

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

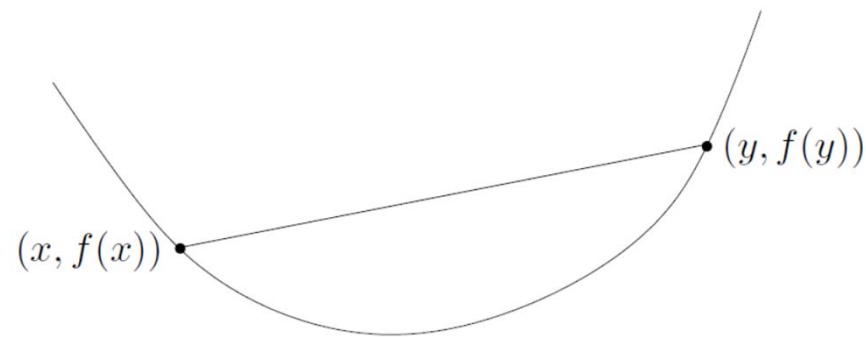
□ $f: \mathbf{R}^n \rightarrow \mathbf{R}$ is convex if

■ $\text{dom } f$ is convex

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

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

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





Convex Function

□ $f: \mathbf{R}^n \rightarrow \mathbf{R}$ is convex if

■ $\text{dom } f$ is convex

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

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

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

□ $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 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

□ $f: \mathbf{R}^n \rightarrow \mathbf{R}$ is convex if

■ $\text{dom } f$ is convex

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

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

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

□ f is concave if $-f$ is convex

■ $\text{dom } f$ is convex

□ Affine functions are both convex and concave, and vice versa.



Extended-value Extensions

□ The extended-value extension of f is

- $\tilde{f}(x) = \begin{cases} f(x) & x \in \text{dom } f \\ \infty & x \notin \text{dom } f \end{cases}$

- $\tilde{f}: \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$

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

- $\text{dom } f = \{x | \tilde{f}(x) < \infty\}$

□ Example

- $f(x) = f_1(x) + f_2(x), \text{dom } f = \text{dom } f_1 \cap \text{dom } f_2$

- $\tilde{f}(x) = \tilde{f}_1(x) + \tilde{f}_2(x)$

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



Extended-value Extensions

□ The extended-value extension of f is

■
$$\tilde{f}(x) = \begin{cases} f(x) & x \in \text{dom } f \\ \infty & x \notin \text{dom } f \end{cases}$$

■
$$\tilde{f}: \mathbf{R}^n \rightarrow \mathbf{R} \cup \{\infty\}$$

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

□ Example

■ Indicator Function of a Set \mathcal{C}

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



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.

- $x \in \text{dom } f, v \in \mathbf{R}^n, t \in \mathbf{R}, x + tv \in \text{dom } f$
- f is convex $\Leftrightarrow g(t) = f(x + tv)$ is convex
- One-dimensional space

