





Пусть $S \subseteq \mathbb{R}^n$ — произвольное непустое множество.

Тогда его сопряжённое множество определяется как:

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

Пусть $S \subseteq \mathbb{R}^n$ — произвольное непустое множество.

Тогда его сопряжённое множество определяется как:

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

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

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

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 $S_1^* = S_2, \ S_2^* = S_1.$

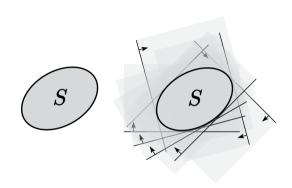


Рис. 1: Выпуклые множества можно описывать двойственным образом — через элементы множества и через множество опорных гиперплоскостей

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 $S^* = S$.

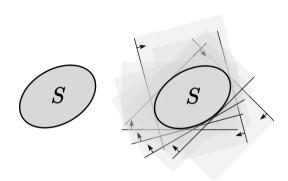


Рис. 1: Выпуклые множества можно описывать двойственным образом — через элементы множества и через множество опорных гиперплоскостей

• Сопряжённое множество всегда замкнуто, выпукло и содержит нуль.



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

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



- 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 $S_2^* \subseteq S_1^*$.



- 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 $S_2^* \subseteq S_1^*$.
- $\bullet \left(\bigcup_{i=1}^m S_i\right)^* = \bigcap_{i=1}^m S_i^*.$

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



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



i Example

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

i Example

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

•
$$S \subset \overline{S} \implies (\overline{S})^* \subset S^*$$
.

i Example

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

- $S \subset \overline{S} \implies (\overline{S})^* \subset S^*$.
- ullet Пусть $p\in S^*$ и $x_0\in \overline{S},\; x_0=\lim_{k o\infty}x_k.$ Тогда в силу непрерывности функции $f(x)=p^Tx$ имеем:

$$p^Tx_k \geq -1 \ \Rightarrow \ p^Tx_0 \geq -1.$$
 Следовательно, $p \in \left(\overline{S}
ight)^*$, и значит $S^* \subset \left(\overline{S}
ight)^*$.

i Example

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

i Example

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

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

i Example

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

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

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

i Example

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

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

$$p^Tx_0 = \sum_{i=1}^k \theta_i \, p^Tx_i \ \geq \ \sum_{i=1}^k \theta_i \cdot (-1) = 1 \cdot (-1) = -1.$$

i Example

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

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

$$p^Tx_0 = \sum_{i=1}^k \theta_i \, p^Tx_i \ \geq \ \sum_{i=1}^k \theta_i \cdot (-1) = 1 \cdot (-1) = -1.$$

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



i Example

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

i Example

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

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

i Example

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

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

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

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

Следовательно, $X^* \subset Y$.

2 Сопряженные множества

i Example

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

вектор $p \in X^*$, тогда для любого $x \in X$: $p^T x \ge -1$.

• Пусть B(0,r) = X, B(0,1/r) = Y. Возьмём

• Среди всех точек шара X возьмём такую $x \in X$, для которой скалярное произведение с p минимально: p^Tx . Это точка $x = -\frac{p}{\|p\|}r$.

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

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

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^{st} такое, что:

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

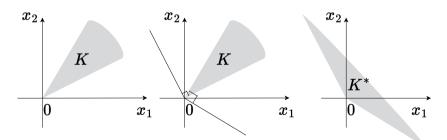
Двойственный конус

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 $\forall \lambda>0.$

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



⊕ 0 @

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



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

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



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

$$\left(S+K\right)^{*}=S^{*}\cap K^{*}$$

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

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

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

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

• Пусть $K_1, ..., K_m$ — 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:

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

i Example

Найдите сопряжённый конус для монотонного неотрицательного конуса:

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



i 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:

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



i 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:

$$y_1 \ge 0$$
, $y_1 + y_2 \ge 0$, ..., $y_1 + \dots + y_n \ge 0$

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:

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

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:

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

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

i Theorem

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

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

является многогранник:

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

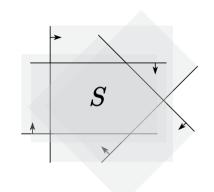


Рис. 3: Polyhedra

Сопряженные множества

Доказательство

ullet Пусть $S=X,\ S^*=Y.$ Возьмём некоторое $p\in X^*$, тогда $\langle p,x_i
angle \geq -1,\ i=\overline{1,k}.$ 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

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

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

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

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

ullet Предположим, наоборот, что $p\in Y$. Для любой точки $x\in X$: $x=\sum_{i=1}^m heta_i x_i, \quad \sum_{i=1}^\kappa heta_i=1, \quad heta_i\geq 0$ Тогда:

