $\bar{x}^*$  имеет место неравенство

$$\bar{f}(x^0, x^0) \geq \bar{f}(\bar{x}^*, x^0).$$

Это следует из того, что  $\bar{x}^*$  является минимумом функции  $\bar{f}(x, x^0)$ . Следовательно, если воспользоваться формулой для  $\bar{f}$ , то получается следующее неравенство:

$$\bar{f}(x^0, x^0) = f(x^0) \geq f(x^0) + \nabla f^T(x^0)(\bar{x}^* - x^0) \quad (6.52)$$

или

$$\nabla f^T(x^0)(\bar{x}^* - x^0) \leq 0. \quad (6.53)$$

Очевидно, что вектор  $(\bar{x}^* - x^0)$  задает направление спуска.

В задаче оптимизации без ограничений применение направления спуска в качестве направления поиска эффективно лишь при использовании специальных правил изменения шага или одномерного поиска. Заметим, что одномерный поиск из точки  $x^0$  в направлении  $(\bar{x}^* - x^0)$  приводит к точке  $\bar{x}^*$ . Поскольку  $\bar{x}^*$  является угловой точкой области  $S$  и точка  $x^0 \in S$ , то все точки на отрезке прямой между ними также допустимы (так как  $S$  – выпуклая область). Таким образом, одномерный поиск должен производиться на отрезке прямой

$$x = x^0 + \alpha(\bar{x}^* - x^0), \quad 0 \leq \alpha \leq 1. \quad (6.54)$$

Следовательно, решение задачи линейного программирования, даже не дающее хорошего приближения к экстремуму, позволяет получить важную информацию: определить направление поиска и точки пересечения соответствующего луча с границей допустимой области  $S$ .

В результате решения задачи

$$f(x^0 + \alpha(\bar{x}^* - x^0)) \rightarrow \min_{\alpha}, \quad 0 \leq \alpha \leq 1, \quad (6.55)$$

находится точка  $x^1 \in S$ , такая, что  $f(x^1) < f(x^0)$ .

Поскольку величина  $\nabla f(x^1) \neq 0$ , то полученная точка  $x^1$  может служить точкой линеаризации для построения следующей аппроксимации. Решение последовательности задач линейного программирования и одномерный поиск продолжают до тех пор, пока расстояние между точками последовательных оптимумов  $x^k$  не станет меньше значения  $\varepsilon$ .

На изложенной информации основан оптимизационный метод Франка – Вульфа (МФВ).

*Алгоритм Франка – Вульфа.* Исходные данные: ЦФ, ограничения, начальная точка  $x^0 \in S$ , точность решения по аргументу и по функции – абсолютная ( $\varepsilon_x$  и  $\varepsilon_y$ ) или относительная ( $\delta_x$  и  $\delta_y$ ).

1. Задать номер итерации  $k = 0$ . Вычислить  $f(x^0)$ .

2. Вычислить  $\nabla f(x^k)$ . Если  $\|\nabla f(x^k)\| \leq \varepsilon_y$ , то положить  $x^* \approx x^k$

и закончить вычисления.

3. Решить следующую задачу линейного программирования:

$$\begin{aligned} \nabla f(x^k) \cdot y &\rightarrow \min, \\ Ay &\leq b, \quad y \geq 0. \end{aligned}$$

Пусть  $y^k$  – оптимальное решение этой задачи.

4. Найти  $\alpha^k$ , представляющее собой решение задачи при

$$\Phi(\alpha) = f\left(x^k + \alpha(y^k - x^k)\right) \rightarrow \min_{\alpha}, \quad 0 \leq \alpha \leq 1.$$

5. Вычислить  $x^{k+1} = x^k + \alpha^k(y^k - x^k)$ .

6. Проверить близость к решению. Если  $\|x^{k+1} - x^k\| \leq \varepsilon_x$  и  $|f(x^{k+1}) - f(x^k)| \leq \varepsilon_y$  (если заданы относительные погрешности, то  $\|x^{k+1} - x^k\| / \|x^{k+1}\| \leq \delta_x$ ,  $|f(x^{k+1}) - f(x^k)| / |f(x^{k+1})| \leq \delta_y$ ), положить  $x^* \approx x^{k+1}$  и закончить вычисления. Иначе выполнить присваивание  $k = k + 1$  и вернуться на шаг 2.

*Замечание.* В формулировке подзадачи линейного программирования опущены постоянные члены целевой функции  $\bar{f}(x, x^k)$ , так как они не влияют на решение.

*Пример.* Решим задачу определения наилучшего положения приемной антенны. Метод Франка – Вульфа предназначен для решения задач с линейными ограничениями, поэтому используем ограничения (6.15), (6.16), (6.17):

$$\begin{aligned} f(x) &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ g(x) &= 4 - x_1 - x_2 \geq 0, \\ x_{1,2} &\geq 0. \end{aligned}$$

Приведем ограничение задачи к виду (6.51):

$$x_1 + x_2 \leq 4,$$

таким образом,

$$A = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad b = 4.$$

Зададим  $x^0 = (1; 1)^T \in S$ ,  $\varepsilon_x = \varepsilon_y = 0,03$ .

Шаг 1. Выполняем присваивание  $k = 0$ . Вычисляем  $f(x^0) = 75$ .

Шаг 2. Вычисляем

$$\nabla f(x^0) = \begin{pmatrix} 2 \cdot 1 - 12 \\ 2 \cdot 1 - 16 \end{pmatrix} = \begin{pmatrix} -10 \\ -14 \end{pmatrix}.$$

Так как  $\|\nabla f(x^0)\| \approx 17,205 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3. Составим ЗЛП:

$$\begin{aligned} -10y_1 - 14y_2 &\rightarrow \min, \\ y_1 + y_2 &\leq 4, \\ y_{1,2} &\geq 0. \end{aligned}$$

Данную задачу можно решить графически либо симплекс-методом. Но очевидно, что минимум достигается при наибольшем возможном значении переменной  $y_2$ , поэтому решением будет  $y^0 = (0; 4)^T$ .

Шаг 4. Выполним одномерный поиск вдоль прямой:

$$\begin{aligned} \Phi(\alpha) &= f(x^0 + \alpha(y^0 - x^0)) \rightarrow \min_{\alpha}, \alpha \in [0, 1] \Rightarrow \\ f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha\left(\begin{pmatrix} 0 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)\right) &= f\left(\begin{pmatrix} 1-\alpha \\ 1+3\alpha \end{pmatrix}\right) = 1 + (1-\alpha-6)^2 + (1+3\alpha-8)^2 \rightarrow \min \Rightarrow \\ \Phi(\alpha) &= 1 + (-\alpha-5)^2 + (3\alpha-7)^2 \rightarrow \min. \end{aligned}$$

Дифференцируем:

$$\begin{aligned} \frac{d\Phi(\alpha)}{d\alpha} &= -2(-\alpha-5) + 6(3\alpha-7) = 0, \alpha \in [0, 1] \Rightarrow \\ &\Rightarrow 20\alpha = 32 \Rightarrow \alpha^0 = 1. \end{aligned}$$

Шаг 5. Находим новое приближение оптимума:

$$x^1 = x^0 + \alpha^0(y^0 - x^0) = \begin{pmatrix} 1-\alpha \\ 1+3\alpha \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

Шаг 6. Вычисляем  $f(x^1) = 53$ . Условия останова итерационного процесса не выполняются, так как

$$\|x^1 - x^0\| \approx 3,162 > \varepsilon_x, \quad |f(x^1) - f(x^0)| = 22 > \varepsilon_y,$$

поэтому полагаем  $k = 1$  и возвращаемся на шаг 2.

Шаг 2<sup>(1)</sup>. Вычисляем

$$\nabla f(x^1) = \begin{pmatrix} 2 \cdot 0 - 12 \\ 2 \cdot 4 - 16 \end{pmatrix} = \begin{pmatrix} -12 \\ -8 \end{pmatrix}.$$

Так как  $\|\nabla f(x^1)\| \approx 14,422 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3<sup>(1)</sup>. Составим ЗЛП:

$$\begin{aligned} -12y_1 - 8y_2 &\rightarrow \min, \\ y_1 + y_2 &\leq 4, \\ y_{1,2} &\geq 0. \end{aligned}$$

В данном случае минимум достигается при наибольшем возможном значении переменной  $y_1$ , поэтому решением будет  $y^1 = (4; 0)^T$ .

Шаг 4<sup>(1)</sup>. Выполним одномерный поиск вдоль прямой:

$$\begin{aligned} \Phi(\alpha) &= f\left(x^1 + \alpha(y^1 - x^1)\right) \rightarrow \min_{\alpha}, \alpha \in [0, 1] \Rightarrow \\ f\left(\begin{pmatrix} 0 \\ 4 \end{pmatrix} + \alpha\left(\begin{pmatrix} 4 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 4 \end{pmatrix}\right)\right) &= f\left(\begin{pmatrix} 4\alpha \\ 4 - 4\alpha \end{pmatrix}\right) = 1 + (4\alpha - 6)^2 + (4 - 4\alpha - 8)^2 \rightarrow \min \Rightarrow \\ \Phi(\alpha) &= 1 + (4\alpha - 6)^2 + (-4\alpha - 4)^2 \rightarrow \min. \end{aligned}$$

Дифференцируем:

$$\begin{aligned} \frac{d\Phi(\alpha)}{d\alpha} &= 8(4\alpha - 6) - 8(-4\alpha - 4) = 0, \alpha \in [0, 1] \Rightarrow \\ &\Rightarrow 64\alpha - 16 = 0 \Rightarrow \alpha^1 = \frac{1}{4}. \end{aligned}$$

Шаг 5<sup>(1)</sup>. Находим новое приближение оптимума:

$$x^2 = x^1 + \alpha^1(y^1 - x^1) = \begin{pmatrix} 4\alpha \\ 4 - 4\alpha \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Шаг 6<sup>(1)</sup>. Вычисляем  $f(x^2) = 51$ . Условия останова итерационного процесса не выполняются:

$$\|x^2 - x^1\| \approx 1,414 > \varepsilon_x, \quad |f(x^2) - f(x^1)| = 2 > \varepsilon_y,$$

поэтому полагаем  $k = 2$  и возвращаемся на шаг 2.

Шаг 2<sup>(2)</sup>. Вычисляем

$$\nabla f(x^2) = \begin{pmatrix} 2 \cdot 1 - 12 \\ 2 \cdot 3 - 16 \end{pmatrix} = \begin{pmatrix} -10 \\ -10 \end{pmatrix}.$$

Так как  $\|\nabla f(x^2)\| \approx 14,142 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3<sup>(2)</sup>. Составим ЗЛП:

$$\begin{aligned} -10y_1 - 10y_2 &\rightarrow \min, \\ y_1 + y_2 &\leq 4, \\ y_{1,2} &\geq 0. \end{aligned}$$

В данном случае углы наклона линий уровня ЦФ и границы ОДР, соответствующей ограничению, совпадают. Поэтому задача имеет множество решений на отрезке  $y_1 + y_2 = 4$ ,  $y_{1,2} \geq 0$ . Выберем любое из них, например  $y^2 = (2; 2)^T$ .

Шаг 4<sup>(2)</sup>. Выполним одномерный поиск вдоль прямой:

$$\begin{aligned} \Phi(\alpha) &= f\left(x^2 + \alpha(y^2 - x^2)\right) \rightarrow \min_{\alpha}, \quad \alpha \in [0, 1] \Rightarrow \\ f\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix} + \alpha\left(\begin{pmatrix} 2 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 3 \end{pmatrix}\right)\right) &= f\begin{pmatrix} 1 + \alpha \\ 3 - \alpha \end{pmatrix} = 1 + (1 + \alpha - 6)^2 + (3 - \alpha - 8)^2 \rightarrow \min \Rightarrow \\ \Phi(\alpha) &= 1 + (\alpha - 5)^2 + (-\alpha - 5)^2 \rightarrow \min. \end{aligned}$$

Дифференцируем:

$$\begin{aligned} \frac{d\Phi(\alpha)}{d\alpha} &= 2(\alpha - 5) - 2(-\alpha - 5) = 0, \quad \alpha \in [0, 1] \Rightarrow \\ &\Rightarrow 4\alpha = 0 \Rightarrow \alpha^2 = 0. \end{aligned}$$

Шаг 5<sup>(2)</sup>. Находим новое приближение оптимума:

$$x^3 = x^2 + \alpha^2 (y^2 - x^2) = \begin{pmatrix} 1 + \alpha \\ 3 - \alpha \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Шаг 6<sup>(2)</sup>. Вычисляем  $f(x^3) = 51$ . Условия останова итерационного процесса выполняются, так как

$$\|x^3 - x^2\| = 0 < \varepsilon_x, \quad |f(x^3) - f(x^2)| = 0 < \varepsilon_y,$$

поэтому решение окончено.

С учетом того что  $\varepsilon_x = \varepsilon_y = 0,03$ , округляем значения аргумента и функции до двух цифр после запятой:  $x^* = x^3 = (1,00; 3,00)^T$ ,  $f^* = f(x^3) = 51$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 7,141$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 19,61 \text{ Вт/м}^2.$$

*Замечание.* В принципе, решение можно было закончить на шаге 4<sup>(2)</sup>. Так как была получена нулевая длина шага  $\alpha^2$ , то понятно, что точка  $x^2$  уже является оптимальной и следующая точка  $x^3$  совпадет с ней.

Дадим графическую интерпретацию решения (рисунок 6.20). Штриховка показывают допустимую область. В случае оптимума, лежащего внутри допустимой области, алгоритм Франка – Вульфа весьма похож на градиентный метод оптимизации без ограничений (зигзагообразная траектория).

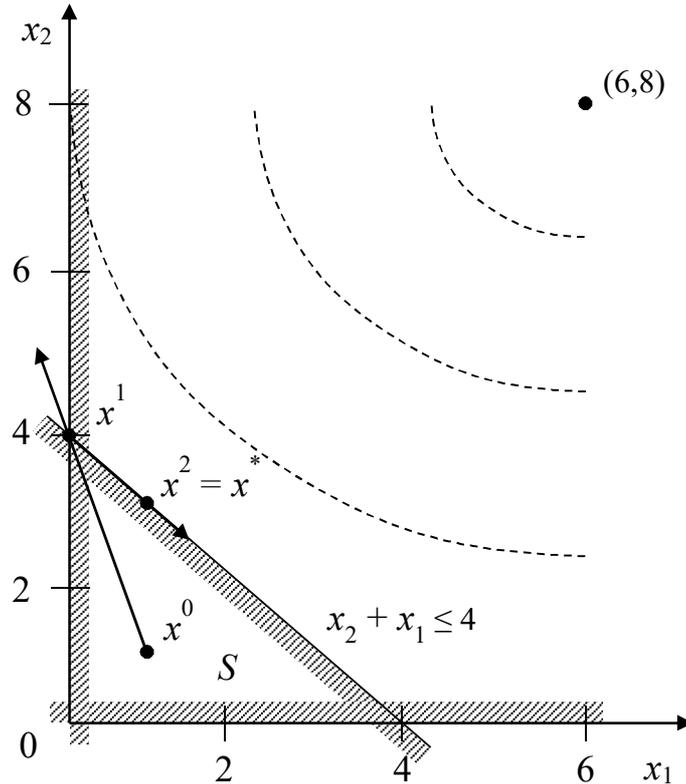


Рисунок 6.20 – Линии уровня целевой функции, область допустимых решений и траектория спуска алгоритма Франка – Вульфа

### 6.6.3. Метод допустимых направлений Зойтендейка

В методе Франка – Вульфа линейная аппроксимация используется *не для определения очередного приближения, а только для выявления подходящего направления поиска в каждой точке*. В этом случае точку экстремума вдоль данного направления можно определить непосредственной проверкой значений целевой функции и функций, входящих в ограничение. Аналогичная стратегия применяется в градиентных методах безусловной оптимизации. Продолжим исследование этого подхода для задач с нелинейными ограничениями.

В задачах с ограничениями линеаризуются и ограничения, и целевые функции, а направления необходимо выбирать так, чтобы они приводили к допустимым точкам  $x \in S$ .

Рассмотрим задачу с ограничениями в виде неравенств:

$$\begin{aligned} f(x) &\rightarrow \min, \\ g_i(x) &\geq 0, \quad i = 1, 2, \dots, m. \end{aligned} \tag{6.56}$$

Пусть  $x^0$  – начальная точка, удовлетворяющая ограничениям

$$g_i(x^0) \geq 0, \quad i = 1, 2, \dots, m.$$

Предположим, что некоторые ограничения являются активными в точке  $x^0$ :

$$I^0 = \{i : g_i(x^0) = 0, i \in I\}, \quad I^0 \subset I = \{1, 2, \dots, m\}.$$

Вектор  $d$  определяет допустимое направление для поиска, если  $d$  – направление спуска, т.е.

$$\nabla f^T(x) \cdot d < 0, \quad (6.57)$$

и точки луча

$$x(\alpha) = x^0 + \alpha d, \quad (6.58)$$

где  $\alpha \geq 0$ , являются допустимыми на небольшом расстоянии от  $x^0$ .

Неравенство  $\nabla f^T(x) \cdot d < 0$  получается следующим образом. Разложим в ряд Тейлора функцию  $f(x)$  в точке  $x^0$ :

$$f(x) = f(x^0) + \nabla f^T(x^0)(x - x^0).$$

В направлении спуска в допустимых точках должно выполняться неравенство  $f(x) < f(x^0)$ , следовательно,  $\nabla f^T(x^0)(x - x^0) < 0$ .

Вводя обозначение  $x - x^0 = d$ , получим искомый результат.

Точки  $x(\alpha)$  будут допустимыми, если для всех активных ограничений выполняется условие

$$\tilde{g}_i(x, x^0) = g_i(x^0) + \nabla g_i^T(x^0)(x - x^0) \geq 0. \quad (6.59)$$

Так как по предположению

$$g_i(x^0) = 0, \quad \forall i \in I^0 \text{ и } x - x^0 = d,$$

то выражение (6.59) эквивалентно следующему условию для  $d$ :

$$\nabla g_i^T(x^0) \cdot d \geq 0, \quad \forall i \in I^0. \quad (6.60)$$

*Основная идея:* на каждом шаге итерации определяется допустимое направление, т.е. вектор  $d$  и скалярный параметр  $\theta > 0$  такие, чтобы выполнялись неравенства

$$\nabla f^T(x^0) \cdot d \leq -\theta, \quad \nabla g_i^T(x^0) \cdot d \geq 0, \quad i \in I^0, \quad (6.61)$$

а значение  $\theta$  выбирается по возможности большим.

При реализации на ЭВМ допустимые направления удобно нормировать, вводя границы

$$-1 \leq d_i \leq 1, \quad i = 1, 2, \dots, n. \quad (6.62)$$

Такой способ выбора вектора  $d$  обеспечивает разумный компромисс между движением внутрь области допустимых решений без нарушения ограничений и движением по направлению наискорейшего спуска.

После того как вектор  $d$  выбран, очередное приближение может быть определено поиском минимального значения  $\alpha$  вдоль прямой

$$x = x^0 + \alpha d^0 \quad (6.63)$$

до тех пор, пока либо ЦФ  $f(x)$  не достигнет экстремума, либо какое-нибудь из ограничений не окажется нарушенным.

Обычно для каждого ограничения  $g_i(x) \geq 0$  находятся значения  $\alpha_i > 0$ , при которых

$$g_i(x^0 + \alpha d^0) = 0, \quad (6.64)$$

а затем определяется  $\bar{\alpha}$  как наименьшее из значений  $\alpha_i$ :

$$\bar{\alpha} = \min_i \{\alpha_i\}.$$

При известном значении  $\bar{\alpha}$  можно использовать любую процедуру одномерного поиска для определения  $\alpha$ , которое минимизирует функцию  $f(x^0 + \alpha d^0)$  на отрезке  $[0, \bar{\alpha}]$ .

Представленный алгоритм называется *методом допустимых направлений Зойтендейка* (МДНЗ).

*Алгоритм допустимых направлений Зойтендейка.* Исходные данные: ЦФ, ограничения, начальная точка  $x^0 \in S$ , точность решения по аргументу и по функции – абсолютная ( $\varepsilon_x$  и  $\varepsilon_y$ ) или относительная ( $\delta_x$  и  $\delta_y$ ), нулевой порог для определения активных ограничений  $\varepsilon_z$ .

1. Задать номер итерации  $k = 0$ . Вычислить  $f(x^0)$ .

2. Вычислить  $\nabla f(x^k)$ ,  $g_i(x^k)$ ,  $\nabla g_i(x^k)$ . Если  $\|\nabla f(x^k)\| \leq \varepsilon_y$ , то положить  $x^* \approx x^k$  и закончить вычисления.

3. В данной допустимой точке  $x^k \in S$  определить множество индексов тех ограничений, которые активны в точке  $x^k$ , т.е.

$$I^k = \{i : g_i(x^k) = 0, i = 1, 2, \dots, m\}.$$

Однако если алгоритм реализуется численно, из-за погрешностей вычислений на ЭВМ может наблюдаться не точное равенство активного ограничения нулю в точке  $x^k$ , а приближенное. С этой проблемой можно также столкнуться при округлении промежуточных результатов расчетов. В этом случае принимают

$$I^k = \{i : |g_i(x^k)| \leq \varepsilon_z, i = 1, 2, \dots, m\},$$

где  $\varepsilon_z$  – заданная погрешность определения активных ограничений.

4. Решить ЗЛП:

$$-\theta \rightarrow \min,$$

$$\nabla f^T(x^k) \cdot d \leq -\theta,$$

$$\nabla g_i^T(x^k) \cdot d \geq \theta, i \in I^k,$$

$$-1 \leq d_j \leq 1, j = 1, 2, \dots, n.$$

Пусть  $d^k$  и  $\theta^k$  – полученное решение. Если  $\theta^k \leq 0$ , то положить  $x^* \approx x^k$  и закончить вычисления, так как дальнейшее улучшение невозможно.

5. Иначе найти

$$\bar{\alpha} = \min \{\bar{\alpha}_i : g_i(x^k + \alpha d^k) = 0, \alpha \geq 0, i = 1, 2, \dots, m\}.$$

Если не существует  $\bar{\alpha}_i \geq 0$ , то положить  $\bar{\alpha}_i = \infty$  (в программной реализации можно взять просто большое число, например  $\bar{\alpha}_i = 10^{10}$ ). При поиске  $\bar{\alpha}$  как  $\min\{\bar{\alpha}_i : g_i(x^k + \alpha d^k) = 0\}$  может оказаться, что не для всех  $i = 1, 2, \dots, m$  будет выполняться равенство  $g_i(x^k + \bar{\alpha} d^k) = 0$ , значит, мы определяем сразу  $I^{k+1}$  – множество индексов, для которых  $g_i(x^{k+1}) = 0$ .

6. Найти  $\alpha^k$ , представляющее собой решение задачи при

$$\Phi(\alpha) = f(x^k + \alpha d^k) \rightarrow \min_{\alpha}, 0 \leq \alpha \leq \bar{\alpha}.$$

7. Положить  $x^{k+1} = x^k + \alpha^k d^k$ .

8. Проверить близость к решению. Если  $\|x^{k+1} - x^k\| \leq \varepsilon_x$  и  $|f(x^{k+1}) - f(x^k)| \leq \varepsilon_y$  (если заданы относительные погрешности, то  $\|x^{k+1} - x^k\| / \|x^{k+1}\| \leq \delta_x$ ,  $|f(x^{k+1}) - f(x^k)| / |f(x^{k+1})| \leq \delta_y$ ), положить  $x^* \approx x^{k+1}$  и закончить вычисления. Иначе выполнить присваивание  $k = k + 1$  и вернуться на шаг 2.

*Пример.* Решим задачу определения наилучшего положения приемной антенны:

$$f(x) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min,$$

$$g_1(x) = 4 - x_1^2 - x_2^2 \geq 0,$$

$$g_2(x) = x_1^2 - x_2 \geq 0.$$

Зададим  $x^0 = (1, 5; 1)^T \in S$ ,  $\varepsilon_x = \varepsilon_y = 0,03$ .

Шаг 1. Выполняем присваивание  $k = 0$ . Вычисляем  $f(x^0) = 70,25$ .

Шаг 2. Вычисляем  $g_1(x^0) = 0,75$ ,  $g_2(x^0) = 1,25$ ,

$$\nabla f(x^0) = \begin{pmatrix} 2 \cdot 1,5 - 12 \\ 2 \cdot 1 - 16 \end{pmatrix} = \begin{pmatrix} -9 \\ -14 \end{pmatrix},$$

$$\nabla g_1(x^0) = \begin{pmatrix} -2 \cdot 1,5 \\ -2 \cdot 1 \end{pmatrix} = \begin{pmatrix} -3 \\ -2 \end{pmatrix}, \quad \nabla g_2(x^0) = \begin{pmatrix} 2 \cdot 1,5 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Так как  $\|\nabla f(x^0)\| \approx 16,643 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3. Так как  $g_1(x^0) \neq 0$  и  $g_2(x^0) \neq 0$ , множество активных ограничений пусто, т.е.  $I^0 = \emptyset$ .

Шаг 4. Составим ЗЛП. Так как  $I^0 = \emptyset$ , в ЗЛП будет только одно ограничение, составленное из коэффициентов градиента ЦФ задачи:

$$\begin{aligned} -\theta &\rightarrow \min, \\ -9d_1 - 14d_2 &\leq -\theta, \\ -1 &\leq d_{1,2} \leq 1. \end{aligned}$$

Данную задачу можно решить графически либо симплекс-методом. Но можно также упростить ее решение. На рисунке 6.21 штриховкой помечена допустимая область, задаваемая ограничениями-неравенствами. При построении области  $S$  переменная  $\theta$  рассматривается как константа. При увеличении значения  $\theta$  (мы минимизируем  $-\theta$ , следовательно, максимизируем  $\theta$ ) линия ограничения  $-9d_1 - 14d_2 \leq -\theta$  будет перемещаться в направлении, указанном стрелкой. Точка  $A(1;1)$  является граничной точкой совместности ограничений. То есть при максимальном значении  $\theta$  ОДР задачи будет состоять только из одной точки  $A$ .

Таким образом, решением задачи будет  $d^0 = (1,1)^T$ , при этом  $\theta^0 = 23$ . Параметр  $\theta^0 > 0$ , следовательно, возможно дальнейшее улучшение решения.

Шаг 5. Теперь на луче  $x^1 = x^0 + \alpha d^0$ , т.е.

$$x^1 = \begin{pmatrix} 1,5 \\ 1 \end{pmatrix} + \alpha \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1,5 + \alpha \\ 1 + \alpha \end{pmatrix}, \quad \alpha > 0,$$

требуется найти точку, в которой он пересекает границу области  $S$ . Проверяем первое ограничение:

$$g_1(\alpha) = 4 - (1,5 + \alpha)^2 - (1 + \alpha)^2 = 0 \Rightarrow -2\alpha^2 - 5\alpha + 0,75 = 0, \quad \alpha \geq 0.$$

Решение данного квадратного уравнения дает один положительный корень

$$\alpha = \frac{5 - \sqrt{31}}{-4} \approx 0,142$$

и один отрицательный, который мы сразу отбрасываем. То есть при движении вдоль луча  $x^1 = x^0 + \alpha d^0$  при значении  $\bar{\alpha}_1 \approx 0,142$  мы достигнем границы ОДР, определяемой ограничением  $g_1(x) \geq 0$ . Проверяем второе ограничение:

$$g_2(\alpha) = (1,5 + \alpha)^2 - (1 + \alpha) = 0 \Rightarrow \alpha^2 + 2\alpha + 1,25 = 0, \alpha \geq 0.$$

Получаем  $g_2(\alpha) \geq 0$  для любого  $\alpha \geq 0$ . То есть при движении вдоль луча  $x^1 = x^0 + \alpha d^0$  мы не можем достичь границы ОДР, определяемой ограничением  $g_2(x) \geq 0$ . Следовательно,  $\bar{\alpha}_2 = \infty$ . Находим  $\bar{\alpha}$ :

$$\bar{\alpha} = \min\{\bar{\alpha}_1, \bar{\alpha}_2\} \approx 0,142.$$

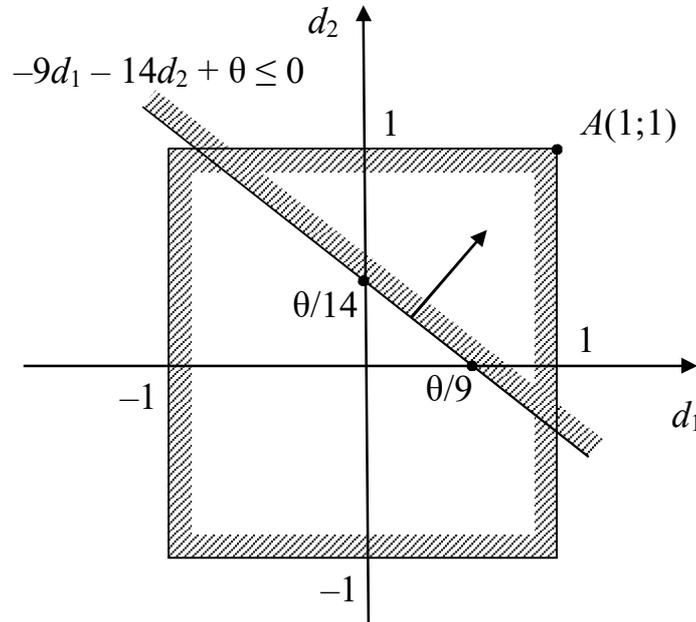


Рисунок 6.21 – Область допустимых решений задачи оптимизации при  $k = 0$

Шаг 6. Исследуем отрезок  $0 \leq \alpha \leq 0,142$  для определения экстремальной точки функции

$$\Phi(\alpha) = 1 + (1,5 + \alpha - 6)^2 + (1 + \alpha - 8)^2 = 1 + (\alpha - 4,5)^2 + (\alpha - 7)^2.$$

Дифференцируем:

$$\frac{d\Phi(\alpha)}{d\alpha} = 2(\alpha - 4,5) + 2(\alpha - 7) = 0 \Rightarrow 4\alpha = 23.$$

Получим  $\alpha = 5,75$ . Так как  $\alpha > 0,142$ , выбирается значение  $\alpha^0 \approx 0,142$ .

*Замечание.* В любой машинной программе поиск граничного значения и одномерный поиск экстремума можно реализовать с помощью известных методов одномерной оптимизации.

Шаг 7. Находим новое приближение оптимума:

$$x^1 = x^0 + \alpha^0 d^0 \approx \begin{pmatrix} 1,5 \\ 1 \end{pmatrix} + 0,142 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \approx \begin{pmatrix} 1,642 \\ 1,142 \end{pmatrix}.$$

Шаг 8. Вычисляем  $f(x^1) \approx 67,026$ . Условия останова итерационного процесса не выполняются:

$$\|x^1 - x^0\| \approx 0,201 > \varepsilon_x, \quad |f(x^1) - f(x^0)| \approx 3,224 > \varepsilon_y,$$

поэтому полагаем  $k = 1$  и возвращаемся на шаг 2.

Шаг 2<sup>(1)</sup>. Вычисляем  $g_1(x^1) = 0$ ,  $g_2(x^1) \approx 1,554$ ,

$$\nabla f(x^1) \approx \begin{pmatrix} -8,716 \\ -13,716 \end{pmatrix}, \quad \nabla g_1(x^1) \approx \begin{pmatrix} -3,284 \\ -2,284 \end{pmatrix}, \quad \nabla g_2(x^0) \approx \begin{pmatrix} 3,284 \\ -1 \end{pmatrix}.$$

Так как  $\|\nabla f(x^1)\| \approx 16,251 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3<sup>(1)</sup>. Так как  $g_1(x^1) = 0$  и  $g_2(x^1) \neq 0$ , первое ограничение является активным, т.е.  $I^1 = \{1\}$ .

Шаг 4<sup>(1)</sup>. Составим ЗЛП. Так как  $I^1 = \{1\}$ , в ЗЛП будут два ограничения, составленные из коэффициентов градиента ЦФ и градиента ограничения  $g_1(x)$ :

$$\begin{aligned}
-\theta &\rightarrow \min, \\
-8,716d_1 - 13,716d_2 &\leq -\theta, \\
-3,284d_1 - 2,284d_2 &\geq \theta, \\
-1 &\leq d_{1,2} \leq 1.
\end{aligned}$$

На рисунке 6.22 показаны оба ограничения, направление их действия, а также координаты точек их пересечения с границами квадрата  $-1 \leq d_{1,2} \leq 1$ . Стрелками показано направление сдвига ограничений при увеличении значения параметра  $\theta$ . ОДР представляет собой треугольник, заключенный между точкой  $A$  (граничной точкой совместности ограничений) и точками пересечения ограничений с границами квадрата  $-1 \leq d_{1,2} \leq 1$  во втором квадранте. Точка  $A$  имеет координаты  $(-0,637 \cdot \theta; 0,477 \cdot \theta)$  – их можно найти, решая СЛАУ, составленную из ограничений ЗЛП, считая параметр  $\theta$  константой:

$$\begin{cases}
-8,716d_1 - 13,716d_2 = -\theta, \\
-3,284d_1 - 2,284d_2 = \theta.
\end{cases}$$

При увеличении значения параметра  $\theta$  и перемещении ограничений в направлении стрелок мы достигнем момента, когда ОДР будет состоять только из одной точки  $A$ . Так как координата  $A_x$  имеет большее (по модулю) значение, чем координата  $A_y$ , она достигнет границ квадрата  $-1 \leq d_{1,2} \leq 1$  раньше, поэтому

$$\theta^1 \approx \frac{-1}{-0,637} \approx 1,571 \Rightarrow d^1 \approx \begin{pmatrix} -0,637 \cdot 1,571 \\ 0,477 \cdot 1,571 \end{pmatrix} \approx \begin{pmatrix} -1 \\ 0,75 \end{pmatrix}.$$

Параметр  $\theta^1 > 0$ , следовательно, возможно дальнейшее улучшение решения.

Шаг 5<sup>(1)</sup>. Теперь на луче  $x^2 = x^1 + \alpha d^1$ , т.е.

$$x^2 \approx \begin{pmatrix} 1,642 \\ 1,142 \end{pmatrix} + \alpha \cdot \begin{pmatrix} -1 \\ 0,75 \end{pmatrix} \approx \begin{pmatrix} 1,642 - \alpha \\ 1,142 + 0,75\alpha \end{pmatrix}, \alpha > 0,$$

требуется найти точку, в которой он пересекает границу области  $S$ . Проверяем первое ограничение:

$$g_1(\alpha) \approx 4 - (1,642 - \alpha)^2 - (1,142 + 0,75\alpha)^2 = 0 \Rightarrow$$

$$-1,563\alpha^2 - 1,571\alpha - 0,000328 = 0, \alpha \geq 0.$$

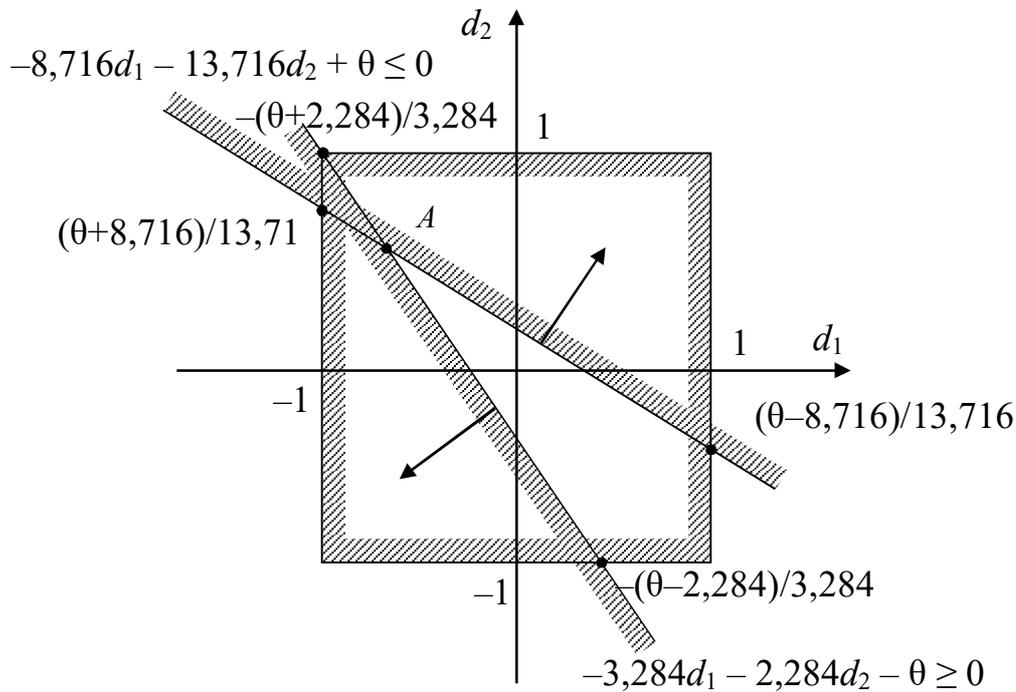


Рисунок 6.22 – Область допустимых решений задачи оптимизации при  $k = 1$

Решение данного квадратного уравнения дает один положительный корень  $\bar{\alpha}_1 \approx 1,005$ . Проверяем второе ограничение:

$$g_2(\alpha) \approx (1,642 - \alpha)^2 - (1,142 + 0,75\alpha) = 0 \Rightarrow$$

$$\alpha^2 - 4,034\alpha + 1,554 = 0, \alpha \geq 0.$$

Решение данного квадратного уравнения дает один положительный корень  $\bar{\alpha}_2 \approx 0,431$ . Определяем  $\bar{\alpha}$ :

$$\bar{\alpha} = \min\{\bar{\alpha}_1, \bar{\alpha}_2\} \approx 0,431.$$

Шаг 6<sup>(1)</sup>. Исследуем отрезок  $0 \leq \alpha \leq 0,431$  для определения экстремальной точки функции

$$\Phi(\alpha) \approx 1 + (1,642 - \alpha - 6)^2 + (1,142 + 0,75\alpha - 8)^2 \approx$$

$$\approx 1 + (-\alpha - 4,358)^2 + (0,75\alpha - 6,858)^2.$$

Дифференцируем:

$$\frac{d\Phi(\alpha)}{d\alpha} = -2(-\alpha - 4,358) + 1,5(0,75\alpha - 6,858) = 0 \Rightarrow -3,125\alpha = -1,571.$$

Получим  $\alpha = 0,503$ . Так как  $\alpha > 0,431$ , выбираем значение  $\alpha^0 \approx 0,431$ .

Шаг 7<sup>(1)</sup>. Находим новое приближение оптимума:

$$x^2 = x^1 + \alpha^1 d^1 \approx \begin{pmatrix} 1,642 \\ 1,142 \end{pmatrix} + 0,431 \cdot \begin{pmatrix} -1 \\ 0,75 \end{pmatrix} \approx \begin{pmatrix} 1,211 \\ 1,465 \end{pmatrix}.$$

Шаг 8<sup>(1)</sup>. Вычисляем  $f(x^2) \approx 66,639$ . Условия останова итерационного процесса не выполняются:

$$\|x^2 - x^1\| \approx 0,539 > \varepsilon_x, \quad |f(x^2) - f(x^1)| \approx 0,387 > \varepsilon_y,$$

поэтому полагаем  $k = 2$  и возвращаемся на шаг 2.

Остальные итерации решения приведены в таблице 6.6.

Таблица 6.6 – Решение задачи методом допустимых направлений

$k$	$x^k$	$f(x^k)$	$I^k$ $d^k$ $\theta^k$	$\bar{\alpha}_1$ $\bar{\alpha}_2$ $\alpha$ $\bar{\alpha}$	$\nabla f(x^k)$			$\ \nabla f(x^k)\ $ $\ x^{k+1} - x^k\ $ $ f(x^{k+1}) - f(x^k) $
		$g_1(x^k)$			$\nabla g_1(x^k)$	$\nabla g_2(x^k)$	$g_2(x^k)$	
1	2	3	4	5	6	7	8	9
0	1,5 1	70,25	$\emptyset$	0,142	-9	-3	3	16,64
		0,75	1	$\infty$	-14	-2	-1	0,201
		1,25	1	5,750				3,224
			23	0,142				
1	1,642 1,142	67,03	{1}	1,005	-8,716	-	3,28	16,25
		0	-1	0,431	-13,72	3,284	4	0,539
		1,554	0,75	0,503		-	-1	0,387
			1,571	0,431		2,284		

Окончание таблицы 6.6

1	2	3	4	5	6	7	8	9
2	1,211	66,64	{2}	0,297	-9,579	-2,421	2,42	16,20
	1,465	0,387	1	$\infty$	-13,07	-2,931	1	0,333
		0	-0,509	1,164			-1	0,759
			2,930	0,297				
3	1,507	65,88	{1}	0,668	-8,985	-3,015	3,015	16,11
	1,314	0	-1	0,274	-13,37	-2,629	-1	0,343
		0,958	0,75	0,334				0,169
			1,043	0,274				
4	1,233	65,71	{2}	0,151	-9,534	-2,466	2,466	16,09
	1,520	0,169	1	$\infty$	-12,96	-3,041	-1	0,169
		0	-0,506	1,183				0,420
			2,972	0,151				
5	1,384	65,29	{1}	0,385	-9,232	-2,768	2,768	16,04
	1,444	0	-1	0,140	-13,11	-2,888	-1	0,175
		0,472	0,75	0,193				0,0536
			0,602	0,140				
6	1,244	65,24	{2}	0,0540	-9,511	-2,489	2,489	16,03
	1,549	0,0536	1	$\infty$	-12,90	-3,097	-1	0,0606
		0	-0,505	1,193				0,158
			2,994	0,0540				
7	1,289	65,08	{1}	0,202	-9,403	-2,597	2,597	16,01
	1,521	0	-1	0,0500	-12,96	-3,043	-1	0,0625
		0,165	0,75	0,101				0,0118
			0,315	0,0500				
8	1,248	65,07	{2}	0,0126	-9,503	-2,497	2,497	16,01
	1,559	0,0118	1	$\infty$	-12,88	-3,117	-1	0,0141
		0	-0,505	1,196				0,0376
			3,002	0,0126				
9	1,261	65,03	{1}	0,125	-9,478	-2,522	2,522	16,00
	1,552	0	-1	0,0117	-12,90	-3,105	-1	0,0146
		0,0380	0,75	0,0620				0,00204
			0,194	0,125				
10	1,249	65,03						
	1,561	0,002						
		0						

С учетом того, что  $\varepsilon_x = \varepsilon_y = 0,03$ , округляем значения аргумента и функции до двух цифр после запятой:  $x^* \approx x^{10} \approx (1,25; 1,56)^T$ ,  $f^* \approx f(x^{10}) \approx 65,03$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 8,064$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 15,38 \text{ Вт/м}^2.$$

Итерационный процесс можно продолжить подобным образом до достижения точки минимума  $x^* \approx (1,25; 1,562)^T$ . Итерации показаны на рисунке 6.23.

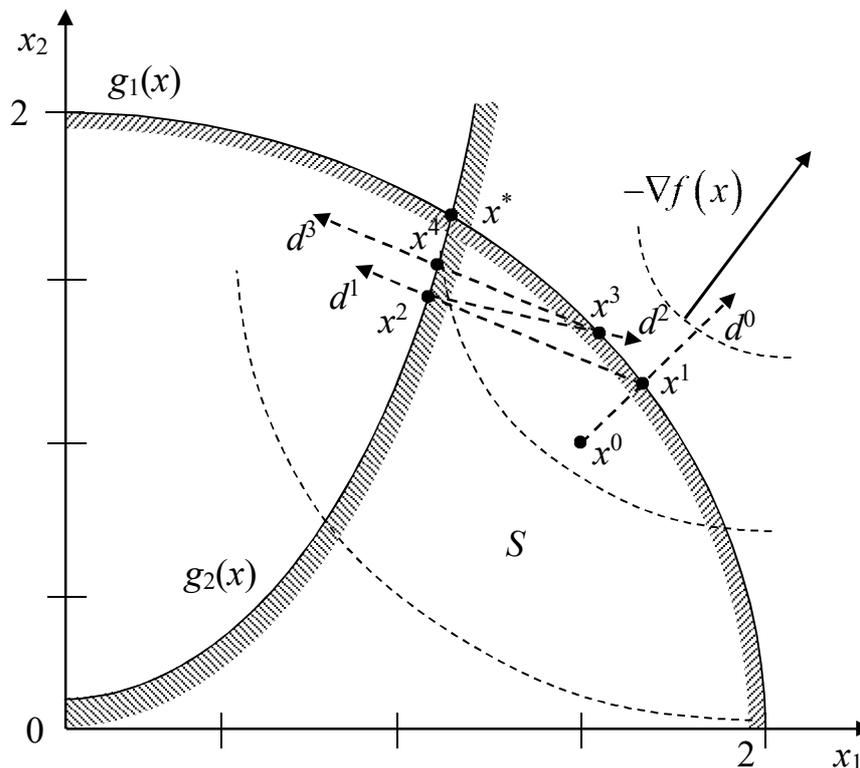


Рисунок 6.23 – Область допустимых решений

В алгоритме Зойтендейка учитываются только активные ограничения в данной допустимой точке, при этом получается зигзагообразный процесс, замедляющий решение, а в некоторых случаях приводящий к «заеданию» – тип ложной сходимости.

К методам, свободным от этих недостатков, можно отнести метод  $\varepsilon$ -возмущений и метод Топкинса и Вейнотта.

Методами допустимых направлений нельзя непосредственно пользоваться для решения задач с нелинейными ограничениями-равенствами  $h_k(x) = 0$ . В этом случае ограничения-равенства необходимо ослабить, допустив ограниченное перемещение вне поверхности, задаваемой ограничениями

$$|h_k(x)| \leq \varepsilon.$$

Однако если значение  $\varepsilon$  мало, то одномерный поиск осуществляется небольшими шагами и скорость процесса невелика. С другой стороны, если допускается ограниченное движение вне границ области допустимых решений  $S$ , облегчающее выбор допустимого направления, то итерационное решение соответствующих ограничениям уравнений нужно проецировать на область  $S$  (рисунок 6.24).

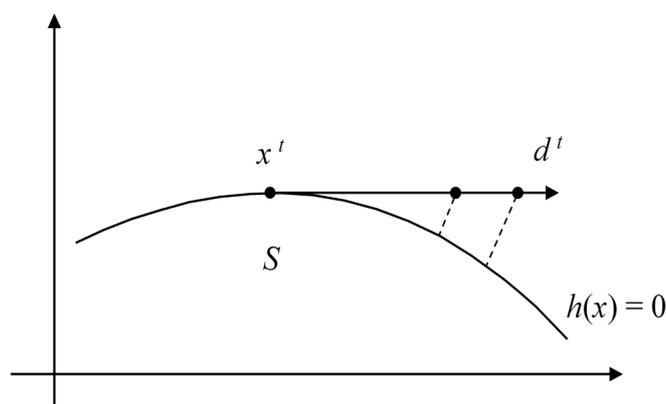


Рисунок 6.24 – Проекция на область допустимых решений задачи оптимизации

Другим существенным недостатком метода допустимых направлений является необходимость решения подзадач ЛП.

Ниже рассматриваются методы, которые не имеют указанных недостатков.

#### 6.6.4. Метод условного градиента

В методе условного градиента (МУГ) рассматривается ЗНП

$$f(x) \rightarrow \min, \quad (6.65)$$

$$g_i(x) \leq 0, \quad i = 1, 2, \dots, m, \quad (6.66)$$

$$x_j \geq 0, \quad j = 1, 2, \dots, n, \quad (6.67)$$

где  $f(x)$  – выпуклая функция.

Пусть  $x^k \in S$  – очередное приближение исходной задачи НП и  $\nabla f(x^k) \neq 0$ . Тогда в окрестности точки  $x^k$  ЦФ  $f(x)$  представима в виде

$$\tilde{f}(x, x^k) = f(x^k) + \nabla f^T(x^k) \cdot (x - x^k) + R_2(x, x^k)$$

и линейная функция

$$f_k(x) = \nabla f^T(x^k) \cdot (x - x^k)$$

является приближением разности  $\tilde{f}(x, x^k) - f(x^k)$  с точностью до величины  $R_2(x, x^k)$  в некоторой окрестности точки  $x^k$ .

Поставим вспомогательную задачу минимизации на множестве  $S$  линейной функции  $f_k(x)$ , т.е.

$$f_k(x) = \nabla f^T(x^k) \cdot (x - x^k) \rightarrow \min \quad (6.68)$$

при тех же ограничениях (6.66), (6.67).

Пусть  $y^k$  – решение этой задачи. Следующее приближение  $x^{k+1}$  к точке минимума  $x^*$  исходной ЦФ  $f(x)$  на множестве  $S$  найдем по формуле

$$x^{k+1} = x^k + \alpha_k (y^k - x^k), \quad \alpha_k \in [0, 1]. \quad (6.69)$$

В силу выпуклости множества  $S$  следует, что  $x^{k+1} \in S$ . Величина  $\alpha_k$  из выражения (6.69) может вычисляться различными способами. Например,

$$\alpha_k = \min \{1, \alpha_k^*\},$$

где  $\alpha_k^*$  находится из условия наискорейшего спуска по направлению  $d_k = y^k - x^k$ :

$$\Phi(\alpha_k^*) = \min_{\alpha} f\left(x^k + \alpha(y^k - x^k)\right).$$

Существует другой способ вычисления значения  $\alpha_k$ . В начале выполнения итерации (6.69) полагают  $\alpha_k = 1$ , после чего проверяют условие

$$f(x^{k+1}) < f(x^k). \quad (6.70)$$

Если это условие нарушается, то  $\alpha_k$  уменьшают в 2 раза (до тех пор, пока неравенство (6.70) не будет выполнено) и переходят к следующей итерации (6.69).

*Алгоритм условного градиента.* Исходные данные: ЦФ, ограничения, начальная точка  $x^0 \in S$ , точность решения по аргументу и по функции – абсолютная ( $\varepsilon_x$  и  $\varepsilon_y$ ) или относительная ( $\delta_x$  и  $\delta_y$ ).

1. Задать номер итерации  $k = 0$ . Вычислить  $f(x^0)$ .

2. Вычислить  $\nabla f(x^k)$ . Если  $\|\nabla f(x^k)\| \leq \varepsilon_y$ , то положить  $x^* \approx x^k$

и закончить вычисления.

3. Решить вспомогательную задачу

$$f_k(x) = \nabla f^T(x^k) \cdot (x - x^k) \rightarrow \min$$

при ограничениях (6.66), (6.67). Пусть  $y^k$  – оптимальное решение этой задачи.

4. Найти  $\alpha^k$ , представляющее собой решение задачи при

$$\Phi(\alpha) = f\left(x^k + \alpha(y^k - x^k)\right) \rightarrow \min_{\alpha}, \quad 0 \leq \alpha \leq 1.$$

5. Вычислить  $x^{k+1} = x^k + \alpha^k(y^k - x^k)$ .

6. Проверить близость к решению. Если  $\|x^{k+1} - x^k\| \leq \varepsilon_x$  и  $|f(x^{k+1}) - f(x^k)| \leq \varepsilon_y$  (если заданы относительные погрешности, то

$\|x^{k+1} - x^k\| / \|x^{k+1}\| \leq \delta_x$ ,  $|f(x^{k+1}) - f(x^k)| / |f(x^{k+1})| \leq \delta_y$ , положить  $x^* \approx x^{k+1}$  и закончить вычисления. Иначе выполнить присваивание  $k = k + 1$  и вернуться на шаг 2.

Отметим, что в общем случае задача (6.68) является, вообще говоря, задачей нелинейного программирования ( $g_i(x)$  – нелинейные функции). Укажем случаи, когда поиск решения  $y^k$  не представляет затруднений.

1. Допустимое множество  $S = \{x \in R^n \mid a_j \leq x_j \leq b_j, j = 1, 2, \dots, n\}$  является  $n$ -мерным параллелепипедом. Тогда

$$y_j^k = \begin{cases} a_j, & \text{если } \partial f(x^k) / \partial x_j > 0, \\ b_j, & \text{если } \partial f(x^k) / \partial x_j < 0, \\ (a_j + b_j) / 2, & \text{если } \partial f(x^k) / \partial x_j = 0. \end{cases} \quad (6.71)$$

Покажем это. Пусть  $\nabla f(x^k) = g$ ,  $\nabla f^T(x^k) \cdot x^k = d$ . В этом случае имеем

$$\begin{cases} g^T x - d \rightarrow \min, \\ x_j - x_{n+j} = a_j, \\ x_j + x_{2n+j} = b_j, \\ j = 1, 2, \dots, n \end{cases}$$

Это ЗЛП. Пусть  $n = 2$ . Найдем начальную угловую точку. Имеем

$$x_1 - x_3 = b_1, \quad x_2 - x_4 = b_2, \quad x_1 + x_5 = a_1, \quad x_2 + x_6 = a_2.$$

Составим симплекс-таблицу:

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	
1	0	-1	0	0	0	$a_1$
0	1	0	-1	0	0	$a_2$
1	0	0	0	1	0	$b_1$
0	1	0	0	0	1	$b_2$

Выделим базис:

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	
1	0	-1	0	0	0	$a_1$
0	1	0	-1	0	0	$a_2$
0	0	1	0	1	0	$b_1 - a_1$
0	0	0	1	0	1	$b_2 - a_2$

Таким образом, получили начальное базисное решение  $x^0 = (b_1, b_2; 0, 0; b_1 - a_1; b_2 - a_2)^T$ . Подставим в ЦФ это значение, получим  $f(x) = g_1 x_3 + g_2 x_4 + p_0$ , где  $p_0 = g_1 b_1 + g_2 b_2 - d$ . Если  $g_1, g_2 > 0$ , то в точке  $x^0$  достигнут минимум, т.е.  $x_1^* = a_1$ ,  $x_2^* = a_2$ . Если  $g_1$  или  $g_2 < 0$ , то решаем задачу дальше:

	$x_3$	$x_4$	
$x_1$	-1	0	$a_1$
$x_2$	0	-1	$a_2$
$x_5$	1	0	$b_1 - a_1$
$x_6$	0	1	$b_2 - a_2$
	$g_1$	$g_2$	$-p_0$

Если  $g_1 < 0$ , то разрешающим столбцом будет первый столбец. В этом столбце разрешающая строка третья. После преобразования получим

$$\begin{aligned} x_1 &= a_1 + (b_1 - a_1) = b_1, \\ x_2 &= b_2. \end{aligned}$$

Если  $g_2 < 0$ , то аналогично получим

$$\begin{aligned} x_2 &= a_2 + (b_2 - a_2) = b_2, \\ x_1 &= b_1. \end{aligned}$$

2. Допустимое множество  $S$  задано линейными ограничениями и условием неотрицательности переменных. Тогда задача (6.68) будет ЗЛП и ее решение можно найти с помощью симплекс-метода. В этом случае МУГ напоминает МФВ.

3. Допустимое множество

$$S = \left\{ x \in R^n \mid \sum (x_j - a_j)^2 \leq R_0^2 \right\}, \quad (6.72)$$

то есть  $S$  – шар радиусом  $R_0$  с центром в точке  $a$ . Тогда

$$y^k = a - R_0 \frac{\nabla f(x^k)}{\|\nabla f(x^k)\|}. \quad (6.73)$$

Покажем это. Пусть  $\nabla f(x^k) = g$ ,  $\nabla f^T(x^k) \cdot x^k = d$ . Тогда имеем следующую задачу:

$$\begin{cases} f(x) = g^T x - d \rightarrow \min, \\ \sum (x_j - a_j)^2 \leq R_0^2. \end{cases}$$

Составляем функцию Лагранжа  $L$ :

$$L = g^T x - d + \lambda \cdot \left[ \sum (x_j - a_j)^2 + b^2 - R_0^2 \right] \rightarrow \min.$$

Здесь слагаемое  $b^2$  добавлено для приведения неравенства к равенству: при  $b^2 \geq 0$  имеем  $\sum (x_j - a_j)^2 + b^2 = R_0^2$ . Ищем минимум:

$$\begin{cases} \frac{\partial L}{\partial x_j} = g_j + 2\lambda(x_j - a_j) = 0, \\ \frac{\partial L}{\partial y} = 2\lambda b = 0 \Rightarrow b = 0, \\ \frac{\partial L}{\partial \lambda} = \sum (x_j - a_j)^2 + y^2 - R_0^2 = 0, \end{cases}$$

$$x_j - a_j = -\frac{1}{2\lambda} g_j, \quad (2\lambda)^2 = \frac{\sum g_j^2}{\sum (x_j - a_j)^2},$$

$$\frac{1}{2\lambda} = \frac{R_0}{\sqrt{\sum g_j^2}}, \quad x_j = a_j - g_j \frac{R_0}{\sqrt{\sum g_j^2}}$$

или

$$y^k = a - \nabla f(x^k) \frac{R_0}{\|\nabla f(x^k)\|}.$$

В этом случае МУГ напоминает метод проекции градиента, который будет рассмотрен далее.

*Пример 1.* Решим задачу определения наилучшего положения приемной антенны с ограничениями в виде прямоугольной области (6.20):

$$f(x) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ 0 \leq x_1 \leq 10, \quad 0 \leq x_2 \leq 2.$$

Зададим  $x^0 = (1; 1)^T \in S$ ,  $\varepsilon_x = \varepsilon_y = 0,1$ .

Шаг 1. Выполняем присваивание  $k = 0$ . Вычисляем  $f(x^0) = 75$ .

Шаг 2. Вычисляем

$$\nabla f(x^0) = \begin{pmatrix} 2 \cdot 1 - 12 \\ 2 \cdot 1 - 16 \end{pmatrix} = \begin{pmatrix} -10 \\ -14 \end{pmatrix}.$$

Так как  $\|\nabla f(x^0)\| \approx 17,205 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3. Составим вспомогательную задачу:

$$-10(x_1 - 1) - 14(x_2 - 1) \rightarrow \min, \\ 0 \leq x_1 \leq 10, \quad 0 \leq x_2 \leq 2.$$

Данную задачу можно решить графически либо симплекс-методом. Но, как было показано выше, проще воспользоваться соотношением (6.71):

$$\nabla f(x^0) = \begin{pmatrix} -10 \\ -14 \end{pmatrix} \Rightarrow y^0 = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 2 \end{pmatrix}.$$

Шаг 4. Выполним одномерный поиск вдоль прямой:

$$\Phi(\alpha) = f\left(x^0 + \alpha(y^0 - x^0)\right) \rightarrow \min_{\alpha}, \alpha \in [0,1] \Rightarrow$$

$$f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha\left(\begin{pmatrix} 10 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)\right) = f\left(\begin{pmatrix} 1+9\alpha \\ 1+\alpha \end{pmatrix}\right) = 1 + (1+9\alpha-6)^2 + (1+\alpha-8)^2 \rightarrow \min \Rightarrow$$

$$\Phi(\alpha) = 1 + (9\alpha-5)^2 + (\alpha-7)^2 \rightarrow \min.$$

Дифференцируем:

$$\frac{d\Phi(\alpha)}{d\alpha} = 18(9\alpha-5) + 2(\alpha-7) = 0, \alpha \in [0,1] \Rightarrow$$

$$\Rightarrow 164\alpha = 104 \Rightarrow \alpha^0 \approx 0,634.$$

Шаг 5. Находим новое приближение оптимума:

$$x^1 = x^0 + \alpha^0(y^0 - x^0) = \begin{pmatrix} 1+9\alpha \\ 1+\alpha \end{pmatrix} \approx \begin{pmatrix} 6,707 \\ 1,634 \end{pmatrix}.$$

Шаг 6. Вычисляем  $f(x^1) \approx 42,02$ . Условия останова итерационного процесса не выполняются:

$$\|x^1 - x^0\| \approx 5,742 > \varepsilon_x, \quad |f(x^1) - f(x^0)| \approx 32,98 > \varepsilon_y,$$

поэтому полагаем  $k = 1$  и возвращаемся на шаг 2.

Шаг 2<sup>(1)</sup>. Вычисляем

$$\nabla f(x^1) \approx \begin{pmatrix} 2 \cdot 6,707 - 12 \\ 2 \cdot 1,634 - 16 \end{pmatrix} \approx \begin{pmatrix} 1,415 \\ -12,73 \end{pmatrix}.$$

Так как  $\|\nabla f(x^1)\| \approx 12,81 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3<sup>(1)</sup>. Составим вспомогательную задачу:

$$1,415(x_1 - 6,707) - 12,73(x_2 - 1,634) \rightarrow \min,$$

$$0 \leq x_1 \leq 10, \quad 0 \leq x_2 \leq 2.$$

Решаем ее:

$$\nabla f(x^1) \approx \begin{pmatrix} 1,415 \\ -12,73 \end{pmatrix} \Rightarrow y^0 = \begin{pmatrix} a_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

Шаг 4<sup>(1)</sup>. Выполним одномерный поиск вдоль прямой:

$$\begin{aligned}\Phi(\alpha) &= f\left(x^1 + \alpha(y^1 - x^1)\right) \rightarrow \min_{\alpha}, \alpha \in [0, 1] \Rightarrow \\ f\left(\begin{pmatrix} 6,707 \\ 1,634 \end{pmatrix} + \alpha\left(\begin{pmatrix} 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 6,707 \\ 1,634 \end{pmatrix}\right)\right) &\approx f\left(\begin{pmatrix} 6,707 - 6,707\alpha \\ 1,634 + 0,366\alpha \end{pmatrix}\right) \approx \\ &\approx 1 + (6,707 - 6,707\alpha - 6)^2 + (1,634 + 0,366\alpha - 8)^2 \rightarrow \min \Rightarrow \\ \Phi(\alpha) &\approx 1 + (-6,707\alpha + 0,707)^2 + (0,366\alpha - 6,366)^2 \rightarrow \min.\end{aligned}$$

Дифференцируем:

$$\begin{aligned}\frac{d\Phi(\alpha)}{d\alpha} &\approx -13,414(-6,707\alpha + 0,707) + 0,732(0,366\alpha - 6,366) = 0, \alpha \in [0, 1] \Rightarrow \\ &\Rightarrow 90,24\alpha = 14,14 \Rightarrow \alpha^1 \approx 0,157.\end{aligned}$$

Шаг 5<sup>(1)</sup>. Находим новое приближение оптимума:

$$x^2 = x^1 + \alpha^1(y^1 - x^1) \approx \begin{pmatrix} 6,707 - 6,707\alpha \\ 1,634 + 0,366\alpha \end{pmatrix} \approx \begin{pmatrix} 5,656 \\ 1,691 \end{pmatrix}.$$

Шаг 6<sup>(1)</sup>. Вычисляем  $f(x^2) \approx 40,92$ . Условия останова итерационного процесса не выполняются:

$$\|x^2 - x^1\| \approx 1,053 > \varepsilon_x, \quad |f(x^2) - f(x^1)| \approx 1,109 > \varepsilon_y,$$

поэтому полагаем  $k = 2$  и возвращаемся на шаг 2.

Результаты вычислений на следующих итерациях приведены в таблице 6.7.

С учетом того что  $\varepsilon_x = \varepsilon_y = 0,1$ , округляем значения аргумента и функции до одной цифры после запятой:  $x^* \approx x^{44} \approx (6,0; 2,0)^T$ ,  $f^* \approx f(x^{44}) \approx 37,5$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 6,124$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 26,67 \text{ Вт/м}^2.$$

Как видно, метод сходится весьма медленно к точке оптимального решения  $x^* = (6; 2)^T$ . Хотя для других начальных точек, например  $x^0 = (2; 2)^T$ , точное решение находится всего за одну итерацию.

Таблица 6.7 – Решение задачи методом условного градиента, пример 1

$k$	$x^k$	$f(x^k)$	$\nabla f(x^k)$	$y^k$	$\alpha^k$	$\ \nabla f(x^k)\ $ $\ x^{k+1} - x^k\ $ $ f(x^{k+1}) - f(x^k) $
0	1 1	75	-10 -14	10 2	0,634	17,21 5,742 32,98
1	6,707 1,634	42,02	1,415 -12,73	0 2	0,157	12,81 1,053 1,109
2	5,656 1,691	40,92	-0,688 -12,62	10 2	0,181	12,64 0,790 0,624
3	6,444 1,747	40,29	0,888 -12,51	0 2	0,107	12,54 0,689 0,474
4	5,756 1,774	39,82	-0,488 -12,45	10 2	0,135	12,46 0,574 0,330
...	...	...	...	...	...	...
43	6,060 1,961	37,47	0,119 -12,08	0 2	0,0163	12,08 0,0985 0,0097
44	5,961 1,962	37,46				

*Пример 2.* Решим задачу определения наилучшего положения приемной антенны с ограничениями (6.15), (6.16), (6.17):

$$\begin{aligned} f(x) &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ g(x) &= 4 - x_1 - x_2 \geq 0, \\ x_{1,2} &\geq 0. \end{aligned}$$

Зададим  $x^0 = (1; 1)^T \in S$ ,  $\varepsilon_x = \varepsilon_y = 0,1$ .

Шаг 1. Выполняем присваивание  $k = 0$ . Вычисляем  $f(x^0) = 75$ .

Шаг 2. Вычисляем

$$\nabla f(x^0) = \begin{pmatrix} 2 \cdot 1 - 12 \\ 2 \cdot 1 - 16 \end{pmatrix} = \begin{pmatrix} -10 \\ -14 \end{pmatrix}.$$

Так как  $\|\nabla f(x^0)\| \approx 17,205 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3. Составим вспомогательную задачу:

$$\begin{aligned} -10(x_1 - 1) - 14(x_2 - 1) &\rightarrow \min, \\ g(x) &= 4 - x_1 - x_2 \geq 0, \\ x_{1,2} &\geq 0. \end{aligned}$$

Это ЗЛП, которую можно решить графически или симплекс-методом. Но, как и ранее в решении данной задачи методом Франка – Вульфа, очевидно, что оптимальным будет решение с максимально возможным значением компонента  $x_2$ , поэтому  $y^0 = (0; 4)^T$ .

Шаг 4. Выполним одномерный поиск вдоль прямой:

$$\begin{aligned} \Phi(\alpha) &= f\left(x^0 + \alpha(y^0 - x^0)\right) \rightarrow \min_{\alpha}, \alpha \in [0, 1] \Rightarrow \\ f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha\left(\begin{pmatrix} 0 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)\right) &= f\begin{pmatrix} 1 - \alpha \\ 1 + 3\alpha \end{pmatrix} = 1 + (1 - \alpha - 6)^2 + (1 + 3\alpha - 8)^2 \rightarrow \min \Rightarrow \\ \Phi(\alpha) &= 1 + (-\alpha - 5)^2 + (3\alpha - 7)^2 \rightarrow \min. \end{aligned}$$

Получим результат, аналогичный МФВ:

$$\frac{d\Phi(\alpha)}{d\alpha} = -2(-\alpha - 5) + 6(3\alpha - 7) = 0, \alpha \in [0,1] \Rightarrow$$

$$\Rightarrow 20\alpha = 32 \Rightarrow \alpha^0 = 1.$$

Шаг 5. Находим новое приближение оптимума:

$$x^1 = x^0 + \alpha^0(y^0 - x^0) = \begin{pmatrix} 1 - \alpha \\ 1 + 3\alpha \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

Шаг 6. Вычисляем  $f(x^1) = 53$ . Условия останова итерационного процесса не выполняются:

$$\|x^1 - x^0\| \approx 3,162 > \varepsilon_x, \quad |f(x^1) - f(x^0)| = 22 > \varepsilon_y,$$

поэтому полагаем  $k = 1$  и возвращаемся на шаг 2.

Дальнейшие итерации решения также практически аналогичны результатам, полученным МФВ, сведем их в таблицу 6.8.

Таблица 6.8 – Решение задачи методом условного градиента, пример 2

$k$	$x^k$	$f(x^k)$	$\nabla f(x^k)$	$y^k$	$\alpha^k$	$\ \nabla f(x^k)\ $ $\ x^{k+1} - x^k\ $ $ f(x^{k+1}) - f(x^k) $
0	1	75	-10	0	1	17,21
	1		-14	4		5,742 32,98
1	0	53	-12	4	0,25	14,422
	4		-8	0		1,444 2
2	1	51	-10	1	1	14,142
	3		-10	3		0 0
3	1 3	51				

С учетом того что  $\varepsilon_x = \varepsilon_y = 0,1$ , округляем значения аргумента и функции до одной цифры после запятой:  $x^* = x^3 = (1,0; 3,0)^T$ ,  $f^* = f(x^3) = 51$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 7,141$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 19,61 \text{ Вт/м}^2.$$

*Пример 3.* Решим задачу определения наилучшего положения приемной антенны с ограничением (6.2):

$$\begin{aligned} f(x) &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ g(x) &= 4 - x_1 - x_2 \geq 0. \end{aligned}$$

Зададим  $x^0 = (1; 1)^T \in S$ ,  $\varepsilon_x = \varepsilon_y = 0,1$ .

Шаг 1. Выполняем присваивание  $k = 0$ . Вычисляем  $f(x^0) = 75$ .

Шаг 2. Вычисляем

$$\nabla f(x^0) = \begin{pmatrix} 2 \cdot 1 - 12 \\ 2 \cdot 1 - 16 \end{pmatrix} = \begin{pmatrix} -10 \\ -14 \end{pmatrix}.$$

Так как  $\|\nabla f(x^0)\| \approx 17,205 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3. Составим вспомогательную задачу:

$$\begin{aligned} 4 - (x_1 - 1)^2 - (x_2 - 1)^2 &\rightarrow \min, \\ g(x) = 4 - x_1^2 - x_2^2 &\geq 0. \end{aligned}$$

Данная задача является нелинейной. Но учитывая, что ограничение имеет вид (6.72), где  $R_0 = 2$ ,  $a = (0; 0)^T$ , она легко решается:

$$y^0 = a - R_0 \frac{\nabla f(x^0)}{\|\nabla f(x^0)\|} = -2 \frac{\nabla f(x^0)}{\|\nabla f(x^0)\|} \approx \frac{-2}{17,205} \begin{pmatrix} -10 \\ -14 \end{pmatrix} \approx \begin{pmatrix} 1,162 \\ 1,627 \end{pmatrix}.$$

Шаг 4. Выполняем одномерный поиск вдоль прямой:

$$\begin{aligned}\Phi(\alpha) &= f\left(x^0 + \alpha(y^0 - x^0)\right) \rightarrow \min_{\alpha}, \alpha \in [0, 1] \Rightarrow \\ & f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha\left(\begin{pmatrix} 1,162 \\ 1,627 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)\right) \approx f\left(\begin{pmatrix} 1+0,162\alpha \\ 1+0,627\alpha \end{pmatrix}\right) \approx \\ & \approx 1 + (1+0,162\alpha - 6)^2 + (1+0,627\alpha - 8)^2 \rightarrow \min \Rightarrow \\ & \Phi(\alpha) \approx 1 + (0,162\alpha - 5)^2 + (0,627\alpha - 7)^2 \rightarrow \min.\end{aligned}$$

Дифференцируем:

$$\begin{aligned}\frac{d\Phi(\alpha)}{d\alpha} &\approx 0,324(0,162\alpha - 5) + 1,254(0,627\alpha - 7) = 0, \alpha \in [0, 1] \Rightarrow \\ &\Rightarrow 0,839\alpha = 10,40 \Rightarrow \alpha^0 = 1.\end{aligned}$$

Шаг 5. Находим новое приближение оптимума:

$$x^1 = x^0 + \alpha^0(y^0 - x^0) \approx \begin{pmatrix} 1+0,162\alpha \\ 1+0,627\alpha \end{pmatrix} \approx \begin{pmatrix} 1,162 \\ 1,627 \end{pmatrix}.$$

Шаг 6. Вычисляем  $f(x^1) \approx 65,01$ . Условия останова итерационного процесса не выполняются:

$$\|x^1 - x^0\| \approx 0,648 > \varepsilon_x, \quad |f(x^1) - f(x^0)| \approx 9,989 > \varepsilon_y,$$

поэтому полагаем  $k = 1$  и возвращаемся на шаг 2.

Остальные итерации решения сведем в таблицу 6.9.

