Методы оптимизации. ДЗ на октябрь.

Прохоров Юрий, 776

1 Conjugate (dual) sets

Опр. Пусть $S \subset \mathbb{R}^n$. *Сопряженным* или *двойственным множеством* ко множеству S называется

$$S^* = \left\{ y \in \mathbb{R}^n \;\middle|\; \langle x, y \rangle \geq -1 \;\; \forall x \in S \right\}$$

Опр. Пусть $S \subset \mathbb{R}^n$ — конус. *Сопряженным конусом* называется

$$K^* = \left\{ y \in \mathbb{R}^n \;\middle|\; \langle x, y \rangle \geq 0 \;\; \forall x \in K \right\}$$

Свойства:

- S^* всегда выпукло, замкнуто и содержит 0.
- $S^* = \bigcap\limits_{x \in S} \left\{ y \mid \langle x, y \rangle \geq -1 \right\}$ — пересечение полупространств.
- $S^* = \left(\overline{S}\right)^*$, $\;S^* = \left(\mathrm{conv} S\right)^*$, $\;S^* = \left(S \cup \{0\}\right)^*$
- $S^{**} = \overline{\mathrm{conv}\left(S \cup \{0\}\right)}$
- $\left(\bigcup\limits_{\alpha} S_\alpha\right)^* = \bigcap\limits_{\alpha} S_\alpha^*$
- Для конуса K и произвольного множества S: $\quad (S + K)^* = S^* \cup K^*$
- Для конусов K_1, K_2, имеющих внутреннюю точку: $\quad (K_1 \cap K_2)^* = K_1 + K_2$

Теорема. Сопряженным ко множеству

$$S = \mathrm{conv}(x_1, \ldots, x_k) + \mathrm{cone}(x_{k+1}, \ldots, x_m)$$

является полиэдр (многогранник)

$$S^* = \left\{ p \;\middle|\; \langle p, x_i \rangle \geq -1, \; i = \overline{1, k}, \quad \langle p, x_j \rangle \geq 0, \; j = \overline{k+1, m} \right\}$$

Задача 1

Найти сопряженное множество к конусу $S = \mathrm{cone}\left\{(-3, 1), (2, 3), (4, 5)\right\}$.

Решение:

Согласно теореме выше, строим перпендикуляры к граням конуса, и тогда сопряженным конусом будет множество между этими перпендикулярами.

На рисунке 1 изображен конус S, а на рисунке 2 — сопряженный к нему конус S^* (см. ниже).

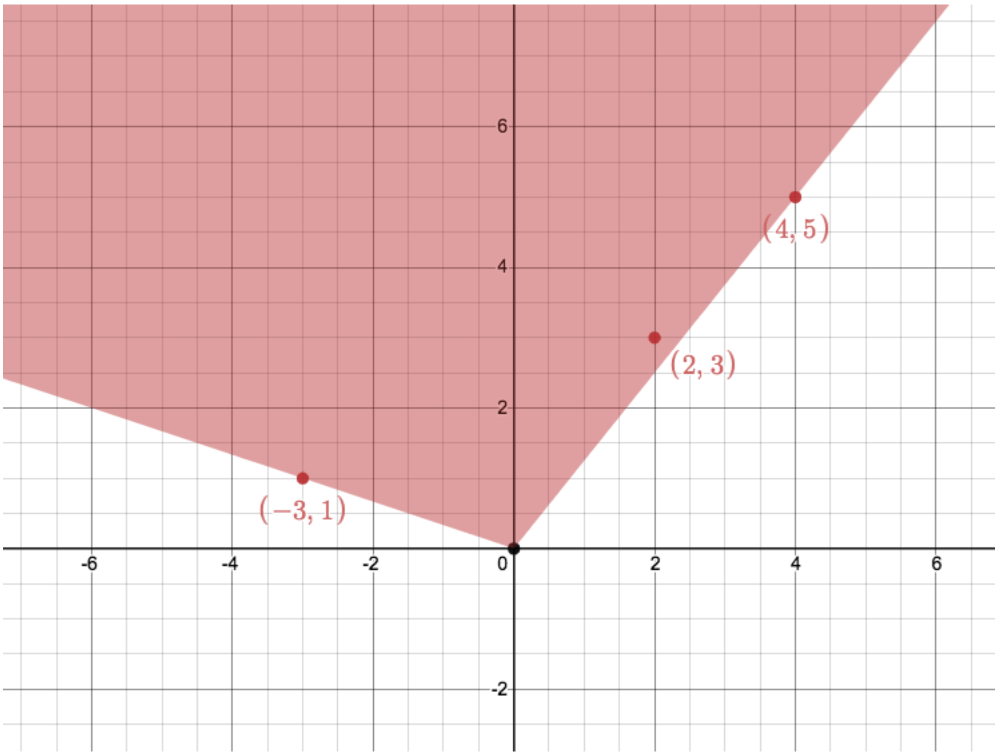


Рис. 1: Конус S

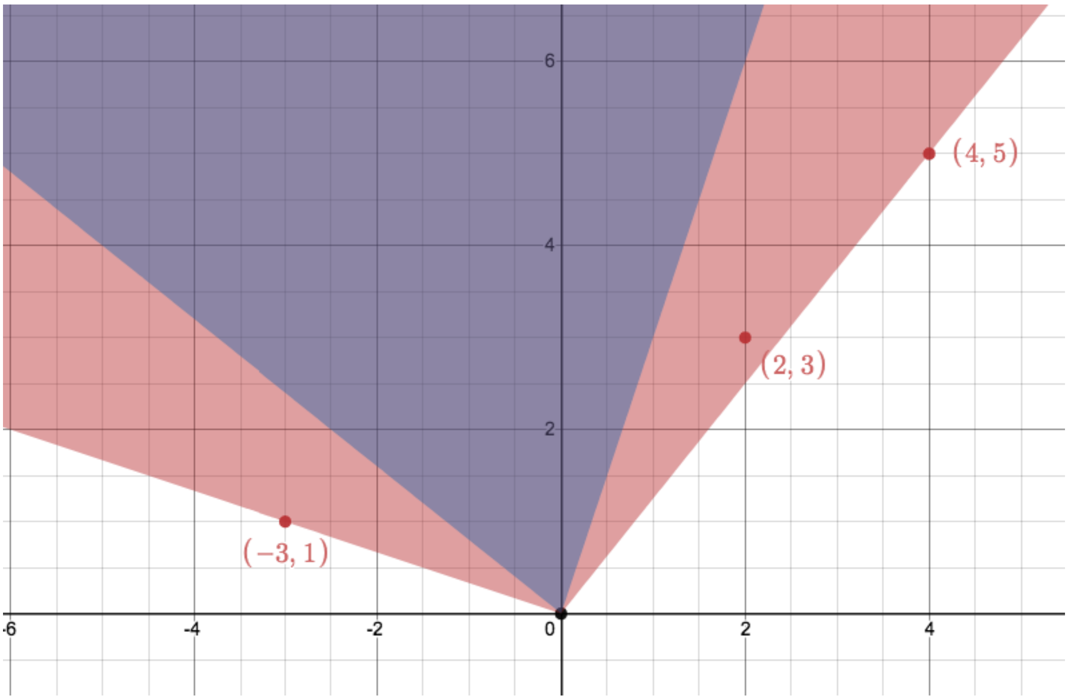


Рис. 2: Конус S (красный), сопряженный конус S^* (синий)

