МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Московский государственный институт электроники и математики
(Технический университет)

Л.А. МАНИТА

УСЛОВИЯ ОПТИМАЛЬНОСТИ В КОНЕЧНОМЕРНЫХ НЕЛИНЕЙНЫХ ЗАДАЧАХ ОПТИМИЗАЦИИ

Утверждено Редакционно-издательским советом института
в качестве Учебного пособия

Москва 2010

УДК 519.85
ББК 22

Рецензенты: д.ф.-м.н., профессор Л. Г. Афанасьева (МГУ им. М.В. Ломоносова, Механико-математический факультет)
д.ф.-м.н., профессор, ведущий научный сотрудник Д. Д. Соколов (Научно-исследовательский вычислительный центр МГУ)

Манита Л.А.
Условия оптимальности в конечномерных нелинейных задачах оптимизации. Учебное пособие — Московский государственный институт электроники и математики. М., 2010. — 84 с.

ISBN

В учебном пособии рассмотрен один из разделов математических методов исследования операций — методы оптимизации в конечномерных пространствах. Изложены необходимые и достаточные условия оптимальности для некоторых классов задач конечномерной оптимизации: задачи без ограничений, задачи с ограничениями в виде равенств и неравенств, выпуклые задачи, в том числе задачи выпуклого программирования. Для каждого класса задач приведены примеры с подробными решениями.

Пособие предназначено для студентов, обучающихся по специальности “Математические методы в экономике”, а также может быть полезно и для студентов других специальностей (например, “Прикладная математика”).

УДК 519.85
ББК 22
ISBN

Предисловие

Настоящее пособие написано на основе читаемого в 5 семестре для студентов специальности “Математические методы в экономике” (группы МЭ-51, МЭ-52, ЭМ-61) курса лекций “Математические методы и модели исследования операций”. Цель пособия — дать представление о методах исследования конечномерных задач оптимизации, выработать навыки математической формализации прикладных задач и использования теории необходимых и достаточных условий оптимальности для нахождения решений полученных оптимизационных моделей.

Пособие состоит из 8 разделов и приложения. В разделе 1 дается определение задачи оптимизации и ее решения, приводятся теоремы о достаточных условиях существования решений. В разделе 2 рассматриваются задачи без ограничений, доказаны необходимые и достаточные условия оптимальности для одномерной и многомерной задач без ограничений. Приводятся необходимые сведения о знакоопределенных и знакопеременных симметрических матрицах, формулируется критерий Сильвестра для исследования симметрических матриц. В разделе 3 определяется задача условной оптимизации и описан метод геометрического решения таких задач. В разделе 4 рассматриваются задачи с ограничениями в виде равенств, в разделе 5 — с ограничениями в виде равенств и неравенств, описан метод Лагранжа решения таких задач, необходимые и достаточные условия оптимальности. В разделе 6 множители Лагранжа характеризуются как оценки изменения оптимального значения целевой функции при малых возмущениях ограничений. В разделе 7 даются необходимые сведения о выпуклых множествах и выпуклых функциях, определяется выпуклая задача оптимизации, приводится критерий оптимальности в выпуклой задаче, изучается частный случай выпуклой задачи — задача выпуклого программирования. В разделе 8 рассматриваются некоторые задачи математической экономики (задача о потреблении, задача об оптимальном выпуске, задача об оптимальном портфеле), показано, как применять условия оптимальности для исследования этих задач. В приложении приведены некоторые важные факты из математического анализа и линейной алгебры, которые используются в данном пособии.

В пособии разбираются задачи, которые демонстрируют правила применения необходимых и достаточных условий оптимальности и дают представление о методах исследования оптимизационных задач разного типа.

Содержание

1 Постановка задачи оптимизации. Теоремы существования решений

Пусть $f_0\,(x)$ — скалярная функция, определенная на некотором множестве $D \subset \mathbf{R}^n$. Рассмотрим задачи нахождения точек минимума или максимума функции f_0 на множестве D. Такие задачи будем называть задачами оптимизации. Задачу на минимум будем записывать в виде

$$f_0\,(x) \to \min, \quad x \in D, \tag{1}$$

задачу на максимум в виде

$$f_0\,(x) \to \max, \quad x \in D.$$

Если же требуется найти и точки минимума, и точки максимума, то будем писать

$$f_0\,(x) \to extr, \quad x \in D. \tag{2}$$

Функцию f_0 будем называть целевой функцией, множество D — множеством допустимых решений (допустимым множеством). Отметим, что задача на максимум сводится к задаче на минимум. А именно, допустимая точка x^* (то есть, $x^* \in D$) является решением задачи на максимум:

$$f_0\,(x) \to \max, \quad x \in D$$

тогда и только тогда, когда x^* — решение задачи на минимум

$$f\,(x) = -f_0\,(x) \to \min, \quad x \in D,$$

при этом $f\,(x^*) = -f_0\,(x^*)$.

Дадим строгое определение решения задачи оптимизации.

Определение 1.1 *Допустимая точка x^* является точкой локального минимума в задаче (2) (в дальнейшем будем обозначать как $x^* \in locmin(2)$), если $\exists \varepsilon > 0$, такое, что для $\forall x \in D \bigcap U_\varepsilon(x^*)$ выполняется*

$$f_0(x) \geq f_0(x^*). \tag{3}$$

Здесь $U_\varepsilon(x^*)$ — ε- окрестность точки x^*:

$$U_\varepsilon(x^*) = \{\, x \in \mathbf{R}^n |\ \parallel x\ - x^* \parallel < \varepsilon\}$$

Если в определении 1.1 неравенство (3) выполнено строго для $x \neq x^*$, то x^* — строгий локальный минимум.

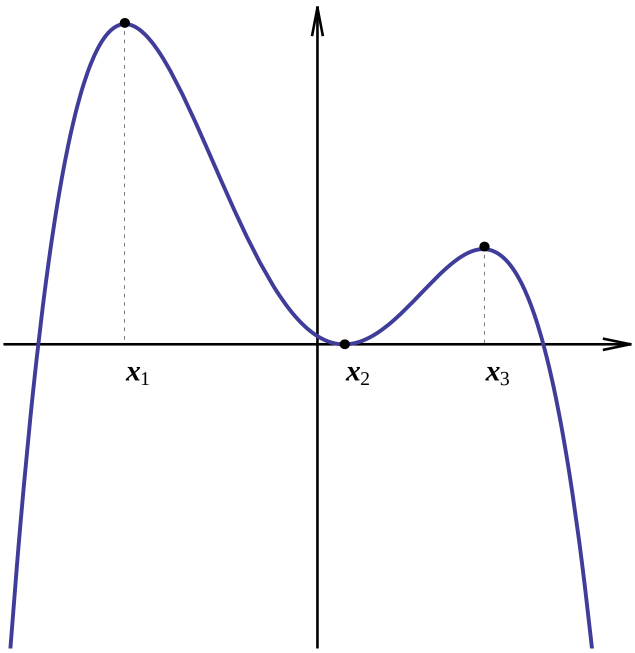

Рис. 1: x_2 — локальный минимум, x_3 — локальный максимум, x_1 — абсолютный максимум, абсолютного минимума нет.

Замечание 1 *Если заменить в определении 1.1 знак неравенства (3) на противоположный, то получим определение локального максимума для задачи (2).*

Определение 1.2 *Допустимая точка x^* является точкой глобального (или абсолютного) минимума в задаче (2) (будем обозначать в дальнейшем как $x^* \in absmin(1)$), если для $\forall x \in D$ выполняется*

$$f_0(x) \geq f_0(x^*). \tag{4}$$

Если неравенство (4) выполнено строго для $x \neq x^*$, то x^* — строгий абсолютный минимум.

Определение 1.3 *Точки локального минимума и локального максимума функции f_0 называют точками локального экстремума.*

Из определений локального и абсолютного минимумов (максимумов) следует, что абсолютный минимум (максимум) является и локальным минимумом (максимумом). Обратное неверно, т.е. не всегда локальный минимум (максимум) является абсолютным (см. рис.1).

В основном, необходимые и достаточные условия оптимальности определяют локальные экстремумы, но не отвечают на вопрос о существовании глобальных экстремумов. Ниже приведем условия, при которых глобальные экстремумы существуют. Тогда при исследовании задач можно воспользоваться следующим соображением: если для задачи выполнены условия существования глобального экстремума, а необходимым условиям локального экстремума удовлетворяет единственная допустимая точка, то она и будет глобальным экстремумом. Напомним важные результаты из курса математического анализа.

Теорема 1 *(теорема Вейерштрасса)*
Пусть D — компакт в $\mathbf{R}^n$ (т.е. замкнутое ограниченное множество), функция $f_0(x)$ непрерывна на D. Тогда точка глобального минимума и точка глобального максимума функции f_0 на D существуют.

Доказательство теоремы 1 можно найти, например, в [8]. Иногда бывает полезной и другая форма данной теоремы.

Теорема 2
Пусть D — замкнутое множество в $\mathbf{R}^n$, $f_0(x)$ — непрерывная функция на D. Если для некоторой точки $a \in \mathbf{R}^n$ множество

$$N\,(a) = \{x \in D \,|\, f_0(x) \leq f_0(a)\}$$

ограничено, то у функции f_0 на D существует точка глобального минимума.

Доказательство. Так как множество D замкнуто, а функция f_0 непрерывна, то множество $N\,(a)$ замкнуто, и кроме того оно ограничено, а следовательно, является компактом. Очевидно, что

$$\inf_{D} f_0(x) = \inf_{N(a)} f_0(x)$$

Из теоремы Вейерштрасса следует, что на множестве $N\,(a)$ существует точка глобального минимума, следовательно, она же будет являться точкой глобального минимума на D. Теорема доказана. □

Приведем без доказательства еще одну полезную теорему существования решений в оптимизационных задачах. Для этого определим понятие бесконечно растущей функции [4].

Определение 1.4 *Функция $f_0(x)$ называется бесконечно растущей на множестве D, если на всякой последовательности $\{x_k\}_{k=1}^{\infty} \subset D$, такой, что либо*

$$\|x_k\| \longrightarrow +\infty, \quad k \to +\infty,$$

либо

$$x_k \to x^0 \in \overline{D} \backslash D,$$

где $\overline{D}$ — замыкание множества D, выполняется

$$\lim_{k\to\infty} f_0(x_k) = +\infty. \tag{5}$$

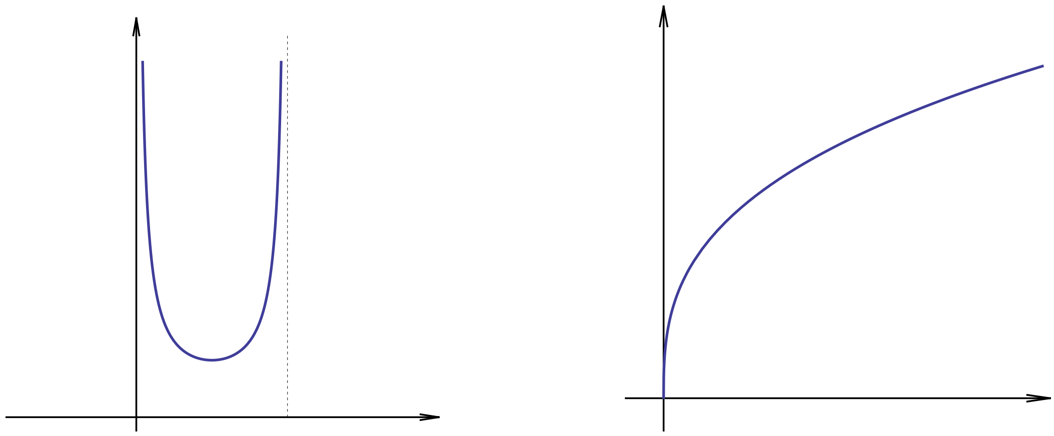

Рис. 2: примеры бесконечно растущих функций

Замечание 2 *Если D — открытое множество, то $\overline{D}\backslash D$ — граничные точки множества D, если D — замкнутое множество, то $\overline{D}\backslash D = \emptyset$.*

На рис. 2 функции бесконечно растущие, слева — на интервале $(0, 1)$, справа — на луче $[0, \infty)$.

Теорема 3 *(существование глобального минимума у бесконечно растущей функции)*
Пусть f_0 — бесконечно растущая непрерывная на D функция, тогда на D существует глобальный минимум функции f_0.

***Пример* 1** *Рассмотрим в $\mathbf{R}^2$ функцию $f_0(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2$. Покажем, что f_0 является бесконечно растущей функцией.*

Решение. Действительно, имеем

$$x_1^2 + x_1x_2 + x_2^2 \;\geq\; x_1^2 + \frac{x_1^2 + x_2^2}{2} + x_2^2 = \frac{3}{2}\left(x_1^2 + x^2\right) = \frac{3}{2}\,\|(x_1, x_2)\|^2 \;\to\; +\infty$$

при $\|(x_1, x_2)\| \to +\infty$.

***Пример* 2** *Покажем, что функция $f_0\,(x_1, x_2) = x_1^2 + 4x_1x_2 + x_2^2$ не является бесконечно растущей в $\mathbf{R}^2$.*

Решение. Имеем

$$x_1^2 + 4x_1x_2 + x_2^2 = (x_1 + 2x_2)^2 - 3x_2^2.$$

Рассмотрим следующую последовательность точек

$$\{x_k = (x_{1k}, x_{2k})\}_{k=1}^{\infty} : \quad x_{1k} = -2k, \quad x_{2k} = k.$$

Тогда

$$\|x_k\|^2 = 5k^2 \to \infty, \quad k \to +\infty,$$

и

$$f_0\left(x_{1k}, x_{2k}\right) = -3k^2 \to -\infty, \quad k \to +\infty,$$

следовательно, f_0 не является бесконечно растущей функцией. Очевидно, что не является бесконечно растущей и функция $-f_0$.

При изучении оптимизационных задач важное значение имеют условия оптимальности, среди которых выделяют необходимые условия оптимальности, то есть те условия, которым должна удовлетворять точка, являющаяся решением задачи, и достаточные условия, то есть те условия, при выполнении которых исследуемая точка обязательно является решением задачи. Условия оптимальности позволяют в некоторых случаях получить полное решение, а также используются при построении и обосновании численных методов решения этих задач.

Оптимизационные задачи можно разделить на несколько классов в зависимости от свойств функции f_0 и вида множества D. Мы рассмотрим некоторые из этих классов.

2 Задачи без ограничений

2.1 Одномерная задача без ограничений

Одномерной задачей без ограничений будем называть следующую задачу

$$f_0(x) \to extr, \quad x \in \mathbf{R}^1. \tag{6}$$

Теорема 4 *(Необходимые условия локального минимума в одномерной задаче без ограничений) [8]*
Пусть функция $f_0(x)$ s раз дифференцируема в точке x^* и

$$f_0'\left(x^*\right) = ... = f_0^{(m-1)}\left(x^*\right) = 0,\ f_0^{(m)}\left(x^*\right) \neq 0$$

для некоторого $m \in \mathbb{N}, m \leq s$. Если m нечетно, то x^* не является экстремумом; если m четно, то x^* является строгим локальным экстремумом, причем $x^* \in locmin(6)$, если $f_0^{(m)}\left(x^*\right) > 0$, и $x^* \in locmax(6)$, если $f_0^{(m)}\left(x^*\right) < 0$.

Доказательство. Используем формулу Тейлора с остаточным членом в форме Пеано:

$$f_0\left(x^* + \triangle x\right) - f_0\left(x^*\right) = \sum_{k=1}^{s} \frac{1}{k!} f_0^{(k)}\left(x^*\right)\left(\triangle x\right)^k + o\left(\left(\triangle x\right)^s\right) =$$

$$= \sum_{k=m}^{s} \frac{1}{k!} f_0^{(k)}\left(x^*\right) \left(\triangle x\right)^k + o\left(\left(\triangle x\right)^s\right) =$$

$$= \frac{f_0^{(m)}\left(x^*\right)}{m!} \left(\triangle x\right)^m \left(1 + \sum_{k=m+1}^{s} \frac{f_0^{(k)}\left(x^*\right) m!}{f_0^{(m)}\left(x^*\right) k!} \left(\triangle x\right)^{k-m} + o\left(\left(\triangle x\right)^{s-m}\right)\right) =$$

$$= \frac{f_0^{(m)}\left(x^*\right)}{m!} \left(\triangle x\right)^m \left(1 + o\left(\triangle x\right)\right).$$

При малых $\triangle x$ выражение в скобке положительно, следовательно, знак $f_0\left(x^* + \triangle x\right) - f_0\left(x^*\right)$ зависит от знака $\frac{f_0^{(m)}(x^*)}{m!}\left(\triangle x\right)^m$. Если m нечетно, то последнее выражение может быть как больше 0, так и меньше 0, так как $\triangle x$ может иметь любой знак. Таким образом, точка x^* не может являться точкой экстремума. Значит, в точке локального экстремума m должно быть четным. В этом случае $\left(\triangle x\right)^m > 0$ при $\triangle x \neq 0$ и при малых $\triangle x$ знак $f_0\left(x^* + \triangle x\right) - f_0\left(x^*\right)$ совпадает со знаком $f_0^{(m)}\left(x^*\right)$. Таким образом, если $f_0^{(m)}\left(x^*\right) > 0$, то x^* — строгий локальный минимум, если $f_0^{(m)}\left(x^*\right) < 0$, то x^* — строгий локальный максимум. $\square$

***Пример* 3** *Исследовать точку $x^* = 0$ для функции $f\left(x\right) = x^3 \exp\left(-x\right)$.*

Решение. Вычислим производные функции f в точке 0:

$$\begin{aligned}
f'\left(x\right) &= 3x^2 \exp\left(-x\right) - x^3 \exp\left(-x\right), \\
& \quad f'\left(0\right) = 0, \\
f''\left(x\right) &= 6x \exp\left(-x\right) - 3x^2 \exp\left(-x\right) - 3x^2 \exp\left(-x\right) + x^3 \exp\left(-x\right) = \\
&= 6x \exp\left(-x\right) - 6x^2 \exp\left(-x\right) + x^3 \exp\left(-x\right), \\
& \quad f''\left(0\right) = 0, \\
f'''\left(x\right) &= 6 \exp\left(-x\right) - 6x \exp\left(-x\right) - 12x \exp\left(-x\right) + \\
& \quad + 6x^2 \exp\left(-x\right) + 3x^2 \exp\left(-x\right) - x^3 \exp\left(-x\right) = \\
&= 6 \exp\left(-x\right) - 18x \exp\left(-x\right) + 9x^2 \exp\left(-x\right) - x^3 \exp\left(-x\right), \\
& \quad f'''\left(0\right) = 6.
\end{aligned}$$

Первая ненулевая производная — третья, нечетная, следовательно, точка $x^* = 0$ не является экстремумом.

Замечание 3 *Если все производные до порядка s включительно в точке x^* равны нулю, то точка x^* может быть как точкой локального экстремума, так и не являться экстремумом.*

***Пример* 4** *Исследовать точку $x^* = 0$ для функций*

$$f_0(x) = \begin{cases} x^{s+\frac{1}{2}}, & x \geq 0, \\ (-x)^{s+\frac{1}{2}}, & x < 0, \end{cases}$$

и

$$f_1(x) = \begin{cases} x^{s+\frac{1}{2}}, & x \geq 0, \\ -(-x)^{s+\frac{1}{2}}, & x < 0. \end{cases}$$

Решение. Функции $f_0\,(x)$и $f_1\,(x)$ s раз непрерывно дифференцируемы в точке $x^* = 0$, и все их производные в нуле равны нулю

$$f_i'\,(0) = ... = f_i^{(s)}\,(0) = 0, \quad i = 1, 2.$$

При этом $x^* = 0$ является точкой глобального минимума функции f_0 и не является точкой экстремума функции f_1.

2.2 Знакоопределенные симметрические матрицы. Критерий Сильвестра

При изучении задач оптимизации в пространствах размерности больше 1 возникает необходимость исследовать знакоопределенность матрицы вторых производных целевой функции. В этом параграфе приведены необходимые сведения о знакоопределенных и знакопеременных симметрических матрицах.

Определение 1
Действительную симметрическую матрицу $A_{n\times n}$ будем называть
- неотрицательно определенной, если $(Az, z) \geq 0 \quad \forall z \in \mathbf{R}^n$;
- положительно определенной, если $(Az, z) > 0 \quad \forall z \in \mathbf{R}^n$, $z \neq 0$;
- неположительно определенной, если $(Az, z) \leq 0 \quad \forall z \in \mathbf{R}^n$;
- отрицательно определенной, если $(Az, z) < 0 \quad \forall z \in \mathbf{R}^n$, $z \neq 0$;
- знакопеременной или знаконеопределенной, если найдутся z_1, $z_2 \in \mathbf{R}^n$, такие, что $(Az_1, z_1) > 0$, $(Az_2, z_2) < 0$.

Знакоопределенность матрицы A связана со знаками собственых чисел матрицы. Напомним кратко необходимые сведения о собственных числах матрицы.

Определение 2 *Число λ является собственным числом матрицы $A_{n\times n}$ с собственным вектором $v \in \mathbf{R}^n$, $v \neq 0$, если $Av = \lambda v$.*

Действительная симметрическая матрица обладает действительным спектром, т.е. все собственные числа матрицы A действительны. Действительная симметрическая матрица A положительно определена, отрицательно определена, неотрицательно определена, неположительно определена, знаконеопределена тогда и только тогда, когда собственные значения λ_j матрицы A соответственно все положительны, все отрицательны, все неотрицательны, все неположительны, имеют различные знаки.

Для проверки знакоопределенности матрицы можно использовать критерий Сильвестра. Для этого приведем определения угловых и главных миноров матрицы.

Определение 3 *Угловым минором $\triangle_k$ порядка k $(1 \le k \le n)$ матрицы $A_{n\times n}$ называется определитель подматрицы, которая получается из матрицы A вычеркиванием строк и столбцов с номерами $k+1, ..., n$.*

По определению считаем, что угловой минор $\triangle_n$ равен определителю матрицы A.

Определение 4 *Главным минором матрицы $A_{n\times n}$ называется определитель подматрицы, которая получается из матрицы A вычеркиванием строк и столбцов с одинаковыми номерами (включается случай, при котором не вычеркивается ничего).*

Будем использовать следующие обозначения для главных миноров: $\delta_{i_1 i_2 \ldots i_k}$ — определитель подматрицы, которая получается из матрицы A вычеркиванием строк и столбцов с номерами $i_1, i_2, \ldots, i_k$

$$(i_1, i_2, \ldots, i_k \in \{1, \ldots, n\}\,,\ 1 \le k \le n-1)\,.$$

Например, δ_2 — определитель подматрицы, которая получается из матрицы A вычеркиванием строки и столбца с номерами 2, δ_{12} — определитель подматрицы, которая получается из матрицы A вычеркиванием строк и столбцов с номерами 1 и 2, δ_{124} — определитель подматрицы, которая получается из матрицы A вычеркиванием строк и столбцов с номерами 1, 2 и 4, и т.д. Через δ_0 будем обозначать главный минор, который совпадает с определителем матрицы A.

Очевидно, что угловые миноры являются главными, а именно:

$$\triangle_k = \delta_{k+1,\ldots,n}, \quad k = 1, \ldots, n-1, \quad \triangle_n = \delta_0.$$

***Пример* 5** *Вычислить угловые и главные миноры матрицы*

$$A = \begin{pmatrix} 4 & 2 & 3 \\ 2 & 1 & 0 \\ 3 & 0 & 2 \end{pmatrix}.$$

Решение. Угловые миноры:

$$\triangle_1 = 4, \quad \triangle_2 = 0, \quad \triangle_3 = -9.$$

Главные миноры:

$$\delta_1 = 2, \quad \delta_2 = -1, \quad \delta_3 = 0, \quad \delta_{12} = 2, \quad \delta_{13} = 1, \quad \delta_{23} = 4, \quad \delta_0 = -9.$$

Здесь $\triangle_1 = \delta_{23}, \quad \triangle_2 = \delta_3, \quad \triangle_3 = \delta_0$.

Теорема 5 *(Критерий Сильвестра)*

Симметрическая матрица A положительно определена тогда и только тогда, когда все ее угловые миноры положительны. Симметрическая матрица A неотрицательно определена тогда и только тогда, когда все ее главные миноры неотрицательны. Симметрическая матрица A отрицательно определена тогда и только тогда, когда ее угловые миноры имеют чередующиеся знаки, начиная с минуса.

Доказательство можно найти, например, в [7].

Критерий Сильвестра приводим только для проверки неотрицательной, положительной и отрицательной определенности матрицы, так как соответствующие условия имеют простой вид. Если нужно проверить матрицу A на неположительную определенность, то можно ввести матрицу $B = -A$ и проверить ее, используя критерий Сильвестра, на неотрицательную определенность. Аналогично можно поступить и при проверке матрицы на отрицательную определенность (т.е. вместо проверки A на отрицательную определенность провести проверку матрицы $B = -A$ на положительную определенность).

Замечание 4 *Так как угловые миноры являются главными минорами, то матрица не может быть неотрицательно определенной, если среди угловых миноров есть отрицательные.*

Неотрицательность угловых миноров не является достаточным условием неотрицательной определенности матрицы.

***Пример* 6** *Проверить знакоопределенность матрицы*

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Решение. Угловые миноры неотрицательны: $0, \quad 0, \quad 1$. Следовательно, матрица A может быть неотрицательно неопределенной. Вычислим главные миноры:

$$0, \, -1, \, 0, \, 0, \, 0, \, -1, \, 1.$$

По критерию Сильвестра матрица A знакопеременная.

2.3 n – мерная задача без ограничений ($n \geq 2$)

Рассмотрим задачу

$$\left\{ \begin{array}{l} f_0(x) \to extr, \\ x \in \mathbf{R}^n, \quad n \geq 2. \end{array} \right. \tag{7}$$

Для функции многих переменных через $f_0'(x)$ будем обозначать вектор частных производных (градиент) функции f_0 в точке x:

$$f_0'(x) = \left(\frac{\partial f_0(x)}{\partial x_1}, ..., \frac{\partial f_0(x)}{\partial x_n} \right).$$

Символом $f_0''(x)$ будем обозначать матрицу вторых производных функции f_0, вычисленную в точке x:

$$f_0''(x) = \left(\frac{\partial^2 f_0(x)}{\partial x_i \partial x_j} \right)_{i,j=1}^n$$

Теорема 6 *(Необходимое условие локального экстремума 1-го порядка в n – мерной задаче без ограничений)*
Пусть $f_0(x)$ дифференцируема в окрестности точки x^*. Если точка x^* является локальным экстремумом в задаче (7), то

$$f_0'(x^*) = 0. \tag{8}$$

Доказательство. Пусть для определенности $x^* \in locmin$ (7). Рассмотрим произвольный вектор $h \in \mathbf{R}^n$. Используем дифференцируемость функции f_0 и определение локального минимума. Тогда при всех достаточно малых α имеем:

$$0 \leq f_0(x^* + \alpha h) - f_0(x^*) = (f_0'(x^*), \alpha h) + o(\alpha)$$

Разделим обе части неравенства на $\alpha > 0$ и перейдем к пределу при $\alpha \to +0$, получим

$$(f_0'(x^*), h) \geq 0.$$

Предположим, что $f_0'(x^*) \neq 0$, положим $h = -f_0'(x^*)$. Тогда

$$(f_0'(x^*), -f_0'(x^*)) = -\|f_0'(x^*)\|^2 \geq 0.$$

Последнее неравенство выполнено только при $f_0'(x^*) = 0$. Противоречие с предположением. Для локального максимума доказательство аналогично. □

Точка x^*, удовлетворяющая условию $f_0'(x^*) = 0$, называется стационарной точкой функции f_0. Стационарная точка необязательно будет являться решением задачи (7). Для дальнейшей проверки стационарных точек можно использовать необходимое условие оптимальности второго порядка.

