Convex Functions

Prof. Daniel P. Palomar

ELEC5470/IEDA6100A - Convex Optimization

The Hong Kong University of Science and Technology (HKUST)

Fall 2020-21

Outline of Lecture

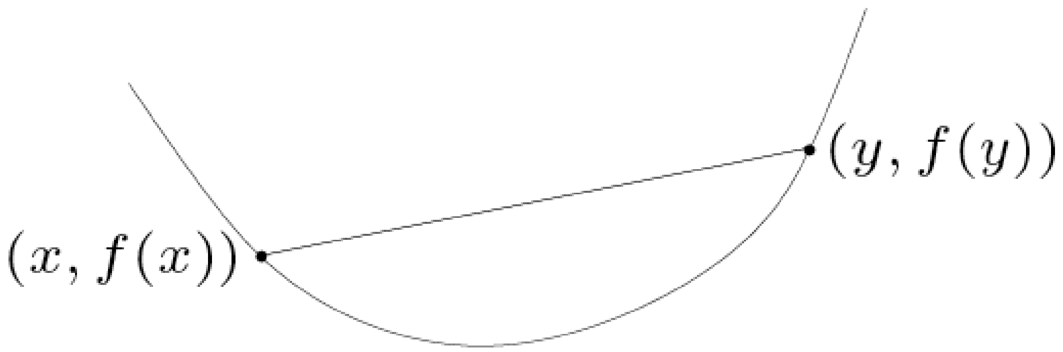
- Definition convex function
- Examples on $\mathbf{R}$, $\mathbf{R}^n$, and $\mathbf{R}^{n\times n}$
- Restriction of a convex function to a line
- First- and second-order conditions
- Operations that preserve convexity
- Quasi-convexity, log-convexity, and convexity w.r.t. generalized inequalities

(Acknowledgement to Stephen Boyd for material for this lecture.)

Definition of Convex Function

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

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



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

- f is concave if $-f$ is convex.

Examples on $\mathbf{R}$

Convex functions:

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

Concave functions:

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

Examples on $\mathbf{R}^n$

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

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

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

$$f(X) = \mathsf{Tr}(AX) + b$$

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

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

$$f(X) = \mathsf{logdet}(X)$$

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

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

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

  is convex on $\mathbf{S}^n$.

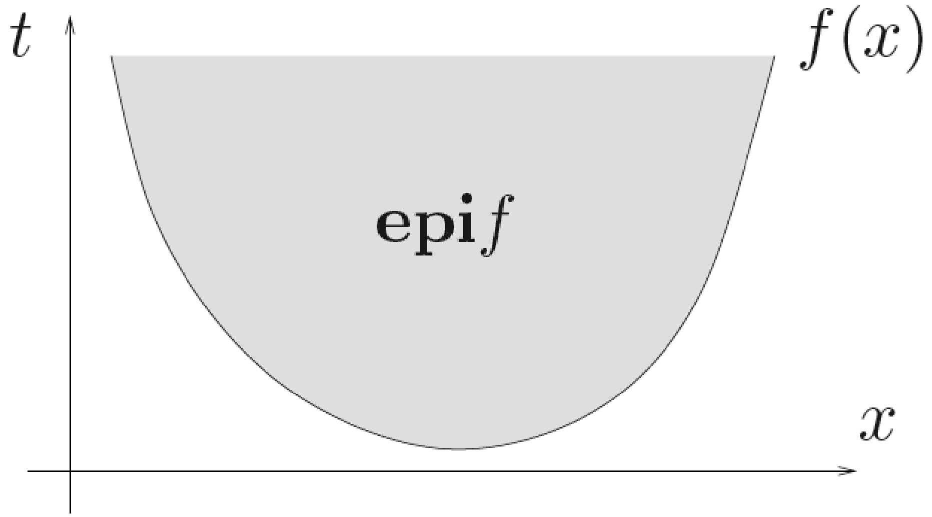
Epigraph

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

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

- Relation between convexity in sets and convexity in functions:

$$f \text{ is convex} \iff \text{epi}\, f \text{ is convex}$$



Restriction of a Convex Function to a Line

- $f : \mathbf{R}^n \longrightarrow \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \longrightarrow \mathbf{R}$

$$g(t) = f(x + tv), \qquad \mathbf{dom}\, g = \{t \mid x + tv \in \mathbf{dom}\, f\}$$

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

- 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 $\log\det(X)$ follows from concavity of $\log(x)$.