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

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

Пример



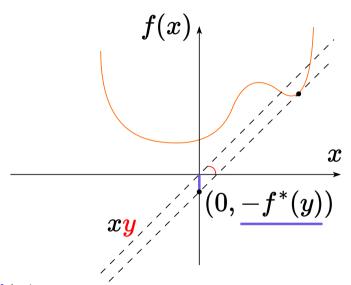


Сопряжённые функции





Сопряжённые функции

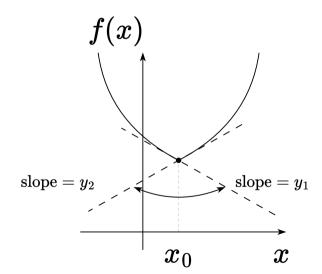


Напомним, что для отображения $f:\mathbb{R}^n \to \mathbb{R}$ функция, определяемая как

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

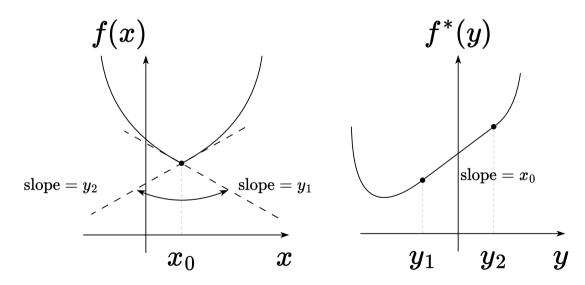
называется его сопряжённой. Выражение выше называется преобразованием Лежандра.

Геометрическая интуиция

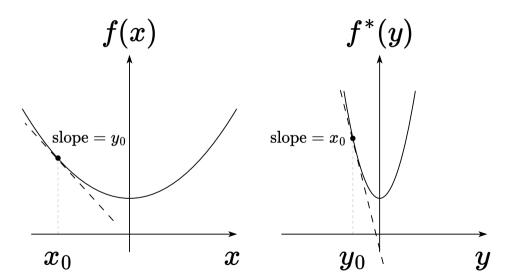




Геометрическая интуиция







 $\min_{v,y,z}$ Сопряжённые функции

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

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 q сильно выпукла с минимайзером x, то

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



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 q сильно выпукла с минимайзером x, то

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

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

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

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



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 q сильно выпукла с минимайзером x, то

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

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

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

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

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

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

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



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

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



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

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

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

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

что и доказывает утверждение.

Свойства сопряжённых функций

Напомним, что для отображения $f:\mathbb{R}^n o \mathbb{R}$ 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

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

ullet 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,

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

• Если f строго выпукла, то

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



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

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

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

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



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

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

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

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

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

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

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

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

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

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

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

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

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

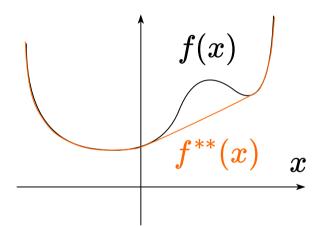
Следовательно. $x \in \partial f^*(y)$.

• 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 $f(z)-y^Tz$ имеет единственный минимизатор по z, и им должна быть $\nabla f^*(y)$.

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



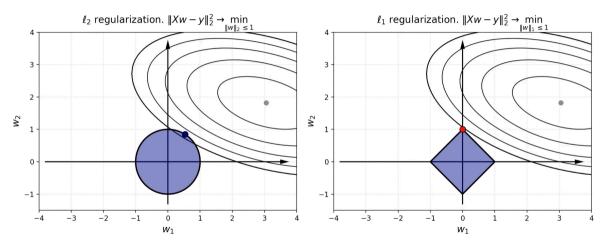
Негладкие задачи





Задача наименьших квадратов с ℓ_1 - регуляризацией

ℓ_1 induces sparsity



@fminxyz



Нормы не являются гладкими

$$\min_{x \in \mathbb{R}^n} f(x),$$

Рассмотрим классическую выпуклую задачу оптимизации. Мы предполагаем, что f(x) является выпуклой функцией, но теперь мы не требуем гладкости.

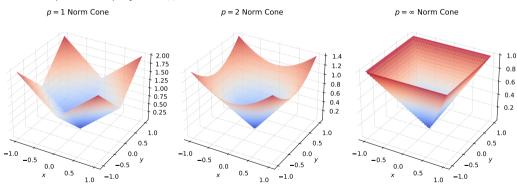


Рис. 5: Нормы конусов для разных p — нормы не являются гладкими



Пример Вульфа

Wolfe's example

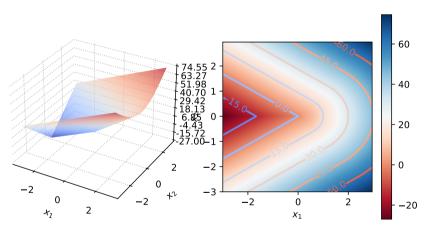


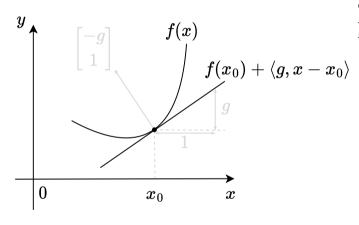
Рис. 6: Пример Вульфа. **�**Открыть в Colab



