

Белорусский государственный университет
Факультет прикладной математики и информатики

**ИЗБРАННЫЕ РАЗДЕЛЫ
МЕТОДОВ ОПТИМИЗАЦИИ**

канд. физ.-мат. наук Войделевич А.С.

КОНСПЕКТ ЛЕКЦИЙ

Минск
2019

СОДЕРЖАНИЕ

§1	Выпуклые множества	3
§2	Выпуклые функции	7

§1 ВЫПУКЛЫЕ МНОЖЕСТВА

Пусть \mathbb{A}^n — n -мерное аффинное пространство над полем \mathbb{R} , $n \in \mathbb{N}$. Зафиксируем некоторую точку $O \in \mathbb{A}^n$ в качестве начала координат. Далее будет отождествлять произвольную точку $P \in \mathbb{A}^n$ с её радиусом вектором \overrightarrow{OP} , а само пространство \mathbb{A}^n — с вещественным n -мерным векторным пространством \mathbb{R}^n . Для обозначения векторов и точек будем использовать строчные буквы, а для обозначения множеств — заглавные.

Опр. 1.1. Точка $\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_m p_m \in \mathbb{R}^n$, где $p_i \in \mathbb{R}^n$, $\alpha_i \geq 0$ и $\sum_{i=1}^m \alpha_i = 1$, называется *выпуклой комбинацией точек* p_1, p_2, \dots, p_m .

Опр. 1.2. Для произвольных точек $x, y \in \mathbb{R}^n$ множество

$$[x, y] \stackrel{\text{def}}{=} \{\alpha x + (1 - \alpha)y : \alpha \in [0, 1]\},$$

состоящее из всех возможных выпуклых комбинаций точек x и y , называется *отрезком* (с концами x, y).

Опр. 1.3. Множество $X \subset \mathbb{R}^n$ называется *выпуклым*, если для произвольных двух точек $x, y \in X$ оно содержит весь отрезок $[x, y]$.

Отметим, что согласно определению 1.3 пустое множество \emptyset и произвольное одноточечное множество $\{p\}$, $p \in \mathbb{R}^n$, являются выпуклыми.

Лемма 1.1. Множество X является выпуклым, если и только если X содержит любую выпуклую комбинацию своих точек.

▷ Если множество X содержит любую выпуклую комбинацию своих точек, то, в частности, для любых двух точек $x, y \in X$ имеем $[x, y] \subset X$, а значит, X — выпуклое множество.

Обратное утверждение доказывается индукцией по количеству m точек $p_1, p_2, \dots, p_m \in X$, входящих в выпуклую комбинацию $\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_m p_m$. База индукции $m = 2$ следует из определения 1.3. Предположим теперь, что множество X содержит всевозможные выпуклые комбинации своих точек размера $m \geq 2$. Докажем, что X также содержит любую выпуклую комбинацию размера $m + 1$. Действительно, пусть $p_1, p_2, \dots, p_{m+1} \in X$ и числа $\alpha_1, \alpha_2, \dots, \alpha_{m+1} \geq 0$, такие что $\sum_{i=1}^{m+1} \alpha_i = 1$. Без нарушения общности будем считать, что $\alpha_1 < 1$ (иначе это выпуклая комбинация, состоящая из одной точки). Тогда

$$p = \alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_{m+1} p_{m+1} = \alpha_1 p_1 + (1 - \alpha_1)q,$$

где $q = \frac{\alpha_2}{1 - \alpha_1} p_2 + \frac{\alpha_3}{1 - \alpha_1} p_3 + \dots + \frac{\alpha_{m+1}}{1 - \alpha_1} p_{m+1}$. Так как $\sum_{i=2}^{m+1} \frac{\alpha_i}{1 - \alpha_1} = 1$, то согласно предположению индукции $q \in X$, а значит, $p \in X$. ◁

Рассмотрим операции над выпуклыми множествами, которые сохраняют выпуклость.

Лемма 1.2. Пусть I — некоторое множество индексов произвольной мощности, а $\{X_i \subset \mathbb{R}^n : i \in I\}$ — семейство выпуклых множеств. Тогда множество $X = \bigcap_{i \in I} X_i$ является выпуклым.