Example: concavity of logdet (X) :

$$g(t) = \text{logdet}(X + tV) = \text{logdet}(X) + \text{logdet}\left(I + tX^{-1/2}VX^{-1/2}\right)$$
$$= \text{logdet}(X) + \sum_{i=1}^{n} \log(1 + t\lambda_i)$$

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

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

First and Second Order Condition

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

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} & \cdots & \frac{\partial f(x)}{\partial x_n} \end{bmatrix}^T \in \mathbf{R}^n.$$

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

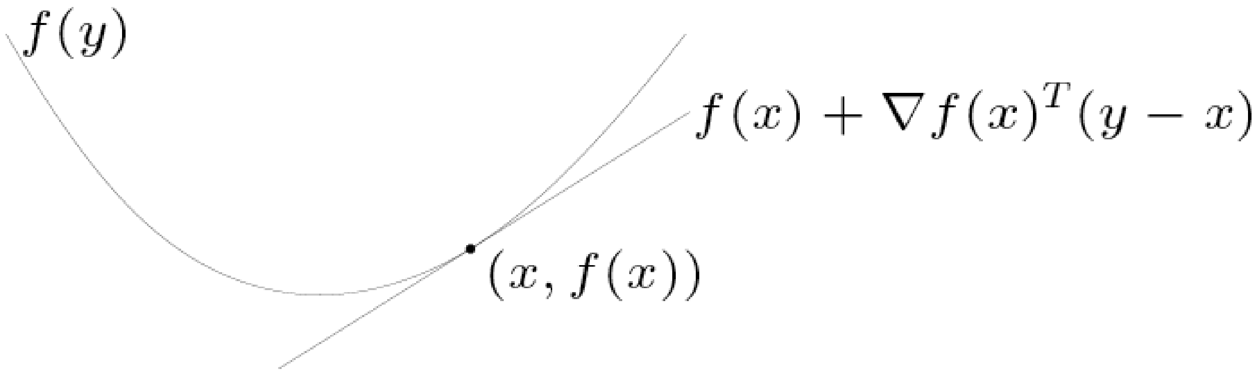
$$\nabla^2 f(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{ij} \in \mathbf{R}^{n \times n}.$$

- Taylor series:

$$f(x+\delta) = f(x) + \nabla f(x)^T \delta + \frac{1}{2}\delta^T \nabla^2 f(x)\,\delta + o\left(\|\delta\|^2\right).$$

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

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



- Interpretation: first-order approximation if a global underestimator.

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

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

Examples

- **Quadratic function**: $f(x) = (1/2)\, x^T P x + q^T x + r$ (with $P \in \mathbf{S}^n$)

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

is convex if $P \succeq 0$.

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

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

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 $y > 0$.

Operations that Preserve Convexity

How do we establish the convexity of a given function?

1. Applying the definition.

2. With first- or second-order conditions.

3. By restricting to a line.

4. 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

- **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 $f(Ax + b)$ is convex (e.g., $\|y - Ax\|$ is convex, logdet $\left(I + HXH^T\right)$ is concave).

- **Pointwise maximum**: if $f_1, \ldots, f_m$ are convex, then $f(x) = \max\{f_1, \ldots, f_m\}$ is convex.

Example: sum of r largest components of $x \in \mathbf{R}^n$: $f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$ where $x_{[i]}$ is the ith largest component of x.

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

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

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

is convex.

Example: distance to farthest point in a set C:

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

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

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

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

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

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

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

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

Example: distance to a set C:

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

is convex if C is convex.

- **Perspective**: if $f(x)$ is convex, then its perspective

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

  is convex.

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

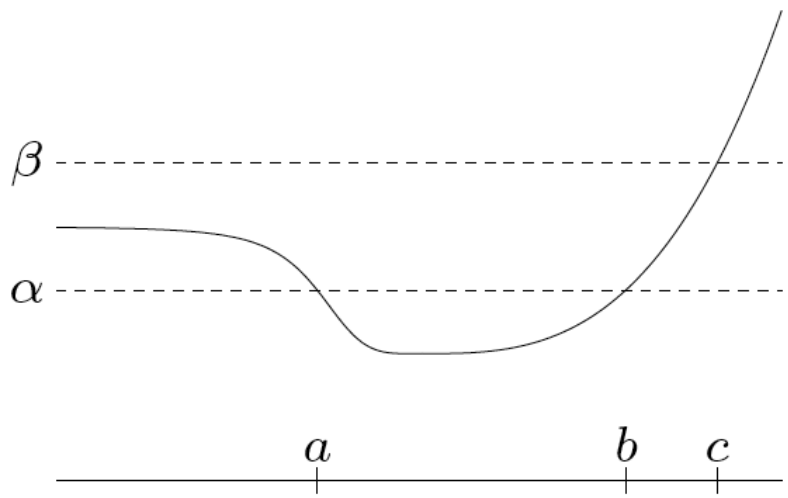
Example: the negative logarithm $f(x) = -\log x$ is convex; hence the relative entropy function $g(x,t) = t\log t - t\log x$ is convex on $\mathbf{R}^2_{++}$.

Quasi-Convexity Functions

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

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

are convex for all α.



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

Examples:

- $\sqrt{|x|}$ is quasiconvex on $\mathbf{R}$
- ceil $(x) = \inf \{z \in \mathbf{Z} \mid z \geq x\}$ is quasilinear
- $\log x$ is quasilinear on $\mathbf{R}_{++}$
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\mathbf{R}^2_{++}$
- the linear-fractional function

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

  is quasilinear

Log-Convexity

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

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

- f is log-convex if $\log f$ is convex.

- Example: x^a on $\mathbf{R}_{++}$ 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

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

$$f\left(\theta x + (1-\theta)\, y\right) \preceq_K \theta f\left(x\right) + \left(1-\theta\right) f\left(y\right).$$

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

References

Chapter 3 of

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

  https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf