4 Выпуклый анализ

Пусть в этом пункте X — линейное нормированное пространство (для простоты понимания можно считать, что $X = \mathbb{R}^n$ — конечномерное пространство).

4.1 Элементы выпуклого анализа

4.1.1 Выпуклые множества

Введем некоторые понятия, которые используются в выпуклом анализе:

- *отрезок* $[a, b] =$
  $= \{x \in X \mid x = a + t(b-a) = (1-t)a + tb,\ 0 \le t \le 1\}$;
- *интервал* $(a, b) = \{x \in X \mid x = (1-t)a + tb,\ 0 < t < 1\}$;
- *выпуклое множество* A, если $\forall\, a, b \in A$ отрезок $[a, b] \subset A$, т. е. $\forall\ a, b \in A$ точка $(1-t)a + tb \in A\ \forall\ 0 \le t \le 1$;
- *конус* K $(K \ne \emptyset)$, если $\forall\ x \in K$ точка $tx \in K\ \forall\ t \ge 0$;
- *аффинное множество* A, если $\forall\ a, b \in A$ точка $(1-t)a + tb \in A\ \forall\ t \in \mathbb{R}$.

Очевидно: аффинное множество выпукло.

Возьмем фиксированные точки $a_1, \dots, a_m \in X$. Пусть $\sum\limits_{i=1}^{m} t_i a_i$ — комбинация этих точек, $t_i \in \mathbb{R}$. Дадим различные определения таких комбинаций и оболочек этих комбинаций:

- *выпуклая комбинация*, если $t_i \ge 0$, $\sum\limits_{i=1}^{m} t_i = 1$;
- *коническая комбинация*, если $t_i \ge 0$;
- *аффинная комбинация*, если $\sum\limits_{i=1}^{m} t_i = 1$;

- *выпуклая оболочка* $\operatorname{conv}\{a_1, \dots, a_m\} := \Big\{ a = \sum\limits_{i=1}^m t_i a_i \,\Big|\, t_i \ge 0,$ $\sum\limits_{i=1}^m t_i = 1 \Big\}$ — совокупность всех выпуклых комбинаций;
- *коническая выпуклая оболочка* $\operatorname{conev}\{a_1, \dots, a_m\} := \Big\{ a = \sum\limits_{i=1}^m t_i a_i \,\Big|\, t_i \ge 0 \Big\}$;
- *аффинная оболочка* $\operatorname{aff}\{a_1, \dots, a_m\} := \Big\{ a = \sum\limits_{i=1}^m t_i a_i \,\Big|\, \sum\limits_{i=1}^m t_i = 1 \Big\}$.

4.1.2 Выпуклые функции

Пусть задана функция (функционал), отображающая линейное нормированное пространство в “расширенную” прямую:

$$f\colon X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}.$$

С каждой такой функцией f связываются два множества:
$\operatorname{dom} f = \{x \in X \mid f(x) < +\infty\}$ — *эффективное множество*
и $\operatorname{epi} f = \{(a, x) \in \mathbb{R} \times X \mid a \ge f(x)\}$ — *надграфик* функции f.

- Функция f называется *выпуклой*, если надграфик f — выпуклое множество.
- Функция f называется *замкнутой*, если надграфик f — замкнутое множество.
- Функция f называется *собственной*, если $f(x) > -\infty \ \forall\, x$ и $f \not\equiv +\infty$.

Мы будем изучать выпуклые собственные функции. Для краткости будем называть их просто “выпуклые” функции.

Из определения выпуклого множества сразу следует, что функция выпукла тогда и только тогда, когда выполнено *неравенство Йенсена*:

$$f(tx_1 + (1-t)x_2) \le t f(x_1) + (1-t) f(x_2) \quad \forall\, x_1, x_2 \in X, \quad \forall\, t \in (0,\, 1).$$

Ясно, что сумма двух выпуклых функций является функцией выпуклой. Но суперпозиция двух выпуклых функций не всегда является выпуклой функцией. Приведите пример.

Выпуклость многих классических функций одной переменной сразу вытекает из следующей теоремы.

Теорема 1 (Р, с. 44)**.** *Пусть функция* $f\colon \mathbb{R} \to \mathbb{R}$ *дважды непрерывно дифференцируема* ($f \in D^2(\mathbb{R})$)*. Тогда она выпукла тогда и только тогда, когда ее вторая производная неотрицательна* ($f''(x) \geq 0 \ \forall \ x \in \mathbb{R}$)*.*

Приведем несколько примеров выпуклых функций.

1. $f(x) = e^{ax}$, $a \in \mathbb{R}$.

2. $f(x) = |x|^p$, $p \geq 1$.

Выпуклость функций из примеров 1-2 сразу следует из теоремы 1 и определения выпуклой функции.

3. *Аффинная функция* (в многомерном случае $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle a, x\rangle + b$, $a \in \mathbb{R}^n$, $b \in \mathbb{R}$).

Аффинная функция является выпуклой по неравенству Йенсена. Модуль аффинной функции $|\langle a, x\rangle + b|$ также является выпуклой функцией, поскольку ее надграфик — пересечение полупространств — выпуклое множество.

Выпуклость функций нескольких переменных можно определять также из следующего многомерного обобщения теоремы 1.

Теорема 2 (Р, с. 44)**.** *Пусть функция* $f\colon \mathbb{R}^n \to \mathbb{R}$ *дважды непрерывно дифференцируема* ($f \in D^2(\mathbb{R}^n)$)*. Тогда она выпукла тогда и только тогда, когда ее матрица вторых производных (гессиан) всюду неотрицательно определена*

$$\left(f''(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{i,j=1}^{n} \geq 0 \ \forall \ x \in \mathbb{R}^n \right).$$

4. *Квадратичная функция* $Q(x) = \langle Ax, x\rangle$ (A — симметричная матрица), является выпуклой тогда и только тогда, когда матрица A неотрицательно определена.

Это сразу вытекает из теоремы 2.

Выпуклыми функциями многих переменных (функционалами) являются также следующие функции:

5. *Функция нормы*

$$f(x) = \|x\|_p := \begin{cases} \left(\sum\limits_{j=1}^{n} |x_j|^p \right)^{\frac{1}{p}}, & 1 \le p < +\infty, \\ \max\{|x_1|, \ldots, |x_n|\}, & p = +\infty. \end{cases}$$

6. *Индикаторная функция* выпуклого множества $A \subset X$

$$\delta A(x) = \begin{cases} 0, & x \in A, \\ +\infty, & x \notin A. \end{cases}$$

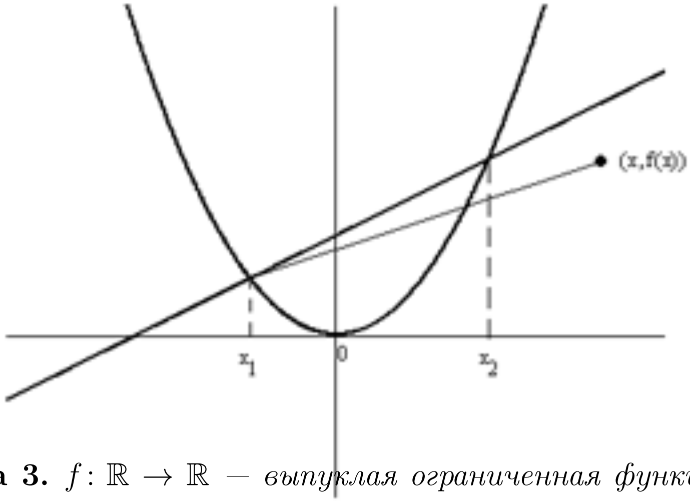
7. *Функция Минковского* выпуклого множества $A \subset X$

$$\mu A(x) = \begin{cases} 0, & \alpha^{-1} x \in A \ \forall\ \alpha > 0, \\ +\infty, & \alpha^{-1} x \notin A \ \forall\ \alpha > 0, \\ \inf\{\alpha \mid \alpha > 0,\ \alpha^{-1} x \in A\}, & \text{иначе.} \end{cases}$$

Функция Минковского означает наименьшее число, в которое надо уменьшить вектор x, чтобы он попал в множество A.

8. *Опорная функция* непустого множества $A \subset X$

$$sA(x^*) = \max_{x \in A} \langle x^*, x \rangle \qquad (sA\colon X^* \to \mathbb{R}).$$



Теорема 3. $f\colon \mathbb{R} \to \mathbb{R}$ — *выпуклая ограниченная функция* $\Rightarrow$ $f(x) = \mathrm{const}$.

Доказательство. Предположим противное, что $f(x) \neq \mathrm{const}$. Тогда существуют точки x_1, x_2 такие, что $x_1 < x_2$ и $f(x_1) \neq f(x_2)$. Поскольку f — выпуклая ограниченная функция, то надграфик функции $\operatorname{epi} f$ — выпуклое множество, следовательно, отрезок $\big[(x_1, f(x_1)), (x_2, f(x_2))\big] \subset \operatorname{epi} f$. Значит, на промежутке $[x_1, x_2]$ график функции f лежит ниже этого отрезка (или совпадает с ним). Не ограничивая общности, будем считать, что $f(x_1) < f(x_2)$. Пусть $l(x)$ — прямая, проходящая через точки $(x_1, f(x_1))$ и $(x_2, f(x_2))$. Покажем, что

$$f(x) \geq l(x) \quad \forall\, x > x_2. \tag{$*$}$$

Если это не так, то существует точка $x > x_2$ такая, что $f(x) < l(x)$, т. е. точка $(x, f(x))$ лежит ниже прямой l. Так как функция f выпукла, то отрезок $\big[(x_1, f(x_1)), (x, f(x))\big] \subset \operatorname{epi} f$ и лежит ниже прямой l, но это не так в точке x_2 (ниже точки $(x_2, f(x_2))$ не может быть точек из $\operatorname{epi} f$). Поэтому неравенство $(*)$ действительно выполняется.

Так как $l(x) \to +\infty$ при $x \to +\infty$, то $f(x) \to +\infty$ при $x \to +\infty$. Получаем противоречие с ограниченностью функции f. Значит, предположение противного неверно и $f(x) = \mathrm{const}$. $\square$

4.1.3 Субдифференциал выпуклой функции

Дадим определение важного понятия выпуклого анализа — субдифференциала функции, обобщающего для выпуклых функций понятие производной в гладком анализе.

Субдифференциалом выпуклой функции f в точке $\hat{x}$ называется следующее множество в сопряженном пространстве X^*:

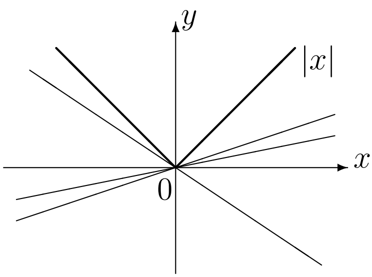
$$\partial f(\hat{x}) = \{x^* \in X^* \mid \langle x^*, x - \hat{x} \rangle \le f(x) - f(\hat{x}) \quad \forall\, x \in X\}.$$

Напомним, что *сопряженным пространством* X^* называется пространство линейных непрерывных функционалов на X. В случае $X = \mathbb{R}^n$ сопряженное пространство $(\mathbb{R}^n)^* = \mathbb{R}^n$.

Из определения сразу вытекает, что субдифференциал — выпуклое множество в X^*. Легко доказать, что оно замкнуто. Субдифференциал дифференцируемой функции совпадает с ее производной.

Для функции одной переменной субдифференциал $\partial f(\hat{x})$ — это совокупность угловых коэффициентов k, при которых прямые $y = kx + b$, проходящие через точку $(\hat{x}, f(\hat{x}))$, лежат под графиком функции $y = f(x)$.

Пример. $f(x) = |x|$.



$$\partial|x| = \begin{cases} \operatorname{sign} x, & x \ne 0, \\ [-1,\, 1], & x = 0. \end{cases}$$

Рис. 6

Для субдифференциала суммы функций имеет место теорема аналогичная теореме о производной суммы функций.

Теорема (Моро–Рокафеллар). *[16, с. 49] Пусть f_1 и f_2 — выпуклые функции на X. Существует точка x_0, в которой функция f_1 конечна ($|f_1(x_0)| < \infty$), а функция f_2 непрерывна ($f \in C(x_0)$). Тогда*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) \quad \forall\ x.$$

При доказательстве необходимых условий экстремума в гладкой задаче с равенствами и неравенствами нам понадобится следующая теорема о субдифференциале максимума.

Теорема (Дубовицкий–Милютин). *[16, с. 51] Пусть f_1 и f_2 — выпуклые функции, непрерывные в точке $\hat{x}$, $f_1(\hat{x}) = f_2(\hat{x})$. Тогда*

$$\partial \max\{f_1, f_2\}(\hat{x}) = \operatorname{conv}\big(\partial f_1(\hat{x}) \cup \partial f_2(\hat{x})\big).$$

Важным примером выпуклой функции является сублинейная функция. Функция p называется *сублинейной*, если ее надграфик является выпуклым конусом с вершиной в нуле.

Из неравенства Йенсена следует, что собственная функция является сублинейной тогда и только тогда, когда
a) $p(\lambda x) = \lambda p(x)$ для любого $\lambda > 0$, для любого $x \in X$;
b) $p(x_1 + x_2) \leq p(x_1) + p(x_2)$ для любых $x_1, x_2 \in X$.

Если X — нормированное пространство, то $x \to \|x\|$ — сублинейная функция.

Сформулируем теорему Моро–Рокафеллара и теорему Дубовицкого–Милютина для сублинейных функций.

Теорема (Моро–Рокафеллар). *Пусть p_1, p_2 — сублинейные функции, функция p_1 непрерывна, функция p_2 замкнута. Тогда в точке $\hat{x} = 0$*

$$\partial(p_1 + p_2) = \partial p_1 + \partial p_2.$$

Теорема (Дубовицкий–Милютин). *Пусть p_1, p_2 — непрерывные сублинейные функции. Тогда*

$$\partial \max\{p_1, p_2\}(0) = \operatorname{conv}(\partial p_1(0) \cup \partial p_2(0)).$$

Субдифференциал модуля функции

Пусть $f(x)$ — дифференцируемая функция n переменных. Рассмотрим функцию $|f(x)|$. Найдем ее субдифференциал.

Если $f(x) < 0$, то $|f(x)| = -f(x)$ и $\partial|f(x)| = -f'(x)$.

Если $f(x) > 0$, то $|f(x)| = f'(x)$ и $\partial|f(x)| = f(x)$.

Если $f(x) = 0$, то запишем функцию $|f(x)|$ в виде $|f(x)| = \max\{-f(x), f(x)\}$. Тогда по теореме Дубовицкого–Милютина

$$\partial|f(x)| = \partial \max\{-f(x), f(x)\} = \operatorname{conv}\big\{\partial(-f)(x) \cup \partial f(x)\big\} =$$

$$= \operatorname{conv}\{-f'(x), f'(x)\} = [-f'(x);\ f'(x)].$$

Субдифференциал является отрезком в пространстве $\mathbb{R}^n$. Таким образом,

$$\partial|f| = \begin{cases} -f'(x), & f(x) < 0, \\ \{\alpha f'(x),\ \alpha \in [-1;\ 1]\}, & f(x) = 0, \\ f'(x), & f(x) > 0. \end{cases}$$

Субдифференциал нормы

Пусть X — линейное нормированное пространство, функционал $f\colon X \to \mathbb{R}$, $f(x) = \|x\|$. Найдем субдифференциал нормы в нуле.

По определению субдифференциала

$$x^* \in \partial f(\hat{x}) \iff \langle x^*, x - \hat{x} \rangle \le f(x) - f(\hat{x}) \ \forall\, x \in X.$$

У нас $\hat{x} = 0$, $f(\hat{x}) = \|0\| = 0$. Следовательно,

$$x^* \in \partial f(0) \iff \langle x^*, x \rangle \le \|x\| \quad \forall\, x \in X. \tag{$*$}$$

Пусть $x^* \in \partial f(0)$. По лемме Банаха существует вектор $x \in X$ такой, что $\|x\| = 1$, $\langle x^*, x \rangle = \|x^*\|$. Подставляя этот вектор x в неравенство $(*)$, получим

$$\|x^*\| = \langle x^*, x \rangle \overset{(*)}{\le} \|x\| = 1 \Longrightarrow \|x^*\| \le 1.$$

Значит, $\partial f(0) \subset B^* := \big\{x^* \in X^* \mid \|x^*\|_{X^*} \le 1\big\}$ — единичный шар сопряженного пространства. Докажем, что $\partial f(0) = B^*$.

Для этого надо доказать обратное включение $\partial f(0) \supset B^*$. Возьмем $x^* \in B^*$, тогда $|x^*| \le 1$. По неравенству Коши–Буняковского

$$\langle x^*, x \rangle \le \|x\| \cdot \|x^*\| \overset{\|x^*\| \le 1}{\le} \|x\| \ \forall\, x \in X.$$

Неравенство $(*)$ выполняется, значит, $x^* \in \partial f(0)$. Таким образом, $\partial f(0) = B^* = \{x^* \in X^* \mid \|x^*\|_{X^*} \le 1\}$.

4.2 Отделимость выпуклых множеств

При выводе необходимых условий экстремума (принципа Лагранжа) в выпуклых задачах и в задачах с равенствами и неравенствами мы будем использовать свойство отделимости непересекающихся выпуклых множеств.

Определение 1. *Множества A и B из пространства X называются отделимыми, если существует линейный непрерывный функционал $\lambda \in X^*$, $\lambda \neq 0$, для которого*

$$\inf_{a\in A} \langle \lambda, a\rangle \geq \sup_{b\in B} \langle \lambda, b\rangle.$$

Из определения следует, что множества являются отделимыми, если можно провести гиперплоскость $H = \{x \in X \mid \langle \lambda, x\rangle = c\}$, $c \in \mathbb{R}$, так что одно из множеств лежит в одном замкнутом полупространстве $H_+ = \{x \in X \mid \langle \lambda, x\rangle \geq c\}$, а другое — в другом замкнутом полупространстве $H_- = \{x \in X \mid \langle \lambda, x\rangle \leq c\}$.

Замечание. Поскольку функционал λ в определении отделимости является непрерывным, то для отделимости (не строгой!) вместо множеств A и B можно рассматривать их внутренности или замыкания.

Определение 2. *Множества A и B называются строго отделимыми, если существует линейный непрерывный функционал $\lambda \in X^*$, для которого*

$$\inf_{a\in A} \langle \lambda, a\rangle > \sup_{b\in B} \langle \lambda, b\rangle.$$

4.2.1 Теоремы отделимости

Приведем результаты об отделимости в конечномерном случае и в случае линейных нормированных пространств.

Теорема 1 (отделимости множеств в конечномерном пространстве)**.** *Пусть A и B — непустые выпуклые множества в $\mathbb{R}^n$,* $\operatorname{int} A \cap B = \emptyset$ (*в частности,* $A \cap B = \emptyset$)*. Тогда множества A и B отделимы.*

Доказательство. Поскольку отделимость множеств $\operatorname{int} A$ и B равносильна отделимости множеств A и B, то, не ограничивая общности, считаем, что $A \cap B = \emptyset$. Обозначим $C := A - B = \{a - b \mid a \in A,\ b \in B\}$ (разность по Минковскому множеств A и B). Поскольку $A \cap B = \emptyset$, то $0 \notin C$. Надо доказать, что существует вектор $\lambda \neq 0$, для которого

$$\inf_{a \in A} \langle \lambda, a \rangle \geq \sup_{b \in B} \langle \lambda, b \rangle \iff \inf_{c \in C} \langle \lambda, c \rangle = \langle \lambda, 0 \rangle \geq 0,$$

т. е., гиперплоскость $\langle \lambda, c \rangle = 0$ отделяет C от начала координат — точки O.

1) Предположим, что $0 \notin \overline{C}$. Рассмотрим задачу

$$f(x) = |x| \to \inf; \quad x \in \overline{C}. \tag{P}$$

Так как f — непрерывная функция и $f(x) \to +\infty$ при $|x| \to \infty$, то по следствию из т. Вейерштрасса абсолютный минимум в задаче (P) достигается. Поэтому существует вектор $\lambda \in \operatorname{absmin} P$, $\lambda \in \overline{C}$. Причем, $\lambda \neq 0$ в силу условия $0 \notin \overline{C}$. Покажем, что функционал λ искомый. Если $\inf\limits_{c \in C} \langle \lambda, c \rangle < 0$, то существует точка $c \in C$, для которой $\langle \lambda, c \rangle < 0$, т. е., векторы λ и c образуют тупой угол. Значит, основание H высоты OH треугольника $Oc\lambda$ лежит на отрезке $[c;\ \lambda]$ (и, в силу выпуклости множества $\overline{C}$, принадлежит $\overline{C}$), и $|OH| < |O\lambda|$, что противоречит выбору точки λ.

2) Предположим, что $0 \in \overline{C}$. Так как $0 \notin C$, то $0 \in \partial C$. Поэтому существует последовательность $y_k \to 0$, $y_k \notin \overline{C}$. В пункте 1) доказано, что для таких точек существуют векторы $\lambda_k \neq 0$, что выполняется неравенство: $\inf\limits_{c \in C} \langle \lambda_k, c \rangle \geq \langle \lambda_k, y_k \rangle$. Не ограничивая общности, считаем, что $|\lambda_k| = 1 \ \forall \ k$. Сфера в $\mathbb{R}^n$ компактна. Значит, можно перейти к подпоследовательности λ_k, сходящейся к некоторой точке λ, $|\lambda| = 1$. При этом $|\langle \lambda_k, y_k \rangle| \leq |\lambda_k| \cdot |y_k| = |y_k| \to 0$. Следовательно,

$$\inf_{c \in C} \langle \lambda, c \rangle = \inf_{c \in C} \lim_{k \to \infty} \langle \lambda_k, c \rangle \geq \lim_{k \to \infty} \langle \lambda_k, y_k \rangle \to 0 \Rightarrow \inf_{c \in C} \langle \lambda, c \rangle \geq 0.$$

То есть, λ – искомый функционал. □

Теорема 2 (отделимости множеств в нормированном пространстве)**.** *[АТФ, с. 124] Пусть A и B — непустые выпуклые множества в X, $\operatorname{int} A \neq \emptyset$, $\operatorname{int} A \cap B = \emptyset$. Тогда множества A и B отделимы.*

В теореме отделимости в конечномерном пространстве $\operatorname{int} A$ может быть пуста. Напомним, что из определения выпуклого множества следует, что пустое множество является выпуклым. И вместо условия $\operatorname{int} A \cap B = \emptyset$ можно писать $A \cap B = \emptyset$. В теореме отделимости в бесконечномерном пространстве это не так. Надо требовать, чтобы $\operatorname{int} A \neq \emptyset$.

Пример двух непересекающихся выпуклых подмножеств, которые нельзя отделить.

Рассмотрим в бесконечномерном гильбертовом пространстве $l_2 = \left\{ x = (x_1, x_2, \ldots) \mid \sum\limits_{k=1}^{\infty} x_k^2 < \infty \right\}$ два множества с пустыми внутренностями: множество $A = \left\{ a \in l_2 \mid \sum\limits_{k=1}^{+\infty} a_k = 1 \right\}$ и точку $b = 0$. Множество A имеет пустую внутренность, поскольку никакой шар не содержится в A. Покажем, что эти множества нельзя отделить. Т. е надо показать, что не существует ненулевого линейного непрерывного функционала $\lambda \in l_2^*$ (напомним, что l_2^* изоморфно l_2) такого, что

$$\inf_{a \in A} \langle \lambda, a \rangle \geq \langle \lambda, b \rangle = 0. \qquad (*)$$

Возьмем произвольный функционал $\lambda \in l_2^*$, $\lambda \neq 0$. Покажем, что $\inf\limits_{a \in A} \langle \lambda, a \rangle = -\infty$, тогда неравенство $(*)$ не будет выполняться. Поскольку $\lambda = (\lambda_1, \lambda_2, \ldots) \neq 0$, то $\exists\ \lambda_i \neq \lambda_j$ (для определенности, $\lambda_i < \lambda_j$). Возьмем точку $a = (a_1, a_2, \ldots)$, у которой следующие компоненты: $a_i = n$, $a_j = 1 - n$, $a_k = 0$ при $k \neq i, j$. Тогда точка $a \in A$ и

$$\langle \lambda, a \rangle = \lambda_i a_i + \lambda_j a_j = \lambda_i n + \lambda_j (1 - n) = n(\lambda_i - \lambda_j) + \lambda_j \to -\infty$$

при $n \to +\infty$. Значит, множество A и точку b нельзя отделить.

4.2.2 Теоремы о строгой отделимости

Теорема 1 (о строгой отделимости точки от множества в нормированном пространстве)**.** *Пусть A — непустое выпуклое замкнутое множество в X, $b \notin A$. Тогда точку b можно строго отделить от множества A.*

Доказательство. Поскольку точка $b \notin A$ — замкнутому множеству, то существует $\varepsilon > 0$, для которого $A \cap (b + \varepsilon U) = \emptyset$, где $U := \{x \in X \mid \|x\| \leq 1\}$ — единичный шар пространства X. Следовательно, по теореме отделимости в нормированном пространстве выпуклые множества A и $B := b + \varepsilon U$ $(\operatorname{int} B \neq \emptyset$, $\operatorname{int} B \cap A = \emptyset)$ можно отделить, т. е., существует линейный непрерывный функционал $\lambda \in X^*$, $\lambda \neq 0$ такой, что

$$\inf_{a\in A}\langle \lambda, a\rangle \geq \sup_{x\in b+\varepsilon U}\langle \lambda, x\rangle = \sup_{u\in U}\langle \lambda, b+\varepsilon u\rangle =$$

$$= \langle \lambda, b\rangle + \varepsilon \sup_{u\in U}\langle \lambda, u\rangle = \langle \lambda, b\rangle + \varepsilon\|\lambda\|_{X^*}.$$

Отсюда $\inf\limits_{a\in A}\langle \lambda, a\rangle > \langle \lambda, b\rangle$. Это и означает, что точка b строго отделена от множества A. □

Рассмотрим следующее обобщение доказанной теоремы.

Теорема 2 (о строгой отделимости компакта)**.** *A и B — непустые выпуклые замкнутые подмножества нормированного пространства X, причем B — компакт и $A \cap B = \emptyset$. Тогда множества A и B строго отделимы.*

Доказательство. Обозначим $U := \{x \in X \mid \|x\| \leq 1\}$ — единичный шар пространства X. Покажем, что существует $\varepsilon > 0$, для которого $A \cap (B + \varepsilon U) = \emptyset$. В противном случае существуют последовательности $\{a_k\} \in A$, $\{b_k\} \in B$ такие, что $\|a_k - b_k\| \to 0$ при $k \to \infty$. Поскольку B — компакт, то из последовательности b_k можно выделить сходящуюся подпоследовательность. Для простоты записи считаем, что это подпоследовательность и есть сама последовательность b_k. Так как последовательности $\{a_k\}$ и $\{b_k\}$ сходятся к одной точке, а последовательность

$b_k \to b \in B$, то и последовательность $a_k \to b$ и $b \in A$ в силу замкнутости A. Таким образом, точка $b \in A \cap B$ — противоречие. Следовательно, $A \cap (B + \varepsilon U) = \emptyset$. Тогда по теореме отделимости ($A \cap \operatorname{int} B' = \emptyset$, $\operatorname{int} B' \neq \emptyset$) множества A и $B' = B + \varepsilon U$ можно отделить, т. е., существует линейный непрерывный функционал $\lambda \in X^*$, $\lambda \neq 0$ такой, что

$$\inf_{a\in A}\langle \lambda, a\rangle \geq \sup_{x\in B+\varepsilon U}\langle \lambda, x\rangle = \sup_{b\in B, u\in U}\langle \lambda, b + \varepsilon u\rangle =$$

$$= \sup_{b\in B}\langle \lambda, b\rangle + \varepsilon \sup_{u\in U}\langle \lambda, u\rangle = \sup_{b\in B}\langle \lambda, b\rangle + \varepsilon\|\lambda\|_{X^*}.$$

Отсюда $\inf\limits_{a\in A}\langle \lambda, a\rangle > \sup\limits_{b\in B}\langle \lambda, b\rangle$. Это и означает, что A и B строго отделимы. $\square$

Приведем еще одно доказательство теоремы о строгой отделимости компакта, использующее лемму о замкнутости суммы (разности) замкнутого множества и компакта, которая также имеет самостоятельный интерес.

Лемма (о замкнутости суммы (разности) замкнутого множества и компакта)**.** *Пусть X — нормированное пространство, множества $A, B \subset X$, A — замкнуто, B — компакт. Тогда алгебраическая сумма $C = A + B$ (разность $A - B$) замкнута.*

Доказательство. Пусть имеется сходящаяся последовательность точек $c_k \in C$, $c_k \to c_0$. Докажем, что $c_0 \in C$. По условию леммы $c_k = a_k + b_k$, где $a_k \in A$, $b_k \in B$. Так как B — компакт, то из последовательности b_k можно выделить сходящуюся подпоследовательность b_{k_m}. Тогда подпоследовательность b_{k_m} сходится к некоторой точке $b_0 \in B$. Значит, подпоследовательность $a_{k_m} = c_{k_m} - b_{k_m}$ будет сходиться к точке $c_0 - b_0 =: a_0$ и точка $a_0 \in A$ в силу замкнутости A. Значит, $c_0 = a_0 + b_0$, $a_0 \in A$, $b_0 \in B$, т. е. $c_0 \in C$. $\square$

Второе доказательство теоремы о строгой отделимости компакта

Доказательство. Обозначим $C := A - B$. Тогда множество C замкнуто по лемме о замкнутости суммы (разности) замкнутого множества и компакта и $0 \notin C$. Следовательно, по теореме о строгой отделимости точки от выпуклого замкнутого множества точку 0 можно строго отделить от множества C, т. е., существует линейный непрерывный функционал $\lambda \in X^*$, $\lambda \neq 0$ такой, что

$$\inf_{c \in C} \langle \lambda, c \rangle > 0 \iff \inf_{a \in A, b \in B} \langle \lambda, a - b \rangle > 0$$

$$\inf_{a \in A} \langle \lambda, a \rangle - \sup_{b \in B} \langle \lambda, b \rangle > 0 \iff \inf_{a \in A} \langle \lambda, a \rangle > \sup_{b \in B} \langle \lambda, b \rangle.$$

Это и означает, что множества A и B строго отделимы. □

Замечание. *Требование компактности одного из множеств в теореме о строгой отделимости существенно.*

Действительно, непересекающиеся замкнутые множества ось x и множество точек над гиперболой $\dfrac{1}{x}$, $x > 0$, не пересекаются, но не являются строго отделимыми.

