Lecture 3: Convexity II and ~~Optimization terminologies~~

Dabeen Lee

Industrial and Systems Engineering, KAIST

IE 539: Convex Optimization

September 6, 2022

Outline

Convex sets

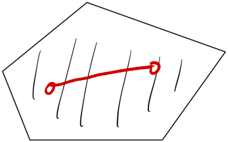
- Convex functions.
- First-order, second-order characterizations of convex functions,
- Operations preserving convexity,
- ~~Optimization terminologies~~

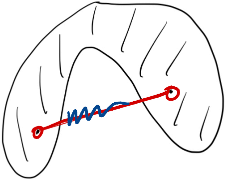
Convex set

A set $X \subseteq \mathbb{R}^d$ is <u>convex</u> if for any $x, y \in X$ and $\lambda \in [0,1]$,

$$\lambda x + (1-\lambda) y \in X$$

In words, the line segment connecting $x, y \in X$ is contained in X.





Convex function

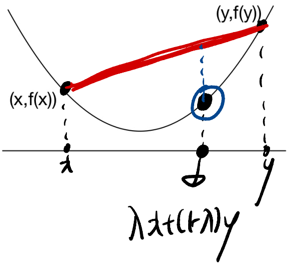
$$\mathrm{dom}(f) = \{x \in \mathbb{R}^d : f(x) < \infty\}.$$

$\mathbb{R}^d$

A function $f : \mathbb{R}^d \to \mathbb{R}$ is convex if the domain $\cancel{\mathrm{dom}}(f)$ is convex and
for any $x, y \in \mathrm{dom}(f)$ and $\lambda \in [0,1]$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

In ~~words~~,
picture



Concave, strictly convex, strongly convex functions

$f(x) = x^2 \to$ convex

$-x^2 \to$ concave.

We say that $f : \mathbb{R}^d \to \mathbb{R}$ is concave if

$-f$ is convex

A function $f : \mathbb{R}^d \to \mathbb{R}$ is strictly convex if $\mathrm{dom}(f)$ is convex and

for any $x, y \in \mathrm{dom}(f)$, $\lambda \in (0,1)$

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)$$

A function $f : \mathbb{R}^d \to \mathbb{R}$ is strongly convex if $\mathrm{dom}(f)$ is convex and

with respect to $\|\cdot\|_2$.

if $f(x) - \alpha \cdot \|x\|_2^2$ is convex.

Exercise: Show that Strongly convexity implies strict convexity.

Examples

Univariate functions $f: \mathbb{R} \to \mathbb{R}$.

- Exponential function

  $f(x) = e^{ax}$ for any $a \in \mathbb{R}$.

- Power function $f(x) = x^a$

  if $2 \geq a \geq 1$ $f(x)$ over $x \in \mathbb{R}$ is convex

  $\sqrt{x}$ ← if $0 \leq a < 1$ $f(x)$ over $x \in \mathbb{R}_+$ is concave

  if $a < 0$ $f(x)$ over $x \in \mathbb{R}_{++}$ is convex.

- Logarithm

  $f(x) = \log x$ $(x \in \mathbb{R}_{++})$

  is concave → $-\log x$ is convex

- Negative entropy

  $-x \log x \to f(x) = x \log(x)$

Examples

Linear function $f(x) = a^T x + b$

$a \in \mathbb{R}^d$. $b \in \mathbb{R}$

$\|Ax\|_2^2 \geq 0$.

Quadratic function

$f(x) = \frac{1}{2} x^T A x + b^T x + c$ is convex if and only if $A \succeq 0$ positive semidefinite

$A \in \mathbb{R}^{d \times d}$. $b \in \mathbb{R}^d$ $c \in \mathbb{R}$.

