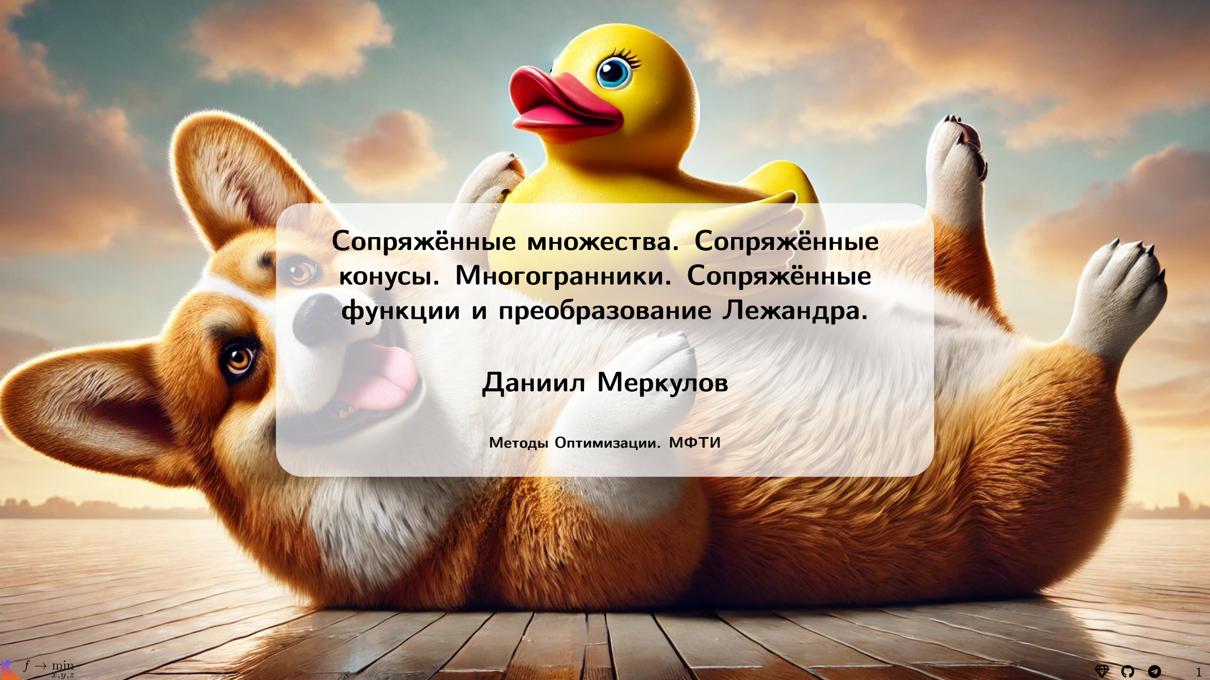


Сопряжённые множества. Сопряжённые конусы. Многогранники. Сопряжённые функции и преобразование Лежандра.

Даниил Меркулов

Методы Оптимизации. МФТИ

Повторение

Виды выпуклости

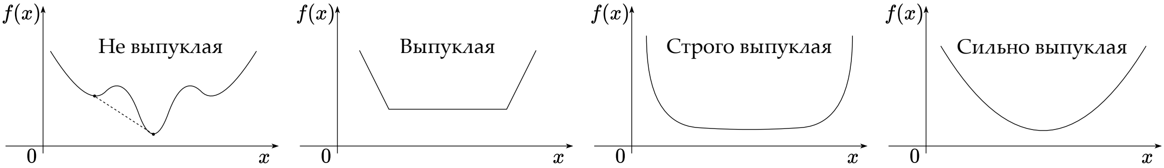


Рис. 1: Примеры выпуклых функций

Гладкость

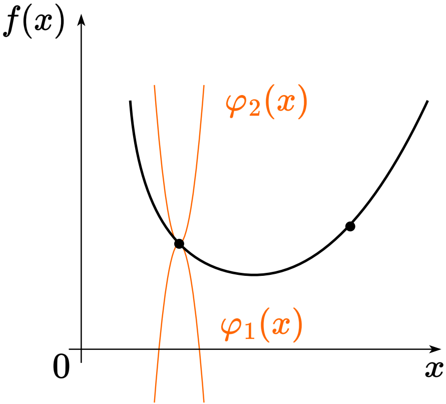


Рис. 2: Иллюстрация Липшицевых парабол, между которыми зажата гладкая функция. Чаще нас интересует мажорирующая из них.

Определение: Будем говорить, что функция $f : \mathbb{R}^n \to \mathbb{R}$ является L -гладкой, если $\forall x, y \in \mathbb{R}^n$ выполнено:

$$\|\nabla f(y) - \nabla f(x)\| \leqslant L\|y - x\|$$

Гладкость

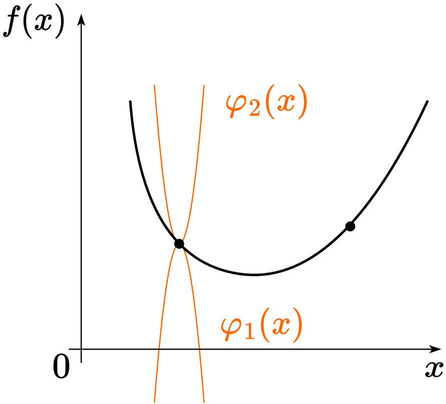


Рис. 2: Иллюстрация Липшицевых парабол, между которыми зажата гладкая функция. Чаще нас интересует мажорирующая из них.

Определение: Будем говорить, что функция $f : \mathbb{R}^n \to \mathbb{R}$ является L -гладкой, если $\forall x, y \in \mathbb{R}^n$ выполнено:

$$\|\nabla f(y) - \nabla f(x)\| \leqslant L\|y - x\|$$

Обратим внимание, что значение константы гладкости (Липшицевости градиента) зависит от выбора нормы.
Если $f : \mathbb{R}^n \to \mathbb{R}$ - непрерывно дифференцируема и градиент Липшицев с константой L, то $\forall x, y \in \mathbb{R}^n$:

$$\|f(y) - f(x) - \langle \nabla f(x), y - x \rangle\| \leqslant \frac{L}{2}\|y - x\|^2$$

Гладкость

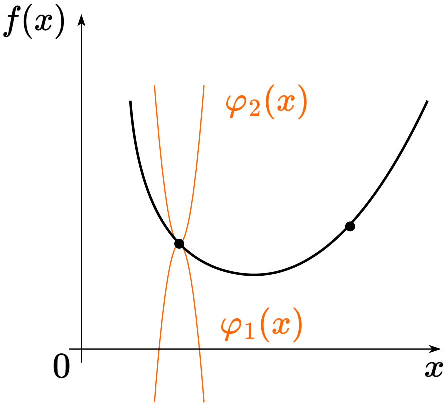


Рис. 2: Иллюстрация Липшицевых парабол, между которыми зажата гладкая функция. Чаще нас интересует мажорирующая из них.

