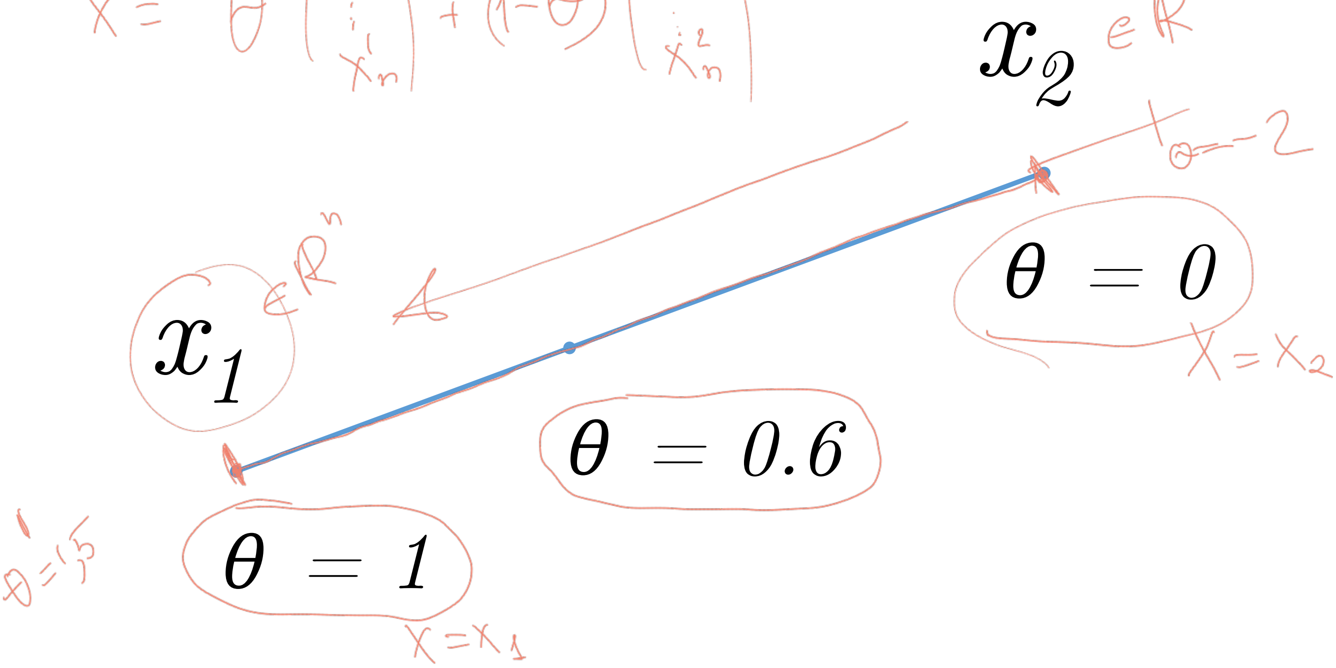
Convex set

Line segment

Suppose x_1, x_2 are two points in $\mathbb{R}^n$. Then the line segment between them is defined as follows:

$$x = \theta x_1 + (1 - \theta)x_2,\ \theta \in [0, 1]$$

$$x = \theta \begin{pmatrix} x_1^1 \\ \vdots \\ x_n^1 \end{pmatrix} + (1-\theta) \begin{pmatrix} x_1^2 \\ \vdots \\ x_n^2 \end{pmatrix}$$



Convex set

The set S is called **convex** if for any x_1, x_2 from S the line segment between them also lies in S, i.e.

$S = \{0\}$

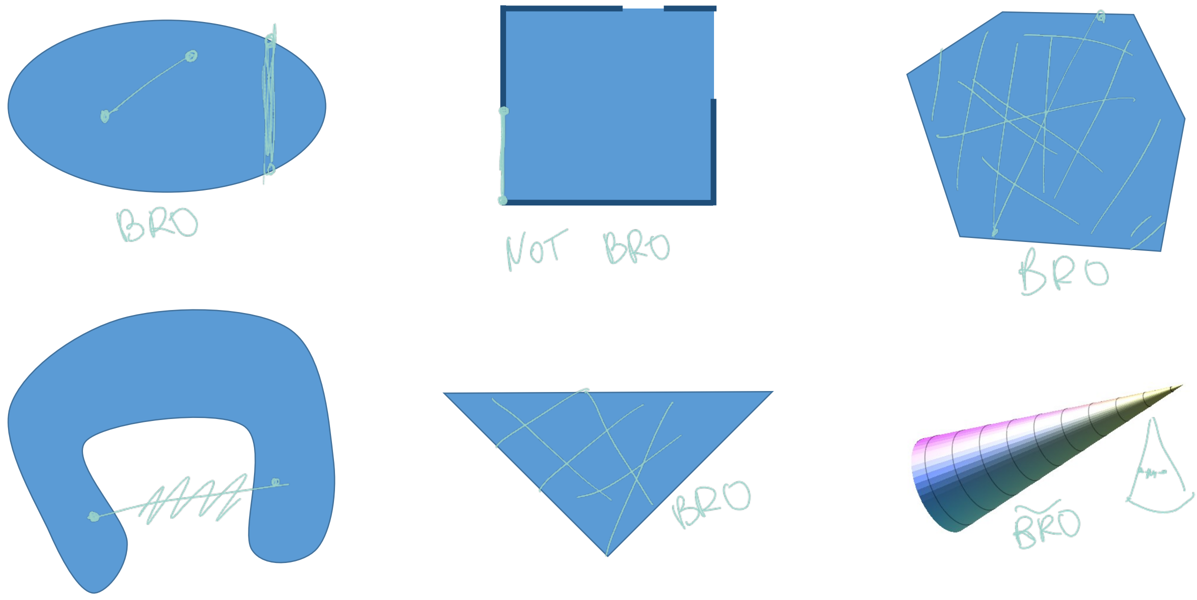
$\varnothing$ - выпуклое по опр.

