МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования «Нижегородский государственный университет им. Н.И. Лобачевского»

В. И. Сумин

Начала математического программирования. Теорема Вейерштрасса. Безусловный экстремум

Учебно-методическое пособие

Рекомендовано методической комиссией механико-математического факультета для студентов ННГУ, обучающихся по направлениям подготовки «Математика», «Математика и компьютерные науки», «Прикладная математика и информатика», «Механика и математическое моделирование»

Нижний Новгород
2015

УДК 519.6
ББК 22.193
С-89

С-89 Сумин В.И. НАЧАЛА МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ. ТЕОРЕМА ВЕЙЕРШТРАССА. БЕЗУСЛОВНЫЙ ЭКСТРЕМУМ: Учебно-методическое пособие. – Нижний Новгород: Нижегородский государственный университет, 2015. – 40 с.

Рецензент: д.ф.-м.н., профессор **М.А. Федоткин**

Учебно-методическое пособие по общему курсу «Методы оптимизации» для студентов бакалавриата ННГУ, обучающихся по направлениям подготовки «Математика», «Математика и компьютерные науки», «Прикладная математика и информатика», «Механика и математическое моделирование». Пособие содержит материал нескольких первых аудиторных занятий по теме «Основы теории математического программирования», а также включает материал для самостоятельного изучения.

Ответственный за выпуск:
председатель методической комиссии
механико-математического факультета ННГУ,
к.ф.-м.н., доцент **Н.А. Денисова**

УДК 519.6
ББК 22.193

Содержание

Предисловие

Данное учебно-методическое пособие по общему курсу "Методы оптимизации", читаемому для студентов механико-математического факультета ННГУ, включает материал нескольких первых занятий по теме "Основы теории математического программирования". На первых занятиях прежде всего рассматриваются те разделы теории конечномерных экстремальных задач, которые в несколько меньшем объеме изучались ранее в курсе "Математический анализ". В частности, это теорема Вейерштрасса о существовании экстремума, необходимые условия экстремума в гладких задачах на безусловный экстремум и достаточные условия экстремума в таких задачах. Таким образом устанавливается мост между двумя общими курсами. Пособие содержит основные положения указанных разделов теории математического программирования, примеры решения соответствующих задач и упражнения.

Данное пособие - заново отредактированный вариант методической разработки [Сумин В.И. Классические разделы теории конечномерных экстремальных задач. Часть 1. Теорема Вейерштрасса. Безусловный экстремум. Методическая разработка для студентов механико-математического факультета ННГУ. Нижний Новгород: Изд-во ННГУ. 2001].

Нумерация формул в пособии сквозная. Нумерация теорем, лемм, следствий, замечаний, примеров, упражнений - двойная, как и нумерация пунктов в параграфах. Например, третья по порядку теорема второго параграфа имеет номер 2.3.

Обозначения и сокращения

соотв. - "соответственно";

$\equiv$ - "равно по определению" или "тождественно равно";

$\forall$ - "для всех", "для любого", "для каждого";

$\Leftrightarrow$ - знак эквивалентности двух утверждений;

$\emptyset$ - пустое множество;

$\{x \in X : \ldots\}$ - совокупность элементов множества X, обладающих свойством "$\ldots$";

R^n - n-мерное пространство действительных векторов-столбцов

$$x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix},$$

наделенное евклидовым скалярным произведением $\langle x, y\rangle_n \equiv \sum\limits_{i=1}^{n} x^i y^i$ $(x, y \in R^n)$ и евклидовой нормой $\|x\|_n \equiv \sqrt{\langle x, x\rangle_n}$;

$0_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in R^n$ - нуль пространства R^n;

$R \equiv R^1$;

$R_+ \equiv \{x \in R : x \geq 0\}$;

$U_\varepsilon(\overline{x}) \equiv \{x : \|x - \overline{x}\|_n < \varepsilon\}$ - ε-окрестность элемента $\overline{x}$ в R^n;

A^T - матрица, траспонированная по отношению к матрице A (например, вектор-столбец $x \in R^n$ получается транспонированием вектора-строки $\left(x^1, x^2, \ldots, x^n\right) : x = \left(x^1, x^2, \ldots, x^n\right)^T$);

$\bigtriangledown f(x)$ - градиент функции $f(x)$ векторного аргумента $x \in R^n$, то есть состоящая из частных производных вектор-строка

$$\bigtriangledown f(x) \equiv \left(f'_{x^1}(x), f'_{x^2}(x), \cdots, f'_{x^n}(x)\right);$$

$\bigtriangledown^2 f(x)$ - матрица вторых производных (матрица Гессе) функции $f(x)$ векторного аргумента $x \in R^n$, то есть $(n \times n)$-матрица

$$\bigtriangledown^2 f(x) \equiv \begin{pmatrix} f''_{x^1x^1}(x) & f''_{x^1x^2}(x) & \cdots & f''_{x^1x^n}(x) \\ f''_{x^2x^1}(x) & f''_{x^2x^2}(x) & \cdots & f''_{x^2x^n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f''_{x^nx^1}(x) & f''_{x^nx^2}(x) & \cdots & f''_{x^nx^n}(x) \end{pmatrix}.$$

§0. Основные определения

Конечномерными экстремальными задачами или *конечномерными задачами на экстремум* называются задачи минимизации и задачи максимизации функций нескольких переменных. Пусть $X \subset R^n$ - некоторое множество, $f(\cdot) : X \to R$ - некоторая функция. Задачу минимизации функции $f(\cdot)$ на множестве X будем записывать в виде

$$f(x) \to \min, \quad x \in X. \tag{1}$$

Соответственно задачу максимизации функции $f(\cdot)$ на множестве X - в виде

$$f(x) \to \max, \quad x \in X. \tag{2}$$

Подчеркнем, что (1) и (2) это просто символы задач минимизации и максимизации, которые в каждом конкретном случае должны сопровождаться разъяснением того, что значит "решить задачу".

Познакомимся с терминологией, принятой в теории экстремальных задач. Рассмотрим, например, задачу минимизации (1). Когда, как мы предположили, множество X принадлежит пространству R^n, задача (1) называется *n-мерной задачей минимизации.* Если при этом X совпадает со всем R^n, то мы имеем дело с *задачей на безусловный минимум* или *задачей безусловной минимизации.* В противном случае задача (1) это *задача на условный минимум* или *задача условной минимизации.* Множество X называется *множеством допустимых элементов* или *ограничением* задачи (1). Задачу безусловной минимизации называют еще *задачей минимизации без ограничений.* Говорят, что точка $\overline{x}$ - точка *локального минимума* (соотв. *строгого локального минимума*) задачи (1) и функции $f(x)$, $x \in X$, если существует число $\varepsilon > 0$ такое, что

$$f(\overline{x}) \le f(x) \quad \forall \quad x \in (X \cap U_\varepsilon(\overline{x}))$$

$$(\text{соотв. } f(\overline{x}) < f(x) \quad \forall \quad x \in (X \cap U_\varepsilon(\overline{x})), \quad x \neq \overline{x}).$$

Если

$$f(\overline{x}) \le f(x) \quad \forall \quad x \in X$$

$$(\text{соотв. } f(\overline{x}) < f(x) \quad \forall \quad x \in X, \quad x \neq \overline{x}),$$

то $\overline{x}$ называют точкой *глобального минимума* (соотв. *строгого глобального минимума*) задачи (1) и функции $f(x)$, $x \in X$. В этом случае пишут

$$f(\overline{x}) = \min_{x \in X} f(x).$$

Если у функции $f(x)$, $x \in X$ есть точки минимума (соотв. локального, глобального, ...), то говорят, что эта *функция имеет минимум* (соотв. локальный, глобальный, ...). Выражение "функция $f(x)$, $x \in X$ имеет глобальный минимум" эквивалентно выражению "*функция* $f(\cdot)$ *достигает наименьшего значения (достигает минимума) на множестве* X ". Вместо слов "глобальный минимум" часто говорят "*абсолютный минимум*". Понятно, что всякая точка глобального (соотв. строгого глобального) минимума является и точкой локального (соотв. строгого локального) минимума. Обратное, вообще говоря, неверно. Множество всех точек глобального минимума задачи (1) будем обозначать через X_*. Иначе говоря,

$$X_* \equiv \left\{ x \in X : f(x) = \min_{y \in X} f(y) \right\}.$$

Точную нижнюю грань множества значений функции

$$f_* \equiv \inf_{x \in X} f(x),$$

которая, очевидно, либо конечна, либо равна $(-\infty)$, называют *значением задачи минимизации* (1). Если $X_* \neq \emptyset$, то $f_* = \min\limits_{x \in X} f(x)$.

Пример 0.1. Пусть $X = R$, $f(x) \equiv e^x$. Тогда $f_* = 0$, $X_* = \emptyset$. То есть функция $f(\cdot)$ ограничена снизу, но не достигает минимума на X. Точек локального минимума нет.

Пример 0.2. Пусть $X = R_+$, $f(x) \equiv e^x$. Тогда $f_* = 1$, $X_* = \{0\}$, множество точек локального минимума совпадает с X_*. Точка $x = 0$ является точкой строгого локального и строгого глобального минимума.

Пример 0.3. Пусть $X = R$, $f(x) \equiv \sin x$. Тогда $f_* = -1$, $X_* = \{(2k - (1/2))\,\pi : k - \text{целое}\}$, множество точек локального минимума совпадает с X_*. Все точки локального минимума являются точками строгого локального минимума. Так как X_* состоит более чем из одной точки, то точки глобального минимума не являются точками строгого глобального минимума.

Проделайте следующие упражнения 0.1 - 0.6, пользуясь при необходимости результатами дифференциального исчисления функций одной переменной.

Упражнение 0.1. Пусть $X = R$, $f(x) \equiv e^x \cdot \sin x$. Докажите, что тогда $f_* = -\infty$, $X_* = \emptyset$, множеством точек локального минимума является множество $\{\pi\,(k - (1/4)) : k - \text{ четное}\}$, каждая точка локального минимума является точкой строгого локального минимума.

Упражнение 0.2. Пусть $X = \{x \in R : x \leq 0\}$, $f(x) \equiv e^x \cdot \sin x$. Докажите, что тогда $f_* = f(-\pi/4)$, $X_* = \{-\pi/4\}$, точка $x = -\pi/4$ является точкой строгого глобального минимума, множеством точек локального минимума является множество $\{\pi\,(k-(1/4)) : k = 0, -2, -4, \ldots\}$, каждая точка локального минимума является точкой строгого локального минимума.

Упражнение 0.3. Докажите, что функция, определенная на действительной оси формулой

$$f\,(x) = \begin{cases} x^2\,(2+\sin\,(1/x)) & , \quad x \neq 0; \\ 0 & , \quad x = 0, \end{cases}$$

имеет в точке $x = 0$ локальный минимум, но в любой сколь угодно малой окрестности $U_\varepsilon\,(0)$ не является ни убывающей слева от точки 0, ни возрастающей справа от этой точки (даже если убывание и возрастание понимать в нестрогом смысле).

