Выпуклый анализ (лекции)

лектор — А.А. Васильева

1 Выпуклые и аффинные множества

Определение 1. *Пусть X — линейное пространство. Множество $M \subset X$ называется выпуклым, если для любых x, $y \in M$ и для любого $\lambda \in [0, 1]$ выполнено $(1-\lambda)x + \lambda y \in M$.*

Например, отрезок $[x, y] = \{(1-\lambda)x + \lambda y : \ \lambda \in [0, 1]\}$ является выпуклым.

Определение 2. *Пусть X — линейное пространство, $x_1, \ldots, x_n \in X$, $\lambda_1, \ldots, \lambda_n \geq 0$, $\sum\limits_{j=1}^{n} \lambda_j = 1$. Тогда сумма $\sum\limits_{j=1}^{n} \lambda_j x_j$ называется выпуклой комбинацией точек $x_1, \ldots, x_n$.*

Предложение 1. *Множество M является выпуклым тогда и только тогда, когда любая выпуклая комбинация точек $x_1, \ldots, x_n \in M$ принадлежит M.*

Доказательство. В одну сторону утверждение тривиально (в определении выпуклого множества участвуют выпуклые комбинации двух точек).

Докажем с другую сторону. Пусть множество M выпукло, $x_1, \ldots, x_n \in M$. Докажем, что выпуклая комбинация этих точек принадлежит M. Проведем индукцию по n. При $n = 1$ и $n = 2$ утверждение тривиально. Пусть оно верно для $n \leq m$. Докажем для $n = m + 1$.

Пусть $x_1, \ldots, x_{m+1} \in M$, $\lambda_1, \ldots, \lambda_{m+1} \geq 0$, $\sum\limits_{j=1}^{m+1} \lambda_j = 1$. Можно считать, что все $\lambda_j > 0$, т.к. иначе получаем выпуклую комбинацию меньшего количества точек. Имеем

$$\sum_{j=1}^{m+1} \lambda_j x_j = \lambda_1 x_1 + (1-\lambda_1) \sum_{j=2}^{m+1} \frac{\lambda_j}{\sum\limits_{i=2}^{m+1} \lambda_i} x_j \in M,$$

так как $\sum\limits_{j=2}^{m+1} \frac{\lambda_j}{\sum\limits_{i=2}^{m+1} \lambda_i} x_j$ является выпуклой комбинацией m точек множества M и по предположению индукции принадлежит M. $\square$

Определение 3. *Выпуклой оболочкой множества M называется множество всех выпуклых комбинаций элементов M, т.е.*

$$\operatorname{conv} M = \left\{ \sum_{j=1}^{n} \lambda_j x_j : \ n \in \mathbb{N}, \ x_j \in M, \ \lambda_j \geq 0, \ 1 \leq j \leq n, \ \sum_{j=1}^{n} \lambda_j = 1 \right\}.$$

Предложение 2. *Выпуклая оболочка множества M является наименьшим по включению выпуклым множеством, содержащим M, т.е. 1) $\operatorname{conv} M$ выпукло и $M \subset \operatorname{conv} M$, 2) если A — выпуклое множество, $A \supset M$, то $A \supset \operatorname{conv} M$.*

Доказательство. Ясно, что $M \subset \operatorname{conv} M$. Докажем, что $\operatorname{conv} M$ выпукло. В самом деле, пусть $x_1, \ldots, x_n \in M$, $y_1, \ldots, y_m \in M$, $\lambda_1, \ldots, \lambda_n \geq 0$, $\mu_1, \ldots, \mu_m \geq 0$, $\sum\limits_{j=1}^{n} \lambda_j = 1$, $\sum\limits_{j=1}^{m} \mu_j = 1$, $x = \sum\limits_{j=1}^{n} \lambda_j x_j$, $y = \sum\limits_{j=1}^{m} \mu_j y_j$. Тогда для любого $t \in [0, 1]$ точка $(1-t)x + ty$ является выпуклой комбинацией элементов множества M и поэтому принадлежит $\operatorname{conv} M$.

Пусть $A \supset M$ — выпуклое множество. Если $x_1, \ldots, x_n \in M$, то эти точки принадлежат A и, значит, любая их выпуклая комбинация принадлежит A. $\square$

Задача 1. 1) Показать, что $\operatorname{conv}(A \cup \operatorname{conv}(B \cup C)) = \operatorname{conv}(A \cup B \cup C)$.
2) Пусть $M_1, \ldots, M_N \subset X$ — выпуклые множества. Показать, что

$$\operatorname{conv}(M_1 \cup \cdots \cup M_N) = \left\{ \sum_{j=1}^{N} \lambda_j x_j : \ x_j \in M_j, \ \lambda_j \geq 0, \ \sum_{j=1}^{N} \lambda_j = 1 \right\}.$$

3) Пусть X — нормированное пространство, $M \subset X$ — выпуклое множество. Показать, что его внутренность и замыкание также выпуклы.

Определение 4. *Пусть X — линейное пространство. Множество $A \subset X$ называется аффинным подпространством, если для любых $x, y \in A$, $\lambda \in \mathbb{R}$ выполнено $(1-\lambda)x + \lambda y \in A$.*

Определение 5. *Пусть X — линейное пространство, $x_1, \ldots, x_n \in X$, $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, $\sum\limits_{j=1}^{n} \lambda_j = 1$. Тогда сумма $\sum\limits_{j=1}^{n} \lambda_j x_j$ называется аффинной комбинацией точек $x_1, \ldots, x_n$.*

Следующее утверждение доказывается так же, как для выпуклого множества.

Предложение 3. *Множество A является аффинным подпространством тогда и только тогда, когда любая аффинная комбинация точек $x_1, \ldots, x_n \in A$ принадлежит A.*

Определение 6. *Аффинной оболочкой множества M называется множество всех аффинных комбинаций элементов M, т.е.*

$$\operatorname{aff} M = \left\{ \sum_{j=1}^{n} \lambda_j x_j : \ n \in \mathbb{N}, \ x_j \in M, \ \lambda_j \in \mathbb{R}, \ 1 \le j \le n, \ \sum_{j=1}^{n} \lambda_j = 1 \right\}.$$

Следующее утверждение доказывается так же, как для выпуклой оболочки.

Предложение 4. *Аффинная оболочка множества M является наименьшим по включению аффинным подпространством, содержащим M, т.е. 1) $\operatorname{aff} M$ является аффинным подпространством и $M \subset \operatorname{aff} M$, 2) если A — аффинное подпространство, $A \supset M$, то $A \supset \operatorname{aff} M$.*

Предложение 5. *Множество $L \subset X$ является аффинным подпространством тогда и только тогда, когда $L = x_0 + Z$, где $x_0 \in L$, Z — линейное подпространство в X.*

Доказательство. Пусть $L = x_0 + Z$, где Z — линейное подпространство. Пусть $x = x_0 + \xi$, $y = x_0 + \eta$, где $\xi, \eta \in Z$, $\lambda \in \mathbb{R}$. Тогда

$$(1-\lambda)x + \lambda y = (1-\lambda)x_0 + (1-\lambda)\xi + \lambda x_0 + \lambda\eta = x_0 + (1-\lambda)\xi + \lambda\eta \in x_0 + Z.$$

Значит, L — аффинное подпространство.

Наоборот, пусть L — аффинное подпространство, $x_0 \in L$, $Z = -x_0 + L$. Покажем, что Z — линейное подпространство.

1) Пусть $\xi, \eta \in Z$, т.е. $\xi = x - x_0$, $\eta = y - x_0$, где $x, y \in L$. Тогда $\xi + \eta = x - x_0 + y - x_0$. Но точка $x - x_0 + y$ является аффинной комбинацией трех точек из L и поэтому сама принадлежит L. Значит, $\xi + \eta \in Z$.

2) Пусть $Z \ni \xi = x - x_0$, $x \in L$, $\alpha \in \mathbb{R}$. Тогда $\alpha\xi = \alpha x - \alpha x_0 = \alpha x + (1-\alpha)x_0 - x_0$; точка $\alpha x + (1-\alpha)x_0$ является аффинной комбинацией двух точек из L и поэтому сама принадлежит L. Значит, $\alpha\xi \in Z$. $\square$

Замечание. Если Z — линейное подпространство, $L = x_0 + Z$, $y_0 \in L$, то $L = y_0 + Z$.

В самом деле, $y_0 - x_0 \in Z$, поэтому $y_0 + Z \subset x_0 + Z$ и наоборот.

Таким образом, в представлении аффинного подпространства L в виде сдвига линейного подпространства можно брать любую точку из L; линейное подпространство получится одно и то же.

Определение 7. *Точки $x_1, \ldots, x_n$ называются аффинно независимыми, если из условия $\sum\limits_{j=1}^{n} \lambda_j x_j = 0$, $\sum\limits_{j=1}^{n} \lambda_j = 0$ следует, что $\lambda_j = 0$, $1 \le j \le n$. В обратном случае точки называются аффинно зависимыми.*

Предложение 6. *Точки $x_0, x_1, \ldots, x_n$ являются аффинно независимыми тогда и только тогда, когда векторы $x_1 - x_0, \ldots, x_n - x_0$ являются линейно независимыми.*

Доказательство. Пусть точки $x_0, x_1, \ldots, x_n$ аффинно независимы. Предположим, что $\sum\limits_{j=1}^{n} \alpha_j (x_j - x_0) = 0$ для некоторых $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Тогда

$$\left(-\sum_{j=1}^{n} \alpha_j\right) x_0 + \sum_{j=1}^{n} \alpha_j x_j = 0.$$

Из аффинной независимости точек x_j следует, что $\alpha_j = 0$, $1 \le j \le n$.

Обратно, пусть векторы $x_1 - x_0, \ldots, x_n - x_0$ линейно независимы. Пусть $\sum\limits_{j=0}^{n} \lambda_j x_j = 0$, $\sum\limits_{j=0}^{n} \lambda_j = 0$. Выразив λ_0, получим

$$\left(-\sum_{j=1}^{n} \lambda_j\right) x_0 + \sum_{j=1}^{n} \lambda_j x_j = 0,$$

или $\sum\limits_{j=1}^{n} \lambda_j (x_j - x_0) = 0$. Из линейной независимости получаем, что $\lambda_j = 0$, $1 \le j \le n$; значит, и $\lambda_0 = 0$. $\square$

Определение 8. *Пусть L — аффинное подпространство. Его размерностью называется размерность линейного пространства Z такого, что $L = x_0 + Z$, $x_0 \in L$.*

Пусть M — выпуклое множество. Его размерностью называется размерность аффинной оболочки M.

Предложение 7. *Пусть L — аффинное подпространство размерности n. Тогда максимальное количество аффинно независимых точек в L равно $n+1$.*

Доказательство. Имеем $L = x_0 + Z$, где $x_0 \in L$, Z — линейное подпространство. Пусть $\xi_1, \ldots, \xi_n$ — базис в Z. Тогда точки x_0, $x_0+\xi_1, \ldots, x_0+\xi_n$ аффинно независимы. Значит, L содержит $n+1$ аффинно независимых точек.

Пусть $y_0, y_1, \ldots, y_m \in L$ — аффинно независимые точки. Тогда $L = y_0 + Z$, при этом $y_1 - y_0, \ldots, y_m - y_0$ линейно независимы. Значит, $m \leq n$. $\square$

Предложение 8. *Пусть M — выпуклое множество размерности n. Тогда M содержит $n+1$ аффинно независимых точек.*

Доказательство. Пусть k — максимальное число аффинно независимых точек в M. Рассмотрим $x_1, \ldots, x_k \in M$ — аффинно независимые точки. Положим $L = \operatorname{aff}\{x_1, \ldots, x_k\}$. Если $L \not\supset M$, то существует точка $x_{k+1} \in M \backslash L$. Тогда $x_1, \ldots, x_k, x_{k+1}$ аффинно независимые (так как $x_2 - x_1, \ldots, x_k - x_1, x_{k+1} - x_1$ линейно независимые) и k не максимальное. Значит, $L \supset M$. Поэтому $L = \operatorname{aff} M$. Значит, $\dim L = n$, $k = n+1$. $\square$

Теорема 1. (Каратеодори). *Пусть $A \subset X$, $\dim \operatorname{conv} A = d$. Тогда*

$$\operatorname{conv} A = \left\{ \sum_{j=1}^{d+1} \lambda_j x_j : \ x_j \in A, \ \lambda_j \geq 0, \ 1 \leq j \leq d+1, \ \sum_{j=1}^{d+1} \lambda_j = 1 \right\}.$$

Доказательство. Пусть $x \in \operatorname{conv} A$, $n \in \mathbb{N}$ — наименьшее число такое, что $x = \sum\limits_{j=1}^{n} \lambda_j x_j$, где $x_j \in A$, $\lambda_j \geq 0$, $1 \leq j \leq n$, $\sum\limits_{j=1}^{n} \lambda_j = 1$. Покажем, что $n \leq d+1$.

Пусть $n > d+1$. Обозначим $L = \operatorname{aff} A$. Заметим, что $L = \operatorname{aff} \operatorname{conv} A$. Значит, $\dim L = d$ и поэтому $x_1, \ldots, x_n$ аффинно зависимы. Поэтому

существуют числа $\mu_1, \ldots, \mu_n \in \mathbb{R}$, одновременнно не равные 0, такие, что $\sum\limits_{j=1}^{n} \mu_j = 0$ и $\sum\limits_{j=1}^{n} \mu_j x_j = 0$.

Для любого $t \in \mathbb{R}$ имеем

$$x = \sum_{j=1}^{n} (\lambda_j - t\mu_j) x_j, \quad \sum_{j=1}^{n} (\lambda_j - t\mu_j) = 1.$$

Подберем t так, чтобы $\lambda_j - t\mu_j \geq 0$ для всех j и $\lambda_{j_0} - t\mu_{j_0} = 0$ для некоторого j_0. Тогда x будет представлена в виде выпуклой комбинации $< n$ точек. Получим противоречие с тем, что n минимальное.

Пусть $I = \{j \in \overline{1, n} : \ \mu_j > 0\}$. Это множество непусто, так как $\sum\limits_{j=1}^{n} \mu_j = 0$. Также заметим, что все λ_j строго положительны. Положим $t = \min_{j \in I} \frac{\lambda_j}{\mu_j}$. Пусть минимум достигается на индексе j_0. Тогда $\lambda_{j_0} - t\mu_{j_0} = 0$; если $j \in I$, то $\lambda_j - t\mu_j \geq 0$ по определению t. Если $j \notin I$, то $\lambda_j - t\mu_j \geq 0$, так как $\mu_j \leq 0$. $\square$

Замечание. Было доказано, что каждая точка из $\operatorname{conv} A$ представима в виде выпуклой комбинации аффинно независимых точек.

Следствие 1. *Пусть $K \subset \mathbb{R}^n$ — компакт. Тогда $\operatorname{conv} K$ компакт.*

Доказательство. Достаточно проверить ограниченность и замкнутость $\operatorname{conv} K$. Так как K ограничено, то и $\operatorname{conv} K$ ограничено.

Проверим замкнутость. Пусть $\{x_m\}_{m \in \mathbb{N}} \subset \operatorname{conv} K$, $x_m \underset{m \to \infty}{\to} x$. Покажем, что $x \in \operatorname{conv} K$. По теореме Каратеодори,

$$x_m = \sum_{j=1}^{n+1} \lambda_{m,j} x_{m,j}, \quad x_{m,j} \in K, \ \lambda_{m,j} \geq 0, \ \sum_{j=1}^{n+1} \lambda_{m,j} = 1.$$

Последовательности $\{x_{m,j}\}_{m \in \mathbb{N}} \subset \mathbb{R}^n$ и $\{\lambda_{m,j}\}_{m \in \mathbb{N}} \subset \mathbb{R}$ ограничены, поэтому из них можно выбрать сходящиеся подпоследовательности. Значит, существует последовательность $\{m_k\}_{k \in \mathbb{N}}$ такая, что $\lambda_{m_k,j} \underset{k \to \infty}{\to} \lambda_j$, $x_{m_k,j} \underset{k \to \infty}{\to} \xi_j$, $1 \leq j \leq n+1$. Отметим, что $\xi_j \in K$, $\lambda_j \geq 0$, $1 \leq j \leq n+1$, $\sum\limits_{j=1}^{n+1} \lambda_j = 1$. Значит,

$$x_{m_k} = \sum_{j=1}^{n+1} \lambda_{m_k,j} x_{m_k,j} \underset{k \to \infty}{\to} \sum_{j=1}^{n+1} \lambda_j \xi_j \in \operatorname{conv} K,$$

так что $x \in \operatorname{conv} K$. $\square$

Задача 2. Привести пример замкнутого множества в $\mathbb{R}^2$ такого, что $\operatorname{conv} A$ незамкнуто.

Задача 3. Привести пример компактного множества в l_2 такого, что $\operatorname{conv} A$ некомпактно.

2 Теоремы отделимости

2.1 Сумма множеств

Пусть X — линейное пространство, A, $B \subset X$. Суммой множеств A и B называется множество

$$A + B = \{a + b : \ a \in A, \, b \in B\}.$$

Если A или B пусто, то $A + B = \varnothing$.

Если X — нормированное пространство, A и B ограничены, то $A + B$ ограничено.

Если X — нормированное пространство, A или B открыто, то $A + B$ открыто.

Предложение 9. *Если A и B выпуклы, то $A + B$ выпукло.*

Доказательство. Пусть a_1, $a_2 \in A$, b_1, $b_2 \in B$, $\lambda \in [0, 1]$. Тогда

$$(1 - \lambda)(a_1 + b_1) + \lambda(a_2 + b_2) =$$

$$= [(1 - \lambda)a_1 + \lambda a_2] + [(1 - \lambda)b_1 + \lambda b_2] \in A + B.$$

$\square$

Предложение 10. *Пусть X — нормированное пространство, $A \subset X$ и $B \subset X$ компактны. Тогда $A + B$ компактно.*

Доказательство. Сначала докажем, что декартово произведение $A \times B$ компактно. Достаточно проверить, что из любой последовательности $\{(a_n, b_n)\}_{n \in \mathbb{N}} \subset A \times B$ можно выбрать сходящуюся подпоследовательность. В самом деле, выбираем сходящуюся подпоследовательность $\{a_{n_k}\}_{k \in \mathbb{N}}$, а затем $\{b_{n_{k_j}}\}_{j \in \mathbb{N}}$. Тогда последовательность $\{(a_{n_{k_j}}, b_{n_{k_j}})\}_{j \in \mathbb{N}}$ сходится к элементу $A \times B$.

Итак, $A \times B$ компактно. Рассмотрим отображение $f : A \times B \to X$, $f(a, b) = a + b$. Оно непрерывно, его образ — это $A + B$. Значит, $A + B$ компактно. $\square$

Задача 4. Привести пример двух выпуклых замкнутых множеств A, $B \subset \mathbb{R}^2$ таких, что $A + B$ незамкнуто.

2.2 Функционал Минковского

Пусть X — линейное пространство, $M \subset X$ — выпуклое множество, $0 \in M$. Предположим, что для любого $x \in X$ существует $\lambda > 0$ такое, что $\frac{x}{\lambda} \in M$. Функционалом Минковского множества M называется отображение $p_M : X \to \mathbb{R}_+$, заданное по формуле

$$p_M(x) = \inf \left\{ \lambda > 0 : \ \frac{x}{\lambda} \in M \right\}.$$

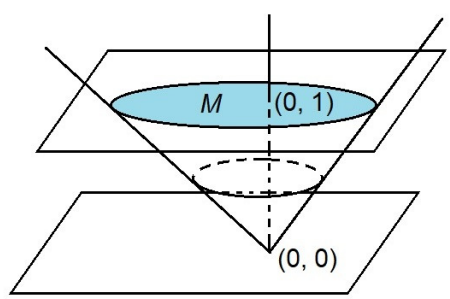

Рис. 1: Построение графика p_M.

Пример. Пусть $(X, \|\cdot\|)$ — нормированное пространство, M — единичный шар в X. Тогда

$$p_M(x) = \inf \left\{ \lambda > 0 : \ \left\| \frac{x}{\lambda} \right\| \le 1 \right\} = \|x\|.$$

Задача 5. Найти функционал Минковского для а) треугольника с вершинами $(1, 1)$, $(-1, 0)$, $(0, -1)$, б) множества $\{(x, y) : \ y > x^2 - 1\}$, в) множества $\{(x, y) : \ y > \sqrt{x^2 + 1} - 2\}$.

Отметим следующие свойства функционала Минковского:

- Пусть $M \subset N$. Тогда $p_M(x) \ge p_N(x)$.

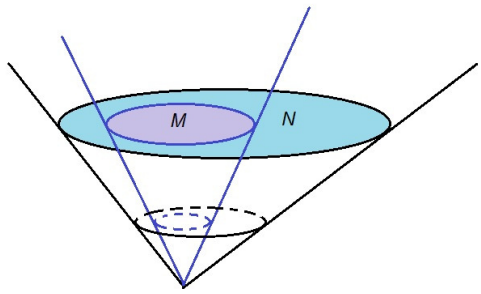

Рис. 2: Графики p_M и p_N.

- Пусть $t > 0$. Тогда $p_{tM}(x) = \frac{p_M(x)}{t}$.

Определение 9. *Функция $f : X \to \mathbb{R}$ называется выпуклой, если для любых x, $y \in X$ и для любого $\lambda \in [0, 1]$ выполнено $f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$.*

Определение 10. *Функция $f : X \to \mathbb{R}$ называется положительно-однородной, если для любых $x \in X$, $\lambda \geq 0$ выполнено $f(\lambda x) = \lambda f(x)$.*

Предложение 11. *Функционал Минковского выпуклый и положительно-однородный.*

Доказательство. Сначала докажем положительную однородность. Легко видеть, что $p_M(0) = 0$. Пусть $x \in X$, $\lambda > 0$. Тогда

$$p_M(\lambda x) = \inf\left\{t > 0 : \ \frac{\lambda x}{t} \in M\right\} = \inf\left\{\lambda s > 0 : \ \frac{x}{s} \in M\right\} = \lambda p_M(x).$$

Пусть x_1, $x_2 \in X$. Покажем, что $p_M(x_1 + x_2) \leq p_M(x_1) + p_M(x_2)$. В самом деле, пусть $t_1 > p_M(x_1)$, $t_2 > p_M(x_2)$. Тогда $\frac{x_1}{t_1} \in M$, $\frac{x_2}{t_2} \in M$. Значит,

$$\frac{x_1 + x_2}{t_1 + t_2} = \frac{t_1}{t_1 + t_2} \cdot \frac{x_1}{t_1} + \frac{t_2}{t_1 + t_2} \cdot \frac{x_2}{t_2} \in M.$$

Поэтому $p_M(x_1 + x_2) \leq t_1 + t_2$; остается сделать предельный переход при $t_1 \to p_M(x_1)$, $t_2 \to p_M(x_2)$.

Теперь легко проверить выпуклость p_M: если x_1, $x_2 \in X$, $\lambda \in [0, 1]$, то

$$p_M((1-\lambda)x_1 + \lambda x_2) \leq p_M((1-\lambda)x_1) + p_M(\lambda x_2) = (1-\lambda)p_M(x_1) + \lambda p_M(x_2).$$

□

На функциональном анализе доказывалась теорема Хана–Банаха:

Теорема 2. *Пусть X — линейное пространство, $L \subset X$ — подпространство, $p : X \to \mathbb{R}$ — выпуклый положительно-однородный функционал, $f_0 : L \to \mathbb{R}$ — линейный функционал такой, что $f_0(x) \leq p(x)$ для любого $x \in L$. Тогда существует линейный функционал $f : X \to \mathbb{R}$ такой, что $f|_L = f_0$ и $f(x) \leq p(x)$ для любого $x \in X$.*

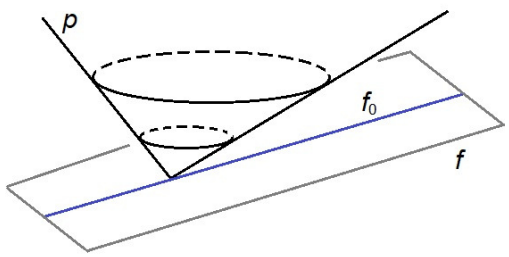

Рис. 3: Графики p, f_0 и f.

Эта теорема будет использоваться для доказательства теорем отделимости.

2.3 Теоремы отделимости

Напомним, что X^* — пространство линейных непрерывных функционалов на X.

Теорема 3. *Пусть X — нормированное пространство, A, $B \subset X$ — непустые выпуклые множества, $\operatorname{int} A \neq \varnothing$, $(\operatorname{int} A) \cap B = \varnothing$. Тогда существует функционал $x^* \in X^* \backslash \{0\}$ такой, что*

$$\sup_{x \in A} x^*(x) \leq \inf_{x \in B} x^*(x).$$

Геометрический смысл теоремы следующий. Пусть

$$\sup_{x \in A} x^*(x) \leq c \leq \inf_{x \in B} x^*(x).$$

Рассмотрим гиперплоскость

$$L = \{x \in X : \ x^*(x) = c\}.$$

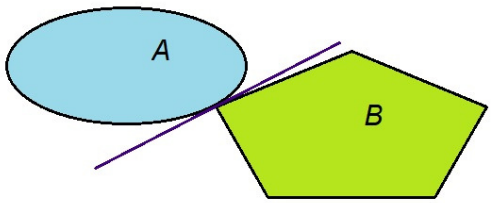

Рис. 4: Разделяющая гиперплоскость.

Она разбивает X на два полупространства: $\Pi_- = \{x \in X : \ x^*(x) \leq c\}$ и $\Pi_+ = \{x \in X : \ x^*(x) \geq c\}$. Тогда $A \subset \Pi_-$, $B \subset \Pi_+$, т.е. гиперплоскость L разделяет множества A и B.

Следующая теорема — о строгой отделимости.

Теорема 4. *Пусть X — нормированное пространство, A, $B \subset X$ — непустые выпуклые множества, A замкнуто, B компактно, $A \cap B = \varnothing$. Тогда существует функционал $x^* \in X^* \backslash \{0\}$ такой, что*

$$\sup_{x \in A} x^*(x) < \inf_{x \in B} x^*(x).$$

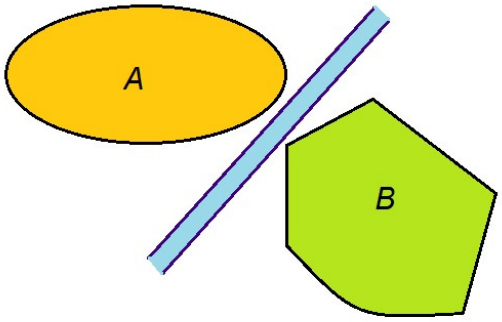

Рис. 5: Строгая отделимость множеств.

Если

$$\sup_{x \in A} x^*(x) < c < \inf_{x \in B} x^*(x),$$

то A содержится в полупространстве $\{x \in X : \ x^*(x) < c - \varepsilon\}$, а B — в полупространстве $\{x \in X : \ x^*(x) > c + \varepsilon\}$ для некоторого $\varepsilon > 0$, т.е. $\{x \in X : \ x^*(x) = c\}$ строго разделяет множества A и B.

Наконец, сформулируем конечномерную теорему отделимости.

Теорема 5. *Пусть A, $B \subset \mathbb{R}^n$ — непустые выпуклые множества, $A \cap B = \varnothing$. Тогда существует функционал $x^* \in (\mathbb{R}^n)^* \backslash \{0\}$ такой, что*

$$\sup_{x \in A} x^*(x) \leq \inf_{x \in B} x^*(x).$$

2.4 Доказательство теоремы 3

Сначала докажем теорему в частном случае.

Лемма 1. *Пусть $M \subset X$ — выпуклое множество, $0 \in \operatorname{int} M$, $x_0 \notin \operatorname{int} M$. Тогда существует функционал $x^* \in X^* \backslash \{0\}$ такой, что*

$$\sup_{x \in M} x^*(x) \leq x^*(x_0).$$

Доказательство. Так как M выпукло и $0 \in \operatorname{int} M$, то для него корректно определен функционал Минковского.

Покажем, что $p_M(x_0) \geq 1$. В самом деле, пусть $p_M(x_0) < 1$. Тогда существует $\kappa > 1$ такое, что $\hat{x} := \kappa x_0 \in M$. Существует $\delta > 0$ такое, что $B_\delta(0) \subset M$. Тогда $C = \operatorname{conv}(B_\delta(0) \cup \{\kappa x_0\}) \subset M$.

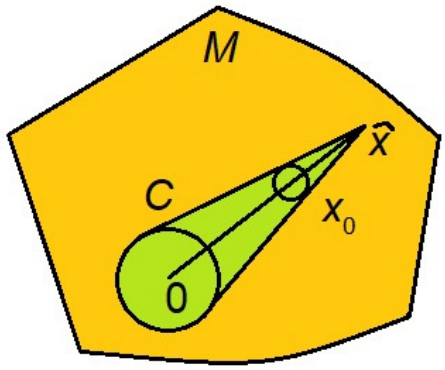

Рис. 6: Множество C.

Покажем, что $x_0 \in \operatorname{int} C$ (тогда получим противоречие с тем, что $x_0 \notin \operatorname{int} M$). В самом деле, пусть $\lambda = 1/\kappa$. Тогда $x_0 = (1-\lambda)0 + \lambda\hat{x}$ и

$$B_{(1-\lambda)\delta}(x_0) = (1-\lambda)B_\delta(0) + \lambda\hat{x} \subset C.$$

Определим функционал f_0 на $\operatorname{span}\{x_0\}$ по формуле $f_0(tx_0) = t$. Если $t < 0$, то $f_0(tx_0) < 0 \le p_M(tx_0)$, а если $t \ge 0$, то

$$f_0(tx_0) = t \le t \cdot p_M(x_0) = p_M(tx_0).$$

По теореме Хана–Банаха, существует линейный функционал f такой, что $f(x) \le p_M(x)$ для любого $x \in X$ и $f|_{\operatorname{span}\{x_0\}} = f_0$. Тогда 1) $f \ne 0$, 2) $f(x_0) = 1$, 3) для любого $x \in M$ выполнено $f(x) \le p_M(x) \le 1$. Из 2) и 3) следует, что $\sup_{x\in M} f(x) \le f(x_0)$.

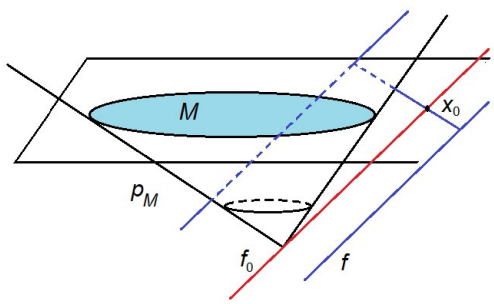

Рис. 7: Построение разделяющей гиперплоскости.

Покажем, что f непрерывен. В самом деле, для любого $x \in X$

$$|f(x)| = \max\{f(-x),\, f(x)\} \le \max\{p_M(-x),\, p_M(x)\} \le$$

$$\le \max\{p_{B_\delta(0)}(x),\, p_{B_\delta(0)}(-x)\} = \frac{\|x\|}{\delta}$$

(напомним, что $B_\delta(0) \subset M$, поэтому $p_M(x) \le p_{B_\delta(0)}(x)$, $x \in X$). □

Следствие 2. *Пусть $M \subset X$ — выпуклое множество с непустой внутренностью, $x_0 \notin \operatorname{int} M$. Тогда существует функционал $x^* \in X^*\backslash\{0\}$ такой, что $x^*(x_0) \ge \sup_{x\in M} x^*(x)$.*

Доказательство. Пусть $\hat{x} \in \operatorname{int} M$. Применим предыдущую лемму к множеству $M - \hat{x}$ и точке $x_0 - \hat{x}$ и получим требуемое утверждение. □

Предложение 12. *Пусть A — выпуклое множество,* $\operatorname{int} A \neq \varnothing$. *Тогда* $A \subset \overline{\operatorname{int} A}$.

Доказательство. Пусть $O_\delta(x_0) \subset A$, $x \in A$. Тогда

$$C = \operatorname{conv}(O_\delta(x_0) \cup \{x\}) \subset A,$$

при этом $(1-\lambda)x_0 + \lambda x \in \operatorname{int} A$ для $0 \leq \lambda < 1$. Значит, $x \in \overline{\operatorname{int} A}$. □

Доказательство теоремы 3. Нам нужно найти функционал $x^* \in X^* \backslash \{0\}$ такой, что для любых $x \in A$, $y \in B$ выполнено $x^*(x) \leq x^*(y)$. Это эквивалентно условиям:

1. $x^*(x) \leq x^*(y)$, $x \in \operatorname{int} A$, $y \in B$ (в силу предложения 12 и непрерывности x^*);
2. $x^*(x + (-y)) \leq 0$, $x \in \operatorname{int} A$, $y \in B$;
3. $x^*(z) \leq 0$ для любого $z \in (\operatorname{int} A) + (-B)$.