Вычисление субградиента







Важное свойство непрерывной выпуклой функции f(x) заключается в том, что для любой выбранной точки x_0 для всех $x\in \mathrm{dom}\, f$ выполняется неравенство:

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle$$

Рис. 7: Линейная аппроксимация Тейлора служит глобальной нижней оценкой для выпуклой функции

എറ ഉ

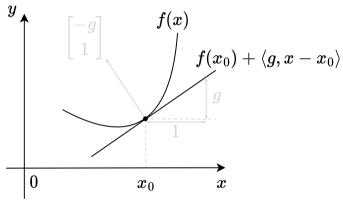


Рис. 7: Линейная аппроксимация Тейлора служит глобальной нижней

Важное свойство непрерывной выпуклой функции f(x) заключается в том, что для любой выбранной точки x_0 для всех $x \in \mathsf{dom}\ f$ выполняется неравенство:

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle$$

для некоторого вектора g, т.е. касательная к графику функции является глобальной нижней оценкой для функции.

• Если f(x) дифференцируема, то $q = \nabla f(x_0)$.

оценкой для выпуклой функции

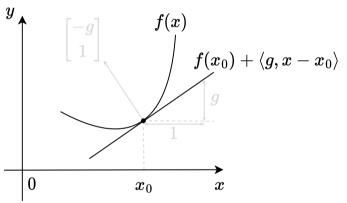


Рис. 7: Линейная аппроксимация Тейлора служит глобальной нижней оценкой для выпуклой функции

Важное свойство непрерывной выпуклой функции f(x) заключается в том, что для любой выбранной точки x_0 для всех $x \in \mathsf{dom}\ f$ выполняется неравенство:

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle$$

для некоторого вектора g, т.е. касательная к графику функции является глобальной нижней оценкой для функции.

- Если f(x) дифференцируема, то $q = \nabla f(x_0)$.
- Не все непрерывные выпуклые функции дифференцируемы.

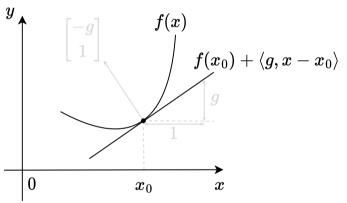


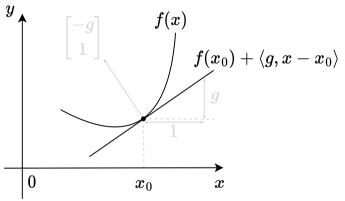
Рис. 7: Линейная аппроксимация Тейлора служит глобальной нижней оценкой для выпуклой функции

Важное свойство непрерывной выпуклой функции f(x) заключается в том, что для любой выбранной точки x_0 для всех $x \in \mathsf{dom}\ f$ выполняется неравенство:

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle$$

для некоторого вектора g, т.е. касательная к графику функции является глобальной нижней оценкой для функции.

- Если f(x) дифференцируема, то $q = \nabla f(x_0)$.
- Не все непрерывные выпуклые функции дифференцируемы.



Важное свойство непрерывной выпуклой функции f(x) заключается в том, что для любой выбранной точки x_0 для всех $x \in \text{dom } f$ выполняется неравенство:

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle$$

для некоторого вектора g, т.е. касательная к графику функции является *глобальной* нижней оценкой для функции.

- Если f(x) дифференцируема, то $g = \nabla f(x_0).$
- Не все непрерывные выпуклые функции дифференцируемы.

Мы не хотим потерять такое удобное свойство.

Рис. 7: Линейная аппроксимация Тейлора служит глобальной нижней оценкой для выпуклой функции

⊕ ი დ

Субградиент и субдифференциал

Вектор g называется **субградиентом** функции $f(x):S \to \mathbb{R}$ в точке x_0 , если $\forall x \in S$:

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle$$

Субградиент и субдифференциал

Вектор g называется **субградиентом** функции $f(x):S\to\mathbb{R}$ в точке x_0 , если $\forall x\in S$:

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle$$

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

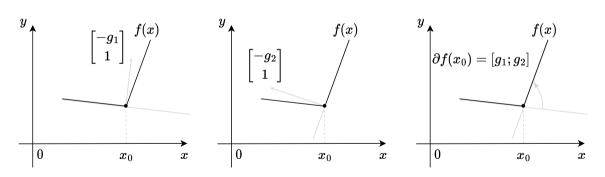


Субградиент и субдифференциал

Вектор g называется **субградиентом** функции $f(x):S\to\mathbb{R}$ в точке x_0 , если $\forall x\in S$:

$$f(x) \geq f(x_0) + \langle g, x - x_0 \rangle$$

Множество всех субградиентов функции f(x) в точке x_0 называется субдифференциалом функции f в точке x_0 и обозначается $\partial f(x_0)$.

