6.253: Convex Analysis and Optimization
Homework 3

Prof. Dimitri P. Bertsekas

Spring 2010, M.I.T.

Problem 1

(a) Show that a nonpolyhedral closed convex cone need not be retractive, by using as an example the cone $C = \{(u, v, w) \mid \|(u, v)\| \leq w\}$, the recession direction $d = (1, 0, 1)$, and the corresponding asymptotic sequence $\{(k, \sqrt{k}, \sqrt{k^2 + k})\}$. (This is the, so-called, second order cone, which plays an important role in conic programming; see Chapter 5.)
(b) Verify that the cone C of part (a) can be written as the intersection of an infinite number of closed halfspaces, thereby showing that a nested set sequence obtained by intersection of an infinite number of retractive nested set sequences need not be retractive.

Solution.
(a) Clearly, $d = (1, 0, 1)$ is the recession direction associated with the asymptotic sequence $\{x_k\}$, where $x_k = (k, \sqrt{k}, \sqrt{k^2 + k})$. On the other hand, it can be verified by straightforward calculation that the vector

$$x_k - d = (k - 1, \sqrt{k}, \sqrt{k^2 + k} - 1)$$

does not belong to C. Indeed, denoting

$$u_k = k - 1, \qquad v_k = \sqrt{k}, \qquad w_k = \sqrt{k^2 + k} - 1,$$

we have

$$\|(u_k, v_k)\|^2 = (k - 1)^2 + k = k^2 - k + 1,$$

while

$$w_k^2 = (\sqrt{k^2 + k} - 1)^2 = k^2 + k + 1 - 2\sqrt{k^2 + k},$$

and it can be seen that

$$\|(u_k, v_k)\|^2 > w_k^2, \qquad \forall\ k \geq 1.$$

(b) Since by the Schwarz inequality, we have

$$\max_{\|(x,y)\|=1} (ux + vy) = \|(u, v)\|,$$

it follows that the cone

$$C = \{(u, v, w) \mid \|(u, v)\| \leq w\}$$

can be written as

$$C = \cap_{\|(x,y)\|=1}\{(u, v, w) \mid ux + vy \leq w\}.$$

Hence C is the intersection of an infinite number of closed halfspaces.

Problem 2

Let C be a nonempty convex set in $\mathbf{R}^n$, and let M be a nonempty affine set in $\mathbf{R}^n$. Show that $M \cap rin(C) = \emptyset$ is a necessary and sufficient condition for the existence of a hyperplane H containing M, and such that $rin(C)$ is contained in one of the open halfspaces associated with H.

Solution.
If there exists a hyperplane H with the properties stated, the condition $M \cap rin(C) = \emptyset$ clearly holds. Conversely, if $M \cap rin(C) = \emptyset$, then M and C can be properly separated. This hyperplane can be chosen to contain M since M is affine. If this hyperplane contains a point in $rin(C)$, then it must contain all of C. This contradicts the proper separation property, thus showing that $rin(C)$ is contained in one of the open halfspaces.

Problem 3

Let C_1 and C_2 be nonempty convex subsets of $\mathbf{R}^n$, and let B denote the unit ball in $\mathbf{R}^n$, $B = \{x \mid \|x\| \leq 1\}$. A hyperplane H is said to *separate strongly* C_1 and C_2 if there exists an $\epsilon > 0$ such that $C_1 + \epsilon B$ is contained in one of the open halfspaces associated with H and $C_2 + \epsilon B$ is contained in the other. Show that:

(a) The following three conditions are equivalent.

(i) There exists a hyperplane separating strongly C_1 and C_2.

(ii) There exists a vector $\alpha \in \mathbf{R}^n$ such that $\inf_{x \in C_1} \alpha' x > \sup_{x \in C_2} \alpha' x$.

(iii) $\inf_{x_1 \in C_1,\, x_2 \in C_2} \|x_1 - x_2\| > 0$, i.e., $0 \notin cl(C_2 - C_1)$.

(b) If C_1 and C_2 are disjoint, any one of the five conditions for strict separation, given in Prop. 1.5.3, implies that C_1 and C_2 can be strongly separated.

Solution.

