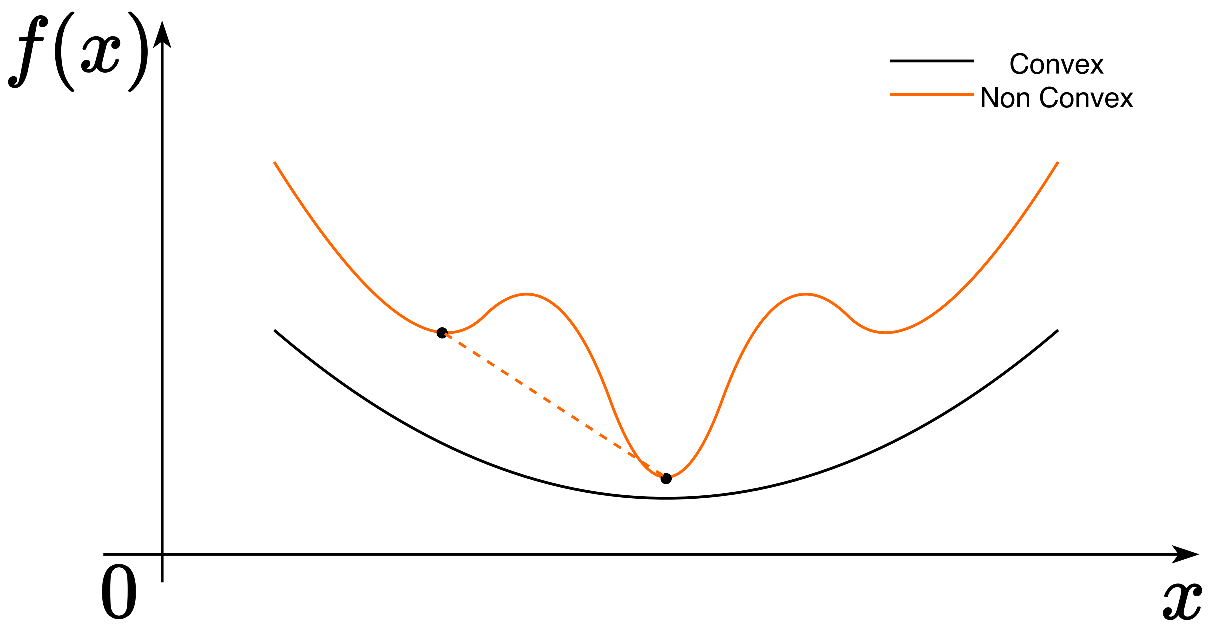
Convex function

The function $f(x)$, **which is defined on the convex set** $S \subseteq \mathbb{R}^n$, is called **convex** on S, if:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

for any $x_1, x_2 \in S$ and $0 \leq \lambda \leq 1$.
If above inequality holds as strict inequality $x_1 \neq x_2$ and $0 < \lambda < 1$, then function is called strictly convex on S.



🧑‍🏫 **EXAMPLE**

- $f(x) = x^p, \quad p > 1, \quad x \in \mathbb{R}_+$
- $f(x) = \|x\|^p, \quad p > 1, x \in \mathbb{R}^n$
- $f(x) = e^{cx}, \quad c \in \mathbb{R}, x \in \mathbb{R}$
- $f(x) = -\ln x, \quad x \in \mathbb{R}_{++}$
- $f(x) = x \ln x, \quad x \in \mathbb{R}_{++}$
- The sum of the largest k coordinates $f(x) = x_{(1)} + \ldots + x_{(k)}, \quad x \in \mathbb{R}^n$

