§2 ВЫПУКЛЫЕ ФУНКЦИИ

Опр. 2.1. *Функция* $f\colon X \to \mathbb{R}$, *заданная на выпуклом множестве* $X \subset \mathbb{R}^n$, *называется выпуклой, если для любых* x, $y \in X$ *и любого* $\alpha \in [0,1]$ *выполнено неравенство* $f\bigl(\alpha x + (1-\alpha)y\bigr) \le \alpha f(x) + (1-\alpha)f(y)$. *Если последнее неравенство строгое при* $\alpha \in (0,1)$, *то функция* f *называется строго выпуклой.*

Лемма 2.1. *Для того, чтобы функция* $f\colon X \to \mathbb{R}$, *определённая на выпуклом множестве* X, *была выпуклой, необходимо и достаточно, чтобы было выпуклым множество* $\operatorname{epi} f \overset{\text{def}}{=} \{(x,y)\colon x \in X, y \ge f(x)\} \subset X \times \mathbb{R}$. (*Множество* $\operatorname{epi} f$ *называется надграфиком функции* f.)

$\triangleright$ Пусть $f\colon X \to \mathbb{R}$ — выпуклая функция. Выбрав произвольные две точки $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2) \in \operatorname{epi} f$ и число $\alpha \in [0,1]$, докажем, что $\alpha z_1 + (1-\alpha)z_2 \in$ $\in \operatorname{epi} f$, т.е. что $f\bigl(\alpha x_1 + (1-\alpha)x_2\bigr) \le \alpha y_1 + (1-\alpha)y_2$. Так как $f(x_1) \le y_1$ и $f(x_2) \le y_2$, то необходимое неравенство следует из выпуклости функции f.

Предположим теперь, что $\operatorname{epi} f$ — выпуклое множество. Очевидно, что для произвольных двух точек x_1, $x_2 \in X$ пары $z_1 = \bigl(x_1, f(x_1)\bigr)$, $z_2 = \bigl(x_2, f(x_2)\bigr)$ принадлежат надграфику функции f. Следовательно, для произвольного числа $\alpha \in [0,1]$ имеем $\alpha z_1 + (1-\alpha)z_2 \in \operatorname{epi} f$, а значит, $f\bigl(\alpha x_1 + (1-\alpha)x_2\bigr) \le \alpha f(x_1) +$ $+ (1-\alpha)f(x_2)$. Другими словами, функция f выпукла. $\triangleleft$

Лемма 2.2 (Неравенство Йенсена). *Пусть* $f\colon X \to \mathbb{R}$ — *выпуклая функция. Тогда для произвольных точек* x_1, x_2, $\ldots$, $x_m \in X$ *и чисел* α_1, α_2, $\ldots$, $\alpha_m \ge 0$, *таких что* $\sum\limits_{i=1}^{m} \alpha_i = 1$, *справедливо неравенство*

$$f(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m) \le \alpha_1 f(x_1) + \alpha_2 f(x_2) + \ldots + \alpha_m f(x_m). \tag{2.1}$$

$\triangleright$ Докажем неравенство (2.1) индукцией по количеству точек m. База индукции $m = 2$ следует из определения выпуклой функции. Предположим, что неравенство (2.1) верно при $m \ge 2$. Пусть x_1, x_2, $\ldots$, $x_{m+1} \in X$ и числа α_1, α_2, $\ldots$, $\alpha_{m+1} \ge 0$, такие что $\sum\limits_{i=1}^{m+1} \alpha_i = 1$. Без нарушения общности будем считать, что $\alpha_1 < 1$. Так как $\sum\limits_{i=2}^{m+1} \dfrac{\alpha_i}{1-\alpha_1} = 1$, то верна цепочка неравенств

$$f\left(\sum_{i=1}^{m+1} \alpha_i x_i\right) \le \alpha_1 f(x_1) + (1-\alpha_1) f\left(\sum_{i=2}^{m+1} \frac{\alpha_i}{1-\alpha_1} x_i\right) \le \sum_{i=1}^{m+1} \alpha_i f(x_i). \quad \triangleleft$$

Напомним, что точка x называется внутренней для множества X, если найдётся такое $r > 0$, что $B_r(x) \subset X$.