Определение: Будем говорить, что функция $f : \mathbb{R}^n \to \mathbb{R}$ является L -гладкой, если $\forall x, y \in \mathbb{R}^n$ выполнено:

$$\|\nabla f(y) - \nabla f(x)\| \leqslant L\|y - x\|$$

Обратим внимание, что значение константы гладкости (Липшицевости градиента) зависит от выбора нормы.
Если $f : \mathbb{R}^n \to \mathbb{R}$ - непрерывно дифференцируема и градиент Липшицев с константой L, то $\forall x, y \in \mathbb{R}^n$:

$$\|f(y) - f(x) - \langle \nabla f(x), y - x \rangle\| \leqslant \frac{L}{2}\|y - x\|^2$$

Если зафиксируем $x_0 \in \mathbb{R}^n$, то:

$$\varphi_1(x) = f(x_0) + \langle f(x_0), x - x_0 \rangle - \frac{L}{2}\|x - x_0\|^2$$

$$\varphi_2(x) = f(x_0) + \langle f(x_0), x - x_0 \rangle + \frac{L}{2}\|x - x_0\|^2$$

Гладкость

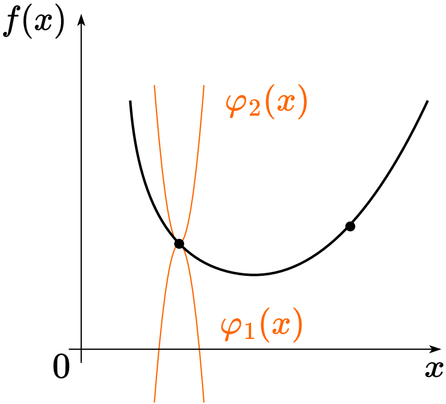


Рис. 2: Иллюстрация Липшицевых парабол, между которыми зажата гладкая функция. Чаще нас интересует мажорирующая из них.

Определение: Будем говорить, что функция $f : \mathbb{R}^n \to \mathbb{R}$ является L -гладкой, если $\forall x, y \in \mathbb{R}^n$ выполнено:

$$\|\nabla f(y) - \nabla f(x)\| \leqslant L\|y - x\|$$

Обратим внимание, что значение константы гладкости (Липшицевости градиента) зависит от выбора нормы.
Если $f : \mathbb{R}^n \to \mathbb{R}$ - непрерывно дифференцируема и градиент Липшицев с константой L, то $\forall x, y \in \mathbb{R}^n$:

$$\|f(y) - f(x) - \langle \nabla f(x), y - x \rangle\| \leqslant \frac{L}{2}\|y - x\|^2$$

Если зафиксируем $x_0 \in \mathbb{R}^n$, то:

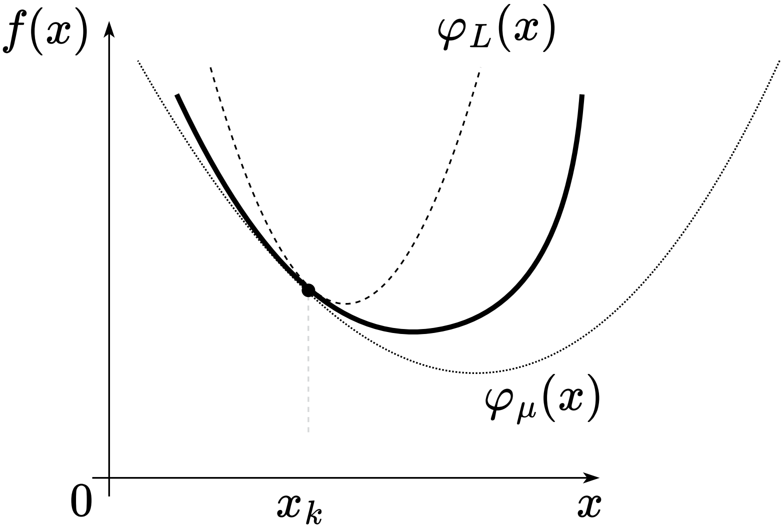
$$\varphi_1(x) = f(x_0) + \langle f(x_0), x - x_0 \rangle - \frac{L}{2}\|x - x_0\|^2$$

$$\varphi_2(x) = f(x_0) + \langle f(x_0), x - x_0 \rangle + \frac{L}{2}\|x - x_0\|^2$$

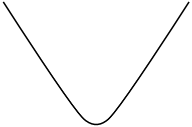
Это две параболы, и для них верно, что

$$\varphi_1(x) \leqslant f(x) \leqslant \varphi_2(x) \ \forall x$$

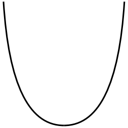
Гладкость и сильная выпуклость



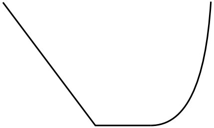
Гладкость и сильная выпуклость



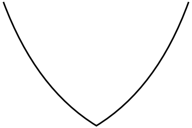
Гладкая
Выпуклая



Гладкая
μ - сильно выпуклая



Негладкая
Выпуклая



Негладкая
μ - сильно выпуклая

Сопряженные множества

Сопряженное множество

Пусть $S \subseteq \mathbb{R}^n$ — произвольное непустое множество.
Тогда его сопряжённое множество определяется как:

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

Сопряженное множество

Пусть $S \subseteq \mathbb{R}^n$ — произвольное непустое множество. Тогда его сопряжённое множество определяется как:

$$S^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \ \ \forall x \in S\}$$

Множество S^{**} называется двойным сопряжённым множеством к S, если:

$$S^{**} = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \ \ \forall x \in S^*\}$$

Сопряженное множество

Пусть $S \subseteq \mathbb{R}^n$ — произвольное непустое множество. Тогда его сопряжённое множество определяется как:

$$S^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \ \ \forall x \in S\}$$

Множество S^{**} называется двойным сопряжённым множеством к S, если:

$$S^{**} = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geq -1 \ \ \forall x \in S^*\}$$

- Множества S_1 и S_2 называются **взаимно сопряжёнными**, если $S_1^* = S_2, \ S_2^* = S_1$.

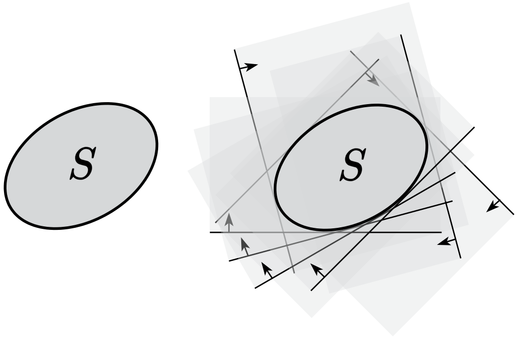


Рис. 3: Выпуклые множества можно описывать двойственным образом — через элементы множества и через множество опорных гиперплоскостей

Сопряженное множество

Пусть $S \subseteq \mathbb{R}^n$ — произвольное непустое множество. Тогда его сопряжённое множество определяется как:

$$
S^* = \{y \in \mathbb{R}^n \mid \langle y, x\rangle \geq -1 \ \ \forall x \in S\}
$$

Множество S^{**} называется двойным сопряжённым множеством к S, если:

$$
S^{**} = \{y \in \mathbb{R}^n \mid \langle y, x\rangle \geq -1 \ \ \forall x \in S^*\}
$$

- Множества S_1 и S_2 называются **взаимно сопряжёнными**, если $S_1^* = S_2,\ S_2^* = S_1$.
- Множество S называется **самосопряжённым**, если $S^* = S$.

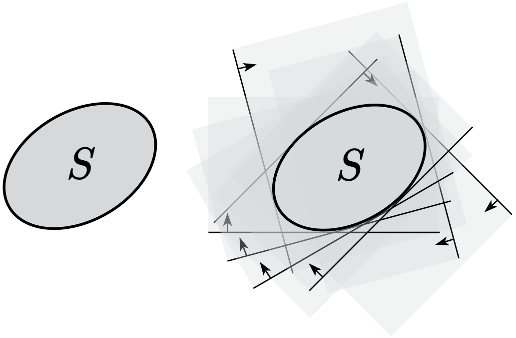


Рис. 3: Выпуклые множества можно описывать двойственным образом — через элементы множества и через множество опорных гиперплоскостей

Свойства сопряжённых множеств

- Сопряжённое множество всегда замкнуто, выпукло и содержит нуль.