С учетом того что  $\varepsilon_x = \varepsilon_y = 0,1$ , округляем значения аргумента и функции до одной цифры после запятой:  $x^* \approx x^2 \approx (1,2; 1,6)^T$ ,  $f^* \approx f(x^2) \approx 65,0$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 8,062$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 15,39 \text{ Вт/м}^2.$$

Таблица 6.9 – Решение задачи методом условного градиента, пример 3

$k$	$x^k$	$f(x^k)$	$\nabla f(x^k)$	$y^k$	$\alpha^k$	$\ \nabla f(x^k)\ $ $\ x^{k+1} - x^k\ $ $ f(x^{k+1}) - f(x^k) $
0	1	75	-10	1,162	1	17,21
	1		-14	1,627		0,648
1	1,162	65,01	-9,675	1,209	1	16,00
	1,627		-12,75	1,593		0,0581
2	1,209	65,00				0,0101
	1,593					

## 6.7. Метод проекции градиента

### 6.7.1. Задачи с линейными ограничениями

Существует две разновидности *метода проекции градиента* (МПГ) – для решения задач с линейными и нелинейными ограничениями.

Рассмотрим ЗО при единственном линейном ограничении в виде равенства

$$f(x) \rightarrow \min, \quad (6.74)$$

$$h(x) = \sum_{j=1}^n a_j x_j = a^T x = b. \quad (6.75)$$

В заданной точке  $x^k$ , в которой  $\nabla f(x^k) \neq 0$ , делается попытка найти направление поиска, которое бы лежало на поверхности ограничения и являлось направлением спуска. Такое направление можно получить геометрически, ортогонально проектируя вектор, противоположный  $\nabla f(x^k)$ , на поверхность ограничения (рисунок 6.25).

На рисунке  $\nabla f_c^k$  – проекция антиградиента на поверхность ограничений, которая приводит в допустимые точки. Действительно,  $\forall \alpha \geq 0$  точки, заданные соотношением

$$x = x^k - \alpha \cdot \nabla f_c^k, \quad (6.76)$$

удовлетворяют линейному ограничению  $a^T x = b$ . Покажем это. Умножим обе части выражения (6.76) слева на  $a^T$ :

$$a^T x = a^T x^k - \alpha \cdot a^T \nabla f_c^k.$$

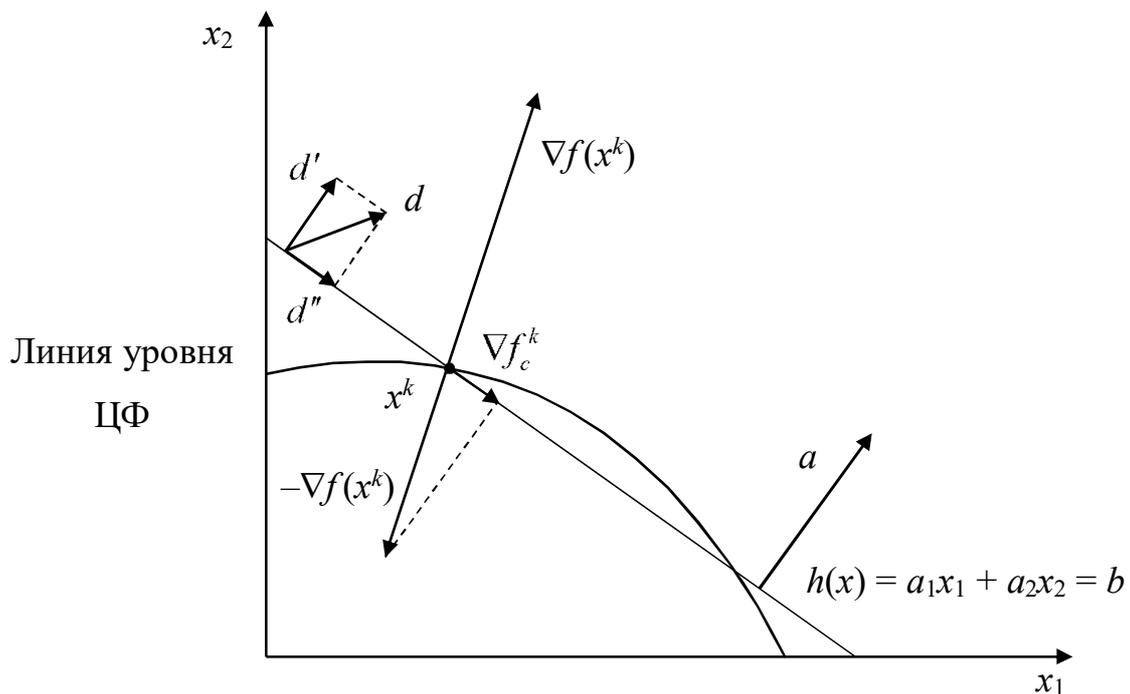


Рисунок 6.25 – Метод проекции градиента

Учитывая, что  $a^T x^k = b$  и  $a^T \nabla f_c^k = 0$  (последнее является следствием того, что  $a \perp \nabla f_c$ ), получим требуемое.

Это направление задает спуск, так как угол между векторами  $\nabla f^k$  и  $-\nabla f_c^k$  больше  $90^\circ$ . Процесс ортогонального проектирования состоит в разложении вектора на две ортогональные компоненты: одна параллельна поверхности, заданной ограничением, вторая перпендикулярна к ней. Параллельная компонента является искомой проекцией градиента.

Пусть вектор  $a$  – нормаль к поверхности ограничения. Отметим, что из выражения  $a^T d = 0$  следует допустимость направления, задаваемого вектором  $d$  (вектор  $d$  параллелен поверхности).

Таким образом, все векторы, перпендикулярные к поверхности ограничения, должны быть параллельны  $a$ . Следовательно, для любого вектора  $d$  его компонента  $d'$ , перпендикулярная к поверхности ограничения, равна значению  $a$ , умноженному на константу.

Обозначим через  $d''$  компоненту  $d$ , параллельную поверхности ограничения. Тогда  $d''$  удовлетворяет соотношению

$$a^T d'' = 0. \quad (6.77)$$

Таким образом, любой вектор можно представить в виде векторной суммы:

$$d = d' + d'', \quad (6.78)$$

где  $d' = \lambda a$ , а  $d''$  удовлетворяет уравнению  $a^T d'' = 0$ .

Найдем коэффициент  $\lambda$ . Рассмотрим скалярное произведение  $a^T d$ . В силу равенств (6.78) и (6.77) имеем

$$a^T d = a^T \lambda a + a^T d'' = \lambda a^T a, \quad (6.79)$$

откуда

$$\lambda = \left(a^T a\right)^{-1} a^T d. \quad (6.80)$$

Из выражения (6.78) найдем  $d'' = d - d' = d - \lambda a$ . Подставим сюда (6.80), получим

$$d'' = d - a \left(a^T a\right)^{-1} a^T d = \left[ E - a \left(a^T a\right)^{-1} a^T \right] d, \quad (6.81)$$

где  $E$  – единичная матрица, порядок которой согласован с  $d$ . Матрица  $P = E - a \left(a^T a\right)^{-1} a^T$  – проекционная матрица. Она проектирует вектор  $d$  на плоскость, задаваемую ограничением  $h(x)$ .

Отметим, что матрица  $P$  является симметрической и положительно полуопределенной. Симметричность матрицы  $P$  очевидна. Для

доказательства положительной полуопределенности рассмотрим произведение  $y^T P y$  для произвольного  $y \neq 0$ . Тогда

$$y^T P y = y^T \left[ E - a(a^T a)^{-1} a^T \right] y = y^T y - \frac{(y^T a)(a^T y)}{a^T a} = \frac{(y^T y)(a^T a) - (y^T a)^2}{a^T a}.$$

Используя неравенство Шварца  $(y^T a)^2 \leq (y^T y) \cdot (a^T a)$ , убеждаемся, что числитель не отрицателен.

*Пример.* Вычислить проекцию вектора  $d = (0; 1)^T$  на поверхность, задаваемую ограничениями  $h(x) = 2x_1 + x_2 = 3$  в точке  $(3/4; 3/2)$  (рисунок 6.26).

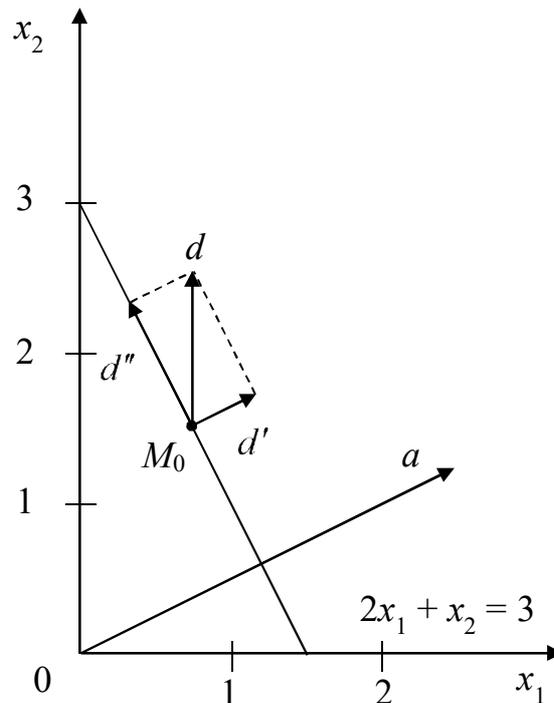


Рисунок 6.26 – Проекция вектора  $d$  на поверхность

Уравнение прямой, проходящей через данную точку  $M_0$  перпендикулярно к данной плоскости, задается уравнением

$$\frac{x - x_0}{A} = \frac{y - y_0}{B} = \frac{z - z_0}{C},$$

где  $Ax + By + Cz + D = 0$  – заданная плоскость;  $M_0(x_0, y_0, z_0)$  – заданная точка. В нашем случае точка  $M_0 = (3/4, 3/2)$ . Плоскость – это прямая, т.е.  $A = 2, B = 1, C = 0, D = -3$ . Таким образом, получим

$$\frac{x_1 - 3/4}{2} = \frac{x_2 - 3/2}{1},$$

откуда следует уравнение прямой  $2x_2 = x_1 + 9/4$ . Перенесем параллельно эту прямую в начало координат, получим  $2x_2 = x_1$ . В качестве нормали возьмем вектор с координатами  $\nabla h(x) = (2; 1)^T$ .

Проекционная матрица задается выражением

$$P = E - a \cdot (a^T a)^{-1} a^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot \frac{1}{2} \cdot (2 \quad 1) = \begin{pmatrix} 0,2 & -0,4 \\ -0,4 & 0,8 \end{pmatrix}.$$

Проекция  $d''$  имеет вид

$$d'' = P \cdot d = \begin{pmatrix} 0,2 & -0,4 \\ -0,4 & 0,8 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -0,4 \\ 0,8 \end{pmatrix}.$$

Проекция  $d' = d - d'' = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} -0,4 \\ 0,8 \end{pmatrix} = \begin{pmatrix} 0,4 \\ 0,2 \end{pmatrix}$ . Очевидно, что векторы  $d'$  и  $d''$  ортогональны, т.е.

$$(d')^T d'' = (0,4 \quad 0,2) \cdot \begin{pmatrix} -0,4 \\ 0,8 \end{pmatrix} = 0,$$

и вектор  $d''$  перпендикулярен вектору  $a$ :

$$a^T d'' = (2 \quad 1) \cdot \begin{pmatrix} -0,4 \\ 0,8 \end{pmatrix} = 0.$$

Способ построения проекционной матрицы  $P$  легко распространяется на случай многих ограничений:

$$h_k(x) = a_k^T x = b_k, \quad k = 1, 2, \dots, m, \quad d' = A^T \lambda,$$

$$d'' = P \cdot d = \left[ E - A^T (AA^T)^{-1} A \right] d,$$

$$\lambda = \left( AA^T \right)^{-1} Ad - \text{вектор},$$

где  $A$  – матрица со строками  $a_k^T$ , т.е.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

где  $a_{i1}, a_{i2}, \dots, a_{in}$  – коэффициенты  $i$ -го ограничения.

Пусть  $d'' = -P \cdot \nabla f(x^k)$  – направление спуска. Если  $d'' = 0$ , то точка  $x^k$  удовлетворяет необходимым условиям Лагранжа. Вектор множителей Лагранжа задается выражением

$$\lambda = \left( AA^T \right)^{-1} A \cdot \nabla f(x). \quad (6.82)$$

Первое утверждение следует из того, что матрица  $P$  симметрическая и положительно полуопределена. Действительно, имеем

$$x^{k+1} = x^k - \alpha P \cdot \nabla f(x^k), \quad \alpha \geq 0;$$

$$f(x^{k+1}) = f(x^k - \alpha P \cdot \nabla f(x^k)) = f(x^k) - \alpha \nabla f^T(x^{k+1}) \cdot P \cdot \nabla f(x^k).$$

Так как  $\nabla f^T(x^{k+1}) \cdot P \cdot \nabla f(x^k) \geq 0$ ,  $\alpha \geq 0$ , то  $f^{k+1} \leq f^k$ , что и требовалось доказать.

Имеем  $d'' = -P \cdot \nabla f$ , т.е.  $d''$  – это проекция вектора  $\nabla f$  на поверхность ограничений. Поэтому если  $d'' = 0$ , значит, градиент  $\nabla f$  перпендикулярен поверхности ограничений.

Рассмотрим второе свойство. Так как  $d'' = -P \cdot \nabla f$ , то произведение  $\nabla f^T d'' = -\nabla f^T \cdot P \cdot \nabla f \leq 0$ . Если  $d'' = -P \cdot \nabla f = 0$ , то  $\nabla f \perp d''$ , т.е. вектор  $\nabla f$  перпендикулярен поверхности ограничений. Тогда из формулы

$$d = d' + d'' = A^T \lambda + d'' = 0$$

следует

$$\nabla f = A^T \lambda. \quad (6.83)$$

Так как строками матрицы  $A$  являются векторы коэффициентов в линейных ограничениях, то выражение (6.83) представляет собой другую форму записи необходимого условия Лагранжа:

$$\nabla f - \sum_{k=1}^m \lambda_k a_k = 0, \quad (6.84)$$

где  $a_k$  –  $k$ -й вектор ограничений. Отсюда мы видим, что  $\lambda_k$  – это множители Лагранжа, т.е. множители функции

$$L = f(x) - \sum_{k=1}^m h_k(x) \lambda_k.$$

Необходимое условие экстремума

$$\frac{\partial L}{\partial x_i} = 0 \Rightarrow \nabla f - \sum \lambda_k \nabla h_k = 0,$$

$$\nabla h_k = a_k.$$

Нетрудно получить выражение для  $\lambda$ :

$$d = d' + d'',$$

$$Ad = AA^T \lambda + \underbrace{Ad''}_{=0} \Rightarrow \lambda = (AA^T)^{-1} Ad = (AA^T)^{-1} A \cdot \nabla f,$$

$$\lambda = (AA^T)^{-1} A \cdot \nabla f, \quad (6.85)$$

где  $(AA^T)^{-1}$  существует только в том случае, если значения  $a_k$  линейно независимы. Если имеются линейно зависимые ограничения, их следует исключить из рассмотрения.

*Алгоритм проекции градиента* для случая линейных ограничений. Исходные данные: ЦФ, ограничения, начальная точка  $x^0 \in S$ , точность решения по аргументу и по функции – абсолютная ( $\varepsilon_x$  и  $\varepsilon_y$ ) или относительная ( $\delta_x$  и  $\delta_y$ ). Ограничения необходимо привести к виду  $a_i x \geq b_i$ ,  $i = 1, 2, \dots, m$ .

1. Задать номер итерации  $k = 0$ . Вычислить  $f(x^0)$ .
2. Вычислить  $\nabla f(x^k)$ ,  $g_i(x^k)$ . Если  $\|\nabla f(x^k)\| \leq \varepsilon_y$ , то положить  $x^* \approx x^k$  и закончить вычисления.
3. Составить матрицу  $A$  ограничений из строк, соответствующих активным ограничениям.
4. Вычислить проекционный оператор  $P = E - A^T(AA^T)^{-1}A$  в предположении, что векторы  $a_k$  линейно независимы. Вычислить вектор проекции  $d^k = -P \cdot \nabla f(x^k)$ .
5. Если  $\|d^k\| \geq \varepsilon_x$ , то перейти на шаг 6. Иначе вычислить множители Лагранжа  $\lambda$  по формуле

$$\lambda = (AA^T)^{-1} \cdot A \cdot \nabla f(x^k).$$

Если все  $\lambda_i \geq 0$ , то положить  $x^* \approx x^k$  и закончить вычисления – оптимальное решение найдено. В противном случае ограничение с наибольшим по модулю множителем Лагранжа исключить из множества активных ограничений и перейти на шаг 4.

6. Определить максимальную длину шага:

$$\alpha_{\max} = \min \left\{ \begin{array}{l} \infty, \text{ если } a_k d^k = 0 \text{ или } \frac{b_k - a_k x^k}{a_k d^k} \leq 0 \\ \frac{b_k - a_k x^k}{a_k d^k}, \text{ если } a_k d^k \neq 0 \text{ и } \frac{b_k - a_k x^k}{a_k d^k} > 0 \end{array} \right\}, \quad k = 1, 2, \dots, m.$$

7. Решить задачу одномерного поиска параметра  $\alpha^k$ :

$$\Phi(\alpha) = f(x^k + \alpha d^k) \rightarrow \min_{\alpha}, \quad 0 \leq \alpha \leq \alpha_{\max}.$$

8. Положить  $x^{k+1} = x^k + \alpha^k d^k$ .
9. Проверить близость к решению. Если  $\|x^{k+1} - x^k\| \leq \varepsilon_x$  и  $|f(x^{k+1}) - f(x^k)| \leq \varepsilon_y$  (если заданы относительные погрешности, то

$\|x^{k+1} - x^k\| / \|x^{k+1}\| \leq \delta_x, \quad |f(x^{k+1}) - f(x^k)| / |f(x^{k+1})| \leq \delta_y$ ), положить  $x^* \approx x^{k+1}$  и закончить вычисления. Иначе выполнить присваивание  $k = k + 1$  и вернуться на шаг 2.

*Замечание.* Мы рассмотрели алгоритм, в котором используются линейные равенства. Но его можно легко распространить на неравенства, используя либо дополнительные переменные, либо активные ограничения, что предпочтительнее, так как второй способ позволяет уменьшить размерность.

При втором способе с помощью получаемой по формуле (6.82) оценки множителей Лагранжа осуществляется поочередное исключение ограничений из множества активных ограничений. Модификация выглядит следующим образом. В заданной точке  $x^k$  для определения активного множества проверяются ограничения в виде неравенств  $a_j^T x \geq b_j, \quad j = 1, 2, \dots, m$ .

*Пример.* Решим задачу определения наилучшего положения приемной антенны с ограничениями (6.15), (6.16), (6.17):

$$\begin{aligned} f(x) &= 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min, \\ g_1(x) &= 4 - x_1 - x_2 \geq 0, \\ g_2(x) &= x_1 \geq 0, \\ g_3(x) &= x_2 \geq 0. \end{aligned}$$

Таким образом, исходные данные:  $a_1 = (-1 \quad -1), \quad a_2 = (1 \quad 0),$   
 $a_3 = (0 \quad 1), \quad b_1 = -4, \quad b_2 = 0, \quad b_3 = 0$ . Зададим  $x^0 = (1; 0)^T \in S, \quad \varepsilon_x = \varepsilon_y = 0,03$ .

Шаг 1. Выполняем присваивание  $k = 0$ . Вычисляем  $f(x^0) = 90$ .

Шаг 2. Вычисляем  $g_1(x^0) = 3, \quad g_2(x^0) = 1, \quad g_3(x^0) = 0$ .

$$\nabla f(x^0) = \begin{pmatrix} 2 \cdot 1 - 12 \\ 2 \cdot 0 - 16 \end{pmatrix} = \begin{pmatrix} -10 \\ -16 \end{pmatrix}.$$

Так как  $\|\nabla f(x^0)\| \approx 18,868 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3. Составляем матрицу  $A$ . Активным является ограничение  $g_3(x^0) = 0$ , поэтому

$$A = a_3 = (0 \ 1)$$

Шаг 4. Находим проекционный оператор и вектор проекции:

$$\begin{aligned} P &= E - A^T (AA^T)^{-1} A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \left( (0 \ 1) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)^{-1} \cdot (0 \ 1) = \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot (0 \ 1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \\ d^0 &= -P \cdot \nabla f(x^0) = -\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -10 \\ -16 \end{pmatrix} = \begin{pmatrix} 10 \\ 0 \end{pmatrix}. \end{aligned}$$

Шаг 5. Так как  $\|d^0\| = 10 > \varepsilon_x$ , переходим на шаг 6.

Шаг 6. Определяем максимальную длину шага:

$$a_1 d^0 = (-1 \ -1) \cdot \begin{pmatrix} 10 \\ 0 \end{pmatrix} = -10, \quad b_1 - a_1 x^0 = -4 - (-1 \ -1) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -3,$$

$$a_2 d^0 = (1 \ 0) \cdot \begin{pmatrix} 10 \\ 0 \end{pmatrix} = 10, \quad b_2 - a_2 x^0 = 0 - (1 \ 0) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -1,$$

$$a_3 d^0 = (0 \ 1) \cdot \begin{pmatrix} 10 \\ 0 \end{pmatrix} = 0, \quad b_3 - a_3 x^0 = 0 - (0 \ 1) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0,$$

$$\Rightarrow \alpha_{\max} = \min \left\{ \begin{array}{l} \frac{-3}{-10} \text{ для } g_1 \\ \infty \text{ для } g_2, \text{ так как } \frac{-1}{10} < 0 \\ \infty \text{ для } g_3, \text{ так как } a_3 d^0 = 0 \end{array} \right\} = 0,3.$$

Шаг 7. Решаем задачу одномерного поиска:

$$\Phi(\alpha) = f(x^0 + \alpha d^0) \rightarrow \min_{\alpha}, \quad 0 \leq \alpha \leq 0,3 \Rightarrow$$

$$f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 10 \\ 0 \end{pmatrix}\right) = f\begin{pmatrix} 1+10\alpha \\ 0 \end{pmatrix} = 1 + (1+10\alpha - 6)^2 + 8^2 \rightarrow \min \Rightarrow$$

$$\Phi(\alpha) = 65 + (10\alpha - 5)^2 \rightarrow \min.$$

Дифференцируем:

$$\frac{d\Phi(\alpha)}{d\alpha} = 20(10\alpha - 5) = 0, \quad 0 \leq \alpha \leq 0,3 \Rightarrow$$

$$\Rightarrow 200\alpha = 100 \Rightarrow \alpha^0 = 0,3.$$

Шаг 8. Находим новое приближение оптимума:

$$x^1 = x^0 + \alpha^0 d^0 = \begin{pmatrix} 1+10\alpha \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}.$$

Шаг 9. Вычисляем  $f(x^1) = 53$ . Условия останова итерационного процесса не выполняются:

$$\|x^1 - x^0\| = 3 > \varepsilon_x, \quad |f(x^1) - f(x^0)| = 21 > \varepsilon_y,$$

поэтому полагаем  $k = 1$  и возвращаемся на шаг 2.

Шаг 2<sup>(1)</sup>. Вычисляем  $g_1(x^1) = 0$ ,  $g_2(x^1) = 4$ ,  $g_3(x^1) = 0$ ,

$$\nabla f(x^1) = \begin{pmatrix} 2 \cdot 4 - 12 \\ 2 \cdot 0 - 16 \end{pmatrix} = \begin{pmatrix} -4 \\ -16 \end{pmatrix}.$$

Так как  $\|\nabla f(x^1)\| \approx 16,492 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3<sup>(1)</sup>. Составляем матрицу  $A$ . Активными являются ограничения  $g_1(x^1) = 0$  и  $g_3(x^1) = 0$ , поэтому

$$A = \begin{pmatrix} a_1 \\ a_3 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Шаг 4<sup>(1)</sup>. Находим проекционный оператор и вектор проекции:

$$\begin{aligned}
 P &= E - A^T (AA^T)^{-1} A = \\
 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix} \cdot \left( \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \right)^{-1} \cdot \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} = \\
 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} = \\
 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \\
 d^1 &= -P \cdot \nabla f(x^1) = -\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ -16 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
 \end{aligned}$$

Шаг 5<sup>(1)</sup>. Так как  $\|d^1\| = 0 < \varepsilon_x$ , вычисляем множители Лагранжа:

$$\lambda = (AA^T)^{-1} \cdot A \cdot \nabla f(x^1) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ -16 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ -16 \end{pmatrix} = \begin{pmatrix} 4 \\ -12 \end{pmatrix}.$$

Наибольший по модулю множитель Лагранжа соответствует ограничению  $g_3(x)$ , исключаем его из множества активных ограничений:

$$A = a_1 = \begin{pmatrix} -1 & -1 \end{pmatrix}.$$

Возвращаемся на шаг 4.

Шаг 4<sup>(1)</sup>. Находим проекционный оператор и вектор проекции:

$$\begin{aligned}
 P &= E - A^T (AA^T)^{-1} A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \end{pmatrix} \cdot \left( \begin{pmatrix} -1 & -1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right)^{-1} \cdot \begin{pmatrix} -1 & -1 \end{pmatrix} = \\
 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \end{pmatrix} \cdot \frac{1}{2} \cdot \begin{pmatrix} -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0,5 & 0,5 \\ 0,5 & 0,5 \end{pmatrix} = \begin{pmatrix} 0,5 & -0,5 \\ -0,5 & 0,5 \end{pmatrix}, \\
 d^1 &= -P \cdot \nabla f(x^1) = -\begin{pmatrix} 0,5 & -0,5 \\ -0,5 & 0,5 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ -16 \end{pmatrix} = \begin{pmatrix} -6 \\ 6 \end{pmatrix}.
 \end{aligned}$$

Шаг 5<sup>(1)</sup>. Так как  $\|d^1\| \approx 8,485 > \varepsilon_x$ , переходим на шаг 6.

Шаг 6<sup>(1)</sup>. Определяем максимальную длину шага:

$$a_1 d^1 = (-1 \ -1) \cdot \begin{pmatrix} -6 \\ 6 \end{pmatrix} = 0, \quad b_1 - a_1 x^1 = -4 - (-1 \ -1) \cdot \begin{pmatrix} 4 \\ 0 \end{pmatrix} = 0,$$

$$a_2 d^1 = (1 \ 0) \cdot \begin{pmatrix} -6 \\ 6 \end{pmatrix} = -6, \quad b_2 - a_2 x^1 = 0 - (1 \ 0) \cdot \begin{pmatrix} 4 \\ 0 \end{pmatrix} = -4,$$

$$a_3 d^1 = (0 \ 1) \cdot \begin{pmatrix} -6 \\ 6 \end{pmatrix} = 6, \quad b_3 - a_3 x^1 = 0 - (0 \ 1) \cdot \begin{pmatrix} 4 \\ 0 \end{pmatrix} = 0,$$

$$\Rightarrow \alpha_{\max} = \min \left\{ \begin{array}{l} \infty \text{ для } g_1, \text{ так как } a_1 d^1 = 0 \\ \frac{-4}{-6} \text{ для } g_2 \\ \infty \text{ для } g_3, \text{ так как } \frac{b_3 - a_3 x^1}{a_3 d^1} = 0 \end{array} \right\} = \frac{2}{3}.$$

Шаг 7<sup>(1)</sup>. Решаем задачу одномерного поиска:

$$\Phi(\alpha) = f(x^1 + \alpha d^1) \rightarrow \min_{\alpha}, \quad 0 \leq \alpha \leq \frac{2}{3} \Rightarrow$$

$$f\left(\begin{pmatrix} 4 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -6 \\ 6 \end{pmatrix}\right) = f\left(\begin{pmatrix} 4-6\alpha \\ 6\alpha \end{pmatrix}\right) = 1 + (4-6\alpha-6)^2 + (6\alpha-8)^2 \rightarrow \min \Rightarrow$$

$$\Phi(\alpha) = 1 + (-6\alpha-2)^2 + (6\alpha-8)^2 \rightarrow \min.$$

Дифференцируем:

$$\frac{d\Phi(\alpha)}{d\alpha} = -12(-6\alpha-2) + 12(6\alpha-8) = 0, \quad 0 \leq \alpha \leq \frac{2}{3} \Rightarrow$$

$$\Rightarrow 144\alpha = -72 \Rightarrow \alpha^0 = 0,5.$$

Шаг 8<sup>(1)</sup>. Находим новое приближение оптимума:

$$x^2 = x^1 + \alpha^1 d^1 = \begin{pmatrix} 4-6\alpha \\ 6\alpha \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

Шаг 9<sup>(1)</sup>. Вычисляем  $f(x^2) = 51$ . Условия останова итерационного процесса не выполняются:

$$\|x^2 - x^1\| \approx 4,243 > \varepsilon_x, \quad |f(x^2) - f(x^1)| = 18 > \varepsilon_y,$$

поэтому полагаем  $k = 2$  и возвращаемся на шаг 2.

Шаг 2<sup>(2)</sup>. Вычисляем  $g_1(x^2) = 0$ ,  $g_2(x^2) = 1$ ,  $g_3(x^2) = 3$ ,

$$\nabla f(x^2) = \begin{pmatrix} 2 \cdot 1 - 12 \\ 2 \cdot 3 - 16 \end{pmatrix} = \begin{pmatrix} -10 \\ -10 \end{pmatrix}.$$

Так как  $\|\nabla f(x^2)\| \approx 14,142 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3<sup>(2)</sup>. Составляем матрицу  $A$ . Активным является ограничение  $g_1(x^2) = 0$ , поэтому

$$A = a_1 = \begin{pmatrix} -1 & -1 \end{pmatrix}.$$

Шаг 4<sup>(2)</sup>. Проекционный оператор для данного активного ограничения мы уже находили выше:

$$\begin{aligned} P &= E - A^T (AA^T)^{-1} A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \end{pmatrix} \cdot \left( (-1 \quad -1) \cdot \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right)^{-1} \cdot (-1 \quad -1) = \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \end{pmatrix} \cdot \frac{1}{2} \cdot (-1 \quad -1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0,5 & 0,5 \\ 0,5 & 0,5 \end{pmatrix} = \begin{pmatrix} 0,5 & -0,5 \\ -0,5 & 0,5 \end{pmatrix}, \end{aligned}$$

находим вектор проекции:

$$d^2 = -P \cdot \nabla f(x^2) = - \begin{pmatrix} 0,5 & -0,5 \\ -0,5 & 0,5 \end{pmatrix} \cdot \begin{pmatrix} -10 \\ -10 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Шаг 5<sup>(2)</sup>. Так как  $\|d^2\| = 0 < \varepsilon_x$ , вычисляем множители Лагранжа:

$$\lambda = (AA^T)^{-1} \cdot A \cdot \nabla f(x^2) = \frac{1}{2} \cdot (-1 \quad -1) \cdot \begin{pmatrix} -10 \\ -10 \end{pmatrix} = 10.$$

Все  $\lambda_i \geq 0$ , таким образом, точка  $x^2$  удовлетворяет условиям Куна – Таккера. Полагаем  $x^* = x^2$  и заканчиваем решение.

С учетом того, что  $\varepsilon_x = \varepsilon_y = 0,03$ , округляем значения аргумента и функции до двух цифр после запятой:  $x^* = x^3 = (1,00; 3,00)^T$ ,  $f^* = f(x^3) = 51$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 7,141$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 19,61 \text{ Вт/м}^2.$$

### 6.7.2. Задачи с нелинейными ограничениями

Один из подходов к решению ЗНП состоит в такой модификации градиентных методов безусловной минимизации, чтобы в процессе построения последовательных приближений к точке минимума учитывались ограничения на допустимое множество  $S$ .

Рассмотрим ЗНП

$$\begin{aligned} f(x) &\rightarrow \min, \\ g_1(x) &\leq 0, \quad i = 1, 2, \dots, m, \\ x &\in S \subset R^n, \end{aligned}$$

где  $S$  – выпуклое замкнутое множество;  $f(x)$  – выпуклая дифференцируемая на множестве  $S$  функция.

На каждой итерации метода предусмотрена процедура возврата очередного приближения градиентного спуска

$$x^{k+1} = x^k - \alpha_k \nabla f(x^k)$$

на допустимое множество  $S$ , если  $x^{k+1} \notin S$ . Такой возврат проводится посредством проектирования  $x^{k+1}$  на  $S$ , т.е. замены  $x^{k+1}$  на ближайшую точку множества  $S$ .

*Определение.* Пусть заданы замкнутое множество  $S \subset R^n$  и точка  $z \in R^n$ . Точка  $z_S = P_S(z)$  называется проекцией точки  $z$  на множество

$S$ , если  $\rho(z_S, z) = \min_{x \in S} \rho(x, z)$ , где  $\rho(x, y) = \|x - y\|$  – расстояние между точкой  $x$  и точкой  $y$  в пространстве  $R^n$ .

Ясно, что для точки  $z \in S$  проекция  $P_S(z)$  совпадает с точкой  $z$ .

Таким образом, в методе проекции градиента последовательные приближения  $x^k$  к точке минимума  $x^*$  целевой функции  $f(x)$  на множестве  $S$  вычисляются по формулам

$$x^{k+1} = P_S \left[ x^k - \alpha^k \nabla f(x^k) \right], \quad k = 0, 1, \dots, \quad x^0 \in S. \quad (6.86)$$

В зависимости от способа вычисления  $\alpha^k$  различают несколько вариантов МПГ, из них приведем два наиболее распространенных:

1)  $\alpha^k$  находятся как и в методе наискорейшего спуска безусловной минимизации, т.е.

$$\Phi(\alpha) = f(x^k - \alpha \nabla f(x^k)) \rightarrow \min_{\alpha};$$

2) в предположении, что  $\nabla f(x)$  удовлетворяет на множестве  $S$  условию Липшица, т.е.

$$\|\nabla f(x') - \nabla f(x'')\| \leq \Lambda \|x' - x''\| \quad (6.87)$$

для  $\forall x', x'' \in S$ , полагают  $\alpha^k \equiv \alpha$ ,  $k = 0, 1, \dots$ , где  $\alpha$  – произвольное число из интервала  $(0, 2/\Lambda]$ . Если известна минимальная константа Липшица  $\Lambda$  из выражения (6.87), то выбирается  $\alpha^k = 1/\Lambda$ .

*Алгоритм проекции градиента* для случая нелинейных ограничений. Исходные данные: ЦФ, ограничения, начальная точка  $x^0 \in S$ , точность решения по аргументу и по функции – абсолютная ( $\varepsilon_x$  и  $\varepsilon_y$ ) или относительная ( $\delta_x$  и  $\delta_y$ ).

1. Задать номер итерации  $k = 0$ . Вычислить  $f(x^0)$ .
2. Вычислить  $\nabla f(x^k)$ . Если  $\|\nabla f(x^k)\| \leq \varepsilon_y$ , то положить  $x^* \approx x^k$  и закончить вычисления.
3. Найти  $\alpha^k$ , представляющее собой решение задачи при

$$\Phi(\alpha) = f(x^k - \alpha \nabla f(x^k)) \rightarrow \min_{\alpha},$$

либо из соотношения (6.87), т.е.  $\alpha^k \in (0, 2/\Lambda]$ , например  $\alpha^k = 1/\Lambda$ .

4. Вычислить следующее приближение по формуле (6.86):

$$x^{k+1} = P_S [x^k - \alpha^k \nabla f(x^k)].$$

5. Проверить близость к решению. Если  $\|x^{k+1} - x^k\| \leq \varepsilon_x$  и  $|f(x^{k+1}) - f(x^k)| \leq \varepsilon_y$  (если заданы относительные погрешности, то  $\|x^{k+1} - x^k\|/\|x^{k+1}\| \leq \delta_x$ ,  $|f(x^{k+1}) - f(x^k)|/|f(x^{k+1})| \leq \delta_y$ ), положить  $x^* \approx x^{k+1}$  и закончить вычисления. Иначе выполнить присваивание  $k = k + 1$  и вернуться на шаг 2.

*Замечание.* Определение проекции  $P_S(z)$  для точки  $z \in S$  является самостоятельной ЗНП

$$\tilde{f}(x) = \|x - z\|^2 = \sum_{j=1}^n (x_j - z_j)^2 \rightarrow \min, \quad (6.88)$$

решение которой может вызвать затруднения. Решить ее можно, например, с помощью метода множителей Лагранжа.

Рассмотрим задачу проецирования, которая решается в явном виде.

*Пример.* Найти проекцию  $P_S(z)$  точки  $z \in R^n$  на множество  $S$ , где

$$S = \left\{ x \in R^n \left| \|x\|^2 = \sum_{j=1}^n x_j^2 \leq R_0^2 \right. \right\} -$$

шар радиусом  $R_0$  в пространстве  $R^n$  с центром в начале координат.

Запишем задачу поиска проекции как ЗНП:

$$\begin{cases} f(x) = \sum_{j=1}^n (x_j - z_j)^2 \rightarrow \min, \\ \sum_{j=1}^n x_j^2 \leq R_0^2. \end{cases} \quad (6.89)$$

Рассмотрим возможные частные случаи.

1. Пусть  $z \in S$ , т.е.  $\sum_{j=1}^n z_j^2 \leq R_0^2$ . Тогда  $P_S(z) = z$ .

2. Пусть  $z \notin S$ , т.е.  $\sum_{j=1}^n z_j^2 = R^2 > R_0^2$ . Запишем неравенство ЗНП

(6.89) в виде равенства, добавив новую переменную  $x_{n+1} = y^2 \geq 0$ . В результате получим задачу на условный экстремум с ограничением-равенством:

$$f(x) = \sum_{j=1}^n (x_j - z_j)^2 \rightarrow \min,$$

$$h(x, y) = \sum x_j^2 + y^2 - R_0^2 = 0.$$

Составляем функцию Лагранжа и минимизируем ее:

$$L(x, y, \lambda) = \sum_{j=1}^n (x_j - z_j)^2 + \lambda \left( \sum_{j=1}^n x_j^2 + y^2 - R_0^2 \right) \Rightarrow$$

$$\begin{cases} \frac{\partial L}{\partial x_j} = 2(x_j - z_j) + 2\lambda x_j = 0, \\ \frac{\partial L}{\partial y} = 2\lambda y = 0, \\ \frac{\partial L}{\partial \lambda} = \sum x_j^2 + y^2 - R_0^2 = 0. \end{cases}$$

Решив эту систему  $n+2$  уравнений с учетом  $\sum_{j=1}^n z_j^2 = R^2$ , находим

$x_j = z_j R_0/R$ ,  $j = 1, \dots, n$ ,  $1 + \lambda = R_0/R$ ,  $y = 0$ , т.е.

$$P_S(z) = \begin{cases} z, & \text{если } \sum_{j=1}^n z_j^2 \leq R_0^2, \\ \frac{z}{\sqrt{\sum_{j=1}^n z_j^2}} R_0, & \text{если } \sum_{j=1}^n z_j^2 > R_0^2. \end{cases}$$

В тех случаях, когда ОДР задачи не выпукла либо выражение для  $P_S(z)$  невозможно получить аналитически (например, если ОДР будет являться не шаром, а эллипсоидом), использование МПГ вызывает затруднения.

*Пример.* Решим задачу определения наилучшего положения приемной антенны с ограничением (6.2):

$$f(x) = 1 + (x_1 - 6)^2 + (x_2 - 8)^2 \rightarrow \min,$$

$$g(x) = 4 - x_1^2 - x_2^2 \geq 0.$$

Зададим  $x^0 = (1; 1)^T \in S$ ,  $\varepsilon_x = \varepsilon_y = 0,1$ . Определим константу Липшица:

$$\Lambda = \frac{\|\nabla f(x') - \nabla f(x'')\|}{\|x' - x''\|} = \frac{\left\| \begin{pmatrix} 2x'_1 - 12 \\ 2x'_2 - 16 \end{pmatrix} - \begin{pmatrix} 2x''_1 - 12 \\ 2x''_2 - 16 \end{pmatrix} \right\|}{\|x' - x''\|} = \frac{\left\| \begin{pmatrix} 2x'_1 - 2x''_1 \\ 2x'_2 - 2x''_2 \end{pmatrix} \right\|}{\|x' - x''\|} =$$

$$= 2 \frac{\|x' - x''\|}{\|x' - x''\|} = 2.$$

Если ограничение задачи переписать в виде  $g(x) = x_1^2 + x_2^2 \leq 4$ , то очевидно, что оно является кругом в пространстве  $R^2$  с центром в точке  $(0;0)$  и радиусом  $R_0 = 2$ .

Шаг 1. Выполняем присваивание  $k = 0$ . Вычисляем  $f(x^0) = 75$ .

Шаг 2. Вычисляем

$$\nabla f(x^0) = \begin{pmatrix} 2 \cdot 1 - 12 \\ 2 \cdot 1 - 16 \end{pmatrix} = \begin{pmatrix} -10 \\ -14 \end{pmatrix}.$$

Так как  $\|\nabla f(x^0)\| \approx 17,205 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3. Выполняем одномерный поиск вдоль прямой:

$$\Phi(\alpha) = f\left(x^0 - \alpha \nabla f(x^0)\right) \rightarrow \min_{\alpha} \Rightarrow$$

$$f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \alpha \begin{pmatrix} -10 \\ -14 \end{pmatrix}\right) = f\begin{pmatrix} 1+10\alpha \\ 1+14\alpha \end{pmatrix} = 1 + (1+10\alpha-6)^2 + (1+14\alpha-8)^2 \rightarrow \min \Rightarrow$$

$$\Phi(\alpha) = 1 + (10\alpha-5)^2 + (14\alpha-7)^2 \rightarrow \min.$$

Дифференцируем:

$$\frac{d\Phi(\alpha)}{d\alpha} = 20(10\alpha-5) + 28(14\alpha-7) = 0 \Rightarrow$$

$$\Rightarrow 592\alpha = 296 \Rightarrow \alpha^0 = 0,5.$$

Отметим, что если воспользоваться вторым способом определения длины шага, получится аналогичное значение:

$$\alpha^0 = \frac{1}{\Lambda} = 0,5.$$

Шаг 4. Находим новое приближение оптимума:

$$x^1 = P_S \left[ x^0 - \alpha^0 \nabla f(x^0) \right] = P_S(1+10\alpha; 1+14\alpha) = P_S(6;8).$$

Так как точка  $z = (6;8)^T \notin S$  (ограничение в этой точке  $g(z) = -96 < 0$ ), находим ее проекцию на ОДР:

$$x^1 = \frac{z}{\sqrt{\sum_{j=1}^n z_j^2}} R_0 = \frac{(6;8)^T}{\sqrt{100}} \cdot 2 = \begin{pmatrix} 1,2 \\ 1,6 \end{pmatrix}.$$

Шаг 5. Вычисляем  $f(x^1) = 65$ . Условия останова итерационного процесса не выполняются:

$$\|x^1 - x^0\| \approx 0,632 > \varepsilon_x, \quad |f(x^1) - f(x^0)| = 10 > \varepsilon_y,$$

поэтому полагаем  $k = 1$  и возвращаемся на шаг 2.

Шаг 2<sup>(1)</sup>. Вычисляем

$$\nabla f(x^1) = \begin{pmatrix} 2 \cdot 1,2 - 12 \\ 2 \cdot 1,6 - 16 \end{pmatrix} = \begin{pmatrix} -9,6 \\ -12,8 \end{pmatrix}.$$

Так как  $\|\nabla f(x^1)\| = 16 > \varepsilon_y$ , продолжаем вычисления.

Шаг 3<sup>(1)</sup>. Выполняем одномерный поиск вдоль прямой:

$$\begin{aligned} \Phi(\alpha) &= f(x^1 - \alpha \nabla f(x^1)) \rightarrow \min_{\alpha} \Rightarrow \\ f\left(\begin{pmatrix} 1,2 \\ 1,6 \end{pmatrix} - \alpha \begin{pmatrix} -9,6 \\ -12,8 \end{pmatrix}\right) &= f\left(\begin{pmatrix} 1,2 + 9,6\alpha \\ 1,6 + 12,8\alpha \end{pmatrix}\right) = \\ &= 1 + (1,2 + 9,6\alpha - 6)^2 + (1,6 + 12,8\alpha - 8)^2 \rightarrow \min \Rightarrow \\ \Phi(\alpha) &= 1 + (9,6\alpha - 4,8)^2 + (12,8\alpha - 6,4)^2 \rightarrow \min. \end{aligned}$$

Дифференцируем:

$$\begin{aligned} \frac{d\Phi(\alpha)}{d\alpha} &= 19,2(9,6\alpha - 4,8) + 25,6(12,8\alpha - 6,4) = 0 \Rightarrow \\ &\Rightarrow 512\alpha = 256 \Rightarrow \alpha^1 = 0,5. \end{aligned}$$

Шаг 4<sup>(1)</sup>. Находим новое приближение оптимума:

$$x^2 = P_S \left[ x^1 - \alpha^1 \nabla f(x^1) \right] = P_S(1,2 + 9,6\alpha; 1,6 + 12,8\alpha) = P_S(6;8).$$

Так как точка  $z = (6;8)^T \notin S$  (ограничение в этой точке  $g(z) = -96 < 0$ ), находим ее проекцию на ОДР:

$$x^2 = \frac{z}{\sqrt{\sum_{j=1}^n z_j^2}} R_0 = \frac{(6;8)^T}{\sqrt{100}} \cdot 2 = \begin{pmatrix} 1,2 \\ 1,6 \end{pmatrix}.$$

Шаг 5<sup>(1)</sup>. Вычисляем  $f(x^2) = 65$ . Условия останова итерационного процесса выполняются:

$$\|x^2 - x^1\| = 0 < \varepsilon_x, \quad |f(x^2) - f(x^1)| = 0 < \varepsilon_y,$$

поэтому полагаем  $x^* = x^2$  и заканчиваем решение.

С учетом того, что  $\varepsilon_x = \varepsilon_y = 0,03$ , округляем значения аргумента и функции до двух цифр после запятой:  $x^* = x^2 = (1,20; 1,60)^T$ ,  $f^* = f(x^2) = 65$ . Расстояние от антенны до вышки составит  $\sqrt{f^*} \approx 8,062$  км. Интенсивность принимаемого сигнала

$$E^* = \frac{1000}{1 + (x_1^* - 6)^2 + (x_2^* - 8)^2} \approx 15,39 \text{ Вт/м}^2.$$

### Вопросы для самопроверки

1. Назовите достоинства и недостатки прямых методов решения ЗНП.
2. В чем суть метода замены переменных при решении ЗНП? Приведите пример.
3. В каких случаях применим метод замены переменных?
4. В чем заключается основная идея метода множителей Лагранжа?
5. Какие ограничения накладываются на множители Лагранжа?
6. В каком случае функция Лагранжа является выпуклой?
7. Что является седловой точкой функции Лагранжа?
8. Сформулируйте необходимые условия экстремума функции Лагранжа.
9. Как с экономической точки зрения интерпретируются множители Лагранжа?
10. Сформулируйте необходимые и достаточные условия первого порядка оптимальности ЗНП.
11. Какие условия накладываются на функции ограничений?
12. Что означают понятия «активное» и «неактивное» ограничение?
13. Что необходимо для выполнения условий Куна – Таккера?

14. Сформулируйте необходимые и достаточные условия второго порядка оптимальности ЗНП.
15. Сформулируйте общий принцип методов штрафов.
16. В чем заключается сходство МЗП, ММЛ и методов штрафов?
17. Что такое внешний штраф? Приведите графическую иллюстрацию.
18. Дайте понятие внутреннего штрафа. Приведите графическую иллюстрацию.
19. Для учета ограничений какого типа используется квадратичный штраф? Приведите графическую иллюстрацию.
20. Запишите выражение для логарифмического штрафа. Приведите графическую иллюстрацию.
21. Запишите выражение для штрафа типа обратной функции. Приведите графическую иллюстрацию.
22. Запишите выражение для штрафа типа квадрата срезки. Приведите графическую иллюстрацию.
23. В чем заключаются достоинства штрафа типа квадрата срезки?
24. При выполнении каких требований целесообразно использовать МШФ?
25. Каким образом можно выбирать штрафной параметр  $R$ ?
26. В чем сущность обобщенного алгоритма МШФ?
27. В чем сущность метода линеаризации для ЗНП общего вида?
28. На основе какого подхода разработан алгоритм Франка – Вульфа?
29. В чем заключается основная идея метода допустимых направлений Зойтендейка?
30. В чем суть проекционного метода решения ЗНП?
31. В чем заключается идея метода условного градиента?
32. Каким образом формируется вспомогательная задача в методе условного градиента?
33. Каким образом определяется направление спуска в методе проекции градиента?
34. В чем заключается суть ортогонального проецирования?
35. Что представляет собой проекция градиента?

36. Что такое проекционная матрица?
37. Перечислите основные свойства проекций.
38. Каким образом можно уменьшить размерность задачи в МПГ?
39. Приведите определение проекции точки на множество.

## Глава 7. КВАДРАТИЧНОЕ ПРОГРАММИРОВАНИЕ

### 7.1. Задача выбора портфеля ценных бумаг

Задачи квадратичного программирования (ЗКП) являются частным случаем ЗНП, поэтому для их решения можно применять рассмотренные ранее методы решения ЗНП. Однако специализированные методы решения таких задач являются более эффективными.

Определение *оптимального портфеля ценных бумаг* представляет собой одну из важнейших задач, с которыми сталкиваются инвестиционные фирмы – банки, фонды, страховые компании и частные лица.

Под портфелем понимаются размеры вложений в различные виды ценных бумаг: обычные облигации, облигации краткосрочных государственных займов, банковские депозитные сертификаты, акции и др.

Рассмотрим финансовую операцию, заключающуюся в покупке рискованных ценных бумаг по известной цене и продаже их в будущем по цене, заранее неизвестной. Предполагается, что инвестор в настоящий момент времени инвестирует некоторую сумму денег в ценные бумаги. Эти деньги будут инвестированы на определенный промежуток времени, который называют *периодом владения*. В конце этого периода инвестор продает ценные бумаги, которые были куплены в начале периода. Таким образом, в момент  $t = 0$  инвестор должен принять решение о покупке ценных бумаг, которые будут находиться в его портфеле до момента  $t = 1$ . Такую задачу называют *задачей выбора инвестиционного портфеля*.

Эффективность рискованной ценной бумаги зависит от трех факторов:

- цены покупки, которая точно известна;
- промежуточных выплат за период владения (дивидендов), которые точно не известны;
- цены продажи, которая неизвестна.

Таким образом, финансовая операция, заключающаяся в покупке ценной бумаги с целью получения определенного дохода в будущем,

является рискованной. Основная гипотеза, которая позволяет анализировать такую операцию, состоит в следующем: предполагаем, что каждое конкретное значение эффективности такой финансовой операции является реализацией случайной величины

$$R = \frac{C_1 - C_0 + D}{C_0},$$

где  $C_0$  – цена покупки;  $C_1$  – цена продажи;  $D$  – дивиденды, выплаченные за период владения.

Формируя портфель ценных бумаг, инвестор хотел бы максимизировать ожидаемую доходность портфеля при минимальном риске. Как правило, эти две цели противоречат друг другу. Принимая решение, инвестор стремится сделать так, чтобы эти цели были сбалансированы.

Доходность портфеля также является случайной величиной:

$$R_p = \frac{W_1 - W_0}{W_0},$$

где  $W_0$  – совокупная цена покупки ценных бумаг, входящих в портфель в момент  $t = 0$ ;  $W_1$  – совокупная рыночная цена ценных бумаг в момент  $t = 1$  и совокупный денежный доход от этих ценных бумаг (дивиденды), который владелец получит за период владения от момента  $t = 0$  до  $t = 1$ .

Любая случайная величина может характеризоваться двумя параметрами: ожидаемым или средним значением (математическим ожиданием) и стандартным отклонением (среднеквадратичным отклонением). Согласно рассматриваемой модели предполагается, что инвестор основывает свое решение по выбору портфеля только на этих двух параметрах. Следовательно, инвестор должен оценить ожидаемую доходность и стандартное отклонение каждого возможного портфеля. Затем он должен выбрать лучший из портфелей, основываясь на соотношении этих двух параметров. При этом ожидаемая доходность рассматривается как мера потенциального вознаграждения, связанная с конкретным портфелем, а стандартное отклонение – как мера риска, связанная с данным портфелем.

Итак, рассмотрим финансовую операцию, которая заключается в покупке ценных бумаг в момент  $t = 0$  по известной цене и в продаже их в момент  $t = 1$  по цене, заранее не известной. При этом инвестор может рассчитывать на получение промежуточных выплат. Обозначим ожидаемое значение эффективности ценной бумаги как  $\mu = M(R)$  – это математическое ожидание случайной величины  $R$  (среднее по всем реализациям (значениям) случайной величины, вычисленное с учетом частоты их возможного появления); дисперсию или вариацию случайной величины как  $V = M\{(R - \mu)^2\}$  – это мера отклонения в среднем случайной величины  $R$  от ее ожидаемого значения. Часто вместо дисперсии используют среднеквадратичное или стандартное отклонение  $\sigma = \sqrt{V}$ .

Ковариация  $V_{12} = M\{(R_1 - m_1)(R_2 - m_2)\}$  характеризует статистическую взаимосвязь случайных величин  $R_1$  и  $R_2$ .

Риск вложений в конкретные ценные бумаги связан с неопределенностью будущих доходов и, следовательно, с неопределенностью эффективности данной операции. Чем больше стандартное отклонение, тем больше в среднем случайная величина может отклониться от своего ожидаемого значения, тем больше неопределенность и выше риск. С другой стороны, если  $\sigma = 0$ , то эффективность не отклоняется от своего ожидаемого значения, она принимает определенные не случайные значения и риск отсутствует. Таким образом, стандартное отклонение характеризует уровень риска, связанный с конкретной ценной бумагой, и принимается в качестве *меры риска*.

Предположим, что инвестор вкладывает деньги не в один вид ценных бумаг, а в несколько. В этом случае говорят, что инвестор диверсифицирует свой портфель. Рассмотрим эффект такой диверсификации.

Пусть  $x_j, j = 1, 2, \dots, n$ , – доля общего вложения, приходящаяся на  $j$ -й вид ценных бумаг;  $n$  – количество видов ценных бумаг, которые инвестор включает в портфель. Очевидно, что должно выполняться равенство

$$\sum_{j=1}^n x_j = 1.$$

Пусть  $R_p$  – эффективность портфеля,  $R_j$  – эффективность  $j$ -й ценной бумаги. Тогда

$$R_p = \sum_{j=1}^n R_j x_j.$$

Ожидаемая эффективность портфеля

$$m_p = M\{R_p\} = \sum_{j=1}^n x_j M\{R_j\} = \sum_{j=1}^n x_j m_j,$$

где  $m_j = M\{R_j\}$  – ожидаемая эффективность  $j$ -й ценной бумаги. Отклонение от ожидаемой эффективности

$$R_p - m_p = \sum_{j=1}^n x_j (R_j - m_j).$$

Дисперсия эффективности портфеля

$$V_p = M\{(R_p - m_p)^2\} = \sum_{i=1}^n \sum_{j=1}^n x_i x_j M\{(R_i - m_i)(R_j - m_j)\} = \sum_{i=1}^n \sum_{j=1}^n V_{ij} x_i x_j,$$

где  $V_{ij} = M\{(R_i - m_i)(R_j - m_j)\}$  – ковариация случайных эффективностей  $R_i$  и  $R_j$   $i$ -го и  $j$ -го видов ценных бумаг. Нетрудно заметить, что

$$V_{jj} = M\{(R_j - m_j)^2\} = \sigma_j^2.$$

Величина  $V_p$  (или  $\sigma_p$ ) характеризует неопределенность портфеля в целом и называется *риском портфеля*.

При этих предположениях можно сформулировать следующую оптимизационную задачу: определить доли вложений  $x_j$ ,  $j = 1, 2, \dots, n$ , минимизирующие вариацию (риск) портфеля

$$V_p = \sum_{i=1}^n \sum_{j=1}^n V_{ij} x_i x_j, \quad (7.1)$$

при условии, что обеспечивается заданное значение ожидаемой эффективности портфеля

$$m_p = \sum_{j=1}^n m_j x_j. \quad (7.2)$$

Кроме того, должны быть выполнены дополнительные ограничения вида

$$\sum_{j=1}^n x_j = 1, \quad (7.3)$$

$$x_j \geq 0, \quad j = 1, 2, \dots, n. \quad (7.4)$$

В этой модели целевая функция квадратичная, а ограничения линейные.

Заметим, что если ограничения (7.4) отсутствуют (на фондовых рынках это означает, что допускаются так называемые короткие продажи), то задача (7.1)–(7.3) может быть решена методом множителей Лагранжа. Аналитическое решение имеет вид

$$x = u + m_p v, \quad (7.5)$$

где  $u$  и  $v$  – векторы, определяемые соотношениями

$$\begin{aligned} u &= \frac{1}{d} [bV^{-1}e - aV^{-1}m], \\ v &= \frac{1}{d} [cV^{-1}m - aV^{-1}e], \end{aligned} \quad (7.6)$$

$$a = e^T V^{-1} m, \quad b = m^T V^{-1} m, \quad c = e^T V^{-1} e, \quad d = bc - a^2, \quad e = (1, 1, \dots, 1)^T.$$

*Пример.* Пусть ожидаемые доходности рискованных ценных бумаг и матрица ковариации составляют

$$m = (0,48 \quad 0,36 \quad 0,54)^T, \quad V = \begin{pmatrix} 0,058 & 0,031 & 0,036 \\ 0,031 & 0,032 & 0,011 \\ 0,036 & 0,011 & 0,073 \end{pmatrix}.$$

Зададим доходность портфеля  $m_p = 0,46$ . Тогда оптимальная структура портфеля, найденная по формулам (7.5), (7.6), будет

$$x = (0,102 \quad 0,410 \quad 0,487)^T.$$

## 7.2. Классификация методов

Как уже было сказано выше, задача квадратичного программирования представляет собой задачу с линейными ограничениями и квадратичной ЦФ и является частным случаем задачи нелинейного программирования.

Общая ЗКП имеет вид

$$f(x) = \sum_{j=1}^n c_j x_j + \sum_{i=1}^n \sum_{j=1}^n q_{ij} x_i x_j \rightarrow \min \quad (7.7)$$

при ограничениях

$$\sum_{j=1}^n a_{ij} x_j = b_i, \quad i = 1, 2, \dots, m, \quad (7.8)$$

$$x_j \geq 0, \quad j = 1, 2, \dots, n, \quad (7.9)$$

или в матричном виде

$$\begin{aligned} f(x) &= c^T x + x^T Q x, \\ Ax &= b, \\ x &\geq 0, \end{aligned} \quad (7.10)$$

где  $A$  – матрица ограничений ( $m \times n$ );  $b$  – вектор ограничений ( $m \times 1$ );  $c$  – вектор ( $1 \times n$ );  $Q$  – квадратная матрица ( $n \times n$ );  $x$  – вектор переменных ( $1 \times n$ ).

Далее рассмотрим следующие методы решения ЗКП:

- условия Куна – Таккера (вообще, они определяют достаточные условия оптимальности решения ЗКП, но также позволяют решить ее);
- метод искусственного базиса;

- метод симплексного преобразования;
- метод решения задачи о дополнителъности;
- метод Мицеля – Хвощевского.

### 7.3. Условие Куна – Таккера для задачи квадратичного программирования

Запишем задачу квадратичного программирования с ограничениями-равенствами:

$$\begin{cases} f(x) = cx + x^T Qx \rightarrow \min, \\ g(x) = x \geq 0, \\ h(x) = Ax - b = 0. \end{cases} \quad (7.11)$$

Функция Лагранжа для этой задачи имеет вид

$$L(x, \mu, \lambda) = f(x) - \mu^T x - \lambda^T h(x), \quad (7.12)$$

где  $\mu \geq 0$ ,  $\lambda$  – параметры (ограничений по знаку параметры  $\lambda$  не имеют).

Условия Куна – Таккера для ЗКП записываются следующим образом:

$$\begin{cases} c + x^T (Q + Q^T) - \mu - \lambda^T A = 0, \\ Ax = b, \quad x \geq 0, \\ \mu^T x = 0, \mu \geq 0, \end{cases} \quad (7.13)$$

где  $\lambda$  – не ограниченная по знаку переменная.

*Пример 1.* Рассмотрим ЗКП

$$\begin{aligned} f(x) &= -6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2 \rightarrow \min, \\ x_1 + x_2 &= 2, \quad x_1, x_2 \geq 0. \end{aligned}$$

Для этой задачи условия Куна – Таккера задаются в виде

$$\begin{aligned}
-6 + 4x_1 - 2x_2 - \mu_1 - \lambda &= 0, \\
-2x_1 + 4x_2 - \mu_2 - \lambda &= 0, \\
x_1 + x_2 &= 2,
\end{aligned} \tag{7.14}$$

$$\mu_1 x_1 = 0; \mu_2 x_2 = 0; x_1, x_2 \geq 0; \mu_1, \mu_2 \geq 0.$$

Величина  $\lambda$  не ограничена по знаку.

Поскольку ограничение  $h(x)$  – линейное, то применима теорема Куна – Таккера. Таким образом, условие Куна – Таккера является необходимым условием оптимальности решения ЗКП.

*Пример 2.* Запишем условия Куна – Таккера для задачи выбора портфеля ценных бумаг. Модель имеет вид

$$\begin{aligned}
f(x) &= V_{11}x_1^2 + V_{22}x_2^2 + V_{33}x_3^2 + 2x_1x_2V_{12} + 2x_1x_3V_{13} + 2x_2x_3V_{23} \rightarrow \min, \\
m_1x_1 + m_2x_2 + m_3x_3 &= m_p, \\
x_1 + x_2 + x_3 &= 1, \\
x_1, x_2 &\geq 0.
\end{aligned}$$

Подставим исходные данные, получим

$$\begin{aligned}
f(x) &= \\
&= 0,058x_1^2 + 0,032x_2^2 + 0,073x_3^2 + 0,062x_1x_2 + 0,072x_1x_3 + 0,022x_2x_3 \rightarrow \min, \\
&0,48x_1 + 0,36x_2 + 0,54x_3 = 0,46, \\
&x_1 + x_2 + x_3 = 1.
\end{aligned}$$

Для этой задачи условия Куна – Таккера задаются следующим образом:

$$\begin{aligned}
0,116x_1 + 0,062x_2 + 0,072x_3 - \mu_1 - 0,48\lambda_1 - \lambda_2 &= 0, \\
0,062x_1 + 0,064x_2 + 0,022x_3 - \mu_2 - 0,36\lambda_1 - \lambda_2 &= 0, \\
0,072x_1 + 0,022x_2 + 0,146x_3 - \mu_3 - 0,54\lambda_1 - \lambda_2 &= 0, \\
0,48x_1 + 0,36x_2 + 0,54x_3 &= 0,46, \\
x_1 + x_2 + x_3 &= 1,
\end{aligned} \tag{7.15}$$

$$\mu_1 x_1 = 0, \mu_2 x_2 = 0, \mu_3 x_3 = 0; x_1, x_2, x_3 \geq 0; \mu_1, \mu_2, \mu_3 \geq 0.$$

Величины  $\lambda_1, \lambda_2$  не ограничены по знаку.

В случае положительной определенности или полуопределенности матрицы  $Q$  целевая функция является выпуклой и условия Куна – Таккера будут также достаточными условиями оптимальности решения ЗКП. Таким образом, в случае  $H_f \geq 0$  достаточно найти точку  $x$ , удовлетворяющую условиям Куна – Таккера.

#### 7.4. Метод решения ЗКП с помощью искусственного базиса

Одним из методов решения ЗКП может быть метод решения системы (7.13) как задачи ЛП с использованием искусственного базиса и первой операции симплекс-метода. При этом условие  $\mu_j x_j = 0$  неявно учитывается в процедуре симплекс-метода при помощи правила ограниченного ввода в базис. Это правило запрещает вводить в базис небазисную переменную  $\mu_j$ , если значение соответствующей базисной переменной  $x_j$  положительное.

*Пример 1.* Рассмотрим этот метод на примере задачи из подраздела 7.3:

$$f(x) = -6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2 \rightarrow \min,$$

$$x_1 + x_2 = 2, x_1, x_2 \geq 0.$$

Введем искусственные переменные  $x_3, x_4, x_5$  в уравнения (7.14). В результате получим целевую функцию

$$z = x_3 + x_4 + x_5 = 8 - 3x_1 - 3x_2 + \mu_1 + \mu_2 + 2\lambda.$$

Расчетные данные приведены в таблицах 7.1–7.4. Рамками выделены опорные элементы.

Таблица 7.1 – Первая итерация

	$x_1$	$x_2$	$\mu_1$	$\mu_2$	$\lambda$	$b$
$x_3$	4	-2	-1	0	-1	6
$x_4$	-2	4	0	-1	-1	0
$x_5$	1	1	0	0	0	2
$z$	-3	-3	1	1	2	-8

Таблица 7.2 – Вторая итерация

	$x_1$	$x_4$	$\mu_1$	$\mu_2$	$\lambda$	$b$
$x_3$	3	1/2	-1	-1/2	-3/2	6
$x_2$	-1/2	1/4	0	-1/4	-1/4	0
$x_5$	3/2	-1/4	0	1/4	1/4	2
$z$	-9/2	3/4	1	1/4	5/4	-8

Таблица 7.3 – Третья итерация

	$x_5$	$x_4$	$\mu_1$	$\mu_2$	$\lambda$	$b$
$x_3$	-2	1	-1	-1	-2	2
$x_2$	1/3	1/6	0	-1/6	-1/6	2/3
$x_1$	2/3	-1/6	0	1/6	1/6	4/3
$z$	3	0	1	1	2	-2

Таблица 7.4 – Четвертая итерация

	$x_5$	$x_4$	$\mu_1$	$\mu_2$	$x_3$	$b$
$\lambda$	1	-1/2	1/2	1/2	-1/2	-1
$x_2$			1/12	-1/12	-1/12	1/2
$x_1$			-1/12	1/12	1/12	3/2
$z$	1	1	0	0	1	0

Таким образом, получили следующее решение:

$$x_1 = 3/2; x_2 = 1/2; \lambda = -1; \mu_1 = 0; \mu_2 = 0.$$

Заметим, в уравнении (7.14) взят знак *минус* в слагаемом с  $\lambda$ . Если бы взяли знак *плюс*, то в таблице 7.1 вместо (-1) получили (+1) в первой и второй строках. В конечном решении было бы  $\lambda = +1$ .

*Пример 2.* Применим метод искусственного базиса к задаче выбора портфеля ценных бумаг. Составим начальную таблицу (таблица 7.5) из коэффициентов системы (7.15).

Таблица 7.5 – Начальная симплекс-таблица

	$x_1$	$x_2$	$x_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
$x_4$	0,116	0,062	0,072	-1	0	0	-0,48	-1	0
$x_5$	0,062	0,064	0,022	0	-1	0	-0,36	-1	0
$x_6$	0,072	0,022	0,146	0	0	-1	-0,54	-1	0
$x_7$	0,48	0,36	0,54	0	0	0	0	0	0,46
$x_8$	1	1	1	0	0	0	0	0	1
$z$	-1,730	-1,508	-1,78	1	1	1	1,38	3	-1,46

Результаты расчетов приведены в таблицах 7.6–7.10.

Таблица 7.6 – Первая итерация

	$x_4$	$x_2$	$x_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
$x_1$	8,621	0,534	0,621	-8,621	0	0	-4,138	-8,621	0
$x_5$	-0,534	0,031	-0,016	0,534	-1	0	-0,103	-0,466	0
$x_6$	-0,621	-0,016	0,101	0,621	0	-1	-0,242	-0,379	0
$x_7$	-4,138	0,103	0,242	4,138	0	0	1,986	4,138	0,46
$x_8$	-8,621	0,466	0,379	8,621	0	0	4,138	8,621	1
$z$	14,91	-0,583	-0,706	-13,91	1	1	-5,779	-11,91	-1,46

Таблица 7.7 – Вторая итерация

	$x_4$	$x_5$	$x_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
$x_1$	17,88	-17,32	0,906	-17,88	17,32	0	-2,346	-0,559	0
$x_2$	-17,32	32,40	-0,534	17,32	-32,40	0	-3,352	-15,08	0
$x_6$	-0,906	0,534	0,093	0,906	-0,534	-1	-0,297	-0,628	0
$x_7$	-2,346	-3,352	0,297	2,346	3,352	0	2,333	5,698	0,46
$x_8$	-0,559	-15,08	0,628	0,559	15,08	0	5,698	15,64	1
$z$	4,811	18,90	-1,018	-3,811	-17,90	1	-7,734	-20,71	-1,46

Таблица 7.8 – Третья итерация

	$x_4$	$x_5$	$x_6$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
$x_1$	26,75	-22,55	-9,795	-26,75	22,55	9,795	0,566	5,592	0
$x_2$	-22,55	35,49	5,773	22,55	-35,49	-5,773	-5,068	-18,71	0
$x_3$	-9,795	5,773	10,81	9,795	-5,773	-10,81	-3,214	-6,788	0
$x_7$	0,566	-5,068	-3,214	-0,566	5,068	3,214	3,289	7,717	0,46
$x_8$	5,592	-18,71	-6,788	-5,592	18,71	6,788	7,717	19,91	1
$z$	-5,158	24,78	11,00	6,158	-23,78	-10,00	-11,01	-27,62	-1,46

Таблица 7.9 – Четвертая итерация

	$x_4$	$x_5$	$x_6$	$\mu_1$	$\mu_2$	$\mu_3$	$x_8$	$\lambda_2$	$b$
$x_1$	26,34	21,18	9,298	26,34	21,18	9,298	0,073	4,132	0,073
$x_2$	-18,88	23,20	1,315	18,88	-23,20	-1,315	0,657	-5,635	0,657
$x_3$	-7,466	-2,019	7,983	7,466	2,019	-7,983	0,417	1,503	0,417
$x_7$	-1,817	2,905	-0,321	1,817	-2,905	0,321	-0,426	-0,766	0,034
$\lambda_1$	0,725	-2,425	-0,880	-0,725	2,425	0,880	0,130	2,580	0,130
$z$	2,817	-1,905	1,321	-1,817	2,905	-0,321	1,426	0,766	-0,034

Таблица 7.10 – Пятая итерация

	$x_4$	$x_5$	$x_6$	$\mu_1$	$\mu_2$	$\mu_3$	$x_8$	$x_7$	$b$
$x_1$	16,54	-5,515	-11,03	-16,54	5,515	11,03	-2,371	5,392	0,109
$x_2$	-5,515	1,838	3,676	5,515	-1,838	-3,676	3,790	-7,353	0,408
$x_3$	-11,03	3,676	7,353	11,03	-3,676	-7,353	-0,419	1,961	0,483
$\lambda_2$	2,371	-3,790	0,419	-2,371	3,790	-0,419	0,556	-1	-0,044
$\lambda_1$	-5,392	7,353	-1,961	5,392	-7,353	1,961	-1,305	3,366	0,243
$z$	1	1	1	0	0	0	1	1	0

Таким образом, получили следующее решение:  $x_1 = 0,109$ ,  $x_2 = 0,408$ ,  $x_3 = 0,483$ . Незначительное расхождение с точным решением обусловлено вычислительными погрешностями.

## 7.5. Решение с помощью симплексного преобразования таблицы коэффициентов уравнений

Эту задачу можно решить гораздо проще с помощью симплексного преобразования таблицы коэффициентов уравнений; так мы ищем начальное базисное решение. Отличием здесь является только то, что переменные  $\mu_i$  и  $x_i$  не могут быть одновременно базисными, поэтому мы делаем базисными только  $x_i$  и  $\lambda_j$ . При этом необходимо учесть, что  $\lambda_j$  могут быть любого знака, в то время как  $x_j \geq 0$ .

*Пример 1.* Применим этот метод решения к рассмотренной выше задаче

$$f(x) = -6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2 \rightarrow \min,$$

$$x_1 + x_2 = 2, x_1, x_2 \geq 0.$$

Решение сведено в таблицы 7.11–7.14.

Таблица 7.11 – Первая итерация

$x_1$	$x_2$	$\mu_1$	$\mu_2$	$\lambda$	$b$
4	-2	-1	0	-1	6
-2	4	0	-1	-1	0
1	1	0	0	0	2

Таблица 7.12 – Вторая итерация

$x_1$	$x_2$	$\mu_1$	$\mu_2$	$\lambda$	$b$
3	0	-1	-1/2	-3/2	6
-1/2	1	0	-1/4	-1/4	0
3/2	0	0	1/4	1/4	2

Таблица 7.13 – Третья итерация

$x_1$	$x_2$	$\mu_1$	$\mu_2$	$\lambda$	$b$
0	0	-1	-1	-2	2
0	1	0	-1/6	-1/6	2/3
1	0	0	1/6	1/6	4/3

Мы получили базисные переменные  $x_1, x_2$ . Теперь введем в базис переменную  $\lambda$ , не обращая внимания на ее знак.

