Convex Functions (I)

Lijun Zhang
zlj@nju.edu.cn
http://cs.nju.edu.cn/zlj

Outline

- Basic Properties
  - Definition
  - First-order Conditions, Second-order Conditions
  - Jensen's inequality and extensions
  - Epigraph
- Operations That Preserve Convexity
  - Nonnegative Weighted Sums
  - Composition with an affine mapping
  - Pointwise maximum and supremum
  - Composition
  - Minimization
  - Perspective of a function
- Summary

Outline

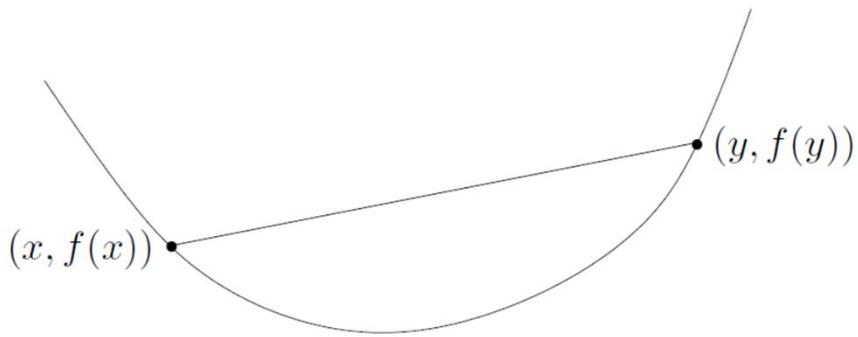
- Basic Properties
  - Definition
  - First-order Conditions, Second-order Conditions
  - Jensen’s inequality and extensions
  - Epigraph
- Operations That Preserve Convexity
  - Nonnegative Weighted Sums
  - Composition with an affine mapping
  - Pointwise maximum and supremum
  - Composition
  - Minimization
  - Perspective of a function
- Summary

Convex Function

- $f: \mathbf{R}^n \to \mathbf{R}$ is convex if
  - $\text{dom}\, f$ is convex

    $\theta x + (1-\theta)y \in \text{dom}\, f, \forall \theta \in [0,1], x, y \in \text{dom}\, f$
  - $\forall \theta \in [0,1],\ x, y \in \text{dom}\ f$

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



Convex Function

- $f: \mathbf{R}^n \to \mathbf{R}$ is convex if
  - $\mathrm{dom}\, f$ is convex

    $\theta x + (1-\theta)y \in \mathrm{dom}\, f, \forall \theta \in [0,1], x, y \in \mathrm{dom}\, f$
  - $\forall \theta \in [0,1],\ x, y \in \mathrm{dom}\ f$

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

- $f: \mathbf{R}^n \to \mathbf{R}$ is strictly convex if
  - $\forall \theta \in (0,1),\ x \neq y$

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

Convex Function

- $f: \mathbf{R}^n \to \mathbf{R}$ is convex if
  - $\mathrm{dom}\, f$ is convex

$$\theta x + (1-\theta) y \in \mathrm{dom}\, f, \forall \theta \in [0,1], x, y \in \mathrm{dom}\, f$$

  - $\forall \theta \in [0,1],\ x, y \in \mathrm{dom}\ f$

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

- f is concave if $-f$ is convex
  - $\mathrm{dom}\, f$ is convex
- Affine functions are both convex and concave, and vice versa.

Extended-value Extensions

- The extended-value extension of f is
  - $\tilde{f}(x) = \begin{cases} f(x) & x \in \mathrm{dom}\, f \\ \infty & x \notin \mathrm{dom}\, f \end{cases}$
  - $\tilde{f}: \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$

    $\tilde{f}(\theta x + (1-\theta)y) \leq \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$
  - $\mathrm{dom}\, f = \{x | \tilde{f}(x) < \infty\}$
- Example
  - $f(x) = f_1(x) + f_2(x), \mathrm{dom}\, f = \mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2$
  - $\tilde{f}(x) = \tilde{f}_1(x) + \tilde{f}_2(x)$

    $\tilde{f}(x) = \infty, \text{if } x \notin \mathrm{dom}\, f_1 \text{ or } x \notin \mathrm{dom}\, f_2$

Extended-value Extensions

- The extended-value extension of f is
  - $$\tilde{f}(x) = \begin{cases} f(x) & x \in \operatorname{dom} f \\ \infty & x \notin \operatorname{dom} f \end{cases}$$
  - $\tilde{f}: \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$
    $$\tilde{f}(\theta x + (1-\theta)y) \le \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$$
- Example
  - Indicator Function of a Set C

$$\tilde{I}_C(x) = \begin{cases} 0 & x \in C \\ \infty & x \notin C \end{cases}$$

Zeroth-order Condition

- Definition
  - High-dimensional space

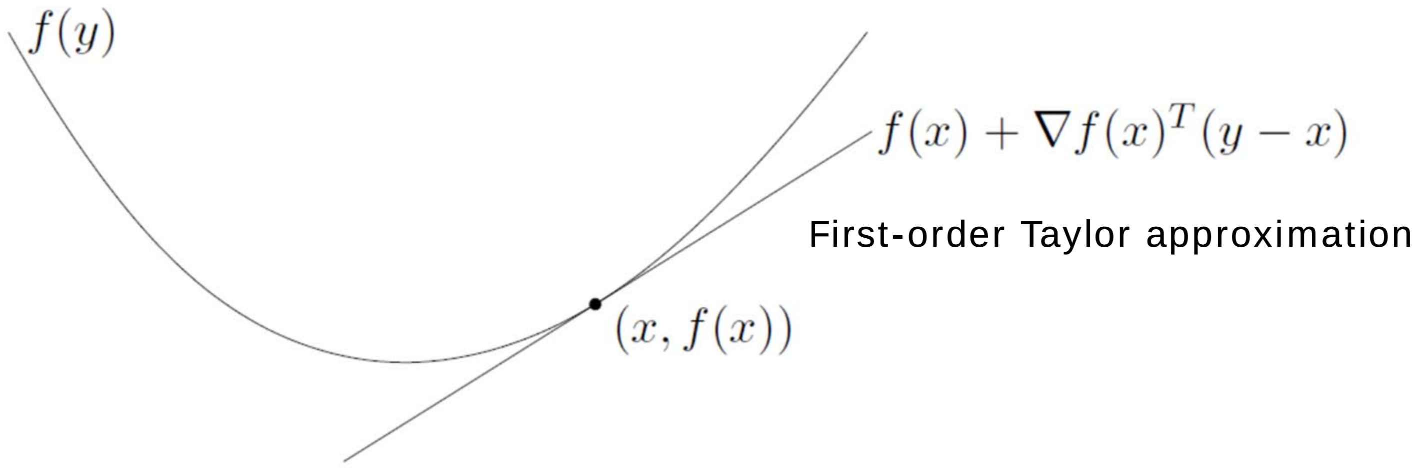
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