Свойства сопряжённых множеств

- Сопряжённое множество всегда замкнуто, выпукло и содержит нуль.
- Для произвольного множества $S \subseteq \mathbb{R}^n$:

$$S^{**} = \overline{\mathbf{conv}(S \cup \{0\})}$$

Свойства сопряжённых множеств

- Сопряжённое множество всегда замкнуто, выпукло и содержит нуль.
- Для произвольного множества $S \subseteq \mathbb{R}^n$:

$$S^{**} = \overline{\mathbf{conv}(S \cup \{0\})}$$

- Если $S_1 \subseteq S_2$, то $S_2^* \subseteq S_1^*$.

Свойства сопряжённых множеств

- Сопряжённое множество всегда замкнуто, выпукло и содержит нуль.
- Для произвольного множества $S \subseteq \mathbb{R}^n$:

$$S^{**} = \overline{\mathbf{conv}(S \cup \{0\})}$$

- Если $S_1 \subseteq S_2$, то $S_2^* \subseteq S_1^*$.
- $\left(\bigcup\limits_{i=1}^m S_i\right)^* = \bigcap\limits_{i=1}^m S_i^*$.

Свойства сопряжённых множеств

- Сопряжённое множество всегда замкнуто, выпукло и содержит нуль.
- Для произвольного множества $S \subseteq \mathbb{R}^n$:

$$S^{**} = \overline{\mathbf{conv}(S \cup \{0\})}$$

- Если $S_1 \subseteq S_2$, то $S_2^* \subseteq S_1^*$.
- $\left(\bigcup\limits_{i=1}^m S_i \right)^* = \bigcap\limits_{i=1}^m S_i^*$.
- Если S замкнуто, выпукло и содержит 0, то $S^{**} = S$.

Свойства сопряжённых множеств

- Сопряжённое множество всегда замкнуто, выпукло и содержит нуль.
- Для произвольного множества $S \subseteq \mathbb{R}^n$:

$$S^{**} = \overline{\mathbf{conv}(S \cup \{0\})}$$

- Если $S_1 \subseteq S_2$, то $S_2^* \subseteq S_1^*$.
- $\left(\bigcup\limits_{i=1}^{m} S_i\right)^* = \bigcap\limits_{i=1}^{m} S_i^*$.
- Если S замкнуто, выпукло и содержит 0, то $S^{**} = S$.
- $S^* = \left(\overline{S}\right)^*$.

Пример 1

Example

Доказать, что $S^* = \left(\overline{S}\right)^*$.

Пример 1

Example

Доказать, что $S^* = \left(\overline{S}\right)^*$.

- $S \subset \overline{S} \;\Rightarrow\; \left(\overline{S}\right)^* \subset S^*$.

Пример 1

Example

Доказать, что $S^* = \left(\overline{S}\right)^*$.

- $S \subset \overline{S} \;\Rightarrow\; \left(\overline{S}\right)^* \subset S^*$.
- Пусть $p \in S^*$ и $x_0 \in \overline{S}$, $x_0 = \lim\limits_{k\to\infty} x_k$. Тогда в силу непрерывности функции $f(x) = p^T x$ имеем: $p^T x_k \geq -1 \;\Rightarrow\; p^T x_0 \geq -1$. Следовательно, $p \in \left(\overline{S}\right)^*$, и значит $S^* \subset \left(\overline{S}\right)^*$.

Пример 2

Example

Доказать, что $(\mathbf{conv}(S))^* = S^*$.

Пример 2

Example

Доказать, что $(\mathbf{conv}(S))^* = S^*$.

- $S \subset \mathbf{conv}(S) \;\; \Rightarrow \;\; (\mathbf{conv}(S))^* \subset S^*$.

Пример 2

Example

Доказать, что $(\mathbf{conv}(S))^* = S^*$.

- $S \subset \mathbf{conv}(S) \quad \Rightarrow \quad (\mathbf{conv}(S))^* \subset S^*$.
- Пусть $p \in S^*$ и $x_0 \in \mathbf{conv}(S)$, то есть

$$
x_0 = \sum_{i=1}^{k} \theta_i x_i, \;\; x_i \in S, \;\; \sum_{i=1}^{k} \theta_i = 1, \;\; \theta_i \geq 0
$$

.

Пример 2

ℹ Example

Доказать, что $\left(\mathbf{conv}(S)\right)^* = S^*$.

- $S \subset \mathbf{conv}(S) \;\; \Rightarrow \;\; \left(\mathbf{conv}(S)\right)^* \subset S^*$.
- Пусть $p \in S^*$ и $x_0 \in \mathbf{conv}(S)$, то есть

$$
x_0 = \sum_{i=1}^k \theta_i x_i, \;\; x_i \in S, \;\; \sum_{i=1}^k \theta_i = 1, \;\; \theta_i \geq 0
$$

.

- Тогда

$$
p^T x_0 = \sum_{i=1}^k \theta_i p^T x_i \;\geq\; \sum_{i=1}^k \theta_i \cdot (-1) = 1 \cdot (-1) = -1.
$$

Пример 2

ℹ Example

Доказать, что $(\mathbf{conv}(S))^* = S^*$.

- $S \subset \mathbf{conv}(S) \quad \Rightarrow \quad (\mathbf{conv}(S))^* \subset S^*$.
- Пусть $p \in S^*$ и $x_0 \in \mathbf{conv}(S)$, то есть

$$x_0 = \sum_{i=1}^k \theta_i x_i, \;\; x_i \in S, \;\; \sum_{i=1}^k \theta_i = 1, \;\; \theta_i \ge 0$$

.

- Тогда

$$p^T x_0 = \sum_{i=1}^k \theta_i p^T x_i \;\ge\; \sum_{i=1}^k \theta_i \cdot (-1) = 1 \cdot (-1) = -1.$$

- Значит, $p \in (\mathbf{conv}(S))^*$, и, следовательно, $S^* \subset (\mathbf{conv}(S))^*$.

Пример 3

Example

Докажите, что если $B(0, r)$ — это шар радиуса r в некоторой норме с центром в нуле, то $(B(0, r))^* = B(0, 1/r)$.

Пример 3

Example

Докажите, что если $B(0, r)$ — это шар радиуса r в некоторой норме с центром в нуле, то $(B(0, r))^* = B(0, 1/r)$.

- Пусть $B(0, r) = X, \ B(0, 1/r) = Y$. Возьмём вектор $p \in X^*$, тогда для любого $x \in X$: $p^T x \geq -1$.

Пример 3

Example

Докажите, что если $B(0, r)$ — это шар радиуса r в некоторой норме с центром в нуле, то $(B(0, r))^* = B(0, 1/r)$.

- Пусть $B(0, r) = X,\ B(0, 1/r) = Y$. Возьмём вектор $p \in X^*$, тогда для любого $x \in X$: $p^T x \geq -1$.
- Среди всех точек шара X возьмём такую $x \in X$, для которой скалярное произведение с p минимально: $p^T x$. Это точка $x = -\dfrac{p}{\|p\|} r$.

$$
p^T x = p^T \left(-\frac{p}{\|p\|} r \right) = -\|p\| r \geq -1
$$

$$
\|p\| \leq \frac{1}{r} \in Y
$$

Следовательно, $X^* \subset Y$.

Пример 3

Example

Докажите, что если $B(0, r)$ — это шар радиуса r в некоторой норме с центром в нуле, то $(B(0, r))^* = B(0, 1/r)$.

