IE 521 Convex Optimization

Lecture 4: Convex Functions

Niao He

16th April 2019

Outline

Which function is different from others?

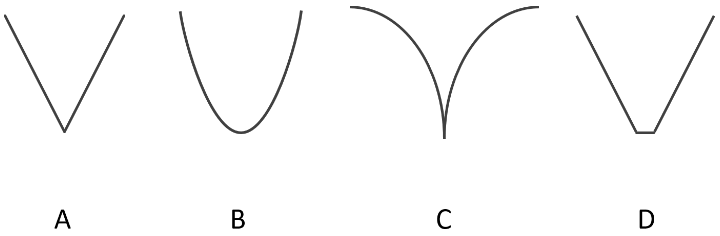

Figure: Functions

Definition of Convex Function

Definition. A function $f(x) : \mathbb{R}^n \to \mathbb{R}$ is convex if

(i) $\text{dom}(f) := \{x \in \mathbb{R}^n : |f(x)| < \infty\}$ is a convex set;

(ii) $\forall x, y \in dom(f)$ and $\lambda \in [0, 1]$,

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

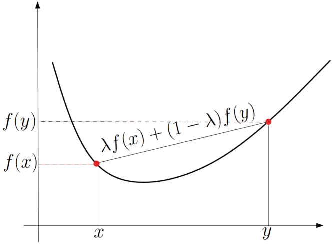

Figure: Convex function

- Geometrically, the line segment between $(x, f(x))$, $(y, f(y))$ sits above the graph of f.

Brotherhood Definitions

Definition. (Strict and Strong Convex)

- A function is called *strictly convex* if (ii) holds with strict sign, i.e.,$\forall\lambda \in (0,1)$,

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y).$$

- A function is called *α-strongly convex* ($\alpha > 0$) if $f(x) - \frac{\alpha}{2}\|x\|_2^2$ is convex, i.e., $\forall\lambda \in [0,1]$,

$$f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)-\frac{\alpha}{2}\lambda(1-\lambda)||x-y||_2^2.$$

Note that strongly convex $\Longrightarrow$ strictly convex $\Longrightarrow$ convex

Definition. (Concave/Strictly Concave/Strongly Concave)

- A function is called concave if $-f(x)$ is convex.
- Similarly for strict concavity and strong concavity.

Examples of Convex Functions

Example 1. Simple univariate functions:

- Even powers: x^p, p is even
- Exponential: $e^{ax}, \forall a \in \mathbb{R}$
- Negative logarithmic: $-\log x$
- Absolute value: $|x|$
- Negative entropy: $x \log(x)$

Example 2. Affine functions:

$$f(x) = a^T x + b$$

- both convex & concave, but not strictly convex/concave

Examples of Convex Functions

Example 3. Norms:

- l_p-norm on $\mathbb{R}^n$:

$$\| x \|_p := (\sum_{i=1}^n |x_i|^p)^{1/p} \qquad (p \geq 1)$$

- Q-norm on $\mathbb{R}^n$:

$$\| x \|_Q := \sqrt{x^T Q x} \qquad (Q \succ 0)$$

- Frobenius norm on $\mathbb{R}^{m\times n}$:

$$\| A \|_F = (\sum_{i=1}^m \sum_{j=1}^n |A_{ij}|^2)^{1/2}$$

- Spectral and nuclear norms on $\mathbb{R}^{m\times n}$:

$$\| A \| = \max_{i=1,\dots,\min\{m.n\}} \sigma_i(A)$$

$$\| A \|_* = \sum_{i=1,\dots,\min\{m.n\}} \sigma_i(A)$$

Examples of Convex Functions

Example 4. Some quadratic functions:

$$f(x) = \frac{1}{2}x^T Qx + b^T x + c$$

- convex if and only if $Q \succeq 0$ is positive semi-definite
- strictly convex if and only if $Q \succ 0$ is positive definite

Examples of Convex Functions

Example 5. Indicator function:

$$I_C(x) = \begin{cases} 0, x \in C \\ \infty, x \notin C \end{cases}$$

- $I_C(x)$ is convex if the set C is a convex set. (why?)

Example 6. Supporting function:

$$I_C^*(x) = \sup_{y \in C} x^T y$$

- $I_C^*(x)$ is always convex for any set C. (why?)