Лемма 2.3. *Выпуклая функция* $f\colon X \to \mathbb{R}$ *непрерывна во всех внутренних точках множества* $X \subset \mathbb{R}^n$.

$\triangleright$ Пусть x_0 — внутренняя точка множества X, а значит, $B_r(x_0) \subset X$ для некоторого $r > 0$. Пусть $\{e_1, e_2, \ldots, e_n\}$ — стандартный базис пространства $\mathbb{R}^n$. Тогда для любого $\alpha \in [0, 1)$ справедливо неравенство

$$f(x_0 \pm \alpha r e_i) \le (1-\alpha) f(x_0) + \alpha f(x_0 \pm r e_i).$$

Следовательно, $\overline{\lim\limits_{t\to+0}} f(x_0 \pm t e_i) \le f(x_0)$, поэтому, $\overline{\lim\limits_{x\to x_0}} f(x) \le f(x_0)$. Так как $f(x_0) \le \dfrac{1}{2} f(x_0 + h) + \dfrac{1}{2} f(x_0 - h)$ для любого $h \in B_r(\mathbf{0})$, то

$$f(x_0) \le \frac{1}{2} \varliminf_{x\to x_0} f(x) + \frac{1}{2} \varlimsup_{x\to x_0} f(x).$$

Следовательно, $f(x_0) = \varliminf\limits_{x\to x_0} f(x) = \varlimsup\limits_{x\to x_0} f(x)$, т.е. функция f непрерывна в x_0. $\triangleleft$

Опр. 2.2. *Вектор $c \in \mathbb{R}^n$ называется субградиентом функции $f\colon X \to \mathbb{R}$ в точке $x_0 \in X \subset \mathbb{R}^n$, если $f(x) \ge f(x_0) + c^{\mathsf{T}}(x - x_0)$ для всех $x \in X$. Множество всевозможных субградиентов функции f в точке x_0 называется субдифференциалом функции f в точке x_0 и обозначается $\partial f(x_0)$.*

Лемма 2.4. *Пусть $f\colon X \to \mathbb{R}$ — выпуклая функция, а x_0 — внутренняя точка множества X. Тогда множество $\partial f(x_0)$ непусто.*

$\triangleright$ Пусть $c^{\mathsf{T}} x + by = d$ — уравнение опорной гиперплоскости к множеству $\operatorname{epi} f$ в точке $\big(x_0, f(x_0)\big)$. Тогда $c^{\mathsf{T}} x + by \ge d$ при $(x, y) \in \operatorname{epi} f$ и $c^{\mathsf{T}} x_0 + b f(x_0) = d$. Докажем, что $b > 0$. Так как $\big(x_0, f(x_0) + 1\big) \in \operatorname{epi} f$, то $c^{\mathsf{T}} x_0 + b f(x_0) + b \ge d$, т.е. $b \ge 0$. Если $b = 0$, то $c^{\mathsf{T}} x \ge d$, $x \in X$, а значит, $c^{\mathsf{T}}(x - x_0) \ge 0$. Так как x_0 — внутренняя точка, то $x_0 - tc \in X$ для некоторого положительного числа $t > 0$. Следовательно, $-t\|c\|^2 \ge 0$, т.е. $c = 0$. Получено противоречие. Таким образом, $b > 0$, а значит,

$$f(x) \ge -\frac{c^{\mathsf{T}} x}{b} + \frac{d}{b} \quad \text{и} \quad f(x_0) = -\frac{c^{\mathsf{T}} x_0}{b} + \frac{d}{b}.$$

Наконец, отнимая последнее равенство от неравенства, получаем, что

$$f(x) - f(x_0) \ge (\widetilde{c}, x - x_0), \quad \text{где} \quad \widetilde{c} = -\frac{c}{b}. \quad \triangleleft$$

Имеет место следующее обобщение неравенства Йенсена.

