S. Boyd and L. Vandenberghe - Convex Optimization
Chapter 3: Convex Functions

C. Molinari

PhD student at UTFSM
Internship at INRIA

05 December 2016

Indice

Definition (Epigraph)

$$\mathbf{epi}\,(f) = \{(x,\alpha) \in \mathbb{R}^n \times \mathbb{R} : \quad f(x) \leq \alpha\}$$

Definition (Sublevel sets)

For $\gamma \in \mathbb{R}$,

$$\mathbf{sub}_\gamma\,(f) = \{x \in \mathbb{R}^n : \quad f(x) \leq \gamma\}$$

Remark

- $\mathbf{dom}\,(f) = \bigcup_{\gamma \in \mathbb{R}} \mathbf{sub}_\gamma\,(f)$;
- $\mathbf{argmin}\,(f) = \bigcap_{\gamma > \inf f} \mathbf{sub}_\gamma\,(f)$

Definition (Quasi-convexity)

A function f is *quasi-convex* if $\mathbf{sub}_\gamma(f)$ is a convex set for every $\gamma \in \mathbb{R}$

Proposition

- f is convex function *if and only* if $\mathbf{epi}(f)$ is a convex set;
- *if* f is a convex function, *then* f is a quasi-convex

Counterexamples: $\mathbf{sub}_\gamma(f)$ is convex $\forall \gamma \in \mathbb{R}$ but f is not convex for

- $f_1(x) = -e^x$ (concave);
- $f_2(x) = \sqrt{|x|}$ (nor convex, nor concave);
- $f_3(x) = x^3$ (nor convex, nor concave)

Definition (Lower-semicontinuity)

A function f is *lower-semicontinuous* if for every $x_0 \in \mathbf{dom}(f)$

$$f(x_0) \leq \liminf_{x \to x_0} f(x)$$

Proposition

The following are *equivalent*

- f is lower-semicontinuous;
- $\mathbf{epi}(f)$ is a closed set in $\mathbb{R}^n \times \mathbb{R}$;
- $\mathbf{sub}_\gamma(f)$ is a closed set in $\mathbb{R}^n$ for every $\gamma \in \mathbb{R}$

Definition (Hyperplane)

An *hyperplane* in $\mathbb{R}^n$ is a subset $\mathcal{H}$ of the form

$$\mathcal{H} = \{x \in \mathbb{R}^n : \ \langle a, \ x \rangle = b\},$$

where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$

Theorem (Representation for Subsets)

$C \subseteq R^n$ is *convex* and *closed* if and only if

$$C = \bigcap \{\mathcal{H} : \ \mathcal{H} \text{ is an hyperplane and } C \subseteq \mathcal{H}\}$$

(which implication is easy?)

Definition (Affine function)

An *affine function* on $\mathbb{R}^n$ is a function g of the form

$$g(x) = \langle a, \ x \rangle + b,$$

where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$

Theorem (Representation for Functions)

f is *convex* and *lower-semicontinuous* if and only if

$$f(x) = \sup \{ g(x) : \ g \text{ is an affine function and } g(z) \leq f(z) \ \ \forall z \in \mathbb{R}^n \} \tag{1}$$

(which implication is easy?)

Example: If f is convex with $\mathbf{dom}(f) = \mathbb{R}^n$, than (1) holds. (why?)

Indice

Definition (Fenchel conjugate)

For $f:\ \mathbb{R}^n \to \overline{\mathbb{R}}$, the Fenchel conjugate of f is the function $f^\star:\ \mathbb{R}^n \to \overline{\mathbb{R}}$ defined as

$$f^\star(x) = \sup_{x\in\mathbb{R}^n} \{\langle y,\ x\rangle - f(x)\}$$

Example: Support function

Given a subset $C \subset \mathbb{R}^n$, we call *support function* of C the Fenchel conjugate of indicator function of C, i.e.

$$\delta_C^\star(y) = \sup_{x\in C}\langle y,\ x\rangle$$

Geometrical interpretation

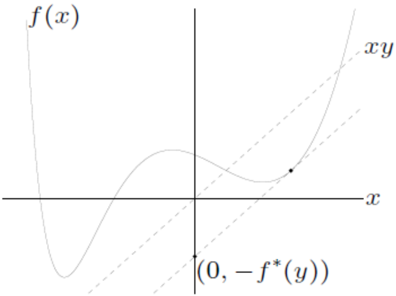

Figura: The conjugate function $f^\star(y)$ is the maximum gap between the linear function $\langle y,\ x\rangle$ and $f(x)$

Proposition

$f^\star$ is a convex and lower semi-continuous (wheter or not f is convex)

Proof: $f^\star$ is the pointwise supremum of affine functions

Proposition (Conjugate of the conjugate)

f is convex and lower-semicontinuos *if and only if*

$$f^{\star\star} = f$$

(which side is easy?)