Least squares loss

$$f(x) = \| Ax - b\|_2^2 = (Ax-b)^T \cdot (Ax-b)$$

$$= \underline{x^T A^T A x} - \underline{2b^T A x} + \underline{b^T b}$$

$A^T A$ is PSD ? $x^T A^T A x \overset{?}{\geq} 0$ for any $x \in \mathbb{R}$

Norm

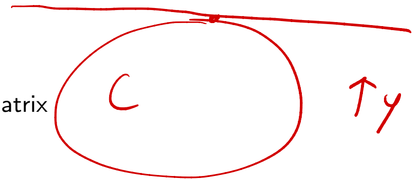
$f(x) = \|x\|$

$$\| \lambda x + (1-\lambda) y \| \leq \|\lambda x\| + \|(1-\lambda) y\| \leq \lambda \cdot \|x\| + (1-\lambda)\|y\|.$$

Examples

★ Maximum eigenvalue of a symmetric matrix

M



Indicator function of a convex set C.

$$I_C(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{if } x \notin C \end{cases}$$

Support function

$$I_C^*(y) = \sup_{x \in C} \langle y, x \rangle .$$

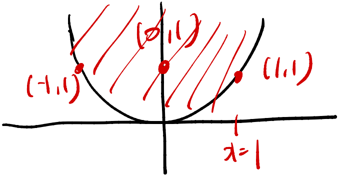
Conjugate function

$$f^*(y) = \sup \{ \langle y, x \rangle - f(x) \}$$

$I_C^*(y)$ $\qquad$ $-I_C(x)$

Epigraph

$f(\lambda) = \lambda^2$



The epigraph of a function $f : \mathbb{R}^d \to \mathbb{R}$ is defined as

$$\{ (\lambda, t) \in \mathrm{dom}(f) \times \mathbb{R} : f(\lambda) \le t \}.$$

Exercise: prove that f is convex if and only if its epigraph is convex

$\Leftarrow)$ $(\lambda_1, t_1), (\lambda_2, t_2) \in \mathrm{Epi}(f)$ $\lambda \in \mathbb{R}$.

$(\lambda \lambda_1 + (1-\lambda)\lambda_2, \lambda t_1 + (1-\lambda) t_2) \overset{?}{\in} \mathrm{Epi}(f)$

Example: norm cone $=$ epigraph of the norm

$f(\lambda\lambda_1 + (1-\lambda)\lambda_2)$

$\le \lambda f(\lambda_1) + (1-\lambda) f(\lambda_2)$

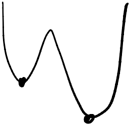
$\le \lambda t_1 + (1-\lambda) t_2$

Given a norm $\|\cdot\|$

the norm cone with respect to $\|\cdot\|$

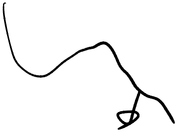
: $\boxed{\{ (\lambda, t) \in \mathbb{R}^d \times \mathbb{R} : \|\lambda\| \le t \}}$

Level set, convex level sets of a nonconvex function

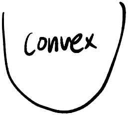


A level set of a function $f : \mathbb{R}^d \to \mathbb{R}$ is defined as

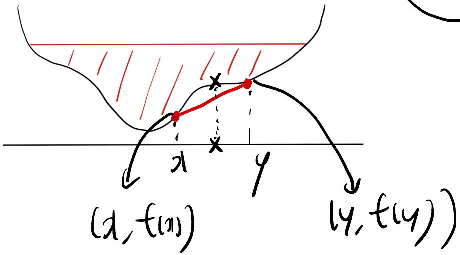
$$\{x \in \mathrm{dom}(f) : \; f(x) \leq \alpha\}$$

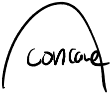


non convex
non concave



Level sets of a nonconvex set can be convex:





First-order characterization I

Theorem

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function. Then f is convex if and only if dom(f) is convex and

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

for all $x, y \in$ dom(f).

Proof

$(\Leftarrow)$ $x, y \in \mathrm{dom}(f)$. $\lambda \in [0,1]$

$z = \lambda x + (1-\lambda) y$

$f(x) \geq f(z) + \nabla f(z)^T (x - z)$ $\times \lambda$

$f(y) \geq f(z) + \nabla f(z)^T (y - z)$ $\times (1-\lambda)$

$\lambda f(x) + (1-\lambda) f(y) \geq f(z) + \nabla f(z)^T (\lambda x + (1-\lambda) y - z)$ 0

$= f(\lambda x + (1-\lambda) y)$

First-order characterization I

Theorem

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function. Then f is convex if and only if dom(f) is convex and

$$f(y) \geq f(x) + \underbrace{\nabla f(x)}_{f'(\cdot)}{}^\top (y - x)$$

for all $x, y \in$ dom(f).