Множество $(\operatorname{int} A) + (-B)$ выпукло, открыто и не содержит 0 (т.к. $(\operatorname{int} A) \cap B = \varnothing$). Применяя предыдущее следствие, находим требуемый функционал x^*. □

2.5 Доказательство теоремы 4

Пусть $M \subset X$ — непустое множество, $x \in X$. Напомним, что расстоянием от x до M называется величина $\operatorname{dist}(x, M) = \inf_{y \in M} \|x - y\|$.

Предложение 13. *Функция* $x \mapsto \operatorname{dist}(x, M)$ *непрерывна.*

Доказательство. Пусть $x_1, x_2 \in X$, $\varepsilon > 0$, $y \in M$,

$$\|x_1 - y\| \leq \operatorname{dist}(x_1, M) + \varepsilon.$$

Тогда

$$\operatorname{dist}(x_2, M) \leq \|x_2 - y\| \leq \|x_1 - y\| + \|x_1 - x_2\| \leq$$
$$\leq \operatorname{dist}(x_1, M) + \varepsilon + \|x_1 - x_2\|.$$

Устремив ε к 0, получим $\operatorname{dist}(x_2, M) \leq \operatorname{dist}(x_1, M) + \|x_1 - x_2\|$. Поменяв x_1 и x_2, получим $\operatorname{dist}(x_1, M) \leq \operatorname{dist}(x_2, M) + \|x_1 - x_2\|$. □

Доказательство теоремы 4. Так как A замкнуто и $A \cap B = \varnothing$, то для любой точки $x \in B$ выполнено $\operatorname{dist}(x, A) > 0$. Из непрерывности функции расстояния и компактности B следует, что $\inf_{x \in B} \operatorname{dist}(x, A) > 0$. Следовательно, существует такое $\varepsilon > 0$, что $B + U_\varepsilon$ не пересекается с A, где U_ε — открытый шар с центром в 0 радиуса ε. Множество $B + U_\varepsilon$ открытое. Применяя теорему 3, получаем, что существует линейный непрерывный функционал $x^* \neq 0$ такой, что

$$\sup_{x \in A} x^*(x) \leq \inf_{x \in U_\varepsilon + B} x^*(x) = \inf_{y \in B,\, z \in U_\varepsilon} x^*(y + z) =$$

$$= \inf_{y \in B} x^*(y) + \inf_{z \in U_\varepsilon} x^*(z) = \inf_{y \in B} x^*(y) - \varepsilon \|x^*\|.$$

Отсюда следует утверждение теоремы. □

Пример. Пусть $X = \mathbb{R}^2$, $A = \{(x, y) : \ x > 0,\, y > 1/x\}$, $B = \{(x, y) : x < 0,\, y > -1/x\}$. Эти множества замкнуты и не пересекаются. Их можно разделить единственной прямой, а строго разделить нельзя.

Напомним, что замкнутое полупространство задается неравенством $\{x \in X : \ x^*(x) \leq c\}$, где $x^* \in X^* \backslash \{0\}$, $c \in \mathbb{R}$.

Следствие 3. *Множество $M \subset X$, $M \neq X$, выпукло и замкнуто тогда и только тогда, когда его можно представить в виде пересечения некоторого семейства замкнутых полупространств.*

Доказательство. Так как замкнутое полупространство — выпуклое замкнутое множество, то пересечение любого их семейства выпукло и замкнуто.

Обратно, пусть $M \subset X$ выпукло и замкнуто. Достаточно рассмотреть случай $M \neq \varnothing$. Тогда любую точку $x \notin M$ можно строго отделить от M, т.е. существуют $f_x \in X^* \backslash \{0\}$, $c_x \in \mathbb{R}$ такие, что $\sup_{y \in M} f_x(y) \leq c_x$, $f_x(x) > c_x$. Пусть

$$\Pi_x = \{y \in X : \ f_x(y) \leq c_x\}.$$

Тогда $M = \cap_{x \notin M} \Pi_x$. □

2.6 Доказательство конечномерной теоремы отделимости

Лемма 2. *Пусть $M \subset \mathbb{R}^n$ — выпуклое множество,* $\dim M = n$. *Тогда* $\operatorname{int} M \neq \varnothing$.

Доказательство. В силу предложения 8, M содержит $n+1$ аффинно независимую точку: $x_0, x_1, \ldots, x_n$. Покажем, что $\operatorname{int} M \neq \varnothing$. В самом деле, векторы $\xi_1 = x_1 - x_0, \ldots, \xi_n = x_n - x_0$ линейно независимы. Имеем

$$M \supset K := \operatorname{conv}\{x_0, \ldots, x_n\} = \left\{ \sum_{i=0}^{n} \lambda_i x_i : \ \lambda_i \geq 0, \ \sum_{i=0}^{n} \lambda_i = 1 \right\} =$$

$$= \left\{ x_0 + \sum_{i=1}^{n} \lambda_i \xi_i : \ \lambda_i \geq 0, \ \sum_{i=1}^{n} \lambda_i \leq 1 \right\}.$$

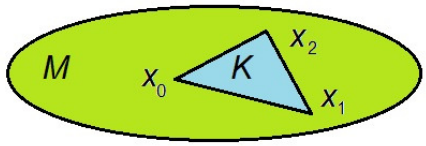

Рис. 8: Множество K.

Покажем, что если $\lambda_i > 0$, $1 \leq i \leq n$, $\sum\limits_{i=1}^{n} \lambda_i < 1$, то $x_0 + \sum\limits_{i=1}^{n} \lambda_i \xi_i \in \operatorname{int} M$.

В самом деле, все нормы на $\mathbb{R}^n$ эквивалентны; в частности, они эквивалентны норме $\left\| \sum\limits_{i=1}^{n} \mu_i \xi_i \right\| = \sum\limits_{i=1}^{n} |\mu_i|$. Остается заметить, что если $\lambda_i > 0$, $1 \leq i \leq n$, $\sum\limits_{i=1}^{n} \lambda_i < 1$, $\sum\limits_{i=1}^{n} |\mu_i - \lambda_i|$ достаточно мало, то $\mu_i > 0$, $1 \leq i \leq n$, $\sum\limits_{i=1}^{n} \mu_i < 1$, так что $x_0 + \sum\limits_{i=1}^{n} \mu_i \xi_i \in K$. $\square$

Доказательство теоремы 5. Так же, как при доказательстве теоремы 3, достаточно проверить, что если $M \subset \mathbb{R}^n$ — выпуклое множество, $\hat{x} \notin M$, то существует линейный непрерывный функционал $x^* \in (\mathbb{R}^n)^* \setminus \{0\}$ такой, что $\sup_{x \in M} x^*(x) \leq x^*(\hat{x})$.

Если $\dim M = n$, то M имеет непустую внутренность (по лемме 2), поэтому мы можем применить теорему 3.

Пусть $\dim M < n$. Положим $L = \operatorname{aff} M$. Если $\hat{x} \notin L$, то можно применить теорему 4. Если $\hat{x} \in L$, то берем гиперплоскость, содержащую L. Она разделяет L и $\hat{x}$ (они лежат в границе соответствующих полупространств). $\square$

Пример. Пусть $A = [(-1,\, 0),\, (1,\, 0)]$, $B = [(0,\, -1), (0,\, 1)]$. Эти множества выпуклы, $\operatorname{int} A$ не пересекается с B (т.к. пустая). Легко видеть, что A и B разделить нельзя.

Предложение 14. *Пусть $M \subset \mathbb{R}^n$ — выпуклое множество, x_0 — граничная точка множества M. Тогда существует функционал $x^* \in X^* \backslash \{0\}$ такой, что $x^*(x_0) \geq \sup_{x \in M} x^*(x)$.*

Доказательство. Это утверждение доказывается так же, как конечномерная теорема отделимости: если $\dim M = n$, то M имеет внутреннюю точку (см. лемму 2) и можно применить теорему 3, а если $\dim M < n$, то M и x_0 лежат в некоторой гиперплоскости. $\square$

Пример. Пусть X — бесконечномерное нормированное пространство, $f : X \to \mathbb{R}$ — линейный неограниченный функционал на X. Пусть $M = \{x \in X :\ f(x) < 0\}$. Тогда $0 \notin M$. Покажем, что M и 0 нельзя разделить замкнутой гиперплоскостью. В самом деле, пусть $g : X \to \mathbb{R}$ — линейный непрерывный функционал, $g \neq 0$, $\sup_{x \in M} g(x) \leq 0$. Тогда для любого x такого, что $f(x) = 0$, выполнено $g(x) \leq 0$. Если $f(x) = 0$, то $f(-x) = 0$, так что $g(x) = 0$. Таким образом, $\ker f \subset \ker g$, а это возможно только в том случае, если $g = \alpha f$, где $\alpha \in \mathbb{R}$. Но это невозможно, т.к. g непрерывный, а f разрывный.

Задача 6. Пусть $M \subset l_2$,

$$M = \left\{ (x_1,\, \ldots,\, x_n,\, \ldots) :\ \sum_{n=1}^{\infty} n^2 x_n^2 < 1 \right\}.$$

Найти точку $\hat{x} = (\hat{x}_1,\, \ldots,\, \hat{x}_n,\, \ldots)$ такую, что $\sum\limits_{n=1}^{\infty} n^2 \hat{x}_n^2 = 1$, при этом не существует линейного непрерывного функционала, разделяющего $\hat{x}$ и M.

Указание. 1) Напомним, что линейный непрерывный функционал на l_2 имеет вид $f(x_1,\, \ldots,\, x_n,\, \ldots) = \sum\limits_{n=1}^{\infty} \alpha_n x_n$, где $(\alpha_1,\, \ldots,\, \alpha_n,\, \ldots) \in l_2$. 2) Воспользоваться тем, что в неравенстве Коши–Буняковского равенство достигается только на сонаправленных векторах.

3 Крайние и выступающие точки

Пусть X — линейное нормированное пространство, $M \subset X$ — выпуклое множество.

Точка $x \in M$ называется крайней, если она не является серединой никакого невырожденного отрезка, содержащегося в M; то есть из условий $x = \frac{y+z}{2}$, y, $z \in M$, следует, что $y = z = x$.

Точка $x \in M$ называется выступающей, если существует функционал $x^* \in X^* \backslash \{0\}$ такой, что $x^*(x) = \sup_{y \in M} x^*(y)$ и из условия $x^*(y) = x^*(x)$ следует, что $y = x$.

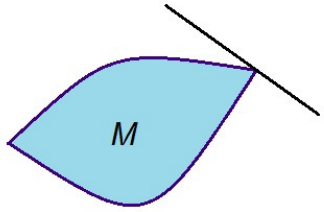

Рис. 9: Выступающая точка.

Пусть $x^* \in X^* \backslash \{0\}$, $x^*(x) = \sup_{y \in M} x^*(y)$. Гиперплоскость $\{y \in X : x^*(y) = x^*(x)\}$ называется опорной гиперплоскостью в точке x. Множество M лежит по одну сторону от этой гиперплоскости.

Таким образом, точка x выступающая, если найдется опорная к M гиперплоскость в точке x, пересекающаяся с M только в x.

Предложение 15. *Выступающая точка является крайней.*

Доказательство. Пусть $x \in M$ — выступающая точка, $x^* \in X^* \backslash \{0\}$, $\{\xi \in X : \ x^*(\xi) = x^*(x)\}$ — опорная гиперплоскость, пересекающаяся с M только в точке x. Пусть $x = \frac{y+z}{2}$, y, $z \in M$. Тогда $x^*(x) = \frac{x^*(y)+x^*(z)}{2}$, $x^*(y) \leq x^*(x)$, $x^*(z) \leq x^*(x)$. Значит, $x^*(y) = x^*(z) = x^*(x)$, поэтому $y = z = x$. $\square$

Пример. Пусть $f(x) = \max\{x^3, 0\}$, $M = \{(x, y) \in \mathbb{R}^2 : \ y \geq f(x)\}$. Тогда M — выпуклое множество. Точка $(0, 0)$ является крайней, но не выступающей.

Задача 7. Показать, что у единичного шара в пространстве c_0 нет крайних точек.

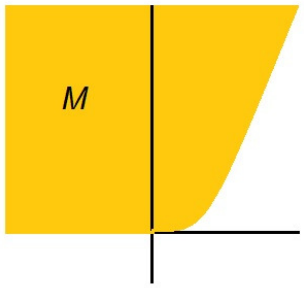

Рис. 10: (0, 0) — крайняя, но не выступающая точка.

Пусть X — нормированное пространство, $A \subset X$. Через $\overline{\operatorname{conv}} A$ будем обозначать замыкание выпуклой оболочки множества A: $\overline{\operatorname{conv}} A = \overline{\operatorname{conv} A}$.

Теорема 6. (Страшевич). *Пусть $M \subset \mathbb{R}^n$ — выпуклый компакт. Тогда $M = \overline{\operatorname{conv}} S$, где S — множество выступающих точек множества M.*

Доказательство. Можно считать, что норма на $\mathbb{R}^n$ евклидова.

Пусть $C = \overline{\operatorname{conv}} S$. Тогда $C \subset M$ — выпуклый компакт. Если $C \neq M$, то существует точка $x_0 \in M \backslash C$. По теореме о строгой отделимости, существует функционал $x^* \in (\mathbb{R}^n)^* \backslash \{0\}$ такой, что $x^*(x_0) > \sup_{x \in C} x^*(x) =: \alpha$. (Если $S = \varnothing$, то считаем, что супремум пустого множества равен $-\infty$ и в качестве x^* берем любой ненулевой функционал.)

Покажем, что M содержит выступающую точку, не принадлежащую C. Для этого мы сначала построим шар, содержащий множество C, у которого x_0 лежит на границе. Затем сделаем его раздутие относительно центра, содержащее M, так, чтобы на границе шара была точка множества M. Она будет выступающей, но не будет принадлежать C.

Пусть $\xi \in \mathbb{R}^n$ — вектор единичной длины, перпендикулярный гиперплоскости $\{x \in \mathbb{R}^n : \ x^*(x) = \alpha\}$, $x^*(\xi) < 0$. Пусть V_R — замкнутый шар радиуса R с центром в точке $x_0 + R\xi$. Покажем, что если R достаточно велико, то $C \subset V_R$.

В самом деле, иначе существует последовательность точек $x_m \in C \backslash V_m$. Из $\{x_m\}_{m \in \mathbb{N}}$ можно выбрать сходящуюся подпоследовательность $x_{m_k} \underset{k \to \infty}{\to} \overline{x}$. Так как $x^*(x_m) \leq \alpha$, то $x^*(\overline{x}) \leq \alpha < x^*(x_0)$. С другой стороны,

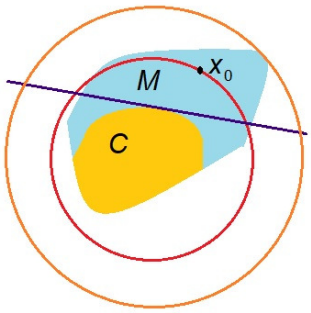

Рис. 11: Построение шаров.

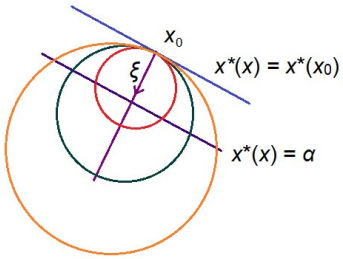

Рис. 12: Шары V_R.

$x_m \in \{x \in \mathbb{R}^n : \ x^*(x) \leq x^*(x_0)\} \backslash \operatorname{int} V_m$. Справа стоят вложенные замкнутые множества, поэтому

$$\overline{x} \in \{x \in \mathbb{R}^n : \ x^*(x) \leq x^*(x_0)\} \backslash \cup_{k \in \mathbb{N}} \operatorname{int} V_{m_k} =$$

$$= \{x \in \mathbb{R}^n : \ x^*(x) = x^*(x_0)\}.$$

Получаем противоречие.

Итак, пусть R таково, что $C \subset V_R$. Пусть

$$\tilde{R} = \inf\{r > 0 : \ B_r(x_0 + R\xi) \supset M\}.$$

Тогда $\tilde{R} \geq R$.

Пусть $\tilde{R} = R$. Тогда $M \subset V_R$.

Пусть L — опорная гиперплоскость к V_R в точке x_0. Так как шар евклидов, то $L \cap V_R = \{x_0\}$. Но L также будет опорной гиперплоскостью к M и $M \cap L = \{x_0\}$, поскольку $x_0 \in M \subset V_R$. Значит, x_0 — выступающая точка множества M. При этом $x_0 \notin C$.

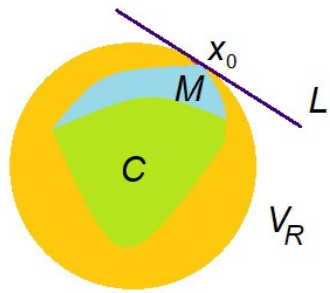

Рис. 13: x_0 — выступающая точка

Пусть $\tilde{R} > R$, B — замкнутый шар радиуса $\tilde{R}$ с центром в $x_0 + R\xi$. Так как M компактно, то существует точка $\hat{x} \in M$, $\|\hat{x} - x_0 - R\xi\| = \tilde{R}$. Так как $C \subset V_R$, то $\hat{x} \notin C$. Дальше аналогично первому случаю получаем, что $\hat{x}$ — выступающая точка множества M. $\square$

Следствие 4. (теорема Крейна–Мильмана). *Пусть $M \subset \mathbb{R}^n$ — выпуклый компакт. Тогда $M = \overline{\operatorname{conv}}\,\Sigma$, где Σ — множество крайних точек множества M.*

Теоремы Страшевича и Крейна–Мильмана обобщаются на выпуклые компакты в произвольных нормированных пространствах, есть и более широкие обобщения. В конечномерном случае теорему Крейна–Мильмана можно усилить: вместо замыкания выпуклой оболочки можно взять выпуклую оболочку.

Теорема 7. *Пусть $M \subset \mathbb{R}^n$ — выпуклый компакт. Тогда $M = \operatorname{conv}\Sigma$, где Σ — множество крайних точек множества M.*

Доказательство. Теорему докажем индукцией по n. Если $n = 1$, то M — точка или отрезок, в этом случае утверждение очевидно.

Пусть $n > 1$, $x \in M$. Проведем через x прямую, она пересекает M по отрезку $[\xi, \eta]$. Тогда $x = (1-\lambda)\xi + \lambda\eta$ для некоторого $\lambda \in [0, 1]$. Если мы докажем, что ξ и η являются выпуклыми комбинациями крайних точек множества M, то утверждение будет доказано.

Так как ξ и η — граничные точки множества M, то существуют опорные гиперплоскости к M в этих точках: L_ξ и L_η.

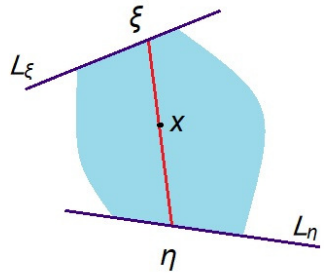

Рис. 14: Плоскости L_ξ и L_η

Пусть $M_\xi = M \cap L_\xi$, $M_\eta = M \cap L_\eta$. Это выпуклые компакты, лежащие в аффинном подпространстве размерности $n-1$. Поэтому можно применить предположение индукции (рассмотрев линейные подпространства $L_\xi - \xi$ и $L_\eta - \eta$ и лежащие в них выпуклые компакты $M_\xi - \xi$ и $M_\eta - \eta$). Получим, что ξ и η являются выпуклыми комбинациями крайних точек множеств M_ξ и M_η соответственно.

Покажем, что если ζ — крайняя точка M_ξ, то она является крайней точкой M (для M_η доказывается аналогично). В самом деле,

$$L_\xi = \{z \in \mathbb{R}^n : \ x^*(z) = x^*(\xi)\}, \ \ \sup_{z \in M} x^*(z) = x^*(\xi),$$

$x^* \in (\mathbb{R}^n)^* \backslash \{0\}$. Пусть $\zeta = \frac{\zeta_1 + \zeta_2}{2}$, ζ_1, $\zeta_2 \in M$. Тогда

$$x^*(\xi) = x^*(\zeta) = \frac{x^*(\zeta_1) + x^*(\zeta_2)}{2},$$

$x^*(\zeta_1) \le x^*(\xi)$, $x^*(\zeta_2) \le x^*(\xi)$. Значит, $x^*(\zeta_1) = x^*(\zeta_2) = x^*(\xi)$, т.е. ζ_1, $\zeta_2 \in M_\xi$. Так как ζ — крайняя точка M_ξ, то $\zeta_1 = \zeta_2 = \zeta$. $\square$

Задача 8. Пусть $M \subset \mathbb{R}^n$ — выпуклый компакт. Верно ли, что $M = \operatorname{conv} S$, где S — множество выступающих точек M?

4 Выпуклые функции

4.1 Определение и основные свойства

Пусть X — линейное пространство, $f : X \to \mathbb{R} \cup \{\pm\infty\}$. Надграфиком функции f называется множество

$$\operatorname{epi} f = \{(x, u) \in X \times \mathbb{R} : \ u \ge f(x)\}.$$

Определение 11. *Функция f называется выпуклой, если множество* $\operatorname{epi} f$ *выпукло.*

Замечание. Если f выпукла, то множества $\{x : \ f(x) < +\infty\}$, $\{x : f(x) = -\infty\}$, $\{x : \ f(x) \le c\}$ выпуклы.

Задача 9. Описать множество всех выпуклых функций $f : \mathbb{R}^2 \to \mathbb{R} \cup \{\pm\infty\}$ таких, что $\operatorname{int}\{(x, y) \in \mathbb{R}^2 : \ f(x, y) = -\infty\} = \{(x, y) : x^2 + y^2 < 1\}$.

Задача 10. Описать множество всех выпуклых функций $f : \mathbb{R}^2 \to \mathbb{R} \cup \{\pm\infty\}$ таких, что $\operatorname{int}\{(x, y) \in \mathbb{R}^2 : \ f(x, y) = -\infty\} = \{(x, y) : \max\{|x|, |y|\} < 1\}$.

Определение 12. *Функция $f : X \to \mathbb{R} \cup \{+\infty\}$ называется собственной, если* $\operatorname{dom} f := \{x : \ f(x) < +\infty\} \neq \varnothing$.

Замечание. В определение собственной функции входит, что она не принимает значения $-\infty$.

Предложение 16. *Пусть $f : X \to \mathbb{R} \cup \{+\infty\}$. Следующие условия эквивалентны:*

1. *f выпукла;*

2. *для любых $x, y \in X$, $\lambda \in [0, 1]$ выполнено*

$$f((1-\lambda)x + \lambda y) \le (1-\lambda)f(x) + \lambda f(y)$$

(полагаем $0 \cdot (+\infty) = 0$).

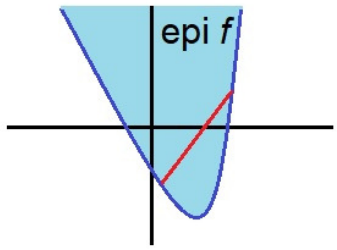

Рис. 15: Надграфик

Доказательство. $1 \Rightarrow 2$. Пусть f выпукла. Достаточно рассмотреть случай, когда $0 < \lambda < 1$, $f(x) < +\infty$, $f(y) < +\infty$. Так как $\operatorname{epi} f$ выпукло, то $[(x,\, f(x)), (y,\, f(y))] \subset \operatorname{epi} f$. Значит, $((1-\lambda)x+\lambda y,\, (1-\lambda)f(x)+\lambda f(y)) \in \operatorname{epi} f$. Это и означает, что $f((1-\lambda)x+\lambda y) \le (1-\lambda)f(x)+\lambda f(y)$.

$2 \Rightarrow 1$. Пусть $(x,\, u)$, $(y,\, v) \in \operatorname{epi} f$, $\lambda \in [0,\, 1]$. Тогда

$$(1-\lambda)u+\lambda v \ge (1-\lambda)f(x)+\lambda f(y) \ge f((1-\lambda)x+\lambda y),$$

так что $((1-\lambda)x+\lambda y,\, (1-\lambda)u+\lambda v) \in \operatorname{epi} f$. □

Предложение 17. *Пусть $f : X \to \mathbb{R} \cup \{+\infty\}$ выпукла. Тогда выполнено неравенство Йенсена:*

$$f\left(\sum_{j=1}^{n} \lambda_j x_j\right) \le \sum_{j=1}^{n} \lambda_j f(x_j)$$

для любых $x_1,\, \ldots,\, x_n \in X$, $\lambda_1 \ge 0,\, \ldots,\, \lambda_n \ge 0$ таких, что $\sum\limits_{j=1}^{n} \lambda_j = 1$.

Доказательство. Докажем утверждение индукцией по n. Для $n = 1,\, 2$ это тривиально. Пусть неравенство Йенсена верно для $n \le m$. Докажем его для $n = m+1$. Можно считать, что $\lambda_j > 0$, $1 \le j \le m+1$. Имеем

$$f\left(\sum_{j=1}^{m+1} \lambda_j x_j\right) = f\left(\lambda_1 x_1 + (1-\lambda_1)\sum_{j=2}^{m+1} \frac{\lambda_j}{\sum\limits_{i=2}^{m+1} \lambda_i} x_j\right) \le$$

$$\le \lambda_1 f(x_1) + (1-\lambda_1) f\left(\sum_{j=2}^{m+1} \frac{\lambda_j}{\sum\limits_{i=2}^{m+1} \lambda_i} x_j\right) \le$$

$$\le \lambda_1 f(x_1) + (1-\lambda_1)\sum_{j=2}^{m+1} \frac{\lambda_j}{\sum\limits_{i=2}^{m+1} \lambda_i} f(x_j) = \sum_{j=1}^{n} \lambda_j f(x_j)$$

(здесь использовалось эквивалентное определение выпуклой функции и предположение индукции). □

Рассмотрим некоторые примеры выпуклых функций.

1. Пусть $M \subset X$ — выпуклое множество, $0 \in M$, для любого $x \in X$ существует $\lambda > 0$ такое, что $x/\lambda \in M$. Тогда функционал Минковского p_M — выпуклая функция. В частности, норма является выпуклой функцией.

2. Пусть $A \subset X$. Индикаторная функция множества A определяется как
$$I_A(x) = \begin{cases} 0, & x \in A, \\ +\infty, & x \notin A. \end{cases}$$
Эта функция выпукла тогда и только тогда, когда A выпукло.

Задача 11. Показать, что функция $f(x_1, \ldots, x_n) = \ln(e^{x_1} + \cdots + e^{x_n})$ выпукла.

Указание. Использовать неравенство Гёльдера.

Предложение 18. *Пусть f, $g : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклые функции. Тогда $f + g$ выпукла.*

Доказательство. Пусть x, $y \in X$, $\lambda \in [0, 1]$. Тогда
$$f((1-\lambda)x + \lambda y) + g((1-\lambda)x + \lambda y) \leq$$
$$\leq (1-\lambda)f(x) + \lambda f(y) + (1-\lambda)g(x) + \lambda g(y) =$$
$$= (1-\lambda)(f+g)(x) + \lambda(f+g)(y).$$
□

Предложение 19. *Пусть $f_n : X \to \mathbb{R} \cup \{\pm\infty\}$ — выпуклые функции, $n \in \mathbb{N}$. Пусть для любого $x \in X$ существует предел $\lim\limits_{n\to\infty} f_n(x) =: f(x)$. Тогда f выпукла.*

Доказательство. Пусть (x, u), $(y, v) \in \operatorname{epi} f$. Возьмем произвольное $\varepsilon > 0$. Тогда для достаточно больших n получаем $(x, u + \varepsilon) \in \operatorname{epi} f_n$, $(y, v + \varepsilon) \in \operatorname{epi} f_n$. Значит, для любого $\lambda \in [0, 1]$
$$((1-\lambda)x + \lambda y, (1-\lambda)u + \lambda v + \varepsilon) \in \operatorname{epi} f_n,$$

или $f_n((1-\lambda)x+\lambda y) \le (1-\lambda)u+\lambda v+\varepsilon$. Переходим к пределу по n и получаем $f((1-\lambda)x+\lambda y) \le (1-\lambda)u+\lambda v+\varepsilon$. В силу произвольности $\varepsilon > 0$, $f((1-\lambda)x+\lambda y) \le (1-\lambda)u+\lambda v$, т.е.

$$((1-\lambda)x+\lambda y,\, (1-\lambda)u+\lambda v) \in \operatorname{epi} f.$$

Значит, f выпукла. □

Предложение 20. *Пусть A — множество, $f_\alpha : X \to \mathbb{R} \cup \{\pm\infty\}$ — выпуклые функции, $\alpha \in A$, $f(x) = \sup_{\alpha\in A} f_\alpha(x)$. Тогда f выпукла.*

Доказательство. Заметим, что $\operatorname{epi} f = \cap_{\alpha\in A} \operatorname{epi} f_\alpha$. Остается использовать выпуклость пересечения выпуклых множеств. □

4.2 Выпуклость гладких функций

Сначала напомним теоремы из математического анализа для функций одной переменной.

1. Если $f : \mathbb{R} \to \mathbb{R}$ всюду дифференцируема, то f выпукла тогда и только тогда, когда функция f' неубывающая.

2. Если $f : \mathbb{R} \to \mathbb{R}$ всюду дважды дифференцируема, то f выпукла тогда и только тогда, когда $f''(x) \ge 0$ для любого $x \in \mathbb{R}$.

Обобщим эти утверждения на произвольные нормированные пространства.

Пусть X — нормированное пространство. Напомним, что $f : X \to \mathbb{R}$ дифференцируема по Гато в точке x_0, если для любого $h \in X$ существует предел

$$\lim_{t\to 0} \frac{f(x_0+th)-f(x_0)}{t} =: f'(x_0)[h],$$

при этом отображение $f'(x_0) : h \mapsto f'(x_0)[h]$ — линейный непрерывный функционал.

Определение 13. *Пусть X — нормированное пространство. Отображение $A : X \to X^*$ называется монотонным, если $(A(x)-A(y))(x-y) \ge 0$ для любых x, $y \in X$.*

Если $X = \mathbb{R}$, то $X^* = \mathbb{R}$; $A : \mathbb{R} \to \mathbb{R}$ монотонно тогда и только тогда, когда A — неубывающая функция.

Предложение 21. *Если функция $f : X \to \mathbb{R}$ всюду дифференцируема по Гато, то f выпукла тогда и только тогда, когда отображение $x \mapsto f'(x)$ монотонно.*

Доказательство. Пусть x, $y \in X$. Положим

$$g_{x,y}(t) := g(t) := f((1-t)x + ty), \quad t \in \mathbb{R}.$$

Заметим, что f выпукла тогда и только тогда, когда $g_{x,y}$ выпукла для любых x, $y \in X$.

Функция g дифференцируема:

$$g'(t) = \lim_{\lambda \to 0} \frac{f((1-t)x + ty + \lambda(y-x)) - f((1-t)x + ty)}{\lambda} =$$

$$= f'((1-t)x + ty)[y - x].$$

Значит, g выпукла тогда и только тогда, когда g' возрастает, т.е.

$$f'((1-s)x + sy)[y-x] - f'((1-t)x + ty)[y-x] \geq 0, \quad t \leq s. \qquad (1)$$

Если g' возрастает, то подставим $s = 1$, $t = 0$ и получим

$$(f'(y) - f'(x))[y - x] \geq 0,$$

т.е. отображение $x \mapsto f'(x)$ монотонно.

Если отображение $x \mapsto f'(x)$ монотонно, то для любых $s > t$

$$0 \leq (f'((1-s)x + sy) - f'((1-t)x + ty))[(1-s)x + sy - (1-t)x - ty] =$$

$$= (f'((1-s)x + sy) - f'((1-t)x + ty))[(s-t)(y-x)];$$

разделив на $s - t$, получим (1). □

Пусть функция $f : X \to \mathbb{R}$ дифференцируема по Гато в окрестности U точки x_0. Пусть $h \in X$. Рассмотрим отображение $x \mapsto f'(x)[h]$, $x \in U$. Если оно дифференцируемо по Гато в точке x_0 для любого $h \in X$, то f называется дважды дифференцируемой по Гато в точке x_0.

Обозначим

$$f''(x_0)[h, \tilde{h}] := (f'(\cdot)[h])'(x_0)[\tilde{h}].$$

Заметим, что $f''(x_0)[h, \tilde{h}]$ билинейна по h, $\tilde{h}$.

Предложение 22. *Если функция $f : X \to \mathbb{R}$ всюду дважды дифференцируема по Гато, то f выпукла тогда и только тогда, когда $f''(x)[h,\, h] \geq 0$ для любых $x \in X$, $h \in X$.*

Доказательство. Как и в предыдущем утверждении, используем функцию $g_{x,y}(t) := g(t) := f((1-t)x+ty)$ и исследуем ее на выпуклость. Имеем

$$g'(t) = f'((1-t)x+ty)[y-x], \quad g''(t) = f''((1-t)x+ty)[y-x,\, y-x]. \tag{2}$$

Пусть $f''(\xi)[h,\, h] \geq 0$ для любого ξ, $h \in X$. Подставим $\xi = (1-t)x+ty$, $h = x-y$ в (2) и получим $g''(t) \geq 0$. Значит, функция $g_{x,y}$ выпукла для любых x, y.