4.3 Выпуклые задачи

4.3.1 Задачи без ограничений

Пусть $f\colon X \to \overline{\mathbb{R}}$ — выпуклая функция, отображающая нормированное пространство X в расширенную прямую. *Выпуклой задачей без ограничений* называется следующая экстремальная задача:

$$f(x) \to \min\,. \tag{P}$$

Теорема (аналог теоремы Ферма). *Для того чтобы точка* $\hat{x}$ *доставляла в выпуклой задаче без ограничений* (P) *абсолютный минимум* $(\hat{x} \in \operatorname{absmin} P)$*, необходимо и достаточно, чтобы выполнялось соотношение*

$$0 \in \partial f(\hat{x}).$$

Доказательство. $\hat{x} \in \operatorname{absmin} P \Leftrightarrow f(x) - f(\hat{x}) \geq 0 = \langle 0, x - \hat{x}\rangle \ \forall\, x \in X \Leftrightarrow 0 \in \partial f(\hat{x})$.

Поскольку из выпуклости функции f не следует, вообще говоря, выпуклость функции $-f$, то существенно, что рассматривается задача на минимум, а не на максимум.

4.3.2 Задачи с ограничением

Пусть $f\colon X \to \overline{\mathbb{R}}$ — выпуклая функция, отображающая нормированное пространство X в расширенную прямую, $D \subset X$ — выпуклое множество. *Выпуклой задачей с ограничением* (или просто *выпуклой задачей*) называется следующая экстремальная задача:

$$f(x) \to \min; \quad x \in D. \tag{P}$$

Теорема. *Пусть $\hat{x} \in locmin\, P$ — доставляет локальный минимум в выпуклой задаче (P). Тогда $\hat{x} \in \operatorname{absmin} P$ — доставляет абсолютный минимум в этой задаче.*

Доказательство. Пусть $\hat{x} \in \operatorname{locmin} P$. Это означает, что существует окрестность U точки $\hat{x}$ такая, что

$$f(\hat{x}) \leq f(x) \quad \forall\, x \in U \cap D. \tag{$*$}$$

Возьмем произвольную точку $x \in D$. Тогда, поскольку x и $\hat{x}$ из D, то выпуклая комбинация $\bar{x} = (1-\alpha)\hat{x} + \alpha x \in U \cap D$ при достаточно малом $\alpha > 0$. Следовательно, $f(\hat{x}) \overset{(*)}{\leq} f(\bar{x}) \overset{\mathrm{d}ef\, \bar{x}}{=} f((1-\alpha)\hat{x}+\alpha x) \leq$ (по неравенству Иенсена) $(1-\alpha)f(\hat{x})+\alpha f(x)$, откуда $\alpha f(\hat{x}) \leq \alpha f(x) \Leftrightarrow f(\hat{x}) \leq f(x)\ \forall\, x \in D$. Значит, $\hat{x}$ доставляет абсолютный минимум в задаче (P). $\triangleright$

Из теоремы следует, что в выпуклой задаче локальный минимум является и абсолютным (глобальным). Поэтому в дальнейшем в выпуклых задачах, говоря “минимум”, имеем в виду абсолютный минимум.

4.3.3 Задача выпуклого программирования

Пусть $f_i\colon X \to \overline{\mathbb{R}}$, $i = 0, 1, \ldots, m$, — выпуклые функции, отображающие линейное нормированное пространство X в расширенную прямую, $A \subset X$ — выпуклое множество. *Задачей выпуклого программирования* называется следующая экстремальная задача:

$$f_0(x) \to \min; \quad f_i(x) \le 0,\ i = 1, \ldots, m,\ x \in A. \qquad (P)$$

Точка x называется *допустимой* в задаче (P), если $x \in A$ и выполняются все заданные ограничения типа неравенств.

Упражнение. Докажите, что задача выпуклого программирования является выпуклой задачей, т. е., что множество допустимых элементов в этой задаче является выпуклым множеством.

При проверке достаточных условий минимума в задаче выпуклого программирования будет использоваться некоторое условие регулярности множества допустимых элементов — условие Слейтера. Множество допустимых элементов в задаче (P) удовлетворяет *условию Слейтера*, если существует точка $\bar{x} \in A$, для которой $f_i(\bar{x}) < 0$, $i = 1, \ldots, m$.

Теорема Каруша-Куна–Таккера. 1. *Если $\hat{x} \in \operatorname{absmin} P$ — решение задачи выпуклого программирования, то найдется ненулевой вектор множителей Лагранжа $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$ такой, что для функции Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ выполняются условия:*

a) принцип минимума функции Лагранжа: $\min\limits_{x \in A} \mathscr{L}(x) = \mathscr{L}(\hat{x})$;

b) дополняющей нежесткости: $\lambda_i f_i(\hat{x}) = 0,\ i = 1, \ldots, m$;

c) неотрицательности: $\lambda_i \ge 0,\ i = 0, 1, \ldots, m$.

2. *Если для допустимой точки $\hat{x}$ выполнены условия a) – c) и $\lambda_0 \ne 0$, то $\hat{x} \in \operatorname{absmin} P$.*

3. *Если для допустимой точки $\hat{x}$ выполнены условия a) – c) и множество допустимых элементов удовлетворяет условию Слейтера, то $\hat{x} \in \operatorname{absmin} P$.*

Доказательство. Пусть $\hat{x} \in \operatorname{absmin} P$. Не ограничивая общности, считаем, что $f_0(\hat{x}) = 0$ — иначе введем новую функцию $\tilde{f}_0(x) = f_0(x) - f_0(\hat{x})$. Положим

$$B = \{b = (b_0, \ldots, b_m) \in \mathbb{R}^{m+1} \mid \exists\, x \in A : f_i(x) \le b_i,\ i = 0, \ldots, m\}.$$

Покажем, что B — *непустое выпуклое множество.* Действительно, неотрицательный октант $\mathbb{R}^{m+1}_+ \subset B$, т. е. любой вектор с неотрицательными компонентами принадлежит B, ибо в определении множества B можно положить $x = \hat{x}$. Докажем выпуклость множества B. Пусть точки b и b' принадлежат множеству B. Надо доказать, что точка $\alpha b + (1-\alpha)b' \in B$ для любого $\alpha \in (0,\,1)$. Поскольку точки $b, b' \in B$, то по определению множества B существуют точки $x, x' \in A$ такие, что $f_i(x) \le b_i$, $f_i(x') \le b'_i$, $i = 0, 1, \ldots, m$. Положим $x_\alpha = \alpha x + (1-\alpha)x'$. Тогда $x_\alpha \in A$, поскольку множество A — выпукло, а ввиду выпуклости функций f_i, $i = 0, 1, \ldots, m$, по неравенству Йенсена

$$f_i(x_\alpha) = f_i\big(\alpha x + (1-\alpha)x'\big) \le \alpha f_i(x) + (1-\alpha) f_i(x') \le \alpha b_i + (1-\alpha) b'_i,$$

т. е. точка $\alpha b + (1-\alpha)b' \in B$.

Докажем, что точка $0 \notin \operatorname{int} B$. Если бы $0 \in \operatorname{int} B$, то при достаточно малом $\varepsilon < 0$ точка $(\varepsilon, 0, \ldots, 0) \in B$. Тогда по определению множества B существует точка $\tilde{x} \in A$ такая, что $f_0(\tilde{x}) \le \varepsilon < 0 = f_0(\hat{x})$, $f_i(\tilde{x}) \le 0$, $i = 1, \ldots, m$. Значит, $\tilde{x}$ является допустимой точкой ($\tilde{x} \in D(P)$) и $f_0(\tilde{x}) < f_0(\hat{x})$, т. е. $\hat{x} \notin \operatorname{absmin} P$ — противоречие. Следовательно, $0 \notin \operatorname{int} B$.

По теореме отделимости в конечномерном пространстве множество B и точку 0 можно отделить, т. е. существует вектор $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \neq 0$, для которого

$$\inf_{b \in B} \sum_{i=0}^{m} \lambda_i b_i \ge \langle \lambda, 0 \rangle = 0.$$

Таким образом,

$$\sum_{i=0}^{m} \lambda_i b_i \ge 0 \quad \forall\, b \in B\ . \qquad (*)$$

Из неравенства $(*)$ будут выведены условия неотрицательности, дополняющей нежесткости и принцип минимума.

1. *Условие неотрицательности.* Поскольку, как мы уже говорили, любой вектор с неотрицательными компонентами принадлежит B, то вектор $(0, \ldots, 0, 1, 0, \ldots, 0) \in B$, где единица стоит на i-ом месте (счет начинаем с нуля). Подставив эту точку в неравенство $(*)$, получим, что $\lambda_i \geq 0$.

Условие дополняющей нежесткости. Нетрудно видеть, что точка $(0, \ldots, 0, f_i(\hat{x}), 0, \ldots, 0) \in B$. Действительно, в определении множества B возьмем $x = \hat{x}$, тогда $x \in A$ и нужные неравенства выполняются: $f_j(\hat{x}) \leq 0$, $j \neq i$, $f_i(\hat{x}) \leq f_i(\hat{x})$. Подставив эту точку в неравенство $(*)$, имеем $\lambda_i f_i(\hat{x}) \geq 0$. Если $\lambda_i f_i(\hat{x}) > 0$, то (так как по уже доказанному условию неотрицательности $\lambda_i \geq 0$) $\lambda_i > 0 \Rightarrow f_i(\hat{x}) > 0$ — это противоречит допустимости точки $\hat{x}$. Значит, $\lambda_i f_i(\hat{x}) = 0$.

Принцип минимума. Возьмем точку $x \in A$, тогда точка $(f_0(x), f_1(x), \ldots, f_m(x)) \in B$. Поэтому по неравенству $(*)$

$$\mathscr{L}(x) = \sum_{i=0}^{m} \lambda_i f_i(x) \geq 0 = \sum_{i=0}^{m} \lambda_i f_i(\hat{x}) = \mathscr{L}(\hat{x}),$$

так как, не ограничивая общности, положили $f_0(\hat{x}) = 0$ и по уже доказанному условию дополняющей нежесткости $\lambda_i f_i(\hat{x}) = 0$, $i = 1, \ldots, m$. Таким образом, принцип минимума для функции Лагранжа доказан.

2. Пусть для допустимой точки $\hat{x}$ выполнены условия a)–c) и $\lambda_0 \neq 0$. Положим $\lambda_0 = 1$. Тогда для любой допустимой точки x будет выполняться неравенство

$$f_0(\hat{x}) \overset{b)}{=} f_0(\hat{x}) + \sum_{i=1}^{m} \lambda_i f_i(\hat{x}) \overset{\mathrm{def}}{=} \mathscr{L}(\hat{x}) \overset{a)}{\leq} \mathscr{L}(x) \overset{\mathrm{def}}{=} f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \overset{c)}{\leq} f_0(x)$$

(в последней сумме все слагаемые $\lambda_i f_i(x)$ неположительны: $\lambda_i \geq 0$, $f_i(x) \leq 0$). Неравенство $f_0(\hat{x}) \leq f_0(x)$ для любой допустимой точки x означает, что $\hat{x} \in \operatorname{absmin} P$.

3. Пусть для допустимой точки $\hat{x}$ выполнены условия a)–c) и множество допустимых элементов удовлетворяет условию Слейтера в точке $\bar{x}$ ($\exists\ \bar{x} \in A : f_i(\bar{x}) < 0,\ i = 1, \ldots, m$). Предположим, что при этом $\lambda_0 = 0$. Так как вектор $\lambda \neq 0$, то в силу условия неотрицательности существует хотя бы одно $\lambda_j > 0$, $j \in \{1, \ldots, m\}$. Следовательно, имеем

$$\mathscr{L}(\bar{x}) = \sum_{i=1}^{m} \lambda_i f_i(\bar{x}) \overset{c)}{\leq} \lambda_j f_j(\bar{x}) < 0 \overset{b)}{=} \sum_{i=1}^{m} \lambda_i f_i(\hat{x}) = \mathscr{L}(\hat{x})$$

(отметим, что в первой сумме все слагаемые $\lambda_i f_i(\bar{x})$ в силу условия c) неположительны: $\lambda_i \geq 0$, а $f_i(\bar{x}) < 0$). Но полученное неравенство $\mathscr{L}(\bar{x}) < \mathscr{L}(\hat{x})$ противоречит выполненному условию a). Значит, наше предположение, что $\lambda_0 = 0$ неверно. Поэтому $\lambda_0 \neq 0$ и по доказанному п. 2 $\hat{x} \in \operatorname{absmin} P$.

Теорема Каруша-Куна–Таккера полностью доказана. $\triangleright$

Из теоремы Каруша-Куна–Таккера можно вывести достаточные условия абсолютного минимума в выпуклой задаче с неравенствами без ограничения типа включения. Рассмотрим задачу

$$f_0(x) \to \min; \quad f_i(x) \leq 0,\ i = 1, \ldots, m. \qquad (P')$$

Теорема (достаточные условия абсолютного минимума в выпуклой задаче с неравенствами без ограничения типа включения). *Пусть $f_i \colon X \to \overline{\mathbb{R}}$, $i = 0, \ldots, m$, — выпуклые функции, $\hat{x}$ — допустимая точка в задаче (P') $(\hat{x} \in D_{P'})$, для функции Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ с множителем Лагранжа $\lambda_0 > 0$ в точке $\hat{x}$ выполняются условия:*

a) условие стационарности функции Лагранжа: $0 \in \partial \mathscr{L}(\hat{x})$;

b) дополняющей нежесткости: $\lambda_i f_i(\hat{x}) = 0,\ i = 1, \ldots, m$;

c) неотрицательности: $\lambda_i \geq 0,\ i = 1, \ldots, m$.

Тогда $\hat{x} \in \operatorname{absmin} P'$.

Доказательство. Поскольку $f_i \colon X \to \overline{\mathbb{R}}$, $i = 0, \ldots, m$, — выпуклые функции, то функция Лагранжа $\mathscr{L}(x) = \sum\limits_{i=0}^{m} \lambda_i f_i(x)$ с неотрицательными множителями Лагранжа является выпуклой функцией. По аналогу теоремы Ферма для выпуклых функций условие $0 \in \partial \mathscr{L}(\hat{x})$ является необходимым и достаточным условием абсолютного минимума функции Лагранжа в точке $\hat{x}$. Значит, условие стационарности функции Лагранжа в настоящей теореме равносильно принципу минимума для функции Лагранжа в теореме Каруша-Куна–Таккера. Таким образом, для нашей задачи выполняются соотношения $a) - c)$ теоремы Каруша-Куна–Таккера с множителем Лагранжа $\lambda_0 \neq 0$. По п. 2 теоремы Каруша-Куна–Таккера следует, что $\hat{x} \in \operatorname{absmin} P'$. $\rhd$

Легко видеть, что теорема остается верной и для задач с ограничениями типа равенств и неравенств, если функции, задающие равенства, являются аффинными. (В этом случае, как мы говорили об этом в п. 3. 2 равенство $f(x) = 0 \Leftrightarrow 0 \leq f(x) \leq 0$ заменяется двумя неравенствами $f(x) \leq 0$, $-f(x) \leq 0$. Из аффинности функции f следует выпуклость функций f и $-f$.)

4.3.4 Пример

$$f(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2 + 3|x_1 + x_2 - 2| \to \min .$$

Решение. Функция $f(x)$ является выпуклой функцией как сумма двух выпуклых функций. Действительно, функция $g(x) = x_1^2 + x_1 x_2 + x_2^2$ выпукла по теореме 2 п 4.1.2, так как по критерию Сильвестра матрица вторых производных

$$g''(x) = \left(\frac{\partial^2 g(x)}{\partial x_i \partial x_j} \right)^2_{i,j=1} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

положительно определена для любого x. Функция $h(x_1, x_2) = |x_1 + x_2 - 2|$, являющаяся модулем афинной функции, также является выпуклой функцией (см. п. 4.1.2).