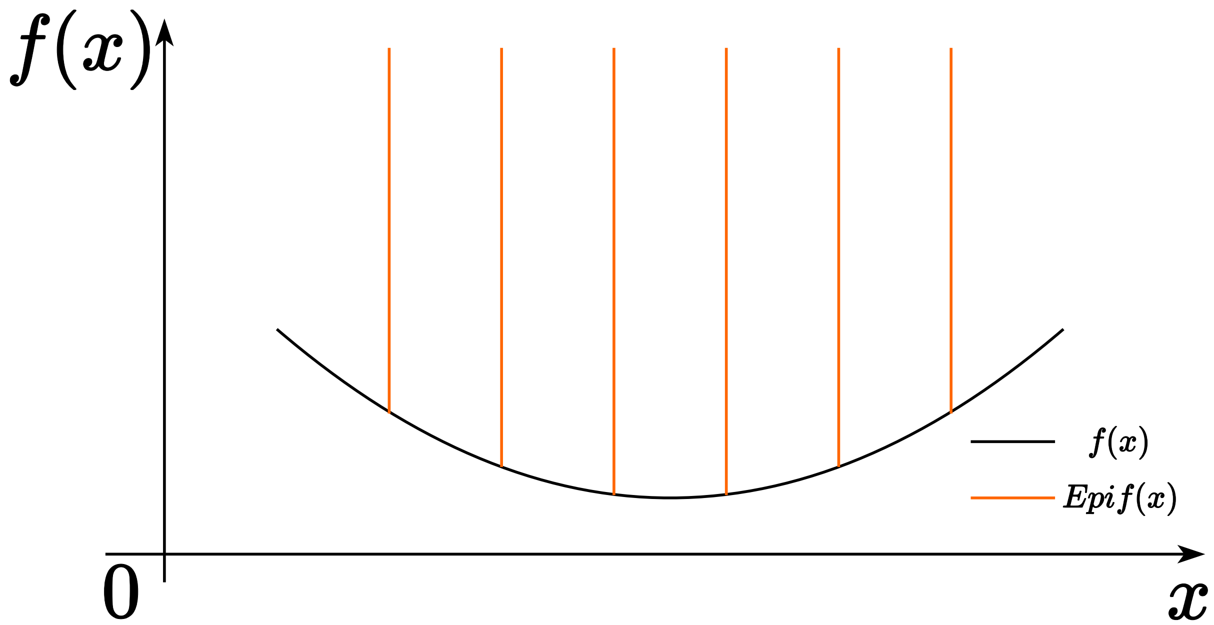
- $f(X) = \lambda_{max}(X), \quad X = X^T$
- $f(X) = -\log \det X, \quad X \in S^n_{++}$

Epigraph

For the function $f(x)$, defined on $S \subseteq \mathbb{R}^n$, the following set:

$$\text{epi } f = \{[x, \mu] \in S \times \mathbb{R} : f(x) \leq \mu\}$$

is called **epigraph** of the function $f(x)$.

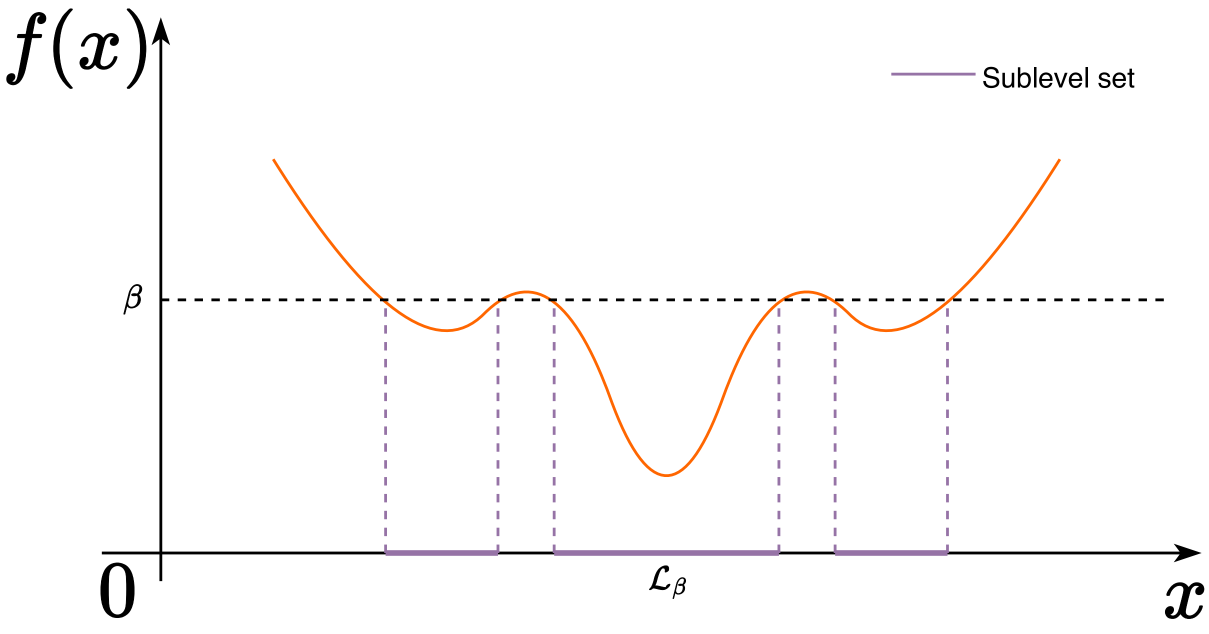


Sublevel set

For the function $f(x)$, defined on $S \subseteq \mathbb{R}^n$, the following set:

$$\mathcal{L}_\beta = \{x \in S : f(x) \leq \beta\}$$

is called **sublevel set** or Lebesgue set of the function $f(x)$.



