Convex Functions

Goran Banjac

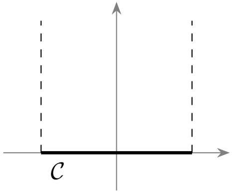
Large-Scale Convex Optimization
ETH Zurich

March 10, 2020

Extended-valued functions and domain

- we consider extended-valued functions $f : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\} =: \overline{\mathbb{R}}$
- **example:** indicator function of a set $\mathcal{C}$

$$\mathcal{I}_{\mathcal{C}}(x) := \begin{cases} 0 & x \in \mathcal{C} \\ +\infty & \text{otherwise} \end{cases}$$



- the (effective) domain of $f : \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ is the set

$$\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$$

- we will always assume $\operatorname{dom} f \neq \emptyset$; such f is called *proper*

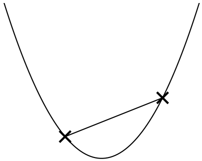
Convex function

- function $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ is **convex** if for all $x, y \in \mathbb{R}^n$ and $\theta \in [0,1]$:

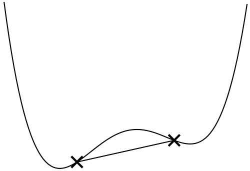
$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$$

  (in extended value arithmetics)

- line connecting $(x, f(x))$ and $(y, f(y))$ is above the graph



convex function



nonconvex function

- f is concave if $-f$ is convex

Examples of convex functions

- examples on $\mathbb{R}$:
  - quadratics: ax^2 for $a \geq 0$
  - exponential: e^{ax} for any $a \in \mathbb{R}$
  - negative logarithm: $-\log x$ on $\mathbb{R}_{++}$
  - negative entropy: $x \log x$ on $\mathbb{R}_{++}$

- examples on $\mathbb{R}^n$:
  - indicator function of a convex set $\mathcal{C}$: $\mathcal{I}_{\mathcal{C}}(x)$
  - norms: $\|x\|$
  - affine function: $f(x) = a^T x + b$

Jensen's inequality

- assume that $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ is convex
- then for all collections $\{x_1, \ldots, x_k\}$ of points

$$f\left(\sum_{i=1}^{k} \theta_i x_i\right) \leq \sum_{i=1}^{k} \theta_i f(x_i)$$

  where $\theta_i \geq 0$ and $\sum_{i=1}^{k} \theta_i = 1$
- for $k = 2$ this reduces to the convexity definition

Epigraph

- the *graph* of f is the set of pairs $(x, f(x)) \in \mathbb{R}^n \times \overline{\mathbb{R}}$
- the **epigraph** of f is the set:

$$\operatorname{epi} f = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq t\}$$

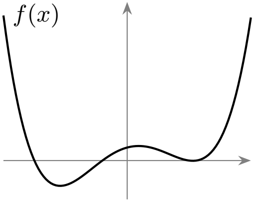
- function $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ is convex if and only if $\operatorname{epi} f$ is a convex set

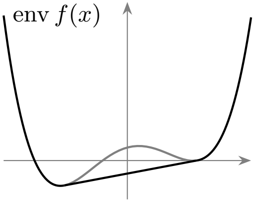
- f is called closed (lower semi-continuous) if $\operatorname{epi} f$ is closed set
- **example:** norm cone

$$\{(x,t) \in \mathbb{R}^{n+1} \mid \|x\| \leq t\}$$

Convex envelope

- **convex envelope** of f is its largest convex minorizer, that is,
  - $\operatorname{env} f$ is convex
  - $\operatorname{env} f \leq f$
  - $\operatorname{env} f \geq g$ for all convex $g \leq f$



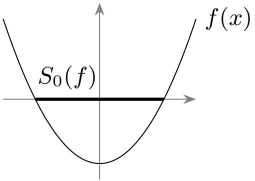


- epigraph of convex envelope of f is convex hull of $\operatorname{epi} f$

Sublevel set

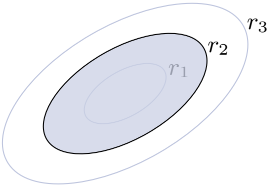
- a **sublevel set** $S_r(f)$ of function $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ is defined as:

$$S_r(f) = \{x \in \mathbb{R}^n \mid f(x) \le r\}$$



- sublevel sets of convex functions are convex
- **example:** norm ball

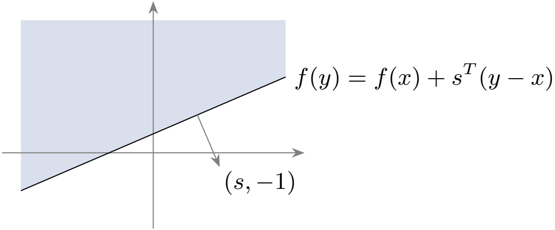
$$\{x \in \mathbb{R}^n \mid \|x\| \le r\}$$