Таблица 7.14 – Четвертая итерация

$x_1$	$x_2$	$\mu_1$	$\mu_2$	$\lambda$	$b$
0	0	1/2	1/2	1	-1
0	1	1/12	-1/12	0	1/2
1	0	-1/12	1/12	0	3/2

Итак, получили базисные переменные

$$x_1 = 3/2, x_2 = 1/2, \lambda = -1, \mu_1 = 0, \mu_2 = 0.$$

*Пример 2.* Применим метод симплексного преобразования к задаче выбора портфеля ценных бумаг. Составим таблицу из коэффициентов системы (7.15).

Таблица 7.15 – Начальная симплекс-таблица

$x_1$	$x_2$	$x_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
0,116	0,062	0,072	-1	0	0	-0,48	-1	0
0,062	0,064	0,022	0	-1	0	-0,36	-1	0
0,072	0,022	0,146	0	0	-1	-0,54	-1	0
0,48	0,36	0,54	0	0	0	0	0	0,46
1	1	1	0	0	0	0	0	1

Результаты расчетов приведены в таблицах 7.16–7.20.

Таблица 7.16 – Первая итерация

$x_1$	$x_2$	$x_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
1	0,534	0,621	-8,621	0	0	-4,138	-8,621	0
0	0,031	-0,016	0,534	-1	0	-0,103	-0,466	0
0	-0,016	0,101	0,621	0	-1	-0,242	-0,379	0
0	0,103	0,242	4,138	0	0	1,986	4,138	0,46
0	0,466	0,379	8,621	0	0	4,138	8,621	1

Таблица 7.17 – Вторая итерация

$x_1$	$x_2$	$x_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
1	0	0,906	-17,88	17,32	0	-2,346	-0,559	0
0	1	-0,534	17,32	-32,40	0	-3,352	-15,08	0
0	0	0,093	0,906	-0,534	-1	-0,297	-0,628	0
0	0	0,297	2,346	3,352	0	2,333	5,698	0,46
0	0	0,628	0,559	15,08	0	5,698	15,64	1

Таблица 7.18 – Третья итерация

$x_1$	$x_2$	$x_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
1	0	0	-26,75	22,55	9,795	0,566	5,592	0
0	1	0	22,55	-35,49	-5,773	-5,068	-18,71	0
0	0	1	9,795	-5,773	-10,81	-3,214	-6,788	0
0	0	0	-0,566	5,068	3,214	3,289	7,717	0,46
0	0	0	-5,592	18,71	6,788	7,717	19,91	1

Таблица 7.19 – Четвертая итерация

$x_1$	$x_2$	$x_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
1	0	0	-26,34	21,18	9,298	0	4,132	-0,073
0	1	0	18,88	-23,20	-1,315	0	-5,635	0,657
0	0	1	7,466	2,019	-7,983	0	1,503	0,417
0	0	0	1,817	-2,905	0,321	0	-0,766	0,034
0	0	0	-0,725	2,425	0,880	1	2,580	0,130

Таблица 7.20 – Пятая итерация

$x_1$	$x_2$	$x_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\lambda_1$	$\lambda_2$	$b$
1	0	0	-16,54	5,515	11,03	0	0	0,109
0	1	0	5,515	-1,838	-3,676	0	0	0,408
0	0	1	11,03	-3,676	-7,353	0	0	0,483
0	0	0	-2,371	3,790	-0,419	0	1	-0,044
0	0	0	5,392	-7,353	1,961	1	0	0,243

Таким образом, имеем следующее решение:

$$x_1 = 0,109, x_2 = 0,408, x_3 = 0,483.$$

Оно отличается от точного решения из-за погрешностей вычислений, но совпадает с решением, полученным методом искусственного базиса.

К сожалению, в случае положительно полуопределенной матрицы  $Q$  (т.е. матрица Гессе  $H_f(x) \geq 0$ ), этот метод может расходиться.

Лемке разработал более эффективный и простой метод решения системы (7.13), называемый *методом решения задач о дополнителности*.

## 7.6. Метод решения задач о дополнителности

Рассмотрим общую задачу отыскания неотрицательного решения системы уравнений следующего вида: найти такие векторы  $\omega$  и  $z$ , что

$$\omega = Mz + q, \quad (7.16)$$

$$\omega \geq 0, \quad z \geq 0, \quad (7.17)$$

$$\omega^T z = 0, \quad (7.18)$$

где  $M$  – матрица размера  $n \times n$ ;  $\omega, z, q$  –  $n$ -мерные векторы.

Эта задача называется *задачей о дополнителности*. Заметим, что в ее формулировке не участвует оптимизируемая целевая функция.

Рассмотрим выпуклую ЗКП

$$f(x) = c^T x + x^T Q x \rightarrow \min,$$

$$Ax \leq b, \quad x \geq 0,$$

где  $Q$  – положительно определенная или полуопределенная матрица. Функция Лагранжа имеет вид

$$L(\mu, \lambda) = f(x) - \mu^T x + \lambda^T (Ax - b).$$

Условия оптимальности Куна – Таккера для этой задачи записываются в следующем виде:

$$2Qx + c - \mu + A^T \lambda = 0,$$

$$s = -(Ax - b),$$

$$\mu^T x + s^T \lambda = 0,$$

$$x, s \geq 0, \mu, \lambda \geq 0.$$

где  $s$  – дополнительный вектор переменных для преобразования ограничения-неравенства в ограничение-равенство  $Ax - b + s = 0$ .

Сравнивая эту систему с задачей о дополнителности, можно заметить, что

$$\omega = \begin{pmatrix} \mu \\ s \end{pmatrix}; \quad z = \begin{pmatrix} x \\ \lambda \end{pmatrix}; \quad M = \begin{pmatrix} 2Q & A^T \\ -A & 0 \end{pmatrix}; \quad q = \begin{pmatrix} c \\ b \end{pmatrix}.$$

Таким образом, можно получить решение выпуклой ЗКП путем решения эквивалентной задачи о дополнителности.

Задачу линейного программирования можно также свести к задаче о дополнителности, положив  $Q = 0$ . Таким образом, метод решения задач о дополнителности можно применить также для решения ЗЛП.

*Пример 1.* Записать модель задачи о дополнителности для следующего уравнения:

$$f(x) = -6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2, \\ x_1 + x_2 \leq 2, \quad x_1, x_2 \geq 0.$$

Для этого примера имеем:

$$A = (1; 1), \quad b = (2), \quad c = \begin{pmatrix} -6 \\ 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Эквивалентная задача о дополнителности имеет вид:

$$M = \begin{pmatrix} Q + Q^T & A^T \\ -A & 0 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 4 & 1 \\ -1 & -1 & 0 \end{pmatrix}, \\ q = \begin{pmatrix} c \\ b \end{pmatrix} = \begin{pmatrix} -6 \\ 0 \\ 2 \end{pmatrix}; \quad z = \begin{pmatrix} x_1 \\ x_2 \\ \lambda \end{pmatrix}; \quad \omega = \begin{pmatrix} \mu_1 \\ \mu_2 \\ s \end{pmatrix},$$

где  $\lambda$  – скалярная величина;  $s$  – скалярная величина;  $s = x_3$  (дополнительная переменная). Таким образом, получим  $\omega = Mz + q$ .

*Определение.* Неотрицательное решение  $(\omega, z)$  системы уравнений  $\omega = Mz + q$  называется допустимым решением задачи о дополнителности.

Допустимое решение  $(\omega, z)$  задачи о дополнителности, удовлетворяющее условию дополняющей нежесткости  $\omega^T z = 0$ , называется решением задачи о дополнителности.

Условие  $\omega^T z = 0$  эквивалентно условиям  $\omega_i z_i = 0, i = 1, 2, \dots, n$ . Если все элементы вектора  $q_i \geq 0$ , то очевидное решение задачи о дополнителности будет  $\omega = q, z = 0$  (так как имеется условие  $\omega_i z_i = 0$ ). Следовательно, задача о дополнителности нетривиальна только в том случае, если имеется хотя бы один элемент  $q_i < 0$ . При этом начальное базисное решение  $\omega = q, z = 0$ , удовлетворяющее условию  $\omega^T z = 0$ , является недопустимым для задачи о дополнителности (так как  $\omega_i = q_i < 0$ ).

Алгоритм задачи о дополнителности начинает работу с недопустимого базисного решения  $\omega = q, z = 0$ . Для получения неотрицательного решения производится следующее преобразование задачи: во все уравнения системы  $\omega - Mz = q$  вводится искусственная переменная  $z_0$ :

$$\begin{aligned} \omega - Mz - e \cdot z_0 &= q, \\ \omega, z &\geq 0; \omega^T z = 0; z_0 \geq 0, \end{aligned}$$

где  $e = (1, 1, \dots, 1)^T$  – вектор, состоящий из единиц. Таким образом, можем составить начальную симплекс-таблицу (таблица 7.21).

Шаг 1. В базис вводится переменная  $z_0$ , заменяющая базисную переменную с наибольшим по абсолютной величине отрицательным значением. Пусть

$$q_s = \min_i q_i < 0,$$

в этом случае  $z_0$  заменяет в базисе переменную  $\omega_s$ . Выполнение итераций метода исключений приводит к таблице 7.22, где

$$q'_s = -q_s, q'_i = q_i - q_s, \forall i \neq s, \quad (7.19)$$

$$m'_{sj} = -m_{sj} / (-1) = m_{sj}, \forall j = 1, 2, \dots, n, \quad (7.20)$$

$$m'_{ij} = -m_{ij} + m_{sj}, \forall j = 1, 2, \dots, n, i \neq s. \quad (7.21)$$

Таблица 7.21 – Начальная симплекс-таблица

Базис	$z_1$	...	$z_s$	...	$z_n$	$z_0$	$q$
$\omega_1$	$-m_{11}$	...	$-m_{1s}$	...	$-m_{1n}$	-1	$q_1$
...	...	...	...	...	...	...	...
$\omega_s$	$-m_{s1}$	...	$-m_{ss}$	...	$-m_{sn}$	-1	$q_s$
...	...	...	...	...	...	...	...
$\omega_n$	$-m_{n1}$	...	$-m_{ns}$	...	$-m_{nn}$	-1	$q_n$

Таблица 7.22 – Новая симплекс-таблица

Базис	$z_1$	...	$z_s$	...	$z_n$	$\omega_s$	$q$
$\omega_1$	$m'_{11}$	...	$m'_{1s}$	...	$m'_{1n}$	0	$q'_1$
...	...	...	...	...	...	...	...
$z_0$	$m'_{s1}$	...	$m'_{ss}$	...	$m'_{sn}$	1	$q'_s$
...	...	...	...	...	...	...	...
$\omega_n$	$m'_{n1}$	...	$m'_{ns}$	...	$m'_{nn}$	0	$q'_n$

Шаг 2. Заметим, что переменные  $\omega_s$  и  $z_s$  в таблице 7.22 небазисные. При введении в базис одной из этих переменных условие дополненности не нарушается. Так как  $\omega_s$  выведена из базиса на предыдущей итерации, то естественно ввести в базис  $z_s$ . Следовательно, существует правило для выбора вводимой в базис небазисной переменной: в базис всегда вводится переменная, дополнительная к базисной переменной, выведенной из базиса на предыдущей итерации (правило дополняющей нежесткости). Переменные  $\omega_i, z_i, \forall i$  называются *парой взаимодополняющих переменных*.

После того как выбрана переменная, вводимая в базис, необходимо определить базисную переменную, которую надо вывести из базиса. При этом используется правило минимального отношения, аналогичное правилу в симплекс-методе ЛП. Применение этого правила гарантирует выполнение условия  $q'_i \geq 0$ . Таким образом, для вывода из базиса переменной используют соотношения

$$\min_{m'_{is} > 0} \frac{q'_i}{m'_{is}} = \frac{q'_k}{m'_{ks}} \quad (7.22)$$

для всех  $i = 1, 2, \dots, n$ , для которых значения  $m'_{is} > 0$ .

Мы выводим  $\omega_k$ , а на ее место ставим  $z_s$ . Новая таблица получается путем выполнения итерации метода исключения с  $m'_{ks}$  в качестве ведущего элемента.

Шаг 3. Поскольку переменная  $\omega_k$  выведена из базиса, то по правилу дополнительной нежесткости в базис необходимо ввести  $z_k$ . Далее изменения базиса проводятся описанным выше способом.

#### *Критерии окончания работы алгоритма*

1. Минимальное отношение достигается в строке  $s$ , а переменная  $z_0$  выводится из базиса. Полученное в результате итерации метода исключения базисное решение представляет собой решение задачи о дополнителности.

2. Правило минимального отношения неприменимо, так как все  $m'_{is} \leq 0$ ,  $i = 1, 2, \dots, n$ . Это значит, что решение задачи о дополнителности не существует. В таком случае говорят, что задача о дополнителности обладает лучевым решением, т. е. не удастся обнаружить даже локальный оптимум.

*Замечание.* Доказана сходимость метода к решению задачи о дополнителности за конечное число шагов при выполнении одного из условий:

- все элементы матрицы  $M$  положительны;
- главные миноры матрицы  $M$  больше нуля (т.е. матрица  $M$  положительно определена).

Наиболее важным применением метода решения задачи о дополнителности является решение выпуклых ЗКП. При этом матрица  $M$

может быть положительно полуопределена. Если существует решение задачи о дополнителности для  $M \geq 0$ , то метод позволяет найти его за конечное число шагов.

*Пример 2.* Для ЗКП

$$f(x) = -6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2,$$

$$x_1 + x_2 \leq 2, x_1, x_2 \geq 0,$$

как уже было показано выше, эквивалентная задача о дополнителности записывается следующим образом:

$$M = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 4 & 1 \\ -1 & -1 & 0 \end{pmatrix}, q = \begin{pmatrix} -6 \\ 0 \\ 2 \end{pmatrix}.$$

Так как в столбце  $q$  имеется отрицательный элемент, то вводим искусственную переменную  $z_0$ . Получим таблицу 7.23.

Таблица 7.23 – Начальная симплекс-таблица

Базис	$z_1$	$z_2$	$z_3$	$z_0$	$q$
$\omega_1$	-4	2	-1	-1	-6
$\omega_2$	2	-4	-1	-1	0
$\omega_3$	1	1	0	-1	2

Начальное базисное решение  $\omega_1 = -6$ ,  $\omega_2 = 0$ ,  $\omega_3 = 2$ ,  $z_1 = z_2 = z_3 = z_0 = 0$ . Получим частичное решение с использованием формул (7.19)–(7.21) и заменой  $\omega_1$  на  $z_0$  (таблица 7.24).

Таблица 7.24 – Первая итерация

Базис	$z_1$	$z_2$	$z_3$	$\omega_1$	$q$
$z_0$	4	-2	1	1	6
$\omega_2$	6	-6	0	0	6
$\omega_3$	5	-1	1	0	8

Частичное базисное решение  $z_0 = 6$ ,  $\omega_2 = 6$ ,  $\omega_3 = 8$ ,  $z_1 = z_2 = z_3 = \omega_1 = 0$ . Так как  $\omega_1$  была выведена на предыдущей итерации, то вводим в базис  $z_1$ . Используя правило наименьшего частного (7.22) для столбца  $z_1$ , получим разрешающей элемент 6. Таким образом, переменная  $z_1$  должна заменить  $\omega_2$  (таблица 7.25).

Таблица 7.25 – Вторая итерация

Базис	$\omega_2$	$z_2$	$z_3$	$\omega_1$	$q$
$z_0$	0	2	1	1	2
$z_1$	1	-1	0	0	1
$\omega_3$	0	4	1	0	3

Так как мы вывели из базиса  $\omega_2$ , то по правилу дополнительной нежесткости надо ввести в базис  $z_2$ . Разрешающий элемент, равный 4, показывает, что надо вывести из базиса  $\omega_3$  (таблица 7.26). Окончательно получаем таблицу 7.27.

Таблица 7.26 – Третья итерация

Базис	$\omega_2$	$\omega_3$	$z_3$	$\omega_1$	$q$
$z_0$	0	0	2/4	1	2/4
$z_1$	1	0	1/4	0	7/4
$z_2$	0	1	1/4	0	3/4

Таблица 7.27 – Итоговая симплекс-таблица

Базис	$\omega_2$	$\omega_3$	$z_0$	$\omega_1$	$q$
$z_3$	0	0	1	2	1
$z_1$	1	0	0	-2/4	6/4
$z_2$	0	1	0	-2/4	2/4

В таблице 7.27 видим, что в базисе остались только  $z_1, z_2, z_3$ , т.е. переменная  $z_0$  вышла из базиса. Следовательно, по первому критерию

окончания решения процесс заканчивается. Решение задачи о дополнителности имеет вид

$$z_1 = 6/4, z_2 = 2/4; z_3 = 1; \omega_1 = \omega_2 = \omega_3 = 0.$$

Таким образом, имеем следующее оптимальное решение ЗКП:

$$x_1^* = 3/2, x_2^* = 1/2, f(x^*) = -11/2.$$

## 7.7. Метод Мицеля – Хващевского

Запишем задачу квадратичного программирования в следующем виде:

$$f(x) = x^T Q x + c^T x \rightarrow \min_x, \quad (7.23)$$

$$Ax \leq b, x \geq 0, \quad (7.24)$$

где  $x$  –  $n$ -мерный вектор, компоненты которого являются неизвестными, подлежащими определению;  $Q$  – матрица размерности  $n \times n$  с компонентами  $Q_{ij}$  (положительно определенная);  $c$  –  $n$ -мерный вектор, определяющий линейный член целевой функции;  $A$  – матрица ограничений размерности  $m \times n$ ;  $b$  –  $m$ -мерный вектор ограничений.

Известно, что необходимые и достаточные условия существования решения задачи (7.23) выражаются условиями Куна – Таккера, которые имеют вид

$$2Qx + c - \mu + A^T \lambda = 0, \quad (7.25)$$

$$Ax - b + s = 0, \quad (7.26)$$

$$x, s \geq 0, \mu, \lambda \geq 0, \quad (7.27)$$

$$\mu^T x + s^T \lambda = 0, \quad (7.28)$$

где  $s$  – вектор дополнительных переменных, используемый для преобразования ограничения-неравенства (7.24) в ограничения-равенства;  $\mu, \lambda$  – векторы неопределенных множителей Лагранжа размерности  $n$  и  $m$  соответственно.

Одним из методов решения ЗКП может быть метод решения (7.25)–(7.28) как задачи линейного программирования с использованием искусственного базиса и модифицированного симплекс-метода. При этом условие (7.28) неявно учитывается в процедуре симплекс-метода при помощи правила ограниченного ввода в базис переменных  $s_j$  и  $\mu_i$ .

Этот метод применим только в случае положительной определенности матрицы  $Q$ . Если матрица  $Q$  положительно полуопределена, то метод может расходиться.

Другим вариантом решения задачи (7.25)–(7.28) может быть метод решения задачи о дополнителности, разработанный Лемке. В процедуре решения этим методом также предусматривается введение искусственной переменной. Недостатком метода является его громоздкость.

Поэтому предлагается новый алгоритм решения ЗКП, на наш взгляд, более простой в численной реализации.

*Описание алгоритма.* Разрешим систему (7.25), (7.26) относительно переменных  $x_i$  и  $\lambda_j$ . Для этого введем составные векторы  $y$ ,  $g$  размерности  $n + m$  и матрицу  $M$  размерности  $(n + m) \times (n + m)$ :

$$y = \begin{pmatrix} x \\ \lambda \end{pmatrix}, g = \begin{pmatrix} -c + \mu \\ b - s \end{pmatrix} = \mathfrak{G} + Ep, M = \begin{pmatrix} 2Q & A^T \\ A & 0 \end{pmatrix}, \quad (7.29)$$

где  $\mathfrak{G} = \begin{pmatrix} -c \\ b \end{pmatrix}$ ;  $p = \begin{pmatrix} \mu \\ -s \end{pmatrix}$ ;  $E$  – единичная матрица размерности  $(n + m) \times (n + m)$ . Размерность векторов  $\mathfrak{G}$  и  $p$  равна  $n + m$ .

В результате получим систему линейных алгебраических уравнений размерности  $n + m$

$$My = g. \quad (7.30)$$

Если определитель матрицы  $\det M$  отличен от нуля, то существует единственное решение системы (7.30). Это решение будет зависеть от параметров  $\mu$  и  $s$ . Так как параметры  $\mu$  и  $s$  не определены, то решить систему (7.30) можно лишь при дополнительных условиях (7.27) и (7.28), которые полностью доопределяют исходную задачу.

Суть ограничений (7.27) и (7.28) состоит в том, что если все компоненты  $x_i \geq 0$ ,  $i = 1, 2, \dots, n$ ,  $\lambda_j \geq 0$ ,  $j = 1, 2, \dots, m$ , то соответствующие им компоненты  $\mu_i$  и  $s_j$  равны нулю. Следовательно, мы можем положить  $p = 0$  и решать систему

$$My = \vartheta \quad (7.31)$$

любым классическим методом (например, методом симплексного преобразования системы либо методом Гаусса). Сформулируем алгоритм решения задачи (7.30).

Шаг 1. Решаем систему (7.31) и находим  $y$ . Если все  $y_i \geq 0$ ,  $i = 1, 2, \dots, n + m$ , то полагаем

$$x_j = y_j, \quad j = 1, 2, \dots, n + m, \quad (7.32)$$

и заканчиваем вычисления.

Если полученное решение (7.31) содержит отрицательные компоненты  $y_i < 0$ , то переход на шаг 2.

Шаг 2. Решаем систему (7.30) методом симплексного преобразования. Подобная процедура используется при решении ЗЛП для поиска одной из угловых точек. Поскольку решение системы (7.30) единственно, то найденное решение и будет искомым. Отличие метода симплексного преобразования при решении данной задачи состоит в необходимости учета ограничений (7.27), (7.28), которые и обеспечивают единственность решения (разумеется, если оно вообще существует). Модифицированный алгоритм симплексного преобразования заключается в следующем. Предполагается, что все компоненты  $\vartheta_i > 0$ . (Если имеются  $\vartheta_i < 0$ , соответствующие строки системы (7.30) умножаем на  $-1$ .) Переписываем систему (7.30) в виде

$$My - Ep = \vartheta \quad (7.33)$$

или

$$Wz = \vartheta, \quad (7.34)$$

где

$$W = (M, -E), \quad z = (y, p)^T. \quad (7.35)$$

Этап 1. Среди столбцов у коэффициентов  $W_{ij}$  матрицы  $W$  выбирается столбец, в котором имеется хотя бы один положительный элемент (если таких столбцов нет, то решение не существует). Если в этом столбце несколько положительных элементов, то выбирается тот, который дает наименьшее частное от деления элементов вектора  $\mathfrak{G}$  на соответствующие элементы этого столбца. Выделенный таким образом разрешающий элемент обозначим как  $W_{ij}^0$ .

Этап 2. Рассчитываются элементы матрицы  $W_{ij}^1$  и столбца  $\mathfrak{G}_i^1$  по следующим формулам:

$$W_{kl}^1 = W_{kl}^0 / W_{kj}^0, l = 1, 2, \dots, L, L = 2(n + m),$$

$$W_{il}^1 = W_{il}^0 - \frac{W_{ij}^0}{W_{kj}^0} \cdot W_{kl}^0, l = 1, 2, \dots, L; i = 1, 2, \dots, n + m; i \neq k,$$

$$\mathfrak{G}_k^1 = \frac{\mathfrak{G}_k^0}{W_{kj}^0},$$

$$\mathfrak{G}_i^1 = \mathfrak{G}_i^0 - \frac{W_{ij}^0}{W_{kj}^0} \cdot \mathfrak{G}_k^0, i = 1, 2, \dots, n + m; i \neq k.$$

Переход на этап 1. Столбец с номером  $j$  в матрице  $W^1$  состоит из нулей и единицы ( $W_{kj}^1 = 1$ ). При поиске разрешающего элемента в матрице  $W^1$  мы не рассматриваем столбец с номером  $j + N$ , где  $N = n + m$ , который соответствует переменной  $p_j$  (условие дополнительной нежесткости (7.28)). Если при поиске разрешающего элемента в матрице  $W^1$  мы окажемся в столбце с номером  $N + r$ , то это значит, что на следующем этапе при работе с матрицей  $W^2$  мы исключаем из рассмотрения столбец с номером  $r$ , соответствующий переменной  $y_r$  (условие (7.28)).

После выполнения  $N = n + m$  преобразований матрицы  $W$  и вектора  $\mathfrak{G}$  мы получим матрицу  $W^N$ , в которой  $N$  столбцов состоят из нулей и единиц, причем единичные элементы расположены в разных

строках и разных столбцах. Номера столбцов, состоящих из нулей и единиц, соответствуют тем номерам переменных, которые совпадают с номерами строк, содержащих единицы на пересечении столбцов, состоящих из нулей и единицы. Значения этих переменных равны значениям компонент вектора  $\mathfrak{G}^N$ , расположенных в указанных строках.

После того как решение  $z$  системы (7.34) найдено, вычисляем компоненты искомого решения:

$$x_j = z_j, j = 1, 2, \dots, n. \quad (7.36)$$

*Пример 1.* Рассмотрим решение следующей задачи квадратичного программирования:

$$\begin{aligned} f(x) &= x_1^2 + 2x_2^2 - 2x_1x_2 - 2x_1 - 6x_2 \rightarrow \min, \\ x_1 + x_2 &\leq 2, \\ -x_1 + 2x_2 &\leq 2, \\ x_1, x_2 &\geq 0. \end{aligned}$$

Ее искомое решение имеет вид

$$\begin{pmatrix} 2 & -2 & 1 & -1 \\ -2 & 4 & 1 & 2 \\ 1 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 2 \\ 2 \end{pmatrix} \Rightarrow y = \begin{pmatrix} 2/3 \\ 4/3 \\ 26/9 \\ -4/9 \end{pmatrix}.$$

Последняя компонента получилась отрицательной, поэтому переходим на шаг 2.

Шаг 2 алгоритма при решении данной задачи выполняется 4 раза. Результаты вычислений приведены в таблице 7.28.

Таблица 7.28 – Результаты вычислений

$x_1$	$x_2$	$\lambda_1$	$\lambda_2$	$\mu_1$	$\mu_2$	$s_1$	$s_2$	$\mathfrak{G}$
<span style="border: 1px solid black; padding: 2px;">2</span>	-2	1	-1	-1	0	0	0	2
-2	4	1	2	0	-1	0	0	6
1	1	0	0	0	0	1	0	2
-1	2	0	0	0	0	0	1	2

Окончание таблицы 7.28

$x_1$	$x_2$	$\lambda_1$	$\lambda_2$	$\mu_1$	$\mu_2$	$s_1$	$s_2$	$\theta$
1	-1	1/2	-1/2	-1/2	0	0	0	1
0	2	2	1	-1	-1	0	0	8
0	2	-1/2	1/2	1/2	0	1	0	1
0	1	1/2	-1/2	-1/2	0	0	1	3
1	0	1/4	-1/4	-1/4	0	1/2	0	3/2
0	0	5/2	1/2	-3/2	-1	-1	0	7
0	1	-1/4	1/4	1/4	0	1/2	0	1/2
0	0	3/4	-3/4	-3/4	0	-1/2	1	5/2
1	0	0	-3/10	-4/10	1/10	6/10	0	0,8
0	0	1	1/5	-3/5	-2/5	-2/5	0	2,8
0	1	0	3/10	3/20	-1/10	4/10	0	1,2
0	0	0	-9/10	-3/10	3/10	-2/10	1	0,4

В результате получим следующее оптимальное решение:

$$x_1^* = 0,8, x_2^* = 1,2, f(x^*) = -7,2.$$

*Пример 2.* Записать модель задачи о дополнителности и решить ее для портфеля ценных бумаг:

$$\begin{aligned}
 & f(x) = \\
 & = 0,058x_1^2 + 0,032x_2^2 + 0,073x_3^2 + 0,062x_1x_2 + 0,072x_1x_3 + 0,022x_2x_3 \rightarrow \min, \\
 & \quad -0,48x_1 - 0,36x_2 - 0,54x_3 \leq -0,46, \\
 & \quad x_1 + x_2 + x_3 \leq 1; x_1, x_2, x_3 \geq 0.
 \end{aligned}$$

Для этой задачи имеем:

$$A = \begin{pmatrix} -0,48 & -0,36 & -0,54 \\ 1 & 1 & 1 \end{pmatrix}, b = \begin{pmatrix} -0,46 \\ 1 \end{pmatrix}, c = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\mathfrak{Q} = \begin{pmatrix} -c_1 \\ -c_2 \\ -c_3 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -0,46 \\ 1 \end{pmatrix}, \quad y = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix}, \quad p = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ -s_1 \\ -s_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{pmatrix},$$

$$M = \begin{pmatrix} 0,116 & 0,062 & 0,72 & -0,48 & 1 \\ 0,062 & 0,064 & 0,022 & -0,36 & 1 \\ 0,072 & 0,022 & 0,146 & -0,54 & 1 \\ -0,48 & -0,36 & -0,54 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}.$$

Рассмотрим решение системы  $My = \mathfrak{Q}$ :

$$\begin{pmatrix} 0,116 & 0,062 & 0,072 & -0,48 & 1 \\ 0,062 & 0,064 & 0,022 & -0,36 & 1 \\ 0,072 & 0,022 & 0,146 & -0,54 & 1 \\ -0,48 & -0,36 & -0,54 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -0,46 \\ 1 \end{pmatrix} \Rightarrow y = \begin{pmatrix} 0,109 \\ 0,408 \\ 0,483 \\ 0,243 \\ 0,044 \end{pmatrix}$$

Так как все компоненты положительные, то мы получили сразу решение нашей задачи:  $x_1 = 0,109$ ,  $x_2 = 0,408$ ,  $x_3 = 0,483$ .

Описанный алгоритм имеет следующие преимущества по сравнению с упомянутыми выше алгоритмами.

Для большинства практических случаев решение задачи (7.23) сводится к решению системы (7.26), которое не представляет трудностей.

При решении задачи (7.26), записанной в общем виде, нет необходимости вводить искусственные переменные и вспомогательную целевую функцию.

## Вопросы для самопроверки

1. Запишите общую задачу квадратичного программирования.
2. Запишите условие Куна – Таккера для задач квадратичного программирования.
3. Как решить задачу квадратичного программирования с помощью симплексного преобразования таблицы коэффициентов уравнений?
4. Как решить задачу квадратичного программирования с помощью искусственного базиса.
5. Запишите задачу о дополнителности.
6. В чем заключается метод Лемке решения задачи о дополнителности?
7. Изложите метод решения задачи квадратичного программирования Мицеля – Хващевского.

## ГЛАВА 8. ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

### 8.1. Общая постановка задачи динамического программирования

*Динамическое программирование* (ДП) – метод оптимизации, приспособленный к операциям, в которых процесс принятия решения может быть разбит на этапы (шаги). Такие операции называются *многошаговыми*. Основоположником ДП является американский математик Р. Беллман, обративший внимание на эти задачи в 50-х годах XX века.

Модели ДП применяются:

- при решении задач управления инвестиционным портфелем;
- в моделях управления банкротством предприятий;
- при управлении малым бизнесом;
- при управлении запасами, в которых необходимо установить моменты их пополнения и размер пополняющего заказа;
- при разработке принципов календарного планирования производства и выравнивания занятости в условиях колеблющегося спроса на продукцию;
- при распределении дефицитных капитальных вложений между возможными новыми направлениями их использования;
- при составлении календарных планов текущего и капитального ремонта сложного оборудования и его замены;
- при разработке долгосрочных правил замены выбывающих из эксплуатации основных фондов и т.д.

Приведем общую постановку задачи ДП. Рассматривается управляемый процесс, например экономический процесс распределения средств между предприятиями. Процесс описывается  $n$ -мерным вектором  $x = (x_1, x_2, \dots, x_n)$ , т.е. точкой в пространстве  $R^n$ , которое называют фазовым пространством. В результате управления система (объект управления)  $S$  переводится из начального состояния  $x^{(0)}$  в состояние  $x^{(N)}$ . Предположим, что управление можно разбить на  $N$  шагов, т.е. решение принимается последовательно на каждом шаге,

а управление, переводящее систему  $S$  из начального состояния в конечное, представляет собой совокупность  $N$  пошаговых управлений.

Обозначим через  $u^{(k)}$  управление на  $k$ -м шаге ( $k = 1, 2, \dots, N$ ). Управление  $u^{(k)}$  может быть числом, точкой в  $m$ -мерном пространстве, качественным признаком. В общем случае  $u$  представляет собой  $m$ -мерный вектор управления, т.е.  $u \in R^m$ .

В результате действия управления на каждом шаге получаем последовательность состояний  $x^{(0)}, x^{(1)}, \dots, x^{(N)}$ , которую изобразим кружками (рисунок 8.1).

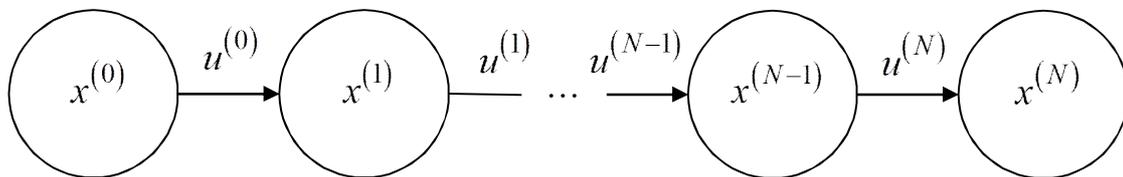


Рисунок 8.1 – Последовательность состояний системы при управлении ею

Показатель эффективности рассматриваемого управляемого процесса характеризуется целевой функцией, которая зависит от начального состояния и управления:

$$J = J(x^{(0)}, u). \quad (8.1)$$

Сделаем несколько предположений.

1. Состояние  $x^{(k)}$  системы в конце  $k$ -го шага зависит только от предшествующего состояния  $x^{(k-1)}$  и управления на  $k$ -м шаге  $u^{(k)}$  и не зависит от более ранних предшествующих состояний и управлений (принцип отсутствия «последствия»). Сформулированное положение записывается в виде уравнений

$$x^{(k)} = f_k(x^{(k-1)}, u^{(k)}), k = 1, 2, \dots, N, \quad (8.2)$$

которые называются *уравнениями состояний*.

2. Целевая функция (8.1) является аддитивной по показателю эффективности каждого шага. Обозначим показатель эффективности  $k$ -го шага как

$$J_k = J_k(x^{(k-1)}, u^{(k)}), \quad k = 1, \dots, N, \quad (8.3)$$

тогда

$$J(x, u) = \sum_{k=1}^N J_k = \sum_{k=1}^N J_k(x^{(k-1)}, u^{(k)}). \quad (8.4)$$

Сформулируем задачу пошаговой оптимизации (задачу ДП): определить такое допустимое управление  $u$ , переводящее систему  $S$  из состояния  $x^{(0)}$  в состояние  $x^{(N)}$ , при котором целевая функция (8.4) принимает наибольшее (наименьшее) значение.

Предположим, что на фазовую траекторию и выбор управлений наложены ограничения

$$x^{(k)} \in X_k, \quad k = 1, \dots, N-1, \quad (8.5)$$

$$u^{(k)} \in U_k(x^{(k-1)}), \quad k = 1, \dots, N, \quad (8.6)$$

где  $X_k$  и  $U_k(x^{(k-1)})$  – заданные множества в пространствах  $R^n$  и  $R^m$  соответственно, причем множество  $U_k(x^{(k-1)})$  зависит, вообще говоря, от предыдущего состояния  $x^{(k-1)}$   $k$ -го шага.

Ограничения на начальное и конечное состояния процесса

$$x^{(0)} \in X_0, \quad x^{(N)} \in X_N$$

называются *начальными* и *конечными условиями*.

Пусть  $u = \{u^{(1)}, u^{(2)}, \dots, u^{(N)}\}$  – управление процессом, удовлетворяющее ограничениям (8.6) и переводящее его из начального состояния  $x^{(0)} \in X_0$  в конечное состояние  $x^{(N)} \in X_N$  в соответствии с уравнением (8.2) и учетом ограничений (8.5). Обозначим множество всех таких управлений буквой  $U$ .

Многошаговая задача оптимизации формулируется следующим образом: среди всех управлений  $u \in U$  выбрать такое  $u_{\text{opt}} = \{u_{\text{opt}}^{(1)}, \dots, u_{\text{opt}}^{(N)}\}$ , для которого целевая функция (8.3) принимает минимальное или максимальное (в зависимости от смысла задачи) значение.

Управление  $u_{\text{opt}}$  и соответствующая ему фазовая траектория  $x_{\text{opt}}$  называются *оптимальными*.

Условие многошаговой задачи оптимизации будем записывать в виде

$$J(x, u) = \sum_{k=1}^N J_k = \sum_{k=1}^N J_k(x^{(k-1)}, u^{(k)}) \rightarrow \underset{u}{\text{extr}}, \quad (8.7)$$

$$x^{(k)} = f_k(x^{(k-1)}, u^{(k)}), \quad k = 1, \dots, N, \quad (8.8)$$

$$u^{(k)} \in U_k(x^{(k-1)}), \quad k = 1, \dots, N, \quad (8.9)$$

$$x^{(0)} \in X_0, \quad x^{(N)} \in X_N.$$

Выделим особенности модели ДП:

- 1) задача оптимизации интерпретируется как  $N$ -шаговый процесс управления;
- 2) целевая функция равна сумме целевых функций каждого шага;
- 3) выбор управления на  $k$ -м шаге зависит только от состояния системы к этому шагу и не влияет на предшествующие шаги (нет обратной связи).

Состояние  $x^{(k)}$  после  $k$ -го шага управления зависит только от предшествующего состояния  $x^{(k-1)}$  и управления  $u^{(k)}$  (отсутствие последствий).

Существуют различные способы решения подобных задач, применяемые в зависимости от вида функций, ограничений, размерности и т.п. Рассмотрим вычислительную схему ДП, которая будет безразличной к способам задания функций и ограничений. Вычислительная схема связана с принципом оптимальности и использует рекуррентные соотношения.

## 8.2. Принцип оптимальности и уравнения Беллмана

Принцип оптимальности впервые был сформулирован Р. Беллманом в 1953 г. Суть этого принципа состоит в следующем: *каково бы ни было состояние  $x$  системы в результате какого-либо числа шагов, на ближайшем шаге нужно выбирать управление так, чтобы оно в совокупности с оптимальным управлением на всех последующих шагах приводило к оптимальному выигрышу на всех оставшихся шагах, включая данный.* Беллман четко сформулировал и условия, при которых принцип верен. Основное требование – процесс управления должен быть без обратной связи, т.е. управление на данном шаге не должно оказывать влияние на предшествующие шаги.

Принцип оптимальности утверждает, что для любого процесса без обратной связи оптимальное управление таково, что оно является оптимальным для любого подпроцесса по отношению к исходному состоянию этого подпроцесса. Поэтому решение на каждом шаге оказывается наилучшим с точки зрения управления в целом. Если изобразить геометрически оптимальную траекторию в виде ломаной линии, то любая часть этой ломаной будет являться оптимальной траекторией относительно начала и конца.

Однако существует и обратная схема решения задачи ДП, в которой процесс решения начинается с последнего шага.

Изложим здесь прямую схему решения задачи ДП, которая отличается тем, что процесс решения начинается с первого шага. При этом смысл переменных состояния  $x^{(k)}$  и функции выигрыша  $J_k(x^{(k-1)}, u^{(k)})$  меняется, управление на  $k$ -м шаге будет зависеть от текущего состояния. Кроме того, функция выигрыша будет зависеть не от начального, а от конечного состояния.

Вместо исходной задачи ДП с фиксированным числом шагов  $N$  и начальным состоянием  $x^{(0)}$  рассмотрим последовательность задач, полагая  $N = 1, 2, \dots$  при различных  $x$  – одношаговую, двухшаговую и т.д., используя принцип оптимальности.

Шаг 1:  $x^{(0)}$  – начальное состояние системы,  $x^{(1)}$  – конечное состояние,  $u^{(1)}$  – управление на 1-м шаге,  $J_1(x^{(1)}, u^{(1)})$  – целевая функция (выигрыш) 1-го шага.

Согласно принципу оптимальности управление  $u^{(1)}$  нужно выбирать так, чтобы для любых состояний  $x^{(0)}$  получить максимум целевой функции на этом шаге.

Обозначим через  $J_1^{\max}(x^{(1)})$  условный максимум целевой функции – показателя эффективности 1-го шага при условии, что к концу этого шага система  $S$  была в произвольном состоянии  $x^{(1)}$ , а управление на этом шаге было оптимальным.

Очевидно, что

$$J_1^{\max}(x^{(1)}) = \max_{u^{(1)} \in U_1(x^{(1)})} (J_1(u^{(1)})). \quad (8.10)$$

Максимизация ведется по всем допустимым управлениям  $U_1$ . Решение  $u^{(1)}$ , при котором достигается максимум (8.10), также зависит от  $x^{(1)}$  и называется *условным оптимальным управлением на 1-м шаге*. Оно обозначается как  $\bar{u}^{(1)}(x^{(1)})$ .

Решив задачу локальной оптимизации (8.10), найдем для всех возможных состояний  $x^{(1)}$  две функции:  $J_1^{\max}(x^{(1)})$  и  $\bar{u}^{(1)}(x^{(1)})$ .

Рассмотрим теперь двухшаговую задачу: присоединим к 1-му шагу 2-й шаг. Значение целевой функции на двух первых шагах равно

$$J_2(x^{(2)}, u^{(2)}) + J_1^{\max}(x^{(1)}). \quad (8.11)$$

Согласно принципу оптимальности для любых  $x^{(2)}$  решение нужно выбирать так, чтобы оно вместе с оптимальным управлением на первом шаге приводило к максимуму целевой функции на двух первых шагах, т.е. нужно найти максимум выражения (8.11) по всем допустимым управлениям  $u^{(2)}$ . Максимум этой суммы зависит от  $x^{(2)}$

и обозначается  $J_2^{\max}(x^{(2)})$ . Он называется *условным максимумом целевой функции при оптимальном управлении на двух первых шагах*. Соответствующее управление на 2-м шаге обозначим  $\bar{u}^{(2)}(x^{(2)})$ , его будем называть *условным оптимальным управлением на 2-м шаге*:

$$J_2^{\max}(x^{(2)}) = \max_{u^{(2)} \in U_2} \left\{ J_2(x^{(2)}, u^{(2)}) + J_1^{\max}(x^{(1)}) \right\}. \quad (8.12)$$

Следует отметить, что выражение, стоящее в фигурных скобках равенства (8.12), зависит только от  $x^{(2)}$  и  $u^{(2)}$ , так как  $x^{(1)}$  можно найти из уравнения состояний (8.2) при  $k = 2$

$$x^{(2)} = f^{(2)}(x^{(1)}, u^{(2)})$$

и подставить вместо  $x^{(1)}$  в функцию  $J_1^{\max}(x^{(1)})$ .

В результате максимизации только по одной переменной  $u^{(2)}$  согласно уравнению (8.12) вновь получаются две функции:  $J_2^{\max}(x^{(2)})$  и  $\bar{u}^{(2)}(x^{(2)})$ .

Далее рассматривается трехшаговая задача: к двум первым шагам присоединяется 3-й шаг и т.д. (рисунок 8.2)

Обозначим через  $J_k^{\max}(x^{(k)})$  условный максимум целевой функции, полученный при оптимальном управлении на первых  $k$  шагах, начиная с 1-го по  $k$ -й, при условии, что к концу  $k$ -го шага система находилась в состоянии  $x^{(k)}$ . Фактически эта функция имеет вид

$$J_k^{\max}(x^{(k)}) = \max_{U_k} \left\{ J_k(x^{(k)}, u^{(k)}) + J_{k-1}^{\max}(x^{(k-1)}) \right\}, k = 1, 2, \dots, N-1. \quad (8.13)$$

Управление  $u^{(k)}$  на  $k$ -м шаге, при котором достигается максимум (8.13), обозначается  $\bar{u}^{(k)}(x^{(k)})$  и называется *условным оптимальным управлением на  $k$ -м шаге* (так же, как и в равенстве (8.12), в правую

часть уравнения (8.13) вместо  $x^{(k-1)}$  следует подставить выражение, найденное из уравнения состояния  $x^{(k)} = f^{(k)}(x^{(k-1)}, u^{(k)})$ .

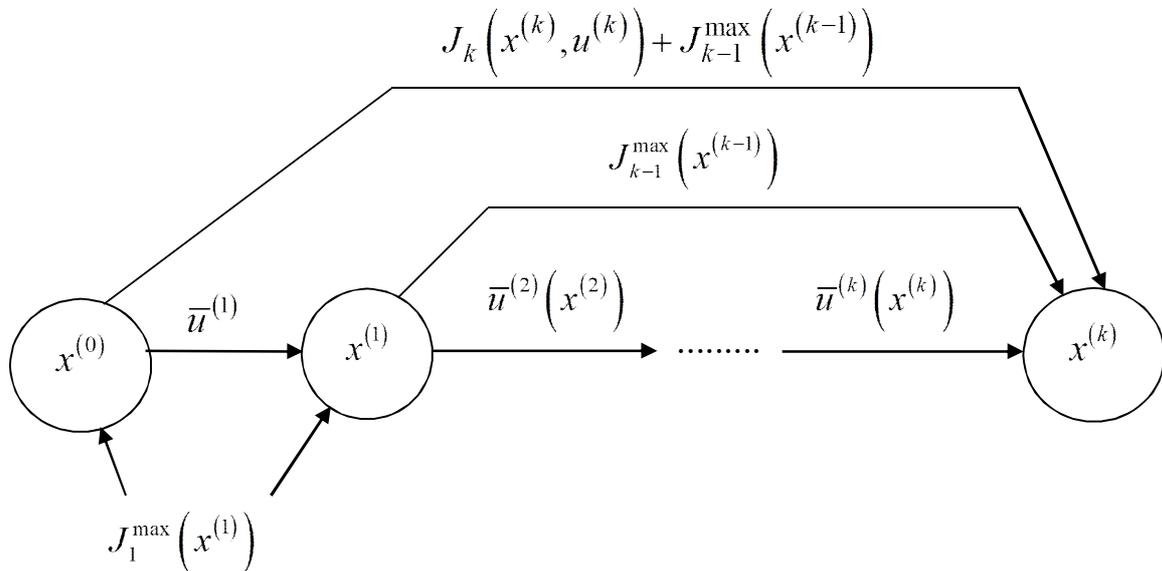


Рисунок 8.2 – Принцип оптимальности Беллмана

Функцию  $J_{k-1}^{\max}(x^{(k-1)})$  будем называть *функцией Беллмана* первых  $(k-1)$  шагов.

Уравнения (8.13) называют *уравнениями Беллмана*. Это рекуррентные соотношения, позволяющие найти последующее значение функции, зная предыдущие. Если из уравнения (8.10) найти  $J_1^{\max}(x^{(1)})$ , то при  $k=2$  из соотношения (8.13) можно определить выражение для  $J_2^{\max}(x^{(2)})$  и соответствующее управление  $\bar{u}^{(2)}(x^{(2)})$ . Далее, зная  $J_2^{\max}(x^{(2)})$ , находим  $J_3^{\max}(x^{(3)})$  и  $\bar{u}^{(3)}(x^{(3)})$ , используя (8.13) и уравнение состояния (8.2).

На последнем шаге получим  $J_N^{\max}(x^{(N)})$ . Зная конечное состояние  $x^{(N)}$ , найдем максимальное значение целевой функции на последнем шаге и условное оптимальное управление  $\bar{u}^{(N)}(x^{(N)})$ . Процесс



### 8.3. Задача о распределении средств между предприятиями

Планируется деятельность четырех промышленных предприятий (системы) на очередной год. Начальные средства  $x^{(0)} = 5$  у.е. Размеры вложения в каждое предприятие кратны 1 у.е. Средства  $u$ , выделенные  $k$ -му предприятию ( $k = 1, 2, 3, 4$ ), приносят в конце года прибыль  $J_k(u)$ . Функции  $J_k(u)$  заданы таблично (таблица 8.1). Принимаются также следующие условия

- а) прибыль  $J_k(u)$  не зависит от вложения средств в другие предприятия;
- б) прибыль от каждого предприятия выражается в одних условных единицах;
- в) суммарная прибыль равна сумме прибылей, полученных от каждого предприятия.

Определить, какое количество средств нужно выделить каждому предприятию, чтобы суммарная прибыль была наибольшей.

Таблица 8.1 – Заданные функции  $J_k(u)$

$u$	1	2	3	4	5
$J_1(u)$	8	10	11	12	18
$J_2(u)$	6	9	11	13	15
$J_3(u)$	3	4	7	11	18
$J_4(u)$	4	6	8	13	16

Обозначим через  $u^{(k)}$  количество средств, выделенных  $k$ -му предприятию. Суммарная прибыль

$$J = \sum_{k=1}^4 J_k(x^{(k)}, u^{(k)}). \quad (8.16)$$

Переменные  $u$  удовлетворяют ограничениям

$$\sum_{k=1}^4 u^{(k)} = 5, u^{(k)} \geq 0, k = 1, 2, 3, 4. \quad (8.17)$$

Требуется найти переменные  $u^{(1)}, u^{(2)}, u^{(3)}, u^{(4)}$ , удовлетворяющие системе ограничений (8.17) и обращающие в максимум функцию (8.16).

Процесс решения задачи распределения средств  $x^{(0)}$  можно рассматривать как 4-шаговый, номер шага совпадает с номером предприятия; переменные  $u^{(1)}, u^{(2)}, u^{(3)}, u^{(4)}$  – управление соответственно на I, II, III, IV шагах. Конечное состояние процесса распределения  $x^{(4)} = 5$  у.е., так как все средства должны быть вложены в производство. Схема распределения средств показана на рисунке 8.3.

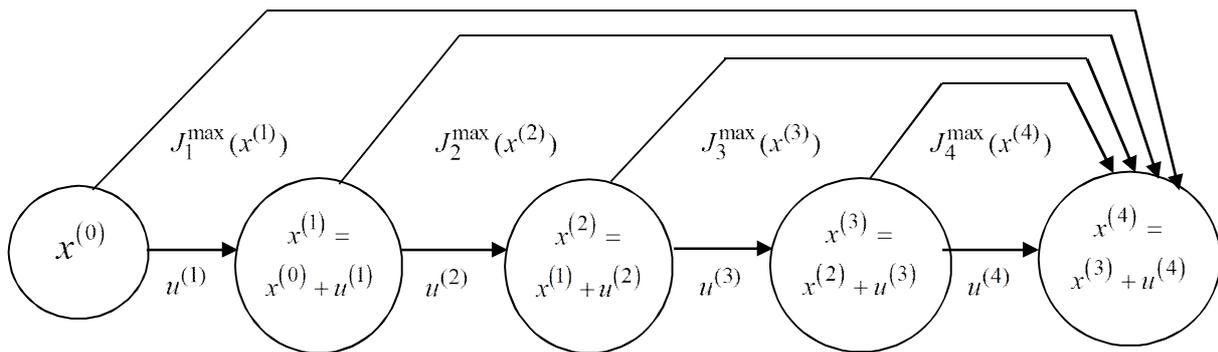


Рисунок 8.3 – Схема распределения средств

Уравнения состояний (8.2) в данной задаче имеют вид

$$x^{(k)} = x^{(k-1)} + u^{(k)}, k = 1, 2, 3, 4, \quad (8.18)$$

где  $x^{(k)}$  – параметр состояния, т.е. количество средств, распределенных после  $k$ -го шага, или средства, которые получают первые  $k$  предприятий.

Введем функцию  $J_k^{\max}(x^{(k)})$ ,  $k = 1, 2, 3, 4$ , – условную оптимальную прибыль, полученную от 1-го предприятия ( $k=1$ ), 1-го и 2-го предприятий ( $k=2$ ), 1-го, 2-го и 3-го предприятий ( $k=3$ ), всех четырех предприятий ( $k=4$ ).

Допустимые управления на  $k$ -м шаге удовлетворяют условию  $0 \leq u^{(k)} \leq x^{(k)}$  – либо  $k$ -му предприятию ничего не выделяем ( $u^{(k)} = 0$ ), либо выделяем не больше того, что выделено на  $k$ -ом шаге ( $u^{(k)} \leq x^{(k)}$ ).

Условная оптимальная прибыль вычисляется в соответствии с уравнениями:

$$J_1^{\max}(x^{(1)}) = \max_{0 \leq u^{(1)} \leq x^{(1)}} J_1(u^{(1)});$$

$$J_2^{\max}(x^{(2)}) = \max_{0 \leq u^{(2)} \leq x^{(2)}} \left( J_2(u^{(2)}) + J_1^{\max}(x^{(1)}) \right);$$

$$J_3^{\max}(x^{(3)}) = \max_{0 \leq u^{(3)} \leq x^{(3)}} \left( J_3(u^{(3)}) + J_2^{\max}(x^{(2)}) \right);$$

$$J_4^{\max}(x^{(4)}) = \max_{0 \leq u^{(4)} \leq x^{(4)}} \left( J_4(u^{(4)}) + J_3^{\max}(x^{(3)}) \right).$$

Последовательно ищем максимумы, проводя оптимизацию каждого шага (см. рисунок 8.3).

Шаг I. В таблице 8.1 прибыли  $J_1(u)$  монотонно возрастают, поэтому все средства, распределенные 1-му предприятию, следует вложить в него. Тогда для возможных значений  $x^{(1)} = 0, 1, \dots, 5$  получим

$$J_1^{\max}(x^{(1)}) = J_1(x^{(1)}) \text{ и } \bar{u}^{(1)}(x^{(1)}) = x^{(1)}.$$

Шаг II. Величина  $x^{(2)}$  может принимать значения 0, 1, 2, 3, 4, 5 (например,  $x^{(2)} = 0$ , если первые два предприятия ничего не получили и все средства отданы 3-му и 4-му предприятиям;  $x^{(2)} = 5$ , если все средства отданы 1-му и 2-му предприятиям, а 3-му и 4-му предприятиям ничего не досталось, и т.д.). В зависимости от этого выбираем управление  $0 \leq u^{(2)} \leq x^{(2)}$  и вычисляем значения суммы  $J_2(u^{(2)}) + J_1^{\max}(x^{(1)})$ . Для каждого  $x^{(2)}$  наибольшее из этих значений

есть оптимальная прибыль  $J_2^{\max}(x^{(2)})$ , полученная при оптимальном распределении средств  $x^{(2)}$  между 1-м и 2-м предприятиями, т.е.

$$J_2^{\max}(x^{(2)}) = \max_{0 \leq u^{(2)} \leq x^{(2)}} \left( J_2(u^{(2)}) + J_1^{\max}(x^{(1)}) \right).$$

Результаты расчетов приведены в таблице 8.2. Для каждого значения  $x^{(2)}$  при  $k=2$  величины  $J_2^{\max}(x^{(2)})$  и  $\bar{u}^{(2)}(x^{(2)})$  помещены в столбцы 5 и 6 соответственно. Первые слагаемые в столбце 4 – значения  $J_2(u^{(2)})$ , взяты из таблицы 8.1 (строка 2), вторые слагаемые – из строки 1 таблицы 8.1.

Шаг III. Результаты оптимизации в соответствии с выражением для  $J_3^{\max}(x^{(3)})$  представлены в таблице 8.2 при  $k=3$  в столбцах 7–9.

Для всевозможных значений  $x^{(3)}$  величины  $J_3^{\max}(x^{(3)})$  и  $\bar{u}^{(3)}(x^{(3)})$  даны в столбцах 8 и 9 соответственно; первые слагаемые в столбце 7 – значения  $J_3(u^{(3)})$ , взяты из таблицы 8.1, а вторые слагаемые взяты из столбца 5 таблицы 8.2 при  $x^{(3)} = x^{(2)} + u^{(3)}$ .

Шаг IV. Оптимизация в соответствии с выражением для  $J_4^{\max}(x^{(4)})$  приведена в таблице 8.2 при  $k=4$ .

Дадим пояснение этим расчетам при  $x^{(4)} = 5$ . Если  $u^{(4)} = 0$ , т.е. 4-му предприятию ничего не даем, то прибыль, полученная от первых трех предприятий при условии, что все средства между ними распределены оптимально, равна  $J_4(0) + J_3^{\max}(5) = 0 + 22 = 22$  (величина  $J_3^{\max}(5)$  взята из столбца 8 таблицы 8.2 при  $x^{(4)} = 5$ ). Если  $u^{(4)} = 1$ , то  $x^{(3)} = 4$ . Суммарная прибыль от четырех предприятий при условии, что 4 у.е. средств между первыми тремя предприятиями распределены оптимально, равна  $J_4(1) + J_3^{\max}(4) = 4 + 20 = 24$  (значение  $J_4(1)$  взято из строки 4 таблицы 8.1, а  $J_3^{\max}(4)$  – из столбца 8 таблицы 8.2).

Аналогично при  $u^{(4)} = 2, x^{(3)} = 3$  будет  $J_4(2) + J_3^{\max}(3) = 6 + 17 = 23$ .

Продолжая этот процесс, получим:

при  $u^{(4)} = 3, x^{(3)} = 2 - J_4(3) + J_3^{\max}(2) = 8 + 14 = 22$ ;

при  $u^{(4)} = 4, x^{(3)} = 1 - J_4(4) + J_3^{\max}(1) = 13 + 8 = 21$ ;

при  $u^{(4)} = 5, x^{(3)} = 0 - J_4(5) + J_3^{\max}(0) = 16 + 0 = 16$ .

Таблица 8.2 – Результаты расчетов

$x^{(k)}$	$u^{(k)}$	$x^{(k-1)}$	$k=2$			$k=3$			$k=4$		
			$J_2 + J_1^{\max}$	$J_2^{\max}$	$\bar{u}^{(2)}$	$J_3 + J_2^{\max}$	$J_3^{\max}$	$\bar{u}^{(3)}$	$J_4 + J_3^{\max}$	$J_4^{\max}$	$\bar{u}^{(4)}$
1	2	3	4	5	6	7	8	9	10	11	12
0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0+8=8	8	0	0+8=8	8	0	0+8=8	8	0
	1	0	6+0=6			3+0=3			4+0=4		
2	0	2	0+10=10			0+14=14	14	0	0+14=14	14	0
	1	1	6+8=14	14	1	3+8=11			4+8=12		
	2	0	9+0=9			4+0=4			6+0=6		
3	0	3	0+11=11			0+17=17	17	0	0+17=17		
	1	2	6+10=16			3+14=17	17	1	4+14=18	18	1
	2	1	9+8=17	17	2	4+8=12			6+8=14		
	3	0	11+0=11			7+0=7			8+0=8		
4	0	4	0+12=12			0+19=19			0+20=20		
	1	3	6+11=17			3+17=20	20	1	4+17=21	21	1
	2	2	9+10=19	19	2	4+14=18			6+14=20		
	3	1	11+8=19	19	3	7+8=15			8+8=16		
	4	0	13+0=13			11+0=11			13+0=13		
5	0	5	0+18=18			0+21=21			0+22=22		
	1	4	6+12=18			3+19=22	22	1	4+20=24	24	1
	2	3	9+11=20			4+17=21			6+17=23		
	3	2	11+10=21	21	3	7+14=21			8+14=22		
	4	1	13+8=21	21	4	11+8=19			13+8=21		
	5	0	15+0=15			18+0=18			16+0=16		

В результате имеем  $J_{\max} = J_4^{\max}(x^{(4)}) = 24$  при  $\bar{u}^{(4)} = 1$ . Далее выполняем безусловную оптимизацию:

$$u_{\text{opt}}^{(4)} = \bar{u}^{(4)}(x^{(4)}) = 1, \quad x^{(3)} = x^{(4)} - u_{\text{opt}}^{(4)} = 5 - 1 = 4;$$

$$u_{\text{opt}}^{(3)} = \bar{u}^{(3)}(x^{(3)}) = 1, \quad x^{(2)} = x^{(3)} - u_{\text{opt}}^{(3)} = 4 - 1 = 3;$$

$$u_{\text{opt}}^{(2)} = \bar{u}^{(2)}(x^{(2)}) = 2, \quad x^{(1)} = x^{(2)} - u_{\text{opt}}^{(2)} = 3 - 2 = 1;$$

$$u_{\text{opt}}^{(1)} = \bar{u}^{(1)}(x^{(1)}) = x^{(1)} = 1.$$

Таким образом, максимум суммарной прибыли равен 24 у.е. средств при условии, что 1-му предприятию выделена 1 у.е., 2-му предприятию – 2 у.е., 3-му предприятию – 1 у.е., 4-му предприятию – 1 у.е.

#### 8.4. Задача об оптимальном распределении ресурсов между отраслями

Планируется деятельность двух отраслей производства на  $N$  лет. Начальные ресурсы  $x^{(0)}$ . Средства  $u$ , вложенные в I отрасль в начале года, дают в конце года прибыль  $J_1(u)$  и возвращаются в размере  $q_1(u) < u$ ; аналогично для II отрасли функция прибыли равна  $J_2(u)$ , а возврата –  $q_2(u) < u$ . В конце года все возвращенные средства заново перераспределяются между I и II отраслями, новые средства не поступают, прибыль в производство не вкладывается.

Требуется распределить имеющиеся средства  $x^{(0)}$  между двумя отраслями производства на  $N$  лет так, чтобы суммарная прибыль от обеих отраслей за  $N$  лет оказалась максимальной.

Необходимо:

- а) построить модель ДП для задачи и вычислительную схему;
- б) решить задачу при условии, что  $x^{(0)} = 10000$  у.е.,  $N = 4$ ,  $J_1(u) = 0,6u$ ,  $q_1(u) = 0,7u$ ,  $J_2(u) = 0,5u$ ,  $q_2(u) = 0,8u$ .

1. Процесс распределения средств между двумя отраслями производства разворачивается во времени, решения принимаются

в начале каждого года, следовательно, осуществляется деление на шаги: номер шага – номер года. Управляемая система – две отрасли производства, а управление состоит в выделении средств каждой отрасли в очередном году. Параметры состояния к концу  $k$ -го года –  $x^{(k)}$ ,  $k = 1, 2, \dots, N$  – количество средств, подлежащих распределению. Переменных управления на каждом шаге две:  $u_1^{(k)}$  – количество средств, выделенных I отрасли;  $u_2^{(k)}$  – количество средств, выделенных II отрасли. Но так как все средства  $x^{(k)}$ , выделенные на  $k$ -м шаге, распределяются, то  $u_2^{(k)} = x^{(k)} - u_1^{(k)}$ , поэтому управление на  $k$ -м шаге зависит от одной переменной  $u_1^{(k)}$ , т.е.  $U^{(k)}(u_1^{(k)}, x^{(k)} - u_1^{(k)})$ .

Уравнения состояний

$$x^{(k-1)} = q_1(u^{(k)}) + q_2(x^{(k)} - u^{(k)}) \quad (8.19)$$

выражают средства, вложенные в конце  $(k - 1)$ -го года.

Показатель эффективности  $k$ -го шага – прибыль, полученная в конце  $k$ -го года от обеих отраслей:

$$J_k = J_1(u^{(k)}) + J_2(x^{(k)} - u^{(k)}). \quad (8.20)$$

Суммарный показатель эффективности – целевая функция – прибыль за  $N$  лет:

$$J = \sum_{k=1}^N (J_1(u^{(k)}) + J_2(x^{(k)} - u^{(k)})). \quad (8.21)$$

Пусть  $J_k^{\max}(x^{(k)})$  – оптимальная прибыль за  $k$  лет, начиная с первого по  $k$ -й год включительно, при условии, что выделенные в 1-м, 2-м, ...,  $(k - 1)$ -м году средства  $x^{(k)}$  распределялись оптимально. Тогда оптимальная прибыль за  $N$  лет  $J_{\max} = J_N^{\max}(x^{(N)})$ .

Уравнения Беллмана приобретают следующий вид:

$$J_1^{\max}(x^{(1)}) = \max_{0 \leq u^{(1)} \leq x^{(1)}} \left\{ J_1(u^{(1)}) + J_2(x^{(1)} - u^{(1)}) \right\}, \quad (8.22)$$

$$J_k^{\max}(x^{(k)}) = \max_{0 \leq u^{(k)} \leq x^{(k)}} \left\{ J_1(u^{(k)}) + J_2(x^{(k)} - u^{(k)}) + J_{k-1}^{\max}(x^{(k-1)}) \right\}, \quad (8.23)$$

$$k = 2, 3, \dots, N.$$

2. Выполним вычисления для конкретных данных. Уравнение состояний (8.19) запишем в виде

$$x^{(k-1)} = 0,7u^{(k)} + 0,8(x^{(k)} - u^{(k)}) = 0,8x^{(k)} - 0,1u^{(k)}. \quad (8.24)$$

Целевую функцию  $k$ -го шага (8.20) – в виде

$$J_k = 0,6u^{(k)} + 0,5(x^{(k)} - u^{(k)}) = 0,5x^{(k)} + 0,1u^{(k)}.$$

Целевую функцию задачи –

$$J = \sum_{k=1}^4 (0,5x^{(k)} + 0,1u^{(k)}). \quad (8.25)$$

Уравнения Беллмана

$$J_1^{\max}(x^{(1)}) = \max_{0 \leq u^{(1)} \leq x^{(1)}} \left\{ 0,5x^{(1)} + 0,1u^{(1)} \right\}, \quad (8.26)$$

$$J_k^{\max}(x^{(k)}) = \max_{0 \leq u^{(k)} \leq x^{(k)}} \left\{ 0,5x^{(k)} + 0,1u^{(k)} + J_{k-1}^{\max}(x^{(k-1)}) \right\}. \quad (8.27)$$

Проводим условную оптимизацию.

Шаг I. Используем уравнение (8.26). Обозначим через  $J_1$  функцию, стоящую в скобках, т.е.  $J_1 = 0,5x^{(1)} + 0,1u^{(1)}$ . Функция  $J_1$  – линейная, возрастающая. Так как максимум достигается на конце интервала  $[0, x^{(1)}]$  (рисунок 8.4), то  $J_1^{\max}(x^{(1)}) = 0,6x^{(1)}$  при  $\bar{u}^{(1)}(x^{(1)}) = x^{(1)}$ .

Шаг II. Запишем уравнение (8.27) для  $k = 2$ :

$$J_2^{\max}(x^{(2)}) = \max_{0 \leq u^{(2)} \leq x^{(2)}} \left\{ 0,5x^{(2)} + 0,1u^{(2)} + 0,6x^{(1)} \right\}.$$

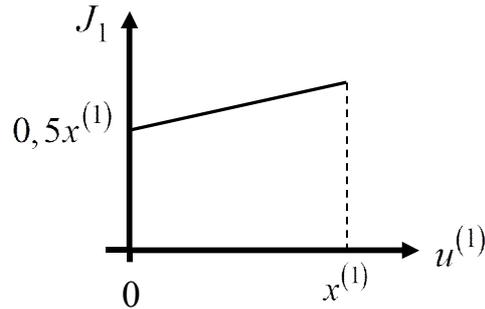


Рисунок 8.4 – Графическое определение функции  $J_1^{\max}(x^{(1)})$

Найдем  $x^{(1)}$  из уравнения состояния (8.24)  $x^{(1)} = 0,8x^{(2)} - 0,1u^{(2)}$  и подставим его в правую часть, получим

$$J_2^{\max}(x^{(2)}) = \max_{0 \leq u^{(2)} \leq x^{(2)}} \left\{ 0,5x^{(2)} + 0,1u^{(2)} + 0,6(0,8x^{(2)} - 0,1u^{(2)}) \right\}, \text{ или}$$

$$J_2^{\max}(x^{(2)}) = \max_{0 \leq u^{(2)} \leq x^{(2)}} \left\{ 0,98x^{(2)} + 0,04u^{(2)} \right\}.$$

Как и в предыдущем случае, максимум достигается при  $\bar{u}^{(k)} = x^{(k)}$ , т.е.  $J_2^{\max}(x^{(2)}) = 1,02x^{(2)}$  при  $\bar{u}^{(2)}(x^{(2)}) = x^{(2)}$ .

Шаг III. Из уравнения состояния находим  $x^{(2)} = 0,8x^{(3)} - 0,1u^{(3)}$ . Уравнение (8.27) при  $k = 3$  примет вид

$$J_3^{\max}(x^{(3)}) = \max_{0 \leq u^{(3)} \leq x^{(3)}} \left\{ 0,5x^{(3)} + 0,1u^{(3)} + 1,02x^{(2)} \right\} =$$

$$= \max_{0 \leq u^{(3)} \leq x^{(3)}} \left\{ 0,5x^{(3)} + 0,1u^{(3)} + 1,02(0,8x^{(3)} - 0,1u^{(3)}) \right\}.$$

После преобразования получим

$$J_3^{\max}(x^{(3)}) = \max_{0 \leq u^{(3)} \leq x^{(3)}} \left\{ 1,316x^{(3)} - 0,002u^{(3)} \right\}.$$

Имеем убывающую по  $u^{(3)}$  функцию на отрезке  $[0, x^{(3)}]$ , поэтому максимум достигается при  $\bar{u}^{(3)} = 0$  (рисунок 8.5):  $B_3(x^{(3)}) = 1,316x^{(3)}$  при  $u^{(3)}(x^{(3)}) = 0$ .

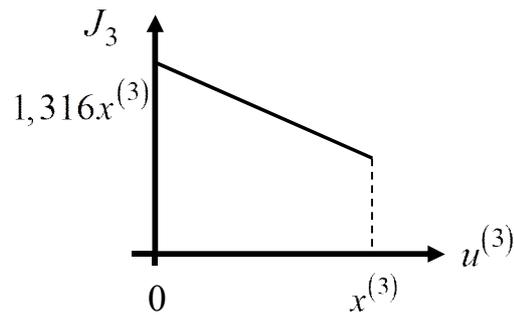


Рисунок 8.5 – Графическое определение функции  $J_3^{\max}(x^{(3)})$

Шаг IV. Из уравнения состояния находим  $x^{(3)} = 0,8x^{(4)} - 0,1u^{(4)}$ .  
Уравнение (8.27) при  $k = 4$  имеет вид

$$J_4^{\max}(x^{(4)}) = \max_{0 \leq u^{(4)} \leq x^{(4)}} \left\{ 0,5x^{(4)} + 0,1u^{(4)} + 1,316(0,8x^{(4)} - 0,1u^{(4)}) \right\}, \text{ или}$$

$$J_4^{\max}(x^{(4)}) = \max_{0 \leq u^{(4)} \leq x^{(4)}} \left\{ 1,5528x^{(4)} - 0,0316u^{(4)} \right\}.$$

Максимум достигается в начале отрезка, т.е.

$$J_4^{\max}(x^{(4)}) = 1,5528x^{(4)} \text{ при } \bar{u}^{(4)}(x^{(4)}) = 0.$$

Так как все выделенные средства должны быть распределены, то  $x^{(4)} = 10000$  у.е. и  $J_{\max} = J_4^{\max}(10000) = 15528$  у.е., при этом в первую отрасль вложено  $u_1^{(4)} = 0$  у.е., во вторую отрасль –  $u_2^{(4)} = 10000 - u_1^{(4)} = 10000$  у.е.;

Вычислим значения вложенных средств (безусловная оптимизация):

– в первые три года  $x^{(3)} = 0,8x^{(4)} - 0,1u^{(4)} = 0,8 \cdot 10000 - 0,1 \cdot 0 = 8000$  у.е.;

– в первую отрасль вложено  $u_1^{(3)} = 0$  у.е.;

– во вторую отрасль –  $u_2^{(3)} = 8000 - u_1^{(3)} = 8000$  у.е.;

– в первые два года  $x^{(2)} = 0,8x^{(3)} - 0,1u^{(3)} = 0,8 \cdot 8000 - 0,1 \cdot 0 = 6400$  у.е.;

- в первую отрасль вложено  $u_1^{(2)} = x^{(2)} = 6400$  у.е.;
- во вторую отрасль –  $u_2^{(2)} = 6400 - u_1^{(2)} = 6400 - 6400 = 0$  у.е.;
- в первый год  $x^{(1)} = 0,8x^{(2)} - 0,1u^{(2)} = 0,7x^{(2)} = 0,8 \cdot 6400 - 0,1 \cdot 6400 = 4480$  у.е.;
- в первую отрасль вложено  $u_1^{(1)} = x^{(1)} = 4480$  у.е.;
- во вторую отрасль –  $u_2^{(1)} = 4480 - u_1^{(2)} = 4480 - 4480 = 0$  у.е.

