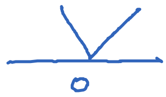
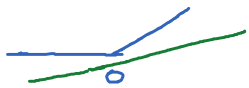
Non-Smooth Convex Functions

2022年3月8日 13:23

Many convex functions are not differentiable

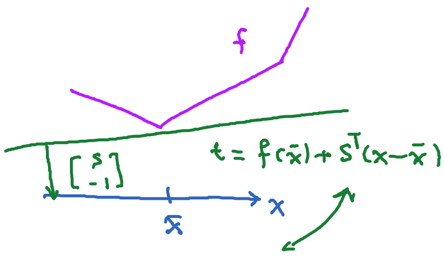
e.g. $x \mapsto |x|$ $\qquad$ $x \mapsto \max\{x, 0\}$ (ReLU)





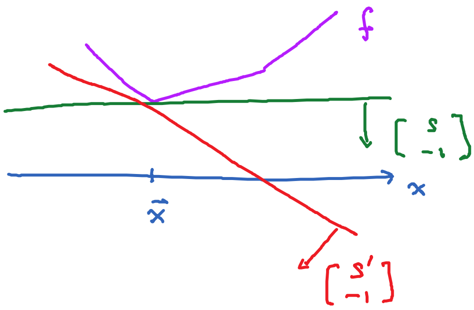
Definition: Let $f: \mathbb{R}^n \to \mathbb{R}$ be a function. A vector $s \in \mathbb{R}^n$ is called a <u>subgradient</u> of f at $\bar{x}$ if

$$f(x) \geq \underbrace{f(\bar{x}) + s^T(x-\bar{x})}_{\text{affine in } x} \quad \forall x \in \mathbb{R}^n$$



$$\begin{bmatrix} s \\ -1 \end{bmatrix}^T \begin{bmatrix} x \\ t \end{bmatrix} = s^T\bar{x} - f(\bar{x})$$

Note: There could be more than one subgradient at $\bar{x}$:



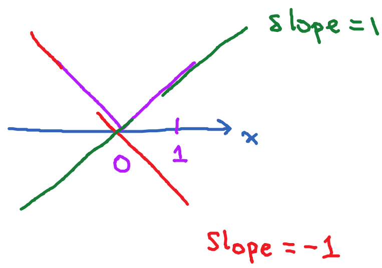
The set of all subgradients of f at $\bar{x}$ is the <u>subdifferential</u> of f at $\bar{x}$, denoted by

$$\partial f(\bar{x}) = \{ s : f(x) \geq f(\bar{x}) + s^T(x-\bar{x}), \forall x \}.$$

Example: $x \mapsto f(x) = |x|$. Take $\bar{x} = 0$. Then,

$$\partial f(0) = \{ s : |x| \geq 0 + sx, \forall x \} = [-1, 1]$$

$$|x| \geq sx \quad \begin{cases} x > 0 : & s \leq 1 \\ x < 0 : & s \geq -1 \end{cases}$$

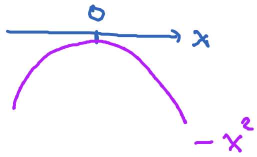


Take $\bar{x} = 1$. Then,

$$\partial f(1) = \{ s : |x| \geq 1 + s(x-1), \forall x \} = \{1\},$$

Take $x = 0$: $s \geq 1$
Take $x = 2$: $s \leq 1$ $\Bigg\} \Rightarrow s = 1$: $|x| \geq x, \ \forall x$ ✓

<u>Example</u>: $x \mapsto f(x) = -x^2$.



$\partial f(0) = \emptyset$ (Exercise)

<u>Theorem</u>: Let $f: \mathbb{R}^n \to \mathbb{R}$ be convex. Then, for any $x \in \mathbb{R}^n$, $\partial f(x)$ is non-empty, convex, compact. Moreover,

$$f \text{ is differentiable at } x \iff \partial f(x) = \{\nabla f(x)\}.$$

<u>Proposition</u>: Let $f: \mathbb{R}^n \to \mathbb{R}$ be any function. Consider

$$(P) \qquad \inf_{x \in \mathbb{R}^n} f(x).$$

Then, $\bar{x}$ is optimal for $(P) \iff 0 \in \partial f(\bar{x})$

<u>Proof</u>: $\bar{x}$ is optimal for (P)

$$\iff f(x) \geq f(\bar{x}) + 0^T(x - \bar{x}) \quad \forall x \in \mathbb{R}^n$$

$$\iff 0 \in \partial f(\bar{x})$$

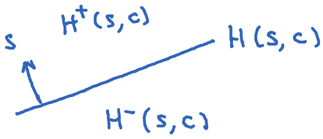
Linear Programming: Basic Objects

2022年3月8日 13:24

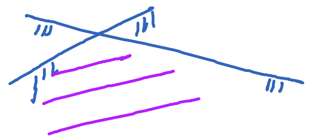
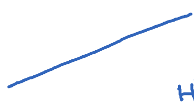
Fundamental objects

