CS257 Linear and Convex Optimization

Lecture 4

Bo Jiang

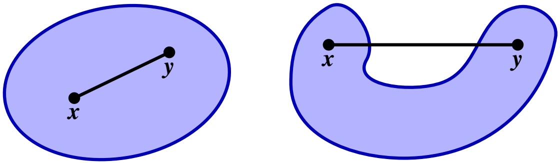
John Hopcroft Center for Computer Science
Shanghai Jiao Tong University

September 28, 2020

Recap: Convex Sets

A set $C \subset \mathbb{R}^n$ is convex if the line segment between any two points $\boldsymbol{x}, \boldsymbol{y} \in C$ lies entirely in C, i.e.

$$\boldsymbol{x} \in C, \boldsymbol{y} \in C, \theta \in [0, 1] \implies \theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} \in C$$



convex nonconvex

Properties.

- The intersection of convex sets is convex.
- The image of a convex set under an affine transformation is convex.
- The inverse image of a convex set under an affine transformation is convex.

Recap: Convex Sets

Convex combination. $\sum_{i=1}^m \theta_i \boldsymbol{x}_i$, where $\boldsymbol{\theta} \geq \boldsymbol{0}, \boldsymbol{1}^T\boldsymbol{\theta} = 1$

Convex hull of S

- smallest convex set containing S
- set of all convex combinations of elements of S

Examples of convex sets.

- $\emptyset$, $\mathbb{R}^n$, singleton (point), line, line segment, ray
- Hyperplane $P = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{w}^T\boldsymbol{x} = b\}$, $\boldsymbol{w} \in \mathbb{R}^n$, $b \in \mathbb{R}$
- Halfspace $H = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{w}^T\boldsymbol{x} \leq b\}$
- Affine space $S = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$
- Polyhedron $P = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$
- Norm ball $\bar{B}(\boldsymbol{x}_0, r) = \{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq r\}$
- Ellipsoid $\mathcal{E} = \{\boldsymbol{x}_0 + \boldsymbol{A}\boldsymbol{u} : \|\boldsymbol{u}\|_2 \leq 1\}$, $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{A} \succ \boldsymbol{O}$.
- Positive semidefinite matrices $\mathcal{S}_+^n = \{\boldsymbol{A} \in \mathbb{R}^{n\times n} : \boldsymbol{A} \succeq \boldsymbol{O}\}$
- Simplex $\Delta = \operatorname{conv}\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m\} = \{\sum_{i=0}^m \theta_i \boldsymbol{x}_i : \boldsymbol{\theta} \geq \boldsymbol{0}, \boldsymbol{1}^T\boldsymbol{\theta} = 1\}$

Contents

Convex Functions

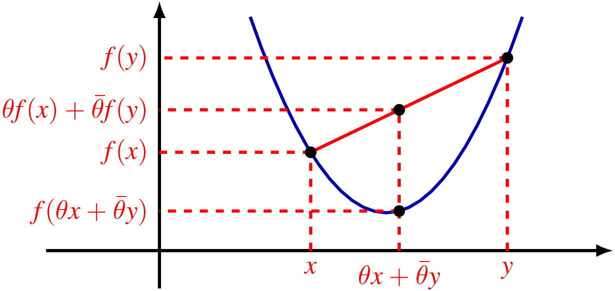
A function $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is convex if

1. its domain $\mathrm{dom} f = S$ is a convex set
2. for any $\boldsymbol{x}, \boldsymbol{y} \in S$ and $\theta \in [0, 1]$,

$$f(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$$

Note. Condition 1 guarantees $\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}$ is in the domain.

Geometrically, the line segment between $(\boldsymbol{x}, f(\boldsymbol{x}))$ and $(\boldsymbol{y}, f(\boldsymbol{y}))$ lies above the graph of f.



Convex Functions (cont'd)

A function $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is strictly convex if

1. its domain $\operatorname{dom} f = S$ is a convex set
2. for any $\boldsymbol{x} \neq \boldsymbol{y} \in S$ and $\theta \in (0, 1)$,

$$f(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) < \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$$

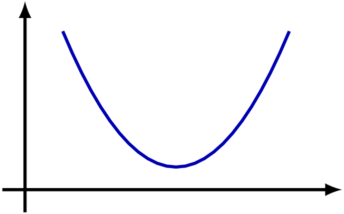
Proposition. Let f be convex. If $f(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$ for some $\theta = \theta_0 \in (0, 1)$, then it holds for any $\theta \in [0, 1]$, i.e. $g(\theta) = f(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y})$ is an affine function for $\theta \in [0, 1]$.

Strict convexity says the restriction of f to any line segment in S is not an affine function.

