МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

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

В. И. Сумин

НАЧАЛА ВЫПУКЛОГО АНАЛИЗА
Часть 2
Выпуклые функции

Учебно-методическое пособие

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

Нижний Новгород
2015

УДК 519.6
ББК 22.193
С-89

С-89 Сумин В.И. НАЧАЛА ВЫПУКЛОГО АНАЛИЗА. Часть 2. ВЫПУКЛЫЕ ФУНКЦИИ: Учебно-методическое пособие. – Нижний Новгород: Нижегородский государственный университет, 2015. – 28 с.

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

Учебно-методическое пособие по общему курсу «Методы оптимизации» для студентов бакалавриата ННГУ, обучающихся по направлениям подготовки «Математика», «Математика и компьютерные науки», «Прикладная математика и информатика», «Механика и математическое моделирование». Пособие включает материал аудиторных занятий по теме «Элементарный выпуклый анализ», посвященных выпуклым функциям, а также содержит материал для самостоятельного изучения.

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

УДК 519.6
ББК 22.193

С О Д Е Р Ж А Н И Е

Введение

Это вторая часть пособия "Начала выпуклого анализа", которое является заново отредактированным вариантом учебно-методического пособия [1], успешно использующегося на мех-мате ННГУ в учебном процессе. Первая часть пособия посвящена выпуклым множествам, а данная, вторая, часть – выпуклым функциям.

Напомним основные обозначения, принятые в первой части. Используются стандартные обозначения: $\mathbf{R}^n$ – n-мерное действительное пространство векторов-столбцов[1]

$$x = \mathrm{col}\{x^1, \ldots, x^n\} \equiv \{x^1, \ldots, x^n\}^T \equiv \begin{Bmatrix} x^1 \\ \vdots \\ x^n \end{Bmatrix}$$

со скалярным произведением $(x, y) \equiv \sum\limits_{i=1}^{n} x^i y^i$, нормой $\|x\| \equiv \sqrt{(x, x)}$ и метрикой $\rho(x, y) \equiv \|x - y\|$; $0_n \equiv \mathrm{col}\{0, \ldots, 0\}$ – нуль в $\mathbf{R}^n$; если $X \subset \mathbf{R}^n$, то $\overset{\circ}{X} \equiv intX$ – внутренность множества X, $\overline{X}$ – замыкание X, ∂X – граница X, $\lambda X \equiv \{y : y = \lambda x, x \in X\}$ $\forall \lambda \in \mathbf{R}$; $X \pm Y \equiv \{z : z = x \pm y,\ x \in X,\ y \in Y\}$ – алгебраические сумма и разность множеств $X \subset \mathbf{R}^n$, $Y \subset \mathbf{R}^n$; $U_\varepsilon(x_0) \equiv \{x : \|x - x_0\| < \varepsilon\}$ – ε-окрестность точки x_0 в $\mathbf{R}^n$, $\varepsilon > 0$; $[x_0, x_1]$ – отрезок прямой, соединяющий точки x_0 и x_1 в $\mathbf{R}^n$,

$$[x_0, x_1] \equiv \{x : \quad x = \lambda x_1 + (1 - \lambda)x_0, \quad 0 \le \lambda \le 1\}.$$

Приняты также *специальные обозначения*: $x_\lambda \equiv \lambda x_1 + (1 - \lambda)x_0$ $(x_0 \in \mathbf{R}^n,\ \ x_1 \in \mathbf{R}^n,\ \ 0 \le \lambda \le 1)$; $\Gamma_{c,\alpha} \equiv \{x : (c, x) = \alpha\}$ – гиперплоскость в $\mathbf{R}^n$ $(\alpha \in \mathbf{R},\ c \in \mathbf{R}^n,\ c \ne 0_n)$; $\Gamma^-_{c,\alpha} \equiv \{x : (c, x) < \alpha\}$, $\Gamma^+_{c,\alpha} \equiv \{x : (c, x) > \alpha\}$ – открытые полупространства, определяемые в $\mathbf{R}^n$ гиперплоскостью $\Gamma_{c,\alpha}$.

Векторное неравенство $x \le y$ для $x, y \in \mathbf{R}^n$ означает: $x^i \le y^i$, $i = \overline{1, n}$; аналогично понимается неравенство $x \ge y$. $\mathbf{R}^n_+ \equiv \{x \in \mathbf{R}^n : x \ge 0_n\}$ – неотрицательный ортант в $\mathbf{R}^n$.

Нумерация теорем, лемм, рисунков, примеров и упражнений в пособии одинарная и сквозная, а нумерация разделов (пунктов) в каждой части своя. Нумерация формул двойная; например, третья по порядку формула второй части имеет номер (2.3).

[1]Здесь и далее используются следующие значки сокращенной записи: $\equiv$ – "тождественно равно"или "равно по определению"; $\forall$ – "для любого", "для каждого" или "для всех"; $\exists$ – "существует"; $\in$ – "принадлежит", "принадлежащий", "принадлежащие"; $\notin$ – "не принадлежит"; $\subset$ – "вложено в", "содержится в".

1. Определение выпуклой функции и его геометрический смысл. Пусть X – непустое выпуклое множество в $\mathbf{R}^n$. Функция $f(.) : X \to \mathbf{R}$ называется *выпуклой на множестве* X, если для любых точек $x_0 \in X$, $x_1 \in X$ имеем

$$f(\lambda x_1 + (1-\lambda)x_0) \quad \leq \lambda f(x_1) + (1-\lambda)f(x_0)\ \forall \lambda \in (0,1). \qquad (2.1)$$

Неравенство (2.1) называется *неравенством выпуклости.* Если выпуклая функция $f(.) : X \to \mathbf{R}$ такова, что для любых точек $x_0 \in X$, $x_1 \in X$, $x_0 \neq x_1$ неравенство выпуклости (2.1) выполняется как строгое неравенство для всех $\lambda \in (0,1)$, то функция $f(.)$ называется *строго выпуклой на* X. Заметим, что в соответствии с приведенным определением любая функция, заданная на одноточечном множестве, выпукла на нем. Для единообразия формулировок нам удобно будет считать такую функцию и строго выпуклой. Функция $f(.) : X \to \mathbf{R}$ называется *вогнутой* (*строго вогнутой*) *функцией на множестве* X, если функция $(-f(.)) : X \to \mathbf{R}$ выпукла (соотв. строго выпукла) на X.

Упражнение 22. Докажите, что: 1) функция $f(x) = \sum\limits_{i=1}^{n} |x^i|$ выпукла, но не строго выпукла и не вогнута на $\mathbf{R}^n$; 2) линейная функция $f(x) = \sum\limits_{i=1}^{n} c^i x^i$ выпукла и вогнута на $\mathbf{R}^n$; 3) если $X \subset \mathbf{R}^n$– выпуклое множество, а $f(.) : X \to \mathbf{R}$ – непрерывная функция, то $f(.)$ выпукла на X тогда и только тогда, когда $f\left(\left(x_0 + x_1\right)/2\right) \leq \left(f\left(x_0\right) + f\left(x_1\right)\right)/2 \ \ \forall x_0, x_1 \in X$; 4) если $X \subset \mathbf{R}^n$– выпуклое множество с непустой внутренностью, а $f(.) : X \to \mathbf{R}$ – непрерывная функция, выпуклая на $\overset{\circ}{X}$, то $f(.)$ выпукла и на X.

Пусть $X \subset \mathbf{R}^n$. *Графиком функции* $f(.) : X \to \mathbf{R}$ называется множество

$$\mathrm{Gr} f \equiv \left\{ \left\{ \begin{array}{c} x \\ y \end{array} \right\} \in \mathbf{R}^{n+1} : x \in X, \quad y = f(x) \right\}.$$

Надграфиком, или *эпиграфом, функции* $f(.) : X \to \mathbf{R}$ называется множество

$$\mathrm{epi} f \equiv \left\{ \left\{ \begin{array}{c} x \\ y \end{array} \right\} \in \mathbf{R}^{n+1} : x \in X, \quad y \geq f(x) \right\}.$$

Это определение проиллюстрировано на рис. 18a для $n = 1$.

Лемма 17. *Функция* $f(.)$, *определенная на выпуклом множестве* X, *выпукла на* X *тогда и только тогда, когда для любых точек* A, $B \in \mathrm{Gr} f$ *соединяющий их отрезок* $[A, B]$ *принадлежит* $\mathrm{epi} f$.

□ [2] Пусть $A = \left\{ \begin{array}{c} x_0 \\ f(x_0) \end{array} \right\}$, $B = \left\{ \begin{array}{c} x_1 \\ f(x_1) \end{array} \right\}$. Отрезок $[A, B]$ имеет следу-

[2]Значок □ открывает доказательство, значок ⊠ означает, что оно закончено.

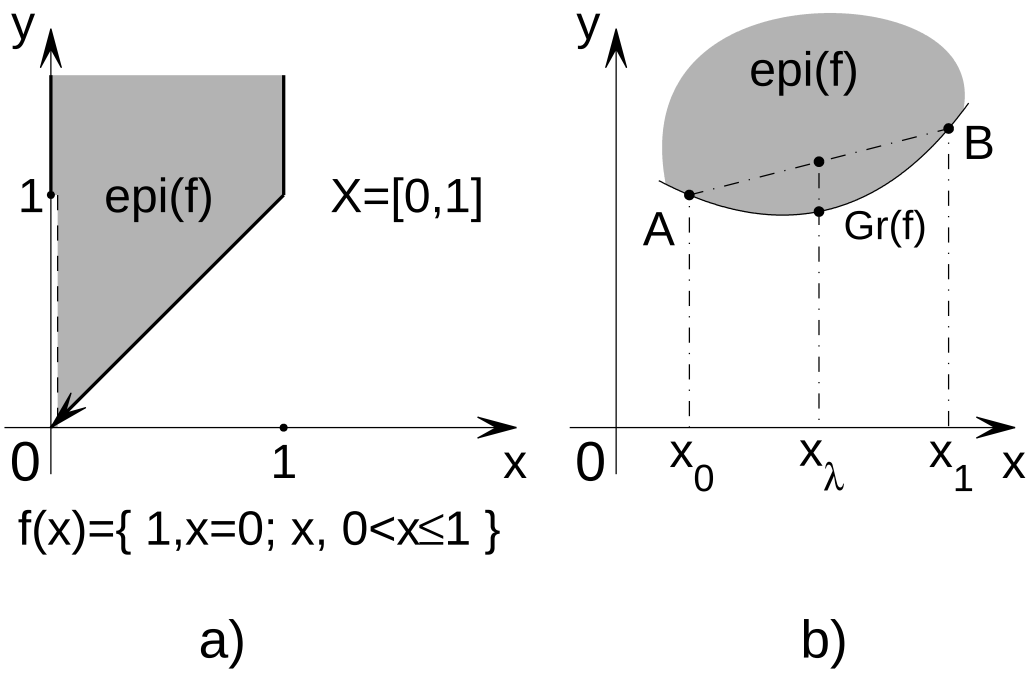

Рис. 1

ющую параметризацию

$$[A, B] = \left\{ \left\{ \begin{array}{c} x \\ y \end{array} \right\} \in \mathbf{R}^{n+1} : \quad \begin{array}{l} x = x_\lambda \equiv \lambda x_1 + (1-\lambda)x_0 \\ y = \lambda f(x_1) + (1-\lambda) f(x_0) \end{array} ; \quad 0 \le \lambda \le 1 \right\}.$$

Очевидно, что $[A, B] \subset \mathrm{epi} f$ тогда и только тогда, когда $f(x_\lambda) \le \lambda f(x_1) + (1-\lambda) f(x_0)$ $\forall \lambda \in [0, 1]$. Следовательно, $f(.)$ выпукла на X тогда и только тогда, когда для любых A, $B \in \mathrm{Gr} f$ имеем $[A, B] \subset \mathrm{epi} f$. ⊠

Лемма 17 проиллюстрирована на рис. 18b для $n = 1$.

