



МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ  
УНИВЕРСИТЕТ имени М. В. ЛОМОНОСОВА

# МЕТОДЫ ОПТИМИЗАЦИИ

УЧЕБНИК И ПРАКТИКУМ  
ДЛЯ БАКАЛАВРИАТА И МАГИСТРАТУРЫ

Под редакцией профессора,  
доктора физико–математических наук **Ф. П. Васильева**

*Рекомендовано Учебно–методическим отделом  
высшего образования в качестве учебника для студентов  
высших учебных заведений, обучающихся по естественнонаучным  
направлениям и специальностям*

Книга доступна в электронной библиотечной системе  
[biblio-online.ru](http://biblio-online.ru)

Москва ■ Юрайт ■ 2016

УДК 51-3(075.8)

ББК 22.1я73

М54

**Ответственный редактор:**

**Васильев Федор Павлович** — профессор, доктор физико-математических наук, профессор кафедры оптимального управления факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова.

**Авторы:**

**Потапов Михаил Михайлович** — профессор, доктор физико-математических наук, профессор кафедры оптимального управления факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова;

**Будак Борис Александрович** — кандидат физико-математических наук, преподаватель кафедры оптимального управления факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова;

**Артемова Людмила Анатольевна** — кандидат физико-математических наук, преподаватель кафедры оптимального управления факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова.

**Рецензенты:**

*Арутюнов А. В.* — доктор физико-математических наук, профессор, заведующий кафедрой нелинейного анализа и оптимизации факультета физико-математических и естественных наук Российского университета дружбы народов;

*Куржанский А. Б.* — академик Российской академии наук, доктор физико-математических наук, заслуженный профессор Московского государственного университета имени М. В. Ломоносова.

**М54** **Методы оптимизации** : учебник и практикум для бакалавриата и магистратуры / под ред. Ф. П. Васильева. — М. : Издательство Юрайт, 2016. — 375 с. — Серия : Бакалавр и магистр. Академический курс.

ISBN 978-5-9916-6157-7

В учебнике излагаются наиболее известные и часто используемые на практике методы поиска экстремума (максимума, минимума) функций конечного числа переменных, объясняются идейные, эвристические соображения, лежащие в основе каждого метода, исследуется их сходимости, обсуждаются вычислительные аспекты, области применимости методов, их достоинства и недостатки. Рассмотрены также методы решения задач оптимального управления процессами, описываемыми системами обыкновенных дифференциальных уравнений. Основной текст книги соответствует типовой программе для бакалавров. Выделенный звездочками дополнительный текст рассчитан на обучающихся в магистратуре и читателей, заинтересованных в углубленном изучении предмета. Учебник содержит достаточно большое количество разобранных примеров и упражнений для самостоятельной работы студентов.

Соответствует актуальным требованиям Федерального государственного образовательного стандарта высшего образования.

*Для бакалавров и магистров технических вузов.*

УДК 51-3(075.8)

ББК 22.1я73



*Все права защищены. Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав. Правовую поддержку издательства обеспечивает юридическая компания «Дельфи».*

ISBN 978-5-9916-6157-7

© Коллектив авторов, 2015

© ООО «Издательство Юрайт», 2016

# Оглавление

<b>Обозначения</b> .....	<b>7</b>
<b>Предисловие</b> .....	<b>11</b>
<b>Глава 1. Введение в оптимизацию. Классический метод</b> .....	<b>15</b>
1.1. Постановка задачи .....	15
1.1.1. Задача минимизации .....	15
1.1.2. Задача максимизации .....	20
1.2. Теоремы Вейерштрасса .....	22
1.2.1. Классическая теорема .....	22
1.2.2. Обобщенная теорема Вейерштрасса .....	25
1.2.3. Теорема Вейерштрасса для некомпактных множеств .....	30
1.3. Задачи безусловной оптимизации. Классический метод .....	32
1.3.1. Описание метода .....	32
1.3.2. Примеры .....	38
1.4. Задачи на условный экстремум. Правило множителей Лагранжа .....	40
1.4.1. Постановка задачи. Необходимое условие оптимальности на выпуклом множестве .....	40
1.4.2. Правило множителей Лагранжа в задачах с ограничениями типа равенств .....	43
1.4.3. Правило множителей Лагранжа в общем случае .....	45
1.4.4. Достаточные условия оптимальности .....	47
1.4.5. Примеры .....	50
1.4.6. Экономический смысл множителей Лагранжа .....	60
1.5. Метод штрафных функций .....	61
1.5.1. Описание метода .....	61
1.5.2. Сходимость метода .....	64
1.5.3. Недостатки и достоинства метода .....	69
1.6. Обоснование правила множителей Лагранжа .....	70
1.7. Вспомогательные предложения .....	73
1.7.1. Формулы конечных приращений .....	73
1.7.2. Леммы о числовых последовательностях .....	75
<i>Упражнения</i> .....	76
<b>Глава 2. Линейное программирование</b> .....	<b>81</b>
2.1. Постановка задачи .....	81
2.1.1. Общая задача линейного программирования .....	81
2.1.2. Примеры .....	83
2.1.3. Каноническая задача .....	85
2.1.4. Основная задача .....	86

2.2. Геометрическая интерпретация задачи линейного программирования .....	87
2.3. Угловые точки в канонической задаче.....	91
2.4. Симплекс-метод .....	94
2.4.1. Приведенная задача .....	94
2.4.2. Описание симплекс-метода.....	97
2.4.3. Примеры .....	107
2.4.4. Конечность метода в невырожденной задаче.....	109
2.5.* Антициклон .....	112
2.5.1. Пример.....	112
2.5.2. Лексикографическое упорядочение векторов.....	114
2.5.3. Лексикографический антициклон.....	116
2.5.4. Симплекс-метод в канонической задаче максимизации.....	121
2.6. Поиск начальной угловой точки .....	122
2.7. Теорема Вейерштрасса в линейном программировании .....	126
2.8. Теоремы двойственности.....	127
<i>Упражнения</i> .....	137
<b>Глава 3. Элементы выпуклого анализа.....</b>	<b>140</b>
3.1. Выпуклые функции .....	140
3.1.1. Понятие выпуклой функции .....	140
3.1.2. Критерии выпуклости.....	143
3.1.3. Свойства точек минимума выпуклых функций.....	146
3.1.4. Другие свойства выпуклых функций .....	148
3.2. Сильно выпуклые функции .....	151
3.2.1. Понятие сильно выпуклой функции .....	151
3.2.2. Критерии сильной выпуклости .....	153
3.3. Метрическая проекция .....	157
3.4. Теорема Куна – Таккера.....	164
3.5. Теорема Куна – Таккера для многогранного множества .....	171
3.6.* Двойственные задачи. Примеры .....	174
<i>Упражнения</i> .....	183
<b>Глава 4. Методы минимизации функций.....</b>	<b>186</b>
4.1. Методы минимизации функций одной переменной.....	186
4.1.1. Метод деления отрезка пополам .....	187
4.1.2. Метод золотого сечения .....	188
4.1.3. Метод ломаных.....	190
4.1.4. Метод покрытия.....	194
4.1.5. Метод касательных .....	198
4.2. Градиентный метод .....	198
4.2.1. Описание процесса. Выбор шага, проблемы старта и остановки.....	198
4.2.2. Сходимость метода скорейшего спуска для сильно выпуклых функций .....	202
4.2.3. Траектория процесса. Непрерывный аналог градиентного метода.....	205
4.3. Метод проекции градиента.....	208
4.3.1. Описание процесса. Варианты выбора шага.....	208
4.3.2. Сходимость метода для сильно выпуклых функций .....	209

4.4. Метод условного градиента.....	211
4.4.1. Идея линейной аппроксимации и варианты ее реализации.....	211
4.4.2. Сходимость метода.....	214
4.5. Метод Ньютона .....	216
4.5.1. Описание процесса. Варианты выбора шага.....	216
4.5.2. Сходимость классического метода Ньютона для сильно выпуклых функций.....	217
4.5.3. Сходимость метода Ньютона с переменным шагом для сильно выпуклых функций.....	220
4.6.* Метод переменной метрики .....	223
4.6.1. Описание процесса. $G$ -проекция.....	223
4.6.2. Сходимость метода переменной метрики для выпуклых функций .....	227
4.6.3. Непрерывный вариант метода переменной метрики.....	230
4.7.* Квазиньютоновские методы .....	233
4.8. Метод покоординатного спуска.....	245
4.8.1. Основная версия метода для задач без ограничений.....	245
4.8.2. Другие варианты метода покоординатного спуска.....	247
4.9. Метод покрытия для функции многих переменных .....	249
4.10.* Метод модифицированных функций Лагранжа .....	252
4.11.* Экстраградиентный метод .....	259
4.11.1. Описание процесса .....	259
4.11.2. Сходимость экстраградиентного метода.....	260
<i>Упражнения</i> .....	264