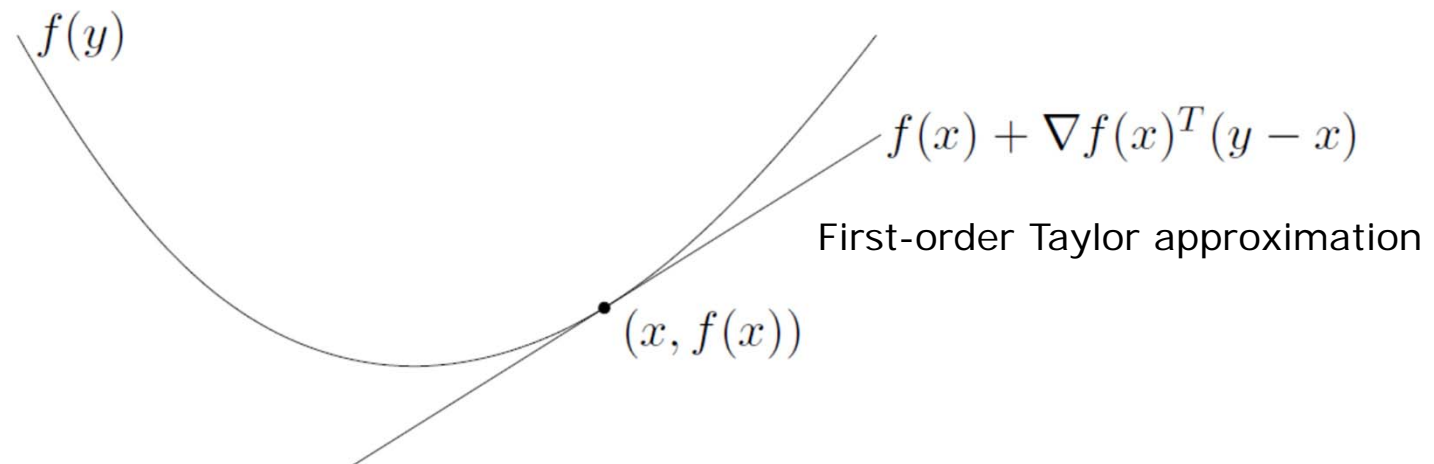


First-order Conditions

□ f is differentiable. Then f is convex if and only if

- $\text{dom } f$ is convex
- For all $x, y \in \text{dom } f$

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





First-order Conditions

□ f is differentiable. Then f is convex if and only if

■ $\text{dom } f$ is convex

■ For all $x, y \in \text{dom } f$

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

■ Local Information \Rightarrow Global Information

■ $\nabla f(x) = 0 \Rightarrow f(y) \geq f(x), \forall y \in \text{dom } f$

□ f is strictly convex if and only if

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



Proof

□ f is convex $\Leftrightarrow f: \mathbf{R} \rightarrow \mathbf{R}, f(y) \geq f(x) + f'(x)(y - x), x, y \in \text{dom } f$

■ Necessary condition:

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

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

$$\xrightarrow{t \rightarrow 0} \Rightarrow f(y) \geq f(x) + f'(x)(y - x)$$

■ Sufficient condition:

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

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



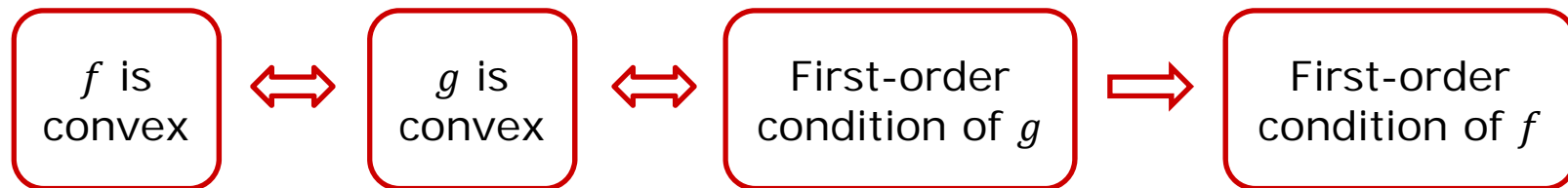
Proof

$$\square \quad f \text{ is convex} \Leftrightarrow f: \mathbf{R} \rightarrow \mathbf{R}, f(y) \geq f(x) + f'(x)(y - x), x, y \in \text{dom } f$$

$$\square \quad f \text{ is convex} \Leftrightarrow f: \mathbf{R}^n \rightarrow \mathbf{R}, f(y) \geq f(x) + \nabla f(x)^\top (y - x), x, y \in \text{dom } f$$