Задача 2

Найти S^*, S^{**}, S^{***} для

$$S = \{x \mid x_1 + x_2 \geq 0, \quad 2x_1 + x_2 \geq -4, \quad -2x_1 + x_2 \geq -4\}$$

Решение:

Заметим, что

$$S = \operatorname{conv}\left\{(-4,4), \left(\frac{4}{3}, -\frac{4}{3}\right)\right\} + \operatorname{cone}\{(-1,2),(1,2)\}$$

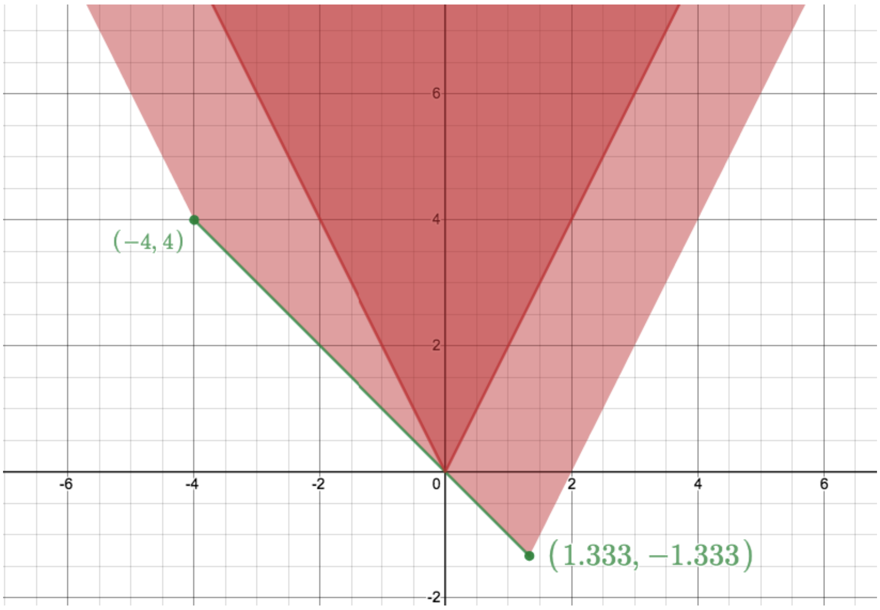


Рис. 3: Множество S

Тогда по теореме:

$$S^* = \{2y - x \geq 0\} \cap \{2y + x \geq 0\} \cap \left\{y \leq x + \frac{3}{4}\right\} \cap \left\{y \geq x - \frac{1}{4}\right\}$$

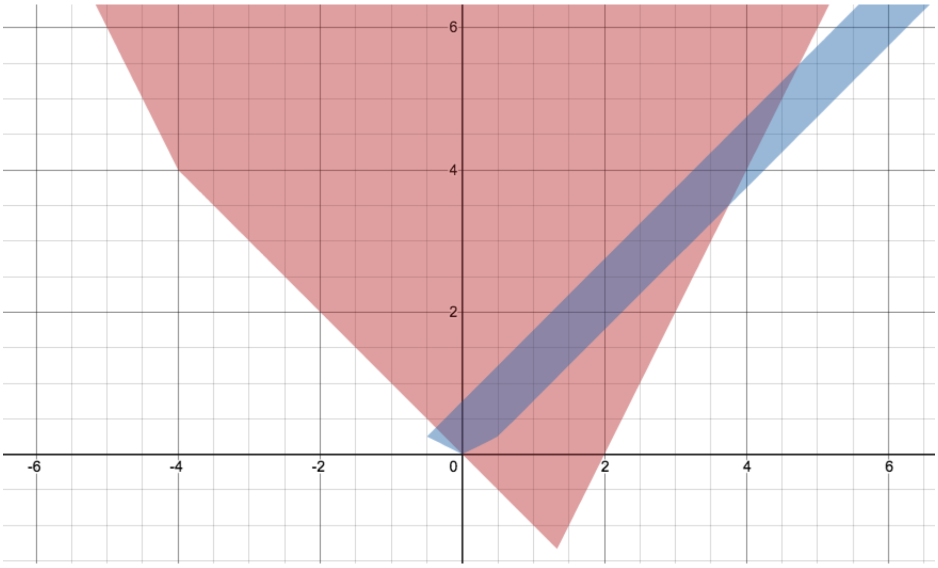


Рис. 4: Множество S (красный), множество S^* (синий)

Множества S и S^* выпуклы, замкнуты и содержат 0, поэтому $S^{**} = S, \quad S^{***} = S^*$.

Задача 3

Найти сопряженный конус к конусу симметричных положительно полуопределенных матриц $\mathbb{S}^n_+$.

Решение:

Источник решения: Stephen Boyd, Convex Optimization, p. 52, example 2.24

В пространстве матриц скалярное произведение задается естественным образом:

$$\langle X, Y\rangle = \sum_i \sum_j x_{ij} y_{ij} = \mathrm{tr}(X^T Y)$$

Покажем, что $\mathbb{S}^n_+$ — самосопряженный конус, то есть $(\mathbb{S}^n_+)^* = \mathbb{S}^n_+$. Другими словами,

$$Y \succeq 0 \qquad \Longleftrightarrow \qquad \mathrm{tr}(X^T Y) \geq 0 \ \ \forall X \succeq 0$$

Покажем два включения.