- A function is convex if and only if it is convex when restricted to any line that intersects its domain.
  - $x \in \operatorname{dom} f, v \in \mathbf{R}^n,\ t \in \mathbf{R}, x + tv \in \operatorname{dom} f$
  - f is convex $\Leftrightarrow g(t) = f(x+tv)$ is convex
  - One-dimensional space

First-order Conditions

- f is differentiable. Then f is convex if and only if
  - $\mathrm{dom}\, f$ is convex
  - For all $x, y \in \mathrm{dom}\, f$

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



First-order Conditions

- f is differentiable. Then f is convex if and only if
  - $\mathrm{dom}\, f$ is convex
  - For all $x, y \in \mathrm{dom}\, f$

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

  - Local Information ⇒ Global Information
  - $\nabla f(x) = 0 \Rightarrow f(y) \geq f(x), \forall\, y \in \mathrm{dom}\, f$
- f is strictly convex if and only if

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

Proof

- $f: \mathbf{R} \to \mathbf{R}$ is convex $\Leftrightarrow$ $f(y) \geq f(x) + f'(x)(y-x), x, y \in \operatorname{dom} f$
  - Necessary condition:

$$f\big(x + t(y-x)\big) \leq (1-t)f(x) + tf(y), 0 \leq t \leq 1$$

$$\Rightarrow f(y) \geq f(x) + \frac{f\big(x+t(y-x)\big) - f(x)}{t}$$

$$\overset{t \to 0}{\Longrightarrow} f(y) \geq f(x) + f'(x)(y-x)$$

  - Sufficient condition:

$$\left.\begin{matrix} z = \theta x + (1-\theta)y \\ f(x) \geq f(z) + f'(z)(x-z) \\ f(y) \geq f(z) + f'(z)(y-z) \end{matrix}\right\} \Rightarrow \left.\begin{matrix} f(x) \geq f(z) + (1-\theta)f'(z)(x-y) \\ f(y) \geq f(z) - \theta f'(z)(x-y) \end{matrix}\right\}$$

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

Proof

- $f: \mathbf{R} \to \mathbf{R}$ is convex $\Leftrightarrow$ $f(y) \geq f(x) + f'(x)(y-x), x, y \in \mathrm{dom}\, f$
- $f: \mathbf{R}^n \to \mathbf{R}$ is convex $\Leftrightarrow$ $f(y) \geq f(x) + \nabla f(x)^\top (y-x), x, y \in \mathrm{dom}\, f$

$$g(t) = f(ty + (1-t)x), \quad g'(t) = \nabla f(ty + (1-t)x)^\top (y-x)$$

  - f is convex $\Rightarrow$ $g(t)$ is convex $\Rightarrow$ $g(1) \geq g(0) + g'(0) \Rightarrow f(y) \geq f(x) + \nabla f(x)^\top (y-x)$

f is convex $\Leftrightarrow$ g is convex $\Leftrightarrow$ First-order condition of g $\Rightarrow$ First-order condition of f

Proof

- $f: \mathbf{R} \to \mathbf{R}$ is convex $\Leftrightarrow$ $f(y) \geq f(x) + f'(x)(y-x), x, y \in \operatorname{dom} f$
- $f: \mathbf{R}^n \to \mathbf{R}$ is convex $\Leftrightarrow$ $f(y) \geq f(x) + \nabla f(x)^\top (y-x), x, y \in \operatorname{dom} f$

$$g(t) = f(ty + (1-t)x), \quad g'(t) = \nabla f(ty + (1-t)x)^\top (y-x)$$

- $$f(ty + (1-t)x) \geq f(\tilde{t}y + (1-\tilde{t})x) + \nabla f(\tilde{t}y + (1-\tilde{t})x)^\top (y-x)(t-\tilde{t})$$

$\Rightarrow g(t) \geq g(\tilde{t}) + g'(\tilde{t})(t-\tilde{t}) \Rightarrow g(t)$ is convex $\Rightarrow$ f is convex

f is convex $\Leftrightarrow$ g is convex $\Leftrightarrow$ First-order condition of g $\Leftarrow$ First-order condition of f

Second-order Conditions

- f is twice differentiable. Then f is convex if and only if
  - $\operatorname{dom} f$ is convex
  - For all $x \in \operatorname{dom} f$, $\nabla^2 f(x) \succcurlyeq 0$

- Attention
  - $\nabla^2 f(x) \succ 0 \Rightarrow f$ is strictly convex
  - f is strict convex $\nRightarrow$ $\nabla^2 f(x) \succ 0$

    $f(x) = x^4$ is strict convex but $f''(0) = 0$
  - $\operatorname{dom} f$ is convex is necessary, $f(x) = 1/x^2$

Examples

- Functions on $\mathbf{R}$
  - e^{ax} is convex on $\mathbf{R}$, $\forall a \in \mathbf{R}$
  - x^a is convex on $\mathbf{R}_{++}$ when $a \geq 1$ or $a \leq 0$, and concave for $0 \leq a \leq 1$
  - $|x|^p$, for $p \geq 1$, is convex on $\mathbf{R}$
  - $\log x$ is concave on $\mathbf{R}_{++}$
  - Negative entropy $x \log x$ is convex on $\mathbf{R}_{++}$