Теорема 7 *(Необходимое условие локального экстремума 2-го порядка в n – мерной задаче без ограничений)*
Пусть функция f_0 дважды непрерывно дифференцируема в окрестности точки $x^* \in \mathbb{R}^n$. Если $x^* \in locmin(7)$, то матрица $f_0''(x^*)$ неотрицательно определена. Если $x^* \in locmax(7)$, то матрица $f_0''(x^*)$ неположительно определена.

Доказательство. Пусть x^* — локальный минимум, тогда, учитывая (8), для любого $h \in \mathbf{R}^n$ при достаточно малых α имеем

$$0 \leq f_0\left(x^* + \alpha h\right) - f_0\left(x^*\right) = \frac{1}{2}\left(f_0''\left(x^*\right)\alpha h, \alpha h\right) + o\left(\alpha^2\right).$$

Разделим обе части неравенства на α^2 и перейдем к пределу при $\alpha \to 0$, получим

$$\left(f_0''\left(x^*\right)h, h\right) \geq 0,$$

то есть, матрица $f_0''(x^*)$ неотрицательно определена. Для локального максимума рассуждения аналогичны. □

Замечание 5 *Если $f_0''\left(x^*\right)$ — знакопеременная матрица, то есть существуют $h^1, h^2 \in \mathbf{R}^n$, такие, что*

$$\left(f_0''\left(x^*\right)h^1, h^1\right) > 0,\ \left(f_0''\left(x^*\right)h^2, h^2\right) < 0,$$

то x^* не является локальным экстремумом.

Достаточное условие оптимальности содержит усиление требований на матрицу $f_0''(x^*)$.

Теорема 8 *(Достаточное условие локального экстремума в n – мерной задаче без ограничений)*
Пусть функция f_0 дважды непрерывно дифференцируема в окрестности точки $x^* \in \mathbb{R}^n$. Если $f_0'(x^*) = 0$, а матрица $f_0''(x^*)$ положительно определена, то x^* — строгий локальный минимум задачи (7). Если $f_0'(x^*) = 0$, а матрица $f_0''(x^*)$ отрицательно определена, то x^* — строгий локальный максимум задачи (7).

Доказательство. Проведем доказательство для минимума, для максимума получается аналогично. Допустим, что утверждение теоремы неверно, то есть x^* не является строгим локальным минимумом. Тогда существует последовательность точек $\left\{x^k\right\}_{k=1}^{\infty}$, удовлетворяющая условиям: $x^k \neq x^*$, $x^k \to x^*$ и

$$f_0\left(x^k\right) \leq f_0\left(x^*\right). \tag{9}$$

Запишем x^k в виде $x^k = x^* + \alpha_k h^k$, где $\alpha_k = \left\|x^k - x^*\right\|$, $h^k = \left(x^k - x^*\right)/\alpha_k$. Так как $\left\|h^k\right\| = 1$, то из ограниченной последовательности $\left\{h^k\right\}_{k=1}^{\infty}$ можно выбрать сходящуюся подпоследовательность $\left\{h^{k_s}\right\}_{s=1}^{\infty}$. В силу громоздкости двойную индексацию использовать не будем, считая, что последовательность $\left\{h^k\right\}_{k=1}^{\infty}$ уже является сходящейся. Обозначим

$$\lim_{k\to\infty} h^k = h^*.$$

Очевидно, что $h^* \neq 0$. Так как по условию теоремы $f_0'(x^*) = 0$, а по предположению выполнено (9), то имеем следующую оценку:

$$0 \geq f_0\left(x^k\right) - f_0\left(x^*\right) = \frac{1}{2}\left(f_0''\left(x^*\right)\alpha_k h^k, \alpha_k h^k\right) + o\left(\alpha_k^2\right).$$

Разделим обе части неравенства на α_k^2 и перейдем к пределу при $k \to \infty$, получим

$$0 \geq \frac{1}{2}\left(f_0''\left(x^*\right)h^*, h^*\right).$$

Противоречие с условием теоремы. Следовательно, x^*— строгий локальный минимум. $\square$

Если функция f_0 достаточно проста, то приведенные необходимые и достаточное условия позволяют полностью решить задачу (7).

Пример 7 *Найти точки экстремума функции*

$$f_0\left(x_1, x_2\right) = x_1^2 + 4x_1x_2 + x_2^2.$$

Решение. Применим теорему о необходимых условиях первого порядка. Найдем стационарные точки функции f_0, т.е. точки, в которых производная (градиент) обращается в нуль:

$$f_0'\left(x_1, x_2\right) = \left(2x_1 + 4x_2, 4x_1 + 2x_2\right);$$

$$f_0'\left(x_1, x_2\right) = 0 \Leftrightarrow x_1 = x_2 = 0.$$

Итак, точка $\left(x_1^*, x_2^*\right) = (0, 0)$ удовлетворяет необходимым условиям первого порядка. Проверим выполнение необходимых условий второго порядка и достаточных условий. Вычислим матрицу вторых производных:

$$f_0''\left(x_1, x_2\right) = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}.$$

Угловые миноры

$$\triangle_1 = 2 > 0,\ \triangle_2 = 4 - 16 = -12 < 0.$$

По критерию Сильвестра матрица вторых производных является знакопеременной. Следовательно, не выполнены необходимые условия второго порядка. Таким образом, точка $(0, 0)$ не является точкой экстремума.

***Пример* 8** *Найти точки экстремума функции*

$$f_0\left(x,y\right)=x^4+y^4-2x^2-2y^2.$$

Решение. Применим теорему о необходимых условиях первого порядка. Найдем стационарные точки функции $f_0\left(x,y\right)$:

$$f_0'\left(x,y\right)=\left(4x^3-4x,4y^3-4y\right);$$

$$f_0'\left(x,y\right)=0\Leftrightarrow\left\{\begin{array}{l}4x\left(x^2-1\right)=0\\4y\left(y^2-1\right)=0\end{array}\right.. \qquad (10)$$

Системе (10) удовлетворяют 9 точек:

$$\left(0,0\right),\ \left(0,\pm1\right),\ \left(1,0\right),\ \left(1,\pm1\right),\ \left(-1,\pm1\right).$$

Проверим выполнение необходимых условий второго порядка и достаточных условий. Вычислим матрицу вторых производных:

$$f_0''\left(x,y\right)=\left(\begin{array}{cc}12x^2-4&0\\0&12y^2-4\end{array}\right).$$

Проверим точку $\left(0,0\right)$:

$$f_0''\left(0,0\right)=\left(\begin{array}{cc}-4&0\\0&-4\end{array}\right).$$

По критерию Сильвестра матрица $f_0''\left(0,0\right)$ является отрицательно определенной, следовательно, $\left(0,0\right)$ — строгий локальный максимум.

Для точек $\left(0,\pm1\right)$:

$$f_0''\left(0,\pm1\right)=\left(\begin{array}{cc}-4&0\\0&8\end{array}\right).$$

По критерию Сильвестра матрица $f_0''\left(0,\pm1\right)$ является знакопеременной, следовательно, точки $\left(0,\pm1\right)$ не являются локальными экстремумами.

Для точек $\left(\pm1,0\right)$:

$$f_0''\left(\pm1,0\right)=\left(\begin{array}{cc}8&0\\0&-4\end{array}\right).$$

По критерию Сильвестра матрица $f_0''\left(\pm1,0\right)$ является знакопеременной, следовательно, $\left(\pm1,0\right)$ не являются локальными экстремумами.

Для точек $\left(\pm1,\pm1\right)$ и $\left(\pm1,\mp1\right)$:

$$f_0''\left(\pm1,\pm1\right)=f_0''\left(\pm1,\mp1\right)=\left(\begin{array}{cc}8&0\\0&8\end{array}\right).$$

По критерию Сильвестра матрицы $f_0''(\pm 1, \pm 1)$ и $f_0''(\pm 1, \mp 1)$ положительно определены, следовательно, $(\pm 1, \pm 1)$ и $(\pm 1, \mp 1)$ — строгие локальный минимумы.

Функция $f_0(x, y)$ является бесконечно растущей. Действительно, покажем, что $f_0(x, y) \to +\infty$, если $x^2 + y^2 \to +\infty$. Можем записать

$$f_0(x, y) = x^4 + y^4 - 2x^2y^2 + 2x^2y^2 - 2x^2 - 2y^2 = \left(x^2 + y^2\right)^2 - 2\left(x^2 + y^2\right) - 2x^2y^2 \geq$$

$$\geq \left(x^2 + y^2\right)^2 - \frac{5}{2}\left(x^2 + y^2\right),$$

при оценке использовали следущий факт: $|ab| \leq \frac{1}{2}\left(a^2 + b^2\right)$. Отсюда имеем $f_0(x, y) \to +\infty$ при $x^2 + y^2 \to +\infty$. Таким образом, у функции f_0 есть глобальный минимум. Выясним, какой из локальных минимумов является глобальным. Вычислим значения функции f_0 в точках $(\pm 1, \pm 1)$ и $(\pm 1, \mp 1)$:

$$f_0(\pm 1, \pm 1) = f_0(\pm 1, \mp 1) = -2,$$

следовательно, $(\pm 1, \pm 1)$ и $(\pm 1, \mp 1)$ — глобальные минимумы.

Локальный максимум будем проверять по определению. Рассмотрим приращение функции f_0 в нуле:

$$\triangle f_0 = f_0(\triangle x, \triangle y) - f_0(0, 0) = (\triangle x)^4 + (\triangle y)^4 - 2(\triangle x)^2 - 2(\triangle y)^2,$$

при малых $\triangle x$, $\triangle y$ имеем: $\triangle f_0 < 0$, но при больших $\triangle x$, $\triangle y$ приращение $\triangle f_0 > 0$, следовательно, $(0, 0)$ — локальный, но не глобальный максимум.

3 Задачи условной оптимизации

Задача (2) называется задачей условной оптимизации, если множество D является собственным подмножеством пространства $\mathbb{R}^n$, т.е. $D \neq \emptyset$ и $D \neq \mathbb{R}^n$. Для задачи условной оптимизации верны утверждения теорем раздела 2, если локальное решение является внутренней точкой множества D. Если же решение лежит на границе D, то эти утверждения перестают быть верными.

Дадим геометрическую интерпретацию задачи условной оптимизации. Напомним определение линии (или поверхности) уровня функции f_0.

Линией (поверхностью) уровня α функции f_0 назовем множество точек в $\mathbb{R}^n$ вида

$$L_\alpha = \{x \,|\, f_0(x) = \alpha\},\ \alpha \in \mathbb{R}.$$

Построим допустимое множество D и несколько линий уровня целевой функции. Полезно учитывать следующее. Если функция f_0 дифференцируема в точке x, и ее градиент в этой точке отличен от нуля, то он (градиент) ортогонален к проходящей через точку x линии уровня и направлен

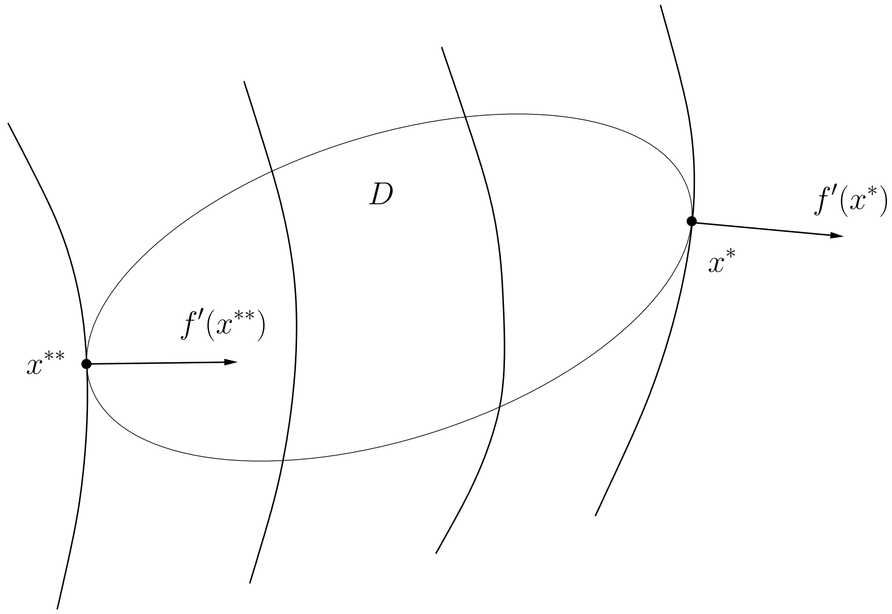

Рис. 3: x^{**} — абсолютный минимум функции $f\,(x)$ на множестве D, x^* — абсолютный максимум.

в сторону возрастания целевой функции. Поиск минимума сводится к нахождению минимального числа α^* среди всех таких α, что линия уровня L_α имеет непустое пересечение с множеством D. При этом минимумом в задаче (2) является любая точка множества $x^* \in L_{\alpha^*} \bigcap D$. Аналогично для максимума: находим максимальное число α^{**} среди всех таких α, что линия уровня L_α имеет непустое пересечение с множеством D. При этом максимумом в задаче (2) является любая точка множества $x^{**} \in L_{\alpha^{**}} \bigcap D$. Заметим, что таким образом мы находим глобальные решения (см.рис.3).

***Пример* 9** [1]

$$\begin{aligned}
&f_0 = -x^2 + 6x + y \to \max,\\
&2x + 3y \le 24, \quad x + 2y \le 15,\\
&3x + 2y \le 24, \quad y \le 4, \quad x \ge 0, \quad y \ge 0.
\end{aligned}$$

Решение. Множество допустимых точек — многоугольник $OABD$ (рис. 4). Чтобы решить задачу, надо найти точки многоугольника $OABD$, в которых целевая функция принимает наибольшее значение. Построим линии уровня целевой функции:

$$f_0(x, y) = -x^2 + 6x + y = C,$$

где C — некоторая константа, и исследуем поведение линий уровня при различных значениях C. При каждом фиксированном значении C получаем параболу; чем больше значение C, тем выше расположена парабола.

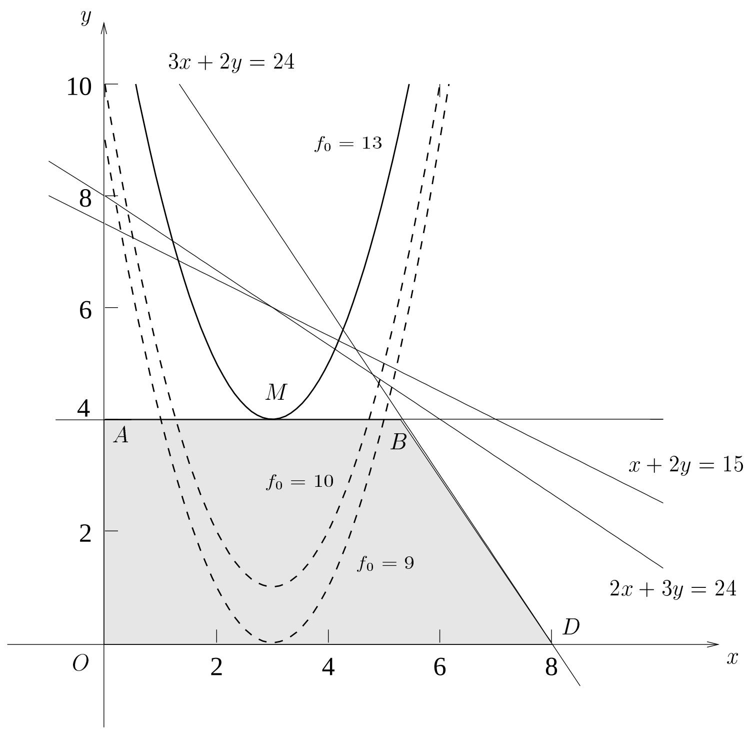

Рис. 4: к примеру 9

Значит, функция принимает максимальное значение в точке касания одной из парабол с границей многоугольника $OABD$. В данном случае это точка M, в которой линия уровня $-x^2+6x+y=13$ касается стороны многоугольника AB. Координаты точки M можно найти из системы уравнений

$$y - x^2 + 6x = 13, \quad y = 4.$$

Решая эту систему, получим $x^* = 3, \quad y^* = 4$ и $f_{0\,\max} = 13$.

***Пример* 10** [1]

$$\begin{aligned}
&f_0 = (x-3)^2 + (y-4)^2 \to \text{extr}\\
&3x + 2y \geq 7, \quad 10x - y \leq 8,\\
&-18x + 4y \leq 12, \quad x \geq 0, \quad y \geq 0.
\end{aligned}$$

Решение. Множество допустимых точек — треугольник ABD. Построим линию уровня целевой функции: $f_0 = (x-3)^2 + (y-4)^2 = C$, $C \geq 0$. Линии уровня — окружности с центром в т. $E(3,4)$ и радиусом $\sqrt{C}$. С увеличением (уменьшением) числа C значения функции f_0 соответственно увеличиваются (уменьшаются).

Видно (рис. 5), что минимальное значение целевая функция принимает в точке $M = (x^*, y^*)$, в которой окружность касается допустимого

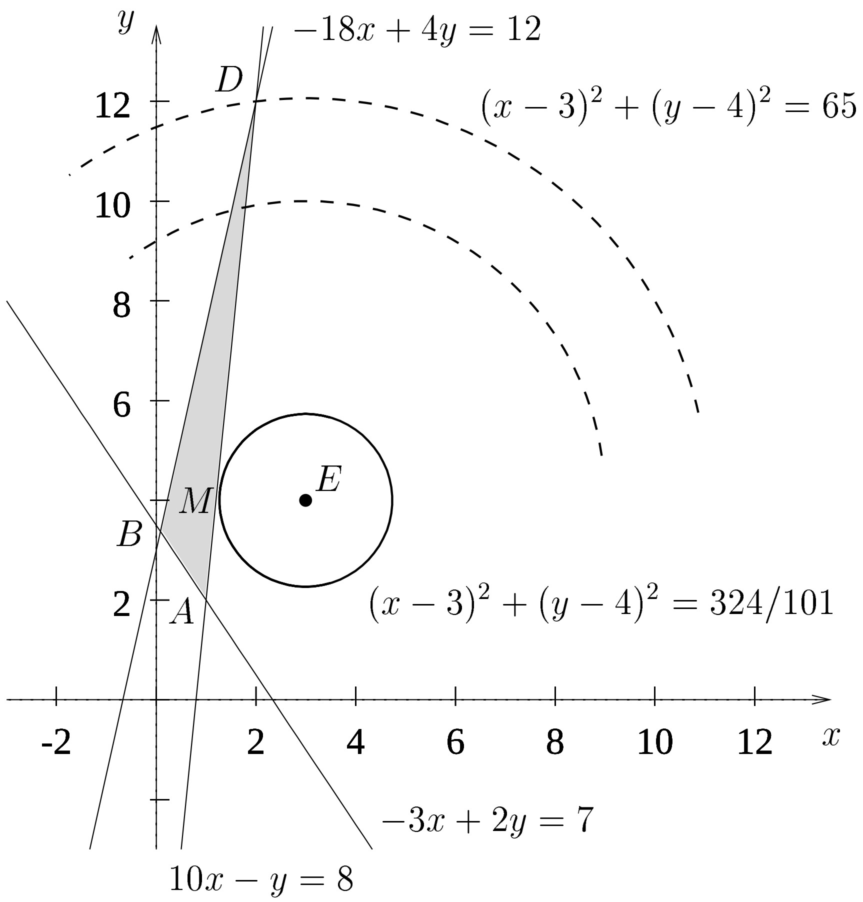

Рис. 5: к примеру 10

множества. Так как окружность касается в точке M прямой $10x - y = 8$, то угловые коэффициенты этой прямой и касательной к окружности в точке M равны. Угловой коэффициент прямой равен 10. Угловой коэффициент касательной к окружности в точке M определим как значение производной функции y от переменной x в этой точке. Считаем, что y является неявной функцией переменной x, дифференцируем уравнение окружности по переменной x, получим $2(x - 3) + 2(y - 4)y' = 0$. Отсюда $y' = -(x - 3)/(y - 4)$ — угловой коэффициент наклона касательной к окружности в точке (x, y). Таким образом, $-(x^* - 3)/(y^* - 4) = 10$. Так как точка M лежит на прямой $10x - y = 8$, то получаем второе уравнение для определения координат (x^*, y^*).

Решая систему

$$\begin{cases} -(x^* - 3)/(y^* - 4) = 10 \\ 10x^* - y^* = 8, \end{cases}$$

находим $x^* = 123/101, \quad y^* = 422/101, \quad f_{0\,\min} = 324/101.$

На рисунке 5 видно, что целевая функция принимает максимальное значение в точке $N = (2, 12); \quad f_{0\,\max} = 65.$

4 Задача с ограничениями в виде равенств

4.1 Метод множителей Лагранжа. Необходимые и достаточные условия оптимальности

В этом разделе мы будем рассматривать задачу условной оптимизации, в которой допустимое множество задается системой конечного числа уравнений: $D = \{x \in \mathbb{R}^n \,|\, f_i(x) = b_i,\, i = 1, ..., m\}\,,\ (m < n)\,.$ Обычно задача с ограничениями в виде равенств записывается в виде

$$\begin{cases} f_0(x) \to extr \\ f_i(x) = b_i,\ i = 1, ..., m. \end{cases} \tag{11}$$

Для исследования задач с ограничениями в виде равенств эффективным является метод Лагранжа, который заключается в следующем. Составляется функция Лагранжа

$$L\,(x, \lambda_0, ..., \lambda_m) = \lambda_0 f_0\,(x) + \sum_{i=1}^{m} \left(f_i\,(x) - b_i\right)$$

и рассматривается задача безусловной оптимизации

$$L\,(x, \lambda_0, ..., \lambda_m) \to extr,$$

в которой переменными являются x, а набор $\lambda = (\lambda_0, ..., \lambda_m)$ — неизвестными параметрами.

Теорема 9 *(Необходимое условие локального экстремума первого порядка в задаче с ограничениями в виде равенств)*

Пусть $f_i : \mathbf{R}^n \to \mathbf{R}^1 \quad (i = 0, ..., m)$ *— непрерывно дифференцируемые в окрестности точки* x^* *функции. Если* x^**— локальный экстремум в задаче (11), то найдется ненулевой набор множителей Лагранжа* $\lambda^* = (\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*) \in \mathbf{R}^{m+1}$*, такой, что для функции Лагранжа выполняется условие стационарности по* x*:*

$$\mathcal{L}'_x(x^*, \lambda^*) = 0,$$

то есть, $\dfrac{\partial}{\partial x_j}\mathcal{L}(x^*, \lambda^*) = 0 \quad (j = 1, ..., n)$.

Если векторы $f'_1(x^*), ..., f'_m(x^*)$ *линейно независимы, то* $\lambda_0^* \neq 0$.

Замечание 6 *Условия, которые гарантируют, что в задаче оптимизации* $\lambda_0^* \neq 0$*, называются условиями регулярности. То есть в задаче с ограничениями в виде равенств условие регулярности заключается в линейной независимости градиентов функций, задающих ограничения. Иной вид условий регулярности приведем далее для задачи выпуклого программирования.*

Доказательство теоремы 9. Рассмотрим векторы $f'_0(x^*)$, $f'_1(x^*)$, $\dots$, $f'_m(x^*)$. Возможны 2 ситуации: либо эти векторы линейно независимы, либо они линейно зависимы. В последнем случае выполнение теоремы следует из определения линейной зависимости векторов. Предположим, что векторы $f'_0(x^*)$, $f'_1(x^*), ..., f'_m(x^*)$ линейно независимы. Составим матрицу

$$\begin{pmatrix} \frac{\partial f_0(x^*)}{\partial x_1} & \cdots & \frac{\partial f_0(x^*)}{\partial x_n} \\ \vdots & \ddots & \\ \frac{\partial f_m(x^*)}{\partial x_1} & \cdots & \frac{\partial f_m(x^*)}{\partial x_n} \end{pmatrix}.$$

Так как векторы $f'_0(x^*)$, $f'_1(x^*), ..., f'_m(x^*)$ линейно независимы, то существует подматрица $M_{(m+1)\times(m+1)}$, определитель которой не равен нулю (см. приложение, теорема о ранге матрицы). Предположим для простоты, что M образована первыми $(m+1)$ столбцами, то есть

$$M = \begin{pmatrix} \frac{\partial f_0(x^*)}{\partial x_1} & \cdots & \frac{\partial f_0(x^*)}{\partial x_{m+1}} \\ \vdots & \ddots & \\ \frac{\partial f_m(x^*)}{\partial x_1} & \cdots & \frac{\partial f_m(x^*)}{\partial x_{m+1}} \end{pmatrix}.$$

Если это не так, то можно перенумеровать переменные таким образом, чтобы это предположение выполнялось. Кроме того, можно считать, что

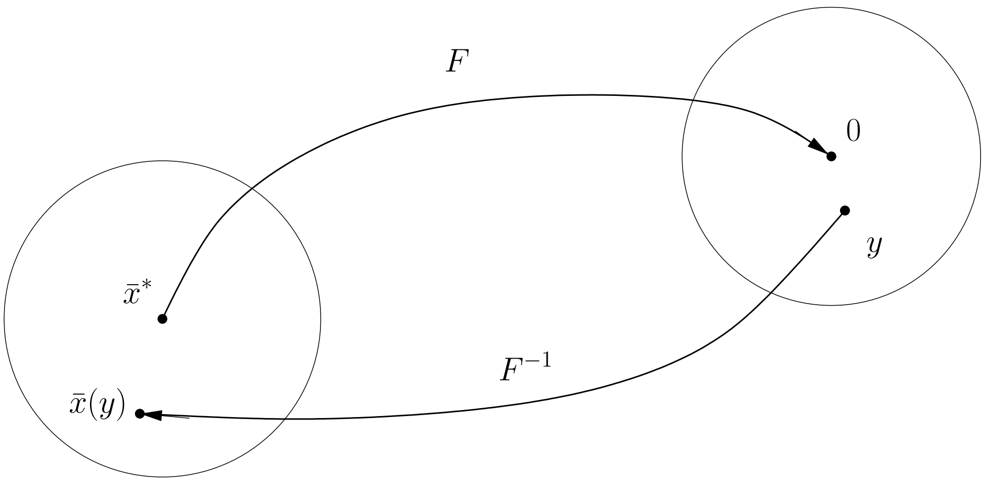

Рис. 6: применение теоремы об обратном отображении

$f_0\left(x^*\right)=0$, так как в противном случае можно рассмотреть новую целевую функцию $f_0\left(x\right)-f_0\left(x^*\right)$. Обозначим $\bar{x}=\left(x_1,...,x_{m+1},x^*_{m+2},...,x^*_n\right)$ и рассмотрим отображение $F\left(\bar{x}\right)$:

$$F\left(\bar{x}\right):\,\mathbf{R}^{m+1}\to\mathbf{R}^{m+1},\quad F\left(\bar{x}\right)=\left(f_0\left(\bar{x}\right),f_1\left(\bar{x}\right)-b_1,...,f_m\left(\bar{x}\right)-b_m\right).$$

Тогда $\bar{x}^*=x^*$ и $F\left(\bar{x}^*\right)=0$. Так как $det\,F'\left(\bar{x}^*\right)=det\,M\neq 0$, то к отображению F можно применить теорему об обратном отображении (см. приложение). Тогда существуют положительные постоянные ε, δ и κ, такие, что для любого y из ε-окрестности нуля в пространстве $\mathbf{R}^{m+1}$ существует единственная точка $\bar{x}\left(y\right)$ из δ-окрестности точки $\bar{x}^*$ в пространстве $\mathbf{R}^{m+1}$, для которой $F(\bar{x}(y))=y$ и $||\bar{x}(y)-\bar{x}^*||\leq\kappa||y||$.