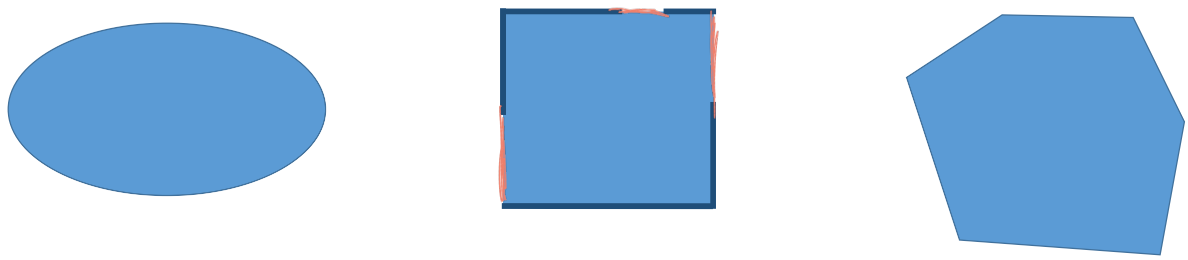
$$\forall \theta \in [0, 1],\ \forall x_1, x_2 \in S:$$
$$\theta x_1 + (1 - \theta)x_2 \in S$$

$S = \varnothing$

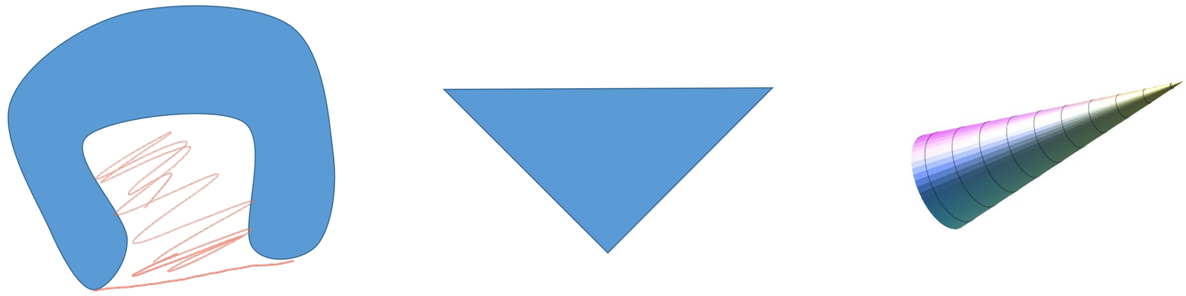
Examples:

- Any affine set
- Ray
- Line segment





BRO NOT BRO BRO

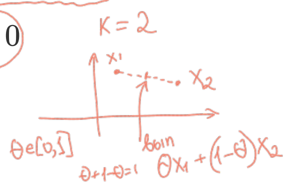


NOT BRO BRO BRO

Related definitions

Convex combination

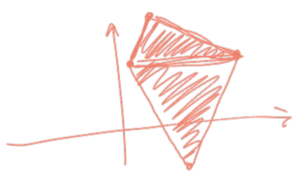
Let $x_1, x_2, \dots, x_k \in S$, then the point $\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k$ is called the convex combination of points $x_1, x_2, \dots, x_k$ if $\sum\limits_{i=1}^k \theta_i = 1$, $\theta_i \geq 0$



Convex hull

The set of all convex combinations of points from S is called the convex hull of the set S.

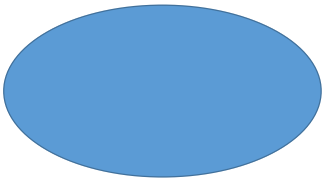
$$\mathbf{conv}(S) = \left\{ \sum\limits_{i=1}^k \theta_i x_i \mid x_i \in S, \sum\limits_{i=1}^k \theta_i = 1, \; \theta_i \geq 0 \right\}$$



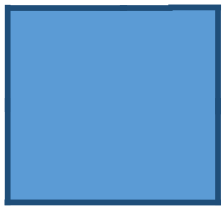
- The set $\mathbf{conv}(S)$ is the smallest convex set containing S.
- The set S is convex if and only if $S = \mathbf{conv}(S)$.

convex hull

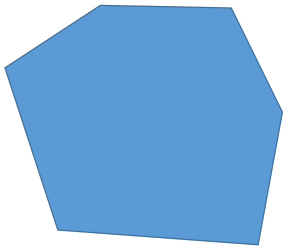
Examples:



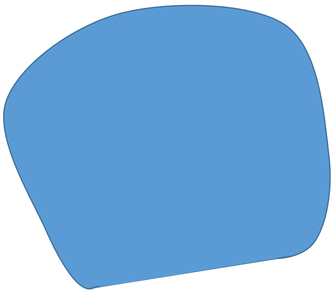
BRO



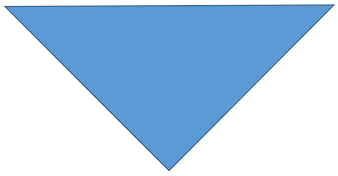
BRO



BRO

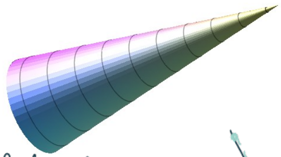


BRO



BRO

НАФИГА?

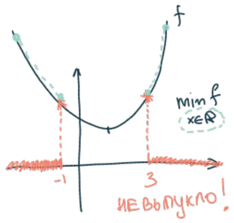


BRO

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

S - выпукло

f - выпукла
- любой локальн. минимум = глобальный
- существуют алгоритмы

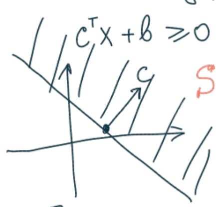


Finding convexity

In practice it is very important to understand whether a specific set is convex or not. Two approaches are used for this depending on the context.

- By definition.
- Show that S is derived from simple convex sets using operations that preserve convexity.