1. $(\mathbb{S}^n_+)^* \subset \mathbb{S}^n_+$

   Пусть $Y \in (\mathbb{S}^n_+)^*$. Допустим, $Y \notin \mathbb{S}^n_+$. Это означает, что

   $$\exists q \in \mathbb{R}^n : \quad q^T Y q = \sum_{i,j=1}^n y_{ij} q_i q_j = \sum_{i,j=1}^n (qq^T)_{ij} y_{ij} = \langle qq^T, Y\rangle < 0 \tag{$*$}$$

   Однако, как известно, матрица $qq^T \in \mathbb{S}^n_+$, значит, так как $Y \in (\mathbb{S}^n_+)^*$:

   $$\langle Y, qq^T\rangle \geq 0 \ \text{ — противоречие} \qquad \Longrightarrow \qquad y \in \mathbb{S}^n_+$$

2. $(\mathbb{S}^n_+)^* \supset \mathbb{S}^n_+$

   Пусть $X, Y \in \mathbb{S}^n_+$. Надо показать, что $\langle X, Y\rangle \geq 0$. Для любой симметричной матрицы существует ОНБ из собственных векторов, в котором она имеет диагональный вид:

   $$X = V\Lambda V^T, \ \ \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \succeq 0 \qquad \Longrightarrow \qquad X = \sum_{i=1}^n \lambda_i v_i v_i^T, \quad v \text{ — столбцы } V$$

   Тогда скалярное произведение матриц можно записать в виде (с учетом симметричности матриц):

   $$\langle X, Y\rangle = \mathrm{tr}\left(Y \sum_{i=1}^n \lambda_i v_i v_i^T\right) = \sum_{i=1}^n \lambda_i \mathrm{tr}(Y v_i v_i^T) = \sum_{i=1}^n \lambda_i v_i^T Y v_i \geq 0$$

   Последнее равенство получается так же, как и цепочка равенств $(*)$. Итак, $Y \in (\mathbb{S}^n_+)^*$.

Задача 4

Найти сопряженный конус для

$$K = \left\{(x, y, z) \ \middle|\ y > 0, \ ye^{x/y} \leq z\right\}$$

Решение:

Надо найти такие (a, b, c), что из

$$\begin{cases} ye^{x/y} \leq z \\ y > 0 \end{cases} \qquad \Longrightarrow \qquad ax + by + cz \geq 0$$

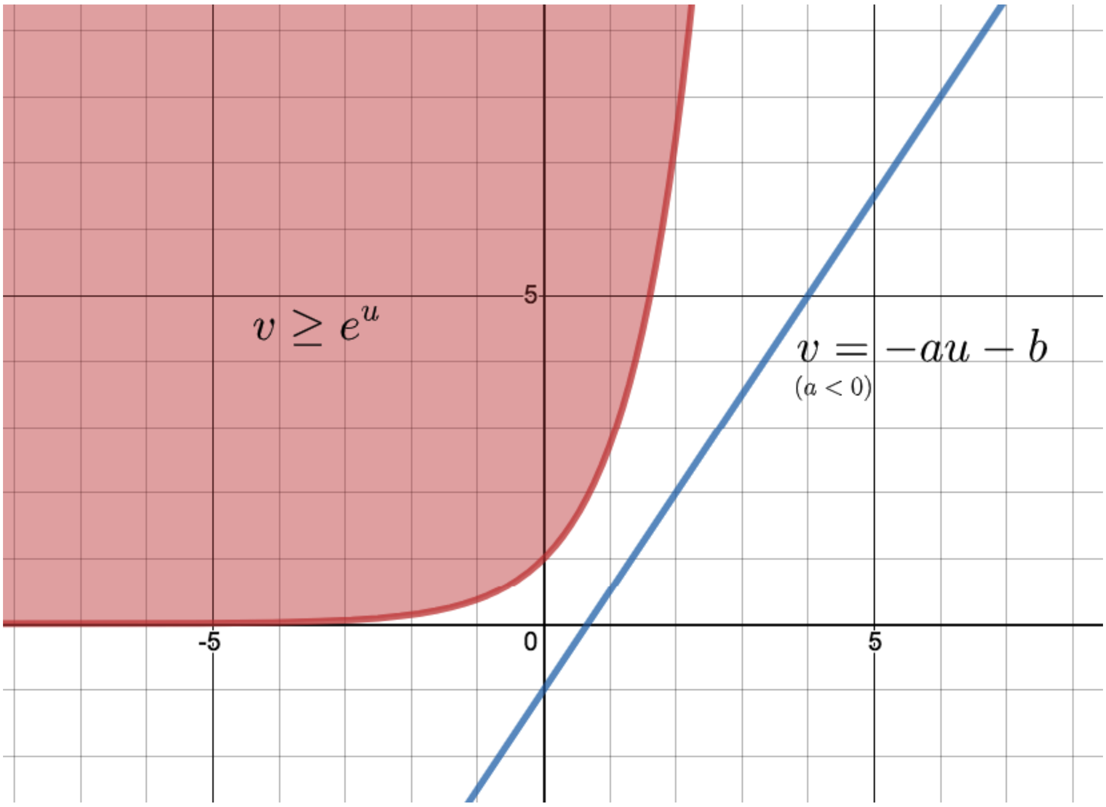
Обозначая $u = \frac{x}{y}, \ v = \frac{z}{y}$, получаем условие:

$$v \geq e^u \qquad \Longrightarrow \qquad au + b + cv \geq 0$$

Так как K^* — конус, то достаточно рассмотреть три случая: $c = -1, \ c = 0, \ c = 1$. Остальные случаи получаются умножением на положительную константу.

1. $c = 1$

   Нужно найти такие константы a и b, чтобы область $\{v \geq e^u\}$ лежала внутри области $\{v \geq -au - b\}$



   Это произойдет тогда и только тогда, когда синяя прямая на графике выше лежит ниже экспоненты. При $a > 0$ это невозможно. При $a = 0$ получается условие $b \geq 0$. При $a < 0$ найдем условие на касание:

$$\begin{cases} e^u = -au - b \\ e^u = -a \end{cases} \implies b = a\big(1 - \ln(-a)\big)$$

   Итак при $\;c = 1: \;\; a < 0: \;\; b \geq a\big(1 - \ln(-a)\big)$

   $a = 0: \;\; b \geq 0$

   $a > 0: \;\;$ нет таких b

2. $c = 0$.

   Нужно, чтобы $\big\{v \geq e^u\big\} \subset \{au + b \geq 0\}$. Это невозможно ни при каких a и b, так как $au + b = 0$ — вертикальная прямая.

3. $c < 0$.

   Аналогично, условие $\big\{v \geq e^u\big\} \subset \{v \leq au + b\}$ не выполнено ни при каких a и b, так как первое множество неограничено сверху при любых u.