Пусть $g_{x,y}$ выпукла для любых x, $y \in X$. Тогда

$$f''((1-t)x+ty)[y-x,\, y-x] \geq 0$$

для любого $t \in \mathbb{R}$. Подставив $t = 0$, $y = x+h$, получим $f''(x)[h,\, h] \geq 0$. $\square$

В частности, если $X = \mathbb{R}^n$, $f : X \to \mathbb{R}$ дважды дифференцируема, то f выпукла тогда и только тогда, когда матрица $\left(f''_{x_i x_j}(x)\right)_{i,j=1}^n$ неотрицательно определена. В самом деле, для $h = (h_1,\, \ldots,\, h_n)$

$$f'(x)[h] = \sum_{i=1}^n f'_{x_i}(x) h_i, \quad f''(x)[h,\, h] = \sum_{i,j=1}^n f''_{x_i x_j}(x) h_i h_j;$$

неотрицательность правой части для любого $h \in \mathbb{R}^n$ и означает неотрицательную определенность матрицы $\left(f''_{x_i x_j}(x)\right)_{i,j=1}^n$.

5 Субдифференциал выпуклой функции

Определение 14. *Пусть X — нормированное пространство, $f : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, $f(x_0) < +\infty$. Субдифференциалом f в точке x_0 называется множество*

$$\partial f(x_0) = \{x^* \in X^* :\ f(x) \geq f(x_0) + x^*(x - x_0), \quad \forall x \in X\}.$$

Геометрический смысл. График функции f лежит не ниже графика функции $l(x) = f(x_0) + x^*(x - x_0)$. График l — опорная гиперплоскость к $\operatorname{epi} f$.

Свойства субдифференциала.

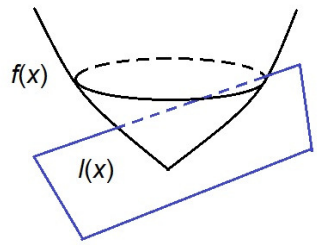

Рис. 16: Расположение графиков f и l

1. $\partial f(x_0)$ — выпуклое множество.

   В самом деле, пусть x^*, $y^* \in \partial f(x_0)$, $\lambda \in [0, 1]$. Тогда $f(x) \geq f(x_0) + x^*(x - x_0)$, $f(x) \geq f(x_0) + y^*(x - x_0)$. Умножим первое неравенство на $1 - \lambda$, второе на λ, сложим и получим

   $$f(x) \geq f(x_0) + ((1 - \lambda)x^* + \lambda y^*)(x - x_0),$$

   так что $(1 - \lambda)x^* + \lambda y^* \in \partial f(x_0)$.

2. $\partial f(x_0)$ — замкнутое множество.

   В самом деле, пусть $x_n^* \in \partial f(x_0)$, $x_n^* \underset{n\to\infty}{\to} x^*$. Тогда $f(x) \geq f(x_0) + x_n^*(x - x_0)$, $x \in X$. Перейдя к пределу в неравенстве, получим $x^* \in \partial f(x_0)$.

3. Если f липшицева с константой L в окрестности точки x_0, то $\partial f(x_0)$ содержится в шаре радиуса L.

   В самом деле, пусть $x^* \in \partial f(x_0)$, $\|x^*\| > L$. Тогда существует вектор h такой, что $\|h\| = 1$, $x^*(h) > L$. Для любого $\delta > 0$ получаем

   $$f(x_0 + \delta h) - f(x_0) \geq x^*(\delta h) > L\delta.$$

   С другой стороны, в силу условия Липшица, при достаточно малых δ имеем $f(x_0 + \delta h) - f(x_0) \leq L\delta$ — противоречие.

Задача 12. Пусть $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, $x_0 \in \operatorname{int} \operatorname{dom} f$. Показать, что $\partial f(x_0) = [f'_-(x_0), f'_+(x_0)]$, где $f'_\pm(x_0)$ — левая и правая производные.

Напомним формулировку теоремы Банаха–Штейнгауза из функционального анализа.

Теорема 8. *Пусть X — банахово пространство, Y — нормированное пространство, $A_\lambda : X \to Y$ — линейные непрерывные операторы, $\lambda \in \Lambda$. Пусть для любого $x \in X$ выполнено $\sup_{\lambda\in\Lambda} \|A_\lambda x\| < \infty$. Тогда*

$$\sup_{\lambda\in\Lambda} \|A_\lambda\| < \infty.$$

Следующее утверждение выводится из теоремы Банаха–Штейнгауза.

Предложение 23. *Пусть X — банахово пространство, $f : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, $x_0 \in \operatorname{dom} f$, для любого $h \in X$ существует $\delta > 0$ такое, что $x_0 + \delta h \in \operatorname{dom} f$. Тогда $\partial f(x_0)$ — ограниченное множество.*

Доказательство. Покажем, что множество $\{x^*(h)\}_{x^*\in\partial f(x_0)}$ ограничено для любого $h \in X$. Так как X полно, то по теореме Банаха–Штейнгауза получим ограниченность $\partial f(x_0)$.

Пусть $x^* \in \partial f(x_0)$. Тогда для любого $t \in \mathbb{R}$ выполнено $f(x_0 + th) \geq f(x_0) + t \cdot x^*(h)$. Положим $g(t) = f(x_0 + th)$. Тогда $g(t) \geq g(0) + t \cdot x^*(h)$. Значит, $x^*(h) \in \partial g(0)$. Так как $0 \in \operatorname{int}\operatorname{dom} g$, то $\partial g(0)$ ограничено по предыдущей задаче. □

Задача 13. Пусть $c_{0,0}$ — подпространство в c_0, состоящее из финитных последовательностей. Построить выпуклую функцию $f : c_{0,0} \to \mathbb{R}$ такую, что $\partial f(0)$ неограничен.

Задача 14. Пусть $f : X \to \mathbb{R}$, $f(x) = \|x\|$. а) Показать, что $\partial f(0) = B_{X^*}$ (единичный шар в X^*). б) Пусть $x_0 \neq 0$. Показать, что $\partial f(x_0) = \{x^* \in X^* : \ \|x^*\| = 1, \ x^*(x_0) = \|x_0\|\}$.

Задача 15. Пусть $f : c_0 \to \mathbb{R}$, $f(x) = \|x\|$, $x_0 \neq 0$. Найти $\partial f(x_0)$.

Указание. Напомним, что $x^* \in c_0^*$ тогда и только тогда, когда

$$x^*(x_1, \ldots, x_n, \ldots) = \sum_{n=1}^{\infty} \alpha_n x_n,$$

где $\sum\limits_{n=1}^{\infty} |\alpha_n| < \infty$. При этом $\|x^*\| = \sum\limits_{n=1}^{\infty} |\alpha_n|$.

5.1 Теоремы о непустоте субдифференциала

Предложение 24. *Пусть $f : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, $x_0 \in \operatorname{int}\operatorname{dom} f$. Пусть f ограничена сверху на некотором непустом открытом множестве. Тогда $\partial f(x_0) \neq \varnothing$.*

Доказательство. Рассмотрим в $X \times \mathbb{R}$ множества $A = \operatorname{epi} f$ и $B = \{(x_0,\, f(x_0))\}$. Они выпуклые, при этом $B \cap \operatorname{int} A = \varnothing$.

По условию теоремы, существует $U \subset X$ — открытое непустое множество, $f(x) \le c$ для любого $x \in U$. Тогда $U \times (c,\, +\infty) \subset A$, так что $\operatorname{int} A \ne \varnothing$.

По теореме отделимости, существует функционал $y^* \in (X \times \mathbb{R})^* \backslash \{0\}$ такой, что

$$\inf_{(x,\,u) \in A} y^*(x,\, u) \ge y^*(x_0,\, f(x_0)). \tag{3}$$

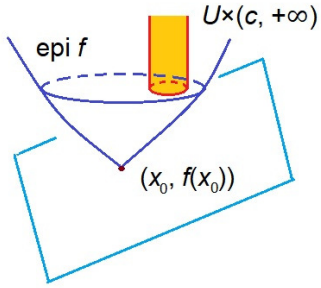

Рис. 17: Непустота субдифференциала

Напомним, что $(X \times \mathbb{R})^* = X^* \times \mathbb{R}$. Значит, существуют $x^* \in X^*$ и $\alpha \in \mathbb{R}$ такие, что $y^*(x,\, u) = x^*(x) + \alpha u$, $x \in X$, $u \in \mathbb{R}$. При этом $(x^*,\, \alpha) \ne (0,\, 0)$.

Неравенство (3) переписывается в виде

$$x^*(x) + \alpha u \ge x^*(x_0) + \alpha f(x_0), \quad (x,\, u) \in \operatorname{epi} f. \tag{4}$$

Пусть $\alpha < 0$. Взяв $x = x_0$, $u > f(x_0)$, получим противоречие.

Пусть $\alpha = 0$. Тогда (4) имеет вид

$$x^*(x) \ge x^*(x_0), \quad (x,\, u) \in \operatorname{epi} f.$$

По условию теоремы, $x_0 \in \operatorname{int} \operatorname{dom} f$, то есть существует $\delta > 0$ такое, что для любого $x \in B_\delta(x_0)$ выполнено $f(x) < +\infty$. Значит, $x^*(x) \ge x^*(x_0)$ для любого $x \in B_\delta(x_0)$. Отсюда следует, что $x^* = 0$. В самом деле, если $\|h\| \le \delta$, то $x^*(x_0 \pm h) \ge x^*(x_0)$ и поэтому $\pm x^*(h) \ge 0$; значит, $x^*(h) = 0$. Итак, мы получили $(x^*,\, \alpha) = (0,\, 0)$, что неверно.

Значит, $\alpha > 0$. Возьмем $x \in \operatorname{dom} f$, $u = f(x)$ и получим

$$x^*(x) + \alpha f(x) \geq x^*(x_0) + \alpha f(x_0),$$

или

$$f(x) \geq f(x_0) - \frac{x^*}{\alpha}(x - x_0).$$

Значит, $-\frac{x^*}{\alpha} \in \partial f(x_0)$. □

В частности, если f непрерывна в точке x_0, то $\partial f(x_0) \neq \varnothing$.

Предложение 25. *Пусть $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, $x_0 \in \operatorname{int} \operatorname{dom} f$. Тогда $\partial f(x_0) \neq \varnothing$.*

Доказательство. Рассуждения такие же, как в предыдущей теореме, только вместо теоремы отделимости применяем теорему об опорной гиперплоскости к выпуклому множеству в конечномерном пространстве. □

Пример. Пусть $\dim X = \infty$, $f : X \to \mathbb{R}$ — линейный неограниченный функционал. Тогда f — выпуклая функция. При этом $\partial f(x_0) = \varnothing$ для любого $x_0 \in X$. В самом деле, если $x^* \in \partial f(x_0)$, то для любого $h \in X$ выполнено $f(x_0+h) \geq f(x_0)+x^*(h)$, или $(f-x^*)(h) \geq 0$. Тогда $f-x^* = 0$, что невозможно, так как f неограниченный, а x^* ограниченный.

Задача 16. Привести пример выпуклой функции $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, у которой субдифференциал в некоторой точке $x_0 \in \operatorname{dom} f$ пуст.

5.2 Теоремы о локальной липшицевости

Пусть $f : X \to \mathbb{R} \cup \{+\infty\}$, $U \subset \operatorname{dom} f$ — открытое множество. Скажем, что f локально липшицева на U, если для любой точки $x \in U$ существует ее окрестность, на которой f липшицева.

Предложение 26. *Пусть X — нормированное пространство, $f : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, ограниченная на $B_R(x_0)$. Пусть $\alpha \in (0, 1)$. Тогда f липшицева на $B_{R\alpha}(x_0)$.*

Доказательство. Пусть f не липшицева на $B_{R\alpha}(x_0)$. Тогда для любого $n \in \mathbb{N}$ существуют точки x_n, $y_n \in B_{R\alpha}(x_0)$ такие, что $f(y_n) - f(x_n) > n\|y_n - x_n\|$. Проведем через x_n, y_n луч с началом в x_n. Его пересечение

с $B_R(x_0)$ — это отрезок $[x_n,\, z_n]$; тогда $y_n = (1-\lambda_n)x_n + \lambda_n z_n$, $\lambda_n \in (0,\, 1)$. Так как f выпукла, то

$$f(y_n) \le (1-\lambda_n)f(x_n) + \lambda_n f(z_n),$$

т.е.

$$n\|y_n - x_n\| < f(y_n) - f(x_n) \le \lambda_n(f(z_n) - f(x_n)).$$

Далее, $y_n - x_n = \lambda_n(z_n - x_n)$. Значит, $\|y_n - x_n\| = \lambda_n\|z_n - x_n\|$ и

$$f(z_n) - f(x_n) > n\|z_n - x_n\| \ge n(\|z_n - x_0\| - \|x_n - x_0\|) \ge nR(1-\alpha).$$

Получаем противоречие с тем, что f ограничена на $B_R(x_0)$. □

Теорема 9. *Пусть X — нормированное пространство, $f : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция. Пусть f ограничена сверху на некотором непустом открытом множестве. Тогда f локально липшицева на* $\operatorname{int}\operatorname{dom} f$.

Доказательство. В силу предложения 26, достаточно показать, что для любой точки из $\operatorname{int}\operatorname{dom} f$ существует ее окрестность, в которой f ограничена.

Пусть $x_0 \in \operatorname{int}\operatorname{dom} f$. Покажем, что f ограничена снизу в некоторой окрестности точки x_0. В самом деле, $\partial f(x_0) \ne \varnothing$ по предложению 24. Пусть $x^* \in \partial f(x_0)$. Тогда для любого $x \in X$ выполнено $f(x) \ge f(x_0) + x^*(x - x_0)$; остается заметить, что правая часть ограничена на шаре.

Теперь докажем ограниченность сверху в окрестности точки x_0. По условию теоремы, существуют шар $B_r(\xi_0)$ и число $c \in \mathbb{R}$ такие, что $f(x) \le c$ для любого $x \in B_r(\xi_0)$. Так как $x_0 \in \operatorname{int}\operatorname{dom} f$, то существует $\varkappa > 1$ такое, что $\hat{x} := \xi_0 + \varkappa(x_0 - \xi_0) \in \operatorname{dom} f$. Тогда $x_0 = \xi_0 + \lambda(\hat{x} - \xi_0)$, $\lambda = 1/\varkappa$.

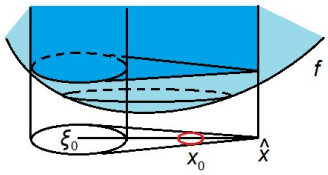

Рис. 18: Ограниченность сверху

Имеем: $B_r(\xi_0) \times [c,\, +\infty) \subset \operatorname{epi} f$, $\{\hat{x}\} \times [f(\hat{x}),\, +\infty) \subset \operatorname{epi} f$. Так как f выпукла, то $B_{(1-\lambda)r}(x_0) \times [\max\{c,\, f(\hat{x})\},\, +\infty) \subset \operatorname{epi} f$. Значит, f ограничена сверху на $B_{(1-\lambda)r}(x_0)$. □

Теорема 10. *Пусть* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *— выпуклая функция. Тогда* f *локально липшицева на* $\operatorname{int} \operatorname{dom} f$.

Доказательство. Снова достаточно проверить локальную ограниченность на $\operatorname{int} \operatorname{dom} f$.

Пусть $x_0 \in \operatorname{int} \operatorname{dom} f$. Тогда $\partial f(x_0) \neq \varnothing$ в силу предложения 25, поэтому f ограничена снизу в окрестности x_0.

Докажем ограниченность сверху в некоторой окрестности x_0. Пусть $\{e_j\}_{j=1}^n$ — стандартные базисные векторы, $\delta > 0$ такие, что $x_0 \pm \delta e_j \in \operatorname{dom} f$, $1 \leq j \leq n$. Тогда $\operatorname{conv}\{x_0 \pm \delta e_j\}_{j=1}^n \subset \operatorname{dom} f$, x_0 — внутренняя точка $\operatorname{conv}\{x_0 \pm \delta e_j\}_{j=1}^n$. Пусть $\lambda_j \geq 0, \mu_j \geq 0, 1 \leq j \leq n$, $\sum\limits_{j=1}^n \lambda_j + \sum\limits_{j=1}^n \mu_j = 1$. Тогда

$$f\left(\sum_{j=1}^n \lambda_j (x_0 + \delta e_j) + \sum_{j=1}^n \mu_j (x_0 - \delta e_j)\right) \leq$$

$$\leq \sum_{j=1}^n \lambda_j f(x_0 + \delta e_j) + \sum_{j=1}^n \mu_j f(x_0 - \delta e_j) \leq \max\{f(x_0 \pm \delta e_j)\}_{j=1}^n.$$

Теорема доказана. □

5.3 Субдифференциал l_p-нормы на $\mathbb{R}^n$

Пусть $1 < p < \infty$, $f : \mathbb{R}^n \to \mathbb{R}$, $f(x_1, \ldots, x_n) = \left(\sum\limits_{j=1}^n |x_j|^p\right)^{1/p}$. Вычислим $\partial f(0)$.

Напомним неравенство Гёльдера: для любых $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$

$$\sum_{j=1}^n x_j y_j \leq \left(\sum_{j=1}^n |x_j|^p\right)^{1/p} \left(\sum_{j=1}^n |y_j|^{p'}\right)^{1/p'},$$

где $\frac{1}{p'} + \frac{1}{p} = 1$.

Это неравенство является точным, т.е. для любого $(y_1, \ldots, y_n) \in \mathbb{R}^n$ существует вектор $(x_1, \ldots, x_n) \in \mathbb{R}^n \backslash \{0\}$ такой, что

$$\sum_{j=1}^n x_j y_j = \left(\sum_{j=1}^n |x_j|^p\right)^{1/p} \left(\sum_{j=1}^n |y_j|^{p'}\right)^{1/p'}. \tag{5}$$

В самом деле, достаточно рассмотреть $(y_1, \ldots, y_n) \neq (0, \ldots, 0)$. Положим $x_j = |y_j|^{p'-1}\operatorname{sgn} y_j$. Тогда

$$\sum_{j=1}^{n} x_j y_j = \sum_{j=1}^{n} |y_j|^{p'},$$

$$\left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} = \left(\sum_{j=1}^{n} |y_j|^{p(p'-1)}\right)^{1/p} = \left(\sum_{j=1}^{n} |y_j|^{p'}\right)^{1-1/p'},$$

откуда следует (5).

Предложение 27. *Выполнено равенство*

$$\partial f(0) = \left\{(y_1, \ldots, y_n) \in \mathbb{R}^n : \sum_{j=1}^{n} |y_j|^{p'} \leq 1\right\}.$$

Доказательство. Пусть $(y_1, \ldots, y_n) \in \mathbb{R}^n$. Условие $(y_1, \ldots, y_n) \in \partial f(0)$ эквивалентно неравенству

$$\left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} \geq \sum_{j=1}^{n} x_j y_j, \quad \forall (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Если $\sum\limits_{j=1}^{n} |y_j|^{p'} \leq 1$, то в силу неравенства Гёльдера

$$\sum_{j=1}^{n} x_j y_j \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} \left(\sum_{j=1}^{n} |y_j|^{p'}\right)^{1/p'} \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p}.$$

Значит, $(y_1, \ldots, y_n) \in \partial f(0)$.

Пусть $\sum\limits_{j=1}^{n} |y_j|^{p'} > 1$. Выберем вектор $(x_1, \ldots, x_n) \in \mathbb{R}^n \backslash \{0\}$ такой, что выполнено (5). Тогда

$$\sum_{j=1}^{n} x_j y_j = \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} \left(\sum_{j=1}^{n} |y_j|^{p'}\right)^{1/p'} > \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p},$$

поэтому $(y_1, \ldots, y_n) \notin \partial f(0)$. $\square$

5.4 Субдифференциал в точке гладкости

Пусть $\varphi : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, x_0, $x \in \operatorname{dom} f$, $0 < t < 1$. Тогда из неравенства $\varphi((1-t)x_0 + tx) \le (1-t)\varphi(x_0) + t\varphi(x)$ следует, что

$$\frac{\varphi((1-t)x_0 + tx) - \varphi(x_0)}{t} \le \varphi(x) - \varphi(x_0). \tag{6}$$

Теорема 11. *Пусть $f : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, $x_0 \in \operatorname{int} \operatorname{dom} f$, при этом f дифференцируема по Гато в точке x_0. Тогда*

$$\partial f(x_0) = \{f'(x_0)\}.$$

Доказательство. Пусть $x^* \in \partial f(x_0)$. Тогда $f(x) \ge f(x_0) + x^*(x - x_0)$ для любого $x \in X$. Перепишем это неравенство в виде $f(x) - x^*(x) \ge f(x_0) - x^*(x_0)$. Положим $g(x) = f(x) - x^*(x)$. Тогда x_0 — точка минимума функции g. Функция g дифференцируема по Гато в точке x_0 и $g'(x_0) = f'(x_0) - x^*$. По теореме Ферма, $g'(x_0) = 0$, так что $x^* = f'(x_0)$. Значит, $\partial f(x_0) \subset \{f'(x_0)\}$.

Докажем, что $f'(x_0) \in \partial f(x_0)$. Пусть $\varphi(x) = f(x) - f'(x_0)[x]$. Это выпуклая функция и $\varphi'(x_0) = 0$. Отсюда следует, что x_0 — точка минимума φ. (В самом деле, если существует точка x такая, что $\varphi(x) < \varphi(x_0)$, то применяем (6), переходим к пределу при $t \to +0$ и получаем $\varphi'(x_0)[x - x_0] \le \varphi(x) - \varphi(x_0) < 0$, т.е. $\varphi'(x_0) \ne 0$). $\square$

Теорема 12. *Пусть $f : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, $x_0 \in \operatorname{int} \operatorname{dom} f$, $\partial f(x_0) = \{x^*\}$. Пусть f ограничена сверху на непустом открытом множестве или X конечномерно. Тогда f дифференцируема по Гато в точке x_0 и $f'(x_0) = x^*$.*

Доказательство. Из (6) следует, что для любого $h \in X$ функция

$$t \mapsto \frac{f(x_0 + th) - f(x_0)}{t}, \quad t > 0,$$

является неубывающей; значит, существует $\lim\limits_{t \to +0} \frac{f(x_0+th)-f(x_0)}{t} =: \varphi(h)$. Заметим, что $\varphi(\lambda h) = \lambda\varphi(h)$, $f(x_0 + \lambda h) - f(x_0) \ge \lambda\varphi(h)$ при $\lambda \ge 0$. Так как $x^* \in \partial f(x_0)$, то $\varphi(h) \ge x^*(h)$.

Пусть $f'(x_0) \ne x^*$ или функция f не дифференцируема по Гато. Тогда существует вектор $h \in X$ такой, что $\varphi(h) > x^*(h)$.

Множества

$$A = \{(x,\, u) \in X \times \mathbb{R} :\ u > f(x)\}$$

и

$$B = \{(x_0 + \lambda h,\, f(x_0) + \lambda\varphi(h)) :\ \lambda \geq 0\}$$

выпуклы и не пересекаются. Применяем одну из теорем отделимости к этим множествам.

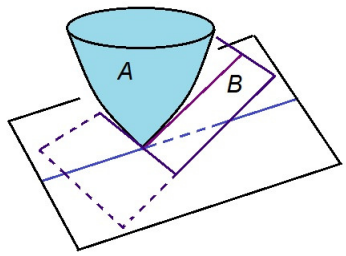

Рис. 19: Множества A и B

Как и в теоремах о непустоте субдифференциала, получаем, что существует функционал $y^* \in X^*$ такой, что

$$f(x) + y^*(x) \geq f(x_0) + y^*(x_0),\ x \in X,$$

$$f(x_0) + \lambda\varphi(h) + y^*(x_0 + \lambda h) \leq f(x_0) + y^*(x_0),\ \lambda \geq 0.$$

Значит, $-y^* \in \partial f(x_0)$, $\varphi(h) \leq -y^*(h)$. Так как $\varphi(h) > x^*(h)$, то $x^* \neq -y^*$, т.е. $\partial f(x_0)$ содержит две различные точки. $\square$

6 Теорема Моро–Рокафеллара

Предложение 28. *Пусть $f : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, $x^* \in X^*$, $x_0 \in \operatorname{dom} f$, $c \in \mathbb{R}$. Пусть $\tilde{f}(x) = f(x) - c - x^*(x)$. Тогда $\partial\tilde{f}(x_0) = -x^* + \partial f(x_0)$.*

Доказательство. Пусть $z^* \in X^*$. Тогда

$$z^* \in \partial\tilde{f}(x_0) \Leftrightarrow \forall x \in X\ \ f(x) - c - x^*(x) \geq f(x_0) - c - x^*(x_0) + z^*(x - x_0) \Leftrightarrow$$

$$\forall x \in X\ \ f(x) \geq f(x_0) + (x^* + z^*)(x - x_0) \Leftrightarrow x^* + z^* \in \partial f(x_0).$$

$\square$

Теорема 13. *Пусть f, $g : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклые функции, $f(x_0) < +\infty$, $g(x_0) < +\infty$, f непрерывна в точке x_0. Тогда $\partial(f+g)(x_0) = \partial f(x_0) + \partial g(x_0)$.*

Доказательство. Докажем включение $\partial(f+g)(x_0) \supset \partial f(x_0) + \partial g(x_0)$. Пусть $x^* \in \partial f(x_0)$, $y^* \in \partial g(x_0)$. Тогда для любого $x \in X$

$$f(x) \geq f(x_0) + x^*(x - x_0), \quad g(x) \geq g(x_0) + y^*(x - x_0).$$

Сложим эти неравенства и получим

$$(f+g)(x) \geq (f+g)(x_0) + (x^* + y^*)(x - x_0),$$

т.е. $x^* + y^* \in \partial(f+g)(x_0)$.

Теперь докажем включение $\partial(f+g)(x_0) \subset \partial f(x_0) + \partial g(x_0)$. Пусть $x^* \in \partial(f+g)(x_0)$. Покажем, что $x^* \in \partial f(x_0) + \partial g(x_0)$.

Положим $\tilde{f}(x) = f(x) - f(x_0) - x^*(x - x_0)$, $\tilde{g}(x) = g(x) - g(x_0)$. В силу предложения 28, $\partial\tilde{f}(x_0) = -x^* + \partial f(x_0)$, $\partial\tilde{g}(x_0) = \partial g(x_0)$. Поэтому достаточно проверить, что $0 \in \partial\tilde{f}(x_0) + \partial\tilde{g}(x_0)$.

Имеем $0 \in \partial(\tilde{f} + \tilde{g})(x_0)$ (снова в силу предложения 28). Кроме того, $\tilde{f}(x_0) = \tilde{g}(x_0) = 0$. Значит, для любого $x \in X$ выполнено $\tilde{f}(x) + \tilde{g}(x) \geq 0$, или $\tilde{f}(x) \geq -\tilde{g}(x)$.

Положим $A = \operatorname{epi} \tilde{f}$, $B = \{(x, u) \in X \times \mathbb{R} : \ u \leq -\tilde{g}(x)\}$. Это выпуклые множества.

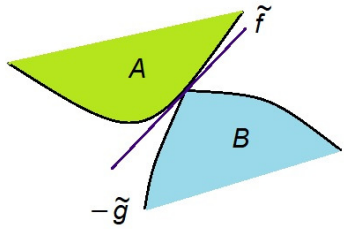

Рис. 20: Множества A и B

Так как $\tilde{f}$ непрерывна в точке x_0, то $\operatorname{int} A \neq \varnothing$. Покажем, что $(\operatorname{int} A) \cap B = \varnothing$. В самом деле, пусть $(x, u) \in (\operatorname{int} A) \cap B$. Тогда $\tilde{f}(x) < u \leq -\tilde{g}(x)$, т.е. $\tilde{f}(x) + \tilde{g}(x) < 0$ — противоречие.

Итак, можно применить теорему отделимости. Получаем, что существует $(z^*, \alpha) \in (X^* \times \mathbb{R}) \backslash \{(0, 0)\}$,

$$\inf_{(x,\,u) \in A} (z^*(x) + \alpha u) \geq \sup_{(x,\,u) \in B} (z^*(x) + \alpha u).$$

Заметим, что $(x_0,\, 0) \in A \cap B$. Поэтому

$$\inf_{(x,\,u)\in A} (z^*(x) + \alpha u) \geq z^*(x_0).$$

Поэтому, как в теореме о непустоте субдифференциала, получаем, что $\alpha > 0$ и $-\frac{z^*}{\alpha} \in \partial \tilde{f}(x_0)$.

Далее, $z^*(x_0) \geq z^*(x) - \alpha \tilde{g}(x)$, или $\tilde{g}(x) \geq \frac{z^*}{\alpha}(x - x_0)$. Значит, $\frac{z^*}{\alpha} \in \partial \tilde{g}(x_0)$.

Таким образом, $0 = -\frac{z^*}{\alpha} + \frac{z^*}{\alpha} \in \partial \tilde{f}(x_0) + \partial \tilde{g}(x_0)$. Теорема доказана. $\square$

7 Выпуклая задача без ограничений

Пусть X — линейное пространство, $f : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция. Выпуклая задача без ограничений имеет вид

$$f(x) \to \inf, \quad x \in X. \tag{7}$$

Из определения субдифференциала следует

Предложение 29. *Пусть X — нормированное пространство. Точка x_0 является решением задачи* (7) *тогда и только тогда, когда $0 \in \partial f(x_0)$.*

Задача Штейнера. Пусть $a = (a_1,\, a_2)$, $b = (b_1,\, b_2)$, $c = (c_1,\, c_2)$ — точки на евклидовой плоскости, не лежащие на одной прямой. Задача Штейнера имеет вид

$$f(x) := |x - a| + |x - b| + |x - c| \to \inf$$

(здесь $|\cdot|$ — стандартная евклидова норма на $\mathbb{R}^2$).

Обозначим

$$f_a(x) = |x - a|, \quad f_b(x) = |x - b|, \quad f_c(x) = |x - c|.$$

Функция f непрерывна, при этом $f(x) \underset{|x|\to\infty}{\to} +\infty$. Значит, f имеет точку минимума. Обозначим ее через $\hat{x} = (\hat{x}_1,\, \hat{x}_2)$.

Рассмотрим два случая.

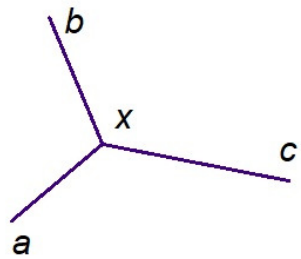

Рис. 21: Минимизация суммы длин отрезков

1. Пусть $\hat{x} \notin \{a,\, b,\, c\}$. Тогда f дифференцируема в точке $\hat{x}$. Значит, $f'(\hat{x}) = 0$. Имеем

$$f(x) = f_a(x) + f_b(x) + f_c(x),$$

$$\nabla f_a(\hat{x}_1,\, \hat{x}_2) = \left(\frac{\hat{x}_1 - a_1}{|\hat{x} - a|},\, \frac{\hat{x}_2 - a_2}{|\hat{x} - a|}\right) =: e_a,$$

$$\nabla f_b(\hat{x}_1,\, \hat{x}_2) = \left(\frac{\hat{x}_1 - b_1}{|\hat{x} - b|},\, \frac{\hat{x}_2 - b_2}{|\hat{x} - b|}\right) =: e_b,$$

$$\nabla f_c(\hat{x}_1,\, \hat{x}_2) = \left(\frac{\hat{x}_1 - c_1}{|\hat{x} - c|},\, \frac{\hat{x}_2 - c_2}{|\hat{x} - c|}\right) =: e_c,$$

то есть $f'(\hat{x}_1,\, \hat{x}_2) = e_a + e_b + e_c$. Векторы e_a, e_b, e_c имеют единичную длину и направлены соответственно от a, b, c к $\hat{x}$.

Получаем: $0 = f'(\hat{x}_1,\, \hat{x}_2) = e_a + e_b + e_c$, или $-e_a - e_b - e_c = 0$. Это эквивалентно тому, что попарные углы между $-e_a$, $-e_b$, $-e_c$ равны $120°$. Значит, если соединить $\hat{x}$ с a, b, c, то углы между отрезками будут равны $120°$. Такая точка называется точкой Торричелли.

2. Пусть $\hat{x} = a$ (другие вершины рассматриваются аналогично). Тогда f_b, f_c дифференцируемы в точке $\hat{x}$, $\nabla f_b(a) = e_b$, $\nabla f_c = e_c$; векторы e_b, e_c имеют единичную длину и направлены соответственно от b, c к a.

Найдем $\partial f_a(a)$. Условие $(y_1,\, y_2) \in \partial f_a(a)$ равносильно неравенству

$$\sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2} \geq y_1(x_1 - a_1) + y_2(x_2 - a_2),\ (x_1,\, x_2) \in \mathbb{R}^2,$$

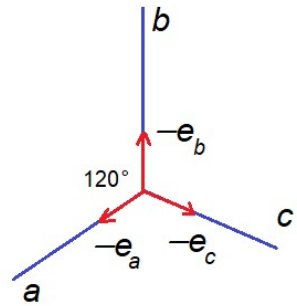

Рис. 22: Точка Торричелли

или

$$\sqrt{z_1^2 + z_2^2} \geq y_1 z_1 + y_2 z_2,\ (z_1,\, z_2) \in \mathbb{R}^2,$$

т.е. $(y_1,\, y_2) \in \partial|\cdot|(0,\, 0) = B$ — единичный евклидов круг.