Рассмотрим точки $y_{\pm s}=(\pm\varepsilon/s,0,...,0)$, где $s\in\mathbb{N}$. По сформулированному выше утверждению найдутся точки $\bar{x}_{\pm s}=\bar{x}\left(y_{\pm s}\right)$, для которых

$$F\left(\bar{x}_{\pm s}\right)=\left(\pm\frac{\varepsilon}{s},0,...,0\right)\tag{12}$$

и

$$\|\bar{x}_{\pm s}-\bar{x}^*\|\leq\kappa\frac{\varepsilon}{s}.\tag{13}$$

Из (12) получаем:

$$\left\{\begin{array}{c} f_0\left(\bar{x}_{\pm s}\right)=\pm\frac{\varepsilon}{s},\\ f_i\left(\bar{x}_{\pm s}\right)=b_i\,,\quad i=1,...,m.\end{array}\right.$$

Таким образом, точки $\bar{x}_{\pm s}=\left(\bar{x}_{\pm s,1},...,\bar{x}_{\pm s,m+1},x^*_{m+2},...,x^*_n\right),\quad s\in\mathbb{N}$, являются допустимыми, а из (13) следует, что они сколь угодно близки к точке x^*. Кроме того, из (12) имеем следующие оценки:

$$f_0\left(\bar{x}_{+s}\right)=\frac{\varepsilon}{s}>0=f_0(x^*)>f_0\left(\bar{x}_{-s}\right)=-\frac{\varepsilon}{s}.$$

Следовательно, x^* не является локальным экстремумом. Получили противоречие с условием теоремы, поэтому предположение о том, что векторы $f_0'(x^*)$, $f_1'(x^*), ..., f_m'(x^*)$ линейно независимы неверно, т.е. найдется ненулевой набор чисел $\lambda_0^*, \lambda_1^*, ..., \lambda_m^*$, такой, что

$$\sum_{j=0}^{m} \lambda_j^* \, f_j\,'(x^*) = 0.$$

Доказательство условия регулярности. Предположим, что векторы $f_1'(x^*), ..., f_m'(x^*)$ линейно независимы и $\lambda_0^* = 0$. Так как $x^* \in locextr$, то имеем

$$\mathcal{L}_x'\left(x^*, \lambda^*\right) = \sum_{j=1}^{m} \lambda_j^* \, f_j\,'(x^*) = 0 \tag{14}$$

и среди чисел $\lambda_1^*, ..., \lambda_m^*$ есть хотя бы одно ненулевое. Но условие (14) означает, что векторы $f_1'(x^*), ..., f_m'(x^*)$ линейно зависимы. Противоречие. Следовательно, $\lambda_0^* \neq 0$. □

Замечание 7 *Отметим, что множители Лагранжа определены с точностью до ненулевого множителя. Именно, если набор $(\lambda_0^*, ..., \lambda_m^*)$ удовлетворяет условиям теоремы, то для любого $\alpha \neq 0$ набор $(\alpha\lambda_0^*, ..., \alpha\lambda_m^*)$ также удовлетворяет условиям теоремы. Поэтому, если доказано, что $\lambda_0^* \neq 0$, то в условиях теоремы 9 можно рассматривать набор множителей Лагранжа с $\lambda_0^* = 1$.*

Теорема 10 *(Необходимое условие локального минимума 2-го порядка в задаче с ограничениями в виде равенств)*

Пусть функции $f_0, f_1, \ldots, f_m$ дважды непрерывно дифференцируемы в окрестности допустимой точки $x^* \in \mathbf{R}^n$. Пусть $f_1^{'}\left(x^*\right), ..., f_m^{'}\left(x^*\right)$ линейно независимы. Если x^*— локальный минимум в задаче (11), то существует набор множителей Лагранжа $\lambda^* = (1, \lambda_1^*, \ldots, \lambda_m^*)$, такой, что $\mathcal{L}_x'\left(x^*, \lambda^*\right) = 0$ и

$$\left(\mathcal{L}_{xx}''\left(x^*, \lambda^*\right) h, h\right) \geq 0 \tag{15}$$

для всех $h \in H\left(x^*\right) = \left\{h \in \mathbf{R}^n \left| (f_i'\left(x^*\right), h) = 0, \quad i = 1, \ldots m\right.\right\}.$

Доказательство. Положим $g_i\left(x\right) = f_i\ \left(x\right) - b_i$, $i\ = 1, ..., m$. Для таким образом определенных функций $g_1\left(x\right), ..., g_m\left(x\right)$ в точке x^* выполнены условия теоремы Люстерника (см. приложение). Тогда для всех достаточно малых по модулю $\alpha \in \mathbf{R}$ существует $r\ (\alpha) \in \mathbf{R}^n$, такое, что для всех $h \in H\left(x^*\right)$ выполняется

$$0 = g_i\left(x^* + \alpha h + r\left(\alpha\right)\right) = f_i\left(x^* + \alpha h + r\left(\alpha\right)\right) - b_i\ \left(i\ = 1, ..., m\right).$$

То есть $x^* + \alpha h + r(\alpha)$ — допустимая точка в задаче (11).

Рассмотрим функцию Лагранжа в точках x^* и $x^* + \alpha h + r(\alpha)$:

$$\mathcal{L}(x^*, \lambda^*) = f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* (f_i(x^*) - b_i) = f_0(x^*),$$

$$\mathcal{L}(x^* + \alpha h + r(\alpha), \lambda^*) =$$

$$= f_0(x^* + \alpha h + r(\alpha)) + \sum_{i=1}^{m} \lambda_i^* (f_i(x^* + \alpha h + r(\alpha)) - b_i) =$$

$$= f_0(x^* + \alpha h + r(\alpha)).$$

Так как $x^* \in locmin(11)$, то для достаточно малых α имеем

$$f_0(x^* + \alpha h + r(\alpha)) \geq f_0(x^*).$$

Таким образом,

$$0 \leq f_0(x^* + \alpha h + r(\alpha)) - f_0(x^*) = \mathcal{L}(x^* + \alpha h + r(\alpha), \lambda^*) - \mathcal{L}(x^*, \lambda^*) =$$

$$= (\mathcal{L}'_x(x^*, \lambda^*), \alpha h + r(\alpha)) + \tfrac{1}{2} (\mathcal{L}''_{xx}(x^*, \lambda^*)(\alpha h + r(\alpha)), (\alpha h + r(\alpha))) + o(\alpha^2) =$$

$$= \tfrac{1}{2} (\mathcal{L}''_{xx}(x^*, \lambda^*)(\alpha h + r(\alpha)), (\alpha h + r(\alpha))) + o(\alpha^2).$$

Разделим все на α^2 и рассмотрим предел при $\alpha \to 0$. Получим

$$(\mathcal{L}''_{xx}(x^*, \lambda^*) h,\, h) \geq 0 \quad \forall h \in \mathbf{R}^n : \ (f_i'(x^*), h) = 0, \quad i = 1, ..., m.$$

Теорема доказана. □

Замечание 8 *Если x^* — локальный максимум, то условие (15) в теореме заменяется на $(\mathcal{L}''_{xx}(x^*, \lambda^*) h, h) \leq 0$.*

Замечание 9 *Если найдутся такие h^1, $h^2 \in H(x^*)$, что*

$$\left(\mathcal{L}''_{xx}(x^*, \lambda^*) h^1, h^1\right) > 0, \quad \left(\mathcal{L}''_{xx}(x^*, \lambda^*) h^2, h^2\right) < 0,$$

то x^* не является локальным экстремумом.

Теорема 11 *(Достаточные условия локального минимума в задаче с ограничениями в виде равенств)*

Пусть функции $f_0, f_1, \ldots, f_m$ дважды непрерывно дифференцируемы в некоторой окрестности допустимой точки $x^* \in \mathbf{R}^n$. Если существует набор множителей Лагранжа $\lambda^* = (1, \lambda_1^*, \ldots, \lambda_m^*)$, для которого выполнено:

а) $\mathcal{L}'_x(x^*, \lambda^*) = 0$;

б) для всех $h \in H(x^*)$, $\quad h \neq 0$,

$$(\mathcal{L}''_{xx}(x^*, \lambda^*) h, h) > 0, \tag{16}$$

то x^* — строгий локальный минимум задачи.

Доказательство. Предположим, что выполнены условия теоремы, но x^* не является строгим локальным минимумом, т.е. существует последовательность допустимых точек $\left\{x^{(k)}\right\}_{k=1}^{\infty}$, такая, что: $x^k \neq x^*$, $x^{(k)} \to x^*$ при $k \to +\infty$ и $f_0\left(x^{(k)}\right) \leq f_0\left(x^*\right)$. Обозначим

$$x^{(k)} - x^* = \alpha_k h^{(k)}, \quad \alpha_k = \left\|x^{(k)} - x^*\right\|, \quad h^{(k)} = \alpha_k^{-1}\left\|x^{(k)} - x^*\right\|,$$

$$\alpha_k \in \mathbf{R}, \quad h^{(k)} \in \mathbf{R}^n.$$

Тогда $h^{(k)} \to h^*$, и так как $\left\|h^{(k)}\right\| = 1$, то $h^* \neq 0$. (Вернее, в силу ограниченности последовательности $\left\{h^{(k)}\right\}_{k=1}^{\infty}$ существует сходящаяся подпоследовательность $\left\{h^{(k_m)}\right\}_{m=1}^{\infty}$: $h^{(k_m)} \to h^*, k \to +\infty$. Но далее мы не будем писать двойной индекс, подразумевая именно сходящуюся подпоследовательность.)

Покажем, что $\left(f_i'\left(x^*\right), h^*\right) = 0 \quad (i = 1, \ldots, m)$. Так как $x^{(k)}$ — допустимая точка для всех $k \geq 1$, то $f_i\left(x^{(k)}\right) = b_i\,(i = 1, \ldots, m)$. Отсюда имеем

$$0 = f_i\left(x^{(k)}\right) - f_0\left(x^{(k)}\right) = \left(f_i'\left(x^*\right), \alpha_k h^{(k)}\right) + o\left(\alpha_k\right)(i = 1, \ldots, m).$$

Разделим равенство на α_k и рассмотрим предел при $k \to +\infty$:

$$\left(f_i'\left(x^*\right), h^{(k)}\right) + \frac{o\left(\alpha_k\right)}{\alpha_k} = 0,$$

$$\left(f_i'\left(x^*\right), h^*\right) = 0, \quad i = 1, \ldots, m.$$

Так как $x^*, \quad x^{(k)}\,(k = 1, \ldots)$ — допустимые точки, то можно показать, что

$$\mathcal{L}\left(x^*, \lambda^*\right) = f_0\left(x^*\right), \quad \mathcal{L}\left(x^{(k)}, \lambda^*\right) = f_0\left(x^{(k)}\right).$$

Отсюда, используя предположение о том, что x^* не является строгим локальным минимумом, имеем

$$0 \geq f_0\left(x^{(k)}\right) - f_0\left(x^*\right) = \mathcal{L}\left(x^{(k)}, \lambda^*\right) - \mathcal{L}\left(x^*, \lambda^*\right) =$$

$$= \left(\mathcal{L}_x'\left(x^*, \lambda^*\right), \alpha_k h^{(k)}\right) + \frac{1}{2}\left(\mathcal{L}_{xx}''\left(x^*, \lambda^*\right)\alpha_k h^{(k)}, \alpha_k h^{(k)}\right) + o\left(\alpha_k^2\right).$$

Разделим обе части неравенства на α_k^2 и рассмотрим предел при $k \to +\infty$:

$$0 \geq \frac{1}{2}\left(\mathcal{L}_{xx}''\left(x^*, \lambda^*\right) h^*, h^*\right).$$

Получили противоречие с условием теоремы. Следовательно, точка x^* — строгий локальный минимум. Теорема доказана. □

Замечание 10 *Если условие (16) в теореме выполнено в виде*

$$(\mathcal{L}_{xx}''(x^*, \lambda^*)h, h) < 0, \ \forall h \in H\left(x^*\right), \ h \neq 0,$$

то x^* — строгий локальный максимум.

4.2 Схема решения конечномерной задачи с ограничениями в виде равенств

1. Составить функцию Лагранжа:
$$\mathcal{L}\left(x,\lambda\right)=\lambda_0 f_0\left(x\right)+\sum_{i=1}^{m}\lambda_i\left(f_i(x)-b_i\right).$$

2. Выписать необходимое условие локального экстремума:
$$\mathcal{L}'_x\left(x,\lambda\right)=0\Leftrightarrow\sum_{i=0}^{m}\lambda_i f'_i\left(x\right)=0.$$

3. Найти стационарные точки, т.е. допустимые решения пункта 2, для которых не все множители Лагранжа равны 0. При этом удобно отдельно рассмотреть случаи $\lambda_0=0$ и $\lambda_0\neq 0$. Во втором случае можно положить λ_0 равным любой положительной константе. (Иногда проще проверить выполнение условия регулярности, чем непосредственно решать систему пункта 2 при $\lambda_0=0$).

4. Определить, используя достаточные условия локального экстремума, какая из найденных стационарных точек является локальным минимумом или локальным максимумом.

5. Выяснить, есть ли глобальные минимум и максимум в данной задаче. Можно применить теоремы существования глобального экстремума или же провести проверку по определению.

***Пример* 11**

$$\begin{gathered}f_0\left(x,y,z\right)=xyz\ \to\ \mathrm{extr}\\ 2xy+yz\ =\ 12,\quad 2x-y\ =\ 8\end{gathered}$$

Решение. Применим теорему о необходимых условиях локального экстремума в задачах с ограничениями в виде равенств. Выпишем функцию Лагранжа

$$\mathcal{L}=\lambda_0 xyz+\lambda_1\left(2xy+yz-12\right)+\lambda_2\left(2x-y-8\right)$$

и условия стационарности:

$$\begin{array}{rcl}\mathcal{L}'_x & = & \lambda_0 yz+2\lambda_1 y+2\lambda_2\ =\ 0\\ \mathcal{L}'_y & = & \lambda_0 xz+2\lambda_1 x+\lambda_1 z-\lambda_2\ =\ 0\\ \mathcal{L}'_z & = & \lambda_0 xy+\lambda_1 y\ =\ 0\end{array}\qquad(17)$$

Проверим условие регулярности, то есть линейную независимость векторов $f_1'(x,y,z)$ и $f_2'(x,y,z)$, где $f_1(x,y,z) = 2xy + yz$, $f_2(x,y,z) = 2x - y$.

$$f_1'(x,y,z) = (2y, 2x+z, y)\,,\ f_2'(x,y,z) = (2,-1,0)\,.$$

Предположим, что $f_1'(x,y,z)$ и $f_2'(x,y,z)$ линейно зависимы, т.е. существуют такие, не равные нулю одновременно, константы C_1 и C_2, что

$$C_1 f_1'(x,y,z) + C_2 f_2'(x,y,z) = 0.$$

$$\left\{\begin{array}{rcl} C_1\cdot 2y + 2C_2 &=& 0\\ C_1\cdot(2x+z) - C_2 &=& 0\\ C_1 y &=& 0\end{array}\right. \iff \left\{\begin{array}{rcl} C_2 &=& 0\\ C_1\cdot(2x+z) &=& 0\\ C_1 y &=& 0\end{array}\right.$$

Из 1-го уравнения $C_2 = 0$, поэтому $C_1 \neq 0 \quad\Rightarrow\quad y = 0$, но на допустимом множестве $y \neq 0$, так как $f_1(x,0,z) = 0 \neq 12 \quad\Rightarrow\quad f_1'(x,y,z)$ и $f_2'(x,y,z)$ линейно независимы в любой точке допустимого множества $\quad\Rightarrow\quad \lambda_0 \neq 0$. Положим $\lambda_0 = 1$. Система (17) примет вид

$$\left\{\begin{array}{rcl} yz + 2\lambda_1 y + 2\lambda_2 &=& 0\\ xz + 2\lambda_1 x + \lambda_1 z - \lambda_2 &=& 0\\ xy + \lambda_1 y &=& 0\\ 2x - y &=& 8\\ 2xy + yz &=& 12\end{array}\right.$$

Так как на допустимом множестве $y \neq 0$, то из 3-го уравнения получаем

$$x = -\lambda_1 \quad\Rightarrow\quad y = 2x - 8 = -2\lambda_1 - 8.$$

Далее

$$\left\{\begin{array}{rcl} z(-2\lambda_1 - 8) + 2\lambda_1(-2\lambda_1 - 8) + 2\lambda_2 &=& 0\\ -\lambda_1 z - 2\lambda_1^2 + \lambda_1 z - \lambda_2 &=& 0\\ -2\lambda_1(-2\lambda_1 - 8) + z(-2\lambda_1 - 8) &=& 12\end{array}\right..$$

Отсюда

$$\left\{\begin{array}{rcl} \lambda_2 &=& -2\lambda_1^2\\ -2\lambda_1 z - 8z - 4\lambda_1^2 - 16\lambda_1 - 4\lambda_1^2 &=& 0\\ 4\lambda_1^2 + 16\lambda_1 - 2\lambda_1 z - 8z - 12 &=& 0\end{array}\right..$$

Вычтем из второго уравнения третье. Тогда

$$12 - 4\lambda_1^2 - 16\lambda_1 = 4\lambda_1^2 + 16\lambda_1 + 4\lambda_1^2,$$

$$12\lambda_1^2 + 32\lambda_1 - 12 = 0.$$

Отсюда имеем

$$\lambda_1^{(1)} = -3, \qquad \lambda_1^{(2)} = \frac{1}{3},$$
$$\lambda_2^{(1)} = -18, \qquad \lambda_2^{(2)} = -\frac{2}{9}.$$

Находим соответственно x, y, z:

$$x^{(1)} = 3, \quad y^{(1)} = -2, \quad z^{(1)} = -12,$$
$$x^{(2)} = -\frac{1}{3}, \quad y^{(2)} = -\frac{26}{3}, \quad z^{(2)} = -\frac{28}{39}\,.$$

Применим теорему о достаточных условиях локального экстремума в задачах с ограничениями в виде равенств. Проверим знак

$$\left(\mathcal{L}''\left(x^*, y^*, z^*\right) h, h\right),$$

где $h \in H\left(x^*, y^*, z^*\right), \quad h \neq 0$, и

$$H\left(x^*, y^*, z^*\right) = \left\{h \in R^3 \mid \left(f_1'\left(x^*, y^*, z^*\right), h\right) = \left(f_2'\left(x^*, y^*, z^*\right), h\right) = 0\right\}.$$

1) Исследуем первую из найденных точек:

$$\left(x^{(1)}, y^{(1)}, z^{(1)}\right) = (3, -2, -12):$$

$$f_1'\left(x^{(1)}, y^{(1)}, z^{(1)}\right) = (-4, -6, -2), \qquad f_2'\left(x^{(1)}, y^{(1)}, z^{(1)}\right) = (2, -1, 0)$$

Для этой точки множество H состоит из векторов $h^{(1)} = \left(h_1^{(1)}, h_2^{(1)}, h_3^{(1)}\right)$, удовлетворяющих системе

$$\left\{\begin{array}{rcl} -4h_1^{(1)} - 6h_2^{(1)} - 2h_3^{(1)} & = & 0 \\ 2h_1^{(1)} - h_2^{(1)} & = & 0 \end{array}\right.$$

Отсюда $h_2^{(1)} = 2h_1^{(1)}$, $h_3^{(1)} = -8h_1^{(1)}$. Итак, $h^{(1)} = \left(h_1^{(1)}, 2h_1^{(1)}, -8h_1^{(1)}\right)$, где $h_1^{(1)} \neq 0$.

$$\mathcal{L}''_{xx}\left(x,\, y,\, z,\, \lambda_1,\, \lambda_2\right) = \begin{pmatrix} 0 & z + 2\lambda_1 & y \\ z + 2\lambda_1 & 0 & x + \lambda_1 \\ y & x + \lambda_1 & 0 \end{pmatrix},$$

$$\mathcal{L}''_{xx}\left(x^{(1)}, y^{(1)}, z^{(1)}, \lambda_1^{(1)}, \lambda_2^{(1)}\right) = \begin{pmatrix} 0 & -18 & -2 \\ -18 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix},$$

$$\mathcal{L}''_{xx}\left(x^{(1)}, y^{(1)}, z^{(1)}, \lambda_1^{(1)}, \lambda_2^{(1)}\right) h^{(1)} = \begin{pmatrix} -20h_1^{(1)} \\ -18h_1^{(1)} \\ -2h_1^{(1)} \end{pmatrix},$$

$$\left(\mathcal{L}''_{xx}\left(x^{(1)}, y^{(1)}, z^{(1)}, \lambda_1^{(1)}, \lambda_2^{(1)}\right) h^{(1)}, h^{(1)}\right) = (-20 - 36 + 16)\left(h_1^{(1)}\right)^2 < 0,$$

следовательно, $(3, -2, -12)$ — строгий $locmax$.

2) Исследуем вторую из найденных точек:

$$\left(x^{(2)}, y^{(2)}, z^{(2)}\right) = \left(-\frac{1}{3}, -\frac{26}{3}, -\frac{28}{39}\right):$$

$$f_1'\left(x^{(2)}, y^{(2)}, z^{(2)}\right) = \left(-\frac{52}{3}, -\frac{54}{39}, -\frac{26}{3}\right), \quad f_2'\left(x^{(2)}, y^{(2)}, z^{(2)}\right) = (2, -1, 0).$$

Для этой точки множество H состоит из векторов $h^{(2)} = \left(h_1^{(2)}, h_2^{(2)}, h_3^{(2)}\right)$, удовлетворяющих системе

$$\begin{cases} -\dfrac{52}{3}h_1^{(2)} - \dfrac{54}{39}h_2^{(2)} - \dfrac{26}{3}h_3^{(2)} &= 0 \\ 2h_1^{(2)} - h_2^{(2)} &= 0 \end{cases}.$$

Отсюда $h_2^{(2)} = 2h_1^{(2)}$, $h_3^{(2)} = -\dfrac{392}{169}h_1^{(2)}$. Итак, $h^{(2)} = \left(h_1^{(2)}, 2h_1^{(2)}, -\dfrac{392}{169}h_1^{(2)}\right)$, где $h_1^{(2)} \neq 0$.

$$\mathcal{L}''_{xx}\left(x^{(2)}, y^{(2)}, z^{(2)}, \lambda_1^{(2)}, \lambda_2^{(2)}\right) = \begin{pmatrix} 0 & -\dfrac{2}{39} & -\dfrac{26}{3} \\ -\dfrac{2}{39} & 0 & 0 \\ -26/3 & 0 & 0 \end{pmatrix}$$

и $\left(\mathcal{L}''_{xx}\left(x^{(2)}, y^{(2)}, z^{(2)}, \lambda_1^{(2)}, \lambda_2^{(2)}\right) h^{(2)}, h^{(2)}\right) = 40\left(h_1^{(2)}\right)^2 > 0$.

Следовательно, $(-1/3, -26/3, -28/39)$ — строгий $locmin$.

Заметим, что глобальных экстремумов нет. Например, последовательность точек $(n, 2n - 8, z)$, где z определяется из условия $2xy + yz = 12$, является допустимой, на которой $f_0(n, 2n - 8, z) = -4n^3 + 16n^2 + 12n \to -\infty$ при $n \to +\infty$. Последовательность точек $(-n, -2n - 8, z)$, где z определяется из условия $2xy + yz = 12$, также является допустимой, на которой $f_0(-n, -2n - 8, z) = 4n^3 + 16n^2 - 12n \to +\infty$ при $n \to +\infty$.

5 Задача с ограничениями в виде равенств и неравенств

5.1 Метод множителей Лагранжа. Необходимые и достаточные условия оптимальности

В этом разделе будем изучать задачу, в которой допустимое множество задается системой уравнений и неравенств:

$$\begin{aligned} & f_0(x) \to \mathrm{extr}, \\ & f_i(x) \le b_i, \quad i = 1, ..., k, \\ & f_i(x) = b_i, \quad i = k+1, ..., m. \end{aligned} \tag{18}$$

Здесь $x \in \mathbf{R}^n, \quad f_i : \mathbf{R}^n \to \mathbf{R} \quad (i = 0, \ldots, m)$.

Аналогично задаче с ограничениями в виде равенств определим функцию Лагранжа для задачи (18):

$$L(x, \lambda_0, ..., \lambda_m) = \lambda_0 f_0(x) + \sum_{i=1}^{m} (f_i(x) - b_i).$$

Сформулируем (без доказательства) необходимые условия оптимальности для задачи (18).

Теорема 12 *(Необходимые условия локального минимума в задаче с ограничениями в виде равенств и неравенств).*

Пусть допустимая точка x^* — точка локального минимума в задаче (18). Пусть функции f_i $(i = 0, ..., m)$ непрерывно дифференцируемы в окрестности точки x^*. Тогда существует ненулевой набор множителей Лагранжа $\lambda^* = (\lambda_0^*, \lambda_1^*, ..., \lambda_m^*) \in \mathbf{R}^{m+1}$, такой, что для функции Лагранжа выполняются условия:

1. *стационарности по x :*

$$\mathcal{L}'_x(x^*, \lambda^*) = 0 \Leftrightarrow \frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_j} = 0, \quad j = 1, ..., n;$$

2. *дополняющей нежесткости:* $\lambda_i^* (f_i(x^*) - b_i) = 0, \quad i = 1, ...k;$

3. *неотрицательности:* $\lambda_i^* \ge 0, \quad i = 0, ...k.$

Замечание 11 *Условие дополняющей нежесткости понимать можно следующим образом. Если в точке x^* неравенство выполнено строго: $f_i(x^*) < b_i$, то при малых $\triangle x$ в силу непрерывности функции f_i неравенство будет также выполнено: $f_i(x^* + \triangle x) < b_i$, а так как мы исследуем локальный минимум (т.е. рассматриваем малую окрестность точки x^*), то i - ое ограничение оказывается несущественным, поэтому в функцию Лагранжа оно входит с нулевым множителем: $\lambda_i^* = 0$.*

Замечание 12 *Набор множителей Лагранжа определен с точностью до положительного множителя, т.е. если набор $(\lambda_0^*, ..., \lambda_m^*)$ удовлетворяет условиям теоремы, то для любого $\alpha > 0$ набор $(\alpha\lambda_0^*, ..., \alpha\lambda_m^*)$ также удовлетворяет условиям теоремы. Поэтому, если $\lambda_0^* \neq 0$, то можно рассматривать набор множителей Лагранжа с $\lambda_0^* = 1$.*

Чтобы сформулировать достаточные условия, введем следующие обозначения. Определим $I(x^*)$ — множество активных индексов в точке x^* (номера тех неравенств, которые в точке x^* обращаются в равенство)

$$I(x^*) = \{i \,|\, 1 \leq i \leq k \,:\, f_i(x^*) = b_i\}$$

и $S(x^*)$ — множество пассивных индексов в точке x^* (номера тех неравенств, которые в точке x^* выполняются как строгие неравенства)

$$S(x^*) = \{i \,|\, 1 \leq i \leq k \,:\, f_i(x^*) < b_i\}.$$

То есть точка x^* лежит на границе области, соответствующей ограничению – неравенству с активным индексом и строго внутри области, соответствующей ограничению – неравенству с пассивным индексом.