Proof ($\Rightarrow$) $d = 1$

$$\boxed{g(\lambda) := f((1-\lambda)x + \lambda y)} \leq (1-\lambda) f(x) + \lambda f(y)$$

$$= f(x + \lambda(y-x))$$

$$\rightarrow \lim_{\lambda \to 0^+} \frac{f(x+\lambda(y-x)) - f(x)}{\lambda} \leq (f(y) - f(x))$$

$$\| \quad (y-x) \cdot f'(x)$$

First-order characterization I

Theorem

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function. Then f is convex if and only if $\mathrm{dom}(f)$ is convex and

$$\boxed{f(y) \geq f(x) + \nabla f(x)^\top (y - x)}$$

for all $x, y \in \mathrm{dom}(f)$.

Proof ($\Rightarrow$) general d.

If f is convex, g: convex $\Rightarrow$ we can use the $d=1$ case.

$$\underset{f(y)}{\underset{\parallel}{g(1)}} \geq \underset{f(x)}{\underset{\parallel}{g(0)}} + \underset{(y-x)^\top \nabla f(x)}{\underset{\parallel}{g'(0)}} \cdot (1-0)$$

$$g(\lambda) = f((1-\lambda)x + \lambda y) = f(x + \lambda \cdot (y-x))$$

$$g'(\lambda) = (y-x)^\top \nabla f(x + \lambda (y-x))$$

First-order characterization I

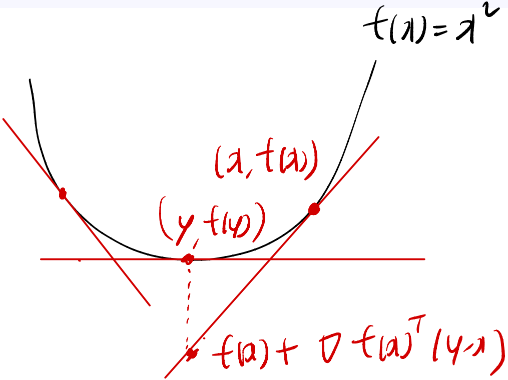
Theorem

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function. Then f is convex if and only if dom(f) is convex and

$$\boxed{f(y) \geq f(x) + \nabla f(x)^\top (y - x)}$$

for all $x, y \in$ dom(f).

~~Proof~~



First-order characterization II

Theorem

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function. Then f is convex if and only if $\mathrm{dom}(f)$ is convex and

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

for all $x, y \in \mathrm{dom}(f)$.

& monotone property

Proof

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

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

$(\Leftarrow)$ $g(\lambda) = f(x + \lambda(y-x)) \to g'(\lambda) = (y-x)^T \nabla f(x+\lambda(y-x))$

$g(1) - g(0) = g(\lambda)\Big]_0^1 = \int_0^1 g'(\lambda)\, d\lambda$

$= \int_0^1 (y-x)^T \nabla f(x + \lambda(y-x))\, d\lambda$

Theorem

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function. Then f is convex if and only if dom(f) is convex and

$$\boxed{\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0}$$

for all $x, y \in$ dom(f).

Proof

$$\underbrace{g(1)}_{f(y)} - \underbrace{g(0)}_{f(x)} = g(\lambda)\Big]_0^1 = \int_0^1 g'(\lambda)\, d\lambda$$

$$= \int_0^1 (y-x)^T \nabla f(x + \lambda(y-x))\, d\lambda$$

$$\boxed{f(y) - f(x) - \nabla f(x)^T (y-x)}$$

$$= \int_0^1 (y-x)^T \left(\nabla f(x + \lambda(y-x)) - \nabla f(x)\right) d\lambda$$

$$= \int_0^1 \frac{1}{\lambda} \langle \nabla f(x+\lambda(y-x)) - \nabla f(x), \lambda(y-x) \rangle\, d\lambda \geq 0$$

Second-order characterization

Theorem

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice differentiable function[1]*. Then f is convex if and only if dom(f) is convex and*

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

for all $x \in$ dom(f).

Proof

Example : Quadratic function.

$$f(x) = \frac{1}{2} x^T A x + b^T x + c.$$

$$\nabla^2 f(x) = \frac{1}{2}(A + A^T) \succeq 0$$

$$\Leftrightarrow \quad A \succeq 0$$

[1] $\nabla^2 f$ exists at any point in dom(f), and dom(f) is open.