Собирая все вместе, имеем

$$K^* = \left\{ \lambda(a, b, 1) \;\middle|\; \lambda \geq 0, \;\; \begin{array}{l} \text{если } a = 0, \text{ то } b \geq 0 \\ \text{если } a < 0, \text{ то } b \geq a\big(1 - \ln(-a)\big) \end{array} \right\}$$

Задача 5

Найти сопряженное множество для

$$S = \operatorname{conv}\big\{(-4,-1),(-2,-1),(-2,1)\big\} + \operatorname{cone}\big\{(1,0),(2,1)\big\}$$

Решение:

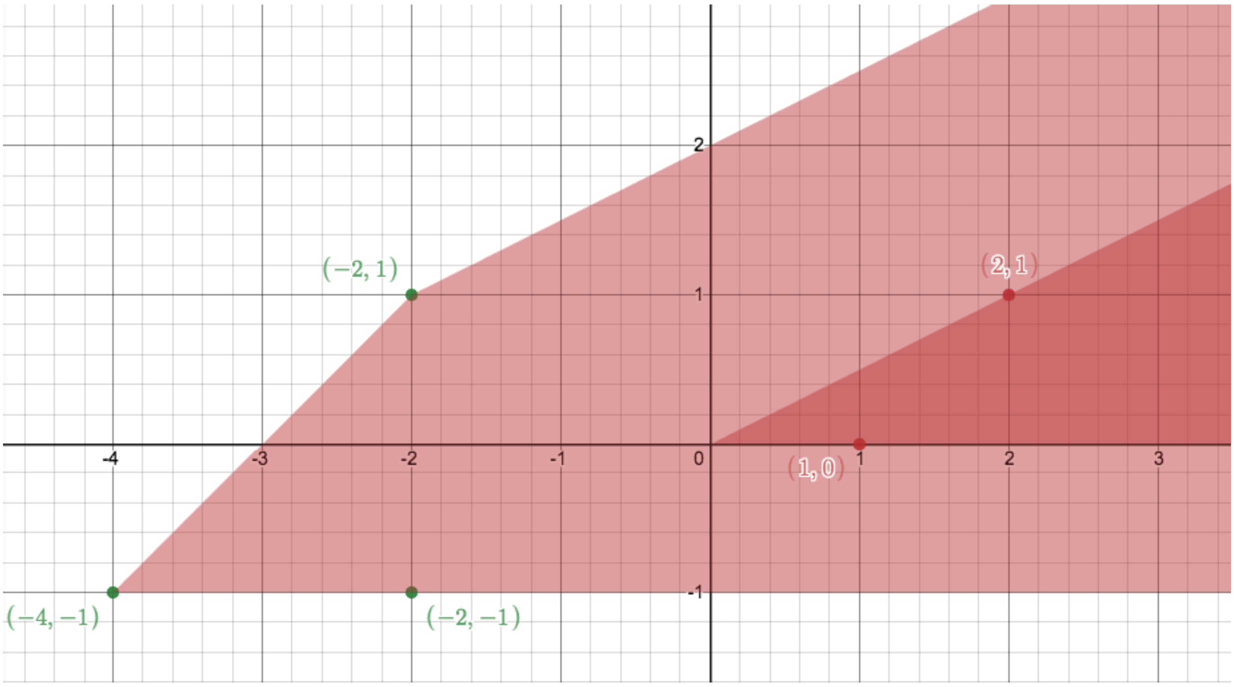


Рис. 5: Множество S

По теореме выше, имеем:

$$S^* = \{y + 4x \le 1\} \cap \{y + 2x \le 1\} \cap \{y - 2x \ge -1\} \cap \{x \ge 0\} \cap \{y + 2x \ge 0\}$$

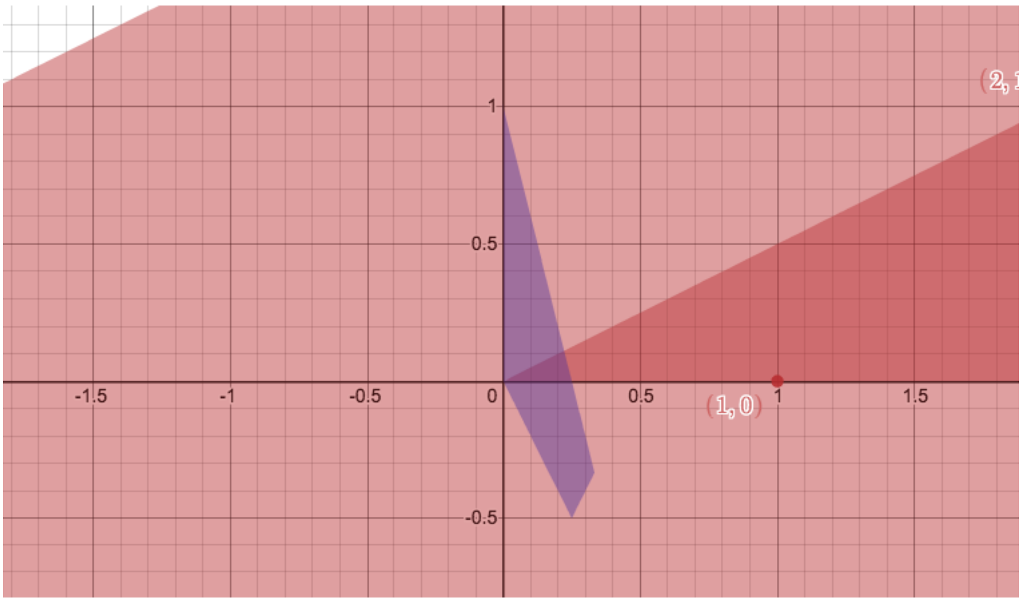


Рис. 6: Множество S^*

Задача 6

Доказать, что если определить сопряженное множество как

$$S^* = \{y \mid \langle x, y\rangle \leq 1 \ \ \forall x \in S\},$$

то $B_1(0) = \{x \mid \langle x, x\rangle \leq 1\}$ — единственное самосопряженное множество.

Решение:

1. Покажем, что $B = B_1(0)$ — самосопряженное множество.

   Пусть $y \in B$. Тогда для любого $x \in B$ по неравенству Коши-Буняковского:

   $$\langle x, y\rangle \leq |\langle x, y\rangle| \leq \|x\|\|y\| \leq 1 \cdot 1 = 1 \qquad \Longrightarrow \qquad y \in B^*$$

   Пусть $y \in B^*$. Тогда для любого $x \in B$ выполнено $\langle x, y\rangle \leq 1$. Если $y = 0$, то $y \in B$. При $y \neq 0$ возьмем $x = \frac{y}{\|y\|} \in B$, и тогда

   $$\langle x, y\rangle = \frac{\langle y, y\rangle}{\|y\|} = \|y\| \leq 1 \qquad \Longrightarrow \qquad y \in B$$

2. Покажем, что если $S \subset S^*$, то $S \subset B_1(0)$.

   Пусть $x \in S \subset S^*$. Тогда $\forall y \in S^* \ \ \langle x, y\rangle \leq 1$. При $y = x$ имеем:

   $$\langle x, y\rangle = \|x\|^2 \leq 1 \qquad \Longrightarrow \qquad x \in B_1(0)$$

3. Покажем, что если $S^* \subset S \subset B_1(0)$, то $B_1(0) \subset S$.

   Имеем, что если $\langle x, y\rangle \ \forall x \in S$, то $y \in S$. Покажем, что произвольный $y \in B_1(0)$ удовлетворяет этому условию, учитывая, что $S \subset B_1(0)$. По неравенству Коши-Буняковского для любого $x \in S$:

   $$\langle x, y\rangle \leq \|x\|\|y\| \leq 1 \qquad \Longrightarrow \qquad y \in S$$

Из пунктов $1, 2, 3$ следует утверждение задачи.