Criteria of convexity

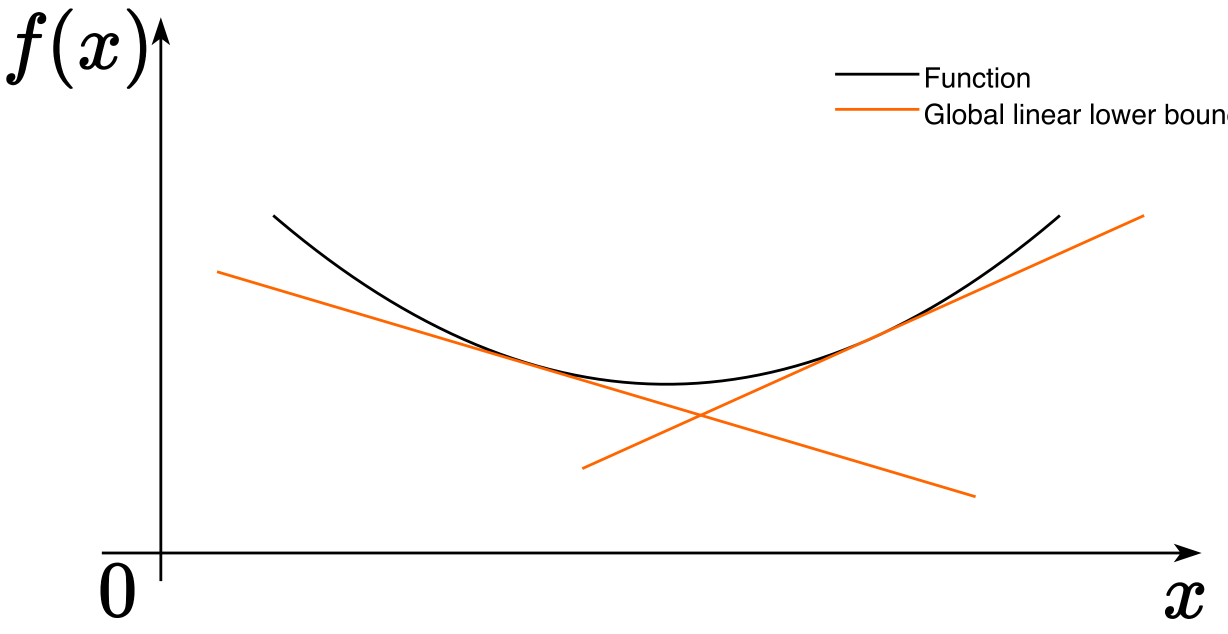
First order differential criterion of convexity

The differentiable function $f(x)$ defined on the convex set $S \subseteq \mathbb{R}^n$ is convex if and only if $\forall x, y \in S$:

$$f(y) \geq f(x) + \nabla f^T(x)(y - x)$$

Let $y = x + \Delta x$, then the criterion will become more tractable:

$$f(x + \Delta x) \geq f(x) + \nabla f^T(x)\Delta x$$



Second order differential criterion of convexity

Twice differentiable function $f(x)$ defined on the convex set $S \subseteq \mathbb{R}^n$ is convex if and only if $\forall x \in \mathbf{int}(S) \neq \emptyset$:

$$\nabla^2 f(x) \succeq 0$$

In other words, $\forall y \in \mathbb{R}^n$:

$$\langle y, \nabla^2 f(x) y \rangle \geq 0$$

Connection with epigraph

The function is convex if and only if its epigraph is a convex set.

🧑‍🏫 **EXAMPLE**

Let a norm $\|\cdot\|$ be defined in the space U. Consider the set:

$$K := \{(x, t) \in U \times \mathbb{R}^+ : \|x\| \leq t\}$$

which represents the epigraph of the function $x \mapsto \|x\|$. This set is called the cone norm. According to statement above, the set K is convex.

In the case where $U = \mathbb{R}^n$ and $\|x\| = \|x\|_2$ (Euclidean norm), the abstract set K transitions into the set:

$$\{(x,t) \in \mathbb{R}^n \times \mathbb{R}^+ : \|x\|_2 \leq t\}$$

Connection with sublevel set

If $f(x)$ - is a convex function defined on the convex set $S \subseteq \mathbb{R}^n$, then for any β sublevel set $\mathcal{L}_\beta$ is convex.

The function $f(x)$ defined on the convex set $S \subseteq \mathbb{R}^n$ is closed if and only if for any β sublevel set $\mathcal{L}_\beta$ is closed.

