Convex function

Definition

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

$$\left(\begin{array}{l} f : \mathbb{R}^n \to \mathbb{R} \\ \mathrm{dom}\, f \text{ is a convex set} \\ x, y \in \mathrm{dom}\, f \\ 0 \leq \theta \leq 1 \end{array} \right)$$

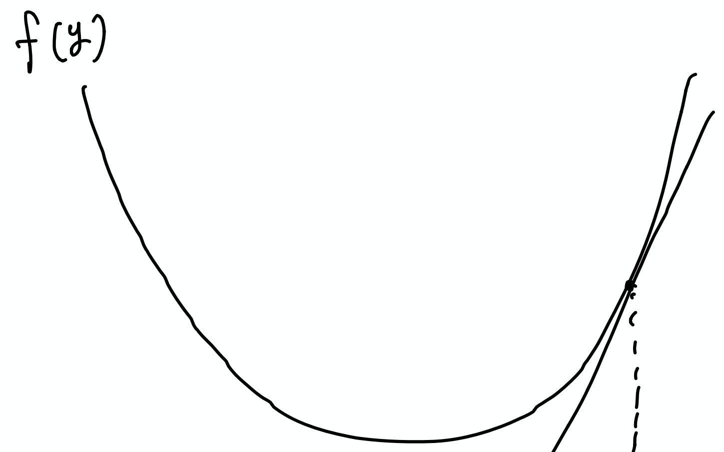
First-order conditions

Suppose f is differentiable.

Then f is convex if and only if

$\mathrm{dom}\, f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad \text{for } \forall x, \forall y \in \mathrm{dom}\, f$$



Second-order conditions

Suppose f is twice differentiable.

Then f is convex if and only if $\mathbf{dom} f$ is convex and

$$\nabla^2 f \succeq 0$$

Examples of convex functions

- e^{ax} is convex on $\mathbb{R}$ for $\forall a \in \mathbb{R}$
- x^a is

  convex when $a \le 0$, $a \ge 1$

  concave when $0 \le a \le 1$

  on $\mathbb{R}_{++}$
- $|x|^p$ for $p \ge 1$ is convex on $\mathbb{R}$
- $\log x$ is concave on $\mathbb{R}_{++}$
- $x \log x$ is convex on $\mathbb{R}_+$ (0 for $x=0$)

- Every norm is convex

- $f(x) = \max\{x_1 \cdots x_n\}$ is convex on $\mathbb{R}^n$

- $f(x, y) = x^2/y$ is convex

- $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$ is convex on $\mathbb{R}^n$

  differentiable approximation of max function

- $f(x) = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$ is concave on

  Geometric mean $\quad \mathrm{dom} f = \mathbb{R}^n_{++}$

- $f(X) = \log \det X$ is concave on

  $\mathrm{dom} f = S^n_{++}$

- Sublevel sets of a convex function

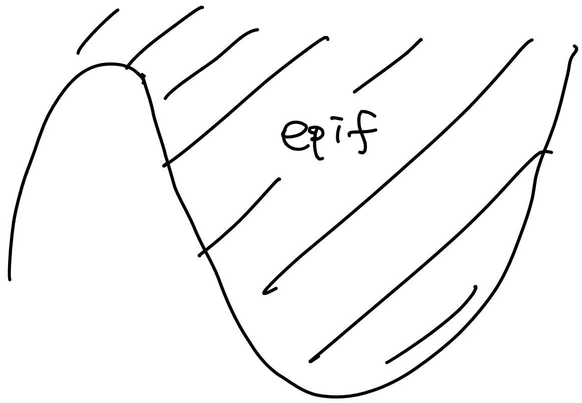
  $$C_\alpha = \{x \in \mathrm{dom} f \mid f(x) \le \alpha\}$$

  is convex

- epigraph

$$\mathrm{epi}\, f = \{(x,t) \mid x \in \mathrm{dom}\, f,\ f(x) \le t\}$$

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



※ 'epi' means 'above'