(a) We first show that (i) implies (ii). Suppose that C_1 and C_2 can be separated strongly. By definition, this implies that for some nonzero vector $a \in \mathbf{R}^n$, $b \in \mathbf{R}$, and $\epsilon > 0$, we have

$$C_1 + \epsilon B \subset \{x \mid a'x > b\},$$

$$C_2 + \epsilon B \subset \{x \mid a'x < b\},$$

where B denotes the closed unit ball. Since $a \neq 0$, we also have

$$\inf\{a'y \mid y \in B\} < 0, \qquad \sup\{a'y \mid y \in B\} > 0.$$

Therefore, it follows from the preceding relations that

$$b \leq \inf\{a'x + \epsilon a'y \mid x \in C_1, y \in B\} < \inf\{a'x \mid x \in C_1\},$$

$$b \geq \sup\{a'x + \epsilon a'y \mid x \in C_2, y \in B\} > \sup\{a'x \mid x \in C_2\}.$$

Thus, there exists a vector $a \in \mathbf{R}^n$ such that

$$\inf_{x \in C_1} a'x > \sup_{x \in C_2} a'x,$$

proving (ii).

Next, we show that (ii) implies (iii). Suppose that (ii) holds, i.e., there exists some vector $a \in \mathbf{R}^n$ such that

$$\inf_{x \in C_1} a'x > \sup_{x \in C_2} a'x,$$

Using the Schwartz inequality, we see that

$$\begin{aligned} 0 &< \inf_{x \in C_1} a'x - \sup_{x \in C_2} a'x \\ &= \inf_{x_1 \in C_1,\ x_2 \in C_2} a'(x_1 - x_2), \\ &\leq \inf_{x_1 \in C_1,\ x_2 \in C_2} \|a\| \|x_1 - x_2\|. \end{aligned}$$

It follows that

$$\inf_{x_1 \in C_1, x_2 \in C_2} \|x_1 - x_2\| > 0,$$

thus proving (iii). Finally, we show that (iii) implies (i). If (iii) holds, we have for some $\epsilon > 0$,

$$\inf_{x_1 \in C_1, x_2 \in C_2} \|x_1 - x_2\| > 2\epsilon > 0.$$

From this we obtain for all $x_1 \in C_1$, all $x_2 \in C_2$, and for all $y_1,\ y_2$ with $\|y_1\| \le \epsilon$, $\|y_2\| \le \epsilon$,

$$\|(x_1 + y_1) - (x_2 + y_2)\| \ge \|x_1 - x_2\| - \|y_1\| - \|y_2\| > 0,$$

which implies that $0 \notin (C_1 + \epsilon B) - (C_2 + \epsilon B)$. Therefore, the convex sets $C_1 + \epsilon B$ and $C_2 + \epsilon B$ are disjoint. By the Separating Hyperplane Theorem, we see that $C_1 + \epsilon B$ and $C_2 + \epsilon B$ can be separated, i.e., $C_1 + \epsilon B$ and $C_2 + \epsilon B$ lie in opposite closed halfspaces associated with the hyperplane that separates them. Then, the sets $C_1 + (\epsilon/2)B$ and $C_2 + (\epsilon/2)B$ lie in opposite open halfspaces, which by definition implies that C_1 and C_2 can be separated strongly.

(b) Since C_1 and C_2 are disjoint, we have $0 \notin (C_1 - C_2)$. Any one of conditions (2)-(5) of Prop. 1.5.3 imply condition (1) of that proposition, which states that the set $C_1 - C_2$ is closed, i.e.,

$$cl(C_1 - C_2) = C_1 - C_2.$$

Hence, we have $0 \notin cl(C_1 - C_2)$, which implies that

$$\inf_{x_1 \in C_1,\, x_2 \in C_2} \|x_1 - x_2\| > 0.$$

From part (a), it follows that there exists a hyperplane separating C_1 and C_2 strongly.

Problem 4

We say that a function $f : \mathbf{R}^n \mapsto (-\infty, \infty]$ is *quasiconvex* if all its level sets

$$V_\gamma = \{x \mid f(x) \leq \gamma\}$$

are convex. Let X be a convex subset of $\mathbf{R}^n$, let f be a quasiconvex function such that $X \cap dom(f) \neq \emptyset$, and denote $f^* = \inf_{x \in X} f(x)$.

(a) Assume that f is not constant on any line segment of X, i.e., we do not have $f(x) = c$ for some scalar c and all x in the line segment connecting any two distinct points of X. Show that every local minimum of f over X is also global.

(b) Assume that X is closed, and f is closed and proper. Let Γ be the set of all $\gamma > f^*$, and denote

$$R_f = \cap_{\gamma \in \Gamma} R_\gamma, \qquad L_f = \cap_{\gamma \in \Gamma} L_\gamma,$$

where R_γ and L_γ are the recession cone and the lineality space of V_γ, respectively. Use the line of proof of Prop. 3.2.4 to show that f attains a minimum over X if any one of the following conditions holds:

(1) $R_X \cap R_f = L_X \cap L_f$.

(2) $R_X \cap R_f \subset L_f$, and X is a polyhedral set.

Solution.