Лемма 18. *Функция* $f(.)$, *определенная на выпуклом множестве* X, *выпукла на* X *тогда и только тогда, когда выпукло множество* $\mathrm{epi} f$.

□ *Необходимость*. Пусть $f(.)$ выпукла на X. Фиксируем произвольно $A = \left\{ \begin{array}{c} x_0 \\ y_0 \end{array} \right\} \in \mathrm{epi} f$ и $B = \left\{ \begin{array}{c} x_1 \\ y_1 \end{array} \right\} \in \mathrm{epi} f$. Покажем, что отрезок $[A, B]$ принадлежит $\mathrm{epi} f$. Отрезок $[A, B]$ имеет параметризацию:

$$[A, B] = \left\{ \left\{ \begin{array}{c} x \\ y \end{array} \right\} \in \mathbf{R}^{n+1} : \quad \begin{array}{l} x = x_\lambda \equiv \lambda x_1 + (1-\lambda)x_0 \\ y = y_\lambda \equiv \lambda y_1 + (1-\lambda) y_0 \end{array} ; \quad 0 \le \lambda \le 1 \right\}.$$

Возьмем в множестве $\mathrm{Gr} f$ точки $A' = \left\{ \begin{array}{c} x_0 \\ f(x_0) \end{array} \right\}$, $B' = \left\{ \begin{array}{c} x_1 \\ f(x_1) \end{array} \right\}$. В силу леммы 17 отрезок $[A', B'] \subset \mathrm{epi} f$ (см. иллюстрацию на рис. 19a). Но отрезок

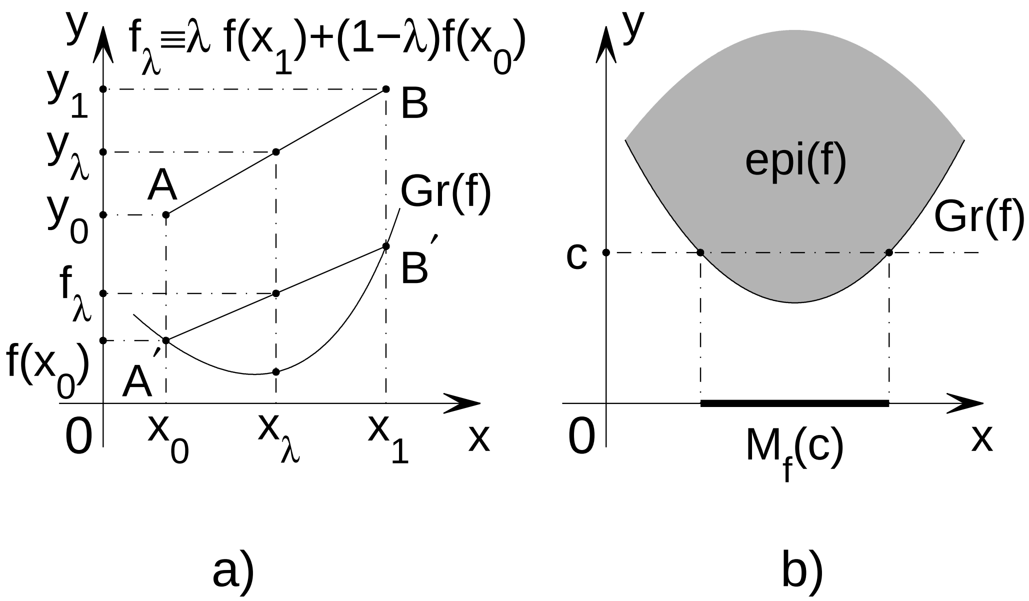

Рис. 2

$[A', B']$ имеет параметризацию

$$[A', B'] = \left\{ \left\{ \begin{array}{c} x \\ y \end{array} \right\} \in \mathbf{R}^{n+1} : \quad \begin{array}{l} x = x_\lambda \equiv \lambda x_1 + (1-\lambda)x_0 \\ y = \lambda f(x_1) + (1-\lambda) f(x_0) \end{array} ; \quad 0 \leq \lambda \leq 1 \right\}.$$

Так как $y_i \geq f(x_i)$, $i = 0, 1$, то $y_\lambda \geq \lambda f(x_1) + (1-\lambda) f(x_0)$ и, следовательно, точка $\left\{ \begin{array}{c} x_\lambda \\ y_\lambda \end{array} \right\}$ отрезка $[A, B]$ принадлежит epi$f$ вместе с точкой

$$\left\{ \begin{array}{c} x_\lambda \\ \lambda f(x_1) + (1-\lambda) f(x_0) \end{array} \right\}$$

отрезка $[A', B']$. Значит, $[A, B] \subset \mathrm{epi} f$.

Достаточность. Пусть epi$f$ – выпуклое множество. Это означает, в частности, что для любых двух точек A, $B \in \mathrm{Gr} f$ отрезок $[A, B] \subset \mathrm{epi} f$. В силу леммы 17 функция $f(.)$ выпукла на X. ☒

2. Простейшие свойства выпуклых функций. В упражнениях 23, 24, 26 и лемме 19 описаны некоторые *внутренние операции в классе выпуклых функций*, то есть операции над функциями, не выводящие из этого класса.

Упражнение 23. Докажите, что если $f_i(.) : X \to \mathbf{R}$, $i = \overline{1, m}$, – выпуклые функции, а λ_i, $i = \overline{1, m}$, – неотрицательные числа, то комбинация $\sum\limits_{i=1}^{m} \lambda_i f_i(.)$ есть выпуклая функция на X.

Упражнение 24. Пусть $f(.) : X \to \mathbf{R}$ – выпуклая функция на X. Докажите, что: 1) для любого фиксированного $y \in \mathbf{R}^n$ функция $F(x) \equiv f(x + y)$

выпукла на множестве $X - \{y\}$; 2) для любого фиксированного $\mu \in \mathbf{R}$, $\mu \neq 0$ функция $\varphi(x) = f(\mu x)$ выпукла на множестве $\mu^{-1}X$; 3) если $f(x) \geq 0$ $\forall x \in X$, то функция $f^2(x)$ выпукла на X.

Лемма 19. *Пусть X – выпуклое множество в $\mathbf{R}^n$, Y – произвольное множество, $f(x,y) : X \times Y \to \mathbf{R}$ – функция, удовлетворяющая условиям*: а) $\forall y \in Y$ *функция $f(.,y)$ выпукла на X*; б) $\forall x \in X$ *величина $F(x) \equiv \sup\limits_{y \in Y} f(x,y)$ конечна. Тогда функция $F(x)$ выпукла на X.*

$\square$ Для любых $x_0 \in X$, $x_1 \in X$ имеем при $\lambda \in [0,1]$:

$$F(x_\lambda) \equiv \sup_{y \in Y} f(x_\lambda, y) \leq \sup_{y \in Y} \{\lambda f(x_1,y) + (1-\lambda) f(x_0,y)\} \leq$$

$$\leq \sup_{y \in Y} \{\lambda f(x_1,y)\} + \sup_{y \in Y} \{(1-\lambda) f(x_0,y)\} = \lambda \sup_{y \in Y} \{f(x_1,y)\} +$$

$$+(1-\lambda) \sup_{y \in Y} \{f(x_0,y)\} = \lambda F(x_1) + (1-\lambda) F(x_0). \quad \boxtimes$$

Упражнение 25. Докажите, что в условиях леммы 19 имеет место равенство:

$$\mathrm{epi} F(.) = \bigcap_{y \in Y} \mathrm{epi} f(.,y).$$

Из этого равенства в силу лемм 1 и 18 вытекает утверждение леммы 19.

Упражнение 26. Докажите, что если $f(.) : X \to \mathbf{R}$ – выпуклая на X функция, то и функция $f^+(x) \equiv \max\{f(x), 0\}$, $x \in X$, выпукла на X.

Лемма 20 (*неравенство Йенсена*). *Пусть $f(.) : X \to \mathbf{R}$ – выпуклая функция, а $\{x_1, \ldots, x_m\}$ – некоторый конечный набор точек множества X. Для любой выпуклой комбинации $\sum\limits_{i=1}^{m} \lambda_i x_i$:*

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i). \tag{2.2}$$

$\square$ Доказательство проведем индукцией по m. При $m = 1$ неравенство (2.2) превращается в равенство, а при $m = 2$ оно есть очевидное следствие выпуклости $f(.)$. Предположив, что доказываемое утверждение верно при $m = k$, докажем его при $m = k+1$. Для выпуклой комбинации $\sum\limits_{i=1}^{k+1} \lambda_i x_i$ может быть: либо $\lambda_{k+1} = 1$, тогда $\lambda_1 = \ldots = \lambda_k = 0$ и неравенство (2.2) тривиально выполняется; либо $\lambda_{k+1} < 1$. В последнем случае имеем в силу справедливости (2.2) при $m = 2$ и $m = k$:

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \equiv f\left(\lambda_{k+1} x_{k+1} + (1-\lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{(1-\lambda_{k+1})} x_i\right) \leq$$

$$\le \lambda_{k+1} f(x_{k+1}) + (1-\lambda_{k+1}) f\left(\sum_{i=1}^{k} \frac{\lambda_i}{(1-\lambda_{k+1})} x_i\right) \le$$

$$\le \lambda_{k+1} f(x_{k+1}) + (1-\lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{(1-\lambda_{k+1})} f(x_i) = \sum_{i=1}^{k+1} \lambda_i f(x_i). \quad \boxtimes$$

Для функции $f(.) : X \to \mathbf{R}$ любое множество вида

$$M_f(c) \equiv \{x \in X : f(x) \le c\},$$

где $c \in \mathbf{R}$, называется *множеством Лебега* (см. рис. 19b).