Таким образом, оптимальная прибыль за 4 года, полученная от двух отраслей производства при начальных средствах 10000 у.е., равна 15528 у.е. при условии, что I отрасль получает по годам (4480, 6400, 0, 0) у.е., а II отрасль – (0, 0, 8000, 10000) у.е. Заметим, что этот ответ не совпадает с ответом, полученным в работе [13], где используется схема поиска оптимального решения «с конца периода»: первая отрасль – (0, 0, 6400, 4480) у.е., вторая отрасль – (10000, 8000, 0, 0) у.е. Следовательно, данная задача имеет два оптимальных решения, при которых максимальная прибыль равна 15528 у.е.

## **8.5. Задача о замене оборудования**

Замена оборудования – важная экономическая проблема. Задача состоит в определении оптимальных сроков замены старого оборудования (станков, производственных зданий и т.п.). Старение оборудования включает его физический и моральный износ, в результате чего растут производственные затраты, затраты на ремонт и обслуживание, снижаются производительность труда, ликвидная стоимость. Критерием оптимальности являются, как правило, либо прибыль от эксплуатации оборудования (задача максимизации), либо суммарные затраты на эксплуатацию в течение планируемого периода (задача минимизации).

При построении модели задачи принято считать, что решение о замене оборудования выносится в начале каждого промежутка эксплуатации (например, в начале года) и что, в принципе, оборудование можно использовать неограниченно долго.

Основная характеристика оборудования – параметр состояния – его возраст  $t$ .

При составлении динамической модели процесс замены рассматривают как  $N$ -шаговый, разбивая весь период эксплуатации на  $N$  шагов. Возможное управление на каждом шаге характеризуется качественными признаками, например  $U^c$  (сохранить оборудование),  $U^3$  (заменить оборудование).

*Пример.* Оборудование эксплуатируется в течение 5 лет, после этого продается. В начале каждого года можно принять решение сохранить оборудование или заменить его новым. Стоимость нового оборудования  $p_0 = 4000$  у.е. После  $t$  лет эксплуатации ( $1 \leq t \leq 5$ ) оборудование можно продать за  $g(t) = p_0 2^{-t}$  у.е. (ликвидная стоимость). Затраты на содержание в течение года зависят от возраста  $t$  оборудования и рассчитываются как  $r(t) = 600(t + 1)$ . Требуется определить оптимальную стратегию эксплуатации оборудования, чтобы суммарные затраты с учетом начальной покупки и заключительной продажи были минимальны.

В данном примере способ деления управления на шаги естественный, по годам,  $N = 5$ . Параметр состояния – возраст оборудования –  $x^{(k)} = t$ ,  $x^{(0)} = 0$  – оборудование новое, в начале первого года эксплуатации. Управление на каждом шаге зависит от двух переменных –  $U^c$  и  $U^3$ .

Уравнения состояний зависят от управления:

$$x^{(k)} = \begin{cases} t + 1, & \text{если } u^{(k)} = U^c, \\ 1, & \text{если } u^{(k)} = U^3, \end{cases} \quad k = 1, 2, 3, 4. \quad (8.28)$$

Поясним соотношение (8.28). Если к началу  $k$ -го шага  $x^{(k-1)} = t$ , то при сохранении оборудования ( $u^{(k)} = U^c$ ) через год его возраст увеличится на 1. Если оборудование заменяется новым ( $u^{(k)} = U^3$ ), то в начале  $k$ -го шага его возраст  $t = 0$ , а после года эксплуатации  $t = 1$ , т.е.  $x^{(k)} = 1$ .

Показатель эффективности  $k$ -го шага

$$J_k(u^{(k)}, t) = \begin{cases} 600(t+1), & \text{если } u^{(k)} = U^c, \\ 4600 - 4000 \cdot 2^{-t}, & \text{если } u^{(k)} = U^3, k = 1, \dots, 4. \end{cases} \quad (8.29)$$

При  $U^c$  затраты только на эксплуатацию оборудования возраста  $t$ , при  $U^3$  оно продается  $(-4000 \cdot 2^{-t})$ , покупается новое (4000) и эксплуатируется в течение первого года (600), общие затраты:  $-4000 \cdot 2^{-t} + 4000 + 600$ .

Пусть  $B_k(t)$  – оптимальные затраты на эксплуатацию оборудования, начиная с 1-го шага по  $k$ -й, при условии, что к началу  $k$ -го шага оно имеет возраст  $t$  лет. Запишем для функций  $B_k(t)$  уравнения Беллмана (8.10) и (8.13), заменив задачу максимизации задачей минимизации:

$$B_1 = 4000 + 600 = 4600, \quad (8.30)$$

$$B_k = \min \begin{cases} 600(t+1) + B_{k-1}(t), & \text{если } u^{(k)} = U^c, \\ 4600 - 4000 \cdot 2^{-t} + B_{k-1}(t), & \text{если } u^{(k)} = U^3, \end{cases} \quad (8.31)$$

$$k = 2, 3, 4.$$

Для последнего шага имеем

$$B_5 = \min \begin{cases} 600(t+1) + B_4(t) - 4000 \cdot 2^{-(t+1)}, & \text{если } u^{(k)} = U^c, \\ 4600 - 4000 \cdot 2^{-t} + B_4(t) - 4000 \cdot 2^{-1}, & \text{если } u^{(k)} = U^3. \end{cases} \quad (8.32)$$

Третьи слагаемые означают, что в конце 5-го года оборудование продается независимо от его состояния.

Из определения функций  $B_k(t)$  следует

$$J_{\min} = B_5. \quad (8.33)$$

Дадим геометрическое решение этой задачи (рисунок 8.6). На оси абсцисс будем откладывать номер шага  $k$ , на оси ординат – возраст  $t$  оборудования. Точка  $(k-1, t)$  на плоскости соответствует началу  $k$ -го

года эксплуатации оборудования возраста  $t$  лет. Перемещение на графике в зависимости от принятого управления на  $k$ -м шаге показано на рисунке 8.6.

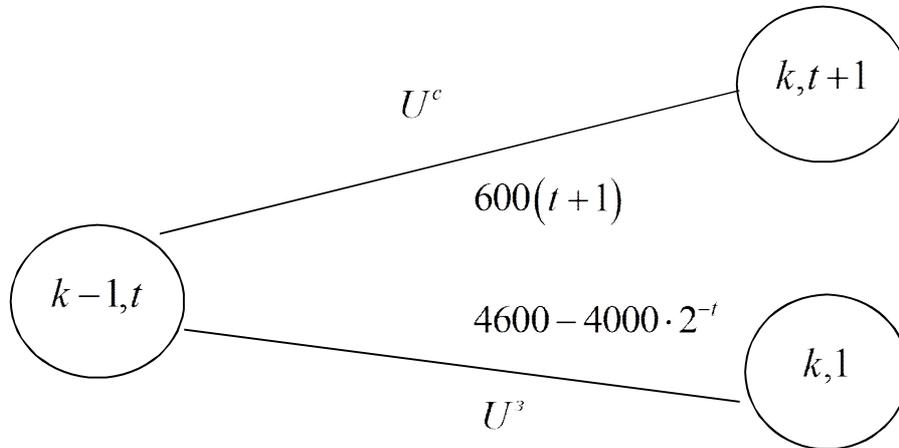


Рисунок 8.6 – Геометрическое решение задачи

Состояние начала эксплуатации оборудования соответствует точке  $x^{(0)}(0,0)$ , конец – точке  $x^{(5)}(6,t)$ . Любая траектория, переводящая точку  $x(k-1,t)$  из  $x^{(0)}(0,0)$  в  $x^{(5)}(6,t)$ , состоит из отрезков – шагов, соответствующих годам эксплуатации. Надо выбрать такую траекторию, при которой затраты на эксплуатацию оборудования окажутся минимальными.

Над каждым отрезком, соединяющим точки  $(k-1,t)$  и  $(k,t+1)$ , запишем затраты, соответствующие управлению  $U^c$ , найденные из соотношения (8.29):  $600(t+1)$ , а над отрезком, соединяющим точки  $(k-1,t)$  и  $(k,1)$ , запишем затраты, соответствующие управлению  $U^s$ , т.е.  $4600 - 4000 \cdot 2^{-t}$ .

Так мы разметим все отрезки, соединяющие точки на графике, соответствующие переходам из любого состояния  $x^{(k-1)}$  в состояние  $x^{(k)}$  (см. рисунок 8.6). Например, над отрезками, соединяющими точки  $(k,2)$  и  $(k+1,3)$ , будет стоять число 1800, что соответствует затратам на эксплуатацию в течение каждого года оборудования

возраста  $t = 2$  лет, а над отрезками, соединяющими  $(k, 2)$  и  $(k + 1, 1)$ , будет стоять число 3600 – сумма затрат на покупку и эксплуатацию нового оборудования в течение года без «затрат» (выручки) за проданное оборудование возраста 2 года. При этом следует учесть, что  $0 \leq t \leq k$ .

Над отрезками, соединяющими  $(k, 3)$  и  $(k + 1, 1)$ , будет стоять число 4100 – сумма затрат на покупку и эксплуатацию нового оборудования в течение года за вычетом выручки за проданное оборудование возраста 3 года. И наконец, над отрезком, соединяющим  $(4, 4)$  и  $(5, 1)$ , будет стоять число 4350 (здесь выручка от продажи оборудования возраста 4 года составляет 250 у.е.).

Проведем на размеченном графе состояний (рисунок 8.7) оптимизацию.

Шаг I. Начальные состояния – точки  $(0, 0)$ , конечные –  $(1, 1)$ . Из состояния  $(0, 0)$  можно попасть только в состояние  $(1, 1)$ , затратив на эксплуатацию оборудования 600 у.е. Следовательно, суммарные затраты включают затраты на покупку оборудования в начале первого года и затраты на эксплуатацию в течение года, т.е.  $4000 + 600 = 4600$  у.е. Отметим это направление жирной стрелкой, в эллипс с точкой  $(1, 1)$  запишем 4600.

Шаг II. Начальные состояния – точки  $(1, 1)$ , конечные –  $(2, t)$ .

*Состояние*  $(1, 1)$ . Из него можно попасть в состояние  $(2, 2)$ , затратив на эксплуатацию оборудования 1200 у.е., а суммарные затраты составят  $4600 + 1200 = 5800$  у.е., и в состояние  $(2, 1)$  с затратами  $4600 + 2600 = 7200$  у.е. Значит, если к этому шагу система находилась в точке  $(2, 1)$ , то следует двигаться в точку  $(2, 2)$  (укажем это направление жирной стрелкой), а неизбежные минимальные затраты, соответствующие этому переходу, равны 5800 у.е. (поместим эту величину в эллипсе точки  $(2, 2)$ ).

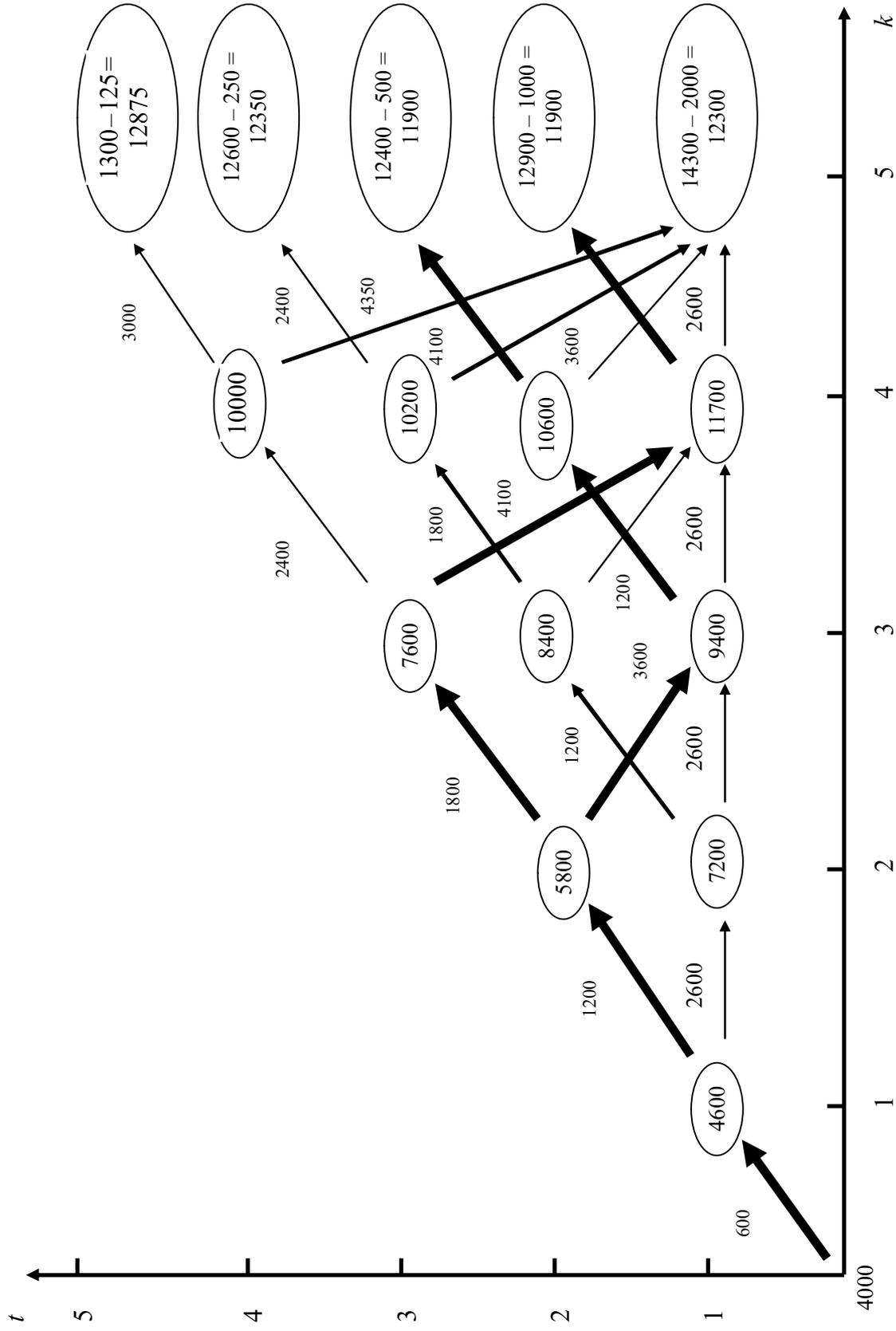


Рисунок 8.7 – Граф состояний

Шаг III. Начальные состояния – точки  $(2,t)$ , конечные –  $(3,t)$ .

*Состояние*  $(2,1)$ . Из него можно попасть в состояние  $(3,2)$ , затратив на эксплуатацию оборудования 1200 у.е., а суммарные затраты составят  $7200 + 1200 = 8400$  у.е., и в состояние  $(3,1)$  с затратами  $7200 + 2600 = 9400$  у.е. Значит, если к этому шагу система находилась в точке  $(2,1)$ , то следует идти в точку  $(3,2)$  (укажем это направление полужирной стрелкой), а неизбежные минимальные затраты, соответствующие этому переходу, равны 8400 у.е. (поместим эту величину в эллипсе точки  $(3,2)$ ).

*Состояние*  $(2,2)$ . Из него можно попасть в точку  $(3,3)$  с затратами  $5800 + 1800 = 7600$  у.е. и в точку  $(3,1)$  с затратами  $5800 + 3600 = 9400$  у.е. Отметим второе управление полужирной стрелкой.

Шаг IV. Начальные состояния – точки  $(3,t)$ , конечные –  $(4,t)$ .

Анализируем, как можно попасть из каждого начального состояния в конечное на IV шаге.

*Состояние*  $(3,1)$ . Из него можно попасть в состояние  $(4,2)$ , затратив на эксплуатацию оборудования 1200 у.е., т.е. суммарные затраты составят  $9400 + 1200 = 10600$  у.е., и в состояние  $(4,1)$  с затратами  $9400 + 2600 = 12000$  у.е. Значит, если к этому шагу система находилась в точке  $(3,1)$ , то следует идти в точку  $(4,2)$  (укажем это направление жирной стрелкой), а неизбежные минимальные затраты, соответствующие этому переходу, равны 10600 у.е. (поместим эту величину в эллипсе точки  $(4,2)$ ).

*Состояние*  $(3,2)$ . Из него можно попасть в точку  $(4,3)$  с затратами  $8400 + 1800 = 10200$  у.е. и в точку  $(4,1)$  с затратами  $8400 + 3600 = 12000$  у.е.

*Состояние*  $(3,3)$ . Из него можно попасть в точку  $(4,4)$  с затратами  $7600 + 2400 = 10000$  у.е. и в точку  $(4,1)$  с затратами  $7600 + 4100 = 11700$  у.е. Выбираем второе управление, отмечаем его полужирной стрелкой. В эллипсе точки  $(4,1)$  записываем затраты при

переходе из точки  $(3,3)$  в точку  $(4,1)$ , равные 11700 у.е., так как они меньше, чем при переходе из точки  $(3,2)$  в точку  $(4,1)$ . Более того, эти затраты являются наименьшими, в связи с чем переход из точки  $(3,3)$  в точку  $(4,1)$  отмечаем жирной стрелкой.

Шаг V. Начальные состояния – точки  $(4,t)$ , конечные –  $(5,t)$ . В состояниях  $(5,t)$  оборудование продается, доход от продажи равен  $4000 \cdot 2^{-t}$ , но поскольку целевая функция связана с затратами, то в кружках точек  $(5,t)$  от суммарных затрат отнимаем выручку от продажи.

Анализируем, как можно попасть из каждого начального состояния в конечное на V шаге.

*Состояние  $(4,1)$ .* Из него можно попасть в состояние  $(5,2)$ , затратив на эксплуатацию оборудования 1200 у.е. и выручив затем от продажи 1000 у.е., т.е. суммарные затраты равны  $11700 + 1200 - 1000$  у.е., и в состояние  $(5,1)$  с затратами  $11700 + 2600 - 2000 = 12300$  у.е. Значит, если к последнему шагу система находилась в точке  $(4,1)$ , то следует идти в точку  $(5,2)$  (укажем это направление жирной стрелкой, так как увидим в дальнейшем, что этот переход соответствует наименьшим затратам), а неизбежные минимальные затраты, соответствующие этому переходу, равны 11900 у.е. (поместим эту величину в эллипсе точки  $(5,2)$ ).

*Состояние  $(4,2)$ .* Из него можно попасть в точку  $(5,3)$  с затратами  $10600 + 1800 - 500 = 11900$  у.е. и в точку  $(5,1)$  с затратами  $10600 + 3600 - 2000 = 12200$  у.е. Выбираем первое управление, отмечаем его жирной стрелкой.

*Состояние  $(4,3)$ .* Из него можно попасть в точку  $(5,4)$  с затратами  $10200 + 2400 - 250 = 12350$  у.е. и в точку  $(5,1)$  с затратами  $10200 + 4100 - 2000 = 12300$  у.е. Выбираем второе управление, отмечаем его полужирной стрелкой.

Состояние (4,4). Из него можно попасть в точку (5,5) с затратами  $10000 + 3000 - 125 = 12875$  у.е. и в точку (5,1) с затратами  $10000 + 4350 - 2000 = 12350$  у.е. Выбираем второе управление, отмечаем его полужирной стрелкой. В эллипсе точки (5,1) записываем затраты при переходе из точки (4,3) в точку (5,1), равные 12300 у.е., так как эти затраты меньше, чем при переходе из точки (4,4) в точку (5,1).

На размеченном графе (см. рисунок 8.7) отмечены жирными стрелками те переходы из состояния в состояние, для которых суммарные затраты минимальны (11900 у.е.). Как видно из графика, оптимальных траекторий две: первая включает набор точек  $\{(0,0), (1,1), (2,2), (3,1), (4,2), (5,3)\}$ , который соответствует оптимальному управлению  $u_{\text{opt}} = (U^c, U^c, U^3, U^c, U^c)$ , вторая – набор точек  $\{(0,0), (1,1), (2,2), (3,3), (4,1), (5,2)\}$ , соответствующий оптимальному управлению  $u_{\text{opt}} = (U^c, U^c, U^c, U^3, U^c)$ .

Оптимальный режим эксплуатации в первом случае состоит в том, чтобы заменить оборудование новым в начале 3-го года, а во втором случае – в начале 4-го года.

Таким образом, размеченный граф (сеть) позволяет наглядно интерпретировать расчетную схему и решить задачу методом ДП.

Заметим, такой же ответ с двумя оптимальными траекториями получен в работе [20] при использовании «обратной схемы».

## 8.6. Задача о кратчайшем пути

На рисунке (8.8) представлено расположение 10 населенных пунктов (номера населенных пунктов указаны в эллипсах), связанных между собой сетью автомобильных дорог. Расстояние между городами в километрах приведено рядом со стрелками. По дорогам можно перемещаться только в направлении стрелок. Требуется проехать из города 1 в город 10 по кратчайшему пути.

Процесс перемещения из города в город происходит за четыре шага. Уравнения состояний выражают маршруты движения:

$$x^{(k)} = \left\{ x^{(k-1)}, u^{(k)} \right\}, k = 1, \dots, 4,$$

где  $x^{(k-1)}$  – номер пункта отправления;  $u^{(k)}$  – управление на  $k$ -м шаге, указывает номер пункта назначения. Положим  $x^{(0)} = 0$ . Целевая функция здесь – это пройденное расстояние.

На рисунке 8.8 приведены результаты расчетов кратчайшего расстояния. Значения  $J_k^{\min}(x^{(k)})$  указаны под эллипсами.

В итоге получили оптимальный маршрут с кратчайшим расстоянием 11 км ( $1 \rightarrow 4 \rightarrow 6 \rightarrow 9 \rightarrow 10$ ). Следует заметить, что он совпал с маршрутом, найденным по «обратной схеме» в работе [22]. Однако здесь присутствует еще один оптимальный маршрут – ( $1 \rightarrow 3 \rightarrow 5 \rightarrow 8 \rightarrow 10$ ) с такой же длиной пути 11 км.

### **8.7. Редукция линейной задачи динамического программирования к задаче линейного программирования**

Многошаговый процесс поиска решения линейной задачи динамического программирования можно свести к эквивалентной задаче линейного программирования.

Рассмотрим задачу динамического программирования, записанную в форме

$$\begin{aligned} J(x, u) &= \sum_{k=1}^N J_k = \sum_{k=1}^N J_k(x^{(k)}, u^{(k)}) \rightarrow \underset{u^{(k)}}{\text{extr}}, \\ x^{(k-1)} &= f_{k-1}(x^{(k)}, u^{(k)}), k = 2, \dots, N, \\ u^{(k)} &\in U_k(x^{(k)}), k = 1, \dots, N, \\ x^{(N)} &\in X_N. \end{aligned} \tag{8.34}$$

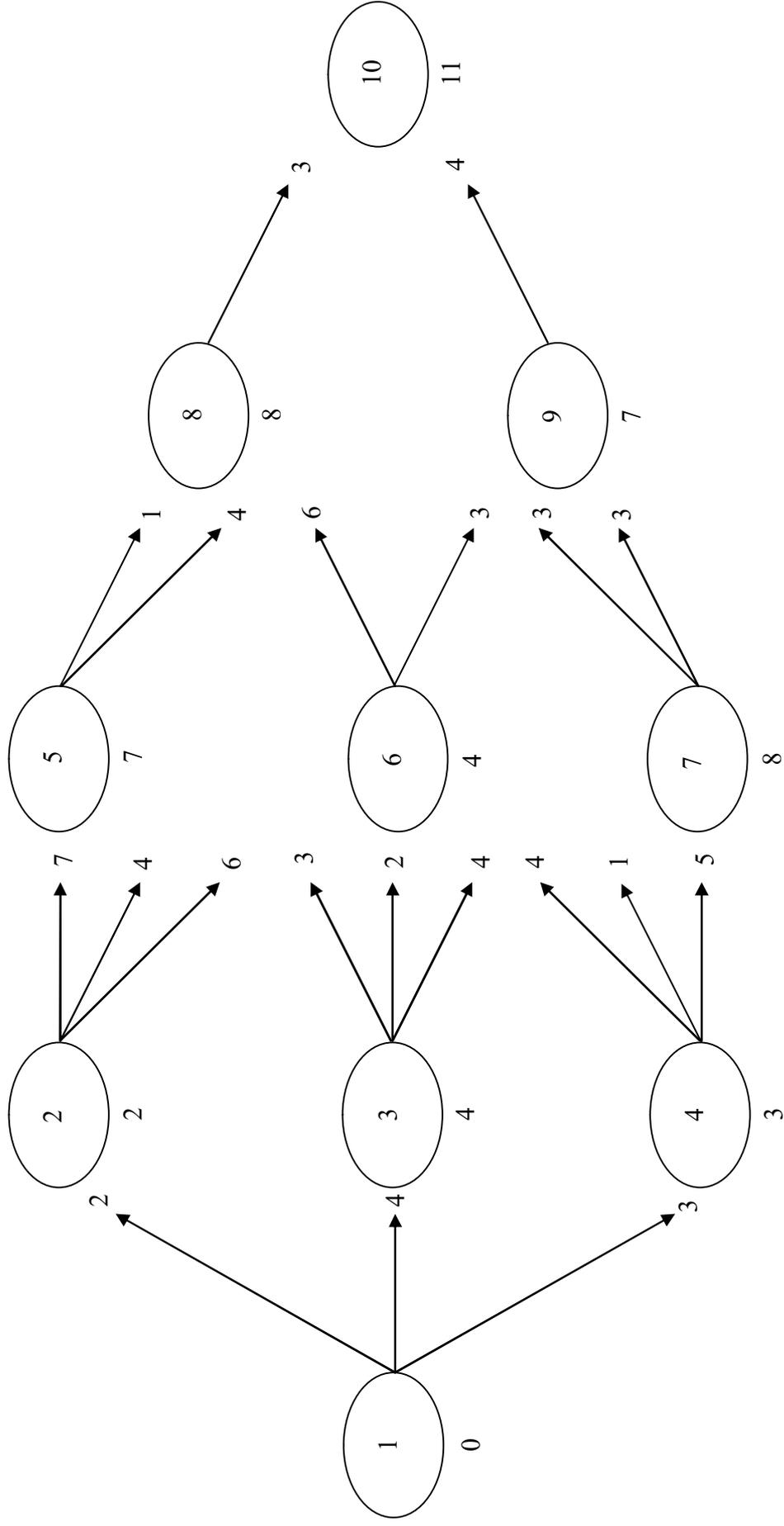


Рисунок 8.8 – Расчет кратчайшего расстояния

Выразим в «прямой схеме» последовательно из уравнения состояния  $x^{(k-1)}$  через  $x^{(k)}$  и затем подставим в целевую функцию. В результате получим

$$\begin{aligned}
 J\left(x^{(N)}, u^{(1)}, u^{(2)}, \dots, u^{(N)}\right) &= \sum_{k=1}^N J_k = \\
 &= J_N\left(x^{(N)}, u^{(N)}\right) + J_{N-1}\left(f_{N-1}\left(x^{(N)}, u^{(N)}\right), u^{(N-1)}\right) + \\
 &+ J_{N-2}\left(f_{N-2}\left(f_{N-1}\left(x^{(N)}, u^{(N)}\right), u^{(N-1)}\right), u^{(N-2)}\right) + \dots + \\
 &+ J_1\left(f_1\left(f_2\left(\dots f_{N-2}\left(f_{N-1}\left(x^{(N)}, u^{(N)}\right), u^{(N-1)}\right), \dots, u^{(3)}\right), u^{(2)}\right), u^{(1)}\right),
 \end{aligned} \tag{8.35}$$

$$u^{(k)} \in U_k\left(x^{(k-1)}\right), k = 1, \dots, N,$$

$$x^{(1)} = f_1\left(f_2\left(\dots f_{N-2}\left(f_{N-1}\left(x^{(N)}, u^{(N)}\right), u^{(N-1)}\right), \dots, u^{(2)}\right), u^{(1)}\right), \tag{8.36}$$

$$x^{(N)} \in X_N.$$

Таким образом, получили задачу линейного программирования относительно  $N$  переменных  $u^{(1)}, u^{(2)}, \dots, u^{(N)}$ . Задачу линейного программирования решаем известным симплексным методом.

*Пример 1.* Применим этот подход к решению задачи о распределении средств между предприятиями (см. подраздел 8.3):

$$\begin{aligned}
 J\left(x^{(4)}, u^{(1)}, u^{(2)}, u^{(3)}, u^{(4)}\right) &= \sum_{k=1}^4 J_k = \\
 &= J_1\left(x^{(4)} - u^{(4)} - u^{(3)} - u^{(2)}, u^{(1)}\right) + J_2\left(x^{(4)} - u^{(4)} - u^{(3)}, u^{(2)}\right) + \\
 &+ J_3\left(x^{(4)} - u^{(4)}, u^{(3)}\right) + J_4\left(x^{(4)}, u^{(4)}\right) \rightarrow \max_u,
 \end{aligned} \tag{8.37}$$

$$\begin{aligned}
u^{(4)} &\leq x^{(4)}, \\
u^{(3)} &\leq x^{(4)} - u^{(4)}, \\
u^{(2)} &\leq x^{(4)} - u^{(4)} - u^{(3)}, \\
u^{(1)} + u^{(2)} + u^{(3)} + u^{(4)} &= x^{(4)}, \\
u^{(j)} &\geq 0, \quad j = 1, \dots, 4.
\end{aligned} \tag{8.38}$$

Здесь  $x^{(4)}$  – количество распределенных средств на 4-м шаге (в нашем случае  $x^{(4)} = 5$  у.е.)

Учитывая, что функции прибыли возрастающие, и часть ограничений (8.38) излишни, можно переписать эту задачу в виде

$$\begin{aligned}
&J(u^{(1)}, u^{(2)}, u^{(3)}, u^{(4)}) = \\
&= J_1(u^{(1)}) + J_2(u^{(2)}) + J_3(u^{(3)}) + J_4(u^{(4)}) \rightarrow \max_u, \\
&u^{(1)} + u^{(2)} + u^{(3)} + u^{(4)} = x^{(4)}, \\
&u^{(j)} \geq 0, \quad j = 1, \dots, 4.
\end{aligned} \tag{8.39}$$

Таким образом, получили задачу линейного программирования с одним ограничением.

*Пример 2.* Рассмотрим задачу о распределении ресурсов между двумя отраслями на 4 года (см. подраздел 8.4).

Имеем следующую модель:

$$\begin{aligned}
J &= \sum_{k=1}^4 (0,5x^{(k)} + 0,1u^{(k)}) \rightarrow \max_{u^{(k)}}, \\
x^{(k-1)} &= 0,8x^{(k)} - 0,1u^{(k)}, \quad k = 1, \dots, 4.
\end{aligned}$$

Эквивалентная задача линейного программирования примет вид

$$\begin{aligned}
J &= 0,1u^{(1)} + 0,05u^{(2)} + 0,01u^{(3)} - 0,022u^{(4)} + 1,076x^{(4)} \rightarrow \max_{u^{(k)}}, \\
u^{(4)} &\leq x^{(4)}, \\
u^{(3)} &\leq 0,8x^{(4)} - 0,1u^{(4)},
\end{aligned}$$

$$u^{(2)} \leq 0,64x^{(4)} - (0,08u^{(4)} + 0,1u^{(3)}),$$

$$u^{(1)} \leq 0,512x^{(4)} - (0,1u^{(2)} + 0,08u^{(3)} + 0,064u^{(4)}),$$

$$u^{(1)}, u^{(2)}, u^{(3)}, u^{(4)} \geq 0;$$

$$x^{(4)} = 10000.$$

### Вопросы для самопроверки

1. Что понимают под динамическим программированием?
2. Запишите условие многошаговой задачи оптимизации.
3. Перечислите особенности модели динамического программирования
4. В чем состоит принцип оптимальности управления при решении задачи динамического программирования?
5. В чем отличие «прямой» и «обратной» схем решения задачи ДП?
6. Запишите уравнения Беллмана для «прямой» схемы решения задачи ДП.
7. Запишите модель задачи о распределении средств между предприятиями методом динамического программирования.
8. Запишите модель задачи об оптимальном распределении ресурсов между отраслями на  $N$  лет методом динамического программирования.
9. Запишите модель задачи о замене оборудования методом динамического программирования.
10. Как преобразовать линейную задачу динамического программирования в задачу линейного программирования?

## Глава 9. ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ

### 9.1. Функционалы. Основные понятия

Далеко не все встречающиеся в приложениях задачи поиска экстремума могут быть решены в форме задач нелинейного программирования. Математические модели многих прикладных задач на экстремум сводятся к оптимизации *функционалов* на множестве функций. Первые задачи подобного типа были сформулированы в математике и решены в XVII–XVIII вв. С тех пор раздел математики, где исследуются задачи на экстремум в бесконечномерных *функциональных пространствах*, называется *вариационным исчислением* (ВИ).

Основная задача ВИ возникла как непосредственное обобщение поставленной Якобом Бернулли (швейцарский ученый, 1654–1705) в 1696 г. *задачи о брахистохроне* (брахистохрона – линия наискорейшего ската материальной точки). Эта задача содержала типичные черты нового класса задач и на протяжении всей истории развития ВИ служила и как объект испытаний новых методов, и как основа для многих интеграционных обобщений.

*Задача о брахистохроне.* В плоскости  $(xOy)$  имеются две заданные точки  $A$  и  $B$ , не лежащие на одной вертикальной прямой (рисунок 9.1). Координаты точек:  $A(0,0)$  и  $B(x_1, y_1)$ . Необходимо найти линию, соединяющую их. При движении вдоль этой линии тяжелый материальный шарик под действием силы тяжести  $P$  должен достигнуть точки  $B$  за кратчайшее время. Силами трения пренебрегаем и полагаем, что начальная скорость равна нулю:  $v_A = 0$ .

Найдем зависимость времени движения от траектории  $y(x)$ . Обозначим через  $ds$  элемент пути, а через  $d\tau$  элемент времени. Тогда можем записать

$$v(x) = \frac{ds}{d\tau} = \sqrt{2gy(x)}, \quad ds = \sqrt{1 + (y'(x))^2} dx,$$

где  $g$  – ускорение свободного падения. Подставив выражение  $ds$  в первое уравнение, получим

$$d\tau = \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{2gy(x)}} dx.$$

Отсюда находим время движения шарика из точки  $A$  в точку  $B$ :

$$T = J(y) = \frac{1}{\sqrt{2g}} \int_0^{x_1} \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{y}} dx.$$

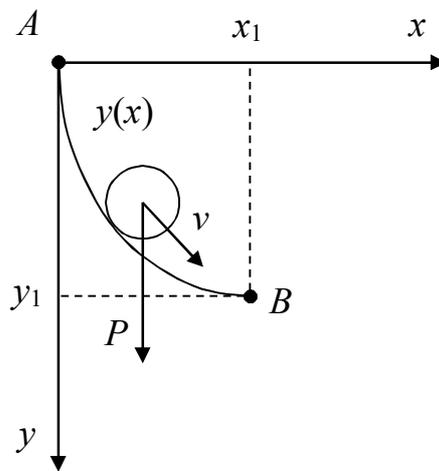


Рисунок 9.1 – Линия наискорейшего спуска в задаче о брахистохроне

Необходимо найти такую кривую  $y(x)$ ,  $x \in [0, x_1]$ , чтобы значение  $J(y)$  было минимальным. Граничные условия задачи следующие:

$$y(0) = 0, y(x_1) = y_1.$$

*Определение.* Если каждому элементу  $y = y(x)$  множества  $G$  из некоторого функционального пространства  $X$  поставлено в соответствие определенное число  $J$ , то говорят, что на множестве  $G \subset X$  задан функционал  $J(y) \equiv J[y(x)]$ .

Задание функционала  $J[y(x)]$  равносильно заданию закона, по которому каждой функции  $y(x)$  ставится в соответствие определенное действительное число. Например:

$$1) J = \int_a^b y(x) dx;$$

$$2) J = \int_a^b \varphi[y(x)] dx;$$

$$3) J(y_1, y_2) = J(y) = \int_a^b \varphi[y_1(x), y_2(x)] dx;$$

$$4) J(y, y') = \int_a^b \varphi[y(x), y'(x)] dx;$$

5) длина дуги плоской кривой, соединяющей две заданные точки  $A(a, y_a)$  и  $B(b, y_b)$ , если задано уравнение кривой  $y = y(x)$ , вычисляется по формуле (рисунок 9.2)

$$l(y) = J(y, y') = \int_a^b \sqrt{1 + (y')^2} dx;$$

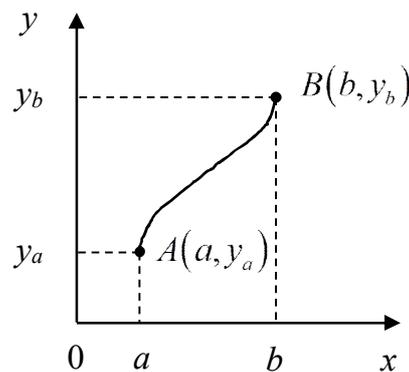


Рисунок 9.2 – К вопросу о вычислении длины дуги кривой

б) функционалом является также площадь поверхности  $y(x_1, x_2)$ :

$$S(y) = J(y, y') = \iint_E \sqrt{1 + \left(\frac{\partial y(x_1, x_2)}{\partial x_1}\right)^2 + \left(\frac{\partial y(x_1, x_2)}{\partial x_2}\right)^2} dx_1 dx_2,$$

где  $E$  – область изменения векторной переменной  $x = (x_1, x_2)$ . Она представляет собой проекцию поверхности  $y(x_1, x_2)$  на плоскость  $(x_1, x_2)$ .

В качестве функциональных пространств  $X$  в ВИ используются пространства  $C_n[a, b]$ ,  $n = 0, 1, \dots$ , определяемые следующим образом:  $C_n[a, b]$  – линейное нормированное пространство, которое состоит из функций  $y(x)$ , имеющих на отрезке  $[a, b]$  непрерывные производные  $y^{(k)}(x)$  до  $n$ -го порядка включительно ( $y^{(0)} = y(x)$ ) с нормой

$$\|y\|_n = \sum_{k=0}^n \max_{x \in [a, b]} |y^{(k)}(x)|.$$

Расстояние  $\rho(y_1, y_2)$  между функциями (кривыми)  $y_1(x)$  и  $y_2(x)$  в пространстве  $C_n[a, b]$  определяется формулой

$$\rho(y_1, y_2) = \|y_1 - y_2\| = \sum_{k=0}^n \max_{x \in [a, b]} |y_1^{(k)}(x) - y_2^{(k)}(x)|, \quad n = 0, 1, \dots$$

Пусть функция  $y^*(x) \in C_n[a, b]$  и  $\varepsilon > 0$  – произвольное число. Множество функций (кривых)  $y(x) \in C_n[a, b]$ , для которых выполняется неравенство

$$\rho(y^*, y)_n < \varepsilon,$$

называется  $\varepsilon$ -окрестностью  $n$ -го порядка кривой  $y^*(x)$ .

## 9.2. Необходимое и достаточное условие существования экстремума функционалов

*Определение.* Функционал  $J[y(x)]$  достигает на кривой  $y^*(x)$  локального экстремума (min или max), если для  $\forall y(x)$  из некоторой  $\varepsilon$ -окрестности кривой  $y^*(x)$  выполняется неравенство

$$J[y^*(x)] \underset{\geq}{\overset{\leq}{}} J[y(x)]. \quad (9.1)$$

Если выражение (9.1) выполняется для  $\forall y(x) \in G \subset C_n[a, b]$ , то говорят, что на кривой  $y^*(x)$  достигается *глобальный экстремум* функционала  $J[y(x)]$  на множестве  $G$ .

Получим необходимое условие экстремума. Если дан ненулевой элемент  $v$  подпространства  $G_v \subset G$  и  $\varepsilon$  – действительное число, то произведение  $\varepsilon \cdot v$  называют *вариацией*. Рассмотрим функционал  $J$  в точке  $y^* + \varepsilon \cdot v$  пространства  $G$ . Нетрудно видеть, что этот функционал является функцией скалярного аргумента  $\varepsilon$ :

$$F(\varepsilon) \equiv J(y^* + \varepsilon \cdot v). \quad (9.2)$$

Разложим  $F(\varepsilon)$  в ряд Тейлора в точке  $\varepsilon$ :

$$F(\varepsilon) \equiv J(y^* + \varepsilon \cdot v) = J(y^*) + \delta J(y^*, v) \cdot \varepsilon + O(\varepsilon), \quad (9.3)$$

где  $\delta J(y^*, v) = dF(0)/d\varepsilon$  – первая вариация  $J$  (дифференциал Гато);

$$\lim_{\varepsilon \rightarrow 0} \frac{O(\varepsilon)}{\varepsilon} \rightarrow 0.$$

*Необходимое условие экстремума функционала  $J$  приобретает вид*

$$\delta J(y^*, v) = \frac{dF(0)}{d\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{J(y^* + \varepsilon \cdot v) - J(y^*)}{\varepsilon} = 0. \quad (9.4)$$

*Пример 1.* Пусть  $\varphi(y)$  – непрерывная дифференцируемая функция на множестве  $R_1$  и  $J(y)$  – функционал, определенный в функциональном пространстве  $G$ :

$$J(y) = \int_0^1 \varphi(y(x)) dx.$$

Рассмотрим

$$\begin{aligned}
 J(y^* + \varepsilon \cdot v) &= \\
 &= \int_0^1 \varphi(y^*(x) + \varepsilon \cdot v(x)) dx = \int_0^1 \left\{ \varphi(y^*(x)) + \varepsilon \cdot \frac{d\varphi}{dy} \Big|_{y=y^*(x)} v(x) \right\} dx = \\
 &= J(y^*) + \varepsilon \int_0^1 \varphi'(y^*(x)) \cdot v(x) dx + O(\varepsilon).
 \end{aligned}$$

Таким образом, первая вариация имеет вид

$$\delta J(y^*) = \int_0^1 \varphi'(y^*(x)) \cdot v(x) dx,$$

где  $v(x)$  – любой элемент пространства  $G$ .

Приравнивая к нулю первую вариацию функционала, получаем уравнение, которому удовлетворяют все подозрительные на экстремум функции  $y \in G$ .

*Пример 2.* Пусть  $\varphi(y_1, y_2)$  – непрерывная дифференцируемая на  $R_2$  функция:

$$y = (y_1(x), y_2(x)) \in G,$$

$$J(y) = J(y_1, y_2) = \int_0^1 \varphi(y_1(x), y_2(x)) dx.$$

*Необходимое условие экстремума* имеет вид

$$\int_0^1 \nabla \varphi(y^*(x)) v(x) dx = 0,$$

где  $v(x) = (v_1(x), v_2(x))$ ,  $\nabla \varphi = \left( \frac{\partial \varphi}{\partial y_1}, \frac{\partial \varphi}{\partial y_2} \right)$ .

*Достаточное условие.* Разложим  $F(\varepsilon)$  в ряд Тейлора в точке  $\varepsilon = 0$  до второго порядка включительно:

$$F(\varepsilon) = J(y^* + \varepsilon \cdot v) = J(y^*) + \varepsilon \cdot \delta J(y^*) + \frac{1}{2} \varepsilon^2 \cdot \delta^2 J(y^*) + O(\varepsilon^2). \quad (9.5)$$

Считаем, что  $\delta J(y^*) = 0$ . Тогда, если

$$J(y^* + \varepsilon \cdot v) - J(y^*) > 0,$$

то в точке  $y^*$  функционал достигает своего минимума. Последнее условие с учетом выражения (9.5) приобретает вид

$$\frac{1}{2} \varepsilon^2 \cdot \delta^2 J(y^*) + O(\varepsilon^2) > 0,$$

или с точностью до членов второго порядка малости по  $\varepsilon$

$$\frac{1}{2} \varepsilon^2 \cdot \delta^2 J(y^*) > 0,$$

или  $\delta^2 J(y^*) > 0$  – достаточное условие.

*Пример 3.* Пусть  $\varphi(y)$  – дважды дифференцируемая непрерывная функция на  $R_1$  и пусть дан функционал

$$J(y) = \int_0^1 \varphi(y(x)) dx.$$

Разложим в ряд Тейлора подынтегральную функцию:

$$\begin{aligned} J(y^* + \varepsilon \cdot v) &= \int_0^1 \varphi(y^*(x) + \varepsilon \cdot v(x)) dx = \\ &= \int_0^1 \left\{ \varphi(y^*(x)) + \varepsilon \cdot \varphi'(y^*(x)) \cdot v(x) + \frac{1}{2} \varepsilon^2 \cdot \varphi''(y^*(x)) \cdot v^2(x) \right\} dx = \\ &= J(y^*) + \varepsilon \int_0^1 \varphi'(y^*(x)) \cdot v(x) dx + \frac{1}{2} \int_0^1 \varphi''(y^*(x)) \cdot v^2(x) dx. \end{aligned}$$

Если  $\delta J(y^*) = 0$  и  $\delta^2 J(y^*) = \int_0^1 \varphi''(y^*(x)) \cdot v^2(x) dx > 0$ , то в этой

точке функционал достигает минимума.

Пример 4. Пусть

$$J(y_1, y_2) = \int_0^1 \varphi(y_1(x), y_2(x)) dx.$$

Можно показать, что достаточные условия (ДУ) экстремума имеют вид

$$\int_0^1 (v_1(x), v_2(x))^T \nabla^2 \varphi(y_1^*, y_2^*) (v_1(x), v_2(x)) dx > 0,$$

где

$$\nabla^2 \varphi = \begin{pmatrix} \frac{\partial^2 \varphi}{\partial y_1^2} & \frac{\partial^2 \varphi}{\partial y_1 \partial y_2} \\ \frac{\partial^2 \varphi}{\partial y_1 \partial y_2} & \frac{\partial^2 \varphi}{\partial y_2^2} \end{pmatrix}.$$

### 9.3. Вариационные задачи с закрепленными концами

Сформулируем основную лемму вариационного исчисления. Рассмотрим функционал  $J = \int \varphi(t, y(t)) dt$ . Необходимое условие (НУ) экстремума имеет вид

$$\int_E (\varphi_y(t), v(t)) dt = 0, \quad (9.6)$$

где  $\varphi_y = \frac{\partial \varphi}{\partial y} = \left\{ \frac{\partial \varphi}{\partial y_1}, \dots, \frac{\partial \varphi}{\partial y_n} \right\}$ ;  $v(t) = \{v_1(t), \dots, v_n(t)\}$  – произвольная векторная непрерывная функция;  $E$  – область изменения аргумента  $t$ ;  $(\varphi_y(t), v(t))$  – скалярное произведение элементов векторного пространства непрерывных функций.

**Лемма.** Если равенство (9.6) выполняется для любых непрерывных функций  $v(t) \in Y(E, R_n)$ , а  $\varphi_y(t)$  также непрерывная векторная функция  $\varphi_y(t) \in Y(E, R_n)$ , то  $\varphi_y(t) \equiv 0$ ,  $t \in E$ .

В общем случае вместо  $\varphi_y(t)$  может быть любая другая вектор-функция  $f(t) = \{f_1(t), \dots, f_n(t)\}$ . Поэтому, если

$$\int_E (f(t), v(t)) dt = 0,$$

то  $f(t) \equiv 0$ .

#### 9.4. Уравнение Эйлера для вариационных задач с закрепленными концами

Уравнения Эйлера представляют собой необходимые условия существования экстремума для определенного вида функционалов (рисунок 9.3).

Получим НУ экстремума функционала

$$J(y, y') = \int_a^b \varphi(x, y(x), y'(x)) dx, \quad (9.7)$$

где  $y(x)$  – скалярная непрерывная функция с непрерывной первой производной  $y'(x)$ ;  $\varphi$  – известная непрерывная дифференцируемая функция своих аргументов, причем

$$y(a) = y_a, y(b) = y_b, \quad (9.8)$$

где  $y_a, y_b$  – заданные числа. Выражения (9.8) есть *граничные условия* (ГУ) вариационной задачи.

Считаем, что экстремум функционала (9.7) достигается в точке  $\left[ y^*, (y^*)' \right]$ . Тогда, приравнявая нулю первую вариацию, получаем

$$\delta J(y, y') = \int_a^b \{ \varphi_y(y, y') \cdot v(x) + \varphi_{y'}(y, y') \cdot v'(x) \} dx = 0, \quad (9.9)$$

где

$$\varphi_y = \frac{\partial \varphi}{\partial y}, \varphi_{y'} = \frac{\partial \varphi}{\partial y'}; \quad (9.10)$$

$v(x)$  – непрерывная с непрерывной производной  $v'(x)$  функция, удовлетворяющая условиям

$$v(a) = v(b) = 0, \text{ так как } (y + \varepsilon \cdot v)|_{x=a}^{x=b} = \begin{cases} y_b, \\ y_a. \end{cases} \quad (9.11)$$

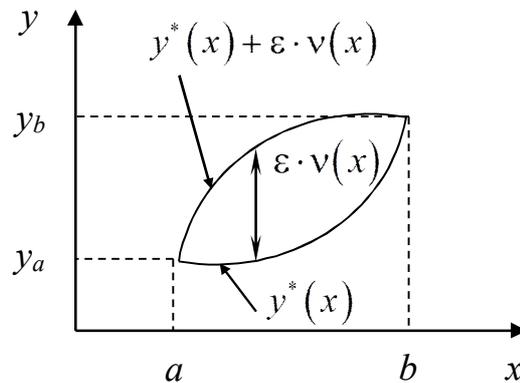


Рисунок 9.3 – Функция и ее вариация для задачи с закрепленными концами

Преобразуем выражение (9.9). Интегрируем второе слагаемое в нем по частям:

$$\int_a^b \varphi_{y'} v' dx = \varphi_{y'} v(x) \Big|_a^b - \int_a^b v(x) \frac{d}{dx} \varphi_{y'} dx,$$

где  $\varphi_{y'} = \frac{\partial \varphi}{\partial y'}$ .

С учетом выражения (9.11) первое слагаемое равно нулю, тогда уравнение (9.9) преобразуется в уравнение

$$\int_a^b \left\{ \varphi_y - \frac{d}{dx} \varphi_{y'} \right\} v(x) dx = 0. \quad (9.12)$$

Так как  $v(x)$  – произвольная функция, то из выражения (9.12) следует

$$\varphi_y - \frac{d}{dx} \varphi_{y'} = 0, \quad (9.13)$$

где  $\varphi_y = \frac{\partial \varphi(x, y, y')}{\partial y}$ .

Уравнение (9.14) называется *уравнением Эйлера*. В развернутом виде это уравнение записывается следующим образом:

$$y''(x)\varphi_{y'y'} + y'(x)\varphi_{yy'} + \varphi_{xy'} - \varphi_y = 0, \quad (9.14)$$

где  $\varphi_{y'y'}$ ,  $\varphi_{yy'}$ ,  $\varphi_{xy'}$  – смешанные частные производные 2-го порядка.

Общее решение уравнения (9.14) зависит от двух констант –  $c_1$  и  $c_2$ , т.е.

$$\Phi(y^*(x), x, c_1, c_2) = 0. \quad (9.15)$$

Определить  $c_1$  и  $c_2$  можно из условия закрепленности концов траектории, т.е. из условия (9.8):

$$\Phi(y_a, a, c_1, c_2) = 0, \quad \Phi(y_b, b, c_1, c_2) = 0. \quad (9.16)$$

Таким образом, поиск экстремума функционала (9.7) сводится к решению уравнения Эйлера (9.13) или (9.14) с двумя краевыми условиями. Решение его представляет самостоятельную сложную задачу.

*Пример.* Пусть  $J(y, y') = \int_0^{\pi/2} [y^2 - (y')^2] dx$ ,  $y(0) = 1$ ,  $y(\pi/2) = 0$ .

Запишем уравнение Эйлера:

$$2(y + y'') = 0.$$

Общее решение имеет вид  $y(x) = c_1 \sin x + c_2 \cos x$ . Из краевых условий находим  $c_1 = 0$ ,  $c_2 = 1$ , тогда  $y^*(x) = \cos x$ .

Уравнение Эйлера (9.14) не всегда интегрируемо, а в ряде случаев его решение может вызвать затруднения. Перечислим частные случаи, в которых решение уравнения Эйлера упрощается по сравнению с общим случаем.

**Случай 1.** Функция  $\varphi$  не зависит от  $y'$ , т.е.  $\varphi = \varphi(x, y(x))$ . Уравнение (9.14) принимает вид

$$\varphi_y(x, y) = 0.$$

В общем случае решение этого алгебраического (недифференциального) уравнения не существует в классе функций, лежащих внутри пространства  $G$ .

*Пример.* Пусть  $\varphi(x, y) = \frac{1}{2}y^2$ .

Уравнение Эйлера  $y = 0$  имеет решение  $y(x) = 0$ , которое не удовлетворяет ГУ (9.8) задачи за исключением случая, когда  $y_a = y_b = 0$ .

Функционал достигает своего экстремума на кривой (рисунок 9.4), удовлетворяющей ограничениям (9.8), но не относящейся к классу непрерывных функций.

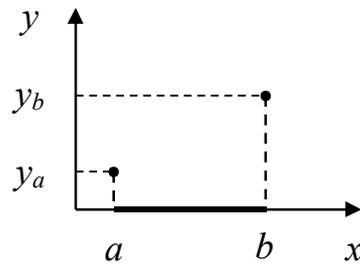


Рисунок 9.4 – Решение уравнения Эйлера для случая 1

**Случай 2.** Функция  $\varphi$  зависит только от  $y'$ :  $\varphi = \varphi(y')$ . Уравнение Эйлера имеет вид

$$\frac{d}{dx}\varphi_{y'} = 0 \text{ или } y''(x)\varphi_{y'y'}^* = 0,$$

а его общее решение –  $y(x) = c_1x + c_2$ . Таким образом, в данном случае экстремум функционала достигается на семействе прямых (двухпараметрическом семействе прямых).

*Справка.* Действительно, из уравнения Эйлера следует либо  $y'' = 0$ , либо  $\varphi_{y'y'}^* = 0$ . В первом случае  $y(x) = c_1x + c_2$ . Во втором случае мы имеем алгебраическое уравнение  $\varphi_{y'y'}(y') = 0$  (здесь обозначение (\*) опускаем) относительно  $y'$ . Пусть  $\varphi_{y'y'}(y')$  является многочленом степени  $m$ , т.е.

$$\Phi_{y'y'}(y') = \prod_{i=1}^m (y' - k_i).$$

Тогда его корнями являются  $k_1, k_2, \dots, k_m$ , т.е.  $y' = k_i, i = 1, 2, \dots, m$ , или  $y(x) = k_i x + c, i = 1, 2, \dots, m$ . Это однопараметрическое семейство прямых линий, входящее в двухпараметрическое семейство  $y(x) = c_1 x + c_2$ .

*Пример 1.* Найти траекторию перемещения из точки  $A(a, y_a)$  в точку  $B(b, y_b)$  за минимальное время, если скорость движения  $v$  зависит лишь от  $y'$ , то есть  $v = v(y')$ . Минимизируемый функционал имеет вид (см. задачу о брахистохроне, где вместо  $v(x) = \sqrt{2g \cdot y(x)}$  мы подставляем  $v = v(y')$ )

$$J(y') = \int_a^b \frac{\sqrt{1 + (y')^2}}{v(y')} dx.$$

Так как функционал  $J$  зависит только от  $y'$ , то экстремум может достигаться лишь на прямых линиях  $y(x) = c_1 x + c_2$ , где  $c_1$  и  $c_2$  определяются краевыми условиями.

*Пример 2.* Из всех кривых, соединяющих две точки  $A$  и  $B$ , наименьшую длину

$$J(y') = \int_a^b \sqrt{1 + (y')^2} dx$$

имеет прямая линия.

**Случай 3.** Подынтегральная функция линейно зависит от  $y'$ :

$$\Phi(x, y, y') = \Phi_1(x, y) + \Phi_2(x, y) y'.$$

Уравнение Эйлера  $\partial\Phi_1/\partial y - \partial\Phi_2/\partial x = 0$  представляет, как и в первом случае, алгебраическое (а не дифференциальное) уравнение и его решение из класса непрерывных с непрерывной первой производной

функции ( $G = C_1$ ) не удовлетворяет ГУ (9.8) за исключением частных случаев.

*Пример 1.* Пусть  $J(y, y') = \int_0^1 (y^2 + x^2 y') dx$ ,  $y(0) = 0$ ,  $y(1) = b$ .

Запишем уравнение Эйлера:  $2y - 2x = 0$ .

Таким образом,  $y(x) = x$  удовлетворяет краевым условиям лишь при  $b = 1$ .

*Пример 2:*  $J(y, y') = \int_a^b (y^2 + 2xyy') dx = \int_{ay_a^2}^{by_b^2} d(xy^2) = by_b^2 - ay_a^2$ .

**Случай 4.** Функция  $\varphi$  не зависит от  $y$ , т.е.  $\varphi = \varphi(x, y')$ . Уравнение Эйлера  $\frac{d}{dx} \varphi_{y'}(x, y') = 0$ , откуда получаем первый интеграл  $\varphi_{y'}(x, y') = C_1$ . То есть имеем дифференциальное уравнение первого порядка. Иногда удается получить его аналитическое решение.

*Пример.* Время, затрачиваемое на перемещение по кривой  $y(x)$  из точки  $A(a, y_a)$  в точку  $B(b, y_b)$ , если скорость движения  $v = x$ , представляет собой функционал

$$J(y') = \int_a^b \frac{\sqrt{1 + (y')^2}}{x} dx.$$

Первый интеграл уравнения Эйлера равен

$$\frac{y'}{x\sqrt{1 + (y')^2}} = \frac{1}{c_1}.$$

Решение его проще искать в параметрической форме. Делаем замену  $y' = \operatorname{tg} \gamma$ , где  $\gamma$  – параметр, тогда

$$x = \frac{c_1 y'}{\sqrt{1 + (y')^2}} = c_1 \cdot \left[ \frac{\sin \gamma / \cos \gamma}{1 / \cos \gamma} \right] = c_1 \sin \gamma.$$

Находим  $y$  как функцию параметра  $\gamma$ :

$$dy = y'dx = c_1 \operatorname{tg} \gamma \cos \gamma d\gamma = c_1 \sin \gamma d\gamma,$$

$$y = -c_1 \cos \gamma + c_2.$$

Итак, экстремум достигается на кривых, заданных в параметрической форме:

$$\left\{ \begin{array}{l} y - c_2 = -c_1 \cos \gamma \\ x = c_1 \sin \gamma \end{array} \right\}.$$

Это семейство окружностей  $(y - c_2)^2 + x^2 = c_1^2$  с центром на оси ординат. Параметры  $c_1, c_2$  определяются из ГУ задачи.

**Случай 5.** Функция  $\varphi$  не зависит от  $x$ , т.е.  $\varphi = \varphi(y, y')$ . В этом случае уравнение Эйлера имеет вид

$$\varphi_y - \varphi_{yy'} y' - \varphi_{y'y'} y'' = 0.$$

Умножим обе части этого равенства на  $y'$ , получим  $\frac{d}{dx}(\varphi - y' \varphi_{y'}) = 0$ , откуда находим первый интеграл уравнения Эйлера:  $\varphi - y' \varphi_{y'} = c_1$ .

Это дифференциальное уравнение первого порядка можно проинтегрировать, разрешив его относительно  $y'$  и разделив переменные, или путем введения параметра.

*Пример.* Решим задачу о брахистохроне:

$$J(y, y') = \frac{1}{2g} \int_0^b \frac{\sqrt{1 + (y')^2}}{\sqrt{y}} dx \rightarrow \min.$$

Первый интеграл

$$\frac{\sqrt{1 + (y')^2}}{\sqrt{y}} - \frac{(y')^2}{\sqrt{y(1 + (y')^2)}} = \frac{1}{\sqrt{c_1}}.$$

После преобразований получим  $y \left[ 1 + (y')^2 \right] = c_1$ . Решение ищем в параметрическом виде, вводя  $y' = \operatorname{ctg} \gamma$ . Тогда

$$y = \frac{c_1}{1 + (y')^2} = \frac{c_1}{1 + \operatorname{ctg}^2 \gamma} = c_1 \sin^2 \gamma = \frac{1}{2} c_1 (1 - \cos 2\gamma),$$

$$dx = \frac{dy}{y'} = \frac{2c_1 \sin \gamma \cos \gamma}{\operatorname{ctg} \gamma} d\gamma = 2c_1 \sin^2 \gamma d\gamma = c_1 (1 - \cos 2\gamma) d\gamma,$$

$$x = c_1 \int (1 - \cos 2\gamma) d\gamma = c_1 \left( \gamma - \frac{1}{2} \sin 2\gamma \right) + c_2 =$$

$$= \frac{1}{2} c_1 (2\gamma - \sin 2\gamma) + c_2, \quad c_2 = 0, \text{ так как } y(0) = 0.$$

При  $\gamma = 0$ ,  $x = c_2 = 0 \rightarrow c_2 = 0$ ,  $v(0) = \sqrt{2g \cdot y(0)} = 0$ . Получим параметрическое уравнение циклоиды:

$$y = \frac{c_1}{2} (1 - \cos 2\gamma), \quad x = \frac{c_2}{2} (2\gamma - \sin 2\gamma).$$

## 9.5. Многомерный случай

Приведенные выше результаты автоматически обобщаются на многомерный случай:

$$J(y, y') = \int_a^b \varphi(x, y(y), y'(x)) dx, \quad y(a) = y_a, \quad y(b) = y_b, \quad (9.17)$$

где

$$y = (y_1, y_2, \dots, y_n), \quad y' = (y'_1, y'_2, \dots, y'_n),$$

$$y_a = \{y_{a1}, y_{a2}, \dots, y_{an}\}, \quad y_b = \{y_{b1}, y_{b2}, \dots, y_{bn}\}.$$

Метод получения уравнений Эйлера полностью сохраняется, единственное отличие – необходимо правильно выполнять соответствующие операции над векторными величинами. В результате для первой вариации получим

$$\delta J(y, y') = \int_a^b \left\{ \varphi_y - \frac{d}{dx} \varphi_{y'} \right\} v(x) dx = 0, \quad (9.18)$$

где  $v(x) = \{v_1(x), v_2(x), \dots, v_n(x)\}$ .

Из соотношения (9.18) получим систему дифференциальных уравнений Эйлера

$$\varphi_y - \frac{d}{dx} \varphi_{y'} = 0, \quad (9.19)$$

где  $\varphi_y = \{\varphi_{y_1}, \varphi_{y_2}, \dots, \varphi_{y_n}\}$ ,  $\varphi_{y'} = \{\varphi_{y'_1}, \varphi_{y'_2}, \dots, \varphi_{y'_n}\}$ . Система (9.19) решается совместно с ГУ  $y(a) = y_a$ ,  $y(b) = y_b$ . Таким образом, приходим к следующей теореме.

**Теорема.** Для того чтобы набор функций  $y_1(x), \dots, y_n(x) \in C_1[a, b]$  доставлял экстремум функционалу (9.17), необходимо, чтобы эти функции удовлетворяли системе дифференциальных уравнений Эйлера

$$\varphi_{y_k} - \frac{d}{dx} \varphi_{y'_k} = 0, \quad k = 1, \dots, n. \quad (9.20)$$

*Пример.* Найти условие экстремума функционала

$$J(y_1, y_2, y'_1, y'_2) = \int_0^{\pi/2} \left( (y'_1)^2 + (y'_2)^2 + 2y_1 y_2 \right) dx,$$

$$y_1(0) = 0, \quad y_2(0) = 0, \quad y_1(\pi/2) = 1, \quad y_2(\pi/2) = -1.$$

Составляем систему уравнений Эйлера:

$$\begin{cases} y_1'' - y_2 = 0 \\ y_2'' - y_1 = 0 \end{cases}$$

Решаем ее:

$$y_1(x) = c_1 e^x + c_2 e^{-x} - c_3 \cos x - c_4 \sin x,$$

$$y_2(x) = c_1 e^x + c_2 e^{-x} + c_3 \cos x + c_4 \sin x.$$

Из ГУ находим константы:  $c_1 = c_2 = c_3 = 0$ ,  $c_4 = -1$ . Получим кривые  $y_1^*(x) = \sin x$ ,  $y_2^*(x) = -\sin x$ , на которых функционал может достигать экстремума.

## 9.6. Уравнения Эйлера – Пуассона

Найдем необходимое условие экстремума функционала

$$J(y, y', y'', \dots, y^{(n)}) = \int_a^b \varphi(x, y, y', \dots, y^{(n)}) dx \quad (9.21)$$

при наличии ограничений

$$\begin{cases} y(a) = y_a, y'(a) = y'_a, \dots, y^{(n-1)}(a) = y_a^{(n-1)}, \\ y(b) = y_b, y'(b) = y'_b, \dots, y^{(n-1)}(b) = y_b^{(n-1)}, \end{cases} \quad (9.22)$$

где  $y(x)$  – скалярная непрерывная функция с непрерывными производными  $y'(x), y''(x), \dots, y^{(n)}(x)$ ;  $\varphi$  – известная непрерывная дифференцируемая функция своих аргументов.

**Теорема.** Для того чтобы функционал (9.21) достигал на функции  $y(x) \in C_n[a, b]$  локального экстремума, необходимо, чтобы эта функция удовлетворяла уравнению Эйлера – Пуассона

$$\varphi_y - \frac{d}{dx} \varphi_{y'} + \frac{d^2}{dx^2} \varphi_{y''} - \dots + (-1)^n \frac{d^n}{dx^n} \varphi_{y^{(n)}} = 0. \quad (9.23)$$

Подобно случаю простейшей задачи ВИ, решения уравнения (9.23) (экстремали функционала (9.21)), удовлетворяющие ГУ (9.22), являются кривыми, на которых функционал (9.21) может достигать экстремума на множестве  $G = \{y(x) \in C_n[a, b] \mid \text{ГУ (9.22)}\}$ .

*Пример.* Найти экстремали функционала

$$J(y) = \int_0^1 \left( 120xy - \frac{1}{2}(y'')^2 \right) dx,$$

ГУ:  $y(0) = y'(0) = 0, y(1) = 1, y'(1) = 6$ .

Запишем уравнение Эйлера – Пуассона:

$$y^{(4)} = 120x.$$

Его общее решение будет

$$y(x) = x^5 + c_1x^3 + c_2x^2 + c_3x + c_4.$$

Отсюда с помощью ГУ получаем систему уравнений для определения констант  $c_i, i = 1, 2, 3, 4$ .

Проверить самостоятельно:

$$\left. \begin{cases} c_4 = 0, \\ c_3 = 0, \\ c_1 + c_2 + c_3 + c_4 = 0, \\ 3c_1 + 2c_2 + c_3 = 1 \end{cases} \right\} \Rightarrow c_1 = 1, c_2 = -1.$$

Ответ:  $y(x) = x^5 + x^3 - x^2$ .

### Вопросы для самопроверки

1. Что такое вариационное исчисление? Дайте понятие функционала.
2. Приведите необходимые и достаточные условия существования экстремума функционала.
3. Сформулируйте основную лемму вариационного исчисления.
4. Что такое вариационные задачи с закрепленными концами?
5. Приведите уравнение Эйлера для случаев 1 и 2 вариационных задач с закрепленными концами.
6. Приведите уравнение Эйлера для случаев 3 и 4 вариационных задач с закрепленными концами.
7. Приведите уравнение Эйлера для случая 5 вариационных задач с закрепленными концами.
8. Приведите уравнение Эйлера для многомерного случая вариационных задач с закрепленными концами.
9. Приведите уравнение Эйлера – Пуассона.

## ГЛАВА 10. ЭВРИСТИЧЕСКИЕ МЕТОДЫ ПОИСКА ЭКСТРЕМУМОВ ФУНКЦИЙ

### 10.1. Понятие и особенности эвристических методов

За последние десятилетия были разработаны новые методы оптимизации, концептуально отличающиеся от традиционных. Большинство из них основано на определенных характеристиках, присущих биологическим, молекулярным, физическим и нейробиологическим системам.

*Эвристические методы* основаны на подсознательном мышлении и характеризуются неосознанным (интуитивным) способом действий для достижения осознанных целей. Эвристические методы еще называют методами инженерного (изобретательского) творчества.

*Эвристический алгоритм* – это алгоритм решения задачи, правильность которого для всех возможных случаев не доказана, но про который известно, что он дает достаточно хорошее решение в большинстве случаев. В действительности может быть даже известно (т.е. доказано) то, что эвристический алгоритм формально неверен. Однако его все равно можно применять, если он дает неверный результат только в достаточно редких и хорошо выделяемых случаях или же дает неточный, но все же приемлемый результат.

Проще говоря, эвристика – это не полностью математически обоснованный (или даже «не совсем корректный»), но при этом практически полезный алгоритм.

Важно понимать, что эвристика, в отличие от корректного алгоритма решения задачи, обладает следующими особенностями:

- она не гарантирует нахождение лучшего решения;
- она не гарантирует нахождение решения, даже если оно заведомо существует (возможен «пропуск цели»);
- она может дать неверное решение в некоторых случаях.

## 10.2. Классификация эвристических методов

Дана целевая функция  $f(x) = f(x_1, \dots, x_n)$ , определенная на множестве допустимых решений  $D \in R^n$ . Требуется найти глобальный условный минимум функции  $f(x)$  на множестве  $D$

$$f(x^*) = \min_{x \in D} (f(x)), \quad (10.1)$$

где  $D = \{x \mid x_i \in [a_i, b_i], i = 1, \dots, n\}$ .

На рисунке 10.1 представлены классы эвристических методов.

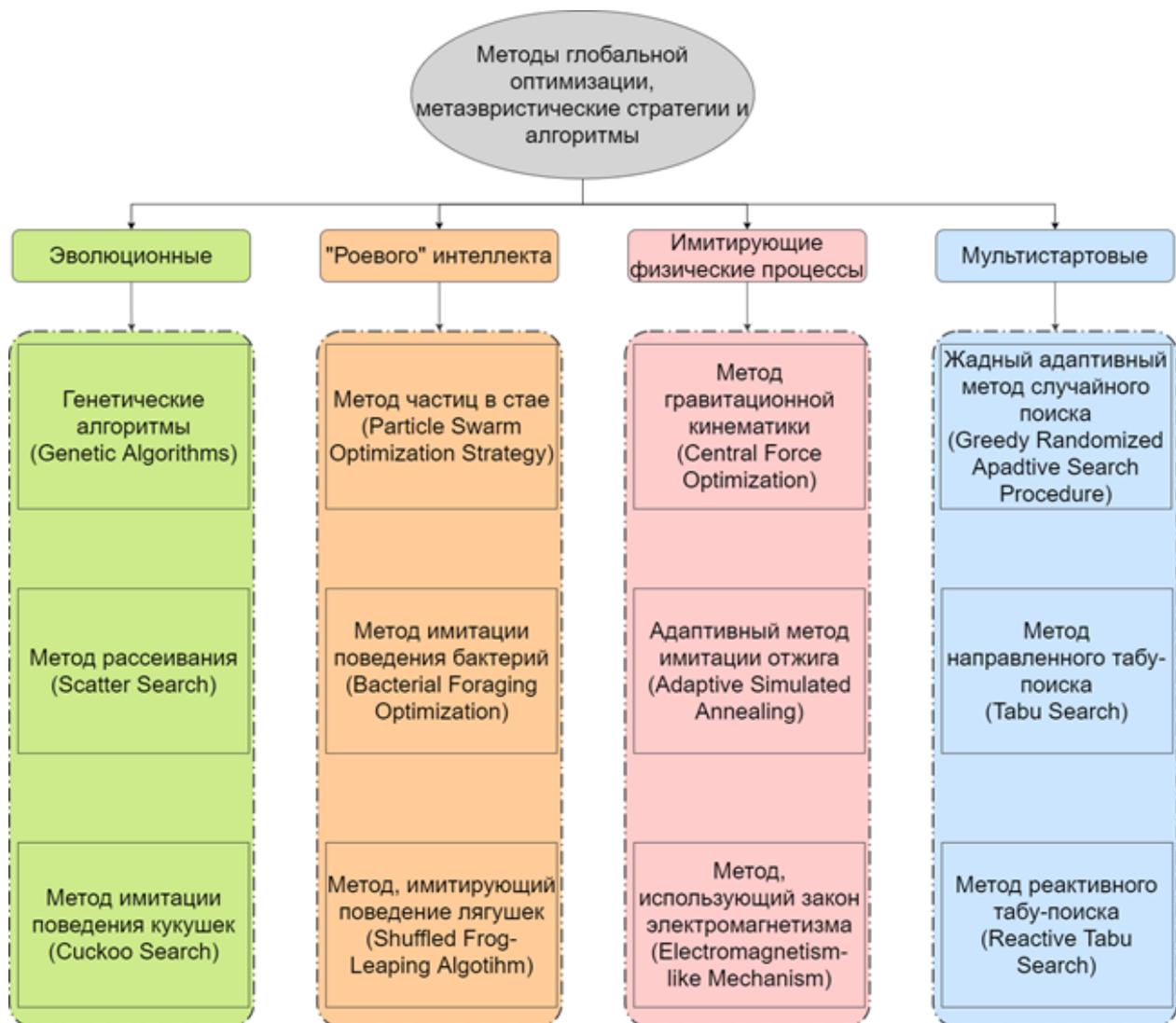


Рисунок 10.1 – Классификация эвристических методов

Приведем более полный перечень эвристических методов.