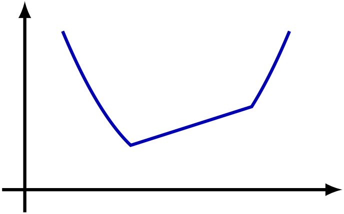
A function f is (strictly) concave if $-f$ is (strictly) convex.

An affine function $f(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{x} + b$ is both convex and concave, but not strictly convex or strictly concave.

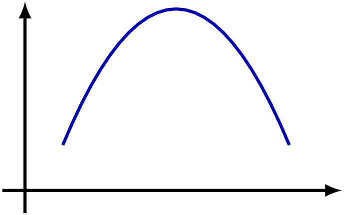
Convex Functions (cont'd)



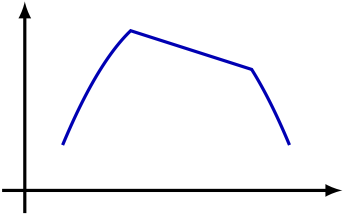
strictly convex function



convex function



strictly concave function



concave function

Examples

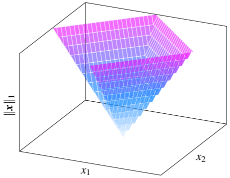
Example. Univariate functions

- $f(x) = e^{ax}$ ($a \in \mathbb{R}$) is convex, and strictly convex for $a \neq 0$
- $f(x) = \log x$ is strictly concave over $(0, \infty)$
- $f(x) = x^a$ is convex over $(0, \infty)$ for $a \geq 1$ or $a \leq 0$
- $f(x) = x^a$ is concave over $(0, \infty)$ for $0 \leq a \leq 1$

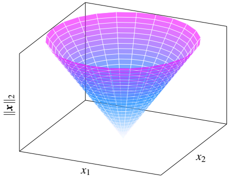
Example. Any norm $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is convex,

$$\|\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}\| \leq \|\theta \boldsymbol{x}\| + \|\bar{\theta} \boldsymbol{y}\| = \theta \|\boldsymbol{x}\| + \bar{\theta} \|\boldsymbol{y}\|$$

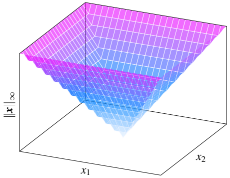
But not strictly convex (why?)



1-norm



2-norm



∞-norm

Restriction to Lines

Proposition. f is convex iff for any $\boldsymbol{x} \in \operatorname{dom} f$ and any direction $\boldsymbol{d}$, the function $g(t) = f(\boldsymbol{x} + t\boldsymbol{d})$ is convex on $\operatorname{dom} g = \{t : \boldsymbol{x} + t\boldsymbol{d} \in \operatorname{dom} f\}$.

Proof. "$\Rightarrow$". Assume f is convex. Fix an arbitrary $\boldsymbol{x} \in \operatorname{dom} f$ and direction $\boldsymbol{d}$. Need to show $g(t) = f(\boldsymbol{x} + t\boldsymbol{d})$ is convex.

Let $t_1, t_2 \in \operatorname{dom} g$, $\theta \in [0, 1]$. Let $\boldsymbol{x}_i = \boldsymbol{x} + t_i\boldsymbol{d}$, $\bar{t} = \theta t_1 + \bar{\theta} t_2$ and $\bar{\boldsymbol{x}} = \boldsymbol{x} + \bar{t}\boldsymbol{d}$.

1. Note $\bar{\boldsymbol{x}} = \boldsymbol{x} + (\theta t_1 + \bar{\theta} t_2)\boldsymbol{d} = \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2$
2. $t_i \in \operatorname{dom} g \implies \boldsymbol{x}_i \in \operatorname{dom} f$
3. $\operatorname{dom} f$ is convex $\implies \bar{\boldsymbol{x}} \in \operatorname{dom} f \implies \bar{t} \in \operatorname{dom} g \implies \operatorname{dom} g$ is convex
4. Since f is convex,

$$g(\bar{t}) = f(\bar{\boldsymbol{x}}) \leq \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2) = \theta g(t_1) + \bar{\theta} g(t_2)$$

   so g is convex.

Restriction to Lines (cont'd)

Proof (cont'd). "$\Leftarrow$". Assume $g(t) = f(\boldsymbol{x} + t\boldsymbol{d})$ is convex for any $\boldsymbol{x} \in \operatorname{dom} f$ and any direction $\boldsymbol{d}$. Need to show f is convex.

Fix $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$, $\theta \in [0, 1]$. Let $\boldsymbol{d} = \boldsymbol{x} - \boldsymbol{y}$, and $g(t) = f(\boldsymbol{y} + t\boldsymbol{d})$.