$0 < r_1 < r_2 < r_3$

Affine function

- affine function $f(x) = s^T x + r$ cuts $\mathbb{R}^n \times \mathbb{R}$ in two halves



- s defines slope of function
- for any fixed $x \in \mathbb{R}^n$, $f(y) = s^T y + r$ can be written as

$$f(y) = f(x) + s^T(y - x)$$

- upper halfspace is epigraph with normal vector $(s, -1)$:

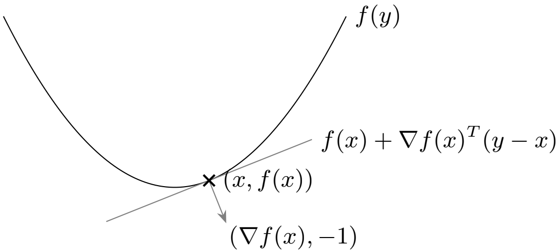
$$\operatorname{epi} f = \{(x, t) \mid s^T x + r \le t\} = \{(x, t) \mid (s, -1)^T (x, t) \le -r\}$$

First-order condition for convexity

- a differentiable function $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ is convex if and only if

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

for all $x, y \in \mathbb{R}^n$



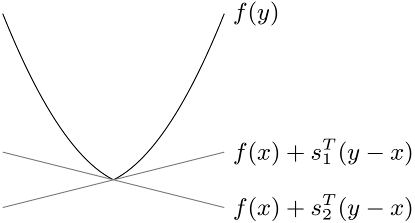
- for all $x \in \mathbb{R}^n$ function f has an affine minorizer that:
  - has slope s given by $\nabla f(x)$
  - coincides with f at x
  - is supporting hyperplane to $\operatorname{epi} f$ with normal $(\nabla f(x), -1)$

Subdifferential

- **subdifferential** of $f : \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ at $x \in \mathbb{R}^n$ is the set of vectors $s \in \mathbb{R}^n$ satisfying

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

  for all $y \in \mathbb{R}^n$
- notation:
  - subdifferential ∂f
  - any element $s \in \partial f(x)$ is called *subgradient* of f at x



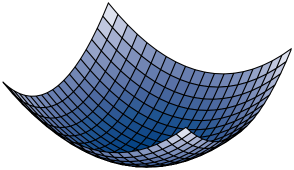
- subgradients define affine minorizers that coincide with f at x
- $s \in \partial f(x)$ if and only if $(s, -1) \in N_{\mathrm{epi}\, f}(x, f(x))$

Second-order condition for convexity

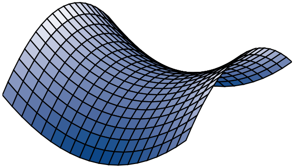
- a twice differentiable function $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ is convex if and only if

$$\nabla^2 f(x) \succeq 0$$

  for all $x \in \mathbb{R}^n$ (*i.e.*, Hessian is positive semidefinite)
- "the function has nonnegative curvature"
- **example:** $f(x) = x^T P x$



$$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



$$P = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

How to conclude convexity

- use convexity definition
- show that epigraph is convex set
- use first or second order condition for convexity
- use convexity preserving operations
  - positive sum
  - composition with affine mapping
  - pointwise supremum
  - partial minimization
  - composition rule

Positive sum and composition with affine mapping

- assume that f_j is convex for $j = \{1, \dots, m\}$
- assume that there exists x such that $f_j(x) < \infty$ for all j
- then **positive sum**

$$f = \sum_{j=1}^{m} t_j f_j$$

  with $t_j > 0$ is convex

- if f is convex, then $f(Ax + b)$ is convex

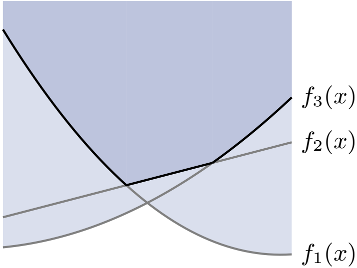
- **example:** log barrier for linear inequalities

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x), \quad \mathbf{dom}\, f = \{x \mid a_i^T x < b_i,\ i = 1, \dots, m\}$$

Pointwise supremum

- pointwise supremum of convex functions from family $\{f_j\}_{j\in J}$:

$$f(x) = \sup_{j\in J} f_j(x)$$



- convex since intersection of convex epigraphs
- **example:** maximum eigenvalue of symmetric matrix: for $X \in \mathbb{S}^n$,

$$\lambda_{\max}(X) = \sup_{\|y\|_2=1} y^T X y$$

Partial minimization

- if f is convex in (x, y) and $\mathcal{C}$ is a nonempty convex set, then

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

  is convex provided $g(x) > -\infty$ for all x

- **examples:**
  - distance to a convex set $\mathcal{C}$

$$\mathrm{dist}(x, \mathcal{C}) = \inf_{y \in \mathcal{C}} \|x - y\|$$

  - image of a convex function under linear mapping

$$\begin{aligned}(Lf)(x) &\coloneqq \inf_y \{f(y) \mid Ly = x\} \\ &= \inf_y \{f(y) + \mathcal{I}_{\mathcal{S}}(x, y)\}\end{aligned}$$

    where $\mathcal{S} = \{(x, y) \mid Ly = x\}$

Composition rule

- consider the function $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ defined as:

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

  where $h\colon \mathbb{R} \mapsto \overline{\mathbb{R}}$ is convex and $g\colon \mathbb{R}^n \mapsto \mathbb{R}$
- suppose that one of the following holds:
  - h is nondecreasing and g is convex
  - h is nonincreasing and g is concave
  - g is affine
- then f is convex
- **examples:**
  - $\exp g$ is convex if g is convex
  - $1/g$ is convex if g is concave and positive
  - norm-squared: $\|x\|^2$

Vector composition rule

- consider the function $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ defined as:

$$f(x) = h(g_1(x), g_2(x), \ldots, g_k(x))$$

  where $h\colon \mathbb{R}^k \mapsto \overline{\mathbb{R}}$ is convex and $g_i\colon \mathbb{R}^n \mapsto \mathbb{R}$
- suppose that for each $i \in \{1, \ldots, k\}$ one of the following holds:
  - h is nondecreasing in the ith argument and g_i is convex
  - h is nonincreasing in the ith argument and g_i is concave
  - g_i is affine
- then f is convex

Conjugate function

- the **conjugate** of a function f is

$$f^*(y) = \sup_x \left\{ y^T x - f(x) \right\}$$

- f^* is convex (even if f is not) since pointwise supremum of affine functions
- f^* is closed since its epigraph is intersection of closed halfspaces
- **examples:**
  - negative logarithm $f(x) = -\log(x)$

$$f^*(y) = \sup_{x>0} \{ yx + \log x \} = \begin{cases} -1 - \log(-y) & y < 0 \\ +\infty & \text{otherwise} \end{cases}$$

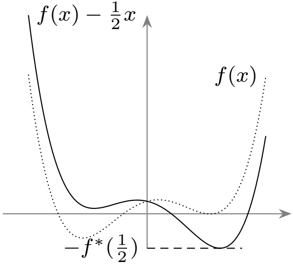
  - indicator function of a convex cone $\mathcal{K}$

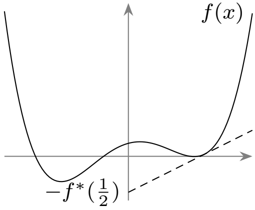
$$\mathcal{I}^*_{\mathcal{K}}(y) = \sup_{x \in \mathcal{K}} \{ y^T x \} = \mathcal{I}_{\mathcal{K}^\circ}(y)$$

Conjugate function – graphical interpretation

$$f^*(y) = \sup_x \left\{ y^T x - f(x) \right\} = -\inf_x \left\{ f(x) - y^T x \right\}$$

- **example:** $f^*(\frac{1}{2})$





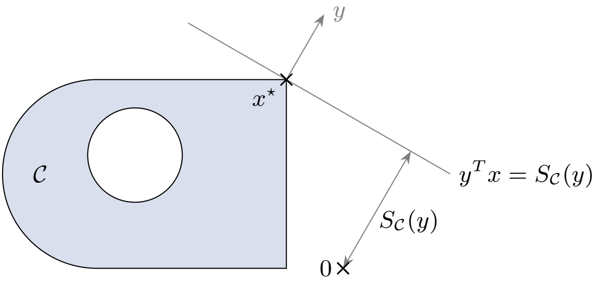
- conjugates of f and $\operatorname{env} f$ are the same, *i.e.*, $f^* = (\operatorname{env} f)^*$
- biconjugate $f^{**} = (f^*)^*$ is the closed convex envelope of f, that is, $f^{**} \leq f$ and $f^{**} = f$ if and only if f is closed and convex

Support function

- the **support function** of a set $\mathcal{C} \subseteq \mathbb{R}^n$ is defined as:

$$S_{\mathcal{C}}(y) \coloneqq \sup_{x \in \mathcal{C}} \{y^T x\}$$

- some properties of support function:
  - positively homogeneous, *i.e.*, $S_{\mathcal{C}}(\theta y) = \theta S_{\mathcal{C}}(y)$ if $\theta > 0$
  - $S_{\mathcal{C}}(y) = \mathcal{I}_{\mathcal{C}}^*(y)$
  - $\operatorname{dom} S_{\mathcal{C}} = (\mathcal{C}^\infty)^\circ$



References

- these lecture notes are based to a large extent on the following material:
  - Stanford EE364a class developed by Stephen Boyd
  - Lund course on Large-Scale Convex Optimization developed by Pontus Giselsson

- the original slides can be downloaded from

  `https://web.stanford.edu/class/ee364a/lectures.html`

  `https://archive.control.lth.se/ls-convex-2015/`