Reduction to a line

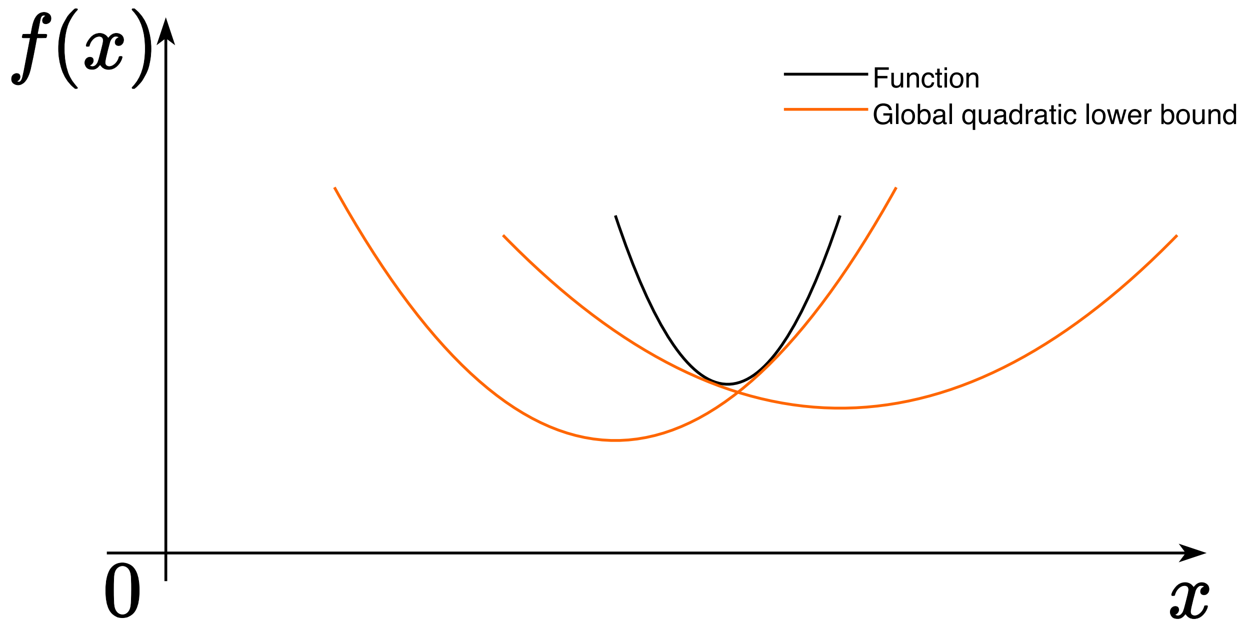
$f : S \to \mathbb{R}$ is convex if and only if S is a convex set and the function $g(t) = f(x + tv)$ defined on $\{t \mid x + tv \in S\}$ is convex for any $x \in S, v \in \mathbb{R}^n$, which allows to check convexity of the scalar function in order to establish convexity of the vector function.

Strong convexity

$f(x)$, **defined on the convex set** $S \subseteq \mathbb{R}^n$, is called μ-strongly convex (strongly convex) on S, if:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) - \mu\lambda(1-\lambda)\|x_1 - x_2\|^2$$

for any $x_1, x_2 \in S$ and $0 \leq \lambda \leq 1$ for some $\mu > 0$.



Criteria of strong convexity

First order differential criterion of strong convexity

Differentiable $f(x)$ defined on the convex set $S \subseteq \mathbb{R}^n$ is μ-strongly convex if and only if $\forall x, y \in S$:

$$f(y) \geq f(x) + \nabla f^T(x)(y - x) + \frac{\mu}{2}\|y - x\|^2$$

Let $y = x + \Delta x$, then the criterion will become more tractable:

$$f(x + \Delta x) \geq f(x) + \nabla f^T(x)\Delta x + \frac{\mu}{2}\|\Delta x\|^2$$

Second order differential criterion of strong convexity

Twice differentiable function $f(x)$ defined on the convex set $S \subseteq \mathbb{R}^n$ is called μ-strongly convex if and only if $\forall x \in \mathbf{int}(S) \neq \emptyset$:

$$\nabla^2 f(x) \succeq \mu I$$

In other words:

$$\langle y, \nabla^2 f(x) y \rangle \geq \mu \|y\|^2$$

Facts

- $f(x)$ is called (strictly) concave, if the function $-f(x)$ - is (strictly) convex.
- Jensen's inequality for the convex functions:

  $$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$

  for $\alpha_i \geq 0; \quad \sum\limits_{i=1}^{n} \alpha_i = 1$ (probability simplex)
  For the infinite dimension case:

  $$f\left(\int\limits_S x p(x) dx\right) \leq \int\limits_S f(x) p(x) dx$$

  If the integrals exist and $p(x) \geq 0, \quad \int\limits_S p(x) dx = 1$

- If the function $f(x)$ and the set S are convex, then any local minimum $x^* = \arg\min\limits_{x \in S} f(x)$ will be the global one. Strong convexity guarantees the uniqueness of the solution.
- Let $f(x)$ - be a convex function on a convex set $S \subseteq \mathbb{R}^n$. Then $f(x)$ is continuous $\forall x \in \mathbf{ri}(S)$.

