Conjugate Functions

Instructor: Jin Zhang

Department of Mathematics
Southern University of Science and Technology

Fall 2023

Contents

Definition (conjugate functions).

Given a function $f : \mathbb{R}^n \to [-\infty, \infty]$ (not necessarily convex), its (Fenchel) conjugate $f^* : \mathbb{R}^n \to [-\infty, \infty]$ is denoted by

$$f^*(y) := \sup_{x \in \mathbb{R}^n} \left\{ \langle y, x \rangle - f(x) \right\}, \qquad y \in \mathbb{R}^n.$$

Example (conjugate of indicator functions).

Let $f = \delta_C$, where $C \subseteq \mathbb{R}^n$ is nonempty. Then

$$f^* = \delta_C^* = \sigma_C.$$

Theorem (convexity and closedness of conjugate functions).

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be extended real-valued. Then the conjugate function f^* is closed and convex.

Example (conjugate of $\frac{1}{2}\|\cdot\|^2 + \delta_C$).

Suppose that $C \subseteq \mathbb{R}^n$ is nonempty. Define $f(x) = \frac{1}{2}\|x\|^2 + \delta_C(x)$. Then

$$f^*(y) = \frac{1}{2}\|y\|^2 - \frac{1}{2}d_C^2(y).$$

- **Theorem (properness of conjugate functions).** Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper convex. Then f^* is proper.
- **Theorem (Fenchel's inequality).** Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper. Then for any $x, y \in \mathbb{R}^n$,
$$f(x) + f^*(y) \geq \langle y, x \rangle.$$

2. The Biconjugate

Definition (biconjugate)

For a function $f : \mathbb{R}^n \to \bar{[} -\infty, \infty]$, we define the biconjugate of f as the conjugate of f^*, i.e.,

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

Lemma ($f^{**} \leq f$).

Let $f : \mathbb{R}^n \to [-\infty, \infty]$ be extended real-valued. Then $f(x) \geq f^{**}(x)$ for any $x \in \mathbb{R}^n$.

Theorem ($f = f^{**}$ for proper closed convex functions).

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper closed and convex. Then $f^{**} = f$.

Example (conjugate of support functions).

Let $C \subset \mathbb{R}^n$ be nonempty, then

$$\sigma_C^* = \delta_{\mathrm{cl}(\mathrm{conv}(C))}.$$

Example (conjugate of the max function).

Consider the function $f : \mathbb{R}^n \to \mathbb{R}$ given by $f(x) = \max\{x_1, \cdots, x_n\}$, then

$$f^* = \delta_{\Delta_n}.$$

Example (conjugate of $\frac{1}{2}\|\cdot\|^2 - d_C^2$).

Let $C \subset \mathbb{R}^n$ be nonempty closed and convex. Define

$$f(x) = \frac{1}{2}\|x\|^2 - \frac{1}{2}d_C^2(x).$$

Then

$$f^*(y) = \frac{1}{2}\|y\|^2 + \delta_C(y).$$

Theorem (conjugate of separable functions).

Let $g : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_p} \to \overline{\mathbb{R}}$ be given by

$$g(x_1, \cdots, x_p) = \sum_{i=1}^{p} f_i(x_i),$$

where $f_i : \mathbb{R}^{n_i} \to \overline{\mathbb{R}}$ is proper for any $i = 1, \cdots, p$. Then

$$g^*(y_1, \cdots, y_p) = \sum_{i=1}^{p} f_i^*(y_i) \text{ for any } y_i \in \mathbb{R}^{n_i},\ i = 1, \cdots, p.$$

Theorem (conjugate of $f\left(\mathcal{A}(x-a)\right)+\langle b,x\rangle+c$).

Let $f:\mathbb{R}^n\to\overline{\mathbb{R}}$, and let $\mathcal{A}:\mathbb{R}^p\to\mathbb{R}^n$ be an invertible linear transformation, $a\in\mathbb{R}^p$, $b\in\mathbb{R}^p$, and $c\in\mathbb{R}$. Then the conjugate of the function

$$g(x)=f\left(\mathcal{A}(x-a)\right)+\langle b,x\rangle+c$$

is given by

$$g^*(y)=f^*\left(\left(\mathcal{A}^T\right)^{-1}(y-b)\right)+\langle a,y\rangle-c-\langle a,b\rangle,\qquad y\in\mathbb{R}^p.$$

Theorem (conjugate of $\alpha f(\cdot)$ and $\alpha f(\cdot/\alpha)$).

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be extended real-valued and let $\alpha \in \mathbb{R}_{++}$.