▷ Действительно, пусть $x, y \in X$. Тогда $x, y \in X_i$ для всех $i \in I$, а значит, $[x, y] \subset X_i$. Таким образом, $[x, y] \subset X$, т.е. множество X является выпуклым. ◁

Выпуклой оболочкой $\text{Conv } X$ произвольного множества $X \subset \mathbb{R}^n$ называется наименьшее (по вложению) выпуклое множество, содержащее X . Из леммы 1.2 следует, в частности, что $\text{Conv } X$ — это пересечение всех выпуклых множеств, содержащих X .

Лемма 1.3. Пусть $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ — аффинное преобразование, т.е. преобразование, действующее по правилу $F: x \mapsto Ax + b$, где $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Тогда для произвольных выпуклых множеств $X \subset \mathbb{R}^n$ и $Y \subset \mathbb{R}^m$ множества $F(X) \stackrel{\text{def}}{=} \{Fx : x \in X\} \subset \mathbb{R}^m$ и $F^{-1}(Y) \stackrel{\text{def}}{=} \{x : Fx \in Y\} \subset \mathbb{R}^n$ также являются выпуклыми.

▷ Пусть $y_1 = Fx_1$, $y_2 = Fx_2$ и $\alpha \in [0, 1]$. Тогда утверждение леммы следует из равенства $\alpha y_1 + (1 - \alpha)y_2 = \alpha Fx_1 + (1 - \alpha)Fx_2 = F(\alpha x_1 + (1 - \alpha)x_2)$, которое выполнено для любого аффинного преобразования F . ◁

Лемма 1.4. Пусть $\{X_i \subset \mathbb{R}^{n_i} : 1 \leq i \leq m\}$ — семейство выпуклых множеств. Тогда прямое произведение

$$X_1 \times X_2 \times \dots \times X_m \stackrel{\text{def}}{=} \{(x_1, x_2, \dots, x_m) : x_i \in X_i, 1 \leq i \leq m\}$$

является выпуклым множеством в пространстве $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \dots \times \mathbb{R}^{n_m}$.

▷ Пусть $x_i, \tilde{x}_i \in X_i$, $1 \leq i \leq m$, и $\alpha \in [0, 1]$. Тогда

$$\begin{aligned} \alpha(x_1, x_2, \dots, x_m) + (1 - \alpha)(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_m) &= \\ &= (\alpha x_1 + (1 - \alpha)\tilde{x}_1, \dots, \alpha x_m + (1 - \alpha)\tilde{x}_m) \in X_1 \times X_2 \times \dots \times X_m, \end{aligned} \quad (1.1)$$

так как $\alpha x_i + (1 - \alpha)\tilde{x}_i \in X_i$, $1 \leq i \leq m$. ◁

Композиция операций, сохраняющих выпуклость, также, очевидно, сохраняет выпуклость. Следовательно, для произвольных выпуклых множеств $X_1, X_2, \dots, X_m \in \mathbb{R}^n$ и вещественных чисел $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{R}$ множество

$$\alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_m X_m \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^m \alpha_i x_i : x_i \in X_i, 1 \leq i \leq m \right\}$$

является выпуклым. Действительно, эту линейную комбинацию множеств можно представить как композицию прямого произведения и аффинного преобразования. Отметим, что линейную композицию вида $A + B$ называют суммой Минковского множеств $A, B \subset \mathbb{R}^n$.