Упражнение 0.4. Приведите пример одномерной задачи минимизации, в которой имеется единственная точка локального минимума, множество точек глобального минимума пусто, причем: а) значение задачи конечно, б) значение задачи равно $(-\infty)$.

Упражнение 0.5. Докажите, что для функции, определенной на плоскости R^2 формулой

$$f\,(x) \equiv \left(x^2 - \left(x^1\right)^2\right)\cdot\left(x^2 - 3\left(x^1\right)^2\right), \quad x \in R^2,$$

точка 0_2 не является точкой локального минимума, хотя сужение этой функции на любую прямую, проходящую через 0_2, имеет в точке 0_2 строгий локальный минимум (см., например, [ГО], С.155).

Упражнение 0.6. Докажите, что для функции, определенной на плоскости R^2 формулой

$$f\,(x) \equiv \begin{cases} \left(x^2 - e^{-1/\left(x^1\right)^2}\right)\cdot\left(x^2 - 3\cdot e^{-1/\left(x^1\right)^2}\right) & , \quad x^1 \neq 0; \\ \left(x^2\right)^2 & , \quad x^1 = 0, \end{cases} \qquad x \in R^2,$$

точка 0_2 не является точкой локального минимума, хотя сужение этой функции на любую кривую вида $x^2 = c\cdot\left(x^1\right)^{m/k}$, где m и k - взаимно простые натуральные числа (k нечетно), $c \in R$, имеет в точке 0_2 строгий локальный минимум (см., например, [ГО], С.156). Докажите, что условие нечетности k здесь можно убрать, сузив область определения функции до полуплоскости $x^1 \geq 0$ (см. [ГО], С.156).

Слова "решить задачу минимизации (1)" в различных конкретных ситуациях, возникающих в приложениях, могут иметь разный смысл. Иногда требуется найти лишь значение задачи f_*, при этом может даже не вставать вопрос о множестве X_*. Иногда необходимо найти как раз множество X_* или хотя бы какой-то его элемент, а может быть и просто установить факт непустоты или пустоты X_*, то есть изучить вопрос о разрешимости задачи (1). В некоторых случаях бывает достаточно ограничиться нахождением лишь какой-либо точки локального минимума. И так далее ... . При этом мы пока говорим лишь о "решении задачи минимизации в точном смысле", оставляя в стороне чрезвычайно важный вопрос о ее приближенном решении.

Далее, для определенности, будем понимать задачу минимизации (1) как задачу нахождения точек минимума функции $f(x)\,,\ x \in X$. Соответственно, *решением задачи* (1) будем называть любую точку минимума, локального или глобального, этой функции. Точки глобального (соотв. локального) минимума функции $f(x)\,,\ x \in X$ будем называть *глобальными* (соотв. *локальными*) *решениями задачи* (1). Точки строгого минимума (локального или глобального) этой функции будем называть *строгими* (*локальными* или *глобальными*) *решениями* (1).

Рассмотрим задачу максимизации (2). Заменяя в приведенных выше определениях точек минимума задачи минимизации (1) и функции $f(x)\,,$ $x \in X$ слово "минимум" словом "максимум", знак неравенства "$\leq$" знаком "$\geq$", знак неравенства "$<$" знаком "$>$", а номер (1) номером (2), получим определения: *точки локального максимума*, *точки строгого локального максимума*, *точки глобального максимума*, *точки строгого глобального максимума задачи максимизации* (2) *и функции* $f(x)\,,\ x \in X$. Точную верхнюю грань множества значений функции

$$f^* \equiv \sup_{x \in X} f(x)\,,$$

которая, очевидно, либо конечна, либо равна $+\infty$, называют *значением задачи максимизации* (2).

Далее задачу максимизации (2) будем понимать как задачу нахождения точек максимума функции $f(x)\,,\ x \in X$. Соответственно, *решением задачи* (2) будем называть любую точку максимума, локального или глобального, этой функции. Точки глобального (соотв. локального) максимума функции $f(x)\,,\ x \in X$ будем называть *глобальными* (соотв. *локальными*) *решениями задачи* (2). Точки строгого максимума (локального или глобального) этой функции будем называть *строгими* (*локальными* или *глобальными*) *решениями* (2). Понимаемая так задача максимизации (2) эквивалентна задаче минимизации

$$(-f(x)) \to \min, \qquad x \in X \tag{3}$$

в том смысле, что совпадают множества глобальных решений задач (2) и (3), а также, соответственно, множества их локальных решений, множества их строгих локальных решений и множества их строгих глобальных решений. Поэтому все основные положения теории экстремальных задач мы запишем далее для задачи минимизации (1). Из этих положений легко вывести аналогичные утверждения для задачи максимизации (2). Читателю будет полезно проделать это самостоятельно (соответствующие задания помещены ниже в упражнениях).

Упражнение 0.7. Докажите, что ограниченная функция

$$f(x) \equiv (arctgx) \cdot \cos x, \qquad x \in R$$

имеет локальные минимумы и максимумы, но абсолютные минимум и максимум ею не достигаются.

Точки минимума и точки максимума называют *точками экстремума*. Соответственно можно говорить о точках *локального экстремума, глобального экстремума, строгого экстремума*, ...

§1. Теорема Вейерштрасса

Важную роль в теории экстремума играют теоремы о достаточных условиях существования решения задач на экстремум (*теоремы существования экстремума*). Наиболее известной из таких теорем является классическая теорема Вейерштрасса[1].

1.1. Теорема Вейерштрасса. Так обычно называется утверждение о том, что функция, непрерывная на компакте в R^n, достигает на этом компакте своих точной нижней и точной верхней граней (или, иначе говоря, принимает на этом компакте минимальное и максимальное значения)[2]. Напомним, что компактом в пространстве R^n называется всякое ограниченное замкнутое множество. То есть теорема Вейерштрасса это следующая теорема существования решения задачи минимизации (1) (см., например, [Ф], п.173).

Теорема 1.1. *Пусть на множестве $X \subset R^n$, $X \neq \emptyset$, задана функция $f(\cdot) : X \to R$, причем выполнены следующие условия:*

1) функция $f(\cdot)$ непрерывна на X;

2) X замкнуто в R^n;

3) X ограничено в R^n.

Тогда множество X_* глобальных решений задачи (1) *непусто и компактно.*

Заметим, что все условия теоремы 1.1 существенны. Поясним сказанное на примере условия 3). Существенность этого условия означает, что вычеркнув его из теоремы 1.1, мы получим неверное утверждение. Существенность условия 3) видна из рассмотренного выше примера 0.1, в котором множество X и функция $f(\cdot)$ удовлетворяют условиям 1) и 2) теоремы 1.1, а условие 3) этой теоремы нарушается, при этом $X_* = \emptyset$.

Упражнение 1.1. Докажите, что каждое из условий 1) и 2) теоремы 1.1 является существенным.

1.2. Следствия теоремы Вейерштрасса. Приведем два полезных следствия теоремы Вейерштрасса, в которых множество X допустимых элементов задачи (1) уже не предполагается ограниченным.

Следствие 1.1. *Утверждение теоремы 1.1 останется справедливым, если заменить в нем условие* 3) *следующим условием:*

$3^{'}$) *существует число* $A \in R$ *такое, что множество*

$$X_A \equiv \{x \in X : f(x) \leq A\}$$

непусто и ограничено.

Упражнение 1.2. Докажите следствие 1.1. Воспользуйтесь тем, что множество глобальных решений задачи (1) совпадает с множеством глобальных решений задачи

$$f(x) \to \min, \qquad x \in X_A,$$

если X_A непусто.

Упражнение 1.3. Докажите, что каждое из условий следствия 1.1 является существенным.

Упражнение 1.4. Приведите пример задачи безусловной минимизации непрерывной функции, для которой нарушаются условия следствия 1.1, а $X_* \neq \emptyset$.

Следствие 1.2. *Утверждение теоремы 1.1 останется справедливым, если заменить в нем условие 3) следующим условием:*

$3^{''}$) $\lim\limits_{x \to \infty (x \in X)} f(x) = +\infty.$

Напомним, что $3^{''}$) означает: $\forall\ \gamma \in R$ найдется $\delta \in R_+$ такое, что при выполнении для $x \in X$ неравенства $\|x\|_n > \delta$ имеем $f(x) > \gamma$. На языке последовательностей условие $3^{''}$) можно сформулировать следующим образом: для любой последовательности $\{x_k\}_{k=1}^{k=\infty}$ элементов множества X, стремящейся к бесконечности, имеем $\lim\limits_{k \to \infty} f(x_k) = (+\infty)$.

Упражнение 1.5. Докажите следствие 1.2.

Упражнение 1.6. Докажите, что каждое из условий следствия 1.2 является существенным.

Упражнение 1.7. Приведите пример задачи безусловной минимизации непрерывной функции, для которой нарушаются условия следствия 1.2, а множество глобальных решений непусто.

Упражнение 1.8. Пусть $X \subset R^n$ - непустое замкнутое множество, y - фиксированная точка в R^n. Пусть $f(x) \equiv \|x-y\|_n$, $x \in X$. В этом случае множество X_* глобальных решений задачи (1) есть семейство ближайших к точке y точек множества X (такие точки множества X называются *проекциями точки y на множество* X). 1) Докажите, что X_* непусто (см., например, [В], С.76, следствие 1). 2) Докажите, что в данном случае условие замкнутости множества X существенно для справедливости вывода о непустоте X_*. 3) Приведите примеры таких: а) X и y, для которых X_* состоит из одной точки; б) X и y, для которых X_* состоит более чем из одной точки; в) X и y, для которых $X_* = X$; г) X и y, для которых задача (1) имеет локальные решения, не являющиеся глобальными. 4) Докажите, что если X - подпространство в R^n, то X_* состоит ровно из одной точки, которая является ортогональной проекцией точки y на подпространство X.

Замечание 1.1. В некоторых важных специальных случаях удается доказать теоремы существования и единственности решения экстремальных задач. См., например, следствие 2.4 в п.2.12.

§2. Безусловный экстремум

Рассмотрим задачу на безусловный минимум

$$f(x) \to \min, \quad x \in R^n. \tag{4}$$

Задача (4) называется *гладкой задачей на безусловный минимум*, если функция $f(\cdot) : R^n \to R$ обладает некоторыми свойствами дифференцируемости (например, она дифференцируема в точке минимума или в некоторой окрестности этой точки). Сформулируем для гладкой задачи (4) *необходимые условия экстремума* (то есть такие условия, которые обязательно выполняются в точке экстремума) и *достаточные условия экстремума* (то есть такие условия, при выполнении которых данная точка заведомо будет точкой экстремума). Заметим, что *порядком условия экстремума* (необходимого или достаточного) называют максимальный порядок производных, участвующих в формулировке этого условия.

2.1. Необходимое условие экстремума первого порядка. Следующая теорема (см., например, [З], С.462; [Ф], п.196), дающая для задачи (4) необходимые условия экстремума первого порядка, при $n = 1$ превращается в утверждение, известное нам из курса математического анализа как "теорема Ферма"[3] (см., например, [Ф], п.109).