Лемма 21. *Все множества Лебега выпуклой функции выпуклы.*

Упражнение 27. Докажите лемму 21.

Упражнение 28. Пусть $\mathbf{P}$ – выпуклое множество в $\mathbf{R}^n$, $g_i(.) : \mathbf{P} \to \mathbf{R}$ – выпуклые на $\mathbf{P}$ функции, $i \in I$, I – произвольное множество. Докажите, что $X \equiv \{x \in \mathbf{R}^n : x \in \mathbf{P}, \quad g_i(x) \le 0, \quad i \in I\}$ выпукло.

3. Лемма об одномерных сечениях. Пусть $f(.) : X \to \mathbf{R}$ – некоторая функция, определенная на выпуклом множестве $X \subset \mathbf{R}^n$, а $X_{x_0,\ell}$ – некоторое непустое одномерное сечение X. Обозначим через $[f]_{x_0,\ell}(.)$ функцию одного переменного, определенную на соответствующем множестве $A_{x_0,\ell}$ (см. п.12 в §1) формулой

$$[f]_{x_0,\ell}(t) \equiv f(x_0 + t\ell), \quad t \in A_{x_0,\ell}.$$

Непосредственно из определений п.1 вытекает

Лемма 22. *Следующие утверждения эквивалентны*:

1) $f(.)$ *выпукла (строго выпукла) на* X;

2) $f(.)$ *выпукла (строго выпукла) на любом непустом одномерном сечении* $X_{x_0,\ell}$;

3) $[f]_{x_0,\ell}(.)$ *выпукла (строго выпукла) на множестве* $A_{x_0,\ell}$ *для любых* $x_0 \in X$, $\ell \in \mathbf{R}^n$, $\ell \ne 0_n$.

Упражнение 29. Докажите лемму 22.

4. Дифференцируемость выпуклой функции по возможным направлениям. Пусть $X \subset \mathbf{R}^n$, $f(.) : X \to \mathbf{R}$. Если $\ell \in \mathbf{R}^n$ – вектор возможного для X в некоторой точке $x_0 \in X$ направления ($\ell \in K(x_0, X)$, см. п.13 в §1), то имеет смысл говорить о *производной функции* $f(.)$ *по вектору* ℓ. Эта производная, напомним, определяется соотношением

$$\frac{\partial f(x_0)}{\partial \ell} \equiv \lim_{t \to +0} \frac{f(x_0 + t\ell) - f(x_0)}{t}, \tag{2.3}$$

если предел справа существует (конечный или нет)[3]. При $\ell = 0_n$ предел справа в формуле (2.3) существует и равен нулю. Поэтому считаем, что при $\ell = 0_n$ производная $\partial f(x_0) / \partial \ell$ равна нулю.

[3]Напомним также, что производная по единичному вектору ℓ называется *производной по направлению.*

Теорема 7. *Если функция $f(.)$ выпукла на X, то для любого $x_0 \in X$ и любого вектора $\ell \in K(x_0, X)$ существует конечная или равная $-\infty$ производная $\partial f(x_0)/\partial \ell$.*

$\square$ Достаточно рассмотреть случай $\ell \in K(x_0, X)$, $\ell \neq 0_n$. Заметим, что множество $A_{x_0,\ell}$ (см. п.12 в §1) содержит некоторый отрезок $[0, \varepsilon]$, $\varepsilon > 0$ (см. п.13 в §1, упражнение 20). Положим для сокращения записи $\Phi(t) \equiv [f]_{x_0,\ell}(t)$, $t \in A_{x_0,\ell}$. В силу леммы 22 функция $\Phi(t)$ выпукла на $A_{x_0,\ell}$. Теорема будет доказана, если мы докажем, что функция

$$\psi(t) \equiv \frac{\Phi(t) - \Phi(0)}{t} \equiv \frac{f(x_0 + t\ell) - f(x_0)}{t}, \quad t \in (0, \varepsilon],$$

имеет при $t \to +0$ конечный или равный $-\infty$ предел. Для этого достаточно доказать, что $\psi(t)$ монотонно убывает при $t \to +0$.

Фиксируем произвольно t_1, $t_2 \in (0, \varepsilon]$, $t_1 < t_2$. Можно записать $t_1 = \lambda t_2 + (1 - \lambda) \cdot 0$, взяв $\lambda = t_1/t_2$. Из выпуклости $\Phi(t)$ следует, что

$$\Phi(t_1) \leq \lambda \Phi(t_2) + (1 - \lambda)\Phi(0)$$

или

$$\Phi(t_1) - \Phi(0) \leq \lambda \left[\Phi(t_2) - \Phi(0)\right],$$

откуда получаем

$$\psi(t_1) \leq \psi(t_2).$$

Монотонное убывание $\psi(t)$ при $t \to +0$ доказано. $\boxtimes$

Геометрический смысл проведенного доказательства иллюстрируется на рис. 20*a*. Следующая теорема уточняет теорему 7.

Теорема 8. *Если $f(.)$ выпукла на X, то для любой точки $x_0 \in \overset{\circ}{X}$ и любого вектора $\ell \in \mathbf{R}^n$, $\ell \neq 0_n$ существует конечная производная $\partial f(x_0)/\partial \ell$.*

$\square$ Воспользуемся обозначениями предыдущего доказательства. Докажем, что теперь $\psi(t)$ не только монотонно убывает при $t \to +0$, но и ограничена снизу. Отсюда будет следовать, что $\lim\limits_{t \to +0} \psi(t)$ конечен и, следовательно, существует конечная производная $\partial f(x_0)/\partial \ell$. В данном случае $A_{x_0,\ell}$ содержит окрестность $(-\varepsilon, \varepsilon)$ точки $t = 0$ (см. п.13 в §1, упражнение 20). Фиксируем произвольно $t_1 \in (-\varepsilon, 0)$ и $t_2 \in (0, \varepsilon)$. Рассматривая $t = 0$ как выпуклую комбинацию t_1 и t_2, запишем

$$0 = \lambda t_1 + (1 - \lambda) t_2, \text{ где } \quad \lambda = \frac{t_2}{t_2 - t_1}, \quad (1 - \lambda) = -\frac{t_1}{t_2 - t_1}.$$

Из выпуклости $\Phi(\cdot)$ получаем

$$\Phi(0) \leq \lambda \Phi(t_1) + (1 - \lambda)\Phi(t_2),$$

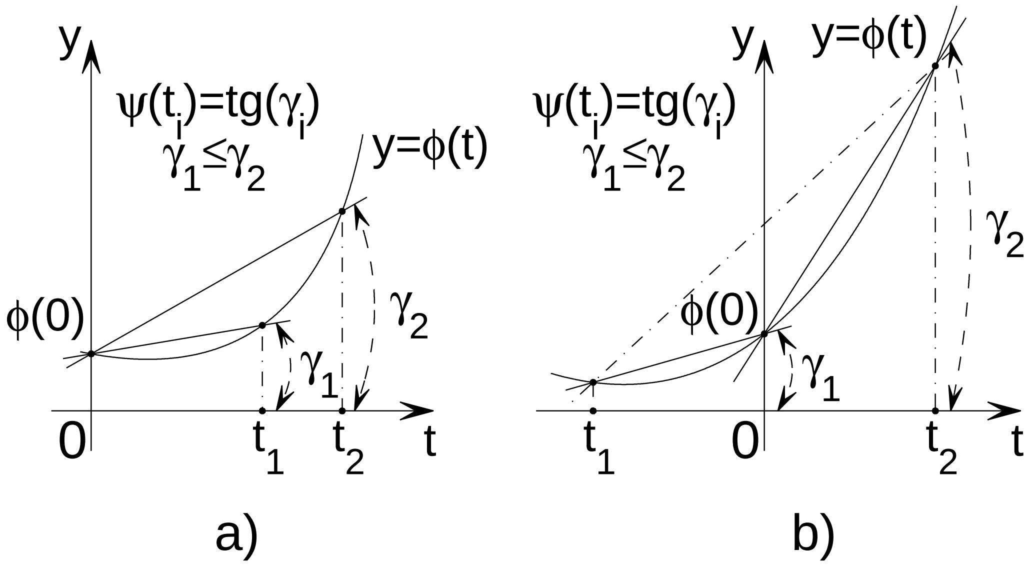

Рис. 3

что дает

$$(1-\lambda)\left[\Phi(t_2)-\Phi(0)\right] \geq \lambda\left[\Phi(0)-\Phi(t_1)\right]$$

и, следовательно,

$$\psi(t_2) \equiv \frac{\Phi(t_2)-\Phi(0)}{t_2} \geq \frac{\Phi(0)-\Phi(t_1)}{-t_1} = \frac{\Phi(t_1)-\Phi(0)}{t_1}.$$

Итак, при фиксированном $t_1 < 0$

$$\forall t_2 \in (0, \varepsilon): \quad \psi(t_2) \geq const \equiv \frac{\Phi(t_1)-\Phi(0)}{t_1}. \quad \boxtimes$$

Геометрический смысл доказательства теоремы 8 показан на рис. $20b$. Простые примеры следующего ниже упражнения 30 показывают, что для выпуклой функции $f(\cdot) : X \to \mathbf{R}$ производные по возможным направлениям в граничных точках множества X действительно могут обращаться в $-\infty$.

Упражнение 30. Пусть $n = 1$, $X = [0,1]$. 1) Докажите, что $\ell = -1$ есть возможное для X направление в точке $x = 1$. 2) Докажите, что разрывная функция

$$f(x) \equiv \begin{cases} 0, & 0 \leq x < 1; \\ 2, & x = 1 \end{cases}$$

выпукла на X. Проверьте, что $\partial f(1)/\partial \ell = -\infty$. 3) То же самое для непрерывной на X функции $f(x) \equiv 1 - \sqrt{1-x^2}, 0 \leq x \leq 1$.