In general: $f^{\star\star} \le f$ (why?)

Proposition (Fenchel-Young inequality)

For every $x, y \in \mathbb{R}^n$,

$$f(x) + f^\star(y) \geq \langle x,\ y \rangle \tag{2}$$

Proof: directly from the definition

Proposition (Legendre transform)

If f is convex and differentiable, then *equality* in (2) holds if and only if

$$y = \nabla f(x)$$

Remark:

In particular, under the hypothesis of the previous Proposition, if for a given y we can solve the equation $y = \nabla f(\bar{x})$, then we can compute $f^\star(y)$ as

$$f^\star(y) = \langle \bar{x}, \nabla f(\bar{x}) \rangle - f(\bar{x})$$

Examples

Norm

Consider $f(x) = \|x\|$; then $f^{\star} = \delta_{B_1^*(0)}$, i.e.

$$f^{\star}(y) = \begin{cases} 0 & \text{if } \|y\|_* \leq 1 \\ +\infty & \text{otherwise} \end{cases}$$

(where $\|\cdot\|_*$ is the dual norm to $\|\cdot\|$)

Norm squared

Consider $f(x) = \frac{1}{2}\|x\|^2$; then

$$f^{\star}(y) = \frac{1}{2}\|y\|_*^2$$

Indice

Definition

Definition (Quasi-convexity)

A function f is *quasi-convex* if $\mathbf{sub}_\gamma(f)$ is a convex set for every $\gamma \in \mathbb{R}$

Definition (Equivalent Def. by Generalized Jensen's Inequality)

A function f is *quasi-convex* if $\mathbf{dom}(f)$ is convex and for every $x, y \in \mathbf{dom}(f)$ and for every $\theta \in [0,1]$

$$f(\theta x + (1-\theta)) \leq \max\{f(x), f(y)\}$$

(Check the equivalence)

Example: the *cardinality* function (also called ℓ^0-"norm"), defined by

$$\mathbf{card}(x) = |\{i \in \{1, \ldots, n\} : \ x_i \neq 0\}|,$$

is quasi-concave on $\mathbb{R}^n_+$

Results

Proposition (Quasi-Convexity in $\mathbb{R}$)

A continuous function on $\mathbb{R}$ is quasi-convex if and only if it is

- non-decreasing; or
- non-increasing; or
- non-increasing in $t \leq c$ and non-decresing in $t \geq c$ for some $c \in \mathbf{dom}\,(f)$

First/Second-Order Conditions

Proposition (First-Order Condition)

Suppose f is differentiable; then f is quasi-convex if and only if $\mathbf{dom}\,(f)$ is convex and for all $x, y \in \mathbf{dom}\,(f)$

$$f(y) \leq f(x) \quad \Rightarrow \quad \langle \nabla f(x),\ y - x \rangle \leq 0$$

Geometrical Interpretation: if $\nabla f(x) \neq 0$, it defines a supporting hyperplane to $\mathbf{sub}_{f(x)}(f)$ at the point x

Remark:

If f is only quasi-convex, $\nabla f(x) = 0$ does *not* imply that x is a global minimizer (example?)

First/Second-Order Conditions

Proposition (Second-Order Condition)

Suppose f is twice-differentiable; if f is quasi-convex, then for all $x, y \in \mathbf{dom}\,(f)$

$$\langle y,\ \nabla f(x)\rangle = 0 \quad \Rightarrow \quad \langle y,\ \nabla^2 f(x)\, y\rangle \geq 0$$

Proposition (Partial Converse)

Suppose f is twice-differentiable; if for all $x, y \in \mathbf{dom}\,(f)$, $y \neq 0$

$$\langle y,\ \nabla f(x)\rangle = 0 \quad \Rightarrow \quad \langle y,\ \nabla^2 f(x)\, y\rangle > 0,$$

then f is quasi-convex

Geometrical Interpretation

The condition

$$\langle y,\ \nabla f(x)\rangle = 0 \quad \Rightarrow \quad \langle y,\ \nabla^2 f(x)\, y\rangle \geq 0$$

means

- <u>for $n = 1$</u>:
$$f'(x) = 0 \quad \Rightarrow \quad f''(x) \geq 0$$
(i.e., at every point with zero slope, the second derivative is non-negative);