## **Глава 5. Принцип максимума Понтрягина..... 266**

5.1. Постановка задачи оптимального управления.....	266
5.1.1. Примеры постановок задач оптимального управления.....	266
5.1.2. Допустимые управления.....	269
5.1.3. Задача Коши.....	270
5.1.4. Постановка задачи.....	274
5.2. Принцип максимума в задаче оптимального управления со свободным правым концом.....	277
5.2.1. Формулировка принципа максимума .....	277
5.2.2. Краевая задача принципа максимума.....	278
5.2.3. Доказательство принципа максимума Понтрягина.....	283
5.3.* Принцип максимума (общий случай) .....	287
5.3.1. Формулировка принципа максимума в общем случае.....	287
5.3.2. Краевая задача принципа максимума в общем случае .....	289
5.3.3. Общие граничные условия.....	291
5.4.* Примеры .....	293
5.5.* Принцип максимума и классическое вариационное исчисление .....	305
<i>Упражнения</i> .....	309

## **Глава 6. Динамическое программирование ..... 313**

6.1. Принцип оптимальности .....	313
6.2. Схема Беллмана .....	315
6.2.1. Разностная аппроксимация задачи оптимального управления .....	315
6.2.2. Уравнение Беллмана.....	316

6.2.3. Вычислительная схема .....	319
6.2.4. Синтезирующая функция.....	321
6.2.5. Оценка погрешности .....	322
6.2.6. Построение минимизирующих последовательностей.....	324
6.2.7. Приложение к задачам конечномерной оптимизации .....	325
6.2.8. Замечания к схеме Беллмана.....	326
6.3.* Проблема синтеза для систем с непрерывным временем .....	327
6.3.1. Задача Коши — Беллмана .....	328
6.3.2. Примеры .....	331
6.3.3. Приближенный метод .....	334
6.3.4. Оценка погрешности .....	337
6.4.* Контрпримеры .....	339
6.5.* Применение метода динамического программирования к задачам дискретной оптимизации .....	342
<i>Упражнения</i> .....	346
<b>Практикум</b> .....	<b>349</b>
<b>Литература</b> .....	<b>363</b>
<b>Ответы к упражнениям</b> .....	<b>366</b>
<b>Предметный указатель</b> .....	<b>372</b>

## Обозначения

$\mathbb{R}$  — числовая ось

$[a; b] = \{x \in \mathbb{R} : a \leq x \leq b\}$  — отрезок

$(a; b) = \{x \in \mathbb{R} : a < x < b\}$  — интервал

$(a; b] = \{x \in \mathbb{R} : a < x \leq b\}$ ,  $[a; b) = \{x \in \mathbb{R} : a \leq x < b\}$  — полуинтервалы

$\mathbb{R}^n$  —  $n$ -мерное линейное пространство векторов-столбцов  $x$  с координатами  $x^i$ ,  $i = 1, \dots, n$

$x^\top = (x^1, \dots, x^n)$  — вектор-строка, полученная транспонированием вектора-столбца  $x$

$$\|x\|_p = \left( \sum_{i=1}^n |x^i|^p \right)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x^i| \quad \text{— нормы вектора } x \text{ в } \mathbb{R}^n$$

$E^n$  —  $n$ -мерное евклидово пространство, представляющее собой пространство  $\mathbb{R}^n$ , оснащенное скалярным произведением  $\langle x, y \rangle = \sum_{i=1}^n x^i y^i$

$$\|x\| = \|x\|_2 = \|x\|_{E^n} = \left( \sum_{i=1}^n |x^i|^2 \right)^{\frac{1}{2}} \quad \text{— евклидова норма вектора}$$

$\{e_1, \dots, e_n\}$  — ортонормированный базис пространства  $E^n$ , где  $e_i = (0, \dots, 0, 1, 0, \dots, 0)^\top$  — вектор-столбец,  $i$ -я координата которого равна единице, а остальные равны нулю,  $i = 1, \dots, n$

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \dots & \dots & \dots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} = \{a_{i,j}, i = 1, \dots, m; j = 1, \dots, n\} = \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix} = (A_1, \dots, A_m) \quad \text{— матрица}$$

размера  $m \times n$  с элементами  $a_{ij}$ , где  $a_i = (a_{i,1}, \dots, a_{i,n})$  —  $i$ -я строка матрицы  $A$ ,  $A_j = (a_{1,j}, \dots, a_{m,j})^\top$  —  $j$ -й столбец матрицы  $A$

$$A^\top = \begin{pmatrix} a_{1,1} & \dots & a_{m,1} \\ \dots & \dots & \dots \\ a_{1,n} & \dots & a_{m,n} \end{pmatrix} = (a_1^\top, \dots, a_n^\top) = \begin{pmatrix} A_1^\top \\ \dots \\ A_m^\top \end{pmatrix} \quad \text{— матрица размера } n \times m, \text{ полу-}$$

ченная транспонированием матрицы  $A$  размера  $m \times n$

$$I_n = (e_1, \dots, e_n) = \begin{pmatrix} e_1^\top \\ \dots \\ e_n^\top \end{pmatrix} \quad \text{— единичная матрица } n\text{-го порядка со столбцами}$$

$e_1, \dots, e_n$

$\det A$  — определитель квадратной матрицы  $A$

$A^{-1}$  — обратная матрица для квадратной матрицы  $A$  с  $\det A \neq 0$

$\text{rank } A$  — ранг матрицы  $A$

$Ax = ((Ax)^1, \dots, (Ax)^m)^\top$ , где  $(Ax)^i = \sum_{j=1}^n a_{ij}x^j$ ,  $i = 1, \dots, m$ , — произведение

матрицы размера  $m \times n$  на вектор  $x \in E^n$

$\|A\| = \max_{\|x\|_{E^n} \leq 1} \|Ax\|_{E^m}$  — норма (операторная норма) матрицы  $A$  размера  $m \times n$

$AB$  — произведение матрицы  $A = \{a_{ij}\}$  размера  $m \times n$  на матрицу  $B = \{b_{ij}\}$  размера  $n \times q$ , его результатом является матрица  $C = \{c_{ij}\}$  размера  $m \times q$

с элементами  $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$ ,  $i = 1, \dots, m, j = 1, \dots, q$

$A \geq 0$  — неотрицательно определенная симметричная матрица, т.е.  $\langle Ax, x \rangle \geq 0$  для всех  $x \in E^n$

$A > 0$  — положительно определенная симметричная матрица, т.е.  $\langle Ax, x \rangle > 0$  для всех  $x \in E^n, x \neq 0$

$A \leq 0$  — неположительно определенная матрица, т.е.  $(-A) \geq 0$

$A < 0$  — отрицательно определенная матрица, т.е.  $(-A) > 0$

$x = (x^1, \dots, x^n)^\top \geq 0$  — неотрицательный вектор, т.е.  $x^i \geq 0$  для всех  $i = 1, \dots, n$

$x = (x^1, \dots, x^n)^\top > 0$  — положительный вектор, т.е.  $x^i > 0$  для всех  $i = 1, \dots, n$

$x = (x^1, \dots, x^n)^\top \geq (y^1, \dots, y^n)^\top$  — вектор  $x$  не меньше вектора  $y$ , т.е.  $x - y \geq 0$

$x = (x^1, \dots, x^n)^\top > (y^1, \dots, y^n)^\top$  — вектор  $x$  больше вектора  $y$ , т.е.  $x - y > 0$

$E_+^n = \{x \in E^n: x \geq 0\}$  — неотрицательный ортант пространства  $E^n$

$x_+ = (x_+^1, \dots, x_+^n)^\top$ , где  $x_+^i = \max\{0; x^i\}$ ,  $i = 1, \dots, n$ , — проекция вектора  $x$  на неотрицательный ортант пространства  $E^n$

$x = (x^1, \dots, x^n)^\top \succ 0$  — лексикографически положительный вектор, т.е.  $x \neq 0$  и первая ненулевая координата этого вектора положительна

$x = (x^1, \dots, x^n)^\top \succ (y^1, \dots, y^n)^\top$  — вектор  $x$  лексикографически больше вектора  $y$ , т.е.  $x - y \succ 0$

$x_* = \text{lex min}_{i \in M} x_i$  — лексикографический минимум множества векторов  $\{x_i, i \in M\}$ , т.е. для каждого номера  $i \in M$  либо  $x_i \succ x_*$ , либо  $x_i = x_*$

$S = S(v, B) = \begin{pmatrix} \Gamma \\ \Delta \end{pmatrix}$  — симплекс-таблица угловой точки  $v$  с базисом  $B$ , где

$\Gamma = (\Gamma_1, \dots, \Gamma_r)^\top$ ,  $\Gamma_i = (\gamma_{i0}, \gamma_{i1}, \dots, \gamma_{in})$ ,  $i = 1, \dots, r$ ,  $\Delta = (\Delta_0, \Delta_1, \dots, \Delta_n)$

$S = \begin{pmatrix} \Gamma \\ \Delta \end{pmatrix} \succ 0$  — лексикографически положительная симплекс-таблица, т.е.