(a). The conjugate of the function $g(x) = \alpha f(x)$ is given by

$$g^*(y) = \alpha f^*\left(\frac{y}{\alpha}\right), \quad y \in \mathbb{R}^n.$$

(b). The conjugate of the function $h(x) = \alpha f\left(\frac{x}{\alpha}\right)$ is given by

$$h^*(y) = \alpha f^*(y), \quad y \in \mathbb{R}^n.$$

- **Exponent.** Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = \exp(x)$. Then for any $y \in \mathbb{R}$,

$$f^*(y) = \begin{cases} y \log y - y, & y \geq 0, \\ \infty, & \text{else.} \end{cases}$$

- **Negative Log.** Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be given by

$$f(x) = \begin{cases} -\log x, & x > 0, \\ \infty, & x \leq 0. \end{cases}$$

Then

$$f^*(y) = \begin{cases} -1 - \log(-y), & y < 0, \\ \infty, & y \geq 0. \end{cases}$$

- **Negative Sum of Logs.** Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be given by

$$f(x) = \begin{cases} -\sum_{i=1}^n \log x_i, & x > 0, \\ \infty, & x \leq 0. \end{cases}$$

Then

$$f^*(y) = \begin{cases} -n - \sum_{i=1}^n \log(-y_i), & y < 0, \\ \infty, & y \geq 0. \end{cases}$$

- **Hinge Loss.** Consider the one-dimensional function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \max\{1 - x, 0\}.$$

Then for any $y \in \mathbb{R}$,

$$f^*(y) = y + \delta_{[-1,0]}(y).$$

- **Strictly Convex Quadratic Functions.**
  Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by

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

  where A is a real symmetric and positive definite matrix, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$. Then

$$f^*(y) = \frac{1}{2} (y - b)^T A^{-1} (y - b) - c.$$

Fenchel's dual problem

Consider the problem

$$(P) \qquad \min_{x \in \mathbb{R}^n} \; f(x) + g(x).$$

Rewriting the problem as

$$\min_{x,z \in \mathbb{R}^n} \; \{f(x) + g(z) \; : \; x = z\}$$

and the Lagrangian

$$\begin{aligned} L(x, z; y) &= f(x) + g(z) + \langle y, z - x \rangle \\ &= -\left[\langle y, x \rangle - f(x)\right] - \left[\langle -y, z \rangle - g(z)\right]. \end{aligned}$$

(continued).

The dual objective function is

$$q(y) = \inf_{x,z\in\mathbb{R}^n} L(x, z; y) = -f^*(y) - g^*(-y).$$

We thus obtain the Fenchel's dual:

$$(D) \quad \max_{y\in\mathbb{R}^n} \{-f^*(y) - g^*(-y)\}$$

Theorem (Fenchel's duality theorem). Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper convex functions. If $\mathrm{ri}\,(\mathrm{dom}(f)) \cap \mathrm{ri}\,(\mathrm{dom}(g)) \neq \varnothing$, then

$$\min_{x\in\mathbb{R}^n} \{f(x) + g(x)\} = \max_{x\in\mathbb{R}^n} \{-f^*(y) - g^*(-y)\},$$

and the maximum in the right-hand problem is attained whenever it is finite.

5. Infimal Convolution and Conjugacy

Recall

Let $h_1, h_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper. The infimal convolution of h_1, h_2 is defined by the following formula:

$$(h_1 \Box h_2)(x) \equiv \inf_{u \in \mathbb{R}^n} \{h_1(u) + h_2(x-u)\}.$$

Theorem (conjugate of infimal convolution).

For two proper functions $h_1, h_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ it holds that

$$(h_1 \Box h_2)^* = h_1^* + h_2^*.$$

Theorem (conjugate of sum).

Let $h_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper convex and $h_2 : \mathbb{R}^n \to \mathbb{R}$ be real-valued convex. Then

$$(h_1 + h_2)^* = h_1^* \square h_2^*.$$

Corollary

Let $h_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper closed convex and $h_2 : \mathbb{R}^n \to \mathbb{R}$ be real-valued convex. Then

$$h_1 + h_2 = (h_1^* \square h_2^*)^* .$$

Theorem (representation of the infimal convolution by conjugates).

Let $h_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper convex and $h_2 : \mathbb{R}^n \to \mathbb{R}$ be real-valued convex. Suppose that $h_1 \Box h_2$ is real-valued. Then

$$h_1 \Box h_2 = (h_1^* + h_2^*)^* .$$

Theorem (conjugate subgradient theorem).
Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper and convex. The following two claims are equivalent for any $x, y \in \mathbb{R}^n$:

(i). $\langle x, y \rangle = f(x) + f^*(y)$.

(ii). $y \in \partial f(x)$.

If in addition f is closed, then (i) and (ii) are equivalent to

(iii). $x \in \partial f^*(y)$.

Corollary (conjugate subgradient theorem—second formulation).
Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper closed and convex. Then for any $x, y \in \mathbb{R}^n$,

$$\partial f(x) = \operatorname*{arg\,sup}_{\bar{y} \in \mathbb{R}^n} \{\langle x, \bar{y}\rangle - f^*(\bar{y})\}$$

and

$$\partial f^*(y) = \operatorname*{arg\,sup}_{\bar{x} \in \mathbb{R}^n} \{\langle y, \bar{x}\rangle - f(\bar{x})\}$$