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

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





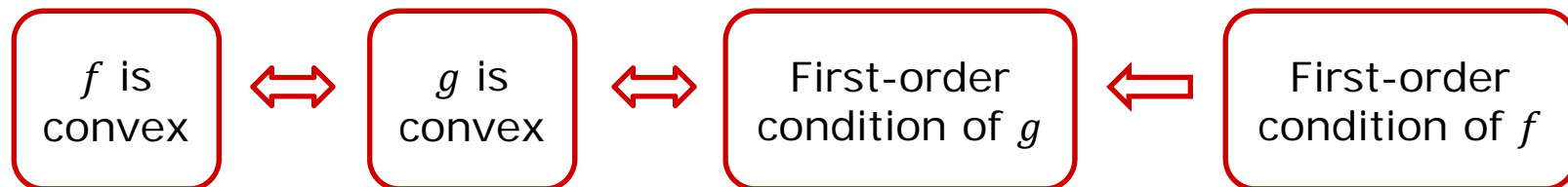
Proof

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

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

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

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





Second-order Conditions

□ f is twice differentiable. Then f is convex if and only if

- $\text{dom } f$ is convex
- For all $x \in \text{dom } f$, $\nabla^2 f(x) \succeq 0$

□ 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 $f''(0) = 0$
- $\text{dom } f$ 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

- $f(x) = \max\{x_1, \dots, x_n\}$

- Quadratic-over-linear: $f(x, y) = \frac{x^2}{y}$

 - ✓ $\text{dom } f = \{(x, y) \in \mathbf{R}^2 \mid y > 0\}$

- $f(x) = \log(e^{x_1} + \dots + e^{x_n})$

$$\max\{x_1, \dots, x_n\} \leq f(x) \leq \max\{x_1, \dots, x_n\} + \log n$$

- $f(x) = (\prod_{i=1}^n x_i)^{1/n}$ 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

$$\begin{aligned} \checkmark \quad f(\theta x + (1 - \theta)y) &\leq f(\theta x) + f((1 - \theta)y) \\ &= \theta f(x) + (1 - \theta)f(y) \end{aligned}$$

■ $f(x) = \max\{x_1, \dots, x_n\} = \max_i x_i$

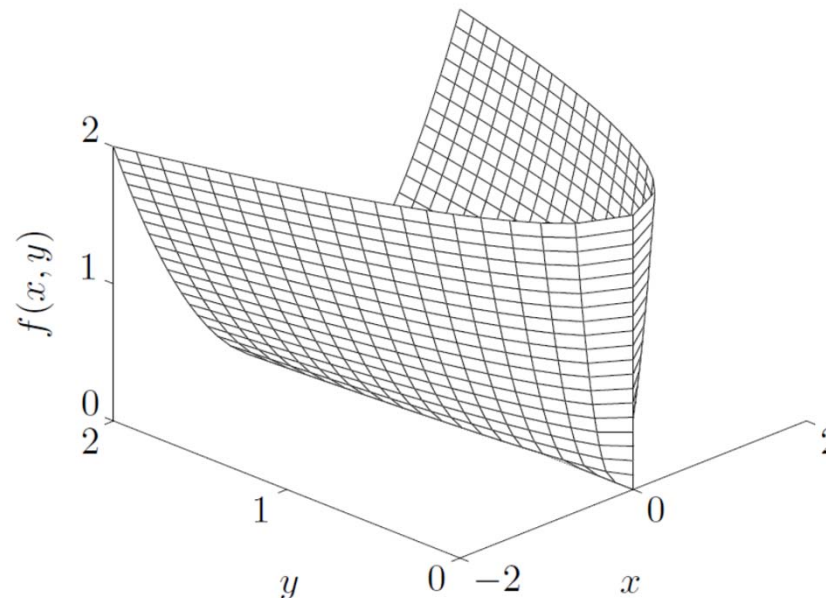
$$\begin{aligned} \checkmark \quad f(\theta x + (1 - \theta)y) &= \max_i \{ \theta x_i + (1 - \theta)y_i \} \\ &\leq \theta \max_i \{x_i\} + (1 - \theta) \max_i \{y_i\} \end{aligned}$$

Examples

□ Functions on \mathbf{R}^n

■ $f(x, y) = \frac{x^2}{y}, \text{ dom } f = \{(x, y) \in \mathbf{R}^2 \mid y > 0\}$