Пример: докажем, что полупространство - выпукло



$c \in \mathbb{R}^n$, $b \in \mathbb{R}$ заданы

1) Пусть $x_1 \in S$: $c^T x_1 + b \geq 0$
$x_2 \in S$: $c^T x_2 + b \geq 0$

$\theta \in [0,1]$

2) Возьмём $\tilde{x} = \theta x_1 + (1-\theta) x_2$

$c^T \tilde{x} = \theta c^T x_1 + (1-\theta) c^T x_2$ $\quad \theta b + (1-\theta) b$

$c^T \tilde{x} + b = \theta c^T x_1 + (1-\theta) c^T x_2 + b =$

$= \underset{\geq 0}{\theta}(\underset{\geq 0}{c^T x_1 + b}) + \underset{\geq 0}{(1-\theta)}(\underset{\geq 0}{c^T x_2 + b}) \geq 0 \quad \in S$

By definition

$$
x_1, x_2 \in S, \; 0 \leq \theta \leq 1 \quad \rightarrow \quad \theta x_1 + (1-\theta) x_2 \in S
$$

Preserving convexity

The linear combination of convex sets is convex

Let there be 2 convex sets S_x, S_y, let the set
$S = \{s \mid s = c_1 x + c_2 y, \; x \in S_x, \; y \in S_y, \; c_1, c_2 \in \mathbb{R}\}$

Take two points from S: $s_1 = c_1 x_1 + c_2 y_1, s_2 = c_1 x_2 + c_2 y_2$ and prove that the segment between them $\theta s_1 + (1-\theta) s_2, \theta \in [0,1]$ also belongs to S:

$$
\theta s_1 + (1-\theta) s_2
$$

$$
\theta(c_1 x_1 + c_2 y_1) + (1-\theta)(c_1 x_2 + c_2 y_2)
$$

$$
c_1(\theta x_1 + (1-\theta) x_2) + c_2(\theta y_1 + (1-\theta) y_2)
$$

$$
c_1 x + c_2 y \in S
$$

The intersection of any (!) number of convex sets is convex

If the desired intersection is empty or contains one point, the property is proved by definition. Otherwise, take 2 points and a segment between them. These points must lie in all intersecting sets, and since they are all convex, the segment between them lies in all sets and, therefore, in their intersection.

The image of the convex set under affine mapping is convex

$$S \subseteq \mathbb{R}^n \text{ convex } \;\rightarrow\; f(S) = \{f(x) \mid x \in S\} \text{ convex } \quad (f(x) = \mathbf{A}x + \mathbf{b})$$

Examples of affine functions: extension, projection, transposition, set of solutions of linear matrix inequality $\{x \mid x_1A_1 + \ldots + x_mA_m \preceq B\}$ Here $A_i, B \in \mathbf{S}^p$ are symmetric matrices $p \times p$.

Note also that the prototype of the convex set under affine mapping is also convex.

$$S \subseteq \mathbb{R}^m \text{ convex } \;\rightarrow\; f^{-1}(S) = \{x \in \mathbb{R}^n \mid f(x) \in S\} \text{ convex } \;\; (f(x) = \mathbf{A}x + \mathbf{b})$$

Example 1

Prove, that ball in $\mathbb{R}^n$ (i.e. the following set $\{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_c\| \leq r\}$) - is convex.

1) $x_1, x_2 \in S$: $\|x_1 - x_c\| \leq r$, $\|x_2 - x_c\| \leq r$

$\|a+b\| \leq \|a\| + \|b\|$

Доказать $\|\tilde{x} - x_c\| \leq r$

2) $\tilde{x} = \theta x_1 + (1-\theta)x_2$

$\|\theta x_1 + (1-\theta)x_2 - x_c\| = \|\theta(x_1 - x_c) + (1-\theta)(x_2 - x_c)\| \leq \theta \underset{r}{\underset{\wedge\!\shortparallel}{\|x_1 - x_c\|}} + (1-\theta)\underset{r}{\underset{\wedge\!\shortparallel}{\|x_2 - x_c\|}} =$

$x_c = \theta x_c + (1-\theta)x_c$

$= \theta \cdot r + (1-\theta) r = r$

Example 2

Which of the sets are convex: 1. Stripe, $\{x \in \mathbb{R}^n \mid \alpha \leq a^\top x \leq \beta\}$ 1. Rectangle, $\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, i = \overline{1,n}\}$ 1. Kleen, $\{x \in \mathbb{R}^n \mid a_1^\top x \leq b_1, a_2^\top x \leq b_2\}$ 1. A set of points closer to a given point than a given set that does not contain a point, $\{x \in \mathbb{R}^n \mid \|x - x_0\|_2 \leq \|x - y\|_2, \forall y \in S \subseteq \mathbb{R}^n\}$ 1. A set of points, which are closer to one set than another, $\{x \in \mathbb{R}^n \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T), S, T \subseteq \mathbb{R}^n\}$ 1. A set of points, $\{x \in \mathbb{R}^n \mid x + X \subseteq S\}$, where $S \subseteq \mathbb{R}^n$ is convex and $X \subseteq \mathbb{R}^n$ is arbitrary. 1. A set of points whose distance to a given point does not exceed a certain part of the distance to another given point is $\{x \in \mathbb{R}^n \mid \|x - a\|_2 \leq \theta\|xb\|_2, a, b \in \mathbb{R}^n, 0 \leq 1\}$

Example 3

Let $x \in \mathbb{R}$ is a random variable with a given probability distribution of $\mathbb{P}(x = a_i) = p_i$, where $i = 1, \ldots, n$, and $a_1 < \ldots < a_n$. It is said that the probability vector of outcomes of $p \in \mathbb{R}^n$ belongs to the probabilistic simplex, i.e.
$P = \{p \mid \mathbf{1}^T p = 1, p \succeq 0\} = \{p \mid p_1 + \ldots + p_n = 1, p_i \ge 0\}$. Determine if the following sets of p are convex: 1. $\alpha < \mathbb{E} f(x) < \beta$, where $\mathbb{E} f(x)$ stands for expected value of $f(x) : \mathbb{R} \to \mathbb{R}$, i.e. $\mathbb{E} f(x) = \sum\limits_{i=1}^{n} p_i f(a_i)$ 1. $\mathbb{E} x^2 \le \alpha$ 1. $\mathbb{V} x \le \alpha$