$\Gamma_i \succ 0$  для всех  $i = 1, \dots, r$

$S_1 = \begin{pmatrix} \Gamma_1 \\ \Delta_1 \end{pmatrix} \succ S_2 = \begin{pmatrix} \Gamma_2 \\ \Delta_2 \end{pmatrix}$  — симплекс-таблица  $S_1$  лексикографически больше

симплекс-таблицы  $S_2$ , т.е.  $\Delta_1 \succ \Delta_2$

$\bar{X}$  — замыкание множества  $X$

$\text{Gr}X$  — граничные точки множества  $X$

$\text{int}X$  — внутренние точки множества  $X$

$\dim X$  — размерность множества  $X$

$X \cup Y$  — объединение двух множеств  $X$  и  $Y$

$X \cap Y$  — пересечение двух множеств  $X$  и  $Y$

$X + Y$  — сумма двух множеств  $X$  и  $Y$



$X \setminus Y$  — разность двух множеств  $X$  и  $Y$

$\alpha X$  — произведение числа  $\alpha$  на множество  $X$

$X \times Y$  — прямое (декартово) произведение двух множеств  $X$  и  $Y$

$\emptyset$  — пустое множество

$\rho(x, X) = \inf_{y \in X} \|x - y\|$  — расстояние от точки  $x$  до множества  $X$

$O(v, \varepsilon) = \{x \in E^n: \|x - v\| < \varepsilon\}$  —  $\varepsilon$ -окрестность точки  $v$ , представляющая собой открытый шар радиуса  $\varepsilon > 0$  с центром в точке  $v$

$S(v, R) = \{x \in E^n: \|x - v\| \leq R\}$  — замкнутый шар радиуса  $R$  с центром в точке  $v$

$L^\perp$  — ортогональное дополнение подпространства  $L \subseteq E^n$

$P_X(z)$  — проекция точки  $z \in E^n$  на множество  $X$

$P_X^G(z)$  —  $G$ -проекция точки  $z \in E^n$  на множество  $X$

$f(x) \rightarrow \min, x \in X$ , — краткая символическая запись задачи минимизации функции  $f(x)$  на множестве  $X$

$f_* = \inf_{x \in X} f(x)$  — точная нижняя грань функции  $f(x)$  на множестве  $X$

$X_* = \{x \in X: f(x) = f_* > -\infty\}$  — множество точек минимума функции  $f(x)$  на  $X$

$f(x) \rightarrow \max, x \in X$  — краткая символическая запись задачи максимизации функции  $f(x)$  на множестве  $X$

$f^* = \sup_{x \in X} f(x)$  — точная верхняя грань функции  $f(x)$  на множестве  $X$

$X^* = \{x \in X: f(x) = f^* < +\infty\}$  — множество точек максимума функции  $f(x)$  на  $X$

$\frac{\partial f(x)}{\partial x^i} = f'_{x^i}(x) = f_{x^i}(x)$  — частная производная функции  $f(x) = f(x^1, \dots, x^n)$

в точке  $x$  по переменной  $x^i, i = 1, \dots, n$

$f'(x) = \left( \frac{\partial f(x)}{\partial x^1}, \dots, \frac{\partial f(x)}{\partial x^n} \right) = \text{grad} f(x)$  — градиент функции  $f(x)$  в точке  $x$

$\frac{\partial^2 f(x)}{\partial x^i \partial x^j} = f''_{x^i x^j}(x) = f_{x^i x^j}(x)$  — частная производная второго порядка

по переменным  $x^i, x^j$  функции  $f(x) = f(x^1, \dots, x^n)$  в точке  $x, i, j = 1, \dots, n$

$f''(x)$  — вторая производная функции  $f(x)$  в точке  $x$ , она является квадратной матрицей  $n$ -го порядка с элементами  $\frac{\partial^2 f(x)}{\partial x^i \partial x^j}, i, j = 1, \dots, n$

$\frac{df(x)}{de} = \lim_{t \rightarrow 0} \frac{f(x + te) - f(x)}{t}$  — производная функции  $f(x)$  в точке  $x$

по направлению  $e, \|e\| = 1$

$\dot{x}(t) = \frac{dx(t)}{dt}$  — первая производная функции  $x(t)$  по времени  $t$

$\ddot{x}(t) = \frac{d^2 x(t)}{dt^2}$  — вторая производная функции  $x(t)$  по времени  $t$

$\lim_{k \rightarrow \infty} x_k$  — предел последовательности  $\{x_k\} = (x_1, x_2, \dots, x_k, \dots)$

$\overline{\lim}_{k \rightarrow \infty} x_k$  — верхний предел последовательности  $\{x_k\}$

$\liminf_{k \rightarrow \infty} x_k$  — нижний предел последовательности  $\{x_k\}$

$\lim_{x \rightarrow a} f(x)$  — предел функции  $f(x)$  при  $x \rightarrow a$

$\overline{\lim}_{x \rightarrow a} f(x)$  — верхний предел функции  $f(x)$  при  $x \rightarrow a$

$\liminf_{x \rightarrow a} f(x)$  — нижний предел функции  $f(x)$  при  $x \rightarrow a$

$f(a+0) = \lim_{x \rightarrow a+0} f(x)$  — предел функции  $f(x)$  одной переменной  $x$

при стремлении  $x$  к точке  $a$  справа

$f(a-0) = \lim_{x \rightarrow a-0} f(x)$  — предел функции  $f(x)$  одной переменной  $x$

при стремлении  $x$  к точке  $a$  слева

$O(t)$  — величина, определенная в окрестности точки  $t = 0$  и такая, что

$\left| \frac{O(t)}{t} \right| \leq C$  при всех достаточно малых  $t$ , где  $C$  — некоторая неотрицательная

постоянная

$o(t)$  — величина, определенная в окрестности точки  $t = 0$  и такая, что

$\lim_{t \rightarrow 0} \frac{o(t)}{t} = 0$

$C(X)$  — пространство непрерывных на замкнутом ограниченном множестве  $X$  функций  $f(x)$  с нормой  $\|f\|_C = \max_{x \in X} |f(x)|$

$C^1(X)$  — пространство непрерывно дифференцируемых на множестве  $X$  функций

$C^2(X)$  — пространство дважды непрерывно дифференцируемых на множестве  $X$  функций

$C^{1,1}(X)$  — пространство непрерывно дифференцируемых на множестве  $X$  функций, градиент  $f'(x)$  которых удовлетворяет условию Липшица на  $X$

$Q(L)$  — множество функций, удовлетворяющих на множестве  $X$  условию Липшица с константой  $L$

$H = H(x, u, t, \psi, a_0)$  — функция Гамильтона — Понтрягина

$l = l(x, y, t, T, a)$  — малый Лагранжиан

$\mathcal{L} = \mathcal{L}(x, \bar{\lambda})$  — функция Лагранжа в общей форме

$L = L(x, \lambda)$  — функция Лагранжа в нормальной форме

$M = M(x, \lambda)$  — модифицированная функция Лагранжа

$\forall$  — квантор общности

$\exists$  — квантор существования

$\in$  — знак принадлежности

$\subset$  — знак строгого включения одного множества в другое

$\subseteq$  — знак нестрогого включения одного множества в другое

def

= — знак определения, используется при введении объектов

## Предисловие

Первые задачи геометрического содержания, связанные с отысканием наименьших и наибольших величин, появились еще в древние времена. Развитие промышленности в XVII—XVIII вв. привело к необходимости исследования более сложных задач на экстремум и к появлению вариационного исчисления. Однако лишь в XX в. при огромном размахе производства и осознании ограниченности ресурсов Земли во весь рост встала задача оптимального использования энергии, материалов, рабочего времени, большую актуальность приобрели вопросы наилучшего в том или ином смысле управления различными процессами физики, техники, экономики и др.

Сюда относятся, например, задача организации производства с целью получения максимальной прибыли при заданных затратах ресурсов, задача оптимальной организации перевозок при ограниченном количестве транспортных средств, задача минимизации издержек (отходов) производства, задача оптимального размещения объектов производства, оптимального использования посевной площади, оптимального управления запасами, оптимальной организации работы пунктов массового обслуживания, выбора оптимальных сроков сезонного производства, задача управления системой гидростанций и водохранилищ с целью получения максимального количества электроэнергии, задача о космическом перелете из одной точки пространства в другую наиболее быстрым образом или с наименьшими затратами энергии, задача о быстрейшем нагреве или остывании металла до заданного температурного режима, задача о наилучшем гашении вибраций и многие другие задачи.

На математическом языке такие задачи, называемые *задачами оптимизации*, могут быть сформулированы как задачи отыскания экстремума (минимума или максимума) некоторой функции (функционала)  $f(x)$  на заданном множестве  $X$  из некоторого пространства. Функцию  $f(x)$  принято называть *целевой функцией*, а ее аргумент  $x$  часто называют *параметром управления* или *допустимой точкой*, множество  $X$  — *допустимым множеством*. Значение целевой функции  $f(x)$  выражает собой качество (цену) допустимой точки  $x$ . Требование принадлежности точки  $x$  множеству  $X$  выражает собой наличие некоторых ограничений, вытекающих, например, из законов сохранения энергии, ограниченности имеющихся ресурсов, возможностей технической реализации управления, нежелательности каких-либо запрещенных (аварийных) состояний и т.п.