Утв. 1 Пусть A — невырожденная матрица, $S \subset \mathbb{R}^n$. Тогда $\big(AS\big)^* = A^{-T}S^*$.

Доказательство. Обозначим $P = AS$. Тогда

$$P^* = \{q \mid q^T p \geq -1 \ \ \forall p \in P\} = \{q \mid q^T Ax \geq -1 \ \ \forall x \in S\} = \{q \mid (A^T q)^T x \geq -1 \ \ \forall x \in S\} =$$
$$= \{A^{-T} y \mid y^T x \geq -1 \ \ \forall x \in S\} = A^{-T} S^* \quad \square$$

Задача 7

Найти сопряженное множество к эллипсоиду

$$S = \left\{ x \in \mathbb{R}^n \ \middle|\ \sum_{i=1}^n a_i^2 x_i^2 \leq 1 \right\}$$

Решение:

Сделаем преобразование с матрицей

$$A = \operatorname{diag}(a_1, \ldots, a_n), \qquad a_i > 0$$

Тогда $P = AS = B_1(0)$ — единичный шар. Известно, что это самосопряженное множество, поэтому $P^* = B_1(0)$. С другой стороны, в силу утверждения 1:

$$P^* = A^{-T} S^* \qquad \Longrightarrow \qquad S^* = A^T P^* = AB_1(0) = A\left\{ x \ \middle|\ \sum_{i=1}^n x_i^2 \leq 1 \right\} = \left\{ y \ \middle|\ \sum_{i=1}^n \frac{y_i^2}{a_i^2} \leq 1 \right\}$$

Выпишу еще какие-то утверждения, которые я получил, пока решал задачи.

Утв. 2 Пусть L — подпространство евклидового пространства X. Тогда $L^* = L^\perp$, где $L^\perp$ — ортогональное дополнение L.

Доказательство.
По определению, в евклидовом пространстве задано некоторое скалярное произведение $\langle x, y\rangle$. Покажем два включения. Сразу заметим, что L — конус.

1. $L^* \subset L^\perp$

   Пусть $y \in L^*$. Тогда $\forall x \in L$, так как $(-x) \in L$:

$$\langle x, y\rangle \geq 0, \quad \langle -x, y\rangle \geq 0 \qquad \Longrightarrow \qquad \langle x, y\rangle = 0 \qquad \Longrightarrow \qquad y \in L^\perp$$

2. $L^* \supset L^\perp$

   Пусть $y \in L^\perp$. Тогда $\forall x \in L$:

$$\langle x, y\rangle = 0 \geq 0 \qquad \Longrightarrow y \in L^*$$

□

Следствие. Пусть $\mathbb{A}_n$ — множество антисимметричных матриц порядка n. Тогда $(\mathbb{A}_n)^* = \mathbb{S}_n$

Доказательство.
В силу утверждения 2, достаточно показать, что $(\mathbb{A}_n)^\perp = \mathbb{S}_n$. Покажем два включения.

1. $(\mathbb{A}_n)^\perp \subset \mathbb{S}_n$

   Пусть $Y \in (\mathbb{A}_n)^\perp$. Тогда возьмем антисимметричную матрицу A такую, что

$$a_{ij} = 0, \text{ кроме } a_{kl} = -a_{lk} = 1 \quad (k \neq l)$$

   Тогда

$$\langle Y, A\rangle = y_{kl} - y_{lk} = 0 \qquad \Longrightarrow \qquad y_{kl} = y_{lk} \quad \forall k \neq l \qquad \Longrightarrow \qquad Y \in \mathbb{S}_n$$

2. $(\mathbb{A}_n)^\perp \supset \mathbb{S}_n$

   Пусть $Y \in \mathbb{S}_n$. Тогда для любой антисимметричной матрицы A:

$$\langle Y, A\rangle = \sum_{i,j=1}^{n} y_{ij}a_{ij} = \sum_{i<j} \big(y_{ij}a_{ij} - y_{ij}a_{ij}\big) = 0 \qquad \Longrightarrow \qquad Y \in (\mathbb{A}_n)^\perp$$

□

2 Conjugate functions

Опр. Пусть $f : \mathbb{R}^n \to \mathbb{R}$. *Сопряженной* к функции f функцией называется

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

Областью определения f^* является множество таких y, что супремум в определении выше конечен.

Свойства:

- f^* — выпуклая функция;
- $f^{**} = f \iff f$ — выпуклая замкнутая функция (*Теорема Фенхеля-Моро*);
- Неравенство Фенхеля-Юнга: $f(x) + f^*(y) \geq \langle y, x \rangle$
- Если $f(x) \leq g(x)$, то $f^*(y) \geq g^*(y)$
- Если $f(x, y) = f_1(x) + f_2(y)$ и функции f_1, f_2 выпуклы, то $f^*(p, q) = f^*(p) + f^*(q)$

Как искать сопряженную функцию к дифференцируемой фукнции $f(x)$:

1. Положить $g(x, y) = \langle y, x \rangle - f(x)$.
2. Определить, при каких y $\sup\limits_{x \in \mathbb{R}} g(x, y)$ конечен — это область определения f^*.
3. Найти максимум $g(x, y)$ по x: $\nabla_x g(x, y) = y - \nabla f(x) = 0$.

   Часто получается (но не всегда), что все значения y, при которых уравнение $y = \nabla f(x)$ разрешимо относительно x, есть область определения f^*.
4. Выразить x через y и подставить в $g(x, y)$ — это выражение для $f^*(y)$.

Опр. Пусть $(X, \|\cdot\|)$ — линейное нормированное пространство. *Сопряженным пространством* X^* называется множество всех линейных непрерывных функционалов на X.

Действие функционала $y \in X^*$ на элементе $x \in X$ обозначается $\langle y, x \rangle$.

Опр. *Сопряженной нормой* (dual norm) на X^* называется функция $\|\cdot\|_* : X^* \to \mathbb{R}$:

$$\|y\|_* = \sup_{x \neq 0} \frac{\big|\langle y, x \rangle\big|}{\|x\|} = \sup_{\|x\| \leq 1} \big|\langle y, x \rangle\big|$$

Свойства:

- $(X^*, \|\cdot\|_*)$ — линейное нормированное пространство.
- Неравенство Коши-Буняковского-Шварца: $\langle y, x \rangle \leq \|y\|_* \cdot \|x\|$.
- Сопряженным пространством ко множеству столбцов $\mathbb{R}^n$ является множество всех строк $\mathbb{R}^n$, а $\langle y, x \rangle$ является обычным скалярным произведением.
- Сопряженной нормой к $\|\cdot\|_p$ на $\mathbb{R}^n$ является $\|\cdot\|_q$, где $\dfrac{1}{p} + \dfrac{1}{q} = 1, \quad p > 1$.
- Сопряженной нормой к $\|\cdot\|_1$ является $\|\cdot\|_\infty$, сопряженной нормой к $\|\cdot\|_\infty$ является $\|\cdot\|_1$.
- Норма $\|\cdot\|_2$ самосопряжена.