1. Эволюционные методы:

- генетические алгоритмы [24];
- методы, имитирующие иммунные системы организмов, – методы искусственных иммунных систем [24];
- метод рассеивания [24];
- эволюционная стратегия преобразования ковариационной матрицы [24];
- метод динамических сеток [31];
- метод дифференциальной эволюции [24];
- метод, имитирующий распространение сорняков [24];
- метод, имитирующий поведение кукушек [24];
- метод растущих деревьев [28];
- метод эволюции разума [28];
- метод бактериальной оптимизации [30];
- культурный метод [30];
- меметический метод [28];
- самообразующийся миграционный метод [28];
- метод прокладки путей [28];
- метод интеллектуальных капель воды [28].

2. Методы «роевого» интеллекта:

- метод частиц в стае [24];
- метод муравьиных колоний [24];
- метод имитации поведения бактерий [24];
- методы пчелиных колоний [24];
- метод, имитирующий поведение стаи рыб в поисках корма [24];
- метод, имитирующий поведение летучих мышей [34];
- метод, имитирующий поведение светлячков [24, 35];
- алгоритм, имитирующий поведение лягушек [36];
- метод, имитирующий динамику формирования рек [28];
- метод, имитирующий поведение котов [37];
- метод, имитирующий поведение обезьян [38].

3. Методы, имитирующие физические процессы:
  - метод гравитационной кинематики [24];
  - метод имитации отжига (и его модификации – параллельный, быстрый, адаптивный метод имитации отжига) [24, 28, 30];
  - метод поиска гармонии [30];
  - метод, использующий закон электромагнетизма [24, 40];
  - метод экстремальной оптимизации [32];
  - метод стохастического диффузионного поиска [39].
4. Мультистартовые методы:
  - жадный адаптивный метод случайного поиска [33];
  - метод направленного табу-поиска [24];
  - метод реактивного табу-поиска [30].

### **Эволюционные методы**

**Эволюционные методы поиска** (Evolutionary Methods) – эвристические методы, в основе которых лежит эволюционная теория Чарльза Дарвина. Способы реализации этих методов значительно отличаются друг от друга. Идея состоит в наличии популяции индивидов, зависимой от внешней среды, что влечет за собой естественный отбор. Каждая особь из популяции соответствует одному из значений целевой функции. Особь эволюционирует за счет мутаций и селекционирования. Получается новая популяция с лучшими характеристиками, то есть с лучшим значением целевой функции.

**Генетический алгоритм** (Genetic Algorithm) – классический эволюционный алгоритм, основанный на принципе естественного эволюционного отбора. Главной идеей является борьба за существование между решениями задачи. Каждое решение записывается в виде вектора значений (аллелей), который называется хромосомой или особью. Совокупность решений называют популяцией. Каждая особь в популяции оценивается значением целевой функции, рассчитанной на основе значений из хромосомы. Более перспективные решения проходят в следующую стадию и оказывают влияние на «потомство», т.е. на вновь генерируемые решения.

**Метод динамических сеток** (Variable Mesh Optimization). Его работа заключается в эволюции начальной популяции, называемой сеткой, во время которой происходит смена одного поколения другим путем расширения и последующего сокращения популяции.

**Метод дифференциальной эволюции** (Differential Evolution) – метод, основанный на формировании случайной популяции, состоящей из векторов (возможных решений). В каждом поколении происходит образование новых векторов путем комбинации трех случайных векторов из прошлого поколения. Если новый вектор оказывается лучше старого, то он заменяется, иначе остается.

**Метод рассеивания** (Scatter Search) – впервые представлен в 1977 г. как эвристический метод для целочисленного программирования. Однако в реальности он практически не применялся и не изучался вплоть до 1990 года, когда был представлен на семинаре EPFL «Исследования и методы поиска с помощью искусственного интеллекта» (Лозанна, Швейцария).

Метод рассеивания активно используется при проектировании процедур решения многих сложных задач комбинаторной оптимизации.

С помощью метода рассеивания решаются задачи для поточного цеха, мастерских, планирования проектов, дизайна территории, финансовые и статистические задачи, производственные задачи, в том числе для производства угля и стали, и другие.

Активное использование метода рассеивания обнаруживает резервы для его дальнейшего усовершенствования.

**Методы искусственных иммунных систем** (Artificial Immune System) – это класс автоматизированных вычислительных систем, которые основаны на принципах и процессах иммунной системы позвоночных. Обычно такие алгоритмы используют память и обучаемость иммунной системы для решения заданных проблем.

Искусственная иммунная система (ИИС) – адаптивная вычислительная система, использующая модели, принципы, механизмы и функции, описанные в теоретической иммунологии, которые применяются для решения прикладных задач.

Упоминания об ИИС появились в середине 1980-х годов в статьях Фармера, Паккарда и Перельсона, а также Берсини и Варела по иммунным сетям. Однако основы ИИС утвердились только к середине 1990-х годов. Форрест (Forrest) и Кепхарт (Kephart) опубликовали первую статью по ИИС (негативный отбор) в 1994 г., а Дасгупта (Dasgupta) провел обширные исследования негативного алгоритма отбора. Хант и Кук (Hunt, Cooke) начали работу над иммунным сетевым алгоритмом в 1995 г.; Тиммис и Нил (Timmis, Neal) продолжили эту работу и внесли некоторые улучшения. Первая книга по искусственным иммунным системам вышла под редакцией Дасгупты (Dasgupta) в 1999 г.

Несмотря на то что природные иммунные системы изучены далеко не полностью, сегодня существуют, по меньшей мере, три теории, объясняющие функционирование иммунной системы и описывающие взаимодействие ее элементов, а именно теория негативного отбора, теория клональной селекции и теория иммунной сети. Они легли в основу создания трех алгоритмов функционирования ИИС.

### **Методы «роевого» интеллекта**

**Алгоритмы роевого интеллекта (Swarm Intelligence)** основаны на изучении поведения колоний живых организмов, реализующих для всей колонии «оптимальное» поведение. Они зарекомендовали себя для решения сложных комбинаторных задач. Алгоритмы предполагают случайность и беспорядочность (аналогия с колониями муравьев, пчел и т.д.). Все начинается со случайной начальной позиции начальной популяции, затем происходит миграция агентов в пространстве поиска таким образом, чтобы приблизиться к искомому экстремуму целевой функции. Механика миграции агентов на каждом шаге идентична. Механизм обучения заимствуется из информации о поведении конкретных видов живых существ (муравьев, пчел, светлячков и пр.).

**Метод муравьиных колоний (Ant Colony Optimization).** Его общий принцип заключается в моделировании поведения муравьиной колонии в естественных условиях. Муравьи выделяют химическое вещество, называемое феромоном, для них этот гормон является чем-то

вроде языка общения. С помощью феромона муравьи, например, объясняют, где искать пищу и т.п. При движении они оставляют за собой след, наткнувшись на который другой муравей принимает решение, следовать по нему или пойти иным путем. Чем больше особей прошло по одному пути, тем больше гормона остается на земле и, следовательно, этот путь становится более предпочтительным для остальных. В методе отсутствуют явные лидеры, каждый член колонии одинаково вносит вклад в поиск решения. Выбор направления (генерация новых координат) определяется в два этапа:

- выбирается случайная плотность вероятности методом рулетки (чем лучше значение, тем больше шансов пойти по лучшему пути);

- формируется новая координата, определяемая параметрами математического ожидания и среднеквадратическим отклонением.

**Методы пчелиных колоний (Bees Algorithms).** Их общий принцип заключается в моделировании действий пчел во время поиска и сбора нектара. Пчелы вылетают на разные участки в поисках нектара, вернувшись в улей, танцуют перед остальными специальный танец. Этим танцем они «объясняют», где находится нектар, как до него добраться и т.д. Группы незанятых летных пчел выдвигаются в лучшие места, следовательно, эти пчелы становятся занятыми. Они производят сбор в окрестностях лучших источников, а возвратившись, передают остальным пчелам свои результаты, и так этот процесс повторяется.

**Искусственный метод пчелиных колоний (Artificial Bee Colony).** Идея метода состоит в том, чтобы смоделировать поведение пчел при поиске и сборе нектара. Пчелы вылетают на разные участки и производят поиск нектара в окрестности этих участков. Причем на один участок вылетает одна пчела. Каждая особь имеет память, она запоминает параметры места, на которое вылетела. Если новый источник нектара оказался лучше, чем предыдущий (меньшее значение ЦФ), то пчела запоминает его. Вернувшись в улей, каждая пчела танцует перед остальными специальный танец, логика танца такая же, что и в методе роя пчел. Незанятые пчелы вероятностным образом выбирают, куда полетят. После этого пчелы повторяют процесс поиска,

запоминания и передачи остальным сведений о позиции. Если в процессе поиска решение не обновляется в течение итераций, то пчела, вылетавшая на этот источник, «забывает» его и выбирает случайную точку в области допустимых решений.

**Метод частиц в стае (Particle Swarm Optimization Strategy).** Идея метода состоит в том, чтобы смоделировать поведение животных (стаи) во время поиска пищи. Каждая особь из стаи рассматривается как отдельная частица, имеющая свою скорость и положение в пространстве. Частицы используют свой прошлый опыт на каждой итерации, делятся своим положением и «качеством решения» со всеми, чтобы скорректировать направление движения, если это необходимо, но получают это не все, так как имеются внешние факторы, например ветер, шумы, погода или чувствительность особи (старость и т.п.). Метод включает такое качество, как лидерство: в случае локальной неудачи отдельной частицы она принимает решение следовать за наблюдаемым лидером или ориентируется на собственный опыт.

**Метод, имитирующий поведение светлячков (Glowworms Swarm Optimization).** Светлячки случайным образом располагаются в области допустимых значений. Они привлекают других особей с помощью излучаемого света. Те, у которых тусклее свет, следуют за более яркой особью. Чем больше расстояние от источника света, тем слабее он виден. Если же в поле зрения не наблюдается более перспективной особи, светлячок перемещается в случайную точку.

**Метод, имитирующий поведение рыб в стае (Fish School Search)** – алгоритм, имитирующий поиск рыбой пищи для роста и размножения. Каждая рыба имеет память и вес, который набирается при питании и сбрасывается во время плавания. Косяк запоминает наилучшие места, способные прокормить большую часть рыб. Эволюционирование косяка происходит путем размножения, обмена опытом и информацией, а также благодаря совместному движению.

**Метод имитации поведения бактерий (Bacterial Foraging Optimization)** – метод, разработанный Кевином Пассино в 2002 году как продолжение алгоритмов на основе роя. Применение стратегии группового кормодобывания бактерий *E. coli* для мультиэкстремальных функций является ключевой идеей данного алгоритма. Бактерии ищут

питательные вещества таким образом, чтобы получить максимальную энергию в единицу времени. Они общаются между собой, посылая сигналы. Бактерия принимает решение о добыче пищи после рассмотрения двух предыдущих факторов. Процесс, в котором бактерия перемещается, делая небольшие шаги в поисках питательных веществ, называется хемотаксисом. Ключевая идея метода – имитировать хемотаксическое движение виртуальных бактерий в пространстве поиска пищи.

### **Методы, имитирующие физические процессы**

В основе этих методов лежат физические законы и явления, такие как реструктуризация кристаллической решетки металла в процессе отжига, замерзание или нагревание жидкости, всемирный закон тяготения и т.д.

**Метод гравитационной кинематики** (Central Force Optimization) – детерминированный метод, в основе которого лежит закон всемирного тяготения. Уникальность его состоит в том, что каждое решение целевой функции рассматривается в качестве массы тела, участвующего в гравитационном взаимодействии с другими телами.

Гравитационный поиск оперирует двумя законами:

1) тяготения – каждая частица притягивает другие, а сила притяжения между двумя частицами прямо пропорциональна произведению их масс и обратно пропорциональна расстоянию между ними;

2) движения – текущая скорость любой частицы равна сумме части скорости в предыдущий момент времени и изменению скорости, равному силе, с которой воздействует система на частицу, деленной на инерциальную массу частицы.

**Метод имитации отжига** (Simulated Annealing) – метод, моделирующий кристаллическую решетку во время нагревания и охлаждения металла. Под действием высокой температуры атомы стараются занять места с наименьшей энергией. По мере охлаждения кристаллическая решетка сжимается, и атомы все меньше и меньше передвигаются на новые места. При высокой температуре больше вероятность принять худшее решение, чтобы проскочить локальный минимум.

**Адаптивный метод имитации отжига** (Adaptive Simulated Annealing) – метод, устраняющий недостатки базового метода. Все алгоритмы имитации отжига допускают с некоторой вероятностью переход в состояние с более высоким значением целевой функции для того, чтобы точка могла покинуть окрестности локального минимума. Для обеспечения этого свойства необходимо снижать температуру очень медленно, что увеличивает время подсчета. В адаптивном методе отжига этот недостаток устраняется путем ввода нового закона уменьшения температуры и возможностью ее изменения по каждой координате отдельно. Данный метод относится к классу стохастических. На каждой итерации разрешается переход в состояние с более высоким значением целевой функции с некоторой вероятностью, рассчитываемой исходя из закона Больцмана – Гиббса. В целом метод сильно схож с обычным методом имитации отжига.

**Метод поиска гармонии** (Harmony Search) – метод, разработанный и предложенный в 2001 году Geem Zong Woo. Некоторые авторы заявляют, что он был навеян игрой джаз-музыкантов, другие говорят, что в его основе лежит просто процесс создания приятной мелодии. Сути это не меняет. Важно лишь то, что опытный музыкант может достаточно быстро как сыграть с другими музыкантами, так и импровизировать что-нибудь прекрасное. Это и легло в основу алгоритма.

Классический алгоритм гармонического поиска оперирует понятием гармонической памяти (по аналогии с родительским пулом генетических алгоритмов). Здесь хранятся векторы, представляющие приближенные решения задачи оптимизации. Изначально они генерируются случайным образом в области, которая исследуется на наличие максимума или минимума (в зависимости от того, что нужно). Размер этой памяти, естественно, ограничен. Далее начинается магия импровизации (итеративная часть алгоритма).

Таким образом, метод имитирует процесс импровизации музыканта-исполнителя, подбирающего нужную ноту для достижения гармонии. Во время поиска каждое решение порождает значение ЦФ с целью достижения глобального экстремума. Используются идеи метода частиц в стае, метода отжига, стохастического градиента.

Сама импровизация заключается в следующем:

- генерируется пробный вектор (либо абсолютно случайно, либо с использованием векторов из памяти);
- пробный вектор модифицируется (происходит небольшой сдвиг в компонентах, если пробный вектор был создан с помощью памяти);
- модифицированный пробный вектор заменяет худший из имеющихся в памяти векторов.

На множестве генерируется заданное количество решений, для каждого из которых вычисляется значение искомой функции. Координаты решения и значения функции помещаются по строкам в матрицу – память гармонии. Среди всех решений выбирается наихудшее. Затем генерируется новое решение и сравнивается с наихудшим. Если по значению функции оно лучше, то его помещают в память вместо наихудшего. Так продолжается до максимального количества итераций.

Для получения очередной координаты с некоторой вероятностью случайным образом из памяти выбирается соответствующая координата решения. Иначе выбирается случайное значение на промежутке. Значение, взятое из памяти, с заданной вероятностью корректируется с помощью небольшого приращения. Если корректировки не было, то используется нескорректированное значение. Вся эта процедура позволяет выйти за область локального минимума, для поиска новой гармонии. Это и есть процесс импровизации.

## **Мультистартовые методы**

**Мультистартовые методы** (Multi-start Methods) – методы, основанные на многократном отыскании локальных минимумов из разных начальных точек. При последующем поиске исключаются повторные спуски в те же локальные минимумы, затем происходит генерирование новых перспективных начальных точек.

Основная сложность метода в том, что для высокой надежности поиска точки глобального минимума нужно взять количество начальных точек для локальных алгоритмов больше, чем число локальных минимумов функции, которое обычно неизвестно.

**Жадный адаптивный метод случайного поиска** (Greedy Randomized Adaptive Search Procedure) – метод, использующий идею мультистарта. На каждой итерации есть две фазы: построение и локальный поиск. В результате первой фазы порождаются решения хорошего качества для второй фазы. Затем после второй фазы полученные точки берутся в качестве начальных для первой фазы и процедура продолжается до нахождения наилучшего приближенного решения. Этот вид жадного метода рандомизированного построения также известен как **полужадный эвристический метод**, впервые описанный J. Pirie Hart и Andrew W. Shogan.

Жадная рандомизированная адаптивная процедура поиска была впервые представлена Фео и Резенде.

Существуют вариации классического алгоритма, например Reactive GRASP. В этом варианте основной параметр, определяющий ограниченность списка кандидатов (restrictive list of candidates, RCL) на этапе строительства, саморегулируется в соответствии с качеством найденных ранее решений. Существуют также методы ускорения поиска, такие как изменение стоимости, функции смещения, запоминание и обучение, а также локальный поиск частично построенных решений.

**Метод направленного табу-поиска** (Taboo Search) – алгоритм поиска с запретами, основоположником которого является Ф. Гловер, предложивший принципиально новую схему локального поиска. Она включает три фазы: исследовательскую, перераспределительную, интенсивно-уточняющую. На исследовательской фазе генерируются точки вблизи текущего решения, в тех областях, где еще не производился поиск. На перераспределительной фазе формируется список посещенных областей. На интенсивно-уточняющей фазе улучшаются хорошие точки, найденные в первых двух фазах, с целью получения результата с заданной точностью.

## 10.3. Эволюционные методы оптимизации

Рассмотрим более детально некоторые эволюционные методы.

### 10.3.1. Генетические алгоритмы

**Генетические алгоритмы с бинарным кодированием.** *Целевая функция  $f(x)$  генетического алгоритма эквивалентна природному понятию приспособленности живого организма.*

Вектор параметров  $x = (x_1, x_2, \dots, x_n)^T$  целевой функции называется *фенотипом*, а отдельные его параметры  $x_i$ ,  $i = 1, \dots, n$  – *признаками*. Любой живой организм может быть представлен своим генотипом и фенотипом.

*Генотип* – это совокупность наследственных признаков, информация о которых заключена в хромосомном наборе.

*Фенотип* – совокупность всех признаков и свойств организма, формирующихся в процессе взаимодействия его генотипа и внешней среды. Каждый ген имеет свое отражение в фенотипе.

ГА ведут поиск решения только на уровне генотипа. Каждую координату  $x_i$  вектора  $x = (x_1, x_2, \dots, x_n)^T \in D$  представляют в некоторой форме  $s_i$ , удобной для использования в ГА и называемой *геном*. Для этого необходимо осуществить преобразование, в общем случае не взаимно однозначное, вектора параметров  $x$  в некоторую структуру  $s = (s_1, s_2, \dots, s_n)^T \in S$ , называемую *хромосомой (генотипом, особью)*:

$$D \xrightarrow{e} S,$$

где  $e$  – функция кодирования;  $S$  – пространство представлений.

Чтобы восстановить по хромосоме решение, необходимо задать обратное преобразование

$$S \xrightarrow{e^{-1}} D,$$

где  $e^{-1}$  – функция декодирования.

В пространстве представлений  $S$  вводится так называемая *функция приспособленности*  $\mu(s)$ :  $S \xrightarrow{\mu} R$ , где  $R$  – множество

вещественных чисел, аналогичная целевой функции  $f(x)$  на множестве  $D$ . Функцией  $\mu(s)$  может быть любая функция, удовлетворяющая следующему условию:

$$\forall x^1, x^2 \in D : s^1 = e(x^1), s^2 = e(x^2), s^1 \neq s^2, \\ \text{если } f(x^1) > f(x^2), \text{ то } \mu(s^1) > \mu(s^2).$$

Решение исходной оптимизационной задачи, например задачи максимизации  $f(x^*) = \max_{x \in D} f(x)$ , сводится к поиску решения  $s^*$  другой задачи оптимизации:

$$\mu(s^*) = \max_{s \in S} \mu(s). \quad (10.2)$$

При решении используются конечные наборы

$$I = \left\{ s^k = (s_1^k, s_2^k, \dots, s_n^k)^T, k = 1, 2, \dots, K \right\} \subset S$$

возможных решений, называемые *популяциями*, где  $s^k$  – хромосома с номером  $k$ ;  $K$  – размер популяции;  $s_i^k$  – ген с номером  $i$ .

Затем осуществляется обратное преобразование

$$x^* = e^{-1}(s^*). \quad (10.3)$$

Существуют различные способы задания функции приспособленности:

- 1) как правило, полагают  $\mu(s) = f(e^{-1}(s))$ ;
- 2) в некоторых случаях необходимо, чтобы функция  $\mu(s)$  принимала только положительные значения. Тогда функцию приспособленности можно представить, например, как

$$\mu(s^k) = f(s^k) + \left| \min_{k=1,2,\dots,m} f(s^k) \right| + 1 \\ \text{или } \mu(s^k) = \left( f(s^k) - \min_{k=1,2,\dots,m} f(s^k) + 1 \right)^2.$$

**Бинарное кодирование.** Существуют следующие способы бинарного кодирования.

1. Координате  $x_i$  ставится в соответствие целое число  $\beta_i \in \{0, 1, \dots, k_i\}$ :

$$\beta_i = \begin{cases} 0, & x_i = a_i, \\ \left[ \frac{x_i - a_i}{h_i} \right] + 1, & x_i \in (a_i, b_i), \\ k_i, & x_i = b_i, \end{cases} \quad (10.4)$$

где  $[\cdot]$  означает операцию выделения целой части числа.

2. Каждому номеру  $\{0, 1, \dots, k_i\}$  ставится в соответствие уникальная бинарная комбинация длиной  $l_i$ . Величина  $l_i$  определяется из условия  $k_i = 2^{l_i} - 1$ . Преобразование в битовую строку осуществляется с помощью обычного двоичного кодирования (позиционного кода) или *рефлексивного кода Грея*, который обладает свойством непрерывности бинарной комбинации: изменению кодируемого числа  $\beta_i$  на единицу соответствует изменение бинарной комбинации только в одном разряде.

**Декодирование.** Значение координаты  $x_i$  при  $\gamma_i = 0$  определяется левым концом промежутка  $[a_i, b_i]$ , при  $\gamma_i = k_i$  – правым концом, а при  $\gamma_i = 1, \dots, k_i - 1$  – серединой соответствующего подынтервала с номером  $\gamma_i$ :

$$x_i = \begin{cases} a_i, & \gamma_i = 0, \\ \frac{(a_i + (\gamma_i - 1)h_i) + (a_i + \gamma_i h_i)}{2}, & \gamma_i = 1, 2, \dots, k_i - 1, \\ b_i, & \gamma_i = k_i. \end{cases} \quad (10.5)$$

Общая схема алгоритма приведена на рисунке 10.2.

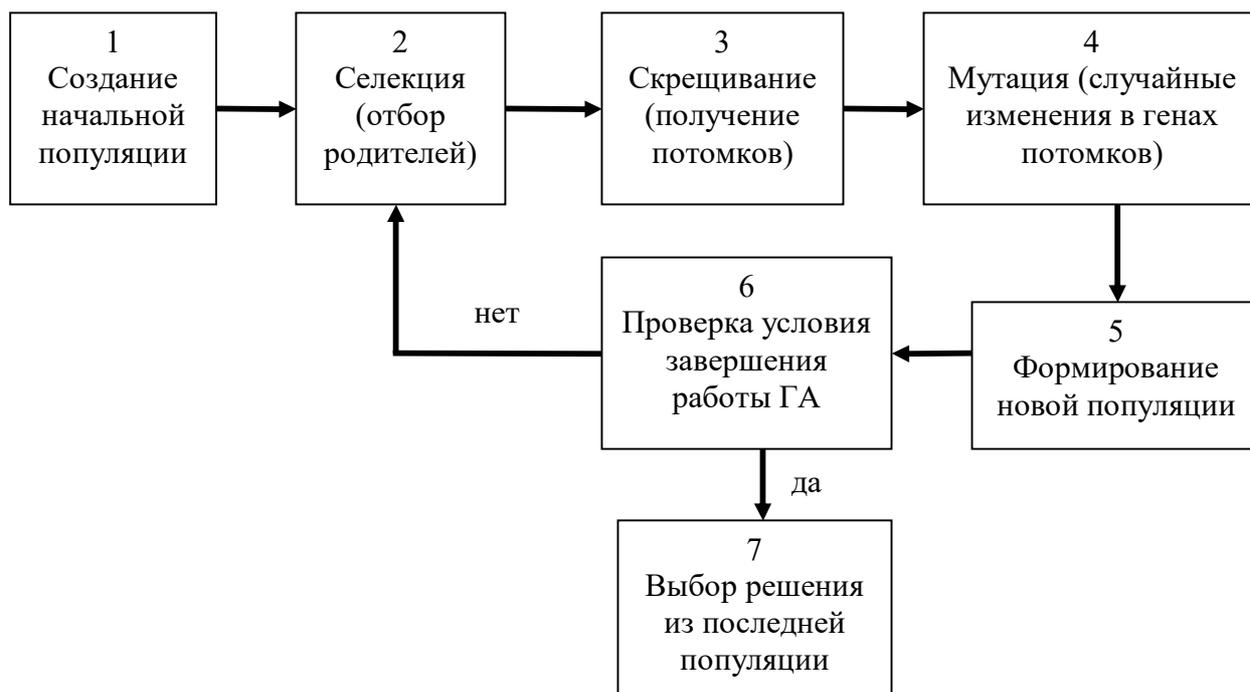


Рисунок 10.2 – Общая схема работы генетического алгоритма

**Генетические алгоритмы с вещественным кодированием.** Рассматриваемая в задаче (10.1) целевая функция  $f(x)$  эквивалентна природному понятию приспособленности живого организма. Вектор переменных  $x = (x_1, x_2, \dots, x_n)^T$  целевой функции называется фенотипом, а отдельные его координаты  $x_i \in R$  – признаками,  $i = 1, \dots, n$ . Любой вектор  $x \in R^n$  считается хромосомой (генотипом, особью), а каждая его координата  $x_i \in R$  представляет собой ген. Фенотип совпадает с генотипом, а поиск фактически производится в пространстве фенотипов. На множестве допустимых решений  $D$  вводится так называемая функция приспособленности  $\mu(x) : D \xrightarrow{\mu} R$ , где  $R$  – множество вещественных чисел. Эта функция аналогична целевой функции  $f(x)$ . Функцией  $\mu(s)$  может быть любая функция, удовлетворяющая следующему условию:

$$\forall x^1, x^2 \in D : s^1 = e(x^1), s^2 = e(x^2), x^1 \neq x^2,$$

$$\text{если } f(x^1) > f(x^2), \text{ то } \mu(x^1) > \mu(x^2).$$

Решение исходной оптимизационной задачи  $f(x^*) = \min_{x \in D} (f(x))$

сводится к поиску решения  $x_\mu^*$  другой оптимизационной задачи:

$$\mu(x_\mu^*) = \min_{x \in D} (\mu(x)). \quad (10.6)$$

В силу выбора функции  $\mu(x)$  решения задач (10.1) и (10.6) (хромосома) совпадают:

$$x^* = \arg \min_{x \in D} f(x) = \arg \min_{x \in D} \mu(x) = x_\mu^*. \quad (10.7)$$

При решении задачи (10.6) используются конечные наборы  $I = \{x^k = (x_1^k, x_2^k, \dots, x_n^k), k = 1, 2, \dots, K\} \subset D$  возможных решений, называемых популяциями, где  $x^k$  – хромосома с номером  $k$ ;  $K$  – размер популяции;  $x_i^k$  – ген с номером  $i$ .

### 10.3.2. Методы иммунных систем

**Метод искусственных иммунных систем.** *Иммунной системой* живого организма называется подсистема, объединяющая органы и ткани, которые защищают организм от заболеваний. Назначение иммунной системы заключается в том, что она идентифицирует и уничтожает чужеродные тела, попавшие в организм, и совершенствуется, накапливая опыт борьбы с ними.

*Антигеном* называется вещество, которое воспринимается живым организмом как чужеродное и от которого организм пытается защититься. Для того чтобы организм смог защититься от антигенов, в нем при помощи специальных иммунных клеток вырабатываются антитела.

*Антителом* называется вещество, которое распознает антиген и способствует его уничтожению. Если иммунные клетки выработали антитела, которые смогли распознать антиген, то информация об этих антителах сохраняется в клетках памяти.

*Клеткой памяти* называется иммунная клетка, которая сохраняет в себе информацию о новых антителах, способных распознать антиген для того, чтобы в следующий раз, когда в организм попадет такой же

или похожий антиген, иммунная система смогла работать эффективнее.

Целевая функция  $f(x)$  метода ИИС эквивалентна природному понятию приспособленности иммунной клетки к борьбе с антигенами, т.е. способности клетки вырабатывать антитела. Поэтому будем называть целевую функцию  $f(x)$  функцией приспособленности. Вектор параметров  $x = (x_1, x_2, \dots, x_n)^T \in D$  целевой функции называется иммунной клеткой, которая вырабатывает антитела.

При решении задачи глобальной минимизации используются конечные наборы  $I = \{x^k = (x_1^k, x_2^k, \dots, x_n^k), k = 1, 2, \dots, K\} \subset D$  возможных решений, называемых популяциями, где  $x^k$  – иммунная клетка с номером  $k$ ;  $K$  – размер популяции. Чем меньше значение целевой функции  $f(x^k)$ , тем более иммунная клетка  $x^k$  приспособлена, т.е. способна вырабатывать антитела и подходит в качестве решения. Общая схема метода искусственных иммунных систем приведена на рисунке 10.3.

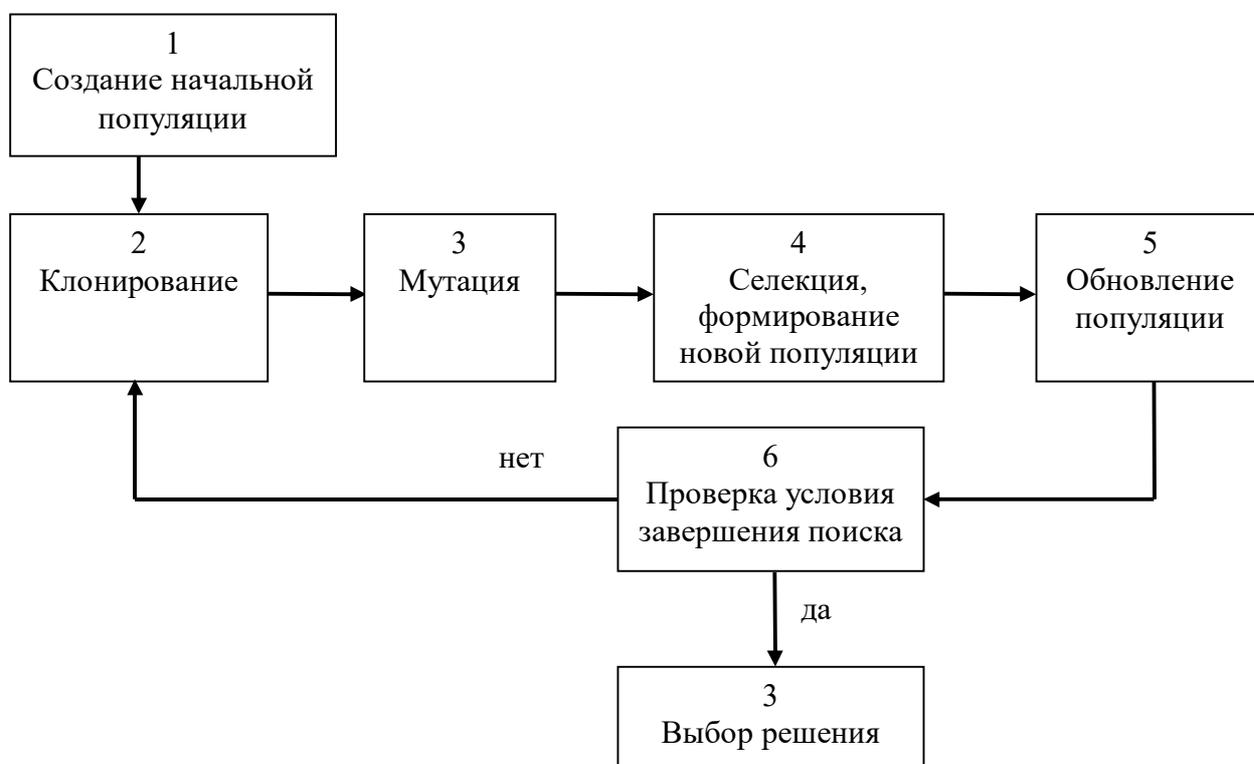


Рисунок 10.3 – Общая схема метода искусственных иммунных систем

**Расширенный метод искусственных иммунных систем.** Особенностью расширенного метода ИИС является сочетание локального и глобального поисков решения.

Общий алгоритм расширенного метода ИИС приведен на рисунке 10.4.

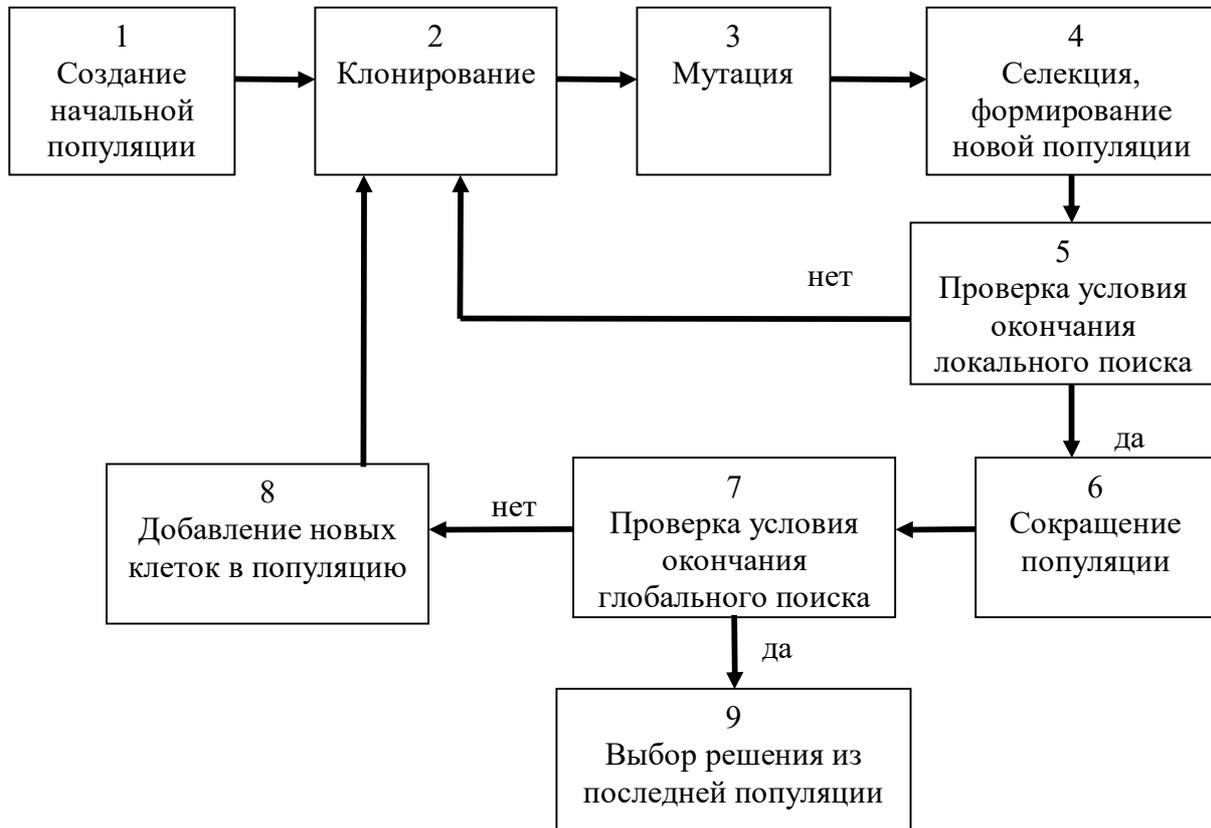


Рисунок 10.4 – Общая схема работы расширенного метода искусственных иммунных систем

*Локальный* поиск реализуется в виде циклического итерационного процесса, во время которого к популяции применяются биологические операторы: клонирование, мутация и селекция. В результате происходит смена популяции на новую, к которой, если условие окончания локального поиска не выполнено, опять применяются биологические операторы, и т.д. до выполнения условия окончания локального поиска. Средняя приспособленность популяции при этом растет. Если условие окончания локального поиска выполнено, то начинается новая итерация глобального поиска.

*Глобальный* поиск реализуется в виде циклического итерационного процесса. Каждая итерация глобального поиска представляет

собой локальный поиск. В конце каждой итерации происходит сокращение популяции с использованием идей кластеризации. Оставшиеся иммунные клетки помечаются как клетки памяти, после чего проверяются условия окончания работы метода. Если условия окончания не выполнены, то к популяции присоединяются новые особи, количество которых пропорционально размеру популяции. Таким образом, размер популяции  $N$  в расширенном методе ИИС является переменной величиной. Если условие окончания работы метода выполнено, то в качестве приближенного решения задачи из последней популяции выбираются иммунные клетки, которым соответствует наибольшее значение функции приспособленности.

### 10.3.3. Метод рассеивания

В методе рассеивания рассматриваемая целевая функция  $f(x)$  называется функцией приспособленности, а вектор параметров  $x = (x_1, x_2, \dots, x_n)^T$  целевой функции – особью. Чем меньше значение целевой функции  $f(x)$ , тем более особь  $x$  приспособлена, т.е. подходит в качестве решения.

Свое название метод берет от способа формирования начального набора возможных решений, называемого базовым множеством особей. Генерация базового множества построена таким образом, чтобы особи в нем были достаточно различны (рассеяны) между собой. Это позволяет эффективно исследовать все множество допустимых решений поставленной задачи.

Процедура поиска начинается с генерации базового множества особей  $A$ . Для этого сначала на отрезке  $[a_i, b_i]$  изменения каждой координаты выделяются  $s$  подынтервалов одинаковой длины. В базовое множество последовательно добавляется по одной особи:

1) для каждого отрезка  $[a_i, b_i]$  генерируется номер  $j_i$  подынтервала с вероятностью, обратно пропорциональной числу раз, которое этот подынтервал уже выбирался;

2) генерируются значения координат  $x_i$ ,  $i = 1, \dots, n$ , случайным образом внутри выбранных подынтервалов;

3) если минимальное расстояние между сформированной особью и особями базового множества больше заданной величины  $\sigma$ , то особь добавляется в базовое множество  $A$ . Иначе – особь не добавляется, процесс рассеивания продолжается с пункта 1.

Процесс формирования базового множества гарантирует разнообразие входящих в него особей и продолжается до тех пор, пока не будет выбрано нужное количество особей.

При решении задачи используются конечные наборы возможных решений  $I = \{x^k = \{x_1^k, x_2^k, \dots, x_n^k\}, k = 1, 2, \dots, Np\} \subset D$ , называемые популяциями, где  $x^k$  – особь с номером  $k$ ;  $Np$  – размер популяции.

Идеи рассеивания применяются также при формировании начальной популяции из базового множества особей  $A$ :  $Np = b_1 + b_2$ , где  $b_1$  – количество особей, выбираемых по качеству (наилучшие особи по величине функции приспособленности);  $b_2$  – количество особей, выбираемых по расстоянию, т.е. суммарное расстояние от них до уже имеющихся в начальной популяции особей должно быть максимально.

Метод рассеивания имитирует эволюцию начальной популяции  $I_0 = \{x^j, j = 1, 2, \dots, Np \mid x^j = (x_1^j, \dots, x_n^j) \in D\}$  и представляет собой итерационный процесс, исследующий множество  $D$ . В процессе работы метода на каждой итерации к популяции применяются биологические операторы: селекция и скрещивание, после чего происходит замена особей с низким уровнем приспособленности на новые. Так формируется новая популяция. Метод заканчивает работу после того, как будет исчерпано заданное количество вычислений значения функции приспособленности. В качестве приближенного решения задачи из последней популяции выбираются особи, которым соответствует наименьшее значение целевой функции. Следует отметить, что размер базового множества особей  $A$  должен быть достаточно большим, чтобы обеспечить работу метода до выполнения условий окончания.

Метод рассеивания сочетает в себе локальный и глобальный поиски решения.

*Глобальный* поиск реализуется в виде циклического итерационного процесса. Каждая итерация глобального поиска представляет

собой локальный поиск. В конце каждой итерации происходит сокращение популяции: удаляется  $b_2$  особей с наихудшей приспособленностью, после чего проверяются условия окончания работы метода. Если условия окончания не выполнены, то к популяции присоединяется  $b_2$  новых особей из базового множества таким же образом, как и при формировании начальной популяции, после чего начинается новая итерация глобального поиска. Если условие окончания работы метода выполнено, то в качестве приближенного решения задачи из последней популяции выбираются особи, которым соответствует наименьшее значение целевой функции.

*Локальный* поиск реализуется в виде циклического итерационного процесса, во время которого к популяции применяются биологические операторы селекция и скрещивание. Во время селекции из особей текущей популяции составляются всевозможные родительские пары, каждая из которых при скрещивании порождает новое решение, называемое потомком. Все потомки помещаются во множество потомков, после чего происходит их сравнение с особями из текущей популяции и принимается решение о замене особей потомками. В результате происходит смена популяции на новую, к которой, если условия окончания локального поиска не выполнены, опять применяются биологические операторы, и т.д. до выполнения условия окончания локального поиска. Средняя приспособленность популяции при этом растет. Если условие окончания локального поиска выполнено, то начинается новая итерация глобального поиска.

На рисунке 10.5 приведен общий алгоритм метода рассеивания.

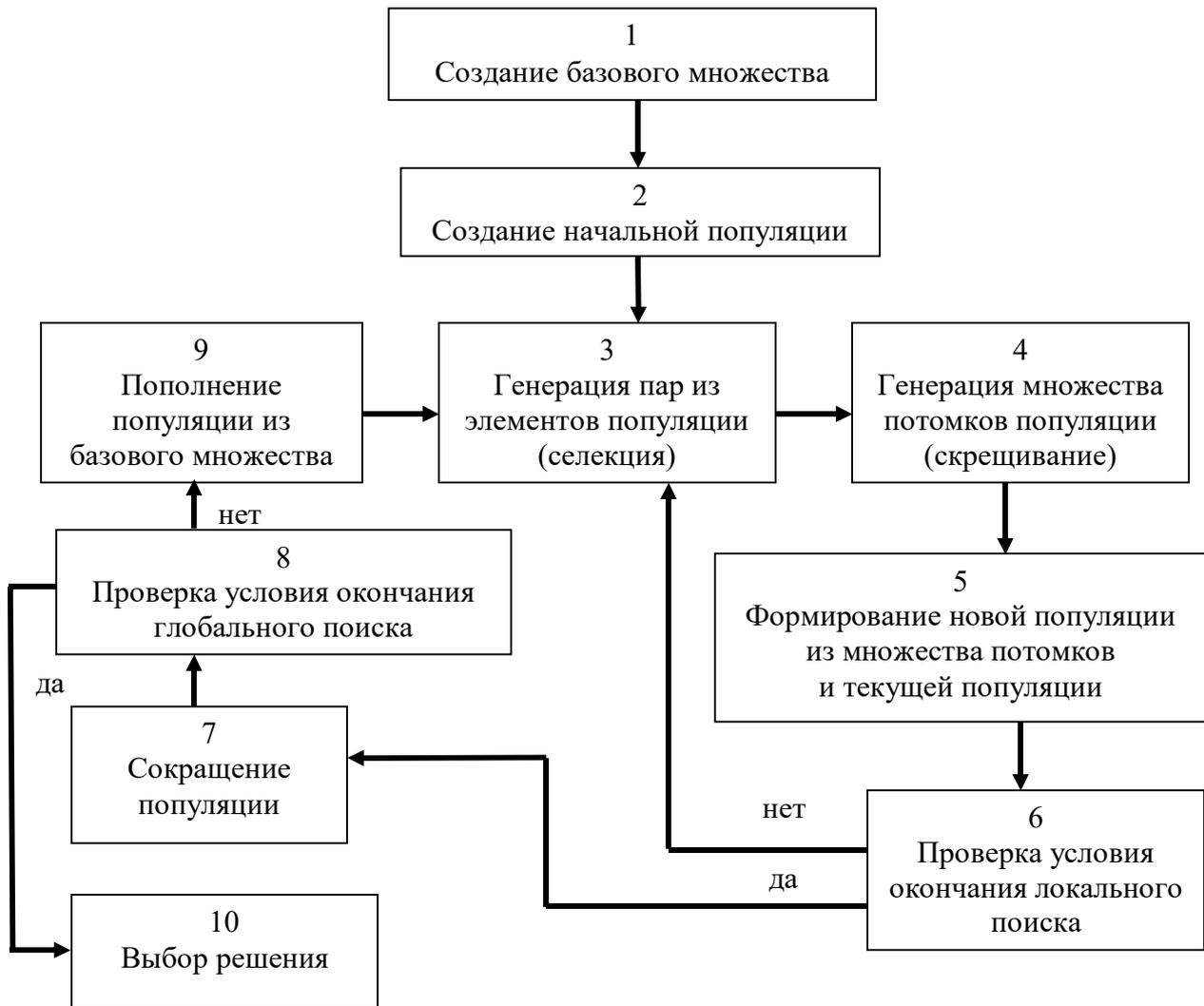


Рисунок 10.5 – Общий алгоритм метода рассеивания

### 10.3.4. Эволюционная стратегия преобразования ковариационной матрицы

Подобно другим эволюционным методам, в эволюционной стратегии на каждой итерации образуется новая популяция. Генерация новых особей, так называемая *мутация*, происходит случайным образом согласно формируемому эволюционной стратегией распределению. Лучшие найденные решения участвуют в дальнейшем эволюционном процессе. Для генерации новых особей (векторов) в эволюционных стратегиях применяется нормальное распределение, параметры которого влияют на эффективность работы реализующего стратегию алгоритма. Особенностью эволюционной стратегии является изменение

параметров распределения на каждой итерации в соответствии с успешностью поиска. Стратегия преобразует текущие параметры распределения таким образом, чтобы они как можно лучше подходили для успешной работы на последующих шагах.

В эволюционной стратегии преобразования ковариационной матрицы целевая функция  $f(x)$  называется функцией приспособленности, а вектор параметров  $x = (x_1, x_2, \dots, x_n)^T$  целевой функции – особью. Чем меньше значение целевой функции  $f(x)$ , тем более особь  $x$  приспособлена, т.е. подходит в качестве решения. Каждый вектор  $x \in D$  является возможным решением поставленной оптимизационной задачи.

При решении задачи глобальной оптимизации используются конечные наборы  $I = \{x^k = (x_1^k, x_2^k, \dots, x_n^k), k = 1, 2, \dots, Np\} \subset D$  возможных решений, называемые популяциями, где  $x^k$  – особь с номером  $k$ ;  $Np$  – размер популяции. Применение эволюционной стратегии сводится к исследованию множества  $D$  при помощи перехода от одной популяции к другой. Чем меньше значение целевой функции  $f(x^k)$ , тем лучше особь  $x^k$  приспособлена.

Рассматриваемая стратегия имитирует эволюцию начальной популяции  $I_0 = \{x^j, j = 1, 2, \dots, Np \mid x^j = (x_1^j, \dots, x_n^j) \in D\}$  и представляет собой итерационный процесс. На каждой итерации генерируется новая популяция согласно сформированному стратегией распределению. После этого производится анализ новой популяции и корректировка параметров распределения: шага  $\mu$  и ковариационной матрицы  $C$ , затем осуществляется переход на новую итерацию. Процедура поиска завершается, когда сформируется заданное количество популяций. В качестве приближенного решения задачи из последней популяции выбирается особь, которой соответствует наименьшее значение функции приспособленности.

Обобщенный алгоритм метода представлен на рисунке 10.6.

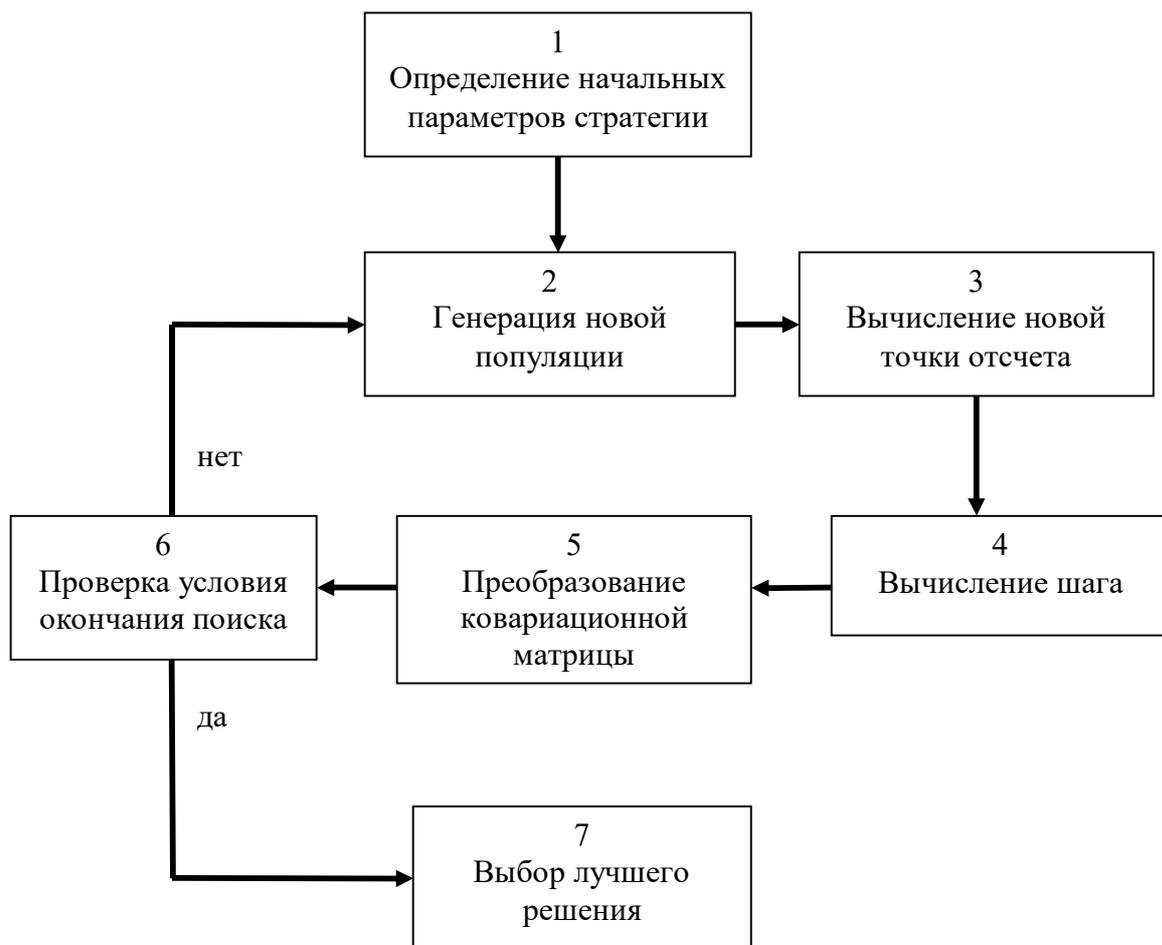


Рисунок 10.6 – Общий алгоритм эволюционной стратегии преобразования ковариационной матрицы

### 10.3.5. Метод динамических сеток

При работе метода динамических сеток происходит эволюция начальной популяции – смена одного поколения другим путем расширения и последующего сокращения популяции. Популяция представляется в виде некоторой сетки, состоящей из набора решений, называемых узлами. В процессе поиска сетка подвергается изменениям: расширению – добавлению новых узлов в сетку, и сокращению – удалению узлов, расположенных слишком близко друг к другу. В методе динамических сеток рассматривается целевая функция  $f(x)$ . Каждому узлу ставится в соответствие вектор параметров  $x = (x_1, x_2, \dots, x_n)^T \in D$  целевой функции. Каждый вектор  $x$  является возможным решением поставленной оптимизационной задачи. При

решении задачи используются конечные наборы  $I = \{x^j = (x_1^j, \dots, x_n^j), j = 1, 2, \dots, Np\} \subset D$  возможных решений, называемые популяциями или сеткой, где  $x^j$  – узел с номером  $j$ ;  $Np$  – количество узлов в сетке. Применение метода динамических сеток сводится к исследованию множества  $D$  при помощи изменения сетки и перехода от одной популяции к другой. Чем меньше значение целевой функции  $f(x^j)$ , тем более узел  $x^j$  подходит в качестве решения.

Метод динамических сеток имитирует эволюцию начальной популяции  $I_0 = \{x^j, j = 1, 2, \dots, Np \mid x^j = (x_1^j, \dots, x_n^j) \in D\}$  и представляет собой итерационный процесс. В процессе работы метода на каждой итерации происходит расширение (локальное, глобальное и дополнительное) и последующее сокращение сетки. Так формируется новая сетка. Критерием окончания поиска является достижение заранее заданного количества  $n$  вычислений целевой функции. В качестве приближенного решения задачи из последней популяции выбирается узел, которому соответствует наименьшее значение целевой функции.

В процессе расширения происходит добавление новых узлов в сетку. Стадия расширения состоит из нескольких этапов:

- 1) локального расширения;
- 2) глобального расширения;
- 3) дополнительного расширения.

Общий алгоритм динамических сеток представлен на рисунке 10.7.

На этапе *локального* расширения в окрестности каждого  $p$ -го узла,  $p = 1, 2, \dots, P$ , выбирается некоторое заранее заданное число  $K$  ближайших по расстоянию узлов, называемых соседними узлами. Среди соседних узлов выбирается наилучший. Если данный соседний узел лучше  $p$ -го узла, то производится генерация нового узла в направлении наилучшего из  $K$  соседних узлов.

На этапе *глобального* расширения для всех узлов сетки, кроме наилучшего узла  $x^{\text{best}}$ , производится генерация нового узла в направлении наилучшего узла сетки.

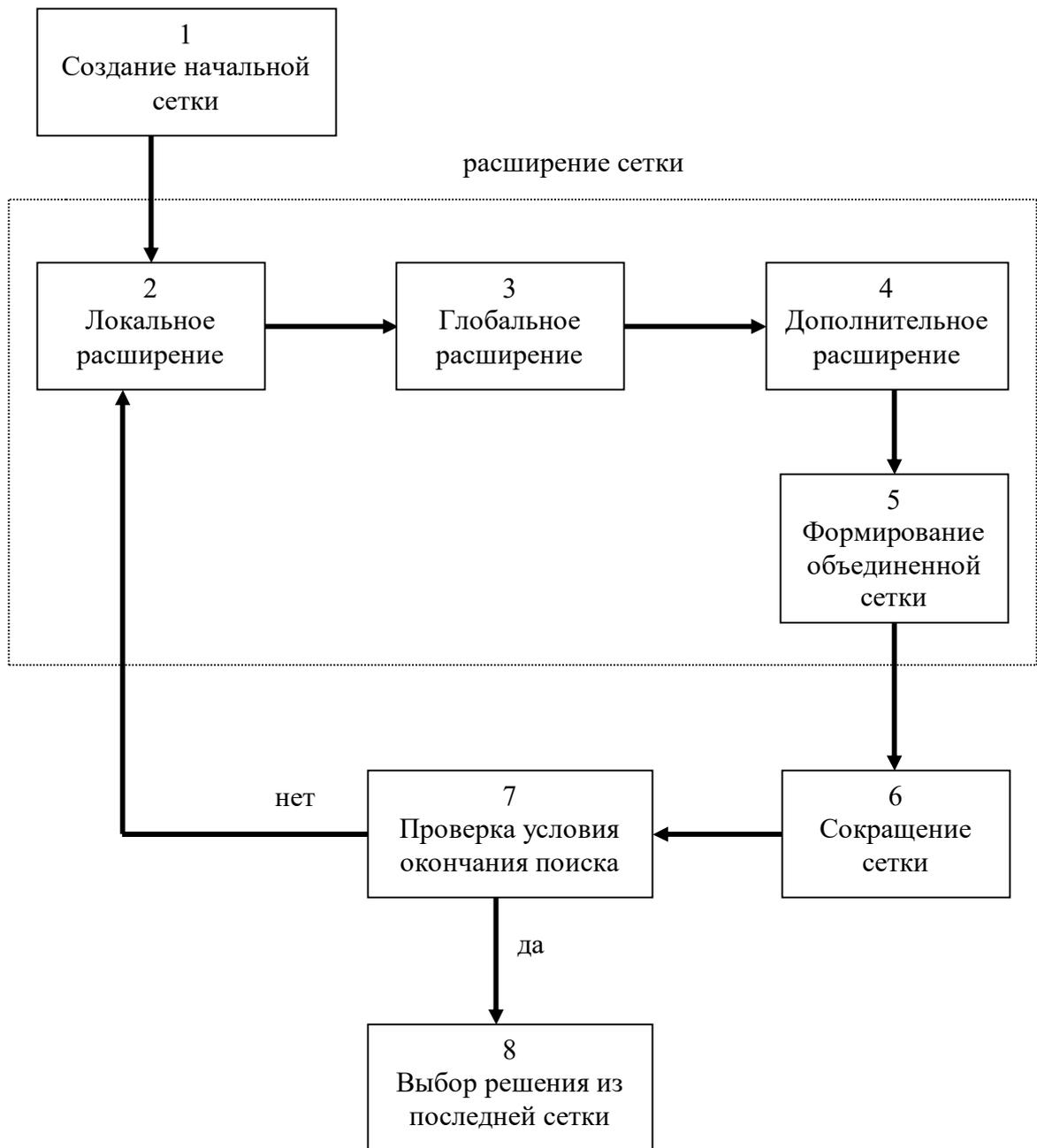


Рисунок 10.7 – Алгоритм метода динамических сеток

Если при локальном и глобальном расширении сгенерировано узлов меньше, чем заранее заданное число  $N$ , то выполняется *дополнительное* расширение сетки. Путем генерации новых узлов производится исследование новых участков множества допустимых решений. На последующей за расширением стадии сокращения в методе динамических сеток происходит удаление слишком близких друг к другу решений. Таким образом, стратегия метода направлена на поддержание достаточного разнообразия узлов в сетке.

## 10.4. Генетические алгоритмы оптимизации

### 10.4.1. Основные понятия генетических алгоритмов

Идея использования принципов биологической эволюции для решения оптимизационных задач возникала в различных вариантах у ряда авторов. Первые публикации на эту тему появились в 1960-х годах, а уже в 1975 г. вышла книга Холланда, в которой был предложен первый генетический алгоритм.

В настоящее время с помощью генетических алгоритмов решено множество различных задач. Генетические алгоритмы нашли широкое практическое применение в менеджменте и управлении для решения задач поиска оптимальных решений, формирования моделей и прогнозирования значений различных показателей, поиска глобального минимума многоэкстремальных функций.

*Генетический алгоритм* (англ. genetic algorithm) – это эвристический алгоритм поиска, используемый для решения задач оптимизации и моделирования путем последовательного подбора, комбинирования и вариации искомых параметров с использованием механизмов, напоминающих биологическую эволюцию. Он является разновидностью эволюционных вычислений (англ. evolutionary computation). Другими словами, генетический алгоритм – это метод перебора решений для тех задач, в которых невозможно найти решение с помощью математических формул. Однако простой перебор решений в сложной многомерной задаче – это бесконечно долго. Поэтому в генетическом алгоритме перебираются не все решения, а только лучшие. Из группы решений ищутся наиболее подходящие. Затем они немного изменяются, получаются новые решения, среди которых снова отбираются лучшие, а худшие отбрасываются. Таким образом, на каждом шаге работы алгоритма отбираются наиболее подходящие решения (проводится селекция), которые на следующем шаге дадут еще более лучшие решения (рисунок 10.8).

Для генетических алгоритмов характерны все черты, присущие классу эвристических методов.

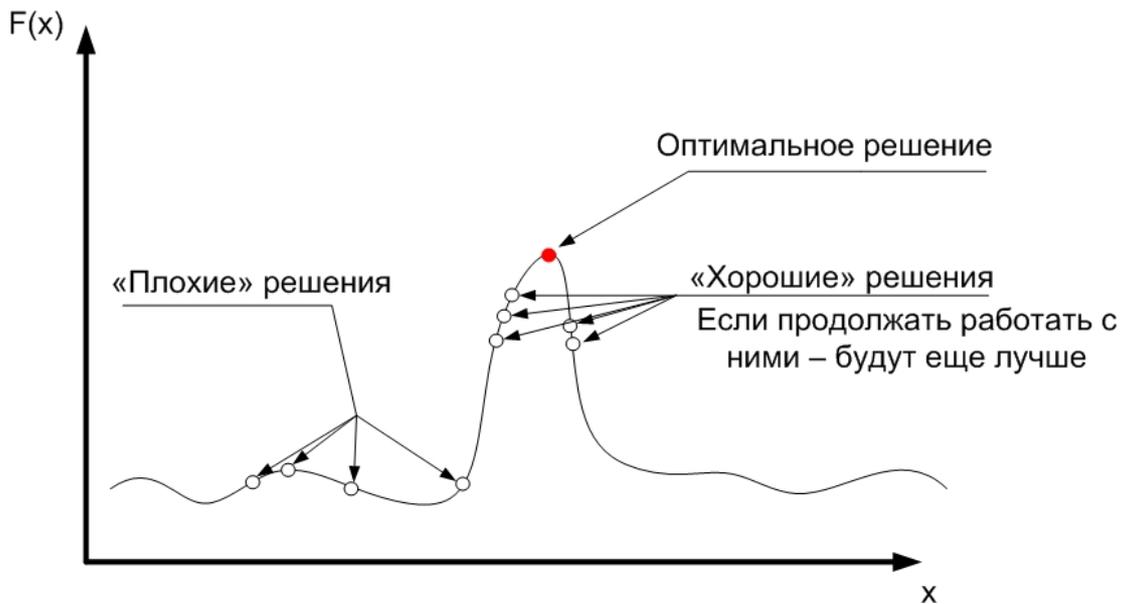


Рисунок 10.8 – Выбор наилучших решений

Как правило, ГА состоит из четырех основных операторов: селекции, скрещивания (кроссинговера, репродукции), мутации, создания нового поколения. Цикл скрещивания и мутации с последующей оценкой приспособленности называется *поколением*. Поэтапно алгоритм процесса формирования нового поколения можно представить следующим образом.

Шаг 1. Создать начальную популяцию из  $K$  хромосом (особей). Для этого случайным образом генерируется конечный набор пробных решений  $X^k = (x_1^k, x_2^k, \dots, x_n^k)$ ,  $x_i^k \in X$  (первое поколение),  $k = 1, 2, \dots, K$ ,  $K$  – размер популяции.

Шаг 2. Оценить степень приспособленности каждой особи, т.е. рассчитать значения целевой функции  $f_k(x) = f_k(x_1^k, x_2^k, \dots, x_n^k)$ .

Шаг 3. Выход, если выполняется критерий останова, иначе переход на шаг 4.

Шаг 4. Выбрать  $K$  родителей из популяции при помощи метода селекции (вероятность выбора родителя должна зависеть от степени его приспособленности).

Шаг 5. Выбрать из родительского пула пару родителей для репродукции. При помощи оператора скрещивания с вероятностью  $p_c$  получить потомка.

Шаг 6. Подвергнуть потомков мутации с вероятностью  $p_m$ .

Шаг 7. Повторять шаги 5–6, пока не будет сгенерировано новое поколение популяции, содержащее  $K$  хромосом.

Шаг 8. Оценить степень приспособленности каждой особи в новой популяции.

Шаг 9. Перейти на шаг 4, если количество поколений не превышает допустимого.

Все операторы могут изменяться от задачи к задаче, поэтому существует множество модификаций ГА, применимых к разным условиям. Перед началом работы алгоритма определяются вероятности скрещивания и мутации. Вероятность скрещивания рекомендуется выбирать равной 80–90 %, а вероятность мутации – 1–3 %. Эти числа могут варьироваться в зависимости от решаемой задачи.

Основные отличия ГА от других методов оптимизации:

- алгоритм начинается с некоторого набора начальных приближений, называемых начальной популяцией, из-за чего у него меньше шансов остаться в ловушке локального экстремума;
- используется только целевая функция;
- применяются вероятностные методы.

Из-за того что ГА основаны на принципе «выживает сильнейший», естественнее всего максимизировать целевую функцию, называемую функцией приспособленности. Поэтому ГА подходят для решения задач безусловной максимизации. Функция приспособленности  $F(x)$  может совпадать с целевой функцией  $f(x)$  безусловной задачи максимизации, то есть  $F(x) = f(x)$ . Задача минимизации может быть легко трансформирована в задачу максимизации, обычно выбирается положительная функция приспособленности. Часто для перехода от задачи минимизации к задаче максимизации используется формула

$$F(x) = \frac{1}{1 + f(x)}. \quad (10.8)$$

Благодаря своей адаптивной природе и возможности определять операторы в зависимости от области применения, ГА используются для решения разнообразных задач:

- поиска глобального экстремума многомерной функции;
- аппроксимации функций;
- создания искусственного интеллекта;
- поиска кратчайшего пути;
- составления расписаний.

*Ген* – атомарный элемент хромосомы (может быть битом, числом или неким другим объектом).

*Аллель* – значение конкретного гена.

*Локус* – положение конкретного гена в хромосоме.

*Хромосома (цепочка)* – упорядоченная последовательность генов, строка из каких-либо чисел. Если эта строка представлена бинарной строкой из нулей и единиц, например 101010, то она получена с использованием либо двоичного кодирования, либо кода Грея. Каждая позиция хромосомы называется геном.

*Генотип (код)* – упорядоченная последовательность хромосом (одно из решений).

*Фенотип* – набор значений, соответствующих генотипу.

*Особь (индивидуум)* – конкретный экземпляр генотипа (вариант решения задачи). Особи представляются хромосомами с закодированными в них множествами параметров задачи, т.е. решений. Обычно особь состоит из одной хромосомы, поэтому в дальнейшем особь и хромосома – идентичные понятия. Генетические алгоритмы оперируют популяциями, т.е. конечным числом особей.

*Популяция* – набор особей (набор решений задачи). В начале алгоритма случайным образом генерируется начальная популяция (набор решений). Эти решения будут становиться лучше (эволюционировать) в процессе работы алгоритма до тех пор, пока не удовлетворят условиям задачи.

*Скрещивание (кроссинговер)* – операция, при которой две хромосомы обмениваются своими частями. Например, 1100&1010 → 1110&1000. Здесь произошел обмен вторым геном (вторым младшим разрядом).

*Мутация* – случайное изменение одной или нескольких позиций в хромосоме. Например, 1010011 → 1010001.

*Инверсия* – изменение порядка следования битов в хромосоме или в ее фрагменте. Например, 1100 → 0011.

*Рекомбинация* – операция, при которой две хромосомы обмениваются своими частями.

*Функция приспособленности* или *функция пригодности* – одно из центральных понятий ГА. Эта функция оценивает, насколько приспособлена данная особь в популяции, другими словами, она определяет качество особей популяции.

*Пример.* Допустим, роботу необходимо объехать шесть контрольных точек за наименьшее время. Расстояния от каждой точки до каждой заданы в виде матрицы (рисунок 10.9).

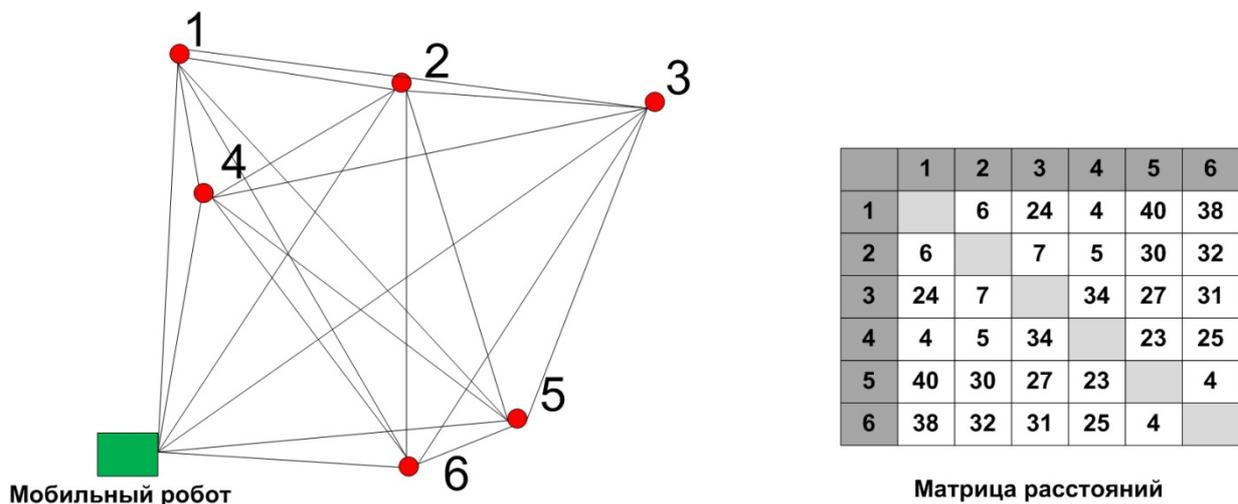


Рисунок 10.9 – Задача о поиске кратчайшего пути

Возьмем несколько возможных решений (особей) – это и есть *начальная популяция* (рисунок 10.10). В нашем примере это будет сумма расстояний от точки до точки в выбранном маршруте. Рассчитаем функции пригодности. Для первой особи  $f_1(x) = 6 + 7 + 34 + 23 + 4 = 74$ . Для остальных особей таким же образом получаем  $f_2(x) = 111$ ,  $f_3(x) = 47$ ,  $f_4(x) = 125$ . Таким образом, особь 3 – лучшая, а 4 – самая плохая.

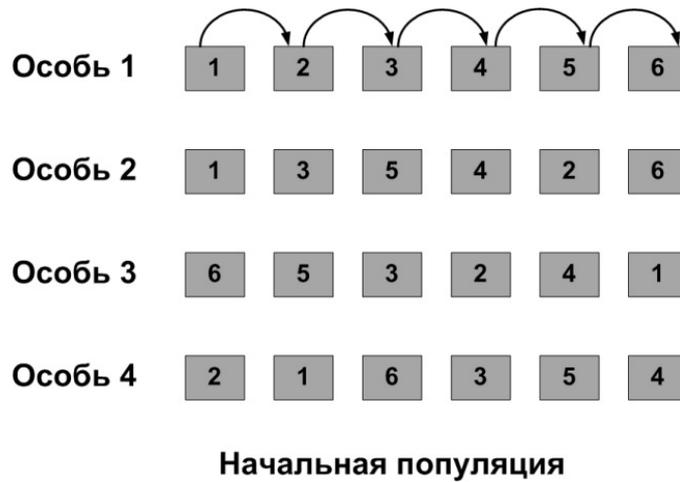


Рисунок 10.10 – Начальная популяция (начальные пробные решения)

Функция приспособленности всегда зависит от задачи и в ее роли часто выступает целевая функция. Иногда оценку приспособленности проводят в две стадии. Первая стадия – это собственно оценка  $f_k(x) = f_k(x_1^k, x_2^k, \dots, x_n^k)$ . Вторая – дополнительные преобразования. Например, ею может быть нормировка к виду

$$F_k = \frac{f_k - f_k^0}{f_k^1 - f_k^0}, \quad (10.9)$$

где  $f_k^1$ ,  $f_k^0$  – лучший и худший показатели в текущей популяции.