1. $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f \implies 1, 0 \in \operatorname{dom} g$
2. $\operatorname{dom} g$ is convex $\implies \theta \in \operatorname{dom} g \implies \boldsymbol{x} + \theta\boldsymbol{d} \in \operatorname{dom} f$
3. Since $\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} = \boldsymbol{y} + \theta\boldsymbol{d}$, $\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y} \in \operatorname{dom} f \implies \operatorname{dom} f$ is convex.
4. Since g is convex and $\theta = \theta \times 1 + \bar{\theta} \times 0$,

$$f(\theta\boldsymbol{x} + \bar{\theta}\boldsymbol{y}) = g(\theta) \leq \theta g(1) + \bar{\theta} g(0) = \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$$

so f is convex.

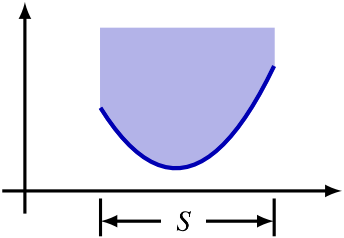
Epigraph

Recall the graph of $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is the set

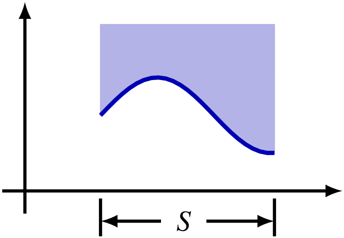
$$\{(\boldsymbol{x}, f(\boldsymbol{x})) \in \mathbb{R}^{n+1} : \boldsymbol{x} \in S\}$$

The epigraph[1] of f is

$$\operatorname{epi} f = \{(\boldsymbol{x}, y) \in \mathbb{R}^{n+1} : \boldsymbol{x} \in S, y \geq f(\boldsymbol{x})\}$$



convex



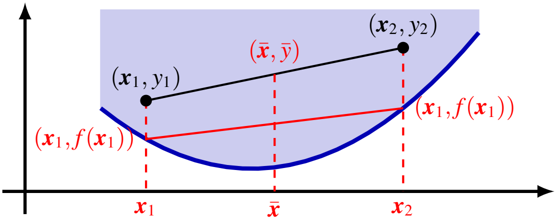
nonconvex

[1] The prefix epi- means “above”, “over”.

Epigraph (cont'd)

Theorem. $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is a convex function iff $\mathrm{epi} f$ is a convex set.

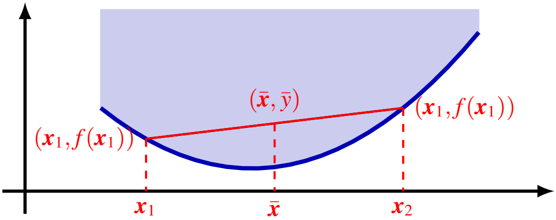
Proof. "$\Rightarrow$". Assume f is convex. Let $(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2) \in \mathrm{epi} f$, $\theta \in [0, 1]$. Need to show $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \triangleq (\theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2, \theta y_1 + \bar{\theta} y_2) \in \mathrm{epi} f$.



1. f convex $\implies \bar{\boldsymbol{x}} \in S$ and $f(\bar{\boldsymbol{x}}) \le \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2)$
2. $(\boldsymbol{x}_i, y_i) \in \mathrm{epi} f \implies f(\boldsymbol{x}_i) \le y_i \implies \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2) \le \theta y_1 + \bar{\theta} y_2 = \bar{y}$
3. By 1 and 2, $\bar{\boldsymbol{x}} \in S$ and $f(\bar{\boldsymbol{x}}) \le \bar{y} \implies (\bar{\boldsymbol{x}}, \bar{y}) \in \mathrm{epi} f$

Epigraph (cont'd)

Proof (cont'd). "$\Leftarrow$". Assume $\mathrm{epi} f$ is convex. Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, $\theta \in [0, 1]$.
Need to show $\bar{\boldsymbol{x}} \triangleq \theta \boldsymbol{x}_1 + \bar{\theta} \boldsymbol{x}_2 \in S$ and $f(\bar{\boldsymbol{x}}) \leq \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2) \triangleq \bar{y}$.



1. $f(\boldsymbol{x}_i) \leq f(\boldsymbol{x}_i) \implies (\boldsymbol{x}_i, f(\boldsymbol{x}_i)) \in \mathrm{epi} f$ by definition
2. $\mathrm{epi} f$ convex $\implies (\bar{\boldsymbol{x}}, \bar{y}) = \theta(\boldsymbol{x}_1, f(\boldsymbol{x}_1)) + \bar{\theta}(\boldsymbol{x}_2, f(\boldsymbol{x}_2)) \in \mathrm{epi} f$
3. $\bar{\boldsymbol{x}} \in S$, $f(\bar{\boldsymbol{x}}) \leq \bar{y} = \theta f(\boldsymbol{x}_1) + \bar{\theta} f(\boldsymbol{x}_2)$ by definition of $\mathrm{epi} f$