- Пусть $B(0, r) = X,\ B(0, 1/r) = Y$. Возьмём вектор $p \in X^*$, тогда для любого $x \in X$: $p^T x \geq -1$.
- Среди всех точек шара X возьмём такую $x \in X$, для которой скалярное произведение с p минимально: $p^T x$. Это точка $x = -\dfrac{p}{\|p\|} r$.

$$p^T x = p^T \left(-\frac{p}{\|p\|} r \right) = -\|p\| r \geq -1$$

$$\|p\| \leq \frac{1}{r} \in Y$$

Следовательно, $X^* \subset Y$.

- Теперь пусть $p \in Y$. Нужно показать, что $p \in X^*$, то есть $\langle p, x \rangle \geq -1$. Достаточно применить неравенство Коши–Буняковского:

$$\| \langle p, x \rangle \| \leq \|p\| \|x\| \leq \frac{1}{r} \cdot r = 1$$

Последнее верно, так как $p \in B(0, 1/r)$ и $x \in B(0, r)$.
Следовательно, $Y \subset X^*$.

Двойственный конус

Сопряжённым (двойственным) к конусу K называется множество K^* такое, что:

$$K^* = \{y \mid \langle x, y\rangle \geq 0 \quad \forall x \in K\}$$

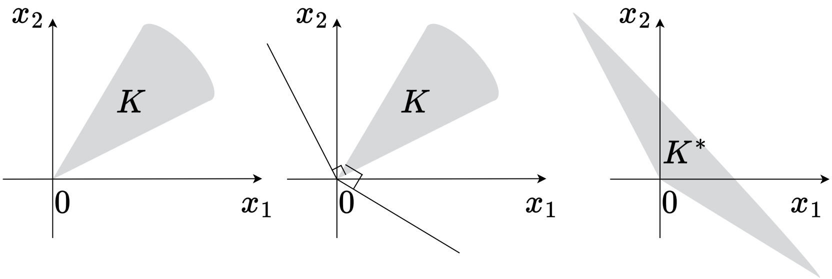
Двойственный конус

Сопряжённым (двойственным) к конусу K называется множество K^* такое, что:

$$K^* = \{y \mid \langle x, y\rangle \geq 0 \quad \forall x \in K\}$$

Чтобы показать, что это определение напрямую вытекает из предыдущих определений, напомним, что такое сопряжённое множество и что такое конус при $\forall \lambda > 0$.

$$\{y \in \mathbb{R}^n \mid \langle y, x\rangle \geq -1 \;\; \forall x \in S\} \;\; \to \;\; \{\lambda y \in \mathbb{R}^n \mid \langle y, x\rangle \geq -\frac{1}{\lambda} \;\; \forall x \in S\}$$



Свойства двойственных конусов

- Пусть K — замкнутый выпуклый конус. Тогда $K^{**} = K$.

Свойства двойственных конусов

- Пусть K — замкнутый выпуклый конус. Тогда $K^{**} = K$.
- Для произвольного множества $S \subseteq \mathbb{R}^n$ и конуса $K \subseteq \mathbb{R}^n$:

$$
(S + K)^* = S^* \cap K^*
$$

Свойства двойственных конусов

- Пусть K — замкнутый выпуклый конус. Тогда $K^{**} = K$.
- Для произвольного множества $S \subseteq \mathbb{R}^n$ и конуса $K \subseteq \mathbb{R}^n$:

$$(S + K)^* = S^* \cap K^*$$

- Пусть $K_1, \ldots, K_m$ — конусы в $\mathbb{R}^n$, тогда:

$$\left(\sum_{i=1}^{m} K_i \right)^* = \bigcap_{i=1}^{m} K_i^*$$

Свойства двойственных конусов

- Пусть K — замкнутый выпуклый конус. Тогда $K^{**} = K$.
- Для произвольного множества $S \subseteq \mathbb{R}^n$ и конуса $K \subseteq \mathbb{R}^n$:

$$(S + K)^* = S^* \cap K^*$$

- Пусть $K_1, \dots, K_m$ — конусы в $\mathbb{R}^n$, тогда:

$$\left(\sum_{i=1}^m K_i \right)^* = \bigcap_{i=1}^m K_i^*$$

- Пусть $K_1, \dots, K_m$ — конусы в $\mathbb{R}^n$. Если их пересечение имеет внутреннюю точку, то:

$$\left(\bigcap_{i=1}^m K_i \right)^* = \sum_{i=1}^m K_i^*$$

Пример

Example

Найдите сопряжённый конус для монотонного неотрицательного конуса:

$$
K = \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \ldots \geq x_n \geq 0\}
$$

Пример

Example

Найдите сопряжённый конус для монотонного неотрицательного конуса:

$$K = \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq ... \geq x_n \geq 0\}$$

Заметим, что:

$$\sum_{i=1}^{n} x_i y_i = y_1(x_1 - x_2) + (y_1 + y_2)(x_2 - x_3) + ... + (y_1 + y_2 + ... + y_{n-1})(x_{n-1} - x_n) + (y_1 + ... + y_n)x_n$$

Пример

Example

Найдите сопряжённый конус для монотонного неотрицательного конуса:

$$
K = \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \ldots \geq x_n \geq 0\}
$$

Заметим, что:

$$
\sum_{i=1}^{n} x_i y_i = y_1(x_1 - x_2) + (y_1 + y_2)(x_2 - x_3) + \ldots + (y_1 + y_2 + \ldots + y_{n-1})(x_{n-1} - x_n) + (y_1 + \ldots + y_n)x_n
$$

Так как во всей представленной сумме второй множитель в каждом слагаемом неотрицателен, то:

$$
y_1 \geq 0, \;\; y_1 + y_2 \geq 0, \;\; \ldots, \;\; y_1 + \ldots + y_n \geq 0
$$

Следовательно, $K^* = \left\{ y \mid \sum\limits_{i=1}^{k} y_i \geq 0, \;\; k = \overline{1, n} \right\}$.

Многогранники

Множество решений системы линейных неравенств и равенств является многогранником:

$$
Ax \preceq b, \quad Cx = d
$$

Здесь $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{p \times n}$, а знак неравенства понимается покомпонентно.

Многогранники

Множество решений системы линейных неравенств и равенств является многогранником:

$$Ax \preceq b, \quad Cx = d$$

Здесь $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{p \times n}$, а знак неравенства понимается покомпонентно.

Theorem

Пусть $x_1, \ldots, x_m \in \mathbb{R}^n$. Тогда сопряжённым к многогранному множеству

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

является многогранник:

$$S^* = \left\{ p \in \mathbb{R}^n \mid \langle p, x_i \rangle \geq -1, \ i = \overline{1,k}; \ \ \langle p, x_i \rangle \geq 0, \ i = \overline{k+1,m} \right\}$$

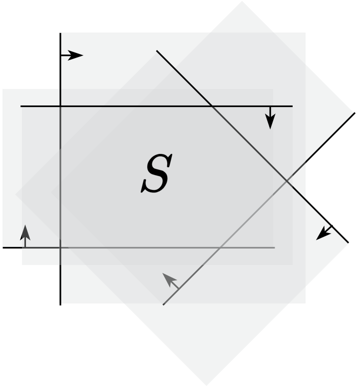


Рис. 5: Polyhedra

Доказательство

- Пусть $S = X,\ S^* = Y$. Возьмём некоторое $p \in X^*$, тогда $\langle p, x_i \rangle \geq -1,\ i = \overline{1,k}$. В то же время, для любого $\theta > 0,\ i = \overline{k+1,m}$:

$$\langle p, x_i \rangle \geq -1 \to \langle p, \theta x_i \rangle \geq -1$$

$$\langle p, x_i \rangle \geq -\frac{1}{\theta} \to \langle p, x_i \rangle \geq 0.$$

Следовательно, $p \in Y \to X^* \subset Y$.