### 10.4.2. Кодирование параметров задачи

Успех в решении задачи с помощью ГА напрямую зависит от способа кодирования ее параметров. Кодирование определяет, каким образом будут выглядеть хромосомы. Хромосомы могут представлять собой битовые строки, вещественные числа, перестановки элементов, список правил и практически любую структуру данных. В классическом ГА, представленном Холландом, производится бинарное кодирование параметров. Например, целые числа из интервала от 0 до 31 можно представить последовательностями нулей и единиц, используя двоичную систему счисления. Число 0 при этом записывается как 00000, а число 31 – как 11111. В данном случае хромосомы

приобретают вид двоичных последовательностей, состоящих из 5 битов, т.е. цепочек длиной 5.

При бинарном кодировании операции мутации и скрещивания достаточно просты: мутация происходит инвертированием бита, а скрещивание – обменом частями битовых строк. Если представлять целые числа двоичным кодом, то многие соседние числа, например 15 и 16 (в двоичном представлении 01111 и 10000), будут различаться между собой несколькими разрядами, поэтому за одну операцию мутации из 15 нельзя получить 16 и наоборот. Эта проблема решается при помощи кода Грея, благодаря которому соседние числа, закодированные в бинарную строку, различаются только на один разряд.

Итак, будем считать, что каждая переменная  $x_i$  кодируется определенным фрагментом хромосомы, состоящим из фиксированного количества генов (рисунок 10.11). Все локусы хромосом диаллельны – то есть в любой позиции фрагмента может стоять как ноль, так и единица. Рядом стоящие фрагменты не отделяют друг от друга какими-либо маркерами, тем не менее при декодировании хромосомы в вектор переменных на протяжении всего моделируемого периода эволюции используется одна и та же маска картирования.

$x_1$				$x_2$				$x_3$				$x_4$				$x_5$				$x_6$			
1	1	0	0	0	1	0	1	1	0	1	0	1	0	0	1	0	0	0	0	1	1	1	1

Рисунок 10.11 – Простейшая маска картирования хромосомы, определяющая план распределения наследственной информации по длине хромосомы

Хромосомы генерируются случайным образом путем последовательного заполнения разрядов (генов) сразу в бинарном виде. Таблица 10.1 воспроизводит в полном объеме процедуру декодирования фрагмента 4-битовой хромосомы в проекцию вектора переменных  $x_i \in [a_i, b_i]$ .

Аналогичную таблицу можно составить для 5-битовых хромосом. В этом случае в четвертом столбце преобразование двоичной

последовательности в вещественные числа будет выполняться по формуле  $a + i \cdot (b - a) / 31, i = 1, \dots, 31$ .

Таблица 10.1 – Декодирование фрагментов хромосом в проекции вектора переменных

Десятичное значение сдвига	Двоично-десятичный код	Код Грея	Вещественное значение координаты
0	0000	0000	$a$
1	0001	0001	$a + 1(b - a) / 15$
2	0010	0011	$a + 2(b - a) / 15$
3	0011	0010	$a + 3(b - a) / 15$
4	0100	0110	$a + 4(b - a) / 15$
5	0101	0111	$a + 5(b - a) / 15$
6	0110	0101	$a + 6(b - a) / 15$
7	0111	0100	$a + 7(b - a) / 15$
8	1000	1100	$a + 8(b - a) / 15$
9	1001	1101	$a + 9(b - a) / 15$
10	1010	1111	$a + 10(b - a) / 15$
11	1011	1110	$a + 11(b - a) / 15$
12	1100	1010	$a + 12(b - a) / 15$
13	1101	1011	$a + 13(b - a) / 15$
14	1110	1001	$a + 14(b - a) / 15$
15	1111	1000	$b$

В общем случае число интервалов будет равно  $2^L - 1$ , где  $L$  – число битов (при  $L = 4$  получим  $2^4 - 1 = 15$ , при  $L = 5$  получим  $2^5 - 1 = 31$  и т.д.).

Строка  $c = (c_1, \dots, c_L)$ ,  $c_j \in \{0, 1\}$ ,  $j = 1, \dots, L$ , закодированная с помощью двоичного кода, представляет целое число  $v \in \{0, \dots, 2^L - 1\}$ :

$$v = \sum_{j=1}^L c_j 2^{(L-j)}. \quad (10.10)$$

Точки разбиения интервала  $[a, b]$  при двоичном кодировании определяются по формуле

$$a + i \cdot \frac{b-a}{2^{L-1}-1}, i = 1, \dots, 2^{L-1} - 1. \quad (10.11)$$

Преобразование строки  $c = (c_1, \dots, c_L)$ ,  $c \in \{0, 1\}$ , закодированной двоичным кодом, в строку Грея  $g = (g_1, \dots, g_L)$ , осуществляется по формуле

$$g_k = \begin{cases} c_1, & \text{если } k = 1, \\ c_{k-1} \oplus c_k, & \text{если } k > 1, \end{cases} \quad (10.12)$$

где  $\oplus$  означает сложение по модулю 2. В случае бинарного и тернарного сложения по модулю 2 результаты приведены в таблице 10.2.

Таблица 10.2 – Сложение по модулю 2

Бинарное сложение		
$x$	$y$	$a \oplus b$
0	0	0
1	0	1
0	1	1
1	1	0

Тернарное сложение			
$x$	$y$	$z$	$a \oplus b$
0	0	0	0
1	0	0	1
0	1	0	1
1	1	0	0
0	0	1	1
1	0	1	0
0	1	1	0
1	1	1	1

Общее правило сложения по модулю 2: результат равен 0, если нет операндов, равных 1, либо число операндов, равных 1, равно четному количеству.

От кода Грея переходим к двоично-десятичному коду с помощью обратного преобразования по формуле

$$c_k = \sum_{i=1}^k (g_i | \otimes), \quad (10.13)$$

где суммирование выполняется по модулю 2. Далее переходим к натуральным целым числам. Отношение полученного числа к максимальному числу, доступному для кодирования данным количеством разрядов фрагмента (в таблице 10.1 число 15), дает искомое значение сдвига переменной относительно левой границы  $a$  допустимого диапазона ее изменения  $(b - a)$ .

Из таблицы 10.1 хорошо видно, почему код Грея имеет явные преимущества по сравнению с двоично-десятичным кодом, который при некотором стечении обстоятельств порождает своеобразные тупики для поискового процесса. В качестве примера рассмотрим любые три рядом стоящие строки из таблицы 10.1, например кодирующие сдвиг в 4, 5 и 6 единиц.

Предположим, фрагменты хромосом, стоящие в пятой строке и кодирующие число 5, принадлежат оптимальному вектору, являющемуся решением некоторой задачи, а лучшая особь из текущей популяции содержит фрагмент хромосомы из строки 4. Такая ситуация благоприятна для обоих кодов. Достаточно выполнить всего одну операцию – заменить в четвертом разряде фрагмента 0 на 1 – и решение будет найдено. Более интересный случай получается, если лучшая особь содержит фрагмент из строки 6. Для кода Грея эта ситуация ничуть не сложнее предыдущей: замена 0 на 1 в третьем разряде опять приведет к успеху. В то же время двоично-десятичный код вынуждает выполнить последовательно две операции – заменить 1 на 0 в третьем разряде и 0 на 1 в четвертом. С какой бы из них мы ни начали, результат не приблизит нас к решению: первый вариант замены переместит нас в четвертую строку, а второй – в седьмую. А ведь это не самый худший пример, так как работать с сочетаниями строк 3–4, 7–8, 11–12 и т.д. в двоично-десятичном коде еще сложнее. Иначе говоря, если привлечь геометрические интерпретации, код Грея гарантирует, что две соседние, принадлежащие одному ребру, вершины гиперкуба, на котором осуществляется поиск, всегда декодируются в две ближайшие точки пространства вещественных чисел  $R^N$ , отстоящие друг от

друга на одну дискрету точности. Двоично-десятичный код подобным свойством не обладает.

### 10.4.3. Оператор селекции

Оператор селекции предназначен для улучшения средней приспособленности популяции в новом поколении. В качестве оператора селекции часто используется **метод рулетки** (roulette-wheel selection). Название этот метод получил благодаря тому, что своим принципом выбора родительских особей напоминает игру в рулетку. Колесо рулетки делится на секторы, после запуска и остановки колеса стрелка с большей вероятностью укажет на сектор, площадь которого будет больше остальных. Хотя и нельзя предсказать, какой сектор будет в итоге выбран, можно найти вероятность выбора каждого сектора. В ГА рулетка делится на количество секторов, равное размеру популяции, а площади секторов сопоставляются со значениями приспособленности каждой особи. Таким образом, отношение приспособленности одной особи к суммарной приспособленности популяции определяет площадь сектора соответствующей особи. Благодаря этому родительские особи выбираются с вероятностью, пропорциональной их функции приспособленности: чем лучше особь, тем больше у нее шансов размножиться. Вероятность выбора каждой особи вычисляется по формуле

$$p_s(x) = \frac{f_s(x)}{\sum_{i=1}^K f_i(x)}, \quad (10.14)$$

где  $f_i(x)$  – значение функции приспособленности  $i$ -й особи;  $K$  – размер популяции.

Площадь сектора выражается в процентах от площади колеса рулетки по формуле

$$v_s(x) = p_s(x) \cdot 100\%, \quad (10.15)$$

где  $p_s(x)$  – вероятность выбора  $i$ -й особи.

При вращении рулетки  $K$  раз получают родительский пул размером  $K$ . Особи с высокой приспособленностью могут быть включены в родительский пул несколько раз, такая схема репродукции способствует сильнейшим жить и размножаться, а менее приспособленным умирать.

Известны и другие способы отбора. Например, отбор может быть ранжированным: все особи ранжируются, т.е. упорядочиваются по приспособленностям и заданная часть лучших особей (например, лучшая половина) отбирается для формирования следующего поколения.

*Пример.* Найти максимум функции  $f(x) = 2x^2 + 1$  для целочисленной переменной  $x \in [0, 31]$ . Закодируем значения переменной в виде двоичных последовательностей. Очевидно, что целые числа из интервала от 0 до 31 можно представить последовательностями нулей и единиц, используя двоичную систему счисления. Число 0 при этом записывается как 00000, а число 31 – как 11111. В данном случае хромосомы приобретают вид двоичных последовательностей, состоящих из 5 битов, т.е. цепочек длиной 5. В роли функции приспособленности выступает целевая функция  $f(x) = 2x^2 + 1$ . Тогда приспособленность хромосомы  $ch_i$ ,  $i = 1, 2, \dots, N$ , будет определяться значением функции  $f(x)$  для  $x_i = ch_i$ , равного фенотипу, соответствующему генотипу  $ch_i$ . Обозначим эти фенотипы  $x_i$ . В таком случае значение функции приспособленности хромосомы  $ch_i$  будет равно  $f(x_i)$ . Выберем случайным образом исходную популяцию, состоящую из шести кодовых последовательностей:

$$\begin{aligned} ch_1 &= 10011, ch_2 = 00011, ch_3 = 00111, \\ ch_4 &= 10101, ch_5 = 01000, ch_6 = 11101. \end{aligned}$$

Соответствующие им фенотипы – это числа

$$x_1 = 19, x_2 = 3, x_3 = 7, x_4 = 21, x_5 = 8, x_6 = 29.$$

*Справка.* Перевод числа из двоичной системы в десятичную выполняют по следующему алгоритму. Нумеруют разряды числа справа налево, начиная с нуля, и вычисляют результат:

$$\begin{aligned}
 111100110_2 &= 1 \cdot 2^8 + 1 \cdot 2^7 + 1 \cdot 2^6 + 1 \cdot 2^5 + 0 \cdot 2^4 + \\
 &\quad + 0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = \\
 &= 256 + 128 + 64 + 32 + 4 + 2 = 486_{10}.
 \end{aligned}$$

По формуле  $f(x) = 2x^2 + 1$  рассчитываем значения функции приспособленности для каждой хромосомы в популяции:

$$\begin{aligned}
 f(x_1) &= 723, f(x_2) = 19, f(x_3) = 99, f(x_4) = 883, \\
 f(x_5) &= 129, f(x_6) = 1683.
 \end{aligned}$$

Среднее число приспособленности составляет  $\bar{f}(x) \approx 589$ .

Селекцию хромосом проводим *методом рулетки*. Для этого выбираем шесть хромосом для репродукции. Колесо рулетки представлено на рисунке 10.12. Размеры секторов пропорциональны числам

$$\frac{\sum_{k=1}^i f(x_k)}{\sum_{k=1}^6 f(x_k)}, i = 1, \dots, 6.$$

Площади секторов для каждой особи соответственно равны 20,4 %; 0,6 %; 2,8 %; 24,9 %; 3,7 %; 47,6 %.

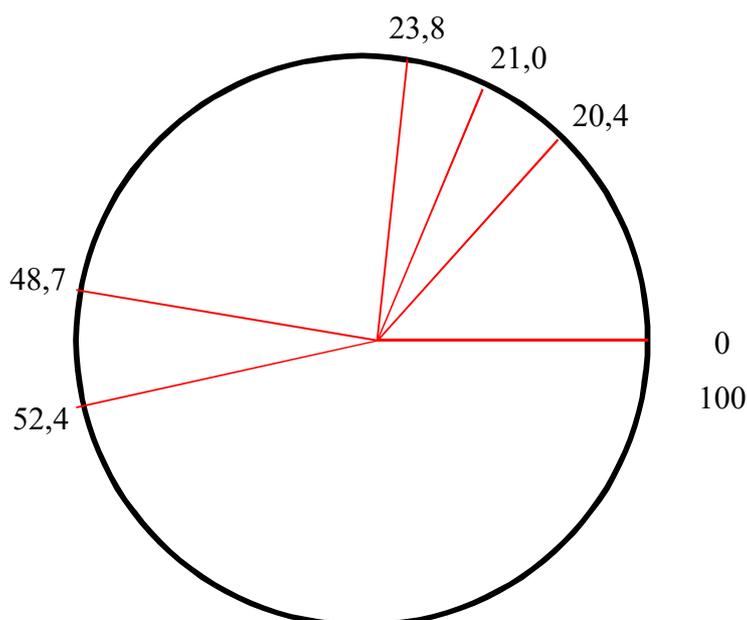


Рисунок 10.12 – Селекция хромосом методом рулетки

Пусть при розыгрыше случайных чисел из интервала от 0 до 100 выпали следующие числа: 97, 26, 54, 13, 31, 88. Это означает выбор хромосом  $ch_6$ ,  $ch_4$ ,  $ch_6$ ,  $ch_1$ ,  $ch_4$ ,  $ch_6$ . Сформируем для скрещивания пары  $(ch_1 - ch_6)$ ,  $(ch_4 - ch_6)$ ,  $(ch_6 - ch_6)$ . Зададим случайным образом для пары  $(ch_1 - ch_6)$  точку скрещивания, равную 4, а для пар  $(ch_4 - ch_6)$ ,  $(ch_6 - ch_6)$  точку скрещивания, равную 2. Результаты скрещивания приведены в таблицах 10.3–10.5.

Таблица 10.3 – Результаты скрещивания для пары  $(ch_1 - ch_6)$

	Родители					Скрещивание		Потомки				
	1	2	3	4	5			1	2	3	4	5
$ch_1$	1	0	0	1	1		$chl_1$	1	0	0	0	1
$ch_6$	1	1	1	0	1		$chl_2$	1	1	1	1	1

Таблица 10.4 – Результаты скрещивания для пары  $(ch_4 - ch_6)$

	Родители					Скрещивание		Потомки				
	1	2	3	4	5			1	2	3	4	5
$ch_4$	1	0	1	0	1		$chl_3$	1	1	1	0	1
$ch_6$	1	1	1	0	1		$chl_4$	1	0	1	0	1

Таблица 10.5 – Результаты скрещивания для пары  $(ch_6 - ch_6)$

	Родители					Скрещивание		Потомки				
	1	2	3	4	5			1	2	3	4	5
$ch_6$	1	1	1	0	1		$chl_5$	1	1	1	0	1
$ch_6$	1	1	1	0	1		$chl_6$	1	1	1	0	1

Итак, в новую популяцию включаются хромосомы

$$chl_1 = 10001, chl_2 = 11111, chl_3 = 11101,$$

$$chl_4 = 10101, chl_5 = 11101, chl_6 = 11101.$$

В результате декодирования получим соответствующие им фенотипы:

$$x_{1_1} = 17, x_{1_2} = 31, x_{1_3} = 29, x_{1_4} = 21, x_{1_5} = 29, x_{1_6} = 29.$$

Соответственно значения функции приспособленности хромосом новой популяции

$$f(x_{1_1}) = 579, f(x_{1_2}) = 1923, f(x_{1_3}) = 1683,$$

$$f(x_{1_4}) = 883, f(x_{1_5}) = 1683, f(x_{1_6}) = 1683.$$

Среднее значение функции приспособленности составляет  $\bar{f}(x_1) \approx 1406$ , т.е. возросло с 589 до 1406. При этом для второго потомка получили максимальное значение функции приспособленности, равное 1923.

**Турнирный метод** также является оператором селекции, отбирающим особей в зависимости от значения функции приспособленности. Главная идея заключается в том, чтобы выбрать особь с самой высокой приспособленностью из некоторого числа особей и позволить ей размножиться (записать в родительский пул). В этом методе из популяции, содержащей  $K$  особей, выбирается случайным образом  $t$  особей и лучшая из них записывается в промежуточный массив (рисунок 10.13). Эта операция повторяется  $K$  раз. Особи в полученном промежуточном массиве затем используются для скрещивания (также случайным образом). Размер группы строк, отбираемых для турнира, обычно равен 2 ( $t = 2$ ) или 3 ( $t = 3$ ).

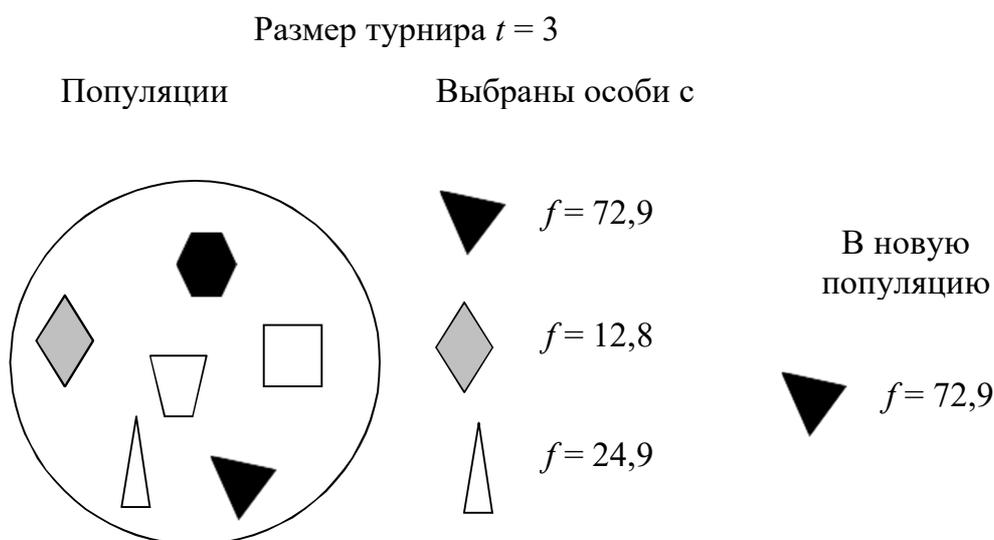


Рисунок 10.13 – Турнирный метод селекции

Достоинство этого метода состоит в том, что отсутствие дополнительных вычислений делает его эффективным и простым в реализации.

#### 10.4.4. Скрещивание

Скрещивание представляет собой обмен генетическим материалом внутри популяции. Подразумевается, что свой генетический материал будут в основном передавать особи с хорошей приспособленностью, потому что цель скрещивания – улучшение среднего качества популяции. Проблема в том, что когда особи найдут локальный экстремум, они будут стремиться занять всю популяцию. Процесс, когда все особи начинают быть похожими друг на друга, называется вырождением популяции. Чтобы не давать алгоритму сходиться к локальному экстремуму, используется оператор мутации.

Пусть  $i$ -я и  $j$ -я особи  $k$ -й популяции представлены как  $p_i^k = (a_1, a_2, \dots, a_n) \in P^k$  и  $p_j^k = (b_1, b_2, \dots, b_n) \in P^k$ . В одноточечном варианте скрещивания возникнут два элемента популяции  $k+1$ :  $p_i^{k+1} = (a_1, \dots, a_c, b_{c+1}, \dots, b_n) \in P^{k+1}$ ,  $p_j^{k+1} = (b_1, \dots, b_c, a_{c+1}, \dots, a_n) \in P^{k+1}$ , где точка  $c$  выбирается случайно.

В двухточечном варианте соответственно точек пересечения будет две, и они также выбираются случайно. Легко расширить эту конструкцию до  $n$  точек. В случае нечетного  $n$  возникает  $(n+1)$ -точечное скрещивание с  $(n+1)$ -й точкой между последней и первой компонентами.

Скрещиванием с маской является результат в виде двух потомков с компонентами, принадлежность которых определяется по битовой маске. То есть результатом скрещивания родителей  $p_i^k, p_j^k$  в  $k$ -й популяции становятся два элемента популяции  $k+1$ , такие что

$$p_i^{k+1} = (c_1, c_2, \dots, c_n) \in P^{k+1}, \quad p_j^{k+1} = (d_1, d_2, \dots, d_n) \in P^{k+1},$$

где  $c_i = a_i$  при  $m_i = 0$ ,  $c_i = b_i$  при  $m_i \neq 0$  и противоположные условия для второго отпрыска. Маска выбирается случайно. Для простоты ею может быть третья особь.

### 10.4.5. Оператор мутации

Мутации необходимы для того, чтобы поддерживать разнообразие особей в популяции и не позволять решению сходиться к локальному оптимуму. Мутации играют роль как в восстановлении утерянного генетического материала, так и в появлении у особей новых признаков. Традиционно оператор мутации рассматривают как простой оператор поиска, в то время как скрещивание использует уже существующие хорошо приспособленные особи для создания потомков. Благодаря оператору мутации случайно изменяются части хромосом. Для каждого способа представления параметров задачи существуют различные операторы мутации. Вероятность возникновения мутации обычно принимают за  $1/L$ , где  $L$  – длина хромосомы. Также возможно осуществлять только те мутации, которые способны увеличить приспособленность особи. Такой вариант может улучшить результаты поиска, но необходимо иметь в виду, что он также может уменьшить разнообразие в популяции и обеспечить сходимость к локальному оптимуму.

*Пример.* Рассмотрим популяцию, полученную в предыдущем примере:

$$\begin{aligned} ch1_1 &= 10001, ch1_2 = 11111, ch1_3 = 11101, \\ ch1_4 &= 10101, ch1_5 = 11101, ch1_6 = 11101. \end{aligned}$$

Примем вероятность мутации, равной  $1/5 = 0,2$ . Для каждого потомка возьмем случайное число на отрезке  $[0,1]$ , если это число меньше  $0,2$ , то инвертируем случайно выбранный ген (заменяем 0 на 1 или наоборот) (таблица 10.6).

Среднее значение функции приспособленности после мутации возросло и составило 1519.

Таблица 10.6 – Мутация потомков

Номер	Особь-потомки	Случайное число	Выбранный ген для мутации	Потомок после мутации	Приспособленность потомка до мутации	Приспособленность потомка после мутации
1	10001	0,6	1	10001	579	579
2	11111	0,1	5	11110	1923	1801
3	11101	0,3	1	11101	1683	1683
4	10101	0,04	2	11101	883	1683
5	11101	0,7	1	11101	1683	1683
6	11101	0,35	3	11101	1683	1683
Среднее значение $f(x)$					1406	1519

#### 10.4.6. Операторы отбора особей в новую популяцию

Классическим способом отбора в новую популяцию является замещение старой популяции своими потомками. Когда потомки получаются менее приспособленными, чем их родители, это становится не очень эффективно, поэтому существует стратегия *элитарного отбора*. Элитаризм бывает двух видов: конкурентный и неконкурентный. Первый представляет собой отбор лучших особей из родительского пула и пула потомков. Второй способ подразумевает переход лучших особей из старой популяции в новую, даже если их приспособленность хуже приспособленности потомков. Стратегия элитаризма позволяет не терять лучшие решения, найденные при мутациях.

#### 10.4.7. Основные отличия генетических алгоритмов от традиционных методов поиска решений

Генетические алгоритмы работают с кодовыми строками, от которых зависят значения аргументов целевой функции и соответственно значение самой целевой функции. Интерпретация этих кодов выполняется только в операторе редукции (декодирование или переход к фенотипу). В остальном работа алгоритма не зависит от смысловой интерпретации кодов.

Для поиска лучшего решения генетический алгоритм на отдельном шаге использует сразу несколько точек поискового пространства (несколько вариантов решения задачи) одновременно, а не переходит от точки к точке, как это делается в традиционных методах. Это позволяет преодолеть один из недостатков ГА – опасность попадания в локальный экстремум целевой функции, если она не является унимодальной (т.е. имеет несколько экстремумов). Использование нескольких точек одновременно значительно снижает такую возможность.

Генетический алгоритм использует как вероятностные правила для порождения новых точек, так и детерминированные правила для перехода от одних точек к другим. Одновременное использование элементов случайности и детерминированности дает значительно больший эффект, чем раздельное.

## 10.5. Поиск минимума функции одной переменной

Пусть требуется найти глобальный минимум функции

$$f(x) = 5 - 24x + 17x^2 - \frac{11}{3}x^3 + \frac{1}{4}x^4$$

на отрезке  $[0, 7]$ . На этом отрезке функция принимает минимальное значение в точке  $x = 1$ . Очевидно, что в точке  $x = 6$  функция попадает в локальный минимум. Если для нахождения глобального минимума использовать градиентные методы, то в зависимости от начального приближения можно попасть в этот локальный минимум.

Рассмотрим на примере данной задачи принцип работы генетических алгоритмов. Для простоты положим, что  $x$  принимает лишь целые значения, т.е.  $x = \{0, 1, 2, 3, 4, 5, 6, 7\}$ . Это предположение существенно упростит изложение, сохранив все основные особенности работы генетического алгоритма.

Выберем случайным образом несколько чисел на отрезке  $[0, 7]$ :  $\{2, 3, 5, 4\}$ . Будем рассматривать эти числа в качестве пробных решений задачи (рисунок 10.14).

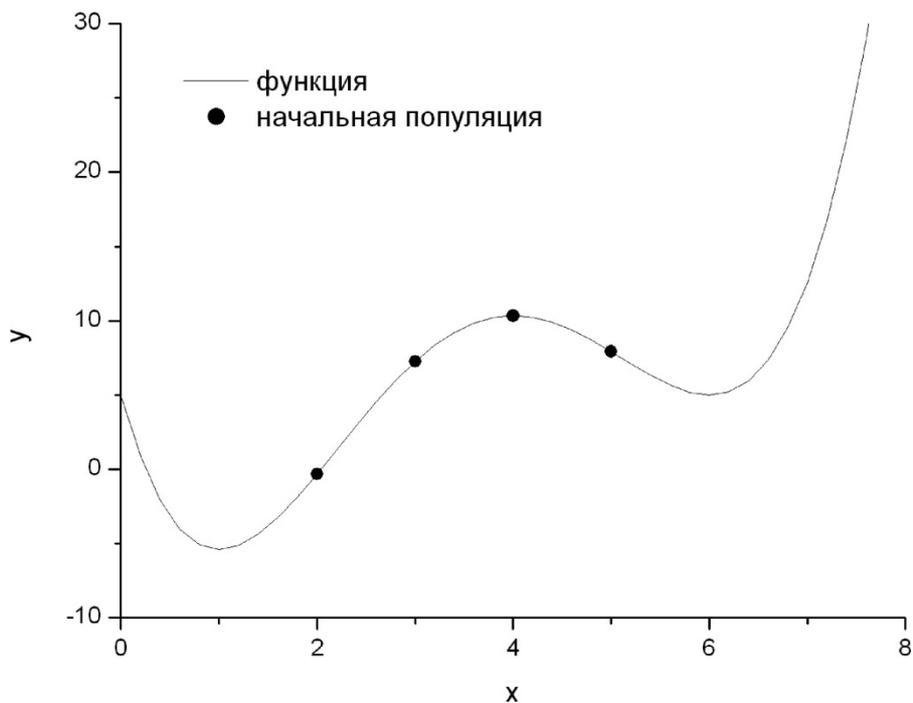


Рисунок 10.14 – Целевая функция с выбранными значениями пробных решений

Запишем пробные решения в двоичной форме: 010, 011, 101, 100. Как известно, принцип естественного отбора заключается в том, что в конкурентной борьбе выживает наиболее приспособленный. В нашем случае приспособленность особи определяется целевой функцией: чем меньше значение целевой функции, тем более приспособленной является особь, т.е. пробное решение, использовавшееся в качестве аргумента целевой функции (таблица 10.7).

Таблица 10.7 – Исходная популяция

Номер	Особи		
	Целое число	Двоичное число	Приспособленность
1	2	010	-0,33
2	3	011	7,25
3	5	101	7,92
4	4	100	10,33

Теперь приступим к процессу размножения: попробуем на основе исходной популяции создать такую новую, чтобы пробные решения ней были ближе к искомому глобальному минимуму целевой

функции. Для этого сформируем из исходной популяции брачные пары для скрещивания. Поставим в соответствие каждой особи исходной популяции случайное целое число из диапазона от 1 до 4. Будем рассматривать эти числа как номера членов популяции. При таком выборе какие-то из членов популяции не будут участвовать в процессе размножения, так как образуют пару сами с собой. Какие-то члены популяции примут участие в процессе размножения неоднократно с различными особями популяции. Процесс размножения (рекомбинация) заключается в обмене участками хромосом между родителями.

Для нашей популяции процесс создания первого поколения потомков показан в таблице 10.8.

Таблица 10.8 – Одноточечный кроссинговер

Номер	Особь популяции	Выбранный номер	Вторая особь-родитель	Точка кроссинговера	Особь-потомки
1	010	1	010	1	010
2	011	4	100	1	000 111
3	101	3	101	2	101
4	100	3	101	2	100 101

Следующим шагом в работе генетического алгоритма являются мутации, т.е. случайные изменения полученных в результате скрещивания хромосом. Пусть вероятность мутации равна 0,3. Для каждого потомка возьмем случайное число на отрезке  $[0,1]$ . Если это число меньше 0,3, то инвертируем случайно выбранный ген (заменяем 0 на 1 или наоборот) (таблица 10.9).

Как видно на примере, мутации способны улучшить (первый и второй потомки) или ухудшить (четвертый потомок) приспособленность особи-потомка. В результате скрещивания хромосомы обмениваются «хвостами», т.е. младшими разрядами в двоичном представлении числа. В результате мутаций изменению может подвергнуться любой разряд, в том числе старший.

Таблица 10.9 – Мутация потомков

Номер	Особи-потомки	Случайное число	Выбранный ген для мутации	Потомок после мутации	Приспособленность потомка до мутации	Приспособленность потомка после мутации
1	000	0,1	3	001	5	-5,42
2	111	0,2	3	110	12,58	5
3	100	0,5	1	100	10,33	10,33
4	101	0,2	2	111	7,92	12,58

Таким образом, если скрещивание приводит к относительно небольшим изменениям пробных решений, то мутации могут их существенно изменить.

Теперь из четырех особей-родителей и четырех полученных особей-потомков необходимо сформировать новую популяцию. В новую популяцию отберем четыре наиболее приспособленные особи из числа «старых» особей и особей-потомков (таблица 10.10).

Таблица 10.10 – Формирование новой популяции из особей-родителей и особей-потомков

Номер	Особи	Приспособленность	Новая популяция	Приспособленность особей в новой популяции
1	010	-0,33	001	-5,42
2	011	7,25	010	-0,33
3	101	7,92	110	5
4	100	10,33	011	7,25
5	001	-5,42		
6	110	5		
7	100	10,33		
8	111	12,58		

В результате получим новое поколение (рисунок 10.15).

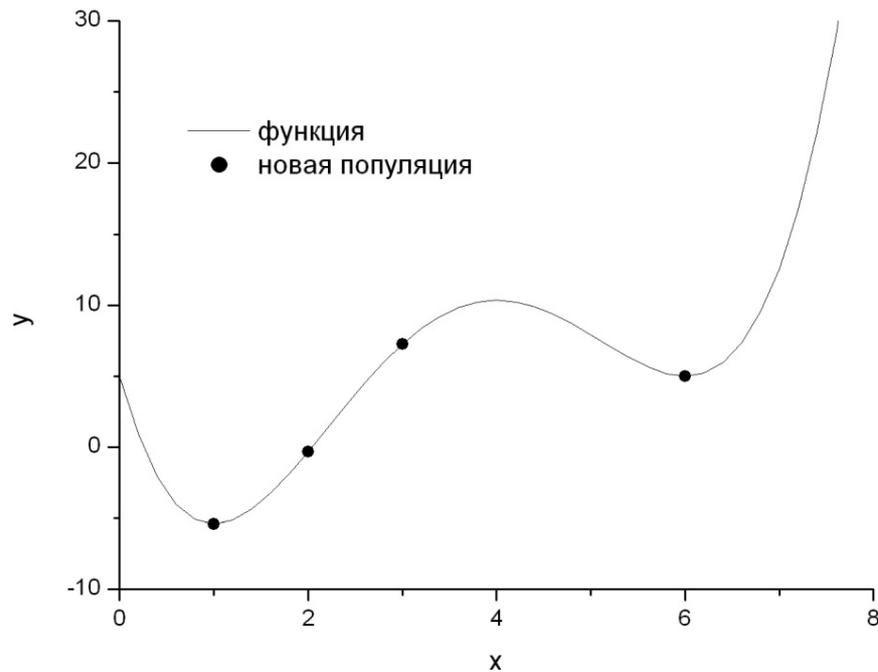


Рисунок 10.15 – Изменение популяции в процессе естественного отбора

Получившуюся популяцию можно будет вновь подвергнуть скрещиванию, мутации и отбору особей в новое поколение. Таким образом, через несколько поколений мы получим популяцию из похожих и наиболее приспособленных особей. Значение приспособленности наиболее «хорошей» особи (или средняя приспособленность по популяции) и будет являться решением задачи. Согласно этому в данном случае, взяв наиболее приспособленную особь 001 во втором поколении, находим, что минимумом целевой функции будет значение 5,42, соответствующее аргументу  $x = 1$ . Тем самым попадания в локальный минимум удалось избежать! На этом примере показан вариант простого генетического алгоритма. При использовании ГА к разным задачам возможно моделирование его основных операторов.

## 10.6. Многомерная безусловная оптимизация при помощи генетических алгоритмов

### 10.6.1. Постановка задачи

Рассмотрим задачу безусловной минимизации

$$f(x) \rightarrow \min \quad (10.16)$$

при условии, что нам известны границы изменения переменной  $x$  :  
 $l \leq x \leq u$ .

Вектор параметров целевой функции  $f(x)$ ,  $x \in D$  имеет компоненты  $x_1, x_2, \dots, x_n$ , которые представляют собой искомые переменные задачи. Для оптимизации многомерных функций удобно использовать непрерывные генетические алгоритмы, т.е. алгоритмы с вещественным кодированием. При таком способе кодирования хромосомы представляют собой вектор вещественных чисел. Размер хромосомы равен длине вектора решения задачи, таким образом, ген представляет собой компоненту вектора решения задачи. Значения генов принимаются из интервалов, на которых определены соответствующие переменные  $x_i$ ,  $i = 1, \dots, n$ , поэтому генетические операторы должны учитывать этот момент.

Применение вещественных чисел позволяет использовать большие домены для переменных, что сложно осуществить при бинарном представлении. Еще одно преимущество в том, что при незначительных изменениях в переменных происходят небольшие изменения значения функции. Использование вещественного кодирования очень близко к естественной формулировке многих задач, что стирает разницу между генотипом и фенотипом. В результате пропадает необходимость в дополнительном кодировании и декодировании, что увеличивает скорость работы алгоритма. Для вещественного кодирования изобретены различные способы кроссинговера и мутации, рассмотрим самые распространенные из них.

## 10.6.2. Селекция в задаче многомерной оптимизации

Селекция является важным фактором в генетическом алгоритме. Этот процесс случайный и позволяет выбрать лучшие решения. Главная цель оператора селекции выделить решения, которые лучше средних, и удалить решения хуже средних из популяции для следующей генерации. Это достигается путем решения следующих задач:

- определения лучших решений в популяции;
- создания нескольких копий хороших решений;
- удаления плохих решений из популяции, чтобы хорошие решения могли заменить их в популяции.

Имеется несколько схем селекции: рулетка, ранговая селекция, элитарная селекция, турнирный метод и другие.

*Ранговая селекция.* В этой схеме селекции порядок ранжирования индивидуумов относительно значений их функции пригодности определяет вероятность в текущей популяции. Популяция сортируется от лучших к худшим особям. Вероятность селекции каждой особи определяется в соответствии с рангом. Известны линейные и нелинейные ранговые методы. В линейном ранговом методе каждому индивидууму присваивается ранг в порядке убывания функции пригодности и вероятность  $i$ -го индивидуума в популяции вычисляется по формуле

$$p_i = \frac{F_i}{\sum_{j=1}^K F_j}, F_i = \frac{1}{K} \left( f_{\max} - (f_{\max} - f_{\min}) \frac{i-1}{K-1} \right), \quad (10.17)$$

где  $K$  – размер популяции;  $f_{\max}$ ,  $f_{\min}$  – максимальное и минимальное значения целевой функции хромосомы (особи) в популяции.

Если целевая функция принимает отрицательные значения, то вероятность можно вычислять по формуле

$$p_i = \frac{1/(1+w_i)}{\sum_{j=1}^K (1/(1+w_j))}, i=1, \dots, K, \quad w_i = \frac{f(x_i)}{\sum_{j=1}^K f(x_j)}, i=1, \dots, K. \quad (10.18)$$

В нелинейном ранговом методе вероятность  $i$ -го индивидуума определяется как

$$p_i = c \cdot (1 - c)^{i-1}, c \in [0, 1]. \quad (10.19)$$

*Турнирная селекция.* В этом методе селекции выбирается группа хромосом (индивидуумов) из популяции случайным образом, затем наилучшие хромосомы из этой группы берутся в качестве родителей для создания следующего поколения с помощью генетических операций, таких как кроссинговер и мутация. Этот процесс повторяется до тех пор, пока не будет достигнута необходимая численность популяции хромосом. Параметром турнирного метода является размер турнира. Количество хромосом в турнире принимает значение от 2 до размера турнира.

Турнирный метод основывается на следующих правилах.

1. Для задачи безусловной оптимизации: среди группы выбранных хромосом выбирается хромосома с лучшим значением пригодности.
2. Для задач условной оптимизации:
  - когда ограничения задачи удовлетворяют все хромосомы (индивидуумы), выбирается хромосома с лучшим значением функции пригодности;
  - когда одна хромосома (индивидуальная) удовлетворяет ограничениям задачи, а другие – не удовлетворяют, выбирается первая;
  - когда все хромосомы (индивидуумы) не удовлетворяют ограничениям-неравенствам задачи, выбирается хромосома с меньшим нарушением ограничения;
  - если все хромосомы (индивидуумы) недопустимы в задаче с ограничениями-равенствами, то выбирается любая хромосома.

### **10.6.3. Кроссинговер в задаче многомерной оптимизации**

После селекции выбранные хромосомы принимают участие в скрещивании для получения лучших потомков с целью улучшения популяции. Эта операция проводится на двух или более родительских хромосомах (решениях) одновременно и создает потомство из родительских генов. Число хромосом, которое будет принимать

участие в скрещивании, равно  $[p_{\text{cros}} \cdot K]$ , где  $p_{\text{cros}}$  – вероятность скрещивания;  $K$  – размер популяции;  $[\cdot]$  – целая часть числа.

Процесс улучшения решения включает:

- выбор двух или более случайных хромосом из текущей популяции;
- создание потомков путем оператора скрещивания.

При описании оператора кроссинговера для обозначения родительских особей, выбранных на этапе селекции, приняты следующие обозначения:

$$\begin{aligned} P_1(x^1) &= (x_1^1, x_2^1, \dots, x_n^1), \\ P_2(x^2) &= (x_1^2, x_2^2, \dots, x_n^2), \end{aligned} \tag{10.20}$$

где  $n$  – размерность задачи.

Для потомков особей  $P_1$  и  $P_2$  приняты обозначения

$$\begin{aligned} C_1(y^1) &= (y_1^1, y_2^1, \dots, y_n^1), \\ C_2(y^2) &= (y_1^2, y_2^2, \dots, y_n^2). \end{aligned} \tag{10.21}$$

Рассмотрим различные операторы скрещивания.

*Простой арифметический кроссинговер.* Пусть на генерации  $t$  для скрещивания случайным образом выбраны две родительские особи

$$\begin{aligned} P_1^t(x^1) &= (x_1^1, x_2^1, \dots, x_n^1), \\ P_2^t(x^2) &= (x_1^2, x_2^2, \dots, x_n^2). \end{aligned}$$

Тогда потомки

$$\begin{aligned} C_1^{t+1}(y^1) &= (y_1^1, y_2^1, \dots, y_n^1), \\ C_2^{t+1}(y^2) &= (y_1^2, y_2^2, \dots, y_n^2) \end{aligned}$$

образуются следующим образом:

$$\begin{aligned} y_i^1 &= a \cdot x_i^1 + (1-a) \cdot x_i^2, \quad i = 1, \dots, n; \\ y_i^2 &= a \cdot x_i^2 + (1-a) \cdot x_i^1, \quad i = 1, \dots, n. \end{aligned} \tag{10.22}$$

В качестве величины  $a$  можно взять

$$a = f_{\max} / (f_{\max} + f_{\min}),$$

где  $f_{\max} = \max_{1 \leq j \leq K} (f_j)$ ;  $f_{\min} = \min_{1 \leq j \leq K} (f_j)$ ;  $f_j$  – значение целевой функции  $j$ -го решения.

Здесь оператор скрещивания использует простой статический параметр  $a \in [0,1]$ . Если  $a$  является константой, то кроссинговер называют *равномерным арифметическим кроссинговером*. В противном случае его называют *неравномерным арифметическим кроссинговером*.

*Кроссинговер смешивания.* Создается случайный потомок в пределах гиперпрямоугольника, определяемого точками родителей. Пусть выбраны родительские хромосомы  $x_i^1$  и  $x_i^2$ . Тогда  $i$ -я хромосома потомка в однопараметрическом операторе смешивания определяется случайно из интервала  $[x_i^1 - \alpha \cdot D_i, x_i^2 + \alpha \cdot D_i]$ , где  $D_i = x_i^2 - x_i^1$ , при этом предполагается  $x_i^1 < x_i^2$ . Результирующие потомки определяются следующим образом:

$$y_i^1 = (1 - \mu_i)x_i^1 + \mu_i x_i^2, y_i^2 = (1 - \mu_i)x_i^2 + \mu_i x_i^1, \quad (10.23)$$

где  $\mu_i = (1 + 2\alpha)r_i - \alpha$ ;  $r_i$  – случайное число из интервала  $[0,1]$ . Наибольший эффект кроссинговер смешивания дает при  $\alpha = 0,5$ . При  $\alpha = 0$  оператор смешивания дает случайное решение из интервала  $[x_i^1, x_i^2]$ .

В двухпараметрическом операторе смешивания потомок создается путем выбора случайного решения из интервала  $[x_i^1 - \alpha \cdot D_i, x_i^2 + \beta \cdot D_i]$ , где  $\alpha$  и  $\beta$  – два положительных вещественных числа.

*Кроссинговер Лапласа.* Оператор скрещивания основан на распределении вероятностей Лапласа. Плотность  $G(x)$  и функция распределения  $F(x)$  вероятностей Лапласа определяются формулами

$$G(x) = \frac{1}{2b} \exp\left(-\frac{|x-a|}{b}\right), -\infty < x < \infty,$$

$$F(x) = \frac{1}{2} \exp\left(\frac{x-a}{b}\right), x \leq a,$$

где  $a \in R$ ,  $b > 0$  – параметры распределения. Случайно выбранные две родительские хромосомы

$$P_1(x^1) = (x_1^1, x_2^1, \dots, x_n^1),$$

$$P_2(x^2) = (x_1^2, x_2^2, \dots, x_n^2)$$

производят двух потомков

$$C_1(y^1) = (y_1^1, y_2^1, \dots, y_n^1),$$

$$C_2(y^2) = (y_1^2, y_2^2, \dots, y_n^2)$$

с компонентами  $y_i^1 = x_i^1 + \alpha \cdot |x_i^1 - x_i^2|$ ,  $y_i^2 = x_i^2 + \alpha \cdot |x_i^1 - x_i^2|$ ,  $i = 1, \dots, n$ , где  $\alpha$  – случайное число из распределения Лапласа, которое можно вычислить с помощью равномерного датчика по формуле  $\alpha = a + b \cdot \ln(2r)$ , где  $r$  – случайное число из интервала  $(0,1]$ .

#### 10.6.4. Мутация в задаче многомерной оптимизации

Основная задача оператора мутации – ввести генетическое разнообразие в популяцию. Этот оператор используется для улучшения тонкой настройки системы, что реализуется в хромосоме с наименьшей вероятностью (плохой хромосоме). Общий алгоритм мутации:

- 1) рассчитать целую часть числа  $p_m \cdot K$ , где  $p_m$  – вероятность мутации;  $K$  – размер популяции;
- 2) случайно выбрать хромосому из популяции и затем случайно выбрать ген этой хромосомы;
- 3) произвести новый ген, соответствующий выбранному гену, посредством операции мутации;
- 4) повторить шаги 1–3  $K$  раз.

Рассмотрим часто используемые операции мутации – неравномерную мутацию и экспоненциальную мутацию.

*Неравномерная мутация.* Пусть выбрана хромосома  $P(x) = (x_1, x_2, \dots, x_n)$  и  $x_k \in [l_k, u_k]$ , где  $l_k, u_k$  – нижняя и верхняя границы интервала соответственно для компоненты  $x_k$ . Мутация на итерации  $t$  происходит по формулам

$$y_k = \begin{cases} x_k + \Delta(t, u_k - x_k), & \text{если } r_k \leq 0,5, \\ x_k - \Delta(t, x_k - l_k), & \text{если } r_k > 0,5, \end{cases} \quad (10.24)$$

где  $r_k$  – случайное число из интервала  $[0,1]$ ;  $\Delta(t, z)$  – функция, определяемая формулой

$$\Delta(t, z) = z \cdot \left( 1 - r^{(1-t/T)^b} \right) \quad (10.25)$$

или

$$\Delta(t, z) = z \cdot r \cdot \left( 1 - (1 - t/T)^b \right), \quad (10.26)$$

где  $t$  – номер текущего поколения;  $T$  – максимальное число поколений;  $b$  – системный параметр, определяющий степень неравномерности;  $r$  – случайное число из интервала  $[0,1]$ .

*Экспоненциальная мутация.* Экспоненциальная мутация была предложена для решения оптимизационных задач в аэродинамике. Мутация  $k$ -го гена происходит следующим образом:

$$y_k = (1 - \tilde{g}) \cdot l_k + \tilde{g} \cdot u_k, \quad (10.27)$$

где

$$\tilde{g} = \begin{cases} g - g \cdot \left( \frac{g-r}{g} \right)^b, & \text{если } r < g, \\ g, & \text{если } r = g, \\ g + (1-g) \cdot \left( \frac{r-g}{1-g} \right)^b, & \text{если } r > g, \end{cases} \quad g = \frac{x_k - l_k}{u_k - x_k}, \quad (10.28)$$

$r$  – случайное число из интервала  $[0,1]$ .

## Вопросы для самопроверки

1. Приведите классификацию эвристических алгоритмов оптимизации.
2. Перечислите эволюционные алгоритмы поиска глобального минимума функции.
3. Назовите методы роевого интеллекта.
4. Перечислите методы, имитирующие физические процессы.
5. С помощью какой операции осуществляется преобразование вектора параметров  $x \in D$  в хромосому пространства представлений  $S$ ?
6. С помощью какой операции осуществляется обратное преобразование хромосомы из пространства представлений  $S$  в решение, заданное на пространстве параметров  $D$ ?
7. Что такое бинарное кодирование и декодирование?
8. Что понимают под вещественным кодированием?
9. Дайте понятие искусственной иммунной системы.
10. Опишите расширенный метод искусственных иммунных систем.
11. Опишите процедуру поиска в методе рассеивания.
12. Приведите особенности эволюционной стратегии преобразования ковариационной матрицы.
13. Назовите особенности метода динамических сетей.
14. Дайте определение генетическому алгоритму.
15. Дайте определение понятиям «ген», «хромосома», «генотип», «особь» и «фенотип».
16. Дайте понятие функции приспособленности.
17. Перечислите основные шаги алгоритма процесса формирования нового поколения.
18. Перечислите операторы генетического алгоритма и поясните их назначение.
19. Опишите операторы селекции.
20. Как вычислить вероятность каждой особи в методе селекции с помощью рулетки?
21. Какие операторы скрещивания вам известны?
22. Перечислите известные вам операторы мутации.

23. Какие из операторов генетического алгоритма выполняются с использованием элементов случайности, а какие – по строго детерминированным правилам?

24. Что может являться критерием остановки работы генетического алгоритма?

25. Перечислите предположения, на которых основан турнирный метод выбора решений в задаче безусловной оптимизации.

26. Перечислите предположения, на которых основан турнирный метод выбора решений в задаче условной оптимизации.

27. Как происходит скрещивание с помощью простого арифметического кроссинговера?

28. Как происходит скрещивание с помощью кроссинговера Лапласа?

29. Перечислите основные шаги алгоритма мутации в вещественных генетических алгоритмах.

# Глава 11. ПРИМЕРЫ ЗАДАЧ, РЕШАЕМЫХ С ИСПОЛЬЗОВАНИЕМ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ

## 11.1. Задача коммивояжера

Задача заключается в том, что коммивояжер должен посетить каждый город на некоторой территории ровно один раз, а затем вернуться к исходной точке. Необходимо так спланировать его маршрут, чтобы стоимость проезда между всеми городами (длина пути всего тура, время поездки) была минимальна. Поиск маршрута состоит в наборе перестановок из  $n$  городов. Оптимальным решением является перестановка, которая дает минимальную стоимость тура. Размерность пространства поиска равна  $(n-1)!$  для асимметричной задачи и  $(n-1)!/2$  для симметричной задачи. Другими словами, размерность тура  $N$  определяется множеством точек  $v = \{v_1, v_2, \dots, v_n\}$ , где  $v_i$  –  $i$ -й город, заданный координатами  $x_i, y_i$ . Длина пути одного из маршрутов (одной из перестановок) рассчитывается по формуле

$$f = \sum_{i=1}^{n-1} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} + \sqrt{(x_n - x_1)^2 + (y_n - y_1)^2}. \quad (11.1)$$

Решением задачи является способ планирования  $T = (T[1], T[2], \dots, T[n], T[1])$ , где  $T[i]$  – перестановка на множестве  $\{1, 2, \dots, N\}$ .

Введем матрицу расстояний между городами  $v_i$  и  $v_j$ :

$$d(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \quad (11.2)$$

тогда формула (11.1) примет вид

$$f(T) = \sum_{i=1}^{n-1} d(T[i], T[i+1]) + d(T[n], T[1]). \quad (11.3)$$

Задача коммивояжера заключается в минимизации функции  $f(T)$  по аргументу  $T$ , где  $T = (T[1], T[2], \dots, T[n], T[1])$ .

**Детерминированные алгоритмы.** Задача коммивояжера может быть редуцирована к задаче целочисленного линейного программирования. Пусть города пронумерованы числами от 1 до  $n$ . Для каждой пары городов зададим расстояние (или стоимость переезда)  $c_{ij}$ ,  $i, j = 1, \dots, n$ . Введем бинарные переменные  $x_{ij}$ :

$$x_{ij} = \begin{cases} 1, & \text{если осуществлен переезд из пункта } i \text{ в пункт } j, \\ 0 & \text{иначе.} \end{cases} \quad (11.4)$$

Тогда имеем следующую оптимизационную задачу. Найти переменные  $x_{ij}$ ,  $i, j = 1, \dots, n$ , обращающие в минимум функцию

$$f(x) = \sum_{i=1}^n \sum_{j=1}^{i-1} c_{ij} x_{ij} \quad (11.5)$$

и удовлетворяющие ограничениям

$$\forall i \in N = \{1, 2, \dots, n\} \quad \sum_{j \neq i}^{n-1} x_{ij} = 2, \quad (11.6)$$

$$\forall S \subset N \quad \sum_{i \in S} \sum_{j \notin S} x_{ij} \geq 2.$$

Первое ограничение означает, что каждая вершина должна сообщаться через пару ребер с остальными вершинами, то есть через входное и выходное ребро. Второе ограничение связано с устранением подмаршрутов, т.е. каждое множество вершин  $S \subset N$  является либо пустым, либо содержит все вершины, сочетающиеся с остальными вершинами через минимум два ребра.

Для решения задачи можно использовать точные методы, такие как комбинаторный метод ветвей и границ и метод отсекающих плоскостей (метод Гомори). Однако с ростом размерности задачи экспоненциально возрастает требуемый объем памяти и времени счета. Поэтому для задач большой размерности точные методы не используются.

В течение последних десятилетий появилось несколько приближенных алгоритмов для аппроксимации оптимального решения:

метод ближайшего соседа, жадный алгоритм, метод включения ближайшего города, метод имитации отжига и др.

**Решение задачи коммивояжера с помощью генетического алгоритма.** Задача коммивояжера относится к категории  $NP$ -полных задач, т.е. задач, решаемых методом полного перебора всех вариантов.

Формализация задачи:

*ген* – это число, характеризующее номер посещаемого города;

*генотип* – строка из чисел длиной  $N$ , описывающая порядок посещения городов;

*особь* – конкретная строка из чисел (допустимый вариант решения задачи).

*Представление тура.* Существуют различные представления тура: бинарное представление, представление путей и представление смежности. Рассмотрим представление путей. Представление путей – это упорядоченный список городов, которые необходимо посетить, причем если город  $i$  является  $j$ -м элементом списка, то город  $i$  необходимо посетить  $j$ -м по счету. Например, тур 3-2-4-1-7-5-8-6 представляется просто как (3 2 4 1 7 5 8 6).

*Генерация начальной популяции.* Начальная популяция обуславливает скорость и сходимость алгоритма.

Для генерации начальной совокупности применяется несколько методов:

- случайная генерация начальной популяции;
- генерация первой особи случайным образом, которая мутирует  $N - 1$  раз с помощью оператора мутации;
- генерация первой особи с использованием эвристического механизма. Приемник первого города расположен на расстоянии, меньшем по сравнению с другими. Затем используется оператор мутации на маршруте, полученном для того, чтобы сгенерировать  $N - 2$  других индивидуума, которые будут составлять начальную популяцию.

*Операторы скрещивания.* Одними из удачных операторов кроссинговера для решения задачи коммивояжера признаны Partially-Mapped Crossover (PMX) и Order Crossover (OX), или, как их еще

называют, частично соответствующий кроссинговер и упорядоченный кроссинговер. Операторы РМХ и ОХ представляют собой обмен частей туров двух родительских особей. Для начала случайно выбираются две текущие точки для выделения частей туров из родительских особей (так называемый двухточечный кроссинговер). При таком обмене могут возникать частичные туры, например тур  $T = (5\ 7\ 1\ 2\ 5\ 4\ 3\ 6)$  содержит частичный тур: коммивояжер, выйдя из города 5, вернется в город 5, не посетив города 4, 3 и 6. Чтобы не возникало частичных туров, в этих операторах существуют механизмы создания правильных туров. Алгоритм создания потомков с помощью оператора РМХ представлен на рисунках 11.1 и 11.2.

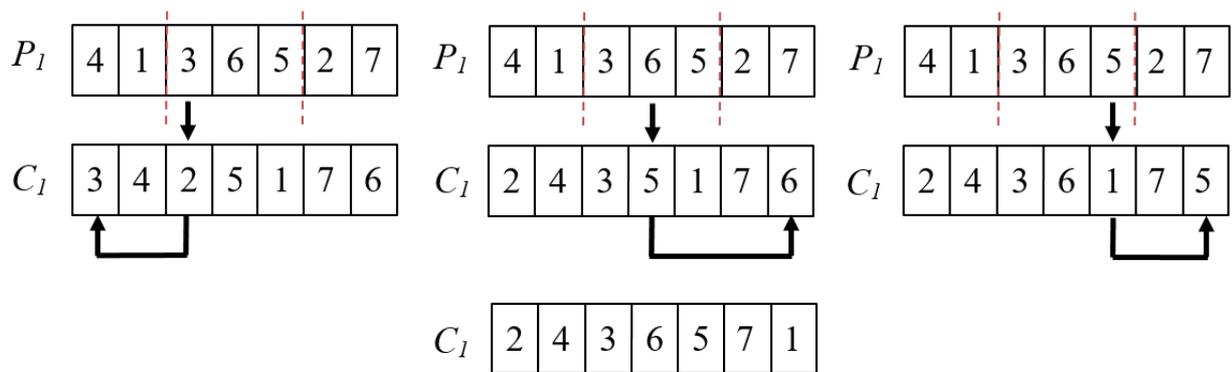


Рисунок 11.1 – Формирование первого потомка с помощью оператора РМХ

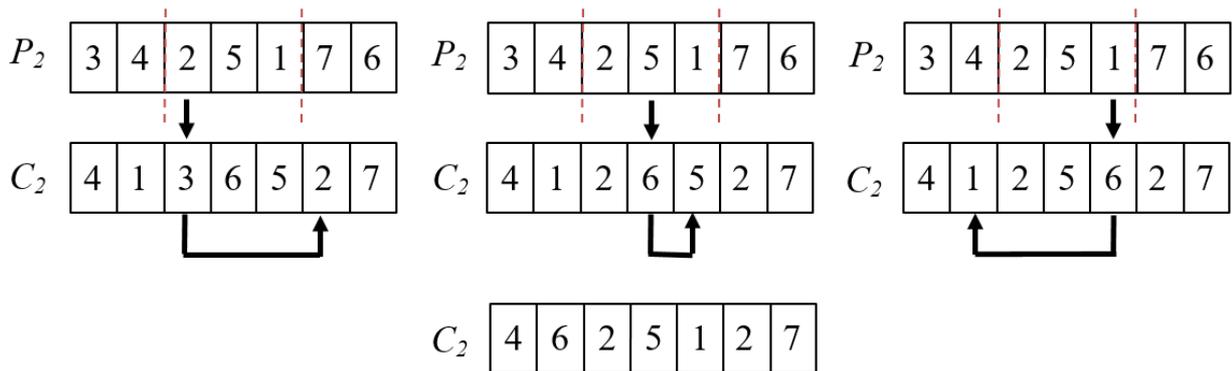


Рисунок 11.2 – Формирование второго потомка с помощью оператора РМХ

Он состоит из следующих шагов.

Шаг 1. Случайно выбрать две текущие точки для выделения части тура из родителя 1.

Шаг 2. Скопировать все города из родителя 2 в потомка.

Шаг 3. Для каждого города  $C$  из передаваемой части тура родителя 1:

1) найти город  $C$  в потомке, на его место поставить тот город, который заменяется городом  $C$ ;

2) скопировать город  $C$  из родителя 1 в потомка на ту же позицию.

Шаг 4. Для создания второго потомка поменять местами родителей и перейти к шагу 1.

Особенность оператора ОХ состоит в том, что он сохраняет порядок следования городов, а не их расположение в туре. Он строит потомка с использованием части тура одного из родителей и сохраняет порядок следования городов, представленный в другом родителе. Алгоритм создания потомков с помощью оператора ОХ представлен на рисунке 11.3.

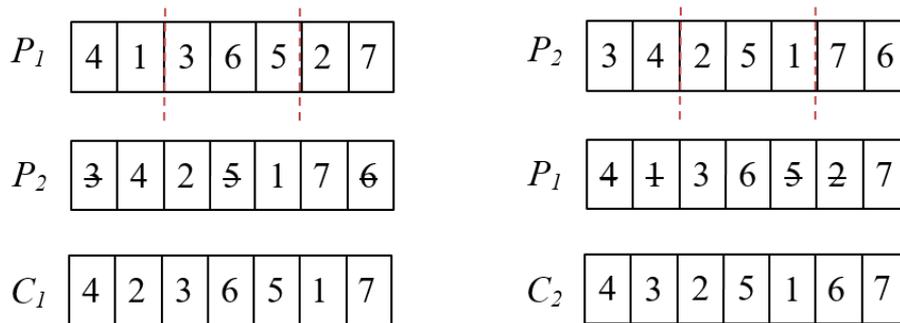


Рисунок 11.3 – Формирование потомков с помощью оператора ОХ

Опишем его.

Шаг 1. Выбрать две текущие точки для выделения части тура из родителя 1.

Шаг 2. Выделить часть тура из родителя 1 и передать потомку. Все города из переданной части тура отметить как использованные в родителе 2.

Шаг 3. Начиная с правой стороны от переданной части тура помещать неиспользованные города родителя 2 в потомок.

Шаг 4. Для создания второго потомка поменять местами родителей и перейти к шагу 1.

Операторы кроссинговера РМХ и ОХ не используют информацию о расстояниях между городами. Эвристический кроссинговер –

пример создания потомков с использованием информации о длине пути между городами. Алгоритм создания потомков с помощью эвристического кроссинговера включает следующие шаги.

Шаг 1. Выбрать начальным городом тура потомка любой из начальных городов туров родительских особей.

Шаг 2. Взять следующий город из родительских особей, оценить расстояние до них и передать потомку тот город, расстояние до которого меньше.

Шаг 3. Если выбранный город образует частичный тур, то проверить город другого родителя, а если и он не создает правильный тур, то выбрать случайный город, не образующий цикл.

Шаг 4. Повторять шаги 2 и 3, пока все города не войдут в тур потомка.

*Операторы мутации в решении задачи коммивояжера.* Существуют различные операторы мутации, ориентированные на решение задачи коммивояжера. Простая мутация перестановки SIM выбирает случайно два города в списке городов и переставляет их, т.е. меняет местами (рисунок 11.4).

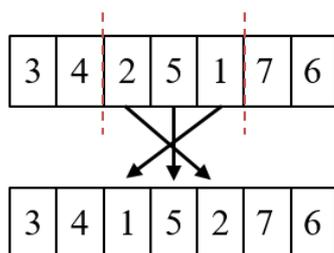


Рисунок 11.4 – Оператор мутации SIM

При мутации перестановки относительно центра CIM хромосома случайным образом разбивается на две части. В каждой части хромосомы гены переставляются в обратном порядке (рисунок 11.5).

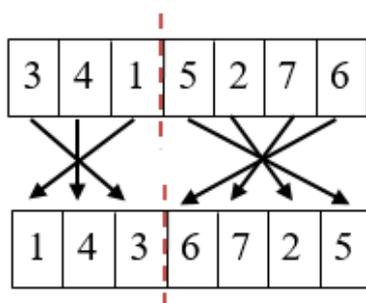


Рисунок 11.5 – Оператор мутации CIM

Мутация вставки IM случайно выбирает ген в хромосоме и переставляет его в другое случайно выбранное место (рисунок 11.6).

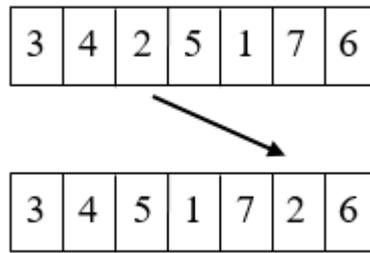


Рисунок 11.6 – Оператор мутации IM

Мутация перестановки SM случайно выбирает два гена и меняет их местами (рисунок 11.7).

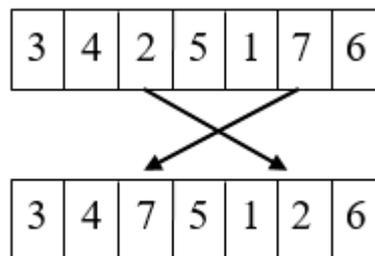


Рисунок 11.7 – Оператор мутации SM

*Пример.* Рассмотрим пример из пункта 10.4.1. Допустим, роботу необходимо объехать шесть контрольных точек за наименьшее время. Расстояния от каждой точки до каждой заданы в виде матрицы расстояний (см. рисунки 10.9 и 10.10). Выпишем их в таблицу 11.1.

Таблица 11.1 – Начальные решения

Номер решения	Маршрут						Длина пути	Вероятность выбора
	1	2	3	4	5	6		
1	1	2	3	4	5	6	74	0,26
2	1	3	5	4	2	6	111	0,23
3	6	5	3	2	4	1	47	0,29
4	2	1	6	3	5	4	125	0,22

Далее используем метод рулетки для отбора лучших решений. Рассчитаем вероятности выбора каждого решения по формуле

$$p_s(x) = \left( 1 - \frac{f_s(x)}{\sum_{i=1}^K f_i(x)} \right) / \sum_{s=1}^K \left( 1 - \frac{f_s(x)}{\sum_{i=1}^K f_i(x)} \right), \quad (11.7)$$

где  $K$  – количество начальных решений (размер начальной популяции). Результаты занесем в таблицу 11.1.

С помощью равномерного датчика генерируем 4 случайных числа из интервала  $[0,1]$ . Пусть эти числа будут 0,75; 0,5; 0,20; 0,15. Тогда пул родителей составят номера решений 3, 3, 1, 1. Для скрещивания выберем 1-е и 3-е решения, так как 3-е с 3-м и 1-е с 1-м не приведут к улучшению.

Для скрещивания используем оператор РМХ. Выберем две текущие точки, например 2 и 4, и сформируем потомки (таблица 11.2).

Таблица 11.2 – Кроссинговер РМХ

1-й родитель						2-й родитель					
1	2	3	4	5	6	6	5	3	2	4	1
↓ 1-й потомок ↓						↓ 2-й потомок ↓					
6	5	3	2	4	1	1	2	3	4	5	6
→						←					
6	5	3	4	2	1	1	4	3	2	5	6

Таким образом, мы получили два допустимых решения (два потомка): (6 5 3 4 2 1) и (1 4 3 2 5 6). Длина пути для первого потомка составит 76, для второго – 79 (см. рисунок 10.9).

Далее применим оператор мутации SIM к каждому из потомков. Пусть это будут 3-й и 5-й гены для 1-го потомка и 1-й и 6-й – для 2-го потомка (таблица 11.3).

Таблица 11.3 – Оператор мутации SIM

1-й потомок до мутации						2-й потомок до мутации					
6	5	3	4	2	1	1	4	3	2	5	6
1-й потомок после мутации						2-й потомок после мутации					
6	5	2	4	3	1	6	4	3	2	5	1

Длина пути для первого потомка после мутации составит 97, для второго – 136 (см. рисунок 10.9). Таким образом, мутация ухудшила решения.

## **11.2. Использование генетических алгоритмов для решения социально-экономических задач**

### **11.2.1. Условия эффективного применения генетических алгоритмов**

*Социальная экономика* – это отрасль научного знания, деятельность и сфера экономики, основной целью которой является достижение личностного развития, а также способствование прогрессу человечества [41].

Как правило, с социальной экономикой связывают дополнительно два понятия.

1. Социальное предпринимательство – использование стартапов и других средств предпринимательства для разработки, финансирования и реализации решений социальных, культурных или экологических проблем [42].

2. Социальное инвестирование – практика распределения ресурсов с целью достижения положительного социального эффекта. Социальное инвестирование может относиться как к разделу отказа от вложения в определенные виды активов (алкогольная, табачная, оружейная продукция), так и к инвестициям в развитие социально полезных активов, регионов или внутрикорпоративное развитие (на повышение квалификации сотрудников). О последних типах далее и пойдет речь [43].

Генетические алгоритмы позволяют осуществлять поиск оптимальных решений для целого ряда практических задач. Широта сферы применения ГА обусловлена прежде всего их универсальностью, а также способностью одновременно оптимизировать решение задачи по нескольким критериям [44].

Анализ результатов использования ГА позволяет выделить условия, при выполнении которых задача решается эффективно:

- 1) большое пространство поиска, ландшафт которого является негладким (содержит несколько экстремумов);
- 2) сложно формализуемая функция степени оценки качества решения;
- 3) многокритериальность поиска;
- 4) поиск по заданным критериям приемлемого, а не единственного оптимального решения.

### **11.2.2. Применение генетических алгоритмов для решения задачи прогнозирования развития города**

Задача – спрогнозировать показатели, соответствующие развитию города, на основе независимых переменных, отражающих производство этого города.

В данном примере будет рассмотрен город Уфа. Пусть имеется некий набор параметров (таблица 11.4). Согласно постановке задачи  $Y_1$ ,  $Y_2$ ,  $Y_3$  – показатели, характеризующие рост доходов в бюджет города, сокращение безработицы, рост заработной платы. Показатель  $Y_1$ , представлен в денежном выражении;  $Y_2$  – число зарегистрированных безработных, чел.;  $Y_3$  представлен также в денежном выражении. Данные показатели рассматриваются как зависимые переменные от тех переменных, которые обозначены  $X_i$  и характеризуют производство основных видов продукции в натуральном выражении.

Таблица 11.4 – Зависимые и независимые показатели города

Показатель	Наименование переменной
Доход в бюджет	$Y_1$
Безработица	$Y_2$
Заработная плата	$Y_3$
Добыча нефти	$X_1$
Производство автомобильного бензина	$X_2$
Производство дизельного топлива	$X_3$
Первичная переработка нефти	$X_4$
Производство мазута	$X_5$

### Окончание таблицы 11.4

Показатель	Наименование переменной
Производство пиломатериалов	$X_6$
Производство стеновых материалов	$X_7$
Производство электроэнергии	$X_8$
Производство теплоэнергии	$X_9$
Производство мяса и мясопродуктов АПК	$X_{10}$
Производство хлеба и хлебобулочных изделий АПК	$X_{11}$
Производство молока и молочных продуктов АПК	$X_{12}$
Производство скота и птицы (сельское хозяйство)	$X_{13}$
Производство молока (сельское хозяйство)	$X_{14}$
Капитальные вложения в производственную сферу	$X_{15}$
Производство услуг	$X_{16}$

Решение задачи проводится в несколько этапов.

Этап 1 – предобработка данных. Аналитическое выявление тех показателей  $X_i$ , от которых зависит  $Y_i$  (например, в периоде 2005–2006 гг. показатель  $Y_1$  представлен переменными  $X_7, X_8, X_{13}, X_{16}$ ), и составление целевой функции  $\varphi$ .

Этап 2 – прогнозирование с помощью ГА. Использование генетического алгоритма для выявления наилучшего показателя  $Y_i$  (рисунок 11.8).