Examples

Functions on $\mathbf{R}^n$

- Every norm on $\mathbf{R}^n$ is convex
- $f(x) = \max\{x_1, \ldots, x_n\}$
- Quadratic-over-linear: $f(x, y) = \dfrac{x^2}{y}$
  - ✓ $\mathrm{dom}\, f = \{(x, y) \in \mathbf{R}^2 \mid y > 0\}$
- $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$

$$\max\{x_1, \ldots, x_n\} \le f(x) \le \max\{x_1, \ldots, x_n\} + \log n$$

- $f(x) = (\prod_{i=1}^{n} x_i)^{1/n}$ is concave on $\mathbf{R}^n_{++}$
- $f(X) = \log\det X$ is concave on $\mathbf{S}^n_{++}$

Examples

- Functions on $\mathbf{R}^n$
  - Every norm on $\mathbf{R}^n$ is convex
    - ✓ $f(x)$ is a norm on $\mathbf{R}^n$
    - ✓ $f(\theta x + (1-\theta)y) \leq f(\theta x) + f\big((1-\theta)y\big)$

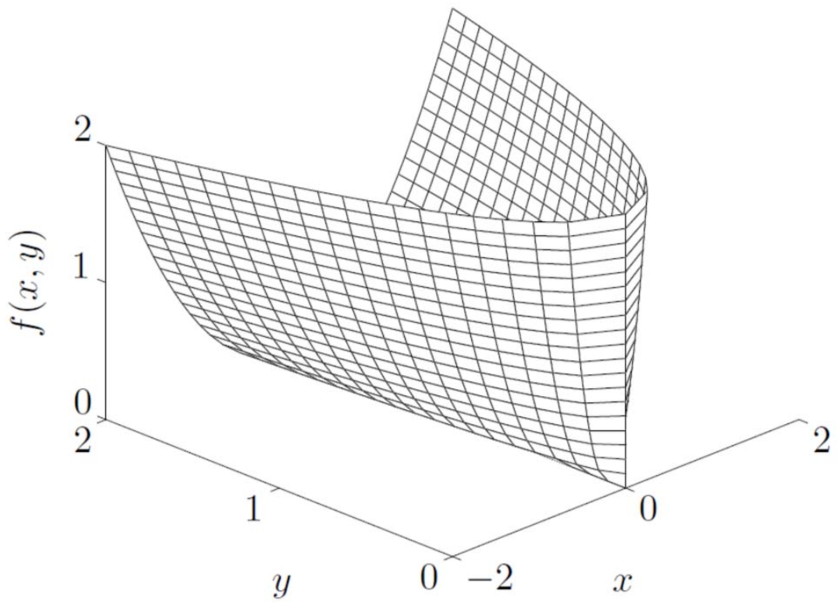
      $= \theta f(x) + (1-\theta) f(y)$
  - $f(x) = \max\{x_1, \ldots, x_n\} = \max_i x_i$
    - ✓ $f(\theta x + (1-\theta)y) = \max_i \{\theta x_i + (1-\theta) y_i\}$

      $\leq \theta \max_i \{x_i\} + (1-\theta) \max_i \{y_i\}$

Examples

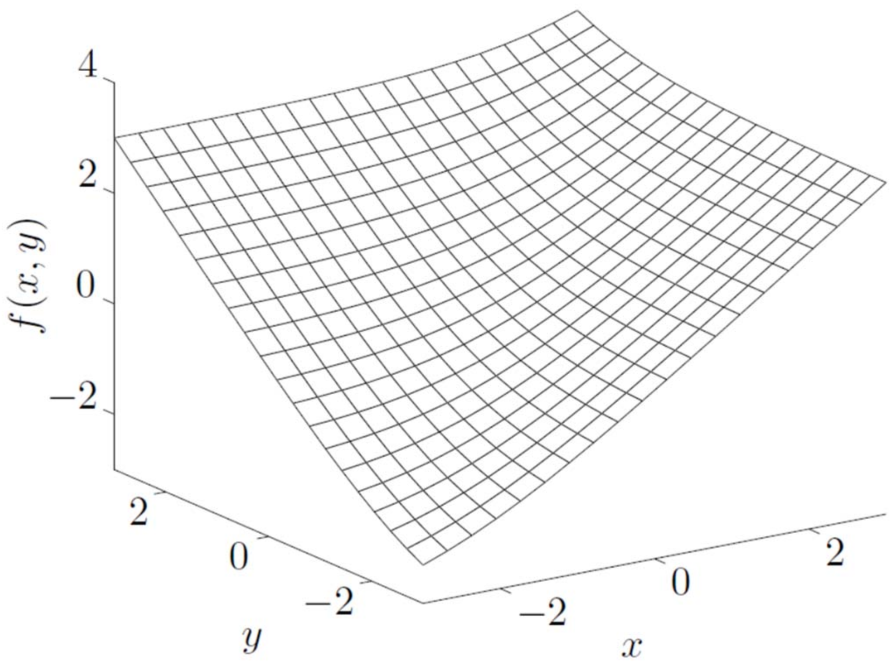
Functions on $\mathbf{R}^n$

- $f(x,y) = \frac{x^2}{y}, \mathrm{dom}\, f = \{(x,y) \in \mathbf{R}^2 \,|\, y > 0\}$
  - ✓ $\nabla^2 f(x,y) = \frac{2}{y^3}\begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix} = \frac{2}{y^3}\begin{bmatrix} y \\ -x \end{bmatrix}\begin{bmatrix} y \\ -x \end{bmatrix}^\top \succcurlyeq 0$



Examples

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



Examples

Functions on $\mathbf{R}^n$

- $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$
  - ✓ $\nabla^2 f(x) = \frac{1}{(\mathbf{1}^\top z)^2}\left((\mathbf{1}^\top z)\mathrm{diag}(z) - zz^\top\right)$
  - ✓ $z = (e^{x_1}, \dots, e^{x_n})$
  - ✓ $v^\top \nabla^2 f(x) v = \frac{1}{(\mathbf{1}^\top z)^2}\left((\sum_{i=1}^n z_i)(\sum_{i=1}^n v_i^2 z_i) - (\sum_{i=1}^n v_i z_i)^2\right) \geq 0$
  - ✓ Cauchy-Schwarz inequality: $(a^\top a)(b^\top b) \geq (a^\top b)^2$

