Fenchel Duality for Machine Learning, Part II: A Bit More Theory

Ryan M. Rifkin

Honda Research Institute USA, Inc.
Human Intention Understanding Group

2007

Epigraphs

Definition (Epigraph)

Given a function $f : \mathbb{R}^n \to [-\infty, \infty]$, the *epigraph* of f, epi f, is defined by

$$\text{epi } f = \{(y, e) : e \geq f(y)\} \subset \mathbb{R}^n \times \mathbb{R}.$$

We say f is *closed* or *convex* if epi f is closed or convex. f is *proper* when epi $f \neq \emptyset$ and $f > -\infty$. We call a closed, convex, proper function a *ccp* function.

Functional Minimizers

Definition

Given $f : \mathbb{R}^n \to (-\infty, \infty]$, we define the set $\operatorname{argmin}_{y \in \mathbb{R}^n} f(y)$ as follows,

$$\operatorname*{argmin}_{y \in \mathbb{R}^n} f(y) = \begin{cases} \mathbb{R}^n & \inf_{y \in \mathbb{R}^n} f(y) = \infty \\ \{y : f(y) = f_0\} & \inf_{y \in \mathbb{R}^n} f(y) = f_0 \in \mathbb{R}, \\ \emptyset & \inf_{y \in \mathbb{R}^n} f(y) = -\infty \end{cases}$$

with symmetrical definitions for argmax when needed.

Even in the second case, $\operatorname{argmin}_y f(y)$ may still be empty. The main case of interest is the second case when $\operatorname{argmin}_y f(y)$ is not empty.

Subgradients

Definition (subgradients and subdifferentials)

If $f : \mathbb{R}^n \to (-\infty, \infty]$ is convex and $y \in \operatorname{dom} f$, then $\phi \in \mathbb{R}^n$ is a *subgradient* of f at y iff it satisfies $\phi^t z \leq f(y+z) - f(y)$ for all $z \in \mathbb{R}^n$. The set of all such ϕ is the *subdifferential* and denoted $\partial f(y)$. By convention, $\partial f(y) = \emptyset$ if $y \notin \operatorname{dom} f$.

Subgradient generalizes the notion of gradient: if f is differentiable, then $\partial f(y)$ contains a single point, the gradient.

Fenchel-Legendre Conjugate

Definition (Fenchel-Legendre conjugate)

Given a function $f : \mathbb{R}^n \rightarrow (-\infty, \infty]$, the *Fenchel-Legendre conjugate* is

$$f^*(z) = \sup_y \{y^t z - f(y)\}.$$

The Fenchel-Legendre Conjugate Is Convex

$$
\begin{aligned}
f^*(z) &= \sup_y \{y^t z - f(y)\} \\
\text{epi } f^* &= \{(z, e) : e \geq f^*(z)\} \\
&= \{(z, e) : e \geq \sup_y \left(y^t z - f(y)\right)\}
\end{aligned}
$$

We see that epi f^* is an arbitrary intersection of closed convex sets, so epi f^* is closed and convex, and f^* *is closed and convex even if* f *is neither.*

The Fenchel-Legendre Conjugate Is Convex, Alt.

$$
\begin{aligned}
& f^*(\lambda z_1 + (1 - \lambda z_2)) \\
= \; & \sup_y \{ y^t (\lambda z_1 + (1-\lambda) z_2) - f(y) \} \\
= \; & \sup_y \{ \lambda \left(y^t z_1 - f(y) \right) + (1-\lambda) \left(y^t z_2 - f(y) \right) \} \\
\leq \; & \sup_y \{ \lambda \left(y^t z_1 - f(y) \right) \} + \sup_y \{ (1-\lambda) \left(y^t z_2 - f(y) \right) \} \\
= \; & \lambda \sup_y \{ y^t z_1 - f(y) \} + (1-\lambda) \sup_y \{ y^t z_2 - f(y) \} \\
= \; & \lambda f^*(z_1) + (1-\lambda) f^*(z_2)
\end{aligned}
$$

Biconjugation

Theorem (biconjugation)

$f : \mathbb{R}^n \to (-\infty, \infty]$ is closed and convex iff $f^{**} = f$.

In particular, conjugation is a bijection between ccp functions.

Fenchel-Young Theorem

Theorem (Fenchel-Young)

Let $f : \mathbb{R}^n \to (-\infty, \infty]$ *be convex.* $\forall y, z \in \mathbb{R}^n$,

$$f(y) + f^*(z) \geq y^t z$$

with equality holding iff $z \in \partial f(y)$.

Fenchel Duality, Main Theorem

Theorem (Fenchel Duality Theorem)

Let $f, g : \mathbb{R}^n \to (-\infty, \infty]$ be ccp with $f + g$ bounded below. If $0 \in int(dom\ f - dom\ g)$ or $0 \in int(dom\ f^* + dom\ g^*)$, then

$$\inf_{y,z} \{f(y) + g(y) + f^*(z) + g^*(-z)\} = 0,$$

and all minimizers y, z satisfy the complementarity equations:

$$\begin{aligned} f(y) - y^t z + f^*(z) &= 0 \\ g(y) + y^t z + g^*(-z) &= 0. \end{aligned}$$

Additionally, if $0 \in int(dom\ f - dom\ g)$ and $0 \in int(dom\ f^* + dom\ g^*)$ then a minimizer (y, z) exists.