Сопряженная норма не является сопряженной функцией для $f(x) = \|x\|$. Сопряженной функцией будет

$$f^*(y) = \begin{cases} 0 & , \ \|y\|_* \leq 1; \\ +\infty & , \ \text{иначе.} \end{cases}$$

Задача 1

Найти $f^*(y)$, если $f(x) = -\frac{1}{x},\ x > 0.$

Решение:

$g(x, y) = yx + \frac{1}{x}.$

При $y > 0: \quad g(x, y) \to +\infty$ при $x \to +\infty.$

При $y \leq 0: \quad g(x, y) \to +\infty$ при $x \to +0.$

Поэтому $f^*(y)$ нигде не определена, либо $f^*(y) = +\infty \quad \forall y \in \mathbb{R}.$

Задача 2

Найти $f^*(y)$, если $f(x) = -\frac{1}{2} - \ln x,\ x > 0.$

Решение:

$g(x, y) = yx + \frac{1}{2} + \ln x.$

При $y \geq 0: \quad g(x, y) \to +\infty$ при $x \to +\infty.$

При $y < 0$ найдем максимум: $\nabla_x g(x, y) = y + \frac{1}{x} = 0 \quad \Longrightarrow \quad x = -\frac{1}{y}.$

$$f^*(y) = -\frac{1}{2} - \ln(-y), \qquad y < 0$$

Задача 3

Найти $f^*(y)$, если $f(x) = \log \sum_{i=1}^{n} e^{x_i}.$

Решение:

$g(x, y) = y^T x - \log \sum_{i=1}^{n} e^{x_i}.$

$\nabla g(x, y) = 0 \quad \Longleftrightarrow \quad y_k = \frac{e^{x_k}}{\sum_{i=1}^{n} e^{x_i}} \quad \forall k = \overline{1, n} \qquad (*)$

Легко видеть, что $(*)$ разрешимо тогда и только тогда, когда

$$y_k > 0 \ \ \forall k, \qquad y_1 + \ldots + y_n = \mathbf{1}^T y = 1$$

Покажем, что множество таких y (но с нестрогим неравенством: $y_k \geq 0$) будет областью определения f^*:

- Пусть $\exists y_k < 0.$

  Тогда $g(x, y) \to +\infty$ при $x_k \to -\infty, \quad x_i = 0 \ \ (i \neq k).$

- Пусть $y \succeq 0$, но $\mathbf{1}^T y \neq 1.$

  Тогда при $x = t\mathbf{1} = (t, \ldots, t)$: $\quad g(x, y) = t \cdot \mathbf{1}^T y - t - \log n$

  Это выражение стремится к $+\infty$. При $\mathbf{1}^T y < 1$ это происходит при $t \to -\infty$, а при $\mathbf{1}^T y > 1$ — при $t \to +\infty$.

Теперь вычислим саму сопряженную функцию:

- $y \succ 0,\ \mathbf{1}^T y = 1$.

  $y_k = \dfrac{e^{x_k}}{\sum_{i=1}^n e^{x_i}} \quad \Longrightarrow \quad x_k = \log y_k + \log \sum\limits_{i=1}^n e^{x_i} \quad \Longrightarrow$

  $f^*(y) = \sum\limits_{k=1}^n y_k x_k - f(x) = \sum\limits_{k=1}^n y_k \log y_k + \mathbf{1}^T y \cdot \log \sum\limits_{i=1}^n e^{x_i} - \log \sum\limits_{i=1}^n e^{x_i} = \sum\limits_{k=1}^n y_k \log y_k$

- $y_1 = \ldots = y_m = 0$ (для определенности), $\mathbf{1}^T y = 1$.

  $g(x, y) = \sum\limits_{k=m+1}^n y_k x_k - f(x)$. Это выражение достигает супремума при

  $$x_1, \ldots, x_m \to -\infty, \quad x_k = \log y_k + \log \sum_{i=m+1}^n e^{x_i}, \quad k = \overline{m+1, n}$$

Итак, если положить $0 \log 0 = 0$, то

$$f^*(y) = \begin{cases} \sum\limits_{k=1}^n y_k \log y_k, & y \succeq 0,\ \mathbf{1}^T y = 1; \\ +\infty, & \text{иначе.} \end{cases}$$

Задача 4

Найти $f^*(y)$, если $f(x) = -\sqrt{a^2 - x^2},\ |x| < a,\ a > 0$.

Решение:

$g(x, y) = yx + \sqrt{a^2 - x^2}$.

$\nabla_x g(x, y) = y - \dfrac{x}{\sqrt{a^2 - x^2}} = 0 \quad \Longrightarrow \quad x = \dfrac{ay}{\sqrt{1 + y^2}}$

$f^*(y) = a\sqrt{1 + y^2}, \quad y \in \mathbb{R}$.

Задача 5

Найти $f^*(Y)$, если $f(X) = -\log\det X,\ X \in \mathbb{S}^n_{++}$.

Решение:

Будем предполагать, что задача рассматривается в линейном пространстве симметричных матриц со скалярным произведением $\langle X, Y \rangle = \mathbf{tr} XY$.

$g(X, Y) = \langle Y, X \rangle + \log\det X$.

$\nabla_X g(X, Y) = Y + \nabla\big(\log\det X\big) = Y + X^{-T} = Y + X^{-1} = 0 \quad \Longrightarrow \quad X = -Y^{-1}$

Найдем, при каких Y значение f^* конечно:

- Пусть $Y \not\prec 0$.

  Тогда $\exists$ собственное значение $\lambda \geq 0$ и собственный вектор e (пусть он нормирован на 1). Рассмотрим $X = I + tee^T$. Тогда

  $$g(X, Y) = \mathbf{tr}Y + t \cdot \mathbf{tr}(Yee^T) + \log\det(I + tuu^T) = \mathbf{tr}Y + \lambda t\|e\| + \log\det(I + tuu^T)$$

  Заметим, что матрица ee^T имеет ранг 1 и что ее единственное ненулевое собственное значение равно 1, а соответствующим собственным вектором является e. Поэтому собственные значения матрицы tee^T — $\{t, 0, \ldots, 0\}$, а матрицы $I + tee^T$ — $\{1 + t, 1, \ldots, 1\}$. Значит, $\det(I + tee^T) = 1 + t$.

  $$g(X, Y) = \mathbf{tr}Y + \lambda t + \log(1 + t) \to +\infty \quad \text{при} \quad t \to +\infty$$

- Пусть $Y \prec 0$.