Лемма 2.5. *Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$ — выпуклая функция, а $\xi\colon \Omega \to \mathbb{R}^n$ — случайный вектор. Тогда справедливо неравенство $f(E\xi) \le Ef(\xi)$, при условии, что соответствующие математические ожидания существуют.*

$\triangleright$ Так как $f(x) - f(y) \ge c_y^{\mathsf{T}}(x - y)$, где $c_y \in \partial f(y)$, то $f(x) = \max\limits_{y\in\mathbb{R}^n}(c_y^{\mathsf{T}} x + d_y)$. Следовательно,

$$f(E\xi) = \max_{y\in\mathbb{R}^n}(c_y^{\mathsf{T}} E\xi + d_y) = \max_{y\in\mathbb{R}^n} E(c_y^{\mathsf{T}} \xi + d_y) \le E \max_{y\in\mathbb{R}^n}(c_y^{\mathsf{T}} \xi + d_y) = Ef(\xi). \quad \triangleleft$$

Лемма 2.6. *Если в точке $x_0 \in X$ выпуклая функция $f\colon X \to \mathbb{R}$ — дифференцируема, то $\nabla f(x_0) \in \partial f(x_0)$.*

$\triangleright$ Пусть $x \in X$ и $t \in (0,1]$. Тогда $f\big(x_0 + t(x - x_0)\big) \le (1-t)f(x_0) + tf(x)$, а значит, $\dfrac{f\big(x_0 + t(x-x_0)\big) - f(x_0)}{t} \le f(x) - f(x_0)$. Устремляя t к 0, получаем, что

$$f(x) \ge f(x_0) + \nabla f(x_0)^{\mathsf{T}}(x - x_0). \quad \triangleleft$$

Лемма 2.7. *Если $f \in C^2(X)$, где X — открытое выпуклое множество, то для выпуклости функции f необходимо и достаточно, чтобы матрица $\nabla^2 f(x)$ была неотрицательно определённой ($\nabla^2 f(x) \succeq 0$). Если матрица $\nabla^2 f(x)$ положительно определена ($\nabla^2 f(x) \succ 0$), то f — строго выпуклая функция.*

$\triangleright$ Выберем произвольные точку $x_0 \in X$ и направление $\ell \ne \mathbf{0}$. Рассмотрим функцию $g(t) = f(x_0 + t\ell)$, заданную на интервале $T \overset{\text{def}}{=} \{t\colon x_0 + t\ell \in X\}$. Очевидно, что функция f (строго) выпукла тогда и только тогда, когда (строго) выпуклы скалярные функции g при всевозможных $x_0 \in X$ и $\ell \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Имеем $g'(t) = \ell^{\mathsf{T}}\nabla f(x_0 + t\ell)$ и $g''(t) = \ell^{\mathsf{T}}\nabla^2 f(x_0 + t\ell)\ell$. Функция g выпукла тогда и только тогда, когда $g''(t) \ge 0$, что равносильно неотрицательной определённости матрицы $\nabla^2 f(x)$. Если $\nabla^2 f(x) \succ 0$, то $g''(t) > 0$, а значит, g — строго выпуклая функция. $\triangleleft$

В заключении рассмотрим операции над функциями, сохраняющие выпуклость. Будем писать, что $x \le y$ для векторов $x, y \in \mathbb{R}^n$, если выполнены неравенства $x_i \le y_i$, $1 \le i \le n$.

Лемма 2.8. *Пусть f, f_1, f_2, …, f_m — выпуклые функции. Тогда следующие функции также являются выпуклыми:*

a) $g(x) = \sum\limits_{i=1}^{m} c_i f_i(x)$, *где* $c_i \ge 0$, $1 \le i \le m$;

b) $g(x) = f(Fx)$, *где* $Fx = Ax + b$ *— аффинное преобразование;*

c) $g(x) = \max\limits_{1 \le i \le m} f_i(x)$;

d) $g(x) = h\big(f_1(x), f_2(x), \ldots, f_m(x)\big)$, *где h — выпуклая монотонно неубывающая функция, т.е. $h(y) \le h(\widetilde{y})$ для всех y и $\widetilde{y}$, таких что $y \le \widetilde{y}$.*