Итак, a является точкой минимума тогда и только тогда, когда $0 \in e_b + e_c + B$, или $-e_b - e_c \in B$. Это равносильно тому, что угол при вершине a не меньше $120°$.

Таким образом, если угол при одной из вершин треугольника больше или равен $120°$, то эта вершина будет решением задачи. Если все углы меньше $120°$, то решением будет точка Торричелли.

8 Задача выпуклого программирования

Пусть X — линейное пространство, $f_0,\, \ldots,\, f_m : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклые функции. Задача

$$\begin{cases} f_0(x) \to \inf, \\ f_j(x) \leq 0, \quad 1 \leq j \leq m, \end{cases} \tag{8}$$

называется задачей выпуклого программирования.

Точка $x \in X$ называется допустимой, если $f_j(x) \leq 0$, $1 \leq j \leq m$. Заметим, что множество допустимых точек выпукло.

Допустимая точка $\hat{x}$ называется точкой минимума в задаче (8), если $f_0(\hat{x}) \leq f_0(x)$ для любого допустимого x.

Теорема Каруша – Куна – Таккера является аналогом принципа Лагранжа.

Теорема 14. (Каруш – Кун – Таккер). *Пусть X — линейное пространство, $f_0, \ldots, f_m : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклые функции.*

1. *(необходимое условие). Пусть $\hat{x}$ — точка минимума в задаче* (8), *$f_0(\hat{x}) < \infty$. Тогда существует ненулевой набор чисел $\lambda_0, \lambda_1, \ldots, \lambda_m$ со следующими свойствами:*

   (a) $\lambda_j \geq 0$, $0 \leq j \leq m$ (условие неотрицательности);

   (b) $\lambda_j f_j(\hat{x}) = 0$, $1 \leq j \leq m$ (условие дополняющей нежесткости);

   (c) $\hat{x}$ является точкой минимума функции $\mathcal{L}(x) := \sum\limits_{j=0}^{m} \lambda_j f_j(x)$ (условие минимума).

2. *(достаточное условие). Пусть $\hat{x}$ — допустимая точка. Пусть существует набор чисел $\lambda_0, \lambda_1, \ldots, \lambda_m$ со свойствами a)–c), при этом $\lambda_0 > 0$. Тогда $\hat{x}$ — точка минимума в задаче* (8).

3. *Пусть существует точка $\overline{x} \in X$ такая, что $f_j(\overline{x}) < 0$, $1 \leq j \leq m$ (условие Слейтера). Тогда, если $\lambda_0, \lambda_1, \ldots, \lambda_m$ — ненулевой набор чисел со свойствами a)–c), то $\lambda_0 > 0$.*

Таким образом, если выполнено условие Слейтера, то $\lambda_0 > 0$ и необходимое условие минимума совпадает с достаточным.

Задача 17. Привести пример задачи выпуклого программирования, в которой для любого набора чисел $(\lambda_0, \ldots, \lambda_m)$, удовлетворяющего a–c из теоремы Куна – Таккера, выполнено $\lambda_0 = 0$.

Задача 18. Привести пример задачи выпуклого программирования такой, что допустимая $\hat{x}$ — не точка минимума, но существует ненулевой набор $(\lambda_0, \ldots, \lambda_m)$, удовлетворяющий a–c из теоремы Куна – Таккера.

Теперь докажем теорему.

Доказательство. Необходимое условие. Пусть $\hat{x}$ — точка минимума. Без ограничения общности можно считать, что $f_0(\hat{x}) = 0$. Рассмотрим множества в $\mathbb{R}^{m+1}$:

$$A = \{(\mu_0, \mu_1, \ldots, \mu_m) : \exists x \in X : \ f_j(x) \leq \mu_j, \ 0 \leq j \leq m\},$$

$$B = \{(\mu_0, \mu_1, \ldots, \mu_m) : \mu_j < 0, \ 0 \leq j \leq m\}.$$

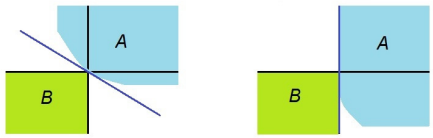

Рис. 23: Множества A и B

Множество B выпукло и открыто. Докажем, что A выпукло. В самом деле, пусть $(\mu_0, \mu_1, \ldots, \mu_m)$, $(\nu_0, \nu_1, \ldots, \nu_m) \in A$, $\lambda \in [0, 1]$. По определению, существуют точки x, $y \in X$ такие, что $f_j(x) \le \mu_j$, $f_j(y) \le \nu_j$, $0 \le j \le m$. Так как функции f_j выпуклы, то $f_j((1-\lambda)x + \lambda y) \le (1-\lambda)f_j(x) + \lambda f_j(y) \le (1-\lambda)\mu_j + \lambda\nu_j$. Значит, $(1-\lambda)(\mu_0, \mu_1, \ldots, \mu_m) + \lambda(\nu_0, \nu_1, \ldots, \nu_m) \in A$.

Докажем, что множества A и B не пересекаются. В самом деле, пусть $(\mu_0, \mu_1, \ldots, \mu_m) \in A \cap B$. Тогда существует $x \in X$ такое, что $f_j(x) \le \mu_j < 0$, $0 \le j \le m$. Значит, точка x допустимая и $f_0(x) < 0$. Это противоречит тому, что минимум в задаче (8) равен 0.

По теореме отделимости, существует набор чисел $(\lambda_0, \lambda_1, \ldots, \lambda_m) \ne (0, 0, \ldots, 0)$ такой, что

$$\inf_{(\mu_0, \mu_1, \ldots, \mu_m) \in A} \sum_{j=0}^{m} \lambda_j \mu_j \ge \sup_{(\mu_0, \mu_1, \ldots, \mu_m) \in B} \sum_{j=0}^{m} \lambda_j \mu_j. \tag{9}$$

Пусть $\lambda_{j_*} < 0$ для некоторого $j_* \in \{0, \ldots, m\}$. Зафиксируем какие-нибудь отрицательные числа μ_j при $j \ne j_*$; μ_{j_*} устремим к $-\infty$ и получим, что правая часть (9) равна $+\infty$. Значит, $\lambda_j \ge 0$, $0 \le j \le m$ (т.е. выполнено условие неотрицательности). Тогда правая часть (9) равна 0, т.е.

$$\sum_{j=0}^{m} \lambda_j \mu_j \ge 0, \quad (\mu_0, \mu_1, \ldots, \mu_m) \in A. \tag{10}$$

В частности, так как $(f_0(x), f_1(x), \ldots, f_m(x)) \in A$ для любого $x \in X$, то

$$\sum_{j=0}^{m} \lambda_j f_j(x) \ge 0. \tag{11}$$

Проверим условие дополняющей нежесткости. Пусть $1 \le j \le m$. Положим $\mu_i = 0$, $i \ne j$, $\mu_j = f_j(\hat{x})$. Тогда $f_0(\hat{x}) = 0 = \mu_0$, $f_i(\hat{x}) \le 0 = \mu_i$,

$i \in \{1, \ldots, m\} \backslash \{j\}$, $f_j(\hat{x}) = \mu_j$. Значит, $(\mu_0, \mu_1, \ldots, \mu_m) \in A$ и поэтому $\lambda_j f_j(\hat{x}) \overset{(10)}{\geq} 0$. С другой стороны, $\lambda_j \geq 0$, $f_j(\hat{x}) \leq 0$. Значит, $\lambda_j f_j(\hat{x}) = 0$.

Из условия дополняющей нежесткости и равенства $f_0(\hat{x}) = 0$ получаем $\sum\limits_{j=0}^{m} \lambda_j f_j(\hat{x}) = 0$. Это вместе с (11) дает условие минимума.

Достаточное условие. Пусть $\hat{x}$ — допустимая точка, $(\lambda_0, \lambda_1, \ldots, \lambda_m)$ — набор чисел со свойствами a)–c), $\lambda_0 > 0$. Покажем, что $f_0(x) \geq f_0(\hat{x})$ для любой допустимой точки x. Из неравенства $\mathcal{L}(x) \geq \mathcal{L}(\hat{x})$ и условия дополняющей нежесткости получаем

$$\sum_{j=0}^{m} \lambda_j f_j(x) \geq \sum_{j=0}^{m} \lambda_j f_j(\hat{x}) = \lambda_0 f_0(\hat{x}).$$

Так как $\lambda_j \geq 0$, $f_j(x) \leq 0$ при $1 \leq j \leq m$, то $\lambda_0 f_0(x) \geq \lambda_0 f_0(\hat{x})$. Остается сократить на $\lambda_0 > 0$.

Условие строгой положительности λ_0. Пусть выполнено условие Слейтера, но $\lambda_0 = 0$. Из условий минимума и дополняющей нежесткости получаем

$$\sum_{j=1}^{m} \lambda_j f_j(\overline{x}) \geq \sum_{j=1}^{m} \lambda_j f_j(\hat{x}) = 0.$$

С другой стороны, левая часть строго отрицательна: $\lambda_j \geq 0$, $f_j(\overline{x}) < 0$ для любого $j = 1, \ldots, m$, при этом λ_j одновременно не равны нулю. Получили противоречие. $\square$

9 Еще об опорных гиперплоскостях, или внешние нормальные конусы

Пусть X — нормированное пространство, $A \subset X$ — выпуклое множество, $\hat{x} \in A$. Внешним нормальным конусом к множеству A в точке $\hat{x}$ называется

$$K_A(\hat{x}) = \{x^* \in X^* : \ x^*(x - \hat{x}) \leq 0, \ x \in A\}.$$

Если $x^* \in K_A(\hat{x}) \backslash \{0\}$, то $\{x \in X : \ x^*(x) = x^*(\hat{x})\}$ — опорная гиперплоскость к A в точке $\hat{x}$.

Свойства внешнего нормального конуса:

1. если x^*, $y^* \in K_A(\hat{x})$, то $x^* + y^* \in K_A(\hat{x})$;

2. если $x^* \in K_A(\hat{x})$, $\lambda \geq 0$, то $\lambda x^* \in K_A(\hat{x})$.

Отсюда следует, что $K_A(\hat{x})$ — выпуклое множество.

Пример. Пусть $f : \mathbb{R}^n \to \mathbb{R}$ — выпуклая функция,

$$A = \{x \in \mathbb{R}^n : \ f(x) \leq c\},$$

$f(\hat{x}) = c$, f дифференцируема по Гато в точке $\hat{x}$, $f'(\hat{x}) \neq 0$. Найдем $K_A(\hat{x})$.

Пусть $x^*(x) = \sum\limits_{j=1}^{n} \alpha_j x_j$. Функционал x^* принадлежит $K_A(\hat{x})$ тогда и только тогда, когда $\hat{x}$ — решение задачи

$$-x^*(x) \to \inf, \quad f(x) - c \leq 0.$$

Это задача выпуклого программирования. Так как $f'(\hat{x}) \neq 0$, то c не является минимальным значением f, т.е. существует $\overline{x} \in \mathbb{R}^n$ такое, что $f(\overline{x}) < c$. Значит, выполнено условие Слейтера. Поэтому в теореме Куна–Таккера можно взять $\lambda_0 = 1$, при этом необходимое условие минимума совпадет с достаточным.

Итак, $x^* \in K_A(\hat{x})$ тогда и только тогда, когда существует $\lambda \geq 0$ такое, что $\lambda(f(\hat{x}) - c) = 0$ и $\hat{x}$ — точка минимума функции $-x^*(x) + \lambda(f(x) - c)$. Так как $f(\hat{x}) - c = 0$ и f дифференцируема в точке $\hat{x}$, то получаем эквивалентное условие: существует $\lambda \geq 0$ такое, что $-x^* + \lambda f'(\hat{x}) = 0$.

Значит, $K_A(\hat{x}) = \{\lambda f'(\hat{x}) : \ \lambda \geq 0\}$ (т.е. это луч, натянутый на вектор внешней нормали к A в точке $\hat{x}$).

Теорема 15. *Пусть X — нормированное пространство, A, $B \subset X$ — выпуклые множества,* $\operatorname{int}(A \cap B) \neq \varnothing$. *Пусть $\hat{x} \in A \cap B$ — граничная точка множеств A и B. Тогда*

$$K_{A \cap B}(\hat{x}) = \operatorname{conv}(K_A(\hat{x}) \cup K_B(\hat{x})) = K_A(\hat{x}) + K_B(\hat{x}).$$

Доказательство. Сделав сдвиг, можно считать, что $0 \in \operatorname{int}(A \cap B)$.

Сначала опишем $K_A(\hat{x})$.

Пусть p_A — функционал Минковского множества A.

Отметим, что так как множество A выпукло и $0 \in \operatorname{int} A$, то функция p_A непрерывна. В самом деле, $B_\delta(0) \subset A$ для некоторого $\delta > 0$, поэтому

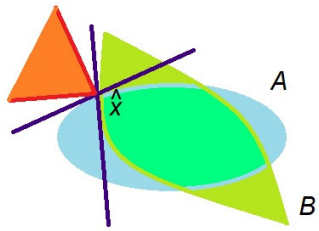

Рис. 24: Внешний нормальный конус

$p_A(x) \leq p_{B_\delta(0)}(x) = \frac{\|x\|}{\delta}$. По свойству функционала Минковского, для любых $x, y \in X$ имеем $p_A(x) = p_A(x-y+y) \leq p_A(x-y) + p_A(y)$; значит, $|p_A(x) - p_A(y)| \leq \max\{p_A(x-y), p_A(y-x)\} \leq \frac{\|x-y\|}{\delta}$.

Итак, p_A непрерывен. Значит,

$$\overline{A} = \{x : \ p_A(x) \leq 1\}. \tag{12}$$

Имеем

$$x^* \in K_A(\hat{x}) \ \Leftrightarrow \ \sup_{x \in A} x^*(x) = x^*(\hat{x}) \ \Leftrightarrow \ \sup_{x \in \overline{A}} x^*(x) = x^*(\hat{x}).$$

В силу (12), последнее условие эквивалентно тому, что $\hat{x}$ — решение задачи

$$-x^*(x) \to \inf, \ p_A(x) - 1 \leq 0.$$

Также из непрерывности p_A следует, что $p_A(\hat{x}) = 1$ (т.к. $\hat{x}$ — граничная точка A). Так как $p_A(0) = 0 < 1$, то выполнено условие Слейтера. Значит, в теореме Куна–Таккера можно взять $\lambda_0 = 1$, при этом необходимое условие минимума совпадает с достаточным.

Таким образом, $x^* \in K_A(\hat{x})$ тогда и только тогда, когда существует $\lambda \geq 0$ такое, что $\lambda(p_A(\hat{x}) - 1) = 0$ и $\hat{x}$ — точка минимума функции $-x^*(x) + \lambda p_A(x)$, или $0 \in \partial(-x^* + \lambda p_A)(\hat{x})$. Так как $p_A(\hat{x}) = 1$, то условие дополняющей нежесткости выполнено при всех $\lambda \geq 0$. Значит,

$$x^* \in K_A(\hat{x}) \ \Leftrightarrow \ x^* \in \{\lambda \partial p_A(\hat{x}) : \ \lambda \geq 0\}.$$

Итак,

$$K_A(\hat{x}) = \{\lambda \partial p_A(\hat{x}) : \ \lambda \geq 0\} \tag{13}$$

и аналогично

$$K_B(\hat{x}) = \{\lambda \partial p_B(\hat{x}) : \ \lambda \geq 0\}. \tag{14}$$

Так как A и B выпуклы, $0 \in \operatorname{int}(A \cap B)$, то $\overline{A \cap B} = \overline{A} \cap \overline{B}$ (если $x \in \overline{A} \cap \overline{B}$, то $p_A(x) \leq 1$, $p_B(x) \leq 1$; значит, $p_A(\lambda x) < 1$, $p_B(\lambda x) < 1$ при $0 < \lambda < 1$ и, следовательно, $\lambda x \in (\operatorname{int} A) \cap (\operatorname{int} B)$). Поэтому

$$x^* \in K_{A \cap B}(\hat{x}) \ \Leftrightarrow \ \sup_{x \in A \cap B} x^*(x) = x^*(\hat{x}) \ \Leftrightarrow \ \sup_{x \in \overline{A} \cap \overline{B}} x^*(x) = x^*(\hat{x}).$$

Последнее условие равносильно тому, что $\hat{x}$ — точка минимума в задаче

$$-x^*(x) \to \inf, \ p_A(x) - 1 \leq 0, \ p_B(x) - 1 \leq 0.$$

В этой задаче снова выполнено условие Слейтера, при этом $p_A(\hat{x}) = p_B(\hat{x}) = 1$. Применяя теорему Куна–Таккера, получаем, что $\hat{x}$ — решение задачи тогда и только тогда, когда существуют $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ такие, что $\hat{x}$ — точка минимума $-x^*(x) + \lambda_1 p_A(x) + \lambda_2 p_B(x)$. Записав условие минимума через субдифференциалы и применив теорему Моро–Рокафеллара, получаем, что

$$K_{A \cap B}(\hat{x}) = \{\lambda_1 \partial p_A(\hat{x}) + \lambda_2 \partial p_B(\hat{x}) : \ \lambda_1 \geq 0, \ \lambda_2 \geq 0\}.$$

Остается заметить, что

$$\{\lambda_1 \partial p_A(\hat{x}) + \lambda_2 \partial p_B(\hat{x}) : \ \lambda_1 \geq 0, \ \lambda_2 \geq 0\} =$$

$$= \{\lambda \partial p_A(\hat{x}) : \ \lambda \geq 0\} + \{\lambda \partial p_B(\hat{x}) : \ \lambda \geq 0\} =$$

$$= \operatorname{conv}\left(\{\lambda \partial p_A(\hat{x}) : \ \lambda \geq 0\} \cup \{\lambda \partial p_B(\hat{x}) : \ \lambda \geq 0\}\right)$$

и воспользоваться (13), (14). Теорема доказана. □

Замечание. Если $\hat{x}$ — граничная для A, но внутренняя для B, то $K_{A \cap B}(\hat{x}) = K_A(\hat{x})$ (в этом случае $p_B(\hat{x}) < 1$ и в теореме Куна–Таккера будет $\lambda_2 = 0$).

Задача 19. Построить пример двух выпуклых замкнутых множеств A и B на плоскости таких, что $\operatorname{int}(A \cap B) = \varnothing$, $\operatorname{int} A \neq \varnothing$, $\operatorname{int} B \neq \varnothing$ и в некоторой точке $\hat{x}$, являющейся граничной для A и B, выполнено $K_{A \cap B}(\hat{x}) \neq K_A(\hat{x}) + K_B(\hat{x})$.

10 Теорема Дубовицкого–Милютина

Теорема 16. (Дубовицкий–Милютин). *Пусть X — нормированное пространство, $f_1,\ f_2 : X \to \mathbb{R} \cup \{+\infty\}$ — выпуклые функции, $f_1(x_0) < +\infty$, $f_2(x_0) < +\infty$, $f_1,\ f_2$ непрерывны в точке x_0. Тогда*

$$\partial \max\{f_1,\ f_2\}(x_0) = \begin{cases} \partial f_1(x_0), & \text{если } f_1(x_0) > f_2(x_0), \\ \partial f_2(x_0), & \text{если } f_2(x_0) > f_1(x_0), \\ \operatorname{conv}(\partial f_1(x_0) \cup \partial f_2(x_0)), & \text{если } f_1(x_0) = f_2(x_0). \end{cases}$$

Доказательство. Обозначим $f(x) = \max\{f_1(x),\ f_2(x)\}$.

Случай $f_1(x_0) > f_2(x_0)$**.** Докажем, что $\partial f_1(x_0) \subset \partial f(x_0)$. Пусть $x^* \in \partial f_1(x_0)$. Тогда для любого $x \in X$ выполнено $f_1(x) \geq f_1(x_0) + x^*(x - x_0)$. Так как $f(x_0) = f_1(x_0)$, то $f(x) \geq f(x_0) + x^*(x - x_0)$, т.е. $x^* \in \partial f(x_0)$.

Докажем, что $\partial f(x_0) \subset \partial f_1(x_0)$. Пусть $x^* \in \partial f(x_0)$. Тогда для любого $x \in X$ выполнено $f(x) \geq f(x_0) + x^*(x - x_0)$. Так как $f_1,\ f_2$ непрерывны, то $f(x) = f_1(x)$ в некоторой окрестности U точки x_0, так что

$$f_1(x) \geq f_1(x_0) + x^*(x - x_0), \quad x \in U. \tag{15}$$

Покажем, что это неравенство выполнено для всех $x \in X$. В самом деле, пусть $f_1(x) < f_1(x_0) + x^*(x - x_0)$ для некоторого $x \in X$. Тогда для любого $t \in (0,\ 1)$ получаем

$$f_1((1-t)x_0 + tx) \leq (1-t)f_1(x_0) + tf_1(x) <$$

$$< (1-t)f_1(x_0) + tf_1(x_0) + tx^*(x - x_0) = f_1(x_0) + x^*(t(x - x_0)).$$

Но $(1-t)x_0 + tx \in U$ при малых t — противоречие с (15).

Случай $f_2(x_0) > f_1(x_0)$ аналогичный.

Случай $f_1(x_0) = f_2(x_0)$**.** Покажем, что $\operatorname{conv}(\partial f_1(x_0) \cup \partial f_2(x_0)) \subset \partial f(x_0)$. В самом деле, пусть $x_1^* \in \partial f_1(x_0)$, $x_2^* \in \partial f_2(x_0)$, $\lambda \in [0,\ 1]$. Тогда для любого $x \in X$

$$f(x) \geq f_1(x) \geq f_1(x_0) + x_1^*(x - x_0) = f(x_0) + x_1^*(x - x_0),$$

$$f(x) \geq f_2(x) \geq f_2(x_0) + x_2^*(x - x_0) = f(x_0) + x_2^*(x - x_0).$$

Первое неравенство домножим на $1 - \lambda$, второе — на λ, затем их сложим. Получим

$$f(x) \geq f(x_0) + ((1-\lambda)x_1^* + \lambda x_2^*)(x - x_0),$$

то есть $(1-\lambda)x_1^* + \lambda x_2^* \in \partial f(x_0)$.

Теперь докажем обратное включение. Пусть $x^* \in \partial f(x_0)$. Положим $g_1(x) = f_1(x) - f_1(x_0) - x^*(x - x_0)$, $g_2(x) = f_2(x) - f_2(x_0) - x^*(x - x_0)$, $g(x) = \max\{g_1(x), g_2(x)\}$. Тогда функции g_1 и g_2 выпуклы, $g(x) = f(x) - f(x_0) - x^*(x - x_0)$, $g_1(x_0) = g_2(x_0) = 0$,

$$0 \in \partial g(x_0). \tag{16}$$

Если мы покажем, что

$$0 \in \operatorname{conv}(\partial g_1(x_0) \cup \partial g_2(x_0)), \tag{17}$$

то получим

$$x^* \in x^* + \operatorname{conv}(\partial g_1(x_0) \cup \partial g_2(x_0)) =$$
$$= \operatorname{conv}((x^* + \partial g_1(x_0)) \cup (x^* + \partial g_2(x_0))) = \operatorname{conv}(\partial f_1(x_0) \cup \partial f_2(x_0)).$$

Докажем (17). Рассмотрим две задачи выпуклого программирования:

$$\begin{cases} g_1(x) \to \inf, \\ g_2(x) \le 0, \end{cases}$$

$$\begin{cases} g_2(x) \to \inf, \\ g_1(x) \le 0. \end{cases}$$

Покажем, что x_0 является решением одной из этих задач. В самом деле, $g_1(x_0) = g_2(x_0) = 0$, то есть x_0 допустимая. Если x_0 не является точкой минимума ни в одной из задач, то существуют точки x_1, $x_2 \in X$ такие, что $g_1(x_1) < 0$, $g_2(x_1) \le 0$, $g_2(x_2) < 0$, $g_1(x_2) \le 0$. Пусть $\tilde{x} = \frac{x_1+x_2}{2}$. Тогда $g_i(\tilde{x}) \le \frac{1}{2}g_i(x_1) + \frac{1}{2}g_i(x_2) < 0$, $i = 1, 2$. Значит, $g(\tilde{x}) < 0$. Поэтому x_0 не является точкой минимума функции g, так что $0 \notin \partial g(x_0)$ — противоречие с (16).

Пусть x_0 — решение первой задачи (случай второй задачи рассматривается аналогично). По теореме Куна–Таккера, существуют числа $\lambda_1 \ge 0$, $\lambda_2 \ge 0$, одновременно не равные 0, такие, что x_0 — точка минимума функции $\lambda_1 g_1(x) + \lambda_2 g_2(x)$. Значит, $0 \in \partial(\lambda_1 g_1 + \lambda_2 g_2)(x_0)$. Так как функции g_i непрерывны в x_0, то по теореме Моро–Рокафеллара $0 \in \lambda_1 \partial g_1(x_0) + \lambda_2 \partial g_2(x_0)$, т.е. $0 = \lambda_1 x_1^* + \lambda_2 x_2^*$, $x_1^* \in \partial g_1(x_0)$, $x_2^* \in \partial g_2(x_0)$. Значит, $0 = \frac{\lambda_1}{\lambda_1+\lambda_2} x_1^* + \frac{\lambda_2}{\lambda_1+\lambda_2} x_2^* \in \operatorname{conv}(\partial g_1(x_0) \cup \partial g_2(x_0))$. $\square$

Субдифференциал максимума конечного множества функций вычисляется по индукции.

Рассмотрим пример на конечномерном пространстве.

Пусть $f_1, \ldots, f_m : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ — выпуклые функции, $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$. Пусть x_0 — точка минимума f, функции f_i непрерывны в точке x_0, $1 \le i \le m$. Тогда $0 \in \partial f(x_0)$. Пусть $I = \{i \in \overline{1, m} : f_i(x_0) = f(x_0)\}$. По теореме Дубовицкого–Милютина, $0 \in \operatorname{conv}(\cup_{i \in I} \partial f_i(x_0))$. По теореме Каратеодори,

$$0 = \sum_{j=1}^{d+1} \lambda_j x_j^*,$$

$$x_j^* \in \cup_{i \in I} \partial f_i(x_0),\ \lambda_j \ge 0,\ 1 \le j \le d+1,\ \sum_{j=1}^{d+1} \lambda_j = 1.$$

Значит, существует множество $J \subset I$ такое, что $\operatorname{card} J \le d+1$, $0 \in \operatorname{conv}(\cup_{i \in J} \partial f_i(x_0)) = \partial(\max_{i \in J} f_i)(x_0)$ (снова по теореме Дубовицкого–Милютина). Значит, x_0 — точка минимума функции $\max_{i \in J} f_i$.

Вывод. Существует множество $J \subset \{1, \ldots, m\}$, $\operatorname{card} J \le d+1$, такое, что $\min_{x \in \mathbb{R}^d} \max_{1 \le i \le m} f_i(x) = \min_{x \in \mathbb{R}^d} \max_{i \in J} f_i(x)$.

Это частный случай теоремы об очистке. В общей форме мы эту теорему докажем позже.

11 Теоремы Радона, Хелли и Дворецкого

Теорема 17. (Радон). *Пусть $x_1, \ldots, x_n \in \mathbb{R}^d$, $n \ge d+2$. Тогда существует разбиение множества $\{1, \ldots, n\}$ на непустые непересекающиеся подмножества I и J такие, что*

$$\operatorname{conv}\{x_i\}_{i \in I} \cap \operatorname{conv}\{x_j\}_{j \in J} \ne \varnothing.$$

Доказательство. Так как $n \ge d+2$, то точки $x_1, \ldots, x_n$ аффинно зависимы. Значит, существуют числа $\lambda_1, \ldots, \lambda_n$, одновременно не равные 0, такие, что $\sum\limits_{j=1}^{n} \lambda_j = 0$, $\sum\limits_{j=1}^{n} \lambda_j x_j = 0$. Пусть $I = \{j \in \overline{1, n} :\ \lambda_j \ge 0\}$, $J = \{j \in \overline{1, n} :\ \lambda_j < 0\}$. Тогда $I \sqcup J = \{1, \ldots, n\}$, $I \ne \varnothing$, $J \ne \varnothing$. Имеем:

$$\sum_{i \in I} \lambda_i = \sum_{j \in J} |\lambda_j| > 0, \quad \sum_{i \in I} \lambda_i x_i = \sum_{j \in J} |\lambda_j| x_j.$$

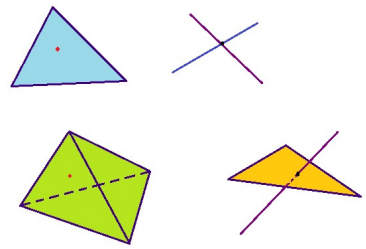

Рис. 25: Примеры: $d = 2$, $d = 3$

Значит,

$$\sum_{i\in I} \frac{\lambda_i}{\sum\limits_{k\in I} \lambda_k} x_i = \sum_{j\in J} \frac{|\lambda_j|}{\sum\limits_{k\in J} |\lambda_k|} x_j.$$

Левая часть равенства принадлежит $\operatorname{conv}\{x_i\}_{i\in I}$, правая часть принадлежит $\operatorname{conv}\{x_j\}_{j\in J}$. $\square$

Теорема 18. (Хелли). *а) Пусть $A_1, \ldots, A_n \in \mathbb{R}^d$ — выпуклые множества. Если $\cap_{i\in I} A_i \neq \varnothing$ для любого $I \subset \{1, \ldots, n\}$ такого, что $\operatorname{card} I = d+1$, то $\cap_{1\le i\le n} A_i \neq \varnothing$.*

б) Пусть $A_\lambda \subset \mathbb{R}^d$ $(\lambda \in \Lambda)$ — выпуклые компакты. Если $\cap_{\lambda\in I} A_\lambda \neq \varnothing$ для любого $I \subset \Lambda$ такого, что $\operatorname{card} I = d+1$, то $\cap_{\lambda\in\Lambda} A_\lambda \neq \varnothing$.

Доказательство. а) Докажем следующее утверждение. Пусть $m \ge d+1$, $A_1, \ldots, A_{m+1} \subset \mathbb{R}^d$ — выпуклые множества. Пусть $\cap_{1\le i\le m+1,\, i\neq k} A_i \neq \varnothing$ для любого $k \in \{1, \ldots, m+1\}$. Тогда $\cap_{i=1}^{m+1} A_i \neq \varnothing$. Из этого утверждения пункт а) теоремы Хелли легко выводится индукцией по n.

Пусть $x_k \in \cap_{1\le i\le m+1,\, i\neq k} A_i$, $1 \le k \le m+1$. Так как $m+1 \ge d+2$, то, по теореме Радона, существует разбиение $\{1, \ldots, m+1\} = I \sqcup J$, $I \neq \varnothing$, $J \neq \varnothing$, такое, что $\operatorname{conv}\{x_i\}_{i\in I} \cap \operatorname{conv}\{x_i\}_{i\in J} \neq \varnothing$.

Пусть $\hat{x} \in \operatorname{conv}\{x_i\}_{i\in I} \cap \operatorname{conv}\{x_i\}_{i\in J}$. Покажем, что $\hat{x} \in \cap_{k=1}^{m+1} A_k$. В самом деле, пусть $k \in J$. Покажем, что $\hat{x} \in A_k$. Для этого проверим, что $\operatorname{conv}\{x_i\}_{i\in I} \subset A_k$. Так как A_k выпукло, то достаточно показать, что $x_i \in A_k$ для любого $i \in I$. Имеем: $x_i \in \cap_{1\le j\le m+1,\, j\neq i} A_j$. Поскольку $k \notin I$, получаем, что $k \in \{1, \ldots, m+1\}\backslash\{i\}$, так что $\cap_{1\le j\le m+1,\, j\neq i} A_j \subset A_k$. Аналогично разбирается случай $k \in I$.