Examples

□ Functions on $\mathbf{R}^n$

- $f(X) = \log\det X$ is concave on $\mathbf{S}^n_{++}$
  - ✓ $g(t) = f(Z + tV), Z + tV \succ 0, Z \succ 0$
  - ✓ $g(t) = \log\det(Z + tV)$

$$= \log\det\left(Z^{\frac{1}{2}}\left(I + tZ^{-\frac{1}{2}}VZ^{-\frac{1}{2}}\right)Z^{\frac{1}{2}}\right)$$

$$= \sum_{i=1}^{n} \log(1 + t\lambda_i) + \log\det Z$$

  - ✓ $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $Z^{-\frac{1}{2}}VZ^{-\frac{1}{2}}$
  - ✓ $g'(t) = \sum_{i=1}^{n} \frac{\lambda_i}{1+t\lambda_i}, g''(t) = -\sum_{i=1}^{n} \frac{\lambda_i^2}{(1+t\lambda_i)^2}$

$\det(AB) = \det(A)\det(B)$ https://en.wikipedia.org/wiki/Determinant

Sublevel Sets

- α-sublevel set

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

  - $f(x)$ is convex $\Rightarrow C_\alpha$ is convex, $\forall \alpha \in \mathbf{R}$
  - C_α is convex, $\forall \alpha \in \mathbf{R} \not\Rightarrow f(x)$ is convex

    e.g., $f(x) = -e^x$

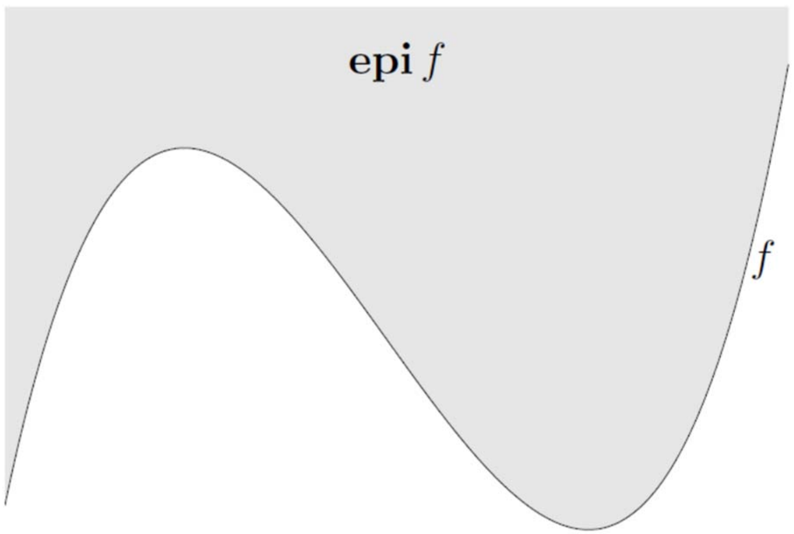
- α-superlevel set

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

  - $f(x)$ is concave $\Rightarrow C_\alpha$ is convex, $\forall \alpha \in \mathbf{R}$
  - $G(x) = \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}, A(x) = \frac{1}{n}\sum_{i=1}^n x_i$
  - $\{x \in \mathbf{R}_+^n \mid G(x) \ge \alpha A(x)\}$ is convex, $\alpha \in [0,1]$

Epigraph

- Graph of function $f\colon \mathbf{R}^n \to \mathbf{R}$
  - $\{(x, f(x)) | x \in \mathrm{dom}\, f\}$
- Epigraph of function $f\colon \mathbf{R}^n \to \mathbf{R}$
  - $\mathrm{epi}\, f = \{(x, t) | x \in \mathrm{dom}\, f, f(x) \le t\}$



Epigraph

- Epigraph of function $f: \mathbf{R}^n \to \mathbf{R}$
  - $\mathbf{epi}\, f = \{(x,t) | x \in \mathrm{dom}\, f, f(x) \le t\}$

- Hypograph
  - $\mathbf{hypo}\, f = \{(x,t) | x \in \mathrm{dom}\, f, t \le f(x)\}$

- Conditions
  - $f(x)$ is convex $\Leftrightarrow$ $\mathbf{epi}\, f$ is convex
  - $f(x)$ is concave $\Leftrightarrow$ $\mathbf{hypo}\, f$ is convex

Example

Matrix Fractional Function

$$f(x, Y) = x^\top Y^{-1} x, \mathrm{dom}\, f = \mathbf{R}^n \times \mathbf{S}^n_{++}$$

- Quadratic-over-linear: $f(x, y) = x^2/y$
- $\mathbf{epi}\, f = \{(x, Y, t) | Y \succ 0, x^\top Y^{-1} x \leq t\}$

$$= \left\{(x, Y, t) \,\middle|\, \begin{bmatrix} Y & x \\ x^\top & t \end{bmatrix} \succcurlyeq 0, Y \succ 0 \right\}$$

  - ✓ Schur complement condition
- $\mathbf{epi}\, f$ is convex
  - ✓ Linear matrix inequality
  - ✓ Recall Example 2.10 in the book

Example

Matrix Fractional Function

$$f(x, Y) = x^\top Y^{-1} x, \mathrm{dom}\, f = \mathbf{R}^n \times \mathbf{S}^n_{++}$$

- Quadratic-over-linear: $f(x, y) = x^2/y$
- $\mathbf{epi}\, f = \{(x, Y, t) | Y \succ 0, x^\top Y^{-1} x \le t\}$

$$= \left\{(x, Y, t) \,\middle|\, \begin{bmatrix} Y & x \\ x^\top & t \end{bmatrix} \succcurlyeq 0, Y \succ 0\right\}$$