Необходимое и достаточное условие экстремума в выпуклой задаче без ограничений:

$$0 \in \partial f(\hat{x}) = \partial g(\hat{x}) + 3\partial h(\hat{x}) \tag{1}$$

(использовали теорему Моро-Рокафеллара о субдифференциале суммы функций). Найдем субдифференциалы функций g и h. Для дифференцируемой функции ее субдифференциал совпадает с производной. Поэтому $\partial g(\hat{x}) = (2\hat{x}_1 + \hat{x}_2, \hat{x}_1 + 2\hat{x}_2)$.

По формуле субдифференциала модуля

$$\partial |h| = \begin{cases} -h'(x), & h(x) < 0, \\ \{\alpha h'(x),\ \alpha \in [-1;\,1]\}, & h(x) = 0, \\ h'(x), & h(x) > 0, \end{cases}$$

для $h(x) = x_1 + x_2 - 2$ имеем:

$$\partial h(x) = \begin{cases} (-1, -1), & x_1 + x_2 - 2 < 0, \\ (\alpha, \alpha),\ |\alpha| \le 1, & x_1 + x_2 - 2 = 0, \\ (1, 1), & x_1 + x_2 - 2 > 0. \end{cases}$$

Значит,

$$\partial f(x) = \partial g(x) + 3\partial h(x) =$$

$$= (2x_1 + x_2, x_1 + 2x_2) + 3 \begin{cases} (-1,-1), & x_1 + x_2 - 2 < 0, \\ (\alpha, \alpha),\ |\alpha| \le 1, & x_1 + x_2 - 2 = 0, \\ (1,1), & x_1 + x_2 - 2 > 0, \end{cases} =$$

$$= \begin{cases} (2x_1 + x_2 - 3, x_1 + 2x_2 - 3), & x_1 + x_2 - 2 < 0, \\ (2x_1 + x_2 + 3\alpha, x_1 + 2x_2 + 3\alpha),\ |\alpha| \le 1, & x_1 + x_2 - 2 = 0, \\ (2x_1 + x_2 + 3, x_1 + 2x_2 + 3), & x_1 + x_2 - 2 > 0. \end{cases}$$

Следовательно, соотношение (1) запишется в виде

$$\begin{cases} 2x_1 + x_2 - 3 = 0, \\ x_1 + 2x_2 - 3 = 0, \end{cases} \quad \text{при} \quad x_1 + x_2 - 2 < 0, \tag{i}$$

$$\begin{cases} 2x_1 + x_2 + 3\alpha = 0, \\ x_1 + 2x_2 + 3\alpha = 0, \end{cases} \quad \text{при} \quad x_1 + x_2 - 2 = 0, \quad (|\alpha| \le 1) \tag{ii}$$

$$\begin{cases} 2x_1 + x_2 + 3 = 0, \\ x_1 + 2x_2 + 3 = 0, \end{cases} \quad \text{при} \quad x_1 + x_2 - 2 > 0. \tag{iii}$$

В случае (i) нет критических точек, так как из уравнений системы следует, что $x_1 + x_2 = 2$ — противоречие с условием $x_1 + x_2 - 2 < 0$.

В случае (iii) также нет критических точек, так как из уравнений системы следует, что $x_1 + x_2 = -2$ — противоречие с условием $x_1 + x_2 - 2 > 0$.

В случае (ii) система из трех уравнений с тремя неизвестными имеет единственное решение $\hat{x}_1 = \hat{x}_2 = 1$, $\alpha = -1$.

Ответ. $\hat{x} = (\hat{x}_1, \hat{x}_2) = (1,1) \in \text{absmin}$, $S_{\text{absmin}} = 3$.

4.4 Выпуклые, конические, аффинные множества

4.4.1 Свойства операций над множествами

1. *Пересечение любого числа выпуклых (конических, аффинных) множеств является выпуклым (коническим, аффинным) множеством.*

$\lhd$ Пусть A_α — выпуклые множества. Докажем, что $A = \cap_\alpha A_\alpha$ — выпуклое множество. Возьмем точки $a, b \in A$. Поскольку $A = \cap_\alpha A_\alpha$, то точки $a, b \in A_\alpha$ для любого α. В силу выпуклости множества A_α точка $ta + (1-t)b \in A_\alpha$ для любого $0 \le t \le 1$ и для любого α. Следовательно, точка $ta + (1-t)b \in A$.

Аналогично свойство доказывается для конических и аффинных множеств. $\rhd$

Упражнение 1. Верно ли утверждение: если для любых двух точек a и b из множества A их полусумма $\frac{a+b}{2} \in A$, то множество A выпукло?

2. $f\colon X \to Y$ *— линейное отображение*[1] $\Rightarrow A$ *— выпуклое (коническое, аффинное) множество* $\Leftrightarrow f(A)$ *— выпуклое (коническое, аффинное) множество.*

$\lhd$ Пусть A — выпуклое множество. Возьмем точки $x_1, x_2 \in f(A)$. Тогда $x_1 = f(a_1)$, $x_2 = f(a_2)$, где $a_1, a_2 \in A$. Следовательно,

$$tx_1 + (1-t)x_2 = tf(a_1) + (1-t)f(a_2) \overset{f\text{ — линейное}}{=} f\big(ta_1 + (1-t)a_2\big).$$

Значит, $tx_1 + (1-t)x_2 \in f(A)$, т.е. множество $f(A)$ — выпукло.

Обратно. Пусть $B := f(A)$ — выпуклое множество. Возьмем точки $a_1, a_2 \in f^{-1}(B) = A$, тогда $x_1 = f(a_1) \in B$, $x_2 = f(a_2) \in B$. В силу выпуклости множества B точки $tx_1 + (1-t)x_2 \in B$ для любого $0 \le t \le 1$. Поскольку f — линейное отображение, то

$$tx_1 + (1-t)x_2 = tf(a_1) + (1-t)f(a_2) = f(ta_1 + (1-t)a_2).$$

[1] f — линейное, если $f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)\ \forall\ a, b \in X, \forall\ \lambda, \mu \in \mathbb{R}$.

Значит, $ta_1 + (1-t)a_2 \in f^{-1}(B) = A$, т. е. множество A — выпукло.

Аналогично свойство доказывается для конических и аффинных множеств. $\triangleright$

3. *Сдвиг $A \to b + A$ выпуклого (аффинного) множества является выпуклым (аффинным) множеством.*

4. *Гомотетия $A \to \lambda A$, $\lambda \in \mathbb{R}$, выпуклого множества является выпуклым множеством.*

5. *Прямое (декартово) произведение выпуклых множеств выпукло ($A \subset X$ и $B \subset Y$ выпуклы $\Rightarrow A \times B$ выпукло в $X \times Y$).*

6. *Алгебраическая сумма $A+B$ и разность $A-B$ выпуклых множеств выпуклы. Линейная комбинация конечного числа выпуклых множеств $\lambda_1 A_1 + \ldots + \lambda_k A_k$ выпукла.*

Свойства 3.-6. легко следуют из определений.

Упражнение 2. Доказать, что для любых $\alpha, \beta > 0$ и выпуклого множества A выполняется равенство: $\alpha A + \beta A = (\alpha + \beta)A$; в частности, $A + A = 2A$. Показать, что для невыпуклых множеств это, вообще говоря, неверно.

4.4.2 Эквивалентные определения выпуклых множеств

Лемма (1-v)**.** *Пусть A — выпуклое множество. Тогда для любых точек $a_1, \dots, a_n \in A$ выпуклая оболочка* $\operatorname{conv}\{a_1, \dots, a_n\} \subset A$.

Доказательство. Индукция по n. При $n = 1$ утверждение очевидно. При $n = 2$ утверждение следует из определения выпуклого множества. Предположим, лемма верна для любой выпуклой оболочки не более, чем $n-1$ точки. Докажем для n точек. Пусть точка a принадлежит выпуклой оболочке n точек: $a \in \operatorname{conv}\{a_1, \dots, a_n\}$. Тогда $a = \sum\limits_{i=1}^{n} t_i a_i$, $\sum\limits_{i=1}^{n} t_i = 1$, $t_i > 0$, $i = 1, \dots, n$. Поскольку $n > 1$, а $t_i > 0$, то $t_1 \neq 1$. Положим $\tilde{t}_k = \frac{t_k}{1-t_1} \geq 0$ при $k = 2, \dots, n$. Так как $\sum\limits_{k=2}^{n} \tilde{t}_k = \frac{\sum_{k=2}^{n} t_k}{1-t_1} = \frac{1-t_1}{1-t_1} = 1$, то точка $\tilde{a} = \sum\limits_{k=2}^{n} \tilde{t}_k a_k$ – выпуклая комбинация, и по предположению индукции $\tilde{a} \in A$. Тогда в силу выпуклости A имеем $a = t_1 a_1 + \sum\limits_{k=2}^{n} t_k a_k = t_1 a_1 + \sum\limits_{k=2}^{n} (1-t_1)\tilde{t}_k a_k = t_1 a_1 + (1-t_1)\tilde{a} \in A$. $\square$

Обозначим $\operatorname{conv} A := \bigcup\limits_{a_1,\dots,a_n \in A,\, n \in \mathbb{N}} \operatorname{conv}\{a_1, \dots, a_n\}$.
Из определения ($n = 1$) следует, что $\operatorname{conv} A \supset A$.

Лемма (2-v)**.** *Множество* $\operatorname{conv} A$ *— выпукло.*

Доказательство. Возьмем точки $a, b \in \operatorname{conv} A$. Тогда по определению выпуклой оболочки множества $a = \sum\limits_{i=1}^{m} \lambda_i a_i$, $b = \sum\limits_{j=1}^{n} \mu_j b_j$, $\sum\limits_{i=1}^{m} \lambda_i = \sum\limits_{j=1}^{n} \mu_j = 1$, $\lambda_i, \mu_j \geq 0$. Для любого $t \in [0; 1]$ имеем

$$(1-t)a + tb = (1-t)\sum_{i=1}^{m} \lambda_i a_i + t\sum_{j=1}^{n} \mu_j b_j = \sum_{i=1}^{m} (1-t)\lambda_i a_i + \sum_{j=1}^{n} t\mu_j b_j.$$

Заметим, что все множители $(1-t)\lambda_i$ и $t\mu_j$ неотрицательны и $\sum\limits_{i=1}^{m} (1-t)\lambda_i + \sum\limits_{j=1}^{n} t\mu_j = (1-t) + t = 1$, значит, получили выпуклую комбинацию точек $a_1, \dots, a_m, b_1, \dots, b_n$, которая принадлежит $\operatorname{conv} A$. $\square$

Теорема 3. *Множество A — выпукло тогда и только тогда, когда* $\operatorname{conv} A = A$.

Доказательство. Пусть A — выпукло. Тогда по лемме 1-v $\operatorname{conv} A \subset A$. Вложение $\operatorname{conv} A \supset A$ следует из определения $\operatorname{conv} A$ $(n = 1)$.

Обратно. Пусть $\operatorname{conv} A = A$. По лемме 2-v $\operatorname{conv} A$ — выпукло, значит, и A — выпукло. □

Обозначим $\widetilde{A} = \cap_\alpha A_\alpha$ — пересечение всех выпуклых множеств A_α, содержащих A $(A_\alpha \supset A)$.
Тогда $\widetilde{A}$ — выпукло как пересечение выпуклых множеств.

Теорема 4. *Имеет место равенство* $\operatorname{conv} A = \widetilde{A}$.

Доказательство. Поскольку по лемме 2-v $\operatorname{conv} A$ — выпукло и $\operatorname{conv} A \supset A$ следует из определения $\operatorname{conv} A$ $(n = 1)$, то $\operatorname{conv} A \supset \cap_\alpha A_\alpha = \widetilde{A}$. Докажем, что $\operatorname{conv} A \subset \widetilde{A}$. Пусть $a \in \operatorname{conv} A$. Тогда $a \in \operatorname{conv}\{a_1, \ldots, a_n\}$, где $a_i \in A \subset A_\alpha$ для любого α. Поскольку A_α — выпукло, то по лемме 1-v $a \in A_\alpha$ для любого α. Так как $\cap_\alpha A_\alpha = \widetilde{A}$, то $a \in \widetilde{A}$. □

Теорема 5. *Пусть* X *— линейное нормированное пространство,* $A \subset X$ *— выпуклое множество. Тогда внутренность* $\operatorname{int} A$ *и замыкание* $\overline{A}$ *множества* A *выпуклы.*

Доказательство. Пусть точки $a, b \in \operatorname{int} A$. Тогда по определению внутренней точки существует $\varepsilon > 0$, для которого шары $a + B_\varepsilon, b + B_\varepsilon \subset A$. В силу выпуклости множества A для любых $t_1, t_2 \geq 0$, $t_1 + t_2 = 1$, множество $t_1(a + B_\varepsilon) + t_2(b + B_\varepsilon) \subset A$. Согласно утверждению из упражнения 2 множество

$$t_1(a + B_\varepsilon) + t_2(b + B_\varepsilon) = t_1 a + t_2 b + t_1 B_\varepsilon + t_2 B_\varepsilon = t_1 a + t_2 b + B_\varepsilon.$$

Значит, множество $t_1 a + t_2 b \in \operatorname{int} A$, т. е. $\operatorname{int} A$ — выпуклое множество.

Пусть $a, b \in \overline{A}$. Покажем, что для любых $t_1, t_2 \geq 0$, $t_1 + t_2 = 1$, точка $t_1 a + t_2 b \in \overline{A}$. Поскольку $a, b \in \overline{A}$, то существуют последовательности $a_k, b_k \in A$ такие, что $a_k \to a$, $b_k \to b$. В силу выпуклости множества A точки $c_k := t_1 a_k + t_2 b_k \in A$. Поэтому $c = \lim c_k \in \overline{A}$. Следовательно, $\overline{A}$ — выпуклое множество. □

4.4.3 Эквивалентные определения выпуклых конусов

Лемма (1-k)**.** *Пусть A — выпуклый конус. Тогда коническая выпуклая оболочка* $\operatorname{conev}\{a_1,\ldots,a_n\} \subset A$*, для любых* $a_1,\ldots,a_n \in A$*, для любого* $n \in \mathbb{N}$*.*

Доказательство. Пусть точка $a \in \operatorname{conev}\{a_1,\ldots,a_n\}$. Значит, точка a представляется в виде $a = \sum\limits_{i=1}^{n} t_i a_i$, $t_i \geq 0$, $i = 1,\ldots,n$. Если $t_1 = \ldots = t_n = 0$, то $a = 0 \in A$ (ноль всегда принадлежит конусу). Если же $\lambda := \sum\limits_{i=1}^{n} t_i > 0$, то рассмотрим точку $b := \sum\limits_{i=1}^{n} \dfrac{t_i}{\lambda} a_i = \dfrac{a}{\lambda}$. Так как $\sum\limits_{i=1}^{n} \dfrac{t_i}{\lambda} = 1$, то b – выпуклая комбинация точек a_i. Следовательно, $b \in A$ по лемме 1-v в силу выпуклости A. Поскольку A — конус, то точка $a = \lambda b \in A$. □

Обозначим $\operatorname{conev} A := \bigcup\limits_{a_1,\ldots,a_n \in A,\, n \in \mathbb{N}} \operatorname{conev}\{a_1,\ldots,a_n\}$.

Из определения ($n = 1$) следует, что $\operatorname{conev} A \supset A$.