б) В силу пункта а), для любого конечного подмножества $I \subset \Lambda$ множество $\cap_{\lambda\in I} A_\lambda$ непусто. Так как множества A_λ компактны, то $\cap_{\lambda\in\Lambda} A_\lambda \neq \varnothing$. (В самом деле, пусть $\lambda_0 \in \Lambda$. Положим $B_\lambda = A_{\lambda_0} \cap A_\lambda$. Тогда $\cap_{\lambda\in I} B_\lambda \neq$

$\varnothing$ для любого конечного множества $I \subset \Lambda$. Если $\cap_{\lambda\in\Lambda} A_\lambda = \varnothing$, то $\cap_{\lambda\in\Lambda} B_\lambda = \varnothing$ и $A_{\lambda_0} = \cup_{\lambda\in\Lambda}(A_{\lambda_0}\backslash B_\lambda)$. Множества $A_{\lambda_0}\backslash B_\lambda$ открыты в A_{λ_0}, поэтому существует конечное множество $I \subset \Lambda$ такое, что $A_{\lambda_0} = \cup_{\lambda\in I}(A_{\lambda_0}\backslash B_\lambda)$, т.е. $\cap_{\lambda\in I} B_\lambda = \varnothing$ — противоречие.) $\square$

Теорема 19. (Дворецкий). *Пусть $A \subset \mathbb{R}^d$ — выпуклый компакт с непустой внутренностью. Тогда существует точка $a \in A$ такая, что для любой прямой l, проходящей через a, выполнено $\frac{\|b-c\|}{\|b-a\|} \le d+1$, где b, c — концы отрезка $l \cap A$.*

Доказательство. Для каждой точки $x \in A$ положим $A_x = \frac{x}{d+1} + \frac{d}{d+1}A$. Покажем, что $\cap_{x\in A} A_x \ne \varnothing$. По теореме Хелли, достаточно проверить, что $\cap_{i=1}^{d+1} A_{x_i} \ne \varnothing$ для любых $x_1, \ldots, x_{d+1} \in A$. Покажем, что $\sum\limits_{i=1}^{d+1} \frac{x_i}{d+1} \in \cap_{k=1}^{d+1} A_{x_k}$. В самом деле, для любого $k \in \{1, \ldots, d+1\}$

$$\sum_{i=1}^{d+1} \frac{x_i}{d+1} = \frac{x_k}{d+1} + \frac{d}{d+1}\sum_{i\ne k}\frac{x_i}{d} \in \frac{x_k}{d+1} + \frac{d}{d+1}A = A_{x_k}.$$

Пусть $a \in \cap_{x\in A} A_x$. Покажем, что это искомая точка. Пусть l — прямая, содержащая точку a, $A \cap l = [b, c]$. Тогда $a \in A_c = \frac{c}{d+1} + \frac{d}{d+1}A$. Значит, $a = \frac{c}{d+1} + \frac{d}{d+1}((1-\lambda)b + \lambda c)$, где $\lambda \in [0, 1]$. Тогда

$$a - b = \frac{c-b}{d+1} - \frac{d}{d+1}b + \frac{d}{d+1}((1-\lambda)b + \lambda c) = \left(\frac{1}{d+1} + \frac{\lambda d}{d+1}\right)(c-b),$$

$$\|a-b\| = \left(\frac{1}{d+1} + \frac{\lambda d}{d+1}\right)\|c-b\| \ge \frac{\|c-b\|}{d+1}.$$

Теорема доказана. $\square$

12 Теорема об очистке

Сначала докажем теорему об очистке для конечного множества функций.

Теорема 20. *Пусть $f_1, \ldots, f_n : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ — выпуклые функции. Тогда существует множество $I \subset \{1, \ldots, n\}$, $\operatorname{card} I \le d+1$, такое, что*

$$\inf_{x\in\mathbb{R}^d} \max_{1\le i\le n} f_i(x) = \inf_{x\in\mathbb{R}^d} \max_{i\in I} f_i(x). \tag{18}$$

Напомним, что в случае, когда минимум в левой части достигается, мы уже доказывали это утверждение, применив теорему Дубовицкого–Милютина и теорему Каратеодори. Здесь мы будем использовать теорему Хелли.

Доказательство. Пусть

$$m = \inf_{x\in\mathbb{R}^d} \max_{1\le i\le n} f_i(x).$$

Тогда $m \ge \inf_{x\in\mathbb{R}^d} \max_{i\in I} f_i(x)$. Если $m = -\infty$, то утверждение теоремы тривиально.

Пусть $m > -\infty$. Будем проводить рассуждения в случае, когда $m < +\infty$ (для $m = +\infty$ рассуждения проводятся так же).

Положим $A_i(\varepsilon) = \{x \in \mathbb{R}^d : \ f_i(x) \le m - \varepsilon\}$, $1 \le i \le n$. Если для любого $J \subset \{1, \ldots, n\}$ такого, что $\operatorname{card} J \le d+1$, выполнено $\cap_{j\in J} A_j(\varepsilon) \ne \varnothing$, то по теореме Хелли $\cap_{1\le j\le n} A_j(\varepsilon) \ne \varnothing$. Это означает, что существует точка $\hat{x} \in \mathbb{R}^d$ такая, что $f_j(\hat{x}) \le m-\varepsilon$ для любого $j \in \{1, \ldots, n\}$. Поэтому $\max_{1\le j\le n} f_j(\hat{x}) \le m - \varepsilon$, что противоречит с определением m.

Итак, для любого $\varepsilon > 0$ существует множество $I(\varepsilon) \subset \{1, \ldots, n\}$ такое, что $\operatorname{card} I(\varepsilon) \le d+1$ и $\cap_{i\in I(\varepsilon)} A_i(\varepsilon) = \varnothing$. Последнее условие означает, что для любого $x \in \mathbb{R}^n$ существует $i \in I(\varepsilon)$ такое, что $f_i(x) > m - \varepsilon$, т.е. $\max_{i\in I(\varepsilon)} f_i(x) > m - \varepsilon$. Возьмем $\varepsilon_k \underset{k\to\infty}{\to} 0$. Существует подпоследовательность $\{k_j\}_{j\in\mathbb{N}}$ такая, что для любого $j \in \mathbb{N}$ множества $I(\varepsilon_{k_j})$ одинаковы: $I(\varepsilon_{k_j}) = I$. Получаем: для любого $x \in \mathbb{R}^d$ выполнено неравенство $\max_{i\in I} f_i(x) > m - \varepsilon_{k_j}$. Устремив $j \to \infty$, получаем $\max_{i\in I} f_i(x) \ge m$ для любого $x \in \mathbb{R}^d$. Значит, $\inf_{x\in\mathbb{R}^d} \max_{i\in I} f_i(x) \ge m$. □

Определение 15. *Пусть S — топологическое пространство, $f : S \to \mathbb{R}\cup\{+\infty\}$. Функция f называется полунепрерывной снизу, если для любого $c \in \mathbb{R}$ множество $\{x \in S : \ f(x) \le c\}$ замкнуто.*

Теорема 21. *Пусть T — метрический компакт, $f : T\times\mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$. Предположим, что*

1. *для любого $t \in T$ функция $f(t, \cdot)$ выпукла и полунепрерывна снизу;*
2. *для любого $x \in \mathbb{R}^d$ функция $f(\cdot, x)$ непрерывна.*

Тогда существует множество $I \subset T$, $\operatorname{card} I = d + 1$, такое, что

$$\inf_{x\in\mathbb{R}^d} \max_{t\in T} f(t, x) = \inf_{x\in\mathbb{R}^d} \max_{t\in I} f(t, x). \tag{19}$$

Доказательство. Положим $m = \inf_{x\in\mathbb{R}^d} \max_{t\in T} f(t,\, x)$.

Пусть для любого $\varepsilon > 0$ существует $\overline{t}(\varepsilon) = (t_1(\varepsilon),\, \ldots,\, t_{d+1}(\varepsilon)) \in T^{d+1}$ такое, что

$$\inf_{x\in\mathbb{R}^d} \max_{1\le i\le d+1} f(t_i(\varepsilon),\, x) \ge m - \varepsilon. \tag{20}$$

Так как T компактно, то существует последовательность $\varepsilon_k \underset{k\to\infty}{\to} +0$ такая, что $\{t_i(\varepsilon_k)\}_{k\in\mathbb{N}}$ сходится для любого $i \in \{1,\, \ldots,\, d+1\}$. Пусть $t_i = \lim\limits_{k\to\infty} t_i(\varepsilon_k)$, $I = \{t_1,\, \ldots,\, t_{d+1}\}$. Покажем, что выполнено (19).

Зафиксируем $x \in \mathbb{R}^d$. В силу (20), существует последовательность $\{i_k\}_{k\in\mathbb{N}} \subset \{1,\, \ldots,\, d+1\}$ такая, что $f(t_{i_k}(\varepsilon_k),\, x) \ge m - \varepsilon_k$. Существует подпоследовательность $\{k_j\}_{j\in\mathbb{N}}$ такая, что $i_{k_j} = i$ — постоянная последовательность. Тогда $f(t_i(\varepsilon_{k_j}),\, x) \ge m - \varepsilon_{k_j}$. Так как f непрерывна по t, то $f(t_i,\, x) \ge m$, т.е. $\max_{1\le j\le d+1} f(t_j,\, x) \ge m$.

Пусть теперь существует $\varepsilon > 0$ такое, что для любого $\overline{t} = (t_1,\, \ldots,\, t_{d+1}) \in T^{d+1}$ существует точка $x(\overline{t}) \in \mathbb{R}^d$ такая, что

$$\max_{1\le i\le d+1} f(t_i,\, x(\overline{t})) < m - \varepsilon. \tag{21}$$

Положим $A(t) = \{x \in \mathbb{R}^d :\ f(t,\, x) \le m - \varepsilon\}$. Это множество выпукло и замкнуто (так как f выпукла и полунепрерывна снизу по x). Если бы множества $A(t)$ были компактны, то можно было бы повторить рассуждения из предыдущей теоремы. Но $A(t)$ могут быть некомпактными.

Пусть $\tilde{A}$ — выпуклый компакт (его построим позже). Положим $\tilde{A}(t) = A(t) \cap \tilde{A}$. Если $\cap_{t\in I}\tilde{A}(t) \ne \varnothing$ для любого $I \subset T$, $\operatorname{card} I = d+1$, то по теореме Хелли $\cap_{t\in T}\tilde{A}(t) \ne \varnothing$, т.е. существует точка $\hat{x} \in \tilde{A}$ такая, что $f(t,\, \hat{x}) \le m - \varepsilon$ для любого $t \in T$. Значит, $\inf_{x\in\mathbb{R}^d} \max_{t\in T} f(t,\, x) \le m - \varepsilon$ — противоречие.

Итак, существует $I = \{t_1,\, \ldots,\, t_{d+1}\} \subset T$ такое, что $\cap_{i=1}^{d+1}\tilde{A}(t_i) = \varnothing$. Значит, для любого $x \in \tilde{A}$ существует $j \in \{1,\, \ldots,\, d+1\}$ такое, что $x \notin A(t_j)$, т.е. $f(t_j,\, x) > m - \varepsilon$. Получаем:

$$\exists\{t_1,\, \ldots,\, t_{d+1}\} \subset T :\ \forall x \in \tilde{A}\ \max_{1\le i\le d+1} f(t_i,\, x) > m - \varepsilon. \tag{22}$$

Построим выпуклый компакт $\tilde{A}$ такой, что для любого набора точек $\{t_1,\, \ldots,\, t_{d+1}\}$ существует точка $x \in \tilde{A}$ такая, что

$$\max_{1\le i\le d+1} f(t_i,\, x) \le m - \varepsilon. \tag{23}$$

Тогда будет противоречие с (22).

Пусть $\bar{t} = (t_1, \ldots, t_{d+1}) \in T^{d+1}$. Напомним, что $x(\bar{t})$ определяется так, чтобы выполнялось (21). Положим

$$V(\bar{t}) = \{\bar{s} = (s_1, \ldots, s_{d+1}) : \max_{1 \le i \le d+1} f(s_i, x(\bar{t})) < m - \varepsilon\}.$$

Это множество открыто в T^{d+1}. Кроме того, $\{V(\bar{t})\}_{\bar{t} \in T^{d+1}}$ образует покрытие компакта T^{d+1}. Выберем конечное подпокрытие $\{V(\bar{t}^{(l)})\}_{l=1}^N$. Положим $\tilde{A} = \operatorname{conv}\{x(\bar{t}^{(l)})\}_{l=1}^N$. В качестве точки $x \in \tilde{A}$, для которой выполнено (23), берем $x(\bar{t}^{(l)})$, где l таково, что $\bar{t} \in V(\bar{t}^{(l)})$. Теорема доказана. $\square$

13 Теорема Чебышева об альтернансе

Пусть $[a, b] \subset \mathbb{R}$. Обозначим через $\mathcal{P}_n[a, b]$ пространство полиномов степени не выше n. Это $n+1$-мерное подпространство в $C[a, b]$. Напомним, что на $C[a, b]$ рассматривается равномерная норма $\|x\| = \max_{t \in [a, b]} |x(t)|$.

Предложение 30. *Для любой функции $x \in C[a, b]$ существует полином $\hat{p} \in \mathcal{P}_n[a, b]$ такой, что*

$$\|x - \hat{p}\|_{C[a, b]} = \operatorname{dist}(x, \mathcal{P}_n[a, b]).$$

Доказательство. Это следует из того, что $\mathcal{P}_n[a, b]$ конечномерно, а функция $p \mapsto \|x - p\|_{C[a, b]}$ непрерывна и стремится к $+\infty$ при $\|p\|_{C[a, b]} \to +\infty$. $\square$

Полином $\hat{p}$ называется ближайшим к функции x.

Теорема 22. *Полином $\hat{p} \in \mathcal{P}_n[a, b]$ является ближайшим к функции $x \in C[a, b]$ тогда и только тогда, когда существуют точки $a \le t_1 < t_2 < \cdots < t_{n+2} \le b$ такие, что $|\hat{p}(t_i) - x(t_i)| = \|x - \hat{p}\|_{C[a, b]}$, $1 \le i \le n+2$, и знаки $\hat{p}(t_i) - x(t_i)$ чередуются: $\hat{p}(t_{i+1}) - x(t_{i+1}) = x(t_i) - \hat{p}(t_i)$, $1 \le i \le n+1$.*

Точки $\{t_i\}_{i=1}^{n+2}$ называются точками альтернанса.

Доказательство. $\Rightarrow$. Пусть $\hat{p}$ — ближайший к x полином, т.е.

$$\alpha := \min_{p \in \mathcal{P}_n[a, b]} \max_{t \in [a, b]} |x(t) - p(t)| = \max_{t \in [a, b]} |x(t) - \hat{p}(t)|.$$

Достаточно рассмотреть случай, когда $\alpha > 0$. По теореме об очистке, существуют точки $a \le t_1 < t_2 < \dots < t_{n+2} \le b$ такие, что

$$\min_{p \in \mathcal{P}_n[a,\,b]} \max_{1 \le i \le n+2} |x(t_i) - p(t_i)| = \alpha.$$

Пусть эти точки не образуют альтернанс, т.е. либо существует $i \in \{1, \dots, n+2\}$ такое, что $|x(t_i) - \hat{p}(t_i)| < \alpha$, либо существует $i \in \{1, \dots, n+1\}$ такое, что $x(t_i) - \hat{p}(t_i) = x(t_{i+1}) - \hat{p}(t_{i+1})$. Покажем, что существует полином $\tilde{p} \in \mathcal{P}_n[a,\,b]$ такой, что

$$\max_{1 \le i \le n+2} |x(t_i) - \tilde{p}(t_i)| < \alpha. \tag{24}$$

Пусть $I = \{i \in \overline{1,\,n+2} : \ |x(t_i) - \hat{p}(t_i)| = \alpha\}$. Если $I = \varnothing$, то достаточно взять $\tilde{p} = \hat{p}$. Если $I \ne \varnothing$ и знаки $x(t_i) - \hat{p}(t_i)$ одинаковы для всех $i \in I$, то в качестве $\tilde{p}$ берем сдвиг $\hat{p}$ на константу.

Пусть $I \ne \varnothing$ и знаки $x(t_i) - \hat{p}(t_i)$ не одинаковые. Точки множества I упорядочиваются как $i_1 < i_2 < \dots < i_m$, $m \le n+2$. Без ограничения общности, $x(t_{i_1}) - \hat{p}(t_{i_1}) > 0$. Разобьем $\overline{1,\,m}$ на целочисленные отрезки в соответствии со знаком $x(t_i) - \hat{p}(t_i)$:

$$x(t_{i_1}) - \hat{p}(t_{i_1}) > 0, \dots, x(t_{i_{s_1}}) - \hat{p}(t_{i_{s_1}}) > 0,$$

$$x(t_{i_{s_1+1}}) - \hat{p}(t_{i_{s_1+1}}) < 0, \dots, x(t_{i_{s_2}}) - \hat{p}(t_{i_{s_2}}) < 0,$$
$$x(t_{i_{s_2+1}}) - \hat{p}(t_{i_{s_2+1}}) > 0, \dots, x(t_{i_{s_3}}) - \hat{p}(t_{i_{s_3}}) > 0,$$

$$\dots$$

$$(-1)^l (x(t_{i_{s_l+1}}) - \hat{p}(t_{i_{s_l+1}})) > 0, \dots, (-1)^l (x(t_{i_m}) - \hat{p}(t_{i_m})) > 0$$

(полагаем $s_0 = 1$, $s_{l+1} = m$). Тогда $l \le n$.

Возьмем точки $\tau_k \in (t_{i_{s_k}},\, t_{i_{s_k+1}})$, $1 \le k \le l$. Положим

$$q(t) = \prod_{k=1}^{l} (\tau_k - t).$$

Это полином степени не выше n, при этом $q(t_{i_j}) > 0$ при $s_k \le j \le s_{k+1}$ для четных k и $q(t_{i_j}) < 0$ при $s_k \le j \le s_{k+1}$ для нечетных k.

Положим $\tilde{p}(t) = \hat{p}(t) + \varepsilon q(t)$, где $\varepsilon > 0$ достаточно мало. Тогда

$$|x(t_{i_j}) - \tilde{p}(t_{i_j})| = |x(t_{i_j}) - \hat{p}(t_{i_j}) - \varepsilon q(t_{i_j})| = \alpha - \varepsilon |q(t_{i_j})| < \alpha,$$

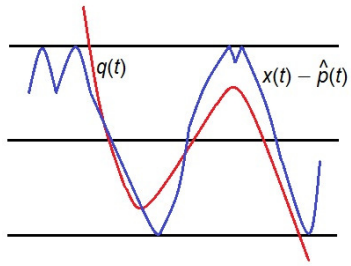

Рис. 26: Функции $x - \hat{p}$ и q.

если ε мало; для $i \in \{1, \ldots, n+2\} \setminus I$ выполнено $|x(t_i) - \hat{p}(t_i)| < \alpha$; поэтому $|x(t_i) - \tilde{p}(t_i)| < \alpha$ при малых $\varepsilon > 0$. Значит, выполнено (24).

$\Leftarrow$. Пусть $t_1 < \cdots < t_{n+2}$ — точки альтернанса у полинома $\hat{p}$. Предположим, что существует полином $p \in \mathcal{P}_n[a, b]$ такой, что $\|x - p\|_{C[a,b]} < \|x - \hat{p}\|_{C[a,b]} =: \alpha$. Тогда знаки $p(t_i) - \hat{p}(t_i) = (x(t_i) - \hat{p}(t_i)) - (x(t_i) - p(t_i))$ чередуются. Значит, существуют точки $\tau_i \in (t_i, t_{i+1})$, $1 \le i \le n+1$, такие, что $p(\tau_i) - \hat{p}(\tau_i) = 0$. Следовательно, полином $p - \hat{p} \in \mathcal{P}_n[a, b]$ имеет $n+1$ нулей и поэтому $p - \hat{p} = 0$ — противоречие. $\square$

14 Поляры

Обозначим через $\langle \cdot, \cdot \rangle$ стандартное скалярное произведение на $\mathbb{R}^n$.

Определение 16. *Пусть $A \subset \mathbb{R}^n$. Полярой множества A называется*

$$A^\circ = \{x \in \mathbb{R}^n : \ \forall y \in A \ \langle x, y \rangle \le 1\}.$$

Замечание. $\varnothing^\circ = \mathbb{R}^n$.

Предложение 31. *Поляра является выпуклым замкнутым множеством, содержащим* 0.

Доказательство. Так как $\langle 0, y \rangle \le 1$, то $0 \in A^\circ$.

Если $\langle x_1, y \rangle \le 1$, $\langle x_2, y \rangle \le 1$ для любого $y \in A$, то для любого $\lambda \in [0, 1]$

$$\langle (1 - \lambda) x_1 + \lambda x_2, y \rangle \le 1, y \in A.$$

Значит, A° выпукло.

Если $\langle x_m, y \rangle \le 1$ для любого $y \in A$, $m \in \mathbb{N}$, $x_m \underset{m \to \infty}{\to} x$, то $\langle x, y \rangle \le 1$ для любого $y \in A$. Значит, A° замкнуто. $\square$

Из определения поляры следует

Предложение 32. *Если $A \subset B$, то $A^\circ \supset B^\circ$.*

Предложение 33. *Пусть $B_R(0)$ — замкнутый евклидов шар радиуса R с центром в 0. Тогда $(B_R(0))^\circ = B_{1/R}(0)$.*

Доказательство. Имеем: $x \in (B_R(0))^\circ$ тогда и только тогда, когда $\langle x,\, y\rangle \le 1$ для любого $y \in B_R(0)$.

Пусть $x \in B_{1/R}(0)$. В силу неравенства Коши–Буняковского, $\langle x,\, y\rangle \le \|x\| \cdot \|y\| \le \frac{1}{R} \cdot R = 1$.

Пусть $x \notin B_{1/R}(0)$. Положим $y = R\frac{x}{\|x\|}$. Тогда $y \in B_R(0)$, но $\langle x,\, y\rangle = R\|x\| > 1$. □

Следствие 5. *Если A ограничено, то A° содержит окрестность 0. Если A содержит окрестность 0, то A° ограничено.*

Доказательство. Если $A \subset B_R(0)$, то $A^\circ \supset (B_R(0))^\circ = B_{1/R}(0)$.

Если $B_\delta(0) \subset A$, то $A^\circ \subset (B_\delta(0))^\circ = B_{1/\delta}(0)$. □

Предложение 34. *Выполнено равенство $A^\circ = (\overline{\operatorname{conv}}\,(A \cup \{0\}))^\circ$.*

Доказательство. По свойству поляр, $A^\circ \supset (\overline{\operatorname{conv}}\,(A \cup \{0\}))^\circ$. Докажем обратное включение.

Пусть $x \in A^\circ$. Так как $\langle x,\, 0\rangle = 0 < 1$, то $x \in (A \cup \{0\})^\circ$.

Пусть $x \in B^\circ$. Если $y = \sum\limits_{j=1}^{m} \lambda_j y_j$, $y_j \in B$, $\lambda_j \ge 0$, $1 \le j \le m$, $\sum\limits_{j=1}^{m} \lambda_j = 1$. Тогда

$$\langle x,\, y\rangle = \sum_{j=1}^{m} \lambda_j \langle x,\, y_j\rangle \le \sum_{j=1}^{m} \lambda_j = 1.$$

Значит, $x \in (\operatorname{conv} B)^\circ$.

Пусть $x \in C^\circ$. Если $y_m \in C$, $y_m \underset{m\to\infty}{\to} y$, то $\langle x,\, y\rangle = \lim\limits_{m\to\infty} \langle x,\, y_m\rangle \le 1$. Значит, $x \in \overline{C}^\circ$. □

Предложение 35. *Выполнено равенство $\left(\cup_{j=1}^k A_j\right)^\circ = \cap_{j=1}^k A_j^\circ$.*

Доказательство. Имеем

$$\left(\cup_{j=1}^k A_j\right)^\circ = \left\{x \in \mathbb{R}^n :\ \langle x,\, y\rangle \le 1 \quad \forall y \in \cup_{j=1}^k A_j\right\} =$$
$$= \{x \in \mathbb{R}^n :\ \langle x,\, y\rangle \le 1 \quad \forall y \in A_j,\ \forall j \in \overline{1,\, k}\} =$$
$$= \cap_{j=1}^k \{x \in \mathbb{R}^n :\ \langle x,\, y\rangle \le 1 \quad \forall y \in A_j\} = \cap_{j=1}^k A_j^\circ.$$

□

Определение 17. *Пусть $A \subset \mathbb{R}^n$. Биполярой множества A называется $A^{\circ\circ} = (A^\circ)^\circ$.*

Замечание. Из свойств поляры следует, что $A^{\circ\circ}$ выпукло, замкнуто и содержит 0.

Теорема 23. *Равенство $A = A^{\circ\circ}$ выполнено тогда и только тогда, когда A выпукло, замкнуто и содержит 0.*

Доказательство. В силу замечания, достаточно показать, что если A выпукло, замкнуто и содержит 0, то $A = A^{\circ\circ}$.

Из определения поляры следует, что $A \subset A^{\circ\circ}$. Если равенство не выполнено, то существует точка $\hat{x} \in A^{\circ\circ} \backslash A$. Так как A выпукло и замкнуто, то по теореме о строгой отделимости существует вектор $z \in \mathbb{R}^n$ такой, что

$$\langle z,\, \hat{x}\rangle > \sup_{x \in A} \langle z,\, x\rangle \geq 0$$

(последнее неравенство верно, так как $0 \in A$). Умножив z на подходящую положительную константу, можно считать, что $\sup_{x\in A}\langle z,\, x\rangle \leq 1$, $\langle z,\, \hat{x}\rangle > 1$. Из первого неравенства следует, что $z \in A^\circ$, а из второго — что $\hat{x} \notin A^{\circ\circ}$. Получили противоречие. □

Следствие 6. *Пусть $A \subset \mathbb{R}^n$. Тогда $A^{\circ\circ} = \overline{\mathrm{conv}}\,(A \cup \{0\})$.*

Доказательство. Используя равенство $A^\circ = (\overline{\mathrm{conv}}\,(A \cup \{0\}))^\circ$ и теорему о биполяре, получаем

$$A^{\circ\circ} = (\overline{\mathrm{conv}}\,(A \cup \{0\}))^{\circ\circ} = \overline{\mathrm{conv}}\,(A \cup \{0\}).$$

□

Предложение 36. *Пусть множества $A_1, \ldots, A_k \subset \mathbb{R}^n$ выпуклы, замкнуты и содержат 0. Тогда $\left(\cap_{j=1}^k A_j\right)^\circ = \overline{\mathrm{conv}}\left(\cup_{j=1}^k A_j^\circ\right)$.*

Доказательство. Имеем

$$\left(\overline{\mathrm{conv}}\left(\cup_{j=1}^k A_j^\circ\right)\right)^\circ = \left(\cup_{j=1}^k A_j^\circ\right)^\circ = \cap_{j=1}^k A_j^{\circ\circ} = \cap_{j=1}^k A_j$$

(здесь использовалось предложение 35 и теорема о биполяре); отсюда получаем

$$\overline{\mathrm{conv}}\left(\cup_{j=1}^k A_j^\circ\right) = \left(\overline{\mathrm{conv}}\left(\cup_{j=1}^k A_j^\circ\right)\right)^{\circ\circ} = \left(\cap_{j=1}^k A_j\right)^\circ.$$

□

Задача 20. Верно ли равенство $(A \cap B)^\circ = \overline{\operatorname{conv}}\,(A^\circ \cup B^\circ)$, если A, B выпуклы, замкнуты и имеют непустое пересечение?

Задача 21. Пусть $A = \{(x_1,\, x_2) \in \mathbb{R}^2 : (x_1 - a)^2 + x_2^2 \le R^2\}$, где $0 < a \le R$. Найти A°.

Задача 22. Пусть $A = \{(x_1,\, x_2) \in \mathbb{R}^2 : (x_1 - a)^2 + x_2^2 \le R^2\}$, где $a > R > 0$. Найти A°.

Задача 23. Найти поляру множеств $A = \{(x_1,\, x_2) \in \mathbb{R}^2 :\ x_2 \ge x_1^2 + c\}$, где $c \in \mathbb{R}$.

Определение 18. *Непустое множество $K \subset \mathbb{R}^n$ называется конусом, если*

1. *$\lambda x \in K$ для любых $x \in K$, $\lambda \ge 0$;*
2. *$x + y \in K$ для любых x, $y \in K$.*

Замечание. Конус является выпуклым множеством.

Например, внешний нормальный конус к множеству является конусом.

Пусть K — конус. Обозначим через $K^\perp$ внешний нормальный конус к K в точке 0.

Предложение 37. *Пусть K — конус. Тогда $K^\circ = K^\perp$.*

Доказательство. Докажем, что $K^\circ \subset K^\perp$. Пусть $\xi \in K^\circ$. Тогда $\langle \xi,\, x\rangle \le 1$ для любого $x \in K$. Если существует $x \in K$ такое, что $\langle \xi,\, x\rangle > 0$, то $\langle \xi,\, \lambda x\rangle \underset{\lambda \to +\infty}{\to} +\infty$. Так как $\lambda x \in K$ при $\lambda \ge 0$, то $\sup_{y \in K} \langle \xi,\, y\rangle = +\infty$. Значит, $\langle \xi,\, x\rangle \le 0$ для любого $x \in K$, при этом $\langle \xi,\, 0\rangle = 0$. Таким образом, $\xi \in K^\perp$.

Докажем обратное включение:

$$K^\perp = \{x \in \mathbb{R}^n :\ \langle x,\, y\rangle \le 0,\ y \in K\} \subset$$

$$\subset \{x \in \mathbb{R}^n :\ \langle x,\, y\rangle \le 1,\ y \in K\} = K^\circ.$$

□

Предложение 38. *Пусть K — замкнутый конус. Тогда*

$$(K^\perp)^\perp = K.$$

Доказательство. Так как K выпукло, замкнуто и содержит 0, то $K = K^{\circ\circ}$. Так как $K^{\perp}$ — конус, то $(K^{\perp})^{\circ} = (K^{\perp})^{\perp}$. Значит, $(K^{\perp})^{\perp} = (K^{\perp})^{\circ} = (K^{\circ})^{\circ} = K$. $\square$

Пусть $\|\cdot\|$ — какая-то норма на $\mathbb{R}^n$. Сопряженная норма определяется равенством $\|x\|^* = \max_{\|y\|\le 1}\langle x,\, y\rangle$. Из определения поляры следует

Предложение 39. *Выполнено равенство*

$$\{y:\ \|y\| \le 1\}^{\circ} = \{x:\ \|x\|^* \le 1\}.$$

Задача 24. Пусть $p(x)$ — выпуклая положительно-однородная функция (т.е. $p(\lambda x) = \lambda p(x)$ для любого $\lambda \ge 0$). Пусть

$$A = \{x \in \mathbb{R}^n : p(x) \le 1\}.$$

Показать, что $A^{\circ} = \operatorname{conv}(\partial p(0) \cup \{0\})$.

Пример. Пусть $1 < p < \infty$,

$$B = \left\{(x_1,\, \ldots,\, x_n):\ \sum_{i=1}^{n} |x_i|^p \le 1\right\}.$$

Тогда

$$B^{\circ} = \left\{(x_1,\, \ldots,\, x_n):\ \sum_{i=1}^{n} |x_i|^{p'} \le 1\right\},$$

где $\frac{1}{p} + \frac{1}{p'} = 1$.

Доказательство. Пусть $\sum\limits_{i=1}^{n} |x_i|^{p'} \le 1$. Если $(y_1,\, \ldots,\, y_n) \in B$, то в силу неравенства Гельдера

$$\sum_{i=1}^{n} x_i y_i \le \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |x_i|^{p'}\right)^{1/p'} \le 1,$$

так что $(x_1,\, \ldots,\, x_n) \in B^{\circ}$.

Пусть $\sum\limits_{i=1}^{n} |x_i|^{p'} > 1$. Так как неравенство Гельдера точное, то существует вектор $(y_1,\, \ldots,\, y_n) \ne 0$ такой, что

$$\sum_{i=1}^{n} x_i y_i = \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |x_i|^{p'}\right)^{1/p'}.$$

При этом можно считать, что $\sum\limits_{i=1}^{n} |y_i|^p = 1$. Тогда правая часть строго больше 1 и $(x_1, \ldots, x_n) \notin B^\circ$. $\square$

15 Выпуклые многогранные множества и многогранники

Определение 19. *Выпуклым многогранным множеством называется*

$$\{x \in \mathbb{R}^n : \ \langle a_j, x\rangle \le b_j, \ 1 \le j \le m\},$$

где $a_1, \ldots, a_m \in \mathbb{R}^n \backslash \{0\}$, $b_1, \ldots, b_m \in \mathbb{R}$.

Т.е. выпуклое многогранное множество — это пересечение конечного числа замкнутых полупространств. Далее слово "выпуклое" будем опускать.

Считаем, что $\mathbb{R}^n$ — тоже многогранное множество.

Определение 20. *Выпуклым многогранником называется* $\operatorname{conv}\{x_1, \ldots, x_m\}$, *где* $x_1, \ldots, x_m \in \mathbb{R}^n$.

Т.е. выпуклый многогранник — это выпуклая оболочка конечного числа точек. Ясно, что выпуклый многогранник компактен.

Предложение 40. *Полярой выпуклого многогранника является многогранное множество.*

Доказательство. Сначала найдем поляру одноточечного множества $\{\xi\}$. Если $\xi = 0$, то $\{\xi\}^\circ = \mathbb{R}^n$. Пусть $\xi \ne 0$. Тогда $\{\xi\}^\circ = \{y \in \mathbb{R}^n : \ \langle \xi, y\rangle \le 1\}$. Это замкнутое полупространство.