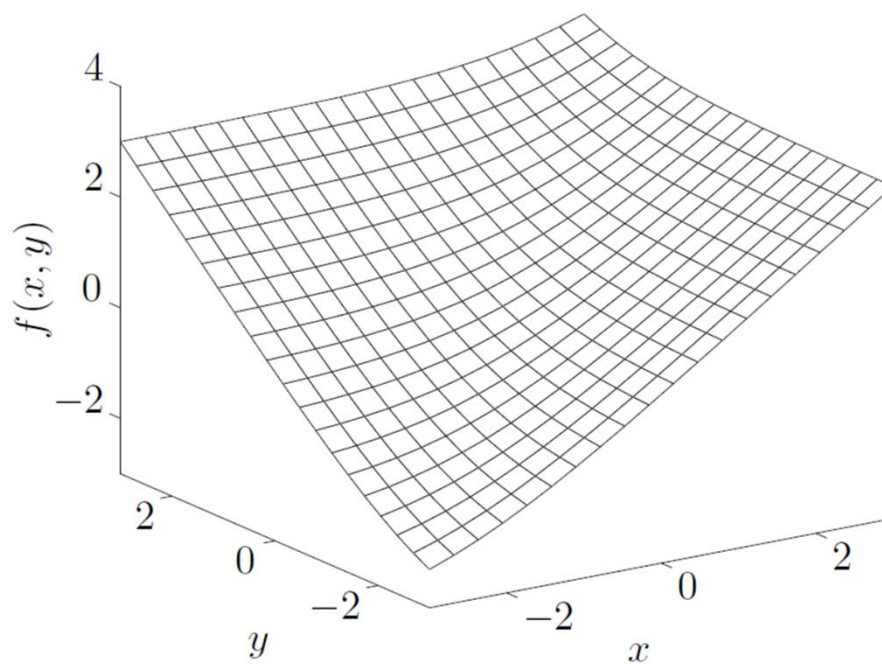
✓ $\nabla^2 f(x, y) = \frac{2}{y^3} \begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix} = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y \\ -x \end{bmatrix}^\top \succcurlyeq 0$



Examples

□ Functions on \mathbf{R}^n

■ $f(x) = \log(e^{x_1} + \dots + e^{x_n})$





Examples

□ Functions on \mathbf{R}^n

■ $f(x) = \log(e^{x_1} + \dots + e^{x_n})$

✓ $\nabla^2 f(x) = \frac{1}{(\mathbf{1}^\top z)^2} ((\mathbf{1}^\top z) \text{diag}(z) - zz^\top)$

✓ $z = (e^{x_1}, \dots, e^{x_n})$

✓
$$v^\top \nabla^2 f(x) v = \frac{1}{(\mathbf{1}^\top z)^2} \left(\left(\sum_{i=1}^n z_i \right) \left(\sum_{i=1}^n v_i^2 z_i \right) - \left(\sum_{i=1}^n v_i z_i \right)^2 \right) \geq 0$$

✓ 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

$$C_\alpha = \{x \in \text{dom } f \mid f(x) \leq \alpha\}$$

- $f(x)$ is convex $\Rightarrow C_\alpha$ is convex, $\forall \alpha \in \mathbf{R}$
- C_α is convex, $\forall \alpha \in \mathbf{R} \nRightarrow f(x)$ is convex
i.e. $f(x) = -e^x$