Лемма 1.5. *Замыкание \overline{X} выпуклого множества $X \subset \mathbb{R}^n$ выпукло.*

▷ Выберем произвольные точки $a, b \in \overline{X}$ и число $\alpha \in [0, 1]$. Необходимо доказать, что $c \stackrel{\text{def}}{=} \alpha a + (1 - \alpha)b \in \overline{X}$. Существуют такие две последовательности точек $(a_k)_{k \in \mathbb{N}}$ и $(b_k)_{k \in \mathbb{N}} \subset X$, что $a_k \rightarrow a$ и $b_k \rightarrow b$. Тогда последовательность точек $(c_k)_{k \in \mathbb{N}} \subset X$, где $c_k = \alpha a_k + (1 - \alpha)b_k$, сходится к c , а значит, $c \in \overline{X}$. ◁

Опр. 1.4. *Множества $X, Y \subset \mathbb{R}^n$ называются отделимыми, если существуют ненулевой вектор c и число d , такие что $c^\top x \geq d \geq c^\top y$ для любых $x \in X$ и $y \in Y$. Если известно, что неравенства строгие $c^\top x > d > c^\top y$, то говорят, что множества X и Y строго отделимы. Гиперплоскость, заданная уравнением $c^\top x = d$, называется разделяющей гиперплоскостью.*

Отметим, что согласно определению, вектор c из уравнения гиперплоскости $c^\top x = d$ ненулевой.

Теорема 1.1. *Если непересекающиеся множества X и $Y \subset \mathbb{R}^n$ выпуклы, замкнуты и одно из них ограничено, то они строго отделимы.*

▷ Пусть X — ограниченное множество и $d_X \stackrel{\text{def}}{=} \sup_{x_1, x_2 \in X} \|x_1 - x_2\|$ — его диаметр. Докажем, что найдутся точки $x_0 \in X$ и $y_0 \in Y$, для которых $\|x_0 - y_0\| = \inf_{x \in X, y \in Y} \|x - y\|$. Действительно, выберем произвольные две точки $x_1 \in X$ и $y_1 \in Y$. Пусть $\tilde{Y} = Y \cap B_r(x_1)$, где $B_r(x_1)$ — шар радиуса $r = d_X + \|x_1 - y_1\|$ с центром в x_1 . Множества X и \tilde{Y} являются компактными, а функция f , действующая по правилу $f: (x, y) \in X \times \tilde{Y} \mapsto \|x - y\|$, — непрерывной. Так как декартово произведение компактных множеств компактно, то функция f достигает свое минимальное значение в некоторых точках $x_0 \in X$ и $y_0 \in \tilde{Y}$. Если $y \in Y \setminus \tilde{Y}$ и $x \in X$, то $\|x - y\| \geq \|x_1 - y\| - \|x_1 - x\| \geq d_X + \|x_1 - y_1\| - d_X = \|x_1 - y_1\|$, а значит, точки x_0, y_0 искомые.

Пусть $\Pi \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n: c^\top x = d\}$ — гиперплоскость, проходящая через середину отрезка $[x_0, y_0]$, перпендикулярно ему. Выберем $c = x_0 - y_0$ и $d = (\|x_0\|^2 - \|y_0\|^2)/2$. Докажем, что множества X и Y не пересекаются с указанной гиперплоскостью, а значит, лежат в разных открытых полупространствах относительно её. Предположим противное, а именно, что некоторая точка $y \in Y$ принадлежит плоскости Π . Треугольник с вершинами x_0, y и y_0 является равнобедренным с основанием $[x_0, y_0]$ и острым углом при вершине y , так как $\|x_0 - y_0\| \leq \|y - x_0\|$. По условию Y — выпуклое множество, а значит, $[y, y_0] \subset Y$. Пусть \tilde{y} — основание перпендикуляра, опущенного из вершины x_0 на сторону $[y, y_0]$. Тогда $\tilde{y} \in [y, y_0]$ и $\|x_0 - \tilde{y}\| < \|x_0 - y_0\|$. Последнее неравенство противоречит выбору точек $x_0 \in X$ и $y_0 \in Y$. Таким образом, $c^\top x > d > c^\top y$. ◁

Опр. 1.5. *Гиперплоскость $\Pi \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n: c^\top x = d\}$ называется опорной к множеству X в точке x_0 , если $x_0 \in \Pi \cap \overline{X}$ и для всех $x \in X$ одновременно выполняется одно из неравенств: $c^\top x \geq d$ или $c^\top x \leq d$.*

Напомним, что точка x называется граничной для множества X , если любая

её окрестность содержит как точки, принадлежащие данному множеству, так и не принадлежащие ему.