Лемма (2-k)**.** *Множество* $\operatorname{conev} A$ *— выпуклый конус.*

Доказательство. То, что множество $\operatorname{conev} A$ является конусом следует из определения, поскольку объединение конусов есть конус. Докажем выпуклость. Возьмем точки $a, b \in \operatorname{conev} A$. Значит, точки a, b представляются в виде $a = \sum\limits_{i=1}^{m} \lambda_i a_i$, $b = \sum\limits_{j=1}^{n} \mu_j b_j$, $a_i, b_j \in A$, $\lambda_i, \mu_j \geq 0$. Тогда для любого $t \in [0;1]$ точка $(1-t)a + tb = (1-t)\sum\limits_{i=1}^{m} \lambda_i a_i + t\sum\limits_{j=1}^{n} \mu_j b_j = \sum\limits_{i=1}^{m} (1-t)\lambda_i a_i + \sum\limits_{j=1}^{n} t\mu_j b_j \in \operatorname{conev} A$, поскольку все множители $(1-t)\lambda_i$ и $t\mu_i$ неотрицательны (получили коническую комбинацию точек $a_1,\ldots,a_m,b_1,\ldots,b_n$). Следовательно, $\operatorname{conev} A$ — выпуклый конус. □

Теорема 6. *Множество A — выпуклый конус тогда и только тогда, когда* $\operatorname{conev} A = A$.

Доказательство. Пусть A — выпуклый конус, тогда по лемме 1-k $\operatorname{conev} A \subset A$. Вложение $\operatorname{conev} A \supset A$ следует из определения $\operatorname{conev} A$ $(n = 1)$.

Обратно. Пусть $\operatorname{conev} A = A$. По лемме 2-k $\operatorname{conev} A$ — выпуклый конус, значит, и A — выпуклый конус. $\square$

Обозначим $\widetilde{A} = \bigcap\limits_{\alpha} A_\alpha$ — пересечение всех выпуклых конусов A_α, содержащих A $(A_\alpha \supset A)$. Тогда $\widetilde{A}$ — выпуклый конус как пересечение выпуклых конусов.

Теорема 7. *Имеет место равенство* $\operatorname{conev} A = \widetilde{A}$.

Доказательство. Поскольку по лемме 2-k $\operatorname{conev} A$ — выпуклый конус и $\operatorname{conev} A \supset A$ по определению $\operatorname{conev} A$ $(n = 1)$, то $\operatorname{conev} A \supset \widetilde{A}$.

Докажем, что $\operatorname{conev} A \subset \widetilde{A}$. Пусть $a \in \operatorname{conev} A$. Тогда $a \in \operatorname{conev}\{a_1, \ldots, a_n\}$, $a_i \in A \subset A_\alpha$ для любого α. Поскольку A_α — выпуклый конус, то по лемме 1-k $a \in A_\alpha$ для любого α. Так как $\cap_\alpha A_\alpha = \widetilde{A}$, то $a \in \widetilde{A}$. $\square$

4.4.4 Эквивалентные определения аффинных множеств

Лемма (1-а)**.** *Пусть A — аффинное множество. Тогда аффинная оболочка конечного число точек* $\operatorname{aff}\{a_1,\ldots,a_n\}\subset A$ *для любых* $a_1,\ldots,a_n\in A$, *для любого* $n\in\mathbb{N}$.

Доказательство. Индукция по n. При $n=1$ утверждение очевидно. Предположим, лемма верна для любой аффинной комбинации не более, чем $n-1$ точки. Докажем для n точек. Пусть точка a принадлежит аффинной оболочке n точек: $a\in\operatorname{aff}\{a_1,\ldots,a_n\}$, где $a_1,\ldots,a_n\in A$. Тогда $a=\sum\limits_{i=1}^{n}t_ia_i$, $\sum\limits_{i=1}^{n}t_i=1$. Если $t_1=1$, то $a=a_1\in A$. Если же $t_1\neq 1$, то полагаем $\widetilde{t}_k=\frac{t_k}{1-t_1}$ при $k=2,\ldots,n$. Так как $\sum\limits_{k=2}^{n}\widetilde{t}_k=\frac{\sum_{k=2}^{n}t_k}{1-t_1}$ $=\frac{1-t_1}{1-t_1}=1$, то $\widetilde{a}=\sum\limits_{k=2}^{n}\widetilde{t}_ka_k$ – аффинная комбинация, и по предположению индукции $\widetilde{a}\in A$. Тогда, в силу аффинности A, имеем $a=t_1a_1+\sum\limits_{k=2}^{n}t_ka_k=t_1a_1+\sum\limits_{k=2}^{n}(1-t_1)\widetilde{t}_ka_k=t_1a_1+(1-t_1)\widetilde{a}\in A$. □

Обозначим $\operatorname{aff}A:=\bigcup\limits_{a_1,\ldots,a_n\in A,\,n\in\mathbb{N}}\operatorname{aff}\{a_1,\ldots,a_n\}$.
Из определения ($n=1$) следует, что $\operatorname{aff}A\supset A$.

Лемма (2-а)**.** *Множество* $\operatorname{aff}A$ *— аффинное.*

Доказательство. Возьмем точки $a,b\in\operatorname{aff}A$. Значит, точки a,b представляются в виде $a=\sum\limits_{i=1}^{m}\lambda_ia_i$, $b=\sum\limits_{j=1}^{n}\mu_jb_j$, $\sum\limits_{i=1}^{m}\lambda_i=\sum\limits_{j=1}^{n}\mu_j=1$. Тогда для любого $t\in\mathbb{R}$ точка $(1-t)a+tb=(1-t)\sum\limits_{i=1}^{m}\lambda_ia_i+t\sum\limits_{j=1}^{n}\mu_jb_j=\sum\limits_{i=1}^{m}(1-t)\lambda_ia_i+\sum\limits_{j=1}^{n}t\mu_jb_j\in\operatorname{aff}A$, так как $\sum\limits_{i=1}^{m}(1-t)\lambda_i+\sum\limits_{j=1}^{n}t\mu_j=(1-t)+t=1$, и, следовательно, получили аффинную комбинацию точек $a_1,\ldots,a_m,b_1,\ldots,b_n$. □

Теорема 8. *Множество A — аффинное тогда и только тогда, когда* $\operatorname{aff} A = A$.

Доказательство. Пусть A — аффинное, тогда по лемме 1-а $\operatorname{aff} A \subset A$. Вложение $\operatorname{aff} A \supset A$ следует из определения $\operatorname{aff} A$ $(n = 1)$.

Обратно. Пусть $\operatorname{aff} A = A$. По лемме 2-а $\operatorname{aff} A$ — аффинное, значит, и A — аффинное. □

Обозначим $\widetilde{A} = \cap_\alpha A_\alpha$ — пересечение всех аффинных множеств A_α, содержащих A $(A_\alpha \supset A)$.
Тогда $\widetilde{A}$ — аффинное как пересечение аффинных множеств.

Теорема 9. *Имеет место равенство* $\operatorname{aff} A = \widetilde{A}$.

Доказательство. Поскольку по лемме 2-а $\operatorname{aff} A$ — аффинное и $\operatorname{aff} A \supset A$ по определению $\operatorname{aff} A$ $(n = 1)$, то $\operatorname{aff} A \supset \widetilde{A}$.

Докажем, что $\operatorname{aff} A \subset \widetilde{A}$. Пусть $a \in \operatorname{aff} A$. Тогда $a \in \operatorname{aff}\{a_1, \ldots, a_n\}$, $a_i \in A \subset A_\alpha$ для любого α. Поскольку A_α — аффинное, то по лемме 1-а $a \in A_\alpha$ для любого α. Так как $\cap_\alpha A_\alpha = \widetilde{A}$, то $a \in \widetilde{A}$. □

4.5 Размерность множества

Пусть X — линейное нормированное пространство.

- *Подпространство* — аффинное множество, проходящее через 0.
- Аффинные множества A и B *параллельны*, если $\exists\, c \in X$: $B = A + c$.
- *Размерность аффинного множества* есть размерность подпространства, параллельного этому аффинному множеству.
- *Размерность множества* M есть размерность $\operatorname{aff} M$.

Размерность точки равна 0; размерность прямой равна 1; размерность гиперплоскости в $\mathbb{R}^n$ равна $n - 1$.

Определение 3. *Точки $a_0, a_1, \ldots, a_n$ — аффинно независимы, если никакая из них не является аффинной комбинацией остальных*[2]*, а* $\dim \operatorname{aff}\{a_0, a_1, \ldots, a_n\} = n$,

Легко доказать, что *точки $a_0, a_1, \ldots, a_n$ — аффинно независимы тогда и только тогда, когда вектора $a_1 - a_0, \ldots, a_n - a_0$ — линейно независимы.*

Пусть точки $a_0, a_1, \ldots, a_n$ — аффинно независимы. Тогда множество $\Sigma^n = \operatorname{conv}\{a_0, a_1, \ldots, a_n\}$ называется n-мерным *симплексом* с вершинами $a_0, a_1, \ldots, a_n$.

В частности, $\Sigma^0 = a_0$ — точка в X; $\Sigma^1 = \operatorname{conv}\{a_0, a_1\} = [a_0, a_1]$ — отрезок; $\Sigma^2 = \operatorname{conv}\{a_0, a_1, a_2\}$ — треугольник, и т. д.

---

[2] т. е. не существует вектора $\lambda = (\lambda_0, \ldots, \lambda_n) \neq 0$, $\sum\limits_{i=0}^{n} \lambda_i = 0$, для которого $\sum\limits_{i=0}^{n} \lambda_i a_i = 0$.

Лемма (о представлении точек симплекса через вершины). *Любая точка симплекса* $\Sigma^n = \mathrm{conv}\{a_0, a_1, \ldots, a_n\}$ *единственным образом представляется в виде выпуклой комбинации его вершин:* $x = \alpha_0 a_0 + \ldots + \alpha_n a_n$*, где* $\alpha_0, \ldots, \alpha_n \geq 0$*,* $\alpha_0 + \ldots + \alpha_n = 1$*.*

Доказательство. Существование такого представления вытекает из определения выпуклой оболочки. Докажем его единственность. Предположим, имеется другое представление: $x = \beta_0 a_0 + \ldots + \beta_n a_n$, где $\beta_0, \ldots, \beta_n \geq 0$, $\beta_0 + \ldots + \beta_n = 1$. Вычитая из первого представления второе, получаем $(\alpha_0 - \beta_0)a_0 + \ldots + (\alpha_n - \beta_n)a_n = 0$, $(\alpha_0 - \beta_0) + \ldots + (\alpha_n - \beta_n) = 0$ (так как $\alpha_0 + \ldots + \alpha_n = 1$, $\beta_0 + \ldots + \beta_n = 1$). Тогда в силу аффинной независимости вершин симплекса $(\alpha_0 - \beta_0) = \ldots = (\alpha_n - \beta_n) = 0$. Значит, оба представления совпадают. $\square$

Коэффициенты $\alpha_0, \ldots, \alpha_n$ называются *барицентрическими координатами* точки $a = \alpha_0 a_0 + \ldots + \alpha_n a_n$ в симплексе Σ^n.

Если считать, что точки a_i являются материальными точками массы α_i, то можно доказать, что центр масс полученной системы материальных точек будет находиться в точке a.

Если все массы одинаковы $\alpha_0 = \ldots = \alpha_n = \dfrac{1}{n+1}$ (так как $\alpha_0 + \ldots + \alpha_n = 1$), то точка $a = \sum\limits_{i=0}^{n} \dfrac{1}{n+1} a_i$ называется *центром симплекса (центром тяжести)*. В частности, для $\Sigma^2 = \mathrm{conv}\{a_0, a_1, a_2\}$ центр тяжести — точка пересечения медиан треугольника с вершинами в точках a_0, a_1, a_2.

Лемма 1. *Пусть* $\dim X = n$, $\operatorname{conv}\{a_1, \dots, a_{n+1}\}$ *— симплекс,* $a = \sum\limits_{i=1}^{n+1} t_i a_i$, $\sum\limits_{i=1}^{n+1} t_i = 1$, $t_i > 0$. *Тогда* $a \in \operatorname{int}\operatorname{conv}\{a_1, \dots, a_{n+1}\}$.

Доказательство. Аффинным преобразованием переведем пространство X в аффинное пространство $L = \{y \in \mathbb{R}^{n+1} \mid \sum\limits_{k=1}^{n+1} y_k = 1\}$ так, чтобы каждая точка a_i перешла в конец базисного вектора $e_i \in \mathbb{R}^{n+1}$. Также можно считать, что норма – евклидова, поскольку все нормы в конечномерном пространстве эквивалентны. Если $a = \sum\limits_{i=1}^{n+1} t_i e_i = (t_1, \dots, t_{n+1})$, то расстояние от a до ближайшей координатной плоскости равно $r = \min t_i$. Поэтому n-мерный шар с центром a радиуса r целиком лежит в симплексе. Это означает, что a – внутренняя точка симплекса. $\square$

Утверждение. *Пусть множество* $A \subset X$, $\dim\operatorname{aff} A = n$. *Тогда* $\operatorname{int}\operatorname{conv} A \neq \emptyset$ *в* $\operatorname{aff} A$.

Доказательство. Поскольку $\dim\operatorname{aff} A = n$, то существуют точки $a_1, \dots, a_{n+1} \in A$, являющиеся аффинно-независимыми. Значит, множество $\operatorname{conv}\{a_1, \dots, a_{n+1}\}$ — n-мерный симплекс и, следовательно, по лемме $\operatorname{int}\operatorname{conv}\{a_1, \dots, a_{n+1}\} \neq \emptyset$. Поскольку $\operatorname{conv}\{a_1, \dots, a_{n+1}\} \subset \operatorname{conv} A$, то и $\operatorname{int}\operatorname{conv} A \neq \emptyset$. $\square$

Следствие 1. *Пусть* A *— выпуклое множество,* $\dim A < \infty$. *Тогда* $\operatorname{int} A \neq \emptyset$ *в* $\operatorname{aff} A$.

Следствие 2. *Пусть* $A \subset \mathbb{R}^n$ *— выпуклое множество,* $\operatorname{int} A = \emptyset$. *Тогда* $\dim A < n$.

Пример. Выпуклое множество с пустой внутренностью, аффинная оболочка которого совпадает со всем пространством.

$\lhd$ Рассмотрим $l_2 = \{x = (x_1, x_2, \dots) \mid \sum\limits_k x_k^2 < \infty\}$ — гильбертово пространство, $A = \{x \in l_2 \mid |x_k| \leq \frac{1}{k},\ k \in \mathbb{N}\}$ — “гильбертов кирпич”. Тогда A — выпуклое множество, $\operatorname{int} A = \emptyset$, $\operatorname{aff} A = l_2$. $\rhd$

Теорема (Каратеодори). *Пусть* $\dim \operatorname{conv} A = d$. *Тогда любой элемент из* $\operatorname{conv} A$ *представляется в виде выпуклой комбинации не более* $d+1$ *элементов из* A.