5. Свойство непрерывности выпуклой функции. Как показывает пример из упражнения 30, функция может терпеть разрыв на границе того

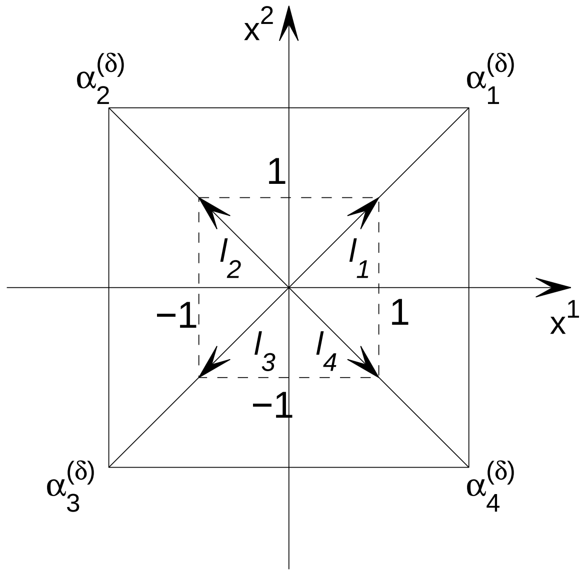

Рис. 4

множества, на котором она выпукла. Оказывается, внутри этого множества функция обязательно непрерывна.

Теорема 9. *Пусть $X \subset \mathbf{R}^n$ – непустое выпуклое множество. Выпуклая на X функция $f(.)$ непрерывна в любой точке x_0, внутренней для множества X.*

☐ Так как сдвиг не нарушает выпуклости множества, то без ограничения общности считаем $x_0 = 0_n$. Пусть $H_\delta \equiv \{x \in \mathbf{R}^n : |x^i| \leq \delta, \quad i = \overline{1,n}\}$ – замкнутая кубическая окрестность 0_n. Так как $x_0 = 0_n$ – внутренняя точка X, то существует $\overline{\delta} > 0$ такое, что $H_\delta \subset X$ $\forall \delta \in (0, \overline{\delta})$. Достаточно показать, что

$$\forall \varepsilon > 0 \quad \exists \delta(\varepsilon) > 0 : \quad |f(x) - f(0_n)| < \varepsilon, \quad \text{если} \quad x \in H_{\delta(\varepsilon)}. \tag{2.4}$$

Воспользуемся теоремой о представлении выпуклого компакта (п.8 в §1, теорема 5), применив ее к компакту H_δ. Угловыми для H_δ являются (см. п.13 в §1, упражнение 21) те и только те точки, для которых $|x^i| = \delta$, $i = \overline{1,n}$. Иначе говоря, это все те точки, каждая из которых лежит на пересечении границы ∂H_δ с одним из лучей, выходящих из точки 0_n в направлении вектора ℓ, удовлетворяющего условию $|\ell^i| = 1$, $i = \overline{1,n}$. Множество таких "угловых"направлений не зависит от величины δ, а число их равно $m \equiv 2^n$, числу угловых точек n-мерного куба. Каким-либо образом пронумеруем указанные векторы: $\ell_1, \ldots, \ell_m$. Пусть $L_i \equiv \{x \in \mathbf{R}^n : x = t\ell_i, t \geq 0\}$ – луч, выходящий из точки 0_n в направлении ℓ_i, $i = \overline{1,m}$. Угловую точку куба H_δ, лежащую на луче L_i, обозначим $a_i^{(\delta)}$, $i = \overline{1,m}$ (см. рис. 21 для $n = 2$).

Для доказательства (2.4) фиксируем $\varepsilon > 0$. В силу теоремы 8 для любого $i = \overline{1,m}$ существует конечная производная $\partial f(0_n)/\partial \ell_i$. Поэтому сужение $f(.)$

на любой луч L_i, $i = \overline{1,m}$, непрерывно в точке 0_n. Это означает, что существует $\delta(\varepsilon) > 0$ такое, что

$$|f(a_i^{(\delta)}) - f(0_n)| < \varepsilon, \quad 0 < \delta \le \delta(\varepsilon), \quad i = \overline{1,m}. \tag{2.5}$$

Покажем, что число $\delta(\varepsilon)$ – искомое, то есть $|f(x) - f(0_n)| < \varepsilon$ при $x \in H_{\delta(\varepsilon)}$. Произвольно фиксируем $x \in H_{\delta(\varepsilon)}$ и любое $\delta \in (0, \delta(\varepsilon)]$ такое, что $x \in H_\delta$. По теореме 5 найдутся числа $\lambda_i \ge 0$, $i = \overline{1,m}$, такие, что

$$x = \sum_{i=1}^{m} \lambda_i a_i^{(\delta)}, \quad \sum_{i=1}^{m} \lambda_i = 1.$$

По неравенству Йенсена

$$f(x) \le \sum_{i=1}^{m} \lambda_i f\left(a_i^{(\delta)}\right)$$

или

$$f(x) - f(0_n) \le \sum_{i=1}^{m} \lambda_i \left\{ f\left(a_i^{(\delta)}\right) - f(0_n) \right\}. \tag{2.6}$$

В силу (2.6) и (2.5)

$$f(x) - f(0_n) < \sum_{i=1}^{m} \lambda_i \varepsilon = \varepsilon. \tag{2.7}$$

Но куб $H_{\delta(\varepsilon)}$ симметричен относительно точки 0_n и, следовательно,

$$f(-x) - f(0_n) < \varepsilon. \tag{2.8}$$

Так как $0_n = \frac{1}{2}x + \frac{1}{2}(-x)$, то из неравенства выпуклости получаем

$$f(0_n) \le \frac{1}{2} f(x) + \frac{1}{2} f(-x)$$

и

$$f(0_n) - f(x) \le f(-x) - f(0_n). \tag{2.9}$$

Неравенства (2.7), (2.8), (2.9) дают:

$$|f(x) - f(0_n)| = \max\{f(x) - f(0_n), \quad f(0_n) - f(x)\} < \varepsilon. \quad \boxtimes$$

6. Критерий выпуклости в классе дифференцируемых функций одной переменной. (См. [4], гл.V, §4, п.3) Пусть функция $f(.)$ определена на промежутке[4] $X \subset \mathbf{R}$ ненулевой длины.

[4]Напомним, что промежуток – это отрезок, интервал или полуинтервал (интервал и полуинтервал не обязательно конечны).

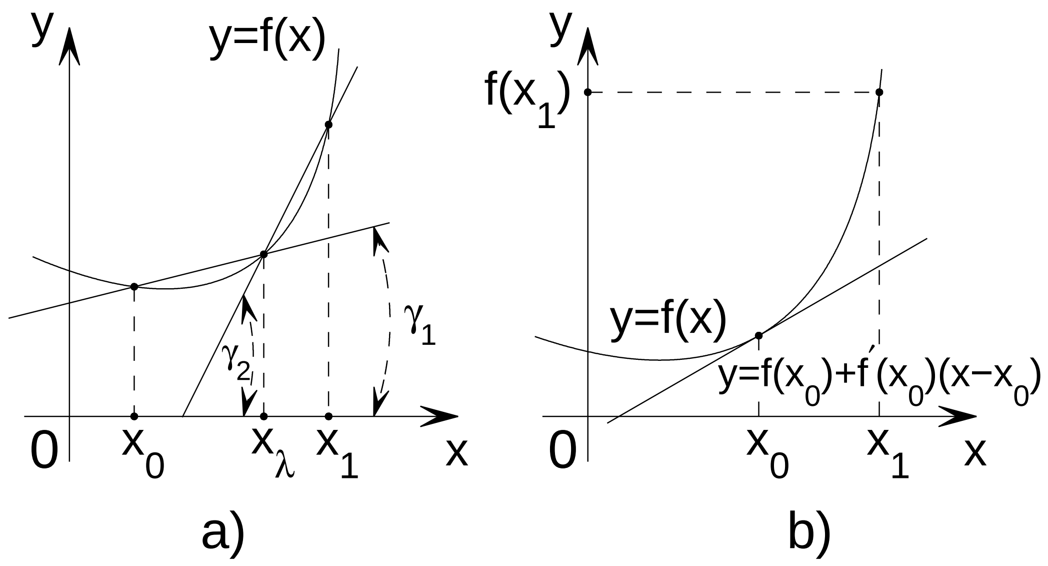

Рис. 5

Лемма 23. *Для того, чтобы $f(.)$ была выпукла (строго выпукла) на X, необходимо и достаточно, чтобы*

$$\frac{f(x_\lambda)-f(x_0)}{x_\lambda-x_0} \leq (\text{соотв.} <)\ \frac{f(x_1)-f(x_\lambda)}{x_1-x_\lambda}\ \forall\ x_0,x_1\in X, x_0<x_1, \lambda\in(0,1). \tag{2.10}$$

□ Определение выпуклости (соотв. строгой выпуклости) $f(.)$ можно записать в виде:

$$\begin{array}{l}\{f(x_\lambda)-f(x_0)\}\,(1-\lambda)\leq(\text{соотв.} <)\ \lambda\,\{f(x_1)-f(x_\lambda)\}\\ \forall\ x_0,x_1\in X, x_0\neq x_1, \lambda\in(0,1).\end{array} \tag{2.11}$$

Из формулы $x_\lambda=\lambda x_1+(1-\lambda)x_0$ имеем:

$$\lambda=\frac{x_\lambda-x_0}{x_1-x_0},\quad 1-\lambda=\frac{x_1-x_\lambda}{x_1-x_0}. \tag{2.12}$$

Из (2.11) и (2.12) получаем (2.10). ⊠

Геометрический смысл леммы 23 ясен из рис. $22a$: $\mathrm{tg}\gamma_1\leq$ (соотв. $<$)$\mathrm{tg}\gamma_2$.

Пусть теперь $f\,(.)$ дифференцируема на промежутке X [5].

Лемма 24. *Для того, чтобы $f(.)$ была выпуклой (строго выпуклой) на X, необходимо и достаточно, чтобы производная $f'(.)$ не убывала (*соотв. *строго возрастала) на X.*

[5]Дифференцируемость в крайней точке означает существование соответствующей односторонней производной.

$\square$ *Необходимость*. Устремляя x_λ сначала к x_0, а затем к x_1, получаем из (2.10) двойное неравенство

$$f'(x_0) \le \frac{f(x_1) - f(x_0)}{x_1 - x_0} \le f'(x_1)\ \forall x_0, x_1 \in X, x_0 < x_1, \tag{2.13}$$

означающее монотонность производной на X. В случае строго выпуклой $f(.)$ зафиксируем некоторые $x_0, x_1 \in X$, $x_0 < x_1$, $\lambda \in (0,1)$. По теореме Лагранжа о конечных приращениях существуют такие $\xi_1 \in (x_0, x_\lambda)$ и $\xi_2 \in (x_\lambda, x_1)$, что

$$\frac{f(x_\lambda) - f(x_0)}{x_\lambda - x_0} = f'(\xi_1), \quad \frac{f(x_1) - f(x_\lambda)}{x_1 - x_\lambda} = f'(\xi_2). \tag{2.14}$$

В силу (2.10): $f'(\xi_1) < f'(\xi_2)$. Доказанная выше монотонность $f'(.)$ дает возможность написать

$$f'(x_0) \le f'(\xi_1) < f'(\xi_2) \le f'(x_1). \tag{2.15}$$

Ввиду произвольности выбора x_0, x_1 неравенство (2.15) означает строгую монотонность $f'(.)$ на X.