□ α -superlevel set

$$C_\alpha = \{x \in \text{dom } f \mid f(x) \geq \alpha\}$$

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

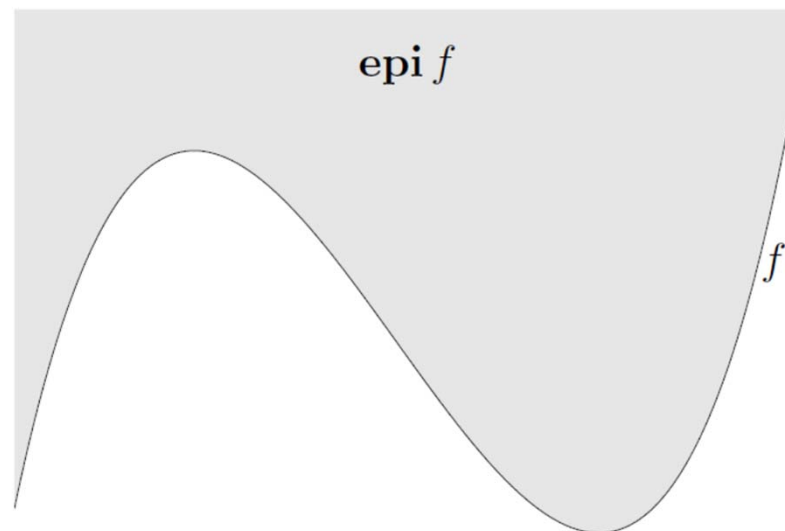
Epigraph

□ Graph of function $f: \mathbf{R}^n \rightarrow \mathbf{R}$

■ $\{(x, f(x)) | x \in \text{dom } f\}$

□ Epigraph of function $f: \mathbf{R}^n \rightarrow \mathbf{R}$

■ $\text{epi } f = \{(x, t) | x \in \text{dom } f, f(x) \leq t\}$





Epigraph

□ Epigraph of function $f: \mathbf{R}^n \rightarrow \mathbf{R}$

- $\text{epi } f = \{(x, t) | x \in \text{dom } f, f(x) \leq t\}$

□ Hypograph

- $\text{hypo } f = \{(x, t) | x \in \text{dom } f, t \leq f(x)\}$

□ Conditions

- $f(x)$ is convex $\Leftrightarrow \text{epi } f$ is convex

- $f(x)$ is concave $\Leftrightarrow \text{hypo } f$ is convex



Example

□ Matrix Fractional Function

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

■ Quadratic-over-linear: $f(x, y) = x^2/y$

■ $\text{epi } f = \{(x, Y, t) | Y \succ 0, x^\top Y^{-1} x \leq t\}$

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

✓ Schur complement condition

■ $\text{epi } f$ is convex



Application of Epigraph

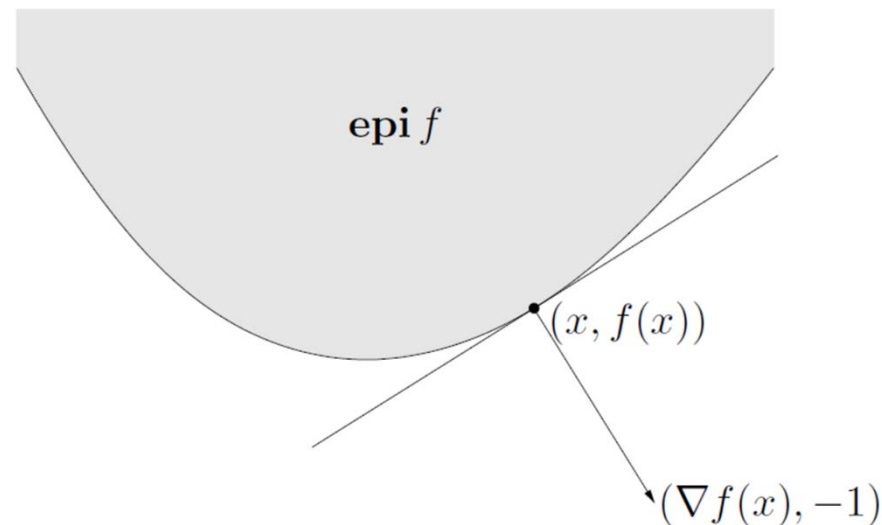
□ First order Condition

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

Application of Epigraph

□ First order Condition

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





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

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

□ Hölder's inequality

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



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

□ 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

- $\mathbf{epi}(wf) = \{(x, t) | wf(x) \leq t\}$
- $\begin{bmatrix} I & 0 \\ 0 & w \end{bmatrix} \mathbf{epi}(f) = \{(x, wt) | f(x) \leq t\}$
- $\mathbf{epi}(wf) = \begin{bmatrix} I & 0 \\ 0 & w \end{bmatrix} \mathbf{epi}(f)$



