Serie 4

Optimal transport, Fall semester

EPFL, Mathematics section, Dr. Xavier Fernández-Real

Given a function $f : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty, \infty\}$ we define the convex conjugate f^* as

$$f^*(y) = \sup_{x \in \mathbb{R}^d} (x \cdot y - f(x)).$$

When $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex, f^* is also known as Legendre transform of f. Notice that, at least informally, if we assume that f is differentiable and that the supremum in the right-hand side is realized at a point $\bar{x}$, then $y = \nabla f(\bar{x})$.

Exercise 4.1. Given two functions $f, g : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ such that $f, g \not\equiv +\infty$. Show the following:

(i) f^* and g^* are convex functions.

(ii) If $f \le g$, then $g^* \le f^*$.

Solution:

(i) Let us prove that f^* is convex. Define $X_f := \{x \in \mathbb{R}^d \mid f(x) < +\infty\}$. Notice that if $x \in X_f$ and $x' \in \mathbb{R}^d \setminus X_f$, then for any $y \in \mathbb{R}^d$ we have

$$x' \cdot y - f(x') < x \cdot y - f(x),$$

and therefore, in the definition of f^* we can take the supremum in X_f instead of $\mathbb{R}^d$:

$$f^*(y) = \sup_{x \in X_f} (x \cdot y - f(x)).$$

Hence f^* is the (pointwise) supremum of a family of affine functions, since scalar products are linear, which implies that it is convex.

(ii) Let $y \in \mathbb{R}^d$. Then, since $f \le g$, we immediately have that $x \cdot y - f(x) \ge x \cdot y - g(x)$ for all $x \in \mathbb{R}^d$, and hence

$$\sup_{x \in \mathbb{R}^d} (x \cdot y - f(x)) \ge \sup_{x \in \mathbb{R}^d} (x \cdot y - g(x)).$$

By definition, $f^* \ge g^*$.

Exercise 4.2. Compute the convex conjugate of

(i) $f(x) = \frac{1}{2}\langle x, x\rangle$ for $x \in X$, $X = \mathbb{R}^d$;

(ii) $f(x) = \langle x, x_0\rangle$, for $x \in X$, where $x_0 \in X$ is a fixed point, $X = \mathbb{R}^d$;

(iii) a function f defined by $f(x_0) = 0$ and for $x \in X$, $x \neq x_0$, $f(x) = +\infty$, where $x_0 \in X$ is a fixed point, $X = \mathbb{R}^d$;

(iv) $f(x) = \frac{1}{p}|x|^p$ if $1 < p < \infty$ and $X = \mathbb{R}$.

Solution:

(i) Notice that

$$f^*(y) = \sup_{x\in X} \langle x, y\rangle - \frac{1}{2}\langle x, x\rangle = \sup_{x\in X} \langle x, y\rangle - \frac{1}{2}\langle x, x\rangle + \frac{1}{2}\langle y, y\rangle - \frac{1}{2}\langle y, y\rangle$$
$$= \sup_{x\in X} \frac{1}{2}\langle y, y\rangle - \frac{1}{2}\langle y - x, y - x\rangle = \frac{1}{2}\langle y, y\rangle - \frac{1}{2}\inf_{x\in X} \langle y - x, y - x\rangle = \frac{1}{2}\langle y, y\rangle,$$

where in the last equality we are using that $\langle y - x, y - x\rangle \geq 0$.

(ii) Observe that $f^*(y) = \sup_{x\in X} \langle x, y\rangle - \langle x, x_0\rangle = \sup_{x\in X} \langle x, y - x_0\rangle$, so that we need to take the supremum of a linear function. In particular, if $y - x_0 \neq 0$, $f^*(y) = +\infty$. Otherwise, if $y = x_0$, $f^*(x_0) = 0$. That is

$$f^*(y) = \begin{cases} 0 & y = x_0, \\ +\infty & y \neq x_0. \end{cases}$$

(iii) If $x \neq x_0$, then $\langle x, y\rangle - f(x) = -\infty$ for all $y \in X$. If instead $x = x_0$, then $\langle x, y\rangle - f(x) = \langle x_0, y\rangle > -\infty$. Hence

$$f^*(y) = \langle x_0, y\rangle$$

(iv) We need to compute $f^*(y) = \sup_{x\in\mathbb{R}} xy - \frac{1}{p}|x|^p$ for all $y \in \mathbb{R}$. We can do it in two ways:

The first way is by using Young's inequality, which states that

$$xy \leq \frac{1}{p}|x|^p + \frac{1}{q}|y|^q \quad \text{for all } x, y \in \mathbb{R},$$

where $1 < q < \infty$ is such that $\frac{1}{p} + \frac{1}{q} = 1$, with equality if and only if $|x|^p = |y|^q$ and $xy \geq 0$. Using this fact, we have $xy - \frac{1}{p}|x|^p \leq \frac{1}{q}|y|^q$ so $\sup_{x\in\mathbb{R}} xy - \frac{1}{p}|x|^p \leq \frac{1}{q}|y|^q$. Since for x such that $|x|^p = |y|^q$ and $xy \geq 0$ we have $xy - \frac{1}{p}|x|^p = \frac{1}{q}|y|^q$, we reach

$$f^*(y) = \sup_{x\in\mathbb{R}} xy - \frac{1}{p}|x|^p = \frac{1}{q}|y|^q.$$

For the second way of computing the supremum, simply consider the function $f_y(x) = xy - \frac{1}{p}|x|^p$ for each $y \in \mathbb{R}$. We can then find the maximum by taking the point where the derivative vanishes, to obtain the desired result.

Exercise 4.3. Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a convex lower semicontinuous function such that $f \not\equiv +\infty$. Prove that $(f^*)^* = f$.

Hint: Prove the two inequalities separately, $f \geq (f^*)^*$ and $f \leq (f^*)^*$. For the latter, use point (ii) and (iii) of Exercise 4.2.

Solution:

(i) Observe that $f(x) \geq \langle x, y\rangle - f^*(y)$. We can take supremum in $y \in \mathbb{R}^d$ on the right-hand side to deduce

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

as we wanted.

(ii) As a consequence of Exercise 4.2 points (ii) and (iii) we already know that the desired conclusion holds for affine functions. That is, if h is an affine function, $(h^*)^* = h$.

On the other hand, since f is convex and lower semicontinuous, we can write it as $f = \sup_{i\in I} h_i$ for some family of affine functions $\{h_i\}_{i\in I}$ such that $h_i \leq f$ for all $i \in I$. In particular, by Exercise 4.1 we have that $f^* \leq h_i^*$ and $(f^*)^* \geq (h_i^*)^* = h_i$ so that

$$(f^*)^* \geq \sup_{i\in I} h_i = f,$$

as we wanted to see.

Exercise 4.4. Given a function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, prove that

(i) $y \in \partial f(x)$ if and only if $f(x) + f^*(y) = \langle x, y\rangle$;

(ii) If f is convex, lower semicontinuous, and $f \not\equiv +\infty$, then $y \in \partial f(x) \iff x \in \partial f^*(y)$.

Solution:

(i) Let $x \in \mathbb{R}^d$ and $y \in \partial f(x)$. By definition, we have that

$$f(x') \geq f(x) + \langle y, x' - x\rangle \quad \text{for all } x' \in \mathbb{R}^d.$$

That is, we equivalently have

$$\langle x, y\rangle - f(x) \geq \langle x', y\rangle - f(x') \quad \text{for all } x' \in \mathbb{R}^d \iff \langle x, y\rangle - f(x) = \sup_{x'\in\mathbb{R}^d} \langle x', y\rangle - f(x') = f^*(y)$$

as we wanted to see.

(ii) Let $x \in \mathbb{R}^d$. Then, for all $x' \in \mathbb{R}^d$, by Exercise 4.3,

$$y \in \partial f(x) \iff f(x) + f^*(y) = \langle x, y\rangle \iff (f^*)^*(x) + f^*(y) = \langle x, y\rangle \iff x \in \partial f^*(y)$$

Exercise 4.5 (💣). Consider a strictly convex C^1 function $f : \mathbb{R}^d \to \mathbb{R}$ such that

$$\lim_{|x|\to\infty} \frac{f(x)}{|x|} = +\infty.$$

Prove that $\nabla f : \mathbb{R}^d \to \mathbb{R}^d$ is a bijection and $f(x) + f^*(y) = \langle x, y\rangle$ if and only if $\nabla f(x) = y$.

Solution: Let us first prove that the gradient is injective. Let $x, y \in \mathbb{R}^d$ with $x \neq y$. Since f is strictly convex we have

$$\begin{aligned} \nabla f(x) \cdot (y - x) &< f(y) - f(x) \\ \nabla f(y) \cdot (x - y) &< f(x) - f(y) \end{aligned} \quad \Rightarrow \quad (\nabla f(x) - \nabla f(y)) \cdot (y - x) < 0.$$

In particular, $\nabla f(x) \neq \nabla f(y)$ for all $x, y \in \mathbb{R}^d$, $x \neq y$.

We now prove surjectivity. Let us fix $y \in \mathbb{R}^d$ and let us consider $g_y(x) = f(x) - x \cdot y$. Notice that

$$\lim_{|x|\to\infty} g_y(x) \geq \lim_{|x|\to\infty} |x| \left(\frac{f(x)}{|x|} - |y| \right) = +\infty$$

and in particular, $g_y(x)$ achieves a minimum for any $y \in \mathbb{R}^d$. That is,

$$f^*(y) = - \inf_{x\in\mathbb{R}^d} g_y(x) = \max_{x\in\mathbb{R}^d} x \cdot y - f(x) = x_0 \cdot y - f(x_0)$$

for some $x_0 \in \mathbb{R}^d$. In particular, by Exercise 4.4, we equivalently have $y \in \partial f(x_0)$, and since f is convex and C^1, $\partial f(x_0) = \{\nabla f(x_0)\}$ so that $y = \nabla f(x_0)$. Notice that, in this case, this is equivalent to $f^*(y) + f(x_0) = x_0 \cdot y$, again, by Exercise 4.4. This shows the surjectivity and proves that ∇f is a bijection.