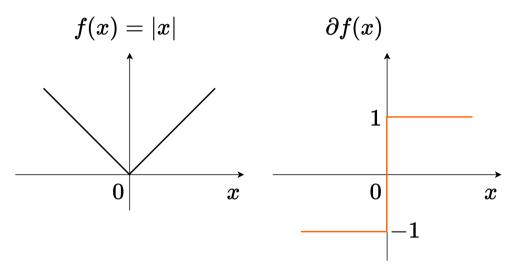


Субградиент и субдифференциал

Найдите $\partial f(x)$, если f(x) = |x|

Субградиент и субдифференциал

Найдите $\partial f(x)$, если f(x) = |x|



ullet Если $x_0\in {f ri}(S)$, то $\partial f(x_0)$ является выпуклым компактным множеством.



- ullet Если $x_0 \in \mathbf{ri}(S)$, то $\partial f(x_0)$ является выпуклым компактным множеством.
- Выпуклая функция f(x) дифференцируема в точке $x_0 \Rightarrow \partial f(x_0) = \{\nabla f(x_0)\}.$



- ullet Если $x_0 \in \mathbf{ri}(S)$, то $\partial f(x_0)$ является выпуклым компактным множеством.
- Выпуклая функция f(x) дифференцируема в точке $x_0\Rightarrow \partial f(x_0)=\{\nabla f(x_0)\}.$
- ullet Если $\partial f(x_0)
 eq \emptyset \quad orall x_0 \in S$, то f(x) выпукла на S.



- ullet Если $x_0 \in \mathbf{ri}(S)$, то $\partial f(x_0)$ является выпуклым компактным множеством.
- Выпуклая функция f(x) дифференцируема в точке $x_0\Rightarrow \partial f(x_0)=\{\nabla f(x_0)\}.$
- ullet Если $\partial f(x_0)
 eq \emptyset \quad orall x_0 \in S$, то f(x) выпукла на S.



- Если $x_0 \in \mathbf{ri}(S)$, то $\partial f(x_0)$ является выпуклым компактным множеством.
- ullet Выпуклая функция f(x) дифференцируема в точке $x_0\Rightarrow \partial f(x_0)=\{\nabla f(x_0)\}.$
- ullet Если $\partial f(x_0)
 eq \emptyset \quad orall x_0 \in S$, то f(x) выпукла на S.

1 Субдифференциал дифференцируемой функции

Пусть $f:S o\mathbb{R}$ — функция, определенная

на множестве S в евклидовом пространстве \mathbb{R}^n . Если $x_0 \in \mathbf{ri}(S)$ и f дифференцируема в точке x_0 , то либо $\partial f(x_0) = \emptyset$ либо $\partial f(x_0) = \{\nabla f(x_0)\}$. Более того, если функция f выпукла, то первая ситуация невозможна.



- Если $x_0 \in \mathbf{ri}(S)$, то $\partial f(x_0)$ является выпуклым компактным множеством
- Выпуклая функция f(x) дифференцируема в точке $x_0 \Rightarrow \partial f(x_0) = \{\nabla f(x_0)\}.$
- Если $\partial f(x_0) \neq \emptyset$ $\forall x_0 \in S$, то f(x) выпукла на S.

і Субдифференциал дифференцируемой функции

Пусть $f:S o\mathbb{R}$ — функция, определенная на множестве S в евклидовом пространстве \mathbb{R}^n . Если $x_0 \in \mathbf{ri}(S)$ и f дифференцируема в точке x_0 , то либо $\partial f(x_0) = \emptyset$ либо $\partial f(x_0) = {\nabla f(x_0)}.$ Более того, если функция f выпукла, то первая ситуация невозможна.

Доказательство

1. Пусть $s \in \partial f(x_0)$ для некоторого $s \in \mathbb{R}^n$ отличного от $\nabla f(x_0)$. Пусть $v \in \mathbb{R}^n$ — единичный вектор. Поскольку x_0 является внутренней точкой множества S, существует $\delta > 0$ такое, что $x_0 + tv \in S$ для всех $0 < t < \delta$. По определению субградиента:

$$f(x_0+tv) \geq f(x_0) + t\langle s,v \rangle$$

Вычисление субградиента



- Если $x_0 \in \mathbf{ri}(S)$, то $\partial f(x_0)$ является выпуклым компактным множеством
- Выпуклая функция f(x) дифференцируема в точке $x_0 \Rightarrow \partial f(x_0) = \{\nabla f(x_0)\}.$
- Если $\partial f(x_0) \neq \emptyset$ $\forall x_0 \in S$, то f(x) выпукла на S.