  Тогда Y невырождена, значит, уравнение $\nabla g(X, Y) = 0$ разрешимо относительно X: $X = -Y^{-1}$.

$$f^*(Y) = \log\det(-Y^{-1}) - n = -\log\det(-Y) - n$$

Итак,

$$f^*(Y) = \begin{cases} -\log\det(-Y) - n & , \ Y \in -\mathbb{S}^n_{++}; \\ +\infty & , \ \text{иначе.} \end{cases}$$

Задача 6

Доказать, что если $f(x) = g(Ax)$, то $f^*(y) = g^*(A^{-T}y)$, где A — невырожденная матрица.

Решение:

Имеем по определению:

$$g^*(q) = \sup_{p \in \mathbb{R}^n} \left(q^T p - g(p) \right)$$

Тогда для функции $f(x) = g(Ax)$:

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left(y^T x - g(Ax) \right) = \Big[p = Ax \Big] = \sup_{p \in \mathbb{R}^n} \left(y^T A^{-1} p - g(p) \right) = g^*(A^{-T} y)$$

3 Subgradient and subdifferential

Опр. Пусть $f : S \to \mathbb{R},\ S \subset \mathbb{R}^n$. Вектор g называется *субградиентом* функции f в точке x_0, если

$$\forall x \in S \quad \to \quad f(x) - f(x_0) \geq \langle g, x - x_0 \rangle$$

Опр. Множество всех субградиентов f в точке x_0 называется *субдифференциалом* функции f в точке x_0 и обозначается $\partial f(x_0) \equiv \partial f_S(x_0)$.

Свойства:

- Если $x_0 \in \mathrm{relint} S$, то $\partial f_S(x_0)$ — выпуклое компактное множество;
- Выпуклая функция f дифференцируема в $x_0 \iff \partial f(x_0) = \{\nabla f(x_0)\}$;
- Если $\forall x \in S\ \ \partial f_S(x_0) \neq \varnothing$, то f выпукла на S;
- Если $\alpha \geq 0$, то $\partial(\alpha f)(x) = \alpha \partial f(x)$;
- Если f выпукла, то $\partial\big(f(Ax+b)\big)(x) = A^T \partial f(Ax+b)$.

Теорема Моро-Рокафеллара.

Пусть $f : E \to \mathbb{R},\ g : G \to \mathbb{R}$ — выпуклые функции, $x_0 \in E \cap G,\ E \cap \mathrm{int} G \neq \varnothing$. Тогда

$$\partial(f+g)(x_0) = \partial f(x_0) + \partial g(x_0)$$

Теорема Дубовицкого-Милютина.

Пусть $f_i : E_i \to \mathbb{R},\ i = \overline{1,m}$ — выпуклые функции, $x_0 \in \mathrm{int}\Big(\bigcap_{i=1}^m E_i\Big)$, а функция $f : \bigcap_{i=1}^m E_i \to \mathbb{R}$:

$$f(x) = \max\big\{f_1(x), \dots, f_m(x)\big\}$$

Тогда

$$\partial f(x_0) = \mathbf{conv}\Bigg(\bigcup_{j \in I} \partial f_j(x_0)\Bigg), \qquad I = \big\{j \ \big|\ f_j(x_0) = f(x_0)\big\}$$

Теорема о субдифференциале сложной функции.

Пусть $g_i : S \to \mathbb{R},\ i = \overline{1,m}$ — выпуклые функции, $\varphi : U \to \mathbb{R}$ — неубывающая (по всем переменным) выпуклая функция, причем $U \supset \big(g_1(S), \dots, g_m(S)\big),\ U \subset \mathbb{R}^m$ — открытое множество. Тогда при $f(x) = \varphi\big(g_1(x), \dots, g_m(x)\big)$:

$$\partial f(x) = \bigcup_{p \in \partial\varphi(u)} \Bigg(\sum_{i=1}^m p_i \partial g_i(x)\Bigg), \qquad u = \big(g_1(x), \dots, g_m(x)\big)$$

В частности, если φ дифференцируема в точке u, то

$$\partial f(x) = \sum_{i=1}^m \frac{\partial \varphi}{\partial u_i}(u) \partial g_i(x)$$

Задача 1

Доказать, что x_0 — точка минимума выпуклой функции f тогда и только тогда, когда $0 \in \partial f(x_0)$.

Решение:

Пусть $f : S \longrightarrow \mathbb{R}$ — выпуклая функция. Тогда $\forall x \in S$:

$$f(x) \geq f(x_0) \qquad \iff \qquad f(x) - f(x_0) \geq \langle 0, x - x_0 \rangle \qquad \iff \qquad 0 \in \partial f(x_0)$$

Задача 2

Найти $\partial f(x)$, если $f(x) = \mathrm{ReLU}(x) = \max\{0, x\}$.

Решение:

$\partial(0)(x) = \{0\}, \quad \partial(x)(x) = \{1\}$. По теореме Дубовицкого-Милютина (обе функции выпуклы):

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

Задача 3

Найти $\partial f(x)$, если $f(x) = \|x\|_p$.

(a) $p = 1$;

(b) $p = 2$;

(c) $p = \infty$;

Решение:

(a) $f(x) = \|x\|_1 = \sum\limits_{i=1}^{n} |x_i| = \sum\limits_{i=1}^{n} \max\{-x_i, x_i\}$.

$\partial(x_i)(x) = \nabla_x(x_i) = (0, \ldots, 1, \ldots, 0)^T$ — единица на i-ом месте. По теореме Дубовицкого-Милютина (все функции выпуклы):

$$\partial|x_i|(x) = \begin{cases} (0, \ldots, -1, \ldots, 0)^T & , \ x_i < 0; \\ \{0\} \times \ldots \times [-1, 1] \times \ldots \times \{0\} & , \ x_i = 0 \\ (0, \ldots, 1, \ldots, 0)^T & , \ x_i > 0. \end{cases}$$

По теореме Моро-Рокафеллара (все функции выпуклы):

$$\partial f(x) = \begin{cases} \big(\mathrm{sign}(x_1), \ldots, \mathrm{sign}(x_n)\big)^T & , \ \text{все } x_i \neq 0; \\ \prod\limits_{i \in J} \big\{\mathrm{sign}(x_i)\big\} \times \prod\limits_{i \notin J} [-1, 1] & , \ x_i \neq 0 \text{ при } i \in J; \\ [-1, 1]^n & , \ x = 0. \end{cases}$$

(b) $f(x) = \|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$.

При $x \neq 0$ функция $f(x)$ дифференцируема, поэтому $\partial f(x) = \dfrac{x}{\|x\|}$.

Покажем, что $\partial f(0) = B_1(0)$ — шар радиуса 1:

- Пусть $y \in B_1(0)$. По неравенству Коши-Буняковского:

$$\langle y, x \rangle \leq \big|\langle y, x \rangle\big| \leq \|y\| \cdot \|x\| \leq 1 \cdot f(x) \qquad \Longrightarrow \qquad y \in \partial f(0)$$

- Пусть $y \in \partial f(0)$. Допустим, $\|y\| > 1$.