Доказательство. Не ограничивая общности, считаем, что $\operatorname{aff} A = X$, $\dim X = d$. Возьмем элемент $a \in \operatorname{conv} A$. Тогда $a = \sum\limits_{i=1}^{n} t_i a_i$, $\sum\limits_{i=1}^{n} t_i = 1$, $t_i > 0$, $a_i \in A$ и предположим, что $n \geq d+2$. Рассмотрим векторы $a_2 - a_1, \ldots, a_n - a_1$. Их не менее $d+1$, и они принадлежат d-мерному линейному пространству и, значит, линейно зависимы, т. е. существуют множители $(\lambda_2, \ldots, \lambda_n) \neq 0$, для которых

$$\sum_{i=2}^{n} \lambda_i (a_i - a_1) = 0 \iff \Big(-\sum_{i=2}^{n} \lambda_i\Big) a_1 + \sum_{i=2}^{n} \lambda_i a_i = 0 \iff \sum_{i=1}^{n} \lambda_i a_i = 0,$$

где $\lambda_1 := -\sum\limits_{i=2}^{n} \lambda_i$, вектор $(\lambda_1, \ldots, \lambda_n) \neq 0$ и $\sum\limits_{i=1}^{n} \lambda_i = 0$. Обозначим $s := \min\big\{-\frac{t_i}{\lambda_i} \mid \lambda_i < 0\big\} = -\frac{t_j}{\lambda_j} > 0$. Тогда $a = \sum\limits_{i=1}^{n} t_i a_i$ (поскольку $\sum\limits_{i=1}^{n} \lambda_i = 0$) $= \sum\limits_{i=1}^{n} (\underbrace{t_i + s\lambda_i}_{\mu_i}) a_i = \sum\limits_{i=1}^{n} \mu_i a_i$, $\sum\limits_{i=1}^{n} \mu_i = 1$, $\mu_i \geq 0$ (в силу выбора s), $\mu_j = 0$. Следовательно, представили точку a как выпуклую комбинацию из не более, чем $n-1$ точки. Уменьшая далее число точек, приходим к выпуклой комбинации из не более, чем $d+1$ точки. $\square$

Теорема 10. *Выпуклая оболочка компакта в конечномерном пространстве является компактом.*

Доказательство. Пусть $\dim X = d$, $A \subset X$ — компакт. Докажем, что $\operatorname{conv} A$ — компакт. Для этого достаточно доказать, что из любой последовательности $\{a_k\} \in \operatorname{conv} A$ можно выбрать подпоследовательность, сходящуюся к некоторому элементу из $\operatorname{conv} A$. По теореме Каратеодори $a_k = \sum\limits_{i=1}^{d+1} t_{ki} a_{ki}$, где $a_{ki} \in A$, $\sum\limits_{i=1}^{d+1} t_{ki} = 1$, $t_{ki} \geq 0$. Поскольку A – компакт, то в последовательности $\{a_{k1}\}$ существует сходящаяся к некоторой точке $b_1 \in A$ подпоследовательность, а в соответствующей подпоследовательности в $\{t_{k1}\}$ существует подпоследовательность, сходящаяся к некоторому числу $t_1 \in [0; 1]$. Не ограничивая общности, считаем, что $a_{k1} \to b_1$ и $t_{k1} \to t_1$ при $k \to \infty$.
Аналогично, переходя к подпоследовательностям, считаем $a_{ki} \to b_i \in A$ и $t_{ki} \to t_i \in [0; 1]$ при $k \to \infty$ для любого $i = 2, \ldots, d+1$. При этом $\sum\limits_{i=1}^{d+1} t_i = 1$, $t_i \geq 0$, и $\lim\limits_{k\to\infty} a_k = \sum\limits_{i=1}^{d+1} t_i b_i = a$. Это означает, что предельная точка принадлежит $\operatorname{conv} A$. $\square$

Пример. Замкнутое множество, выпуклая оболочка которого, является незамкнутым множеством.

$\lhd$ $A = \{(x, y) \in \mathbb{R}^2 \mid xy = 1,\ x > 0\} \cup (0, 0)$ — объединение ветви гиперболы и начала координат — замкнутое множество. Тогда $\operatorname{conv} A$ — открытая первая четверть, включая точку ноль — незамкнутое множество. $\rhd$

Упражнение. Привести пример компакта, выпуклая оболочка которого не является компактом.

4.6 Поляры множеств

4.6.1 Определение поляр множеств

Определение 4. *Линейные пространства* X *и* Y *называются пространствами в двойственности, если определена билинейная форма* $\langle \cdot, \cdot \rangle \colon Y \times X \to \mathbb{R}$ *такая, что если* $\langle y, x \rangle = 0 \ \forall \ y \in Y$, *то* $x = 0$, *а если* $\langle y, x \rangle = 0 \ \forall \ x \in X$, *то* $y = 0$.

Пример. Если X — линейное нормированное пространство, а $Y = X^*$, то форма $\langle y, x \rangle$ является билинейной, а пространства X и $Y = X^*$ являются пространствами в двойственности.

Далее для образности представления рассматриваем именно этот случай.

Определение 5. Пусть X и Y — пространства в двойственности. *Полярой множества* $A \subset X$ называется множество $A^\circ = \{y \in Y \mid \sup\limits_{x \in A} \langle y, x \rangle \leq 1\}$.

4.6.2 Свойства поляры

1. *Поляра множества в линейном нормированном пространстве является выпуклым замкнутым подмножеством, содержащим ноль.*

$\lhd$ Множество A° является пересечением выпуклых замкнутых множеств (полупространств) $L_x = \{y \in Y \mid \langle y, x\rangle \le 1\}$ по всем $x \in A$, а значит – выпукло и замкнуто. Очевидно, поляра содержит ноль. $\rhd$

2. $B \supset A \Rightarrow B^\circ \subset A^\circ$.

$\lhd$ Включение $B^\circ \subset A^\circ$ следует из неравенства

$$\sup_{x\in A} \langle y, x\rangle \overset{A\subset B}{\le} \sup_{x\in B} \langle y, x\rangle. \qquad \rhd$$

3. $(\lambda A) = \lambda^{-1} A^\circ \ \forall\ \lambda > 0$.

$\lhd$ Поскольку $\sup\limits_{x\in\lambda A} \langle y, x\rangle = \sup\limits_{x\in A} \langle y, \lambda x\rangle = \lambda \sup\limits_{x\in A} \langle y, x\rangle$, то $\sup\limits_{x\in\lambda A} \langle y, x\rangle \le 1 \Leftrightarrow \sup\limits_{x\in A} \langle y, x\rangle \le \frac{1}{\lambda}$. Из последней эквивалентности вытекает свойство 3. $\rhd$

4. $(\,\overline{A}\,)^\circ = A^\circ$.

$\lhd$ Поскольку $\overline{A} \supset A$, то по свойству 2 $(\,\overline{A}\,)^\circ \subset A^\circ$.

Докажем обратное включение, что $(\,\overline{A}\,)^\circ \supset A^\circ$. Пусть $y \in A^\circ$. Тогда для точки $\overline{x} \in \overline{A}$ существует последовательность точек $x_k \in A$, $x_k \to \overline{x}$, для которой $\langle y, x_k\rangle \le 1$. Переходя в неравенстве к пределу по $x_k \to \overline{x} \Rightarrow \langle y, \overline{x}\rangle \le 1$, получим $\sup\limits_{\overline{x}\in\overline{A}} \langle y, \overline{x}\rangle \le 1$. Это означает, что $y \in (\,\overline{A}\,)^\circ$. $\rhd$

5. $(A \cup \{0\})^\circ = A^\circ$.

$\lhd$ Поскольку множество $A \cup \{0\} \supset A$, то по свойству 2 множество $(A \cup \{0\})^\circ \subset A^\circ$.

Докажем обратное включение, что $(A \cup \{0\})^\circ \supset A^\circ$. Пусть $y \in A^\circ$. Тогда $\sup\limits_{x\in A} \langle y, x\rangle \le 1$. Отсюда $\sup\limits_{x\in A\cup\{0\}} \langle y, x\rangle \le 1$. Это означает, что $y \in (A \cup \{0\})^\circ$. $\rhd$

6. $(\operatorname{conv} A)^\circ = A^\circ$.

$\lhd$ Поскольку множество $\operatorname{conv} A \supset A$, то по свойству 2 множество $(\operatorname{conv} A)^\circ \subset A^\circ$.

Докажем обратное включение, что $(\operatorname{conv} A)^\circ \supset A^\circ$. Пусть $y \in A^\circ$. Тогда для $x \in \operatorname{conv} A$ имеем разложение:
$x = \sum\limits_{j=1}^{n} t_j x_j$, $x_j \in A$, $\sum\limits_{j=1}^{n} t_j = 1$, $t_j \geq 0$, $j = 1, \ldots, n$. Поэтому

$$\langle y, x\rangle = \left\langle y, \sum_{j=1}^{n} t_j x_j \right\rangle = \sum_{j=1}^{n} t_j \langle y, x_j\rangle \overset{\langle y, x_j\rangle \leq 1}{\leq} \sum_{j=1}^{n} t_j = 1.$$

Следовательно, $y \in (\operatorname{conv} A)^\circ$. $\rhd$

Из свойств 4-6 следует, что

7. $\left(\overline{\operatorname{conv} A \cup \{0\}}\right)^\circ \overset{4}{=} (\operatorname{conv} A \cup \{0\})^\circ \overset{5}{=} (\operatorname{conv} A)^\circ \overset{6}{=} A^\circ$.

8. $(A \cup B)^\circ = A^\circ \cap B^\circ$, $\big(\operatorname{conv}(A \cup B)\big)^\circ = A^\circ \cap B^\circ$.

$\lhd$ Поскольку $A \cup B \supset A$, то по свойству 2 $(A \cup B)^\circ \subset A^\circ$, аналогично, $A \cup B \supset B$, то по свойству 2 $(A \cup B)^\circ \subset B^\circ$. Отсюда $(A \cup B)^\circ \subset A^\circ \cap B^\circ$.

Докажем обратное включение. Пусть функционал $y \in A^\circ \cap B^\circ$. Из того, что $y \in A^\circ$ следует, что $\sup\limits_{x \in A} \langle y, x\rangle \leq 1$; аналогично из того, что $y \in B^\circ$ следует, что $\sup\limits_{x \in B} \langle y, x\rangle \leq 1$. Отсюда $\sup\limits_{x \in A \cup B} \langle y, x\rangle \leq 1$. Значит, $y \in (A \cup B)^\circ$. $\rhd$

9. $(A \cap B)^\circ = \overline{\operatorname{conv}(A^\circ \cup B^\circ)}$ *для любых выпуклых замкнутых множеств A, B, содержащих ноль.*

$\lhd$ Обозначим $C = A^\circ$, $D = B^\circ$. Тогда C, D — выпуклые замкнутые множества, содержащие ноль (как поляры множеств). Следовательно, $C^\circ = A^{\circ\circ} = A$ по теореме о биполяре. Аналогично, $D^\circ = B^{\circ\circ} = B$. Тогда $(A \cap B)^\circ \overset{\text{def}}{=} (C^\circ \cap D^\circ)^\circ \overset{8}{=} \big((C \cup D)^\circ\big)^\circ \overset{7}{=}$
$= \left(\left(\overline{\operatorname{conv}(C \cup D)}\right)^\circ\right)^\circ = \overline{\operatorname{conv}(C \cup D)} = \overline{\operatorname{conv}(A^\circ \cup B^\circ)}$. $\rhd$

Формулы 8-9 обобщаются на случай n множеств.

Лемма 2. *X — линейное нормированное пространство, $A \subset X$ — выпуклое множество. Тогда $0 \in \operatorname{int} A$ тогда и только тогда, когда A° — ограниченное множество.*

Доказательство. Необходимость. Пусть $0 \in \operatorname{int} A$. Докажем, что тогда A° — ограниченное множество. Действительно, условие $0 \in \operatorname{int} A$ означает, что существует радиус $r > 0$, для которого шар $B_r = \{x \in X \mid \|x\| \leq r\} \subset A$. Значит, если $y \in A^\circ$, то по определению поляры $1 \geq \sup\limits_{x \in A} \langle y, x\rangle \geq \sup\limits_{x \in B_r} \langle y, x\rangle = r\|y\|_{X^*}$. Отсюда $\|y\|_{X^*} \leq \frac{1}{r}$, т. е. A° — ограниченное множество.

Обратно. Пусть A° — ограниченное множество. Докажем, что отсюда следует, что $0 \in \operatorname{int} A$. Предположим противное, что $0 \notin \operatorname{int} A$. Тогда по теореме отделимости множество A можно отделить от точки 0, т. е. существует линейно непрерывный функционал $\hat{y} \neq 0$ такой, что $\sup\limits_{a \in A} \langle \hat{y}, a\rangle \leq \langle \hat{y}, 0\rangle = 0$. Из этого неравенства следует, что $\sup\limits_{a \in A} \langle t\hat{y}, a\rangle \leq 0$ для любого $t > 0$. Значит, функционал $t\hat{y} \in A^\circ$ для любого $t > 0$ — противоречие с ограниченностью множества A°. □

Лемма 3. *X — линейное нормированное пространство. Тогда $A \subset X$ — ограниченное множество тогда и только тогда, когда $0 \in \operatorname{int} A^\circ$.*

Доказательство. Пусть A — ограниченное множество. Тогда оно содержится в некотором шаре радиуса R, т. е. существует R такое, что $\|a\|_X \leq R$ для любого $a \in A$. Тогда

$$\sup_{a \in A} \langle y, a\rangle \leq \sup_{a \in A} \|y\|_{X^*} \cdot \|a\|_X \leq R\|y\|_{X^*} \leq 1 \text{ при } \|y\|_{X^*} \leq \frac{1}{R}.$$

Это означает, что шар $B_{\frac{1}{R}} \subset A^\circ$. Следовательно, $0 \in \operatorname{int} A^\circ$.

Обратно. Пусть $0 \in \operatorname{int} A^\circ$. Это означает, что существует радиус $r > 0$, для которого шар $B_r \subset A^\circ$. Тогда $\langle y, x\rangle \leq 1$ для любого $x \in A$, для любого $y \in A^\circ$. Отсюда

$$1 \geq \sup_{y \in A^\circ} \langle y, x\rangle \geq \sup_{y \in B_r} \langle y, x\rangle = r\|x\|_X.$$

Значит, $\|x\|_X \leq \frac{1}{r}$, т. е. A — ограниченное множество. □

4.6.3 Примеры поляр множеств

- $A = a \neq 0 \Rightarrow A^\circ = \{y \in Y \mid \langle y, a\rangle \leq 1\}$ — полупространство.
- $A = \{a_1, \ldots, a_n\} \Rightarrow A^\circ$ — пересечение n полупространств, т. е. многогранник.
- $A = \operatorname{conv}\{a_1, \ldots, a_n\}$ — многогранник $\Rightarrow A^\circ$ — многогранник.

Число граней многогранника A° равно числу вершин A, и наоборот, число вершин многогранника A° равно числу граней A.

- $A = \{x \in \mathbb{R}^n \mid \langle a, x\rangle = 0,\ a \neq 0\}$ — гиперплоскость $\Rightarrow A^\circ = \{\lambda a \in \mathbb{R}^n \mid \lambda \in \mathbb{R}\}$ — прямая.

$\lhd$ Представим функционал $y \in A^\circ$ в виде $y = \lambda a + b$, где $b \perp a$.
Тогда $\sup\limits_{\langle a,x\rangle=0} \langle \lambda a + b, x\rangle = \sup\limits_{\langle a,x\rangle=0} \langle b, x\rangle$ конечен, если $b = 0$ (если $b \neq 0$, то можно взять $x = \mu b \in A$ и $\sup\limits_{\mu} \langle b, \mu b\rangle = +\infty$). Следовательно, $y = \lambda a$. Поэтому $\sup\limits_{\langle a,x\rangle=0} \langle \lambda a, x\rangle = 0$ и, значит, $\lambda a \in A^\circ$. $\rhd$

- $A = \{x \in \mathbb{R}^n \mid \langle a, x\rangle \leq 0\}$ — полупространство $\Rightarrow A^\circ = \{y = \lambda a \in \mathbb{R}^n \mid \lambda \geq 0\} = \operatorname{cone} a$ — луч.