і Субдифференциал дифференцируемой функции

Пусть $f:S o\mathbb{R}$ — функция, определенная на множестве S в евклидовом пространстве \mathbb{R}^n . Если $x_0 \in \mathbf{ri}(S)$ и f дифференцируема в точке x_0 , то либо $\partial f(x_0) = \emptyset$ либо $\partial f(x_0) = {\nabla f(x_0)}.$ Более того, если функция f выпукла, то первая ситуация невозможна.

Доказательство

1. Пусть $s \in \partial f(x_0)$ для некоторого $s \in \mathbb{R}^n$ отличного от $\nabla f(x_0)$. Пусть $v \in \mathbb{R}^n$ — единичный вектор. Поскольку x_0 является внутренней точкой множества S, существует $\delta > 0$ такое, что $x_0 + tv \in S$ для всех $0 < t < \delta$. По определению субградиента:

$$f(x_0+tv) \geq f(x_0) + t\langle s,v \rangle$$

Вычисление субградиента



• Если $x_0 \in \mathbf{ri}(S)$, то $\partial f(x_0)$ является выпуклым компактным множеством

• Выпуклая функция f(x) дифференцируема в точке $x_0 \Rightarrow \partial f(x_0) = \{\nabla f(x_0)\}.$

• Если $\partial f(x_0) \neq \emptyset$ $\forall x_0 \in S$, то f(x) выпукла на S. і Субдифференциал дифференцируемой функции

 Π усть $f:S o \mathbb{R}$ — функция, определенная на множестве S в евклидовом пространстве \mathbb{R}^n . Если $x_0 \in \mathbf{ri}(S)$ и f дифференцируема в точке x_0 , то либо $\partial f(x_0) = \emptyset$ либо $\partial f(x_0) = {\nabla f(x_0)}.$

Более того, если функция f выпукла, то первая ситуация невозможна.

Доказательство

1. Пусть $s \in \partial f(x_0)$ для некоторого $s \in \mathbb{R}^n$ отличного от $\nabla f(x_0)$. Пусть $v \in \mathbb{R}^n$ — единичный вектор. Поскольку x_0 является внутренней точкой множества S, существует $\delta > 0$ такое, что $x_0 + tv \in S$ для всех $0 < t < \delta$. По определению субградиента:

 $f(x_0 + tv) \ge f(x_0) + t\langle s, v \rangle$

что влечёт:

$$\frac{f(x_0+tv)-f(x_0)}{t} \ge \langle s,v \rangle$$

для всех $0 < t < \delta$. Переходя к пределу при $t \to 0$ и используя определение градиента, получаем:

$$\langle \nabla f(x_0), v \rangle = \lim_{t \to 0; 0 < t < \delta} \frac{f(x_0 + tv) - f(x_0)}{t} \geq \langle s, v \rangle$$

2. Отсюда $\langle s - \nabla f(x_0), v \rangle \geq 0$. В силу произвольности v можно выбрать

$$v = -\frac{s - \nabla f(x_0)}{\|s - \nabla f(x_0)\|},$$

которое приводит к $s = \nabla f(x_0)$.



2. Отсюда $\langle s - \nabla f(x_0), v \rangle \geq 0$. В силу произвольности v можно выбрать

$$v = -\frac{s - \nabla f(x_0)}{\|s - \nabla f(x_0)\|},$$

которое приводит к $s = \nabla f(x_0)$.

3. Более того, если функция f выпукла, то согласно дифференциальному условию выпуклости $f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle$ для всех $x \in S$. Но по определению это означает, что $\nabla f(x_0) \in \partial f(x_0)$.

Моро Теорема Роккафеллара (субдифференциал линейной комбинации)

Пусть $f_i(x)$ — выпуклые функции на выпуклых множествах $S_i,\; i=\overline{1,n}.$ Тогда если $\bigcap_{i=1}^n \mathbf{ri}(S_i)
eq$

$$\emptyset$$
, то функция $f(x)=\sum\limits_{i=1}^n a_if_i(x),\ a_i>0$ имеет субдифференциал $\partial_S f(x)$ на множестве

$$S=\bigcap\limits_{i=1}^{n}S_{i}$$
 и

$$\partial_S f(x) = \sum_{i=1}^n a_i \partial_{S_i} f_i(x).$$

 $f \to \min_{x,y,z}$ Вычисление субградиента

Моро 1 Теорема Роккафеллара (субдифференциал линейной комбинации)

Пусть $f_i(x)$ — выпуклые функции на выпуклых множествах $S_i,\; i=\overline{1,n}.$ Тогда если $\bigcap_{i=1}^n \mathbf{ri}(S_i)
eq$

 \emptyset , то функция $f(x) = \sum\limits_{i=1}^n a_i f_i(x), \quad a_i > 0$ имеет субдифференциал $\partial_S f(x)$ на множестве $S = \bigcap_{i=1}^n S_i$ и

$$\partial_S f(x) = \sum_{i=1}^n a_i \partial_{S_i} f_i(x).$$

і Теорема Дубовицкого Милютина (субдифференциал поточечного максимума)