Operations that preserve convexity

- Non-negative sum of the convex functions: $\alpha f(x) + \beta g(x), (\alpha \geq 0, \beta \geq 0)$.
- Composition with affine function $f(Ax + b)$ is convex, if $f(x)$ is convex.
- Pointwise maximum (supremum): If $f_1(x), \ldots, f_m(x)$ are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex.
- If $f(x, y)$ is convex on x for any $y \in Y$: $g(x) = \sup\limits_{y \in Y} f(x, y)$ is convex.
- If $f(x)$ is convex on S, then $g(x, t) = t f(x/t)$ - is convex with $x/t \in S, t > 0$.
- Let $f_1 : S_1 \to \mathbb{R}$ and $f_2 : S_2 \to \mathbb{R}$, where $\mathrm{range}(f_1) \subseteq S_2$. If f_1 and f_2 are convex, and f_2 is increasing, then $f_2 \circ f_1$ is convex on S_1.

Other forms of convexity

- Log-convex: $\log f$ is convex; Log convexity implies convexity.
- Log-concavity: $\log f$ concave; **not** closed under addition!
- Exponentially convex: $[f(x_i + x_j)] \succeq 0$, for $x_1, \ldots, x_n$
- Operator convex: $f(\lambda X + (1-\lambda)Y) \preceq \lambda f(X) + (1-\lambda) f(Y)$
- Quasiconvex: $f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}$
- Pseudoconvex: $\langle \nabla f(y), x - y \rangle \geq 0 \longrightarrow f(x) \geq f(y)$
- Discrete convexity: $f : \mathbb{Z}^n \to \mathbb{Z}$; "convexity + matroid theory."

🧑‍🏫 **EXAMPLE**

Show, that $f(x) = c^\top x + b$ is convex and concave.

▼ Solution

🧑‍🏫 **EXAMPLE**

Show, that $f(x) = x^\top A x$, where $A \succeq 0$ - is convex on $\mathbb{R}^n$.

▼ Solution

🧑‍🏫 **EXAMPLE**

Show, that $f(A) = \lambda_{max}(A)$ - is convex, if $A \in S^n$.

▼ Solution

🧑‍🏫 **EXAMPLE**

PL inequality holds if the following condition is satisfied for some $\mu > 0$,

$$\|\nabla f(x)\|^2 \geq \mu(f(x) - f^*) \forall x$$

The example of function, that satisfy PL-condition, but is not convex. $f(x, y) = \dfrac{(y - \sin x)^2}{2}$