Кодирование переменных: пусть длина каждого гена 16 бит и для целевой функции перебираются параметры  $X_1$  и  $X_2$ , тогда получим хромосомную нить из 32 нулей и единиц:

$$N = \underbrace{10001011011101100011000111111011}_{X_1}$$

$$\underbrace{\phantom{10001011011101100011000111111011}}_{X_2}$$

Вероятность отбора родителя пропорциональна функции пригодности:

$$P_i^{\text{отбора}} = \frac{F_i}{\sum_{i=1}^n F_i}$$

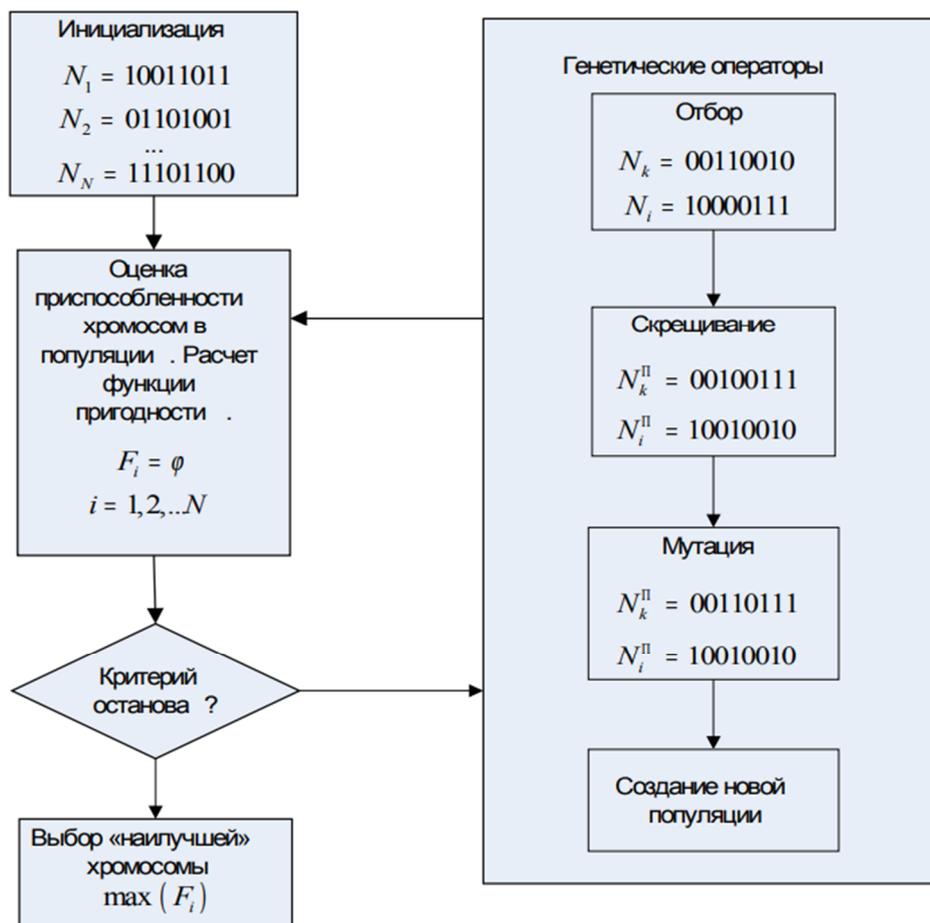


Рисунок 11.8 – Генетический алгоритм в подборе наилучших параметров  $X_i$

Далее работает оператор кроссинговера в случайной точке, например

$$N_1 = 10001110111001001010.010110001101$$

$$N_2 = 01110010101000111101.111111000110$$

Скрещенные особи:

$$N_1^{\Pi} = 10001110111001001010.111111000110$$

$$N_2^{\Pi} = 01110010101000111101.010110001101$$

Мутация:

$$N_1^{\Pi} = 1000111\boxed{0}111001001010.111111000110$$

Мутированная особь:

$$N_1^{\Pi} = 10001111111001001010.111111000110$$

### 11.3. Применение генетических алгоритмов для решения задачи оценки эффективности региональной промышленной политики

Пусть планирование экономической деятельности в виде совокупности инвестиционных проектов (ИП) в регионе осуществляется агентами, заинтересованными в эффективном функционировании региональной экономики. К таким агентам следует отнести производственный сектор, население региона, управляющий их взаимодействием региональный (налоговый) центр и потенциального инвестора – физическое или юридическое лицо любой формы собственности, готовое вложить в развитие региона свои свободные денежные средства (капитал) [45].

*Модель.* Пусть имеется некая сложная модель с ограничениями, описывающая инвестиции:

$$J_{\text{inv}} = \frac{\sum_{k=1}^n \sigma_k x_k + \gamma \sum_{k=1}^n x_{n+k}}{1+r} - x_{2n+1} - x_{2n+2} - \varepsilon_2 x_{2n+4} \rightarrow \max,$$

$$J_{\text{cons}} = \frac{\beta \sum_{k=1}^n x_{n+k}}{1+r} - \varepsilon_1 x_{2n+3} \rightarrow \max,$$

$$J_{\text{tax}} = \frac{-\sum_{k=1}^n \sigma_k x_k + (\alpha_3 - \beta(\alpha_3 - \alpha_4)) \sum_{k=1}^n x_{n+k}}{1+r} - (1 - \varepsilon_1) x_{2n+3} - (1 - \varepsilon_2) x_{2n+4} \rightarrow \max,$$

$$\sum_{k=1}^n \gamma_k x_k + \gamma \sum_{k=1}^n x_{n+k} + x_{2n+1} + x_{2n+2} + x_{2n+4} \geq 0,$$

$$-\sum_{k=1}^n \theta_k x_k + \rho \sum_{k=1}^n x_{n+k} \geq 0,$$

$$\beta \sum_{k=1}^n x_{n+k} + x_{2n+1} + M_0 - \sum_{k=1}^n q_k \geq 0,$$

$$\begin{aligned}
& -\sum_{k=1}^n \sigma_k x_k + (\alpha_3 - \beta(\alpha_3 - \alpha_4)) \sum_{k=1}^n x_{n+k} + \\
& + (1 - \varepsilon_1) x_{2n+3} + (1 - \varepsilon_2) x_{2n+4} + N_0 \geq 0, \\
& x_{n+k} \leq q_k, x_{n+k} \leq \delta_k x_k \quad (k = 1, \dots, n), x_{2n+k} \leq I_0, \\
& x_{2n+2} \leq K_0, x_{2n+3} + x_{2n+4} \leq L_0, x_k \geq 0 \quad (k = 1, \dots, 2n+4), \\
& W = (1 - \varepsilon_3)(1 - \beta)R - (1 - \varepsilon_3)(Am + N_2) = \\
& = (1 - \varepsilon_3) \left[ -\sum_{k=1}^n \theta_k x_k + (1 - \beta) \sum_{k=1}^n x_{n+k} \right] \geq 0,
\end{aligned}$$

где  $J_{\text{inv}}$ ,  $J_{\text{const}}$ ,  $J_{\text{tax}}$  – критерии инвестора, потребителя и управляющего центра;

$\delta_k = P_k V_k / c_k$  – фондоотдача основных производственных фондов (ОПФ)  $k$ -го направления деятельности;

$c_k$  – стоимость ОПФ  $k$ -го типа;

$V_k$  – проектная производительность ОПФ  $k$ -го типа;

$P_k$  – стоимость единицы продукции  $k$ -го типа;

$x_k = c_k m_k$  ( $k = 1, \dots, n$ ) – общая стоимость ОПФ, приобретаемых в ИП;

$m_k$  – количество приобретаемых ОПФ  $k$ -го типа;

$x_{n+k} = P_k m_k y_k$  ( $k = 1, \dots, n$ ) – выручка от реализации продукции  $k$ -го типа;

$y_k$  – объем выпуска  $k$ -го вида продукции;

$R = \sum_{k=1}^n P_k m_k x_k = x_{n+k}$  – выручка от реализации по всем видам про-

дукции;

$q_k$  – прогнозный спрос на продукцию  $k$ -го типа;

$Am$  – амортизационные отчисления;

$$Am(T) = T \sum_{k=1}^n \frac{c_k m_k}{T_k} = \sum_{k=1}^n \frac{T}{T_k} x_k - \text{сумма амортизационных отчисле-$$

ний по всем видам ОПФ;

$T_k$  – срок службы ОПФ  $k$ -го типа;

$T$  – срок действия ИП;

$K_0$  – начальный собственный капитал предприятия;

$M_0$  – начальные собственные средства потребителей;

$N_0$  – начальные собственные средства налогового центра;

$N_2 = \alpha_2 S^0$  – налог на имущество;

$$S^0 = \sum_{k=1}^n \left(1 - \frac{T}{T_k}\right) c_k m_k = \sum_{k=1}^n \left(1 - \frac{T}{T_k}\right) x_k - \text{остаточная стоимость}$$

ОПФ;

$L_0$  – максимальная сумма дотаций за весь период действия ИП;

$I$  – внешние инвестиции;

$\bar{I}$  – внутренние инвестиции;

$I_0$  – максимальная сумма внешних инвестиций;

$W$  – прибыль;

$\alpha_1, \alpha_2, \alpha_3, \alpha_4$  – ставки налогов на добавленную стоимость (НДС), на имущество (НИ), на прибыль (НП), на доходы физических лиц (НДФЛ);

$\varepsilon_1, \varepsilon_2$  – проценты возврата дотаций производителям и потребителям;

$\beta$  – часть выручки, поступающая в фонд оплаты труда;

$r$  – ставка дисконтирования, учитывающая инфляцию, требования инвестора по доходности и различные риски проекта;

$x_{2n+1} = I$  – внешние (возвратные) инвестиции (инвестора);

$x_{2n+2} = \bar{I}$  – внутренние инвестиции производителя;

$x_{2n+3} = \text{Dot}_1$  – дотации потребителям;

$x_{2n+4} = \text{Dot}_2$  – дотации производителям;

$$\gamma_k = (2\alpha_2 + \alpha_3 - \alpha_2\alpha_3) \frac{T}{T_k} + \alpha_2\alpha_3 - 2\alpha_2 - 1;$$

$$\sigma_k = (\alpha_2 + \alpha_3 - \alpha_2\alpha_3) \frac{T}{T_k} + \alpha_2(1 - \alpha_3);$$

$$\theta_k = (1 - \alpha_2) \frac{T}{T_k} - \alpha_2, k = 1, \dots, n;$$

$$\rho = 1 - \alpha_1 - \beta, \gamma = (1 - \alpha_3)\rho;$$

$$\tau = \rho\alpha_3 + \beta\alpha_4.$$

Требуется найти такой общий объем инвестиций, объем инвестиций в приобретение ОПФ указанных видов (отраслей) экономической деятельности, а также планируемую выручку от реализации продукции, чтобы чистый приведенный доход (NPV) инвестиционного проекта по развитию экономической деятельности в регионе был максимальным.

С использованием региональной экономической статистики были рассмотрены четыре отрасли экономической деятельности Красноярского края [45]:

- 1) добыча полезных ископаемых;
- 2) обрабатывающие производства;
- 3) производство и распределение электроэнергии, газа и воды;
- 4) строительство.

Данная задача была решена симплекс-методом и ГА (таблица 11.5) [46].

В случае линейных ограничений ГА и симплекс-метод показывают примерно одинаковые результаты.

Симплекс-метод применим только в случае присутствия линейных ограничений. Если ограничение вида

$$x_{n+k} \leq q_k, k = 1, \dots, n,$$

заменяется на

$$x_{n+k} \leq \frac{k_i}{1 + b_i e^{\alpha_i / (\beta_i x_{n+k})}},$$

то задача этим методом уже не решается и необходимо применять другие методы, одним из которых является генетический алгоритм SPEA [46].

Таблица 11.5 – Результаты расчетов симплекс-методом и модифицированным алгоритмом SPEA для линейных ограничений

$x$	Симплекс-метод	ГА (SPEA)
$x_1$	72 848,31	73 007,44
$x_2$	84 774,68	85 331,24
$x_3$	129 242,2	135 649,19
$x_4$	57 115,08	58 170,4
$x_5$	28 046,6	28 016,63
$x_6$	438 878,5	438 763,51
$x_7$	26 623,9	27 934,7099
$x_8$	24 616,6	24 575,99
$x_9$	0	0
$x_{10}$	0	0
$x_{11}$	0	0
$x_{12}$	0	0
Значение критерия	326	326
	669,590	208,370
	6	8

#### 11.4. Формирование системы прогнозирующих правил в деятельности страховых компаний

Генетические алгоритмы могут использоваться для нахождения оптимального набора правил, позволяющих прогнозировать страховой риск с учетом ряда определяющих его факторов. Для решения этой задачи необходимо иметь базу данных, содержащую фактические значения переменных, влияющих на страховой риск.

Рассмотрим пример оптимизации экспертных правил в сфере страхования с помощью генетического алгоритма. Допустим, что компания, занимающаяся страхованием автомобилей, использует базу данных (БД), которая включает следующие факторы: максимальную скорость автомобиля (км/ч), возраст автомобиля (лет), возраст водителя (лет), риск, определенный экспертно по некоторой шкале на основе анализа обращений клиентов по страховым случаям. Правила,

задающие оценку страхового риска, имеют вид «ЕСЛИ максимальная скорость автомобиля лежит в диапазоне  $[A_i]$  И возрастной диапазон автомобиля  $[B_i]$  И возраст водителя находится в диапазоне  $[C_i]$ , ТО страховой риск имеет значение  $[D_i]$ ».

Для конкретной выборки из БД это правило может иметь следующий вид «ЕСЛИ максимальная скорость  $[91-100 \text{ км/ч}]$  И возраст автомобиля  $[11-15 \text{ лет}]$  И возраст водителя  $[31-40 \text{ лет}]$ , ТО риск  $[3]$ ». Здесь уровень риска отображается на интервал  $[1, 5]$ , при этом высокие значения соответствуют большим страховым рискам.

Подобные правила, основанные на фактических значениях переменных, случайным образом выбранных из БД, составляют исходную популяцию. Для каждой из переменных, входящих в популяцию, предварительно задается диапазон состояний. Например, переменная «возраст автомобиля» может иметь пять возможных состояний: 1–5, 6–10, 11–15, 16–20, 21–25 лет. Далее сформированная популяция обрабатывается генетическими операторами с учетом специфики рассматриваемой задачи. Целевая функция должна показывать, насколько точно сгенерированные правила описывают реальные страховые случаи, хранящиеся в БД. Например, если какое-то правило описывает 4 случая из 5, то значение целевой функции будет  $4/5$ , или 80 %.

Новые члены популяции образуются в результате скрещивания и мутации начального набора правил. В данном случае при скрещивании двух правил происходит обмен парами «атрибут – значение» на участке строки после точки кроссинговера.

*Справка.* В базах данных *атрибут* представляет собой свойство сущности. Так, атрибутами сущности «Ассортимент товара» являются его код, модель, название, поставщик, стоимость и т.д.

В результате образуются два новых правила, жизнеспособность которых оценивается по тому, насколько удачно они описывают страховые случаи, имевшие место в прошлом. Мутация правил обеспечивает необходимое разнообразие признаков и заключается в изменении значений атрибутов с заданной вероятностью. Таким образом, первоначально сформированный набор правил преобразуется случайно

направленным способом в другой набор, который лучше остальных описывает накопленную статистику страховых случаев. Результирующая система правил в дальнейшем используется для прогнозирования страховых рисков.

Следует отметить, что такой подход к формированию системы правил может приводить к некорректным правилам продукций. В то же время он освобождает разработчиков и экспертов от трудоемкой работы по формулированию и оценке правил, так как некорректные результаты отбрасываются при сопоставлении сгенерированных продукций с реальными страховыми ситуациями. Привлечение прошлого опыта для оценки пригодности прогнозирующих правил не позволяет предвидеть новые ситуации. Поэтому при решении задач описанным способом очень важно следить за своевременным пополнением и модификацией информации в БД, чтобы учитывать появление новых фактов, атрибутов и тенденций.

### **Вопросы для самопроверки**

1. В чем заключается задача коммивояжера?
2. Детерминированные алгоритмы решения задачи коммивояжера?
3. Опишите генетический алгоритм решения задачи.

## Заключение

Несмотря на то что методы оптимизации как самостоятельное направление сложились еще в середине 1960-х гг., их развитие продолжается по сей день. Поэтому в одном издании невозможно изложить все существующие классы задач и методы их решения.

В данном учебном пособии, к примеру, не рассматривается теория оптимального управления, а среди всего многообразия эвристических методов оптимизации подробно представлены только генетические алгоритмы.

Оптимизационные задачи встречаются повсюду – в программировании, экономике, транспорте – практически во всех отраслях человеческой деятельности. Специалист должен уметь распознавать, когда стоящая перед ним задача относится к классу оптимизационных, и выбирать адекватные методы ее решения. Коллектив авторов надеется, что предлагаемое учебное пособие поможет в этом.

## Список сокращений и условных обозначений

ВИ	– вариационное исчисление
ГА	– генетический алгоритм
ГУ	– граничные условия
ДП	– динамическое программирование
ДУ	– достаточные условия
ЗКП	– задача квадратичного программирования
ЗЛП	– задача линейного программирования
ЗНП	– задача нелинейного программирования
ЗО	– задача оптимизации
ЗЦП	– задача целочисленного программирования
ИИС	– искусственная иммунная система
КП	– квадратичное программирование
КФ	– квадратичная форма
ЛП	– линейное программирование
МДНЗ	– метод допустимых направлений Зойтендейка
МЛ	– метод линеаризации
ММЛ	– метод множителей Лагранжа
МО	– методы оптимизации
МПГ	– метод проекции градиента
МУГ	– метод условного градиента
МФВ	– метод Франка – Вульфа
МШФ	– методы штрафных функций
НП	– нелинейное программирование
НУ	– необходимые условия
ОДР	– область допустимых решений
ТАУ	– теория автоматического управления
ЦФ	– целевая функция
ШФ	– штрафная функция
ЭА	– эвристический алгоритм
ЭМ	– эвристический метод
СИМ	– перестановки относительно центра
ИМ	– мутация вставки
ОХ	– Order Crossover, упорядоченный кроссинговер

PMX – Partially-Mapped Crossover, частично соответствующий кроссинговер

SIM – простая мутация перестановки

SM – мутация перестановки

$A( )$  – декартовы координаты точки, например  $A(x, y)$  – точка  $A$  с декартовыми координатами  $(x, y)$

const – постоянная величина, константа

$E$  – единичная матрица

extr – экстремум

$\bar{f}$  – значение функции в критической или стационарной точке

$f^*$  – экстремальное значение функции  $f$

$f'(x) = df / dx$  – производная функции скалярного аргумента  $f(x)$

$f''(x) = d^2 f / dx^2$  – вторая производная функции скалярного аргумента  $f(x)$

$f^{(n)}(x) = d^n f / dx^n$  – производная функции скалярного аргумента  $f(x)$  порядка  $n$

$f'_{x_i} = \partial f / \partial x_i$  – частная производная функции векторного аргумента  $f(x)$  по аргументу  $x_i$

$f''_{x_i} = \partial^2 f / \partial x_i^2$  – вторая частная производная функции векторного аргумента  $f(x)$  по аргументу  $x_i$

$f''_{x_i x_j} = \partial^2 f / \partial x_i \partial x_j$  – вторая частная производная функции векторного аргумента  $f(x)$  по аргументам  $x_i$  и  $x_j$

$\text{grad } f(x) = \nabla f(x)$  – вектор-градиент функции

$H_f(x) = \nabla^2 f(x)$  – матрица Гессе (матрица вторых производных функции)

inf – точная нижняя граница множества

$L$  – функция Лагранжа

max – максимум

min – минимум

$Q[a,b]$  – множество унимодальных функций, заданных на отрезке  $[a,b]$

$R^n$  – линейное пространство размерности  $n$

$S$  – область допустимых решений

$\sup$  – точная верхняя граница множества

$\bar{x}$  – критическая или стационарная точка

$x^*$  – экстремальное значение аргумента  $x$

$Z$  – множество целых чисел

$\Delta$  – приращение ( $\Delta x$  – приращение аргумента  $x$ )

$\Sigma$  – сумма

$\Lambda$  – константа Липшица

$\in$  – знак принадлежности ( $a \in A$  – элемент  $a$  принадлежит множеству  $A$ ,  $x \in [a,b]$  – точка  $x$  принадлежит отрезку  $[a,b]$ )

$\notin$  – не принадлежит

$\subset$  – знак включения, подмножества ( $A \subset B$  – множество  $A$  является подмножеством  $B$ )

$( )$  – открытый интервал ( $(a,b)$  – интервал, которому принадлежат все значения от  $a$  до  $b$ , но не включая  $a$  и  $b$ )

$[ ]$  – закрытый интервал или отрезок ( $[a,b]$  – отрезок, которому принадлежат все значения от  $a$  до  $b$  включительно)

$\Rightarrow$  – знак следования ( $a \Rightarrow b$  – из  $a$  следует  $b$ )

$\rightarrow$  – стремится ( $a \rightarrow b$  –  $a$  стремится к  $b$ )

$\{ \}$  – знак множества ( $\{a,b\}$  – множество, состоящее из элементов  $a$  и  $b$ )

$\emptyset$  – пустое множество

$\forall$  – квантор общности ( $\forall x$  – для всех/любого/каждого  $x$ )

$\exists$  – квантор существования ( $\exists x$  – найдется/существует  $x$ )

$| |$  – модуль или определитель ( $|a|$  – модуль скалярной величины  $a$ ;  $|A|$  – определитель матрицы  $A$ )

$\| \|$  – норма матрицы или вектора

$T$  – оператор транспонирования матрицы или вектора

$+\infty$  – бесконечно большое положительное число

$-\infty$  – бесконечно большое отрицательное число

## Список литературы

1. Реклейтис Г. Оптимизация в технике : в 2 кн. / Г. Реклейтис, А. Рейвиндран, К. Рэгсдел. – М. : Мир, 1986. – 320 с.
2. Мину М. Математическое программирование. Теория и алгоритмы / М. Мину. – М. : Наука, 1990. – 488 с.
3. Рубан А.И. Оптимизация систем : учеб. пособие / А.И. Рубан. – Томск : Том. гос. ун-т, 1984.
4. Дэннис Дж. Численные методы безусловной оптимизации и решения нелинейных уравнений / Дж. Дэннис, Р. Шнабель. – М. : Мир, 1988. – 440 с.
5. Сборник задач по математике для втузов. Ч. 4. Методы оптимизации / Э.А. Вуколов [и др.] ; под ред. А.В. Ефимова. – М. : Наука, 1990. – 302 с.
6. Жермен Лакур П. Математика и САПР : в 2 кн. / Лакур П. Жермен, П.Л. Жорж, Ф. Пистр, П. Безье. – М. : Мир, 1989. – 262 с.
7. Химмельблау Д. Прикладное нелинейное программирование / Д. Химмельблау. – М. : Мир, 1991.
8. Полак Э. Численные методы оптимизации. Единый подход / Э. Полак. – М. : Мир, 1994.
9. Уайлд Д.Дж. Методы поиска оптимума / Д.Дж. Уайлд. – М. : Наука, 1997.
10. Коршунов Ю.М. Математические основы кибернетики : учеб. пособие для втузов / Ю.М. Коршунов. – М. : Энергоатомиздат, 1987. – 496 с.
11. Рубан А.И. Методы оптимизации : учеб. пособие / А.И. Рубан. – Изд. 2-е, испр. и доп. – Красноярск : НИИ ИПУ, 2001.
12. Беллман Р. Динамическое программирование / Р. Беллман. – М. : Иностранная литература, 1960. – 400 с.
13. Кремер Н.Ш. Исследование операций и методы оптимизации : учеб. пособие для вузов / Н.Ш. Кремер. – М. : Банки и биржи, ЮНИТИ, 1997.
14. Романовская А.М. Динамическое программирование: учеб. пособие / А.М. Романовская, М.В. Мендзив. – Омск : Изд-во Омского института (филиала) РГТЭУ, 2010. – 58 с.

15. Васин А.А. Исследование операций : учеб. пособие для студентов вузов / А.А. Васин, П.С. Краснощеков, В.В. Морозов. – М. : Академия, 2008. – 464 с.
16. Акулич Л.И. Математическое программирование в примерах и задачах : учеб. пособие для студентов экон. специальностей вузов / Л.И. Акулич. – М. : Высшая школа, 1986.
17. Базара М. Нелинейное программирование. Теория и алгоритмы / М. Базара, К. Шетти. – М. : Мир, 1982. – 584 с.
18. Есипов Б.А. Методы исследования операций: учеб. пособие / Б.А. Есипов. – СПб. : Лань, 2010. – 256 с.
19. Генетические алгоритмы, искусственные нейронные сети и проблемы виртуальной реальности / Г.К. Вороновский, К.В. Махотило, С.Н. Петрашев, С.А. Сергеев. – Х. : Основа, 1997. – 112 с.
20. Емельянов В.В. Теория и практика эволюционного моделирования / В.В. Емельянов, В.В. Курейчик, В.М. Курейчик. – М. : Физматлит, 2003. – 432 с.
21. Батищев Д.А. Генетические алгоритмы решения экстремальных задач / Д.А. Батищев. – Воронеж : ВГТУ, 1995. – 69 с.
22. Кафаров В.В. Основы автоматизированного проектирования химических производств / В.В. Кафаров, В.Н. Ветехин. – М. : Наука, 1987. – 624 с.
23. Sörensen Kenneth, Glover, Fred. Metaheuristics // Encyclopedia of Operations Research and Management Science. – 2013. – P. 960–970.
24. Пантелеев А.В. Методы глобальной оптимизации. Метаэвристические стратегии и алгоритмы / А.В. Пантелеев, Д.В. Метлицкая, Е.А. Алешина. – М. : Вузовская книга, 2013. – 244 с.
25. Курейчик В.М. Эволюционные методы принятия решений / В.М. Курейчик, В.В. Курейчик // Известия ЮФУ. Технические науки. – 2002. – № 1. – С. 86–89.
26. Зайцев А.А. Обзор эволюционных методов оптимизации на основе роевого интеллекта / А.А. Зайцев, В.В. Курейчик, А.А. Полупанов // Известия ЮФУ. Технические науки. – 2010. – № 12. – С. 7–12.

27. Kennedy J. Swarm Intelligence / J. Kennedy // Handbook of Nature-Inspired and Innovative Computing. – Springer, Boston, MA. – 2006. – P. 187–219.

28. Матренин П.В. Методы стохастической оптимизации : учеб. пособие / П.В. Матренин, М.Г. Гриф, В.Г. Секаев. – Новосибирск : Изд-во НГТУ, 2016. – 67 с.

29. Панченко Т.В. Генетические алгоритмы: учебно-методическое пособие / Т.В. Панченко; под ред. Ю. Ю. Тарасевича. – Астрахань : Астраханский университет, 2007. – 87 с.

30. Brownlee J. Clever Algorithms: Nature-Inspired Programming Recipes / J. Brownlee. – Lulu.com, 2011. – 438 p.

31. Multiobjective Variable Mesh Optimization / Salgueiro Yamisleydi, Toro Jorge, Bello Rafael, Falcon Rafael // Annals of Operations Research. – 2017. – No. 258(2). – Pp. 869–893.

32. Boettcher Stefan. Extremal optimization: an evolutionary local-search algorithm / Stefan Boettcher, G. Percus Allon // Computational Modeling and Problem Solving in the Networked World. – Springer, Boston, MA, 2003. – P. 61–77.

33. Resende M.G.C. Greedy Randomized Adaptive Search Procedures / M.G.C. Resende, C.C. Ribeiro // Handbook of Metaheuristics. International Series in Operations Research & Management Science. – Springer, Boston, MA. – 2003. – Vol. 57. – P. 219–249.

34. Кулиев Э.В. Роевой алгоритм поисковой оптимизации на основе моделирования поведения летучих мышей / Э.В. Кулиев, А.А. Лежебоков, Ю.А. Кравченко // Известия ЮФУ. Технические науки. – 2016. – № 7. – С. 53–62.

35. Yang X.S. Firefly algorithms for multimodal optimization / X.S. Yang // The 5th Symposium on Stochastic Algorithms, Foundations and Applications. – 2009. – P. 169–178.

36. Research on Improved Strategy of Shuffled Frog Leaping Algorithm / Zhen Wang, Danhong Zhang, Biao Wang, Wenwen Chen // Chinese Association of Automation (YAC) 34rd Youth Academic Annual Conference. – 2019. – P. 265–268.

37. Chu S.-A. Cat swarm optimization / S.-A. Chu, P.-W. Tsai, J.-S. Pan // Lecture Notes in Computer Science (including subseries

Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). – 2006. – P. 854–858.

38. Mucherino A. Monkey search: a novel metaheuristic search for global optimization / A. Mucherino, O. Seref // Data Mining, Systems Analysis and Optimization in Biomedicine. – 2007. – Vol. 953. – P. 162–173.

39. Bishop J.M.. Stochastic searching networks / J.M. Bishop // Artificial Neural Networks, First IEE International Conference. – 1989. – P. 329–331.

40. Circle detection using electromagnetism optimization / E. Cuevas, D. Oliva, D. Zaldivar, M.P. Cisneros [et al.] // Information Sciences. – 2012. – No. 182(1). – P. 40–55.

41. Социальная экономика [Электронный ресурс]. – Режим доступа: <https://utmagazine.ru/posts/9255-socialnaya-ekonomika>.

42. Социальное предпринимательство [Электронный ресурс]. – Режим доступа: [https://ru.wikipedia.org/wiki/Социальное\\_предпринимательство](https://ru.wikipedia.org/wiki/Социальное_предпринимательство).

43. Социальное инвестирование [Электронный ресурс]. – Режим доступа: [https://ru.wikipedia.org/wiki/Социальное\\_инвестирование](https://ru.wikipedia.org/wiki/Социальное_инвестирование).

44. Белобородова Н.А. Модели и методы прогнозирования развития муниципального образования в нефтегазовом районе России / Н.А. Белобородова // Нефтегазовое дело. – 2010. – № 2. – С. 77–92.

45. Горбунов М.А. Применение генетических алгоритмов для решения задачи оценки эффективности региональной политики промышленности / М.А. Горбунов, А.В. Медведев, Е.С. Семенкин // Вестн. Сибир. гос. ун-та им. акад. М.Ф. Решетнева. – 2011. – № 4. – С. 8–12.

46. Zitzler E. Muptiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach / E. Zitzler, L. Thiele // IEEE Trans. on Evolutionary Computation. – San Diego. – 1999. – P. 257–271.

# Приложение А

## Краткий математический справочник

### П.А.1. Приближенные числа

**Точные и приближенные числа.** Числа, встречающиеся на практике, бывают *точными* и *приближенными*. Точное число  $A$  дает истинное значение величины, приближенное число  $a$  – только приближительное значение. При выполнении различных действий над числами записать точное число иногда не представляется возможным (например, когда результат вычислений является периодической дробью или иррациональным числом). В этом случае удобно пользоваться приближенными числами.

**Округление.** Одним из способов получения приближенных чисел является округление. Округлять можно как точные, так и приближенные числа.

*Округлением* данного числа до разряда  $n$  называют замену его новым числом, которое получается из данного путем отбрасывания всех его цифр, записанных правее цифры этого разряда, или путем замены их нулями. При округлении учитывают следующее:

- если первая (слева) из отбрасываемых цифр менее 5, то последнюю оставленную цифру не изменяют (округление с *недостатком*);
- если первая отбрасываемая цифра больше или равна 5, то последнюю оставленную цифру увеличивают на 1 (округление с *избытком*).

#### *Примеры*

1. Округлим до десятых число  $A = 12,34$ . Ответ:  $a \approx 12,3$ .
2. Округлим до сотых число  $A = 3,2465$ . Ответ:  $a \approx 3,25$ .
3. Округлим до тысячных число  $A = 3,4335$ . Ответ:  $a \approx 3,434$ .
4. Округлим до тысяч число  $A = 12347$ . Ответ:  $a \approx 12000$ .

**Абсолютная и относительная погрешности.** Разность между точным числом и его приближенным значением называется *абсолютной погрешностью* (АП) приближенного числа:

$$\Delta = |A - a|. \quad (\text{П.А.1})$$

Например, если точное число  $A = 1,214$  округлить до десятых, получим приближенное число  $a = 1,2$ . В данном случае абсолютная погрешность приближенного числа  $|1,214 - 1,2| = 0,014$ . Но в большинстве случаев известно только приближенное значение рассматриваемой величины. Тогда абсолютную погрешность определить невозможно, в связи с чем указывают границу, которую она не превышает. Это число называют *граничной абсолютной погрешностью* (ГАП). ГАП приближенного числа  $a$  обозначают символом  $\Delta a$  ( $\Delta a \geq 0$ ). Говорят, что точное значение числа равно его приближенному значению с погрешностью, меньшей, чем ГАП:

$$A \approx a \pm \Delta a. \quad (\text{П.А.2})$$

Таким образом, точное значение величины  $A$  находится в промежутке между числами  $a - \Delta a$  и  $a + \Delta a$ , которые называют соответственно *нижней* и *верхней границами* числа  $A$ .

Если АП приближенного числа  $a$  не превышает единицы разряда, выражаемого  $n$ -й значащей цифрой в десятичной записи этого числа, то  $a$  называется числом, имеющим  $n$  *верных знаков в широком смысле*. Если же АП не превышает половины единицы разряда, выражаемого  $n$ -й значащей цифрой, то приближенное число  $a$  называется числом, имеющим  $n$  *верных знаков в узком смысле*.

*Пример.* Точная запись величины

$$A = \sqrt{5} = 2,236067978\dots$$

является иррациональным числом. Пусть  $a_1 = 2,236$  – приближенное значение числа  $A$ . Тогда в узком смысле число  $a_1$  имеет четыре *верных знака*:

$$A = a_1 \pm 0,0005.$$

Пусть  $a_2 = 2,24$ , тогда в широком смысле число  $a_2$  имеет три *верных знака*:

$$A = a_2 \pm 0,01.$$

АП или ГАП не характеризует качество выполненного измерения. Например, если найдено приближенное значение  $a = 11,999$  точного числа  $A = 12$ , то АП  $\Delta = |12 - 11,999| = 0,001$  имеет достаточно малый порядок. Но если с такой же погрешностью найдено приближенное значение  $a = 0,002$  точного числа  $A = 0,003$ , то очевидно, что в этом случае погрешность слишком велика. В худшем случае величина погрешности может превышать значение измеряемой величины. Поэтому чаще всего целесообразно использовать относительную оценку погрешности.

*Относительной погрешностью* (ОП) приближенного числа называется отношение АП к величине приближенного числа:

$$\delta = \frac{\Delta}{|a|} = \left| \frac{A - a}{a} \right|. \quad (\text{П.А.3})$$

Отношение ГАП к приближенному числу называют *границей относительной погрешностью* (ГОП):

$$\delta = \frac{\Delta a}{|a|}. \quad (\text{П.А.4})$$

ОП и ГОП чаще всего выражают в процентах. Для этого величину ОП или ГОП необходимо умножить на 100 %. Например, если  $A = 12$ ,  $a = 11,999$ , то  $\delta = 0,008 \%$ , а для  $A = 0,003$ ,  $a = 0,002$  получим  $\delta = 50 \%$ . Хотя, как было показано выше, АП этих измерений совпадают.

*Пример.* Пусть  $a_1 = 2,236$  – приближенное значение числа  $A = \sqrt{5}$ . Тогда ОП

$$\delta = \left| \frac{A - a}{a} \right| = \left| \frac{2,236067978... - 2,236}{2,236} \right| \cdot 100 \% \approx 0,003 \%$$

Для ГАП в узком смысле ГОП составит

$$\delta = \frac{\Delta a}{|a|} = \frac{0,0005}{2,236} \cdot 100 \% \approx 0,022 \%$$

Найдем ОП для  $a_2 = 2,24$ :

$$\delta = \left| \frac{A - a}{a} \right| = \left| \frac{2,236067978... - 2,24}{2,24} \right| \cdot 100 \% \approx 0,176 \%.$$

Для ГАП в широком смысле ГОП составит

$$\delta = \frac{\Delta a}{|a|} = \frac{0,01}{2,24} \cdot 100 \% \approx 0,446 \%.$$

Для относительной погрешности справедливы следующие неравенства. В узком смысле

$$\delta \leq \left( \frac{1}{\alpha_m} \right) \cdot 10^{-n+1}, \quad (\text{П.А.5})$$

где  $\alpha_m$  – первая значащая цифра числа  $a$ ;  $n$  – количество верных знаков. В широком смысле

$$\delta \leq \left( \frac{2}{\alpha_m} \right) \cdot 10^{-n+1}. \quad (\text{П.А.6})$$

Для решения обратной задачи – определения количества верных знаков  $n$  числа  $a$ , если известна его относительная погрешность  $\delta$ , можно воспользоваться выражением (П.А.5) в форме

$$N = 1 - \lg(\alpha_m \cdot \delta) \quad (\text{П.А.7})$$

и в качестве  $n$  взять ближайшее к  $N$  целое число.

## П.А.2. Уравнения и неравенства

### Решение уравнений

Два выражения, соединенные знаком « $\Leftrightarrow$ », образуют *равенство*. В математике равенства употребляются в двух случаях.

1. Когда утверждают, что данные выражения равны. При этом равенство называется *тождеством*. Если тождество выполняется при любых значениях переменных, входящих в выражения, либо выражения не содержат переменных, равенство можно заменить знаком тождественного равенства « $\Leftrightarrow$ ».

2. Когда хотят выяснить, при каких условиях выражения равны. В этом случае имеем *уравнение*. Переменные, значения которых нужно определить, решая уравнение, называются *неизвестными*. Переменные, не являющиеся неизвестными, называются *параметрами*. Числа, входящие в уравнение, называются *константами*.

**Понятие решения уравнения.** Значения неизвестных, при которых уравнение обращается в тождество, называются *решением* уравнения. Если в уравнении одно неизвестное, то решение также называется *корнем* уравнения.

### *Примеры*

1. Если в уравнение  $3x + 7 = 13$  вместо неизвестного  $x$  подставить его значение 2, то получим тождество  $3 \cdot 2 + 7 = 13$ . Следовательно,  $x = 2$  – решение (корень) данного уравнения.

2. Уравнение  $2x + c = 3$ , где  $x$  – неизвестное;  $c$  – параметр, имеет решение  $x = 0,5(3 - c)$ , так как  $2 \cdot 0,5 \cdot (3 - c) + c = 3$  – тождество.

3. Уравнение  $x^2 + 2y = 7$  с двумя неизвестными  $x$  и  $y$  имеет множество решений. Пара значений  $x = 1$  и  $y = 3$  удовлетворяет этому уравнению, так как  $1^2 + 2 \cdot 3 = 7$ . То есть эта пара является одним из решений данного уравнения. Существуют и другие решения, например  $x = 1$ ,  $y = 3$ ;  $x = 3$ ,  $y = -1$  и т.д.

4. Уравнение  $x + 8 = x + 5$  не имеет решений (корней), так как при любых значениях неизвестного  $x$  левая часть больше правой части.

5. Уравнение  $5(x - 3) + 2 = 3(x - 4) + 2x - 1$  является тождеством при любом значении неизвестного  $x$ . Следовательно, его решением является любое рациональное значение  $x$ .

**Способы решения уравнений.** Способы решения различных уравнений приведены в [18]. Рассмотрим некоторые из них.

1. Решение (корень) линейного уравнения  $ax + b = 0$  есть

$$x = -\frac{b}{a}. \quad (\text{П.А.8})$$

2. Квадратное уравнение  $ax^2 + bx + c = 0$  имеет решения (корни)

$$x = \frac{-b \pm \sqrt{D}}{2a}, \quad (\text{П.А.9})$$

где  $D = b^2 - 4ac$  – так называемый *дискриминант*. Если  $D > 0$ , уравнение имеет два решения, если  $D = 0$ , решение одно, если  $D < 0$ , то уравнение не имеет рациональных решений.

3. Кубическое уравнение, а также уравнения с полиномами более высоких степеней в правой части, т.е. уравнения вида

$$f(x) = 0, \quad (\text{П.А.10})$$

где  $f(x)$  – полином степени  $n$ :

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1} + a_nx^n, \quad a_n \neq 0,$$

в общем случае не имеют аналитического решения.

Разделим обе части уравнения (П.А.10) на  $a_n$ :

$$\frac{a_0}{a_n} + \frac{a_1}{a_n}x + \frac{a_2}{a_n}x^2 + \dots + \frac{a_{n-1}}{a_n}x^{n-1} + x^n = 0$$

или

$$b_0 + b_1x + b_2x^2 + \dots + b_{n-1}x^{n-1} + x^n = 0, \quad (\text{П.А.11})$$

где  $b_i = a_i/a_n$ ,  $i = 0, 1, 2, \dots, n-1$ . Пусть уравнение (П.А.11) имеет решения (корни)  $x_1, x_2, \dots, x_n$ . Тогда уравнение (П.А.10) можно переписать в виде

$$(x - x_1)(x - x_2) \dots (x - x_n) = 0. \quad (\text{П.А.12})$$

Раскроем скобки:

$$\begin{aligned} x^n - \left( \sum_i x_i \right) x^{n-1} + \left( \sum_{i \neq j} x_i x_j \right) x^{n-2} - \left( \sum_{i \neq j \neq k} x_i x_j x_k \right) x^{n-3} + \\ + \dots + (-1)^{n-1} \left( \sum_i \prod_{i \neq j} x_j \right) x + (-1)^n \left( \prod_i x_i \right) = 0. \end{aligned} \quad (\text{П.А.13})$$

Например,

$$(x-p)(x-q) = x^2 - (p+q)x + pq,$$

$$(x-p)(x-q)(x-r) = x^3 - (p+q+r)x^2 + (pq+pr+qr)x - pqr,$$

и т.д. В таком случае, приравнявая коэффициенты выражений (П.А.11) и (П.А.13) при одинаковых степенях неизвестного  $x^i$ , получим систему уравнений

$$\left\{ \begin{array}{l} \sum_i x_i = -b_{n-1}, \\ \sum_{i \neq j} x_i x_j = +b_{n-2}, \\ \sum_{i \neq j \neq k} x_i x_j x_k = -b_{n-3}, \\ \dots \\ \sum_i \prod_{i \neq j} x_j = (-1)^{n-1} b_1, \\ \prod_i x_i = (-1)^n b_0. \end{array} \right. \quad (\text{П.А.14})$$

Корни уравнения (П.А.10) являются решениями системы (П.А.14). Однако решить ее аналитически при  $n > 2$  в общем случае не получится, так как это снова приведет к решению уравнения степени  $n$ .

*Пример.* Решить уравнение

$$-x^3 + x^2 + 5x + 3 = 0.$$

Сначала приведем его к виду (П.А.11), разделив на  $a_n = a_3 = -1$ :

$$x^3 - x^2 - 5x - 3 = 0.$$

Корни данного уравнения можно получить из решения системы

$$\left\{ \begin{array}{l} x_1 + x_2 + x_3 = 1, \\ x_1 x_2 + x_1 x_3 + x_2 x_3 = -5, \\ x_1 x_2 x_3 = 3. \end{array} \right.$$

В общем виде эта система не решается. Если решить исходное уравнение численно (используя методы вычислительной математики), найдем два корня:  $x = -1$  и  $x = 3$ . Если численно решить полученную систему уравнений, найдем решение  $x_1 = -1$ ,  $x_2 = -1$ ,  $x_3 = 3$ . Это только одно из решений системы, так как от перестановки компонентов решения системы местами они не перестают быть корнями исходного уравнения.

Заметим, что численное решение исходного уравнения дает два корня, а численное решение системы уравнений – три. Причина в том, что корень  $x = -1$  является *кратным*. В общем случае если корень  $x_i$  уравнения (П.А.10) имеет кратность  $k$ , это означает, что в уравнении (П.А.12) множитель  $(x - x_i)$  встречается  $k$  раз.

*Пример.* Так,

$$x^3 - x^2 - 5x - 3 = (x + 1)(x + 1)(x - 3) = (x + 1)^2(x - 3).$$

Чтобы разложить на множители полином в исходном уравнении, необходимо полином в уравнении (П.А.12) умножить на  $a_n$ :

$$-x^3 + x^2 + 5x + 3 = -(x + 1)^2(x - 3).$$

В данном случае корень  $x = -1$  имеет кратность  $k = 2$ , а корень  $x = 3$  – кратность  $k = 1$ .

Рассмотрим частные случаи, когда уравнение (П.А.10) можно решить аналитически. Его еще называют *алгебраическим* уравнением. Линейные и квадратные уравнения – частные случаи алгебраических уравнений.

Случай 1. Если в кубическом уравнении  $ax^3 + bx^2 + cx + d = 0$  коэффициент  $d = 0$ , то

$$ax^3 + bx^2 + cx = x(ax^2 + bx + c) = 0.$$

То есть имеем одно решение  $x = 0$ , а остальные решения получаем из квадратного уравнения  $ax^2 + bx + c = 0$ . Таким образом, исходное уравнение может иметь от одного до трех корней.

Случай 2. Если в уравнении четвертой степени  $ax^4 + bx^3 + cx^2 + dx + e = 0$  коэффициенты  $b = 0$  и  $d = 0$ , т.е. имеем *биквадратное* уравнение

$$ax^4 + cx^2 + e = 0,$$

то, введя обозначение  $y = x^2$ , получим квадратное уравнение

$$ay^2 + cy + e = 0.$$

Решая его, найдем значения  $y$  (от нуля до двух корней). Если данное уравнение имеет хотя бы один действительный корень, то найдем значения неизвестных  $x$ :

$$x = \pm\sqrt{y}.$$

Это частный случай решения квадратного уравнения. Действительно, если уравнение  $y = x^2$  переписать в виде  $x^2 + 0x - y = 0$ , где  $y$  – параметр, то его решением будет

$$x = \frac{-0 \pm \sqrt{0 - 4 \cdot 1 \cdot (-y)}}{2} = \pm\sqrt{y}.$$

Таким образом, исходное уравнение  $ax^4 + cx^2 + e = 0$  может иметь от нуля до четырех решений (корней).

Случай 3. Если в уравнении (П.А.10) каким-то способом было найдено одно решение  $x = x'$ , то остальные решения можно найти из уравнения

$$\frac{f(x)}{x - x'} = 0. \quad (\text{П.А.15})$$

Для деления полинома  $f(x)$  на полином 1-й степени  $x = x'$  целесообразно воспользоваться алгоритмом Евклида.

*Пример.* Найти решения уравнения

$$-x^3 + x^2 + 5x + 3 = 0.$$

Ранее было показано, что произведение корней данного уравнения равно 3. Следовательно, если корни целые, то ими могут быть

только числа  $\pm 1$  и  $\pm 3$ . Пусть методом подбора одно решение найдено, это  $x = -1$ . Согласно уравнению (П.А.15) разделим левую часть уравнения на полином  $(x + 1)$ , используя алгоритм Евклида:

$$\begin{array}{r}
 -x^3 + x^2 + 5x + 3 \quad |x+1 \\
 -x^3 - x^2 \phantom{+ 5x + 3} \\
 \hline
 2x^2 + 5x \phantom{+ 3} \\
 2x^2 + 2x \phantom{+ 3} \\
 \hline
 3x + 3 \\
 3x + 3 \\
 \hline
 0
 \end{array}$$

Таким образом,

$$-x^3 + x^2 + 5x + 3 = (x + 1)(-x^2 + 2x + 3).$$

Решая далее квадратное уравнение  $-x^2 + 2x + 3$ , найдем еще два решения исходного уравнения:  $x = -1$  и  $x = 3$ . Следовательно,

$$(x + 1)(-x^2 + 2x + 3) = -(x + 1)^2(x - 3),$$

что соответствует полученному ранее результату.

4. Решение иррациональных и трансцендентных уравнений. *Иррациональными* называются уравнения, в которых хотя бы один член содержит иррациональную операцию (например, квадратный или кубический корень). *Трансцендентными* называются уравнения, в которых хотя бы один член содержит трансцендентную операцию (например, логарифм, синус, тангенс и т.д.).

В общем случае такие уравнения не имеют аналитического решения. Однако в некоторых частных случаях аналитическое решение существует.

Случай 4. Например, если иррациональный член в уравнении всего один, т.е. уравнение можно переписать в виде

$$f_1(x) = f_2(x), \quad (\text{П.А.16})$$

где выражение  $f_1(x)$  содержит единственный иррациональный член, а выражение  $f_2(x)$  иррациональных и трансцендентных членов не имеет, то можно попытаться избавиться от иррациональности в левой части уравнения.

*Пример.* Решить уравнение

$$\sqrt{x-7} - x + 13 = 0.$$

Так как под квадратным корнем находится выражение  $x-7$ , то любой корень данного уравнения должен быть  $x \geq 7$ . Перепишем уравнение в виде (П.А.16):

$$\sqrt{x-7} = x - 13.$$

Результат вычисления корня не может быть отрицательным, отсюда получаем еще одно ограничение для решений этой системы:

$$x - 13 \geq 0 \Rightarrow x \geq 13.$$

Возведем обе части уравнения в квадрат:

$$x - 7 = x^2 - 26x + 169.$$

Таким образом, корни исходного уравнения находятся из решения задачи

$$x^2 - 27x + 176 = 0, \quad x \geq 7, \quad x \geq 13.$$

Так как условие  $x \geq 13$  более жесткое, чем  $x \geq 7$ , можно оставить только его. Решение квадратного уравнения  $x^2 - 27x + 176 = 0$  дает два корня:  $x = 11$  и  $x = 16$ . Первый из них отбрасываем, потому что он не удовлетворяет условию  $x \geq 13$ . Следовательно, решением исходного уравнения будет  $x = 16$ .

Случай 5. Если в левой части уравнения находится только один трансцендентный член, а правая часть является константой, т.е. уравнение имеет вид

$$f(x) = \text{const}, \quad (\text{П.А.17})$$

где выражение  $f(x)$  содержит единственный трансцендентный член, то такие уравнения имеют известные аналитические решения:

- 1) уравнение  $\ln g(x) = \text{const}$  имеет решение  $g(x) = e^{\text{const}}$ ;
- 2) уравнение  $\sin g(x) = \text{const}$  при условии  $|\text{const}| \leq 1$  имеет решение  $g(x) = \arcsin \text{const} + 2\pi k$ , где  $k \in Z$ ;
- 3) уравнение  $n^{g(x)} = \text{const}$ ,  $\text{const} > 0$ , где  $n$  – константа или параметр, имеет решение  $g(x) = \log_n \text{const}$  и т.д.

*Пример.* Решить уравнение

$$\ln(x+1) - 2 = 0.$$

Приведем его к виду (П.А.17):

$$\ln(x+1) = 2.$$

Тогда

$$x+1 = e^2 \Rightarrow x = e^2 - 1.$$

Случай 6. Если заменой иррациональных или трансцендентных членов уравнения его удастся преобразовать к алгебраическому уравнению, то оно решается одним из рассмотренных ранее способов.

*Пример 1.* Решить уравнение

$$\ln^2 x + \ln x^2 = 3.$$

Приведем уравнение к виду

$$\ln^2 x + 2 \ln x - 3 = 0.$$

Сделаем замену  $y = \ln x$ :

$$y^2 + 2y - 3 = 0.$$

Решение этого квадратного уравнения дает  $y_1 = 1$ ,  $y_2 = -3$ . Найдем решение исходного уравнения:

$$\ln x_1 = 1 \Rightarrow x_1 = e,$$

$$\ln x_2 = -3 \Rightarrow x_2 = e^{-3}.$$

*Пример 2.* Решить уравнение

$$\sin x + \cos x = 1.$$

Из тригонометрии известно, что  $\cos x = \sqrt{1 - \sin^2 x}$ , поэтому получим

$$\begin{aligned}\sin x + \sqrt{1 - \sin^2 x} &= 1 \Rightarrow \\ \sqrt{1 - \sin^2 x} &= 1 - \sin x.\end{aligned}$$

Это частный случай уравнения (П.А.16). Избавимся от корня, возведя обе части уравнения в квадрат (так как выражение  $1 - \sin x \geq 0$  является тождеством, дополнительные ограничения в решении не появляются):

$$\begin{aligned}1 - \sin^2 x &= (1 - \sin x)^2 \Rightarrow \\ 1 - \sin^2 x &= 1 - 2\sin x + \sin^2 x.\end{aligned}$$

Далее делаем замену  $y = \sin x$ , находим  $y$ , затем  $x$  по известной схеме.

5. Уравнения, содержащие неизвестные в знаменателе. Если уравнение имеет вид

$$\frac{f(x)}{g(x)} = 0, \quad (\text{П.А.18})$$

то его решение совпадает с решением уравнения  $f(x) = 0$ , за исключением того, что отбрасываются корни, являющиеся решениями уравнения  $g(x) = 0$ .

*Пример.* Решить уравнение

$$\frac{\sin 2x}{x^2} = 0.$$

Решение уравнения  $\sin 2x = 0$  дает корни

$$\begin{aligned}2x &= \arcsin 0 + 2\pi k = 2\pi k \Rightarrow \\ x &= \pi k, \quad k \in Z \quad (k = 0, \pm 1, \pm 2, \dots).\end{aligned}$$

Таким образом, уравнение имеет множество решений:  $x = 0$ ,  $x = \pm\pi$ ,  $x = \pm 2\pi$  и т.д. Однако решение  $x = 0$  является также решением уравнения  $x^2 = 0$ , поэтому его отбрасываем. Окончательно имеем  $x = \pi k, k = \pm 1, \pm 2, \dots$

Аналогичным образом поступаем, когда дробей в уравнении несколько:

$$\sum_{i=1}^n \frac{f_i(x)}{g_i(x)} = 0, \quad (\text{П.А.19})$$

где выражения  $f_i(x)$  и  $g_i(x)$  могут содержать константы, параметры и неизвестные. Сначала выражение нужно свести к одной дроби:

$$\sum_{i=1}^n \left[ \frac{f_i(x) \cdot \prod_{j \neq i} g_j(x)}{g_i(x) \cdot \prod_{j \neq i} g_j(x)} \right] = 0 \Rightarrow \frac{\sum_{i=1}^n \left[ f_i(x) \cdot \prod_{j \neq i} g_j(x) \right]}{\prod_{i=1}^n g_i(x)} = 0. \quad (\text{П.А.20})$$

Или сначала привести их к наименьшему общему знаменателю и сложить, если это возможно. Тогда корни уравнения (П.А.19) будут являться решениями задачи

$$\sum_{i=1}^n \left[ f_i(x) \cdot \prod_{j \neq i} g_j(x) \right] = 0, \quad g_i(x) \neq 0, \quad i = 1, 2, \dots, n. \quad (\text{П.А.21})$$

*Пример.* Решить уравнение

$$\frac{2}{x-1} + \frac{5}{x+2} = \frac{13}{x^2+x-2}.$$

Приведем его к виду (4.19):

$$\frac{2}{x-1} + \frac{5}{x+2} - \frac{13}{x^2+x-2} = 0.$$

Согласно уравнению (4.20)

$$\frac{2(x+2)(x^2+x-2) + 5(x-1)(x^2+x-2) - 13(x-1)(x+2)}{(x-1)(x+2)(x^2+x-2)} = 0,$$

поэтому решения исходного уравнения можно найти из условия

$$2(x+2)(x^2+x-2) + 5(x-1)(x^2+x-2) - 13(x-1)(x+2) = 0,$$

$$x-1 \neq 0, \quad x+2 \neq 0, \quad x^2+x-2 \neq 0.$$

Однако если учесть, что  $(x-1)(x+2) = x^2 + x - 2$ , то можно сначала уравнение привести к общему знаменателю:

$$\frac{2(x+2) + 5(x-1) - 13}{x^2 + x - 2} = 0.$$

Получим более простую задачу:

$$2(x+2) + 5(x-1) - 13 = 0, \quad x^2 + x - 2 \neq 0.$$

Решение данного линейного уравнения  $x = 2$  удовлетворяет условию  $x^2 + x - 2 \neq 0$ , поэтому является решением исходного уравнения.

### Решение систем уравнений

Если находят общее решение двух и более уравнений, то говорят, что эти уравнения образуют *систему*. Уравнения, образующие систему, обычно записывают одно под другим и объединяют фигурной скобкой:

$$\begin{cases} f_1(x) = 0, \\ f_2(x) = 0, \\ \dots \\ f_n(x) = 0, \end{cases} \quad (\text{П.А.22})$$

или кратко

$$F(x) = 0. \quad (\text{П.А.23})$$

Если в уравнения  $f_i(x)$  входит константа, то иногда их записывают в виде

$$\begin{cases} f_1(x) = \text{const}_1, \\ f_2(x) = \text{const}_2, \\ \dots \\ f_n(x) = \text{const}_n. \end{cases}$$

Каждый набор неизвестных  $x_1, x_2, \dots, x_m$ , который одновременно удовлетворяет всем уравнениям системы, называется *решением системы*. В общем случае система имеет решения, если количество

неизвестных в ней равно количеству уравнений ( $n = m$ ). Если уравнений меньше, чем неизвестных ( $n < m$ ), то такая система называется *недоопределенной* и обычно имеет бесконечное множество решений. Если уравнений больше, чем неизвестных ( $n > m$ ), то система называется *переопределенной* и часто решений не имеет. Система, не имеющая решений, называется *несовместимой*.

При решении переопределенной системы любые  $n$  ее неизвестных (переменных) можно считать *базисными* (или *зависимыми*). Тогда остальные  $m - n$  неизвестных (переменных) являются *независимыми*. Независимым неизвестным можно присваивать любые значения (не противоречащие уравнениям системы), а значения базисных неизвестных находят, решая систему.

*Пример 1.* Система уравнений

$$\begin{cases} xy = -5, \\ \sqrt{x} + \sqrt{y} = 2 \end{cases}$$

является *несовместимой*, так как второе уравнение требует, чтобы выполнялись условия  $x \geq 0$ ,  $y \geq 0$ . Но в таком случае произведение неизвестных не может быть отрицательным, как этого требует первое уравнение системы.

*Пример 2.* Система уравнений

$$\begin{cases} \sin x = 1, \\ \cos x = 1 \end{cases}$$

переопределена, так как в ней  $n = 2$  уравнения и  $m = 1$  неизвестное. В итоге она не имеет решений, потому что синус и косинус не могут быть равны единице для одного и того же значения неизвестного  $x$ .

*Пример 3.* Система уравнений

$$\begin{cases} xyz = -5, \\ \sqrt{x} + \sqrt{y} = 2 \end{cases}$$

недоопределена, так как в ней  $n = 2$  уравнения и  $m = 3$  неизвестных. Примем за базисные переменные  $x$  и  $y$ , а  $z$  будет свободная переменная. Очевидно, что ей можно присваивать любые значения в диапазоне  $(-\infty, 0)$ , тогда система будет иметь решения. Например, если принять  $z = -5$ , система

$$\begin{cases} -5xy = -5, \\ \sqrt{x} + \sqrt{y} = 2 \end{cases}$$

будет иметь решение  $x = 1$ ,  $y = 1$ . Следовательно, значения неизвестных  $x = 1$ ,  $y = 1$ ,  $z = -5$ , являются одним из решений исходной системы. В целом она, как отмечалось выше, имеет бесконечное множество решений.

При решении систем уравнений используют разные подходы.

1. *Способ алгебраического сложения.* Учитывая, что умножение любого уравнения системы (П.А.22) на произвольную константу  $k \neq 0$  не влияет на ее решение, можно сформулировать правило: если в двух уравнениях системы есть два члена, которые совпадают с точностью до коэффициента пропорциональности  $k$ , например

$$\begin{cases} f_i(x) = f_i'(x) + a \cdot g(x) = 0, \\ f_j(x) = f_j'(x) + b \cdot g(x) = 0, \end{cases}$$

то, складывая эти два уравнения с умножением одного из них на константу  $-k$ , от этого члена можно избавиться. В данном случае либо первое уравнение умножаем на  $k = -b/a$ , тогда

$$\begin{cases} -\frac{b}{a} \cdot f_i'(x) - b \cdot g(x) = 0, \\ f_j'(x) + b \cdot g(x) = 0, \end{cases} \Rightarrow -\frac{b}{a} \cdot f_i'(x) + f_j'(x) = 0,$$

либо второе уравнение умножаем на  $k = -a/b$ , тогда

$$\begin{cases} f_i'(x) + a \cdot g(x) = 0, \\ -\frac{a}{b} \cdot f_j'(x) - a \cdot g(x) = 0, \end{cases} \Rightarrow f_i'(x) - \frac{a}{b} \cdot f_j'(x) = 0.$$

2. *Способ подстановки.* Способ заключается в том, чтобы из одного уравнения системы выразить какое-нибудь неизвестное (или член, встречающийся в других уравнениях) и подставить полученное выражение в остальные уравнения.

Рассмотренные способы используются, чтобы избавиться от иррациональных или трансцендентных членов уравнений либо уменьшить в некоторых уравнениях количество неизвестных. Для избавления от иррациональных и трансцендентных членов можно также использовать описанные ранее приемы.

**Решение систем линейных уравнений.** Если все функции  $f_i(x)$  в системе (П.А.22) имеют линейный вид, т.е.

$$f_i(x) = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n + b_i,$$

то такая система уравнений называется *линейной*. Для решения систем линейных уравнений используются либо специальные численные методы (Гаусса, декомпозиции и т.п.), либо методы линейной алгебры.

*Пример 1.* Решить систему линейных уравнений

$$\begin{cases} x + y = 1, \\ 2x - 3y = 7. \end{cases}$$

Используем способ алгебраического сложения. Ко второму уравнению системы добавим первое уравнение, умноженное на коэффициент 3:

$$\begin{aligned} 2x - 3y + 3(x + y) &= 7 + 3 \cdot 1 \Rightarrow \\ 5x &= 10 \Rightarrow x = 2. \end{aligned}$$

Теперь способом подстановки из первого уравнения находим

$$x + y = 1 \Rightarrow y = 1 - x = 1 - 2 = -1.$$

*Пример 2.* Решим эту же систему

$$\begin{cases} x + y = 1, \\ 2x - 3y = 7, \end{cases}$$

сразу применив способ подстановки. Из первого уравнения выразим  $x = 1 - y$  и подставим во второе уравнение:

$$2(1-y) - 3y = 7 \Rightarrow -5y = 5 \Rightarrow y = -1.$$

Затем из уравнения  $x = 1 - y$  находим значение неизвестной переменной  $x$ :

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

*Пример 3.* Решить систему линейных уравнений

$$\begin{cases} 2x_1 + 3x_2 - 7x_3 = 16, \\ x_1 + x_2 - 2x_3 = 6, \\ 5x_1 + 2x_2 + x_3 = 16. \end{cases}$$

Используем метод Гаусса. Сначала первое уравнение разделим на 2, а затем прибавим его ко второму уравнению, умножив на  $-1$ , и к третьему уравнению, умножив на  $-5$ :

$$\begin{cases} x_1 + \frac{3}{2}x_2 - \frac{7}{2}x_3 = 8, \\ x_1 + x_2 - 2x_3 - \left(x_1 + \frac{3}{2}x_2 - \frac{7}{2}x_3\right) = 6 - 8, \\ 5x_1 + 2x_2 + x_3 - 5\left(x_1 + \frac{3}{2}x_2 - \frac{7}{2}x_3\right) = 16 - 5 \cdot 8, \end{cases} \Rightarrow \begin{cases} x_1 + \frac{3}{2}x_2 - \frac{7}{2}x_3 = 8, \\ -\frac{1}{2}x_2 + \frac{3}{2}x_3 = -2, \\ -\frac{11}{2}x_2 + \frac{37}{2}x_3 = -24. \end{cases}$$

Затем второе уравнение разделим на  $-1/2$  и прибавим его к третьему уравнению, умножив на  $11/2$ :

$$\begin{cases} x_1 + \frac{3}{2}x_2 - \frac{7}{2}x_3 = 8, \\ x_2 - 3x_3 = 4, \\ -\frac{11}{2}x_2 + \frac{37}{2}x_3 + \frac{11}{2}(x_2 - 3x_3) = -24 + \frac{11}{2} \cdot 4, \end{cases} \Rightarrow \begin{cases} x_1 + \frac{3}{2}x_2 - \frac{7}{2}x_3 = 8, \\ x_2 - 3x_3 = 4, \\ 2x_3 = -2. \end{cases}$$

Теперь последовательно получим

$$2x_3 = -2 \Rightarrow x_3 = -1,$$

$$x_2 - 3x_3 = 4 \Rightarrow x_2 = 4 + 3x_3 = 1,$$

$$x_1 + \frac{3}{2}x_2 - \frac{7}{2}x_3 = 8 \Rightarrow x_1 = 8 - \frac{3}{2}x_2 + \frac{7}{2}x_3 = 8 - \frac{3}{2} - \frac{7}{2} = 3.$$

**Решение систем нелинейных уравнений.** Если хотя бы одна функция  $f_i(x)$  в системе (П.А.22) имеет нелинейный вид, то такую систему уравнений называют *нелинейной*. Для решения систем нелинейных уравнений используются либо специальные численные методы (Ньютона, итераций и т.д.), либо рассмотренные выше способы. Следует отметить, что, в отличие от систем линейных уравнений, численные методы решения систем нелинейных уравнений дают не точное решение, а приближенное.

*Пример 1.* Решить систему нелинейных уравнений

$$\begin{cases} x^2 + \sin y - 0,5 = 0, \\ 4x - 2 \sin y - 5 = 0. \end{cases}$$

Используем способ алгебраического сложения. Ко второму уравнению добавим первое уравнение, умноженное на коэффициент 2:

$$\begin{aligned} 4x - 2 \sin y - 5 + 2 \cdot (x^2 + \sin y - 0,5) &= 0 \Rightarrow \\ 2x^2 + 4x - 6 &= 0. \end{aligned}$$

Решение этого квадратного уравнения дает  $x_1 = 1$ ,  $x_2 = -3$ . Теперь из первого уравнения найдем значение неизвестной переменной  $y$ :

$$\begin{aligned} x_1^2 + \sin y_1 - 0,5 = 0 &\Rightarrow \sin y_1 = 0,5 - x_1^2 = -0,5 \\ \Rightarrow y_1 = \arcsin(-0,5) + 2\pi k &= -\frac{\pi}{6} + 2\pi k, \quad k \in Z. \end{aligned}$$

$$\begin{aligned} x_2^2 + \sin y_2 - 0,5 = 0 &\Rightarrow \sin y_2 = 0,5 - x_2^2 = 3,5 \\ &\Rightarrow \text{решений нет.} \end{aligned}$$

Система имеет множество решений  $x = 1$ ,  $y = -\pi/6 + 2\pi k$ ,  $k \in Z$ .

*Пример 2.* Решим эту же систему

$$\begin{cases} x^2 + \sin y - 0,5 = 0, \\ 4x - 2 \sin y - 5 = 0, \end{cases}$$

сразу применив способ подстановки. Из второго уравнения выразим

$$x = \frac{1}{4}(5 + 2 \sin y)$$

и подставим в первое уравнение:

$$\frac{1}{16}(5 + 2 \sin y)^2 + \sin y - 0,5 = 0,$$

дальнейшее решение очевидно.

### Доказательство неравенств

Два выражения, соединенные знаками «>», «<», «≥» или «≤» образуют *неравенство*. Причем знаки «>» и «<» образуют *строгое неравенство*, а знаки «≥» и «≤» – *нестрогое неравенство*. Знаки «>» и «<» называются *противоположными*. Знаки «≥» и «≤» также противоположны. Если два неравенства имеют противоположные знаки, то они называются *неравенствами противоположного смысла*.

Кроме того, неравенство образуют два выражения, соединенные знаком «≠». Далее о подобных неравенствах говорить не будем, так как решением неравенства

$$f(x) \neq 0 \quad (\text{П.А.24})$$

являются все значения неизвестного  $x$ , за исключением решений уравнения  $f(x) = 0$ .

Всякое неравенство, справедливое при всех допустимых значениях входящих в него неизвестных, называется *тождественным неравенством*.

Свойства тождественных неравенств:

- 1) если  $a \odot b$  и  $b \odot c$ , то  $a \odot c$ , где « $\odot$ » – любой знак неравенства из «>», «<», «≥» или «≤»;
- 2) если  $a \odot b$ , то  $a + c \odot b + c$ , и наоборот;
- 3) из  $a + b \odot c$  следует, что  $a \odot c - b$ ,  $a + b - c \odot 0$  и т.д.;
- 4) из  $a \odot b$  и  $c \odot d$  следует  $a + c \odot b + d$ ;
- 5) из  $a \odot b$  и  $c \otimes d$  следует  $a - c \odot b - d$ , где « $\otimes$ » – неравенство противоположного смысла;
- 6) если  $a \odot b$  и  $c > 0$ , то  $ac \odot bc$  и  $a/c \odot b/c$ ;

7) если  $a \odot b$  и  $c < 0$ , то  $ac \otimes bc$  и  $a/c \otimes b/c$ ;

8) если  $a \odot b$ ,  $a > 0$ ,  $b > 0$ , то  $a^n \odot b^n$ , и т.д.

Убеждаются в тождественности неравенств с помощью *доказательств*. Доказательство – это обоснование, что неравенство выполняется при любых значениях входящих в него неизвестных. Рассмотрим способы доказательства неравенств.

1. Доказательство неравенства, содержащего только одно неизвестное, можно свести к решению уравнения. Пусть имеется неравенство с одним неизвестным

$$f_1(x) \odot f_2(x). \quad (\text{П.А.25})$$

Преобразуем его к виду

$$f_1(x) - f_2(x) \odot 0$$

или

$$f(x) \odot 0, \quad (\text{П.А.26})$$

где  $f(x) = f_1(x) - f_2(x)$ . Тогда, решив уравнение  $f(x) = 0$ , мы найдем значения  $x$ , в которых тождественность неравенства может меняться. Для проверки необходимо подставить значение, лежащее слева или справа от значения  $x$  (но не далее чем соседнее решение уравнения  $f(x) = 0$ , если решений несколько).

*Пример 1.* Доказать неравенство

$$\sqrt{x} \leq x \quad \forall x \geq 1.$$

Приведем неравенство к виду (П.А.26). Так как левая и правая части неравенства не могут быть отрицательными, их можно возвести в квадрат:

$$x \leq x^2 \Rightarrow x^2 - x \geq 0.$$

Решим уравнение  $x^2 - x = 0$ . Его корнями являются  $x = 0$  и  $x = 1$ . Исследуем тождественность неравенства  $x^2 - x \geq 0$  на образованных корнями интервалах:

$(-\infty, 0]$	$[0, 1]$	$[1, +\infty)$
$(-1)^2 - (-1) = 2 \geq 0$ – выполняется	$(0,5)^2 - 0,5 = -0,25 < 0$ – не выполняется	$2^2 - 2 = 2 \geq 0$ – выполняется

Таким образом, для любого значения  $x \geq 1$  неравенство тождественно.

*Пример 2.* Доказать неравенство

$$\sqrt{x} \leq x + 1 \quad \forall x \geq 0.$$

Приведем неравенство к виду (П.А.26). Так как левая и правая части неравенства не могут быть отрицательными, их можно возвести в квадрат:

$$x \leq (x+1)^2 \Rightarrow x^2 + x + 1 \geq 0.$$

Уравнение  $x^2 + x + 1 = 0$  действительных решений не имеет. Следовательно, неравенство  $x^2 + x + 1 \geq 0$  либо выполняется для любого значения  $x$ , либо не выполняется. Проверим:

$$x = 1 \Rightarrow 1^2 + 1 + 1 = 3 \geq 0.$$

Таким образом, неравенство тождественно  $\forall x$ , а исходное неравенство тождественно  $\forall x \geq 0$  (так как имеется квадратный корень).

*Пример 3.* Доказать неравенство

$$\sqrt{x^2 + 1} \leq x^2 + 1.$$

Приведем неравенство к виду (П.А.26). Так как левая и правая части неравенства не могут быть отрицательными, их можно возвести в квадрат:

$$x^2 + 1 \leq (x^2 + 1)^2 \Rightarrow x^4 + x^2 \geq 0.$$

Уравнение  $x^4 + x^2 = 0$  имеет единственное решение  $x = 0$ . Проверяем тождественность неравенства  $x^4 + x^2 \geq 0$ :

$(-\infty, 0]$	$[0, +\infty)$
$(-1)^4 + (-1)^2 = 2 \geq 0$ – выполняется	$(1)^4 + (1)^2 = 2 \geq 0$ – выполняется

Следовательно, исходное неравенство тождественно  $\forall x$ .

2. Неравенство можно доказать, сведя его к неравенству, тождественность которого очевидна, или к серии таких неравенств.

*Пример 1.* Доказать неравенство

$$\frac{a+b}{2} \geq \sqrt{ab} \quad \forall a > 0, b > 0.$$

Приведем неравенство к виду (П.А.26). Так как левая и правая части неравенства не могут быть отрицательными, их можно возвести в квадрат:

$$\begin{aligned} \left(\frac{a+b}{2}\right)^2 &\geq (\sqrt{ab})^2 \Rightarrow \frac{1}{4}a^2 + \frac{1}{2}ab + \frac{1}{4}b^2 \geq ab \Rightarrow \\ \frac{1}{4}a^2 - \frac{1}{2}ab + \frac{1}{4}b^2 &\geq 0 \Rightarrow \left(\frac{a-b}{2}\right)^2 \geq 0. \end{aligned}$$

Квадрат любого выражения не может быть отрицательным, следовательно, тождественность исходного неравенства доказана.