Сразу же заметим, что на практике формулировка задач оптимизации, составление целевой функции  $f(x)$  и множества  $X$  являются сложной проблемой, требующей больших знаний в предметной области, материальных

заграт по сбору и анализу информации, постановке необходимых экспериментов и т.п. Обсуждение возникающих здесь вопросов является предметом особой области знаний, называемой *математическим моделированием*, и выводит нас за рамки настоящего учебного пособия. Чтобы читатель имел некоторое представление об этой сложной работе по формулировке задач оптимизации, в гл. 2 будут приведены примеры линейных задач оптимизации.

В общем случае, говоря о задаче оптимизации, мы будем подразумевать, что известны целевая функция  $f(x)$ , допустимое множество  $X$  и цель, т.е. ищется ли минимум или максимум функции. Далее возникает проблема: как искать экстремум функции  $f(x)$  на множестве  $X$ ? Если множество  $X$  содержит небольшое число точек, то задачу оптимизации можно решить относительно просто: достаточно вычислить значения целевой функции в этих точках и, сравнивая их по величине, простым перебором найти экстремальное (минимальное или максимальное) значение и указать точку, где этот экстремум достигается. Однако если число точек в  $X$  велико и даже, возможно, бесконечно, то такой метод простого перебора не под силу современным вычислительным средствам. Как искать экстремум в такой ситуации? Ясно, что нужны специальные методы оптимизации, нужна теория таких методов, нужна наука.

Читатель отчасти уже знаком с классическим методом оптимизации, основанном на дифференциальном исчислении, который излагается в основных курсах по математическому анализу. В настоящее время теория оптимизации обогатилась фундаментальными результатами, полученными по таким направлениям, как линейное программирование, выпуклое программирование, динамическое программирование, стохастическое программирование, дискретная оптимизация, теория оптимального управления процессами, описываемыми дифференциальными уравнениями, и др. Потребности практики способствовали бурному развитию методов решения задач оптимизации. Появление современных вычислительных средств, компьютеров сделало возможным эффективное решение многих важных прикладных задач оптимизации, которые ранее из-за своей сложности казались недоступными, появились пакеты программ оптимизации.

В настоящем учебнике излагаются элементы теории оптимизации, а также основы наиболее часто используемых на практике методов приближенного решения задач оптимизации и краткая характеристика этих методов. Усилия авторов были направлены на поиск наиболее экономных схем изложения, упрощение доказательств, чтобы сделать материал более доступным, не снижая уровня строгости.

Руководствуясь этими методическими соображениями, авторы вынесли в первую главу метод штрафных функций, для изложения и понимания которого, по сути, достаточно знания свойств непрерывных функций многих переменных из классического математического анализа. Это дало нам возможность в первой же главе с помощью аппарата штрафных функций дать простое доказательство важнейшего в теории оптимизации правила множителей Лагранжа. Во второй главе дается элементарное изложение основ линейного программирования с привлечением минимального

математического аппарата, опираясь лишь на самые первичные понятия из линейной алгебры, такие как матрица, определитель, линейная зависимость и независимость векторов, ранг матрицы, произведение матриц, обратная матрица, не предполагая знания довольно сложной теории выпуклых многогранных множеств. И в последующих главах авторы старались уделить методическим аспектам большое внимание, стремясь полнее и проще излагать основные факты теории и методов оптимизации.

К сожалению, ограниченный объем книги не позволяет уделить внимание таким важным разделам теории и методов оптимизации, как дискретная оптимизация, стохастическая оптимизация, вопросам устойчивости вычислительных алгоритмов, методам решения неустойчивых задач, задач большой размерности и др.

Материал книги для своего понимания требует лишь знания основ математического анализа, линейной алгебры, элементов теории обыкновенных дифференциальных уравнений в объеме стандартных курсов бакалавриата и магистратуры технических вузов. В книге выделены основной и дополнительный тексты. Основной текст описывает материал действующей типовой программы «бакалавр-магистр». Дополнительный текст, служащий для более углубленного усвоения материала и предназначенный магистрам, в оглавлении и в заголовках внутри книги помечен символом \*.

О содержании книги можно судить по оглавлению. В книге значительное место занимает иллюстративный материал, примеры задач с разбором решений. В конце каждой главы приведены вопросы и упражнения теоретического характера, а в конце учебника отдельным блоком представлен практикум — задачи для самостоятельной работы студентов.

При составлении заданий практикума к гл. 1, 2 мы пользовались упражнениями из задачника [5], к гл. 5, 6 идеи заданий заимствованы из книги [15]. При работе над книгой авторы опирались на свой опыт работы со студентами на факультете ВМК МГУ имени М. В. Ломоносова. В книге широко использованы материалы монографии Ф. П. Васильева [12], касающиеся конечномерных задач и методов их решения, а также задач оптимального управления процессами, описываемыми обыкновенными дифференциальными уравнениями. Эту книгу можно считать продолжением настоящего учебника.

По рассматриваемой в учебнике проблематике имеется обширная литература. Список литературы в конце учебника отражает лишь некоторые источники, которые были непосредственно использованы при написании книги или близко примыкают к ней, расширяя и дополняя ее содержание.

Авторы выражают глубокую благодарность А. В. Тимохову за материалы, которые были существенно использованы при составлении заданий по практикуму к гл. 1–3, а также В. Ю. Королеву за полезные советы по содержанию книги.

В результате освоения дисциплины у студента должны формироваться следующие профессиональные компетенции:

**знать**

- классификацию задач оптимизации (задачи безусловной и условной оптимизации; классическую задачу на условный экстремум; задачи матема-

тического, выпуклого, линейного программирования; задачи оптимального управления);

- основные положения теории оптимизации (теория необходимых и достаточных условий оптимальности для конечномерных задач, правило множителей Лагранжа, теория двойственности задач линейного программирования, принцип максимума Понтрягина в оптимальном управлении и др.);

- элементы выпуклого анализа как основы современной теории оптимизации;

- классификацию численных методов оптимизации (конечношаговые и бесконечношаговые методы, методы условной и безусловной оптимизации, методы нулевого, первого и второго порядков);

- понятийный аппарат численных методов оптимизации (итерация метода, начальная точка, направление убывания, критерий останова, сходимость метода);

### ***уметь***

- аналитически решать простейшие нелинейные задачи оптимизации, используя теорию необходимых и достаточных условий оптимальности;

- аналитически решать простейшие задачи линейного программирования, используя симплекс-метод;

- строить прикладные модели, сводящиеся к задачам оптимизации;

- анализировать характеристики прикладной задачи оптимизации (целевую функцию, допустимое множество) на предмет ее отнесения к конкретному классу задач;

- выбирать адекватный численный метод решения прикладной задачи оптимизации (с последующим обращением к соответствующим библиотекам компьютерных программ);

### ***владеть***

- навыками решения задач оптимизации;

- материалом дисциплины на уровне, позволяющем формулировать, исследовать и решать прикладные задачи оптимизации.

Авторы будут признательны читателям за советы и критические замечания по содержанию книги.

# Глава 1

## ВВЕДЕНИЕ В ОПТИМИЗАЦИЮ. КЛАССИЧЕСКИЙ МЕТОД

### 1.1. Постановка задачи

#### 1.1.1. Задача минимизации

Перед тем как начать изложение теории оптимизации, введем некоторые обозначения и напомним определения из линейной алгебры и математического анализа [9, 22, 29, 31]. Через  $\mathbb{R}^n$  будем обозначать  $n$ -мерное вещественное линейное пространство, состоящее из вектор-столбцов

$$x = \begin{pmatrix} x^1 \\ \dots \\ x^n \end{pmatrix}, \quad y = \begin{pmatrix} y^1 \\ \dots \\ y^n \end{pmatrix}, \quad z = \begin{pmatrix} z^1 \\ \dots \\ z^n \end{pmatrix}, \dots$$

с действительными координатами  $x^i, y^i, z^i, i = 1, \dots, n$ ; сумма  $x + y$  двух вектор-столбцов и произведение  $\alpha x$  вектор-столбца  $x$  на действительное число  $\alpha$  в  $\mathbb{R}^n$  определяются обычным образом:

$$x + y = \begin{pmatrix} x^1 + y^1 \\ \dots \\ x^n + y^n \end{pmatrix}, \quad \alpha x = \begin{pmatrix} \alpha x^1 \\ \dots \\ \alpha x^n \end{pmatrix}.$$

Вектор-столбец

$$\mathbf{0} = \begin{pmatrix} 0 \\ \dots \\ 0 \end{pmatrix}$$

называется *нулевым*. Вектор-строку, полученную транспонированием вектор-столбца  $x$ , обозначим через  $x^\top = (x^1, \dots, x^n)$ . Там, где не могут возникнуть недоразумения, вектор-столбец из  $\mathbb{R}^n$  или соответствующую ему вектор-строку для краткости мы часто будем называть просто вектором или точкой, а знак транспонирования « $\top$ » будем опускать.