Jensen's Inequality

For convex function f, $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$, $\theta \in [0, 1]$

$$f(\theta \boldsymbol{x} + \bar{\theta} \boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + \bar{\theta} f(\boldsymbol{y})$$

More generally, for $\boldsymbol{x}_i \in \operatorname{dom} f$, $\theta_i \geq 0$, and $\sum_{i=1}^m \theta_i = 1$,

$$f\left(\sum_{i=1}^m \theta_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^m \theta_i f(\boldsymbol{x}_i)$$

Example. $f(x) = x^2$ is convex over $\mathbb{R}$.

$$\left(\sum_{i=1}^n \frac{1}{n} x_i\right)^2 \leq \sum_{i=1}^n \frac{1}{n} x_i^2 \implies \frac{1}{n} \sum_{i=1}^n x_i \leq \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$$

Example. $f(x) = \log x$ is concave over $(0, \infty)$. For $x_i > 0$,

$$\log\left(\sum_{i=1}^n \frac{1}{n} x_i\right) \geq \sum_{i=1}^n \frac{1}{n} \log x_i \implies \frac{1}{n} \sum_{i=1}^n x_i \geq \sqrt[n]{\prod_{i=1}^n x_i}$$

Hölder's Inequality

Let $p, q \in (1, \infty)$ be conjugate exponents, i.e. $p^{-1} + q^{-1} = 1$. For $\boldsymbol{x} = (x_1, \ldots, x_n)^T$, $\boldsymbol{y} = (y_1, \ldots, y_n)^T$, Hölder's inequality holds,

$$\sum_{i=1}^{n} |x_i y_i| \leq \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q$$

Proof. Assume $\boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{y} \neq \boldsymbol{0}$; otherwise trivial. Let $\tilde{\boldsymbol{x}} = \boldsymbol{x}/\|\boldsymbol{x}\|_p$ and $\tilde{\boldsymbol{y}} = \boldsymbol{y}/\|\boldsymbol{y}\|_q$. The above inequality is equivalent to $\sum_{i=1}^{n} |\tilde{x}_i \tilde{y}_i| \leq 1$.

1. Show $x^{\frac{1}{p}} y^{\frac{1}{q}} \leq \frac{1}{p}x + \frac{1}{q}y$ for $x, y \geq 0$.
   - trivial if $xy = 0$
   - if $xy > 0$, $\log x$ is concave $\implies$ $\log\left(\frac{1}{p}x + \frac{1}{q}y\right) \geq \frac{1}{p}\log x + \frac{1}{q}\log y$
2. Let $x = |\tilde{x}_i|^p$ and $y = |\tilde{y}_i|^q$ in the inequality in 1,

$$|\tilde{x}_i| \cdot |\tilde{y}_i| \leq p^{-1}|\tilde{x}_i|^p + q^{-1}|\tilde{y}_i|^q$$

3. Sum over i and note $\|\tilde{\boldsymbol{x}}\|_p = \|\tilde{\boldsymbol{y}}\|_q = 1$,

$$\sum_{i=1}^{n} |\tilde{x}_i \tilde{y}_i| \leq \frac{1}{p}\|\tilde{\boldsymbol{x}}\|_p^p + \frac{1}{q}\|\tilde{\boldsymbol{y}}\|_q^q = \frac{1}{p} + \frac{1}{q} = 1$$

Minkowski's Inequality

For $1 < p < \infty$,

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p \le \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$$

Proof. Only need to consider case $\|\boldsymbol{x}+\boldsymbol{y}\|_p > 0$.

- $\|\boldsymbol{x}+\boldsymbol{y}\|_p^p = \sum_i |x_i+y_i|^p \le \sum_i |x_i| \cdot |x_i+y_i|^{p-1} + \sum_i |y_i| \cdot |x_i+y_i|^{p-1}$
- Let $p^{-1}+q^{-1}=1$. By Hölder, and note $(p-1)q=p$,

$$\sum_i |x_i| \cdot |x_i+y_i|^{p-1} \le \|\boldsymbol{x}\|_p \left(\sum_i |x_i+y_i|^{(p-1)q}\right)^{1/q} = \|\boldsymbol{x}\|_p \|\boldsymbol{x}+\boldsymbol{y}\|_p^{p/q}$$

- Interchange x and y, $\sum_i |y_i| \cdot |x_i+y_i|^{p-1} \le \|\boldsymbol{y}\|_p \|\boldsymbol{x}+\boldsymbol{y}\|_p^{p/q}$
- Combining above inequalities,

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p^p \le (\|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p)\|\boldsymbol{x}+\boldsymbol{y}\|_p^{p/q}$$

- Cancel $\|\boldsymbol{x}+\boldsymbol{y}\|_p^{p/q}$ and note $p-p/q=1$.