  Известно, что норма $\|\cdot\|_2$ является самосопряженной нормой и что сопряженная норма определяется выражением

$$\|q\|_* = \sup_{\|p\| \leq 1} \langle q, p \rangle$$

  Тогда имеем

$$\|y\| = \sup_{\|x\| \leq 1} \langle y, x \rangle > 1 \quad \Longrightarrow \quad \exists x_0 : \ (\|x_0\| \leq 1) \ \ \langle y, x_0 \rangle > 1 \quad \Longrightarrow \quad \|x_0\| \leq 1 < \langle y, x_0 \rangle$$

  Это противоречит тому, что $y \in \partial f(0)$.

$$\partial f(x) = \begin{cases} \dfrac{x}{\|x\|} & , \ x \neq 0; \\ \\ B_1(0) & , \ x = 0. \end{cases}$$

(c) $f(x) = \max\limits_i |x_i| = \max\limits_i \big(\max\{-x_i, x_i\} \big)$.

$\partial |x_i|(x)$ такой же, как и в пункте (a). По теореме Дубовицкого-Милютина (все функции выпуклы):

$$\partial f(x) = \begin{cases} (0, \ldots, \operatorname{sign}(x_k), \ldots, 0)^T & , \ |x_i| < |x_k|, \ i = \overline{1, n}, \ i \neq k; \\ \\ \mathop{\mathbf{conv}}\limits_{j \in J} \begin{pmatrix} \ldots \\ \operatorname{sign}(x_j) \\ \ldots \end{pmatrix} & , \ x \neq 0, \ \max \text{ достигается на мн-ве индексов } J; \\ \\ [-1, 1]^n & , \ x = 0. \end{cases}$$

Задача 4

Найти $\partial f(x)$, если $f(x) = \big\|Ax - b\big\|_1^2$.

Решение:

1. Из задачи 3 мы знаем, что

$$\partial \|x\|_1(x) = \begin{cases} \big(\operatorname{sign}(x_1), \ldots, \operatorname{sign}(x_n)\big)^T & , \ \text{все } x_i \neq 0; \\ \prod\limits_{i \in J} \big\{\operatorname{sign}(x_i)\big\} \times \prod\limits_{i \notin J} [-1, 1] & , \ x_i \neq 0 \text{ при } i \in J; \\ [-1, 1]^n & , \ x = 0. \end{cases}$$

2. Для функции $\|x\|_1^2 = \varphi\big(\|x\|_1\big), \ \varphi(u) = u^2$ имеем по теореме о субдифференциале композиции:

   Для того, чтобы строго применить теорему, положим $\varphi(u) = \begin{cases} u^2 & , \ u \geq 0 \\ 0 & , \ u \leq 0 \end{cases}$. Тогда φ монотонно возрастает, выпукла и непрерывно дифференцируема.

$$\partial \|x\|_1^2(x) = 2\|x\|_1 \cdot \partial \|x\|_1(x)$$

3. Для функции $f(x) = \|Ax - b\|_1^2 = g(Ax - b), \ g(x) = \|x\|_1^2$ имеем по одному из свойств (функция g выпукла):

$$\partial f(x) = A^T \partial g(Ax - b) = A^T \cdot 2\|Ax - b\|_1 \cdot \partial \|\cdot\|_1(Ax - b)$$

Итак,

$$\partial f(x) = 2\|Ax - b\| \cdot A^T \partial \|\cdot\|_1(Ax - b)$$

Интересно, что функция f оказывается дифференцируемой при $Ax = b$, так как ее субдифференциал состоит из одной точки.

Утв. 3 Пусть $\|\cdot\|$ — норма на X, $\|\cdot\|_*$ — сопряженная норма на X^*. Тогда

$$\partial \|\cdot\|(x) = \big\{ v \in X^* \ \big| \ \langle v, x \rangle = \|x\|, \ \|v\|_* \leq 1 \big\}$$

Доказательство.
Обозначим множество в правой части равенства как $V(x)$.

1. $V(x) \subset \partial \|\cdot\|(x)$

   Пусть $y \in X$ — произвольный элемент. Тогда по неравенству Коши-Буняковского:

$$\|x\| + \langle v, y - x \rangle = \|x\| + \langle v, y \rangle - \langle v, x \rangle = \langle v, y \rangle \leq \|v\|_* \cdot \|y\| \leq 1 \cdot \|y\| \qquad \Longrightarrow \qquad v \in \partial \|\cdot\|(x)$$

2. $V(x) \supset \partial\|\cdot\|(x)$

   Имеем, что
$$\langle v, y\rangle - \|y\| \le \langle v, x\rangle - \|x\| \qquad \forall y \in X$$
   В левой части неравенства перейдем к супремуму по y. Получаем там сопряженную функцию к норме:
$$\sup_{y \in X} \big(\langle v, y\rangle - \|y\|\big) = \|\cdot\|^*(v) \le \langle v, x\rangle - \|x\| \tag{$*$}$$
   Как известно, сопряженная функция к норме — "единичный шар" в двойственной норме:
$$\|\cdot\|^*(v) = \begin{cases} 0 & , \ \|v\|_* \le 1; \\ +\infty & , \ \text{иначе.} \end{cases}$$
   Правая часть неравенства $(*)$ конечна, поэтому конечна и левая. Значит, $\|v\|_* \le 1$. При этом $(*)$ принимает вид:
$$\langle v, x\rangle \ge \|x\|$$
   С другой стороны, по неравенству Коши-Буняковского, $\langle v, x\rangle \le \|v\|_* \cdot \|x\| \le \|x\|$. Значит, $\langle v, x\rangle = \|x\|$.

   Итак, $\|v\|_* \le 1,\ \langle v, x\rangle = \|x\| \implies v \in V(x)$. □

Задача 5

Найти $\partial f(x)$, если $f(x) = \exp\big(\|x\|\big)$.

Решение:

Будем считать, что $\|\cdot\|$ — произвольная норма на некотором линейном нормированном пространстве X. По теореме о субдифференциале композиции (e^u возрастает и выпукла) и из утверждения 3:
$$\partial f(x) = e^{\|x\|} \cdot \partial\|\cdot\|(x) = e^{\|x\|} \cdot \big\{v \in X^* \ \big|\ \langle v, x\rangle = \|x\|,\ \|v\|_* \le 1\big\}$$
В частности, если $X = \mathbb{R}^n$, а $\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$ — евклидова норма (она самосопряжена), $X^* = \mathbb{R}^n$. Тогда
$$\partial f(x) = \begin{cases} e^{\|x\|} \dfrac{x}{\|x\|} & , \ x \ne 0; \\ \\ B_1(0) & , \ x = 0. \end{cases}$$