Если в  $\mathbb{R}^n$  ввести скалярное произведение двух векторов  $\langle x, y \rangle = \sum_{i=1}^n x^i y^i$ ,  $x, y \in \mathbb{R}^n$ , то  $\mathbb{R}^n$  превращается в  $n$ -мерное евклидово пространство, которое будем обозначать через  $E^n$ . Если  $\langle x, y \rangle = 0$ , то векторы  $x$  и  $y$  называются *ортогональными*. Это скалярное произведение порождает *норму* (длину) вектора в  $E^n$ , которую вычисляют по формуле

$$\|x\| = \langle x, x \rangle^{1/2} = \left( \sum_{i=1}^n |x^i|^2 \right)^{1/2}.$$

Величину

$$\rho(x, y) = \|x - y\| = \left( \sum_{i=1}^n |x^i - y^i|^2 \right)^{1/2}$$

называют *евклидовым расстоянием* между точками  $x, y \in E^n$ .

Для любых трех точек  $x, y, z \in E^n$  справедливо неравенство

$$\|x - y\| \leq \|x - z\| + \|z - y\|,$$

называемое *неравенством треугольника*. Когда важно подчеркнуть, что скалярное произведение, норма, расстояние взяты именно в  $E^n$ , мы будем писать  $\langle x, y \rangle_{E^n}$ ,  $\|x\|_{E^n}$ ,  $\|x - y\|_{E^n}$ .

В  $E^n$  справедливо *неравенство Коши — Буняковского*

$$\langle x, y \rangle \leq \|x\| \cdot \|y\| \text{ для } \forall x, y \in E^n,$$

причем неравенство превращается в равенство тогда и только тогда, когда векторы  $x, y$  коллинеарны, т.е.  $x = \alpha y$  при некотором  $\alpha$ .

Иногда мы будем пользоваться и другими нормами векторов из  $\mathbb{R}^n$ , такими как

$$\|x\|_p = \left( \sum_{i=1}^n |x^i|^p \right)^{1/p}, \quad 1 \leq p < \infty, \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x^i|.$$

Как известно, в конечномерных пространствах все нормы *эквивалентны*. Это значит, что если  $\|\cdot\|_I, \|\cdot\|_{II}$  — две нормы в  $\mathbb{R}^n$ , то существуют такие положительные числа  $m_1$  и  $m_2$ , что  $m_1 \|x\|_I \leq \|x\|_{II} \leq m_2 \|x\|_I$  для любого  $x \in \mathbb{R}^n$ . Отсюда следует, что если последовательность  $\{x_k\}$  сходится к точке  $x$  в какой-либо норме  $\|\cdot\|_I$ , т.е.  $\|x_k - x\|_I \rightarrow 0$  при  $k \rightarrow \infty$ , то эта последовательность будет сходиться к  $x$  и в любой другой норме  $\|\cdot\|_{II}$ , в частности  $\|x_k - x\|_\infty = \max_{1 \leq i \leq n} |x_k^i - x^i| \rightarrow 0$ , что равносильно покоординатной сходимости.

В основном мы будем работать в евклидовом пространстве  $E^n$ . Перейдем к постановке задач оптимизации.

Пусть  $X$  — некоторое известное множество из  $n$ -мерного евклидова пространства  $E^n$ , а  $f(x) = f(x^1, \dots, x^n)$  — функция, определенная на этом множестве. Кратко задачу оптимизации будем записывать следующим образом:

$$f(x) \rightarrow \min, \quad x \in X. \quad (1.1)$$

Поясним, что подразумевается под задачей (1.1). Начнем с того, что напомним некоторые определения из классического математического анализа [22, 29].

**Определение 1.1.** Точку  $x_* \in X$  будем называть *точкой минимума* функции  $f(x)$  на множестве  $X$ , если  $f(x_*) \leq f(x)$  для всех  $x \in X$ ; величину  $f(x_*)$  назовем наименьшим значением  $f(x)$  на  $X$  и обозначим  $\min_{x \in X} f(x) = f_* = f(x_*)$ . Множество всех точек минимума функции  $f(x)$  на  $X$  будем обозначать через  $X_*$ .



В зависимости от свойств множества  $X$  и функции  $f(x)$  множество  $X_*$  может содержать одну, несколько и даже бесконечно много точек, а также возможны случаи, когда множество  $X_*$  пусто. Поясним это на примерах.

### Пример 1.1

Пусть

$$f(x) = \frac{2\|x\|}{1 + \|x\|^2}, \quad x = (x^1, \dots, x^n) \in E^n.$$

График этой функции при  $n = 1$  изображен на рис. 1.1.

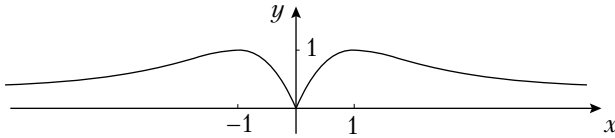


Рис. 1.1

На множестве  $X_1 = E^n$  минимальное значение этой функции равно нулю:  $f_* = f(0) = 0$ , множество  $X_*$  точек минимума состоит из единственной точки  $x_* = 0$ . На множестве  $X_2 = \{x \in E^n : \|x\| \geq 1\}$  функция  $f(x)$  не имеет минимального значения, так как  $f(x) > 0$  при всех  $x \in X_2$ , причем на последовательности  $x_k, k = 1, 2, \dots$ , функция стремится к нулю:  $\lim_{k \rightarrow \infty} f(x_k) = 0$ , но равенство  $f(x) = 0$  невозможно ни в одной точке  $x \in X_2$ , так что  $X_{2*} = \emptyset$ .

### Пример 1.2

Пусть

$$f(x) = \begin{cases} \frac{1}{\|x\|}, & \text{если } x \neq 0, \\ 1, & \text{если } x = 0; \end{cases} \quad x \in E^n.$$

График этой функции при  $n = 1$  изображен на рис. 1.2.

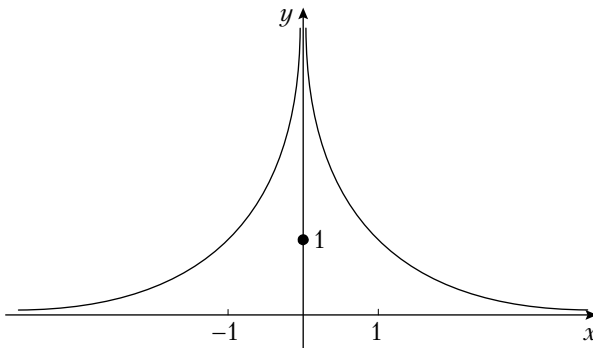


Рис. 1.2

На множествах  $X_1 = E^n$  или  $X_2 = \{x \in E^n : \|x\| \geq 1\}$  функция  $f(x)$  принимает строго положительные значения, множество точек ее минимума пусто.

Если  $X = X_3 = \{x \in E^n : \|x\| \leq 1/2\}$ , то  $f_* = 1$ ,  $X_{3*} = \{x = 0\}$ .

Если  $X = X_4 = \{x \in E^n : \|x\| \leq 2\}$ , то  $f_* = \frac{1}{2}$ ,  $X_{4*} = \{x \in X_4 : \|x\| = 2\}$  —  $n$ -мерная сфера

с центром в точке  $x = 0$  (при  $n = 1$   $X_{4*}$  — это две точки  $x_* = 2$  и  $x_* = -2$ ).

Если  $X = X_5 = \{x \in E^n : \|x\| \leq 1\}$ , то  $f_* = 1$ ,  $X_{5*} = \{x \in X_5 : \|x\| = 1\} \cup \{x = 0\}$ .

### Пример 1.3

Пусть

$$f(x) = \begin{cases} \sin^2\left(\frac{\pi}{x}\right) & \text{при } x > 0, \\ 0 & \text{при } x = 0. \end{cases}$$

График этой функции изображен на рис. 1.3.

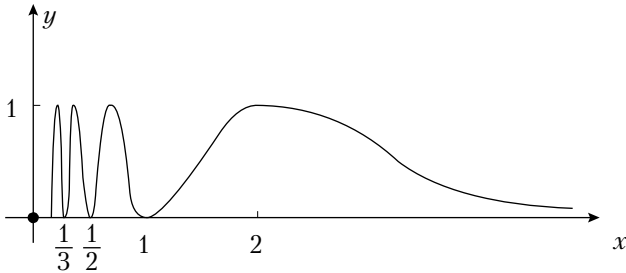


Рис. 1.3

На множестве  $X = X_1 = \{x : 1 \leq x \leq 2\}$   $f_* = 0$ ,  $X_{1*} = \{x_* = 1\}$ .

Если  $X = X_2 = \left\{x : \frac{1}{3} \leq x \leq 1\right\}$ , то  $f_* = 0$ ,  $X_{2*}$  состоит из трех точек  $x_1 = \frac{1}{3}$ ,  $x_2 = \frac{1}{2}$ ,  $x_3 = 1$ .

Если  $X = X_3 = \{x : 0 \leq x \leq 1\}$ , то  $f_* = 0$ ,  $X_{3*} = \left\{x \in X_3 : x = \frac{1}{n}, n \in \mathbb{N}\right\} \cup \{x = 0\}$ .