Пусть $M = \operatorname{conv}\{x_j\}_{j=1}^m$. В силу предложений 34, 35,

$$M^\circ = (\{x_j\}_{j=1}^n)^\circ = \cap_{j=1}^m \{x_j\}^\circ = \cap_{j=1}^m \{y \in \mathbb{R}^n : \ \langle x_j, y\rangle \le 1\}.$$

Это многогранное множество. $\square$

Отметим, что если $\xi \ne 0$, то

$$[0, \xi]^\circ = \{\xi\}^\circ = \{x \in \mathbb{R}^n : \ \langle \xi, x\rangle \le 1\}. \tag{25}$$

Предложение 41. *Пусть M — многогранное множество, $0 \in \operatorname{int} M$. Тогда M° — выпуклый многогранник.*

Доказательство. Если $M = \mathbb{R}^n$, то $M^\circ = \{0\}$. В остальных случаях $M = \cap_{j=1}^m \{x \in \mathbb{R}^n : \langle a_j, x\rangle \le b_j\}$, где $a_j \in \mathbb{R}^n \backslash \{0\}$. Так как $0 \in \operatorname{int} M$, то $b_j > 0$. В силу предложения 36,

$$M^\circ = \overline{\operatorname{conv}}\left(\cup_{j=1}^m \{x \in \mathbb{R}^n : \langle a_j, x\rangle \le b_j\}^\circ\right).$$

Пусть $a \in \mathbb{R}^n \backslash \{0\}$, $b > 0$. Тогда

$$\{x \in \mathbb{R}^n : \langle a, x\rangle \le b\}^\circ = \{x \in \mathbb{R}^n : \langle a/b, x\rangle \le 1\}^\circ \overset{(25)}{=} [0, a/b]^{\circ\circ} = [0, a/b].$$

Поэтому

$$M^\circ = \overline{\operatorname{conv}}(\cup_{j=1}^m [0, a_j/b_j]) = \operatorname{conv}\{0, a_1/b_1, \ldots, a_m/b_m\}.$$

Значит, M° — выпуклый многогранник. □

Заметим, что если многогранное множество ограничено и 0 — его внутренняя точка, то его полярой будет выпуклый многогранник, у которого 0 — внутренняя точка. Обратно, если у выпуклого многогранника 0 — внутренняя точка, то его поляра — ограниченное многогранное множество, у которого 0 — внутренняя точка.

Охарактеризуем крайние точки у многогранного множества.

Предложение 42. *Пусть $M = \{x \in \mathbb{R}^n : \langle a_j, x\rangle \le b_j,\ 1 \le j \le m\}$, $\hat{x} \in M$. Обозначим $I(\hat{x}) = \{j \in \overline{1, m} : \langle a_j, \hat{x}\rangle = b_j\}$. Тогда следующие условия эквивалентны:*

1. *$\hat{x}$ — крайняя точка M;*

2. *система равенств*

$$\langle a_j, x\rangle = b_j, \quad j \in I(\hat{x})$$

имеет единственное решение.

Доказательство. Пусть система имеет единственное решение, $\hat{x} = \frac{y+z}{2}$, $y, z \in M$. Для любого $j \in I(\hat{x})$ выполнено $b_j = \langle a_j, \hat{x}\rangle = \frac{1}{2}\langle a_j, y\rangle + \frac{1}{2}\langle a_j, z\rangle$, $\langle a_j, y\rangle \le b_j$, $\langle a_j, z\rangle \le b_j$. Значит, $\langle a_j, y\rangle = \langle a_j, z\rangle = b_j$ для любого $j \in I(\hat{x})$ и поэтому $y = z = \hat{x}$. Таким образом, $\hat{x}$ — крайняя точка.

Обратно, пусть $\hat{x}$ — крайняя точка. Если решение системы неединственно, то существует вектор $h \neq 0$ такой, что $\langle a_j, h\rangle = 0$, $j \in I(\hat{x})$. Тогда для любого $t \in \mathbb{R}$, $j \in I(\hat{x})$ выполнено $\langle a_j, \hat{x} + th\rangle = b_j$. Если $j \notin I(\hat{x})$, то $\langle a_j, \hat{x}\rangle < b_j$, так что при малых t получаем $\langle a_j, \hat{x} + th\rangle < b_j$. Значит, если $\varepsilon > 0$ достаточно мало, то $\hat{x} \pm \varepsilon h \in M$ и $\hat{x}$ — не крайняя точка. □

Значит, если $\hat{x}$ — крайняя точка, то существует подмножество $J \subset \{1, \ldots, m\}$ такое, что $\operatorname{card} J = n$ и векторы $\{a_j\}_{j\in J}$ линейно независимы. При этом $\hat{x}$ — решение системы $\langle a_j, x\rangle = b_j$, $j \in J$. Семейство подмножеств в $\{1, \ldots, m\}$ мощности n конечно, откуда получаем

Следствие 7. *У многогранного множества конечное число крайних точек.*

Следствие 8. *Если многогранное множество ограничено, то это многогранник.*

Доказательство. Ограниченное многогранное множество M выпукло и компактно. По теореме Крейна–Мильмана в $\mathbb{R}^n$, $M = \operatorname{conv}\Sigma$, где Σ — множество крайних точек M. Так как Σ конечно, то M — многогранник. □

Предложение 43. *Выпуклый многогранник является многогранным множеством.*

Доказательство. Пусть M — выпуклый многогранник. Можно считать, что $0 \in M$. Если $\dim M < n$, то $M \subset L$, где L — аффинное подпространство размерности $\dim M$. Так как $0 \in M$, то L — линейное подпространство и поэтому имеет вид

$$L = \{x \in \mathbb{R}^n : \langle a_j, x\rangle = 0, \quad 1 \leq j \leq s\} =$$

$$= \{x \in \mathbb{R}^n : \langle a_j, x\rangle \leq 0, \langle a_j, x\rangle \geq 0, \quad 1 \leq j \leq s\}.$$

Перейдя к подпространству L, можем теперь считать, что $\operatorname{int} M \neq \varnothing$. Снова сделав сдвиг, считаем, что $0 \in \operatorname{int} M$.

Итак, M — выпуклый многогранник, $0 \in \operatorname{int} M$. Тогда M° — ограниченное многогранное множество, поэтому M° — выпуклый многогранник. Значит, $M^{\circ\circ}$ — многогранное множество. По теореме о биполяре, $M^{\circ\circ} = M$. □

Пусть $x_1, \ldots, x_{n+1} \in \mathbb{R}^n$ — аффинно независимые точки. Множество $\operatorname{conv}\{x_j\}_{j=1}^{n+1}$ называется симплексом.

Отметим, что симплекс имеет непустую внутренность.

Предложение 44. *Пусть $K \subset \mathbb{R}^n$ — симплекс, $0 \in \operatorname{int} K$. Тогда K° — симплекс.*

Доказательство. Мы уже знаем: K° — выпуклый многогранник, $0 \in \operatorname{int} K^\circ$. Значит, $K^\circ = \operatorname{conv}\{\xi_j\}_{j=1}^m$, где ξ_j — крайние точки K°, и $\dim K^\circ = n$. Поэтому среди $\{\xi_j\}_{j=1}^m$ есть по крайней мере $n+1$ аффинно независимые точки (см. предложение 8).

С другой стороны, $K = \operatorname{conv}\{x_j\}_{j=1}^{n+1}$, $K^\circ = \{x \in \mathbb{R}^n : \ \langle x_j, x\rangle \le 1\}$. Если точка ξ является крайней для K°, то существует $J \subset \{1, \ldots, n+1\}$, $\operatorname{card} J = n$, такое, что ξ является единственным решением системы равенств $\langle x_j, x\rangle = 1$, $j \in J$. Число подмножеств в $\{1, \ldots, n+1\}$ мощности n будет ровно $n+1$. Значит, K° имеет не более $n+1$ крайних точек.

Таким образом, $m = n+1$ и K° — симплекс. □

Применяя теорему о биполяре, получаем

Следствие 9. *Симплекс является пересечением $n+1$ замкнутых полупространств.*

16 Строгая выпуклость и гладкость

Определение 21. *Пусть $A \subset \mathbb{R}^n$ — выпуклый компакт, $0 \in \operatorname{int} A$. Множество A называется строго выпуклым, если для любых различных точек x_1, $x_2 \in \partial A$ выполнено $\frac{x_1+x_2}{2} \in \operatorname{int} A$.*

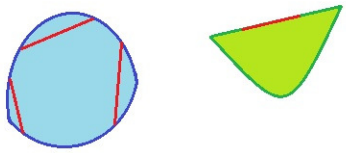

Рис. 27: Строго выпуклое и не строго выпуклое множества.

Предложение 45. *Если множество A строго выпуклое, то каждая его граничная точка является выступающей.*

Доказательство. Пусть $x \in \partial A$ — не выступающая точка. По теореме отделимости, существует вектор $\xi \in \mathbb{R}^n \backslash \{0\}$ такой, что $\langle \xi, x \rangle = \sup_{z \in A} \langle \xi, z \rangle$. Пусть $y \in A$, $\langle \xi, y \rangle = \langle \xi, x \rangle$. Если $y \neq x$, то $\frac{x+y}{2} \in \operatorname{int} A$. Значит, $\left\langle \xi, \frac{x+y}{2} \right\rangle < \sup_{z \in A} \langle \xi, z \rangle$. Но $\left\langle \xi, \frac{x+y}{2} \right\rangle = \frac{1}{2}\langle \xi, x \rangle + \frac{1}{2}\langle \xi, y \rangle = \langle \xi, x \rangle = \sup_{z \in A} \langle \xi, z \rangle$ — противоречие. $\square$

Пусть $A \subset \mathbb{R}^n$ — выпуклый компакт, $0 \in \operatorname{int} A$, $\hat{x} \in \partial A$. По теореме отделимости, существует вектор $\hat{y} \in \mathbb{R}^n \backslash \{0\}$ такой, что $\langle \hat{y}, \hat{x} \rangle = \sup_{x \in A} \langle \hat{y}, x \rangle$. Правая часть положительна, так как $0 \in \operatorname{int} A$. Значит, можно считать, что $\langle \hat{y}, \hat{x} \rangle = 1$. В этом случае $\hat{y} \in A^\circ$.

Предложение 46. *Пусть $A \subset \mathbb{R}^n$ — выпуклый компакт, $0 \in \operatorname{int} A$, $\hat{y} \in A^\circ$. Следующие условия равносильны:*

1. $\hat{y} \in \partial A^\circ$,
2. $\sup_{x \in A} \langle \hat{y}, x \rangle = 1$.

Доказательство. Если $\sup_{x \in A} \langle \hat{y}, x \rangle = 1$, то для любого $\lambda > 1$ получаем $\lambda \hat{y} \notin A^\circ$. Значит, $\hat{y} \in \partial A^\circ$.

Пусть $\sup_{x \in A} \langle \hat{y}, x \rangle = 1 - \varepsilon < 1$. Если $\|y - \hat{y}\| < \delta$, то

$$\sup_{x \in A} \langle y, x \rangle \leq \sup_{x \in A} \langle \hat{y}, x \rangle + \sup_{x \in A} \langle y - \hat{y}, x \rangle \leq 1 - \varepsilon + \delta \max_{x \in A} \|x\| \leq 1$$

при малых $\delta > 0$. Значит, $\hat{y} \in \operatorname{int} A^\circ$. $\square$

Определение 22. *Пусть $A \subset \mathbb{R}^n$ — выпуклый компакт, $0 \in \operatorname{int} A$. Множество A называется гладким, если для любого $x \in \partial A$ опорная гиперплоскость к A в точке x единственна.*

Это эквивалентно тому, что для любого $x \in \partial A$ существует единственный вектор y такой, что $\langle y, x \rangle = 1 = \sup_{z \in A} \langle y, z \rangle$.

Теорема 24. *Пусть $A \subset \mathbb{R}^n$ — выпуклый компакт, $0 \in \operatorname{int} A$. Тогда*

1. *A гладкое $\Leftrightarrow$ A° строго выпуклое,*
2. *A строго выпуклое $\Leftrightarrow$ A° гладкое.*

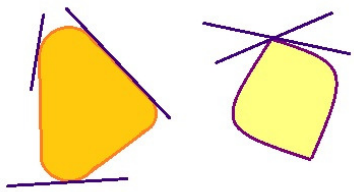

Рис. 28: Гладкое и негладкое множества.

Доказательство. В силу теоремы о биполяре, достаточно проверить первый пункт.

Пусть A не является гладким, то есть существует точка $\xi \in \partial A$, в которой есть две различных опорных гиперплоскости. Значит, существуют $y_1 \neq y_2$ такие, что $\langle y_j, \xi\rangle = 1 = \sup_{x\in A}\langle y_j, x\rangle$, $j = 1, 2$. Тогда для $\frac{y_1+y_2}{2}$ выполнены те же равенства. Значит, $y_1, y_2, \frac{y_1+y_2}{2} \in \partial A^\circ$ в силу предложения 46 и A° не строго выпуклое.

Пусть A° не строго выпуклое, то есть существуют различные $y_1, y_2 \in \partial A^\circ$ такое, что $\frac{y_1+y_2}{2} \in \partial A^\circ$. По предложению 46, $\sup_{x\in A}\left\langle \frac{y_1+y_2}{2}, x\right\rangle = 1$. Так как A компактно, то максимум достигается в некоторой точке ξ, при этом $\xi \in \partial A$. Так как $y_1, y_2 \in A^\circ$, то $\sup_{x\in A}\langle y_j, x\rangle \leq 1$, $j = 1, 2$; в частности, $\langle y_j, \xi\rangle \leq 1$. Но $\left\langle \frac{y_1+y_2}{2}, \xi\right\rangle = 1$. Значит, $\sup_{x\in A}\langle y_j, x\rangle = \langle y_j, \xi\rangle = 1$, $j = 1, 2$, т.е. A не гладкое. $\square$

17 Приложение: задача Чаплыгина

Рассмотрим задачу оптимального управления, в которой определение формы траектории сводится к вычислению поляры некоторого множества.

Сначала напомним необходимые условия минимума (будем рассматривать задачу с ограничениями только типа равенств и с неподвижными концами $t_0 < t_1$).

Пусть $x(t) = (x_1(t), \ldots, x_n(t))$, $u(t) = (u_1(t), \ldots, u_s(t))$. Рассматривается задача

$$\int\limits_{t_0}^{t_1} L_0(t, x(t), u(t))\, dt + l_0(x(t_0), x(t_1)) \to \inf,$$

$$\int\limits_{t_0}^{t_1} L_j(t,\, x(t),\, u(t))\, dt + l_j(x(t_0),\, x(t_1)) = 0,\ 1 \le j \le m,$$

$$\dot{x}(t) = \varphi(t,\, x(t),\, u(t)),$$
$$u(t) \in U.$$

Функция Лагранжа записывается в виде

$$\mathcal{L}(x,\, u) = \sum_{k=0}^{m} \lambda_k \left(\int\limits_{t_0}^{t_1} L_k(t,\, x(t),\, u(t))\, dt + l_k(x(t_0),\, x(t_1)) \right) +$$

$$+ \int\limits_{t_0}^{t_1} p(t)(\dot{x}(t) - \varphi(t,\, x(t),\, u(t)))\, dt =$$

$$= \int\limits_{t_0}^{t_1} L(t,\, x(t),\, \dot{x}(t),\, u(t))\, dt + l(x(t_0),\, x(t_1)).$$

Напомним необходимое условие сильного минимума: существуют числа $\lambda_0,\, \ldots,\, \lambda_m$ и вектор-функция $p(\cdot)$, одновременно не равные 0, такие, что

1. $\lambda_0 \ge 0$ (условие неотрицательности),
2. $-\frac{d}{dt} L_{\dot{x}} + L_x = 0$ (уравнение Эйлера),
3. $L_{\dot{x}}(t_i) = (-1)^i \frac{\partial l}{\partial x(t_i)}$, $i = 0,\, 1$ (условие трансверсальности),
4. $\min_{v \in U} L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t),\, v) = L(t,\, \hat{x}(t),\, \dot{\hat{x}}(t),\, \hat{u}(t))$ (принцип максимума Понтрягина).

Лагранжиан L имеет вид

$$L(t,\, x,\, \dot{x},\, u) = \sum_{k=0}^{m} \lambda_k L_k(t,\, x,\, u) - p(t)\varphi(t,\, x,\, u) + p(t)\dot{x}.$$

Последнее слагаемое не зависит от u. Обозначим

$$H(t,\, x,\, u) = p(t)\varphi(t,\, x,\, u) - \sum_{k=0}^{m} \lambda_k L_k(t,\, x,\, u);$$

H называется функцией Понтрягина. Принцип максимума Понтрягина переписывается в эквивалентном виде

$$\max_{v \in U} H(t,\, \hat{x}(t),\, v) = H(t,\, \hat{x}(t),\, \hat{u}(t)).$$

Следующую теорему сформулируем без доказательства.

Теорема 25. *Если функции L_j, φ непрерывно дифференцируемы по совокупности переменных и явно не зависят от t, то $H(t,\, \hat{x}(t),\, \hat{u}(t)) = \mathrm{const}$ (здесь $\hat{x}(t)$, $\hat{u}(t)$ — допустимая экстремаль, т.е. удовлетворяет ограничениям задачи и пунктам 1–4 необходимого условия минимума).*

Теперь рассмотрим задачу

$$\frac{1}{2}\int\limits_0^{T_0} (xv - yu)\, dt \to \inf,$$

$$\dot{x} = u,\ \dot{y} = v,$$

$$x(0) = x(T_0), \quad y(0) = y(T_0),$$

$$(u,\, v) \in U,$$

где $U \subset \mathbb{R}^2$ — выпуклый компакт, $(0,\, 0) \in \operatorname{int} U$.

Физический смысл задачи следующий: нужно за фиксированное время по замкнутой траектории обойти область с максимальной площадью, при этом ограничение на вектор скорости имеет вид $(u(t),\, v(t)) \in U$. Если U — круг с центром в нуле, то задача эквивалентна задаче Дидоны (в этом случае траекторией будет окружность).

Здесь мы не будем обсуждать вопрос существования решения, только выпишем необходимые условия и определим форму траектории (и увидим, как в этом участвует поляра множества).

Функция Лагранжа имеет вид

$$\mathcal{L} = \frac{\lambda_0}{2}\int\limits_0^{T_0} (xv - yu)\, dt + \lambda_1(x(0) - x(T_0)) + \lambda_2(y(0) - y(T_0)) +$$

$$+ \int\limits_0^{T_0} (p(t)(\dot{x} - u) + q(t)(\dot{y} - v))\, dt.$$

Условие неотрицательности: $\lambda_0 \geq 0$.

Уравнения Эйлера: $-\dot{p} + \frac{\lambda_0}{2} v = 0$, $-\dot{q} - \frac{\lambda_0}{2} u = 0$.

Условия трансверсальности: $p(0) = \lambda_1$, $p(T_0) = \lambda_1$, $q(0) = \lambda_2$, $q(T_0) = \lambda_2$.

Принцип максимума Понтрягина:

$$\max_{(u,\,v)\in U} \left(pu + qv - \frac{\lambda_0}{2}\hat{x}(t)v + \frac{\lambda_0}{2}\hat{y}(t)u \right) =$$

$$= \left(p\hat{u}(t) + q\hat{v}(t) - \frac{\lambda_0}{2}\hat{x}(t)\hat{v}(t) + \frac{\lambda_0}{2}\hat{y}(t)\hat{u}(t) \right).$$

Если $\lambda_0 = 0$, то p и q — константы. Принцип максимума Понтрягина записывается в виде $\max_{(u,\,v)\in U}(pu+qv) = p\hat{u}(t)+q\hat{v}(t)$. При этом $(p,\, q) \neq (0,\, 0)$ (иначе все множители Лагранжа зануляются). Так как $0 \in \operatorname{int} U$, то максимум будет положительным числом (не зависящим от t). Применив формулу Ньютона–Лейбница, получим

$$0 = p(\hat{x}(T_0) - \hat{x}(0)) + q(\hat{y}(T_0) - \hat{y}(0)) = \int\limits_0^{T_0} (p\hat{u}(t) + q\hat{v}(t))\, dt > 0$$

— противоречие.

Пусть $\lambda_0 > 0$. Можно считать, что $\lambda_0 = 1$. Из уравнений Эйлера получаем:

$$p(t) = \frac{1}{2}\hat{y}(t) + a, \quad q(t) = -\frac{1}{2}\hat{x}(t) + b,$$

где a, b — некоторые константы. Тогда

$$pu + qv - \frac{1}{2}\hat{x}(t)v + \frac{1}{2}\hat{y}(t)u = (\hat{y}(t) + a)u - (\hat{x}(t) - b)v.$$

По теореме 25, $(\hat{y}(t) + a)\hat{u}(t) - (\hat{x}(t) - b)\hat{v}(t) = c$ (c — некоторая константа), при этом является максимумом $(\hat{y}(t) + a)u - (\hat{x}(t) - b)v$ по $(u,\, v) \in U$. Минимум в задаче не может достигаться на константе, поэтому $(\hat{x},\, \hat{y}) \not\equiv (b,\, -a)$, так что $c > 0$. Значит,

$$\max_{(u,\,v)\in U} \left(\frac{1}{c}(\hat{y}(t) + a)u - \frac{1}{c}(\hat{x}(t) - b)v \right) = 1.$$

Поэтому $\frac{1}{c}(\hat{y}(t) + a,\, -\hat{x}(t) + b)$ — граничная точка U° (см. предложение 46).

Значит, траектория строится следующим образом: находится поляра U, поворачивается на $90°$, затем делается ее раздутие в c раз; траекторию можно сдвигать на любой вектор. Константа c находится из равенства

$$cT_0 = \int\limits_0^{T_0} ((\hat{y}(t) + a)\hat{u}(t) - (\hat{x}(t) - b)\hat{v}(t))\, dt = 2S_{cU^\circ} = 2c^2 S_{U^\circ},$$

где S_{U° — площадь U°.

Примером такой задачи является задача Чаплыгина; множество U задано неравенством $(u - u_0)^2 + v^2 \le R^2$, где $u_0 < R$ (т.е. U — сдвинутый круг, $(0,\, 0)$ — его внутренняя точка). Здесь R — максимально возможная собственная скорость самолета, u_0 — скорость ветра. В этом случае U° — эллипс (см. одну из задач); значит, и траектория будет эллипсом (повернутым на $90°$).

18 Сопряженные функции (преобразование Лежандра–Юнга–Фенхеля)

Пусть $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$. Положим

$$f^*(y) = \sup_{x \in \mathbb{R}^n} (\langle x,\, y\rangle - f(x)).$$

Функция f^* называется сопряженной к f функцией или ее преобразованием Лежандра–Юнга–Фенхеля.

Если существует точка x_0 такая, что $f(x_0) = -\infty$, то $f^* \equiv +\infty$. Если $f \equiv +\infty$, то $f^* \equiv -\infty$.

Далее рассматриваем случай, когда f собственная (т.е. f нигде не равна $-\infty$ и хотя бы в одной точке конечна).

Определение 23. *Функция $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ называется замкнутой, если* $\operatorname{epi} f$ *замкнут.*

Предложение 47. *Функция замкнута тогда и только тогда, когда она полунепрерывна снизу.*

Доказательство. Пусть f замкнута. Покажем, что для любого $c \in \mathbb{R}$ множество $\{x \in \mathbb{R}^n :\ f(x) \le c\}$ замкнуто. В самом деле, пусть $f(x_0) >$

c. Тогда $(x_0,\, c) \notin \operatorname{epi} f$. Значит, существует окрестность $B_\varepsilon(x_0) \times (c - \varepsilon,\, c + \varepsilon)$, не пересекающаяся с $\operatorname{epi} f$. В частности, для любого $x \in B_\varepsilon(x_0)$ выполнено $f(x) > c$.

Пусть f полунепрерывна снизу. Покажем, что надграфик f замкнут. Пусть $(x_0,\, c) \notin \operatorname{epi} f$. Тогда $c < f(x_0)$; значит, $c+\varepsilon < f(x_0)$ для некоторого $\varepsilon > 0$. Так как f полунепрерывна снизу, то существует окрестность U точки x_0 такая, что $c+\varepsilon < f(x)$ для любого $x \in U$. Значит, $U \times (c-\varepsilon,\, c+\varepsilon)$ не пересекается с $\operatorname{epi} f$. $\square$

Предложение 48. *Пусть $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ — собственная выпуклая замкнутая функция. Тогда существует аффинная функция $h : \mathbb{R}^n \to \mathbb{R}$ такая, что $f(x) \geq h(x)$ для любого $x \in \mathbb{R}^n$. (Аффинная функция — линейная плюс константа.)*

Доказательство. Пусть $x_0 \in \mathbb{R}^n$, $f(x_0) < +\infty$. Тогда $(x_0,\, f(x_0) - 1) \notin \operatorname{epi} f$. Так как $\operatorname{epi} f$ — замкнутое множество, то существует выпуклая окрестность U точки $(x_0,\, f(x_0)-1)$, не пересекающаяся с $\operatorname{epi} f$. Применив теорему отделимости к выпуклым множествам U и $\operatorname{epi} f$, получаем, что существуют $\xi \in \mathbb{R}^n$ и $\alpha \in \mathbb{R}$, одновременно не равные нулю, такие, что

$$\inf_{(x,\, u) \in \operatorname{epi} f} (\langle \xi,\, x \rangle + \alpha u) \geq \sup_{(x,\, u) \in U} (\langle \xi,\, x \rangle + \alpha u).$$

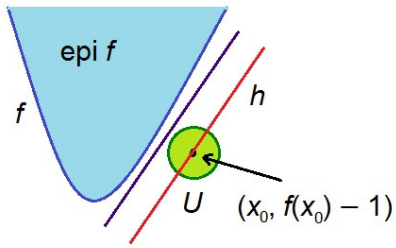

Рис. 29: Разделяющая гиперплоскость

Если $\alpha < 0$, то inf левой части равен $-\infty$, так что этот случай не подходит. Если $\alpha = 0$, то $\langle \xi,\, x_0 \rangle \geq \sup_{(x,\, u) \in U} \langle \xi,\, x \rangle$. Это возможно только если $\xi = 0$, так что снова получаем противоречие. Значит, $\alpha > 0$. Можно считать, что $\alpha = 1$. Тогда $\langle \xi,\, x \rangle + f(x) \geq \langle \xi,\, x_0 \rangle + f(x_0) - 1$. Остается положить $h(x) = -\langle \xi,\, x - x_0 \rangle + f(x_0) - 1$. $\square$

Замечание. *В таком виде предыдущее утверждение верно и в бесконечномерном случае. Для функций на* $\mathbb{R}^n$ *замкнутость несущественна: если* $\operatorname{dom} f$ *имеет непустую внутренность, то можно применить теорему о непустоте субдифференциала. Если внутренность пуста, то рассматриваем сужение функции на аффинную оболочку* $\operatorname{dom} f$*, затем построенную на ней аффинную функцию продолжаем на* $\mathbb{R}^n$*.*

Предложение 49. *Пусть* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *— выпуклая замкнутая собственная функция. Тогда* f^* *— выпуклая замкнутая собственная функция.*

Доказательство. Докажем, что f^* собственная. Имеем

$$f^*(y) = \sup_{x \in \mathbb{R}^n} (\langle x,\, y\rangle - f(x)).$$

Так как существует точка $x_0 \in \mathbb{R}^n$ такая, что $f(x_0) < +\infty$, то $f^*(y) \geq \langle x_0,\, y\rangle - f(x_0) > -\infty$. По предложению 48, существуют $\xi_0 \in \mathbb{R}^n$ и $c \in \mathbb{R}$ такие, что $f(x) \geq \langle \xi_0,\, x\rangle + c$, поэтому

$$f^*(\xi_0) \leq \sup_{x \in \mathbb{R}^n} (\langle x,\, \xi_0\rangle - \langle \xi_0,\, x\rangle - c) = -c.$$

Докажем, что f^* выпуклая и замкнутая. Положим $h_x(\cdot) = \langle x,\, \cdot\rangle - f(x)$. Это выпуклая замкнутая функция. Имеем

$$(y,\, v) \in \operatorname{epi} f^* \;\Leftrightarrow\; v \geq f^*(y) \;\Leftrightarrow\; v \geq \sup_{x \in \mathbb{R}^n} (\langle x,\, y\rangle - f(x)) \;\Leftrightarrow$$

$$\Leftrightarrow\; \forall x \in \mathbb{R}^n\; v \geq \langle x,\, y\rangle - f(x) \;\Leftrightarrow$$

$$\Leftrightarrow\; \forall x \in \mathbb{R}^n\; (y,\, v) \in \operatorname{epi} h_x \;\Leftrightarrow\; (y,\, v) \in \cap_{x \in \mathbb{R}^n} \operatorname{epi} h_x.$$

Так как $\operatorname{epi} h_x$ выпуклы и замкнуты, то $\operatorname{epi} f^*$ выпуклый и замкнутый. $\square$

Пример. Пусть $f(x) = \|x\|$ — некоторая норма на $\mathbb{R}^n$, $\|\cdot\|^*$ — сопряженная норма. Вычислим

$$f^*(y) = \sup_{x \in \mathbb{R}^n} (\langle x,\, y\rangle - \|x\|).$$

Если $\|y\|^* > 1$, то существует $x \in \mathbb{R}^n$ такое, что $\langle x,\, y\rangle > \|x\|$. Значит, $\lim\limits_{\lambda \to +\infty} (\langle \lambda x,\, y\rangle - \|\lambda x\|) = +\infty$. Поэтому $f^*(y) = +\infty$. Если $\|y\|^* \leq 1$, то

$\langle x,\, y\rangle - \|x\| \le 0$ для любого $x \in \mathbb{R}^n$, при этом 0 достигается в точке $x = 0$. Значит, $f^*(y) = 0$.

Таким образом, $f^* = I_{\{x:\, \|x\|^* \le 1\}}$.

Пример из механики. Пусть $L(t,\, x,\, u)$ — гладкая функция, выпуклая по u, $\frac{L(t,x,u)}{|u|} \underset{|u|\to\infty}{\to} +\infty$. Тогда

$$\sup_{u\in\mathbb{R}^n} (\langle p,\, u\rangle - L(t,\, x,\, u))$$

конечный и достигается в некоторой точке $\hat{u}$. Получаем $p = L_u(t,\, x,\, \hat{u})$. Пусть это уравнение однозначно разрешимо относительно третьей переменной: $\hat{u} = \varphi(t,\, x,\, p)$. Положим $H(t,\, x,\, p) := \langle p,\, \hat{u}\rangle - L(t,\, x,\, \hat{u}) = \langle p,\, \varphi(t,\, x,\, p)\rangle - L(t,\, x,\, \varphi(t,\, x,\, p))$. В механике L называется функцией Лагранжа, H — функцией Гамильтона. В частности, если $L(t,\, x,\, \dot{x}) = \frac{m\dot{x}^2}{2} - V(x)$ (разность кинетической и потенциальной энергии), то $p = m\dot{x}$ (импульс), $H(t,\, x,\, p) = \frac{p^2}{2m} + V(x)$ (сумма кинетической и потенциальной энергии).

Неравенство Юнга. Из определения сопряженной функции следует, что для любых $x,\, y \in \mathbb{R}^n$ выполнено $f^*(y) \ge \langle x,\, y\rangle - f(x)$, или

$$\langle x,\, y\rangle \le f(x) + f^*(y),\ x,\, y \in \mathbb{R}^n.$$

Это соотношение называется неравенством Юнга.

Пример. Пусть $1 < p < \infty$, $f(x) = \frac{|x|^p}{p}$, $x \in \mathbb{R}$. Функция $xy - \frac{|x|^p}{p}$ достигает максимума; в точке максимума получаем $y - |x|^{p-1}\operatorname{sgn} x = 0$, откуда $x = |y|^{\frac{1}{p-1}}\operatorname{sgn} y$, $f^*(y) = |y|^{p'} - \frac{1}{p}|y|^{p'} = \frac{|y|^{p'}}{p'}$. Отсюда получаем

$$xy \le \frac{|x|^p}{p} + \frac{|y|^{p'}}{p'}.$$

Лемма 3. (Минковский). *Пусть $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ — выпуклая замкнутая собственная функция. Тогда для любого $(x_0,\, u_0) \notin \operatorname{epi} f$ существует аффинная функция $h : \mathbb{R}^n \to \mathbb{R}$ такая, что $h(x_0) > u_0$ и $h(x) \le f(x)$ для любого $x \in \mathbb{R}^n$.*

Доказательство. По теореме о строгой отделимости, существуют $x^* \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ и $c \in \mathbb{R}$ такие, что

$$\inf_{(x,\, u)\in\operatorname{epi} f} (\langle x^*,\, x\rangle + \alpha u) \ge c > \langle x^*,\, x_0\rangle + \alpha u_0. \tag{26}$$

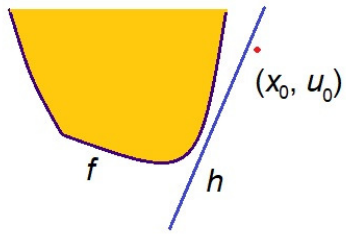

Рис. 30: Разделяющая гиперплоскость

Если $\alpha < 0$, то левая часть равна $-\infty$ (рассматриваем $\overline{x} \in \operatorname{dom} f$ и $u \to +\infty$). Значит, этот случай невозможен.