$\lhd$ Представим вектор $x \in A$ в виде $x = \mu a + b$, где $b \perp a$. Тогда $x \in A \Leftrightarrow \langle a, x\rangle \leq 0 \Leftrightarrow \langle a, \mu a + b\rangle \leq 0 \Leftrightarrow \mu\langle a, a\rangle \leq 0 \Leftrightarrow \mu \leq 0$. Представим $y \in A^\circ$ в виде $y = \lambda a + c$, где $c \perp a$. Тогда $\sup\limits_{x\in A} \langle y, x\rangle = \sup\limits_{\substack{\mu\leq 0\\ b\perp a}} \langle \lambda a + c, \mu a + b\rangle = \sup\limits_{\substack{\mu\leq 0\\ b\perp a}} (\lambda\mu\langle a, a\rangle + \langle c, b\rangle)$ конечен, если $c = 0$ (если $c \neq 0$, то можно взять $x = tc \in A$ и $\sup\limits_{t} \langle c, tc\rangle = +\infty$), $\lambda \geq 0$. В этом случае, $y = \lambda a$, $\lambda \geq 0$. Поэтому $\sup\limits_{x\in A} \langle y, x\rangle = \sup\limits_{\mu\leq 0} \lambda\mu\langle a, a\rangle \leq 0$ и, значит, $y \in A^\circ$. $\rhd$

- $A = \{x \in \mathbb{R}^n \mid \langle a, x\rangle \leq 1\}$, $a \neq 0 \Rightarrow A^\circ = \{y = \lambda a \in \mathbb{R}^n \mid 0 \leq \lambda \leq 1\} = [0, a]$ — отрезок.

$\lhd$ Представим $x \in A$ и $y \in A^\circ$ в виде $x = \mu a + b$, $y = \lambda a + c$, где $b, c \perp a$. Тогда $x \in A \Leftrightarrow \langle a, x\rangle \leq 1 \Leftrightarrow \langle a, \mu a + b\rangle \leq 1 \Leftrightarrow \mu\langle a, a\rangle \leq 1$, $\sup\limits_{x\in A} \langle y, x\rangle = \sup\limits_{\substack{\mu\langle a,a\rangle \leq 1 \\ b \perp a}} \langle \lambda a + c, \mu a + b\rangle = \sup\limits_{\substack{\mu\langle a,a\rangle \leq 1 \\ b \perp a}} (\lambda\mu\langle a, a\rangle + \langle c, b\rangle)$ конечен, если $c = 0$ (если $c \neq 0$, то можно взять $x = tc \in A$ и $\sup\limits_t \langle c, tc\rangle = +\infty$). В этом случае $y = \lambda a$. Следовательно, $\sup\limits_{x\in A} \langle y, x\rangle = \sup\limits_{\mu\langle a,a\rangle \leq 1} \lambda\mu\langle a, a\rangle \leq 1 \Leftrightarrow 0 \leq \lambda \leq 1$ и, значит, $y = \lambda a \in A^\circ$. $\rhd$

- $A = \{x \in X \mid \|x\|_X \leq 1\} \Rightarrow A^\circ = \big\{y \in X^* \mid \|y\|_{X^*} \leq 1\big\}$.

$\lhd$ $\sup\limits_{x\in A} \langle y, x\rangle = \sup\limits_{\|x\|_X \leq 1} \langle y, x\rangle = \|y\|_{X^*} \leq 1 \Leftrightarrow \|y\|_{X^*} \leq 1 \Rightarrow A^\circ = \big\{y \in X^* \mid \|y\|_{X^*} \leq 1\big\}$. $\rhd$

- $A = \{x \in \mathbb{R}^n \mid |x| \leq C\}$, $C > 0$, $|x| := \left(\sum\limits_{j=1}^n x_j^2\right)^{\frac{1}{2}} \Rightarrow A^\circ = \big\{y \in \mathbb{R}^n \mid |y| \leq \frac{1}{C}\big\}$.

$\lhd$ Поскольку $A = CB$, где $B = \{x \in \mathbb{R}^n \mid |x| \leq 1\}$, и по предыдущему $B^\circ = \big\{y \in \mathbb{R}^n \mid |y| \leq 1\big\}$, то $A^\circ = (CB)^\circ \overset{4}{=} \frac{1}{C} B^\circ = \big\{y \in \mathbb{R}^n \mid |y| \leq \frac{1}{C}\big\}$. $\rhd$

4.6.4 Биполяра. Теорема о биполяре

Поляра от поляры называется *биполярой*:
$A^{\circ\circ} := (A^\circ)^\circ = \{x \in X \mid \sup\limits_{y \in A^\circ} \langle y, x\rangle \le 1\}$.

Теорема (о биполяре). *Множество A совпадает со своей биполярой $A^{\circ\circ}$ ($A = A^{\circ\circ}$) тогда и только тогда, когда A — выпуклое замкнутое множество, содержащее ноль.*

Доказательство. Если $A = A^{\circ\circ} = (A^\circ)^\circ$, то множество A само является полярой, а поляра всегда является выпуклым замкнутым множеством, содержащим ноль.

Обратно. Пусть A — выпуклое замкнутое множество, содержащее ноль. Вложение $A \subset A^{\circ\circ}$ — очевидно. Действительно, если $x \in A$, то выполняется неравенство $\langle y, x\rangle \le 1$ для любого $y \in A^\circ$. Тогда по определению поляры $x \in (A^\circ)^\circ = A^{\circ\circ}$.

Докажем вложение $A \supset A^{\circ\circ}$. Предположим противное, что $A \not\supset A^{\circ\circ}$, значит, существует $\hat{x} \in A^{\circ\circ}$, $\hat{x} \notin A$. По теореме о строгой отделимости точку $\hat{x}$ можно строго отделить от множества A, т. е. существует линейный непрерывный функционал $\hat{y}$ такой, что

$$\langle \hat{y}, \hat{x}\rangle > \sup_{a \in A} \langle \hat{y}, a\rangle \ge 0 \text{ (так как } 0 \in A) \Longrightarrow \langle \hat{y}, \hat{x}\rangle > 0.$$

Следовательно, найдется такое $t > 0$, что будет выполняться двойное неравенство $\langle t\hat{y}, \hat{x}\rangle > 1 \ge \sup\limits_{a \in A} \langle t\hat{y}, a\rangle$. Это означает, что $t\hat{y} \in A^\circ$, $\hat{x} \notin A^{\circ\circ}$ — противоречие. □

Следствие 3. *Если A — выпуклое замкнутое множество и* $0 \in A$, *то множество A является полярой некоторого выпуклого замкнутого множества B, содержащего* 0 *($A = B^\circ$).*

Доказательство. По теореме о биполяре $A = A^{\circ\circ} = (A^\circ)^\circ = B^\circ$, где $B = A^\circ$ — выпуклое замкнутое множество и $0 \in B$, так как поляра всегда является выпуклым замкнутым подмножеством, содержащим ноль. □

Следствие 4. *Если A — произвольное множество, то биполяра $A^{\circ\circ} = \overline{\mathrm{conv}(A \cup \{0\})}$.*

Доказательство. В силу свойства 7 $\bigl(A^{\circ} = \bigl(\overline{\mathrm{conv}(A \cup \{0\})}\,\bigr)^{\circ}\bigr)$ имеем: $A^{\circ\circ} = (A^{\circ})^{\circ} \overset{7}{=} \bigl((\,\overline{\mathrm{conv}(A \cup \{0\})}\,)^{\circ}\bigr)^{\circ} = \overline{\mathrm{conv}(A \cup \{0\})}$ (по теореме о биполяре для выпуклого замкнутого множества $\overline{\mathrm{conv}(A \cup \{0\})}$, содержащего ноль). $\square$

4.7 Сопряженный конус

4.7.1 Определение сопряженного конуса

Для наглядности представления будем рассматривать пространство $\mathbb{R}^n$.

Определение. *K — конус в $\mathbb{R}^n$. Сопряженным конусом называется множество $K^* := \{y \in \mathbb{R}^n \mid \inf\limits_{x\in K}\langle y, x\rangle \geq 0\}$ или, что то же самое, множество $\{y \in \mathbb{R}^n \mid \langle y, x\rangle \geq 0 \ \forall\ x \in K\}$.*

Нетрудно видеть, что поскольку для конуса K точка $0 \in K$, то $\inf\limits_{x\in K}\langle y, x\rangle = -\infty$ или 0.

Замечание. *Для конусов условия $\inf\limits_{x\in K}\langle y, x\rangle \geq 0$ и $\inf\limits_{x\in K}\langle y, x\rangle \geq c$, где $c \leq 0$, равносильны тому, что $\langle y, x\rangle$ не принимает отрицательных значений.*

Доказательство. Действительно, если $\langle y, x\rangle < 0$ при некотором $x \in K$, то $\lambda x \in K$ для любого $\lambda > 0$, а $\langle y, x\rangle$ будет принимать любые отрицательные значения. Тогда оба неравенства $\langle y, x\rangle \geq 0$ и $\langle y, x\rangle \geq c$ не будут выполняться. □

Замечание. *Для конусов условия $\langle y, x\rangle \geq 0$ для любого $x \in K$ и $\langle y, x\rangle \geq c$ для любого $x \in K$, где $c \leq 0$, равносильны тому, что $\langle y, x\rangle$ не принимает отрицательных значений.*

Доказательство. Действительно, если $\langle y, x\rangle < 0$ при некотором $x \in K$, то $\lambda x \in K$ для любого $\lambda > 0$, а $\langle y, x\rangle$ будет принимать любые отрицательные значения. Тогда оба неравенства $\langle y, x\rangle \geq 0$ и $\langle y, x\rangle \geq c$ не будут выполняться. □

Лемма 4. *Сопряженный конус совпадает с минус полярой: $K^* = -K^\circ$.*

Доказательство. Имеем $y \in K^* \Leftrightarrow \langle y, x\rangle \geq 0$ для любого $x \in K$. Согласно замечанию это неравенство эквивалентно неравенству $\langle y, x\rangle \geq -1 \ \forall\ x \in K \Leftrightarrow \langle -y, x\rangle \leq 1 \ \forall\ x \in K \Leftrightarrow -y \in K^\circ$ $\Leftrightarrow y \in -K^\circ$. □

Из леммы можно выводить свойства сопряженного конуса.

4.7.2 Примеры сопряженных конусов

- L — *линейное подпространство* $\Rightarrow L^* = L^\perp$.

$\lhd$ Условие $y \in L^* \Leftrightarrow$ означает, что выполняется неравенство

$$\langle y, x \rangle \geq 0 \quad \forall\, x \in L. \tag{$*$}$$

Поскольку линейное подпространство L вместе с вектором x содержит и вектор $-x$, то неравенство $(*)$ выполняется в виде равенства

$$\langle y, x \rangle = 0 \quad \forall\, x \in L.$$

Это означает, что вектор $x \in L^\perp$. $\rhd$

- $K = \{x \in \mathbb{R}^n \mid \langle a, x \rangle \leq 0\} \Rightarrow K^* = \{-\lambda a \mid \lambda \geq 0\}$.

$\lhd$ Ранее в примере для множества $K = \{x \in \mathbb{R}^n \mid \langle a, x \rangle \leq 0\}$ вычислили $K^\circ = \{\lambda a \mid \lambda \geq 0\}$. Поэтому $K^* = -K^\circ = -\{\lambda a \mid \lambda \geq 0\} = \{-\lambda a \mid \lambda \geq 0\}$. $\rhd$

4.7.3 Теорема о втором сопряженном конусе

Теорема (о втором сопряженном конусе)**.** *K — выпуклый замкнутый конус тогда и только тогда, когда $K = K^{**}$.*

Доказательство. Следует из теоремы о биполяре. Условие $0 \in K$ выполняется по определению конуса. $\square$

Следствие (1)**.** *Пусть K — выпуклый замкнутый конус. Тогда K является сопряженным к некоторому выпуклому замкнутому конусу B ($K = B^*$).*

Доказательство. По теореме о втором сопряженном конусе $K = K^{**} = (K^*)^* = B^*$, где $B := K^*$. $\square$

Следствие (2)**.** *Пусть K — произвольный конус. Тогда $K^{**} = \overline{\operatorname{conv} K}$.*

Доказательство. В силу свойства 5 $K^* = (\,\overline{\operatorname{conv} K}\,)^*$ имеем

$$K^{**} = (K^*)^* \overset{5}{=} \left((\,\overline{\operatorname{conv} K}\,)^*\right)^* = \overline{\operatorname{conv} K}.$$

Последнее равенство выполняется по т. о втором сопряженном конусе для конуса $\overline{\operatorname{conv} K}$. $\square$

4.7.4 Свойства сопряженного конуса

1. *Сопряженный конус является выпуклым замкнутым конусом.*

2. $K_1 \supset K_2 \Rightarrow K_1^* \subset K_2^*$.

$\lhd$ Включение $K_1^* \subset K_2^*$ следует из неравенства

$$\inf_{x \in K_2} \langle y, x \rangle \overset{K_2 \subset K_1}{\geq} \inf_{x \in K_1} \langle y, x \rangle. \qquad \rhd$$

3. $(\,\overline{K}\,)^* = K^*$.

$\lhd$ $(\,\overline{K}\,)^* = -(\,\overline{K}\,)^\circ = -K^\circ = K^*$. $\rhd$

4. $(\operatorname{conv} K)^* = K^*$.

$\lhd$ $(\operatorname{conv} K)^* = -(\operatorname{conv} K)^\circ = -(K)^\circ = K^*$. $\rhd$

Из свойств 3-4 следует, что

5. $(\,\overline{\operatorname{conv} K}\,)^* \overset{3}{=} (\operatorname{conv} K)^* \overset{4}{=} K^*$. 6. K_1, K_2 — *произвольные конусы* $\Rightarrow (K_1 + K_2)^* = K_1^* \cap K_2^*$.

$\lhd$ Возьмем точки $x_1 \in K_1$, $x_2 \in K_2$. Поскольку $\langle y, x_1 + x_2 \rangle = \langle y, x_1 \rangle + \langle y, x_2 \rangle$, то

$$\inf_{\substack{x_1 \in K_1 \\ x_2 \in K_2}} \langle y, x_1 + x_2 \rangle = \inf_{\substack{x_1 \in K_1 \\ x_2 \in K_2}} (\langle y, x_1 \rangle + \langle y, x_2 \rangle) = \inf_{x_1 \in K_1} \langle y, x_1 \rangle + \inf_{x_2 \in K_2} \langle y, x_2 \rangle.$$

Отсюда, поскольку для конуса K $\inf\limits_{x \in K} \langle y, x \rangle = -\infty$ или 0, то

$$\inf_{\substack{x_1 \in K_1 \\ x_2 \in K_2}} \langle y, x_1 + x_2 \rangle \geq 0 \iff \begin{cases} \inf\limits_{x_1 \in K_1} \langle y, x_1 \rangle \geq 0, \\ \inf\limits_{x_2 \in K_2} \langle y, x_2 \rangle \geq 0. \end{cases}$$

Это и означает, выполнение доказываемой формулы.