### Пример 1.4

Пусть  $f(x) = \ln x$ ,  $X = \{x \in E^1 : 0 < x \leq 1\}$ .

Здесь  $X_* = \emptyset$ , для последовательности  $x_k = \frac{1}{k}$ ,  $k = 1, 2, \dots$ , имеем  $\lim_{k \rightarrow \infty} f(x_k) = -\infty$  (рис. 1.4).

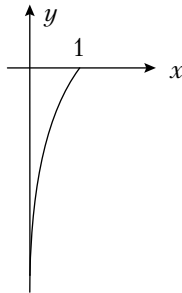


Рис. 1.4

**Определение 1.2.** Функция  $f(x)$  называется *ограниченной снизу* на множестве  $X$ , если существует такое число  $M_0$ , что  $f(x) \geq M_0$  для всех  $x \in X$ . Функция  $f(x)$  не ограничена снизу на  $X$ , если существует последовательность  $\{x_k\} \in X$ , для которой  $\lim_{k \rightarrow \infty} f(x_k) = -\infty$ .

В примерах 1.1–1.3 функции ограничены снизу на рассматриваемых множествах. В примере функция не ограничена снизу, хотя и принимает конечные значения во всех точках из  $X$ .

В тех случаях, когда  $X_* = \emptyset$ , естественным обобщением понятия наименьшего значения функции на множестве является понятие точной нижней грани функции.

**Определение 1.3.** Пусть функция  $f(x)$  ограничена снизу на множестве  $X$ . Тогда число  $f_*$  называют *точной нижней гранью* или *инфимумом*  $f(x)$  на  $X$ , если

- 1)  $f_* \leq f(x)$  при всех  $x \in X$ ;
- 2) для любого сколь угодно малого числа  $\varepsilon > 0$  найдется точка  $x_\varepsilon \in X$  такая, что  $f(x_\varepsilon) < f_* + \varepsilon$ .

Точную нижнюю грань  $f(x)$  на  $X$  будем обозначать через  $\inf_{x \in X} f(x) = f_*$ .

Если функция  $f(x)$  не ограничена снизу на  $X$ , то в качестве точной нижней грани  $f(x)$  на  $X$  принимается  $f_* = -\infty$ .

В примерах 1.1 и 1.2  $f_* = 0$ , а множество  $X_*$  может быть пустым или непустым в зависимости от множества  $X$ ; в примере 1.4  $f_* = -\infty$ ,  $X_* = \emptyset$ .

Если  $X_* \neq \emptyset$ , то, очевидно, точная нижняя грань  $f(x)$  на  $X$  совпадает с наименьшим значением этой функции на  $X$ , т.е.  $\inf_{x \in X} f(x) = \min_{x \in X} f(x)$ . В этом случае говорят, что функция  $f(x)$  *достигает своей точной нижней грани* в каждой из точек  $x_* \in X_*$ . Подчеркнем, что  $\inf_{x \in X} f(x) = f_*$  всегда существует [22, 29],

причем если  $f(x)$  ограничена снизу на  $X$  числом  $M_0$ , то  $f_* \geq M_0 > -\infty$ , и  $f_* = -\infty$ , если  $f(x)$  не ограничена снизу  $X$ . Заметим, что если функция достигает некоторой своей нижней грани, то эта нижняя грань необходимо будет и точной нижней гранью функции, поэтому в дальнейшем в тексте мы в некоторых случаях для краткости будем говорить «функция достигает своей нижней грани».

После введенных определений можем уточнить постановку задачи (1.1). Здесь обычно различают задачи двух типов. К *первому типу* относят задачи, в которых требуется определить величину  $f_* = \inf_{x \in X} f(x)$  — точную нижнюю грань функции  $f(x)$  на множестве  $X$ . В таких задачах возможно, что  $X_* = \emptyset$ . В задачах *второго типа* ищется не только величина  $f_*$ , но и какая-либо точка  $x_* \in X_*$ , в которой достигается точная нижняя грань, — здесь уже подразумевается, что  $X_* \neq \emptyset$ . Если множество  $X_*$  состоит более чем из одной точки, то здесь возможны уточнения постановки задачи (1.1). Например, нужно искать точку  $x_* \in X_*$ , которая ближе всего расположена к заданной точке  $x_0$ .

Возможна и более широкая постановка задачи минимизации второго типа, когда ищутся не только точки минимума в смысле определения 1.1, но и точки так называемого локального минимума.

**Определение 1.4.** Точка  $x_* \in X$  называется *точкой локального минимума* функции  $f(x)$  на множестве  $X$  со значением  $f_* = f(x_*)$ , если существует такая малая  $\varepsilon$ -окрестность  $O(x_*, \varepsilon) = \{x \in E^n : \|x - x_*\| < \varepsilon\}$  точки  $x_*$ , что  $f(x_*) \leq f(x)$  для всех  $x \in X \cap O(x_*, \varepsilon)$ . Если при некотором  $\varepsilon > 0$  равенство  $f(x_*) = f(x)$  для  $x \in X \cap O(x_*, \varepsilon)$  возможно только при  $x = x_*$ , то  $x_*$  называют точкой строгого локального минимума.

Для функции, график которой изображен на рис. 1.5, точки  $x_0, x_2, x_4$  являются точками строгого локального минимума, а в точках  $x$ , удовлетворяющих неравенствам  $x_5 < x \leq x_6$  и  $x_8 \leq x \leq x_9$ , реализуется нестрогий локальный минимум. Других точек локального минимума у данной функции нет.

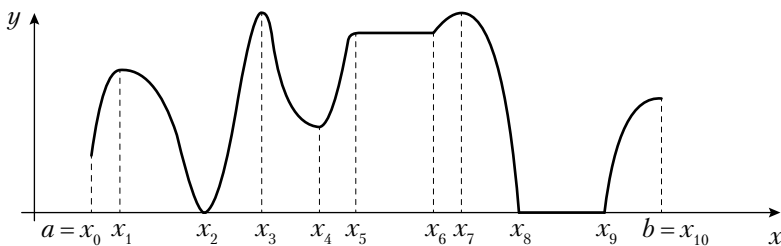


Рис. 1.5

Точки локального минимума, в которых минимум достигается в смысле определения 1.1, в литературе часто называют *точками глобального минимума* или *точками абсолютного минимума* функции  $f(x)$  на множестве  $X$ . Для функции, изображенной на рис. 1.5,  $X_* = \{x_2\} \cup [x_8; x_9]$  — точки глобального минимума.

### 1.1.2. Задача максимизации

Перейдем к обсуждению постановки задачи максимизации, которую кратко будем записывать так:

$$f(x) \rightarrow \max, x \in X. \quad (1.2)$$

Для пояснения постановки задачи (1.2), как и в случае задачи минимизации (1.1), нам понадобятся некоторые определения.

**Определение 1.5.** Точку  $x^* \in X$  будем называть *точкой максимума* функции  $f(x)$  на множестве  $X$ , если  $f(x^*) \geq f(x)$  для всех  $x \in X$ ; величину  $f(x^*)$  назовем наибольшим или максимальным значением  $f(x)$  на  $X$  и обозначим  $\max_{x \in X} f(x) = f(x^*) = f^*$ . Множество всех точек максимума  $f(x)$  на  $X$  будем обозначать через  $X^*$ .

**Определение 1.6.** Функция  $f(x)$  называется *ограниченной сверху* на множестве  $X$ , если существует такое число  $M_1$ , что  $f(x) \leq M_1$  для всех  $x \in X$ . Функция  $f(x)$  не ограничена сверху на  $X$ , если найдется последовательность  $\{x_k\} \in X$  такая, что  $\lim_{k \rightarrow \infty} f(x_k) = +\infty$ . Если функция ограничена на  $X$  и сверху, и снизу, то она называется *ограниченной на  $X$* .

**Определение 1.7.** Пусть функция  $f(x)$  ограничена сверху на множестве  $X$ . Тогда число  $f^*$  называется *точной верхней гранью* или *супремумом*  $f(x)$  на  $X$ , если:

1)  $f^* \geq f(x)$  при всех  $x \in X$ ;

2) для любого сколь угодно малого числа  $\varepsilon > 0$  найдется точка  $x_\varepsilon \in X$  такая, что  $f(x_\varepsilon) > f^* - \varepsilon$ .

Точную верхнюю грань  $f(x)$  на  $X$  будем обозначать через  $\sup_{x \in X} f(x) = f^*$ .

Если функция  $f(x)$  не ограничена сверху на  $X$ , то в качестве точной верхней грани  $f(x)$  на  $X$  принимается  $f^* = +\infty$ .

Если  $X^* \neq \emptyset$ , то, очевидно, точная верхняя грань  $f(x)$  на  $X$  совпадает с максимальным значением этой функции на  $X$ , т.е.  $\sup_{x \in X} f(x) = \max_{x \in X} f(x)$ .