$\triangleright$ Доказательства утверждений тривиальным образом следуют из определения выпуклости. Для примера докажем выпуклость функции g из пункта d). Пусть $x, y \in X$ и $\alpha \in [0,1]$. Так как функции f_i, $1 \le i \le m$, выпуклы по условию, то $f_i\big(\alpha x + (1-\alpha)y\big) \le \alpha f_i(x) + (1-\alpha)f_i(y)$. Положим $u = \big(f_1(x), \ldots, f_m(x)\big)$ и $v = \big(f_1(y), \ldots, f_m(y)\big)$, тогда в силу монотонности функции h верно неравенство

$$h\Big(f_1\big(\alpha x + (1-\alpha)y\big), \ldots, f_m\big(\alpha x + (1-\alpha)y\big)\Big) \le h\big(\alpha u + (1-\alpha)v\big).$$

Так как функция h выпукла, то $h\big(\alpha u + (1-\alpha)v\big) \le \alpha h(u) + (1-\alpha)h(v)$. Наконец, в силу определения функции g имеем

$$g\big(\alpha x + (1-\alpha)y\big) = h\Big(f_1\big(\alpha x + (1-\alpha)y\big), \ldots, f_m\big(\alpha x + (1-\alpha)y\big)\Big),$$

$g(x) = h(u)$ и $g(y) = h(v)$, а значит, $g\big(\alpha x + (1-\alpha)y\big) \leq \alpha g(x) + (1-\alpha)g(y)$. Следовательно, g — выпуклая функция. $\lhd$

Упражнения

4. Докажите, что непрерывная выпуклая функция $f\colon [a,b] \to \mathbb{R}$ удовлетворяет неравенству
$$f\left(\frac{a+b}{2}\right)(b-a) \leq \int\limits_a^b f(x)\,\mathrm{d}x \leq \frac{f(a)+f(b)}{2}(b-a).$$
5. Докажите, что субдифференциал $\partial f(x_0)$ произвольной выпуклой функции f в точке x_0 является замкнутым выпуклым множеством.
6. Пусть $f(x) \overset{\mathrm{def}}{=} \max\limits_{1\leq i\leq m} f_i(x)$, где $f_i(x)$ — выпуклые функции, и пусть c_i — субградиент функции f_i в точке x_0. Докажите, что вектор $c \overset{\mathrm{def}}{=} \sum\limits_{i=1}^m \alpha_i c_i$, $\sum\limits_{i=1}^m \alpha_i = 1$, где $\alpha_i \geq 0$ и $\alpha_i = 0$, если $f_i(x_0) < f(x_0)$, является субградиентом функции $f(x)$.
7. (Неравенство Караматы) Пусть даны два упорядоченных по невозрастанию набора из n действительных чисел $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ и $\mathbf{b} = (b_1, b_2, \ldots, b_n)$. Говорят, что набор $\mathbf{a}$ *мажорирует* набор $\mathbf{b}$, и пишут $\mathbf{a} \succ \mathbf{b}$, если $a_1 \geqslant b_1$, $a_1 + a_2 \geqslant b_1 + b_2$, $\ldots$, $a_1 + a_2 + \ldots + a_{n-1} \geqslant b_1 + b_2 + \ldots + b_{n-1}$, $a_1 + a_2 + \ldots + a_n = b_1 + b_2 + \ldots + b_n$. Докажите, что для любой выпуклой функции $y = f(x)$, определённой на некотором промежутке I, и любых двух наборов $\mathbf{a} = (a_1, a_2, \ldots, a_n)$, $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ из этого промежутка, удовлетворяющих условию $\mathbf{a} \succ \mathbf{b}$, справедливо неравенство
$$f(a_1) + f(a_2) + \ldots + f(a_n) \geqslant f(b_1) + f(b_2) + \ldots + f(b_n).$$
8. Пусть $f(x)$ и $g(x)$ — выпуклая и вогнутая функции соответственно, определённые на выпуклом множестве X, причём для любого $x \in X$ выполняется неравенство $f(x) \geq g(x)$. Докажите, что существует линейная функция $h(x)$, такая что
$$f(x) \geq h(x) \geq g(x) \quad \text{для каждого} \quad x \in X.$$