Пусть $f_i(x)$ — выпуклые функции на открытом выпуклом множестве $S\subseteq\mathbb{R}^n,\ x_0\in S$, и поточечный максимум определяется как f(x) = $\max f_i(x)$. Тогда:

$$\partial_S f(x_0) = \mathbf{conv} \left\{ \bigcup_{i \in I(x_0)} \partial_S f_i(x_0) \right\},$$

$$I(x) = \{i \in [1:m] : f_i(x) = f(x)\}$$

 $f \to \min_{x,y,z}$ Вычисление субградиента

•
$$\partial(\alpha f)(x) = \alpha \partial f(x)$$
, для $\alpha \geq 0$





•
$$\partial(\alpha f)(x) = \alpha \partial f(x)$$
, для $\alpha \geq 0$

•
$$\partial(\sum f_i)(x) = \sum \partial f_i(x)$$
, f_i — выпуклые функции



- $\partial(\alpha f)(x) = \alpha \partial f(x)$, для $\alpha \geq 0$
- $\partial(\sum f_i)(x) = \sum \partial f_i(x), f_i$ выпуклые функции
- $\partial (\overline{f}(Ax+b))(\overline{x}) = A^T \partial f(Ax+b)$, f выпуклая функция



- $\partial(\alpha f)(x) = \alpha \partial f(x)$, для $\alpha \geq 0$
- $\partial(\sum f_i)(x) = \sum \partial f_i(x)$, f_i выпуклые функции
- $\partial(\overline{f}(Ax+b))(\overline{x})=A^T\partial f(Ax+b)$, f выпуклая функция
- $z \in \partial f(x)$ тогда и только тогда, когда $x \in \partial f^*(z)$.



Связь с выпуклой геометрией

Для выпуклого множества $S \subseteq \mathbb{R}^n$, рассмотрим индикаторную функцию $I_S: \mathbb{R}^n \to \mathbb{R}$,

$$I_S(x) = I\{x \in S\} = \begin{cases} 0 & \text{if } x \in S \\ \infty & \text{if } x \notin S \end{cases}$$

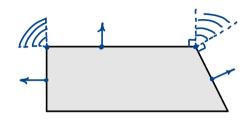
Для $x \in S$, $\partial I_S(x) = \mathcal{N}_S(x)$, нормальный конус для S в x:

$$\mathcal{N}_S(x) = \{g \in \mathbb{R}^n : g^Tx \geq g^Ty \text{ for any } y \in S\}$$

Почему? По определнию субградиента q,

$$I_S(y) \geq I_S(x) + g^T(y-x) \quad \text{for all } y$$

• При $y \notin S$. $I_S(y) = \infty$



Вычисление субградиента

Связь с выпуклой геометрией

Для выпуклого множества $S\subseteq\mathbb{R}^n$, рассмотрим индикаторную функцию $I_S:\mathbb{R}^n\to\mathbb{R}$,

$$I_S(x) = I\{x \in S\} = \begin{cases} 0 & \text{if } x \in S \\ \infty & \text{if } x \notin S \end{cases}$$

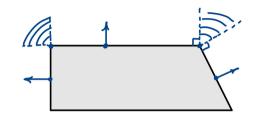
Для $x \in S$, $\partial I_S(x) = \mathcal{N}_S(x)$, нормальный конус для S в x:

$$\mathcal{N}_S(x) = \{g \in \mathbb{R}^n : g^Tx \geq g^Ty \text{ for any } y \in S\}$$

Почему? По определнию субградиента g,

$$I_S(y) \geq I_S(x) + g^T(y-x) \quad \text{for all } y$$

- ullet При y
 otin S, $I_S(y) = \infty$
- ullet При $y \in S$, this means $0 \geq g^T(y-x)$



എറമ

Условия Оптимальности

Для любой f (выпуклой или нет).

$$f(x^\star) = \min_{x} f(x) \quad \Longleftrightarrow \quad 0 \in \partial f(x^\star)$$

To есть, x^{\star} является точкой минимума тогда и только тогда, когда 0 является субградиентом функции f в точке x^* . Это утверждение называется субградиентное условие оптимальности

Почему? Легко: если q=0 является субградиентом, это значит что для всех y

$$f(y) \geq f(x^\star) + 0^T (y - x^\star) = f(x^\star)$$

Отметим, что для выпуклой и дифференцируемой функций f верно

$$\partial f(x) = \{\nabla f(x)\}$$



Получение условия оптимальности первого порядка

Пример мощи субградиентов: мы можем использовать всё, что узнали ранее, для получения условия оптимальности первого порядка. Вспомним,

решением задачи

$$\min_{x} f(x)$$
 subject to $x \in S$

является точка x, для выпуклой и дифференцируемой f, в том и только в том случае, если

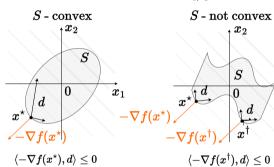
$$\nabla f(x)^T(y-x) \geq 0 \quad \text{for all } y \in S$$

Интуитивно: написанное выше означает, что градиент увеличивается по мере движения от точки x. Как это доказать? Во-первых, перепишем задачу в следующем виде:

$$\min_{x} f(x) + I_S(x)$$

Теперь воспользуемся условием оптимальности в субградиентной форме:

$$f(x)=x_1+x_2 o \min_{x_1,x_2\in \mathbb{R}^2}$$



 x^* - optimal

 x^{\dagger} - not optimal

Получение условия оптимальности первого порядка

Observe

$$0\in\partial(f(x)+I_S(x))$$

$$\Leftrightarrow 0\in\{\nabla f(x)\}+\mathcal{N}_S(x)$$

$$\Leftrightarrow -\nabla f(x)\in\mathcal{N}_S(x)$$

$$\Leftrightarrow -\nabla f(x)^Tx\geq -\nabla f(x)^Ty \text{ for all }y\in S$$

$$\Leftrightarrow \nabla f(x)^T(y-x)\geq 0 \text{ for all }y\in S$$
 что и требовалось.
Замечание: условие $0\in\partial f(x)+\mathcal{N}_S(x)$ является **общим условием** оптимальности для выпуклых задач.
Однако с ним не всегда удобно работать (ККТ удобнее.

 $f(x)=x_1+x_2 o \min_{x_1,x_2\in \mathbb{R}^2}$ S - convex S - not convex $_{^{\uparrow}}x_{2}$ 0 $\langle -\nabla f(x^{\star}), d \rangle \leq 0$ $\langle abla f(x^\dagger), d
angle \leq 0$ x^{\dagger} - not optimal x^* - optimal

про них позже).

Вычисление субградиента

i Example

Найти $\partial f(x)$, if f(x) = |x-1| + |x+1|

i Example

Найти $\partial f(x)$, if f(x) = |x-1| + |x+1|

$$\partial f_1(x) = \begin{cases} -1, & x < 1 \\ [-1;1], & x = 1 \\ 1, & x > 1 \end{cases} \qquad \partial f_2(x) = \begin{cases} -1, & x < -1 \\ [-1;1], & x = -1 \\ 1, & x > -1 \end{cases}$$

So

$$\partial f(x) = \begin{cases} -2, & x < -1 \\ [-2;0], & x = -1 \\ 0, & -1 < x < 1 \\ [0;2], & x = 1 \\ 2, & x > 1 \end{cases}$$

Найти $\partial f(x)$ if $f(x) = \left[\max(0, f_0(x))\right]^q$. Здесь $f_0(x)$ - выпуклая функция на открытом множестве S, и $q \geq 1$.

 $f \to \min_{x,y,z}$ Вычисление субградиента

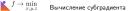
Найти $\partial f(x)$ if $f(x) = \left[\max(0, f_0(x))\right]^q$. Здесь $f_0(x)$ - выпуклая функция на открытом множестве S, и $q \ge 1$.

Согласно теореме о производной композиции функций (функция $\varphi(x) = x^q$ дифференцируема) и обозначая $g(x) = \max(0, f_0(x))$, имеем:

$$\partial f(x) = q(g(x))^{q-1} \partial g(x)$$

По теореме о субдифференциале поточечного максимума

$$\partial g(x) = \begin{cases} \partial f_0(x), & f_0(x) > 0, \\ \{0\}, & f_0(x) < 0, \\ \{a \mid a = \lambda a', \ 0 \leq \lambda \leq 1, \ a' \in \partial f_0(x)\}, & f_0(x) = 0 \end{cases}$$



Пусть V - конечномерное евклидово пространство, и $x_0 \in V$. Пусть $\|\cdot\|$ - произвольная норма в пространстве V (не обязательно заданная при помощи скалярного произведения), и пусть $\lVert \cdot
Vert_*$ - соответствующая сопряженная норма. Тогда,

Let V be a finite-dimensional Euclidean space, and $x_0 \in V$. Let $\|\cdot\|$ be an arbitrary norm in V (not necessarily induced by the scalar product), and let $\|\cdot\|_*$ be the corresponding conjugate norm. Then,

$$\partial \|\cdot\|(x_0) = \begin{cases} B_{\|\cdot\|_*}(0,1), & \text{if } x_0 = 0, \\ \{s \in V: \|s\|_* \leq 1; \langle s, x_0 \rangle = \|x_0\|\} = \{s \in V: \|s\|_* = 1; \langle s, x_0 \rangle = \|x_0\|\}, & \text{otherwise}. \end{cases}$$

Где $B_{\|.\|.}(0,1)$ есть замкнутый единичный относительно сопряженной нормы шар с центром в нуле. Другими словами, вектор $s\in V$ с $\|s\|_*=1$ является субградиентом нормы $\|\cdot\|$ в точке $x_0\neq 0$ тогда и только тогда, когда неравенство Гёльдера $\langle s, x_0 \rangle < \|x_0\|$ переходит в равенство.



Пусть V - конечномерное евклидово пространство, и $x_0 \in V$. Пусть $\|\cdot\|$ - произвольная норма в пространстве V (не обязательно заданная при помощи скалярного произведения), и пусть $\|\cdot\|_*$ - соответствующая сопряженная норма. Тогда,

Let V be a finite-dimensional Euclidean space, and $x_0 \in V$. Let $\|\cdot\|$ be an arbitrary norm in V (not necessarily induced by the scalar product), and let $\|\cdot\|_*$ be the corresponding conjugate norm. Then,

$$\partial \|\cdot\|(x_0) = \begin{cases} B_{\|\cdot\|_*}(0,1), & \text{if } x_0 = 0, \\ \{s \in V: \|s\|_* \leq 1; \langle s, x_0 \rangle = \|x_0\|\} = \{s \in V: \|s\|_* = 1; \langle s, x_0 \rangle = \|x_0\|\}, & \text{otherwise}. \end{cases}$$

Где $B_{\|\cdot\|_*}(0,1)$ есть замкнутый единичный относительно сопряженной нормы шар с центром в нуле. Другими словами, вектор $s\in V$ с $\|s\|_*=1$ является субградиентом нормы $\|\cdot\|$ в точке $x_0\neq 0$ тогда и только тогда, когда неравенство Гёльдера $\langle s,x_0\rangle\leq \|x_0\|$ переходит в равенство.

Пусть $s \in V$. По определению $s \in \partial \lVert \cdot \rVert(x_0)$ если и только если

$$\langle s,x\rangle - \|x\| \leq \langle s,x_0\rangle - \|x_0\|, \text{ for all } x \in V,$$

что равносильно

$$\int_{s,y,z} \sup_{b \mapsto \mathrm{Mix}} \{\langle s,x \rangle - \|x\|\} \leq \langle s,x_0 \rangle - \|x_0\|.$$

Пусть V - конечномерное евклидово пространство, и $x_0 \in V$. Пусть $\|\cdot\|$ - произвольная норма в пространстве V (не обязательно заданная при помощи скалярного произведения), и пусть $\|\cdot\|_*$ - соответствующая сопряженная норма. Тогда,

Let V be a finite-dimensional Euclidean space, and $x_0 \in V$. Let $\|\cdot\|$ be an arbitrary norm in V (not necessarily induced by the scalar product), and let $\|\cdot\|_*$ be the corresponding conjugate norm. Then,

$$\partial \|\cdot\|(x_0) = \begin{cases} B_{\|\cdot\|_*}(0,1), & \text{if } x_0 = 0, \\ \{s \in V: \|s\|_* \leq 1; \langle s, x_0 \rangle = \|x_0\|\} = \{s \in V: \|s\|_* = 1; \langle s, x_0 \rangle = \|x_0\|\}, & \text{otherwise}. \end{cases}$$

Где $B_{\mathbb{L}\mathbb{L}}$ (0,1) есть замкнутый единичный относительно сопряженной нормы шар с центром в нуле. Другими словами, вектор $s\in V$ с $\|s\|_*=1$ является субградиентом нормы $\|\cdot\|$ в точке $x_0\neq 0$ тогда и только тогда, когда неравенство Гёльдера $\langle s, x_0 \rangle \leq \|x_0\|$ переходит в равенство.

Важно отметить, что выражение слева есть супремум

Пусть
$$s \in V$$
. По определению $s \in \partial \|\cdot\|(x_0)$ если и только если из определения сопряженной функции по Фенхелю для нормы, которая, как известно, записывается так:
$$\langle s,x \rangle - \|x\| \leq \langle s,x_0 \rangle - \|x_0\|, \text{ for all } x \in V,$$

что равносильно

$$\sup_{x \in V} \{ \langle s, x \rangle - \|x\| \} = \begin{cases} 0, & \text{if } \|s\|_* \le 1, \\ +\infty, & \text{otherwise.} \end{cases}$$

Таким образом, выражение равносильно $\|s\|_{*} \leqslant 1$

Следовательно, остаётся заметить, что для $x_0 \neq 0$ неравенство $\|s\|_* \leq 1$ должно переходить в равенство, поскольку при $||s||_* < 1$ неравенство Гёльдера влечёт $\langle s, x_0 \rangle < ||s||_* ||x_0|| < ||x_0||$.

Сопряженная норма в примере выше появилась не случайно. Оказывается, что совершенно аналогичным образом для произвольной функции f (не только для нормы) её субдифференциал может быть описан в терминах двойственного объекта - сопряженной по Фенхелю функции.