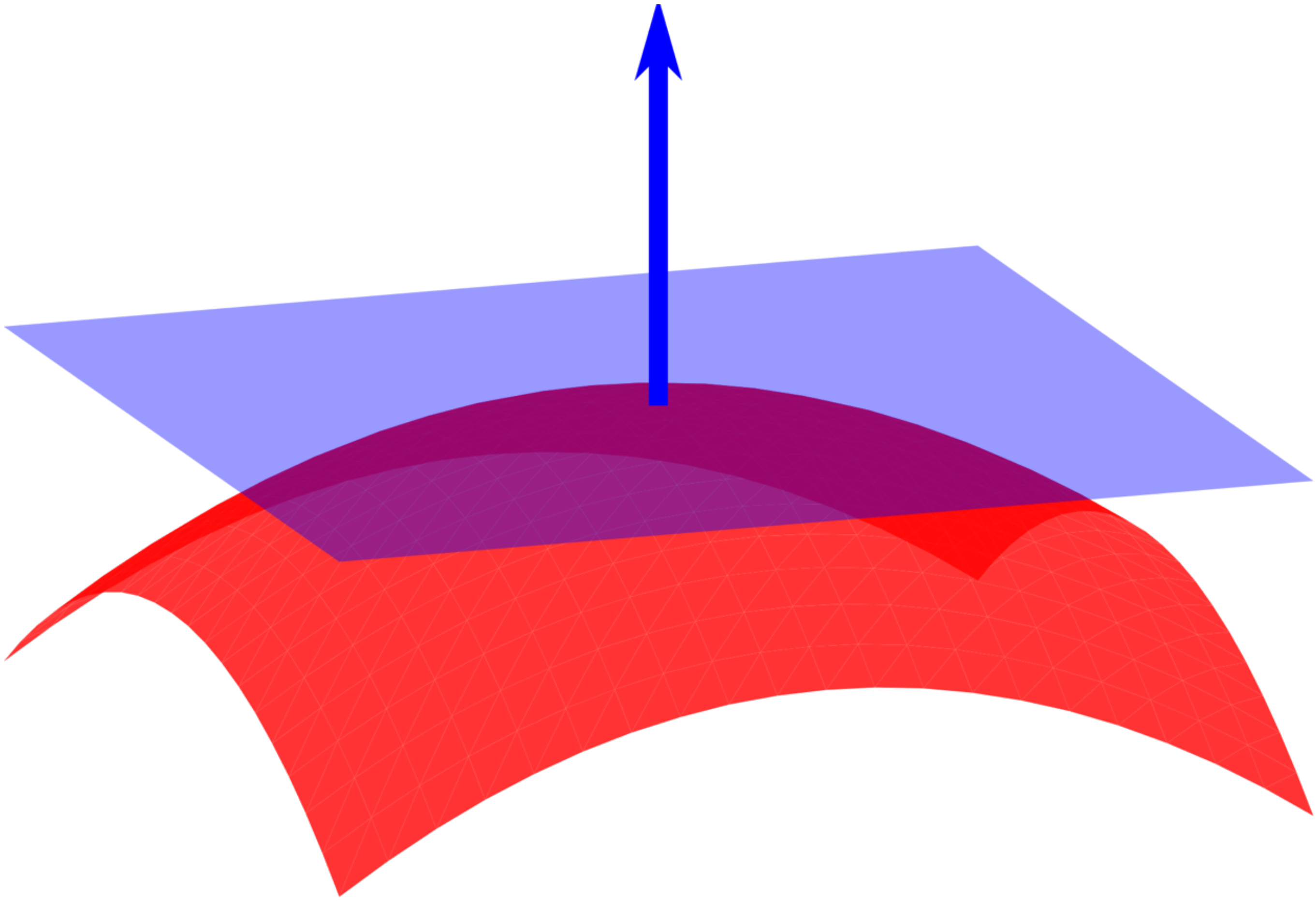
References

- Steven Boyd lectures
- Suvrit Sra lectures
- Martin Jaggi lectures
- Example pf Pl non-convex function Open in Colab

Conjugate (dual) set

Пусть $S \subseteq \mathbb{R}^n$ - произвольное непустое множество. Тогда сопряженное к нему множество определяется, как:

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



Double conjugate set

Множество S^{**} называется вторым сопряженным к множеству S, если:

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

Inter-conjugate and self-conjugate sets

- Множества S_1 и S_2 называются **взаимосопряженными**, если $S_1^* = S_2, S_2^* = S_1$.
- Множество S называется **самосопряженным**, если $S^* = S$.

Properties

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

Examples

1

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

Решение:

- $$S \subset \overline{S} \to \left(\overline{S}\right)^* \subset S^*$$
- 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: $p^T x_k \geq -1 \to p^T x_0 \geq -1$. Значит, $p \in \left(\overline{S}\right)^*$, отсюда $S^* \subset \left(\overline{S}\right)^*$

2

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

Решение:

- $$S \subset \mathbf{conv}(S) \to (\mathbf{conv}(S))^* \subset S^*$$
- Пусть $p \in S^*$, $x_0 \in \mathbf{conv}(S)$, т.е. $x_0 = \sum\limits_{i=1}^k \theta_i x_i \mid x_i \in S, \sum\limits_{i=1}^k \theta_i = 1, \theta_i \geq 0.$

  Значит, $p^T x_0 = \sum\limits_{i=1}^k \theta_i p^T x_i \geq \sum\limits_{i=1}^k \theta_i (-1) = 1 \cdot (-1) = -1$. Значит, $p \in (\mathbf{conv}(S))^*$, отсюда $S^* \subset (\mathbf{conv}(S))^*$

3

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

Решение:

- Пусть $B(0, r) = X, B(0, 1/r) = Y$. Возьмем вектор нормали $p \in X^*$, тогда для любого $x \in X : p^T x \geq -1$.
- Из всех точек шара X возьмем такую $x \in X$, что скалярное произведение её на p: $p^T x$ было бы минимально, тогда это точка $x = -\frac{p}{\|p\|}r$

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

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

  Значит, $X^* \subset Y$.

- Теперь пусть $p \in Y$. Нам надо показать, что $p \in X^*$, т.е. $\langle p, x\rangle \geq -1$. Достаточно применить неравенство Коши - Буняковского:

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

  Последнее исходит из того, что $p \in B(0, 1/r)$, а $x \in B(0, r)$.

  Значит, $Y \subset X^*$.

