

Convex Functions

Yuanming Shi

ShanghaiTech University

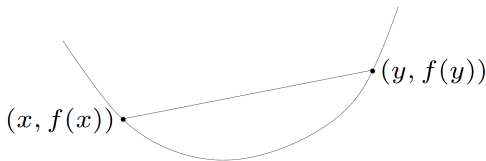
Outline

- 1 Definition of Convex Function
- 2 Restriction of a Convex Function to a Line
- 3 First and Second Order Conditions
- 4 Operations that Preserve Convexity
- 5 Quasi-Convexity, Log-Convexity, and Convexity w.r.t. Generalized Inequalities

Definition of Convex Function

- A function $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ is said to be **convex** if the domain, $\text{dom } f$, is convex and for any $x, y \in \text{dom } f$ and $0 \leq \theta \leq 1$,

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



- f is **strictly convex** if the inequality is strict for $0 < \theta < 1$
- f is **concave** if $-f$ is convex

Examples on \mathbb{R}

Convex functions:

- affine: $ax + b$ on \mathbb{R}
- powers of absolute value: $|x|^p$ on \mathbb{R} , for $p \geq 1$ (e.g., $|x|$)
- powers: x^p on \mathbb{R}_{++} , for $p \geq 1$ or $p \leq 0$ (e.g., x^2)
- exponential: e^{ax} on \mathbb{R}
- negative entropy: $x \log x$ on \mathbb{R}_{++}

Concave functions:

- affine: $ax + b$ on \mathbb{R}
- powers: x^p on \mathbb{R}_{++} , for $0 \leq p \leq 1$
- logarithm: $\log x$ on \mathbb{R}_{++}

Examples on \mathbb{R}^n

- **Affine functions** $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b$ are convex and concave on \mathbb{R}^n
- **Norms** $\|\mathbf{x}\|$ are convex on \mathbb{R}^n (e.g., $\|\mathbf{x}\|_\infty, \|\mathbf{x}\|_1, \|\mathbf{x}\|_2$)
- **Quadratic functions** $f(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x} + 2\mathbf{q}^T \mathbf{x} + r$ are convex \mathbb{R}^n if and only if $\mathbf{P} \succeq \mathbf{0}$
- The **geometric mean** $f(\mathbf{x}) = (\prod_{i=1}^n x_i)^{1/n}$ is concave on \mathbb{R}_{++}^n
- The **log-sum-exp** $f(\mathbf{x}) = \log \sum_i e^{x_i}$ is convex on \mathbb{R}^n (it can be used to approximate $\max_{i=1, \dots, n} x_i$)
- **Quadratic over linear:** $f(\mathbf{x}, y) = \mathbf{x}^T \mathbf{x} / y$ is convex on $\mathbb{R}^n \times \mathbb{R}_{++}$

Examples on $\mathbb{R}^{n \times n}$

- **Affine functions:** (prove it!)

$$f(\mathbf{X}) = \text{Tr}(\mathbf{A}\mathbf{X}) + b$$

are convex and concave on $\mathbb{R}^{n \times n}$

- **Logarithmic determinant function:** (prove it!)

$$f(\mathbf{X}) = \log \det(\mathbf{X})$$

is concave on $\mathbb{S}^n = \{\mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{X} \succeq \mathbf{0}\}$

- **Maximum eigenvalue function:** (prove it!)

$$f(\mathbf{x}) = \lambda_{\max}(\mathbf{X}) = \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\mathbf{y}^T \mathbf{X} \mathbf{y}}{\mathbf{y}^T \mathbf{y}}$$

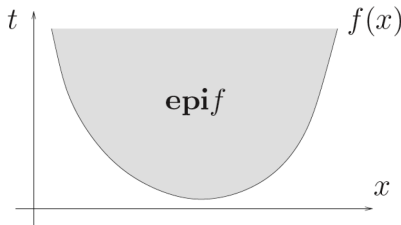
is convex on \mathbb{S}^n

Epigraph

- The **epigraph** of f is the set

$$\text{epi } f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \mathbf{x} \in \text{dom } f, f(\mathbf{x}) \leq t\}$$

- Relation between convexity in sets and convexity in functions:
 f is convex \iff $\text{epi } f$ is convex



Outline

- 1 Definition of Convex Function
- 2 Restriction of a Convex Function to a Line**
- 3 First and Second Order Conditions
- 4 Operations that Preserve Convexity
- 5 Quasi-Convexity, Log-Convexity, and Convexity w.r.t. Generalized Inequalities

Restriction of a Convex Function to a Line

🔔 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if and only if the function $g: \mathbb{R} \rightarrow \mathbb{R}$

$$g(t) = f(\mathbf{x} + t\mathbf{v}), \quad \text{dom } g = \{t \mid \mathbf{x} + t\mathbf{v} \in \text{dom } f\}$$

is convex for any $\mathbf{x} \in \text{dom } f, \mathbf{v} \in \mathbb{R}^n$

- 🔔 In words: a function is convex if and only if it is convex when restricted to an arbitrary line.
- 🔔 Implication: we can check convexity of f by checking convexity of functions of one variable!
- 🔔 Example: concavity of $\log \det(\mathbf{X})$ follows from concavity of $\log(x)$

