Convex Optimization and Applications

3 - Convex functions

Guillaume Sagnol

Technische Universität Berlin

Outline

Convex funtion

Definition (Convex funtion).

Let $S \subseteq \mathbb{R}^n$. A function $f : S \to \mathbb{R}$ is *convex* if

- $\textbf{dom}\, f = S$ is convex;
- $\forall \boldsymbol{x}, \boldsymbol{y} \in S,\ \forall \alpha \in [0, 1]$,

$$f((1-\alpha)\boldsymbol{x} + \alpha \boldsymbol{y}) \leq (1-\alpha) f(\boldsymbol{x}) + \alpha f(\boldsymbol{y}).$$

Moreover, f is called *strictly* convex if the above inequality holds strictly for all $\boldsymbol{x} \neq \boldsymbol{y} \in S,\ \alpha \in (0, 1)$.

The function f is (strictly) *concave* if $-f$ is (strictly) convex.

Extended function

Definition (Function extension).

Let f be a function defined over $\mathbf{dom}\, f = S \subseteq \mathbb{R}^n$. The extension of f is $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$,

$$\tilde{f}(\boldsymbol{x}) = \begin{cases} f(x) & \text{if } \boldsymbol{x} \in \mathbf{dom}\, f; \\ +\infty & \text{otherwise.} \end{cases}$$

Working with function extensions is a convenient simplification, e.g.:

Proposition

f is convex iff $\quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n,\ \forall \alpha \in [0, 1],$

$$\tilde{f}((1-\alpha)\boldsymbol{x} + \alpha \boldsymbol{y}) \leq (1-\alpha)\tilde{f}(\boldsymbol{x}) + \alpha \tilde{f}(\boldsymbol{y}).$$

Extended function

- When the extension of a function is given, we recover $\mathbf{dom}\, f := \{\boldsymbol{x} \in \mathbb{R}^n : \tilde{f}(\boldsymbol{x}) < \infty\}$.

Example

Instead of writing

$$f := f_1 + f_2, \text{ with } \mathbf{dom}\, f := \mathbf{dom}\, f_1 \cap \mathbf{dom}\, f_2,$$

we can simply define

$$\tilde{f}(\boldsymbol{x}) := \tilde{f}_1(\boldsymbol{x}) + \tilde{f}_2(\boldsymbol{x}).$$

- We will often replace convex functions by their extension, without writing the "~".
- **Note**: For concave functions, the extension takes the value $-\infty$ outside the domain.

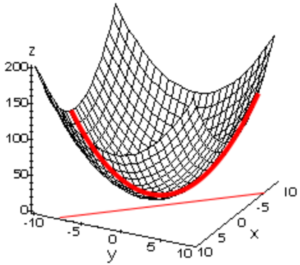
Restriction to a line

Proposition

Let f be a function with **dom** $f \subseteq \mathbb{R}^n$. Then, f is convex if and only if its restriction to any line is convex, i.e., the function

$$g : t \mapsto f(\boldsymbol{x}_0 + t\boldsymbol{u})$$

is convex for all $\boldsymbol{x}_0, \boldsymbol{u} \in \mathbb{R}^n$.



Level sets & Epigraph

Definition

Let f be real-valued with **dom** $f \subseteq \mathbb{R}^n$.

- Its $\alpha-$*sublevel* set is $C_\alpha(f) = \{x \in \mathbb{R}^n : \ f(x) \leq \alpha\}$
- Its $\alpha-$*superlevel* set is $C^\alpha(f) = \{x \in \mathbb{R}^n : \ f(x) \geq \alpha\}$
- The *epigraph* and *hypograph* of f are

$$\begin{aligned} \textbf{epi}\, f &= \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \ f(x) \leq t\} \subseteq \mathbb{R}^{n+1}. \\ \textbf{hypo}\, f &= \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \ f(x) \geq t\} \subseteq \mathbb{R}^{n+1}. \end{aligned}$$

Proposition

f convex $\implies$ $C_\alpha(f)$ convex.

f convex $\iff$ **epi** f convex.

Jensen's inequality

Theorem (Jensen's inequality).

Let f be convex and $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \textbf{dom} f$. Then, for all $\boldsymbol{\lambda} \in \mathbb{R}_+^n$ with $\mathbf{1}^T\boldsymbol{\lambda} = 1$,

$$f\left(\sum_{i=1}^n \lambda_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^n \lambda_i f(\boldsymbol{x}_i).$$

More generally, let X be an integrable random variable with support in $\textbf{dom} f$, i.e., $\mathbb{P}[X \in \textbf{dom} f] = 1$. Then,

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

First order condition

Theorem (1st order condition for convexity).

Let f be differentiable at all points of its domain,and assume that $\mathbf{dom}\, f \subseteq \mathbb{R}^n$ is convex. Then, f is convex iff

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

Recall:

- 1st order Taylor approx = global underestimator
- Local information gives bounds everywhere !