Composition with an affine mapping

□ $f: \mathbf{R}^n \rightarrow \mathbf{R}$

□ $A \in \mathbf{R}^{n \times m}, b \in \mathbf{R}^n$

□ 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 $\text{dom } f = \text{dom } f_1 \cap \text{dom } f_2$

$$\begin{aligned} & \blacksquare f(\theta x + (1 - \theta)y) \\ &= \max\{f_1(\theta x + (1 - \theta)y), f_2(\theta x + (1 - \theta)y)\} \\ &\leq \max\{\theta f_1(x) + (1 - \theta)f_1(y), \theta f_2(x) + (1 - \theta)f_2(y)\} \\ &\leq \theta \max\{f_1(x), f_2(x)\} + (1 - \theta) \max\{f_1(y), f_2(y)\} \\ &= \theta f(x) + (1 - \theta)f(y) \end{aligned}$$

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



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

- $x \in \mathbf{R}^n, x_{[1]} \geq x_{[2]} \geq \dots \geq x_{[n]}$

- $f(x) = \sum_{i=1}^r x_{[i]}$ is convex
$$= \max\{x_{i_1} + \dots + x_{i_r} \mid 1 \leq i_1 < \dots < i_r \leq n\}$$

- 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 $\text{dom } g = \{x | (x, y) \in \text{dom } f, \forall y \in \mathcal{A}, \sup_{y \in \mathcal{A}} f(x, y) < \infty\}$

□ Epigraph interpretation

■ $\text{epi } g = \bigcap_{y \in \mathcal{A}} \text{epi } f(\cdot, y)$

■ Intersection of convex sets is convex

□ Pointwise infimum of a set of concave functions is concave



Examples

□ Support function of a set

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

□ 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

- $f(X) = \lambda_{\max}(X)$, $\text{dom } f = \mathbf{S}^m$
- $f(X) = \sup\{y^\top X y \mid \|y\|_2 = 1\}$

□ Norm of a matrix

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

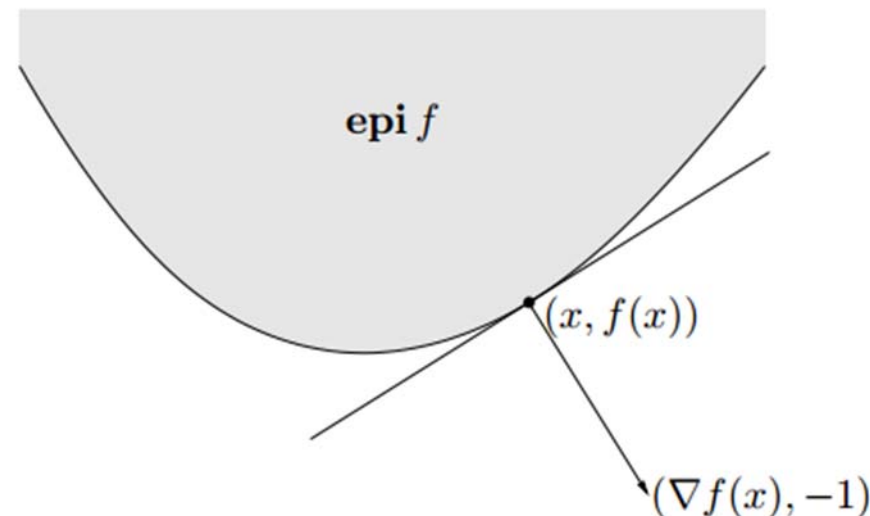


Representation

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

$f: \mathbf{R}^n \rightarrow \mathbf{R}$ is convex and $\text{dom } f = \mathbf{R}^n$