Example

Example: concavity of $\log\det(\mathbf{X})$:

$$\begin{aligned} g(t) = \log\det(\mathbf{X} + t\mathbf{V}) &= \log\det(\mathbf{X}) + \log\det(\mathbf{I} + t\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2}) \\ &= \log\det(\mathbf{X}) + \sum_{i=1}^n \log(1 + t\lambda_i) \end{aligned}$$

where λ_i 's are the eigenvalues of $\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2}$.

The function g is concave in t for any choice of $\mathbf{X} \succ \mathbf{0}$ and \mathbf{V} ; therefore, f is concave.

Outline

- 1 Definition of Convex Function
- 2 Restriction of a Convex Function to a Line
- 3 First and Second Order Conditions**
- 4 Operations that Preserve Convexity
- 5 Quasi-Convexity, Log-Convexity, and Convexity w.r.t. Generalized Inequalities

First and Second Order Conditions I

• **Gradient** (for differentiable f):

$$\nabla f(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1} \quad \dots \quad \frac{\partial f(\mathbf{x})}{\partial x_n} \right]^T \in \mathbb{R}^n$$

• **Hessian** (for twice differentiable f):

$$\nabla^2 f(\mathbf{x}) = \left(\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \right)_{ij} \in \mathbb{R}^{n \times n}$$

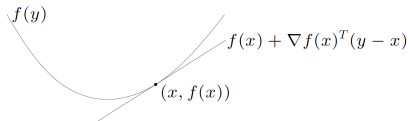
• **Taylor series:**

$$f(\mathbf{x} + \boldsymbol{\delta}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T \boldsymbol{\delta} + \frac{1}{2} \boldsymbol{\delta}^T \nabla^2 f(\mathbf{x}) \boldsymbol{\delta} + o(\|\boldsymbol{\delta}\|^2)$$

First and Second Order Conditions II

- **First-order condition:** a differentiable f with convex domain is convex if and only if

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



- Interpretation: first-order approximation is a global under estimator
- **Second-order condition:** a twice differentiable f with convex domain is convex if and only if

$$\nabla^2 f(\mathbf{x}) \succeq \mathbf{0} \quad \forall \mathbf{x} \in \text{dom } f$$

Examples

• **Quadratic function:** $f(x) = \frac{1}{2}x^T Px + q^T x + r$ (with $P \in \mathbb{S}^n$)

$$\nabla f(x) = Px + q, \quad \nabla^2 f(x) = P$$

is convex if $P \succeq 0$.

• **Least-squares objective:** $f(x) = \|Ax - b\|_2^2$

$$\nabla f(x) = 2A^T(Ax - b), \quad \nabla^2 f(x) = 2A^T A$$

is convex.

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

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

is convex for $y > 0$.

Outline

- 1 Definition of Convex Function
- 2 Restriction of a Convex Function to a Line
- 3 First and Second Order Conditions
- 4 Operations that Preserve Convexity**
- 5 Quasi-Convexity, Log-Convexity, and Convexity w.r.t. Generalized Inequalities

Operations that Preserve Convexity I

How to establish the convexity of a given function?

- Applying the definition
- With first- or second-order conditions
- By restricting to a line
- Showing that the functions can be obtained from simple functions by operations that preserve convexity:
 - nonnegative weighted sum
 - composition with affine function (and other compositions)
 - pointwise maximum and supremum, minimization
 - perspective

Operations that Preserve Convexity II

- **Nonnegative weighted sum:** if f_1, f_2 are convex, then $\alpha_1 f_1 + \alpha_2 f_2$ is convex, with $\alpha_1, \alpha_2 \geq 0$.
- **Composition with affine functions:** if f is convex, then $f(\mathbf{Ax} + \mathbf{b})$ is convex (e.g., $\|\mathbf{y} - \mathbf{Ax}\|$ is convex, $\log\det(\mathbf{I} + \mathbf{H}\mathbf{X}\mathbf{H}^T)$ is concave).
- **Pointwise maximum:** $f := \max\{f_1, \dots, f_m\}$ is convex, if f_1, \dots, f_m are convex

Example: sum of r largest components of $\mathbf{x} \in \mathbb{R}^n$:

$$f(\mathbf{x}) = x_{[1]} + x_{[2]} + \dots + x_{[r]}$$

where $x_{[i]}$ is the i th largest component of \mathbf{x} .

Proof: $f(\mathbf{x}) = \max\{x_{i_1} + x_{i_2} + \dots + x_{i_r} \mid 1 \leq i_1 < i_2 < \dots < i_r \leq n\}$.