  - ✓ Schur complement condition

Linear Matrix Inequality

$$A(x) = x_1 A_1 + \cdots + x_n A_n$$

$$\{x | A(x) \preccurlyeq B\} = \{x | B - A(x) \in \mathbf{S}^m_+\}$$

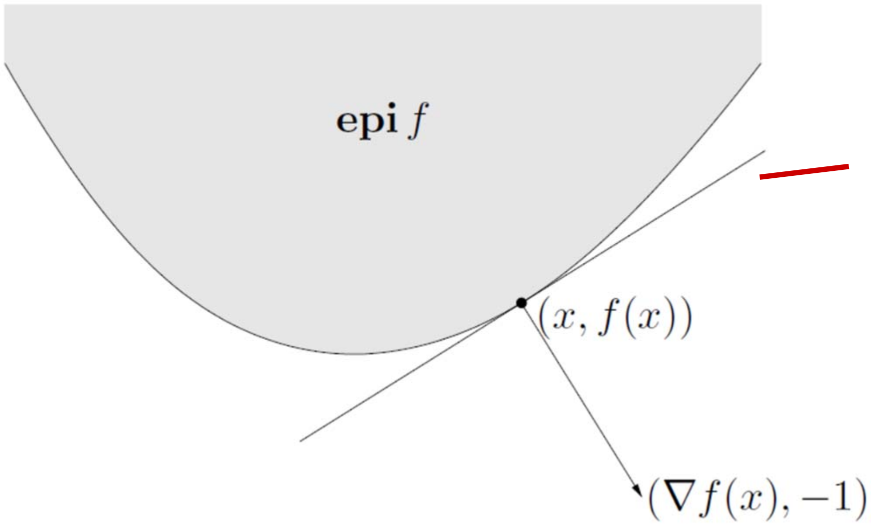
Application of Epigraph

- First order Condition
  - $f(y) \geq f(x) + \nabla f(x)^{\top}(y - x)$
  - $(y, t) \in \textbf{epi}\, f \Rightarrow t \geq f(y) \geq f(x) + \nabla f(x)^{\top}(y - x)$

Application of Epigraph

First order Condition

- $f(y) \geq f(x) + \nabla f(x)^\top (y - x)$
- $(y, t) \in \textbf{epi}\, f \Rightarrow t \geq f(x) + \nabla f(x)^\top (y - x)$
- $(y, t) \in \textbf{epi}\, f \Rightarrow \begin{bmatrix} \nabla f(x) \\ -1 \end{bmatrix}^\top \left(\begin{bmatrix} y \\ t \end{bmatrix} - \begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \leq 0$



Support **epi** f at point $(x, f(x))$

Jensen's Inequality

- Basic inequality
  - $\theta \in [0,1]$
  - $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$

- k points
  - $\theta_i \in [0,1], \theta_1 + \cdots + \theta_k = 1$
  - $f(\theta_1 x_1 + \cdots + \theta_k x_k) \leq \theta_1 f(x_1) + \cdots + \theta_k f(x_k)$

Jensen's Inequality

Infinite points

- $p(x) \geq 0, S \subseteq \operatorname{dom} f, \int_S p(x)\, dx = 1$
- $f\left(\int_S p(x) x\, dx\right) \leq \int_S f(x) p(x)\, dx$
- $f(\mathbf{E}x) \leq \mathbf{E}f(x)$
  - ✓ $f(x) \leq \mathbf{E}f(x+z), z$ is a zero-mean noisy

Hölder's inequality

- $\frac{1}{p} + \frac{1}{q} = 1, p > 1$
- $\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}}$

Outline

- Basic Properties
  - Definition
  - First-order Conditions, Second-order Conditions
  - Jensen's inequality and extensions
  - Epigraph
- Operations That Preserve Convexity
  - Nonnegative Weighted Sums
  - Composition with an affine mapping
  - Pointwise maximum and supremum
  - Composition
  - Minimization
  - Perspective of a function
- Summary

Nonnegative Weighted Sums

- Finite sums
  - $w_i \geq 0, f_i$ is convex
  - $f = w_1 f_1 + \cdots + w_m f_m$ is convex

The set of convex functions is itself a convex cone

- Infinite sums
  - $f(x,y)$ is convex in x, $\forall y \in \mathcal{A}, w(y) \geq 0$
  - $g(x) = \int_{\mathcal{A}} f(x,y) w(y)\, dy$ is convex
- Epigraph interpretation
  - $\textbf{epi}\,(wf) = \{(x,t) | wf(x) \leq t\}$
  - $\begin{bmatrix} I & 0 \\ 0 & w \end{bmatrix} \textbf{epi}\,(f) = \{(x, wt) | f(x) \leq t\}$
  - $\textbf{epi}\,(wf) = \begin{bmatrix} I & 0 \\ 0 & w \end{bmatrix} \textbf{epi}\,(f)$

Composition with an affine mapping

- $f: \mathbf{R}^n \to \mathbf{R}$
- $A \in \mathbf{R}^{n\times m}, b \in \mathbf{R}^n$
- Affine Mapping

$$g(x) = f(Ax + b)$$

  - If f is convex, so is g.

  - If f is concave, so is g.

Pointwise Maximum

- f_1, f_2 is convex
$$f(x) = \max\{f_1(x), f_2(x)\}$$
is convex with $\mathrm{dom}\, f = \mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2$
  - $f(\theta x + (1-\theta)y)$
$$= \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\}$$
$$\leq \max\{\theta f_1(x) + (1-\theta) f_1(y), \theta f_2(x) + (1-\theta) f_2(y)\}$$
$$\leq \theta \max\{f_1(x), f_2(x)\} + (1-\theta) \max\{f_1(y), f_2(y)\}$$
$$= \theta f(x) + (1-\theta) f(y)$$
  - $f_1, \ldots, f_m$ is convex $\Rightarrow f(x) = \max\{f_1(x), \ldots, f_m(x)\}$

Examples

- Piecewise-linear functions
  - $f(x) = \max\{a_1^\top x + b_1, \dots, a_L^\top x + b_L\}$

- Sum of r largest components
  - $x \in \mathbf{R}^n, x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$
  - $f(x) = \sum_{i=1}^{r} x_{[i]}$ is convex

    $= \max\{x_{i_1} + \cdots + x_{i_r} | 1 \leq i_1 < \cdots < i_r \leq n\}$
  - Pointwise maximum of $\frac{n!}{r!(n-r)!}$ linear functions

Pointwise Supremum

- $\forall y \in \mathcal{A}, f(x, y)$ is convex in x

$$g(x) = \sup_{y \in \mathcal{A}} f(x, y)$$

is convex with $\operatorname{dom} g = \{x \mid (x, y) \in \operatorname{dom} f \text{ for all } y \in \mathcal{A}, \sup_{y \in \mathcal{A}} f(x, y) < \infty\}$