Достаточность. Фиксируем произвольно x_0, $x_1 \in X$, $x_0 < x_1$, $x_\lambda \in (x_0, x_1)$. Существуют $\xi_1 \in (x_0, x_\lambda)$ и $\xi_2 \in (x_\lambda, x_1)$, для которых выполняются равенства (2.14). Если $f'(.)$ не убывает (строго возрастает) на X, то

$$f'(\xi_1) \quad \le \quad (\text{ соотв. } <) \quad f'(\xi_2),$$

что вместе с (2.15) дает (2.10). $\boxtimes$

Лемма 25. *Для того, чтобы $f(.)$ была выпуклой (строго выпуклой) на X, необходимо и достаточно, чтобы для всех x_0, $x_1 \in X$, $x_0 \ne x_1$ выполнялось неравенство*

$$f(x_1) \quad \ge \quad (\text{ соотв. } >) \quad f(x_0) + f'(x_0)(x_1 - x_0). \tag{2.16}$$

$\square$ *Необходимость*. Фиксируем произвольно x_0, $x_1 \in X$, $x_0 < x_1$. Рассмотрим (2.10). Перейдем к пределу в (2.10) при $x_\lambda \to x_0 + 0$. Левая часть (2.10) стремится к $f'(x_0)$, а правая – к отношению $\{f(x_1) - f(x_0)\}/(x_1 - x_0)$, которое по теореме Лагранжа равно некоторому значению $f'(\xi)$, где $\xi \in (x_0, x_1)$. Для выпуклой (строго выпуклой) $f(.)$ с учетом леммы 24 имеем

$$f'(x_0) \quad \le \quad (\text{ соотв. } <) \quad f'(\xi) = \{f(x_1) - f(x_0)\}/(x_1 - x_0),$$

что дает (2.16) при $x_1 > x_0$. В случае $x_1 < x_0$ тем же рассуждением придем к (2.16), поменяв в (2.10) x_1 и x_0 местами.

Достаточность. Фиксируем произвольно x_0, $x_1 \in X$, $x_0 < x_1$, $x_\lambda \in (x_0, x_1)$. Заменив в (2.16) x_0 на x_λ, получим

$$f(x_1) \quad \ge \quad (\text{ соотв. } >) \quad f(x_\lambda) + f'(x_\lambda)(x_1 - x_\lambda),$$

что можно переписать как

$$f'(x_\lambda) \quad \le \quad (\text{ соотв.} <) \quad \frac{f(x_1)-f(x_\lambda)}{x_1-x_\lambda}. \tag{2.17}$$

Заменив в (2.16) x_1 на x_0, а x_0 – на x_λ, получим

$$f(x_0) \quad \ge \quad (\text{ соотв.} >) \quad f(x_\lambda)+f'(x_\lambda)(x_0-x_\lambda),$$

что перепишем в виде

$$f'(x_\lambda) \quad \ge \quad (\text{ соотв.} >) \quad \frac{f(x_0)-f(x_\lambda)}{x_0-x_\lambda} \equiv \frac{f(x_\lambda)-f(x_0)}{x_\lambda-x_0}. \tag{2.18}$$

Из (2.17) и (2.18) следует (2.10). ⊠

Геометрический смысл леммы 25 состоит в том, что $f(.)$ выпукла (строго выпукла) на X тогда и только тогда, когда график функции $f(.) : X \to \mathbf{R}$ всеми своими точками, кроме точки касания, лежит не ниже (соотв. строго выше) любой проведенной к нему касательной (см. рис. 22b). Лемма об одномерных сечениях позволит нам обобщить результат леммы 25 на случай функций нескольких переменных (см. ниже п.8).

7. Критерий выпуклости в классе дважды дифференцируемых функций одной переменной. (См. [4], гл.V, §4, п.3) Пусть теперь функция $f(.)$ дважды дифференцируема на промежутке $X \subset \mathbf{R}$ ненулевой длины [6]. Непосредственно из леммы 24 получаем следующие результаты.

Лемма 26. *Для того, чтобы* $f(.)$ *была выпукла на* X, *необходимо и достаточно, чтобы для всех* $x \in X$ *было* $f''(x) \ge 0$.

Лемма 27. *Если* $f''(x) > 0$ для всех $x \in X$, *то* $f(.)$ *строго выпукла на* X.

Лемма об одномерных сечениях позволит нам обобщить результаты лемм 26, 27 на случай функций нескольких переменных (см. ниже п.9).

8. Критерий выпуклости в классе дифференцируемых функций нескольких переменных. Напомним, что градиент функции $f(x)$ векторного аргумента $x \in \mathbf{R}^n$ это состоящая из частных производных вектор-строка

$$\nabla f(x) \equiv \left(f'_{x^1}(x), f'_{x^2}(x), \ldots, f'_{x^n}(x)\right).$$

Пусть $X \subset \mathbf{R}^n$ – выпуклое множество, состоящее более чем из одной точки, и функция $f(.)$ дифференцируема на множестве X [7].

[6] То есть функция $f(.)$ дважды дифференцируема в каждой точке промежутка X. Дважды дифференцируемость в крайней точке означает существование первой и второй соответствующих односторонних производных.

[7] То есть функция $f(.)$ дифференцируема в каждой точке множества X. Отсюда следует, что функция $f(.)$ определена в некоторой окрестности каждой точки множества X.

Теорема 10. *Для того, чтобы* $f(.)$ *была выпукла (строго выпукла) на* X, *необходимо и достаточно, чтобы выполнялось неравенство*

$$\begin{array}{l} f(x_1) \quad \geq \quad (\text{ соотв. } >) \quad f(x_0) + \nabla f(x_0)(x_1 - x_0) \\ \forall\ x_0, x_1 \in X, \quad x_0 \neq x_1. \end{array} \tag{2.19}$$

$\square$ По лемме 22 функция $f(.)$ выпукла (строго выпукла) на X тогда и только тогда, когда для всех $\overline{x} \in X$, $\ell \in \mathbf{R}^n$, $\ell \neq 0_n$, выпукла (соотв. строго выпукла) на $A_{\overline{x},\ell}$ функция $[f]_{\overline{x},\ell}(.)$. В рассматриваемом случае $[f]_{\overline{x},\ell}(.)$ дифференцируема в точках множества $A_{\overline{x},\ell}$, причем

$$\{[f]_{\overline{x},\ell}(t)\}'_t \equiv \{f(\overline{x} + t\ell)\}'_t = \nabla f(\overline{x} + t\ell)\ell, \quad t \in A_{\overline{x},\ell}.$$

Так как на одноточечном множестве любая функция строго выпукла, то в силу леммы 25 функция $f(.)$ выпукла (соотв. строго выпукла) на X тогда и только тогда, когда выполняется условие: для любого множества $A_{\overline{x},\ell}$ длины, большей нуля, имеем:

$$\begin{array}{l} [f]_{\overline{x},\ell}(t_1) \ \geq \ (\text{соотв.} >) \ [f]_{\overline{x},\ell}(t_0) + \{[f]_{\overline{x},\ell}\}'_t(t_0) \cdot (t_1 - t_0) \\ \forall\ t_0, t_1 \in A_{\overline{x},\ell}, t_0 \neq t_1, \end{array}$$

или, что одно и то же,

$$\begin{array}{l} f(\overline{x} + t_1\ell) \ \geq \ (\text{соотв.} >) \ f(\overline{x} + t_0\ell) + \nabla f(\overline{x} + t_0\ell)(t_1\ell - t_0\ell) \\ \forall\ t_0, t_1 \in A_{\overline{x},\ell}, t_0 \neq t_1. \end{array}$$

Последнее условие можно переписать в виде:

$$\left.\begin{array}{l} \text{для любого сечения } X_{\overline{x},\ell} \text{ ненулевой длины имеем:} \\ f(x_1) \ \geq \ (\text{соотв.} >) \ f(x_0) + \nabla f(x_0)(x_1 - x_0), \\ \text{если } x_0, x_1 \in X_{\overline{x},\ell}, \ x_0 \neq x_1. \end{array}\right\} \tag{2.20}$$

Так как любые две точки $x_0, x_1 \in X$ попадают на некоторое сечение $X_{\overline{x},\ell}$, то (2.20) эквивалентно (2.19). $\boxtimes$

9. Критерий выпуклости в классе дважды дифференцируемых функций нескольких переменных. Нам потребуется понятие положительной определенности матриц. Пусть A – симметричная $(n \times n)$-матрица. Матрица A называется *неотрицательно* (*положительно*) *определенной*, если

$$(A\xi, \xi) \geq 0 \quad (\text{ соотв. } (A\xi, \xi) > 0 \ ) \ \forall \xi \in \mathbf{R}^n, \quad \xi \neq 0_n. \tag{2.21}$$

Заметим, что в приведенном определении строку (2.21) можно заменить эквивалентной строкой

$$(A\xi, \xi) \geq 0 \quad (\text{ соотв. } (A\xi, \xi) > 0 \ ) \ \forall \xi \in \mathbf{R}^n, \quad \|\xi\| = 1. \tag{2.22}$$

Тот факт, что матрица A является неотрицательно (положительно) определенной, будем обозначать неравенством $A \geq 0$ (соотв. $A > 0$). Сформулируем удобный критерий положительной определенности, использующий понятие главного минора матрицы. Минор матрицы A называется *главным минором*, если множество номеров строк матрицы A, содержащих элементы данного минора, совпадает с множеством номеров столбцов, содержащих его элементы. Главный минор называется *ведущим главным минором* или *угловым минором*, если указанное множество номеров имеет вид $\{1, 2, \ldots, k\}$, где $k \leq n$.

Лемма 28 (*критерий Сильвестра*, см. [6], гл.10, §4).
$\{A \geq 0\} \Leftrightarrow \{$ *все главные миноры* A *неотрицательны* $\}$;
$\{A > 0\} \Leftrightarrow \{$ *все ведущие главные миноры* A *положительны* $\}$.

Пусть $X \subset \mathbf{R}^n$ – выпуклое множество и функция $f(.)$ дважды дифференцируема на X [8]. В этом случае матрица вторых производных (*матрица Гессе* или *гессиан*) функции $f(.)$

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

при всех $x \in X$ симметрична [9].

Теорема 11. *Если X открыто, то для того, чтобы функция $f(.)$ была выпукла на множестве X, необходимо и достаточно, чтобы матрица $\nabla^2 f(x)$ была неотрицательно определенной для всех $x \in X$* .