Теорема 2.1. *Пусть* $\overline{x}$ *- локальное решение задачи* (4) *и функция* $f(\cdot)$ *имеет в точке* $\overline{x}$ *частные производные по каждой из переменных* $x^1, x^2, \ldots, x^n$. *Тогда*

$$(\bigtriangledown f(\overline{x}))^T = 0_n. \tag{5}$$

Точка $\overline{x}$, в которой выполняется условие (5), называется *стационарной точкой функции* $f(\cdot)$ *и задачи* (4), а само условие (5) - *условием стационарности* точки $\overline{x}$.

Упражнение 2.1. 1) Докажите теорему 2.1. Воспользуйтесь определением частной производной и тем, что если $\overline{x}$ - локальное решение задачи (4), то $\forall$ номера $i \in \{1, 2, \ldots, n\}$ число $\overline{x}^i$ - локальное решение одномерной задачи минимизации

$$f\left(\overline{x}^1, \ldots, \overline{x}^{i-1}, x^i, \overline{x}^{i+1}, \ldots, \overline{x}^n\right) \to \min, \quad x^i \in R.$$

2) Приведите пример, показывающий, что условие стационарности (5), вообще говоря, не является достаточным условием локального экстремума в задаче безусловной минимизации гладкой функции.

Замечание 2.1. В классе выпуклых функций, важном как для приложений, так и в теоретическом плане, условие стационарности (5) превращается из необходимого условия в критерий безусловного экстремума. Подробнее об этом см., например, [ЭВА].

Упражнение 2.2. Выведите из теоремы 2.1 ее аналог для случая задачи на безусловный максимум.

Упражнение 2.3. Ферма предложил следующий прием нахождения точек экстремума функций одной переменной (см. примечание [3]): "Чтобы отыскать максимум или минимум количества $f(x)$, надо составить выражение $f(x+h) - f(x)$, где h есть неопределенное число. Затем ... нужно разделить полученное выражение на это неопределенное h. Полагая в оставшихся членах $h = 0$, имеем уравнение, содержащее букву x, корнями которого являются точки максимумов и минимумов." (См. [АТФ], С.44-45; [Ш], гл.7, С.249.) Докажите, что если $f(x)$ - многочлен, то указанным приемом находятся все стационарные точки $f(x)$ и только они.

Чтобы сформулировать для задачи (4) условия экстремума второго порядка, нам потребуется понятие *определенной (определенной по знаку, знакоопределенной) матрицы.*

2.2. Определенные матрицы. Пусть A - симметричная действительная $(n \times n)-$матрица. Она называется *положительно определенной*, если

$$\langle Ah, h\rangle_n > 0 \quad \forall \quad h \in R^n, \quad h \neq 0_n.$$

Если выполняется более слабое условие

$$\langle Ah, h\rangle_n \geq 0 \quad \forall \quad h \in R^n,$$

то матрица A называется *неотрицательно определенной.* Факт положительной определенности (соотв. неотрицательной определенности) матрицы A будем обозначать неравенством $A > 0$ (соотв. $A \geq 0$). Сформулируем удобный критерий положительной определенности и неотрицательной определенности, использующий понятие главного минора матрицы. Минор матрицы получается пересечением некоторого количества ее строк и такого же количества ее столбцов. Если множество номеров этих строк совпадает с множеством номеров этих столбцов, то минор называется *главным минором* матрицы. Главный минор называется *ведущим главным* (или *угловым*) *минором*, если указанное множество номеров имеет вид $\{1, 2, \ldots, k\}$. Следующее утверждение называется критерием Сильвестра[4] (см. например, [Г1], гл.X, §4; [Б1], гл.5, §2).

Лемма 2.1.

$$\{A > 0\} \;\Leftrightarrow\; \{\textit{все ведущие главные миноры} \quad A \quad \textit{положительны}\},$$
$$\{A \geq 0\} \;\Leftrightarrow\; \{\textit{все главные миноры} \quad A \quad \textit{неотрицательны}\}.$$

В качестве примера рассмотрим матрицу $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$. У нее три главных минора, два из которых ведущие главные. Ведущие главные миноры равны 0, а главный минор, не являющийся ведущим, равен (-1). Таким образом, по лемме 2.1 данная матрица не является неотрицательно определенной. Этот пример показывает, что неотрицательность всех ведущих главных миноров еще не обеспечивает неотрицательной определенности матрицы.

Матрица A называется *отрицательно определенной* (соотв. *неположительно определенной*), если является положительно определенной (соотв. неотрицательно определенной) матрица $(-A)$. Матрица A называется *неопределенной* (*неопределенной по знаку* или *знаконеопределенной*), если она не является ни неотрицательно определенной, ни неположительно определенной.

Упражнение 2.4. 1) Выведите из леммы 2.1 ее аналог для отрицательно определенных и неположительно определенных матриц. 2) Используя понятие главного минора, сформулируйте и докажите критерий неопределенности симметричной действительной матрицы.

Упражнение 2.5. Сформулируйте условия на a, b, c, необходимые и достаточные для того, чтобы действительная матрица $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ была неотрицательно определенной (положительно определенной, отрицательно определенной, неположительно определенной, неопределенной).

В теории определенных матриц и квадратичных форм важную роль играет теорема о возможности приведения любой действительной симметричной матрицы к диагональной форме с помощью невырожденного преобразования. Сформулируем точный результат (см., например, [Б1], С.79). Напомним, что действительная $(m \times n)$ −матрица B называется *ортогональной матрицей*, если $B^T B = E$, где $E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ . & . & \ddots & . \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ - единичная $(n \times n)$ −матрица (это означает, что система столбцов матрицы B это ортонормированная система векторов в R^m); определитель ортогональной матрицы равен $+1$ или -1.

Лемма 2.2. *Для любой действительной симметричной $(n \times n)$ − матрицы A существует ортогональная $(n \times n)$ −матрица B такая, что*

$$B^T A B = \begin{pmatrix} \nu_1 & 0 & \cdots & 0 \\ 0 & \nu_2 & \cdots & 0 \\ . & . & \ddots & . \\ 0 & 0 & \cdots & \nu_n \end{pmatrix},$$

где $\nu_1, \nu_2, \ldots, \nu_n$ - собственные значения матрицы A. Это равносильно тому, что квадратичная форма $\langle Ax, x \rangle_n$, $x \in R^n$ заменой переменных $y = B^T x$, $y \in R^n$ приводится к каноническому виду

$$\langle Ax, x \rangle_n = \sum_{i=1}^{n} \nu_i \left(y^i\right)^2 .$$

Непосредственно из леммы 2.2 получаем следующий критерий положительной определенности и неотрицательной определенности матрицы.

Лемма 2.3. *Матрица A является положительно определенной (соотв. неотрицательно определенной) тогда и только тогда, когда все ее собственные значения положительны (соотв. неотрицательны).*

Упражнение 2.6. 1) Выведите из леммы 2.3 ее аналог для отрицательно определенных и неположительно определенных матриц. 2) Используя понятие собственного значения, сформулируйте и докажите критерий неопределенности симметричной действительной матрицы.

2.3. Необходимое условие экстремума второго порядка. Напомним (см., например, [ИСС], теоремы 13.13, 13.14): если функция $f(\cdot) : R^n \to R$ дважды дифференцируема в точке $\overline{x}$ (отсюда следует, что она один раз дифференцируема в некоторой окрестности $\overline{x}$), то для каждой пары чисел $i,j \in \{1,2,\ldots,n\}$ производные $\partial^2 f/\partial x^i \partial x^j$ и $\partial^2 f/\partial x^j \partial x^i$ в точке $\overline{x}$ совпадают. В этом случае матрица вторых производных $\bigtriangledown^2 f\left(\overline{x}\right)$ симметрична. Сформулируем необходимое условие минимума второго порядка в задаче (4).

Теорема 2.2. *Пусть $\overline{x}$ - локальное решение задачи* (4) *и функция $f\left(\cdot\right): R^n \to R$ дважды дифференцируема в точке $\overline{x}$. Тогда*

$$\bigtriangledown^2 f\left(\overline{x}\right) \geq 0. \tag{6}$$

Доказательство теоремы 2.2. Произвольно фиксируем некоторый ненулевой вектор $h \in R^n$. Функция $\varphi\left(\cdot\right): R \to R$, определенная формулой $\varphi\left(t\right) \equiv f\left(\overline{x}+th\right),\ t \in R$, дважды дифференцируема в нуле. Применяя к ней формулу Тейлора[5] второго порядка с центром в нуле и остаточным членом в форме Пеано[6] (см., например, [Ф], п.124), с учетом (5) получим

$$\begin{aligned} f\left(\overline{x}+th\right) &= f\left(\overline{x}\right) + 2^{-1}\cdot t^2 \cdot \left\langle \bigtriangledown^2 f\left(\overline{x}\right)h, h\right\rangle_n + r\left(t\right), \\ \text{где } r\left(t\right) &= o\left(t^2\right) \text{ при } \quad t \to 0. \end{aligned}$$

Так как для всех достаточно малых t, отличных от нуля, разность

$$\left(f\left(\overline{x}+th\right) - f\left(\overline{x}\right)\right)$$

неотрицательна, то для тех же t неотрицательно выражение

$$2^{-1}\cdot t^2 \cdot \left\langle \bigtriangledown^2 f\left(\overline{x}\right)h, h\right\rangle_n + r\left(t\right).$$

Деля это выражение на $\left(2^{-1}\cdot t^2\right)$ и переходя к пределу при $t \to 0$, получаем, что $\left\langle \bigtriangledown^2 f\left(\overline{x}\right)h, h\right\rangle_n \geq 0$. В силу произвольности h это и означает выполнение (6). Теорема 2.2 доказана.

Простые примеры показывают, что совместного выполнения условий (5) и (6), вообще говоря, не достаточно для того, чтобы точка $\overline{x}$ была точкой локального минимума в задаче безусловной минимизации дважды непрерывно дифференцируемой функции.

Упражнение 2.7. Приведите соответствующий пример.

Замечание 2.2. В то же время в классе часто встречающихся в приложениях квадратичных функций пара условий (5) и (6) дает критерий безусловного минимума, причем глобального. Мы рассмотрим этот вопрос подробнее в п.2.12.

Упражнение 2.8. 1) Выведите из теоремы 2.2 ее аналог для случая задачи на безусловный максимум. 2) Используя понятие неопределенной матрицы, сформулируйте достаточное условие отсутствия локального экстремума в стационарной точке дважды дифференцируемой в этой точке функции.

2.4. Достаточное условие экстремума второго порядка. Следующая теорема дает достаточное условие минимума второго порядка в задаче (4).

Теорема 2.3. *Пусть функция $f(\cdot)$ дважды дифференцируема в точке $\overline{x}$. Если выполняется условие (5) и, кроме того,*

$$\bigtriangledown^2 f(\overline{x}) > 0, \tag{7}$$

то $\overline{x}$ - строгое локальное решение задачи (4).