Second order condition

Theorem (Second order conditions).

Let f be twice differentiable at all points of its domain, and assume that $\mathbf{dom}\, f \subseteq \mathbb{R}^n$ is convex. Then, f is convex iff

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

Moreover, if $\nabla^2 f(\boldsymbol{x}) \succ 0$ for all $\boldsymbol{x} \in \mathbf{dom}\, f$, then f is strictly convex.

Note: The converse of the second statement is not necessarily true: $f : x \mapsto x^4$ is strictly convex over $\mathbb{R}$, but the second derivative $f''(x) = 12x^2$ vanishes at $x = 0$.

Examples

1. affine $\implies$ both convex and concave;
2. $x \mapsto e^{ax}$ is convex on $\mathbb{R}$, for all $a \in \mathbb{R}$;
3. $x \mapsto x^a$ is
   - convex on $\mathbb{R}_+$, for all $a \geq 1$;
   - concave on $\mathbb{R}_+$ for all $a \in (0, 1]$;
   - convex on $\mathbb{R}_{++}$ for all $a \leq 0$.
4. $x \mapsto \log(x)$ is concave over $\mathbb{R}_{++}$.
5. $x \mapsto x \log(x)$ is convex over $\mathbb{R}_+$ (with $0 \log 0 := 0$).
6. $\boldsymbol{x} \mapsto \|\boldsymbol{x}\|$ is convex over $\mathbb{R}^n$ (for *ANY* norm !).
7. The squared norm $\boldsymbol{x} \mapsto \|\boldsymbol{x}\|^2$ is also convex;
8. $\boldsymbol{x} \mapsto \max(x_1, \ldots, x_n)$ is convex over $\mathbb{R}^n$;

Examples

9 The log-sum-exp function $\boldsymbol{x} \mapsto \log\left(e^{x_1} + \ldots + e^{x_n}\right)$ is convex over $\mathbb{R}^n$.

10 The quadratic function $\boldsymbol{x} \mapsto \boldsymbol{x}^T Q \boldsymbol{x} + \boldsymbol{a}^T \boldsymbol{x} + b$ is convex over $\mathbb{R}^n$ if and only if Q is positive semidefinite.

11 The geometric mean $\boldsymbol{x} \mapsto \prod_{i=1}^{n} x_i^{1/n}$ is concave over $\mathbb{R}^n_+$.

12 The harmonic mean $\boldsymbol{x} \mapsto m \left(\sum_{i=1}^{n} x_i^{-1} \right)^{-1}$ is concave over $\mathbb{R}^n_{++}$.

13 The function $X \mapsto \log\det X$ is concave over $\mathbb{S}^n_{++}$.

Simple operations that preserve convexity

- Nonnegative scaling
  $\alpha \geq 0$, f convex $\implies \alpha f$ convex
- Sum
  f_1, f_2 convex $\implies f_1 + f_2$ convex
- Composition with affine mapping
  f convex $\implies \boldsymbol{x} \mapsto f(A\boldsymbol{x} + \boldsymbol{b})$ convex
- Perspective: If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then

$$g : (\boldsymbol{x}, t) \mapsto t\, f\left(\frac{\boldsymbol{x}}{t}\right)$$

  is convex over $\mathbf{dom}\, g = \mathbf{dom}\, f \times \mathbb{R}_{++}$
  (Note that g is a function of $n+1$ variables);

Pointwise maximum and Partial minimization

- Pointwise maximum
  $f_1, \ldots, f_n$ convex $\implies f : x \mapsto \max(f_1(x), \ldots, f_n(x))$ convex.

Pointwise maximum and Partial minimization

- Pointwise maximum
  $f_1, \ldots, f_n$ convex $\implies f : \boldsymbol{x} \mapsto \max(f_1(\boldsymbol{x}), \ldots, f_n(\boldsymbol{x}))$ convex.
  More generally, let $f : X \times Y \to \mathbb{R}$.
  If $\boldsymbol{x} \mapsto f(\boldsymbol{x}, \boldsymbol{y})$ is convex over X, for all $\boldsymbol{y} \in Y$, then

  $$g : \boldsymbol{x} \mapsto \sup_{\boldsymbol{y} \in Y} f(\boldsymbol{x}, \boldsymbol{y})$$

  is convex.

Pointwise maximum and Partial minimization

- Pointwise maximum

  $f_1, \ldots, f_n$ convex $\implies f : \boldsymbol{x} \mapsto \max(f_1(\boldsymbol{x}), \ldots, f_n(\boldsymbol{x}))$ convex.

