Convex Optimization for Machine Learning and Computer Vision

Tutorial

28.11.2018

Contents

Duality Theory

Theorem (Fenchel-Rockafellar duality)

Assume $\exists \bar{u} \in \mathrm{dom}\, G$ s.t. F is continuous at $K\bar{u}$. Then the **strong duality** holds: $\mathcal{P}^* = \mathcal{D}^*$. Moreover, (u^*, p^*) is the optimal solution pair iff

$$\begin{cases} Ku^* \in \partial F^*(p^*), \\ -K^\top p^* \in \partial G(u^*). \end{cases}$$

	$\inf_u$	$G(u) + F(Ku)$	"Primal"
$=$	$\inf_u \sup_p$	$G(u) + \langle p, Ku \rangle - F^*(p)$	"Saddle point"
$=$	$\sup_p \inf_u$	$G(u) + \langle p, Ku \rangle - F^*(p)$	
$=$	$\sup_p$	$-G^*(-K^\top p) - F^*(p)$	"Dual"

Duality gap: $G(u, p) = F(Ku) + G(u) + G^*(-K^\top p) + F^*(p)$

Proximal operator

Definition

Given a proper, convex, lsc function $J : \mathbb{E} \to \bar{\mathbb{R}}$, we define the **proximal operator** of J by

$$\operatorname{prox}_{\tau J}(v) = \arg\min_{u} J(u) + \frac{1}{2\tau}\|u - v\|^2.$$

- proximal gradient: gradient step on differentiable part and proximal step on non-differentiable.
- resolvent: $\operatorname{prox}_{\tau J} = (I + \tau \partial J)^{-1}$.
- fixed point (minimizer), Moreau identity.

Moreau envelope

Definition

The **Moreau envelope** of a proper, convex, lsc function $J : \mathbb{E} \to \bar{\mathbb{R}}$ is defined for each $u \in \mathbb{E}$ by

$$\begin{aligned} \mathrm{env}_{\tau J}(u) &:= \left(J \,\Box\, \frac{1}{2\tau} \|\cdot\|^2 \right)(u) \qquad (1) \\ &= \inf_{v \in \mathbb{E}} \left\{ J(v) + \frac{1}{2\tau} \|v - u\|^2 \right\} \\ &= J(\mathrm{prox}_{\tau J}(u)) + \frac{1}{2\tau} \|\mathrm{prox}_{\tau J}(u) - u\|^2. \qquad (2) \end{aligned}$$

- A smoothed form of original function.
- Moreau envelope has the same minimizers of J.
- prox can be viewed as a gradient step on Moreau envelope.