Introduction to Optimization - Midterm 3

Unless otherwise stated, functions are assumed to be closed, convex and not identically $+\infty$ in all that follows.

Moreau's Identity.
Moreau's Identity states that $\text{prox}_f(x) + \text{prox}_{f^*}(x) = x$ for every $x \in \mathbb{R}^N$.

1. Let $\bar{x}, \bar{y} \in \mathbb{R}^N$. Show that the following statements are equivalent:

   i) $f(\bar{x}) + f^*(\bar{y}) = \bar{y} \cdot \bar{x}$;
   ii) $\bar{y} \in \partial f(\bar{x})$; and
   iii) $\bar{x} \in \partial f^*(\bar{y})$.

   i) holds if, and only if $\bar{x}$ is a minimizer of the convex function $F_{\bar{y}}(x) := f(x) + f^*(\bar{y}) - \bar{y} \cdot x$. This is equivalent to $0 \in \partial F_{\bar{y}}(\bar{x}) = \partial f(\bar{x}) - \bar{y}$, which is precisely ii). The other equivalence is similar.

2. Use the equivalence ii) $\Leftrightarrow$ iii) from above to prove Moreau's Identity.
   Set $p := \text{prox}_f(x)$, so that $x - p \in \partial f(p)$, which is equivalent to $p \in \partial f^*(x - p)$. Write $u := x - p$, so that $x - u \in \partial f^*(u)$, which is nothing more than $u = \text{prox}_{f^*}(x)$. Substituting back the value of u, we get $x - p = \text{prox}_{f^*}(x)$, which is the desired inequality.

Functions with an additive structure.
Let $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ have the structure

$$f(x) = f_1(x_1) + \cdots f_N(x_N),$$

where $f_i : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ for $i = 1, \ldots, N$.

3. Prove that

$$f^*(y) = f_1^*(y_1) + \cdots + f_N^*(y_N).$$

$$f^*(y) = \sup_{x \in \mathbb{R}^N} \{y \cdot x - f(x)\} = \sup_{x \in \mathbb{R}^N} \left\{ \sum_{i=1}^N y_i x_i - f_i(x_i) \right\} = \sum_{i=1}^N \left[\sup x_i \in \mathbb{R} \{y_i x_i - f_i(x_i)\} \right] = \sum_{i=1}^N f_i^*(x_i).$$

The conjugate of the 1-norm.

4. Compute the conjugate of $\phi : \mathbb{R} \to \mathbb{R}$, defined by $\phi(x) = |x|$.
   We have $\phi^*(y) = \sup\{yx - |x|\}$. The expression in the braces is $(y-1)x$ for $x \geq 0$ and $(y+1)x$ for $x < 0$. If either $y > 1$ or $y > -1$, then $\phi^*(y) = +\infty$. Otherwise, $\phi^*(y) = 0$. In other words, ϕ^* is the indicator function of the interval $[-1, 1]$, that is $\iota_{[-1,1]}$.

5. Deduce a formula for $\Phi^*(y)$, where $\Phi(x) = \|x\|_1$.
   From 3 and 4, $\Phi^*(y) = 0$ if all the components of y are in $[-1, 1]$, and $+\infty$ otherwise. In other words, $\Phi^*(y) = \iota_C(y)$, where $C = [-1, 1]^N$.

6. Compute $\text{prox}_{\Phi^*}(x)$ for $x \in \mathbb{R}^N$.
   Set $p = \text{prox}_{\Phi^*}(x)$. By definition,

$$p = \operatorname*{argmin}_{z \in \mathbb{R}^N} \left\{ \iota_C(z) + \frac{1}{2}\|x - z\|^2 \right\} = \operatorname*{argmin}_{z \in C} \left\{ \frac{1}{2}\|x - z\|^2 \right\} = \text{Proj}_C(x).$$

   In other words, $p_i = 1$ if $x_i \geq 1$, $p_i = -1$ if $x_i \leq -1$ and $p_i = x_i$ otherwise.

7. Compute $\text{prox}_{\Phi}(x)$ using Moreau's Identity.
   From Moreau's Identity, $q := \text{prox}_{\Phi}(x) = x - \text{prox}_{\Phi^*}(x)$. If $x_i \geq 1$, then $q_i = x_i - 1$. If $x_i \leq -1$, then $q_i = x_i + 1$. Otherwise, $q_i = x_i - x_i = 0$. This is the *shrink* operator introduced in Lecture 11.

Recall that the (Fenchel) conjugate of $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is the function $f^* : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$, defined by $f^*(y) := \sup_{x \in \mathbb{R}^N} \{y \cdot x - f(x)\}$.

Recall also, that f^* is closed, convex and not identically $+\infty$ if f is, and that the Fenchel-Young inequality $f(x) + f^*(y) \geq y \cdot x$ holds for every $x, y \in \mathbb{R}^N$.

Finally, recall that $\text{prox}_f(x) = \text{argmin}_{z \in \mathbb{R}^N} \left\{ f(z) + \frac{1}{2}\|x - z\|^2 \right\}$.