Convex function

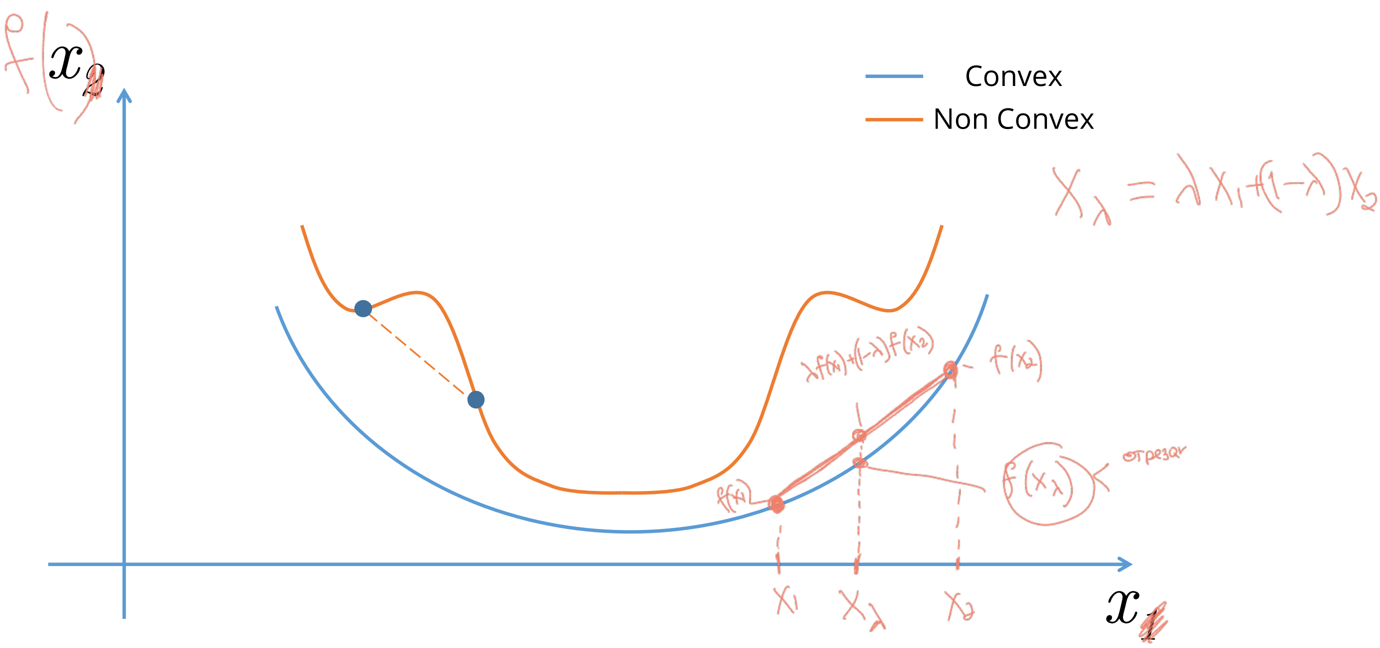
Convex function

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

$$\boxed{f(\lambda x_1 + (1 - \lambda) x_2) \le \lambda f(x_1) + (1 - \lambda) f(x_2)}$$

Неравенство
Йенсена

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



Examples

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

$\ln x$ $\quad \ln x \quad x \quad -\ln x \quad x\ln x$

$f(x) = x^2$ $\quad \ominus x^2 + 0$

вогнутая (впуклая)