В этом случае говорят, что функция  $f(x)$  на  $X$  *достигает своей точной верхней грани* в точке  $x^* \in X^*$ . Подчеркнем, что  $\sup_{x \in X} f(x) = f^*$  всегда существует

[22, 29], причем если  $f(x)$  ограничена сверху на  $X$  числом  $M_1$ , то  $f^* \leq M_1 < +\infty$ , и  $f^* = +\infty$ , если  $f(x)$  не ограничена сверху на  $X$ . Выражение  $\max_{x \in X} f(x)$

не всегда имеет смысл. Справедливы соображения, аналогичные высказанным ранее: если функция достигает какой-то своей верхней грани, то эта верхняя грань необходимо будет точной верхней гранью.

Для иллюстрации сказанного обратимся к функциям  $f(x)$  и множествам  $X$  из примеров 1.1–1.4, но задачи минимизации в них заменим задачами максимизации.

В примере 1.1 для обоих множеств  $X_1 = E^n$ ,  $X_2 = \{\|x\| \geq 1\}$  максимальное значение функции  $f(x)$  равно  $f^* = 1$ ,  $X^* = \{\|x\| = 1\}$  —  $n$ -мерная сфера радиуса 1 с центром  $x = 0$  (при  $n = 1$  это две точки  $x^* = 1$  и  $x^* = -1$ ), что следует из неравенства

$$1 - f(x) = 1 - \frac{2\|x\|}{1 + \|x\|^2} = \frac{(\|x\| - 1)^2}{1 + \|x\|^2} \geq 0 \quad \forall x \in E^n,$$

которое при  $\|x\| = 1$  обращается в равенство  $f(x) = 1 = f^*$ .

В примере 1.2 на множестве  $X = X_2 = \{\|x\| \geq 1\}$  максимальное значение равно  $f^* = 1$ , множество  $X^* = \{\|x\| = 1\}$ ; для остальных множеств  $f^* = +\infty$ ,  $X^* = \emptyset$ .

В примере 1.3 на множестве  $X = [0; +\infty)$  имеем  $f^* = 1$ ,  $X^* = \left\{ x = \frac{2}{2k+1}, k = 0, 1, \dots \right\}$ .

В примере 1.4  $f^* = 0$ ,  $X^* = \{x = 1\}$ .

В задачах максимизации (1.2) также можно различать задачи двух типов: в задаче *первого типа* ищется величина  $f^*$ , а в задачах *второго типа* ищутся  $f^*$  и какая-либо точка максимума  $x^*$  такая, что  $f(x^*) = f^*$ .

Задачи (1.1), (1.2) тесно связаны между собой равенством

$$\sup_{x \in X} f(x) = - \inf_{x \in X} (-f(x)), \quad (1.3)$$

из которого следует, что всякая задача максимизации функции  $f(x)$  на  $X$  равносильна задаче минимизации

$$g(x) \rightarrow \min, \quad x \in X, \quad g(x) \stackrel{\text{def}}{=} -f(x). \quad (1.4)$$

Отсюда ясно, что методы, разработанные для решения задач минимизации, нетрудно переформулировать и использовать для решения задач максимизации. Поэтому в дальнейшем мы в основном будем исследовать задачу минимизации (1.1).

Сказанное относится и к задачам поиска точек локального максимума функции  $f(x)$  на множестве  $X$ .

**Определение 1.8.**  $x^* \in X$  называется точкой *локального максимума* функции  $f(x)$  на множестве  $X$  со значением  $f^* = f(x^*)$ , если существует такая малая  $\varepsilon$ -окрестность  $O(x^*, \varepsilon) = \{x \in E^n : \|x - x^*\| < \varepsilon\}$  точки  $x^*$ , что  $f(x^*) \geq f(x)$  для всех  $x \in X \cap O(x^*, \varepsilon)$ . Если при некотором  $\varepsilon > 0$  равенство  $f(x) = f(x^*)$  для  $x \in X \cap O(x^*, \varepsilon)$  возможно только при  $x = x^*$ , то  $x^*$  называют точкой *строгого локального максимума*.

Точки локального максимума, в которых максимум достигается в смысле определения 1.5, часто называют *точками глобального максимума* или *точками абсолютного максимума* функции  $f(x)$  на множестве  $X$ .

Для функции, график которой изображен на рис. 1.5, точки  $x_1, x_3, x_7, x_{10}$  являются точками строгого локального максимума, а в точках  $x$ , удовлетворяющих неравенствам  $x_5 \leq x < x_6$  и  $x_8 < x < x_9$ , реализуется нестрогий локальный максимум;  $x_3, x_7$  — точки глобального максимума. Других точек локального максимума у нее нет. Обращаем внимание читателя на то, что все точки  $x$  из интервалов  $(x_5, x_6)$  и  $(x_8, x_9)$  одновременно являются как точками локального минимума, так и точками локального максимума.

Множество всех точек локального минимума или максимума функции на множестве  $X$  принято называть *точками локального экстремума* функции на этом множестве или, проще, *точками экстремума*. Если мы хотим найти все точки экстремума (как минимума, так и максимума), то, объединяя постановки задач (1.1), (1.2), кратко будем писать

$$f(x) \rightarrow \text{extr}, \quad x \in X.$$

Функцию  $f(x)$  в задачах (1.1), (1.2) часто называют *целевой функцией*, множество  $X$  — *допустимым множеством*, точки  $x \in X$  — *допустимыми точками*.

## 1.2. Теоремы Вейерштрасса

### 1.2.1. Классическая теорема

Как мы видели в простейших примерах 1.1–1.4, функция  $f(x)$  может быть ограниченной или неограниченной сверху или снизу на множестве  $X$ , множество точек ее экстремума может быть пустым или непустым. Приведем *теорему Вейерштрасса из классического математического анализа*, в которой даются достаточные условия на целевую функцию  $f(x)$  и допустимое множество  $X$ , гарантирующие конечность точной нижней и точной верхней граней и непустоту множества точек экстремума  $f(x)$  на  $X$ . Сна-

чала напомним некоторые определения и факты из математического анализа [20, 22, 29].

**Определение 1.9** Пусть  $\{x_k\} = (x_1, x_2, \dots, x_k, \dots)$  — некоторая последовательность точек из  $E^n$ . Последовательность  $\{x_k\}$  называется *ограниченной*, если существует такое число  $M \geq 0$ , что  $\|x_k\| \leq M$  для всех  $k = 1, 2, \dots$ . Последовательность  $\{x_k\}$  называется *неограниченной*, если существует ее подпоследовательность  $\{x_{k_l}\}$  такая, что  $\lim_{l \rightarrow \infty} \|x_{k_l}\| = +\infty$ . Точка  $v$  называется *предельной точкой* последовательности  $\{x_k\}$ , если существует ее подпоследовательность  $\{x_{k_m}\}$ , сходящаяся к  $v$ .

Согласно теореме Больцано — Вейерштрасса [20, 22, 29] всякая ограниченная последовательность  $\{x_k\} \in E^n$  имеет хотя бы одну предельную точку.

**Определение 1.10.** Множество  $X \subset E^n$  называется *ограниченным*, если существует число  $M \geq 0$  такое, что  $\|x\| \leq M$  для всех  $x \in X$ . Множество  $X$  называется *неограниченным*, если существует последовательность  $\{x_k\} \in X$  такая, что  $\lim_{k \rightarrow \infty} \|x_k\| = +\infty$ .

**Определение 1.11.** Точка  $v \in E^n$  называется *предельной точкой* множества  $X$ , если любая ее  $\varepsilon$ -окрестность  $O(v, \varepsilon) = \{x \in E^n : \|x - v\| < \varepsilon\}$  содержит точку  $w \in X$ , отличную от  $v$  ( $w \neq v$ ).

Из этого определения следует, что точка  $v$  будет предельной точкой множества  $X$  тогда и только тогда, когда существует последовательность  $\{x_k\} \in X$ ,  $x_k \neq v$ ,  $k = 1, 2, \dots$ , сходящаяся к точке  $v$ . Предельная точка множества  $X$  может как принадлежать  $X$ , так и не принадлежать ему.

**Определение 1.12.** Точка  $v$  называется *внутренней точкой* множества  $X$ , если существует ее  $\varepsilon$ -окрестность  $O(v, \varepsilon) = \{x \in E^n : \|x - v\| < \varepsilon\}$ , которая целиком принадлежит множеству  $X$ . Если найдется  $\varepsilon$ -окрестность точки  $v$ , которая не содержит ни одной точки из  $X$ , то  $v$  назовем *внешней точкой* множества  $X$ . Если любая  $\varepsilon$ -окрестность точки  $v$  содержит как точки из  $X$ , так и точки, не принадлежащие  $X$ , то  $v$  называется *граничной точкой* множества  $X$ . Наконец, возможно, что точка  $v$  принадлежит множеству  $X$ , но существует  $\varepsilon$ -окрестность этой точки, не содержащая никакой другой точки этого множества. Такая точка  $v$  называется *изолированной точкой* множества  $X$ .

Совокупность всех внутренних точек множества  $X$  будем обозначать через  $\text{int} X$ , совокупность всех граничных точек — через  $\text{Gr} X$ .