7. K_1, K_2 — *произвольные конусы* $\Rightarrow (K_1 \cap K_2)^* \supset K_1^* + K_2^*$.

$\lhd$ Пусть функционал $y \in K_1^* + K_2^*$. Тогда $y = y_1 + y_2$, где $y_1 \in K_1^*$, $y_2 \in K_2^*$. Поскольку функционал $y_1 \in K_1^*$, то $\langle y_1, x \rangle \geq 0$ для любых точек $x \in K_1$ (в том числе и для $x \in K_1 \cap K_2$); аналогично, поскольку функционал $y_2 \in K_2^*$, то $\langle y_2, x \rangle \geq 0$ для любых точек $x \in K_2$ (в том числе и для $x \in K_1 \cap K_2$). Значит, $\langle y_1 + y_2, x \rangle \geq 0$ для любых точек $x \in K_1 \cap K_2$. Это и означает, что функционал $y \in (K_1 \cap K_2)^*$. $\rhd$

8. $\underline{K_1,K_2}$ — *выпуклые замкнутые конусы* $\Rightarrow$ $(K_1 \cap K_2)^* = \overline{K_1^* + K_2^*}$.

$\triangleleft$ По теореме о втором сопряженном конусе $K_1 = K_1^{**}$, $K_2 = K_2^{**}$. Следовательно,

$$(K_1 \cap K_2)^* = (K_1^{**} \cap K_2^{**})^* = \big((K_1^*)^* \cap (K_2^*)^*\big)^* \stackrel{6}{=}$$

(в силу свойства 6: $A^* \cap B^* = (A+B)^*$ для конусов $A = K_1^*$ и $B = K_2^*$ и следствия 2: $K^{**} = \overline{\operatorname{conv} K}$ для конуса $K = K_1^* + K_2^*$)

$$= \big((K_1^* + K_2^*)^*\big)^* = (K_1^* + K_2^*)^{**} \stackrel{\text{сл. } 2}{=} \overline{\operatorname{conv}(K_1^* + K_2^*)} = \overline{K_1^* + K_2^*},$$

так как сумма выпуклых множеств K_1^*, K_2^* есть выпуклое множество. $\triangleright$

Отметим, что убрать замыкание в последней формуле нельзя. Согласно теореме о втором сопряженном конусе формула 8 может быть переписана в виде: $8'$. $(K_1^* \cap K_2^*)^* = \overline{K_1 + K_2}$.

Пример. Два выпуклых замкнутых конусов, сумма которых есть незамкнутый конус.

$\triangleleft$ Пусть $K_1 = \operatorname{conev}\big\{\{x \in \mathbb{R}^3 \mid x_1^2 + (x_3 - 1)^2 \le 1, x_2 = 1\} \cup 0\big\}$ — выпуклый замкнутый конус в $\mathbb{R}^3$, образованный кругом и точкой 0; K_2 — ось x_2 — выпуклый замкнутый конус в $\mathbb{R}^3$. Тогда $K_1 + K_2 = \{x \in \mathbb{R}^3 \mid x_3 > 0\}$ — открытое полупространство, объединенное с осью x_2 — незамкнутый конус. $\triangleright$

Слева в формуле $8'$ стоит сопряженный конус — замкнутое множество, поэтому без замыкания множества $K_1 + K_2$ формула неверна.

4.7.5 Теорема Дубовицкого-Милютина о непересечении конусов

Из определения отделимости множеств следует, что если два непустых выпуклых конуса K_1 и K_2 отделимы, то существуют линейно непрерывный функционал $\lambda \in X^*$, $\lambda \neq 0$ такой, что

$$\inf_{a \in K_1} \langle \lambda, a \rangle \geq \sup_{b \in K_2} \langle \lambda, b \rangle.$$

Поскольку $0 \in K_1 \cap K_2$, то $\inf\limits_{a \in K_1} \langle \lambda, a \rangle \geq 0 \geq \sup\limits_{b \in K_2} \langle \lambda, b \rangle$. Это означает, что разделяющая гиперплоскость обязательно проходит через ноль.

Лемма. *X — нормированное пространство, $A \subset X$, $\lambda \neq 0$ — линейный функционал на X такой, что $\inf\limits_{a \in A} \langle \lambda, a \rangle \geq c$. Тогда $\langle \lambda, a \rangle > c$ для любой точки $a \in \operatorname{int} A$.*

Иными словами, ненулевой линейный функционал не может достигать минимума (аналогично максимума) во внутренней точке множества.

Доказательство. Предположим, что существует точка $a_0 \in \operatorname{int} A$, для которой $\langle \lambda, a_0 \rangle = c$. Возьмем любой вектор $v \in X$ такой, что $\langle \lambda, v \rangle < 0$. Такой вектор существует, ибо $\lambda \neq 0$. Тогда при малом $\varepsilon > 0$ точка $a_0 + \varepsilon v \in A$, и $\langle \lambda, a_0 + \varepsilon v \rangle = \langle \lambda, a_0 \rangle + \varepsilon \langle \lambda, v \rangle = c + \varepsilon \langle \lambda, v \rangle < c$ — противоречие. $\square$

Теорема (Дубовицкого-Милютина о непересечении конусов). *Пусть* $K_0, K_1, \ldots, K_m$ *— непустые выпуклые конусы, причем множества* $K_1 \setminus \{0\}, \ldots, K_m \setminus \{0\}$ *открыты. Тогда* $K_0 \cap K_1 \cap \ldots \cap K_m = 0$ *тогда и только тогда, когда существует линейно непрерывный функционал* $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in K_0^* \times K_1^* \times \ldots \times K_m^*$, $\lambda \neq 0$ *такой, что*

$$\lambda_0 + \lambda_1 + \ldots + \lambda_m = 0. \tag{$*$}$$

Доказательство. Достаточность. Пусть существуют линейно непрерывный функционал $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in K_0^* \times K_1^* \times \ldots \times K_m^*$, $\lambda \neq 0$, для которого выполнено уравнение $(*)$. Докажем, что тогда $K_0 \cap K_1 \cap \ldots \cap K_m = 0$. Предположим противное, что существует $x \in K_0 \cap K_1 \cap \ldots \cap K_m$, $x \neq 0$ и, значит, $x \in \operatorname{int} K_i$, $i = 0, 1, \ldots, m$. Если все функционалы $\lambda_i = 0$, $i = 1, \ldots, m$, то из уравнения $(*)$ следует, что $\lambda_0 = 0$ — противоречие с условием $\lambda \neq 0$. Следовательно, существует функционал $\lambda_s \neq 0$, $s \in \{1, \ldots, m\}$. Поскольку $\inf\limits_{x \in K_s} \langle \lambda_s, x \rangle \geq 0$, то по лемме ? для $\langle \lambda_s, x \rangle > 0$ для точки $x \in \operatorname{int} K_s$. Таким образом, для функционалов $\lambda_i \in K_i^*$ выполняется строгое неравенство $\sum\limits_{i=0}^{m} \langle \lambda_i, x \rangle > 0$ (все слагаемые неотрицательны, и по крайней мере одно положительно). Отсюда $\left\langle \sum\limits_{i=0}^{m} \lambda_i, x \right\rangle > 0$ — противоречие с неравенством $(*)$.

Необходимость. Пусть $K_0 \cap K_1 \cap \ldots \cap K_m = 0$. Докажем, что существует линейно непрерывный функционал $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m) \in K_0^* \times K_1^* \times \ldots \times K_m^*$, $\lambda \neq 0$ такой, что выполняется неравенство $(*)$. Рассмотрим конусы $K := K_1 \times \ldots \times K_m$ и $D := \{\overline{x} = (x, \ldots, x) \in \underbrace{K_0 \times \ldots \times K_0}_{m}\}$ — “диагональ” конуса K_0^m. Тогда $D \cap \operatorname{int} K = \emptyset$ (если существует $(x, \ldots, x) \in D \cap \operatorname{int} K$, то $x \in K_0 \cap \operatorname{int} K_1 \cap \ldots \cap \operatorname{int} K_m$ — что невозможно в силу заданного условия $K_0 \cap K_1 \cap \ldots \cap K_m = 0$).

По теореме отделимости существует линейно непрерывный функционал $\lambda = (\lambda_1, \ldots, \lambda_m) \neq 0$, для которого

$$\inf_{\overline{x}\in K} \langle \lambda, \overline{x} \rangle \overset{\text{для конусов}}{\geq 0 \geq} \sup_{\overline{x}\in D} \langle \lambda, \overline{x} \rangle \iff \inf_{\overline{x}\in K} \sum_{i=1}^{m} \langle \lambda_i, x_i \rangle \overset{a}{\geq} 0 \overset{b}{\geq} \sup_{x\in K_0} \sum_{i=1}^{m} \langle \lambda_i, x \rangle.$$

Подставляя точки $\overline{x} = (0, \ldots, \underset{i}{x_i}, \ldots, 0)$ в неравенство (a) получаем, что $\inf\limits_{x_i\in K_i} \langle \lambda_i, x_i \rangle \geq 0$. Это означает, что функционалы $\lambda_i \in K_i^*$, $i = 1, \ldots, m$. Обозначим $\sum\limits_{i=1}^{m} \lambda_i = -\lambda_0$. Тогда $\sum\limits_{i=0}^{m} \lambda_i = 0$.

Из неравенства (b) следует, что

$$0 \geq \sup_{x\in K_0} \left\langle \sum_{i=1}^{m} \lambda_i, x \right\rangle \iff 0 \geq \sup_{x\in K_0} \langle -\lambda_0, x \rangle$$

$$\iff 0 \geq -\inf_{x\in K_0} \langle \lambda_0, x \rangle \iff \inf_{x\in K_0} \langle \lambda_0, x \rangle \geq 0.$$

Это означает, что функционал $\lambda_0 \in K_0^*$. $\square$

Теорема 11. *Пусть $K_0, K_1, \ldots, K_m$ — непустые выпуклые конусы, причем $K_1 \backslash \{0\}, \ldots, K_m \backslash \{0\}$ открыты, $K_0 \cap K_1 \cap \ldots \cap K_m \neq 0$. Тогда $(K_0 \cap K_1 \cap \ldots \cap K_m)^* = K_0^* + K_1^* + \ldots + K_m^*$.*

Доказательство. Докажем вложение $(K_0 \cap K_1 \cap \ldots \cap K_m)^* \supset K_0^* + K_1^* + \ldots + K_m^*$. Пусть $\lambda \in K_0^* + \ldots + K_m^*$. Тогда $\lambda = \sum\limits_{i=0}^{m} y_i$, где $\lambda_i \in K_i^*$, $i = 0, \ldots, m$. Поскольку $\lambda_i \in K_i^*$, то $\langle \lambda_i, x \rangle \geq 0$ для любой точки $x \in K_i$ (в том числе и для точки $x \in K_0 \cap \ldots \cap K_m$). Значит, $\left\langle \sum\limits_{i=0}^{m} \lambda_i, x \right\rangle \geq 0$ для любой точки $x \in K_0 \cap \ldots \cap K_m$. Это означает, что функционал $\lambda \in (K_0 \cap \ldots \cap K_m)^*$.

Докажем обратное вложение $(K_0 \cap K_1 \cap \ldots \cap K_m)^* \subset K_0^* + K_1^* + \ldots + K_m^*$. Пусть функционал $\lambda \in (K_0 \cap \ldots \cap K_m)^*$. Надо представить его в виде суммы $\lambda = \sum\limits_{i=0}^{m} \lambda_i$, где $\lambda_i \in K_i^*$, $i = 0, \ldots, m$. Считаем функционал $\lambda \neq 0$ (иначе имеется тривиальное разложение: $0 = 0 + \ldots + 0$). Рассмотрим конус $K_{m+1} := \{x \in X \mid \langle \lambda, x \rangle < 0\} \cup \{0\}$ — открытое полупространство и ноль. Тогда $K_0 \cap \ldots \cap K_{m+1} = 0$ (иначе существует точка $x \in K_0 \cap \ldots \cap K_m$, для которой $\langle \lambda, x \rangle \geq 0$, и $x \in K_{m+1}$ и по определению конуса K_{m+1} будет выполняться неравенство $\langle \lambda, x \rangle < 0$ — противоречие). По теореме Дубовицкого-Милютина существуют линейно непрерывные функционалы $\lambda_i \in K_i^*$, $i = 0, \ldots, m+1$, такие, что $(\lambda_0, \ldots, \lambda_{m+1}) \neq 0$ и $\sum\limits_{i=0}^{m+1} \lambda_i = 0$. Поскольку $K_{m+1}^* = \{\lambda_{m+1} = -\alpha\lambda \mid \alpha \geq 0\}$ (проверьте!), то $\lambda_{m+1} = -\alpha\lambda$ $(\alpha \geq 0)$, т. е. $\alpha\lambda_{m+1} = \sum\limits_{i=0}^{m} \lambda_i$. Если $\alpha = 0$, то $\sum\limits_{i=0}^{m} \lambda_i = 0$ и хотя бы один из этих функционалов ненулевой. Но тогда опять по теореме Дубовицкого-Милютина получаем, что $K_0 \cap \ldots \cap K_m = 0$ — противоречие. Значит, $\alpha > 0$ и $\lambda_{m+1} = \sum\limits_{i=0}^{m} \frac{\lambda_i}{\alpha} \in K_0^* + \ldots + K_m^*$. $\square$

4.8 Некоторые формулы выпуклого анализа

Формулы для поляр $A^\circ := \{y \mid \sup\limits_{x\in A} \langle y, x\rangle \le 1\}$**.**

1. A° — выпуклое замкнутое множество, $0 \in A^\circ$.
2. $B \supset A \Rightarrow B^\circ \subset A^\circ$.
3. $(\lambda A)^\circ = \lambda^{-1} A^\circ \quad \forall\ \lambda > 0$.
4. $\left(\overline{A}\right)^\circ = A^\circ$.
5. $(A \cup \{0\})^\circ = A^\circ$.
6. $(\operatorname{conv} A)^\circ = A^\circ$.
7. $\left(\overline{\operatorname{conv} A \cup \{0\}}\right)^\circ = A^\circ$.
8. $(A \cup B)^\circ = A^\circ \cap B^\circ, \quad \big(\operatorname{conv}(A \cup B)\big)^\circ = A^\circ \cap B^\circ$.
9. $(A \cap B)^\circ = \overline{\operatorname{conv}(A^\circ \cup B^\circ)}$ $\forall$ выпуклых замкнутых множеств A, B, содержащих ноль.

Поляры некоторых множеств

1. $A = \{a \in \mathbb{R}^n\} \Rightarrow A^\circ = \{y \in \mathbb{R}^n \mid \langle y, a\rangle \le 1\}$.
2. $A = \{x \in \mathbb{R}^n \mid \langle a, x\rangle = 0\} \Rightarrow A^\circ = \{\lambda a \mid \lambda \in \mathbb{R}\}$.
3. $A = \{x \in \mathbb{R}^n \mid \langle a, x\rangle \le 0\} \Rightarrow A^\circ = \{\lambda a \mid \lambda \ge 0\}$.
4. $A = \{x \in \mathbb{R}^n \mid \langle a, x\rangle \ge 0\} \Rightarrow A^\circ = \{\lambda a \mid \lambda \le 0\}$.
5. $A = \{x \in \mathbb{R}^n \mid \langle a, x\rangle \le 1\} \Rightarrow A^\circ = \{\lambda a \mid 0 \le \lambda \le 1\}$.
6. $A = \{x \in \mathbb{R}^n \mid \langle a, x\rangle \ge 1\} \Rightarrow A^\circ = \{\lambda a \mid \lambda \le 0\}$.

Формулы для сопряженных конусов

1. K^* — выпуклый замкнутый конус, $0 \in K^*$.
2. $K_1 \supset K_2 \Rightarrow K_1^* \subset K_2^*$.
3. $\left(\overline{K}\right)^* = K^*$.
4. $(\operatorname{conv} K)^* = K^*$.
5. $\left(\overline{\operatorname{conv} K}\right)^* = K^*$.