Examples of Convex Functions

Example 7. More examples

- Piecewise linear functions: $\max(a_1^T x + b_1, ..., a_k^T x + b_k)$
- Log of exponential sums: $\log(\sum_{i=1}^k e^{a_i^T x + b_i})$
- Negative log of determinant: $-\log(\det(X))$

Q. How to show convexity of these functions?

Convexity-Preserving Operators

- Taking conic combination;
- Taking affine composition;
- Taking pointwise maximum and supremum;
- Taking convex monotone composition;
- Taking partial minimization;
- Taking the perspective transformation;

Taking Conic Combination

Proposition. If $f_i(x), i \in I$ are convex functions and $\alpha_i \geq 0, \forall i \in I$, then so is

$$g(x) = \sum_{i \in I} \alpha_i f_i(x).$$

Remark. (Extension to integrals) If $f(x, \omega)$ is convex in x and $\alpha(\omega) \geq 0, \forall \omega \in \Omega$, then so is

$$g(x) = \int_\Omega \alpha(\omega) f(x, \omega) d\omega$$

Example 8. If η is a well-defined random variable on Ω, and $f(x, \eta(\omega))$ is convex, $\forall \omega \in \Omega$, then $\mathbb{E}_\eta [f(x, \eta)]$ is convex.

Taking Affine Composition

Proposition. If $f(x)$ is convex and $\mathcal{A}(y) : y \mapsto Ay + b$ is an affine mapping, then so is

$$g(y) := f(Ay + b).$$

Example 9. The following functions are convex:

- $f(x) = \| Ax - b \|_2^2$,
- $f(x) = \sum_i e^{a_i^T x - b_i}$,
- $f(x) = -\sum_{i=1}^n \log(a_i^T x - b_i)$.

Taking Pointwise Maximum and Supremum

Proposition. If $f_i(x), i \in I$ are convex, then so is

$$g(x) := \max_{i \in I} f_i(x).$$

Remark. (Extension to pointwise supremum) If $f(x, \omega)$ is convex in x, for $\omega \in \Omega$, then so is

$$g(x) := \sup_{\omega \in \Omega} f(x, \omega).$$

Example 10. The following functions are convex:

- $g(x) = max(a_1^T x + b_1, ..., a_k^T x + b_k)$
- $I_C^*(x) = \sup_{y \in C} x^T y$
- $d_{max}(x, C) = \max_{y \in C} \| y - x \|_2$
- $\lambda_{max}(X) = \max_{\|y\|_2 = 1} y^T X y$

Taking Convex Monotone Composition

Proposition. If $f_i(x), i = 1, \ldots, m$ are convex and $F(y_1, \ldots, y_m)$ is convex and component-wise non-decreasing, then so is

$$g(x) = F(f_1(x), \ldots, f_m(x)).$$

Remark. Taking pointwise maximum is a special case of the above rule by setting $F(y_1, ..., y_m) = \max(y_1, ..., y_m)$,

$$F(f_1(x), ..., f_m(x)) = \max_{i=1,...,m} f_i(x).$$

Example 11.

- $e^{f(x)}$ is convex if f is convex
- $-\log f(x)$ is convex if f is concave
- $\log(\sum_{i=1}^k e^{f_i})$ is convex if f_i are convex.

Taking Convex Monotone Composition

Proposition. If $f_i(x), i = 1, \ldots, m$ are convex and $F(y_1, \ldots, y_m)$ is convex and component-wise non-decreasing, then so is

$$g(x) = F(f_1(x), \ldots, f_m(x)).$$

Proof. By convexity of f_i, we have

$$f_i(\lambda x + (1-\lambda)y) \leq \lambda f_i(x) + (1-\lambda)f_i(y), \forall i, \forall \lambda \in [0,1].$$

Hence, we have for any $x, y \in dom(g)$, $\lambda \in [0,1]$,

$$\begin{aligned} g(\lambda x + (1-\lambda)y) &= F(f_1(\lambda x + (1-\lambda)y), \ldots, f_m(\lambda x + (1-\lambda)y)) \\ &\leq F(\lambda f_1(x) + (1-\lambda)f_1(y), \ldots, \lambda f_m(x) + (1-\lambda)f_m(y)) \\ &\leq \lambda F(f_1(x), \ldots, f_m(x)) + (1-\lambda)F(f_1(x), \ldots, f_m(x)) \\ &= \lambda g(x) + (1-\lambda)g(y) \end{aligned}$$

Taking Partial Minimization

Proposition. If $f(x, y)$ is convex in $(x, y) \in \mathbb{R}^n$ and Y is a convex set, then so is

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

Example 12. The following are convex:

- $d(x, C) = \min_{y \in C} \| x - y \|_2$, where C is convex;
- $g(x) = \inf_y \{h(y) | Ay = x\}$, where h is convex.