Operations that Preserve Convexity III

• **Pointwise supremum:** if $f(\mathbf{x}, \mathbf{y})$ is convex in \mathbf{x} for each $\mathbf{y} \in \mathcal{A}$, then

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

is convex.

Example: distance to farthest point in a set C :

$$f(\mathbf{x}) = \sup_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|$$

Example: maximum eigenvalue of symmetric matrix: for $\mathbf{X} \in \mathbb{S}^n$,

$$\lambda_{\max}(\mathbf{X}) = \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\mathbf{y}^T \mathbf{X} \mathbf{y}}{\mathbf{y}^T \mathbf{y}}$$

Operations that Preserve Convexity IV

- **Composition with scalar functions:** let $g : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R} \rightarrow \mathbb{R}$, then the function $f(\mathbf{x}) = h(g(\mathbf{x}))$ satisfies:

$$f(\mathbf{x}) \text{ is convex if } \begin{cases} g \text{ convex, } h \text{ convex nondecreasing} \\ g \text{ concave, } h \text{ convex nonincreasing} \end{cases}$$

- **Minimization:** if $f(\mathbf{x}, \mathbf{y})$ is convex in (\mathbf{x}, \mathbf{y}) and C is a convex set, then

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

is convex (e.g., distance to a convex set).

Example: distance to a set C :

$$f(\mathbf{x}) = \inf_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|$$

is convex if C is convex.

Operations that Preserve Convexity V

• **Perspective:** if $f(\mathbf{x})$ is convex, then its perspective

$$g(\mathbf{x}, t) = tf(\mathbf{x}/t), \quad \text{dom } g = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} | \mathbf{x}/t \in \text{dom } f, t > 0\}$$

is convex.

Example: $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$ is convex; hence $g(\mathbf{x}, t) = \mathbf{x}^T \mathbf{x}/t$ is convex for $t > 0$.

Example: the negative logarithm $f(\mathbf{x}) = -\log \mathbf{x}$ is convex; hence the relative entropy function $g(\mathbf{x}, t) = t \log t - t \log \mathbf{x}$ is convex on \mathbb{R}_{++}^2 .

Outline

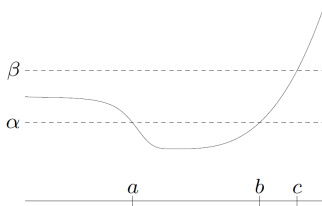
- 1 Definition of Convex Function
- 2 Restriction of a Convex Function to a Line
- 3 First and Second Order Conditions
- 4 Operations that Preserve Convexity
- 5 Quasi-Convexity, Log-Convexity, and Convexity w.r.t. Generalized Inequalities**

Quasi-Convexity Functions

- A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is quasi-convex if $\text{dom } f$ is convex and the sublevel sets

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

are convex for all α .



- f is quasiconcave if $-f$ is quasiconvex.

Examples

• $\sqrt{|x|}$ is quasiconvex on \mathbb{R}

• $\text{ceil}(x) = \inf\{z \in \mathbb{Z} \mid z \geq x\}$ is quasilinear

• $\log x$ is quasilinear on \mathbb{R}_{++}

• $f(x_1, x_2) = x_1 x_2$ is quasiconcave on \mathbb{R}_{++}^2

• the linear-fractional function

$$f(\mathbf{x}) = \frac{\mathbf{a}^T \mathbf{x} + b}{\mathbf{c}^T \mathbf{x} + d}, \quad \text{dom } f = \{\mathbf{x} \mid \mathbf{c}^T \mathbf{x} + d > 0\}$$

is quasilinear

Log-Convexity

- A positive function f is log-concave if $\log f$ is concave:

$$f(\theta \mathbf{x} + (1 - \theta) \mathbf{y}) \geq f(\mathbf{x})^\theta f(\mathbf{y})^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$

- f is log-convex if $\log f$ is convex.
- Example: x^a on \mathbb{R}_{++} is log-convex for $a \leq 0$ and log-concave for $a \geq 0$
- Example: many common probability densities are log-concave

Convexity w.r.t. Generalized Inequalities

- $\mathbf{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is K -convex if $\text{dom } \mathbf{f}$ is convex and for any $\mathbf{x}, \mathbf{y} \in \text{dom } \mathbf{f}$ and $0 \leq \theta \leq 1$,

$$\mathbf{f}(\theta \mathbf{x} + (1 - \theta) \mathbf{y}) \preceq_K \theta \mathbf{f}(\mathbf{x}) + (1 - \theta) \mathbf{f}(\mathbf{y})$$

- Example: $\mathbf{f} : \mathbb{S}^m \longrightarrow \mathbb{S}^m$, $\mathbf{f}(\mathbf{X}) = \mathbf{X}^2$ is \mathbb{S}_+^m -convex

Reference

Chapter 3 of:

- Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

Book:

- Petersen, Kaare Brandt, and Michael Syskind Pedersen. "The matrix cookbook." Technical University of Denmark 7 (2008): 15.