□ В силу леммы об одномерных сечениях: $\{f(.)$ выпукла на $X\} \Leftrightarrow$

$$\Leftrightarrow \{\quad [f]_{x_0,\ell}(.) \text{ выпукла на } A_{x_0,\ell}\ \forall x_0 \in X,\ \forall \ell \in \mathbf{R}^n, \ell \neq 0_n\}.$$

Но функция $[f]_{x_0,\ell}(.)$ дважды дифференцируема в точках множества $A_{x_0,\ell}$, которое в силу выпуклости X есть промежуток. По лемме 26, если $A_{x_0,\ell}$ имеет длину большую нуля, то

$$\{[f]_{x_0,\ell}(.) \text{ выпукла на } A_{x_0,\ell}\} \Leftrightarrow \{\{[f]_{x_0,\ell}(t)\}''_{tt} \geq 0, \quad t \in A_{x_0,\ell}\}\,.$$

В силу открытости X промежуток $A_{x_0,\ell}$ имеет ненулевую длину при любых $x_0 \in X$, $\ell \in \mathbf{R}^n$, $\ell \neq 0_n$. Легко проверяется, что

$$\{[f]_{x_0,\ell}(t)\}''_{tt} \equiv \{f(x_0 + t\ell)\}''_{tt} = \left(\nabla^2 f(x_0 + t\ell)\ell, \ell\right), \quad t \in A_{x_0,\ell}. \qquad (2.23)$$

[8] То есть функция $f\,(.)$ дважды дифференцируема в каждой точке множества X. Отсюда следует, что $f\,(.)$ определена и дифференцируема в некоторой окрестности каждой точки множества X.

[9] Напомним (см., например, [7], теоремы 13.13, 13.14): если функция $f\,(\cdot)$, определенная в окрестности точки $\overline{x} \in \mathbf{R}^n$, дважды дифференцируема в точке $\overline{x}$ (отсюда следует, что она один раз дифференцируема в некоторой окрестности $\overline{x}$), то для каждой пары чисел $i, j \in \{1, 2, \ldots, n\}$ производные $\partial^2 f/\partial x^i \partial x^j$ и $\partial^2 f/\partial x^j \partial x^i$ в точке $\overline{x}$ совпадают. Это значит, что матрица вторых производных $\bigtriangledown^2 f\,(\overline{x})$ симметрична.

Следовательно, $f(.)$ выпукла на X тогда и только тогда, когда

$$\big(\nabla^2 f(x_0 + t\ell)\ell, \ell\big) \geq 0 \quad \forall t \in A_{x_0,\ell} \text{ при любых } x_0 \in X, \ell \in \mathbf{R}^n, \ell \neq 0_n. \quad (2.24)$$

Взяв в (2.24) $t = 0$, получаем

$$\big(\nabla^2 f(x_0)\ell, \ell\big) \geq 0 \text{ при любых } x_0 \in X, \ell \in \mathbf{R}^n, \ell \neq 0_n$$

что можно записать в виде: $\nabla^2 f(x_0) \geq 0$ при любых $x_0 \in X$. ⊠

Упражнение 31. 1) Проверьте формулу (2.23). 2) Покажите, что условие открытости множества X в теореме 11 существенно. Рассмотрите пример: $n = 2$, $X = \left\{x \in \mathbf{R}^2 : x^1 = x^2\right\}$, функция определена на $\mathbf{R}^2$ формулой $f\,(x) \equiv x^1 x^2$.

Используя лемму 27, легко доказать следующее достаточное условие строгой выпуклости.

Теорема 12. *Если матрица $\nabla^2 f(x)$ является положительно определенной при всех $x \in X$, то функция $f(.)$ строго выпукла на X.*

Упражнение 32. 1) Докажите теорему 12. 2) Пусть $X \subset \mathbf{R}^n$ – выпуклое множество с непустой внутренностью, а функция $f(.)$ непрерывна на X и дважды дифференцируема на $\overset{\circ}{X}$. Докажите, что тогда условие выпуклости $f(.)$ на X эквивалентно условию $\left\{\nabla^2 f(x) \geq 0 \ \forall x \in \overset{\circ}{X}\right\}$, а условие $\left\{\nabla^2 f(x) > 0 \ \forall x \in \overset{\circ}{X}\right\}$ достаточно для строгой выпуклости $f(.)$ на X. (См. упражнение 22, п.п. 3), 4).)

Упражнение 33. Проверьте на выпуклость и строгую выпуклость на указанных множествах X следующие функции:
1) $f(x) = x^2$, $X = \mathbf{R}$; 2) $f(x) = \exp(x)$, $X = \mathbf{R}$; 3) $f(x) = -\ln x$, $X = \{x \in \mathbf{R} : x > 0\}$; 4) $f(x) = x^4$, $X = \mathbf{R}$; 5) $f(x) = (x^1)^2 + x^1x^2 + (x^2)^2$, $X = \mathbf{R}^2$; 6) $f(x) = (x^1)^2 - (x^2)^2$, $X = \mathbf{R}^2$; 7) $f(x) = \sqrt{(x^1)^2 + (x^2)^2}$, $X = \mathbf{R}^2$; 8) $f(x) = \|x\|$, $X = \mathbf{R}^n$; 9) $f(x) = \sqrt{1 + (x^1)^2 + (x^2)^2}$, $X = \mathbf{R}^2$.

Упражнение 34. Пусть A – симметричная $(n \times n)$-матрица, $f(x) = \frac{1}{2}(Ax, x)$, $x \in \mathbf{R}^n$, – симметричная квадратичная форма n переменных (напомним, что к такому виду может быть приведена любая квадратичная форма n переменных). 1) Докажите, что $\nabla f(x) \equiv Ax$, $\nabla^2 f(x) \equiv A$. 2) Докажите, что $f(.)$ выпукла на $\mathbf{R}^n$ тогда и только тогда, когда $A \geq 0$. 3) Докажите, что $f(.)$ строго выпукла на $\mathbf{R}^n$ тогда и только тогда, когда $A > 0$.

10. Точки минимума выпуклых функций. Пусть $X \subset \mathbf{R}^n$ – непустое выпуклое множество, $f(.) : X \to \mathbf{R}$ – выпуклая функция. Обозначим через X_* множество точек глобального минимума в задаче минимизации $f(.)$ на X, то есть множество всех тех точек $\overline{x} \in X$, для каждой из которых

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

Теорема 13. 1) *Множество X_* выпукло.* 2) *Всякая точка локального минимума функции $f(.)$ на множестве X является и ее точкой глобального минимума на X, то есть принадлежит множеству X_*.*

☐ 1) Случаи пустого и одноточечного X_* тривиальны. Пусть X_* состоит более чем из одной точки, а x_0, x_1 – две различные точки из X_*. В силу выпуклости X точка $x_\lambda \equiv \lambda x_1 + (1-\lambda)x_0 \in X \ \forall \lambda \in [0,1]$. Выпуклость $f(.)$ дает $\forall \lambda \in [0,1]$

$$f(x_\lambda) \leq \lambda f(x_1) + (1-\lambda) f(x_0) \equiv \lambda \min_{x\in X} f(x) + (1-\lambda) \min_{x\in X} f(x) = \min_{x\in X} f(x).$$

Поэтому $f(x_\lambda) = \min\limits_{x\in X} f(x) \ \forall \lambda \in [0,1]$. То есть $x_\lambda \in X_* \ \forall \lambda \in [0,1]$.Значит, X_* выпукло.

2) Предположим, что x_0 – точка локального минимума $f(.)$ на X, но $x_0 \notin X_*$. Тогда существует точка $x_1 \in X$ такая, что $f(x_1) < f(x_0)$. В силу выпуклости X точка $x_\lambda \equiv \lambda x_1 + (1-\lambda)x_0 \in X \ \forall \lambda \in [0,1]$. При $\lambda \in (0,1)$ имеем

$$f(x_\lambda) \leq \lambda f(x_1) + (1-\lambda) f(x_0) < \lambda f(x_0) + (1-\lambda) f(x_0) = f(x_0). \qquad (2.25)$$

Так как в любой окрестности $U_\varepsilon(x_0)$ найдутся точки вида x_λ, $\lambda \in (0,1)$, то (2.25) противоречит тому, что x_0 – точка локального минимума. Значит, предположение о том, что $x_0 \notin X_*$ неверно. ⊠

Таким образом, для выпуклой функции $f(.) : X \to \mathbf{R}$ множество X_* точек ее глобального минимума совпадает с множеством ее точек локального минимума. Поэтому в случае выпуклых функций вместо терминов "точка глобального минимума"и "точка локального минимума"будем использовать термин "точка минимума".

Теорема 14. *Если $f(.)$ строго выпукла на X, то множество ее точек минимума X_* состоит не более чем из одной точки.*

☐ Пусть X_* состоит более чем из одной точки, а x_0, x_1 – две различные точки из X_*. Тогда $\forall \lambda \in (0,1) : x_\lambda \equiv (\lambda x_1 + (1-\lambda)x_0) \in X_*$ и $\forall \lambda \in (0,1) :$

$$f(x_\lambda) < \lambda f(x_1) + (1-\lambda) f(x_0) = \lambda \min_{x\in X} f(x) + (1-\lambda) \min_{x\in X} f(x) = \min_{x\in X} f(x),$$

что противоречит предположению x_0, $x_1 \in X_*$. ⊠

Пример экспоненты $f(x) \equiv \exp(x), x \in X \equiv \mathbf{R}$ показывает, что у строго выпуклой функции множество точек минимума X_* может быть пустым.

Теорема 15 (*критерий точки минимума выпуклой функции*). *Пусть $x_0 \in X$. Тогда*

$$\{x_0 \in X_*\} \Leftrightarrow \{\partial f(x_0)/\partial \ell \geq 0 \ \forall \ell \in K(x_0, X)\}. \qquad (2.26)$$

☐ *Необходимость*. Пусть $x_0 \in X_*$ и $\ell \in K(x_0, X)$. Тогда для всех $t > 0$ таких, что $(x_0 + t\ell) \in X$, имеем: $f(x_0 + t\ell) \geq f(x_0)$. Поэтому $\partial f(x_0)/\partial \ell \geq 0$.