Доказательство теоремы 2.3. При указанных условиях можно в окрестности точки $\overline{x}$ записать для функции $f(\cdot)$ формулу Тейлора второго порядка (см., например, [ИСС], теорема 13.15):

$$f(x) = f(\overline{x}) + \langle \bigtriangledown f(\overline{x}), x - \overline{x} \rangle_n + 2^{-1} \left\langle \bigtriangledown^2 f(\overline{x})(x - \overline{x}), (x - \overline{x}) \right\rangle_n + r(x),$$

$$\text{где } r(x) = o\left(\|x - \overline{x}\|_n^2\right) \text{ при } \|x - \overline{x}\|_n \to 0.$$

Рассмотрим квадратичную форму

$$\Phi(h) \equiv \left\langle \bigtriangledown^2 f(\overline{x})\, h, h \right\rangle_n, \; h \in R^n.$$

По теореме Вейерштрасса форма $\Phi(h)$ достигает на единичной сфере $\{h \in R^n : \|h\|_n = 1\}$ минимального значения α. Число α положительно, противное означало бы нарушение (7). Следовательно, при $x \in R^n$, $x \neq \overline{x}$ имеем

$$\begin{aligned} \left\langle \bigtriangledown^2 f(\overline{x})(x - \overline{x}), (x - \overline{x}) \right\rangle_n &= (\|x - \overline{x}\|_n)^2 \cdot \Phi\left((x - \overline{x}) / \|x - \overline{x}\|_n\right) \geq \\ &\geq \alpha \cdot (\|x - \overline{x}\|_n)^2. \end{aligned}$$

Поэтому, в силу (5)

$$f(x) - f(\overline{x}) \geq 2^{-1} \cdot \alpha \cdot (\|x - \overline{x}\|_n)^2 + r(x), \quad x \in R^n, \quad x \neq \overline{x}.$$

Правая часть последнего неравенства строго положительна, если норма $\|x - \overline{x}\|_n$ достаточно мала. Значит, в некоторой окрестности точки $\overline{x}$ имеем $f(x) > f(\overline{x})$ при $x \neq \overline{x}$. Теорема 2.3 доказана.

Нетрудно привести пример, показывающий, что в точке безусловного строгого локального минимума дважды непрерывно дифференцируемой функции условие (7) может и не выполняться.

Упражнение 2.9. Приведите соответствующий пример.

Замечание 2.3. Однако, например, в классе квадратичных функций пара условий (5) и (7) дает критерий безусловного строгого глобального минимума, см. п.2.12.

Упражнение 2.10. Выведите из теоремы 2.3 ее аналог для случая задачи на безусловный максимум.

2.5. Теорема о локальной единственности экстремума. Точка локального минимума задачи (1) называется *локально единственной точкой минимума*, если в некоторой окрестности этой точки нет других точек локального минимума этой задачи.

Упражнение 2.11. Рассмотрите одномерную задачу безусловной минимизации для функции, приведенной в упражнении 0.3. Докажите, что точка $x = 0$ это точка строгого локального минимума, не являющаяся локально единственной.

Точка $\overline{x}$ локального минимума задачи (4) называется *невырожденной точкой минимума,* если в этой точке выполняется условие (7). Справедлива следующая теорема о единственности точки минимума (см., например, [А], п.2.7; [П], гл.1, §3, п.2).

Теорема 2.4. *Невырожденная точка минимума локально единственна.*

Доказательству теоремы 2.4 предпошлем следующее упражнение.

Упражнение 2.12. Пусть функция $f\left(\cdot\right) : R^n \to R$ дважды дифференцируема в точке $\overline{x}$. Докажите, что тогда

$$\left(\bigtriangledown f\left(x\right)\right)^T - \left(\bigtriangledown f\left(\overline{x}\right)\right)^T = \bigtriangledown^2 f\left(\overline{x}\right) \cdot \left(x - \overline{x}\right) + r\left(x\right), \quad x \in R^n,$$

где $r\left(x\right)$ - вектор-функция со значениями в R^n такая, что

$$\left\|r\left(x\right)\right\|_n = o\left(\left\|x - \overline{x}\right\|_n\right) \text{ при } \left\|x - \overline{x}\right\|_n \to 0.$$

Доказательство теоремы 2.4. (См. [А], п.2.7.) Допустим, что $\overline{x}$ не является локально единственной точкой минимума. Это означает, что существует сходящаяся к точке $\overline{x}$ последовательность $\{x_k\}_{k=1}^{\infty}$ точек локального минимума функции $f\left(\cdot\right)$. Можно считать, что $x_k \neq \overline{x}$, $k = 1, 2, \ldots$. Применим формулу из упражнения 2.12:

$$\begin{aligned} \left(\bigtriangledown f\left(x_k\right)\right)^T - \left(\bigtriangledown f\left(\overline{x}\right)\right)^T &= \bigtriangledown^2 f\left(\overline{x}\right) \cdot \left(x_k - \overline{x}\right) + r\left(x_k\right), \\ \left\|r\left(x_k\right)\right\|_n &= o\left(\left\|x_k - \overline{x}\right\|_n\right) \text{ при } k \to \infty. \end{aligned}$$

Так как $\left(\bigtriangledown f\left(\overline{x}\right)\right)^T = 0_n$ и $\left(\bigtriangledown f\left(x_k\right)\right)^T = 0_n$ при $k = 1, 2, \ldots$, то

$$\bigtriangledown^2 f\left(\overline{x}\right) \cdot \left(x_k - \overline{x}\right) + r\left(x_k\right) = 0_n \;, \; k = 1, 2, \ldots.$$

Положим

$$y_k \equiv \left(x_k - \overline{x}\right) / \left\|x_k - \overline{x}\right\|_n \quad \left(k = 1, 2, \ldots\right).$$

Последовательность $\{y_k\}_{k=1}^{\infty}$ лежит в компактном в R^n множестве - на сфере единичного радиуса с центром в 0_n. Поэтому эта последовательность имеет сходящуюся подпоследовательность. Без ограничения общности будем считать, что сама последовательность $\{y_k\}_{k=1}^{\infty}$ сходится к некоторому элементу y_0 указанной сферы. Так как

$$\bigtriangledown^2 f(\overline{x}) \cdot y_k = -r(x_k) / \|x_k - \overline{x}\|_n,$$

то в пределе, при $k \to \infty$, получаем

$$\bigtriangledown^2 f(\overline{x}) \cdot y_0 = 0_n.$$

Это противоречит условию (7), ведь $y_0 \neq 0_n \quad (\|y_0\|_n = 1)$. Значит наше предположение о том, что $\overline{x}$ не является локально единственной точкой минимума, неверно. Теорема 2.4 доказана.

2.6. Об условиях экстремума высокого порядка. Используя формулу Тейлора аналогично тому, как это сделано в п.п. 2.1, 2.3, 2.4, можно для задачи (4) вывести некоторые необходимые, а также достаточные условия экстремума третьего и более старших порядков. При $n = 1$ это делается в курсе математического анализа. Справедливо следующее утверждение (см., например, [З], гл. V, §4, п.2; [Ф], п.138).

Теорема 2.5. *Пусть $n = 1$ и функция $f(\cdot) : R \to R$ дифференцируема k раз в точке $\overline{x}$, $k > 1$, причем*

$$f'(\overline{x}) = 0, f''(\overline{x}) = 0, \ldots, f^{(k-1)}(\overline{x}) = 0.$$

Тогда:

1) если $\overline{x}$ - локальное решение задачи (4), *то $f^{(k)}(\overline{x}) = 0$ при нечетном k и $f^{(k)}(\overline{x}) \geq 0$ при четном k;*

2) *если k четно и $f^{(k)}(\overline{x}) > 0$, то $\overline{x}$ - строгое локальное решение задачи* (4).

Упражнение 2.13. Докажите теорему 2.5, воспользовавшись тем, что при указанных в теореме условиях справедлива следующая формула Тейлора $k-$го порядка с остаточным членом в форме Пеано (см., например, [Ф], п.124)

$$f(\overline{x} + h) = \sum_{i=0}^{k} (i!)^{-1} f^{(i)}(\overline{x}) h^i + o\left(h^k\right) \text{ при } h \to 0.$$

Однако, при $n > 1$ условия экстремума старших порядков (выше второго порядка) для задачи (4) довольно громоздки, в частности, из-за того, что участвующие в их формулировках матрицы старших производных многомерны. Так, например, матрица третьих производных

$f'''_{x^ix^jx^k}(x)$, $1 \leq i,j,k \leq n$, - трехмерная, кубическая $(n \times n \times n)$ $-$матрица. Мы ограничимся здесь формулировкой следующего необходимого условия третьего порядка, с подробным доказательством которого можно познакомиться, прочитав в [AT] задачи 2.7, 2.8 и их решения.

Теорема 2.6. *Пусть функция $f(\cdot): R^n \to R$ трижды дифференцируема в некоторой окрестности точки $\overline{x}$, причем все ее третьи производные непрерывны в самой точке $\overline{x}$. Если $\overline{x}$ - локальное решение задачи* (4)*, то кубическая форма n переменных*

$$\nabla^3 f(\overline{x})[h] \equiv \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} f'''_{x^ix^jx^k}(\overline{x})\, h^i h^j h^k$$

обращается в нуль при всех $h \equiv \left(h^1, h^2, \ldots, h^n\right)^T \in R^n$, при которых обращается в нуль квадратичная форма

$$\nabla^2 f(\overline{x})[h] \equiv \sum_{i=1}^{n}\sum_{j=1}^{n} f''_{x^ix^j}(\overline{x})\, h^i h^j.$$

Упражнение 2.14. ([AT], задача 2.12.) Найдите все точки локальных экстремумов функции двух переменных

$$f(x) \equiv \left(x^1 + x^2\right)\cdot\left(x^1 - a\right)\cdot\left(x^2 - b\right), \quad x = \left(x^1, x^2\right)^T \in R^2$$

($a, b \in R$ - параметры задачи).

2.7. Решение задач. Результаты, приведенные выше в §1 и §2 , позволяют предложить, например, следующий план действий при решении задач вида (4) с дифференцируемой в R^n функцией $f(\cdot)$.

1) Выписываем условие (5) и находим стационарные точки минимизируемой функции $f(\cdot)$. Если стационарных точек нет, то в задаче (4) нет точек минимума. Можно попробовать оценить значение задачи.

2) Изучаем вопрос существования глобального решения задачи (4) (например, пытаемся применить следствия теоремы Вейерштрасса). Если глобальное решение существует, то, сравнивая значения функции $f(\cdot)$ в стационарных точках, находим все глобальные решения задачи. (Если функция $f(\cdot)$ дважды дифференцируема в R^n, то отыскание точек глобального минимума можно отложить до того момента, когда после применения условий экстремума 2-го порядка некоторые стационарные точки будут исключены из числа точек, "подозрительных на точки минимума", см. п.4).)