Доказательство

- Пусть $S = X,\ S^* = Y$. Возьмём некоторое $p \in X^*$, тогда $\langle p, x_i \rangle \geq -1,\ i = \overline{1, k}$. В то же время, для любого $\theta > 0,\ i = \overline{k+1, m}$:

$$\langle p, x_i \rangle \geq -1 \to \langle p, \theta x_i \rangle \geq -1$$

$$\langle p, x_i \rangle \geq -\frac{1}{\theta} \to \langle p, x_i \rangle \geq 0.$$

  Следовательно, $p \in Y \to X^* \subset Y$.

- Предположим, наоборот, что $p \in Y$. Для любой точки $x \in X$: $x = \sum\limits_{i=1}^m \theta_i x_i, \quad \sum\limits_{i=1}^k \theta_i = 1, \quad \theta_i \geq 0$

  Тогда:

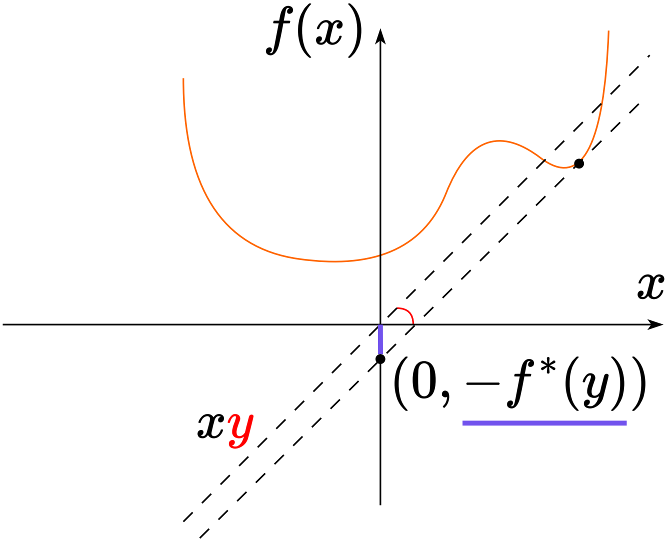
$$\langle p, x \rangle = \sum_{i=1}^m \theta_i \langle p, x_i \rangle = \sum_{i=1}^k \theta_i \langle p, x_i \rangle + \sum_{i=k+1}^m \theta_i \langle p, x_i \rangle \geq \sum_{i=1}^k \theta_i (-1) + \sum_{i=1}^k \theta_i \cdot 0 = -1.$$

  Следовательно, $p \in X^* \to Y \subset X^*$.

Пример

Сопряжённые функции

Сопряжённые функции

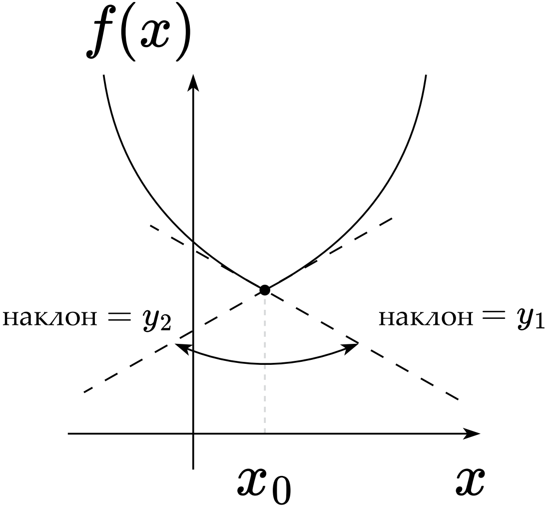


Напомним, что для отображения $f : \mathbb{R}^n \to \mathbb{R}$ функция, определяемая как

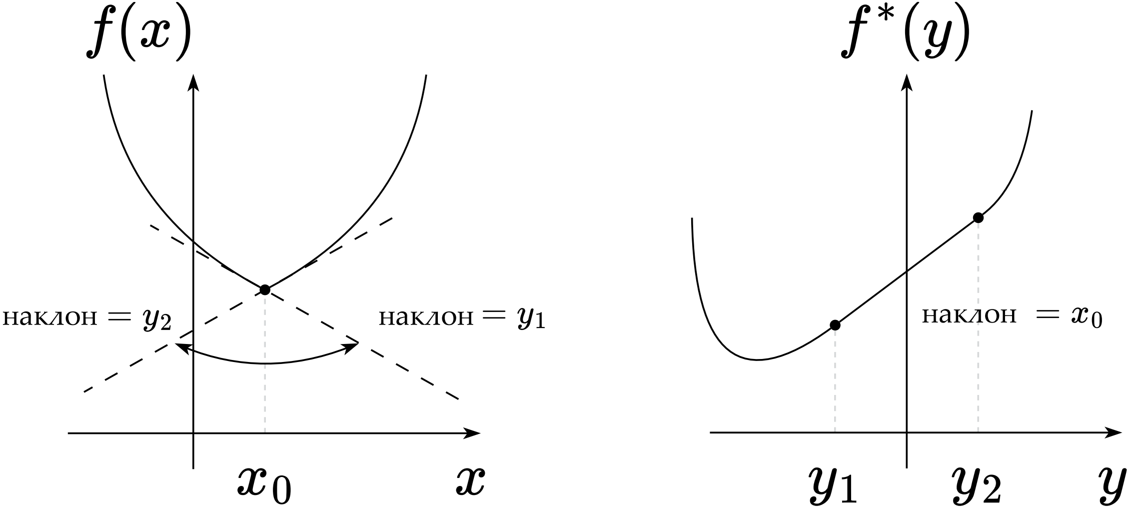
$$f^*(y) = \max_x \left[y^T x - f(x)\right]$$

называется его сопряжённой. Выражение выше называется преобразованием Лежандра.

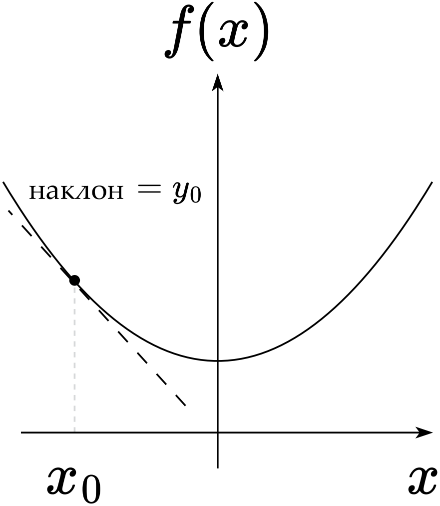
Геометрическая интуиция



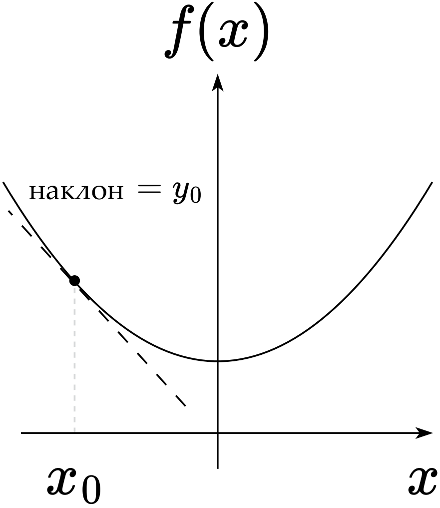
Геометрическая интуиция

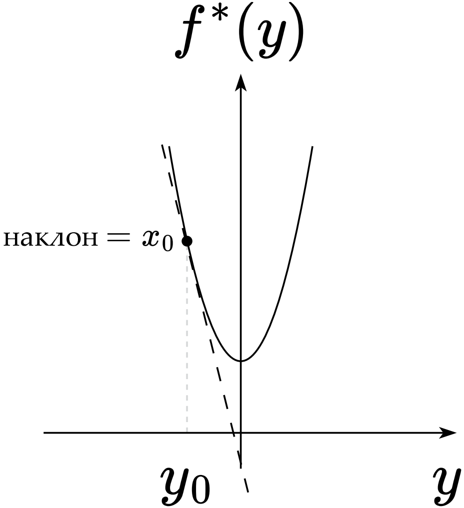


Наклон f и f^*



Наклон f и f^*





Наклон f и f^*

Предположим, что f — замкнутая и выпуклая функция. Тогда f сильно выпукла с параметром $\mu \Leftrightarrow \nabla f^*$ является липшицевой с параметром $1/\mu$.