Если $\alpha > 0$, то можно считать, что $\alpha = 1$. Тогда для любого $x \in \mathbb{R}^n$ выполнено $\langle x^*, x\rangle + f(x) \geq c > \langle x^*, x_0\rangle + u_0$, или $f(x) \geq -\langle x^*, x\rangle + c$, $u_0 < -\langle x^*, x_0\rangle + c$. Остается положить $h(x) = -\langle x^*, x\rangle + c$.

Пусть $\alpha = 0$ (так как x_0 может не принадлежать $\operatorname{dom} f$, то этот случай возможен). В силу (26), существует $\varepsilon > 0$ такое, что $\langle x^*, x\rangle \geq c$ для любых $(x, u) \in \operatorname{epi} f$ и $\langle x^*, x_0\rangle < c - \varepsilon$.

Из предложения 48 следует, что существуют $z^* \in \mathbb{R}^n$ и $b \in \mathbb{R}$ такие, что $f(x) \geq \langle z^*, x\rangle + b$.

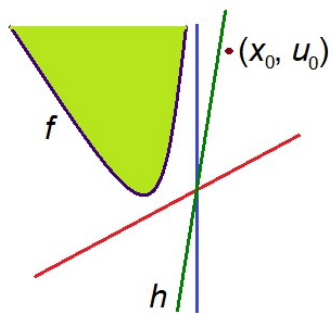

Рис. 31: Построение функции h

Положим $h(x) = \langle z^*, x\rangle + b + R(c - \langle x^*, x\rangle)$, где $R > 0$ — достаточно большое число. Так как $c - \langle x^*, x\rangle \leq 0$ для любого $x \in \operatorname{dom} f$, то $f(x) \geq h(x)$ для любого $x \in \mathbb{R}^n$. С другой стороны,

$$h(x_0) = \langle z^*, x_0\rangle + b + R(c - \langle x^*, x_0\rangle) \geq \langle z^*, x_0\rangle + b + R\varepsilon > u_0,$$

если R достаточно велико. $\square$

Второй сопряженной функцией называется $f^{**} = (f^*)^*$.

Из неравенства Юнга следует, что $\langle x,\, y\rangle - f^*(y) \le f(x)$ для любых $x,\, y \in \mathbb{R}^n$, поэтому

$$f^{**}(x) = \sup_{y\in\mathbb{R}^n}(\langle x,\, y\rangle - f^*(y)) \le f(x).$$

Теорема 26. (Фенхель–Моро). *Пусть $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ — выпуклая замкнутая собственная функция. Тогда $f^{**} = f$.*

Доказательство. Пусть существует точка $x_0 \in \mathbb{R}^n$ такая, что $f^{**}(x_0) < f(x_0)$. По лемме Минковского, существуют аффинная функция h и $\varepsilon > 0$ такие, что $f(x) \ge h(x)$ для любого $x \in \mathbb{R}^n$ и $h(x_0) > f^{**}(x_0) + \varepsilon$. Значит, существуют $c \in \mathbb{R}$ и $\xi \in \mathbb{R}^n$ такие, что

$$f(x) \ge \langle \xi,\, x\rangle - c,\ x \in \mathbb{R}^n, \quad \langle \xi,\, x_0\rangle - c > f^{**}(x_0) + \varepsilon.$$

Из первого неравенства следует, что $f^*(\xi) \le c$. Значит,

$$c + \varepsilon < \langle \xi,\, x_0\rangle - f^{**}(x_0) \le f^*(\xi) \le c.$$

Получили противоречие. □

Следствие 10. *Пусть f — собственная функция. Тогда следующие условия эквивалентны:*

1. *f выпуклая и замкнутая,*
2. *$f^{**} = f$.*

Как найти f^{**}, если собственная функция f не является выпуклой или замкнутой?

Пусть $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ — собственная функция. Обозначим $V_f = \overline{\operatorname{conv}}\operatorname{epi} f$.

Покажем, что если $(x,\, u) \in V_f$, $v \ge u$, то $(x,\, v) \in V_f$. В самом деле, $(x,\, u) = \lim\limits_{n\to\infty}(x_n,\, u_n)$, $x_n = \sum\limits_{j=1}^{N_n}\lambda_{n,j}x_{n,j}$, $u_n = \sum\limits_{j=1}^{N_n}\lambda_{n,j}u_{n,j}$, $\sum\limits_{j=1}^{N_n}\lambda_{n,j} = 1$, $\lambda_{n,j} \ge 0$, $(x_{n,j},\, u_{n,j}) \in \operatorname{epi} f$. Так как $(x_{n,j},\, u_{n,j} + v - u) \in \operatorname{epi} f$, $u_n + v - u = \sum\limits_{j=1}^{N_n}\lambda_{n,j}(u_{n,j} + v - u)$, то $(x,\, v) = \lim\limits_{n\to\infty}(x_n,\, u_n + v - u) \in V_f$.

Положим

$$\overline{\operatorname{conv}} f(x) = \inf\{u \in \mathbb{R} :\ (x,\, u) \in V_f\}$$

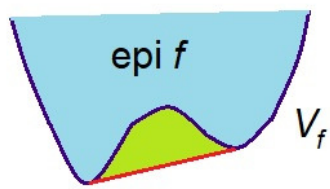

Рис. 32: Множество V_f

(считаем, что $\inf \varnothing = +\infty$). Тогда $V_f = \operatorname{epi} \overline{\operatorname{conv}} f$. Значит, $\overline{\operatorname{conv}} f$ — выпуклая замкнутая функция. Далее, $\overline{\operatorname{conv}} f(x) \le f(x)$ для любого $x \in \mathbb{R}^n$.

Если существует аффинная функция h такая, что $f(x) \ge h(x)$ для любого $x \in \mathbb{R}^n$, то $\operatorname{epi} f \subset \operatorname{epi} h$ и поэтому $V_f \subset \operatorname{epi} h$. Значит, $\overline{\operatorname{conv}} f(x) \ge h(x)$ и $\overline{\operatorname{conv}} f$ — собственная.

Наоборот, если $\overline{\operatorname{conv}} f$ собственная, то в силу предложения 48 существует аффинная функция h такая, что $\overline{\operatorname{conv}} f(x) \ge h(x)$. Значит, $f(x) \ge h(x)$.

Отметим также, что если f^* конечна в точке y, то в силу неравенства Юнга $f(x) \ge \langle x, y\rangle - f^*(y)$ для любого $x \in \mathbb{R}^n$.

Предложение 50. *Выполнено равенство* $f^* = (\overline{\operatorname{conv}} f)^*$.

Доказательство. В самом деле,

$$(\overline{\operatorname{conv}} f)^*(y) = \sup(\langle x, y\rangle - \overline{\operatorname{conv}} f(x)) = \sup_{(x,\,u)\in \operatorname{epi} \overline{\operatorname{conv}} f} (\langle x, y\rangle - u) =$$

$$= \sup_{(x,\,u)\in \operatorname{epi} f} (\langle x, y\rangle - u) = \sup(\langle x, y\rangle - f(x)) = f^*(y).$$

Утверждение доказано. □

Предложение 51. *Пусть* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *— собственная функция. Предположим, что существует аффинная функция* h *такая, что* $f(x) \ge h(x)$ *для любого* $x \in \mathbb{R}^n$. *Тогда* $f^{**} = \overline{\operatorname{conv}} f$.

Доказательство. По теореме Фенхеля–Моро, $(\overline{\operatorname{conv}} f)^{**} = \overline{\operatorname{conv}} f$. В силу предложения 50, $f^* = (\overline{\operatorname{conv}} f)^*$. Значит, $f^{**} = (\overline{\operatorname{conv}} f)^{**} = \overline{\operatorname{conv}} f$. □

Что будет, если аффинной функции h из предыдущего утверждения не существует? Тогда функция $\overline{\operatorname{conv}} f$ не будет собственной и не равна тождественно $+\infty$. Значит, она в какой-то точке равна $-\infty$, поэтому $f^* = (\overline{\operatorname{conv}} f)^* \equiv +\infty$, $f^{**} \equiv -\infty$.

Предложение 52. *Пусть* $f_1, \ldots, f_k : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$,

$$f(x) = \min\{f_1(x), \ldots, f_k(x)\}.$$

Тогда $f^*(y) = \max\{f_1^*(y), \ldots, f_k^*(y)\}$.

Доказательство. Имеем

$$f^*(y) = \sup_{x \in \mathbb{R}^n} (\langle x, y\rangle - \min\{f_1(x), \ldots, f_k(x)\}) =$$

$$= \max_{1 \le j \le k} \sup_{x \in \mathbb{R}^n} \{\langle x, y\rangle - f_j(x)\} = \max\{f_1^*(y), \ldots, f_k^*(y)\}.$$

□

Предложение 53. *Пусть* $f_1, \ldots, f_k : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *— выпуклые замкнутые собственные функции,*

$$f(x) = \max\{f_1(x), \ldots, f_k(x)\}.$$

Пусть функция f *собственная. Тогда* $f^*(y) = \overline{\mathrm{conv}}\, \min\{f_1^*(y), \ldots, f_k^*(y)\}$.

Доказательство. Проверим, что для функции $\min\{f_1^*, \ldots, f_k^*\}$ выполнены условия предложения 51. В самом деле, так как f собственная, то существует точка $x_0 \in \mathbb{R}^n$ такая, что $f(x_0) < +\infty$. В силу неравенства Юнга,

$$f_j^*(y) \ge \langle x_0, y\rangle - f_j(x_0) \ge \langle x_0, y\rangle - f(x_0),\ j = 1, \ldots, k.$$

Значит,

$$\min\{f_1^*(y), \ldots, f_k^*(y)\} \ge \langle x_0, y\rangle - f(x_0).$$

В силу предложений 51, 52 и теоремы Фенхеля–Моро для $f_1, \ldots, f_k$,

$$\overline{\mathrm{conv}}\, \min\{f_1^*, \ldots, f_k^*\} = \min\{f_1^*, \ldots, f_k^*\}^{**} =$$

$$= \max\{f_1^{**}, \ldots, f_k^{**}\}^* = \max\{f_1, \ldots, f_k\}^*.$$

□

Предложение 54. *Пусть* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *— выпуклая функция,* $x \in \mathrm{dom}\, f$. *Следующие условия эквивалентны:*

1. $y \in \partial f(x)$,

2. $\langle x,\, y\rangle = f(x) + f^*(y)$.

Доказательство. 1) $\Rightarrow$ 2). Пусть $y \in \partial f(x)$. Тогда для любого $z \in \mathbb{R}^n$ выполнено $f(z) - f(x) \geq \langle y,\, z - x\rangle$. Значит, $\langle y,\, x\rangle - f(x) \geq \langle y,\, z\rangle - f(z)$. Взяв sup по z, получим $\langle y,\, x\rangle - f(x) \geq f^*(y)$. В силу неравенства Юнга, $\langle x,\, y\rangle = f(x) + f^*(y)$.

2) $\Rightarrow$ 1). Пусть $\langle x,\, y\rangle = f(x) + f^*(y)$. Тогда для любого $z \in \mathbb{R}^n$ получаем $\langle x,\, y\rangle \geq f(x) + \langle z,\, y\rangle - f(z)$. Значит, $f(z) - f(x) \geq \langle y,\, z - x\rangle$. $\square$

Теорема 27. *Пусть $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ — выпуклая замкнутая собственная функция, при этом $f\left(\frac{x+y}{2}\right) < \frac{f(x)+f(y)}{2}$ для любых различных $x,\, y \in \mathbb{R}^n$ таких, что $\frac{x+y}{2} \in \operatorname{dom} f$, $\partial f\left(\frac{x+y}{2}\right) \neq \varnothing$. Тогда*

1. $\operatorname{card} \partial f^*(x^*) \leq 1$ *для любого* $x^* \in \operatorname{dom} f^*$;

2. f^* *дифференцируема по Гато в каждой точке* $x \in \operatorname{int} \operatorname{dom} f^*$.

Доказательство. Докажем утверждение 1. По теореме Фенхеля–Моро, $f^{**} = f$. Значит, в силу предложения 54,

$$\xi \in \partial f^*(x^*) \;\Leftrightarrow\; \langle x^*,\, \xi\rangle = f^*(x^*) + f(\xi). \tag{27}$$

Пусть $x^* \in \operatorname{dom} f^*$, $x,\, y \in \partial f^*(x^*)$ — две различные точки. Покажем, что $f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2}$, $\frac{x+y}{2} \in \operatorname{dom} f$ и $\partial f\left(\frac{x+y}{2}\right) \neq \varnothing$. Тогда получим противоречие с условиями теоремы.

Так как $\partial f^*(x^*)$ — выпуклое множество, то $\frac{x+y}{2} \in \partial f^*(x^*)$. В силу (27),

$$\begin{aligned} \langle x^*,\, x\rangle &= f(x) + f^*(x^*), \\ \langle x^*,\, y\rangle &= f(y) + f^*(x^*), \\ \left\langle x^*,\, \tfrac{x+y}{2}\right\rangle &= f\left(\tfrac{x+y}{2}\right) + f^*(x^*). \end{aligned} \tag{28}$$

Отсюда следует, что $x,\, y,\, \frac{x+y}{2} \in \operatorname{dom} f$; в силу предложения 54, $x^* \in \partial f\left(\frac{x+y}{2}\right)$; в частности, $\partial f\left(\frac{x+y}{2}\right) \neq \varnothing$.

Сложим первые два равенства (28) и поделим на 2:

$$\left\langle x^*,\, \frac{x+y}{2}\right\rangle = \frac{f(x) + f(y)}{2} + f^*(x^*).$$

Отсюда и из третьего равенства (28) следует, что $f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2}$.

Утверждение 2 следует из утверждения 1 в силу предложения 25 и теоремы 12. $\square$

Теорема 28. *Пусть* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *— выпуклая замкнутая собственная функция,* $\operatorname{card} \partial f(x) \le 1$ *для любого* $x \in \operatorname{dom} f$*. Тогда для любых различных* x^*, $y^* \in \mathbb{R}^n$ *таких, что* $\frac{x^*+y^*}{2} \in \operatorname{dom} f^*$, $\partial f^*\left(\frac{x^*+y^*}{2}\right) \ne \varnothing$, *выполнено* $f^*\left(\frac{x^*+y^*}{2}\right) < \frac{f^*(x^*)+f^*(y^*)}{2}$.

Доказательство. Пусть $x \in \partial f^*\left(\frac{x^*+y^*}{2}\right)$. В силу предложения 54,

$$\left\langle \frac{x^*+y^*}{2},\, x \right\rangle = f^*\left(\frac{x^*+y^*}{2}\right) + f(x).$$

Если $f^*\left(\frac{x^*+y^*}{2}\right) = \frac{f^*(x^*)+f^*(y^*)}{2}$, то

$$\frac{\langle x^*,\, x\rangle + \langle y^*,\, x\rangle}{2} = \frac{f^*(x^*)+f^*(y^*)}{2} + f(x). \tag{29}$$

С другой стороны, в силу неравенства Юнга

$$\langle x^*,\, x\rangle \le f^*(x^*) + f(x), \quad \langle y^*,\, x\rangle \le f^*(y^*) + f(x).$$

Отсюда и из (29) следует, что

$$\langle x^*,\, x\rangle = f^*(x^*) + f(x), \quad \langle y^*,\, x\rangle = f^*(y^*) + f(x),$$

поэтому $x^* \in \partial f(x)$, $y^* \in \partial f(x)$ в силу предложения 54. Так как $\operatorname{card} \partial f(x) \le 1$, то $x^* = y^*$. □

Определение 24. *Функция* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *называется строго выпуклой, если для любых различных точек* x, $y \in \operatorname{dom} f$ *и любого* $\lambda \in (0,\, 1)$ *выполнено* $f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y)$.

Строгая выпуклость эквивалентна следующему условию: для любых различных точек x, $y \in \operatorname{dom} f$ выполнено $f\left(\frac{x+y}{2}\right) < \frac{f(x)+f(y)}{2}$. В самом деле, пусть $\lambda \in (0,\, 1)$. Достаточно рассмотреть $\lambda < \frac{1}{2}$. Тогда

$$f((1-\lambda)x+\lambda y) = f\left((1-2\lambda)x + 2\lambda \cdot \frac{x+y}{2}\right) \le (1-2\lambda)f(x) + 2\lambda f\left(\frac{x+y}{2}\right) <$$

$$< (1-2\lambda)f(x) + 2\lambda \cdot \frac{f(x)+f(y)}{2} = (1-\lambda)f(x) + \lambda f(y).$$

Теорема 29. *Пусть* f *— выпуклая замкнутая функция на* $\mathbb{R}^n$, f *и* f^* *всюду конечны. Тогда*

1. *f строго выпукла тогда и только тогда f^* всюду дифференцируема по Гато,*

2. *f всюду дифференцируема по Гато тогда и только тогда, когда f^* строго выпукла.*

Доказательство. Утверждение следует из теорем 27, 28 с использованием предложения 25 и теорем 11, 12. □

19 Теорема о седловой точке

Теорема 30. *Пусть X, Y — конечномерные пространства, $K_X \subset X$, $K_Y \subset Y$ — выпуклые компакты. Пусть $f : K_X \times K_Y \to \mathbb{R}$ — непрерывная функция, при этом $f(\cdot,\, y)$ выпукла для любого $y \in K_Y$, $f(x,\, \cdot)$ вогнута для любого $x \in K_X$ (т.е. $-f(x,\, \cdot)$ выпукла). Тогда существует точка $(\hat{x},\, \hat{y}) \in K_X \times K_Y$ такая, что*

$$f(x,\, \hat{y}) \geq f(\hat{x},\, \hat{y}) \geq f(\hat{x},\, y) \quad \forall x \in K_X,\ y \in K_Y. \tag{30}$$

Точка $(\hat{x},\, \hat{y})$ называется седловой.

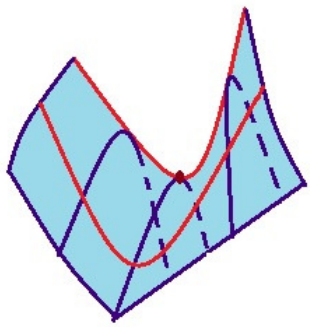

Рис. 33: Седловая точка

Условие (30) переписывается в следующем виде: $\hat{x} \in \operatorname{Argmin} f(\cdot,\, \hat{y})$ (множество всех точек, где достигается минимум $f(\cdot,\, \hat{y})$), $\hat{y} \in \operatorname{Argmax} f(\hat{x},\, \cdot)$ (множество всех точек, где достигается максимум $f(\hat{x},\, \cdot)$), или

$$(\hat{x},\, \hat{y}) \in \operatorname{Argmin} f(\cdot,\, \hat{y}) \times \operatorname{Argmax} f(\hat{x},\, \cdot).$$

Рассмотрим отображение Φ, которое каждой точке $(x,\, y) \in K_X \times K_Y$ сопоставляет множество $\operatorname{Argmin} f(\cdot,\, y) \times \operatorname{Argmax} f(x,\, \cdot)$. Нужно показать, что существует точка $(\hat{x},\, \hat{y})$ такая, что $(\hat{x},\, \hat{y}) \in \Phi(\hat{x},\, \hat{y})$.

Для этого мы докажем теорему Какутани:

Теорема 31. *Пусть $K \subset \mathbb{R}^d$ — выпуклый компакт; отображение Φ сопоставляет каждой точке $x \in K$ выпуклое замкнутое множество $\Phi(x) \subset K$. Предположим, что для любых $x_n \in K$, $y_n \in \Phi(x_n)$ таких, что $(x_n,\, y_n) \underset{n\to\infty}{\to} (x,\, y)$ выполнено $y \in \Phi(x)$. Тогда существует точка $\hat{x} \in K$ такая, что $\hat{x} \in \Phi(\hat{x})$.*

Перед тем, как доказывать теорему Какутани, сформулируем вспомогательные утверждения.

Теорема 32. (Брауэр). *Пусть $K \subset \mathbb{R}^d$ — выпуклый компакт, $f : K \to K$ — непрерывное отображение. Тогда существует точка $\hat{x} \in K$ такая, что $\hat{x} = f(\hat{x})$.*

Теорему Брауэра доказывать не будем.

Барицентрическое разбиение. Пусть $K = \operatorname{conv}\{x_j\}_{j=1}^{d+1}$ — d-мерный симплекс. Пусть $J \subset \{1,\, \ldots,\, d{+}1\}$, $\operatorname{card} J = k{+}1$. Множество $\operatorname{conv}\{x_j\}_{j\in J}$ называется k-мерной гранью K.

Определим барицентрическое разбиение симплекса на $(d+1)!$ симплексов индукцией по d.

1. Если $d = 1$, то симплекс — это отрезок. В этом случае берется его разбиение на две равные части.

2. Пусть барицентрическое разбиение определено для k-мерных симплексов, $k < d$ (число элементов разбиения равно $(k+1)!$). Обозначим $d-1$-мерные грани K через K_j $(1 \le j \le d+1)$, положим $x_0 = \sum\limits_{j=1}^{d+1} \frac{x_j}{d+1}$. По предположению индукции, K_j разбиваем на симплексы $\Delta_{j,i}$, $1 \le i \le d!$. Тогда $\{\operatorname{conv}(\{x_0\} \cup \Delta_{j,i})\}_{1\le j\le d+1,\, 1\le i\le d!}$ будет барицентрическим разбиением K.

Барицентрическое разбиение симплекса обладает следующими свойствами:

- Любые два симплекса из разбиения пересекаются по некоторой k-мерной грани.

- Если Δ — симплекс из барицентрического разбиения, то $\operatorname{diam}\Delta \leq \frac{d}{d+1}\operatorname{diam} K$.

Построим для K барицентрическое разбиение, затем для каждого получившегося симплекса опять построим барицентрическое разбиение, и т.д. n раз. Тогда у каждого элемента разбиения диаметр будет не больше $\left(\frac{d}{d+1}\right)^n \operatorname{diam} K$; это число можно сделать сколь угодно малым за счет выбора n. При этом любые два симплекса из разбиения либо не пересекаются, либо пересекаются по некоторой k-мерной грани. Пусть $\{x_j\}_{j=1}^{N_n}$ — множество всех вершин симплексов, $y_j \in \mathbb{R}^d$, $1 \leq j \leq N_n$. Определим функцию f_n на K следующим образом: если $x = \sum\limits_{j\in J} \lambda_j x_j$, $\operatorname{conv}\{x_j\}_{j\in J}$ — один из симплексов разбиения, $\lambda_j \geq 0$, $\sum\limits_{j\in J} \lambda_j = 1$, то полагаем $f_n(x) = \sum\limits_{j\in J} \lambda_j y_j$. Эта функция непрерывна.

Метрическая проекция. Пусть на $\mathbb{R}^d$ введена евклидова метрика. Рассмотрим непустое выпуклое замкнутое множество $M \subset \mathbb{R}^d$. Для каждой точки $x \in \mathbb{R}^d$ существует единственная ближайшая точка $p(x) \in M$, т.е. такая, что

$$\|x - p(x)\| = \min_{y\in M} \|x - y\|.$$

Доказательство. Так как $y \mapsto \|x - y\|$ непрерывна и стремится к $+\infty$ при $\|y\| \to \infty$, а пересечение M с любым замкнутым шаром компактно, то ближайшая точка существует.

Если $p(x)$, $q(x)$ — две различные ближайшие точки, $\|x - p(x)\| = \|x - q(x)\| = r$, то $\left\|x - \frac{p(x)+q(x)}{2}\right\| < r$ и точки ближайшими не будут. $\square$

Если M — прямая, то $p(x)$ — ортогональная проекция. Значит, $\|p(x) - p(y)\| \leq \|x - y\|$, x, $y \in \mathbb{R}^d$. Отсюда следует, что то же неравенство верно, если M — отрезок. Значит, это неравенство выполнено и в случае, если M — произвольное выпуклое замкнутое множество.

Таким образом, отображение p липшицево с константой 1.

Доказательство теоремы Какутани. Сначала сведем утверждение к случаю, когда K — симлекс. Пусть Δ — симплекс, содержащий компакт K. Для каждой точки $x \in \Delta$ берем точку $p(x)$ — ближайшую из K. Положим $\tilde{\Phi}(x) = \Phi(p(x))$. Это отображение будет продолжением Φ на Δ. Если $(x_n, y_n) \underset{n\to\infty}{\to} (x, y)$, $y_n \in \tilde{\Phi}(x_n)$, то $(p(x_n), y_n) \underset{n\to\infty}{\to} (p(x), y)$, $y_n \in \tilde{\Phi}(x_n) =$

$\Phi(p(x_n))$. Значит, $y \in \Phi(p(x)) = \tilde{\Phi}(x)$. Если теорема доказана для симплекса, то существует точка $\hat{x} \in \Delta$ такая, что $\hat{x} \in \tilde{\Phi}(\hat{x}) = \Phi(p(\hat{x})) \subset K$. Значит, $p(\hat{x}) = \hat{x}$ и $\hat{x} \in \Phi(\hat{x})$.

Теперь докажем теорему для симплекса. Построим барицентрическое разбиение K, затем для каждого построенного симплекса строим барицентрическое разбиение и т.д. Пусть $\{x_{n,j}\}_{j=1}^{N_n}$ — множество всех вершин построенных на n-м шаге симплексов. Возьмем $y_{n,j} \in \Phi(x_{n,j})$ и построим кусочно-аффинную непрерывную функцию f_n такую, что $f_n(x_{n,j}) = y_{n,j}$. Так как $y_{n,j} \in \Phi(x_{n,j}) \subset K$ и K выпукло, то $f_n : K \to K$. По теореме Брауэра, существует точка $x_n \in K$ такая, что $f_n(x_n) = x_n$. Тогда $x_n = \sum\limits_{j \in J_n} \lambda_{n,j} x_{n,j}$, $f_n(x_n) = \sum\limits_{j \in J_n} \lambda_{n,j} y_{n,j}$ (здесь J_n — множество вершин одного из симплексов, полученным на n-м шаге). Перенумеровав вершины, считаем, что $J_n = \{1, \ldots, d+1\}$.

Существует подпоследовательность $\{n_k\}_{k \in \mathbb{N}}$ такая, что $x_{n_k} \underset{k\to\infty}{\to} \hat{x}$, $\lambda_{n_k,j} \underset{k\to\infty}{\to} \lambda_j$, $y_{n_k,j} \underset{k\to\infty}{\to} \hat{y}_j$. Так как диаметр разбиения стремится к 0 при $k \to \infty$, то $x_{n_k,j} \underset{k\to\infty}{\to} \hat{x}$. Значит, $\hat{y}_j \in \Phi(\hat{x})$. Так как $\Phi(\hat{x})$ выпукло, то $\sum\limits_{j=1}^{d+1} \lambda_j y_j \in \Phi(\hat{x})$. С другой стороны,

$$\sum_{j=1}^{d+1} \lambda_j y_j = \lim_{k\to\infty} \sum_{j=1}^{d+1} \lambda_{n_k,j} y_{n_k,j} = \lim_{k\to\infty} f_{n_k}(x_{n_k}) = \lim_{k\to\infty} x_{n_k} = \hat{x}.$$

Поэтому $\hat{x} \in \Phi(\hat{x})$. □

Доказательство теоремы о седловой точке. Мы уже определили отображение $(x, y) \mapsto A_y \times C_x$,

$$A_y = \operatorname{Argmin} f(\cdot, y) \subset K_X, \quad C_x = \operatorname{Argmax} f(x, \cdot) \subset K_Y.$$

Так как f непрерывна, выпукла по первому аргументу и вогнута по второму, то A_y и C_x выпуклы и замкнуты. Пусть $x_n \in K_X$, $y_n \in K_Y$, $u_n \in A_{y_n}$, $v_n \in C_{x_n}$, $x_n \underset{n\to\infty}{\to} x$, $u_n \underset{n\to\infty}{\to} u$, $y_n \underset{n\to\infty}{\to} y$, $v_n \underset{n\to\infty}{\to} v$. Покажем, что $u \in A_y$ (аналогично доказывается, что $v \in C_x$). В самом деле, иначе $f(u, y) > \min_{\xi \in K_X} f(\xi, y) = f(\hat{\xi}, y)$. Тогда при больших n получаем $f(u_n, y_n) > f(\hat{\xi}, y_n)$ — противоречие с тем, что $u_n \in A_{y_n}$. Значит, к построенному отображению можно применить теорему Какутани и получить седловую точку. □

Пусть $(\hat{x},\, \hat{y})$ — седловая точка, т.е. $f(x,\, \hat{y}) \geq f(\hat{x},\, \hat{y}) \geq f(\hat{x},\, y)$ для любых $x \in K_X$, $y \in K_Y$. Тогда

$$\inf_{x \in K_X} f(x,\, \hat{y}) \geq f(\hat{x},\, \hat{y}) \geq \sup_{y \in K_Y} f(\hat{x},\, y),$$

$$\sup_{y \in K_Y} \inf_{x \in K_X} f(x,\, y) \geq f(\hat{x},\, \hat{y}) \geq \inf_{x \in K_X} \sup_{y \in K_Y} f(x,\, y).$$

Всегда верно противоположное неравенство:

$$\sup_{y \in K_Y} \inf_{x \in K_X} f(x,\, y) \leq \inf_{x \in K_X} \sup_{y \in K_Y} f(x,\, y).$$

Значит,

$$\sup_{y \in K_Y} \inf_{x \in K_X} f(x,\, y) = f(\hat{x},\, \hat{y}) = \inf_{x \in K_X} \sup_{y \in K_Y} f(x,\, y).$$

20 Задача линейного программирования

Пусть $c \in \mathbb{R}^n$, $a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $1 \leq i \leq m$. Задача линейного программирования имеет вид

$$\langle c,\, x \rangle \to \inf, \quad \langle a_i,\, x \rangle \leq b_i, \quad 1 \leq i \leq m. \tag{31}$$

Обозначим через A матрицу со строками a_i, $1 \leq i \leq m$, и положим $b = (b_1,\, \ldots,\, b_m)$. Для векторов $\xi = (\xi_1,\, \ldots,\, \xi_m)$, $\eta = (\eta_1,\, \ldots,\, \eta_m) \in \mathbb{R}^m$ пишем $\xi \leq \eta$, если $\xi_i \leq \eta_i$, $1 \leq i \leq m$. Тогда (31) переписывается в виде

$$\langle c,\, x \rangle \to \inf, \quad Ax \leq b. \tag{32}$$

Напомним, что множество $G = \{x \in \mathbb{R}^n :\ Ax \leq b\}$ называется многогранным множеством. Вершинами G называются его крайние точки.

Для $x \in G$ мы обозначали $I(x) = \{i \in \overline{1,\, m} :\ \langle a_i,\, x \rangle = b_i\}$. Было доказано (см. предложение 42), что точка $\overline{x} \in G$ является вершиной тогда и только тогда, когда система уравнений

$$\langle a_i,\, x \rangle = b_i, \quad i \in I(\overline{x}),$$

имеет единственное решение. Отсюда сделали вывод, что у многогранного множества конечное число вершин.

Предложение 55. *Пусть* $G = \{x \in \mathbb{R}^n : \ Ax \le b\}$ *— непустое многогранное множество. Тогда следующие условия эквивалентны:*

1. *G не имеет вершин;*

2. *G содержит прямую;*

3. *существует вектор $h \ne 0$ такой, что $Ah = 0$;*

4. $\operatorname{rk} A < n$.

Доказательство. Эквивалентность 3 и 4 — факт из линейной алгебры.

$3 \Rightarrow 2$. Пусть $x_0 \in G$, $h \ne 0$, $Ah = 0$. Тогда для любого $t \in \mathbb{R}$ выполнено $\langle a_i, x_0 + th\rangle = \langle a_i, x_0\rangle \le b_i$, $1 \le i \le m$; значит, $\{x_0 + th : \ t \in \mathbb{R}\} \subset G$.

$2 \Rightarrow 3$. Пусть прямая $\{x_0 + th : \ t \in \mathbb{R}\} \subset G$. Покажем, что $Ah = 0$. В самом деле, иначе существует i такое, что $\langle a_i, h\rangle \ne 0$. Тогда, если выбрать t достаточно большим по модулю и с подходящим знаком, получим $\langle a_i, x_0 + th\rangle > b_i$ и $x_0 + th \notin G$.

$3 \Rightarrow 1$. Пусть $h \ne 0$, $Ah = 0$. Если $\overline{x}$ — вершина G, то система $\langle a_i, x\rangle = b_i$, $i \in I(\overline{x})$, имеет единственное решение. Но $\langle a_i, \overline{x} + h\rangle = \langle a_i, \overline{x}\rangle = b_i$, $i \in I(\overline{x})$. Значит, $\overline{x}$ — не вершина.

$1 \Rightarrow 3$. Пусть система $Ah = 0$ имеет только нулевое решение. Покажем, что G имеет крайнюю точку.

Пусть $\overline{x}$ — точка из G, для которой $I(\overline{x})$ имеем максимальную мощность. Покажем, что $\overline{x}$ — вершина.

В самом деле, пусть $\overline{x} \pm h \in G$ для некоторого $h \ne 0$. Тогда существует $i \in \overline{1, m}$ такое, что $\langle a_i, h\rangle \ne 0$.

Если $j \in I(\overline{x})$, то $\langle a_j, \overline{x} \pm h\rangle \le b_j = \langle a_j, \overline{x}\rangle$. Значит, $\pm\langle a_j, h\rangle \le 0$, откуда $\langle a_j, h\rangle = 0$.