$\sum_{i=1}^{k} a_i f_i(x)$ не вып

$k = 1 \quad a_i = -1 \quad -x^2$

выпуклая

много локальных минимумов

$f(x)$ выпукла convex

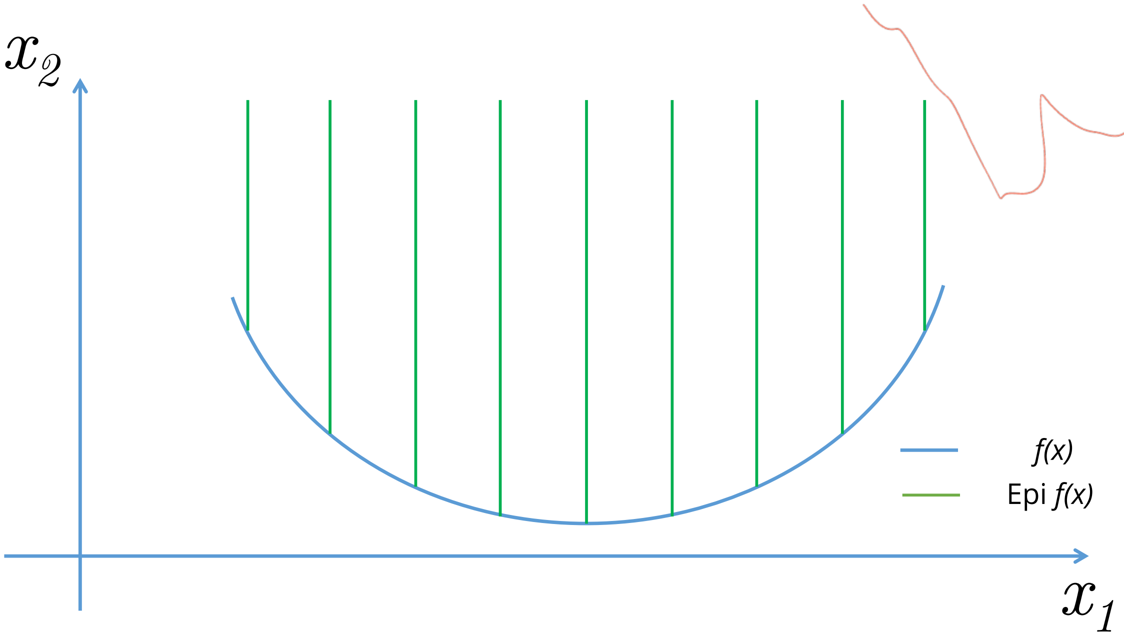
$-f(x)$ вогнута concave

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 $f(x)$

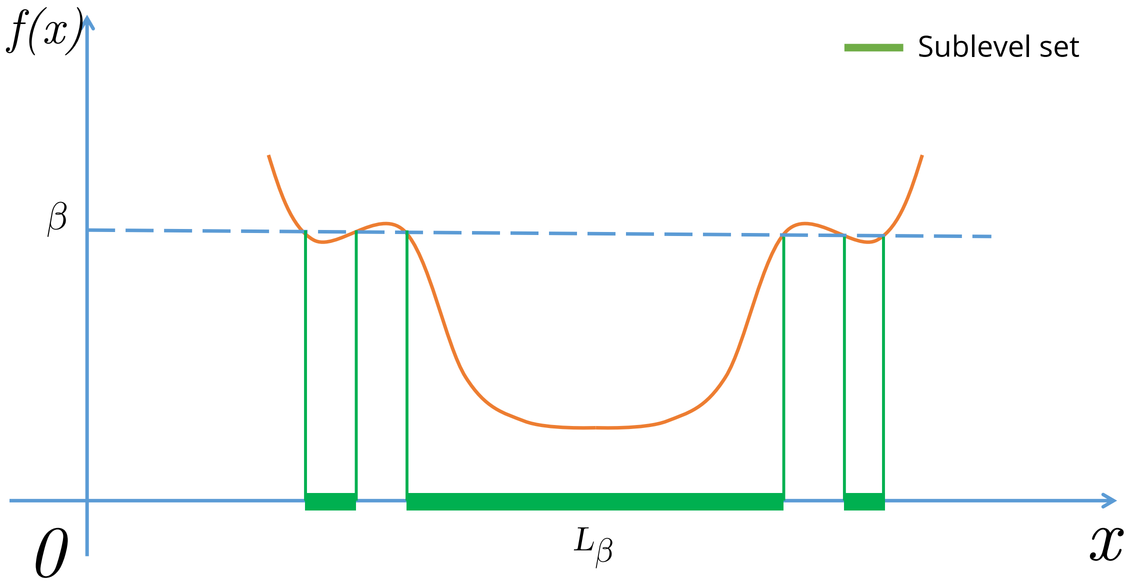


Sublevel set

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 $f(x)$

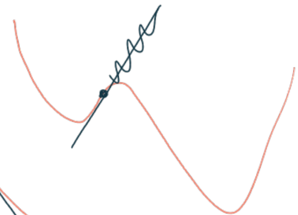


Criteria of convexity

First order differential criterion of convexity

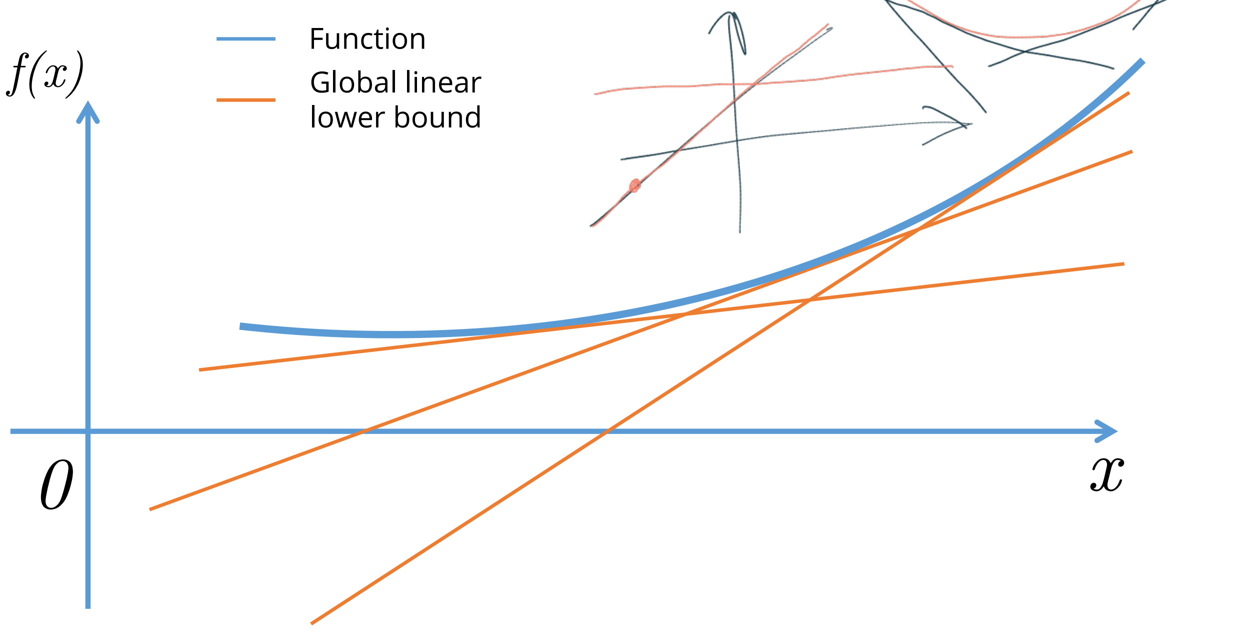
The differentiable function $f(x)$ defined on the convex set $S \subseteq \mathbb{R}^n$ is convex if and only if $\forall x, y \in S$:

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



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 $\forall x \in \textbf{int}(S) \neq \emptyset$:

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

$f''(x) \geq 0$

$x^T \nabla^2 f x \geq 0$

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 convex set.

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 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 covexity of the vector function.

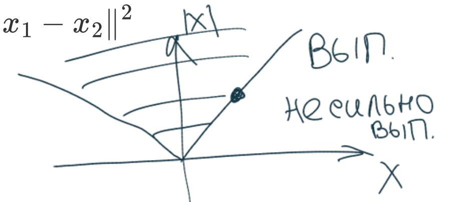
① ГЛОБАЛЬНЫЙ МИНИМУМ ОДИН !

Strong convexity

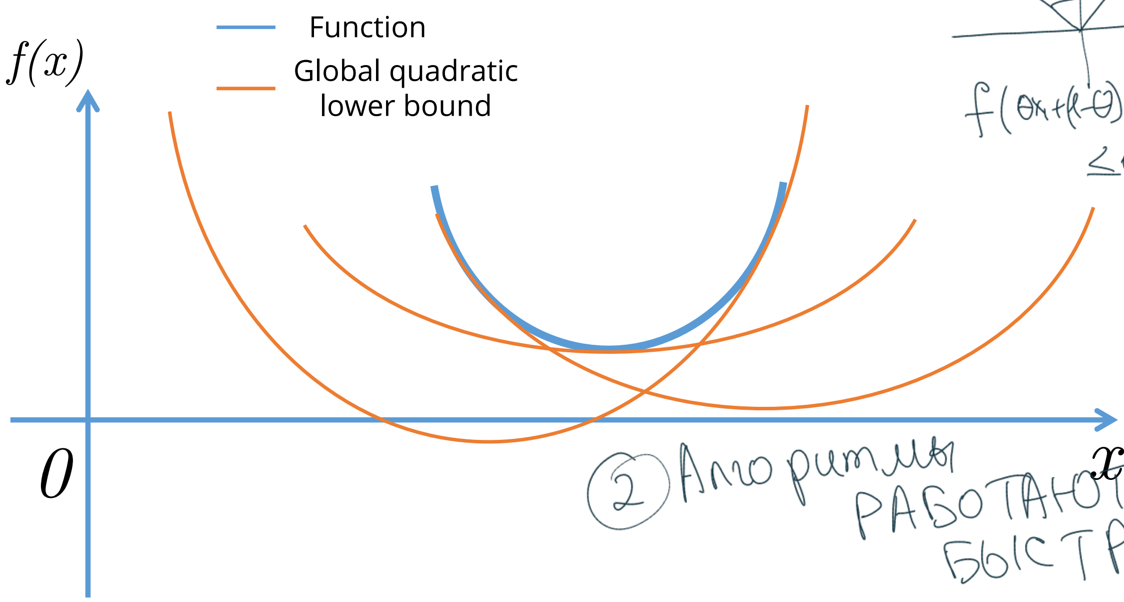
$f(x)$, **defined on the convex set** $S \subseteq \mathbb{R}^n$, is called μ-strongly convex (strogly convex) on S, if:

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

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



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



② Алгоритмы РАБОТАЮТ БЫСТРЕЕ

Criteria of strong convexity

First order differential criterion of strong convexity

Differentiable $f(x)$ defined on the convex set $S \subseteq \mathbb{R}^n$ 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 $\forall x \in \mathbf{int}(S) \neq \emptyset$:

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

$\nabla^2 f - \mu I \succeq 0$ если $\exists \mu > 0$

$\mu > 0$

$x^T(\nabla^2 f - \mu I)x > 0$

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 $-f(x)$ - (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.

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.”

References

- Steven Boyd lectures
- Suvrit Sra lectures
- Martin Jaggi lectures

Example 4

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

$\boxed{\nabla f = c}$

$d(c^T x + b) = \langle c, dx \rangle$

выпукла

$\nabla^2 f = 0_{n \times n}$

$x^T \nabla f x \geq 0$

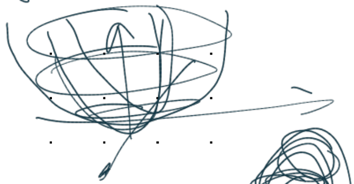
ч.т.д

$x^T 0 x \geq 0 = 0 = 0$

Example 5

$-A \succeq 0$

$f(x) = x^T A x$



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

$f = \langle x, Ax \rangle$

$g = df = \langle dx, Ax \rangle + \langle x, A\,dx \rangle =$

$= \langle (A + A^T)x, \underbrace{dx}_{dx_1} \rangle \qquad \nabla f = (A + A^T)x$

$dg = \langle (A + A^T)dx, dx_1 \rangle = \langle (A + A^T)dx_1, dx \rangle$

$x^T A x \geq 0$

$\langle x, Ax \rangle \geq 0$

$\langle A^T x, x \rangle \geq 0$

$\langle x, A^T x \rangle \geq 0$

$x^T A^T x \geq 0$

$\boxed{\nabla^2 f = \underset{\succeq 0}{A} + \underset{\succeq 0}{A^T}} \succeq 0$

Example 6

Show, that $f(x)$ is convex, using first and second order criteria, if $\boxed{f(x) = \sum\limits_{i=1}^n x_i^4.}$

поэлементное умн.

$\nabla f = 4x^{(3)}$

$\nabla^2 f = H = \begin{pmatrix} 12x_1^2 & & & \\ & 12x_2^2 & & \\ & & \ddots & \\ & & & 12x_n^2 \end{pmatrix}$

критерий Сильвестра

$\det \nabla^2 f \geq 0$

$\det \dots \geq 0$

$\vdots$

$\det \geq 0$

Example 7

Find the set of $x \in \mathbb{R}^n$, where the function $f(x) = \dfrac{-1}{2(1 + x^\top x)}$ is convex, strictly convex, strongly convex?

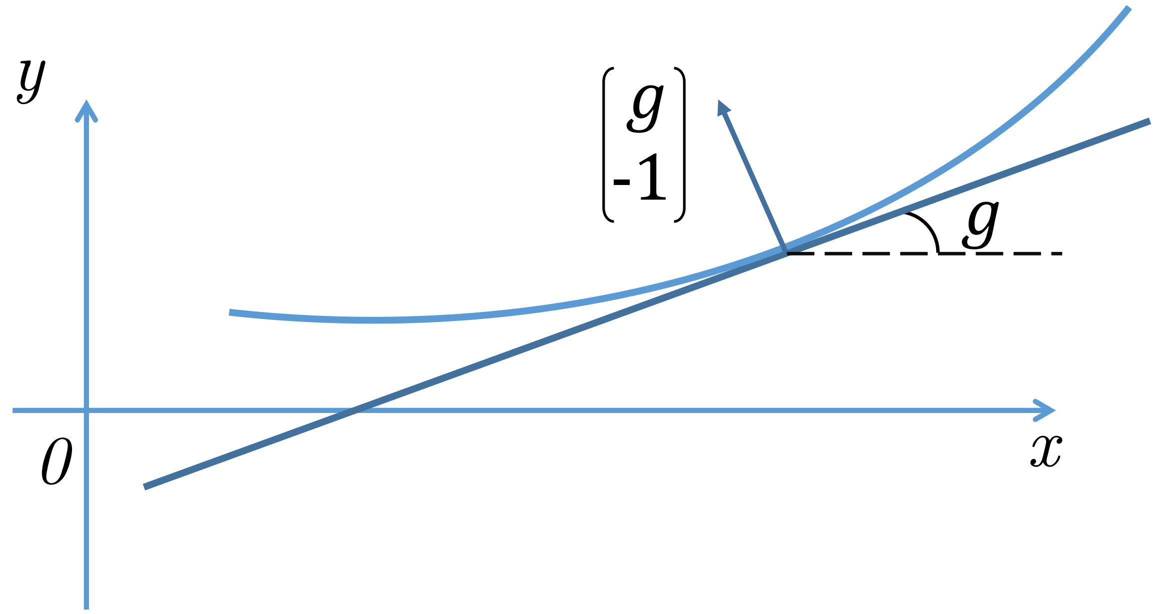
Subgradient and subdifferential

Motivation

Важным свойством непрерывной выпуклой функции $f(x)$ является то, что в выбранной точке x_0 для всех $x \in \text{dom } f$ выполнено неравенство:

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

для некоторого вектора g, то есть касательная к графику функции является *глобальной* оценкой снизу для функции.



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

Не хочется лишаться такого вкусного свойства.