- <u>for generic n</u>:
  - whenever $\nabla f(x) = 0$, then $\nabla^2 f(x) \succcurlyeq 0$;
  - when $\nabla f(x) \neq 0$, then $\nabla^2 f(x)$ is positive semi-definite on the $(n-1)$-dimensional subspace $\nabla f(x)^{\perp}$

Proposition (Representation by family of convex functions)

Suppose that f is *quasi-convex*; then it exists a family of *convex* functions $\phi_t : \mathbb{R}^n \to \overline{\mathbb{R}}$, indexed by $t \in \mathbb{R}$, such that

$$\textbf{sub}_t (f) = \textbf{sub}_0 (\phi_t) \qquad \forall t \in \mathbb{R} \tag{3}$$

Indeed, we can always choose

$$\phi_t (x) = \delta_{\textbf{sub}_t(f)} (x) = \begin{cases} 0 & \text{if } f(x) \leq t \\ +\infty & \text{else} \end{cases}$$

Remark 1:

The representation is not unique, in general

Remark 2:

A necessary condition for ϕ_t to satisfy (3) is to be non-increasing in t for every $x \in \mathbb{R}^n$, i.e. $\phi_s (x) \leq \phi_t (x)$ whenever $s \geq t$

Indice

Exercises (d)

- (**Example 3.21**) Compute the conjugate of the following functions:
  A. *Affine function*: $f(x) = \langle a,\ x \rangle + b$, for $x \in \mathbb{R}^n$;
  B. *Negative Logarithm*: $f(x) = -\log(x)$, with $\mathbf{dom}(f) = \mathbb{R}_{++}$;
  C. *Exponential*: $f(x) = e^x$, for $x \in \mathbb{R}$;
  D. *Negative Entropy*: $f(x) = x\log(x)$, with $\mathbf{dom}(f) = \mathbb{R}_+$

- (**Exercise 3.36, 3.41**) Compute the conjugate of
  A. the *Negative Normalized Entropy* :
  $$f(x) = \sum_{i=1}^{n} x_i \log\left(\frac{x_i}{\langle \mathbf{1},\ x \rangle}\right), \quad \mathbf{dom}(f) = \mathbb{R}^n_{++};$$
  B. the *Max-Function*:
  $$f(x) = \max_{i \in \{1,\dots,n\}} x_i;$$
  C. the *Sum of Largest Elements*:
  $$f(x) = \sum_{i=1}^{r} x_{[i]}$$

Exercises (e)

- (**Exercise 3.38**) Prove *Young's inequality*:

$$xy \leq F(x) + G(y),$$

where f is an incresing function with $f(0) = 0$, $g = f^{-1}$ is its inverse and F, G are defined by

$$F(x) = \int_0^x f(s)\ ds; \quad \text{and}$$
$$G(y) = \int_0^y g(s)\ ds$$

<u>*Hint:*</u> notice that F and G are conjugates

Exercises (f)

- (**Exercise 3.39**) Show the following properties of the *Fenchel conjugate*:

A. *Conjugate and Minimization, 1:* defining

$$g(x) = \inf_{z} f(x, z),$$

where $f(x, z)$ is a convex function in (x, z), express $g^{\star}$ in terms of $f^{\star}$;

B. *Conjugate and Minimization, 2*: for h convex, express the conjugate of

$$g(x) = \inf_{z} \{h(z) : Az + b = x\}$$

in terms of $h^{\star}$, A and b;

C. *Conjugate of Conjugate*: given a function f convex and lower semi-continuous, show that $f^{\star\star} = f$

Exercises (g)

- (**Exercise 3.43**) Show the *First-Order Condition for quasi-convexity*: suppose f is differentiable; then f is quasi-convex if and only if $\mathbf{dom}\,(f)$ is convex and for all $x, y \in \mathbf{dom}\,(f)$

$$f\,(y) \leq f\,(x) \quad \Rightarrow \quad \langle \nabla f\,(x)\,,\; y - x \rangle \leq 0$$

- (**Exercise 3.46**) Show that
  - a function $f: \ \mathbb{R} \to \mathbb{R}$ is *quasi-linear* (i.e., quasi-convex and quasi-concave) if and only if it is *monotone* (non-decreasing or non-increasing);
  - a function $f: \ \mathbb{R}^n \to \mathbb{R}$ is *quasi-linear* if and only if it can be expressed as
$$f\,(x) = g\,(\langle a,\; x \rangle)\,,$$
where $a \in \mathbb{R}^n$ and $g: \ \mathbb{R} \to \mathbb{R}$ is *monotone*