Наклон f и f^*

Предположим, что f — замкнутая и выпуклая функция. Тогда f сильно выпукла с параметром $\mu \Leftrightarrow \nabla f^*$ является липшицевой с параметром $1/\mu$.

Доказательство "$\Rightarrow$": напомним, если g сильно выпукла с минимайзером x, то

$$g(y) \geq g(x) + \frac{\mu}{2}\|y - x\|^2, \quad \text{для всех } y$$

Наклон f и f^*

Предположим, что f — замкнутая и выпуклая функция. Тогда f сильно выпукла с параметром $\mu \Leftrightarrow \nabla f^*$ является липшицевой с параметром $1/\mu$.

Доказательство "$\Rightarrow$": напомним, если g сильно выпукла с минимайзером x, то

$$
g(y) \geq g(x) + \frac{\mu}{2}\|y - x\|^2, \quad \text{для всех } y
$$

Введём обозначения $x_u = \nabla f^*(u)$ и $x_v = \nabla f^*(v)$, тогда

$$
\begin{aligned}
f(x_v) - u^T x_v &\geq f(x_u) - u^T x_u + \frac{\mu}{2}\|x_u - x_v\|^2 \\
f(x_u) - v^T x_u &\geq f(x_v) - v^T x_v + \frac{\mu}{2}\|x_u - x_v\|^2
\end{aligned}
$$

Наклон f и f^*

Предположим, что f — замкнутая и выпуклая функция. Тогда f сильно выпукла с параметром $\mu \Leftrightarrow \nabla f^*$ является липшицевой с параметром $1/\mu$.

Доказательство "$\Rightarrow$": напомним, если g сильно выпукла с минимайзером x, то

$$g(y) \geq g(x) + \frac{\mu}{2}\|y - x\|^2, \quad \text{для всех } y$$

Введём обозначения $x_u = \nabla f^*(u)$ и $x_v = \nabla f^*(v)$, тогда

$$f(x_v) - u^T x_v \geq f(x_u) - u^T x_u + \frac{\mu}{2}\|x_u - x_v\|^2$$
$$f(x_u) - v^T x_u \geq f(x_v) - v^T x_v + \frac{\mu}{2}\|x_u - x_v\|^2$$

Сложив эти неравенства, применяя неравенство Коши–Буняковского-Шварца и преобразуя, получаем:

$$\|x_u - x_v\|^2 \leq \frac{1}{\mu}\|u - v\|^2$$

Наклон f и f^*

Доказательство "$\Leftarrow$": для простоты обозначим $g = f^*$ и $L = \frac{1}{\mu}$. Так как ∇g является липшицевой с константой L, то и $g_x(z) = g(z) - \nabla g(x)^T z$ также липшицева, следовательно

$$
g_x(z) \leq g_x(y) + \nabla g_x(y)^T(z - y) + \frac{L}{2}\|z - y\|_2^2
$$

Наклон f и f^*

Доказательство "$\Leftarrow$": для простоты обозначим $g = f^*$ и $L = \frac{1}{\mu}$. Так как ∇g является липшицевой с константой L, то и $g_x(z) = g(z) - \nabla g(x)^T z$ также липшицева, следовательно

$$g_x(z) \leq g_x(y) + \nabla g_x(y)^T (z - y) + \frac{L}{2}\|z - y\|_2^2$$

Минимизируя обе части по z и преобразуя, получаем

$$\frac{1}{2L}\|\nabla g(x) - \nabla g(y)\|^2 \leq g(y) - g(x) + \nabla g(x)^T (x - y)$$

Наклон f и f^*

Доказательство "$\Leftarrow$": для простоты обозначим $g = f^*$ и $L = \frac{1}{\mu}$. Так как ∇g является липшицевой с константой L, то и $g_x(z) = g(z) - \nabla g(x)^T z$ также липшицева, следовательно

$$
g_x(z) \leq g_x(y) + \nabla g_x(y)^T (z - y) + \frac{L}{2} \|z - y\|_2^2
$$

Минимизируя обе части по z и преобразуя, получаем

$$
\frac{1}{2L} \|\nabla g(x) - \nabla g(y)\|^2 \leq g(y) - g(x) + \nabla g(x)^T (x - y)
$$

Меняя местами x, y и складывая, получаем

$$
\frac{1}{L} \|\nabla g(x) - \nabla g(y)\|^2 \leq (\nabla g(x) - \nabla g(y))^T (x - y)
$$

Наклон f и f^*

Доказательство "$\Leftarrow$": для простоты обозначим $g = f^*$ и $L = \frac{1}{\mu}$. Так как ∇g является липшицевой с константой L, то и $g_x(z) = g(z) - \nabla g(x)^T z$ также липшицева, следовательно

$$
g_x(z) \leq g_x(y) + \nabla g_x(y)^T (z - y) + \frac{L}{2} \|z - y\|_2^2
$$

Минимизируя обе части по z и преобразуя, получаем

$$
\frac{1}{2L} \|\nabla g(x) - \nabla g(y)\|^2 \leq g(y) - g(x) + \nabla g(x)^T (x - y)
$$

Меняя местами x, y и складывая, получаем

$$
\frac{1}{L} \|\nabla g(x) - \nabla g(y)\|^2 \leq (\nabla g(x) - \nabla g(y))^T (x - y)
$$

Положим $u = \nabla f(x)$, $v = \nabla g(y)$; тогда $x \in \partial g^*(u)$, $y \in \partial g^*(v)$, и выше получаем

$$
(x - y)^T (u - v) \geq \frac{\|u - v\|^2}{L},
$$

что и доказывает утверждение.

Свойства сопряжённых функций

Напомним, что для отображения $f : \mathbb{R}^n \to \mathbb{R}$ функция, определяемая как

$$f^*(y) = \max_x \left[y^T x - f(x)\right]$$

называется его сопряжённой.

- Сопряжённые функции часто возникают в двойственных задачах, так как

$$-f^*(y) = \min_x \left[f(x) - y^T x\right]$$

Свойства сопряжённых функций

Напомним, что для отображения $f : \mathbb{R}^n \to \mathbb{R}$ функция, определяемая как

$$f^*(y) = \max_x \left[y^T x - f(x)\right]$$

называется его сопряжённой.