- linear functions : $x \mapsto b^T x$
- hyperplanes / halfspaces : $H(s,c) = \{x : s^T x = c\}$

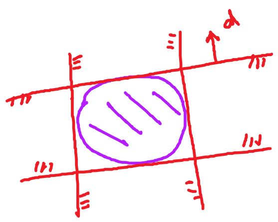
$$H^+(s,c) = \{x : s^T x \geq c\}$$
$$H^-(s,c) = \{x : s^T x \leq c\}$$



Definition: A polyhedron is an intersection of a finite number of halfspaces.


,


, $\mathbb{R}^n = \{x : \vec{0}^T x \geq 0\}$

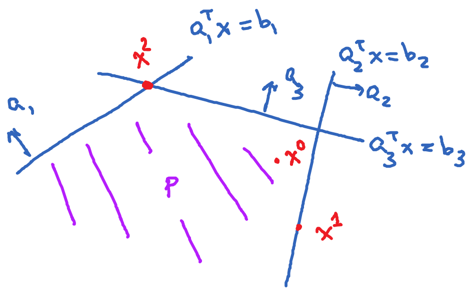


$$B(0,1) = \bigcap_{\substack{d \in \mathbb{R}^n \\ \|d\|_2 = 1}} \{x : d^T x \leq 1\}$$ not a polyhedron

Representation:

$$A = \begin{bmatrix} — a_1^T — \\ \vdots \\ — a_m^T — \end{bmatrix} \in \mathbb{R}^{m \times n}$$

$$P = \{x \in \mathbb{R}^n : a_i^T x \leq b_i \ ;\ i = 1, \cdots, m\} = \{x \in \mathbb{R}^n : Ax \leq b\}$$



$x^0, x^1, x^2 \in P$

1) $a_i^T x_0 < b_i \quad \forall i$ $\quad I(x^0) = \emptyset$

2) $a_1^T x^1 < b_1$, $a_3^T x^1 < b_3$, $\quad a_2^T x^1 = b_2$ (active constraint) $\quad I(x^1) = \{1\}$

3) $a_1^T x^2 = b_1$, $a_3^T x^2 = b_3$, $\quad a_2^T x^2 < b_2$

(active constraints) $\quad I(x^2) = \{1, 3\}$

Consider $\bar{x} \in P$. Define $I(\bar{x}) = \{i : a_i^T \bar{x} = b_i\}$ to be the active index set of $\bar{x}$.

Theorem: Let $\bar{x} \in P \subseteq \mathbb{R}^n$, $I(\bar{x})$ be its active index set. Then, the following are equivalent:

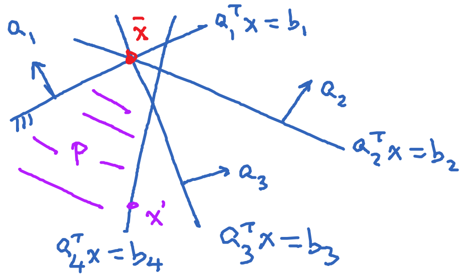
(1) $\exists$ n linearly independent vectors in $\{a_i : i \in I(\bar{x})\}$

(2) $\bar{x}$ is the unique solution to the linear system

$$a_i^T x = b_i \quad , \quad i \in I(\bar{x})$$

e.g.

$\mathbb{R}^2$:



$I(\bar{x}) = \{1,2,3\}$

$\{a_1, a_2, a_3\}$. any 2 are linearly independent

$$\begin{cases} a_1^T x = b_1 \\ a_2^T x = b_2 \\ a_3^T x = b_3 \end{cases} \longrightarrow \bar{x} \text{ is the unique solution}$$

$\Leftrightarrow \quad a_i^T \bar{x} = b_i \ , \ i \in I(\bar{x})$

$I(x') = \{4\}$

$\{a_4\} \leftarrow \nexists$ 2 linearly indep. vectors

x' is a solution to

$$a_i^T x = b_i \quad , \quad i \in I\{x'\},$$

but is not unique.