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





Compositions

□ Definition

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

- $f = h \circ g: \mathbf{R}^n \rightarrow \mathbf{R}$

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

- $\text{dom } f = \{x \in \text{dom } g \mid g(x) \in \text{dom } h\}$

□ Chain Rule

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

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



Scalar Composition

□ $h: \mathbf{R} \rightarrow \mathbf{R}, g: \mathbf{R} \rightarrow \mathbf{R}$

■ h and g are twice differentiable

■ $\text{dom } g = \text{dom } h = \mathbf{R}$

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

■ 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

□ $h: \mathbf{R} \rightarrow \mathbf{R}, g: \mathbf{R} \rightarrow \mathbf{R}$

■ h and g are twice differentiable

■ $\text{dom } g = \text{dom } h = \mathbf{R}$

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

■ 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

□ $h: \mathbf{R} \rightarrow \mathbf{R}, g: \mathbf{R}^n \rightarrow \mathbf{R}$

- Without differentiability assumption
 - Without domain condition
 - $h(x) = 0$ with $\text{dom } h = [1, 2]$, which is convex and non-decreasing
 - $g(x) = x^2$ with $\text{dom } g = \mathbf{R}$, which is convex
- $$f(x) = h(g(x)) = 0$$
- $\text{dom } f = [-\sqrt{2}, -1] \cup [1, \sqrt{2}]$