3) Если функция $f(\cdot)$ дважды дифференцируема в R^n, то вычисляем ее матрицу вторых производных в каждой стационарной точке. Исследуя эту матрицу на знакоопределенность, проверяем выполнение в стационарной точке достаточных условий минимума 2-го порядка (теорема 2.3). Если эти условия выполняются, то данная стационарная точка является строгим локальным решением задачи. Если достаточные условия экстремума 2-го порядка не выполняются, то проверяем в данной стационарной точке выполнение необходимых условий экстремума 2-го порядка (теорема 2.2). Если и они не выполняются, то стационарная точка не является решением задачи (4). Если же необходимые условия 2-го порядка выполняются, то данная точка остается "подозрительной на точку локального минимума" и требует дополнительного исследования.

4) Если глобальные решения задачи (4) существуют и ранее не найдены, то, сравнивая значения функции $f(\cdot)$ в найденных точках локального минимума и в оставшихся "подозрительными на точки минимума" стационарных точках, находим все глобальные решения.

5) Продолжаем исследование на локальный минимум стационарных точек, удовлетворяющих необходимым условиям минимума 2-го порядка, но не удовлетворяющих достаточным условиям минимума 2-го порядка. Такое исследование можно проводить как с помощью условий высоких порядков, так и непосредственно изучая поведение функциии $f(\cdot)$ в окрестности "подозрительной точки".

Замечание 2.4. Приведенный план исследования функции на экстремум, конечно, является условным. В нем не используются никакие другие свойства функции $f(\cdot)$ кроме свойств дифференцируемости. При наличии дополнительной информации этот план должен быть скорректирован. См., например, замечания 2.1 - 2.3, а также п.2.12.

Замечание 2.5. Если функция $f(\cdot) : R^n \to R$ не является дифференцируемой в некоторых точках пространства R^n, то ситуация усложняется. Такие точки, вместе со стационарными точками функции $f(\cdot)$, должны быть включены в число "подозрительных на точки минимума". Их необходимо подвергнуть отдельному исследованию на локальный минимум и включить в число "подозрительных на глобальный минимум".

Замечание 2.6. Решать задачу на безусловный максимум можно, переходя к эквивалентной задаче на безусловный минимум.

2.8. Пример: метод наименьших квадратов. Одним из широко применяемых методов оценки неизвестных величин по результатам измерений является так называемый *метод наименьших квадратов*, предложенный Гауссом и Лежандром [7]. Опишем простейший вариант метода (см., например, [Ф], С.438-440 и [АТ], С.30-31; подробнее см. [МЭ], С.875-

882; [K2]; [C2]).

Пример 2.1. Пусть n неизвестных величин $x^1, x^2, \dots, x^n$ связаны m линейными соотношениями

$$\sum_{j=1}^{n} a_{ij} x^j = b^i, \quad i = 1, 2, \dots, m, \tag{8}$$

где $m > n$, a_{ij} - точно известные коэффициенты, b^i - приближенно известные величины, полученные из измерений; все соотношения (8) равноправны. Считаем, что ранг $(m \times n)$ −матрицы

$$(a_{ij}) \equiv \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

максимален, то есть равен n. Система линейных уравнений (8), вообще говоря, несовместна и может оказаться, что невязки

$$\Delta_i = \Delta_i(x) \equiv \sum_{j=1}^{n} a_{ij} x^j - b^i, \quad x \equiv \left(x^1, x^2, \dots, x^n\right)^T \in R^n, \quad i = 1, 2, \dots, m$$

этой системы не обращаются одновременно в нуль ни для какого вектора $x \in R^n$. Согласно методу наименьших квадратов *наилучшими, наиболее согласующимися с результатами измерений (наиболее вероятными) считаются значения* $x^1, x^2, \dots, x^n$, *минимизирующие сумму квадратов всех невязок*, то есть доставляющие минимум в задаче (4) при

$$f(x) \equiv \sum_{i=1}^{m} \{\Delta_i(x)\}^2, \quad x \in R^n. \tag{9}$$

Сформулированное правило метода наименьших квадратов вытекает из закона Гаусса для распределения вероятностей случайных погрешностей. Подробнее об этом см., например, [K2].

В данном случае условие (5) может быть записано в виде

$$\sum_{j=1}^{n} x^j \langle a_i, a_j \rangle_m = \langle a_i, b \rangle_m, \quad i = 1, 2, \dots, n, \tag{10}$$

где $b \equiv \left(b^1, b^2, \dots, b^m\right)^T$, $a_i \equiv (a_{1i}, a_{2i}, \dots, a_{mi})^T$ - i−ый столбец матрицы (a_{ij}). Матрица K коэффициентов линейной системы (10) является, очевидно, произведением $(n \times m)$ $-$матрицы $(a_{ij})^T$ на $(m \times n)$ $-$матрицу (a_{ij}) :

$$K \equiv \begin{pmatrix} \langle a_1, a_1 \rangle_m & \langle a_1, a_2 \rangle_m & \cdots & \langle a_1, a_n \rangle_m \\ \langle a_2, a_1 \rangle_m & \langle a_2, a_2 \rangle_m & \cdots & \langle a_2, a_n \rangle_m \\ \cdots & \cdots & \cdots & \cdots \\ \langle a_n, a_1 \rangle_m & \langle a_n, a_2 \rangle_m & \cdots & \langle a_n, a_n \rangle_m \end{pmatrix} =$$

$$= \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix} \cdot \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \equiv (a_{ij})^T \cdot (a_{ij}) \,.$$

Воспользуемся формулой Бине - Коши[8] для определителя квадратной матрицы, равной произведению двух прямоугольных ([Г1], С.20). По этой формуле определитель матрицы K равен сумме произведений всевозможных миноров максимального порядка матрицы $(a_{ij})^T$ на соответствующие миноры матрицы (a_{ij}). Так как $(a_{ij})^T = (a_{ij})$, то определитель матрицы K равен сумме квадратов всевозможных миноров $n-$го порядка матрицы (a_{ij}). Так как ранг (a_{ij}) равен n, то среди этих миноров есть отличный от нуля и потому определитель матрицы K не равен нулю. Таким образом, система (10) имеет единственное решение $\overline{x} \equiv \left(\overline{x}^1, \overline{x}^2, \ldots, \overline{x}^n\right)^T$. Точка $\overline{x}$ - единственная стационарная точка функции (9). Это означает, что рассматриваемая задача (4) не может иметь более одного решения, а если она имеет решение (локальное или глобальное), то это решение и есть точка $\overline{x}$.

На самом деле у этой задачи существует глобальное решение. Это вытекает из следствия 1.2, так как для функции (9)

$$\lim_{x \to \infty} f(x) = +\infty \tag{11}$$

(в данном случае $X = R^n$). Для доказательства (11) возьмем некоторый минор $n-$го порядка матрицы (a_{ij}), отличный от нуля. Пусть $i_1, i_2, ..., i_n$ - номера строк матрицы (a_{ij}), из которых составлен этот минор, расположенные по возрастанию. Формулы

$$\Delta_i = \sum_{j=1}^{n} a_{ij} x^j - b^i, \quad i = i_1, i_2, \ldots, i_n$$

устанавливают взаимно-однозначное соответствие между пространством R^n векторов $x \equiv \left(x^1, x^2, \ldots, x^n\right)^T$ (обозначим его $R^n(x)$) и пространством R^n векторов $\Delta \equiv (\Delta_{i_1}, \Delta_{i_2}, \ldots, \Delta_{i_n})^T$ (обозначим это пространство $R^n(\Delta)$). При этом соответствии каждое ограниченное множество пространства $R^n(x)$ переходит в ограниченное множество пространства $R^n(\Delta)$ и наоборот. Произвольно фиксируем $\varepsilon > 0$. При указанном соответствии между векторами Δ и x шар $U_\varepsilon(0_n)$ пространства $R^n(\Delta)$ переходит в некоторое ограниченное множество P пространства $R^n(x)$. Возьмем в пространстве $R^n(x)$ любой шар $U_\delta(0_n)$, содержащий множество P. Если вектор x лежит вне $U_\delta(0_n)$, то соответствующий ему вектор Δ находится вне шара $U_\varepsilon(0_n)$ и, следовательно, $f(x) > \varepsilon^2$. Предельное соотношение (11) доказано.

Таким образом, метод наименьших квадратов рекомендует взять в качестве "решения" системы (8) (которая, напомним, может оказаться и несовместной) вектор $\overline{x}$, являющийся решением системы (10). Такой вектор, причем только один, всегда существует. Если система (8) совместна, то $\overline{x}$ - ее точное решение.

Упражнение 2.14. Проверьте, что для функции (9) условие (5) может быть записано в виде (10).

Замечание 2.7. Рассмотренная в примере 2.1 задача о нахождении неизвестных величин $x^1, x^2, \ldots, x^n$, удовлетворяющих известным приближенно соотношениям (8), допускает следующую интерпретацию (см., например, [АТ], С.30-31). Проводится серия опытов, в ходе каждого из которых задается некоторый набор параметров $y^1, y^2, \ldots, y^n$ (входные параметры) и наблюдается некоторый параметр z (выходной параметр). Точная зависимость между входными параметрами $y^1, y^2, \ldots, y^n$ и выходным параметром z неизвестна, но предполагается, что она линейна:

$$z = \sum_{j=1}^{n} x^j \cdot y^j; \tag{12}$$

коэффициенты $x^1, x^2, \ldots, x^n$ линейной функции (12) нужно найти. Пусть проведена серия из m опытов, которым присвоены номера $1, 2, \ldots, m$, причем в i—ом опыте получены данные:

$$y^1 = a_{i1}, \quad y^2 = a_{i2}, \ldots, \quad y^n = a_{in}, \quad z = b^i,$$

$i = 1, 2, \ldots, m$. Это означает, что неизвестные коэффициенты $x^1, x^2, \ldots$ $\ldots, x^n$ подчиняются соотношениям (8). В качестве наилучшего набора этих коэффициентов метод наименьших квадратов предлагает брать решение задачи на безусловный минимум функции (9).

Упражнение 2.15. 1) Для оценки некоторой неизвестной скалярной величины (константы) x проведено m независимых измерений при одинаковых условиях и получены данные:

$$x = b^i, \quad i = 1, 2, \ldots, m.$$

Таким образом, получена система (8), в которой $n = 1$, $x^1 = x \in R$, $a_{i1} = 1$ $(i = 1, 2, \ldots, m)$. Покажите, что по методу наименьших квадратов наилучшим приближением к искомой величине x (наиболее вероятным значением этой величины) является среднее арифметическое $m^{-1}(b^1 + b^2 + \ldots + b^m)$ данных измерения.

2) Для опытных данных каждой из следующих таблиц найдите наилучшую в смысле метода наименьших квадратов зависимость вида (12):

а) $n = 1$, $m = 2$,

Номер опыта	y^1	z
1	1	2
2	2	3,5

;

б) $n = 1$, $m = 3$,

Номер опыта	y^1	z
1	1	2
2	2	3,5
3	3	6,5

;

в) $n = 2$, $m = 3$,

Номер опыта	y^1	y^2	z
1	1	1	3
2	1	0	0,5
3	0	1	2,5

;

г) $n = 2$, $m = 4$,

Номер опыта	y^1	y^2	z
1	1	1	3
2	1	0	0,5
3	0	1	2,5
4	-1	1	1,5

.