Достаточность. Пусть $x_0 \in X$ такова, что

$$\partial f(x_0)/\partial\ell \geq 0 \ \forall \ell \in K(x_0, X).$$

Предположим, что $x_0 \notin X_*$. Это означает существование точки $x_1 \in X$ такой, что $f(x_1) < f(x_0)$. В силу выпуклости X точка $x_\lambda \equiv \lambda x_1 + (1-\lambda)x_0 = x_0 + \lambda(x_1 - x_0) \in X \ \forall\lambda \in [0,1]$. В силу выпуклости $f(\cdot)$ имеем

$$f(x_\lambda) \leq \lambda f(x_1) + (1-\lambda) f(x_0) \equiv f(x_0) + \lambda \{f(x_1) - f(x_0)\} \ \ \forall\lambda \in [0,1]$$

и, следовательно,

$$\frac{f(x_0 + \lambda(x_1 - x_0)) - f(x_0)}{\lambda} \leq \{f(x_1) - f(x_0)\} \quad \forall\lambda \in (0,1]. \tag{2.27}$$

Так как $\{f(x_1) - f(x_0)\} < 0$, то (2.27) означает, что в возможном для X в точке x_0 направлении вектора $\ell = (x_1 - x_0)$ производная $\partial f(x_0)/\partial\ell < 0$. Полученное противоречие доказывает, что $x_0 \in X_*$. ☒

Пусть теперь, в дополнение к сказанному выше, функция $f(.)$ дифференцируема на множестве X.Тогда для любого вектора $\ell \in K\,(x_0, X)$ производная $\partial f(x_0)/\partial\ell$, очевидно, конечна и ее можно записать в виде

$$\frac{\partial f(x_0)}{\partial \ell} \equiv \nabla f(x_0)\ \ell. \tag{2.28}$$

Следствие (*критерий точки минимума дифференцируемой выпуклой функции*). *Пусть $f(.)$ дифференцируема и выпукла на X, а $x_0 \in X$. Тогда*

$$\{x_0 \in X_*\} \Leftrightarrow \{\nabla f(x_0)\,(x - x_0) \geq 0 \quad \forall x \in X\}\,. \tag{2.29}$$

☐ *Необходимость.* Пусть $x_0 \in X_*$. Произвольно фиксируем точку $x \in X$, $x \neq x_0$. Вектор $\ell \equiv (x - x_0)$ принадлежит конусу $K(x_0, X)$ и, следовательно, по теореме 15 производная $\partial f(x_0)/\partial\ell \geq 0$. В силу (2.28) имеем $\nabla f(x_0)\,(x - x_0) \equiv \partial f(x_0)/\partial\ell \geq 0$. Необходимость доказана.

Достаточность. Пусть $\forall x \in X$ выполнено условие $\nabla f(x_0)\,(x - x_0) \geq 0$. Возьмем любое $\ell \in K(x_0, X)$. Если $\ell = 0_n$, то $\partial f(x_0)/\partial\ell = 0$. Если же $\ell \neq 0_n$, то обязательно найдутся точка $x_1 \in X$, $x_1 \neq x_0$ и число $t > 0$ такие, что $x_1 = x_0 + t\ell$, то есть $\ell = t^{-1}(x_1 - x_0)$. С помощью (2.28) получаем

$$\frac{\partial f(x_0)}{\partial \ell} = \nabla f(x_0) \left(\frac{1}{t}(x_1 - x_0)\right) = \frac{1}{t} \nabla f(x_0)\,(x_1 - x_0) \geq 0.$$

Таким образом $\partial f(x_0)/\partial\ell \geq 0 \ \forall\ell \in K\,(x_0, X)$ и $x_0 \in X_*$ по теореме 15. ☒

Упражнение 35. Определите, при каких значениях параметров a, b, c функция $f(x) \equiv (x^1)^2 + 2ax^2x^2 + b(x^2)^2 + c(x^3)^2$, $x \in \mathbf{R}^3$: 1) имеет точки минимума в $\mathbf{R}^3$; 2) имеет единственную точку минимума в $\mathbf{R}^3$.

Упражнение 36. Пусть $X = \mathbf{R}^2_+ \equiv \{x \in \mathbf{R}^2 : x^i \geq 0, i = 1,2\}$, $f(.)$ – выпуклая на X и дифференцируемая на X функция, $x_0 \in X$. Докажите, что: 1) если $x_0^1 > 0$, $x_0^2 > 0$, то $\{x_0 \in X_*\} \Leftrightarrow \{f'_{x^1}(x_0) = 0,\ f'_{x^2}(x_0) = 0\}$; 2) если $x_0^1 = 0$, $x_0^2 > 0$, то $\{x_0 \in X_*\} \Leftrightarrow \{f'_{x^1}(x_0) \geq 0, f'_{x^2}(x_0) = 0\}$; 3) если $x_0 = 0_2$, то $\{x_0 \in X_*\} \Leftrightarrow \{f'_{x^i}(x_0) \geq 0, i = 1,2\}$.

11. Сильно выпуклые функции. Для функции, непрерывной на непустом ограниченном замкнутом в $\mathbf{R}^n$ множестве X, множество точек глобального минимума X_* непусто (теорема Вейерштрасса). Условие ограниченности X при этом существенно. Как показывает пример экспоненты $f(x) \equiv \exp(x), x \in X \equiv \mathbf{R}$, даже у непрерывной строго выпуклой на замкнутом (но не ограниченном) множестве функции множество X_* точек минимума может оказаться пустым. Подобного не может быть у так называемых сильно выпуклых функций, играющих важную роль в теории оптимизации.

Прежде чем дать определение сильно выпуклой функции, договоримся о следующем обозначении: для симметричных $(n \times n)$-матриц A и B будем писать $A \geq B$, если $A - B \geq 0$. Везде ниже для простоты будем считать, что $X \subset \mathbf{R}^n$ – выпуклое множество с непустой внутренностью. Для простоты сначала мы рассмотрим понятие сильно выпуклой функции, ограничившись классом $C_2(X)$ дважды непрерывно дифференцируемых на X функций, а потом покажем как это понятие распространить на более широкие классы функций.

Случай дважды непрерывно дифференцируемых функций. Функцию $f(.) \in C_2(X)$ назовем сильно выпуклой функцией на множестве X, если существует число $\boldsymbol{\kappa} > 0$ (называемое константой сильной выпуклости функции $f(.)$ на X) такое, что

$$\nabla^2 f(x) \geq \boldsymbol{\kappa} E \quad \forall x \in X, \tag{2.30}$$

где E – единичная $(n \times n)$-матрица.

Упражнение 37. 1) Докажите, что следующие функции являются сильно выпуклыми на указанных множествах X : $f(x) = x^2$, $X = \mathbf{R}$; $f(x) = \exp(x), x \in \mathbf{R}$, $X = \mathbf{R}_+$; $f(x) = \|x\|^2, X = \mathbf{R}^n$. 2) Докажите, что функция $f(x) = x^4$ не является сильно выпуклой на $X = \mathbf{R}$. 3) Пусть A – симметричная $(n \times n)$-матрица. Докажите, что квадратичная форма $f(x) = (Ax, x)$, $x \in \mathbf{R}^n$ является сильно выпуклой функцией на $\mathbf{R}^n$ тогда и только тогда, когда $A > 0$.

Класс сильно выпуклых функций является подклассом класса строго выпуклых функций.

Теорема 16. *Сильно выпуклая функция является строго выпуклой.*

Упражнение 38. Докажите теорему 16.

Важным является следующее свойство сильно выпуклых функций.

Теорема 17. *Если $f(\cdot)$ – сильно выпуклая функция на X, то любое ее множество Лебега $M_f(c) \equiv \{x \in X : f(x) \leq c\}$ ограничено в $\mathbf{R}^n$.*

$\square$ Без ограничения общности будем считать, что $0_n \in X$. Рассмотрим сначала случай $n = 1$. В этом случае X – промежуток в $\mathbf{R}$ и дважды непрерывно дифференцируемую функцию $f(\cdot)$ можно представить на нем в виде

$$f(x) = f(0) + f'(0)x + \int\limits_0^x dt \int\limits_0^t f''(\xi) d\xi, \quad x \in X. \tag{2.31}$$

Условие сильной выпуклости $f(.)$ на промежутке X означает, что $f''(x) \geq \kappa > 0 \ \forall x \in X$. Вместе с (2.31) это дает неравенство

$$f(x) \geq g(x) \equiv f(0) + f'(0)x + 2^{-1}\kappa x^2, \quad x \in X,$$

из которого следует, что $M_f(c) \subset M_g(c) \quad \forall c \in \mathbf{R}$. В силу выпуклости функций $f(.)$ и $g(.)$ множества Лебега $M_f(c)$ и $M_g(c)$ это промежутки. Очевидно, что при любом $c \in \mathbf{R}$ длина промежутка $M_g(c)$, а вместе с ней и длина $M_f(c)$, конечна. Длина промежутка $M_g(c)$ вполне определяется коэффициентами квадратного трехчлена $g(.)$ и значением c, то есть величинами c, κ, $f(0)$, $f'(0)$. Более того, длина $M_g(c)$ непрерывно зависит от этих величин, то есть существует непрерывная функция $\Psi(.,.,.,.) : \mathbf{R} \times \overset{\circ}{\mathbf{R}}_+ \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ такая, что

$$\{\text{длина } M_g(c)\} = \Psi(c, \kappa, f(0), f'(0)).$$

Значит,

$$\{\text{длина } M_f(c)\} \leq \Psi(c, \kappa, f(0), f'(0)). \tag{2.32}$$

Перейдем к случаю $n > 1$. Зафиксировав произвольно $\ell \in \mathbf{R}^n$, $\|\ell\| = 1$, рассмотрим сечение $X_{0_n,\ell}$. Введем новые функции Φ и φ формулами:

$$\begin{aligned} &\Phi(x) \equiv f(x), x \in X_{0_n,\ell}; \\ &\varphi(t) \equiv [f]_{0_n,\ell}(t), t \in A_{0_n,\ell}. \end{aligned}$$

Ясно, что $M_f(c) \cap X_{0_n,\ell} = M_\Phi(c)$. Очевидно, $\Phi(.)$ – сильно выпуклая функция на $X_{0_n,\ell}$, а потому и $\varphi(.)$ – сильно выпуклая функция на $A_{0_n,\ell}$, причем:

$$\varphi''(t) = \left(\nabla^2 f(t\ell)\ell, \ell\right) \geq (\kappa E\ell, \ell) = \kappa, \ t \in A_{\theta_n,\ell}. \tag{2.33}$$

Так как $\|\ell\| = 1$, то

$$\{\text{длина } M_\Phi(c)\} = \{\text{ длина } M_\varphi(c)\}.$$

Поэтому, в силу (2.32) и (2.33): $\{$ длина $M_\Phi(c)\} \leq \Psi(c, \kappa, \varphi(0), \varphi'(0))$. Так как $\varphi(0) = f(0_n)$, $\varphi'(0) = \nabla f(0_n)\ell$, то

$$\{\text{длина } M_\Phi(c)\} \leq \gamma(c, \kappa) \equiv \max_{\|\ell\|=1} \Psi\left(c, \kappa, f(0_n), \nabla f(0_n)\ell\right).$$

Следовательно, $M_f(c) \subset \overline{U_\delta(0_n)}$ при $\delta = \gamma(c, \kappa)$. ⊠

Теперь мы можем доказать основное свойство сильно выпуклых функций.