Отсюда следует, что $I(\overline{x} + th) \supset I(\overline{x})$ для любого $t \in \mathbb{R}$ и $\{j : \ \langle a_j, h\rangle \ne 0\} \subset \{1, \ldots, m\} \backslash I(\overline{x})$. Мы уже знаем, что прямая $\{\overline{x} + th : \ t \in \mathbb{R}\}$ не содержится в G (см. доказательство $2 \Rightarrow 3$). Значит, $\{\overline{x} + th : \ t \in \mathbb{R}\} \cap G$ — это луч или отрезок. Пусть $\overline{x} + t_0 h$ — один из его концов. Отметим, что $t_0 \ne 0$. Существует $j_0 \notin I(\overline{x})$ такое, что $\langle a_{j_0}, \overline{x} + t_0 h\rangle = b_{j_0}$ (иначе, если $\langle a_j, \overline{x} + t_0 h\rangle < b_j$, $j \notin I(\overline{x})$, то при малых $\varepsilon > 0$ выполнено $\langle a_j, \overline{x} + (t_0 \pm \varepsilon)h\rangle < b_j$, $j \notin I(\overline{x})$, так что $\overline{x} + (t_0 \pm \varepsilon)h \in G$). Значит, $I(\overline{x} + t_0 h) \supset I(\overline{x}) \cup \{j_0\}$, так что $I(\overline{x})$ не имеет максимальную мощность. $\square$

Пусть $G \neq \varnothing$ не имеет вершины. Обозначим $L = \{h \in \mathbb{R}^n : \ Ah = 0\}$.

Пусть существует $h \in L$ такое, что $\langle c, h\rangle \neq 0$. Можно считать, что $\langle c, h\rangle < 0$. Пусть $x_0 \in G$. Тогда $\{x_0 + th : \ t \in \mathbb{R}\} \subset G$, при этом $\lim\limits_{t\to+\infty} \langle c, x_0 + th\rangle = -\infty$. Значит, $\inf\{\langle c, x\rangle : \ x \in G\} = -\infty$.

Пусть $\langle c, h\rangle = 0$ для любого $h \in L$. Рассмотрим задачу

$$\langle c, x\rangle \to \inf, \ Ax \leq b, \ \ x \in L^{\perp}.$$

Это задача линейного программирования на пространстве $L^{\perp}$. Множество $G \cap L^{\perp}$ уже будет иметь вершину: если $h \in L^{\perp}$ и $Ah = 0$, то $h \in L^{\perp} \cap L = \{0\}$.

В этой задаче inf такой же, как в (32). В самом деле, пусть $x \in G$. Тогда $x = \tilde{x} + h$, $\tilde{x} \in L^{\perp}$, $h \in L$. Значит, $A\tilde{x} = Ax - Ah = Ax \leq b$, $\langle c, \tilde{x}\rangle = \langle c, x\rangle - \langle c, h\rangle = \langle c, x\rangle$.

Таким образом, (32) сводится к задаче на многогранном множестве с вершиной. Далее будем считать, что G имеет хотя бы одну вершину.

20.1 Теорема о точечном конусе

Всюду здесь рассматриваем только замкнутые конусы в $\mathbb{R}^n$.

Определение 25. *Конус называется точечным, если* 0 *— его крайняя точка.*

Предложение 56. *Пусть* $M \subset \mathbb{R}^n$ *— выпуклый компакт,* $0 \notin M$. *Тогда* $K = \{tx : \ x \in M, \ t \geq 0\}$ *является замкнутым точечным конусом (более того,* 0 *является выступающей точкой).*

Доказательство. Ясно, что если $x \in K$, $\lambda \geq 0$, то $\lambda x \in M$. Пусть x, $y \in K$. Покажем, что $x + y \in K$. Имеем $x = t\xi$, $y = s\eta$, t, $s \geq 0$, ξ, $\eta \in M$. Достаточно рассмотреть случай, когда $t > 0$, $s > 0$. Получаем

$$x + y = t\xi + s\eta = (t + s)\left(\frac{t}{t+s}\xi + \frac{s}{t+s}\eta\right) \in K$$

(здесь использовалась выпуклость M).

Докажем, что K замкнутый. Пусть $\{t_k x_k\}_{k\in\mathbb{N}} \subset K$ — сходящаяся последовательность. Так как $0 \notin M$, то $\inf_{x\in M} \|x\| > 0$. Значит, последовательность $\{t_k\}_{k\in\mathbb{N}}$ ограничена. Поэтому существует подпоследовательность $\{k_j\}_{j\in\mathbb{N}}$ такая, что $\exists \lim\limits_{j\to\infty} x_{k_j} =: x$, $\exists \lim\limits_{j\to\infty} t_{k_j} =: t$. Тогда $x \in M$, $t \geq 0$, $t_k x_k \underset{k\to\infty}{\to} tx \in K$.

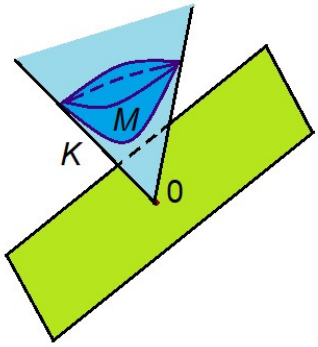

Рис. 34: Построение точечного конуса

Покажем, что 0 — выступающая точка множества K. По теореме о строгой отделимости, существует функционал x^* такой, что $\inf_{x\in M} x^*(x) > 0$. Тогда, если $x \in M$, $t > 0$, то $x^*(tx) > 0$. Поэтому $K \subset \{z:\ x^*(z) \geq 0\}$, $K \cap \{z:\ x^*(z) = 0\} = \{0\}$. □

Предложение 57. *Пусть $M \subset \mathbb{R}^n$ — выпуклый компакт, $0 \notin M$, $K = \{tx:\ t \geq 0,\ x \in M\}$. Пусть $x^* \in (\mathbb{R}^n)^*$, $\inf_{x\in M} x^*(x) > 0$. Тогда множество $\tilde{M} = \{\xi \in K:\ x^*(\xi) = 1\}$ компактно.*

Доказательство. Так как K замкнут, то $\tilde{M}$ замкнуто. Покажем, что $\tilde{M}$ ограничено. В самом деле, пусть $t \geq 0$, $x \in M$, $x^*(tx) = 1$. Тогда $t = 1/x^*(x) \leq \frac{1}{\inf_{z\in M} x^*(z)}$. Значит, $\|tx\| \leq \frac{\sup_{y\in M}\|y\|}{\inf_{z\in M} x^*(z)}$. □

Определение 26. *Пусть K — конус, $x_0 \in K\backslash\{0\}$. Луч $r = \{tx_0:\ t \geq 0\}$ называется образующей K, если для любых x, $y \in K$ таких, что $x+y \in r$ следует, что x, $y \in r$.*

Предложение 58. *Пусть $x^* \in (\mathbb{R}^n)^*\backslash\{0\}$, $M \subset \{x \in \mathbb{R}^n:\ x^*(x) = 1\}$ — выпуклый компакт, $K = \{tx:\ x \in M,\ t \geq 0\}$, $x_0 \in M$, $r = \{tx_0:\ t \geq 0\}$. Тогда следующие условия эквивалентны: 1) x_0 — крайняя точка M, 2) r — образующая K.*

Доказательство. $1 \Rightarrow 2$. Пусть x_0 — крайняя точка M, $tx_0 = t_1x_1+t_2x_2$, x_1, $x_2 \in M$, t_1, t_2, $t \geq 0$. Сразу можно считать, что $t_1 > 0$, $t_2 > 0$. Имеем $t = tx^*(x_0) = t_1x^*(x_1)+t_2x^*(x_2) = t_1+t_2$. Тогда $x_0 = \frac{t_1}{t_1+t_2}x_1 + \frac{t_2}{t_1+t_2}x_2$. Так как x_0 — крайняя точка множества M, то $x_1 = x_2 = x_0$, поэтому $t_1x_1 \in r$, $t_2x_2 \in r$.

$2 \Rightarrow 1$. Пусть r — образующая. Если $x_0 = \frac{x_1+x_2}{2}$, x_1, $x_2 \in M$, то x_1, $x_2 \in r$, т.е. $x_1 = t_1 x_0$, $x_2 = t_2 x_0$. Так как $x^*(x_1) = x^*(x_2) = x^*(x_0) = 1$, то $t_1 = t_2 = 1$. $\square$

Теорема 33. *Пусть $K \subset \mathbb{R}^n$, $K \neq \{0\}$ — точечный конус. Тогда*

$$K = \operatorname{conv}\{r : \ r \text{ — образующая}\}.$$

Доказательство. Пусть $S = K \cap \{x : \ \|x\| = 1\}$, $M = \operatorname{conv} S$. По следствию из теоремы Каратеодори (см. следствие 1), M — компакт. Покажем, что $0 \notin M$. Пусть $0 = \sum\limits_{j=1}^{m} \lambda_j x_j$, $\lambda_j > 0$, $x_j \in S$, $1 \leq j \leq m$, $\sum\limits_{j=1}^{m} \lambda_j = 1$. Так как $0 \notin S$, то $m \geq 2$. Значит,

$$0 = \lambda_1 x_1 + (1-\lambda_1) \sum_{j=2}^{m} \frac{\lambda_j}{\sum\limits_{i=2}^{m} \lambda_i} x_j =: \lambda_1 x_1 + (1-\lambda_1) y,$$

при этом $y \neq x_1$, $y \in M \subset K$. Поэтому 0 не является крайней точкой конуса K — противоречие.

Заметим, что

$$K = \{tx : \ x \in M, \ t \geq 0\}.$$

По теореме о строгой отделимости, существует функционал $x^* \in (\mathbb{R}^n)^* \backslash \{0\}$ такой, что $\inf_{x \in M} x^*(x) > 0$. В силу предложения 57, множество

$$\tilde{M} = \{x \in K : \ x^*(x) = 1\}$$

компактно. По теореме Крейна–Мильмана, $\tilde{M} = \operatorname{conv} \Sigma$, где Σ — множество крайних точек M.

Имеем: $K = \{tx : \ x \in \tilde{M}, \ t \geq 0\}$. Из предложения 58 следует, что объединение множества образующих K имеет вид $\{tx : \ x \in \Sigma, \ t \geq 0\}$.

Пусть $x \in \tilde{M}$, $t \geq 0$. Имеем $x = \sum\limits_{j=1}^{k} \lambda_j \xi_j$, $\xi_j \in \Sigma$, $\lambda_j \geq 0$, $1 \leq j \leq k$, $\sum\limits_{j=1}^{k} \lambda_j = 1$. Значит, $tx = \sum\limits_{j=1}^{k} \lambda_j t \xi_j$ является выпуклой комбинацией точек, принадлежащих образующим K. $\square$

Замечание. Если точечный конус имеет вид $K = \{x : \ Ax \leq 0\}$, где A — матрица, то $\tilde{M}$ из доказательства теоремы 33 является компактным многогранным множеством (а значит, многогранником). Поэтому у $\tilde{M}$ конечное число вершин (см. следствие 7), а у K — конечное число образующих.

Замечание. Если K — точечный конус, то 0 — его выступающая точка. (Это следует из доказательства теоремы 33 и предложения 56.)

20.2 Основная теорема линейного программирования

Рассматриваем задачу (32). Предположим, что $G = \{x \in \mathbb{R}^n : \ Ax \leq b\}$ непусто. Пусть $\operatorname{rk} A = n$. Тогда G имеет вершину.

Пусть $\overline{x}$ — некоторая вершина G. Сделав сдвиг, можно считать, что $\overline{x} = 0$. Тогда $b_i = 0$ при $i \in I(0)$, $b_i > 0$ при $i \notin I(0)$.

Рассмотрим множество

$$K = \{x \in \mathbb{R}^n : \ \langle a_j, \, x \rangle \leq 0, \ j \in I(0)\}.$$

Тогда K — замкнутый конус, содержащий G.

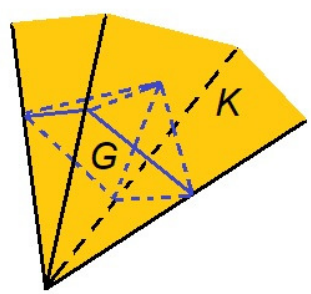

Рис. 35: Множества G и K

Применим предложение 42. Так как 0 — вершина G, то система $\langle a_j, \, x \rangle = 0$, $j \in I(0)$, имеет единственное решение. Значит, 0 — вершина K, т.е. конус K точечный.

Замечание. Так как 0 является выступающей точкой точечного конуса, то вершина многогранного множества также является выступающей точкой.

Пусть $K \neq \{0\}$. В силу теоремы о точечном конусе и замечания после нее,

$$K = \operatorname{conv}\{r_j\}_{j=1}^{s},$$

где r_j — образующие.

Имеем $r_i = \{t\xi_i : \ t \geq 0\}$, где ξ_i — некоторые ненулевые векторы $(1 \leq i \leq s)$. Если $j \in I(0)$, то $\langle a_j, t\xi_i\rangle \leq 0$. Если $j \notin I(0)$, то $b_j > 0$, поэтому $\langle a_j, t\xi_i\rangle < b_j$ при малых $t \geq 0$. Значит, $r_i \cap G$ — невырожденный отрезок или луч. Это множество называется ребром G, выходящим из $\overline{x} = 0$. Если вершина $\overline{x}$ ненулевая, то рассматриваем $G - \overline{x}$, определяем его ребра l_i $(1 \leq i \leq s)$, выходящие из 0; ребрами G назовем множества $l_i + \overline{x}$ $(1 \leq i \leq s)$.

Теорема 34. *Пусть $S = \inf\{\langle c, x\rangle : \ Ax \leq b\}$. Пусть $G := \{x \in \mathbb{R}^n : Ax \leq b\} \neq \varnothing$ имеет вершину. Тогда $S = -\infty$ или минимум достигается в некоторой вершине множества G. При этом минимум достигается в вершине $\overline{x}$ тогда и только тогда, когда для любого ребра l, выходящего из $\overline{x}$, выполнено $\langle c, \overline{x}\rangle \leq \langle c, x\rangle$, $x \in l$.*

Доказательство. Рассмотрим некоторую вершину $\overline{x} \in G$. Без ограничения общности, $\overline{x} = 0$. Пусть $K = \{x \in \mathbb{R}^n : \ \langle a_j, x\rangle \leq 0, \ j \in I(0)\}$, $r_i = \{t\xi_i : \ t \geq 0\}$ $(1 \leq i \leq s)$ — образующие K, $l_i \subset r_i$ — ребра G, выходящие из $\overline{x}$.

1. Пусть $\langle c, \xi_i\rangle \geq 0$ для любого $i \in \{1, \ldots, s\}$. Тогда $\langle c, x\rangle \geq 0$ для любого $x \in \cup_{i=1}^{s} r_i$. Так как $K = \operatorname{conv}(\cup_{i=1}^{s} r_i)$, то $\langle c, x\rangle \geq 0$ для любого $x \in K$ и, тем более, для любого $x \in G$. Значит, вершина $\overline{x} = 0$ — точка минимума.

2. Пусть существует $i \in \{1, \ldots, s\}$ такое, что $\langle c, \xi_i\rangle < 0$. Тогда 0 — не точка минимума. Если l_i неограниченно, то

$$S \leq \lim_{t \to +\infty} \langle c, t\xi_i\rangle = -\infty.$$

   Пусть l_i ограничено, т.е. $l_i = \{t\xi_i : \ 0 \leq t \leq t_0\}$, $t_0 > 0$. Тогда $\langle c, t_0\xi_i\rangle < 0$. Покажем, что $t_0\xi_i$ — вершина G. В самом деле, пусть $t_0\xi_i = \frac{x+y}{2}$, $x, y \in G$. Тогда $x, y \in K$; так как $\xi_i \in r_i$ и r_i — образующая, то $x, y \in r_i$, т.е. $x = t\xi_i$, $y = s\xi_i$; поскольку $x, y \in G$, то $0 \leq t \leq t_0$, $0 \leq s \leq t_0$. Значит, $x = y = t_0\xi_i$.

Таким образом, нашлась вершина G, в которой значение функционала $\langle c, \cdot\rangle$ оказалось строго меньше нуля. Применяем к этой вершине описанную выше процедуру. Получаем, что либо она точка минимума, либо $S = -\infty$, либо есть следующая вершина, в которой значение функционала $\langle c, \cdot\rangle$ строго меньше, и т.д. Так как в G конечное число вершин, то в какой-то момент процесс останавливается: либо находится точка минимума, либо оказывается, что $S = -\infty$.

Теорема доказана. □

Итак, получили алгоритм поиска точки минимума (симплекс-метод):

1. Пусть нам дана вершина $\overline{x} \in G$ (как ее искать, расскажем позже). Найдем все ребра l_i $(1 \le i \le s)$, выходящие из $\overline{x}$, и сравним значение $\langle c, \cdot\rangle$ в точке $\overline{x}$ и на ребрах.

2. Если $\langle c, \overline{x}\rangle \le \langle c, x\rangle$ для любого $i \in \{1, \ldots, s\}$, $x \in l_i$, то $\overline{x}$ — точка минимума.

3. Пусть $\langle c, \overline{x}\rangle > \langle c, x\rangle$ для некоторого $x \in l_i$, $i \in \{1, \ldots, s\}$. Если l_i неограниченно, то $S = -\infty$. Если l_i ограничено, то находим его ненулевой конец. Это будет новая вершина, в которой значение $\langle c, \cdot\rangle$ строго меньше, чем в $\overline{x}$.

4. Делаем то же самое для новой вершины. Процесс закончится за конечное число шагов.

20.3 Реализация симплекс-метода в невырожденном случае

Пусть $\overline{x}$ — вершина G, $\operatorname{card} I(\overline{x}) = n$. Сначала опишем множество ребер, выходящих из $\overline{x}$.

Снова можем считать, что $\overline{x} = 0$. Также без ограничения общности считаем, что $I(\overline{x}) = \{1, \ldots, n\}$. Как и раньше,

$$K = \{x \in \mathbb{R}^n : \ \langle a_i, x\rangle \le 0, \ 1 \le i \le n\}.$$

Предложение 59. *Точка $\xi \in K\backslash\{0\}$ принадлежит образующей конуса K тогда и только тогда, когда существует $j \in \{1, \ldots, n\}$ такое, что $\langle a_i, \xi\rangle = 0$ при $i \ne j$ и $\langle a_j, \xi\rangle < 0$.*

Доказательство. Система уравнений $\langle a_i,\, \xi\rangle = 0$, $1 \le i \le n$, имеет только нулевое решение.

Множество

$$\{x \in \mathbb{R}^n :\ \langle a_i,\, x\rangle = 0,\ i \in \overline{1,\, n},\ i \ne j\}$$

является одномерным подпространством.

Пусть $\langle a_i,\, \xi\rangle = 0$ при $i \ne j$ и $\langle a_j,\, \xi\rangle < 0$. Докажем, что $l = \{t\xi :\ t \ge 0\}$ является образующей. Пусть $\xi = x + y$, $x,\ y \in K$. Тогда $\langle a_i,\, x\rangle \le 0$, $\langle a_i,\, y\rangle \le 0$. Если $i \ne j$, то $\langle a_i,\, x\rangle + \langle a_i,\, y\rangle = \langle a_i,\, \xi\rangle = 0$ и поэтому $\langle a_i,\, x\rangle = \langle a_i,\, y\rangle = 0$. Значит, $x,\, y \in l$.

Пусть существуют два различных индекса $j,\, k \in \{1,\, \ldots,\, n\}$ таких, что $\langle a_j,\, \xi\rangle < 0$, $\langle a_k,\, \xi\rangle < 0$. Докажем, что $l = \{t\xi :\ t \ge 0\}$ не является образующей. В самом деле, система $\langle a_i,\, x\rangle = 0$, $i \notin \{j,\, k\}$, имеет решение h, не пропорциональное ξ. При этом $\xi \pm \varepsilon h \in K$ для достаточно малых $\varepsilon > 0$. $\square$

Пусть $\tilde{A}$ — подматрица A, образованная первыми n строками.

Мы доказали, что образующие K имеют вид $r_j = \{t\xi_j :\ t \ge 0\}$, где ξ_j — решение уравнения $\tilde{A}x = -e_j$, где e_j — стандартные базисные векторы.

Найдем ненулевой конец ребра в G, лежащего в r_j, или покажем, что оно неограниченно.

Пусть $\langle a_i,\, \xi_j\rangle \le 0$, $n+1 \le i \le m$. Тогда для любого $i \in \{1,\, \ldots,\, m\}$, $t \ge 0$ выполнено $\langle a_i,\, t\xi_j\rangle \le b_i$. Значит, ребро неограниченно.

Пусть $J = \{i \in \overline{n+1,\, m} :\ \langle a_i,\, \xi_j\rangle > 0\}$ непусто. Пусть t_0 — максимум из положительных чисел t таких, что $\langle a_i,\, t\xi_j\rangle \le b_i$ для всех $i \in J$, т.е. $t_0 = \min_{i\in J} \frac{b_i}{\langle a_i, \xi_j\rangle}$. Тогда $t_0\xi_j$ — ненулевой конец ребра.

Теперь выясним, когда $\langle c,\, \xi_j\rangle \ge 0$ для всех $j \in \{1,\, \ldots,\, n\}$ (в этом случае 0 — точка минимума), а когда хотя бы одно из скалярных произведений отрицательно (тогда 0 — не точка минимума).

Если $z \in \mathbb{R}^n$ — такой вектор, что $\langle c,\, \xi_j\rangle = \langle z,\, e_j\rangle$ для любого $j = 1,\, \ldots,\, n$, то достаточно исследовать знаки координат z. Напомним, что $e_j = -\tilde{A}\xi_j$, т.е. $\langle c,\, \xi_j\rangle = -\langle z,\, \tilde{A}\xi_j\rangle = -\langle \tilde{A}^* z,\, \xi_j\rangle$. Значит, в качестве z можно взять решение уравнения $-\tilde{A}^* x = c$.

В итоге действуем так: находим решение z уравнения $-\tilde{A}^* x = c$ и исследуем знаки его координат. Если у z все координаты неотрицательны, то $\langle c,\, \xi_j\rangle \ge 0$ для всех $j \in \{1,\, \ldots,\, n\}$ и вершина является точкой минимума. Если j-я координата строго отрицательна, то проверяем, будет ли

соответствующее ребро ограниченным и если да, то ищем его ненулевой конец. Это будет следующая вершина.

20.4 Поиск первой вершины

Предположим, что $\operatorname{rk} A = n$.

Пусть $\overline{x} \in G$. Находим множество $I(\overline{x})$. Если система $\langle a_i,\, x\rangle = b_i$, $i \in I(\overline{x})$, имеет единственное решение, то $\overline{x}$ — вершина.

Пусть решение системы $\langle a_i,\, x\rangle = b_i$ $(i \in I(\overline{x}))$ неединственно. Найдем точку $x_0 \in G$ такую, что $\operatorname{card} I(x_0) > \operatorname{card} I(\overline{x})$. (Дальше продолжаем эту процедуру до тех пор, пока не сможем увеличить мощность; так как $\operatorname{card} I(x) \le m$, то в какой-то момент получим вершину.)

Точка x_0 находится следующим образом.

Пусть h — ненулевое решение системы $\langle a_i,\, x\rangle = 0$, $i \in I(\overline{x})$. Рассмотрим прямую $\{\overline{x} + th :\ t \in \mathbb{R}\}$. Она не содержится в G (так как G имеет вершину; см. предложение 55). Можно считать, что существует $t > 0$ такое, что $\overline{x} + th \notin G$. Пусть $J = \{j \in \overline{1,\, m} \backslash I(\overline{x}) :\ \langle a_i,\, h\rangle > 0\}$. Это множество непусто, иначе $\overline{x} + th \in G$ для любого $t \ge 0$. Пусть t_0 — максимум из положительных t таких, что $\langle a_j,\, \overline{x} + th\rangle \le b_j$ для любого $j \in J$, т.е. $t_0 = \min_{j\in J} \frac{b_j - \langle a_j,\, \overline{x}\rangle}{\langle a_j,\, h\rangle}$, минимум достигается на j_0. Тогда $\overline{x} + t_0 h \in G$, $I(\overline{x} + t_0 h) \supset I(\overline{x}) \cup \{j_0\}$.

20.5 Поиск точки, принадлежащей G

Пусть $x = (x_1,\, \ldots,\, x_n) \in \mathbb{R}^n$. Рассмотрим вспомогательную задачу в $\mathbb{R}^{n+1}$, добавив координату x_0:

$$x_0 \to \inf,\ \ \langle a_j,\, x\rangle \le b_j + x_0,\ 1 \le j \le m,\ \ x_0 \ge 0. \tag{33}$$

Если минимум равен 0, то $G \ne \varnothing$, иначе $G = \varnothing$.

В этой задаче множество допустимых точек непусто (можно взять $x = 0$ и достаточно большое x_0).

Итак, сначала решаем задачу (33): находим вершину и применяем симплекс-метод; либо оказывается, что $G = \varnothing$, либо находим точку из G. Затем ищем вершину G, а после этого решаем (31) симплекс-методом.

21 Двойственная задача

Пусть $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция. Исследуется задача

$$f(x) \to \inf, \quad x \in \mathbb{R}^n.$$

Пусть $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ — выпуклая функция, $f(x) = F(x, 0)$. Положим

$$S(y) = \inf_{x \in \mathbb{R}^n} F(x, y).$$

В частности, $S(0) = \inf_{x \in \mathbb{R}^n} f(x)$.

Предложение 60. *Функция S выпукла.*

Доказательство. Докажем выпуклость надграфика. Пусть $(y_1, u_1), (y_2, u_2) \in \operatorname{epi} S$, $\lambda \in [0, 1]$. Покажем, что

$$((1-\lambda)y_1 + \lambda y_2, (1-\lambda)u_1 + \lambda u_2) \in \operatorname{epi} S. \tag{34}$$

Для любого $\varepsilon > 0$

$$u_1 + \varepsilon > S(y_1), \ u_2 + \varepsilon > S(y_2).$$

Значит, существуют $x_1, x_2 \in \mathbb{R}^n$ такие, что

$$u_1 + \varepsilon > F(x_1, y_1), \ u_2 + \varepsilon > F(x_2, y_2).$$

Так как F выпукла, то для любого $\lambda \in [0, 1]$

$$F((1-\lambda)x_1 + \lambda x_2, (1-\lambda)y_1 + \lambda y_2) \le (1-\lambda)u_1 + \lambda u_2 + \varepsilon;$$

отсюда $S((1-\lambda)y_1 + \lambda y_2)) \le (1-\lambda)u_1 + \lambda u_2 + \varepsilon$. В силу произвольности $\varepsilon > 0$, получаем $S((1-\lambda)y_1 + \lambda y_2)) \le (1-\lambda)u_1 + \lambda u_2$, т.е. выполнено (34). $\square$

Пусть функция S еще замкнутая и собственная. Тогда

$$S(0) = S^{**}(0) = \sup_{y^* \in \mathbb{R}^m} (-S^*(y^*)).$$

Имеем

$$S^*(y^*) = \sup_y (\langle y^*, y\rangle - S(y)) = \sup_y (\langle 0, x\rangle + \langle y^*, y\rangle - \inf_x F(x, y)) =$$

$$= \sup_{x,\,y}(\langle 0,\, x\rangle + \langle y^*,\, y\rangle - F(x,\, y)) = F^*(0,\, y^*).$$

Таким образом,

$$\inf_x f(x) = S(0) = \sup_{y^*}(-S^*(y^*)) = \sup_{y^*}(-F^*(0,\, y^*)).$$

Задача

$$-F^*(0,\, y^*) \to \sup$$

называется двойственной.

Рассмотрим задачу линейного программирования вида

$$\langle c,\, x\rangle \to \inf, \quad Ax \geq b, \quad x \geq 0 \tag{35}$$

($x \in \mathbb{R}^n$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A : \mathbb{R}^n \to \mathbb{R}^m$ — линейный оператор).

Положим

$$f(x) = \begin{cases} \langle c,\, x\rangle, & Ax \geq b,\ x \geq 0, \\ +\infty, & \text{иначе.} \end{cases}$$

Тогда (35) эквивалентна задаче $f(x) \to \inf$, $x \in \mathbb{R}^n$.

Пусть

$$F(x,\, y) = \begin{cases} \langle c,\, x\rangle, & Ax \geq b + y,\ x \geq 0, \\ +\infty, & \text{иначе} \end{cases}$$

(здесь $y \in \mathbb{R}^m$).

Найдем $F^*(0,\, y^*)$:

$$F^*(0,\, y^*) = \sup\{\langle y^*,\, y\rangle - \langle c,\, x\rangle :\ Ax \geq b + y,\ x \geq 0\} =$$

$$= \sup\{\langle y^*,\, y\rangle - \langle c,\, x\rangle :\ Ax = b + y + z,\ x \geq 0,\ z \geq 0\} =$$

$$= \sup\{\langle y^*,\, Ax - b - z\rangle - \langle c,\, x\rangle :\ x \geq 0,\ z \geq 0\} =$$

$$= \sup\{\langle A^*y^* - c,\, x\rangle - \langle y^*,\, z\rangle - \langle y^*,\, b\rangle :\ x \geq 0,\ z \geq 0\} =$$

$$= \begin{cases} -\langle y^*,\, b\rangle, & y^* \geq 0,\ A^*y^* \leq c, \\ +\infty, & \text{иначе.} \end{cases}$$

Значит, двойственная задача имеет вид

$$\langle y^*,\, b\rangle \to \sup, \quad A^*y^* \leq c, \quad y^* \geq 0.$$

Следующее утверждение приводим без доказательства.

Предложение 61. *Функция S замкнутая.*

Рассмотрим случаи.

1. S собственная. Тогда $\inf_x f(x) = \sup_{y^*}(-F^*(0,\, y^*))$.

   Если $S(0) < +\infty$, то в исходной и двойственной задаче экстремальное значение конечно.

   Если $S(0) = +\infty$, то $f \equiv +\infty$ и множество допустимых точек пусто.

2. Пусть S несобственная.

   Если $S \equiv +\infty$, то $S(0) = +\infty$ и снова множество допустимых точек в исходной задаче пусто.

   Если $S(y) = -\infty$ для некоторого y, то $S^* \equiv +\infty$ и в двойственной задаче множество допустимых точек пусто.

Вывод. Либо в обеих задачах экстремальные значения конечны и совпадают, либо хотя бы в одной их них множество допустимых точек пусто.

22 Различные формы задач линейного программирования. Примеры из экономики.

Рассмотрим три задачи:

$$\langle c,\, x\rangle \to \inf, \quad Ax \le b, \tag{36}$$

$$\langle c,\, x\rangle \to \inf, \quad Ax \ge b,\ x \ge 0, \tag{37}$$

$$\langle c,\, x\rangle \to \inf, \quad Ax = b,\ x \ge 0. \tag{38}$$

Покажем, что эти задачи сводятся друг к другу.

Задача (37) является частным случаем задачи (36).

Задача (36) эквивалентна следующей:

$$\langle c,\, x'\rangle - \langle c,\, x''\rangle \to \inf, \quad Ax' - Ax'' \le b,\ x' \ge 0,\ x'' \ge 0;$$

она имеет форму (37).

Задача (38) переписывается в виде

$$\langle c,\, x\rangle \to \inf, \quad Ax \geq b,\ Ax \leq b,\ x \geq 0;$$

она имеет форму (37).

Задача (37) переписывается в виде

$$\langle c,\, x\rangle \to \inf, \quad Ax - z = b,\ x \geq 0,\ z \geq 0;$$

она имеет форму (38).

Рассмотрим два примера из экономики.

Задача оптимального планирования производства. Производится n видов продукции (в количестве $x_1, \ldots, x_n$ единиц). Пусть $r_i > 0$ — цена i-го вида продукции. На производство тратится m видов сырья; при этом на единицу i-го продукта — $a_{i,j}$ единиц j-го сырья. В запасе есть b_j единиц j-го сырья. Мы хотим максимизировать прибыль. Получаем задачу

$$\sum_{i=1}^{n} r_i x_i \to \max,$$

$$\sum_{i=1}^{n} a_{i,j} x_i \leq b_j,\ 1 \leq j \leq m,$$

$$x_i \geq 0,\ 1 \leq i \leq n.$$

Транспортная задача. Есть n потребителей и m карьеров с песком. Из i-го карьера j-му потребителю перевозим $x_{i,j}$ тонн песка, это стоит $a_{i,j}$ рублей. С i-го карьера увозим всего c_i тонн песка, j-му потребителю привозим всего b_j тонн песка. Нужно минимизировать стоимость перевозки.

Получаем задачу:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} x_{i,j} \to \min,$$

$$\sum_{j=1}^{m} x_{i,j} = c_i, \quad 1 \leq i \leq n,$$

$$\sum_{i=1}^{n} x_{i,j} = b_j, \quad 1 \leq j \leq m,$$

$$x_{i,j} \geq 0, \quad 1 \leq i \leq n,\ 1 \leq j \leq m.$$