2.9. Пример: гладкая функция двух переменных, имеющая единственную точку локального экстремума и не имеющая точек глобального экстремума. Пример функции двух переменных с указанными в заголовке этого пункта свойствами представляется интересным, так как интуитивно ясно, что функция одного действительного переменного подобными свойствами обладать не может. А именно: не существует определенной на действительной оси гладкой функции, имеющей единственную точку безусловного локального экстремума, не являющуюся точкой безусловного глобального экстремума (аккуратно доказать этот факт предоставляется читателю, см. ниже упражнение 2.16, а)). Перейдем к объявленному в заголовке этого пункта примеру.

Пример 2.2. Пусть $n = 2$. Рассмотрим для функции

$$f(x) \equiv \left(x^1\right)^3 - 3\left(x^1\right)^2 - \left(x^2\right)^2, \quad x \in R^2 \tag{13}$$

задачу (4) на безусловный минимум. В данном случае

$$\nabla f(x) = \left(3\left(x^1\right)^2 - 6x^1 \quad , \quad -2x^2\right), \quad x \in R^2$$

и функция $f(\cdot)$ имеет две стационарные точки: $x_1 = (0, 0)^T$, $x_2 = (2, 0)^T$.

Вычислим матрицу Гессе[9]

$$\bigtriangledown^2 f(x) = \begin{pmatrix} 6x^1 - 6 & 0 \\ 0 & -2 \end{pmatrix}$$

в стационарных точках: $\bigtriangledown^2 f(x_1) = \begin{pmatrix} -6 & 0 \\ 0 & -2 \end{pmatrix}$, $\bigtriangledown^2 f(x_2) = \begin{pmatrix} 6 & 0 \\ 0 & -2 \end{pmatrix}$. По критерию Сильвестра ни в одной из этих точек матрица Гессе не является неотрицательно определенной матрицей. Поэтому по теореме 2.2 у функции (13) нет точек минимума. Значение рассматриваемой задачи (4) равно $(-\infty)$, так как $f(x) = -\left(x^2\right)^2$ при $x^1 = 0$ и $f(x) \to (-\infty)$ при $x^2 \to \infty$, $x^1 = 0$.

Рассмотрим теперь для той же функции (13) задачу на безусловный максимум

$$f(x) \to \max, \quad x \in R^2. \tag{14}$$

Она эквивалентна задаче на минимум

$$(-f(x)) \to \min, \quad x \in R^2. \tag{15}$$

Стационарные точки в задаче (15) те же, x_1, x_2, x_3. Вычисляем матрицу Гессе минимизируемой функции в этих точках: $\left(-\bigtriangledown^2 f(x_1)\right) = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}$, $\left(-\bigtriangledown^2 f(x_2)\right) = \begin{pmatrix} -6 & 0 \\ 0 & 2 \end{pmatrix}$. В точке x_1 она является положительно определенной, в точке x_2 - знаконеопределенной. Значит, по теореме 2.2 точка x_2 не является точкой локального минимума в задаче (15), а, следовательно, не является точкой локального максимума в задаче (14). По теореме 2.3 точка x_1 дает строгий локальный минимум в задаче (15) и, следовательно, дает строгий локальный максимум в задаче (14). Однако, точка x_1 не является точкой глобального максимума в задаче (14). Действительно, значение этой задачи равно $(+\infty)$, так как при $x^2 = 0$ функция $f(x) = \left(\left(x^1\right)^3 - 3\left(x^1\right)^2\right)$ и $f(x) \to (+\infty)$, если $x^1 \to (+\infty)$, $x^2 = 0$.

Таким образом, рассматриваемая в данном примере *гладкая функция двух переменных* (13) *имеет единственную точку безусловного локального экстремума, не являющуюся точкой безусловного глобального экстремума.* То, что, как указано выше, функция одной переменной не может обладать подобными свойствами, непосредственно вытекает из приведенного в следующем ниже упражнении 2.16 свойства а) непрерывной функции $f(\cdot) : R \to R$.

Упражнение 2.16. Пусть $f(\cdot) : R \to R$ - непрерывная функция. Используя теорему Вейерштрасса, покажите, что:

а) если $f(\cdot)$ имеет единственную точку локального экстремума, то эта точка является и точкой глобального экстремума $f(\cdot)$;

б) между любыми двумя точками локального минимума (соотв. максимума) функции $f(\cdot)$ лежит по крайней мере одна точка локального максимума (соотв. минимума) этой функции (см. [AT], задача 2.28).

2.10. Пример: гладкая функция двух переменных, обладающая счетным числом точек строгого локального максимума, у которой нет точек локального минимума. Из приведенного в упражнении 2.16 свойства б) непрерывной функции одного действительного переменного $f(\cdot): R \to R$ следует, что если такая функция имеет $k > 1$ точек локального максимума (соотв. минимума), то точек локального минимума (соотв. максимума) у нее не менее, чем $(k-1)$; если же точек локального максимума (соотв. минимума) бесконечное число, то и точек локального минимума (соотв. максимума) также бесконечное число. Функция двух переменных $f(\cdot): R^2 \to R$ может уже не обладать такими свойствами. Приведем простой пример.

Пример 2.3 ([ГТ], п.1.4, пример 12). Пусть

$$f(x) \equiv \sin\left(x^2\right) - \left(x^1\right)^2, \quad x \in R^2. \tag{16}$$

Функция $f(\cdot)$, а вместе с ней и функция $(-f(\cdot))$ имеет счетное число стационарных точек:

$$x_m = (0 \quad , \quad (\pi/2) + m \cdot \pi)^T, \quad m = 0, \pm 1, \pm 2, \ldots .$$

Матрица Гессе $\bigtriangledown^2 f(x) = \begin{pmatrix} -2 & 0 \\ 0 & -\sin\left(x^2\right) \end{pmatrix}, \quad x \in R^2$ не является неотрицательно определенной матрицей ни в одной из стационарных точек. Поэтому функция (16) не имеет точек локального минимума. В то же время матрица Гессе $\bigtriangledown^2(-f(x)) = -\bigtriangledown^2 f(x)$ является положительно определенной в точках x_m при $m = 0, \pm 2, \pm 4, \ldots$. Поэтому эти точки являются точками строгого локального минимума функции $(-f(\cdot))$ и, следовательно, точками строгого локального максимума функции $f(\cdot)$. Таким образом, *функция* (16) *обладает счетным числом точек строгого локального максимума, но у нее нет ни одной точки локального минимума.* Все точки локального максимума функции $f(\cdot)$ являются ее точками глобального максимума. Значение задачи безусловной минимизации (4) для функции (16) равно $(-\infty)$.

2.11. Пример: задача Штейнера. Рассмотрим простейший вариант так называемой задачи Штейнера[10] (см. [П], С.333-334; [Т], С.34-36; [Ф], С.434-435).

Пример 2.4. В плоскости требуется найти точку, сумма расстояний которой до трех заданных точек x_1, x_2, x_3 этой плоскости, не лежащих на одной прямой, минимальна. Это означает, что нужно найти глобальное решение задачи на безусловный минимум

$$f(x) \equiv \sum_{i=1}^{3} \|x - x_i\|_2 \to \min, \quad x \in R^2. \tag{17}$$

Задача (17) имеет глобальное решение по следствию 1.2 (здесь $X = R^2$). Функция (17) дифференцируема везде в R^2 за исключением точек x_1, x_2, x_3, при этом

$$\nabla f(x) = \sum_{i=1}^{3} \left((x - x_i) / \|x - x_i\|_2\right), \quad x \in R^2, \quad x \neq x_i \quad (i = 1, 2, 3).$$

Условие (5) имеет вид

$$\sum_{i=1}^{3} \left((x - x_i) / \|x - x_i\|_2\right) = 0_2, \quad x \neq x_i \quad (i = 1, 2, 3)$$

и выполняется в том и только в том случае, когда единичные векторы $\left((x - x_i) / \|x - x_i\|_2\right) \quad (i = 1, 2, 3)$ образуют между собой равные углы, то есть все стороны треугольника $\triangle x_1x_2x_3$ видны из точки x под одинаковыми углами:

$$\angle x_1xx_2 = \angle x_2xx_3 = \angle x_1xx_3 = 2\pi/3 \tag{18}$$

(мы не различаем углы $\angle x_ixx_j$ и $\angle x_jxx_i$). Точка $x \in R^2$, удовлетворяющая условию (18), называется *точкой Торричелли*[11] треугольника $\triangle x_1x_2x_3$. Точка Торричелли треугольника $\triangle x_1x_2x_3$, обозначим ее x_T, существует тогда и только тогда, когда все углы этого треугольника меньше чем $2\pi/3$, при этом такая точка единственна и лежит строго внутри треугольника $\triangle x_1x_2x_3$ (см. ниже упражнение 2.17).

Таким образом, единственной стационарной точкой задачи (17) является, если она существует, точка Торричелли x_T треугольника $\triangle x_1x_2x_3$. Может быть одно из двух: либо треугольник $\triangle x_1x_2x_3$ имеет тупой угол и этот тупой угол больше или равен $2\pi/3$, либо все углы треугольника $\triangle x_1x_2x_3$ строго меньше $2\pi/3$. В первом случае функция (17) не имеет стационарных точек и, следовательно, глобальным решением задачи (17) является одна из точек x_1, x_2, x_3 (эти точки остаются подозрительными на точки минимума, так как функция (17) в них недифференцируема).

Сравнивая значения $f(x_1)$, $f(x_2)$, $f(x_3)$, находим, что глобальным решением задачи (17) является вершина тупого угла треугольника $\triangle x_1x_2x_3$ (величина $f(x_i)$ равна сумме длин сторон треугольника $\triangle x_1x_2x_3$, выходящих из вершины x_i; известно, что в треугольнике напротив бо́льшего угла лежит бо́льшая сторона). Во втором случае подозрительными на точки минимума являются точки x_1, x_2, x_3, x_T (x_T - единственная стационарная точка функции (17), а x_1, x_2, x_3 - точки, в которых эта функция недифференцируема). Применим к треугольнику $\Delta x_1x_Tx_2$, в котором $\angle x_1x_Tx_2 = 2\pi/3$, "теорему косинусов":

$$\|x_1 - x_2\|_2^2 = \|x_T - x_1\|_2^2 + \|x_T - x_2\|_2^2 + \|x_T - x_1\|_2 \cdot \|x_T - x_2\|_2 .$$

Отсюда $\|x_1 - x_2\|_2^2 > ((1/2)\|x_T - x_1\|_2 + \|x_T - x_2\|_2)^2$ и, следовательно,

$$\|x_1 - x_2\|_2 > (1/2)\|x_T - x_1\|_2 + \|x_T - x_2\|_2 .$$

Аналогично

$$\|x_1 - x_3\|_2 > (1/2)\|x_T - x_1\|_2 + \|x_T - x_3\|_2 .$$

Поэтому

$$\begin{gathered} f(x_1) = \|x_1 - x_2\|_2 + \|x_1 - x_3\|_2 > \\ > \|x_T - x_1\|_2 + \|x_T - x_2\|_2 + \|x_T - x_3\|_2 = f(x_T) . \end{gathered}$$

Точно так же $f(x_2) > f(x_T)$ и $f(x_3) > f(x_T)$. Следовательно, в рассматриваемом случае единственной точкой глобального минимума является точка x_T.