- Сопряжённые функции часто возникают в двойственных задачах, так как

$$-f^*(y) = \min_x \left[f(x) - y^T x\right]$$

- Если f замкнута и выпукла, то $f^{**} = f$. Кроме того,

$$x \in \partial f^*(y) \Leftrightarrow y \in \partial f(x) \Leftrightarrow x \in \arg\min_z \left[f(z) - y^T z\right]$$

Свойства сопряжённых функций

Напомним, что для отображения $f : \mathbb{R}^n \to \mathbb{R}$ функция, определяемая как

$$f^*(y) = \max_x \left[y^T x - f(x)\right]$$

называется его сопряжённой.

- Сопряжённые функции часто возникают в двойственных задачах, так как

$$-f^*(y) = \min_x \left[f(x) - y^T x\right]$$

- Если f замкнута и выпукла, то $f^{**} = f$. Кроме того,

$$x \in \partial f^*(y) \Leftrightarrow y \in \partial f(x) \Leftrightarrow x \in \arg\min_z \left[f(z) - y^T z\right]$$

- Если f строго выпукла, то

$$\nabla f^*(y) = \arg\min_z \left[f(z) - y^T z\right]$$

Свойства сопряжённой функции (доказательства)

Покажем, что $x \in \partial f^*(y) \Leftrightarrow y \in \partial f(x)$, предполагая, что f выпукла и замкнута.

- **Доказательство** $\Leftarrow$: Пусть $y \in \partial f(x)$. Тогда $x \in M_y$ — множество точек максимума $y^T z - f(z)$ по z. Но

$$
f^*(y) = \max_z \{y^T z - f(z)\} \quad \text{and} \quad \partial f^*(y) = \text{cl}(\text{conv}(\bigcup_{z \in M_y} \{z\})).
$$

  Следовательно, $x \in \partial f^*(y)$.

Свойства сопряжённой функции (доказательства)

Покажем, что $x \in \partial f^*(y) \Leftrightarrow y \in \partial f(x)$, предполагая, что f выпукла и замкнута.

- **Доказательство** $\Leftarrow$: Пусть $y \in \partial f(x)$. Тогда $x \in M_y$ — множество точек максимума $y^T z - f(z)$ по z. Но

$$
f^*(y) = \max_z \{y^T z - f(z)\} \quad \text{and} \quad \partial f^*(y) = \text{cl}(\text{conv}(\bigcup_{z \in M_y} \{z\})).
$$

  Следовательно, $x \in \partial f^*(y)$.

- **Доказательство** $\Rightarrow$: Из показанного выше, если $x \in \partial f^*(y)$, то $y \in \partial f^*(x)$, но $f^{**} = f$.

Свойства сопряжённой функции (доказательства)

Покажем, что $x \in \partial f^*(y) \Leftrightarrow y \in \partial f(x)$, предполагая, что f выпукла и замкнута.

- **Доказательство** $\Leftarrow$: Пусть $y \in \partial f(x)$. Тогда $x \in M_y$ — множество точек максимума $y^T z - f(z)$ по z. Но

$$f^*(y) = \max_z \{y^T z - f(z)\} \quad \text{and} \quad \partial f^*(y) = \text{cl}(\text{conv}(\bigcup_{z \in M_y} \{z\})).$$

  Следовательно, $x \in \partial f^*(y)$.

- **Доказательство** $\Rightarrow$: Из показанного выше, если $x \in \partial f^*(y)$, то $y \in \partial f^*(x)$, но $f^{**} = f$.

Свойства сопряжённой функции (доказательства)

Покажем, что $x \in \partial f^*(y) \Leftrightarrow y \in \partial f(x)$, предполагая, что f выпукла и замкнута.

- **Доказательство** $\Leftarrow$: Пусть $y \in \partial f(x)$. Тогда $x \in M_y$ — множество точек максимума $y^T z - f(z)$ по z. Но

$$f^*(y) = \max_z \{y^T z - f(z)\} \quad \text{and} \quad \partial f^*(y) = \text{cl}(\text{conv}(\bigcup_{z \in M_y} \{z\})).$$

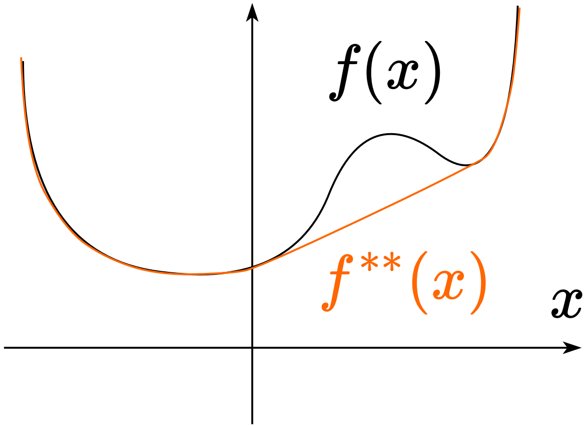
  Следовательно, $x \in \partial f^*(y)$.

- **Доказательство** $\Rightarrow$: Из показанного выше, если $x \in \partial f^*(y)$, то $y \in \partial f^*(x)$, но $f^{**} = f$.

Очевидно, $y \in \partial f(x) \Leftrightarrow x \in \arg\min_z \{f(z) - y^T z\}$

Наконец, если f строго выпукла, то мы знаем, что $f(z) - y^T z$ имеет единственный минимизатор по z, и им должна быть $\nabla f^*(y)$.

Прямое следствие нерваенства Фенхеля-Юнга $f(x) \geq f^{\star\star}(x)$.



Пример

Example

Найдите $f^*(y)$, если $f(x) = ax + b$.

1. По определению:

$$
f^*(y) = \sup_{x\in\mathbb{R}}[yx - f(x)] = \sup_{x\in\mathbb{R}} g(x,y) \quad \mathbf{dom}\ f^* = \{y \in \mathbb{R} : \sup_{x\in\mathbb{R}} g(x,y) \text{ is finite}\}
$$

Пример

Example

Найдите $f^*(y)$, если $f(x) = ax + b$.

1. По определению:

$$f^*(y) = \sup_{x \in \mathbb{R}} [yx - f(x)] = \sup_{x \in \mathbb{R}} g(x, y) \quad \mathbf{dom}\ f^* = \{y \in \mathbb{R} : \sup_{x \in \mathbb{R}} g(x, y) \text{ is finite}\}$$

2. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = yx - f(x) = yx - ax - b = x(y - a) - b.$$

Пример

Example

Найдите $f^*(y)$, если $f(x) = ax + b$.

1. По определению:

$$f^*(y) = \sup_{x \in \mathbb{R}} [yx - f(x)] = \sup_{x \in \mathbb{R}} g(x, y) \quad \mathbf{dom}\ f^* = \{y \in \mathbb{R} : \sup_{x \in \mathbb{R}} g(x, y) \text{ is finite}\}$$

2. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = yx - f(x) = yx - ax - b = x(y - a) - b.$$

3. Зададим область определения функции (т.е. те y, для которых $\sup$ конечен). Это единственная точка, $y = a$. В противном случае можно выбрать такое x, что супремум будет бесконечен.

Пример

Example

Найдите $f^*(y)$, если $f(x) = ax + b$.

1. По определению:

$$f^*(y) = \sup_{x \in \mathbb{R}} [yx - f(x)] = \sup_{x \in \mathbb{R}} g(x, y) \quad \mathbf{dom}\ f^* = \{y \in \mathbb{R} : \sup_{x \in \mathbb{R}} g(x, y) \text{ is finite}\}$$

2. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = yx - f(x) = yx - ax - b = x(y - a) - b.$$

3. Зададим область определения функции (т.е. те y, для которых $\sup$ конечен). Это единственная точка, $y = a$. В противном случае можно выбрать такое x, что супремум будет бесконечен.
4. Таким образом, имеем: $\mathbf{dom}\ f^* = \{a\}; f^*(a) = -b$

Пример

Question

Найдите $f^*(y)$, если $f(x) = \dfrac{1}{x}, \;\; x \in \mathbb{R}_{++}$.