*Пример 2.* Доказать неравенство

$$\sqrt{6 + \sqrt{6 + \dots + \sqrt{6 + \sqrt{6 + \sqrt{6}}}}} < 3.$$

Заменим под знаком последнего корня число 6 числом 9, тогда

$$\begin{aligned} \sqrt{6 + \sqrt{6 + \dots + \sqrt{6 + \sqrt{6 + \sqrt{9}}}}} &= \sqrt{6 + \sqrt{6 + \dots + \sqrt{6 + \sqrt{9}}}} = \\ &= \sqrt{6 + \sqrt{6 + \dots + \sqrt{9}}} = \dots = \sqrt{9} = 3. \end{aligned}$$

Но  $\sqrt{6} < \sqrt{9}$ , следовательно,

$$\begin{aligned} \sqrt{6 + \sqrt{6 + \dots + \sqrt{6 + \sqrt{6 + \sqrt{6}}}}} &< \sqrt{6 + \sqrt{6 + \dots + \sqrt{6 + \sqrt{6 + \sqrt{9}}}}} \Rightarrow \\ \sqrt{6 + \sqrt{6 + \dots + \sqrt{6 + \sqrt{6 + \sqrt{6}}}}} &< 3. \end{aligned}$$

При доказательстве тождественности неравенств используют *замечательные неравенства*. Приведем некоторые из них:

- 1)  $|a_1 + a_2 + \dots + a_n| \leq |a_1| + |a_2| + \dots + |a_n|$ ;
- 2)  $\sqrt[n]{a_1 a_2 \dots a_n} \leq \frac{a_1 + a_2 + \dots + a_n}{n} \quad \forall a_1, a_2, \dots, a_n \geq 0$ ;
- 3)  $\frac{a_1 + a_2 + \dots + a_n}{n} \leq \sqrt[n]{\frac{a_1^2 + a_2^2 + \dots + a_n^2}{n}} \quad \forall a_1, a_2, \dots, a_n \geq 0$ ;
- 4)  $a_1 b_2 + a_2 b_2 + \dots + a_n b_n \leq \sqrt{a_1^2 + a_2^2 + \dots + a_n^2} \cdot \sqrt{b_1^2 + b_2^2 + \dots + b_n^2}$ ;
- 5)  $\sqrt{a_1^2 + a_2^2 + \dots + a_n^2} \leq |a_1| + |a_2| + \dots + |a_n|$ ;
- 6)  $\sqrt{a_1^2 + a_2^2 + \dots + a_n^2} - \sqrt{b_1^2 + b_2^2 + \dots + b_n^2} \leq$   
 $\leq |a_1 - b_1| + |a_2 - b_2| + \dots + |a_n - b_n|$ ;
- 7)  $\sqrt{(a_1 - b_1)^2 + (a_2 - b_1)^2 + \dots + (a_n - b_n)^2} \leq$   
 $\leq \sqrt{a_1^2 + a_2^2 + \dots + a_n^2} + \sqrt{b_1^2 + b_2^2 + \dots + b_n^2}$ ;
- 8)  $(a_1 b_2 + a_2 b_2 + \dots + a_n b_n)^2 \leq (a_1^2 + a_2^2 + \dots + a_n^2) \cdot (b_1^2 + b_2^2 + \dots + b_n^2)$ ;
- 9)  $(1 + a)^n > 1 + na \quad \forall a \geq 0, n > 1$  и т.д.

### П.А.3. Векторы и матрицы

Матрица – это прямоугольная таблица чисел. Матрицы обычно обозначаются заглавными буквами ( $A, B$  и т.д.), а их элементы – соответствующими строчными буквами с индексами, т.е.  $a_{ij}, b_{ij}$  (или  $a_{i,j}, b_{i,j}$ ) и т.д. Первый индекс означает номер строки, а второй – столбца. Поэтому матрицу  $A$  можно записать в виде

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \quad (\text{П.А.27})$$

или  $A = \{a_{ij}, i = 1, 2, \dots, n, j = 1, 2, \dots, m\}$ . Здесь пара чисел  $n$  и  $m$  называется размерностью матрицы ( $n$  – количество строк матрицы,  $m$  – количество столбцов) и обозначается  $n \times m$ .

Матрица называется *нулевой*, если все ее элементы равны нулю:

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

или

$$A = 0.$$

Матрица называется *квадратной*, если количество ее строк равно количеству столбцов, т.е.  $A = \{a_{ij}, i = 1, 2, \dots, n, j = 1, 2, \dots, n\}$  или

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Число  $n$  называется *порядком* квадратной матрицы. Элементы квадратной матрицы, у которых индексы строки и столбца совпадают ( $a_{ii}$ ), называются элементами главной диагонали матрицы (или ее диагональными элементами):

$$\begin{pmatrix} \boxed{a_{11}} & a_{12} & \dots & a_{1n} \\ a_{21} & \boxed{a_{22}} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \boxed{a_{nn}} \end{pmatrix}.$$

Квадратная матрица называется *треугольной (верхней треугольной)*, если ее элементы, лежащие ниже диагональных, равны нулю:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Квадратная матрица называется *нижней треугольной*, если ее элементы, лежащие выше диагональных, равны нулю:

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Треугольная матрица называется *диагональной*, если все ее элементы, кроме диагональных, равны нулю:

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Если ввести обозначение  $d_i = a_{ii}$ , то диагональную матрицу можно записать следующим образом:

$$A = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{pmatrix}$$

или

$$A = \text{diag } d_i, \quad i = 1, 2, \dots, n.$$

Диагональная матрица называется *единичной*, если все ее элементы равны единице. Для единичной матрицы вводятся специальные обозначения –  $E$  или  $I$  (от англ. identity). Таким образом,  $I = E = \text{diag} 1$

или

$$I = E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

*Вектор* – это частный случай матрицы, состоящей из одного столбца. Чтобы это подчеркнуть, иногда говорят «*вектор-столбец*».

Векторы обычно обозначаются строчными буквами ( $x$ ,  $y$  и т.д.), а их элементы – буквами с индексами, т.е.  $x_i$ ,  $y_i$  и т.д. Поэтому вектор  $x$  можно записать как

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix},$$

где  $n$  – размер вектора. Если вектор необходимо записать в виде строки, то нужно обязательно использовать знак транспонирования:

$$x = (x_1 \quad x_2 \quad \dots \quad x_n)^T.$$

Иногда при строчной записи компоненты вектора отделяют друг от друга запятой или точкой с запятой.

В некоторых случаях приходится иметь дело с векторами, которые представляют собой строку. Такие векторы называют *векторами-строками*. Но по умолчанию любой вектор является столбцом. Если вектор – строка, то это нужно специально оговаривать.

Вектор называется *нулевым*, если все его элементы равны нулю:

$$x = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}$$

или

$$x = 0.$$

Вектор называется *единичным*, если он является столбцом единичной матрицы. Такой вектор обозначается  $e_j$ , где  $j$  – номер столбца единичной матрицы:

$$e_j = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

### Основные операции с матрицами

**Умножение и деление матрицы на скаляр.** При умножении или делении матрицы  $A = \{a_{ij}\}$  размера  $n \times m$  на скаляр  $k$  все ее элементы соответственно умножаются или делятся на эту константу, т.е. результатом будет матрица  $B = \{b_{ij} = a_{ij} \cdot k\}$  или  $B = \{b_{ij} = a_{ij} / k\}$  размера  $n \times m$ :

$$B = A \cdot k = \begin{pmatrix} a_{11} \cdot k & a_{12} \cdot k & \dots & a_{1m} \cdot k \\ a_{21} \cdot k & a_{22} \cdot k & \dots & a_{2m} \cdot k \\ \dots & \dots & \dots & \dots \\ a_{n1} \cdot k & a_{n2} \cdot k & \dots & a_{nm} \cdot k \end{pmatrix}, \quad (\text{П.А.28})$$

$$B = A / k = \begin{pmatrix} a_{11} / k & a_{12} / k & \dots & a_{1m} / k \\ a_{21} / k & a_{22} / k & \dots & a_{2m} / k \\ \dots & \dots & \dots & \dots \\ a_{n1} / k & a_{n2} / k & \dots & a_{nm} / k \end{pmatrix}. \quad (\text{П.А.29})$$

Приведем некоторые леммы без доказательств:

- 1) операция умножения коммутативна, т.е.  $A \cdot k = k \cdot A$ ;
- 2)  $A \cdot 1 = 1 \cdot A = A$ ;
- 3)  $A \cdot 0 = 0 \cdot A = 0$ ;
- 4)  $A / 1 = A$

**Сложение и вычитание матриц.** При сложении и вычитании матриц  $A = \{a_{ij}\}$  и  $B = \{b_{ij}\}$  одинакового размера  $n \times m$  все их элементы складываются или вычитаются попарно, т.е. результатом будет матрица  $C = \{c_{ij} = a_{ij} \pm b_{ij}\}$  размера  $n \times m$ :

$$C = A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2m} + b_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \dots & a_{nm} + b_{nm} \end{pmatrix}, \quad (\text{П.А.30})$$

$$C = A - B = \begin{pmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \dots & a_{1m} - b_{1m} \\ a_{21} - b_{21} & a_{22} - b_{22} & \dots & a_{2m} - b_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} - b_{n1} & a_{n2} - b_{n2} & \dots & a_{nm} - b_{nm} \end{pmatrix}. \quad (\text{П.А.31})$$

Приведем некоторые леммы без доказательств:

- 1) операция сложения коммутативна, т.е.  $A + B = B + A$ ;
- 2)  $A + 0 = 0 + A = A$ , где  $0$  – нулевая матрица;
- 3)  $A - 0 = A$ , где  $0$  – нулевая матрица;
- 4)  $0 - A = -A = (-1) \cdot A$ , где  $0$  – нулевая матрица;
- 5)  $A - A = 0$ , где  $0$  – нулевая матрица;
- 6)  $k_1 A \pm k_2 A = (k_1 \pm k_2) A$ , где  $k_1, k_2$  – скаляры.

**Умножение матриц.** При умножении матрицы  $A = \{a_{ik}\}$  размера  $n \times p$  на матрицу  $B = \{b_{kj}\}$  размера  $p \times m$  результатом будет матрица  $C = \{c_{ij}\}$  размера  $n \times m$ , каждый элемент  $c_{ij}$  которой является суммой произведений элементов  $i$ -й строки матрицы  $A$  и  $j$ -го столбца матрицы  $B$ :

$$C = A \cdot B = \begin{pmatrix} \sum_{k=1}^p a_{1k} b_{k1} & \sum_{k=1}^p a_{1k} b_{k2} & \dots & \sum_{k=1}^p a_{1k} b_{km} \\ \sum_{k=1}^p a_{2k} b_{k1} & \sum_{k=1}^p a_{2k} b_{k2} & \dots & \sum_{k=1}^p a_{2k} b_{km} \\ \dots & \dots & \dots & \dots \\ \sum_{k=1}^p a_{nk} b_{k1} & \sum_{k=1}^p a_{nk} b_{k2} & \dots & \sum_{k=1}^p a_{nk} b_{km} \end{pmatrix} \quad (\text{П.А.32})$$

ИЛИ

$$c_{ij} = \sum_{k=1}^p a_{ik} b_{kj}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

*Пример.* Даны матрицы

$$A = \begin{pmatrix} 1 & -2 & 2 & 4 \\ -3 & 7 & -1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 4 & -4 \\ -1 & 3 & 3 \\ -5 & 0 & 8 \\ 1 & -2 & 0 \end{pmatrix}.$$

Матрица  $A$  размера  $2 \times 4$ , матрица  $B$  размера  $4 \times 3$ , т.е.  $n = 2$ ,  $m = 3$ ,  $p = 4$ . Значит, можно найти их произведение:  $C = AB$ . Матрица  $C$  будет иметь размер  $2 \times 3$ :

$$c_{11} = 1 \cdot 2 + (-2) \cdot (-1) + 2 \cdot (-5) + 4 \cdot 1 = 2 + 2 - 10 + 4 = -2;$$

$$c_{12} = 1 \cdot 4 + (-2) \cdot 3 + 2 \cdot 0 + 4 \cdot (-2) = 4 - 6 + 0 - 8 = -10;$$

$$c_{13} = 1 \cdot (-4) + (-2) \cdot 3 + 2 \cdot 8 + 4 \cdot 0 = -4 - 6 + 16 + 0 = 6;$$

$$c_{21} = -3 \cdot 2 + 7 \cdot (-1) + (-1) \cdot (-5) + (-1) \cdot 1 = -6 - 7 + 5 - 1 = -9;$$

$$c_{22} = -3 \cdot 4 + 7 \cdot 3 + (-1) \cdot 0 + (-1) \cdot (-2) = -12 + 21 + 0 + 2 = 11;$$

$$c_{23} = -3 \cdot (-4) + 7 \cdot 3 + (-1) \cdot 8 + (-1) \cdot 0 = 12 + 21 - 8 + 0 = 25.$$

Следовательно,

$$C = \begin{pmatrix} -2 & -10 & 6 \\ -9 & 11 & 25 \end{pmatrix}.$$

Приведем некоторые леммы без доказательств:

- 1)  $A \cdot 0 = 0 \cdot A$ , где  $0$  – нулевая матрица;
- 2)  $A \cdot E = E \cdot A = A$ , где  $E$  – единичная матрица.

*Замечание.* Операция умножения матриц не коммутативна, т.е.  $AB \neq BA$ . Исключения составляют случаи, когда одна из матриц является единичной или нулевой.

**Транспонирование матрицы.** При транспонировании матрицы  $A = \{a_{ij}\}$  размера  $n \times m$  результатом будет матрица  $B = \{b_{ji}\}$  размера  $m \times n$ , каждый элемент  $b_{ji}$  которой равен элементу  $a_{ij}$  исходной матрицы:

$$B = A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}.$$

*Пример.* Транспонируем матрицу

$$C = \begin{pmatrix} -2 & -10 & 6 \\ -9 & 11 & 25 \end{pmatrix},$$

получим

$$C^T = \begin{pmatrix} -2 & -9 \\ -10 & 11 \\ 6 & 25 \end{pmatrix}.$$

Приведем некоторые леммы без доказательств:

- 1)  $E^T = E$ , где  $E$  – единичная матрица;
- 2)  $(A^T)^T = A$ ;
- 3)  $(\text{diag} d_i)^T = \text{diag} d_i$ .

**Определение ранга матрицы.** Рангом матрицы  $A = \{a_{ik}\}$  размера  $n \times m$  называется максимальное число *линейно независимых* строк (столбцов) этой матрицы. Ранг матрицы обозначается

$$\text{rank } A \text{ или } \text{rang } A.$$

Несколько строк (столбцов) являются линейно независимыми, если ни одна (один) из них не выражается линейной комбинацией других строк (столбцов). Причем неважно, что именно проверять – строки или столбцы, так как

$$\text{rang } A^T = \text{rang } A.$$

Очевидно, что ранг матрицы не может быть больше, чем количество ее строк или столбцов. Так, для матрицы размера  $n \times m$

$$\text{rang } A \leq \min(n, m).$$

Данное правило является следствием того, что если векторов в наборе больше, чем их размерность, то они обязательно линейно зависимы. Поэтому в матрице размера  $n \times m$  строка с номером  $n + 1$

обязательно будет линейно зависеть от остальных  $n$  строк. Аналогично столбец с номером  $m + 1$  обязательно будет линейно зависеть от остальных  $m$  столбцов.

*Пример.* Найти ранг матрицы

$$A = \begin{pmatrix} 1 & 2 & -3 & -1 \\ -2 & 5 & 1 & 4 \\ 4 & -1 & -7 & -6 \end{pmatrix}.$$

Ранг не может быть больше, чем  $\min(3,4) = 3$ . Однако из трех строк матрицы только две являются линейно независимыми, так как

$$a_1 = \frac{1}{2}a_3 + \frac{1}{2}a_2,$$

$$a_2 = 2a_1 - a_3,$$

$$a_3 = 2a_1 - a_2,$$

где  $a_1, a_2, a_3$  – соответствующие строки матрицы.

Следовательно,  $\text{rang } A = 2$

Для определения ранга матрицы можно использовать метод Гаусса. Известно, что при преобразовании матрицы методом Гаусса ее ранг не меняется. Ранг матрицы будет равен полученному количеству базисных переменных (т.е. количеству ненулевых строк).

*Пример 1.* Найти ранг матрицы

$$A = \begin{pmatrix} 1 & 2 & -3 & -1 \\ -2 & 5 & 1 & 4 \\ 4 & -1 & -7 & -6 \end{pmatrix}.$$

Используем метод Гаусса:

$$\text{rang} \begin{pmatrix} 1 & 2 & -3 & -1 \\ -2 & 5 & 1 & 4 \\ 4 & -1 & -7 & -6 \end{pmatrix} = \text{rang} \begin{pmatrix} 1 & 2 & -3 & -1 \\ 0 & 9 & -5 & 2 \\ 0 & -9 & 5 & 6 \end{pmatrix} =$$

$$= \text{rang} \begin{pmatrix} 1 & 2 & -3 & -1 \\ 0 & 9 & -5 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Таким образом, получили две базисные переменные, поэтому  $\text{rang } A = 2$ .

*Пример 2.* Найти ранг транспонированной матрицы из предыдущего примера:

$$A = \begin{pmatrix} 1 & -2 & 4 \\ 2 & 5 & -1 \\ -3 & 1 & -7 \\ -1 & 4 & -6 \end{pmatrix}.$$

Используем метод Гаусса:

$$\text{rang} \begin{pmatrix} 1 & -2 & 4 \\ 2 & 5 & -1 \\ -3 & 1 & -7 \\ -1 & 4 & -6 \end{pmatrix} = \text{rang} \begin{pmatrix} 1 & -2 & 4 \\ 0 & 9 & -9 \\ 0 & -5 & 5 \\ 0 & 2 & -2 \end{pmatrix} = \text{rang} \begin{pmatrix} 1 & -2 & 4 \\ 0 & 9 & -9 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Как и ожидалось, после транспонирования ранг матрицы не изменился.

**Норма матрицы.** *Норма* – это некоторая скалярная характеристика матрицы. Норма матрицы  $A$  обозначается  $\|A\|$ .

Существует несколько способов вычисления нормы матрицы. Например, *евклидова норма* вычисляется следующим образом:

$$\|A\| = \sqrt{\sum_{i=1}^n \sum_{j=1}^m a_{ij}^2}. \quad (\text{П.А.33})$$

### Операции с квадратными матрицами

**Определитель.** *Определитель* (или *детерминант*) – это многочлен, комбинирующий элементы квадратной матрицы таким образом, что его значение сохраняется при транспонировании и линейных комбинациях строк или столбцов.

Для обозначения определителя используется символ  $\Delta$ . Если вычисляются определители разных матриц, то в обозначение определителя добавляется обозначение матрицы. Например, определитель матрицы  $A$  обозначается

$$\Delta_A = \det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Для квадратной матрицы размера  $n \times n$  определитель вычисляется следующим образом:

$$\Delta = \sum_{i_1, i_2, \dots, i_n} (-1)^{N(i_1, i_2, \dots, i_n)} a_{i_1, 1} a_{i_2, 2} \dots a_{i_n, n}, \quad (\text{П.А.34})$$

где  $i_1, i_2, \dots, i_n$  – перестановка чисел от 1 до  $n$ ;  $N(i_1, i_2, \dots, i_n)$  – число инверсий в такой перестановке. Общее количество перестановок равно  $n!$ , т.е. в сумме будет  $n!$  слагаемых.

Рассмотрим частные случаи. При  $n = 1$  получим

$$\Delta = |a_{11}| = (-1)^{N(1)} a_{11} = a_{11},$$

так как здесь нет инверсий и  $N(1) = 0$ . При  $n = 2$  получим

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (-1)^{N(1,2)} a_{11} a_{22} + (-1)^{N(2,1)} a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21}.$$

Здесь в первом слагаемом нет инверсий,  $N(1,2) = 0$ , а во втором слагаемом одна инверсия,  $N(2,1) = 1$  (т.е. нужно совершить одну инверсию, чтобы из перестановки 2,1 получить 1,2). При  $n = 3$

$$\begin{aligned} \Delta &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\ &= (-1)^{N(1,2,3)} a_{11} a_{22} a_{33} + (-1)^{N(1,3,2)} a_{11} a_{23} a_{32} + \\ &+ (-1)^{N(2,1,3)} a_{12} a_{21} a_{33} + (-1)^{N(2,3,1)} a_{12} a_{23} a_{31} + \end{aligned}$$

$$\begin{aligned}
& +(-1)^{N(3,2,1)} a_{13} a_{22} a_{31} + (-1)^{N(3,1,2)} a_{13} a_{21} a_{32} = \\
& = a_{11} a_{22} a_{33} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} + a_{12} a_{23} a_{31} - a_{13} a_{22} a_{31} + a_{13} a_{21} a_{32}.
\end{aligned}$$

Однако с увеличением  $n$  формула (П.А.34) становится менее эффективной. Так, при  $n = 4$  получим  $4! = 24$  перестановки, при  $n = 5$  – уже  $5! = 120$  и т.д.

Приведем некоторые леммы без доказательств.

1. При перестановке двух строк или столбцов матрицы местами определитель меняет знак.

2. При добавлении к любой строке или столбцу матрицы линейной комбинации других строк или столбцов определитель не меняется.

3. Если какая-либо строка или столбец матрицы является линейной комбинацией других строк или столбцов, то определитель матрицы равен нулю. Соответственно более простые случаи: если в матрице есть совпадающие строки, или совпадающие столбцы, или нулевая строка, или нулевой столбец, то ее определитель также равен нулю. Другими словами, если  $\text{rang } A < n$ , то  $\det A = 0$ .

4. Определитель треугольной матрицы равен произведению ее диагональных элементов.

$$5. \det(A \cdot k) = k \cdot \det A.$$

$$6. \det(A \cdot B) = \det A \cdot \det B.$$

$$7. \det(\text{diag } d_i) = \prod d_i.$$

$$8. \det A^T = \det A.$$

$$9. \det E = 1, \text{ где } E \text{ – единичная матрица.}$$

$$10. \det A = \prod \lambda_i, \text{ где } \lambda_i \text{ – собственные числа матрицы.}$$

**Миноры.** Предложим другой способ вычисления определителя.

*Минором*  $M_{j_1, j_2, \dots, j_k}^{i_1, i_2, \dots, i_k}$  порядка  $k$ ,  $k = 1, 2, \dots, n$ , квадратной матрицы  $A$  размера  $n \times n$  называется определитель матрицы, элементы которой стоят на пересечении строк с номерами  $i_1, i_2, \dots, i_k$  и столбцов с номерами  $j_1, j_2, \dots, j_k$  в исходной матрице  $A$ :

$$M_{j_1, j_2, \dots, j_k}^{i_1, i_2, \dots, i_k} = \begin{vmatrix} a_{i_1, j_1} & a_{i_1, j_2} & \dots & a_{i_1, j_k} \\ a_{i_2, j_1} & a_{i_2, j_2} & \dots & a_{i_2, j_k} \\ \dots & \dots & \dots & \dots \\ a_{i_k, j_1} & a_{i_k, j_2} & \dots & a_{i_k, j_k} \end{vmatrix}. \quad (\text{П.А.35})$$

Минор старшего порядка  $k = n$  равен определителю всей матрицы:

$$M_{1, 2, \dots, n}^{1, 2, \dots, n} = |A|.$$

Учитывая, что общее количество сочетаний  $k$  элементов из  $n$  определяется по комбинаторной формуле

$$C_n^k = \frac{n!}{k!(n-k)!}, \quad (\text{П.А.36})$$

общее количество миноров порядка  $k$  будет равно  $(C_n^k)^2$ .

*Пример.* Пусть дана матрица

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 4 & -3 & 7 \\ -1 & -1 & -1 \end{pmatrix}.$$

Из нее можно получить девять миноров порядка  $k = 1$ , девять миноров порядка  $k = 2$  и один минор порядка  $k = 3$ . Например:

$$M_1^1 = |1|,$$

$$M_2^1 = |-2|,$$

$$M_{1,2}^{1,2} = \begin{vmatrix} 1 & -2 \\ 4 & -3 \end{vmatrix},$$

$$M_{2,3}^{1,2} = \begin{vmatrix} -2 & 2 \\ -3 & 7 \end{vmatrix}$$

и т.д.

Если номера отмеченных строк совпадают с номерами отмеченных столбцов, то минор называется *главным*.

Очевидно, что количество таких миноров будет определяться формулой (П.А.36), т.е. будет равно  $C_n^k$ .

*Пример.* Для матрицы

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 4 & -3 & 7 \\ -1 & -1 & -1 \end{pmatrix}$$

главными минорами первого порядка будут

$$M_1^1 = |1|, \quad M_2^2 = |-3|, \quad M_3^3 = |-1|.$$

Главными минорами второго порядка будут

$$M_{1,2}^{1,2} = \begin{vmatrix} 1 & -2 \\ 4 & -3 \end{vmatrix}, \quad M_{1,3}^{1,3} = \begin{vmatrix} 1 & 2 \\ -1 & -1 \end{vmatrix}, \quad M_{2,3}^{2,3} = \begin{vmatrix} -3 & 7 \\ -1 & -1 \end{vmatrix}.$$

И единственным главным минором третьего порядка будет определитель всей матрицы

$$M_{1,2,3}^{1,2,3} = \begin{vmatrix} 1 & -2 & 2 \\ 4 & -3 & 7 \\ -1 & -1 & -1 \end{vmatrix}.$$

Если в главном миноре отмечены только  $k$  первых строк и столбцов, то он называется *угловым* или *ведущим главным* минором. Для углового минора  $M_{1,2,\dots,k}^{1,2,\dots,k}$  вводится обозначение  $\Delta_k$ :

$$\Delta_k = M_{1,2,\dots,k}^{1,2,\dots,k} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix}. \quad (\text{П.А.37})$$

Очевидно, что для любой матрицы существует только один угловой минор порядка  $k$ , а всего таких миноров  $n$ . Так, в последнем примере угловыми минорами являются  $\Delta_1 = M_1^1$ ,  $\Delta_2 = M_{1,2}^{1,2}$  и  $\Delta_3 = M_{1,2,3}^{1,2,3}$ .

Дополнительный минор  $\bar{M}_{ij}$  элемента квадратной матрицы  $a_{ij}$  размера  $n \times n$  есть определитель порядка  $n - 1$  для матрицы, полученной путем вычеркивания из матрицы  $A$  строки с номером  $i$  и столбца с номером  $j$ :

$$\bar{M}_{ij} = M_{1,2,\dots,i-1,i+1,\dots,n; 1,2,\dots,j-1,j+1,\dots,n}^{1,2,\dots,i-1,i+1,\dots,n} = \begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,j-1} & a_{2,j+1} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{n,n} \end{vmatrix}.$$

(П.А.38)

Очевидно, что количество дополнительных миноров будет равно количеству элементов матрицы –  $n^2$ .

*Пример.* Для матрицы

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 4 & -3 & 7 \\ -1 & -1 & -1 \end{pmatrix}$$

дополнительными минорами будут

$$\bar{M}_{11} = M_{2,3}^{2,3} = \begin{vmatrix} -3 & 7 \\ -1 & -1 \end{vmatrix},$$

$$\bar{M}_{12} = M_{1,3}^{2,3} = \begin{vmatrix} 4 & 7 \\ -1 & -1 \end{vmatrix}$$

и т.д.

Алгебраическим дополнением  $A_{ij}$  к элементу матрицы  $a_{ij}$  называется дополнительный минор  $\bar{M}_{ij}$ , взятый с обратным знаком, если сумма индексов  $i + j$  нечетная:

$$A_{ij} = (-1)^{i+j} \bar{M}_{ij}. \quad (\text{П.А.39})$$

С помощью алгебраических дополнений определитель можно вычислить, разложив его по произвольной строке  $i$ :

$$\Delta = \sum_{j=1}^n a_{ij} A_{ij} = \sum_{j=1}^n (-1)^{i+j} a_{ij} \bar{M}_{ij}. \quad (\text{П.А.40})$$

То же самое можно сделать и по столбцу. Обычно определитель расписывают по первой строке ( $i = 1$ ). Например, для матрицы порядка  $n = 3$  и  $i = 1$  получим

$$\begin{aligned} \Delta &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \\ &= (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{1+2} a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{1+3} a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}). \end{aligned}$$

Таким образом, получаем рекурсивную формулу вычисления определителя: определитель порядка  $n$  вычисляется с помощью миноров порядка  $n - 1$ , которые сами являются определителями, но уже меньшего порядка.

*Пример.* Найдем определитель матрицы

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 4 & -3 & 7 \\ -1 & -1 & -1 \end{pmatrix}.$$

Распишем его по строке  $i = 2$ :

$$\begin{aligned} \Delta &= (-1)^{2+1} \cdot 4 \cdot \begin{vmatrix} -2 & 2 \\ -1 & -1 \end{vmatrix} + (-1)^{2+2} \cdot (-3) \cdot \begin{vmatrix} 1 & 2 \\ -1 & -1 \end{vmatrix} + (-1)^{2+3} \cdot 7 \cdot \begin{vmatrix} 1 & -2 \\ -1 & -1 \end{vmatrix} = \\ &= -4 \cdot (2 + 2) - 3 \cdot (-1 + 2) - 7 \cdot (-1 - 2) = -16 - 3 + 21 = 2. \end{aligned}$$

*Базисным минором* матрицы называется любой ее ненулевой минор максимального порядка. Чтобы минор был базисным, необходимо

и достаточно, чтобы все окаймляющие его миноры (то есть содержащие его миноры на единицу большего порядка) были равны нулю.

Соответственно порядок базисного минора будет являться рангом квадратной матрицы.

**Обратные матрицы.** Квадратная матрица  $A$  называется *невырожденной*, если она имеет единственную *обратную* матрицу  $A^{-1}$ , определяемую из условия

$$A \cdot A^{-1} = A^{-1} \cdot A = I.$$

Обратная матрица существует не для всех матриц. Необходимым и достаточным условием невырожденности является  $\det A \neq 0$  или  $\text{rang } A = n$ .

Обратная матрица определяется следующим соотношением:

$$a_{ij}^{(-1)} = \frac{1}{\det A} \bar{A}^T, \quad (\text{П.А.41})$$

или

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}^T =$$

$$= \frac{1}{\det A} \begin{pmatrix} \bar{M}_{11} & -\bar{M}_{12} & \dots & (-1)^{1+n} \bar{M}_{1n} \\ -\bar{M}_{21} & \bar{M}_{22} & \dots & (-1)^{2+n} \bar{M}_{2n} \\ \dots & \dots & \dots & \dots \\ (-1)^{n+1} \bar{M}_{n1} & (-1)^{n+2} \bar{M}_{n2} & \dots & \bar{M}_{nn} \end{pmatrix}^T,$$

где  $\bar{A}^T$  – транспонированная матрица алгебраических дополнений.

Так, для квадратной матрицы порядка  $n = 2$  можно записать частную формулу

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} \bar{M}_{11} & -\bar{M}_{12} \\ -\bar{M}_{21} & \bar{M}_{22} \end{pmatrix}^T = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}^T =$$

$$= \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Для матрицы порядка  $n = 3$  получим

$$\begin{aligned} A^{-1} &= \frac{1}{\det A} \begin{pmatrix} \bar{M}_{11} & -\bar{M}_{12} & \bar{M}_{13} \\ -\bar{M}_{21} & \bar{M}_{22} & -\bar{M}_{23} \\ \bar{M}_{31} & -\bar{M}_{32} & \bar{M}_{33} \end{pmatrix}^T = \\ &= \frac{1}{\det A} \begin{pmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\ \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{pmatrix}^T = \\ &= \frac{1}{\det A} \begin{pmatrix} a_{22}a_{33} - a_{23}a_{32} & -(a_{21}a_{33} - a_{23}a_{31}) & a_{21}a_{32} - a_{22}a_{31} \\ -(a_{12}a_{33} - a_{13}a_{32}) & a_{11}a_{33} - a_{13}a_{31} & -(a_{11}a_{32} - a_{12}a_{31}) \\ a_{12}a_{23} - a_{13}a_{22} & -(a_{11}a_{23} - a_{13}a_{21}) & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix}^T = \\ &= \frac{1}{\det A} \begin{pmatrix} a_{22}a_{33} - a_{23}a_{32} & a_{13}a_{32} - a_{12}a_{33} & a_{12}a_{23} - a_{13}a_{22} \\ a_{23}a_{31} - a_{21}a_{33} & a_{11}a_{33} - a_{13}a_{31} & a_{13}a_{21} - a_{11}a_{23} \\ a_{21}a_{32} - a_{22}a_{31} & a_{12}a_{31} - a_{11}a_{32} & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix}, \\ &= \frac{1}{\det A} \begin{pmatrix} a_{22}a_{33} - a_{23}a_{32} & a_{13}a_{32} - a_{12}a_{33} & a_{12}a_{23} - a_{13}a_{22} \\ a_{23}a_{31} - a_{21}a_{33} & a_{11}a_{33} - a_{13}a_{31} & a_{13}a_{21} - a_{11}a_{23} \\ a_{21}a_{32} - a_{22}a_{31} & a_{12}a_{31} - a_{11}a_{32} & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix} \end{aligned}$$

и т.д.

Приведем некоторые леммы без доказательств:

- 1)  $(A \cdot B)^{-1} = B^{-1}A^{-1}$ ;
- 2)  $(A^T)^{-1} = (A^{-1})^T$ ;
- 3)  $(k \cdot A)^{-1} = \frac{1}{k}A^{-1}$ , где  $k$  – скаляр;

$$4) \det A^{-1} = \frac{1}{\det A};$$

$$5) E^{-1} = E.$$

*Пример.* Обратить матрицу

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 4 & -3 & 7 \\ -1 & -1 & -1 \end{pmatrix}.$$

Ее определитель мы уже нашли ранее:  $\det A = 2$ . Таким образом,

$$\begin{aligned} A^{-1} &= \frac{1}{2} \begin{pmatrix} \begin{vmatrix} -3 & +7 \\ -1 & -1 \end{vmatrix} & -\begin{vmatrix} +4 & +7 \\ -1 & -1 \end{vmatrix} & \begin{vmatrix} +4 & -3 \\ -1 & -1 \end{vmatrix} \\ -\begin{vmatrix} -2 & +2 \\ -1 & -1 \end{vmatrix} & \begin{vmatrix} +1 & +2 \\ -1 & -1 \end{vmatrix} & -\begin{vmatrix} +1 & -2 \\ -1 & -1 \end{vmatrix} \\ \begin{vmatrix} -2 & +2 \\ -3 & +7 \end{vmatrix} & -\begin{vmatrix} +1 & +2 \\ +4 & +7 \end{vmatrix} & \begin{vmatrix} +1 & -2 \\ +4 & -3 \end{vmatrix} \end{pmatrix}^T = \\ &= \frac{1}{2} \begin{pmatrix} 10 & -3 & -7 \\ -4 & 1 & 3 \\ -8 & 1 & 5 \end{pmatrix}^T = \begin{pmatrix} 5,0 & -2,0 & -4,0 \\ -1,5 & 0,5 & 0,5 \\ -3,5 & 1,5 & 2,5 \end{pmatrix}. \end{aligned}$$

Если использовать полученную выше формулу, результат будет аналогичным.

### Операции с векторами

Большинство операций над векторами аналогично операциям над матрицами и выполняется по тем же правилам – умножение и деление вектора на скаляр, сложение и вычитание векторов, умножение вектора на матрицу или матрицы на вектор, вычисление нормы вектора.

Матричное умножение вектора на вектор обычно лишено смысла, потому что согласно правилам матричного произведения найти произведение  $x \cdot y$ , где  $x$  и  $y$  – векторы-столбцы, можно только в том случае, если оба этих вектора имеют размер  $n = 1$ , т.е., по сути, являются скалярами.

Вместо этого для векторов вводят другие операции – *векторное произведение* и *скалярное произведение*. Мы будем рассматривать только операцию скалярного произведения.

Скалярное произведение есть операция над двумя векторами, результатом которой является число (скаляр), не зависящее от системы координат и являющееся характеристикой этих векторов.

Скалярное произведение двух векторов  $x$  и  $y$ , имеющих одинаковый размер  $n$ , обозначается как  $(x, y)$  и вычисляется следующим образом:

$$(x, y) = \sum_{i=1}^n x_i y_i. \quad (\text{П.А.42})$$

Через матричное произведение данную операцию можно записать в виде

$$x^T y = (x_1 \quad x_2 \quad \dots \quad x_n) \cdot \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \sum_{i=1}^n x_i y_i.$$

*Пример 1.* Вычислить произведение  $xAx$ , где

$$x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix}.$$

Данное произведение вычислить невозможно, так как в произведении  $xA$  первый компонент имеет матричный размер  $2 \times 1$ , а второй –  $2 \times 2$  и  $1 \neq 2$ .

*Пример 2.* Вычислить произведение  $x^T Ax$ , где

$$x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix}.$$

Вычисляем по правилам матричного произведения:

$$\begin{aligned} x^T Ax &= (1 \quad 2) \cdot \begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = (1 \cdot 1 + 2 \cdot 3 \quad 1 \cdot (-2) + 2 \cdot 1) \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \\ &= (7 \quad 0) \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 7 \cdot 1 + 0 \cdot 2 = 7. \end{aligned}$$

Если вектор  $x$  размера  $n$  задает направление в пространстве  $R^n$ , то он имеет *длину*. То есть если начало вектора расположено в точке  $A(a_1, a_2, \dots, a_n)$ , а конец – в точке  $B(b_1, b_2, \dots, b_n)$ , то

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ \dots \\ b_n - a_n \end{pmatrix}.$$

Тогда длина отрезка  $AB$  и будет длиной (или *модулем*) вектора:

$$|x| = \sqrt{\sum_{i=1}^n x_i^2}. \quad (\text{П.А.43})$$

Очевидно, что одну и ту же координату вектора  $x_i$  можно представить в виде разности множеством способов, поэтому разным точкам  $A$  и  $B$  может соответствовать один и тот же вектор.

*Нормировка вектора* – это операция деления вектора на его длину:

$$y = \frac{x}{|x|}, \quad |x| \neq 0. \quad (\text{П.А.44})$$

После этого вектор  $y$  называется *нормированным*. Длина нормированного вектора равна 1:

$$|y| = 1.$$

*Пример.* Выполнить нормировку вектора

$$x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

Найдем длину вектора  $x$ :

$$|x| = \sqrt{3^2 + 4^2} = 5.$$

Следовательно,

$$y = \frac{x}{|x|} = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}.$$

Проверка:

$$|y| = \sqrt{\left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2} = \sqrt{\frac{9+16}{25}} = 1.$$

Результат скалярного произведения вектора самого на себя соотносится с длиной вектора:

$$|x| = \sqrt{(x, x)} = \sqrt{\sum_{i=1}^n x_i^2}. \quad (\text{П.А.45})$$

Если исходить из определения нормы матрицы (П.А.33), то длина (модуль) и евклидова норма вектора совпадают:

$$|x| = \|x\|.$$

Но не следует забывать, что это лишь одно из определений нормы, существуют и другие, не гарантирующие выполнение данного равенства.

#### П.А.4. Дифференцирование функций

Если две переменные величины связаны между собой так, что каждому значению одной из них соответствует определенное значение другой, то говорят, что между ними существует *функциональная зависимость*. При этом переменная, которая может принимать произвольные допустимые значения, называется *независимой переменной* или *аргументом*. Другая величина, значения которой зависят от значений аргумента, называется *зависимой переменной* или *функцией*. Для обозначения функции, аргумента и функциональной зависимости можно использовать любые символы, например

$$y = f(x), \quad (\text{П.А.46})$$

где  $x$  – значение аргумента;  $y$  – значение функции;  $f$  – обозначение функциональной зависимости.

Функция может иметь несколько аргументов, например

$$z = f(x, y).$$

Это трактуется следующим образом: аргумент функции может быть как скалярной величиной ( $x \in R^1$ ), так и векторной ( $x \in R^n$ ).

Тогда

$$y = f(x), \quad x = (x_1, x_2, \dots, x_n)^T. \quad (\text{П.А.47})$$

В противовес функциональной существует *параметрическая зависимость* переменных величин, когда одному значению какой-либо из них может соответствовать множество значений другой:

$$f(x_1, x_2, \dots, x_n) = \text{const} \text{ или } f(x_1, x_2, \dots, x_n) = 0.$$

Например, зависимость

$$x^2 + y^2 = R^2$$

описывает точки, лежащие на окружности радиуса  $R$ . При этом каждому значению переменной  $x$  соответствуют два значения переменной  $y$  и наоборот:

$$y = \pm\sqrt{R^2 - x^2}, \quad x = \pm\sqrt{R^2 - y^2}.$$

Такую функциональную зависимость можно записать в виде

$$f(x, y) = \text{const},$$

где  $f(x, y) = x^2 + y^2$ ,  $\text{const} = R^2$ , или

$$f(x, y) = 0,$$

где  $f(x, y) = x^2 + y^2 - R^2$ .

### Производная функции скалярного аргумента

Производной функции  $f(x)$  скалярного аргумента  $x$  в точке  $x_0$  называется число, к которому стремится отношение приращения функции  $\Delta f(x)$  к приращению аргумента  $\Delta x$  при стремлении  $\Delta x$  к нулю:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta f(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Производная функции – это тоже функция. Производную функции скалярного аргумента  $f(x)$  записывают следующим образом:

$$\frac{df}{dx} \text{ или } f'(x).$$

В произвольной точке  $x$  имеем

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}. \quad (\text{П.А.48})$$

В таблице П.А.1 даны производные некоторых элементарных функций.

Таблица П.А.1 – Производные некоторых элементарных функций

Функция	Производная
const	0
$x$	1
$x^n$	$nx^{n-1}$
$\frac{1}{x^n} = x^{-n}$	$-nx^{-n-1} = -\frac{n}{x^{n+1}}$
$\sqrt[n]{x} = x^{\frac{1}{n}}$	$\frac{1}{n}x^{\frac{1}{n}-1} = \frac{1}{n}x^{\frac{1-n}{n}} = \frac{1}{n}\sqrt[n]{x^{1-n}} = \frac{1}{n\sqrt[n]{x^{n-1}}}$
$a^x$	$\ln a \cdot a^x$
$\log_n x$	$\frac{1}{x \ln n}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$

Для определения производной сложной функции используются соотношения из таблицы П.А.2. Здесь  $u(x)$  и  $v(x)$  – произвольные функции,  $n$  – константа.

Некоторые формулы являются частными случаями других. Все формулы таблицы П.А.1 можно получить из таблицы П.А.2, если учесть, что  $x' = 1$ , а  $\text{const}' = 0$ .

Таблица П.А.2 – Производные сложных функций

Номер	Функция	Производная
1	$n \cdot u(x)$	$n \cdot u'(x)$
2	$u(x) + v(x)$	$u'(x) + v'(x)$
3	$u(x) - v(x)$	$u'(x) - v'(x)$
4	$u(x) \cdot v(x)$	$u'(x) \cdot v(x) + u(x) \cdot v'(x)$
5	$\frac{u(x)}{v(x)}$	$\frac{u'(x) \cdot v(x) - u(x) \cdot v'(x)}{v^2(x)}$
6	$u^n(x)$	$n \cdot u^{n-1}(x) \cdot u'(x)$
7	$n^{u(x)}$	$\ln n \cdot n^{u(x)} \cdot u'(x)$
8	$\log_n u(x)$	$\frac{u'(x)}{u(x) \cdot \ln n}$
9	$\sin u(x)$	$u'(x) \cdot \cos u(x)$
10	$\cos u(x)$	$-u'(x) \cdot \sin u(x)$
11	$u(x)^{v(x)}$	$u(x)^{v(x)} \left( u'(x) \cdot \frac{v(x)}{u(x)} + v'(x) \cdot \ln u(x) \right)$
12	$u(v(x))$	$u'(v) \cdot v'(x)$

*Пример.* Найдем производную функции

$$f(x) = \sqrt{x}.$$

Если представить эту функцию в виде  $f(x) = u(x)^{v(x)}$ , где  $u(x) = x$ ,  $v(x) = 1/2$ , то согласно пункту 11 таблицы П.А.2 получим

$$f'(x) = \sqrt{x} \cdot \left( x' \cdot \frac{1/2}{x} + (1/2)' \cdot \ln x \right).$$

С учетом того что  $x' = 1$ ,  $(1/2)' = 0$ , находим

$$f'(x) = \sqrt{x} \cdot \frac{1/2}{x} = \frac{1}{2\sqrt{x}}.$$

Конечно, проще эту формулу можно получить согласно пункту 6 таблицы П.А.2, если принять  $u(x) = x$ ,  $n = 1/2$ :

$$f'(x) = \sqrt{x} \cdot \left( x' \cdot \frac{1/2}{x} + (1/2)' \cdot \ln x \right).$$

Согласно пункту 12 таблицы П.А.2, чтобы найти производную сложной функции  $u(v(x))$ , необходимо сначала вычислить производную функции  $u(v)$  так, как будто это функция аргумента  $v$ , а затем результат умножить на производную функции  $v(x)$ .

*Пример 1.* Найти производную сложной функции

$$f(x) = \sin \ln x.$$

Сначала представим функцию в виде  $f(y) = \sin y$ , где  $y(x) = \ln x$ . Тогда

$$f'(y) = \cos y, \quad y'(x) = \frac{1}{x} \Rightarrow$$

$$f'(x) = \frac{1}{x} \cos y(x) = \frac{\cos \ln x}{x}.$$

*Пример 2.* Найти производную сложной функции

$$f(x) = \ln \sqrt{x^2 + 1}.$$

Сначала представим функцию в виде  $f(y) = \ln y$ , где  $y(x) = \sqrt{x^2 + 1}$ . Тогда

$$f'(y) = \frac{1}{y} \Rightarrow f'(x) = \frac{y'(x)}{y}.$$

В свою очередь  $y(x)$  тоже сложная функция, т.е.  $y(z) = \sqrt{z}$ , где  $z(x) = x^2 + 1$ . Тогда

$$y'(z) = \frac{1}{2\sqrt{z}} \Rightarrow y'(x) = \frac{z'(x)}{2\sqrt{z}}.$$

Согласно пункту 2 таблицы П.А.2

$$z'(x) = (x^2 + 1)' = (x^2)' + (1)' = 2x.$$

Окончательно имеем

$$f'(x) = \frac{y'(x)}{y} = \frac{z'(x)}{y(x) \cdot 2\sqrt{z(x)}} = \frac{2x}{\sqrt{x^2 + 1} \cdot 2 \cdot \sqrt{x^2 + 1}} = \frac{x}{x^2 + 1}.$$

Физический смысл производной – это скорость изменения функции. Так, например, если функция  $S(t)$  описывает длину пути, пройденного телом за время  $t$ , то производная этой функции  $v(t) = S'(t)$  будет соответствовать скорости тела в любой момент времени  $t$ . Скорость изменения скорости в физике – это ускорение, поэтому  $a(t) = v'(t)$ . То есть ускорение – это производная от производной функции  $S(t)$ , или *вторая производная*.

Таким образом, производную любого порядка  $n$  можно получить как производную от производной более низкого порядка  $n-1$ :

$$f^{(n)}(x) = \left( f^{(n-1)}(x) \right)', \quad n > 1. \quad (\text{П.А.49})$$

Для второй и третьей производных обычно используют обозначения  $f''(x)$  и  $f'''(x)$ . Для производных более высокого порядка используют обозначение  $f^{(n)}(x)$ , где  $n$  – порядок производной.

Геометрический смысл производной заключается в том, что значение производной функции  $f(x)$  в точке  $x_0$  равно тангенсу угла касательной к функции  $f(x)$  в точке  $x_0$  (рисунок П.А.1):

$$f'(x_0) = \operatorname{tg} \alpha.$$

Как известно, критическая точка – это точка, в которой функция  $f(x)$  непрерывна, а производная равна нулю или терпит разрыв. Соответственно для поиска критических точек функции  $f(x)$  необходимо решить уравнение

$$f'(x) = 0, \quad (\text{П.А.50})$$

а также найти точки разрыва производной  $f'(x)$ .

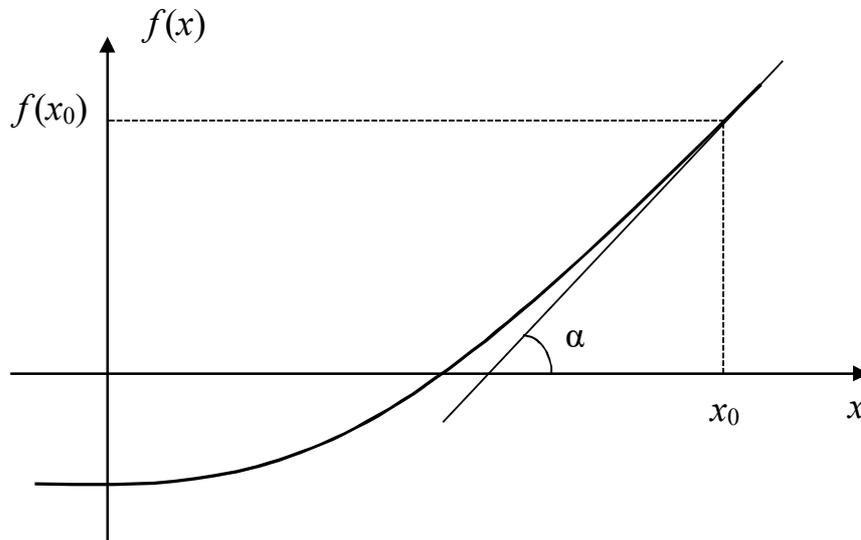


Рисунок П.А.1 – Геометрический смысл производной

Для методов оптимизации важно, что если производная положительна, то функция растет, а если отрицательна – убывает. Это используется при поиске экстремумов для определения направления поиска.

Если слева от критической точки производная положительна, а справа – отрицательна, имеем точку *максимума*, если наоборот, – точку *минимума*. Если производная существует и не меняет свой знак, то в точке наблюдается *перегиб* функции.

Для поиска точек перегиба необходимо решить уравнение

$$f''(x) = 0. \quad (\text{П.А.51})$$

Если слева и справа от какого-либо корня уравнения (П.А.51) знаки второй производной различаются, это будет точка перегиба.

*Пример 1.* Найти и классифицировать точки экстремума функции, а также ее точки перегиба:

$$f(x) = 2x^3 + 3x^2 - 12x.$$

Разрывов функция не имеет. Решим уравнение (П.А.50):

$$f'(x) = 6x^2 + 6x - 12 = 0.$$

Его решениями являются точки  $x=1$  и  $x=-2$ . Исследуем их характер:

Интервал	$(-\infty, -2]$	$[-2, 1]$	$[1, +\infty)$
Знак $f'(x)$	+	-	+

Следовательно, в точке  $x=-2$  функция имеет максимум, а в точке  $x=1$  – минимум. Используем необходимое и достаточное условие. Для этого найдем вторую производную:

$$f''(x) = 12x + 6.$$

В точке  $x=-2$  имеем  $f''(-2) = 18 < 0$ , следовательно, это точка максимума. В точке  $x=1$  имеем  $f''(1) = 6 > 0$ , следовательно, это точка минимума. Получили аналогичные результаты.

Для поиска точек перегиба решим уравнение (П.А.51):

$$f''(x) = 12x + 6 = 0.$$

Его решением является точка  $x = -1/2$ . Исследуем ее характер:

Интервал	$(-\infty, -1/2]$	$[-1/2, +\infty)$
Знак $f''(x)$	-	+

Знаки  $f''(x)$  отличаются, следовательно,  $x = -1/2$  – точка перегиба.

*Пример 2.* Найти и классифицировать точки экстремума функции, а также ее точки перегиба:

$$f(x) = \sqrt{|x|} = \begin{cases} \sqrt{x}, & x \geq 0, \\ \sqrt{-x}, & x < 0. \end{cases}$$

Уравнение (П.А.50)

$$f'(x) = \frac{\text{sign } x}{2\sqrt{|x|}} = 0, \quad \text{или} \quad \begin{cases} \frac{1}{2\sqrt{x}} = 0, & x \geq 0, \\ -\frac{1}{2\sqrt{-x}} = 0, & x < 0, \end{cases}$$

не имеет решений, однако в точке  $x = 0$  функция имеет разрыв. То есть точка  $x = 0$  является критической. Проверим, экстремум ли это:

Интервал	$(-\infty, 0)$	0	$(0, +\infty)$
Знак $f'(x)$	-	разрыв	+

Следовательно  $x = 0$  – точка минимума. Для поиска точек перегиба решим уравнение (П.А.51):

$$f''(x) = \begin{cases} -\frac{1}{4\sqrt{x^3}} = 0, & x \geq 0, \\ -\frac{1}{4\sqrt{-x^3}} = 0, & x < 0. \end{cases}$$

Данное уравнение корней не имеет, следовательно, у функции нет точек перегиба.

Решение уравнений вида (4.50) используется также для одномерной оптимизации при поиске длины шага в методах безусловной и условной оптимизации функции векторного аргумента. Примеры были рассмотрены ранее.

Неравенства с производными используются для анализа выпуклости и вогнутости функций скалярного аргумента. Известно, что функция выпукла на отрезке  $[a, b]$ , если на этом отрезке:

1) выполняется неравенство Йенссена

$$f[\mu x_1 + (1 - \mu)x_2] \leq \mu f(x_1) + (1 - \mu)f(x_2) \quad \forall x_1, x_2 \in [a, b], \quad \forall \mu \in [0, 1];$$

2) первая производная функции не убывает, т.е.

$$f'(x_1) \leq f'(x_2) \quad \forall x_1 < x_2, \quad x_1, x_2 \in [a, b];$$

3) вторая производная функции положительна, т.е.

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

Если отрезок не ограничен, доказывать неравенство нужно для  $\forall x \in R^1$ . Первый критерий является необходимым и достаточным, второй и третий критерии – достаточными.

*Пример.* Доказать выпуклость функции

$$f(x) = x^2 + 1$$

тремя описанными выше способами.

1. Используем неравенство Йенссена:

$$\begin{aligned}(\mu x_1 + (1-\mu)x_2)^2 + 1 &\leq \mu(x_1^2 + 1) + (1-\mu)(x_2^2 + 1) \Rightarrow \\ \mu^2 x_1^2 + 2\mu(1-\mu)x_1x_2 + (1-\mu)^2 x_2^2 + 1 &\leq \mu x_1^2 + \mu + (1-\mu)x_2^2 + 1 - \mu \Rightarrow \\ (\mu^2 - \mu)x_1^2 + 2\mu(1-\mu)x_1x_2 + [(1-\mu)^2 - (1-\mu)]x_2^2 &\leq 0 \Rightarrow \\ -\mu(1-\mu)x_1^2 + 2\mu(1-\mu)x_1x_2 + (1-2\mu + \mu^2 - 1 + \mu)x_2^2 &\leq 0 \Rightarrow \\ \mu(1-\mu)x_1^2 - 2\mu(1-\mu)x_1x_2 + \mu(1-\mu)x_2^2 &\geq 0 \Rightarrow \\ \mu(1-\mu)(x_1^2 - 2x_1x_2 + x_2^2) &\geq 0 \Rightarrow \mu(1-\mu)(x_1 - x_2)^2 \geq 0.\end{aligned}$$

Так как при любом  $\mu \in [0,1]$  первые два множителя не будут отрицательными, а третий множитель – в квадрате, т.е. тоже не может быть отрицательным, неравенство доказано.

2. Используем критерий с первой производной:

$$\begin{aligned}f'(x) &= 2x, \\ 2x_1 &\leq 2x_2 \Rightarrow 2(x_2 - x_1) \geq 0.\end{aligned}$$

Для любого  $x_2 > x_1$  выполнение этого неравенства очевидно.

3. Используем критерий со второй производной:

$$f''(x) = 2.$$

Очевидно, что  $f''(x) > 0$  для любого значения  $x$ .

Применение критерия Йенссена выглядит более трудоемким, но это следствие его универсальности. Критерии на основе производных нельзя применить, если производные функции не существуют или имеют разрывы на интервале  $[a,b]$ .

## Частные производные

Если функция имеет несколько аргументов, а производная берется только по одному из них, то вводится понятие *частной производной*. Например, если имеется функция двух аргументов  $f(x, y)$ , то ее частная производная по аргументу  $x$  обозначается

$$\frac{\partial f}{\partial x}, \text{ или } f'_x(x, y), \text{ или } \nabla_x f(x, y).$$

Соответственно для частной производной по аргументу  $y$  имеем

$$\frac{\partial f}{\partial y}, \text{ или } f'_y(x, y), \text{ или } \nabla_y f(x, y).$$

При взятии частной производной по какому-либо аргументу все остальные аргументы функции считаются константами.

*Пример.* Найти частные производные функции

$$f(x, y) = x^2 - xy^3.$$

При взятии производной по аргументу  $x$  считаем аргумент  $y$  константой:

$$\frac{\partial f}{\partial x} = 2x - y^3.$$

При взятии производной по аргументу  $y$  считаем константой аргумент  $x$ :

$$\frac{\partial f}{\partial y} = -3xy^2.$$

**Градиент.** Для функции векторного аргумента (П.А.47) существует  $n$  частных производных:

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}.$$

Вектор-столбец, составленный из таких частных производных, называется *градиентом* и обозначается

$$\text{grad } f(x) = \nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = \left( \frac{\partial f}{\partial x_1}; \frac{\partial f}{\partial x_2}; \dots; \frac{\partial f}{\partial x_n} \right)^T. \quad (\text{П.А.52})$$

*Пример.* Найти градиент функции

$$f(x) = \sqrt{x_1^2 + x_2^2}.$$

По правилам взятия производной сложных функций получаем

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \frac{\partial(x_1^2 + x_2^2)}{\partial x_1} \cdot \frac{1}{2\sqrt{x_1^2 + x_2^2}} \\ \frac{\partial(x_1^2 + x_2^2)}{\partial x_2} \cdot \frac{1}{2\sqrt{x_1^2 + x_2^2}} \end{pmatrix} = \begin{pmatrix} \frac{x_1}{\sqrt{x_1^2 + x_2^2}} \\ \frac{x_2}{\sqrt{x_1^2 + x_2^2}} \end{pmatrix}.$$

Градиент используется во всех методах безусловной оптимизации 1-го и 2-го порядков функции векторного аргумента, для проверки НУ и ДУ оптимальности, а также при решении ЗНП. При решении ЗНП используется не только градиент ЦФ, но и градиенты ограничений задачи.

Также градиент используется для аналитического поиска критических точек функций векторного аргумента. Согласно определению в критической точке непрерывной функции  $f(x)$  ее градиент равен нулю или не существует (это необходимое условие экстремума). Соответственно для определения критических точек необходимо решить систему уравнений

$$\nabla f(x) = 0, \quad (\text{П.А.53})$$

а также найти все точки разрыва градиента  $\nabla f(x)$ . Данную систему можно решать как численно, так и аналитически.

*Пример.* Найти аналитически критические точки функции Розенброка

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Найдем градиент ЦФ:

$$\begin{aligned} \nabla f(x) &= \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 100 \cdot 2 \frac{\partial(x_2 - x_1^2)}{\partial x_1} (x_2 - x_1^2) + 2 \frac{\partial(1 - x_1)}{\partial x_1} (1 - x_1) \\ 100 \cdot 2 \frac{\partial(x_2 - x_1^2)}{\partial x_2} (x_2 - x_1^2) + 0 \end{pmatrix} = \\ &= \begin{pmatrix} -400x_1(x_2 - x_1^2) - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{pmatrix}. \end{aligned}$$

Разрывов градиент не имеет. Решим систему уравнений (П.А.53):

$$\begin{cases} -400x_1(x_2 - x_1^2) + 2(1 - x_1) = 0, \\ 200(x_2 - x_1^2) = 0. \end{cases}$$

Из второго уравнения получим  $x_2 = x_1^2$ . Подставим в первое уравнение:

$$-400x_1(x_1^2 - x_1^2) - 2(1 - x_1) = 0 \Rightarrow 2(1 - x_1) = 0 \Rightarrow \bar{x}_1 = 1,$$

$$x_2 = x_1^2 \Rightarrow \bar{x}_2 = 1.$$

Таким образом,  $\bar{x} = (1; 1)^T$ ,  $\bar{f} = 0$ . Далее будет выполнена классификация данной критической точки.

**Матрица Якоби.** Пусть имеется система уравнений (П.А.23). Тогда для каждой функции  $f_i(x)$ , входящей в систему, можно вычислить частную производную по каждому неизвестному  $x_j$ :

$$w_{ij} = \frac{\partial f_i}{\partial x_j}.$$

Матрица  $W(x)$ , составленная из элементов  $w_{ij}$ , называется *матрицей Якоби* или *якобианом*:

$$W(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}. \quad (\text{П.А.54})$$

Матрица Якоби используется, в частности, при численном решении систем нелинейных уравнений, например системы (П.А.53) и др.

*Пример.* Найти матрицу Якоби системы нелинейных уравнений

$$\begin{cases} xy = -5, \\ \sqrt{x} + \sqrt{y} = 2. \end{cases}$$

Перепишем систему в виде

$$\begin{cases} f_1(x, y) = xy + 5 = 0, \\ f_2(x, y) = \sqrt{x} + \sqrt{y} - 2 = 0. \end{cases}$$

Тогда

$$W(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} y & x \\ \frac{1}{2\sqrt{x}} & \frac{1}{2\sqrt{y}} \end{pmatrix}.$$

**Матрица Гессе.** Как градиент функции векторного аргумента является аналогом первой производной функции скалярного аргумента, так *матрица Гессе* (или *гессиан*) функции векторного аргумента является аналогом второй производной функции скалярного аргумента.

Чтобы найти матрицу Гессе  $H_f(x)$  функции  $f(x)$ , необходимо взять частную производную каждого компонента вектора градиента  $\partial f / \partial x_i$  по каждому неизвестному  $x_j$ :

$$\begin{aligned}
 H_f(x) &= \begin{pmatrix} \frac{\partial}{\partial x_1} \left( \frac{\partial f}{\partial x_1} \right) & \frac{\partial}{\partial x_2} \left( \frac{\partial f}{\partial x_1} \right) & \cdots & \frac{\partial}{\partial x_n} \left( \frac{\partial f}{\partial x_1} \right) \\ \frac{\partial}{\partial x_1} \left( \frac{\partial f}{\partial x_2} \right) & \frac{\partial}{\partial x_2} \left( \frac{\partial f}{\partial x_2} \right) & \cdots & \frac{\partial}{\partial x_n} \left( \frac{\partial f}{\partial x_2} \right) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial}{\partial x_1} \left( \frac{\partial f}{\partial x_n} \right) & \frac{\partial}{\partial x_2} \left( \frac{\partial f}{\partial x_n} \right) & \cdots & \frac{\partial}{\partial x_n} \left( \frac{\partial f}{\partial x_n} \right) \end{pmatrix} \Rightarrow \\
 H_f(x) &= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}. \quad (\text{П.А.55})
 \end{aligned}$$

Эта матрица симметричная, т.е.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

*Замечание.* Понятие производных определено только для функции скалярного аргумента. Для функции векторного аргумента вместо этого используются понятия градиента и гессиана.

*Пример.* Найти матрицу Гессе функции

$$f(x) = \sqrt{x_1^2 + x_2^2}.$$

Ранее мы уже находили градиент данной функции:

$$\nabla f(x) = \begin{pmatrix} \frac{x_1}{\sqrt{x_1^2 + x_2^2}} \\ \frac{x_2}{\sqrt{x_1^2 + x_2^2}} \end{pmatrix}.$$

Соответственно

$$\begin{aligned}
 H_f(x) &= \begin{pmatrix} \frac{\partial}{\partial x_1} \left( \frac{x_1}{\sqrt{x_1^2 + x_2^2}} \right) & \frac{\partial}{\partial x_2} \left( \frac{x_1}{\sqrt{x_1^2 + x_2^2}} \right) \\ \frac{\partial}{\partial x_1} \left( \frac{x_2}{\sqrt{x_1^2 + x_2^2}} \right) & \frac{\partial}{\partial x_2} \left( \frac{x_2}{\sqrt{x_1^2 + x_2^2}} \right) \end{pmatrix} = \\
 &= \begin{pmatrix} \frac{1}{\sqrt{x_1^2 + x_2^2}} - \frac{1}{2} \cdot \frac{x_1 \cdot 2x_1}{\sqrt{(x_1^2 + x_2^2)^3}} & -\frac{1}{2} \cdot \frac{x_1 \cdot 2x_2}{\sqrt{(x_1^2 + x_2^2)^3}} \\ -\frac{1}{2} \cdot \frac{x_2 \cdot 2x_1}{\sqrt{(x_1^2 + x_2^2)^3}} & \frac{1}{\sqrt{x_1^2 + x_2^2}} - \frac{1}{2} \cdot \frac{x_2 \cdot 2x_2}{\sqrt{(x_1^2 + x_2^2)^3}} \end{pmatrix} = \\
 &= \begin{pmatrix} \frac{x_2^2}{\sqrt{(x_1^2 + x_2^2)^3}} & -\frac{x_1 x_2}{\sqrt{(x_1^2 + x_2^2)^3}} \\ -\frac{x_1 x_2}{\sqrt{(x_1^2 + x_2^2)^3}} & \frac{x_1^2}{\sqrt{(x_1^2 + x_2^2)^3}} \end{pmatrix}.
 \end{aligned}$$

Матрица Гессе используется в методах безусловной оптимизации 2-го порядка функции векторного аргумента, для проверки НУ и ДУ оптимальности второго порядка, а также при классификации экстремальных точек функции векторного аргумента.

Так, если найдена критическая точка  $\bar{x}$  функции  $f(x)$ , т.е. точка, в которой градиент  $\nabla f(\bar{x}) = 0$  или имеет разрыв, то она может быть как точкой минимума или максимума, так и седловой точкой. Точка  $\bar{x}$  является точкой локального минимума, если в некоторой ее окрестности ЦФ  $f(x)$  является выпуклой. Это можно доказать следующими способами.

1. Доказать, что в некоторой  $\varepsilon$ -окрестности точки  $\bar{x}$  выполняется неравенство Иенссена:

$$f\left[\mu x^1 + (1-\mu)x^2\right] \leq \mu f(x^1) + (1-\mu)f(x^2)$$

$$\forall \|\bar{x} - x^1\| \leq \varepsilon, \quad \forall \|\bar{x} - x^2\| \leq \varepsilon, \quad \forall \mu \in [0,1].$$

2. Доказать, что в некоторой  $\varepsilon$ -окрестности точки  $\bar{x}$  минимальное значение ЦФ принимает в точке  $\bar{x}$ :

$$f(\bar{x}) \leq f(x) \quad \forall \|\bar{x} - x\| \leq \varepsilon, \quad \varepsilon > 0.$$

3. Показать, что в точке  $\bar{x}$  гессиан ЦФ  $H_f(\bar{x})$  положительно определен. Если гессиан лишь положительно полуопределен, то нужны дополнительные проверки – первая или вторая.

Точка  $\bar{x}$  является точкой локального максимума, если в некоторой ее окрестности ЦФ  $f(x)$  является вогнутой. Для доказательства можно использовать описанные выше способы, если функция  $f(x)$  отрицательна. В противном случае необходимо:

1) доказать, что в некоторой  $\varepsilon$ -окрестности точки  $\bar{x}$  выполняется неравенство Йенссена противоположного смысла:

$$f\left[\mu x^1 + (1-\mu)x^2\right] \geq \mu f(x^1) + (1-\mu)f(x^2)$$

$$\forall \|\bar{x} - x^1\| \leq \varepsilon, \quad \forall \|\bar{x} - x^2\| \leq \varepsilon, \quad \forall \mu \in [0,1];$$

2) доказать, что в некоторой  $\varepsilon$ -окрестности точки  $\bar{x}$  максимальное значение ЦФ принимает в точке  $\bar{x}$ :

$$f(\bar{x}) \geq f(x) \quad \forall \|\bar{x} - x\| \leq \varepsilon, \quad \varepsilon > 0;$$

3) показать, что в точке  $\bar{x}$  гессиан ЦФ  $H_f(\bar{x})$  отрицательно определен. Если гессиан лишь отрицательно полуопределен, то нужны дополнительные проверки – первая или вторая.

В противном случае имеем седловую точку. Если  $\bar{x}$  не седловая точка, а локальный минимум (максимум), то он одновременно будет глобальным минимумом (максимумом) при выполнении следующих условий.

1. Функция  $f(x)$  выпукла (вогнута) на всей области определения  $x \in D$ . То есть нужно показать, что неравенство Йенссена выполняется  $\forall x^1, x^2 \in D$ , либо  $f(\bar{x}) \leq f(x)$  ( $f(\bar{x}) \geq f(x)$ ) в любой точке  $x \in D$ , либо гессиан  $H_f(x)$  положительно (отрицательно) определен  $\forall x \in D$ . Это достаточный критерий.

2. Значение  $f(\bar{x})$  совпадает с точной нижней (верхней) границей значений функции  $f(x)$  на множестве аргументов  $x \in D$ :

$$f(\bar{x}) = \inf_{x \in D} (\sup) f(x).$$

Это необходимый и достаточный критерий. При поиске нижней и верхней границ значений функции проверяются ее точки разрывов, критические точки, а также предельные значения аргументов.

*Пример.* Выполнить классификацию критической точки  $\bar{x} = (1; 1)^T$  функции Розенброка

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Ранее уже был найден ее градиент:

$$\nabla f(x) = \begin{pmatrix} -400x_1(x_2 - x_1^2) - 2 \cdot (1 - x_1) \\ 200(x_2 - x_1^2) \end{pmatrix},$$

теперь найдем гессиан:

$$\begin{aligned} H_f(x) &= \\ &= \begin{pmatrix} \frac{\partial}{\partial x_1}(-400x_1(x_2 - x_1^2) - 2(1 - x_1)) & \frac{\partial}{\partial x_2}(-400x_1(x_2 - x_1^2) - 2(1 - x_1)) \\ \frac{\partial}{\partial x_1}(200(x_2 - x_1^2)) & \frac{\partial}{\partial x_2}(200(x_2 - x_1^2)) \end{pmatrix} = \\ &= \begin{pmatrix} -400(x_2 - x_1^2) + 800x_1^2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix} = \begin{pmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}. \end{aligned}$$

Вычислим значение гессиана в точке  $\bar{x}$ :

$$H_f(\bar{x}) = \begin{pmatrix} 1200 - 400 + 2 & -400 \\ -400 & 200 \end{pmatrix} = \begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix}.$$

Найдем угловые миноры:

$$\Delta_1 = 802 > 0, \quad \Delta_2 = \begin{vmatrix} 802 & -400 \\ -400 & 200 \end{vmatrix} = 400 > 0.$$

Гессиан положительно определен, следовательно,  $\bar{x} = (1; 1)^T$  — точка локального минимума. Установим, является ли он глобальным минимумом. Для этого нужно показать, что в любой точке  $x$  гессиан

$$H_f(x) = \begin{pmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}$$

положительно определен. Находим угловые миноры:

$$\begin{aligned} \Delta_1 &= 1200x_1^2 - 400x_2 + 2, \\ \Delta_2 &= \begin{vmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{vmatrix} = 200(1200x_1^2 - 400x_2 + 2) - 400^2 x_1^2 = \\ &= 80000x_1^2 - 80000x_2 + 400. \end{aligned}$$

Видно, что оба угловых минора незначительны, т.е. гессиан является неопределенным. Достаточный критерий не выполняется. Используем второй, необходимый и достаточный, критерий. Функция непрерывна, следовательно, своих экстремальных значений она может достичь в критической точке, а также при  $x_i \rightarrow \pm\infty$ . Имеем:

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

$$\lim_{x_1 \rightarrow \pm\infty} f(x) = \lim_{x_1 \rightarrow \pm\infty} 100(x_2 - x_1^2)^2 + (1 - x_1)^2 = +\infty,$$

$$\lim_{x_2 \rightarrow \pm\infty} f(x) = \lim_{x_2 \rightarrow \pm\infty} 100(x_2 - x_1^2)^2 + (1 - x_1)^2 = +\infty,$$

следовательно,

$$\inf f(x) = 0, \quad \sup f(x) = +\infty.$$

Так как  $f(\bar{x}) = \inf f(x)$ , то  $\bar{x} = (1; 1)^T$  — точка глобального минимума.

Учебное издание

**Мицель** Артур Александрович  
**Шелестов** Александр Андреевич  
**Романенко** Владимир Васильевич

**МЕТОДЫ ОПТИМИЗАЦИИ**

Учебное пособие

В двух частях

**Часть 2**

Формат 60×84/16. Усл. печ. л. 20,93.

Подписано в печать 10.11.2022. Тираж 100 экз. Заказ № 220.

Федеральное государственное бюджетное образовательное учреждение  
высшего образования «Томский государственный университет  
систем управления и радиоэлектроники»

634050, г. Томск, пр. Ленина, 40.  
Тел. (3822) 533018.