Определим множество допустимых направлений в точке x^*:

$$\overline{H}(x^*) = \{h \in \mathbf{R}^n : (f_0'(x^*), h) \leq 0, \, (f_i'(x^*), h) \leq 0, \quad i \in I(x^*),$$
$$(f_i'(x^*), h) = 0, \quad i = k+1, ..., m\}. \tag{19}$$

Теорема 13 *(Достаточные условия локального минимума в задаче с ограничениями в виде равенств и неравенств)*

Пусть x^*— допустимая точка, $f_i(x)$ — дважды непрерывно дифференцируемые в окрестности точки x^* функции. Пусть существует набор множителей Лагранжа $\lambda^* = (1, \lambda_1^*, ..., \lambda_m^*)$, такой, что:

1. *выполняются условия теоремы о небходимых условиях локального минимума;*
2. $(\mathcal{L}_{xx}''(x^*, \lambda^*)\, h,\, h) > 0 \quad \forall h \in \overline{H}(x^*),\ h \neq 0.$

Тогда x^* — строгий локальный минимум в задаче (18).

Доказательство. Предположим, что выполнены условия теоремы, но x^* не является строгим локальным минимумом, т.е. существует последовательность $\left\{x^{(s)}\right\}_{s=1}^{\infty}$ допустимых точек, такая, что $x^{(s)} \neq x^*$ $(s = 1, \ldots)$, $x^{(s)} \to x^*$ при $s \to +\infty$ и $f_0\left(x^{(s)}\right) \leq f_0\left(x^*\right)$ для всех $s \in \mathbb{N}$. Обозначим

$$x^{(s)} - x^* = \alpha_s h^{(s)}, \quad \alpha_s = \left\|x^{(s)} - x^*\right\|, \quad h^{(s)} = \left\|x^{(s)} - x^*\right\| / \alpha_s,$$

$$\alpha \in \mathbf{R}, \quad h^{(s)} \in \mathbf{R}^n.$$

Тогда $h^{(s)} \to h^*$, и так как $\left\|h^{(s)}\right\| = 1$, то $h^* \neq 0$. (Точнее, аналогично замечанию в доказательстве теоремы о достаточных условиях для задач с ограничениями в виде равенств, в силу ограниченности последовательности $\left\{h^{(s)}\right\}_{s=1}^{\infty}$ существует сходящаяся подпоследовательность $\left\{h^{(s_l)}\right\}_{l=1}^{\infty}$: $h^{(s_l)} \to h^*$. Но для простоты записи не будем писать двойной индекс, подразумевая именно сходящуюся подпоследовательность.) Покажем,что $h^* \in \overline{H}(x^*)$. По предположению имеем

$$0 \geq f_0\left(x^{(s)}\right) - f_0\left(x^*\right) = \left(f_0'(x^*), \alpha_s h^{(s)}\right) + o(\alpha_s).$$

Разделим неравенство на α_s:

$$0 \geq \left(f_0'(x^*), h^{(s)}\right) + \frac{o(\alpha_s)}{\alpha_s}$$

и рассмотрим предел при $s \to +\infty$:

$$0 \geq \left(f_0'(x^*), h^*\right).$$

Пусть $i \in I\left(x^*\right)$, тогда $f_i\left(x^*\right) = b_i$ и $f_i\left(x^{(s)}\right) \leq b_i$, отсюда

$$0 \geq f_i\left(x^{(s)}\right) - f_i\left(x^*\right) = \left(f_i'(x^*), \alpha_s h^{(s)}\right) + o(\alpha_s).$$

Разделим неравенство на α_s и рассмотрим предел при $s \to +\infty$:

$$0 \geq \left(f_i'(x^*), h^{(s)}\right) + \frac{o(\alpha_s)}{\alpha_s},$$

$$0 \geq \left(f_i'(x^*), h^*\right), \ i \in I\left(x^*\right).$$

Так как точки $x^{(s)}$ являются допустимыми для всех $s \geq 1$, то

$$f_i\left(x^{(s)}\right) = b_i \quad (i = k+1, \ldots, m)$$

для всех $s \geq 1$. Отсюда имеем

$$0 = f_i\left(x^{(s)}\right) - f_i\left(x^*\right) = \left(f_i'(x^*), \alpha_s h^{(s)}\right) + o(\alpha_s)\left(i = k+1, \ldots, m\right).$$

Разделим равенство на α_k и рассмотрим предел при $s \to +\infty$:

$$0 = \left(f_i'(x^*), h^{(s)}\right) + \frac{o(\alpha_s)}{\alpha_s} \Rightarrow (f_i'(x^*), h^*) = 0,\ i = k+1, ..., m.$$

Итак, доказали, что $h^* \in \overline{H}(x^*)$. Покажем теперь, что матрица вторых производных не является положительно определенной на $\overline{H}(x^*)$, а именно, покажем, что

$$(\mathcal{L}''_{xx}(x^*, \lambda^*)\, h^*, h^*) \leq 0.$$

Рассмотрим функцию Лагранжа в точках (x^*, λ^*) и $\left(x^{(s)}, \lambda^*\right)$.

$$\mathcal{L}(x^*, \lambda^*) = f_0(x^*) + \sum_{i=1}^{k} \lambda_i^* (f_i(x^*) - b_i) + \sum_{i=k+1}^{m} \lambda_i^* (f_i(x^*) - b_i) = f_0(x^*).$$

Первая сумма равна нулю в силу условий дополняющей нежесткости в точке x^*, вторая сумма равна нулю, так как точка x^* является допустимой и удовлетворяет ограничениям равенствам. Для точки $x^{(s)}$ аналогичные вычисления, только необязательно выполняются условия дополняющей нежесткости:

$$\mathcal{L}\left(x^{(s)}, \lambda^*\right) = f_0\left(x^{(s)}\right) + \sum_{i=1}^{k} \lambda_i^* \left(f_i\left(x^{(s)}\right) - b_i\right) + \sum_{i=k+1}^{m} \lambda_i^* \left(f_i\left(x^{(s)}\right) - b_i\right) =$$

$$= f_0\left(x^{(s)}\right) + \sum_{i=1}^{k} \lambda_i^* \left(f_i\left(x^{(s)}\right) - b_i\right) = f_0\left(x^{(s)}\right) + S,$$

где $S = \sum_{i=1}^{k} \lambda_i^* \left(f_i\left(x^{(s)}\right) - b_i\right) \leq 0$. Отсюда, используя предположение о том, что x^* не является строгим локальным минимумом, имеем

$$0 \geq S + f_0\left(x^{(s)}\right) - f_0(x^*) = \mathcal{L}\left(x^{(s)}, \lambda^*\right) - \mathcal{L}(x^*, \lambda^*) =$$

$$= \left(\mathcal{L}'_x(x^*, \lambda^*), \alpha_s h^{(s)}\right) + \frac{1}{2}\left(\mathcal{L}''_{xx}(x^*, \lambda^*)\, \alpha_s h^{(s)}, \alpha_s h^{(s)}\right) + o\left(\alpha_s^2\right).$$

Разделим на α_s^2 и рассмотрим предел при $s \to +\infty$. Получим:

$$0 \geq \frac{1}{2}\left(\mathcal{L}''_{xx}(x^*, \lambda^*)\, h^*, h^*\right),$$

где $h^* \in \overline{H}(x^*)$, $h^* \neq 0$. Противоречие с условием теоремы. Следовательно, x^* — строгий локальный минимум. Теорема доказана. □

Для того, чтобы применить теоремы о необходимых и достаточных условиях локального минимума к задаче на максимум, надо свести ее к задаче на минимум. Рассмотрим, какие произуйдут при этом изменения. Итак, имеем задачу на максимум:

$$\begin{aligned} & f_0(x) \to \max, \\ & f_i(x) \le b_i, \quad i = 1, ..., k, \\ & f_i(x) = b_i, \quad i = k+1, ..., m. \end{aligned} \tag{20}$$

Перейдем к эквивалентной задаче на минимум:

$$\begin{aligned} & -f_0(x) \to \min, \\ & f_i(x) \le b_i, \quad i = 1, ..., k, \\ & f_i(x) = b_i, \quad i = k+1, ..., m \end{aligned} \tag{21}$$

Запишем для задачи (21) функцию Лагранжа

$$\mathcal{L}(x, \lambda) = -\lambda_0 f_0(x) + \sum_{i=1}^{m} \lambda_i (f_i(x) - b_i) \tag{22}$$

где $\lambda_0 \ge 0$. Заметим, что можно сразу же записывать функцию Лагранжа для задачи (20), но с $\lambda_0 \le 0$. И это равносильно переходу к задаче на минимум (21). Далее мы применяем теоремы о необходимых и достаточных условиях локального минимума для задачи (21) с функцией Лагранжа (22). Найденный локальный минимум будет являться локальным максимумом для исходной задачи (20).

5.2 Схема решения конечномерной задачи с ограничениями в виде равенств и неравенств

1. Составить функцию Лагранжа:

$$\mathcal{L}(x, \lambda) = \lambda_0 f_0(x) + \sum_{i=1}^{m} \lambda_i (f_i(x) - b_i).$$

2. Выписать необходимые условия локального экстремума:

$$\mathcal{L}'_x(x^*, \lambda^*) = 0 \Leftrightarrow \sum_{i=0}^{m} \lambda_i^* f_i'(x^*) = 0,$$

$$\lambda_i^* \left(f_i(x^*) - b_i\right) = 0, \quad i = 1, ..., k;$$

$$\lambda_i^* \ge 0, \quad i = 1, ..., k;$$

$\lambda_0^* \ge 0$, если находим min, и $\lambda_0^* \le 0$, если находим max .

3. Найти стационарные точки, т.е. допустимые решения пункта 2, для которых не все множители Лагранжа равны 0. При этом удобно отдельно рассмотреть случаи $\lambda_0^* = 0$ и $\lambda_0^* \neq 0$. Во втором случае можно положить λ_0^* равным какой-нибудь константе (положительной в задаче на минимум, отрицательной в задаче на максимум).

4. Определить, какая из найденных стационарных точек является локальным минимумом или локальным максимумом. Для этого можно проверить, удовлетворяют ли найденные точки достаточным условиям оптимальности или провести проверку по определению локального экстремума.

5. Выяснить, есть ли глобальные минимум и максимум в данной задаче.

***Пример* 12**

$$\begin{aligned} &29x^2 - 54xy + 26y^2 \;\to\; \text{extr} \\ &3x - 4y = 1, \quad 2x - y \le 0. \end{aligned}$$

Решение. Применим теорему о необходимых условиях локального экстремума в задачах с ограничениями в виде равенств и неравенств. Выпишем функцию Лагранжа и условия стационарности:

$$\begin{aligned} \mathcal{L} &= \lambda_0 \left(29x^2 - 54xy + 26y^2\right) + \lambda_1 (3x - 4y - 1) + \lambda_2 (2x - y) \\ \mathcal{L}'_x &= \lambda_0 (58x - 54y) + \lambda_1 \cdot 3 + \lambda_2 \cdot 2 \;=\; 0 \\ \mathcal{L}'_y &= \lambda_0 (-54x + 52y) - \lambda_1 \cdot 4 - \lambda_2 \;=\; 0, \end{aligned}$$

условие дополняющей нежесткости:

$$\lambda_2(2x - y) = 0$$

и условия неотрицательности: $\lambda_2 \ge 0$, $\lambda_0 \ge 0$ (в задаче на минимум) и $\lambda_2 \ge 0$, $\lambda_0 \le 0$ (в задаче на максимум).

Рассмотрим случай $\lambda_0 = 0$. Но тогда из условия $\mathcal{L}'_x = \mathcal{L}'_y = 0 \quad \Rightarrow$ $\lambda_1 = \lambda_2 = 0$. Т.о., при $\lambda_0 = 0$ экстремумов нет.

I) Положим $\lambda_0 = 1$ (будем искать локальные минимумы). Имеем

$$\begin{cases} 58x - 54y + 3\lambda_1 + 2\lambda_2 \;=\; 0 \\ -54x + 52y - 4\lambda_1 - \lambda_2 \;=\; 0 \\ \lambda_2 (2x - y) = 0, \quad \lambda_2 \ge 0, \\ 3x - 4y = 1, \quad 2x - y \le 0. \end{cases}$$

Складываем 1-ое и 2-ое уравнения:

$$4x - 2y - \lambda_1 + \lambda_2 = 0 \qquad \Rightarrow \qquad y = 2x - \frac{\lambda_1}{2} + \frac{\lambda_2}{2},$$

подставим полученное для y выражение во 2-ое уравнение:

$$
\begin{aligned}
-54x+52\left(2x-\frac{\lambda_1}{2}+\frac{\lambda_2}{2}\right)-4\lambda_1-\lambda_2 &= 0\\
-54x+104x-26\lambda_1+26\lambda_2-4\lambda_1-\lambda_2 &= 0\\
50x-30\lambda_1+25\lambda_2 &= 0,\\
10x-6\lambda_1+5\lambda_2 &= 0.
\end{aligned}
$$

Отсюда

$$
\begin{aligned}
x &= \frac{3}{5}\lambda_1-\frac{1}{2}\lambda_2\\
y &= 2\cdot\frac{3}{5}\lambda_1-2\cdot\frac{1}{2}\lambda_2-\frac{\lambda_1}{2}+\frac{\lambda_2}{2} =\\
&= \frac{6}{5}\lambda_1-\frac{\lambda_1}{2}-\lambda_2+\frac{\lambda_2}{2} \quad \Rightarrow\\
y &= \frac{12-5}{10}\lambda_1-\frac{\lambda_2}{2}=\frac{7}{10}\lambda_1-\frac{\lambda_2}{2}.
\end{aligned}
$$

Из условия дополняющей нежесткости следует, что либо $2x-y=0$, либо $\lambda_2=0$. Рассмотрим случай

$$
\text{i)}\ \ \lambda_2=0 \qquad \Rightarrow \qquad \begin{cases} x = \frac{3}{5}\lambda_1\\ \\ y = \frac{7}{10}\lambda_1 \end{cases}
$$

Из условия задачи $3x-4y=1 \quad \Rightarrow$

$$
\Rightarrow \quad \frac{9}{5}\lambda_1-\frac{14}{5}\lambda_1=1 \quad \Rightarrow \quad -\frac{5}{5}\lambda_1=1 \quad \Rightarrow \quad \lambda_1=-1.
$$

Находим, что $x=-3/5$, $y=-7/10$.

Проверим выполнение условия $2x-y\le 0$:

$$
-\frac{6}{5}+\frac{7}{10}=-\frac{12}{10}+\frac{7}{10}<0.
$$

Таким образом, точка $(-3/5,-7/10)$ является допустимой и удовлетворяет необходимым условиям локального минимума.

Рассмотрим случай

ii)

$$
2x-y=0 \quad \Rightarrow \quad \begin{cases} 2x-y = 0\\ 3x-4y = 1 \end{cases} \quad \Rightarrow
$$

$$
\begin{cases} y = 2x\\ 3x-8x = 1 \end{cases} \quad \Rightarrow \quad x=-\frac{1}{5}, \quad y=-\frac{2}{5}.
$$

Проверим выполнение условия неотрицательности, найдем λ_1, λ_2.

$$\begin{cases} \dfrac{3}{5}\lambda_1 - \dfrac{1}{2}\lambda_2 & = & -\dfrac{1}{5} \\ \dfrac{7}{10}\lambda_1 - \dfrac{1}{2}\lambda_2 & = & -\dfrac{2}{5} \end{cases}$$

Вычтем из 1-го уравнения 2-ое:

$$\left(\frac{3}{5} - \frac{7}{10}\right)\lambda_1 = -\frac{1}{5} + \frac{2}{5} \quad \Rightarrow \quad \lambda_1 = -2.$$

Найдем λ_2:

$$\frac{3}{5}\cdot(-2) - \frac{1}{2}\lambda_2 = -\frac{1}{5} \quad \Rightarrow \quad \lambda_2 = -2\ < 0.$$

Таким образом, не выполняется условие неотрицательности, т.е. точка $(-1/5, -2/5)$ не удовлетворяет необходимым условиям локального минимума.

II) Положим $\lambda_0 = -1$ (будем искать локальные максимумы).

$$54y - 58x + 3\lambda_1 + 2\lambda_2 = 0$$

$$54x - 52y - 4\lambda_1 - \lambda_2 = 0$$

$$\begin{cases} 2y - 4x - \lambda_1 + \lambda_2 & = & 0 \\ 54x - 52y - 4\lambda_1 - \lambda_2 & = & 0 \end{cases}$$

Отсюда $y = 2x + (\lambda_1 - \lambda_2)/2$, подставим полученное для y выражение во 2-ое уравнение:

$$\begin{aligned} 54x - 52\left(2x + \frac{\lambda_1}{2} - \frac{\lambda_2}{2}\right) - 4\lambda_1 - \lambda_2 &= 0 \\ 54x - 104x - 26\lambda_1 + 26\lambda_2 - 4\lambda_1 - \lambda_2 &= 0 \\ -50x - 30\lambda_1 + 25\lambda_2 &= 0 \\ -10x - 6\lambda_1 + 5\lambda_2 &= 0. \end{aligned}$$

Отсюда

$$\begin{aligned} x &= -\frac{3}{5}\lambda_1 + \frac{1}{2}\lambda_2 \\ y &= -\ 2\cdot\frac{3}{5}\lambda_1 + 2\cdot\frac{1}{2}\lambda_2 + \frac{\lambda_1}{2} - \frac{\lambda_2}{2} = \\ &= -\frac{6}{5}\lambda_1 + \frac{\lambda_1}{2} + \lambda_2 - \frac{\lambda_2}{2} = -\frac{7}{10}\lambda_1 + \frac{\lambda_2}{2}. \end{aligned}$$

Рассмотрим случай

i) $\lambda_2 = 0, x = -3\lambda_1/5, y = -7\lambda_1/10$. Из условия $3x - 4y = 1$ получаем, что

$$-\frac{9}{5}\lambda_1 + \frac{28}{10}\lambda_1 = 1 \quad \Rightarrow \quad \lambda_1 = 1.$$

$$x = -\frac{3}{5}, \quad y = -\frac{7}{10}.$$

Осталось рассмотреть случай

ii) $2x - y = 0 \quad \Rightarrow \quad x = -1/5, \quad y = -2/5$. Вычислим соответствующие значения λ_1 и λ_2 :

$$\begin{cases} -\dfrac{3}{5}\lambda_1 + \dfrac{1}{2}\lambda_2 = -\dfrac{1}{5} \\[2ex] -\dfrac{7}{10}\lambda_1 + \dfrac{1}{2}\lambda_2 = -\dfrac{2}{5} \end{cases}$$

Решая систему, получим $\lambda_1 = \lambda_2 = 2 > 0$. Таким образом, точка $(-1/5, -2/5)$ удовлетворяет необходимым условиям локального максимума.

Проверим, являются ли найденные точки экстремумами в задаче. Проведем проверку двумя способами: по определению и по теореме о достаточных условиях.

1. Исследуем найденные точки, используя теорему о достаточных условиях локального минимума. Проверим знак выражения

$$\left(\mathcal{L}''\left(x^*, y^*, \lambda_0, \lambda_1, \lambda_2\right) h, h\right),$$

где $h \in \overline{H}\left(x^*, y^*\right)$ и $\overline{H}\left(x^*, y^*\right)$ определяется формулой (19). Здесь

$$f_0\left(x, y\right) = 29x^2 - 54xy + 26y^2, \quad f_1\left(x, y\right) = 3x - 4y, \quad f_2\left(x, y\right) = 2x - y.$$

Вычислим градиенты функций $f_0\left(x, y\right)$, $f_1\left(x, y\right)$ и $f_2\left(x, y\right)$.

$$f_0'\left(x, y\right) = \left(58x - 54y, -54x + 52y\right), \quad f_1'\left(x, y\right) = \left(3, -4\right), \quad f_2' = \left(2, -1\right).$$

Исследуем точку $(-3/5, -7/10)$.

В точке $(-3/5, -7/10)$ множество активных индексов пусто, так как ограничение-неравенство выполнено строго: $2x - y < 0$. Множество $\overline{H}\left(-\frac{3}{5}, -\frac{7}{10}\right)$ имеет вид:

$$\overline{H}\big(-\tfrac{3}{5}, -\tfrac{7}{10}\big) = \left\{h \in \mathbf{R}^2 : \left(f_0'\big(-\tfrac{3}{5}, -\tfrac{7}{10}\big), h\right) \le 0,\ \left(f_1'\big(-\tfrac{3}{5}, -\tfrac{7}{10}\big), h\right) = 0\right\}$$

или

$$\overline{H}\big(-\tfrac{3}{5}, -\tfrac{7}{10}\big) = \left\{h = (h_1, h_2) :\ 3h_1 - 4h_2 \le 0,\ 3h_1 - 4h_2 = 0\right\} \Rightarrow$$

$$\overline{H}\big(-\tfrac{3}{5}, -\tfrac{7}{10}\big) = \left\{h = \left(h_1, 3h_1/4\right),\ h_1 \in \mathbf{R}\right\}.$$

Матрица вторых производных функции Лагранжа не зависит от x, y и имеет вид:

$$\mathcal{L}''_{xx}(x^*, y^*, \lambda_0, \lambda_1, \lambda_2) = \begin{pmatrix} 58\lambda_0 & -54\lambda_0 \\ -54\lambda_0 & 52\lambda_0 \end{pmatrix}.$$

Для $\lambda_0 = 1$ матрица $\mathcal{L}''_{xx}(x^*, y^*, \lambda_0, \lambda_1, \lambda_2)$ является положительно определенной, поэтому

$$\left(\mathcal{L}''_{xx}\left(\left(-\tfrac{3}{5}, -\tfrac{7}{10}, 1, \lambda_1, \lambda_2\right)\right) h, h\right) > 0$$

для любого ненулевого вектора h, в том числе и для $h \in \overline{H}\left(-\frac{3}{5}, -\frac{7}{10}\right)$. По теореме о достаточных условиях $(-3/5, -7/10)$ — строгий локальный минимум.

Исследуем точку $(-1/5, -2/5)$.

В точке $(-1/5, -2/5)$ множество активных индексов непусто, так как ограничение-неравенство выполнено как равенство.

Множество $\overline{H}\left(-\frac{1}{5}, -\frac{2}{5}\right)$ имеет вид:

$$\overline{H}\left(-\tfrac{1}{5}, -\tfrac{2}{5}\right) = \left\{h \in \mathbf{R}^2 : \left(f'_0\left(-\tfrac{1}{5}, -\tfrac{2}{5}\right), h\right) \le 0,\right.$$

$$\left.\left(f'_2\left(-\tfrac{1}{5}, -\tfrac{2}{5}\right), h\right) \le 0,\ \left(f'_1\left(-\tfrac{1}{5}, -\tfrac{2}{5}\right), h\right) = 0\right\}$$

или

$$\overline{H}\left(-\tfrac{1}{5}, -\tfrac{2}{5}\right) = \{h = (h_1, h_2) :\ 10h_1 - 10h_2 \le 0,\ 2h_1 - h_2 \le 0,\ 3h_1 - 4h_2 = 0\} \Rightarrow$$

$$\overline{H}\left(-\tfrac{1}{5}, -\tfrac{2}{5}\right) = \left\{h = \left(h_1, \tfrac{3}{4}h_1\right),\ h_1 \le 0\right\}.$$

Для $\lambda_0 = -1$ матрица $\mathcal{L}''_{xx}(x^*, y^*, \lambda_0, \lambda_1, \lambda_2)$ является отрицательно определенной, поэтому

$$\left(\mathcal{L}''_{xx}\left(\left(-\tfrac{1}{5}, -\tfrac{2}{5}, -1, \lambda_1, \lambda_2\right)\right) h, h\right) < 0$$

для любого ненулевого вектора h, в том числе и для $h \in \overline{H}\left(-\frac{1}{5}, -\frac{2}{5}\right)$. По теореме о достаточных условиях $(-1/5, -2/5)$ — строгий локальный максимум.

В этой задаче целевая функция выпуклая, допустимое множество выпукло, поэтому локальный минимум является глобальным (см. свойства решений в выпуклых задачах). Локальный максимум глобальным не является. Например, последовательность точек $\left(-n, -\dfrac{3}{4}n - \dfrac{1}{4}\right)$) является допустимой, и

$$f_0\left(n, -\frac{3}{4}n - \frac{1}{4}\right) \to +\infty, \quad n \to +\infty$$

2. Исследуем найденные точки, используя определение локального экстремума. Для этого проверим знак выражения

$$\triangle f_0 = f_0(x^* + \Delta x, y^* + \Delta y) - f_0(x^*, y^*),$$

где $f_0(x, y) = 29x^2 - 54xy + 26y^2$.

$$\begin{aligned}\triangle f_0 &= 29\left(x^*+\Delta x\right)^2 - 54\left(x^*+\Delta x\right)\left(y^*+\Delta y\right) + 26\left(y^*+\Delta y\right)^2\\ &\quad -29\left(x^*\right)^2 + 54x^*y^* - 26\left(y^*\right)^2 =\\ &= 58x^*\Delta x + 29\Delta x^2 - 54y^*\Delta x - 54x^*\Delta y - 54\Delta x\Delta y\\ &\quad +52y^*\Delta y + 26\left(\Delta y\right)^2\end{aligned}$$

1) $x^* = -3/5, \quad y^* = -7/10.$

$$\begin{aligned}\triangle f_0 &= -\frac{58}{5}\cdot 3\Delta x + 29\Delta x^2 + \frac{54\cdot 7}{10}\Delta x + \frac{54\cdot 3}{5}\Delta y - 54\Delta x\Delta y\\ &\quad -\frac{52\cdot 7}{10}\Delta y + 26\left(\Delta y\right)^2 =\\ &= 3\Delta x - 4\Delta y + 29\left(\Delta x\right)^2 + 26\left(\Delta y\right)^2 - 54\Delta x\Delta y.\end{aligned}$$

Так как точка $(x^* + \Delta x,\ y^* + \Delta y)$ должна быть допустимой, то необходимо выполнение следующих условий:

$$3(-\tfrac{3}{5} + \Delta x) - 4(-\tfrac{7}{10} + \Delta y) = 1 \quad \Rightarrow \quad 3\Delta x - 4\Delta y = 0 \quad \Rightarrow \quad \Delta y = 3\Delta x/4,$$

$$2(-\tfrac{3}{5} + \Delta x) - (-\tfrac{7}{10} + \Delta y) \le 0 \quad \Rightarrow \quad -\tfrac{1}{2} + 2\Delta x - \Delta y \le 0 \quad \Rightarrow \quad \Delta x \le 2/5.$$

Тогда

$$\triangle f_0 = 3\Delta x - 19\Delta x/4 + 29\left(\Delta x\right)^2 + 26\left(3\Delta x/4\right)^2 - 54\Delta x\left(3\Delta x/4\right) = 25\left(\Delta x\right)^2/8 \ge 0.$$

То есть

$$f_0\left(x^* + \Delta x, y^* + \Delta y\right) \ge f_0\left(x^*, y^*\right)$$

для любых допустимых $(\Delta x,\ \Delta y)$, следовательно, $(-3/5, -7/10) - absmin$.

2) $x^* = -1/5$, $y^* = -2/5$.

$$\begin{aligned}\triangle f_0 &= -\frac{58}{5}\Delta x + 29\left(\Delta x\right)^2 + \frac{108}{5}\Delta x + \frac{54}{5}\Delta y -\\ &\quad -54\Delta x\Delta y - \frac{104}{5}\Delta y + 26\left(\Delta y\right)^2 =\\ &= \frac{1}{5}\left[\Delta x\left(-58+108\right) + \Delta y\left(54-104\right)\right] +\\ &\quad +29\left(\Delta x\right)^2 - 54\Delta x\Delta y + 26\left(\Delta y\right)^2 =\\ &= 10\cdot\left[\Delta x - \Delta y\right] + 29\left(\Delta x\right)^2 - 54\Delta x\Delta y + 26\left(\Delta y\right)^2 =\\ &= 10\cdot\left[\Delta x - 3\Delta x/4\right] + 29\left(\Delta x\right)^2 - 54\Delta x\left(3\Delta x/4\right) + 26\left(3\Delta x/4\right)^2\end{aligned}$$

В последнем равенстве опять воспользовались тем, что точка должна быть допустимой, т.е.

$$3(-\tfrac{1}{5}+\Delta x)-4(-\tfrac{2}{5}+\Delta y)=1 \quad\Rightarrow\quad 3\Delta x-4\Delta y=0 \quad\Rightarrow\quad \Delta y=3\Delta x/4$$

и

$$2(-\tfrac{1}{5}+\Delta x)-(-\tfrac{2}{5}+\Delta y)\le 0 \quad\Rightarrow\quad 2\Delta x-\Delta y\le 0 \quad\Rightarrow\quad 2\Delta x-3\Delta x/4\le 0$$

$\Rightarrow \quad \Delta x \le 0$. Если при этом $\Delta x = 0$, то и $\Delta y = 0$, т.е. приращение нулевое.