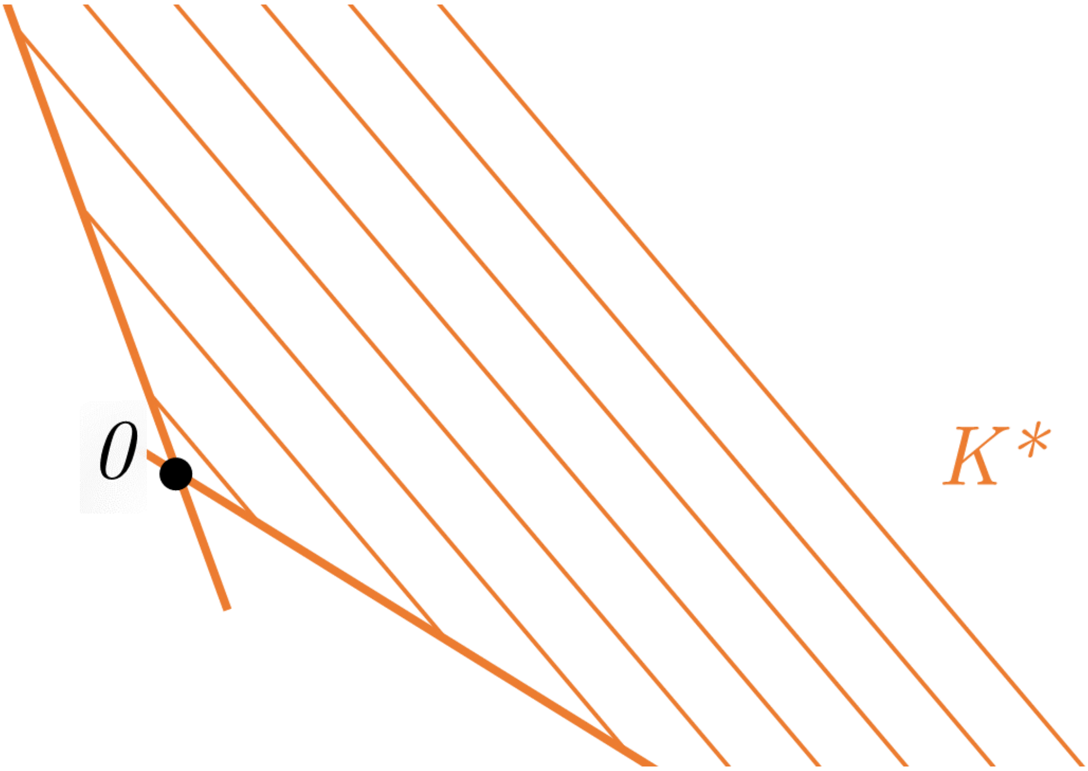
Dual cones

Сопряженным конусом к конусу K называется такое множество K^*, что:

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

Чтобы показать, что это определение непосредственно следует из теории выше, вспомним, что такое сопряженное множество и что такое конус $\forall \lambda > 0$

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



Dual cones properties

- 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:

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

- Пусть $K_1, \ldots, K_m$ - конусы в $\mathbb{R}^n$. Пусть также их пересечение имеет внутреннюю точку, тогда:

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

Examples

Найти сопряженнй конус для монотонного неотрицательного конуса:

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

Решение:

Заметим, что:

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

Так как в представленной сумме в каждом слагаемом второй множитель неотрицательный, то:

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

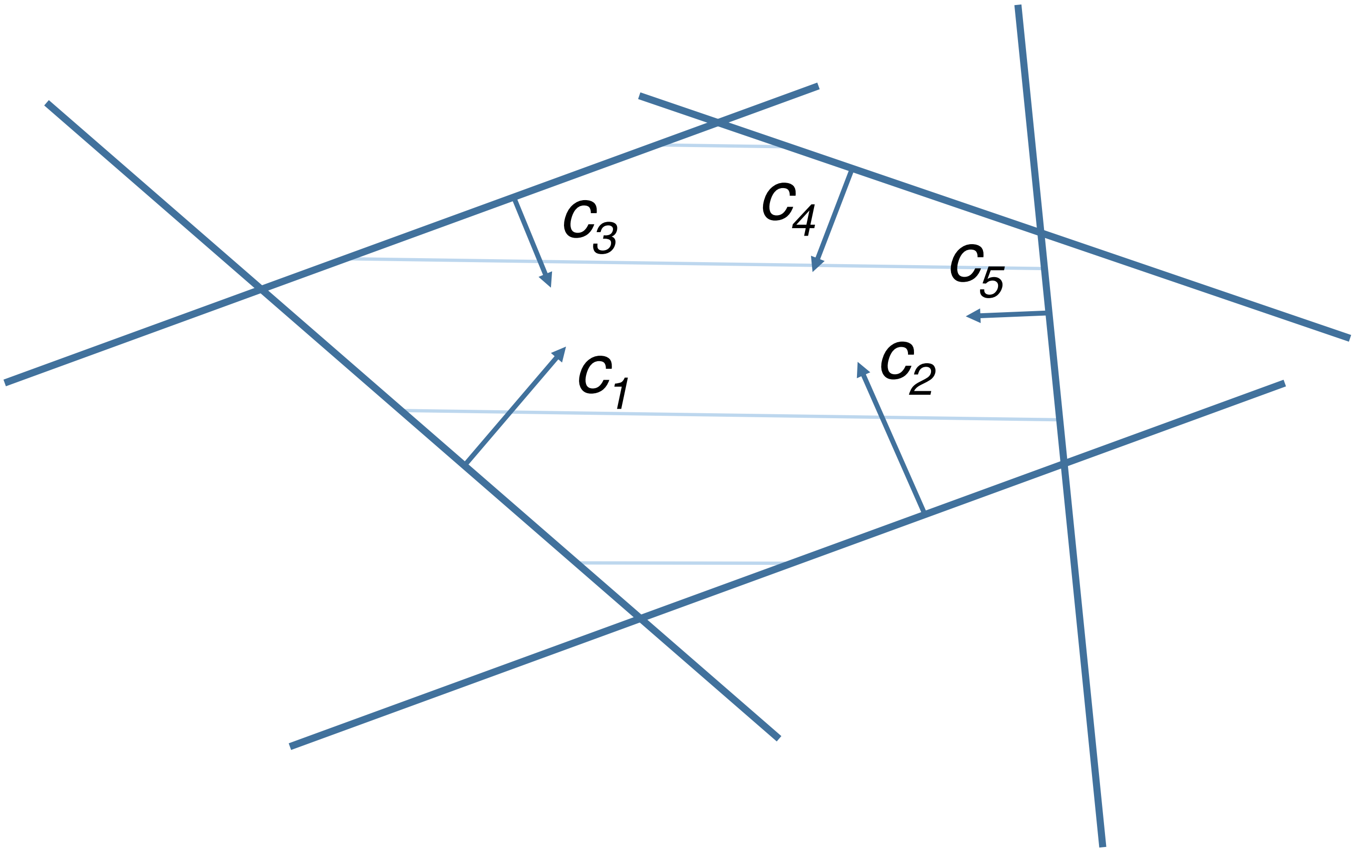
Значит, $K^* = \left\{ y \mid \sum\limits_{i=1}^{k} y_i \geq 0, k = \overline{1, n} \right\}$

Polyhedra

Множество решений системы линейных неравенств и равенств представляет собой многогранник:

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

Здесь $A \in \mathbb{R}^{m \times n}, C \in \mathbb{R}^{p \times n}$, а неравенство - поэлементное.



ТЕОРЕМА:

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

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

является полиэдр (многогранник):

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

ДОКАЗАТЕЛЬСТВО:

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

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

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

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

- Пусть, напротив, $p \in Y$. Для любой точки $x \in X$:

  $$x = \sum_{i=1}^m \theta_i x_i \qquad \sum_{i=1}^k \theta_i = 1, \theta_i \geq 0$$

  Значит:

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

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

5

Найти и изобразить на плоскости множество, сопряженное к многогранному конусу:

$$S = \mathbf{cone}\left\{(-3,1), (2,3), (4,5)\right\}$$

Решение:

Используя теорему выше:

$$S^* = \{-3p_1 + p_2 \geq 0, 2p_1 + 3p_2 \geq 0, 4p_1 + 5p_2 \geq 0\}$$

Лемма (теорема) Фаркаша (Фаркаша - Минковского)

Пусть $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Тогда имеет решение одна и только одна из следующих двух систем:

$$1)\ Ax = b, x \geq 0$$

$$2)\ p^\top A \geq 0, \langle p, b \rangle < 0$$

$Ax = b$ при $x \geq 0$ означает, что b лежит в конусе, натянутым на столбцы матрицы A.

$pA \geq 0,\ \langle p, b \rangle < 0$ означает, что существует разделяющая гиперплоскость между вектором b и конусом из столбцов матрицы A.

СЛЕДСТВИЕ:
Пусть $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Тогда имеет решение одна и только одна из следующих двух систем:

$$1)\ Ax \leq b$$

$$2)\ p^\top A = 0, \langle p, b \rangle < 0, p \geq 0$$

Если в задаче линейного программирования на минимум допустимое множество непусто и целевая функция ограничена на нём снизу, то задача имеет решение.