Subgradient

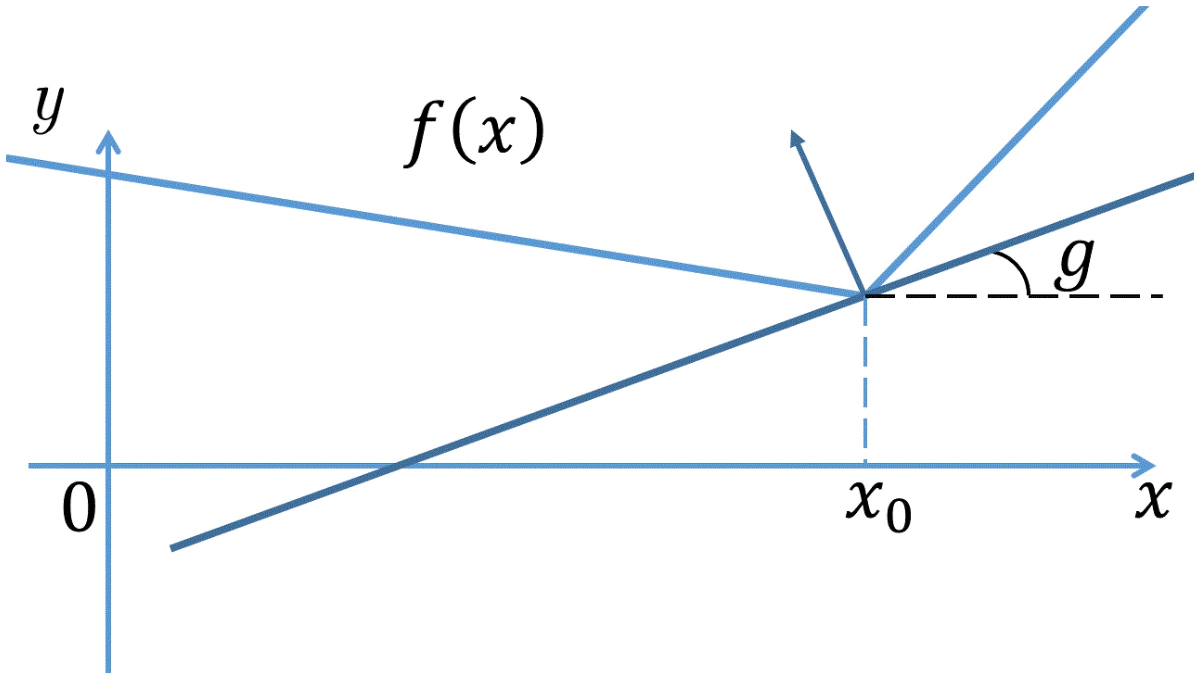
Вектор g называется **субградиентом** функции $f(x) : S \to \mathbb{R}$ в точке x_0, если $\forall x \in S$:

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

Subdifferential

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

- Если $x_0 \in \mathbf{ri} S$, то $\partial f(x_0)$ выпуклое компактное множество.
- Выпуклая функция $f(x)$ дифференцируема в точке $x_0 \iff \partial f(x_0) = \nabla f(x_0)$
- Если $\partial f(x_0) \neq \emptyset \quad \forall x_0 \in S$, то $f(x)$ - выпукла на S.



Moreau - Rockafellar theorem (subdifferential of a linear combination)

Пусть $f_i(x)$ - выпуклые функции на выпуклых множествах $S_i,\ i = \overline{1,n}$.
Тогда, если $\bigcap\limits_{i=1}^n \mathbf{ri} S_i \neq \emptyset$ то функция $f(x) = \sum\limits_{i=1}^n a_i f_i(x),\ a_i > 0$ имеет субдифференциал $\partial_S f(x)$ на множестве $S = \bigcap\limits_{i=1}^n S_i$ и

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

Dubovitsky - Milutin theorem (subdifferential of a point-wise maximum)

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

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

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

Chain rule for subdifferentials

Пусть $g_1, \ldots, g_m$ - выпуклые функции на открытом выпуклом множестве $S \subseteq \mathbb{R}^n$, $g = (g_1, \ldots, g_m)$ - образованная из них вектор - функция, φ - монотонно неубывающая выпуклая функция на открытом выпуклом множестве $U \subseteq \mathbb{R}^m$, причем $g(S) \subseteq U$. Тогда субдифференциал функции $f(x) = \varphi\left(g(x)\right)$ имеет вид:

$$
\partial f(x) = \bigcup_{p \in \partial \varphi(u)} \left(\sum_{i=1}^m p_i \partial g_i(x) \right),
$$

где $u = g(x)$

В частности, если функция φ дифференцируема в точке $u = g(x)$, то формула запишется так:

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

Subdifferential calculus

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

Examples

Концептуально, различают три способа решения задач на поиск субградиента:

- Теоремы Моро - Рокафеллара, композиции, максимума
- Геометрически
- По определению

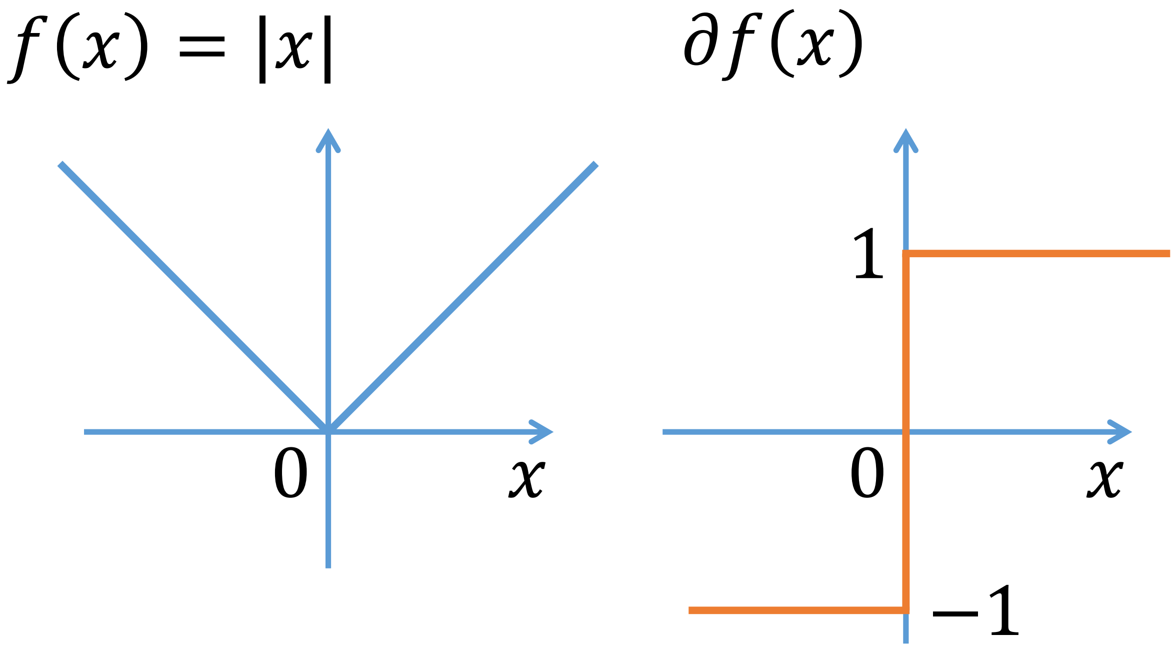
1

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

Решение:

Решить задачу можно либо геометрически (в каждой точке числовой прямой указать угловые коэффициенты прямых, глобально подпирающих функцию снизу), либо по теореме Моро - Рокафеллара, рассмотрев $f(x)$ как композицию выпуклых функций:

$$f(x) = \max\{-x, x\}$$



2

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

Решение:

Совершенно аналогично применяем теорему Моро - Рокафеллара, учитывая следующее:

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

Таким образом:

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

3

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

Решение:

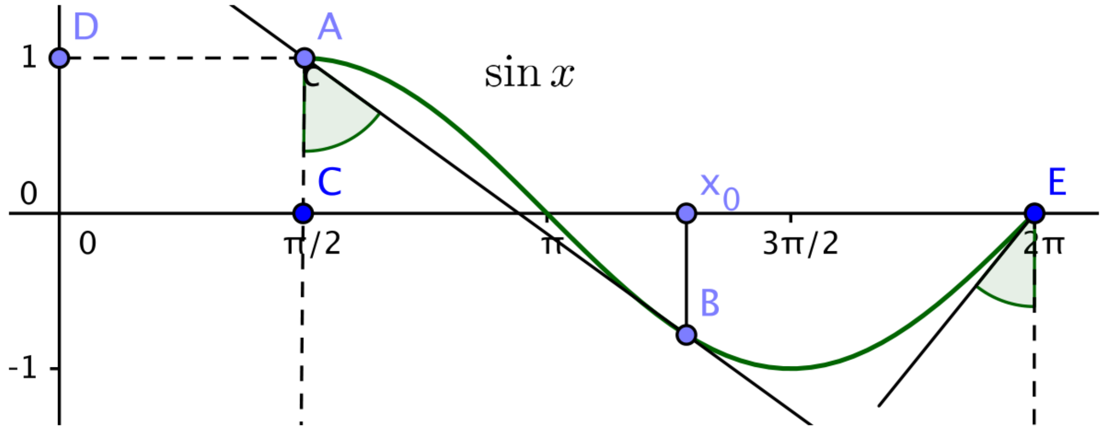
Согласно теореме о композиции (функция $\varphi(x) = x^q$ - дифференцируема), а $g(x) = \max(0, f_0(x))$ имеем: $\partial f(x) = q(g(x))^{q-1} \partial g(x)$

По теореме о поточечном максимуме:

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

4

Найти $\partial f(x)$, если $f(x) = \sin x, x \in [\pi/2; 2\pi]$



$$
\partial f_G(x) = \begin{cases} (-\infty, \cos x_0], & x = \pi/2; \\ \varnothing, & x \in (\pi/2, x_0); \\ \cos x, & x \in [x_0, 2\pi); \\ [1, +\infty), & x = 2\pi. \end{cases}
$$

5

Найти $\partial f(x)$, если $f(x) = |c_1^\top x| + |c_2^\top x|$

Решение: Пусть $f_1(x) = |c_1^\top x|$, а $f_2(x) = |c_2^\top x|$. Так как эти функции выпуклы, субдифференциал их суммы равен сумме субдифференциалов. Найдем каждый из них:

$$
\partial f_1(x) = \partial \left(\max\{c_1^\top x, -c_1^\top x\} \right) = \begin{cases} -c_1, & c_1^\top x < 0 \\ \mathbf{conv}(-c_1; c_1), & c_1^\top x = 0 \\ c_1, & c_1^\top x > 0 \end{cases}
$$
$$
\partial f_2(x) = \partial \left(\max\{c_2^\top x, -c_2^\top x\} \right) = \begin{cases} -c_2, & c_2^\top x < 0 \\ \mathbf{conv}(-c_2; c_2), & c_2^\top x = 0 \\ c_2, & c_2^\top x > 0 \end{cases}
$$

Далее интересными представляются лишь различные взаимные расположения векторов c_1 и c_2, рассмотрение которых предлагается читателю.

6

Найти $\partial f(x)$, если $f(x) = \|x\|_1$

Решение: По определению

$$
\|x\|_1 = |x_1| + |x_2| + \ldots + |x_n| = s_1 x_1 + s_2 x_2 + \ldots + s_n x_n
$$

Рассмотрим эту сумму как поточечный максимум линейных функций по x: $g(x) = s^\top x$, где $s_i = \{-1, 1\}$. Каждая такая функция однозначно определяется набором коэффициентов $\{s_i\}_{i=1}^n$.

Тогда по теореме Дубовицкого - Милютина, в каждой точке $\partial f = \mathbf{conv}\left(\bigcup\limits_{i \in I(x)} \partial g_i(x)\right)$

Заметим, что $\partial g(x) = \partial\left(\max\{s^\top x, -s^\top x\}\right) = \begin{cases} -s, & s^\top x < 0 \\ \mathbf{conv}(-s; s), & s^\top x = 0. \\ s, & s^\top x > 0 \end{cases}$

Причем, правило выбора "активной" функции поточечного максимума в каждой точке следующее:

- Если j-ая координата точки отрицательна, $s_i^j = -1$
- Если j-ая координата точки положительна, $s_i^j = 1$
- Если j-ая координата точки равна нулю, то подходят оба варианта коэффициентов и соответствующих им функций, а значит, необходимо включать субградиенты этих функций в объединение в теореме Дубовицкого - Милютина.

В итоге получаем ответ:

$$\partial f(x) = \left\{g \;:\; \|g\|_\infty \leq 1, \quad g^\top x = \|x\|_1\right\}$$

References

- Lecture Notes for ORIE 6300: Mathematical Programming I by Damek Davis