Лемма 1.6. *Выпуклое множество $X \subset \mathbb{R}^n$ в каждой граничной точке имеет опорную гиперплоскость.*

▷ Пусть x_0 — граничная точка множества X . Так как X — выпуклое множество, то x_0 — граничная точка замыкания \bar{X} . Следовательно, найдётся такая последовательность точек $(y_k)_{k \in \mathbb{N}} \subset \mathbb{R}^n \setminus \bar{X}$, что $y_k \rightarrow x_0$. Согласно, теореме 1.1 для множеств $\{y_k\}$ и \bar{X} (выпуклость \bar{X} следует из леммы 1.5) найдётся такая гиперплоскость, заданная уравнением $c_k^\top x = d_k$, что $c_k^\top x > d_k > c_k^\top y_k$, $x \in \bar{X}$. Без нарушения общности будем считать, что $\|c_k\| = 1$. Тогда, в силу построения, последовательность $(d_k)_{k \in \mathbb{N}}$ является ограниченной. Наконец, без нарушения общности будем считать, что $c_k \rightarrow c$ и $d_k \rightarrow d$. Тогда $c^\top x \geq d$, $x \in X$, и $c^\top x_0 = d$. ◁

Теорема 1.2. *Произвольное выпуклое множество $X \subset \mathbb{R}^n$ можно отделить от точки y , ему не принадлежащей.*

▷ Действительно, если $y \notin \bar{X}$, то доказательство следует из теоремы 1.1, иначе — из леммы 1.6. ◁

Теорема 1.3. *Множества X и $Y \subset \mathbb{R}^n$ отделимы тогда и только тогда, когда множество $X - Y$ и точка $\{0\}$ отделимы.*

▷ Пусть множества X и Y отделимы. Тогда существуют такие $c \in \mathbb{R}^n \setminus \{0\}$ и $d \in \mathbb{R}$, что $c^\top y \leq d \leq c^\top x$ для всех $x \in X$, $y \in Y$. Следовательно, $c^\top(x - y) \geq 0$, а значит, гиперплоскость $\Pi = \{x: c^\top x = 0\}$ отделяет множество $X - Y$ от нуля.

Предположим теперь, что множества $X - Y$ и $\{0\}$ отделимы. Тогда существуют такие $c \in \mathbb{R}^n \setminus \{0\}$ и $d \in \mathbb{R}$, что $0 \leq d \leq c^\top z$ для всех $z \in X - Y$. Следовательно, $c^\top y \leq c^\top x$ для всех $x \in X$, $y \in Y$, а значит, $\sup_{y \in Y} c^\top y \leq \inf_{x \in X} c^\top x$. Выберем такое

число $\tilde{d} \in \mathbb{R}$, что $\sup_{y \in Y} c^\top y \leq \tilde{d} \leq \inf_{x \in X} c^\top x$. Тогда гиперплоскость $\Pi = \{x: c^\top x = \tilde{d}\}$ отделяет множества X и Y . ◁

Следствие 1.1. *Пусть X, Y — непустые выпуклые непересекающиеся множества. Тогда X и Y отделимы.*

Через $\text{Int } X$ обозначим внутренность множества $X \subset \mathbb{R}^n$, т.е. множество всех внутренних точек X . Не сложно видеть, что, если X — выпуклое множество, то $\text{Int } X$ также выпукло.

Следствие 1.2. *Пусть X, Y — выпуклые множества с непустой внутренностью, при этом $\text{Int } X \cap \text{Int } Y = \emptyset$. Тогда X и Y отделимы.*

Упражнения

1. Пусть $X \subset \mathbb{R}^n$ — непустое множество. Докажите, что любую точку p , принадлежащую выпуклой оболочке множества X , можно представить в виде выпуклой линейной комбинации не более чем $n + 1$ точек множества X .
2. Пусть в \mathbb{R}^n заданы точки p_1, p_2, \dots, p_s , где $s \geq n + 2$. Докажите, что точки можно разбить на два непересекающихся множества так, что выпуклые оболочки этих двух множеств будут иметь непустое пересечение.

