Convex Optimization

Part 4: Duality III

Namhoon Lee

POSTECH

7 Nov 2022

Admin

Assignment 3

- about Frank-Wolfe method
- due by Friday 25 November

Midterm course assessment (15 respondents)

- positive overall – thank you!
- feedback on assignments, TA-ing, English, review

Conjugate function

The **conjugate** of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$f^*(y) = \sup_{x \in \operatorname{dom} f} \left(y^\top x - f(x) \right)$$

- f^* is convex (even when f is not), since it is the pointwise supremum of a family of convex functions of y
- a.k.a. Legendre-Fenchel transformation, Fenchel transformation, Fenchel conjugate

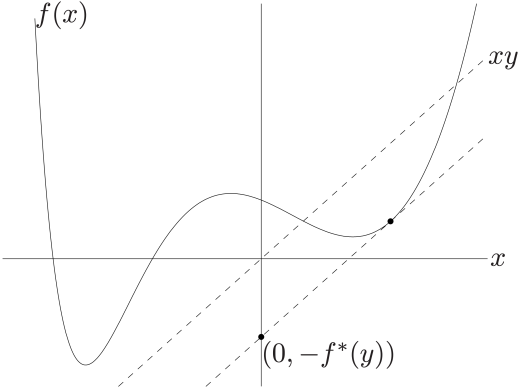

Figure: The conjugate function $f^*(y)$ is the maximum gap between linear function yx and f(x) as shown by the dashed line; figure from BV

Quadratic function

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

Strictly convex case ($A \succ 0$)

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

- notice ∇f^* is the inverse of ∇f

General convex case ($A \succeq 0$)

$$f^*(y) = \frac{1}{2}(y-b)^\top A^\dagger(y-b) - c, \quad \operatorname{dom} f^* = \mathcal{R}(A) + b$$

Negative entropy and negative logarithm

Negative entropy

$$f(x) = \sum_{i=1}^{n} x_i \log x_i \qquad f^*(y) = \sum_{i=1}^{n} e^{y_i - 1}$$

Negative logarithm

$$f(x) = -\sum_{i=1}^{n} \log x_i \qquad f^*(y) = -\sum_{i=1}^{n} \log(-y_i) - n$$

Norm

Norm

$$f(x) = \|x\| \qquad f^*(y) = \begin{cases} 0 & \|y\|_* \leq 1 \\ \infty & \|y\|_* > 1 \end{cases}$$

i.e., the conjugate of a norm is the indicator function of the dual norm unit ball

Proof.

Recall the definition of dual norm

$$\|y\|_* = \sup_{\|x\| \le 1} x^\top y$$

To evaluate $f^*(y) = \sup_x \left(y^\top x - \|x\|\right)$ consider two cases:

- if $\|y\|_* \le 1$, then $y^\top x \le \|x\|$ for all x; thus $f^*(y) = \sup_x(y^\top x - \|x\|) = 0$ with $x = 0$
- if $\|y\|_* > 1$, it means there exists an x with $\|x\| \le 1$, $x^\top y > 1$; taking $x = tz$ and letting $t \to \infty$ we have

  $$y^\top x - \|x\| = t(y^\top z - \|z\|) \to \infty$$

  which shows that $f^*(y) = \infty$

□

Fenchel's inequality

from the definition of conjugate function

$$f(x) + f^*(y) \geq x^\top y \quad \text{for all } x, y$$

- proof is straightforward by the definition of Legendre-Fenchel transform
- can be thought as an extension to non-quadratic convex f of the inequality; for example, for $f(x) = (1/2)x^\top x$ we have

$$\frac{1}{2}x^\top x + \frac{1}{2}y^\top y \geq x^\top y$$

or more generally

$$\frac{1}{2}x^\top Qx + \frac{1}{2}y^\top Q^{-1}y \geq x^\top y \quad \text{where } Q \succ 0$$

Conjugate of the conjugate

$$f^{**}(x) = \sup_{y \in \mathrm{dom}\, f^*} \left(x^\top y - f^*(y)\right)$$

- f^{**} is closed and convex
- from Fenchel's inequality, $x^\top y - f^*(y) \le f(x)$ for all y and x; therefore

  $$f^{**}(x) \le f(x) \quad \text{for all } x$$

  equivalently, $\mathrm{epi}\, f \subseteq \mathrm{epi}\, f^{**}$ (for any f)
- if f is closed and convex, then

  $$f^{**} = f \quad \text{for all } x$$

  equivalently, $\mathrm{epi}\, f = \mathrm{epi}\, f^{**}$ (if f is closed and convex); proof on next page
- (understanding in illustration)

Lagrange dual and conjugate functions

To give one simple connection between Lagrange dual and conjugate function, consider the following problem (which is not very interesting though)

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & x = 0 \end{aligned}$$

Lagrange dual function

$$g(\nu) = \inf_x \left(f(x) + \nu^\top x \right) = -\sup_x \left((-\nu)^\top x - f(x) \right) = -f^*(-\nu)$$

More generally, consider a problem with linear inequality and equality constraints

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b \\ & Cx = d \end{array}$$

Lagrange dual function

$$\begin{aligned} g(\lambda, \nu) &= \inf_x \left(f_0(x) + \lambda^\top (Ax - b) + \nu^\top (Cx - d)\right) \\ &= -b^\top \lambda - d^\top \nu + \inf_x \left(f_0(x) + (A^\top \lambda + C^\top \nu)^\top x\right) \\ &= -b^\top \lambda - d^\top \nu - f_0^*(-A^\top \lambda - C^\top \nu) \end{aligned}$$

- simplifies derivation of dual if conjugate of f_0 is known

Equality constrained norm minimization

$$\begin{array}{ll} \text{minimize} & \|x\| \\ \text{subject to} & Ax = b \end{array}$$

recall the conjugate of the norm function is the indicator function of the dual norm unit ball

$$f_0^*(y) = \begin{cases} 0 & \|y\|_* \leq 1 \\ \infty & \text{otherwise} \end{cases}$$

from the relationship between the conjugate function and Lagrange dual function

$$g(\nu) = -b^\top \nu - f_0^*(-A^\top \nu) = \begin{cases} -b^\top \nu & \|A^\top \nu\|_* \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

Entropy maximization

$$\begin{array}{ll} \text{minimize} & f_0(x) = \sum_{i=1}^{n} x_i \log x_i \\ \text{subject to} & Ax \preceq b \\ & \mathbb{1}^\top x = 1 \end{array}$$

recall the conjugate of the negative entropy function

$$f_0^*(y) = \sum_{i=1}^{n} e^{y_i - 1}$$

from the relationship between the conjugate function and Lagrange dual function

$$g(\lambda, \nu) = -b^\top \lambda - \nu - \sum_{i=1}^{n} e^{-a_i^\top \lambda - \nu - 1} = -b^\top \lambda - \nu - e^{-\nu - 1} \sum_{i=1}^{n} e^{-a_i^\top \lambda}$$

where a_i is the ith column of A

Any questions?