6.253: Convex Analysis and Optimization
Homework 4

Prof. Dimitri P. Bertsekas

Spring 2010, M.I.T.

Problem 1

Let $f : \mathbf{R}^n \mapsto \mathbf{R}$ be the function

$$f(x) = \frac{1}{p} \sum_{i=1}^{n} |x_i|^p$$

where $1 < p$. Show that the conjugate is

$$f^\star(y) = \frac{1}{q} \sum_{i=1}^{n} |y_i|^q,$$

where q is defined by the relation

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Solution.
Consider first the case $n = 1$. Let x and y be scalars. By setting the derivative of $xy - (1/p)|x|^p$ to zero, and we see that the supremum over x is attained when $\text{sgn}(x)|x|^{p-1} = y$, which implies that $xy = |x|^p$ and $|x|^{p-1} = |y|$. By substitution in the formula for the conjugate, we obtain

$$f^*(y) = |x|^p - \frac{1}{p}|x|^p = (1 - \frac{1}{p})|x|^p = \frac{1}{q}|y|^{\frac{p}{p-1}} = \frac{1}{q}|y|^q.$$

We now note that for any function $f : R^n \mapsto (-\infty, \infty]$ that has the form

$$f(x) = f_1(x_1) + \cdots + f_n(x_n),$$

where $x = (x_1, \ldots, x_n)$ and $f_i : R \mapsto (-\infty, \infty]$, $i = 1, \ldots, n$, the conjugate is given by

$$f^*(y) = f_1^*(y_1) + \cdots + f_n^*(y_n),$$

where $f_i^* : R \mapsto (-\infty, \infty]$ is the conjugate of f_i, $i = 1, \ldots, n$. By combining this fact with the result above, we obtain the desired result.

Problem 2

(a) Show that if $f_1 : \mathbf{R}^n \mapsto (-\infty, \infty]$ and $f_2 : \mathbf{R}^n \mapsto (-\infty, \infty]$ are closed proper convex functions, with conjugates denoted by $f_1{}^\star$ and $f_2{}^\star$, respectively, we have

$$f_1(x) \le f_2(x), \qquad \forall\, x \in \mathbf{R}^n,$$

if and only if

$$f_1{}^\star(y) \ge f_2{}^\star(y), \qquad \forall\, y \in \mathbf{R}^n.$$

(b) Show that if C_1 and C_2 are nonempty closed convex sets, we have

$$C_1 \subset C_2,$$

if and only if

$$\sigma_{C_1}(y) \le \sigma_{C_2}(y), \qquad \forall\, y \in \mathbf{R}^n.$$

Construct an example showing that closedness of C_1 and C_2 is a necessary assumption.

Solution.
(a) If $f_1(x) \le f_2(x)$ for all x, we have for all $y \in R^n$,

$$f_1^*(y) = \sup_{x \in R^n} \{x'y - f_1(x)\} \ge \sup_{x \in R^n} \{x'y - f_1(x)\} = f_2^*(y).$$

The reverse implication follows from the fact that f_1 and f_2 are the conjugates of f_1^* and f_2^*, respectively.
(b) Consider the indicator functions δ_{C_1} and δ_{C_2} of C_1 and C_2. We have

$$C_1 \subset C_2 \qquad \text{if and only if} \qquad \delta_{C_1}(x) \ge \delta_{C_2}(x), \quad \forall\, x \in R^n.$$

Since σ_{C_1} and σ_{C_2} are the conjugates of δ_{C_1} and δ_{C_2}, respectively, the result follows from part (a).

To see that the assumption of closedness of C_1 and C_2 is needed, consider two convex sets that have the same closure, but none of the two is contained in the other, such as for example $(0, 1]$ and $[0, 1)$.

Problem 3

Let $X_1, \ldots, X_r$, be nonempty subsets of $\mathbf{R}^n$. Derive formulas for the support functions for $X_1 + \cdots + X_r$, $conv(X_1) + \cdots + conv(X_r)$, $\cup_{j=1}^r X_j$, and $conv\left(\cup_{j=1}^r X_j\right)$.

Solution.
Let $X = X_1 + \cdots + X_r$. We have for all $y \in R^n$,

$$\begin{aligned}
\sigma_X(y) &= \sup_{x \in X_1 + \cdots + X_r} x'y \\
&= \sup_{x_1 \in X_1, \ldots, x_r \in X_r} (x_1 + \cdots + x_r)'y \\
&= \sup_{x_1 \in X_1} x_1'y + \cdots + \sup_{x_r \in X_r} x_r'y \\
&= \sigma_{X_1}(y) + \cdots + \sigma_{X_r}(y).
\end{aligned}$$

Since X_j and $conv(X_j)$ have the same support function, it follows that

$$\sigma_{X_1}(y) + \cdots + \sigma_{X_r}(y)$$

is also the support function of

$$conv(X_1) + \cdots + conv(X_r).$$

Let also $X = \cup_{j=1}^r X_j$. We have

$$\sigma_X(y) = \sup_{x \in X} y'x = \max_{j=1,\ldots,r} \sup_{x \in X_j} y'x = \max_{j=1,\ldots,r} \sigma_{X_j}(y).$$

This is also the support function of $conv(\cup_{j=1}^r X_j)$.

Problem 4

Consider a function ϕ of two real variables x and z taking values in compact intervals X and Z, respectively. Assume that for each $z \in Z$, the function $\phi(\cdot, z)$ is minimized over X at a unique point denoted $\hat{x}(z)$. Similarly, assume that for each $x \in X$, the function $\phi(x, \cdot)$ is maximized over Z at a unique point denoted $\hat{z}(x)$. Assume further that the functions $\hat{x}(z)$ and $\hat{z}(x)$ are continuous over Z and X, respectively. Show that ϕ has a saddle point (x^*, z^*). Use this to investigate the existence of saddle points of $\phi(x, z) = x^2 + z^2$ over $X = [0, 1]$ and $Z = [0, 1]$.

Solution.
We consider a function ϕ of two real variables x and z taking values in compact intervals X and Z, respectively. We assume that for each $z \in Z$, the function $\phi(\cdot, z)$ is minimized over X at a unique point denoted $\hat{x}(z)$, and for each $x \in X$, the function $\phi(x, \cdot)$ is maximized over Z at a unique point denoted $\hat{z}(x)$,

$$\hat{x}(z) = \arg\min_{x \in X} \phi(x, z), \qquad \hat{z}(x) = \arg\max_{z \in Z} \phi(x, z).$$

Consider the composite function $f : X \mapsto X$ given by

$$f(x) = \hat{x}(\hat{z}(x)),$$

which is a continuous function in view of the assumption that the functions $\hat{x}(z)$ and $\hat{z}(x)$ are continuous over Z and X, respectively. Assume that the compact interval X is given by $[a, b]$. We now show that the function f has a fixed point, i.e., there exists some $x^* \in [a, b]$ such that

$$f(x^*) = x^*.$$

Define the function $g : X \mapsto X$ by

$$g(x) = f(x) - x.$$

Assume that $f(a) > a$ and $f(b) < b$, since otherwise [in view of the fact that $f(a)$ and $f(b)$ lie in the range $[a, b]$], we must have $f(a) = a$ and $f(b) = b$, and we are done. We have

$$g(a) = f(a) - a > 0,$$

$$g(b) = f(b) - b < 0.$$

Since g is a continuous function, the preceding relations imply that there exists some $x^* \in (a, b)$ such that $g(x^*) = 0$, i.e., $f(x^*) = x^*$. Hence, we have

$$\hat{x}(\hat{z}(x^*)) = x^*.$$

Denoting $\hat{z}(x^*)$ by z^*, we obtain

$$x^* = \hat{x}(z^*), \qquad z^* = \hat{z}(x^*).$$

By definition, a pair $(\bar{x}, \bar{z})$ is a saddle point if and only if

$$\max_{z \in Z} \phi(\bar{x}, z) = \phi(\bar{x}, \bar{z}) = \min_{x \in X} \phi(x, \bar{z}),$$

or equivalently, if $\bar{x} = \hat{x}(\bar{z})$ and $\bar{z} = \hat{z}(\bar{x})$. Therefore, we see that (x^*, z^*) is a saddle point of ϕ.

We now consider the function $\phi(x, z) = x^2 + z^2$ over $X = [0, 1]$ and $Z = [0, 1]$. For each $z \in [0, 1]$, the function $\phi(\cdot, z)$ is minimized over $[0, 1]$ at a unique point $\hat{x}(z) = 0$, and for each $x \in [0, 1]$, the function $\phi(x, \cdot)$ is maximized over $[0, 1]$ at a unique point $\hat{z}(x) = 1$. These two curves intersect at $(x^*, z^*) = (0, 1)$, which is the unique saddle point of ϕ.

Problem 5

In the context of Section 4.2.2, let $F(x,u) = f_1(x) + f_2(Ax+u)$, where A is an $m \times n$ matrix, and $f_1 : \mathbf{R}^n \mapsto (-\infty, \infty]$ and $f_2 : \mathbf{R}^m \mapsto (-\infty, \infty]$ are closed convex functions. Show that the dual function is

$$q(\mu) = -f_1^\star(A'\mu) - f_2^\star(-\mu),$$

where $f_1^\star$ and $f_2^\star$ are the conjugate functions of f_1 and f_2, respectively. *Note:* This is the Fenchel duality framework discussed in Section 5.3.5.

Solution.
From Section 4.2.1, the dual function is

$$q(\mu) = -p^\star(-\mu),$$

where $p^\star$ is the conjugate of the function

$$p(u) = inf_{x \in R^n} F(x,u).$$

The max crossing value is

$$q^* = \sup_\mu \{-p^\star(-\mu)\}.$$

By using the change of variables $z = Ax + u$ in the following calculation, we have

$$\begin{aligned} p^\star(-\mu) &= \sup_u \{-\mu' u - \inf_x \{f_1(x) + f_2(Ax+u)\}\} \\ &= \sup_{z,x} \{-\mu'(z - Ax) - f_1(x) - f_2(z)\} \\ &= f_1^\star(A'\mu) + f_2^\star(-\mu), \end{aligned}$$

where $f_1^\star$ and $f_2^\star$ are the conjugate functions of f_1 and f_2, respectively. Thus,

$$q(\mu) = -f_1^\star(A'\mu) - f_2^\star(-\mu).$$