Упражнение 2.17. Докажите, что точка Торричелли существует у треугольника тогда и только тогда, когда все углы треугольника меньше чем $2\pi/3$, при этом такая точка единственна и лежит строго внутри треугольника (см. [T], С.35).

2.12. Квадратичные функции. Существуют важные для приложений классы функций, в которых некоторые из сформулированных выше общих необходимых условий экстремума превращаются в критерии экстремума. К таким классам функций относятся, например, класс выпуклых функций и класс квадратичных функций. Остановимся на классе квадратичных функций (о выпуклых функциях см., например, [ЭВА]). Квадратичной называется функция вида

$$f(x) = \sum_{i,j=1}^{n} a_{ij}x^i x^j - \sum_{i=1}^{n} b^i x^i + c, \quad x \in R^n, \tag{19}$$

где a_{ij} $(i,j=1,2,\ldots,n)$, b^i $(i=1,2,\ldots,n)$, c - действительные коэффициенты. Заметим, что именно квадратичной является, например, функция (9), рассматриваемая в методе наименьших квадратов.

Рассмотрим задачу (4) с квадратичной функцией (19). Покажем, что если в некоторой точке $\overline{x}\in R^n$ для этой функции выполняются условия (5) и (6), то точка $\overline{x}$ является глобальным решением задачи (4). Действительно, для квадратичной функции $f(x)$ формула Тейлора второго порядка с центром в точке $\overline{x}$ имеет вид

$$f(x)=f(\overline{x})+\langle\bigtriangledown f(\overline{x}),x-\overline{x}\rangle_n+2^{-1}\left\langle\bigtriangledown^2 f(\overline{x})(x-\overline{x}),(x-\overline{x})\right\rangle_n, x\in R^n,$$

что вместе с формулой (5) дает

$$f(x)-f(\overline{x})=2^{-1}\left\langle\bigtriangledown^2 f(\overline{x})(x-\overline{x}),(x-\overline{x})\right\rangle_n,\qquad x\in R^n.$$

Отсюда, в силу (6), имеем $\forall\, x\in R^n$ неравенство $f(x)-f(\overline{x})\geq 0$, означающее, что $\overline{x}$ - точка глобального минимума в задаче (4). Таким образом, из теорем 2.1 и 2.2 получаем следующие утверждения.

Следствие 2.1. *Для квадратичной функции* $f(\cdot)$ *всякое локальное решение задачи* (4) *является глобальным решением этой задачи.*

Поэтому в случае квадратичной функции $f(\cdot)$ будем говорить не "глобальное решение" или "локальное решение", а просто "решение" задачи (4).

Следствие 2.2. *Для квадратичной функции* $f(\cdot)$ *точка* $\overline{x}$ *является решением задачи* (4) *тогда и только тогда, когда выполняются условия* (5) *и* (6).

С учетом того, что матрица Гессе $\bigtriangledown^2 f(x)$ квадратичной функции $f(\cdot)$ является постоянной матрицей, не зависящей от x, удобно записать критерий минимума из следствия 2.2 несколько иначе. Не ограничивая общности, можно считать, что квадратичная функция имеет вид

$$f(x)=2^{-1}\langle Ax,x\rangle_n-\langle b,x\rangle_n+c,\quad x\in R^n,\tag{20}$$

где A - симметричная действительная $(n\times n)$ $-$матрица; к форме (20) представления квадратичной функции от формы (19) легко перейти, воспользовавшись очевидным тождеством

$$\sum_{i,j=1}^{n}a_{ij}x^ix^j\equiv 2^{-1}\sum_{i,j=1}^{n}(a_{ij}+a_{ji})\,x^ix^j,\quad x\in R^n.$$

В случае представления квадратичной функции в форме (20) имеем:

$$\bigtriangledown f(x)=Ax-b,\quad \bigtriangledown^2 f(x)=A,\quad x\in R^n.$$

Упражнение 2.18. Докажите указанные формулы для градиента и матрицы Гессе квадратичной функции (20).

Для квадратичной функции (20) условие (5) принимает вид

$$A\overline{x} = b, \tag{21}$$

а условие (6) - вид

$$A \geq 0. \tag{22}$$

Критерий минимума из следствия 2.2 может быть переписан следующим образом.

Следствие 2.3. *Точка* $\overline{x}$ *является решением задачи* (4) *для квадратичной функции* (20) *тогда и только тогда, когда выполняются условия* (21) *и* (22).

Форма (20) представления квадратичной функции позволяет легко доказать следующую теорему существования решения задачи (4) (см. [AT], теорема 1.3).

Лемма 2.4. *Если задача* (4) *для квадратичной функции* $f(\cdot)$ *имеет конечное значение, то эта задача разрешима. Отсюда следует, что если решения не существует, то значение задачи равно* $(-\infty)$.

Упражнение 2.19. Докажите лемму 2.4 с помощью замены переменных $x = By$ из леммы 2.2. (См., например, [AT], задача 1.21).

Упражнение 2.20. 1) Приведите пример квадратичной функции вида (20), для которой задача (4) не имеет решения, причем условие (22) выполняется. 2) Приведите пример квадратичной функции, у которой есть стационарная точка, не являющаяся точкой безусловного минимума этой функции.

Упражнение 2.21. Покажите, что в классе квадратичных функций критерием точки безусловного строгого локального (и глобального) минимума является совместное выполнение условий (5) и (7).

Условие (7) для формы (20) представления квадратичной функции принимает вид

$$A > 0. \tag{23}$$

При выполнении условия (23) матрица A невырождена, так как по лемме 2.1 ее определитель больше нуля. Следовательно, существует единственная точка $\overline{x}$, удовлетворяющая равенству (21). Таким образом, если в задаче (4) функция $f(\cdot)$ квадратична и, представив эту функцию в форме (20), мы нашли, что выполняется условие (23), то можно утверждать, что задача (4) имеет единственное решение и это решение $\overline{x}$ находится из (21). Нами доказана следующая теорема существования и единственности решения задачи (4).

Следствие 2.4. *Задача* (4) *для квадратичной функции* (20) *при выполнении условия* (23) *имеет единственное решение.*

Замечание 2.8. Более общее утверждение см. в [ЭВА], §2, п.10.

Упражнение 2.22. Выделите в пространстве параметров a, b, c, k, m множество тех параметров, при которых квадратичная функция

$$f(x) = a\left(x^1\right)^2 + 2bx^1x^2 + c\left(x^2\right)^2 + kx^1 + mx^2, \quad x \in R^2$$

имеет: 1) точку безусловного глобального минимума; 2) точку безусловного строгого глобального минимума.

Примечания

[1] (С.9) *Вейерштрасс* (Weierstraβ) Карл Теодор Вильгельм (1815 - 1897) - немецкий математик, с 1856 г. проф. Берлинского ун-та. Продолжил начатую Больцано и Коши (см. примечание [8]) "реформу наведения строгости" в математическом анализе. Создал хорошо известный всем изучающим математический анализ "язык ε - δ". Этот язык позволил дать определениям основных понятий анализа численное представление и существенно уменьшить в них роль интуитивных соображений. Тем самым Вейерштрасс придал математическому анализу форму, очень близкую к современной. В частности, им было дано строгое доказательство основных свойств функции, непрерывной на отрезке, введено в общее употребление понятие равномерной сходимости функционального ряда (признак Вейерштрасса). Сильное впечатление на современников произвела его теорема о том, что любая непрерывная на отрезке функция является пределом равномерно сходящейся последовательности многочленов. Вейерштрасс - один из основных создателей теории аналитических функций, основоположником которой был Коши. В вариационном исчислении Вейерштрассу принадлежит открытие носящего его имя условия экстремума и изучение разрывных решений вариационных задач, в линейной алгебре - построение теории элементарных делителей. ([ДП]; [МЭС]; [К1])

[2] (С.9) Теорема о том, что непрерывная на отрезке действительной оси функция ограничена и достигает на нем наименьшего и наибольшего значений, была найдена Вейерштрассом около 1860 г. ([Ш], Гл.5, С.210).

[3] (С.11) *Ферма* (Fermat) Пьер (1601 - 1665) - французский математик. По профессии юрист, с 1631 г. советник парламента в Тулузе - заведующий отделом прошений (парламент - судебный орган во Франции того времени). Математике посвящал все свободное от службы время. Разработал (1629) первый общий аналитический прием решения экстремальных задач, сыгравший заметную роль в подготовке современных методов дифференциального исчисления. На языке математического анализа, основы которого были заложены Исааком Ньютоном (1643 - 1727) и Готфридом Лейбницем (1646 - 1716) лишь в конце XVII в., этот прием (для случая многочленов) как раз и сводится к указанной в п.2.1 "теореме Ферма". Прием Ферма коротко описан в упражнении 2.3. Этот прием непосредственно связан с предложенным Ферма методом построения касательных к кривым (Ньютон при создании дифференциального исчисления опирался и на этот метод Ферма).

Имя Ферма широко известно благодаря следующему высказанному им (1637) утверждению, известному как "Великая теорема Ферма" : для любого натурального числа $n > 2$ не существует натуральных чисел x, y, z, удовлетворяющих уравнению $x^n + y^n = z^n$. В своих записках Ферма указал, что "открыл этому поистине чудесное доказательство", но не привел его. Общее доказательство этого утверждения Ферма (или его опровержение), несмотря на колоссальные усилия многих математиков, в том числе самых знаменитых, не удавалось получить в течение трех с половиной веков. Теорема была доказана в 1995 г. английским математиком Эндрю Уайлсом (род. в 1953 г., с 1980-ых годов живет в США, проф. Принстонского ун-та). ([С1]; [ДП]; [МЭС]; [АТФ], С.44-45.)