• Matrix fractional function

$$f(x, Y) = x^T Y^{-1} x$$

is convex on $\mathrm{dom}\, f = \mathbb{R}^n \times S^n_{++}$

generalization of $f(x,y) = \frac{x^2}{y}$

Jensen's inequality

$$f(\theta x + (1-\theta) y) \le \theta f(x) + (1-\theta) f(y)$$

↓ induction

$$f(\theta_1 x_1 + \cdots + \theta_k x_k) \le \theta_1 f(x_1) + \cdots + \theta_k f(x_k)$$

$(\theta_1 + \cdots + \theta_k = 1)$

$$\Leftrightarrow \; f\left(\sum_i \theta_i x_i\right) \le \sum_i \theta_i f(x_i)$$

$\downarrow$ extend to infinite sums

$$f\left(\int_S p(x)\, x \, dx\right) \le \int_S p(x) f(x) \, dx$$

$$f(\mathbb{E}[x]) \le \mathbb{E}[f(x)]$$

$\Downarrow$

Many famous inequalities can be derived by applying Jensen's inequality to some appropriate convex function

Examples

- arithmetic-geometric mean inequality

$$\sqrt{ab} \le \frac{a+b}{2}$$

$$f(\theta x + (1-\theta) y) \le \theta f(x) + (1-\theta) f(y)$$

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

$x = a$

$y = b$

$\theta = \frac{1}{2}$

$$-\log \frac{a+b}{2} \le -\frac{1}{2}\log a - \frac{1}{2}\log b$$

$$\Leftrightarrow \quad \frac{a+b}{2} \ge \sqrt{ab}$$

III Operations that preserve convexity.

① Nonnegative weighted sums

- $f = w_1 f_1 + \cdots + w_m f_m$ is convex

  $(w_1 \ge 0, \ \ldots, \ w_m \ge 0)$

$\Rightarrow$ if $f(x, y)$ is convex in x for each $y \in A$, and $w(y) \ge 0$, then

- $g(x) = \int_A w(y) f(x, y)\, dy$ is convex on y

② Composition with an affine mapping

Suppose $f : \mathbb{R}^n \to \mathbb{R}$, $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$, then

convex

$$g(x) = f(Ax + b) \text{ is convex}$$

③ Pointwise maximum and supremum

If $f_1 \cdots f_m$ are convex, then

$$f(x) = \max\{f_1(x), \cdots f_m(x)\} \text{ is convex}$$

$\Rightarrow$ if $f(x, y)$ is convex in x for each $y \in A$,

$$g(x) = \sup_{y \in A} f(x, y) \text{ is convex in } x$$

④ Scalar composition

$$f(x) = h(g(x))$$

$$\begin{pmatrix} g : \mathbb{R}^n \to \mathbb{R} \\ h : \mathbb{R} \to \mathbb{R} \\ f : \mathbb{R} \to \mathbb{R} \end{pmatrix}$$

$$f'(x) = h'(g(x))\, g'(x)$$

$$f''(x) = h''(g(x))\, g'(x)^2 + h'(g(x))\, g''(x)$$

	g: convex $g''(x) \geq 0$	g: concave $g''(x) \leq 0$
h: convex $h''(x) \geq 0$	convex (if $h'(x) \geq 0$)	convex (if $h'(x) \leq 0$)
h: concave $h''(x) \leq 0$	concave (if $h'(x) \leq 0$)	concave (if $h'(x) \geq 0$)

⑤ Minimization

If f is convex in (x, y), and C is a nonempty convex set, then

$$g(x) = \inf_{y \in C} f(x, y)$$

is convex in x, provided $g(x) > -\infty$ for all x.

⑥ Perspective of a function

If $f: \mathbb{R}^n \to \mathbb{R}$ is convex, then

$$g(x, t) = t f\left(\frac{x}{t}\right)$$

$$\operatorname{dom} g = \left\{ (x, t) \;\middle|\; \frac{x}{t} \in \operatorname{dom} f,\ t > 0 \right\}$$

is convex.