- Epigraph interpretation
  - $\operatorname{epi} g = \bigcap_{y \in \mathcal{A}} \operatorname{epi} f(\cdot, y)$
  - Intersection of convex sets is convex
- Pointwise infimum of a set of concave functions is concave

Examples

- Support function of a set
  - $C \subseteq \mathbf{R}^n, C \neq \emptyset$
  - $S_C(x) = \sup\{x^\top y | y \in C\}$
  - $\mathrm{dom}\, S_C = \{x | \sup_{y \in C} x^\top y < \infty\}$

- Distance to farthest point of a set
  - $C \subseteq \mathbf{R}^n$
  - $f(x) = \sup_{y \in C} \|x - y\|$

Examples

- Maximum eigenvalue of a symmetric matrix
  - $f(X) = \lambda_{\max}(X), \operatorname{dom} f = \mathbf{S}^m$
  - $f(X) = \sup\{y^\top X y \mid \|y\|_2 = 1\}$

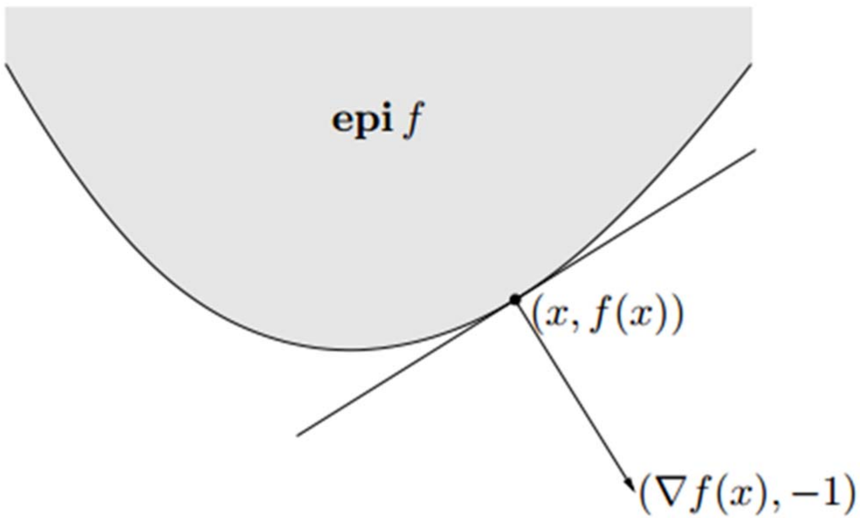
- Norm of a matrix
  - $f(X) = \|X\|_2$ is maximum singular value of X
  - $\operatorname{dom} f = \mathbf{R}^{p\times q}$
  - $f(X) = \sup\{u^\top X v \mid \|u\|_2 = 1, \|v\|_2 = 1\}$

Representation

- Almost every convex function can be expressed as the pointwise supremum of a family of affine functions.

$f: \mathbf{R}^n \to \mathbf{R}$ is convex and $\mathrm{dom}\, f = \mathbf{R}^n$

$$\Longrightarrow f(x) = \sup\{g(x) \mid g \text{ affine}, g(z) \le f(z)\ \forall z\}$$



Compositions

Definition

- $h: \mathbf{R}^k \to \mathbf{R}, g: \mathbf{R}^n \to \mathbf{R}^k$
- $f = h \circ g: \mathbf{R}^n \to \mathbf{R}$

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

- $\mathrm{dom}\, f = \{x \in \mathrm{dom}\, g | g(x) \in \mathrm{dom}\, h\}$

Chain Rule

- $h: \mathbf{R} \to \mathbf{R}, g: \mathbf{R}^n \to \mathbf{R}$

$$\nabla^2 f(x) = h'(g(x))\nabla^2 g(x) + h''\big(g(x)\big)\nabla g(x)\nabla g(x)^\top$$

Scalar Composition

- $h: \mathbf{R} \to \mathbf{R}, g: \mathbf{R} \to \mathbf{R}$
  - h and g are twice differentiable
  - $\mathrm{dom}\, g = \mathrm{dom}\, h = \mathbf{R}$

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

  - f is convex, if $f''(x) \geq 0$
  - $h'' \geq 0, h' \geq 0, g'' \geq 0$
    - ✓ h is convex and nondecreasing, g is convex
  - $h'' \geq 0, h' \leq 0, g'' \leq 0$
    - ✓ h is convex and nonincreasing, g is concave

Scalar Composition

- $h: \mathbf{R} \to \mathbf{R}, g: \mathbf{R} \to \mathbf{R}$
  - h and g are twice differentiable
  - $\operatorname{dom} g = \operatorname{dom} h = \mathbf{R}$

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

  - f is concave, if $f''(x) \leq 0$
  - $h'' \leq 0, h' \geq 0, g'' \leq 0$
    - ✓ h is concave and nondecreasing, g is concave
  - $h'' \leq 0, h' \leq 0, g'' \geq 0$
    - ✓ h is concave and nonincreasing, g is convex

Scalar Composition

- $h: \mathbf{R} \to \mathbf{R}, g: \mathbf{R}^n \to \mathbf{R}$
  - Without differentiability assumption
  - Without domain condition

  - $h(x) = 0$ with $\mathrm{dom}\, h = [1,2]$, which is convex and nondecreasing
  - $g(x) = x^2$ with $\mathrm{dom}\, g = \mathbf{R}$, which is convex

$$f(x) = h\big(g(x)\big) = 0$$

  - $\mathrm{dom}\, f = \left[-\sqrt{2}, -1\right] \cup [1, \sqrt{2}]$

Scalar Composition

- $h: \mathbf{R} \rightarrow \mathbf{R}, g: \mathbf{R}^n \rightarrow \mathbf{R}$
  - Without differentiability assumption
  - Without domain condition
  - h is convex, $\tilde{h}$ is nondecreasing, and g is convex $\Rightarrow f$ is convex
  - h is convex, $\tilde{h}$ is nonincreasing, and g is concave $\Rightarrow f$ is convex
  - The conditions for concave are similar