Теорема 18. *Если $f(.)$ – сильно выпуклая функция на замкнутом множестве $X \subset \mathbf{R}^n$, то множество X_* ее точек минимума на X состоит ровно из одной точки.*

□ Так как по теореме 16 функция $f(.)$ строго выпукла на X, то X_* состоит не более чем из одной точки (теорема 14). Достаточно доказать, что $X_* \neq \varnothing$. Выберем $c \in \mathbf{R}$ так, что $M_f(c) \neq \varnothing$. Тогда $\{M_f(c)\}_* = X_*$. Множество $M_f(c)$ ограничено (теорема 17) и замкнуто ($f(.)$ непрерывна на X). По теореме Вейерштрасса $\{M_f(c)\}_* \neq \varnothing$. ⊠

Упражнение 39. Пусть $f(.) \in C_2(\mathbf{R}^n)$. Известно (см., например, [4], С.463), что если в точке $x_0 \in \mathbf{R}^n$ имеем $\nabla f(x_0) = \{0, \ldots, 0\}$, $\nabla^2 f(x_0) > 0$, то x_0 – точка локального минимума $f(.)$; точки локального минимума, удовлетворяющие указанным условиям, называют невырожденными. Докажите, что в некоторой окрестности любой своей невырожденной точки минимума функция $f(.)$ является сильно выпуклой функцией.

Теперь покажем, как можно распространить определение сильно выпуклой функции на классы более широкие, чем класс дважды непрерывно дифференцируемых функций. Начнем с класса дифференцируемых функций, а потом рассмотрим общий случай.

Случай дифференцируемых функций. Пусть $f(.)$ – некоторая дифференцируемая на X функция. Введем обозначения:

$$g(\kappa; x, x_0) \equiv f(x_0) + \nabla f(x_0)\,(x - x_0) + 2^{-1}\kappa\|x - x_0\|^2, \quad x, x_0 \in X, \kappa > 0;$$
$$G(\kappa; x, x_0) \equiv f(x) - g(\kappa; x, x_0), \quad x, x_0 \in X, \kappa > 0.$$

Заметим, что для дважды непрерывно дифференцируемой на X функции $f(.)$ условие (2.30) можно записать также в виде:

$$\forall x_0 \in X: \quad \nabla_x^2 G(\kappa; x, x_0) \geq 0, \quad x \in X. \tag{2.34}$$

В силу теоремы 11 условие (2.34) эквивалентно условию:

$$\forall x_0 \in X \quad \text{функция} \quad G(\kappa; ., x_0) \text{ выпукла на } \quad X. \tag{2.35}$$

В свою очередь условие (2.35) эквивалентно условию

$$G(\kappa; x, x_0) \geq 0, \quad x, x_0 \in X. \tag{2.36}$$

Упражнение 40. С помощью теоремы 10 покажите, что в случае дифференцируемой на X функции $f(.)$ условие (2.35) эквивалентно условию (2.36).

Таким образом, условие (2.30), определяющее понятие сильной выпуклости в классе дважды непрерывно дифференцируемых на X функций, эквивалентно в этом классе функций условию (2.36) или, что то же самое, условию

$$f(x) \geq f(x_0) + \nabla f(x_0)\,(x - x_0) + 2^{-1}\kappa\|x - x_0\|^2, \quad x, x_0 \in X. \tag{2.37}$$

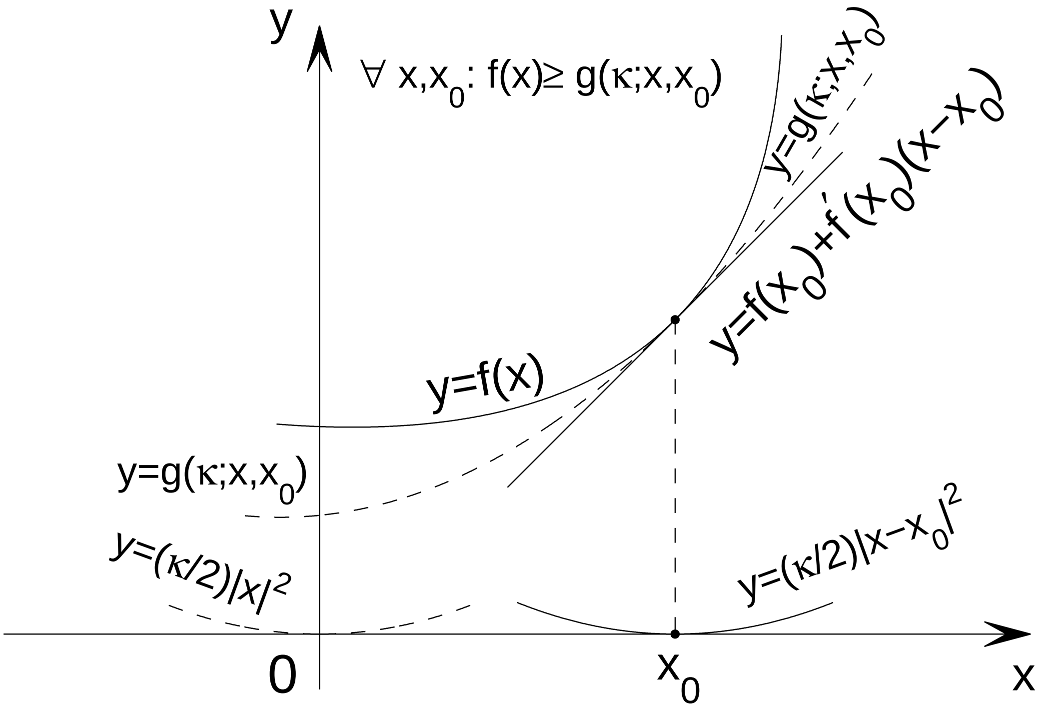

Рис. 6

Так как в классе функций, дважды непрерывно дифференцируемых на множестве X, условие (2.30) эквивалентно не содержащему вторых производных условию (2.37), то дают следующее определение сильно выпуклой функции в классе дифференцируемых на X функций. Функцию $f(.)$, дифференцируемую на множестве $X \subset \mathbf{R}^n$, называют сильно выпуклой функцией на X, если существует $\kappa > 0$ такое, что имеет место (2.37). Это определение проиллюстрировано для $n = 1$ на рис. 23.

Общий случай. Пусть $f(.) : X \to \mathbf{R}$ – некоторая функция. Положим

$$\sigma\,(\kappa, x) \equiv f(x) - 2^{-1}\kappa(Ex, x), \quad x \in X, \ \kappa > 0.$$

Заметим, что для дважды непрерывно дифференцируемой на X функции $f(.)$ условие (2.30) можно записать также в виде:

$$\nabla_x^2 \sigma\,(\kappa, x) \geq 0, \quad x \in X,$$

а так как $\overset{\circ}{X} \neq \varnothing$ и вторые производные функции $f\,(.)$ непрерывны на X, то и в виде

$$\nabla_x^2 \sigma\,(\kappa, x) \geq 0, \quad x \in \overset{\circ}{X}.$$

Последнее эквивалентно тому (теорема 11), что функция $\sigma\,(\kappa, .)$ выпукла на $\overset{\circ}{X}$, а следовательно, в силу ее непрерывности на X, выпукла и на всем множестве X (см. упражнение 22, п.4)). В свою очередь, выпуклость функции

$\sigma(\kappa, .)$ на множестве X эквивалентна условию

$$f(x_\lambda) \leq \lambda f(x_1) + (1-\lambda) f(x_0) - 2^{-1}\kappa\lambda(1-\lambda)\|x_1 - x_0\|^2 \quad \text{при всех } x_0, x_1 \in X, \lambda \in (0,1). \tag{2.38}$$

Упражнение 41. Пользуясь неравенством выпуклости, покажите, что выпуклость функции $\sigma(\kappa, .)$ на X эквивалентна условию (2.38).

Так как в классе функций, дважды непрерывно дифференцируемых на множестве X, условие (2.30) эквивалентно не содержащему производных условию (2.38), то дают следующее общее определение сильно выпуклой функции. Функцию $f(.) : X \to \mathbf{R}$ называют сильно выпуклой функцией на X, если найдется число $\kappa > 0$ такое, что имеет место (2.38).

Литература

Цитированная литература

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

2. Васильев Ф.П. Численные методы решения экстремальных задач. - М.: Наука, 1980 (1-ое издание), 1988 (2-ое издание).

3. Сухарев А.Г., Тимохов А.В., Федоров В.В. Курс методов оптимизации. - М.: Наука, 1986.

4. Зорич В.А. Математический анализ. Часть I. - М.: Наука, 1981.

5. Карманов В.Г. Математическое программирование. - М.: Наука, 1986.

6. Гантмахер Ф.Р. Теория матриц. - М.: Наука, 1966.

7. Ильин В.А., Садовничий В.А., Сендов Бл.Х. Математический анализ. - М.: Наука, 1979.

Дополнительная литература

8. Алексеев В.М., Тихомиров В.М., Фомин С.В. Оптимальное управление. - М.: Наука, 1979.

9. Пшеничный Б.Н. Выпуклый анализ и экстремальные задачи. - М.: Наука, 1980.

10. Алексеев В.М., Галеев Э.М., Тихомиров В.М. Сборник задач по оптимизации. Теория. Примеры. Задачи. Учебное пособие. - М.: Наука, 1984.

11. Галеев Э.М., Тихомиров В.М. Краткий курс теории экстремальных задач. - М.: Изд-во Моск. ун-та, 1989.

12. Магарил-Ильяев Г.Г., Тихомиров В.М. Выпуклый анализ и его приложения. - М.: Эдиториал УРСС, 2000.

13. Галеев Э.М., Тихомиров В.М. Оптимизация: теория, примеры, задачи. - М.: Эдиториал УРСС, 2000.

14. Васильев Ф.П. Методы оптимизации: В 2-х кн. - Новое изд., перераб. и доп. - М.: МЦНМО, 2011.

15. Дмитрук А.В. Выпуклый анализ. Элементарный вводный курс: Учебное пособие. - М.: Издательский отдел факультета ВМиК МГУ; МАКС Пресс, 2012.

Владимир Иосифович **Сумин**

НАЧАЛА ВЫПУКЛОГО АНАЛИЗА.
Часть 1.
ВЫПУКЛЫЕ МНОЖЕСТВА

Учебно-методическое пособие

Нижегородский государственный университет им. Н.И. Лобачевского

Механико-математический факультет

Кафедра математической физики