$\because$ $(x, t, s) \in \operatorname{epi} g$

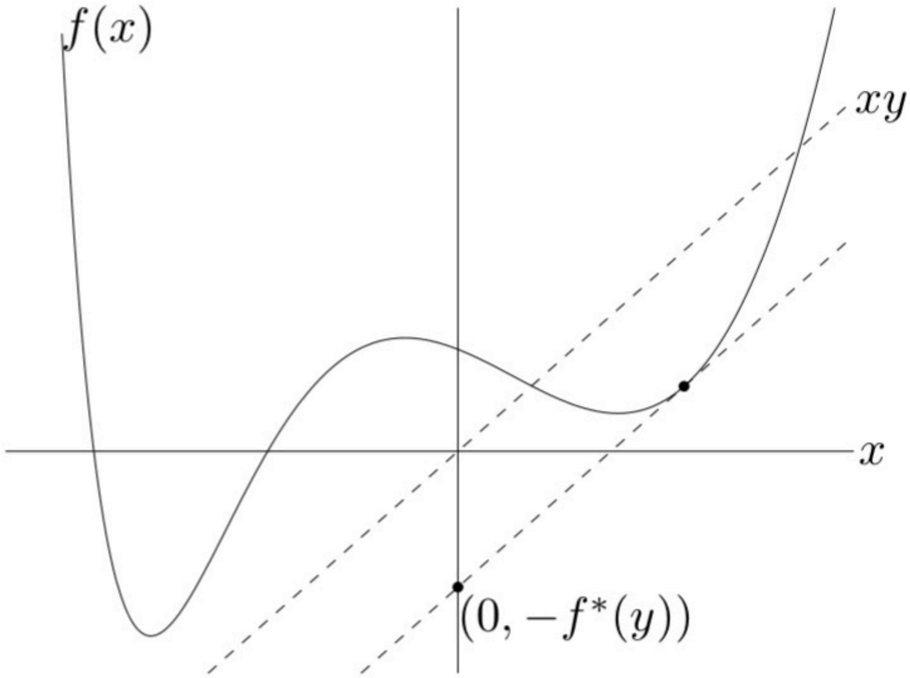
$\iff$ $t f\left(\frac{x}{t}\right) \le s$

$\iff$ $f\left(\frac{x}{t}\right) \le \frac{s}{t}$

$\iff$ $\left(\frac{x}{t}, \frac{s}{t}\right) \in \operatorname{epi} f$ $\blacksquare$

f^* is convex whether or not f is convex !!!

(1) Conjugate function



Let $f: \mathbb{R}^n \to \mathbb{R}$, The function $f^*: \mathbb{R}^n \to \mathbb{R}$, defined as

$$f^*(y) = \sup_{x \in \text{dom} f} \left(y^T x - f(x) \right)$$

is called the conjugate of the function of f.

(2) Examples of conjugate functions.

(1) Affine function

$$f(x) = ax + b. \qquad (\mathbb{R} \to \mathbb{R})$$

$$f^*(y) = \sup_{x \in \mathbb{R}} (yx - ax - b)$$

$$= -b \quad (\text{if } y = a) \qquad (\mathbb{R} \to \mathbb{R})$$

② Negative logarithm

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

$$f^*(y) = \sup_{x \in \mathbb{R}_{++}} (yx + \log x)$$

$$= -\log(-y) - 1 \quad (y < 0)$$

③ Exponential

$$f(x) = e^x$$

$$f^*(y) = \sup_{x \in \mathbb{R}} (yx - e^x)$$

$$= y \log y - y \quad (y > 0)$$

④ Negative entropy

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

$$f^*(y) = \sup_{x \in \mathbb{R}_{++}} (yx - x\log x)$$

$$= e^{y-1}$$

⑤ Inverse

$$f(x) = \frac{1}{x} \quad (x > 0)$$

$$f^*(y) = \sup_{x \in \mathbb{R}_{++}} \left(yx - \frac{1}{x} \right)$$

$$= -2(-y)^{\frac{1}{2}} \quad (y < 0)$$

⑥ Strictly convex quadratic function

$$f(x) = \frac{1}{2} x^T Q x \quad (Q \in S^n_{++})$$

$$f^*(y) = \sup \left(y^T x - \frac{1}{2} x^T Q x \right)$$

$$= \frac{1}{2} y^T Q^{-1} y$$

⑭ Fenchel's inequality

$$f^*(y) = \sup_x \left(y^T x - f(x) \right)$$

$$\Leftrightarrow \quad f^*(y) \geq y^T x - f(x)$$

$$\Leftrightarrow \quad \underbrace{f(x) + f^*(y) \geq x^T y}_{\text{Fenchel's inequality}}$$