3. (Теорема Хелли) Пусть I — произвольное семейство индексов и $\{X_i\}_{i \in I}$ — семейство замкнутых выпуклых множеств в \mathbb{R}^n , из которых хотя бы одно компактно. Докажите, что если любое подсемейство из $n + 1$ множеств имеет непустое пересечение, то и всё семейство имеет непустое пересечение.

§2 ВЫПУКЛЫЕ ФУНКЦИИ

Опр. 2.1. Функция $f: X \rightarrow \mathbb{R}$, 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 $f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$. Если последнее неравенство строгое при $\alpha \in (0, 1)$, то функция f называется строго выпуклой.

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

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

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

Лемма 2.2 (Неравенство Йенсена). Пусть $f: X \rightarrow \mathbb{R}$ — выпуклая функция. Тогда для произвольных точек $x_1, x_2, \dots, x_m \in X$ и чисел $\alpha_1, \alpha_2, \dots, \alpha_m \geq 0$, таких что $\sum_{i=1}^m \alpha_i = 1$, справедливо неравенство

$$f(\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m) \leq \alpha_1 f(x_1) + \alpha_2 f(x_2) + \dots + \alpha_m f(x_m). \quad (2.1)$$

▷ 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 $m \geq 2$. Пусть $x_1, x_2, \dots, x_{m+1} \in X$ и числа $\alpha_1, \alpha_2, \dots, \alpha_{m+1} \geq 0$, такие что $\sum_{i=1}^{m+1} \alpha_i = 1$. Без нарушения общности будем считать, что

$\alpha_1 < 1$. Так как $\sum_{i=2}^{m+1} \frac{\alpha_i}{1 - \alpha_1} = 1$, то верна цепочка неравенств

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

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 $f: X \rightarrow \mathbb{R}$ непрерывна во всех внутренних точках множества $X \subset \mathbb{R}^n$.*

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

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

Следовательно, $\overline{\lim}_{t \rightarrow +0} f(x_0 \pm t e_i) \leq f(x_0)$, поэтому, $\overline{\lim}_{x \rightarrow x_0} f(x) \leq f(x_0)$. Так как $f(x_0) \leq \frac{1}{2}f(x_0 + h) + \frac{1}{2}f(x_0 - h)$ для любого $h \in B_r(0)$, то

$$f(x_0) \leq \frac{1}{2} \lim_{x \rightarrow x_0} f(x) + \frac{1}{2} \overline{\lim}_{x \rightarrow x_0} f(x).$$

Следовательно, $f(x_0) = \lim_{x \rightarrow x_0} f(x) = \overline{\lim}_{x \rightarrow x_0} f(x)$, т.е. функция f непрерывна в x_0 . ◁

Опр. 2.2. *Вектор $c \in \mathbb{R}^n$ называется субградиентом функции $f: X \rightarrow \mathbb{R}$ в точке $x_0 \in X \subset \mathbb{R}^n$, если $f(x) \geq f(x_0) + c^\top(x - x_0)$ 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 $f: X \rightarrow \mathbb{R}$ — выпуклая функция, а x_0 — внутренняя точка множества X . Тогда множество $\partial f(x_0)$ непусто.*

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

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

Наконец, отнимая последнее равенство от неравенства, получаем, что

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

Имеет место следующее обобщение неравенства Йенсена.

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

▷ Так как $f(x) - f(y) \geq c_y^\top(x - y)$, где $c_y \in \partial f(y)$, то $f(x) = \max_{y \in \mathbb{R}^n} (c_y^\top x + d_y)$. Следовательно,

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

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

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

$$f(x) \geq f(x_0) + \nabla f(x_0)^\top(x - x_0). \quad \triangleleft$$

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.

