Optimization methods.
Seminar 9. Conjugate functions

Alexandr Katrutsa

Moscow Institute of Physics and Technology
Department of Control and Applied Mathematics

August 19, 2018

Reminder

- Feasible direction cone
- Tangent cone
- Sharp extremum

Definition

Conjugacy again?

- Previously we introduced conjugate (dual) sets and in particular conjugate cones
- Today we consider conjugate (dual) functions
- Further we will introduce dual (conjugate) optimization problem

Definition

A function $f^* : \mathbb{R}^n \to \mathbb{R}$ is called conjugate function of function f and is defined as

$$f^*(\mathbf{y}) = \sup_{\mathbf{x} \in dom\ f} (\mathbf{y}^\mathsf{T}\mathbf{x} - f(\mathbf{x})).$$

Domain of f^* is a set of $\mathbf{y}$, such that the supremum is finite.

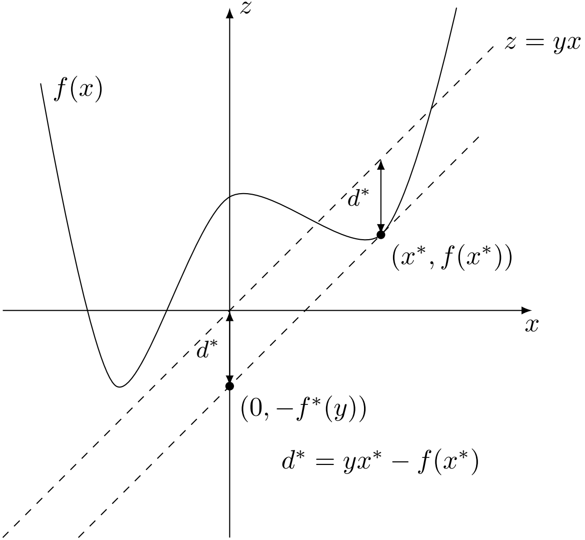
Properties

- Conjugate function f^* is always convex as supremum of linear functions independently of convexity of f
- Young-Fenchel inequality:

$$\mathbf{y}^\intercal \mathbf{x} \leq f(\mathbf{x}) + f^*(\mathbf{y})$$

- If f is differentiable, then $f^*(\mathbf{y}) = \nabla f^\intercal(\mathbf{x}^*)\mathbf{x}^* - f(\mathbf{x}^*)$, where $\mathbf{x}^*$ is a supremum point.

Geometrical interpretation



$$d^* = yx^* - f(x^*)$$

Examples

1. Linear function: $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} + b$
2. Negative entropy: $f(x) = x \log x$
3. Indicator function of the set S: $I_S(x) = 0$ iff $x \in S$
4. Norm: $f(\mathbf{x}) = \|\mathbf{x}\|$.
5. Squared norm: $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$

Calculus rules

- Separable sum: $f(x_1, x_2) = g(x_1) + h(x_2)$ and $f^*(y_1, y_2) = g^*(y_1) + h^*(y_2)$
- Translation of argument: $f(\mathbf{x}) = g(\mathbf{x} - \mathbf{a})$ and $f^*(\mathbf{y}) = \mathbf{a}^\top \mathbf{y} + g^*(\mathbf{y})$
- Composition with linear invertible mapping: $f(\mathbf{x}) = g(\mathbf{A}\mathbf{x})$ and $f^*(\mathbf{y}) = g^*(\mathbf{A}^{-\top}\mathbf{y})$
- Infimal convolution: $f(x) = (h \square g)(x) = \inf\limits_{u+v=x} (h(u) + g(v))$ and $f^*(y) = h^*(y) + g^*(y)$

Moreau-Yosida envelope

- $f(\mathbf{x})$ is convex, but *non-smooth*
- Moreau-Yosida envelope ($\lambda > 0$)

$$M_{\lambda f}(\mathbf{x}) = \inf_{\mathbf{u}}(f(\mathbf{u}) + \frac{1}{2\lambda}\|\mathbf{x} - \mathbf{u}\|_2^2) = \left(f \Box \frac{1}{2\lambda}\| \cdot \|_2^2\right)(\mathbf{x})$$

- Huber function – $M_{\lambda f}$ for module
  - $f(x) = |x|$
  - $M_{\lambda f}(x) = \begin{cases} \frac{x^2}{2\lambda} & |x| \leq \lambda \\ |x| - \lambda/2 & |x| \geq \lambda \end{cases}$

Exercise

- Draw in the one figure $f(x)$ and $M_{\lambda f}(x)$
- Derive expression of $M_{\lambda f}$ for $f(\mathbf{x}) = \|\mathbf{x}\|_1$

Why do we get smooth function?

- $M_{\lambda f}(\mathbf{x})$ – convex
- $M^*_{\lambda f}(\mathbf{y}) = f^*(\mathbf{y}) + \frac{\lambda}{2}\|\mathbf{y}\|_2^2$ – strongly convex with parameter λ
- $M_{\lambda f} = M^{**}_{\lambda f} = (f^* + \frac{\lambda}{2}\|\cdot\|_2^2)^*$
- Conjugate function of the strongly convex function is smooth $\Rightarrow M_{\lambda f}$ – smooth and

$$M'_{\lambda f}(\mathbf{x}) = \frac{1}{\lambda}(\mathbf{x} - \mathbf{u}^*), \quad \mathbf{u}^* = \arg\min_{\mathbf{u}} \left(f(\mathbf{u}) + \frac{1}{2\lambda}\|\mathbf{x} - \mathbf{u}\|_2^2 \right)$$

Important property

Sets of minimizers of f and $M_{\lambda f}$ are the same.

Recap

- Conjugate functions
- Young-Fenchel inequality and other properties
- Smoothing of non-smooth functions
- Examples