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

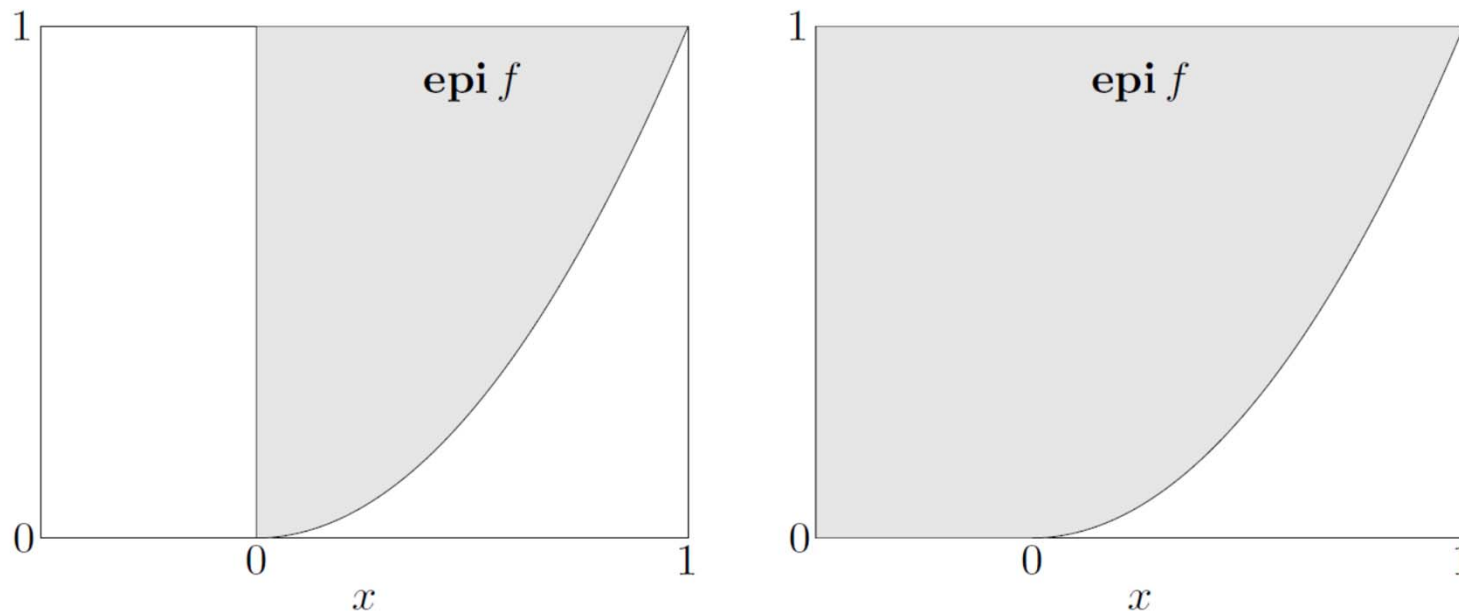


Figure 3.7 *Left.* The function x^2 , with domain \mathbf{R}_+ , is convex and nondecreasing on its domain, but its extended-value extension is *not* nondecreasing. *Right.* The function $\max\{x, 0\}^2$, with domain \mathbf{R} , is convex, and its extended-value extension is nondecreasing.



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 concave
- g is convex $\Rightarrow -\log(-g(x))$ is convex on $\{x | g(x) < 0\}$



Vector Composition

□ $h: \mathbf{R}^k \rightarrow \mathbf{R}, g_i: \mathbf{R} \rightarrow \mathbf{R}$

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

■ h and g are twice differentiable

■ $\text{dom } g = \mathbf{R}, \text{dom } h = \mathbf{R}^k$

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

■ 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

□ $h: \mathbf{R}^k \rightarrow \mathbf{R}, g_i: \mathbf{R} \rightarrow \mathbf{R}$

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

■ h and g are twice differentiable

■ $\text{dom } g = \mathbf{R}, \text{dom } h = \mathbf{R}^k$

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

■ f is concave, if $f''(x) \leq 0$

✓ h is concave, h is nondecreasing in each argument, and g_i are concave

□ The general case is similar



Examples

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



Minimization

□ f is convex in (x, y) , \mathcal{C} is convex ($\mathcal{C} \neq \emptyset$)

■ $g(x) = \inf_{y \in \mathcal{C}} f(x, y)$ is convex if $g(x) > -\infty, \forall x \in \text{dom } g$

■ $\text{dom } g = \{x | (x, y) \in \text{dom } f \text{ for some } y \in \mathcal{C}\}$

□ Proof by Epigraph

■ $\text{epi } g = \{(x, t) | (x, y, t) \in \text{epi } f \text{ for some } y \in \mathcal{C}\}$

■ The projection of a convex set is convex.



Minimization

□ Definition

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

□ Proof

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



Examples

□ Schur complement

- $f(x, y) = x^T A x + 2x^T B y + y^T C y$
- $\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succcurlyeq 0, A, C \text{ is symmetric} \Rightarrow f(x, y) \text{ is convex}$
- $g(x) = \inf_y f(x, y) = x^T (A - B C^\dagger B^T) x$ is convex
 $\Rightarrow A - B C^\dagger B^T \succcurlyeq 0, C^\dagger \text{ is the pseudo-inverse of } C$

□ Distance to a set

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



Examples

□ Affine domain

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

□ Proof

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



Perspective of a function

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

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

is the perspective of f

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

□ Proof

$$\begin{aligned}(x, t, s) \in \text{epi } g &\Leftrightarrow tf\left(\frac{x}{t}\right) \leq s \\ &\Leftrightarrow f\left(\frac{x}{t}\right) \leq \frac{s}{t} \\ &\Leftrightarrow (x/t, s/t) \in \text{epi } f\end{aligned}$$

- Perspective operation preserve convexity



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

- $f: \mathbf{R}^m \rightarrow \mathbf{R}$ is convex

- $A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m, c \in \mathbf{R}^n, d \in \mathbf{R}$

- $\text{dom } g = \left\{ x \mid c^\top x + d > 0, \frac{Ax+b}{c^\top x+d} \in \text{dom } f \right\}$

- $g(x) = (c^\top x + d)f\left(\frac{Ax+b}{c^\top x+d}\right)$ 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