▷ Выберем произвольные точку $x_0 \in X$ и направление $\ell \neq 0$. Рассмотрим функцию $g(t) = f(x_0 + t\ell)$, заданную на интервале $T \stackrel{\text{def}}{=} \{t: x_0 + t\ell \in X\}$. Очевидно, что функция f (строго) выпукла тогда и только тогда, когда (строго) выпуклы скалярные функции g при всевозможных $x_0 \in X$ и $\ell \in \mathbb{R}^n \setminus \{0\}$. Имеем $g'(t) = \ell^\top \nabla f(x_0 + t\ell)$ и $g''(t) = \ell^\top \nabla^2 f(x_0 + t\ell)\ell$. Функция g выпукла тогда и только тогда, когда $g''(t) \geq 0$, что равносильно неотрицательности 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 $x \leq y$ для векторов $x, y \in \mathbb{R}^n$, если выполнены неравенства $x_i \leq y_i$, $1 \leq i \leq n$.

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

a) $g(x) = \sum_{i=1}^m c_i f_i(x)$, где $c_i \geq 0$, $1 \leq i \leq m$;

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

c) $g(x) = \max_{1 \leq i \leq m} f_i(x)$;

d) $g(x) = h(f_1(x), f_2(x), \dots, f_m(x))$, где h — выпуклая монотонно неубывающая функция, т.е. $h(y) \leq h(\tilde{y})$ для всех y и \tilde{y} , таких что $y \leq \tilde{y}$.

▷ 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 , $1 \leq i \leq m$, выпуклы по условию, то $f_i(\alpha x + (1 - \alpha)y) \leq \alpha f_i(x) + (1 - \alpha)f_i(y)$. Положим $u = (f_1(x), \dots, f_m(x))$ и $v = (f_1(y), \dots, f_m(y))$, тогда в силу монотонности функции h верно неравенство

$$h(f_1(\alpha x + (1 - \alpha)y), \dots, f_m(\alpha x + (1 - \alpha)y)) \leq h(\alpha u + (1 - \alpha)v).$$

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

$$g(\alpha x + (1 - \alpha)y) = h\left(f_1(\alpha x + (1 - \alpha)y), \dots, f_m(\alpha x + (1 - \alpha)y)\right),$$

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

Упражнения

4. Докажите, что непрерывная выпуклая функция $f: [a, b] \rightarrow \mathbb{R}$ удовлетворяет неравенству

$$f\left(\frac{a+b}{2}\right)(b-a) \leq \int_a^b f(x) dx \leq \frac{f(a)+f(b)}{2}(b-a).$$

5. Докажите, что субдифференциал $\partial f(x_0)$ произвольной выпуклой функции f в точке x_0 является замкнутым выпуклым множеством.

6. Пусть $f(x) \stackrel{\text{def}}{=} \max_{1 \leq i \leq m} f_i(x)$, где $f_i(x)$ — выпуклые функции, и пусть c_i — субградиент

функции f_i в точке x_0 . Докажите, что вектор $c \stackrel{\text{def}}{=} \sum_{i=1}^m \alpha_i c_i$, $\sum_{i=1}^m \alpha_i = 1$, 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 $\mathbf{a} = (a_1, a_2, \dots, a_n)$ и $\mathbf{b} = (b_1, b_2, \dots, b_n)$. Говорят, что набор \mathbf{a} *мажорирует* набор \mathbf{b} , и пишут $\mathbf{a} \succ \mathbf{b}$, если $a_1 \geq b_1$, $a_1 + a_2 \geq b_1 + b_2$, \dots , $a_1 + a_2 + \dots + a_{n-1} \geq b_1 + b_2 + \dots + b_{n-1}$, $a_1 + a_2 + \dots + a_n = b_1 + b_2 + \dots + b_n$. Докажите, что для любой выпуклой функции $y = f(x)$, определённой на некотором промежутке I , и любых двух наборов $\mathbf{a} = (a_1, a_2, \dots, a_n)$, $\mathbf{b} = (b_1, b_2, \dots, b_n)$ из этого промежутка, удовлетворяющих условию $\mathbf{a} \succ \mathbf{b}$, справедливо неравенство

$$f(a_1) + f(a_2) + \dots + f(a_n) \geq f(b_1) + f(b_2) + \dots + f(b_n).$$

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

$$f(x) \geq h(x) \geq g(x) \quad \text{для каждого } x \in X.$$