Пример

Example

Найдите $f^*(y)$, если $f(x) = -\log x, \quad x \in \mathbb{R}_{++}$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = yx + \log x.$$

Пример

Example

Найдите $f^*(y)$, если $f(x) = -\log x, \quad x \in \mathbb{R}_{++}$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = yx + \log x.$$

2. Эта функция неограничена сверху при $y \geq 0$. Следовательно, область определения f^* является $\mathbf{dom}\ f^* = \{y < 0\}$.

Пример

Example

Найдите $f^*(y)$, если $f(x) = -\log x, \quad x \in \mathbb{R}_{++}$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = yx + \log x.$$

2. Эта функция неограничена сверху при $y \geq 0$. Следовательно, область определения f^* является $\mathbf{dom}\ f^* = \{y < 0\}$.
3. Эта функция выпукла и её максимум достигается в точке с нулевым градиентом:

$$\frac{\partial}{\partial x}(yx + \log x) = \frac{1}{x} + y = 0.$$

Таким образом, имеем $x = -\dfrac{1}{y}$ и сопряжённая функция:

$$f^*(y) = -\log(-y) - 1.$$

Пример

Example

Найдите $f^*(y)$, если $f(x) = e^x$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = yx - e^x.$$

Пример

Example

Найдите $f^*(y)$, если $f(x) = e^x$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = yx - e^x.$$

2. Эта функция неограничена сверху при $y < 0$. Следовательно, область определения f^* является $\mathbf{dom}\ f^* = \{y \geq 0\}$.

Пример

Example

Найдите $f^*(y)$, если $f(x) = e^x$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x\rangle - f(x) = yx - e^x.$$

2. Эта функция неограничена сверху при $y < 0$. Следовательно, область определения f^* является $\mathbf{dom}\ f^* = \{y \geq 0\}$.
3. Максимум этой функции достигается при $x = \log y$. Следовательно:

$$f^*(y) = y \log y - y,$$

полагая $0 \log 0 = 0$.

Пример

Example

Найдите $f^*(y)$, если $f(x) = x \log x, x \neq 0$, и $f(0) = 0, \quad x \in \mathbb{R}_+$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = xy - x \log x.$$

Пример

ℹ Example

Найдите $f^*(y)$, если $f(x) = x\log x, x \neq 0$, и $f(0) = 0, \quad x \in \mathbb{R}_+$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$
g(x, y) = \langle y, x \rangle - f(x) = xy - x\log x.
$$

2. Эта функция ограничена сверху для всех y. Следовательно, $\mathbf{dom}\ f^* = \mathbb{R}$.

Пример

Example

Найдите $f^*(y)$, если $f(x) = x \log x, x \neq 0$, и $f(0) = 0, \quad x \in \mathbb{R}_+$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = xy - x \log x.$$

2. Эта функция ограничена сверху для всех y. Следовательно, $\mathbf{dom}\ f^* = \mathbb{R}$.
3. Максимум этой функции достигается при $x = e^{y-1}$. Следовательно:

$$f^*(y) = e^{y-1}.$$

Пример

Example

Найдите $f^*(y)$, если $f(x) = \frac{1}{2}x^T A x, \quad A \in \mathbb{S}^n_{++}$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = y^T x - \frac{1}{2}x^T A x.$$

Пример

Example

Найдите $f^*(y)$, если $f(x) = \frac{1}{2}x^TAx, \quad A \in \mathbb{S}^n_{++}$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x,y) = \langle y, x\rangle - f(x) = y^Tx - \frac{1}{2}x^TAx.$$

2. Эта функция ограничена сверху для всех y. Следовательно, $\textbf{dom}\ f^* = \mathbb{R}$.

Пример

Example

Найдите $f^*(y)$, если $f(x) = \frac{1}{2}x^TAx, \quad A \in \mathbb{S}^n_{++}$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x,y) = \langle y, x \rangle - f(x) = y^Tx - \frac{1}{2}x^TAx.$$

2. Эта функция ограничена сверху для всех y. Следовательно, $\mathbf{dom}\ f^* = \mathbb{R}$.
3. Максимум этой функции достигается при $x = A^{-1}y$. Следовательно:

$$f^*(y) = \frac{1}{2}y^TA^{-1}y.$$

Пример

Example

Найдите $f^*(y)$, если $f(x) = \max\limits_i x_i, \quad x \in \mathbb{R}^n$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = y^T x - \max_i x_i.$$

Пример

Example

Найдите $f^*(y)$, если $f(x) = \max_i x_i, \quad x \in \mathbb{R}^n$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = y^T x - \max_i x_i.$$

2. Обратим внимание, что если вектор y имеет хотя бы одну отрицательную компоненту, эта функция не ограничена по x.

Пример

Example

Найдите $f^*(y)$, если $f(x) = \max_i x_i, \quad x \in \mathbb{R}^n$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:
$$g(x, y) = \langle y, x \rangle - f(x) = y^T x - \max_i x_i.$$
2. Обратим внимание, что если вектор y имеет хотя бы одну отрицательную компоненту, эта функция не ограничена по x.
3. Если $y \succeq 0$ и $1^T y > 1$, то $y \notin \mathbf{dom}\ f^*(y)$.

Пример

ℹ Example

Найдите $f^*(y)$, если $f(x) = \max_i x_i, \quad x \in \mathbb{R}^n$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = y^T x - \max_i x_i.$$

2. Обратим внимание, что если вектор y имеет хотя бы одну отрицательную компоненту, эта функция не ограничена по x.
3. Если $y \succeq 0$ и $1^T y > 1$, то $y \notin \textbf{dom } f^*(y)$.
4. Если $y \succeq 0$ и $1^T y < 1$, то $y \notin \textbf{dom } f^*(y)$.

Пример

Example

Найдите $f^*(y)$, если $f(x) = \max_i x_i, \quad x \in \mathbb{R}^n$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:
$$g(x, y) = \langle y, x \rangle - f(x) = y^T x - \max_i x_i.$$
2. Обратим внимание, что если вектор y имеет хотя бы одну отрицательную компоненту, эта функция не ограничена по x.
3. Если $y \succeq 0$ и $1^T y > 1$, то $y \notin \mathbf{dom}\ f^*(y)$.
4. Если $y \succeq 0$ и $1^T y < 1$, то $y \notin \mathbf{dom}\ f^*(y)$.
5. Остается только случай $y \succeq 0$ и $1^T y = 1$. В этом случае, $x^T y \leq \max_i x_i$.

Пример

Example

Найдите $f^*(y)$, если $f(x) = \max_i x_i, \quad x \in \mathbb{R}^n$.

1. Рассмотрим функцию, поточечным супремумом которой является сопряжённая:

$$g(x, y) = \langle y, x \rangle - f(x) = y^T x - \max_i x_i.$$

2. Обратим внимание, что если вектор y имеет хотя бы одну отрицательную компоненту, эта функция не ограничена по x.
3. Если $y \succeq 0$ и $1^T y > 1$, то $y \notin \mathbf{dom}\ f^*(y)$.
4. Если $y \succeq 0$ и $1^T y < 1$, то $y \notin \mathbf{dom}\ f^*(y)$.
5. Остается только случай $y \succeq 0$ и $1^T y = 1$. В этом случае, $x^T y \leq \max_i x_i$.
6. Следовательно, $f^*(y) = 0$.