При малых Δx знак $\triangle f_0$ определяется линейным членом $5\Delta x/2$, поэтому $\triangle f_0 = 5\Delta x/2 + 25\left(\Delta x\right)^2/8 \le 0$ при малых Δx; при больших по абсолютной величине допустимых Δx знак $\triangle f_0$ определяется квадратичным членом $25\left(\Delta x\right)^2/8$, поэтому приращение $\triangle f_0 > 0$ при больших по абсолютной величине Δx. Следовательно, $(-1/5, -2/5)$ — локальный максимум, но не глобальный.

Заметим, что при проверке по определению сразу же проверяется и глобальность найденных экстремумов. Если используем достаточные условия, то необходимы дополнительные рассуждения.

Итак: $(-3/5, -7/10) - absmin$, $(-1/5, -2/5) - locmax$.

6 Интерпретация множителей Лагранжа

Рассмотрим задачу с ограничениями в виде равенств

$$f_0(x) \to \min, \quad f_i(x) = b_i \quad (i = 1, \ldots, m) \tag{23}$$

и задачу

$$f_0(x) \to \min, \quad f_i(x) = b_i + \varepsilon_i \quad (i = 1, \ldots, m), \tag{24}$$

которую будем называть возмущением задачи (23).

Теорема 14
Пусть функции $f_i(x)$ $(i = 0, \ldots, m)$ дважды непрерывно дифференцируемы и задача (23) является регулярной, т.е. $\lambda_0^* = 1$. Пусть допустимая точка x^* – решение задачи (23), $\lambda^* = (1, \lambda_1^*, \ldots, \lambda_m^*)$ – соответствующий набор множителей Лагранжа. Предположим, что в точке x^* необходимое условие 2 порядка выполнено строго, то есть

$$\left(\mathcal{L}''_{xx}\left(x^*, 1, \lambda^*\right) h, h\right) > 0 \quad \forall h \in \mathbf{R}^n : \quad \left(f_i'\left(x^*\right), h\right) = 0, \quad i = 1, \ldots, m.$$

Тогда найдется такое $\beta > 0$, что для любого $\varepsilon \in \mathbf{R}^m$, $\|\varepsilon\| < \beta$, существуют единственное решение $x(\varepsilon)$ задачи (24) и соответствующий *ему набор множителей Лагранжа* $\lambda(\varepsilon) = (1, \lambda_1(\varepsilon), ..., \lambda_m(\varepsilon))$, *при этом* $x(0) = x^*, \quad \lambda(0) = \lambda^*$ *и*

$$\frac{\partial \varphi_0(\varepsilon)}{\partial \varepsilon_i} = -\lambda_i(\varepsilon), \quad i = 1, \ldots, m, \tag{25}$$

где $\varphi_0(\varepsilon) = f_0\left(x(\varepsilon)\right)$ *— оптимальное значение целевой функции в задаче (24). В частности,*

$$\left.\frac{\partial \varphi_0(\varepsilon)}{\partial \varepsilon_i}\right|_{\varepsilon=0} = -\lambda_i^*. \tag{26}$$

(Доказательство утверждения можно найти, например, в [4]).

Таким образом, множители Лагранжа показывают, как изменяется оптимальное значение целевой функции при малых изменениях констант в правых частях ограничений.

Пусть возмущение ε в задаче (24) произвольно, но его норма ограничена: $\|\varepsilon\| \le r \quad (r < \beta)$. Требуется определить вектор ε таким образом, чтобы при данном r получить наибольшее уменьшение целевой функции:

$$f_0\left(x(\varepsilon)\right) \to \min, \quad \|\varepsilon\| \le r. \tag{27}$$

Из (25) имеем

$$f_0\left(x(\varepsilon)\right) - f_0\left(x^*\right) = \varphi_0(\varepsilon) - \varphi_0(0) = (\varphi_0'(0), \varepsilon) + o\left(\varepsilon\right) = -\left(\lambda^*, \varepsilon\right) + o\left(\varepsilon\right). \tag{28}$$

Отсюда следует, что задача (27) эквивалентна задаче

$$(\lambda^*, \varepsilon) \to \max \tag{29}$$

$$\|\varepsilon\| \le r, \quad \varepsilon \in \mathbf{R}^m$$

Из геометрических соображений очевидно, что оптимальным решением задачи (29) является вектор, пропорциональный вектору λ^*:

$$\varepsilon^* = r\,\frac{\lambda^*}{\|\lambda^*\|}. \tag{30}$$

Если допускается изменять только одно ограничение в задаче (23), т.е. в задаче (24) допустимые ε имеют только одну ненулевую компоненту, то для наибольшего уменьшения целевой функции нужно изменять то ограничение, которому соответствует наибольший множитель Лагранжа λ_i^*.

Аналогичное свойство множителей Лагранжа имеет место и в задачах с ограничениями в виде равенств и неравенств. А именно, рассмотрим задачу

$$f_0(x) \to \text{extr}$$

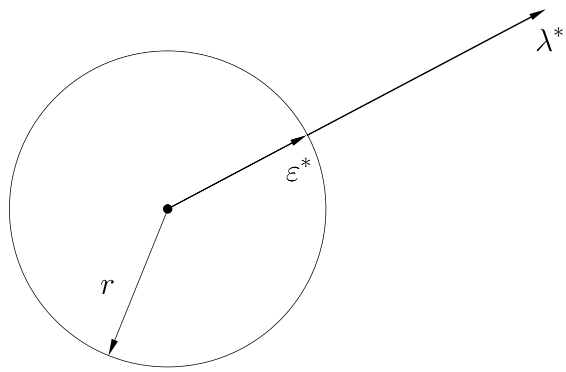

Рис. 7: иллюстрация к формуле (30)

$$f_i(x) \le b_i \quad (i = 1, ..., k)\,, \quad f_i(x) = b_i \quad (i = k+1, ..., m) \tag{31}$$

и соответствующую ей возмущенную задачу

$$f_0(x) \to \mathrm{extr}$$

$$f_i(x) \le b_i + \varepsilon_i, \quad (i = 1, ..., k)\,, \quad f_i(x) = b_i + \varepsilon_i \quad (i = k+1, ..., m)\,. \tag{32}$$

Пусть выполнены все предположения теоремы 14 и, кроме того, условия дополняющей нежесткости выполняются в строгой форме, т.е. если неравенство в точке x^* обращается в равенство (точка x^* лежит на границе ограничения), то соответствующий множитель Лагранжа строго больше нуля:

$$f_j\left(x^*\right) = b_j \quad (1 \le j \le k) \quad \Rightarrow \quad \lambda_j^* > 0.$$

Тогда теорема 14 верна и для задач (31)–(32). Свойство (26) заменяется на

$$\left.\frac{\partial \varphi_0(\varepsilon)}{\partial \varepsilon_j}\right|_{\varepsilon=0} = -\lambda_j^*, \quad j \in I\left(x^*\right) \bigcup \{k+1, ..., m\}\,,$$

где $I\left(x^*\right)$ — множество активных в точке x^* индексов.

7 Выпуклые задачи

7.1 Выпуклые множества. Выпуклые функции

Определение 7.1 *Множество D называется выпуклым, если для всех s, $v \in D$ и любого $\alpha \in [0,1]$ выполнено $\alpha s + (1-\alpha)v \in D$.*

Геометрически выпуклость множества означает, что отрезок, соединяющий любые две точки из множества D, целиком лежит в множестве D.

Замечание 13 *Отрезок, соединяющий точки s и v, это множество точек вида* $\{y \in \mathbf{R}^n : \ y = \alpha s + (1-\alpha)v,\ 0 \le \alpha \le 1\}$.

Пустое и одноточечное множества считаются выпуклыми по определению.

***Пример* 13** *Показать, что множество точек вида*

$$\Gamma = \{ x \in \mathbf{R}^n |\, (c, x) = \gamma \}, \quad c \in \mathbf{R}^n, \quad c \neq 0, \quad \gamma \in \mathbf{R},$$

называемое гиперплоскостью в $\mathbf{R}^n$*, выпукло.*

Решение. Действительно, для всех $\alpha \in [0,1]$ и s, $v \in \Gamma$ справедливо:

$$(c, \alpha s + (1-\alpha)\, v) = \alpha\, (c, s) + (1-\alpha)\, (c, v) = \alpha\gamma + (1-\alpha)\,\gamma = \gamma \quad \Rightarrow$$

$$\alpha s + (1-\alpha)\, v \in \Gamma.$$

***Пример* 14** *Показать, что замкнутые полупространства, порожденные гиперплоскостью* Γ*,*

$$\Gamma^+ = \{ x \in \mathbf{R}^n |\, (c, x) \ge \gamma \}, \quad \Gamma^- = \{ x \in \mathbf{R}^n |\, (c, x) \le \gamma \},$$

являются выпуклыми множествами в $\mathbf{R}^n$*.*

Определение 7.2 *Множества вида*

$$X = \left\{ x \in \mathbf{R}^n |\, \left(c^i, x\right) \le \gamma_i, \quad i = 1, \ldots, m \right\},$$

где $c^i = (c^i_1, \ldots, c^i_n)$, $\gamma_i \in \mathbf{R}$ $(i = 1, \ldots, m)$ *называются полиэдральными или полиэдрами.*

То есть полиэдр — это пересечение конечного числа полупространств. Из следующей теоремы будет следовать, что полиэдр — выпуклое множество.

Теорема 15
Пересечение (конечное или бесконечное) выпуклых множеств — выпуклое множество.

Доказательство. Пусть множества W_a выпуклы в $\mathbf{R}^n$ для всех $a \in A$. Обозначим

$$W = \bigcap_{a \in A} W_a.$$

Пусть w_1, $w_2 \in W$, тогда w_1, $w_2 \in W_a$ для всех $a \in A$. В силу выпуклости каждого W_a имеем

$$\alpha w_1 + (1-\alpha)\, w_2 \in W_a, \quad \forall a \in A, \quad \forall \alpha \in [0,1].$$

Отсюда следует, что $\alpha w_1 + (1-\alpha)\, w_2 \in W$, то есть множество W выпукло. $\square$

Объединение выпуклых множеств не обязательно выпукло.

Определение 7.3 *Функция $f(x)$, определенная на выпуклом множестве D, называется выпуклой, если $\forall\, s,\ v \in D$ и $\forall\, \alpha \in [0,1]$ выполнено неравенство Йенсена:*

$$f(\alpha s + (1-\alpha)v) \leq \alpha f(s) + (1-\alpha) f(v). \tag{33}$$

Если неравенство (33) выполнено строго для всех $s, v \in D,\ s \neq v$, и для всех $0 < \alpha < 1$, то функция f называется строго выпуклой на множестве D.

Если неравенство (33) выполнено в противоположную сторону, т.е. $\geq$, то функция $f(x)$ называется вогнутой. Если $f(x)$ — вогнутая функция, то $-f(x)$ — выпуклая функция.

Далее будем рассматривать собственные выпуклые функции, то есть такие, что $f\left(x\right) > -\infty, \quad x \in D$, и $f \not\equiv +\infty$.

***Пример* 15** *Показать, что функция $f\ : \mathbf{R}^n \to \mathbf{R}$, определенная формулой $f\left(x\right) = (c, x)$, $c \in \mathbf{R}^n$, выпукла.*

Решение. Действительно, для всех $s,\ v \in \mathbf{R}^n$ и для любого $\alpha \in [0,1]$ имеем

$$f\left(\alpha s + (1-\alpha)v\right) = (c, \alpha s + (1-\alpha)v) =$$

$$\alpha\left(c, s\right) + \left(1-\alpha\right)\left(c, v\right) = \alpha f\left(s\right) + \left(1-\alpha\right) f\left(v\right),$$

то есть неравенство Йенсена выполняется как равенство. Таким образом, функция f является выпуклой и вогнутой одновременно.

Теорема 16
Пусть функция $f\ : D \to \mathbf{R}$, где $D \subset \mathbf{R}^n$ — выпуклое множество, выпукла. Тогда для любых

$$s_i \in D, \quad \alpha_i \geq 0 \quad \left(i = 1, \ldots, m\right), \quad \sum_{i=1}^{m} \alpha_i = 1 \quad (m = 1, 2, \ldots)$$

имеет место неравенство (обобщенное неравенство Йенсена)

$$f\left(\sum_{i=1}^{m} \alpha_i s_i\right) \leq \sum_{i=1}^{m} \alpha_i f\left(s_i\right). \tag{34}$$

Доказательство. Проведем доказательство по индукции. При $m = 2$ неравенство (34) есть обычное неравенство Йенсена. Пусть неравенство (34) выполнено для $m = k$, $k \geq 2$, в частности, тогда $\sum_{i=1}^{k} \alpha_i s_i \in D$

для любых $s_i \in D$, $\alpha_i \geq 0 \, (i = 1, \ldots, k)$, $\sum_{i=1}^{k} \alpha_i = 1$. Можно считать, что $\alpha_{k+1} \neq 1$. Тогда

$$f\left(\sum_{i=1}^{k+1} \alpha_i s_i\right) = f\left(\sum_{i=1}^{k} \alpha_i s_i + \alpha_{k+1} s_{k+1}\right) =$$

$$= f\left(\left(\sum_{i=1}^{k} \alpha_i\right) \sum_{i=1}^{k} \frac{\alpha_i}{\left(\sum_{i=1}^{k} \alpha_i\right)} s_i + \alpha_{k+1} s_{k+1}\right).$$

Обозначим

$$\beta_i = \frac{\alpha_i}{\left(\sum_{i=1}^{k} \alpha_i\right)}, \; i = 1, \ldots, k.$$

Очевидно, что $\beta_i \geq 0 \, (i = 1, \ldots, k)$ и $\sum_{i=1}^{k} \beta_i = 1$. Тогда $\sum_{i=1}^{k} \beta_i s_i \in D$. Обозначим $\delta = \left(\sum_{i=1}^{k} \alpha_i\right)$, $v = \sum_{i=1}^{k} \beta_i s_i$. Тогда имеем по условию теоремы $\delta + \alpha_{k+1} = 1$. Используя выпуклость функции f и предположение индукции, получим

$$f\left(\delta v + \alpha_{k+1} s_{k+1}\right) \leq \delta f\left(v\right) + \alpha_{k+1} f\left(s_{k+1}\right) =$$

$$= \delta f\left(\sum_{i=1}^{k} \beta_i s_i\right) + \alpha_{k+1} f\left(s_{k+1}\right) \leq \delta\left(\sum_{i=1}^{k} \beta_i f\left(s_i\right)\right) + \alpha_{k+1} f\left(s_{k+1}\right) =$$

$$= \sum_{i=1}^{k} \delta \beta_i f\left(s_i\right) + \alpha_{k+1} f\left(s_{k+1}\right) = \sum_{i=1}^{k+1} \alpha_i f\left(s_i\right).$$

Теорема доказана. □

Теорема 17

Функция $f : D \to \mathbf{R}$, где $D \subset \mathbf{R}^n$ — выпуклое множество, выпукла тогда и только тогда, когда для любых s, $v \in D$, функция $\varphi : [0,1] \to \mathbf{R}$, определенная формулой $\varphi\left(\varepsilon\right) = f\left(\varepsilon s + \left(1 - \varepsilon\right) v\right) \;\; \left(\varepsilon \in [0,1]\right)$, выпукла.

Доказательство. $\Rightarrow$ Пусть функция f выпукла на множестве D. Для любых ε_1, ε_2, $\alpha \in [0,1]$, s, $v \in D$ имеем

$$\varphi\left(\alpha \varepsilon_1 + \left(1 - \alpha\right) \varepsilon_2\right) = f\left(\left(\alpha \varepsilon_1 + \left(1 - \alpha\right) \varepsilon_2\right) s + \left(\alpha \varepsilon_1 + \left(1 - \alpha\right) \varepsilon_2\right) v\right) =$$

$$= f\left(\alpha\left(\varepsilon_1 s + \left(1 - \varepsilon_1\right) v\right) + \left(1 - \alpha\right)\left(\varepsilon_2 s + \left(1 - \varepsilon_2\right) v\right)\right) \leq$$

$$\leq \alpha f\left(\varepsilon_1 s + \left(1 - \varepsilon_1\right) v\right) + \left(1 - \alpha\right) f\left(\varepsilon_2 s + \left(1 - \varepsilon_2\right) v\right) = \alpha \varphi\left(\varepsilon_1\right) + \left(1 - \alpha\right) \varphi\left(\varepsilon_2\right),$$

то есть функция $\varphi\left(\varepsilon\right)$ выпукла.

$\Leftarrow$ Пусть для любой пары точек s, v из множества D функция φ выпукла. Тогда для всех $s,\, v \in D,\ \ \alpha \in [0,1]$ имеем

$$f\left(\alpha s + (1-\alpha)v\right) = \varphi\left(\alpha\right) = \varphi\left(\alpha \cdot 1 + (1-\alpha) \cdot 0\right) \leq$$

$$\leq \alpha\varphi\left(1\right) + \left(1-\alpha\right)\varphi\left(0\right) = \alpha f\left(s\right) + \left(1-\alpha\right) f\left(v\right),$$

то есть функция f выпукла. Теорема доказана. $\square$

Дадим эквивалентное определение выпуклой функции через ее надграфик.

Определение 7.4 *Рассмотрим функцию $f : D \to \mathbf{R}$, где $D \subset \mathbf{R}^n$. Надграфиком функции f назовем множество*

$$epi\, f = \left\{(x,\beta) \in \mathbf{R}^{n+1}\ : x \in D,\, \beta \in \mathbf{R},\, \beta \geq f\left(x\right)\right\}.$$

Примеры надграфиков на рис. 7.

Теорема 18
Пусть D — выпуклое множество, $f : D \to \mathbf{R}$. Функция f выпукла тогда и только тогда, когда $epi\, f$ — выпуклое множество в $\mathbf{R}^{n+1}$.

Доказательство. $\Rightarrow$ Пусть $f\left(x\right)$ — выпуклая функция на выпуклом множестве D. Пусть $y = (s,\beta) \in epi\, f$, $z = (v,\gamma) \in epi\, f$. Покажем, что $\alpha y + (1-\alpha)\, z \in epi\, f$ для всех $\alpha \in [0,1]$. Рассмотрим

$$\alpha y + (1-\alpha)\, z = \alpha\left(s,\beta\right) + \left(1-\alpha\right)\left(v,\gamma\right) =$$

$$= \left(\alpha s + (1-\alpha)\, v, \alpha\beta + (1-\alpha)\,\gamma\right).$$

Точка $\alpha s + (1-\alpha)\, v \in D$, так как D — выпуклое множество. Надо проверить, что

$$\alpha\beta + (1-\alpha)\,\gamma \geq f\left(\alpha s + (1-\alpha)\, v\right).$$

Имеем

$$f\left(\alpha s + (1-\alpha)\, v\right) \leq \alpha f\left(s\right) + \left(1-\alpha\right) f\left(v\right) \leq \alpha\beta + (1-\alpha)\,\gamma \quad \forall \alpha \in [0,1].$$

Первое неравенство верно в силу выпуклости функции $f\left(x\right)$, второе неравенство следует из определения надграфика. Выпуклость надграфика доказана.

$\Leftarrow$ Рассмотрим точки $\left(s, f\left(s\right)\right),\, \left(v, f\left(v\right)\right) \in epi\, f$. Так как $epi\, f$ — выпуклое множество в $\mathbf{R}^{n+1}$, то

$$\alpha\left(s, f\left(s\right)\right) + \left(1-\alpha\right)\left(v, f\left(v\right)\right) \in epi\, f \quad \forall \alpha \in [0,1].$$

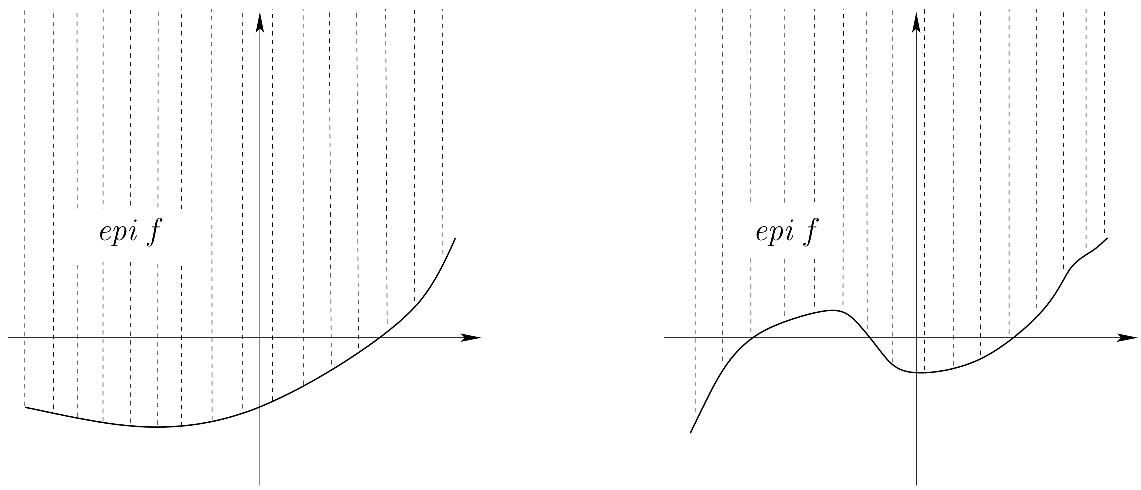

Рис. 8: слева надграфик выпуклой функции, справа надграфик функции, которая не является выпуклой

Отсюда имеем для всех $\alpha \in [0,1]$

$\alpha\,(s, f\,(s)) + (1-\alpha)\,(v, f\,(v)) = (\alpha s + (1-\alpha)\,v, \alpha f\,(s) + (1-\alpha)\,f\,(v)) \in epi\,f.$

Из определения надграфика получаем: $\alpha s + (1-\alpha)\,v \in D$ и

$$\alpha f\,(s) + (1-\alpha)\,f\,(v) \geq f\,(\alpha s + (1-\alpha)\,v)$$

для всех $\alpha \in [0,1]$. Таким образом, множество D — выпуклое множество, и выполнено неравенство Йенсена для функции f. Теорема доказана. $\square$

Теорема 19
Пусть $D \subset \mathbf{R}^n$ — выпуклое множество, функции $f_i : D \to \mathbf{R}$ выпуклы для любого $i = 1, \ldots, m$. Тогда функция

$$f(x) = \sum_{i=1}^{m} c_i\, f_i(x)$$

является выпуклой на D для всех $c_i \geq 0$ $(i = 1, \ldots, m)$.

Доказательство. Рассмотрим любые точки $s, v \in D$ и произвольное $\alpha \in [0,1]$. Тогда

$$f\,(\alpha s + (1-\alpha)v) = \sum_{i=1}^{m} c_i\, f_i\,(\alpha s + (1-\alpha)\,v) \leq_{(1)}$$

$$\sum_{i=1}^{m} c_i \left(\alpha f_i(s) + (1-\alpha) f_i(v)\right) = \alpha \sum_{i=1}^{m} c_i f_i(s) + (1-\alpha) \sum_{i=1}^{m} c_i f_i(v) =$$

$$= \alpha f(s) + (1-\alpha) f(v).$$

Таким образом, для функции $f(x)$ выполнено неравенство Йенсена. Заметим, что в неравенстве (1) использовали выпуклость функций f_i и неотрицательность c_i. □

Теорема 20
Пусть на выпуклом множестве D определены выпуклые функции $f_i(x), i \in I$ (I — некоторое множество индексов). Тогда функция

$$f(x) = \sup_{i \in I} f_i(x)$$

является выпуклой на D.

Доказательство. Заметим, что

$$epi \sup_{i \in I} f_i = \bigcap_{i \in I} epi\, f_i.$$

Так как $epi\, f_i$ — выпуклое множество для любого $i \in I$, а пересечение выпуклых множеств является выпуклым множеством, то $epi \sup_{i \in I} f_i = epi\, f$ — выпуклое множество, следовательно, $f(x)$ — выпуклая функция.

Это утверждение можно доказать, используя только определение выпуклой функции. А именно, рассмотрим любые точки $s, v \in D$ и произвольное $\alpha \in [0,1]$. Тогда

$$f_i(\alpha s + (1-\alpha) v) \leq_{(1)} \alpha f_i(s) + (1-\alpha) f_i(v) \leq_{(2)} \alpha f(s) + (1-\alpha) f(v), \tag{35}$$

здесь неравенство (1) выполнено в силу выпуклости функций f_i, неравенство (2) следует из определения функции $f(x)$. Возьмем sup по i от левой и правой частей (35):

$$\sup_{i \in I} f_i(\alpha s + (1-\alpha) v) \leq \alpha f(s) + (1-\alpha) f(v),$$

то есть имеем неравенство Йенсена для функции f:

$$f(\alpha s + (1-\alpha) v) \leq \alpha f(s) + (1-\alpha) f(v).$$

□

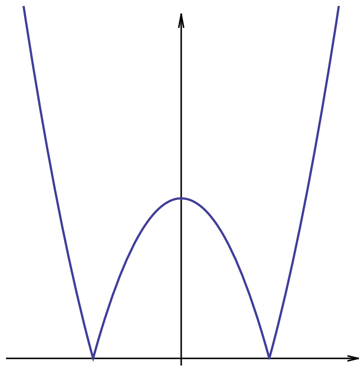

Рис. 9: к примеру 16, график функции $\varphi(x) = |x^2 - 1|$.

Теорема 21
Пусть на выпуклом множестве D_1 определена выпуклая функция $f_1(x)$, на выпуклом множестве D_2 $(f_1(D_1) \subset D_2)$ определена неубывающая выпуклая скалярная функция $f_2(y)$. Тогда функция $f(x) = f_2(f_1(x))$ является выпуклой на D_1.

Доказательство. Проверим выполнение неравенства Йенсена:

$$f\ (\alpha s + (1-\alpha) v) = f_2 (f_1 (\alpha s + (1-\alpha) v)) \leq f_2 (\alpha f_1 (s) + (1-\alpha) f_1 (v)) \leq$$
$$\leq \alpha f_2 (f_1 (s)) + (1-\alpha) f_2 (f_1 (v)) = \alpha f (s) + (1-\alpha) f (v).$$

В этой цепочке первое неравенство выполняется в силу выпуклости $f_1(x)$ и неубывания $f_2(y)$ $(y \leq \bar{y} \Rightarrow f_2(y) \leq f_2(\bar{y}))$, второе неравенство выполняется в силу выпуклости $f_2(x)$. Таким образом, функция $f(x)$ является выпуклой. □

***Пример* 16** *Заданы функции $f(x) = x^2 - 1$ и $g(y) = |y|$. Рассмотрим их композицию:*

$$\varphi(x) = g(f(x)) = \left|x^2 - 1\right|.$$

Функции $f(x)$ и $g(y)$выпуклые, однако их композиция выпуклой функцией не является. Теорема 21 не применима, так как внешняя функция $g(y)$не является неубывающей.