Extended-value Extensions

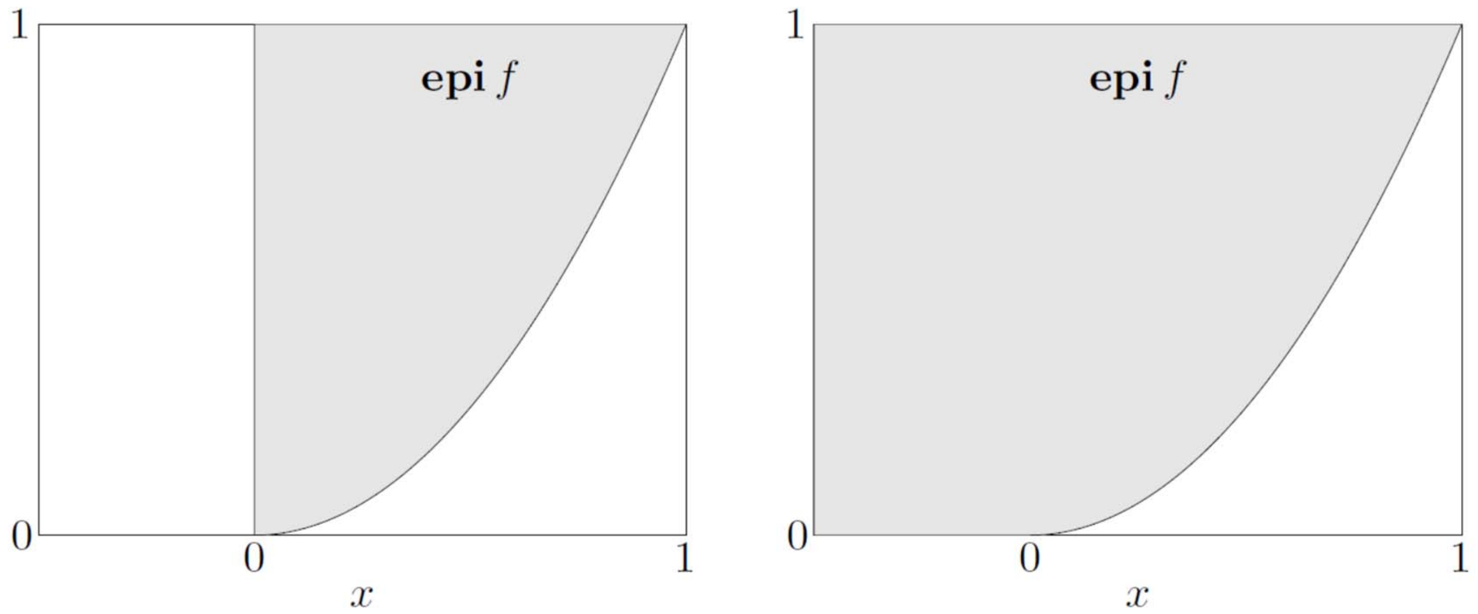


Figure 3.7 *Left.* The function x^2, with domain $\mathbf{R}_+$, is convex and nondecreasing on its domain, but its extended-value extension is *not* nondecreasing. *Right.* The function $\max\{x, 0\}^2$, with domain $\mathbf{R}$, is convex, and its extended-value extension is nondecreasing.

Examples

- g is convex $\Rightarrow \exp g(x)$ is convex
- g is concave and positive $\Rightarrow \log g(x)$ is concave
- g is concave and positive $\Rightarrow 1/g(x)$ is convex
- g is convex and nonnegative and $p \geq 1 \Rightarrow g(x)^p$ is convex
- g is convex $\Rightarrow -\log(-g(x))$ is convex on $\{x | g(x) < 0\}$

Vector Composition

- $h: \mathbf{R}^k \to \mathbf{R}, g_i: \mathbf{R} \to \mathbf{R}$

$$f = h \circ g = h(g_1(x), \dots, g_k(x))$$

  - h and g are twice differentiable
  - $\operatorname{dom} g_i = \mathbf{R}, \operatorname{dom} h = \mathbf{R}^k$

$$f'(x) = \nabla h\big(g(x)\big)^\top g'(x)$$

$$f''(x) = g'(x)^\top \nabla^2 h\big(g(x)\big) g'(x) + \nabla h\big(g(x)\big)^\top g''(x)$$

Vector Composition

- $h: \mathbf{R}^k \to \mathbf{R}, g_i: \mathbf{R} \to \mathbf{R}$

$$f = h \circ g = h(g_1(x), \ldots, g_k(x))$$

  - h and g are twice differentiable
  - $\operatorname{dom} g_i = \mathbf{R}, \operatorname{dom} h = \mathbf{R}^k$

$$f''(x) = g'(x)^\top \nabla^2 h\big(g(x)\big) g'(x) + \nabla h\big(g(x)\big)^\top g''(x)$$

  - f is convex, if $f''(x) \geq 0$
    - ✓ h is convex, h is nondecreasing in each argument, and g_i are convex
    - ✓ h is convex, h is nonincreasing in each argument, and g_i are concave

Vector Composition

- $h: \mathbf{R}^k \to \mathbf{R}, g_i: \mathbf{R} \to \mathbf{R}$

$$f = h \circ g = h(g_1(x), \dots, g_k(x))$$

  - h and g are twice differentiable
  - $\operatorname{dom} g_i = \mathbf{R}, \operatorname{dom} h = \mathbf{R}^k$

$$f''(x) = g'(x)^\top \nabla^2 h\big(g(x)\big) g'(x) + \nabla h\big(g(x)\big)^\top g''(x)$$

  - f is concave, if $f''(x) \le 0$
    - ✓ h is concave, h is nondecreasing in each argument, and g_i are concave

- The general case is similar

Examples

- $h(z) = z_{[1]} + \cdots + z_{[r]}, z \in \mathbf{R}^k, g_1, \ldots, g_k$ are convex $\Rightarrow$ $h \circ g$ is convex
- $h(z) = \log\left(\sum_{i=1}^k e^{z_i}\right), g_1, \ldots, g_k$ are convex $\Rightarrow h \circ g$ is convex
- $h(z) = \left(\sum_{i=1}^k z_i^p\right)^{1/p}$ on $\mathbf{R}_+^k$ is concave for $0 \leq p \leq 1$, and its extension is nondecreasing. If g_i is concave and nonnegative $\Rightarrow h \circ g = \left(\sum_{i=1}^k g_i(x)^p\right)^{1/p}$ is concave
- Suppose $p \geq 1$, and $g_1, \ldots, g_k$ are convex and nonnegative. Then the function $\left(\sum_{i=1}^k g_i(x)^p\right)^{1/p}$ is convex

