Labs

Optimization for Machine Learning Spring 2023

EPFL

School of Computer and Communication Sciences

Martin Jaggi & Nicolas Flammarion

github.com/epfml/OptML_course

Problem Set 10 — Solutions (Convex conjugate)

For a function $f: \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ (which is not necessarily convex !), we consider its **convex conjugate** which for $y \in \mathbb{R}^d$ is defined as

$$f^*(y) = \sup_{x \in \mathbb{R}^d} (\langle x, y \rangle - f(x)) \in \mathbb{R} \cup \{+\infty\}$$

Prove the following properties.

1. Show that f^* is convex.

Proof: Note that f^* is the pointwise supremum of **affine functions** $y \mapsto \langle x, y \rangle - f(x)$. As seen in the first class, the pointwise supremum of convex functions is convex. Therefore f^* is convex.

2. Show that for $x,y\in\mathbb{R}^d$, $f(x)+f^*(y)\geq\langle x,y\rangle$. This is known as the Fenchel inequality.

Proof: For $y \in \mathbb{R}^d$, $f^*(y) = \sup_{x \in \mathbb{R}^d} (\langle x, y \rangle - f(x)) \ge \langle x, y \rangle - f(x)$ for all $x \in \mathbb{R}^d$.

3. Show that the biconjugate f^{**} (the conjugate of the conjugate) is such that $f^{**} \leq f$.

Proof: From the previous inequality we have that for all $x,y\in\mathbb{R}^d$, $f(x)\geq \langle x,y\rangle-f^*(y)$, we can therefore take the supremum over y of the left hand side: $f(x)\geq \sup_{y\in\mathbb{R}^d}(\langle y,x\rangle-f^*(y))=f^{**}(x)$

The Fenchel-Moreau theorem (which we will not prove here) states that $f = f^{**}$ if and only if f is convex and closed. It will turn out to be useful to show the following property.

4. Assume that f is closed and convex. Then show that for any $x, y \in \mathbb{R}^d$,

$$y \in \partial f(x) \Leftrightarrow x \in \partial f^*(y)$$

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

Proof that $y \in \partial f(x) \Rightarrow f(x) + f^*(y) = \langle x,y \rangle$: Assume that $y \in \partial f(x)$, then we have that for all $z \in \mathbb{R}^d$, $f(z) \geq f(x) + \langle y,z-x \rangle$. Therefore for all $z \in \mathbb{R}^d$, $\langle y,x \rangle - f(x) \geq \langle z,y \rangle - f(z)$. We can therefore take the supremum of the left hand size which gives that $\langle y,x \rangle - f(x) \geq \sup_z (\langle z,y \rangle - f(z))$ which also means that $\langle y,x \rangle - f(x) = \sup_z \langle z,y \rangle - f(z) = f^*(y)$ which proves the first part of the result.

Proof that $f(x)+f^*(y)=\langle x,y\rangle\Rightarrow y\in\partial f(x)$: We basically do the previous reasoning the other way round. Let $x,y\in\mathbb{R}^d$ such that $f(x)+f^*(y)=\langle x,y\rangle$. Therefore $\langle x,y\rangle-f(x)=f^*(y)=\sup_z(\langle z,y\rangle-f(z))\geq \langle z,y\rangle-f(z)$ for all $z\in\mathbb{R}^d$. Rearranging we get that for all $z\in\mathbb{R}^d$, $f(z)\geq f(x)+\langle y,z-x\rangle$ which means that $y\in\partial f(x)$.

Hence we have shown that $y \in \partial f(x) \Leftrightarrow f(x) + f^*(y) = \langle x,y \rangle$. Now we can apply this same result to f^* : $x \in \partial f^*(y) \Leftrightarrow f^*(y) + f^{**}(x) = \langle y,x \rangle$. Since f is closed and convex, by the Fenchel-Moreau theorem we have that $f = f^{**}$, hence $x \in \partial f^*(y) \Leftrightarrow f^*(y) + f(x) = \langle y,x \rangle$. Therefore all the implications are proven.