(a) Let x^* be a local minimum of f over X and assume, to arrive at a contradiction, that there exists a vector $\bar{x} \in X$ such that $f(\bar{x}) < f(x^*)$. Then, $\bar{x}$ and x^* belong to the set $X \cap V_{\gamma^*}$, where $\gamma^* = f(x^*)$. Since this set is convex, the line segment connecting x^* and $\bar{x}$ belongs to the set, implying that

$$f(\alpha\bar{x} + (1 - \alpha)x^*) \leq \gamma^* = f(x^*), \qquad \forall\ \alpha \in [0, 1].$$

For each integer $k \geq 1$, there must exist an $\alpha_k \in (0, 1/k]$ such that

$$f(\alpha_k\bar{x} + (1 - \alpha_k)x^*) < f(x^*), \qquad \text{for some } \alpha_k \in (0, 1/k]$$

otherwise, we would have that $f(x)$ is constant for x on the line segment connecting x^* and $(1/k)\bar{x} + (1 - (1/k))x^*$. This contradicts the local optimality of x^*.

(b) We consider the level sets

$$V_\gamma = \{x \mid f(x) \leq \gamma\}$$

for $\gamma > f^*$. Let $\{\gamma_k\}$ be a scalar sequence such that $\gamma_k \downarrow f^*$. Using the fact that for two nonempty closed convex sets C and D such that $C \in D$, we have $R_C \in R_D$, it can be seen that

$$R_f = \cap_{\gamma \in \Gamma} R_\gamma = \cap_{k=1}^\infty R_{\gamma_k}.$$

Similarly, L_f can be written as

$$L_f = \cap_{\gamma \in \Gamma} L_\gamma = \cap_{k=1}^\infty L_{\gamma_k}.$$

Under each of the conditions (1) and (2), we will show that the set of minima of f over X, which is given by

$$X^* = \cap_{k=1}^\infty (X \cap V_{\gamma_k})$$

is nonempty.

Let condition (1) hold. The sets $X \cap V_{\gamma_k}$ are nonempty, closed, convex, and nested. Furthermore, for each k, their recession cone is given by $R_X \cap R_{\gamma_k}$ and their lineality space is given by $L_X \cap L_{\gamma_k}$. We have that

$$\cap_{k=1}^\infty (R_X \cap R_{\gamma_k}) = R_X \cap R_f,$$

and

$$\cap_{k=1}^{\infty}(L_X \cap L_{\gamma_k}) = L_X \cap L_f,$$

while by assumption $R_X \cap R_f = L_X \cap L_f$. Then it follows that X^* is nonempty.

Let condition (2) hold. The sets V_{γ_k} are nested and the intersection $X \cap V_{\gamma_k}$ is nonempty for all k. We also have by assumption that $R_X \cap R_f \in L_f$ and X is a polyhedral set. It follows that X^* is nonempty.

Problem 5

Let $F : \mathbf{R}^{n+m} \mapsto (-\infty, \infty]$ be a closed proper convex function of two vectors $x \in \mathbf{R}^n$ and $z \in \mathbf{R}^m$, and let

$$X = \left\{ x \mid \inf_{z \in \mathbf{R}^m} F(x, z) < \infty \right\}.$$

Assume that the function $F(x, \cdot)$ is closed for each $x \in X$. Show that:

(a) If for some $\bar{x} \in X$, the minimum of $F(\bar{x}, \cdot)$ over $\mathbf{R}^m$ is attained at a nonempty and compact set, the same is true for all $x \in X$.

(b) If the functions $F(x, \cdot)$ are differentiable for all $x \in X$, they have the same asymptotic slopes along all directions, i.e., for each $d \in \mathbf{R}^m$, the value of $\lim_{\alpha \to \infty} \nabla_z F(x, z + \alpha d)' d$ is the same for all $x \in X$ and $z \in \mathbf{R}^m$.

Solution.
The recession cone of F has the form

$$R_F = \{(d_x, d_z) \mid (d_x, d_z, 0) \in R_{epi(F)}\}.$$

The (common) recession cone of the nonempty level sets of $F(x, \cdot)$, $x \in X$, has the form

$$\{d_z \mid (0, d_z) \in R_F\},$$

for all $x \in X$, where R_F is the recession cone of F. Furthermore, the recession function of $F(x, \cdot)$ is the same for all $x \in X$.

(a) By the compactness hypothesis, the recession cone of $F(\bar{x}, \cdot)$ consists of just the origin, so the same is true for the recession cones of all $F(x, \cdot)$, $x \in X$. Thus the nonempty level sets of $F(x, \cdot)$, $x \in X$, are all compact.

(b) This is a consequence of the fact that the recession function of $F(x, \cdot)$ is the same for all $x \in X$, and the comments following Prop. 1.4.5