Taking Partial Minimization

Proposition. If $f(x, y)$ is convex in $(x, y) \in \mathbb{R}^n$ and Y is a convex set, then so is

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

Proof.

- $\text{dom}(g) = \{x : (x, y) \in \text{dom}(f) \text{ and } y \in C\}$ is a projection of $\text{dom}(f)$, hence is convex.
- Given any x_1, x_2, by definition, for any $\epsilon > 0$, $\exists y_1, y_2 \in Y$ s.t.

$$f(x_1, y_1) \leq g(x_1) + \epsilon/2, \qquad f(x_2, y_2) \leq g(x_2) + \epsilon/2$$

  By convexity of $f(x, y)$, this implies $f(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \leq \lambda g(x_1) + (1-\lambda)g(x_2) + \epsilon.$
- $\forall \epsilon > 0$, $g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g(x_1) + (1-\lambda)g(x_2) + \epsilon$. Letting $\epsilon \to 0$ leads to the convexity of g.

Taking Perspective Function

Proposition. If f is convex, then so is the perspective function

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

where $\mathrm{dom}(g) = \{(x,t) : x/t \in \mathrm{dom}(f), t > 0\}$.

Example 13.

- $g(x,t) = x^T x/t$ is convex on $\mathbb{R}^n \times \mathbb{R}_{++}$;
- $g(x,t) = t \log t - t \log x$ is convex on $\mathbb{R}^2_{++}$;
- $D_{\mathrm{KL}}(x,y) = \sum_i [x_i \log(\frac{x_i}{y_i}) - x_i + y_i]$ on $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$.

Taking Perspective Function

Proposition. If f is convex, then so is the perspective function

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

where $\text{dom}(g) = \{(x,t) : x/t \in \text{dom}(f), t > 0\}$.

Proof.

- $\text{dom}(g)$ is the inverse image of $\text{dom}(f)$ under the perspective function $P(x,t) := x/t$ for $t > 0$.
  So it is convex. (why?)
- Consider $(x,t),(y,s) \in \text{dom}(g)$, and $\lambda \in (0,1)$.

$$\begin{aligned} & g(\lambda x + (1-\lambda)y, \lambda t + (1-\lambda)s) \\ =& (\lambda t + (1-\lambda)s) f\left(\frac{\lambda x + (1-\lambda)y}{\lambda t + (1-\lambda)s}\right) \\ \leq& \lambda t f(x/t) + (1-\lambda) s f(y/s) \\ =& \lambda g(x,t) + (1-\lambda) g(y,s). \end{aligned}$$

Quick Check

Which of the following function is convex:

A. $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}^2_+$

B. $f(x_1, x_2) = \min(x_1, x_2)$ on $\mathbb{R}^2$

C. $f(x_1, x_2) = \frac{x_1}{x_2}$ on $\mathbb{R}^2_{++}$

D. $f(x_1, x_2) = \frac{x_1^2}{x_2}$ on $\mathbb{R} \times \mathbb{R}_{++}$

Quick Check

Which of the following function is not convex:

A. $f(x) = \|x\|$
B. $f(x) = \|x\|^2$
C. $f(x) = \|x\|^3$
D. $f(x) = -\log(\|x\|)$

Application: Inventory Model

- Consider a single period inventory system.
- Let x denote the inventory level and d denote the random demand in that period following distribution $\mathcal{D}$.
- Suppose that the vendor suffers either a holding cost of h dollars per unit for excess inventory or a penalty cost of p dollars per unit for lost demands.
- What's the expected total cost $f(x)$ as a function of x? Is it a convex function?

The cost function

$$f(x) = \mathbb{E}_{d\sim\mathcal{D}}[h \cdot \max(x-d,0) + p \cdot \max(d-x,0)]$$

is a convex function.

References

- Boyd & Vandenberghe, Chapter 3.1-3.2