***Пример* 17** *Заданы функции $f(x_1, x_2) = x_1^2 + x_2^2$ и $g(y) = e^y$. Рассмотрим их композицию:*

$$\varphi(x) = g(f(x)) = e^{x_1^2 + x_2^2}.$$

Функции $f(x)$ и $g(y)$ выпуклые, кроме того, $g(y)$ — возрастающая функция, следовательно, выполнены условия теоремы 21, композиция $\varphi(x) = e^{x_1^2 + x_2^2}$ — выпуклая функция.

Теорема 22
Пусть $f\ : D \to \mathbf{R}$ — дифференцируемая функция, $D \subset \mathbf{R}^n$ — выпуклое множество. Функция $f\left(x\right)$ является выпуклой тогда и только тогда, когда

$$f\left(x\right) - f\left(y\right) \geq \left(f'\left(y\right), x - y\right) \,\forall\, x,\, y \in D.$$

Доказательство. $\Rightarrow$ Пусть $f\left(x\right)$— выпуклая дифференцируемая функцияю Рассмотрим точки $x, y \in D$. Обозначим $h = x - y$. Применим формулу Тейлора с остаточным членом в форме Пеано:

$$f\left(y + \alpha h\right) - f\left(y\right) = \left(f'\left(y\right), \alpha h\right) + o(\alpha) \tag{36}$$

С другой стороны, в силу выпуклости $f(x)$:

$$f\left(y + \alpha h\right) - f\left(y\right) = f\left(\left(1 - \alpha\right) y + \alpha x\right) - f\left(y\right) \leq$$

$$\leq \left(1 - \alpha\right) f\left(y\right) + \alpha f\left(x\right) - f\left(y\right) = \alpha \left(f\left(x\right) - f\left(y\right)\right).$$

Отсюда, используя (36), получим

$$\alpha \left(f\left(x\right) - f\left(y\right)\right) \geq \left(f'\left(y\right), \alpha h\right) + o(\alpha).$$

Разделим обе части последнего неравенства на α и рассмотрим предел при $\alpha \to 0$, получим утверждение теоремы.

$\Leftarrow$ Пусть $\forall\, x,\, y \in D$ выполнено

$$f\left(x\right) - f\left(y\right) \geq \left(f'\left(y\right), x - y\right),$$

то есть при каждом фиксированном x имеем оценку

$$f\left(x\right) \geq f\left(y\right) + \left(f'\left(y\right), x - y\right) \quad \forall y.$$

Отсюда

$$f\left(x\right) = \sup_{y \in D} \left\{f\left(y\right) + \left(f'\left(y\right), x - y\right)\right\}.$$

При фиксированном y функция

$$g\left(x\right) = f\left(y\right) + \left(f'\left(y\right), x - y\right)$$

является аффинной, а следовательно, выпуклой по x. В силу теоремы 20 функция $f\left(x\right)$ выпукла. $\square$

Теорема 23
Пусть $f\ : D \to \mathbf{R}$, где $D \subset \mathbf{R}^n$ — выпуклое множество, дважды дифференцируемая функция. Функция f выпукла тогда и только тогда, когда матрица вторых производных $f''\left(x\right)$ неотрицательно определена $\forall x \in D$.

Доказательство. $\Rightarrow$ Пусть f — выпуклая функция. Рассмотрим произвольные точки $y, x \in D$, обозначим $h = x - y$, тогда по определению выпуклого множества точка $y + \alpha h \in D \quad (0 \le \alpha \le 1)$, так как $y + \alpha h = y + \alpha\,(x - y) = \alpha x + (1 - \alpha)\,y$. В силу теоремы 22

$$\forall\, y,\, x \in D \quad f\,(y + \alpha h) - f\,(y) - (f'\,(y)\,, \alpha h) \ge 0.$$

С другой стороны, применяя формулу Тейлора, имеем :

$$f\,(y + \alpha h) - f\,(y) - (f'\,(y)\,, \alpha h) = \frac{1}{2}\,(f''\,(y)\,\alpha h, \alpha h) + o\left(\alpha^2\right).$$

Отсюда

$$\frac{1}{2}\,(f''\,(y)\,\alpha h, \alpha h) + o\left(\alpha^2\right) \ge 0.$$

Разделим обе части неравенства на α^2 и рассмотрим предел при $\alpha \to +0$, получим $(f''\,(y)\,h, h) \ge 0$.

$\Leftarrow$ Пусть $f''\,(s)$ — неотрицательно определенная матрица $\forall\, s \in D$. Рассмотрим точки $y,\, x \in D$, обозначим $h = x - y$. Запишем для функции f формулу Тейлора с остаточным членом в форме Лагранжа

$$f\,(x) - f\,(y) = (f'\,(y)\,, h) + \frac{1}{2}\,(f''\,(y + \alpha h)\,h, h)\,,$$

где $0 < \alpha < 1$. Точка $y + \alpha h \in D$ в силу выпуклости множества D. Отсюда в силу неотрицательной определенности матрицы $f''\,(y + \alpha h)$ имеем для всех $x,\, y \in D$

$$f\,(x) - f\,(y) - (f'\,(y)\,, x - y) \ge 0,$$

следовательно, по теореме 22 функция $f\,(x)$ является выпуклой. $\square$

***Пример* 18** *Исследовать на выпуклость функцию*

$$f\,(x, y, z) = x^2 + xy + y^2 + 2z + \exp\left(z^2 + 3x + 4y\right).$$

Решение. Обозначим $f_1\,(x, y, z) = x^2 + xy + y^2 + 2z$, $f_2\,(x, y, z) = \exp\,(z^2 + 3x + 4y)$. Проверим выпуклость функций $f_1\,(x, y, z)$ и $f_2\,(x, y, z)$, используя теорему 23. Вычислим матрицы вторых производных f_i'' $(i = 1, 2)$. Для функции $f_1\,(x, y, z)$:

$$f_1''\,(x, y, z) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Применим критерий Сильвестра для проверки матрицы $f_1''\,(x, y, z)$. Угловые миноры:

$$\triangle_1 = 2 > 0, \quad \triangle_2 = 2 \cdot 2 - 1 \cdot 1 = 3 > 0, \quad \triangle_3 = 0.$$

Матрица $f_1''(x,y,z)$ может быть неотрицательно определенной, поэтому вычислим главные миноры:

$$\delta_1 = 0, \quad \delta_2 = 0, \quad \delta_3 = 3, \quad \delta_{12} = 0, \quad \delta_{13} = 2, \quad \delta_{23} = 2, \quad \delta_0 = 0.$$

Так как все главные миноры неотрицательны, то матрица $f_1''(x,y,z)$ неотрицательно определена, а, следовательно, функция $f_1(x,y,z)$ выпуклая.

Для функции $f_2(x,y,z)$:

$$f_2''(x,y,z) = \begin{pmatrix} 9e^{\left(z^2+3x+4y\right)} & 12e^{\left(z^2+3x+4y\right)} & 6ze^{\left(z^2+3x+4y\right)} \\ 12e^{\left(z^2+3x+4y\right)} & 16e^{\left(z^2+3x+4y\right)} & 8ze^{\left(z^2+3x+4y\right)} \\ 6ze^{\left(z^2+3x+4y\right)} & 8ze^{\left(z^2+3x+4y\right)} & 2e^{\left(z^2+3x+4y\right)} + 4z^2e^{\left(z^2+3x+4y\right)} \end{pmatrix} =$$

$$= e^{\left(z^2+3x+4y\right)} \begin{pmatrix} 9 & 12 & 6z \\ 12 & 16 & 8z \\ 6z & 8z & 2+4z^2 \end{pmatrix}.$$

Обозначим

$$A = \begin{pmatrix} 9 & 12 & 6z \\ 12 & 16 & 8z \\ 6z & 8z & 2+4z^2 \end{pmatrix}.$$

Так как $e^{\left(z^2+3x+4y\right)} > 0$ для любых (x,y,z), то матрицы $f_2''(x,y,z)$ и A одинаково определены. Применим критерий Сильвестра для проверки матрицы A. Угловые миноры:

$$\triangle_1 = 9 > 0, \quad \triangle_2 = 9 \cdot 16 - 12^2 = 0,$$

$$\triangle_3 = 9 \cdot 16 \cdot \left(2+4z^2\right) + 12 \cdot 8z \cdot 6z + 12 \cdot 8z \cdot 6z - 16 \cdot 36z^2 - 9 \cdot 64z^2 - 144 \cdot \left(2+4z^2\right) = 0$$

Матрица A может быть неотрицательно определенной, поэтому вычислим главные миноры:

$$\delta_1 = 16\left(2+4z^2\right) - 64z^2 = 32, \quad \delta_2 = 9\left(2+4z^2\right) - 36z^2 = 18,$$

$$\delta_3 = \triangle_2 = 0, \quad \delta_{12} = 2+4z^2, \quad \delta_{13} = 16,$$

$$\delta_{23} = \triangle_1 = 9, \quad \delta_0 = \triangle_3 = 0.$$

Таким образом, все главные миноры неотрицательны, следовательно, матрица A — неотрицательно определенная, функция f_2 — выпуклая.

Так как функции $f_1(x,y,z)$ и $f_2(x,y,z)$ выпуклые, то и функция $f(x,y,z) = f_1(x,y,z) + f_2(x,y,z)$ является выпуклой (см. теорему 19).

Замечание 14 *Заметим, что к функции $f_2(x,y,z)$ можно применить теорему о композиции выпуклых функций. Действительно, функция e^t является выпуклой возрастающей, функция $z^2+3x+4y$ выпуклая (для проверки, например, можно выписать матрицу вторых производных), отсюда получаем, что $f_2(x,y,z) = \exp\left(z^2+3x+4y\right)$ — выпуклая функция (см. теорему 21).*

7.2 Понятие выпуклой задачи. Свойства решений

Задачу оптимизации
$$f_0\,(x) \to \min, \quad x \in D, \tag{37}$$
будем называть выпуклой, если допустимое множество D — выпуклое множество и целевая функция $f_0(x)$ — выпуклая функция на D.

Выпуклые задачи обладают замечательным свойством: любое локальное решение является и глобальным.

Теорема 24
Если допустимая точка x^* — локальный минимум в выпуклой задаче (37), то x^* — абсолютный минимум в задаче (37).

Доказательство. Предположим, что x^* не является абсолютным минимумом, то есть существует допустимая точка $\bar{x}$, для которой
$$f_0\,(\bar{x}) < f_0\,(x^*)\,.$$
Так как $x^* \in locmin(37)$, то существует ε – окрестность $O_\varepsilon\,(x^*)$ точки x^*, такая, что:
$$f_0\,(x) \geq f_0\,(x^*) \quad \forall x \in O_\varepsilon\,(x^*) \cap D.$$
Так как D — выпуклое множество, то $x^\alpha = \alpha\bar{x} + (1-\alpha)\,x^* \in D$ для всех $\alpha \in (0,1)$. При достаточно малых α точка $x^\alpha \in O_\varepsilon\,(x^*)$, поэтому имеем следующую оценку:
$$f_0\,(x^*) \leq f_0\,(x^\alpha) \leq \alpha f_0\,(\bar{x}) + (1-\alpha)\,f_0\,(x^*)$$
или $\alpha f_0\,(x^*) \leq \alpha f_0\,(\bar{x})$. Отсюда получаем $f_0\,(x^*) \leq f_0\,(\bar{x})$. Противоречие с предположением. Следовательно, x^*— абсолютный минимум. □

Замечание 15 *В дальнейшем будем говорить просто о решении выпуклой задачи, имея в виду, что это решение обязательно является глобальным.*

Теорема 25 *(критерий оптимальности в выпуклой задаче).*

Пусть в выпуклой задаче целевая функция дифференцируема в точке $x^* \in D$. Точка x^* является решением выпуклой задачи (37) тогда и только тогда, когда
$$(f_0'(x^*), x - x^*) \geq 0 \quad \forall\, x \in D \tag{38}$$

Доказательство. $\Rightarrow$ Пусть x^* — решение выпуклой задачи, т.е. $f_0(x) \geq f_0(x^*)$ для всех $x \in D$. Предположим, что теорема неверна, то есть существует допустимая точка $\hat{x}$, для которой $(f_0'(x^*), \hat{x} - x^*) < 0$. Обозначим $h = \hat{x} - x^*$. В силу выпуклости множества D точка $x^* + \alpha h \in D$ для всех $\alpha \in [0, 1]$. Имеем при малых α:

$$f_0(x^* + \alpha h) - f_0(x^*) = (f_0'(x^*), \alpha h) + o(\alpha) = \alpha (f_0'(x^*), \hat{x} - x^*) + o(\alpha) < 0,$$

противоречие с тем, что x^*— точка минимума. Следовательно, наше предположение неверно и

$$(f_0'(x^*), x - x^*) \geq 0 \quad \forall x \in D.$$

$\Leftarrow$ Пусть $x^* \in D$ и $(f_0'(x^*), x - x^*) \geq 0 \quad \forall x \in D$. По теореме 22

$$f_0(x) - f_0(x^*) \geq (f_0'(x^*), x - x^*) \quad \forall x \in D.$$

Отсюда следует, что $f_0(x) - f_0(x^*) \geq 0 \quad \forall x \in D$, т.е. x^*— решение задачи (37).

Теорема доказана. $\square$

Следствие 1. *Пусть $f_0(x)$ — выпуклая на $\mathbf{R}^n$ функция. Если $f_0(x)$ дифференцируема в точке x^* и $f_0'(x^*) = 0$, то x^* является точкой глобального минимума $f_0(x)$ на $\mathbf{R}^n$, а следовательно, и на любом множестве $D \subset \mathbf{R}^n$, содержащем точку x^*.*

Следствие 2. *Пусть множество D имеет вид*

$$D = \{x \in \mathbf{R}^n \,|x_j \geq 0, \quad j = 1, \ldots, s\},$$

где $0 \leq s \leq n$ (случай $s = 0$ соответствует $D = \mathbf{R}^n$). Тогда (38) эквивалентно следующим условиям:

$$\frac{\partial f_0(x^*)}{\partial x_j} \geq 0, \quad x_j^* \frac{\partial f_0(x^*)}{\partial x_j} = 0, \quad j = 1, \ldots, s,$$

$$\frac{\partial f_0(x^*)}{\partial x_j} = 0, \quad j = s+1, \ldots, n.$$

Замечание 16 *При доказательстве первой части теоремы не используется выпуклость целевой функции. То есть имеет место следующее утверждение.*

Теорема 26
Рассмотрим задачу

$$f_0(x) \to \min, \ x \in D, \tag{39}$$

где D — выпуклое множество, а функция f_0 дифференцируема в допустимой точке x^*. Если $x^* \in locmin(39)$, то

$$\left(f_0'(x^*), x - x^*\right) \geq 0 \quad \forall\, x \in D.$$

Обозначим D^* — множество всех решений в выпуклой задаче (37).

Теорема 27
В выпуклой задаче множество D^* выпукло. Если целевая функция является строго выпуклой, то D^* содержит не более одной точки.

Доказательство. 1. Пустое множество и множество, состоящее из одной точки, являются выпуклыми. Пусть D^* содержит не менее двух точек. Возьмем точки s, $t \in D^*$. Тогда в силу выпуклости множества D точка $x^\alpha = \alpha s + (1-\alpha)t \in D$ для всех $\alpha \in [0,1]$. Обозначим

$$f_0^* = \min_{x \in D} f_0(x).$$

Тогда $f_0(s) = f_0(t) = f_0^*$. Используя выпуклость $f_0(x)$, получим:

$$f_0(x^\alpha) \leq \alpha f_0(s) + (1-\alpha) f_0(t) = f_0^*.$$

Так как f_0^* — минимальное значение, то

$$f_0(x^\alpha) \geq f_0^* \Rightarrow f_0(x^\alpha) = f_0^* \quad \forall\, \alpha \in [0,1].$$

Следовательно, $x^\alpha \in D^*$ для всех $\alpha \in [0,1]$, т.е. D^* — выпуклое множество.

2. Пусть $f_0(x)$ — строго выпуклая функция. Предположим, что в множестве D^* найдется не менее двух различных точек. Пусть s, $t \in D^*$, тогда

$$f_0(\alpha s + (1-\alpha)t) < \alpha f_0(s) + (1-\alpha) f_0(t) = f_0^* \quad \forall\, \alpha \in (0,1).$$

Получили противоречие с тем, что f_0^* — минимальное значение функции $f_0(x)$ на множестве D. Следовательно, D^* состоит не более чем из одной точки. $\square$

7.3 Задача выпуклого программирования.

Рассмотрим следующую задачу

$$\begin{gathered} f_0(x) \to \min \\ f_i(x) \le b_i,\ i = 1, ..., k \\ f_i(x) = b_i,\ i = k+1, ..., m \\ x \in P \subset \mathbf{R}^n. \end{gathered} \tag{40}$$

Задача (40) называется задачей выпуклого программирования (ЗВП), если P — выпуклое множество, $f_i(x)$ $(i = 0, \ldots, k)$ — выпуклые на P функции, $f_i(x)$ $(i = k+1, \ldots, m)$ линейные на P функции, т.е. имеют вид

$$f_i(x) = a_1^i x_1 + ... + a_n^i x_n = \left(a^i, x\right)(i = k+1, \ldots, m).$$

Таким образом, допустимым множеством в задаче (40) является множество

$$D = \{x \in P \subset \mathbf{R}^n : f_i(x) \le b_i,\ i = 1, ..., k,\ f_i(x) = b_i,\ i = k+1, ..., m\}. \tag{41}$$

Можно показать, что для задачи выпуклого программирования допустимое множество D является выпуклым, т.е. задача выпуклого программирования есть частный случай выпуклой задачи и для нее верны утверждения предыдущего параграфа.

Для решения ЗВП применяют метод множителей Лагранжа. Функция Лагранжа для задачи (40) имеет следующий вид:

$$\mathcal{L}(x, \lambda_0, \lambda_1, ..., \lambda_m) = \lambda_0 f_0(x) + \sum_{i=1}^{m} \lambda_i (f_i(x) - b_i).$$

Так как ЗВП является выпуклой задачей, то любой ее локальный минимум является и глобальным, поэтому мы будем говорить просто о минимуме.

Теорема 28 *(Теорема Куна-Таккера)*

I. (Необходимые условия оптимальности)
Пусть x^*— решение ЗВП. Тогда найдется такой ненулевой набор множителей Лагранжа $(\lambda_0^*, \lambda_1^*, ..., \lambda_m^*)$, что выполняется следующее:

1. *условие минимума*

$$\mathcal{L}(x, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*) \ge \mathcal{L}(x^*, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*) \quad \forall x \in P;$$

2. *условие дополняющей нежесткости*

$$\lambda_i^* (f_i(x^*) - b_i) = 0,\ i = 1, ..., k;$$

3. условие неотрицательности $\lambda_i^* \geq 0,\ i = 0, ..., k$.

II. (Достаточные условия оптимальности)
Пусть в точке $x^* \in D$ *выполнены условия 1)-3), и при этом* $\lambda_0^* \neq 0$. *Тогда* x^* *— решение ЗВП.*

Нам понадобится вспомогательное утверждение, которое приведем без доказательства.

Теорема 29 *(Конечномерная теорема отделимости)*
Пусть C *— непустое выпуклое множество,* $C \subset \mathbf{R}^N$ *и* $0 \notin C$. *Тогда найдутся числа* $a_1, ..., a_N$, *не все равные* 0, *такие, что*

$$\sum_{i=1}^{N} a_i z_i \geq 0 \quad \forall z = (z_1, ..., z_N) \in C.$$

Доказательство теоремы Куна-Таккера.
I. Пусть x^*— решение ЗВП. Можно считать, что $f_0\left(x^*\right) = 0$, иначе будем рассматривать новую целевую функцию $\tilde{f}\left(x\right) = f_0\left(x\right) - f_0\left(x^*\right)$.

Определим множество $C \subset \mathbf{R}^{m+1}$:

$$C = \{\, \mu = (\mu_0, ..., \mu_m)|\ \exists x^\mu \in P : f_0\left(x^\mu\right) < \mu_0,$$

$$f_i\left(x^\mu\right) \leq b_i + \mu_i,\ i = 1, ..., k; f_i\left(x^\mu\right) = b_i + \mu_i,\ i = k + 1, ..., m\}\,. \qquad (42)$$

А) Множество C непусто и выпукло. Действительно, для любого $\delta > 0$ вектор

$$\mu\left(\delta\right) = (\delta, \delta, ...\delta, 0, ...0) \in C,$$

т.к. в качестве x^μ можно взять x^*, тогда все условия из определения (42) будут выполнены. Покажем, что C выпукло. Возьмем $\mu^1, \mu^2 \in C$. Вектору μ^i соответствует точка $x^i \in P$. Рассмотрим вектор $\mu(\alpha) = \alpha\mu^1 + (1 - \alpha)\mu^2, \alpha \in [0, 1]$. Покажем, что для вектора $\mu(\alpha)$ и точки $x(\alpha) = \alpha x^1 + (1 - \alpha)x^2$ выполняются все условия из определения (42). Действительно, $x(\alpha) \in P$ для всех $\alpha \in [0, 1]$ в силу выпуклости P. Так как функции $f_i(x)\,(i = 0, \ldots, k)$ выпуклы, то

$$f_0(x(\alpha)) \leq \alpha f_0\left(x^1\right) + (1 - \alpha) f_0\left(x^2\right) < \alpha\mu_0^1 + (1 - \alpha)\mu_0^2 = \mu_0(\alpha)$$

$$f_i(x(\alpha)) \leq \alpha f_i\left(x^1\right) + (1 - \alpha) f_i\left(x^2\right) \leq \alpha\left(b_i + \mu_i^1\right) + (1 - \alpha)\left(b_i + \mu_i^2\right) =$$

$$= b_i + \mu_i(\alpha),\ i = 1, ..., k.$$

Функции $f_i(x)\,(i = k + 1, \ldots, m)$ линейны, поэтому

$$f_i\left(x\left(\alpha\right)\right) = \alpha f_i\left(x^1\right) + (1 - \alpha)\, f_i\left(x^2\right) = \alpha\left(b_i + \mu_i^1\right) + (1 - \alpha)\left(b_i + \mu_i^2\right) =$$

$$= b_i + \mu_i(\alpha), \; i = k+1, \ldots, m.$$

Следовательно, $\mu(\alpha) \in C$ для всех $\alpha \in [0,1] \quad \Rightarrow C$ — выпуклое множество.

Б) Точка $0 \in \mathbf{R}^{m+1}$ не принадлежит множеству C. Действительно, если $0 \in C$, то найдется точка $x^0 \in P$, для которой

$$f_0\left(x^0\right) < 0, \quad f_i(x^\mu) \le b_i, \, i = 1, \ldots, k; \quad f_i(x^\mu) = b_i, \, i = k+1, \ldots, m.$$

Отсюда, x^0 — допустимая точка и $0 = f_0(x^*) > f_0(x^0)$. Противоречие с тем, что x^*— решение ЗВП. Следовательно, $0 \notin C$. Таким образом, к множеству C, определенному условиями (42), применима конечномерная теорема отделимости. Т.е., найдется ненулевой набор чисел $(\lambda_0^*, ..., \lambda_m^*)$, удовлетворяющий неравенству

$$\sum_{i=0}^{m} \lambda_i^* \mu_i \ge 0 \quad \forall \mu \in C. \tag{43}$$

Покажем, что набор $(\lambda_0^*, \lambda_1^*, ..., \lambda_m^*)$ удовлетворяет пунктам 1)-3) части I теоремы.

В) *Условие неотрицательности.* Рассмотрим произвольный индекс $i_0 \, (0 \le i_0 \le k)$ и вектор μ:

$$\mu_{i_o} = 1, \quad \mu_i = \delta, \, i = 0, ..., k, i \ne i_0, \quad \delta > 0, \quad \mu_i = 0, i = k+1, ..., m,$$

то есть

$$\mu = \left(\delta, .., \delta, \underbrace{1}_{i_0}, \delta, ..., \delta, \underbrace{0, \ldots, 0}_{m-k}\right).$$

Для μ и точки $x^\mu = x^*$ выполняются условия (42), т.е. $\mu \in C$. Применим (43):

$$\sum_{i=0, i\ne i_0}^{k} \lambda_i^* \delta + \lambda_{i_0}^* \ge 0 \Rightarrow \lambda_{i_0}^* \ge -\delta \sum_{i=0, i \ne i_0}^{k} \lambda_i^*.$$

Так как δ произвольно, то $\lambda_{i_0}^* \ge 0 \, (0 \le i_0 \le k)$.

Г) *Условие дополняющей нежесткости.* Рассмотрим произвольный индекс $i_0 \, (1 \le i_0 \le k)$. Если $f_{i_0}(x^*) = b_{i_0}$, то условие дополняющей нежесткости выполнено. Рассмотрим случай $f_{i_0}(x^*) < b_{i_0}$. Покажем, что тогда $\lambda_{i_0}^* = 0$. Рассмотрим вектор μ:

$$\mu_0 = \delta, \quad \delta > 0, \quad \mu_i = 0, \, i = 1, ..., m, i \ne i_0, \quad \mu_{i_0} = f_{i_0}(x^*) - b_{i_0},$$

то есть

$$\mu = \left(\delta, 0, \ldots, \underbrace{f_{i_0}(x^*) - b_{i_0}}_{i_o}, 0, \ldots 0\right).$$

Тогда $\mu \in C$, а соответствующая точка $x^{\mu} = x^*$. Применим (43):

$$\lambda_0^* \cdot \delta + \lambda_{i_0}^* \left(f_{i_0}\left(x^*\right) - b_{i_0}\right) \geq 0 \Rightarrow \lambda_0^* \cdot \delta \geq \lambda_{i_0}^* \left(b_{i_0} - f_{i_0}(x^*)\right).$$

Так как δ — произвольное положительное число, то имеем

$$0 \geq \lambda_{i_0}^* \left(b_{i_0} - f_{i_0}(x^*)\right),$$

а по предположению $b_{i_0} - f_{i_0}(x^*) > 0$, следовательно, $0 \geq \lambda_{i_0}^*$. В силу условия неотрицательности получаем $\lambda_{i_0}^* = 0$.

Д) *Условие минимума.* Рассмотрим для $x \in P$ вектор μ:

$$\mu_0 = f_0(x) + \delta, \quad \delta > 0, \quad \mu_i = f_i(x) - b_i,\ i = 1, \ldots, m,$$

то есть

$$\mu = \left(f_0(x) + \delta, f_1(x) - b_1, \ldots, f_m(x) - b_m\right).$$

Очевидно, что $\mu \in C$. Применим (43):

$$\lambda_0^* \left(f_0(x) + \delta\right) + \sum_{i=1}^{m} \lambda_i^* \left(f_i(x) - b_i\right) \geq 0 \quad \Rightarrow$$

$$\lambda_0^* f_0(x) + \sum_{i=1}^{m} \lambda_i^* \left(f_i(x) - b_i\right) \geq -\lambda_0^* \delta.$$

В силу произвольности δ имеем

$$\lambda_0^* f_0(x) + \sum_{i=1}^{m} \lambda_i^* \left(f_i(x) - b_i\right) \geq 0 \quad \forall x \in P. \tag{44}$$

Вычислим $\mathcal{L}\left(x^*, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right)$ и $\mathcal{L}\left(x, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right)$. Имеем

$$\mathcal{L}\left(x^*, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right) = \lambda_0^* f_0\left(x^*\right) + \sum_{i=1}^{m} \lambda_i^* \left(f_i\left(x^*\right) - b_i\right) = 0,$$

и в силу (44) для всех $x \in P$

$$\mathcal{L}\left(x, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right) = \lambda_0^* f_0(x) + \sum_{i=1}^{m} \lambda_i^* (f_i(x) - b_i) \geq 0.$$

То есть $\mathcal{L}\left(x, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right) \geq \mathcal{L}\left(x^*, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right)$ для всех $x \in P$. Таким образом, условие минимума доказано.

II. Пусть x^*— допустимая точка, для которой выполнены условия 1)-3) пункта I теоремы, причем $\lambda_0^* \neq 0$. Докажем, что x^* — решение ЗВП.

Так как $\lambda_0^* \neq 0$, то можно считать, что $\lambda_0^* = 1$. Пусть $x \in D$. Тогда в силу условия неотрицательности и допустимости точки x имеем

$$\mathcal{L}\left(x, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right) = f_0(x) + \sum_{i=1}^{k} \lambda_i^*(f_i(x) - b_i) + \sum_{i=k+1}^{m} \lambda_i^*(f_i(x) - b_i) = f_0(x) + M,$$

где

$$M = \sum_{i=1}^{k} \lambda_i^*(f_i(x) - b_i) \leq 0.$$

Учитывая условие дополняющей нежесткости и допустимость точки x^*, получим

$$\mathcal{L}\left(x^*, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right) = f_0\left(x^*\right) + \sum_{i=1}^{m} \lambda_i^*\left(f_i\left(x^*\right) - b_i\right) = f_0\left(x^*\right).$$

Применим условие минимума:

$$\mathcal{L}\left(x, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right) \geq \mathcal{L}\left(x^*, \lambda_0^*, \lambda_1^*, ..., \lambda_m^*\right) \quad \forall x \in P.$$

Отсюда для всех допустимых x получаем: $f_0(x) + M \geq f_0\left(x^*\right)$. Так как $M \leq 0$, то

$$f_0(x) \geq f_0\left(x^*\right) \quad \forall x \in D.$$

Таким образом, x^* — решение ЗВП. Теорема доказана. □

Приведем некоторые из условий регулярности для задачи выпуклого программирования.

Теорема 30

Пусть в задаче (40) выполнено хотя бы одно из дополнительных условий:

1. ограничения равенства отсутствуют, и существует такая точка $\bar{x} \in P$, что $f_i\left(\bar{x}\right) < b_i$ для всех $i = 1, \ldots, k$;

2. множество P — полиэдр, функции $f_i\left(i = 1, \ldots, k\right)$ являются линейными;

3. множество P — полиэдр, функции $f_i\left(i = s, \ldots, k\right)$ являются линейными (где $1 \leq s \leq k$), и существует такая допустимая в задаче (40) точка $\tilde{x}$, что $f_i\left(\tilde{x}\right) < b_i$ для всех $i = 1, \ldots, s$.

Если x^* — решение задачи (40), то существует набор множителей Лагранжа с $\lambda_0^* \neq 0$, удовлетворяющий необходимым условиям оптимальности в теореме Куна-Таккера.

Приведем доказательство одного из условий регулярности, а именно, условия 1, которое называется условием Слейтера.

Доказательство. Предположим, что выполнено условие Слейтера и $\lambda_0^* = 0$. Так как для $i = 1, \ldots, k$

$$f_i(\bar{x}) - b_i < 0, \quad \lambda_i^* \geq 0$$

и среди множителей Лагранжа $\lambda_1^*, \ldots, \lambda_k^*$ есть хотя бы один не ноль, то

$$\mathcal{L}(\bar{x}, 0, \lambda_1^*, ..., \lambda_m^*) = \sum_{i=1}^{k} \lambda_i^* (f_i(\bar{x}) - b_i) < 0.$$

В точке минимума x^* в силу условия дополняющей нежесткости имеем

$$\mathcal{L}(x^*, 0, \lambda_1^*, ..., \lambda_m^*) = \sum_{i=1}^{k} \lambda_i^* (f_i(x^*) - b_i) = 0.$$

Отсюда $\mathcal{L}(\bar{x}, 0, \lambda_1^*, ..., \lambda_m^*) < \mathcal{L}(x^*, 0, \lambda_1^*, ..., \lambda_m^*)$. Нарушено условие минимума для функции Ланранжа. Следовательно, $\lambda_0^* \neq 0$. □

7.4 Теорема Куна-Таккера в форме теоремы о седловой точке для функции Лагранжа

Рассмотрим еще один вариант теоремы Куна-Таккера, который используется в теории двойственности, в частности, для задач линейного программирования.

Определение 7.5 *Рассмотрим функцию $F : A \times B \to \mathbf{R}$. Точка $(a^*, b^*) \in A \times B$ — седловая точка функции F на множестве $A \times B$, если*

$$F(a^*, b) \leq F(a^*, b^*) \leq F(a, b^*) \quad \forall (a, b) \in A \times B.$$

Пусть задача выпуклого программирования регулярна, то есть можно считать, что $\lambda_0^* = 1$. Рассмотрим множество

$$Q = \{\lambda = (\lambda_1, ..., \lambda_m) : \ \lambda_i \geq 0, i = 1, \ldots, k\} \subset R^m.$$

Теорема 31
Допустимая точка x^* является решением регулярной задачи выпуклого программирования $\Leftrightarrow$ существует $\lambda^* \in Q$, такое, что (x^*, λ^*) является седловой точкой функции Лагранжа на множестве $P \times Q$.

Доказательство. $\Rightarrow$ Пусть x^*— решение ЗВП и $\lambda_0^* = 1$. Тогда из достаточных условий теоремы Куна–Таккера имеем, что x^* дает минимум функции Лагранжа по переменной x на множестве P при фиксированной переменной $\lambda = \lambda^*$. Нужно показать, что по переменной λ функция Лагранжа достигает максимума в λ^* при фиксированной переменной $x = x^*$. Вычислим $L(x^*, \lambda^*)$ и $L(x^*, \lambda)$. Имеем

$$\mathcal{L}(x^*, \lambda^*) = f_0(x^*) + \sum_{i=1}^{k} \lambda_i^* (f_i(x^*) - b_i) + \sum_{i=k+1}^{m} \lambda_i^* (f_i(x^*) - b_i) = f_0(x^*) .$$

Первая сумма равна нулю в силу условия дополняющей нежесткости, вторая сумма равна нулю, так как x^*допустимая точка. Далее,

$$\mathcal{L}(x^*, \lambda) = f_0(x^*) + \sum_{i=1}^{k} \lambda_i (f_i(x^*) - b_i) + \sum_{i=k+1}^{m} \lambda_i (f_i(x^*) - b_i) =$$

$$= f_0(x^*) + \sum_{i=1}^{k} \lambda_i (f_i(x^*) - b_i) .$$

Вторая сумма обратилась в ноль аналогично выкладке выше, а оставшаяся сумма неположительна, так как $\lambda_i \geq 0$, а $f_i(x^*) - b_i \leq 0$ для всех $i = 1, \ldots, k$. Отсюда

$$\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*)$$

для всех $\lambda \in Q$. Таким образом, (x^*, λ^*) — седловая точка функции Лагранжа на множестве $P \times Q$.

$\Leftarrow$ Пусть $x^* \in P$, $\lambda^* \in Q$, и (x^*, λ^*) — седловая точка функции Лагранжа на множестве $P \times Q$. Покажем, что тогда точка x^* — допустимая точка в ЗВП и выполнены достаточные условия оптимальности теоремы Куна–Таккера. Воспользуемся определением седловой точки

$$\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x, \lambda^*) \quad \forall (x, \lambda) \in P \times Q.$$

Вычислим $\mathcal{L}(x^*, \lambda)$ и $\mathcal{L}(x^*, \lambda^*)$:

$$\mathcal{L}(x^*, \lambda) = f_0(x^*) + \sum_{i=1}^{m} \lambda_i (f_i(x^*) - b_i) ,$$

$$\mathcal{L}(x^*, \lambda^*) = f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* (f_i(x^*) - b_i) .$$

Так как $\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*)$, то

$$\sum_{i=1}^{m} \lambda_i (f_i(x^*) - b_i) \leq \sum_{i=1}^{m} \lambda_i^* (f_i(x^*) - b_i)$$

или

$$\sum_{i=1}^{m} (\lambda_i^* - \lambda_i) \left(f_i(x^*) - b_i \right) \geq 0 \quad \forall \lambda \in Q. \tag{45}$$

Рассмотрим произвольный индекс $j, \quad 1 \leq j \leq m$, и определим вектор λ':

$$\lambda'_j = \lambda^*_j + 1, \quad \lambda'_i = \lambda^*_i, \quad i = 1, \ldots, m, \quad i \neq j,$$

то есть

$$\lambda' = \left(\lambda_1^*, \ldots, \lambda_{j-1}^*, \underbrace{\lambda_j^* + 1}_{j}, \lambda_{j+1}^*, \ldots, \lambda_m^* \right).$$

Очевидно, что вектор $\lambda' \in Q$. Подставим λ' в неравенство (45), получим

$$(-1) \left(f_j(x^*) - b_j \right) \geq 0$$

или

$$\left(f_j(x^*) - b_j \right) \leq 0, \quad j = 1, \ldots, m.$$

Таким образом, ограничения – неравенства выполнены. Покажем, что ограничения – равенства также выполнены. Рассмотрим произвольный индекс $l, \quad k+1 \leq l \leq m$, и построим вектор $\lambda'' \in Q$ следующим образом:

$$\lambda''_l = \lambda^*_l + \left(f_l(x^*) - b_l \right), \quad \lambda''_i = \lambda^*_i, \quad i = 1, \ldots, m, \quad i \neq l,$$

то есть

$$\lambda'' = \left(\lambda_1^*, \ldots, \lambda_{l-1}^*, \underbrace{\lambda_l^* + \left(f_l(x^*) - b_l \right)}_{l}, \lambda_{l+1}^*, \ldots, \lambda_m^* \right).$$

Подставим λ'' в неравенство (45), получим

$$(-1) \left(f_l(x^*) - b_l \right)^2 \geq 0$$

или

$$\left(f_l(x^*) - b_l \right) = 0, \quad l = k+1, \ldots, m.$$

Таким образом, точка x^* удовлетворяет ограничениям – неравенствам и равенствам, а так как $x^* \in P$, то точка x^* является допустимой в задаче выпуклого программирования. Проверим условия теоремы Куна–Таккера. Условие минимума функции Лагранжа выполнено в силу определения седловой точки. Условие неотрицательности выполнено в силу определения множества Q. Проверим условие дополняющей нежесткости. Зафиксируем некоторый индекс $s_0, \quad 1 \leq s_0 \leq k$. Рассмотрим следующий вектор $\tilde{\lambda} \in Q$:

$$\tilde{\lambda}_{s_0} = 0, \qquad \lambda_i = \lambda_i^*, \quad = 1, \ldots, m, \quad i \neq s_0.$$

Подставим вектор $\tilde{\lambda}$ в неравенство (45), получим

$$\lambda_{s_0}^* \left(f_{s_0}(x^*) - b_{s_0}\right) \geq 0.$$

Так как $(f_{s_0}(x^*) - b_{s_0}) \leq 0$ и $\lambda_{s_0}^* \geq 0$, то

$$\lambda_{s_0}^* \left(f_{s_0}(x^*) - b_{s_0}\right) = 0, \quad s_0 = 1, \ldots, k.$$

Итак, выполнены пункты 1) – 3) условия теоремы Куна-Таккера, и $\lambda_0^* = 1$, то есть выполнены достаточные условия оптимальности, следовательно, точка x^* — решение задачи выпуклого программирования.

Теорема доказана. □

8 Приложения выпуклого анализа

8.1 Задача об оптимальном выпуске

Рассмотрим предприятие, деятельность которого описывается следующими факторами: предприятие выпускает n видов продукции, используя для этого m различных ресурсов. Обозначим выпуск продукции через $x = (x_1, ..., x_n)$, b_i — запас i – го ресурса на предприятии $(i = 1, ..., m)$. При выпуске объема x предприятие расходует i – ый ресурс в объеме $f_i(x)$, доход предприятия при выпуске x составит $f_0\,(x)$. Кроме того, предполагаем, что предприятие может выпускать продукцию только в рамках технологически возможных выпусков: $x \in P \subset \mathbf{R}_+^n$. Будем считать, что предприятие может ничего не производить, при этом ресурсы не расходуются, доход равен 0, т.е.

$$0 \in P, \quad f_i(0) = 0, \, i = 0, 1, \ldots, m.$$

Цель предприятия состоит в том, чтобы при заданных ресурсах и в рамках технологических возможностей определить размер выпуска, при котором доход максимален.

Математическая постановка задачи:

$$\begin{array}{l} f_0(x) \to \max \\ f_i(x) \leq b_i, \; i = 1, ..., m \\ x \in P \end{array}$$

Будем считать, что P — выпуклое множество, $f_0(x)$ — вогнутая функция, $f_i(x)\,(i = 1, \ldots, m)$ — выпуклые функции. Рассмотрим требование вогнутости функции $f_0(x)$. Если $f_0(x)$ — непрерывная функция, то можно показать, что $f_0(x)$ вогнута на множестве P тогда и только тогда, когда выполняется следующее неравенство:

$$f_0\,(x + 2\triangle x) - f_0\,(x + \triangle x) \leq f_0\,(x + \triangle x) - f_0\,(x)$$

для всех $x \in P$ и $\triangle x \in \mathbf{R}^n$, таких, что $x + 2\triangle x \in P$. При $\triangle x \geq 0$ это неравенство означает, что с ростом масштабов производства прирост дохода снижается (например, в связи с проблемой реализации большого количества производимого продукта). Аналогично, условие выпуклости функций $f_i\,(x)\,(i = 1, \ldots, m)$ можно интерпретировать следующим образом: с ростом производства затраты ресурсов возрастают на каждую дополнительную единицу выпуска (например, если ресурсы неоднородны по качеству, то в начале производства расходуются ресурсы высокого качества, а при большом производстве начинают расходоваться ресурсы более низкого качества, то есть растет их расход на единицу выпуска).

Перейдем к эквивалентной задаче на минимум.

$$\begin{array}{l} -f_0(x) \to \min \\ f_i(x) \leq b_i,\ i = 1, ..., m \\ x \in P \end{array} \qquad (46)$$

Задача (46) является задачей выпуклого программирования. Можно применить теорему Куна–Таккера. Так как выполнено условие регулярности Слейтера (в качестве точки $\bar{x}$ можно взять 0), то положим $\lambda_0^* = 1$.

Пусть x^*— решение задачи (46), т.е. x^*— оптимальный выпуск. Тогда выполнено условие дополняющей нежесткости

$$\lambda_i^*(f_i(x^*) - b_i) = 0,\ i = 1, ..., m. \qquad (47)$$

Если $f_{i_0}(x^*) < b_{i_0}$, то из (47) следует, что $\lambda_{i_0}^* = 0$. Говорят, что i_0 ресурс не является дефицитным. Если $\lambda_{i_0}^* > 0$, то $f_{i_0}(x^*) = b_{i_0}$, т.е. ресурс расходуется полностью, и говорят, что i_0 ресурс является дефицитным.

Пусть $G\,(b) = -f_0\,(x^*)$ — оптимальное значение целевой функции в задаче (46) при ресурсах $b = (b_1, \ldots, b_m)$, тогда, согласно разделу 6,

$$G'\,(b) = \lambda^*.$$

Если предприятие может получить (в небольшом объеме) дополнительно один из ресурсов, то наиболее выгодно (с точки зрения увеличения дохода) получить тот ресурс, которому соответствует наибольший множитель Лагранжа. Если предприятие может получить все ресурсы в небольших объемах, то получать их надо пропорционально множителям Лагранжа, т.е. $\triangle b = k\lambda^*$, где $k > 0$. Таким образом, множители Лагранжа λ_i^* показывают степень дефицитности ресурсов.

Предположим теперь, что за приобретение дополнительных ресурсов нужно платить, и, кроме того, предприятие может продавать ненужные ресурсы. Обозначим через p_i — стоимость единицы i-го ресурса, h_i — объем i-го ресурса, который предприятие покупает или продает: $h_i > 0$ — ресурс покупается, $h_i < 0$ — ресурс продается.

Математическая постановка задачи такова:

$$\begin{aligned}
&f_0(x) - (p,h) \to \max \\
&f_i(x) \le b_i + h_i, \quad i = 1,\dots,m \\
&x \in P, \quad h_i \ge -b_i, \quad i = 1,\dots,m.
\end{aligned}$$

Условие $h_i \ge -b_i$ означает, что предприятие не может продать ресурса больше, чем оно имеет.

Перейдем к задаче на минимум.

$$\begin{aligned}
&(p,h) - f_0(x) \to \min \\
&f_i(x) \le b_i + h_i,\ i = 1,\dots,m \\
&x \in P,\ -h \le b
\end{aligned} \tag{48}$$

Задача (48) является регулярной задачей выпуклого программирования (выполнено условие Слейтера). Составим функцию Лагранжа:

$$L(x,h,\lambda) = (p,h) - f_0(x) + \sum_{i=1}^{m} \lambda_i (f_i(x) - b_i + h_i)$$

Пусть (x^*, h^*)— решение задачи (48), λ^* — соответствующий множитель Лагранжа. Выпишем условия теоремы Куна-Таккера:

1. условие минимума

$$L(x,h,,\lambda^*) \ge L(x^*,h^*,\lambda^*), \quad x \in P, \quad h \ge -b;$$

2. условие дополняющей нежесткости

$$\lambda_i^*(f_i(x^*) - b_i + h_i) = 0,\ i = 1,...,m;$$

3. условие неотрицательности $\lambda_i^* \ge 0,\ i = 1,...,m.$

Предположим, что предприятию невыгодно продавать ресурсы в полном объеме, т.е. $h^* > -b$. Тогда из условия минимума следует, что

$$L'_h(x^*,h^*,\lambda^*) = 0.$$

Отсюда получаем, что $\lambda^* = p$, то есть в этой задаче множитель Лагранжа совпадает с ценой соответствующего ресурса.

8.2 Неоклассическая задача потребления

Рассмотрим некоторого потребителя, который приобретает n товаров, цена i-го товара равна p_i. Бюджет потребителя составляет M^* денежных единиц. Предпочтения потребителя описывает некоторая функция полезности $u\,(x_1,...,x_n)$, где x_i — количество i-го товара, т.е. полезность набора $x = (x_1,...,x_n)$ для потребителя равна $u\,(x_1,...,x_n)$.

Предполагаем, что функция полезности обладает следующими свойствами:

1. строго возрастает по каждой переменной: $\frac{\partial u(x)}{\partial x_i} > 0$, $i = 1, ..., n$;

2. u''_{xx}— отрицательно определенная матрица, то есть функция полезности — вогнутая функция.

Цель потребителя состоит в том, чтобы в рамках бюджета приобрести набор товаров $(x_1, ..., x_n)$ с максимальной полезностью.

Математическая постановка задачи такова:

$$\begin{array}{l} u(x) \to \max \\ (p, x) \le M^*, \quad x \ge 0 \end{array}$$

Здесь $p = (p_1, \ldots, p_n)$ и запись $(p, x) \le M^*$ эквивалентна $\sum_{i=1}^n p_i \cdot x_i \le M^*$. Условие $x \ge 0$ означает, что каждая координата вектора x неотрицательна.

Перейдем к задаче на минимум, получим задачу выпуклого программирования, для которой выполнено условие Слейтера, то есть можно положить $\lambda_0^* = 1$. Заметим, что в этом случае точки, удовлетворяющие необходимым условиям оптимальности, будут являться глобальными решениями в задаче.

Запишем функцию Лагранжа и условия теоремы Куна-Таккера

$$L(x, \lambda_1) = -u(x) + \lambda_1 \left((p, x) - M^* \right)$$

1. $L(x^*, \lambda_1^*) \le L(x, \lambda_1^*)$ для всех $x \ge 0$;
2. $\lambda_1^* \ge 0$;
3. $\lambda_1^* \left((p, x) - M^* \right) = 0$.

Очевидно, что на оптимальном решении выполняется $\sum_{i=1}^n p_i \cdot x_i^* = M^*$, так как в противном случае потребитель может приобрести на оставшуюся сумму денег дополнительное количество товаров, и в силу свойства возрастания функции полезности увеличить свою полезность.

Так как $L(x, \lambda_1^*)$ — выпуклая по x функция, то можно применить следствие 2 теоремы 38 для определения точки минимума функции Лагранжа. Получим следующие условия:
если $x_j^* = 0$, то $\frac{\partial L}{\partial x_j}(x^*, \lambda_1^*) \ge 0$;
если $x_j^* > 0$, то $\frac{\partial L}{\partial x_j}(x^*, \lambda_1^*) = 0$.

Предположим, что хотя бы один товар куплен, т.е. $\exists j_0 : x_{j_0}^* > 0$, тогда $\frac{\partial L}{\partial x_{j_0}}(x^*, \lambda_1^*) = 0$ или $\frac{\partial u(x^*)}{\partial x_{j_0}} = \lambda_1^* p_{j_0}$. Отсюда

$$\frac{\partial u(x^*)}{\partial x_{j_0}} / p_{j_0} = \lambda_1^* \quad \forall j_0 : x_{j_0}^* > 0. \tag{49}$$

Значение $\partial u(x^*)/\partial x_{j_0}$ представляет полезность единицы товара j_0, p_{j_0} — цена j-го товара, то есть количество денежных единиц на единицу товара j_0, таким образом, множитель Лагранжа λ_1^*, равный отношению этих величин, характеризует полезность на одну денежную единицу.

Кроме того, из соотношения (49) следует вывод: если i-ый товар в a раз дороже товара j, то уменьшение товара i на одну единицу компенсируется (с точки зрения функции полезности) увеличением товара j на a единиц:

$$\frac{\partial u(x^*)}{\partial x_i} = a\frac{\partial u(x^*)}{\partial x_j}.$$

***Пример* 19** *Рассмотрим функцию полезности $u\left(x_1, x_2\right) = Ax_1^{\alpha}x_2^{1-\alpha}$, где $\alpha \in (0,1)$ (функцию такого вида называют функцией Кобба-Дугласа). При заданных ценах p_1, p_2 и бюджете M^* найти набор товаров (x_1^*, x_2^*), максимизирующий полезность потребителя.*

Решение. Составим функцию Лагранжа

$$L(x,\lambda_1) = -Ax_1^{\alpha}x_2^{1-\alpha} + \lambda_1\left(p_1x_1 + p_2x_2 - M^*\right).$$

Очевидно, что для данной функции полезности оптимальное решение строго положительно, поэтому мы можем дифференцировать функцию Лагранжа в этой точке. Условия стационарности имеют вид:

$$L'_{x_1}\left(x,\lambda_1\right) = -\alpha Ax_1^{\alpha-1}x_2^{1-\alpha} + \lambda_1p_1 = 0,$$

$$L'_{x_2}\left(x,\lambda_1\right) = -\left(1-\alpha\right)Ax_1^{\alpha}x_2^{-\alpha} + \lambda_1p_2 = 0.$$

Отсюда

$$\alpha A\left(\frac{x_2}{x_1}\right)^{1-\alpha} = \lambda_1p_1,$$

$$\left(1-\alpha\right)A\left(\frac{x_2}{x_1}\right)^{-\alpha} = \lambda_1p_2.$$

Так как $\lambda_1 \neq 0$, то разделим первое уравнение на второе:

$$\frac{x_2}{x_1} = \frac{\left(1-\alpha\right)p_1}{\alpha p_2} \Rightarrow x_2 = x_1\frac{\left(1-\alpha\right)p_1}{\alpha p_2}.$$

Подставим в бюджетное ограничение, которое выполнено, как показано выше, в виде равенства:

$$p_1x_1 + p_2x_1\frac{\left(1-\alpha\right)p_1}{\alpha p_2} = M^*.$$

Находим оптимальное решение:

$$x_1^* = \frac{\alpha M^*}{p_1}, \quad x_2^* = \frac{(1-\alpha) M^*}{p_2}. \tag{50}$$

Из (50) получаем, что оптимальный объем покупки i-го товара прямо пропорционален бюджету M^* и параметру, характеризующему важность товара для потребителя (α — для 1-го товара и $(1-\alpha)$ — для 2-го товара), и обратно пропорционален цене.

8.3 Задача об оптимальном портфеле ценных бумаг

На рынке есть n видов ценных бумаг (акций). Игрок (инвестор) на рынке ценных бумаг желает приобрести бумаги на сумму K^0. Обозначим через x_i долю начального капитала K^0 игрока, которая идет на приобретение бумаг i-го типа:

$$\sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0, \quad i = 1, ...n.$$

Вектор $x = (x_1, ..., x_n)$ назовем портфелем ценных бумаг. Предположим, что доход, который приносят бумаги за некоторый промежуток времени, есть случайная величина, обозначим доходность i-ой бумаги через ξ_i. Тогда доходность портфеля $x = (x_1, ..., x_n)$ — случайная величина $\xi(x)$:

$$\xi(x) = K^0 \left(\sum_{i=1}^{n} \xi_i x_i \right).$$

Предположим, что известны математическое ожидание (среднее значение) μ_i доходностей ξ_i и матрица ковариаций $C_{n\times n} : c_{ij} = cov(\xi_i, \xi_j) = E(\xi_i - \mu_i)(\xi_j - \mu_j)$. Заметим, что ковариация c_{ij} характеризует зависимость случайных величин ξ_i и ξ_j. (Вообще говоря, случайные величины ξ_i $(i = 1, \ldots, n)$ являются зависимыми, увеличение или уменьшение доходности одной из бумаг оказывает влияние на изменение доходностей остальных бумаг). Известно, что матрица ковариаций C симметрическая и неотрицательно определенная (далее будем для удобства считать, что матрица C является положительно определенной).

Вычислим математическое ожидание и дисперсию доходности $\xi(x)$ портфеля $x = (x_1, \ldots, x_n)$:

$$E\xi(x) = K^0 \left(\sum_{i=1}^{n} E(\xi_i x_i) \right) = K^0 \left(\sum_{i=1}^{n} \mu_i x_i \right) = K^0 (\mu, x),$$

где $\mu = (\mu_1, ..., \mu_n)$,

$$D\xi(x) = \left(K^0\right)^2 D \left(\sum_{i=1}^{n} \xi_i x_i \right) = \left(K^0\right)^2 (Cx, x).$$

Так как дисперсия есть среднее значение квадрата уклонения случайной величины от ее среднего значения, то дисперсия $D\xi$ выступает как мера риска портфеля.

Задача игрока состоит в том, чтобы, с одной стороны, максимизировать среднее значение доходности, с другой, минимизировать риск. Следовательно, игрок решает следующую двухкритериальную задачу:

$$\begin{array}{l} E\xi\,(x) \to \max \\ D\xi\,(x) \to \min, \quad x \in D \end{array}$$

или эквивалентную ей задачу

$$\begin{array}{l} (\mu, x) \to \max \\ (Cx, x) \to \min, \quad x \in D \end{array}$$

где допустимое множество D имеет вид

$$D = \left\{ x \in \mathbf{R}^n : \ \sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0, \quad i = 1, ...n \right\}$$

или

$$D = \left\{ x \in \mathbf{R}^n : \ \sum_{i=1}^{n} x_i = 1 \right\}.$$

Поясним, какой смысл имеют отрицательные значения координат портфеля x. Если $x_i > 0$, то это означает рекомендацию вложить сумму $x_i K^0$ в бумаги i-го вида. Если $x_i < 0$, то это означает рекомендацию взять в долг сумму $x_i K^0$, направить эту сумму на покупку бумаг другого типа, а затем игрок возвращает заемщику сумму $x_i K^0$ плюс тот доход, который бы принесли бумаги i-го типа (то есть, заемщик получит величину $x_i K^0\,(1 + \xi_i)$)(такую операцию называют “short-sale” [9, 10]).

Решение двухкритериальной задачи существует достаточно редко, поэтому возникает необходимость рассматривать задачи только с одним критерием. Мы рассмотрим следующие стандартные задачи:

1) задача минимизации риска при фиксированной средней доходности

$$\begin{array}{l} (Cx, x) \to \min \\ (\mu, x) = \mu^* \\ x \in D \end{array} \tag{51}$$

2) задача оптимизации смешанного функционала

$$\begin{array}{l} \frac{1}{2\alpha}\,(Cx, x) - (\mu, x) \to \min \\ x \in D, \end{array}$$

где α — некоторый параметр, $\alpha > 0$.

8.3.1 Портфель без ограничений на среднюю доходность

Рассмотрим сначала задачу (51) в упрощенной форме, а именно, будем минимизировать дисперсию (риск) без ограничений на среднюю доходность с возможностью брать денежные суммы в долг:

$$\begin{aligned} &(Cx, x) \to \min \\ &(I, x) = 1. \end{aligned} \tag{52}$$

Здесь I — вектор, состоящий из единиц:

$$I = \left(\underbrace{1, \ldots, 1}_{n} \right).$$

Целевая функция в задаче (52) является строго выпуклой, а значит, бесконечно растущей на допустимом множестве $(I, x) = 1$. Следовательно, в задаче (52) существует абсолютный минимум. Применим условия оптимальности для задач с ограничениями в виде равенств. Выполнено условие регулярности, следовательно, $\lambda_0 \neq 0$. Выпишем условие стационарности для функции Лагранжа $L = (Cx, x) + \lambda_1 \left((I, x) - 1\right)$:

$$L'_x = 2Cx + \lambda_1 I = 0.$$

Отсюда

$$x^* = -\frac{\lambda_1}{2} C^{-1} I.$$

Подставим в ограничение, получим

$$\left(I, -\frac{\lambda_1}{2} C^{-1} I \right) = 1 \quad \Rightarrow \quad -\frac{\lambda_1}{2} = \frac{1}{(I, C^{-1} I)} \quad \Rightarrow$$

$$x^* = \frac{C^{-1} I}{(I, C^{-1} I)}. \tag{53}$$

Так как мы показали, что в задаче есть абсолютный минимум, а необходимым условиям минимума удовлетворяет единственная точка, то она и дает минимум. Вычислим значение целевой функции и среднюю доходность для оптимального портфеля x^*.

$$(Cx^*, x^*) = \frac{(CC^{-1} I, C^{-1} I)}{(I, C^{-1} I)^2} = \frac{1}{(I, C^{-1} I)},$$

$$(\mu, x^*) = \left(\mu, \frac{C^{-1} I}{(I, C^{-1} I)} \right) = \frac{1}{(I, C^{-1} I)} \left(\mu, C^{-1} I \right).$$

Замечание 17 *Показать, что* $x^* \in absmin(52)$ *можно, используя достаточные условия в задаче с ограничениями в виде равенств и выпуклость целевой функции.*

Пусть доходности бумаг — некоррелированные случайные величины, т.е.

$$C = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n^2 \end{pmatrix}$$

Тогда C^{-1} имеет вид

$$C^{-1} = \begin{pmatrix} \sigma_1^{-2} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n^{-2} \end{pmatrix}$$

Найдем оптимальный портфель, используя формулу (53):

$$C^{-1}I = \begin{pmatrix} \sigma_1^{-2} \\ \vdots \\ \sigma_n^{-2} \end{pmatrix}, \quad \big(I, C^{-1}I\big) = \sum_{i=1}^{n} \sigma_i^{-2} \quad \Rightarrow$$

$$x^* = \frac{1}{\sum_{i=1}^{n} \sigma_i^{-2}} \begin{pmatrix} \sigma_1^{-2} \\ \vdots \\ \sigma_n^{-2} \end{pmatrix}.$$

Обозначим

$$\bar{\sigma} = \max_{i=1,\ldots,n} \sigma_i.$$

Тогда можно оценить оптимальное значение целевой функции (минимальную дисперсию портфеля):

$$(Cx^*, x^*) = \frac{1}{\sum_{i=1}^{n} \sigma_i^{-2}} \le \frac{\bar{\sigma}^2}{n}.$$

Таким образом, дисперсия портфеля уменьшается с увеличением n и в пределе стремится к нулю:

$$\lim_{n\to\infty} (Cx^*, x^*) = \lim_{n\to\infty} \frac{\bar{\sigma}^2}{n} = 0.$$

8.3.2 Портфель с ограничением на среднюю доходность

Рассмотрим теперь задачу минимизация риска при фиксированной средней доходности:

$$\begin{aligned} &(Cx, x) \to \min \\ &(\mu, x) = \mu^* \\ &(I, x) = 1 \end{aligned} \tag{54}$$

Существование абсолютного минимума в задаче (54) следует из той же теоремы, что и для задачи (52).

Проверим условия регулярности. Градиенты ограничений равны μ и I. Условие регулярности не выполнено тогда и только тогда, когда $\mu_i = \mu\,(i = 1, \dots, n)$, т.е. тогда и только тогда, когда доходности всех бумаг одинаковы. Но в этом случае для любого допустимого портфеля x имеем

$$(\mu, x) = \mu\,(I, x) = \mu.$$

Следовательно, ограничение на среднюю доходность выполнено тогда и только тогда, когда доходности всех бумаг равны μ^*. Таким образом, если условие регулярности не выполнено, то задача (54) сводится к задаче (52).

Выпишем функцию Лагранжа с $\lambda_0 = 1$ и необходимые условия локального минимума для задачи с ограничениями в виде равенств:

$$L = (Cx, x) + \lambda_1\left((\mu, x) - \mu^*\right) + \lambda_2\left((I, x) - 1\right),$$

$$L'_x = 2Cx + \lambda_1\mu + \lambda_2 I = 0 \Rightarrow x^* = -\frac{\lambda_1}{2}C^{-1}\mu - \frac{\lambda_2}{2}C^{-1}I.$$

Подставим выражение для x^* в ограничения:

$$\begin{cases} \left(\mu, -\frac{\lambda_1}{2}C^{-1}\mu - \frac{\lambda_2}{2}C^{-1}I\right) = \mu^* \\ \left(I, -\frac{\lambda_1}{2}C^{-1}\mu - \frac{\lambda_2}{2}C^{-1}I\right) = 1 \end{cases}$$

или

$$\begin{cases} -\frac{\lambda_1}{2}\left(\mu, C^{-1}\mu\right) - \frac{\lambda_2}{2}\left(\mu, C^{-1}I\right) = \mu^* \\ -\frac{\lambda_1}{2}\left(I, C^{-1}\mu\right) - \frac{\lambda_2}{2}\left(I, C^{-1}I\right) = 1. \end{cases} \tag{55}$$

Так как матрица C положительно определена, то можно определить скалярное произведение

$$(u, v)_C = \left(u, C^{-1}v\right).$$

Тогда

$$\|u\|_C^2 = \left(u, C^{-1}u\right).$$

Обозначим

$$\triangle = \left(I, C^{-1}I\right)\left(\mu, C^{-1}\mu\right) - \left(I, C^{-1}\mu\right)^2 = \|I\|_C^2\,\|\mu\|_C^2 - (I, \mu)_C^2\,.$$

Из неравенства Коши–Буняковского следует, что $\triangle \geq 0$ и $\triangle = 0 \quad \Leftrightarrow \mu_i = \mu \quad \forall i = 1, ..., n$. В этом случае, как уже было показано, задача (54) сводится к задаче (52). Итак, можно считать, что $\triangle > 0$. Из системы (55) выразим λ_1 и λ_2 и подставим в выражение для x^*

$$x^* = \mu^* \frac{1}{\triangle} C^{-1} \left(\mu \left\| I \right\|_C^2 - I \left(I, \mu \right)_C \right) + \frac{1}{\triangle} C^{-1} \left(I \left\| \mu \right\|_C^2 - \mu \left(I, \mu \right)_C \right).$$

Таким образом, оптимальный портфель линейно зависит от среднего значения доходности μ^*. Вычислим значение дисперсии на оптимальном портфеле:

$$\sigma^{*2} \left(\mu^* \right) = \left(C x^*, x^* \right) = \frac{\left\| I \right\|_C^2 \left(\mu^* \right)^2 - 2 \mu^* \left(I, \mu \right)_C + \left\| \mu \right\|_C^2}{\triangle}.$$

Заметим, что дисперсия является квадратичной функцией от средней доходности μ^*, графиком этой функции является парабола, ветви которой направлены вверх. Найдем минимум дисперсии. Для этого достаточно найти вершину параболы и вычислить значение σ^{*2} в этой точке:

$$\frac{d}{d\mu^*} \left(\sigma^{*2} \left(\mu^* \right) \right) = \frac{1}{\triangle} \left(2 \left\| I \right\|_C^2 \mu^* - 2 \left(I, \mu \right)_C \right),$$

$$\frac{d}{d\mu^*} \left(\sigma^{*2} \left(\mu^* \right) \right) = 0 \Rightarrow \mu^* = \frac{\left(I, \mu \right)_C}{\left\| I \right\|_C^2}.$$

Обозначим

$$\mu_0^* = \frac{\left(I, \mu \right)_C}{\left\| I \right\|_C^2} = \frac{1}{\left(I, C^{-1} I \right)} \left(\mu, C^{-1} I \right).$$

Вычислим

$$\sigma^{*2} \left(\mu_0^* \right) = \frac{1}{\left\| I \right\|_C^2} = \frac{1}{\left(I, C^{-1} I \right)}.$$

Отметим, что получили такие же результаты, как и в задаче (52).

***Пример* 20** *Найти портфель заданной доходности μ^* с минимальной дисперсией, составленный из 3-х видов ценных бумаг: доходность бумаг 1 вида — детерминированная величина $\mu_1 = 2$, две другие доходности — некоррелированные случайные величины со средними значениями и дисперсиями*

$$\mu_2 = 4, \quad \sigma_2^2 = 4$$

и

$$\mu_3 = 10, \quad \sigma_3^2 = 16$$

соответственно.

Решение. Математическая формализация:

$$4x_2^2 + 16x_3^2 \to \min \tag{56}$$

$$\begin{gathered} 2x_1 + 4x_2 + 10x_3 = \mu^*, \\ x_1 + x_2 + x_3 = 1. \end{gathered} \tag{57}$$

Очевидно, что решение задачи существует при $2 \le \mu^* \le 10$. Легко видеть, что при отсутствии ограничений (57) минимальное значение целевой функции (56) равно 0 и достигается при $x_1 = x_2 = 0$. Поэтому, если $\mu^* = 2$, то допустимый портфель единственный, соответственно, он и будет являться оптимальным: $x^* = (1, 0, 0)$. Если $\mu^* = 10$, то аналогично, допустимый портфель единственен и являтся оптимальным: $x^* = (0, 0, 1)$.

Рассмотрим теперь случай $2 < \mu^* < 10$. Составим функцию Лагранжа. Так как выполнено условие регулярности, то можно считать, что $\lambda_0 = 1$:

$$L(x, \lambda) = 4x_2^2 + 16x_3^2 + \lambda_1 (x_1 + x_2 + x_3 - 1) + \lambda_2 (2x_1 + 4x_2 + 10x_3 - \mu^*).$$

Запишем условие стационарности:

$$\begin{gathered} L'_{x_1}(x, \lambda) = \lambda_1 + 2\lambda_2, \\ L'_{x_2}(x, \lambda) = 8x_2 + \lambda_1 + 4\lambda_2, \\ L'_{x_3}(x, \lambda) = 32x_3 + \lambda_1 + 10\lambda_2. \end{gathered}$$

Отсюда

$$\lambda_1 = -2\lambda_2, \quad x_2 = -\frac{1}{4}\lambda_2, \quad x_3 = -\frac{1}{4}\lambda_2.$$

Подставим в ограничения (57), получим

$$x_1^* = \frac{7 - \mu^*}{5}, \quad x_2^* = x_3^* = \frac{\mu^* - 2}{10}. \tag{58}$$

Так как $\mu^* > 2$, то $x_2^* > 0, \quad x_3^* > 0$. Если $\mu^* \le 7$, то $x_1^* \ge 0$, то есть при $2 < \mu^* < 7$ инвестор приобретает бумаги всех 3-х видов (хотя бумаги 1-го вида имеют низкую доходность, но с точки зрения целевой функции (56) они являются выгодными, так как их дисперсия равна нулю). Если $10 > \mu^* > 7$, то в (58) $x_1^* < 0$, то есть инвестору рекомендуется операция “short-sale”, описанная выше. Если такая операция запрещена, то в оптимальном портфеле $x_1^* = 0$ и из ограничений (57) находим x_2^* и x_3^*:

$$x_2^* = \frac{10 - \mu^*}{8}, \quad x_3^* = \frac{\mu^* - 2}{8}.$$

***Пример* 21** *Найти портфель с минимальной дисперсией и фиксированной средней доходностью $\mu^* = 1.5$, составленный из бумаг 2-х видов, если $\mu_1 = 1, \quad \mu_2 = 2,$*

$$C = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix},$$

и операции “short-sale” запрещены.

Решение. Математическая постановка задачи:

$$D\left(x_1, x_2\right) = x_1^2 - 2x_1x_2 + 2x_2^2 \to \min$$

$$x_1 + 2x_2 = 1.5, \quad x_1 + x_2 = 1, \quad x_1, x_2 \geq 0.$$

В задаче выполнено условие регулярности, поэтому составляем функцию Лагранжа с $\lambda_0 = 1$:

$$L\left(x, \lambda\right) = \left(Cx, x\right) + \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 \left(\left(I, x\right) - 1\right) + \lambda_4 \left(\left(\mu, x\right) - \mu^*\right) =$$

$$= x_1^2 - 2x_1x_2 + 2x_2^2 - \lambda_1 x_1 - \lambda_2 x_2 + \lambda_3 \left(x_1 + x_2 - 1\right) + \lambda_4 \left(x_1 + 2x_2 - 1.5\right).$$

Условия стационарности:

$$\begin{array}{c} L'_{x_1}\left(x, \lambda\right) = 2x_1 - 2x_2 - \lambda_1 + \lambda_3 + \lambda_4 = 0, \\ L'_{x_2}\left(x, \lambda\right) = -2x_1 + 4x_2 - \lambda_2 + \lambda_3 + 2\lambda_4 = 0; \end{array} \tag{59}$$

условия дополняющей нежесткости:

$$\lambda_1 x_1 = \lambda_2 x_2 = 0,$$

условия неотрицательности:

$$\lambda_1 \geq 0, \quad \lambda_2 \geq 0.$$

Попробуем найти оптимальный портфель с двумя положительными компонентами. Тогда из условий дополняющей нежесткости имеем:

$$\lambda_1 = \lambda_2 = 0.$$

Из (59) получим

$$2x_1 - 2x_2 + \lambda_3 + \lambda_4 = 0, \quad -2x_1 + 4x_2 + \lambda_3 + 2\lambda_4 = 0.$$

Отсюда

$$x_1 = -1.5\lambda_3 - 2\lambda_4, \quad x_2 = -\lambda_3 - 1.5\lambda_4.$$

Подставим в ограничения:

$$-1.5\lambda_3 - 2\lambda_4 - \lambda_3 - 1.5\lambda_4 = 1, \quad -1.5\lambda_3 - 2\lambda_4 + 2\left(-\lambda_3 - 1.5\lambda_4\right) = 1.5.$$

Найдем λ_3 и λ_4:

$$\lambda_3 = -\frac{5}{9}, \quad \lambda_4 = \frac{1}{9},$$

отсюда

$$x_1^* = \frac{11}{18}, \quad x_2^* = \frac{7}{18}.$$

8.3.3 Оптимальный портфель для смешанного функционала

Рассмотрим задачу со смешанным функционалом:

$$\left\{ \begin{array}{c} (Cx, x) \,/\, (2\alpha) - (\mu, x) \to \min \\ (I, x) = 1. \end{array} \right. \tag{60}$$

Так как $\alpha > 0$, то целевая функция $f_0(x) = (Cx, x) \,/\, (2\alpha) - (\mu, x)$ является выпуклой, бесконечно растущей функцией, поэтому в задаче (60) существует глобальное решение. Найдем решение, для этого составим функцию Лагранжа

$$L(x, \lambda) = (Cx, x) \,/\, (2\alpha) - (\mu, x) + \lambda_1 ((I, x) - 1),$$

здесь $\lambda_0 = 1$, так как условие регулярности выполнено. Необходимое условие оптимальности первого порядка:

$$L'_x = Cx/\alpha - \mu + \lambda_1 I = 0.$$

Отсюда

$$x^* = \alpha C^{-1} (\mu - \lambda_1 I).$$

Подставим в ограничение, получим

$$\lambda_1 = \left(\left(I, C^{-1}\mu\right) - \frac{1}{\alpha} \right) / \left(I, C^{-1} I\right) \quad \Rightarrow$$

$$x^*(\alpha) = \alpha C^{-1} \left(\mu - I \frac{(I, C^{-1}\mu)}{(I, C^{-1}I)} \right) + \frac{C^{-1} I}{(I, C^{-1} I)}.$$

Отметим, что оптимальное решение линейно зависит от параметра α, который можем рассматривать как меру осторожности (или склонности к риску) игрока. Чем меньше α, тем осторожнее игрок. Заметим, что

$$x^*(\alpha) \to \frac{C^{-1} I}{(I, C^{-1} I)}, \quad \alpha \to 0,$$

то есть решение задачи (60) при $\alpha \to 0$ сходится к решению задачи (52).

Мы рассмотрели некоторые из возможных постановок задачи об оптимальном портфеле. Другие постановки можно найти, например, в [3, 9, 10]

9 Приложение

I. Формула Тейлора для функции многих переменных [8].

Рассмотрим функцию $f : U \to \mathbf{R}$, где $U \subset \mathbf{R}^n$, $f \in C^{(m)}(U, \mathbf{R})$. Пусть отрезок $[x, x+\delta]$ $(\delta \in \mathbf{R}^n)$ полностью содержится в U, тогда имеет место формула Тейлора

а) с остаточным членом в форме Лагранжа:

$$f(x_1+\delta_1, \ldots, x_n+\delta_n) - f(x_1, \ldots, x_n) =$$

$$= \sum_{k=1}^{m-1} \frac{1}{k!} (\delta_1\partial_1 + \ldots + \delta_n\partial_n)^k f(x) + \frac{1}{n!} (\delta_1\partial_1 + \ldots + \delta_n\partial_n)^m f(x+\theta\delta). \quad (61)$$

где $0 < \theta < 1$, а $\partial_i f(x)$ $(i = 1, \ldots, n)$ — частная производная функции f в точке x;

б) с остаточным членом в форме Пеано:

$$f(x_1+\delta_1, \ldots, x_n+\delta_n) - f(x_1, \ldots, x_n) =$$

$$= \sum_{k=1}^{m} \frac{1}{k!} (\delta_1\partial_1 + \ldots + \delta_n\partial_n)^k f(x) + o(\|\delta\|^m). \quad (62)$$

Если функция $f \in C^{(2)}(U, \mathbf{R})$, то формулу (61) можно записать в виде

$$f(x_1+\delta_1, \ldots, x_n+\delta_n) - f(x_1, \ldots, x_n) =$$

$$= \delta_1\partial_1 f(x) + \ldots + \delta_n\partial_n f(x) + \frac{1}{2} (\delta_1\partial_1 + \ldots + \delta_n\partial_n)^2 f(x+\theta\delta)$$

или

$$f(x+\delta) - f(x) = (f'(x), \delta) + \frac{1}{2} (f''(x+\theta\delta)\,\delta, \delta),$$

а формулу (62) в виде

$$f(x_1+\delta_1, \ldots, x_n+\delta_n) - f(x_1, \ldots, x_n) =$$

$$= \sum_{k=1}^{2} \frac{1}{k!} (\delta_1\partial_1 + \ldots + \delta_n\partial_n)^k f(x) + o\left(\|\delta\|^2\right)$$

или

$$f(x+\delta) - f(x) = (f'(x), \delta) + \frac{1}{2} (f''(x)\,\delta, \delta) + o\left(\|\delta\|^2\right).$$

II. Векторзначные функции многих переменных. Дифференцируемость. Теорема об обратном отображении [8].

Рассмотрим отображение $F:\mathbf{R}^n\to\mathbf{R}^m$, т.е.

$$F\left(x_1,\ldots,x_n\right)=\left(f_1\left(x_1,\ldots,x_n\right),\ldots,f_m\left(x_1,\ldots,x_n\right)\right).$$

Отображение $F:\mathbf{R}^n\to\mathbf{R}^m$ называется дифференцируемым в точке $x\in\mathbf{R}^n$, если

$$F\left(x+h\right)-F\left(x\right)=F'\left(x\right)h+\alpha\left(x;h\right),$$

где $F'\left(x\right):\mathbf{R}^n\to\mathbf{R}^m$ — линейное отображение относительно h, которое называется производным отображением отображения F в точке x; а $\left\|\alpha\left(x;h\right)\right\|_{\mathbf{R}^m}=o\left(\|h\|_{\mathbf{R}^n}\right)$ при $h\to 0$. Можно выписать покоординатное представление $F'\left(x\right)$:

$$F'\left(x\right)=\begin{pmatrix}\frac{\partial f_1(x)}{\partial x_1} & \ldots & \frac{\partial f_1(x)}{\partial x_n}\\ & \ddots & \\ \frac{\partial f_m(x)}{\partial x_1} & \ldots & \frac{\partial f_m(x)}{\partial x_n}\end{pmatrix}.$$

Матрица $\left(\partial f_j\left(x\right)/\partial x_i\right)\left(i=1,\ldots,n,j=1,\ldots,m\right)$ из частных производных координатных функций отображения F в точке $x\in\mathbf{R}^n$ называется матрицей Якоби отображения F в точке x. Якобианом отображения F в точке $x\in\mathbf{R}^n$ называется определитель матрицы Якоби (когда она квадратная).

Отображение $F:\mathbf{R}^n\to\mathbf{R}^m$ непрерывно дифференцируемо, если матрица Якоби $\left(\partial f_j\left(x\right)/\partial x_i\right)\left(i=1,\ldots,n,j=1,\ldots,m\right)$ существует и непрерывна.

Теорема 32 *(Теорема об обратном отображении) Рассмотрим непрерывно дифференцируемое отображение F из пространства $\mathbf{R}^k$ в пространство $\mathbf{R}^k$, $F\left(u^*\right)=v^*$. Если якобиан F в точке u^*не равен 0, то $\exists\varepsilon>0$, $\delta>0$, $\kappa>0$, такие, что*

$$\forall v\in V_\varepsilon(v^*)\quad\exists!\ u=u(v)\in U_\delta(u^*),$$

при этом

$$u(v^*)=v^*,\quad F(u(v))=v$$

и

$$\|u(v)-u^*\|\le\kappa\left\|v-v^*\right\|.$$

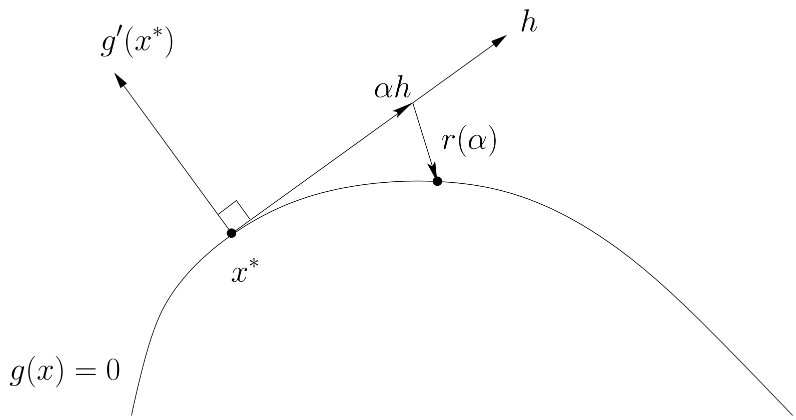

Рис. 10: к теореме Люстерника ($m = 1$).

III. Теорема Люстерника [2, 6].

Теорема 33 *Пусть* $g_1(x), \ldots, g_m(x)$ *— функции, непрерывно дифференцируемые в окрестности точки* $x^* \in \mathbf{R}^n$ *;* $g_1(x^*) = ... = g_m(x^*) = 0$ *и* $g_1'(x^*), \ldots, g_m'(x^*)$ *линейно независимы. Рассмотрим векторы* $h \in \mathbf{R}^n$*, такие, что* $(g_i'(x^*), h) = 0 \quad \forall i = 1, \ldots, m$*. Тогда для всех достаточно малых по модулю* $\alpha \in \mathbf{R}$ *существует* $r(\alpha) \in \mathbf{R}^n$*, такое, что*

1. $g_i(x^* + \alpha h + r(\alpha)) = 0 \quad \forall i = 1, \ldots, m$*;*

2. $\lim_{\alpha \to 0} \frac{\|r(\alpha)\|}{\alpha} = 0$*.*

IV. Ранг матрицы. Теорема о ранге [7].

Матрица $A_{n \times s}$ (ненулевая) имеет ранг p, если A имеет по меньшей мере одну невырожденную подматрицу порядка p (т.е. определитель которой отличен от 0), а все квадратные подматрицы A более высоких порядков вырождены.

Теорема 34 *(Теорема о ранге матрицы)*

1. Ранг матрицы равен максимальному числу линейно независимых вектор-строк матрицы A.

2. Ранг матрицы A не меняется при следующих преобразованиях:

а) при перемене местами двух строк;

б) при умножении одной строки на число $c \neq 0$*;*

в) при сложении одной строки с другой;

г) при транспонировании матрицы A.

Так как ранг A равен рангу матрицы A^T, то утверждения, сформулированные в теореме для строк, справедливы и для столбцов.

Список литературы

[1] Акулич И.Л. Математическое программирование в примерах и задачах. — М., Лань, 2009. — 352 с.

[2] Алексеев В.М., Галеев Э.М., Тихомиров В.М. Сборник задач по оптимизации. — М., Физматлит, 2005. — 256 с.

[3] Афанасьева Л.Г. Очерки исследования операций. — М., Изд-во Центра прикл. исследований при механико-матем. ф-те МГУ, 2007. — 176 с.

[4] Ашманов С.А., Тимохов А.В. Теория оптимизации в задачах и упражнениях. — М., Наука, 1991. — 448 с.

[5] Вентцель Е.С. Исследование операций. Задачи, принципы, методология. — М., КноРус, 2010. — 192 с.

[6] Галеев Э.М. Оптимизация: теория, примеры, задачи. — М., Эдиториал УРСС, 2010. — 320 с.

[7] Гантмахер Ф.Р. Теория матриц. — М., Физматлит, 2010. — 560 с.

[8] Зорич В.А. Математический анализ. — М., Наука, 1981. — Т. 1. — 544 с.

[9] Математические методы и модели исследования операций. Под редакцией В.А. Колемаева. — М., Юнити–Дана, 2008. — 592 с.

[10] Новоселов А.А. Математическое моделирование финансовых рисков. — Новосибирск, Наука, 2001. — 102 с.

[11] Сухарев А.Г., Тимохов А.В., Федоров В.В. Курс методов оптимизации. М., Физматлит, 2005. — 368 с.

МАНИТА Лариса Анатольевна

Условия оптимальности в конечномерных нелинейных задачах оптимизации

Редактор Е.С. Резникова
Технический редактор О.Г. Завьялова

Лиц. №020304 от 28.11.96. Подписано в печать
Формат 60×84/16. Бумага офсетная №2. Ризография. Усл.-печ. л. 5,2.
Уч.-изд.л. . Изд. № . Тираж 75 экз. Заказ .
Московский государственный институт электроники и математики.

109028, Москва, Б.Трехсвятительский пер., 3.

Отдел оперативной полиграфии Московского государственного института электроники и математики. 113054, ул. М. Пионерская, 12.