Уточним, что изолированная точка  $v$  множества  $X$  является его граничной точкой, причем она является предельной точкой единственной последовательности элементов множества  $X$ , а именно  $\{x_k = v\}$ ,  $k = 1, 2, \dots$ . Внутренние точки множества и его граничные точки, не являющиеся изолированными, будут предельными точками множества  $X$ , а внешние и изолированные точки множества  $X$  не являются его предельными точками.

**Определение 1.13.** Множество  $X \subseteq E^n$  называется *замкнутым*, если оно содержит все свои предельные точки.

Из вышесказанного вытекает, что замкнутое множество является объединением всех своих внутренних и граничных (включая изолированные) точек. В частности, множество, состоящее только из изолированных точек, является замкнутым.

**Определение 1.14.** Множество  $X \subseteq E^n$  такое, что каждая последовательность  $\{x_k\} \in X$  обладает хотя бы одной предельной в смысле определения 1.11 точкой, лежащей в  $X$ , называется *компактным*.

Из теоремы Больцано — Вейерштрасса и определений 1.9—1.12 следует, что в  $E^n$  компактными являются все замкнутые ограниченные множества и только они (т.е. любое компактное множество из  $E^n$  ограничено и содержит все свои предельные точки).

Примерами компактных множеств в пространстве  $E^n$  являются следующие:

- $n$ -мерный шар радиуса  $R$  с центром в точке  $x_0 \in E^n$

$$X = \{x = (x^1, \dots, x^n)^\top \in E^n : \|x - x_0\| = [(x^1 - x_0^1)^2 + \dots + (x^n - x_0^n)^2]^{1/2} \leq R\};$$

- $n$ -мерная сфера  $X = \{x \in E^n : \|x - x_0\| = R\}$ ;
- $n$ -мерный параллелепипед  $X = \{x \in E^n : a_i \leq x^i \leq b_i, i = 1, \dots, n\}$ , где  $a_i, b_i, a_i \leq b_i, i = 1, \dots, n$  — заданные числа;
- отрезок  $X = [u, v] = \{x \in E^n : x = x_\alpha = \alpha u + (1 - \alpha)v, 0 \leq \alpha \leq 1\}$ , соединяющий точки  $u, v \in E^n$ ;
- множество  $X = \{x \in E^n : \|x\|(\|x\| - 1)(\|x\| - 2) \leq 0\}$ , являющееся объединением кольца  $1 \leq \|x\| \leq 2$  и изолированной точки 0.

Примерами некомпактных множеств являются следующие замкнутые, но неограниченные множества:

- все пространство  $X = E^n$ ;
- неотрицательный ортант  $X = E_+^n \stackrel{\text{def}}{=} \{x \in E^n : x^1 \geq 0, \dots, x^n \geq 0\}$ ;
- гиперплоскость  $X = \{x \in E^n : \langle c, x \rangle = \gamma\}$ , где  $\gamma$  — заданное число,  $c \in E^n, c \neq 0$  — заданный вектор, который называют *нормальным вектором* гиперплоскости;
- полупространство  $X = \{x \in E^n : \langle c, x \rangle \leq \gamma\}$  или  $X = \{x \in E^n : \langle c, x \rangle \geq \gamma\}$ ;
- прямая  $X = \{x \in E^n : x = x(t) = x_0 + td, -\infty < t < +\infty\}$ , проходящая через точку  $x_0$ , с направляющим вектором  $d \in E^n, d \neq 0$ ;
- луч  $X = \{x \in E^n : x = x(t) = x_0 + td, t \geq 0\}$ .

Открытый луч  $X = \{x \in E^n : x = x(t) = x_0 + td, t > 0\}$  и открытый  $n$ -мерный шар  $X = \{x \in E^n : \|x - x_0\| < R\}$  также являются некомпактными множествами.

**Определение 1.15.** Пусть функция  $f(x)$  определена на множестве  $X \subset E^n$ . Говорят, что функция  $f(x)$  *непрерывна в точке*  $x \in X$ , если для любой последовательности  $\{x_k\} \in X$ , сходящейся к точке  $x$ , справедливо равенство  $\lim_{k \rightarrow \infty} f(x_k) = f(x)$ . Функцию  $f(x)$  называют *непрерывной на множестве*  $X$ , если она непрерывна в каждой точке этого множества.

Приведем формулировку теоремы Вейерштрасса для непрерывных функций в том виде, как она дается в обычных курсах математического анализа [20, 22, 29].

**Теорема 1.1.** Пусть  $X$  — компактное множество из  $E^n$ , функция  $f(x)$  определена и непрерывна на  $X$ . Тогда

$$f_* = \inf_{x \in X} f(x) > -\infty, \quad X_* = \{x \in X : f(x) = f_*\} \neq \emptyset;$$

$$f^* = \sup_{x \in X} f(x) < +\infty, \quad X^* = \{x \in X : f(x) = f^*\} \neq \emptyset.$$



### 1.2.2. Обобщенная теорема Вейерштрасса

Теорема 1.1 дает достаточные условия разрешимости одновременно как задачи минимизации (1.1), так и задачи максимизации (1.2).

Между тем на простейших примерах 1.1–1.4 мы видели, что эти задачи не всегда разрешимы одновременно. Это обстоятельство наводит на мысль о том, что условия теоремы 1.1, гарантирующие одновременную разрешимость задач (1.1), (1.2), по-видимому, являются слишком жесткими. Внимательный анализ теоремы 1.1 показывает, что условие непрерывности целевой функции  $f(x)$  можно существенно ослабить. Оказывается, для разрешимости задачи минимизации достаточно полунепрерывности снизу функции  $f(x)$ , для задачи максимизации — ее полунепрерывности сверху. Дадим строгое определение этих понятий. Предварительно введем понятие нижнего и верхнего пределов последовательности.

**Определение 1.16.** Число  $a$  называется *нижним (верхним) пределом* ограниченной снизу (сверху) числовой последовательности  $\{a_k\}$ , если:

- 1) существует хотя бы одна подпоследовательность  $\{a_{k_m}\}$ , сходящаяся к  $a$ ;
- 2) все предельные точки последовательности  $\{a_k\}$  не меньше (не больше) числа  $a$ , т.е. число  $a$  является наименьшей (наибольшей) предельной точкой последовательности  $\{a_k\}$ .

Нижний и верхний пределы будем обозначать  $\varliminf_{k \rightarrow \infty} a_k$  и  $\overline{\lim}_{k \rightarrow \infty} a_k$  соответственно.

Иначе говоря,  $a$  — нижний (верхний) предел последовательности  $\{a_k\}$ , если для любого  $\varepsilon > 0$ :

- 1) существует номер  $N$  такой, что  $a_k \geq a - \varepsilon$  ( $a_k \leq a + \varepsilon$ ) для всех  $k \geq N$ ;
- 2) для любого номера  $m$  найдется номер  $k_m > m$  такой, что  $a_{k_m} \leq a + \varepsilon$  ( $a_{k_m} \geq a - \varepsilon$ ).

В том случае, когда  $\{a_k\}$  не ограничена снизу (сверху), то по определению принимают  $\varliminf_{k \rightarrow \infty} a_k = -\infty$  ( $\overline{\lim}_{k \rightarrow \infty} a_k = +\infty$ ); в частности, если  $\lim_{k \rightarrow \infty} a_k = -\infty$ , то  $\varliminf_{k \rightarrow \infty} a_k = -\infty$ ; (если  $\lim_{k \rightarrow \infty} a_k = +\infty$ , то  $\overline{\lim}_{k \rightarrow \infty} a_k = +\infty$ ).

Например:

- если  $a_k = (-1)^k$ ,  $k = 1, 2, \dots$ , то  $\varliminf_{k \rightarrow \infty} a_k = -1$ ,  $\overline{\lim}_{k \rightarrow \infty} a_k = 1$ ;
- если  $a_k = (-1)^k k$ ,  $k = 1, 2, \dots$ , то  $\varliminf_{k \rightarrow \infty} a_k = -\infty$ ,  $\overline{\lim}_{k \rightarrow \infty} a_k = +\infty$ ;
- если  $a_k = [1 + (-1)^k]k$ ,  $k = 1, 2, \dots$ , то  $\varliminf_{k \rightarrow \infty} a_k = 0$ ,  $\overline{\lim}_{k \rightarrow \infty} a_k = +\infty$ ;
- если  $a_k = k^{-1}$ ,  $k = 1, 2, \dots$ , то  $\varliminf_{k \rightarrow \infty} a_k = \lim_{k \rightarrow \infty} a_k = 0$ .

Важно заметить, что нижний и верхний пределы (конечные или бесконечные), в отличие от обычного предела, у любой последовательности всегда существуют. Для того чтобы последовательность  $\{a_k\}$  имела обычный предел, необходимо и достаточно, чтобы  $\varliminf_{k \rightarrow \infty} a_k = \overline{\lim}_{k \rightarrow \infty} a_k = a$ ; тогда  $\lim_{k \rightarrow \infty} a_k = a$ .

**Определение 1.17.** Пусть функция  $f(x)$  определена на множестве  $X \subseteq E^n$ . Говорят, что функция  $f(x)$  *полунепрерывна снизу (сверху) в точке*  $x \in X$ ,