[4] (С.13) *Сильвестр* (Sylvester) Джеймс Джозеф (1814 - 1897) - английский математик, проф. Королевской академии в Вулидже (1855 - 1870), ун-та Джона Хопкинса в Балтиморе (США) (1876 - 1883), Оксфордского ун-та (с 1883). Основал (1878) первый американский математический журнал "The American Journal of Mathematics". [МЭС]

[5] (С.15) *Тейлор* (Taylor) Брук (1685 - 1731) - английский математик. В 1712 г. нашел общую формулу (формула Тейлора) для разложения функций в степенные ряды (ряды Тейлора), опубликовал ее в 1715 г.. Однако, лишь Коши (см. примечание [8]) в 1823 г. установил точные условия сходимости ряда Тейлора и провел отчетливое различие между сходимостью этого ряда вообще и сходимостью к данной функции. ([МЭС]; [Ш], Гл.8, С.299)

[6] (С.15) *Пеано* (Peano) Джузеппе (1858 - 1932) - итальянский математик, с 1890 г. проф. Туринского ун-та. [МЭС]

[7] (С.20) *Гаусс* (Gauss) Карл Фридрих (1777 - 1855) - немецкий математик, внесший фундаментальный вклад также в астрономию и геодезию. С 1807 г. директор Геттингенской астрономической обсерватории и проф. математики и астрономии Геттингенского ун-та. Удивительна широта проблематики исследований Гаусса, от сугубо теоретических в самых разных областях математики (теория чисел, анализ, геометрия, теория вероятностей, методы вычислений и др.) до самых конкретных прикладных (например, он принимал непосредственное практическое участие в организации и проведении целого ряда геодезических съемок местности, а в 1833 г. совместно с физиком Вильгельмом Вебером (1804 - 1891) сконструировал первый в Германии электромагнитный телеграф). Теоретические и прикладные исследования в творчестве Гаусса органически связаны (так, геометрию Гаусс относил к экспериментальным наукам, ставя ее на одну ступень с механикой; открыв неевклидову геометрию, пытался с помощью тщательных геодезических измерений проверить, какая геометрия соответствует окружающему нас пространству). Глубочайшее

проникновение в исследуемый материал сочеталось у Гаусса с необычайно высокой требовательностью к завершенности публикуемых им трудов (в частности и по этой причине многие из выдающихся открытий Гаусса не были им опубликованы и остались неизвестными его современникам). По влиянию на дальнейшее развитие науки Гаусса часто ставят в один ряд с такими титанами науки как Архимед (ок. 287 - 212 до н.э.) и Ньютон. ([МЭС]; [К1], гл.1; [С3], гл.8; [Г2]; [ДП], С.214; [СЭС])

Лежандр (Legendre) Адриен Мари (1752 - 1833) - французский математик. Занимал различные официальные должности: экзаменатор Политехнической школы (см. ниже), инспектор геодезических работ и др.. Обосновал и развил теорию геодезических измерений. Открыл носящие его имя специальные функции и условия экстремума в вариационном исчислении. Написал ряд выдающихся учебников по анализу и геометрии, которые широко использовались в течение многих десятилетий. ([МЭС]; [С3], гл.8)

Политехническая школа (Ecole Politechnique) - основанное в 1794 г. в Париже специальное высшее учебное заведение военной направленности, которое в начале XIX в. стало ведущим учебным заведением по подготовке инженеров вообще и образцом для технических и военных школ Европы и Америки. Обучение в Политехнической школе длилось два года и являлось единственным путем к занятию высших технических государственных должностей во Франции первой половины XIXв.. (В течение следующих двух лет выпускник этой школы продолжал обучение в одном из нескольких специальных учебных заведений. Эти заведения были неодинаковы по своему статусу и выбор выпускника не был свободным, а зависел от отметок в выпускном свидетельстве; наиболее престижным являлся Институт путей сообщения (см. примечание [8]).) Важной составной частью учебного плана Политехнической школы было преподавание, причем в неразрывном единстве, теоретической и прикладной математики. Почти все крупные французские математики первой половины XIX в. либо привлекались к работе в этой школе в качестве преподавателей или экзаменаторов, либо сами закончили ее. Основным организатором школы и долгое время ее директором был математик и инженер Гаспар Монж (1746 - 1818), создатель начертательной геометрии. С Политехнической школой связано появление нового типа учебной литературы. Кроме ученых трактатов для подготовленных читателей, характерных для высшего образования в XVIII в., потребовались специальные учебные руководства. Был издан закон, обязывающий публиковать читающиеся в Политехнической школе лекции. Таким образом возникло подавляющее большинство ведущих учебников по математике начала XIX в.. Влияние этих учебников можно проследить до наших дней. Они сыграли значительную роль в развитии математики. ([К1], гл.2; [С3], гл.8)

Метод наименьших квадратов был открыт независимо Гауссом и Лежандром. В 1809 г. вышла (законченная в 1807 г.) знаменитая книга Гаусса "Теория движения небесных тел, обращающихся вокруг Солнца по коническим сечениям". В ней Гаусс изложил свои методы вычисления орбит. Был изложен и метод наименьших квадратов. Гаусс указал, что знает этот метод с 1794 г., а с 1802 г. систематически им пользуется. За два года до выхода книги Гаусса метод наименьших квадратов был опубликован Лежандром; открыт им в 1805 - 1806 г.г.. Основные публикации Гаусса о методе наименьших квадратов относятся к 1820-ым годам и связаны с обработкой результатов проводившихся им по заказу правительства королевства Ганновер геодезических измерений. ([Г2], С.179; [МЭ], С.875.)

[8] (С.22) *Бине* (Binet) Жак Филипп Мари (1786 - 1856) - французский математик и астроном. [МЭС]

Коши (Cauchy) Огюстен Луи (1789 - 1857) - французский математик. В 1807 г. окончил Политехническую школу (см. примечание[7]), в 1810г. - Институт путей сообщения (см. ниже). В 1816 - 1830 г.г. преподавал в Политехнической школе и Коллеж де Франс (см. ниже), с 1848 г. - в Сорбонне (Сорбонна (Sorbonne) - второе, распространенное название Парижского ун-та) и Коллеж де Франс (после революции 1830 г. был в эмиграции до 1838 г.). Первым, рассматривая предел как чисто арифметическое понятие, Коши дал строгое определение бесконечно малой величины (1821), а на его основе - определения непрерывной функции и производной, близкие к современным. Нужно отметить, что чешский математик Бернард Больцано (1781 - 1848) уже к 1817 г. владел точным понятием непрерывной функции, но его работы долгое время оставались неизвестными; именно Коши стал первым "проводником строгости" в анализ. В 1821 г. вышел знаменитый "Курс анализа (алгебраический анализ)", написанный Коши на основе лекций, читавшихся им в Политехнической школе. Основанный на систематическом использовании понятия предела, он служил образцом для большинства курсов позднейшего времени. Коши первым дал точные определения сходимости ряда (1821) и интеграла как предела интегральных сумм (1823), первым доказал существование решения "задачи Коши" для дифференциального уравнения (1844). Он - основоположник теории функций комплексного переменного, а также математической теории упругости (на основе понятия предела ввел точные понятия напряжения и деформации в точке). ([К1], гл.2; [МЭС]; [СЭС]; [ДП]; [Ш]; [Б2].)

Институт путей сообщения (Ecole des Ponts et Chausses) - специальное высшее учебное заведение военной направленности, созданное в Париже и предназначенное прежде всего для дальнейшего образования лучших выпускников Политехнической школы (см. примечание [7]). Основные направления обучения: прикладная механика, практика конструирования, архитектура, иностранные языки. Значительная часть обучения осуществлялась в военных лагерях. ([К1]; [Б2])

Коллеж де Франс (College de France) - основанное в 1530 г. в Париже одно из старейших научно-исследовательских и учебных заведений Франции. Работу в нем ведут выдающиеся ученые по математике и естественным наукам. Принимает для обучения лиц с высшим образованием. [СЭС]

[9] (С.24) *Гессе* (Hesse) Людвиг Отто (1811 - 1874) - немецкий математик, с 1869 г. проф. Политехнической школы в Мюнхене. [МЭС]

[10] (С.26) *Штейнер* (Steiner) Якоб (1796 - 1863) - швейцарский математик, с 1835 г. проф. Берлинского ун-та. [МЭС]

[11] (С.27) *Торричелли* (Torricelli) Эванджелиста (1608 - 1647) - итальянский математик и физик, ученик Галилео Галилея (1564 - 1642), с 1642 г. проф. математики Флорентийского ун-та. Изобрел ртутный барометр, открыл существование атмосферного давления и вакуума (торричеллиева пустота). ([МЭС], [СЭС].)

Литература

[А] Ахмеров Р.Р. Методы оптимизации гладких функций. Учебное пособие. Новосибирск: Новосибирский университет. 1992.

[АТ] Ашманов С.А., Тимохов А.В. Теория оптимизации в задачах и упражнениях. М.: Наука. 1991.

[АТФ] Алексеев В.М., Тихомиров В.М., Фомин С.В. Оптимальное управление. М.: Наука. 1979.

[Б1] Беллман Р. Введение в теорию матриц. М.: Наука. 1969.

[Б2] Белхост Б. Огюстен Коши. М.: Наука. Физматлит. 1997.

[В] Васильев Ф.П. Численные методы решения экстремальных задач. М.: Наука. 1988.

[Г1] Гантмахер Ф.Р. Теория матриц. М.: Наука. 1966.

[Г2] Гиндикин С.Г. Рассказы о физиках и математиках. М.: Наука. 1981.

[ГО] Гелбаум Б., Олмстед Дж. Контрпримеры в анализе. М.: Мир. 1967.

[ГТ] Галеев Э.М., Тихомиров В.М. Оптимизация: теория, примеры, задачи. М.: Эдиториал УРСС. 2000.

[ДП] Даан-Дальмедико А., Пейффер Ж. Пути и лабиринты. Очерки по истории математики. М.: Мир. 1986.

[З] Зорич В.А. Математический анализ. Часть 1. М.: Наука. 1981.

[ИСС] Ильин В.А., Садовничий В.А., Сендов Бл.Х. Математический анализ. М.: Наука. 1979.

[К1] Клейн Ф. Лекции о развитии математики в XIX столетии: В 2-х томах. Т.1. М.: Наука. 1989.

[К2] Крылов А.Н. Лекции о приближенных вычислениях. М.: ГИТТЛ. 1954.

[МЭ] Математическая энциклопедия: Гл. ред. Виноградов И.М.. Т.3. М: Сов. энциклопедия. 1982.

[МЭС] Математический энциклопедический словарь: Гл. ред. Прохоров Ю.В.. М: Сов. энциклопедия. 1988.

[П] Поляк Б.Т. Введение в оптимизацию. М.: Наука. 1983.

[С1] Сингх С. Великая теорема Ферма. М.: Изд-во Московского центра непрерывного математического образования. 2000.

[С2] Стренг Г. Линейная алгебра и ее применения. М.: Мир. 1980.

[С3] Стройк Д.Я. Краткий очерк истории математики. М.: Наука. 1984.

[СЭС] Советский энциклопедический словарь / Научно-редакционный совет: А.М.Прохоров (пред.). - М.: Сов. энциклопедия. 1981.

[Т] Тихомиров В.М. Рассказы о максимумах и минимумах. М.: Наука. 1986.

[Ф] Фихтенгольц Г.М. Курс дифференциального и интегрального исчисления. Т.1. М.: Физматгиз. 1962.

[Ш] Шилов Г.Е. Математический анализ (функции одного переменного). Части 1-2. М.: 1969; Часть 3. М.: 1970.

[ЭВА] Сумин В.И. Элементарный выпуклый анализ. Учебно-методическое пособие. Нижний Новгород: ННГУ, 2007.

Владимир Иосифович **Сумин**

НАЧАЛА МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ. ТЕОРЕМА ВЕЙЕРШТРАССА. БЕЗУСЛОВНЫЙ ЭКСТРЕМУМ

Учебно-методическое пособие

Нижегородский государственный университет им. Н.И. Лобачевского

Механико-математический факультет

Кафедра математической физики