Minimization

- f is convex in (x, y), C is convex $(C \neq \emptyset)$
  - $g(x) = \inf_{y \in C} f(x, y)$ is convex if $g(x) > -\infty, \forall\, x \in \mathrm{dom}\, g$
  - $\mathrm{dom}\, g = \{x | (x, y) \in \mathrm{dom}\, f \text{ for some } y \in C\}$

- Proof by Epigraph
  - $\mathrm{epi}\, g = \{(x, t) | (x, y, t) \in \mathrm{epi}\, f \text{ for some } y \in C\}$
  - The projection of a convex set is convex.

Minimization

- f is convex in (x, y), C is convex $(C \neq \emptyset)$
  - $g(x) = \inf_{y \in C} f(x, y)$ is convex if $g(x) > -\infty, \forall\, x \in \operatorname{dom} g$
  - $\operatorname{dom} g = \{x | (x, y) \in \operatorname{dom} f \text{ for some } y \in C\}$

Pointwise Supremum

- $\forall y \in \mathcal{A}, f(x, y)$ is convex in x

$$g(x) = \sup_{y \in \mathcal{A}} f(x, y)$$

is convex with dom $g = \{x | (x, y) \in \operatorname{dom} f \text{ for all } y \in \mathcal{A}, \sup_{y \in \mathcal{A}} f(x, y) < \infty\}$

Examples

Schur complement

- $f(x,y) = x^\top Ax + 2x^\top By + y^\top Cy$
- $\begin{bmatrix} A & B \\ B^\top & C \end{bmatrix} \succcurlyeq 0, A, C$ is symmetric $\Rightarrow f(x,y)$ is convex
- $g(x) = \inf_y f(x,y) = x^\top \left(A - BC^\dagger B^\top\right) x$ is convex

$\Rightarrow A - BC^\dagger B^\top \succcurlyeq 0$, $C^\dagger$ is the pseudo-inverse of C

Distance to a set

- S is a convex nonempty set, $f(x,y) = \|x - y\|$ is convex in (x,y)
- $g(x) = \mathrm{dist}(x,S) = \inf_{y \in S} \|x - y\|$

Examples

- Distance to farthest point of a set
  - $C \subseteq \mathbf{R}^n$
  - $f(x) = \sup_{y \in C} \|x - y\|$

- Distance to a set
  - S is a convex nonempty set, $f(x, y) = \|x - y\|$ is convex in (x, y)
  - $g(x) = \mathrm{dist}(x, S) = \inf_{y \in S} \|x - y\|$

Examples

Affine domain

- $h(y)$ is convex
- $g(x) = \inf \{h(y) | Ay = x\}$ is convex

Proof

- $f(x,y) = \begin{cases} h(y) & \text{if } Ay - x = 0 \\ \infty & \text{otherwise} \end{cases}$
- $f(x,y)$ is convex in (x,y)
- g is the minimum of f over y

Perspective of a function

- $f: \mathbf{R}^n \to \mathbf{R}$, $g: \mathbf{R}^{n+1} \to \mathbf{R}$ defined as

$$g(x,t) = tf(x/t)$$

  is the perspective of f

  - $\mathrm{dom}\, g = \{(x,t) | x/t \in \mathrm{dom}\, f, t > 0\}$
  - f is convex $\Rightarrow g$ is convex

- Proof

$$(x,t,s) \in \mathrm{epi}\, g \Leftrightarrow tf\left(\frac{x}{t}\right) \le s$$

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

$$\Leftrightarrow (x/t, s/t) \in \mathrm{epi}\, f$$

  - Perspective mapping preserve convexity

Perspective Functions

- Perspective function $P: \mathbf{R}^{n+1} \to \mathbf{R}^n$

$$P(z,t) = \frac{z}{t}, \operatorname{dom} P = \mathbf{R}^n \times \mathbf{R}_{++}$$

- If $C \subseteq \operatorname{dom} P$ is convex, then its image

$$P(C) = \{P(x) | x \in C\}$$

is convex

- If $C \subseteq \mathbf{R}^n$ is convex, the inverse image

$$P^{-1}(C) = \left\{(x,t) \in \mathbf{R}^{n+1} \middle| \frac{x}{t} \in C, t > 0\right\}$$

is convex

Example

- Euclidean norm squared
  - $f(x) = x^\top x$
  - $g(x,t) = t\left(\frac{x}{t}\right)^\top\left(\frac{x}{t}\right) = \frac{x^\top x}{t}, t > 0$

- Composition with an Affine function
  - $f: \mathbf{R}^m \to \mathbf{R}$ is convex
  - $A \in \mathbf{R}^{m\times n}, b \in \mathbf{R}^m, c \in \mathbf{R}^n, d \in \mathbf{R}$
  - $\operatorname{dom} g = \left\{x \middle| c^\top x + d > 0, \frac{Ax+b}{c^\top x+d} \in \operatorname{dom} f\right\}$
  - $g(x) = (c^\top x + d) f\left(\frac{Ax+b}{c^\top x+d}\right)$ is convex

Outline

- Basic Properties
  - Definition
  - First-order Conditions, Second-order Conditions
  - Jensen's inequality and extensions
  - Epigraph
- Operations That Preserve Convexity
  - Nonnegative Weighted Sums
  - Composition with an affine mapping
  - Pointwise maximum and supremum
  - Composition
  - Minimization
  - Perspective of a function
- Summary

Summary

- Basic Properties
  - Definition
  - First-order Conditions, Second-order Conditions
  - Jensen's inequality and extensions
  - Epigraph
- Operations That Preserve Convexity
  - Nonnegative Weighted Sums
  - Composition with an affine mapping
  - Pointwise maximum and supremum
  - Composition
  - Minimization
  - Perspective of a function