  More generally, let $f : X \times Y \to \mathbb{R}$.
  If $\boldsymbol{x} \mapsto f(\boldsymbol{x}, \boldsymbol{y})$ is convex over X, for all $\boldsymbol{y} \in Y$, then

  $$g : \boldsymbol{x} \mapsto \sup_{\boldsymbol{y} \in Y} f(\boldsymbol{x}, \boldsymbol{y})$$

  is convex.

- Partial minimization

  If $f : \mathbb{R}^{n+m} \to \mathbb{R}$ is convex on **dom** f (i.e., $f(\boldsymbol{x}, \boldsymbol{y})$ is *jointly* convex in $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{R}^m$). Then,

  $$g : \boldsymbol{x} \mapsto \inf_{\boldsymbol{y} \in \mathbb{R}^m} f(\boldsymbol{x}, \boldsymbol{y})$$

  is convex.

Composition rules

Let $h : \mathbb{R}^k \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^k$, and define $f = h \circ g$.

(i) If h is convex, h is nondecreasing in each argument, and g_i is convex ($\forall i \in [k]$), then f is convex.

(ii) If h is convex, h is nonincreasing in each argument, and g_i is concave ($\forall i \in [k]$), then f is convex.

(iii) If h is concave, h is nondecreasing in each argument, and g_i is concave ($\forall i \in [k]$), then f is concave.

(iv) If h is concave, h is nonincreasing in each argument, and g_i is convex ($\forall i \in [k]$), then f is concave.

Examples

We can use the above rules to show the convexity of:

- $f(x) = -\sum_{i=1}^{m} \log(b_i - \boldsymbol{a}_i^T x)$
- $f(X) = \lambda_{\max}(X)$
- $x \mapsto f(x)^2$, where $f : \mathbb{R} \mapsto \mathbb{R}_+$ is convex
- $f : (x, t) \mapsto \frac{\|x\|^2}{t}$
- $f : x \mapsto \operatorname{dist}(x, S) := \inf_{\boldsymbol{y} \in S} \|x - \boldsymbol{y}\|$ for some convex set S

Convex conjugate

Definition (Convex conjugate).

Let $f : \mathbb{R}^n \to \mathbb{R}$. The *convex conjugate function* of f, also known as *Fenchel conjugate*, is

$$f^* : \boldsymbol{y} \mapsto \sup_{\boldsymbol{x} \in \mathbf{dom}\, f} \langle \boldsymbol{y}, \boldsymbol{x} \rangle - f(\boldsymbol{x}).$$

f^* is implicitly defined with values in $\mathbb{R} \cup \{\infty\}$, so we have

$$\mathbf{dom}\, f^* := \{\boldsymbol{y} \in \mathbb{R}^n : f^*(\boldsymbol{y}) < \infty\}.$$

Properties

Proposition (Properties of convex conjugates).

Let $f : \mathbb{R}^n \to \mathbb{R}$. Then,

- f^* is convex (*even if f is not*).
- If f is convex and the epigraph of f is closed, then

$$f = f^{**}.$$

- *Fenchel-Young inequality*:

$$\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n,\ \langle \boldsymbol{x}, \boldsymbol{y} \rangle \leq f(\boldsymbol{x}) + f^*(\boldsymbol{y}).$$

- If f is differentiable, and x^* solves the equation $\boldsymbol{y} = \nabla f(\boldsymbol{x}^*)$, then

$$f^*(\boldsymbol{y}) = \boldsymbol{x}^{*T} \nabla f(\boldsymbol{x}^*) - f(\boldsymbol{x}^*).$$

Examples

- Let $f(x) = \boldsymbol{a}^T \boldsymbol{x} + b$ be an affine function. The function $\langle \boldsymbol{y}, \boldsymbol{x} \rangle - f(\boldsymbol{x}) = \langle \boldsymbol{y} - \boldsymbol{a}, \boldsymbol{x} \rangle - \boldsymbol{b}$ is unbounded over $\mathbb{R}^n$, unless $\boldsymbol{y} = \boldsymbol{a}$. Hence,

$$\mathbf{dom}\, f^* = \{\boldsymbol{a}\}, \text{ with } f^*(\boldsymbol{a}) = -\boldsymbol{b}.$$

- Let f be the strictly quadratic function $\boldsymbol{x} \mapsto \frac{1}{2}\boldsymbol{x}^T Q \boldsymbol{x}$, where $Q \succ 0$ (Q is positive definite). Then, $\forall \boldsymbol{y} \in \mathbb{R}^n$, $\boldsymbol{x} \mapsto \boldsymbol{x}^T \boldsymbol{y} - \frac{1}{2}\boldsymbol{x}^T Q \boldsymbol{x}$ is minimized over $\boldsymbol{x} \in \mathbb{R}^n$ for $\boldsymbol{x} = Q^{-1}\boldsymbol{y}$. Hence,

$$f^*(\boldsymbol{y}) = \frac{1}{2}\boldsymbol{y}^T Q^{-1} \boldsymbol{y}.$$