6.253: Convex Analysis and Optimization
Homework 2

Prof. Dimitri P. Bertsekas

Spring 2010, M.I.T.

Problem 1

(a) Let C be a nonempty convex cone. Show that $cl(C)$ and $ri(C)$ is also a convex cone.
(b) Let $C = cone(\{x_1, \dots, x_m\})$. Show that

$$ri(C) = \{\sum_{i=1}^{m} a_i x_i | a_i > 0, i = 1, \dots, m\}.$$

Solution.
(a) Let $x \in cl(C)$ and let α be a positive scalar. Then, there exists a sequence $\{x_k\} \in C$ such that $x_k \to x$, and since C is a cone, $\alpha x_k \in C$ for all k. Furthermore, $\alpha x_k \to \alpha x$, implying that $\alpha x \in cl(C)$. Hence, $cl(C)$ is a cone, and it also convex since the closure of a convex set is convex.

By Prop.1.3.2, the relative interior of a convex set is convex. To show that $rin(C)$ is a cone, let $x \in rin(C)$. Then, $x \in C$ and since C is a cone, $\alpha x \in C$ for all $\alpha > 0$. By the Line Segment Principle, all the points on the line segment connecting x and αx, except possibly αx, belong to $rin(C)$. Since this is true for every $\alpha > 0$, it follows that $\alpha x \in rin(C)$ for all $\alpha > 0$, showing that $rin(C)$ is a cone.

(b) Consider the linear transformation A that maps $(\alpha_1, \dots, \alpha_m) \in \mathbf{R}^m$ into $\sum_{i=1}^{m} \alpha_i x_i \in \mathbf{R}^n$. Note that C is the image of the nonempty convex set

$$\{(\alpha_1, \dots, \alpha_m) \mid \alpha_1 \geq 0, \dots, \alpha_m \geq 0\}$$

under A. Therefore, we have

$$\begin{aligned} rin(C) &= rin(A \cdot \{(\alpha_1, \dots, \alpha_m) \mid \alpha_1 \geq 0, \dots, \alpha_m \geq 0\}) \\ &= A \cdot rin(\{(\alpha_1, \dots, \alpha_m) \mid \alpha_1 \geq 0, \dots, \alpha_m \geq 0\}) \\ &= A \cdot \{(\alpha_1, \dots, \alpha_m) \mid \alpha_1 > 0, \dots, \alpha_m > 0\} \\ &= \left\{ \sum_{i=1}^{m} \alpha_i x_i \mid \alpha_1 > 0, \dots, \alpha_m > 0 \right\}. \end{aligned}$$

Problem 2

Let C_1 and C_2 be convex sets. Show that

$$C_1 \cap ri(C_2) \neq \emptyset \qquad \text{if and only if} \qquad ri(C_1 \cap aff(C_2)) \cap ri(C_2) \neq \emptyset.$$

Solution.
Let $x \in C_1 \cap rin(C_2)$ and $\bar{x} \in rin(C_1 \cap aff(C_2))$. Let L be the line segment connecting x and $\bar{x}$. Then L belongs to $C_1 \cap aff(C_2)$ since both of its endpoints belong to $C_1 \cap aff(C_2)$. Hence, by the Line Segment Principle, all points of L except possibly x, belong to $rin(C_1 \cap aff(C_2))$. On the other hand, by the definition of relative interior, all points of L that are sufficiently close to x belong to $rin(C_2)$, and these points, except possibly for x belong to $rin(C_1 \cap aff(C_2)) \cap rin(C_2)$. The other direction is obvious.

Problem 3

(a) Consider a vector x^* such that a given function $f : \mathbf{R}^n \mapsto \mathbf{R}$ is convex over a sphere centered at x^*. Show that x^* is a local minimum of f if and only if it is a local minimum of f along every line passing through x^* [i.e., for all $d \in \mathbf{R}^n$, the function $g : \mathbf{R} \mapsto \mathbf{R}$, defined by $g(\alpha) = f(x^* + \alpha d)$, has $\alpha^* = 0$ as its local minimum].
(b) Consider the nonconvex function $f : \mathbf{R}^2 \mapsto \mathbf{R}$ given by

$$f(x_1, x_2) = (x_2 - px_1^2)(x_2 - qx_1^2),$$

where p and q are scalars with $0 < p < q$, and $x^* = (0, 0)$. Show that $f(y, my^2) < 0$ for $y \neq 0$ and m satisfying $p < m < q$, so x^* is not a local minimum of f even though it is a local minimum along every line passing through x^*.

Solution.
(a) If x^* is a local minimum of f, evidently it is also a local minimum of f along any line passing through x^*.

Conversely, let x^* be a local minimum of f along any line passing through x^*. Assume, to arrive at a contradiction, that x^* is not a local minimum of f and that we have $f(\bar{x}) < f(x^*)$ for some $\bar{x}$ in the sphere centered at x^* within which f is assumed convex. Then, by convexity of f, for all $\alpha \in (0, 1)$, we have

$$f(\alpha x^* + (1 - \alpha)\bar{x}) \leq \alpha f(x^*) + (1 - \alpha) f(\bar{x}) < f(x^*),$$

so f decreases monotonically along the line segment connecting x^* and $\bar{x}$. This contradicts the hypothesis that x^* is a local minimum of f along any line passing through x^*.

(b) We first show that the function $g : \mathbf{R} \mapsto \mathbf{R}$ defined by $g(\alpha) = f(x^* + \alpha d)$ has a local minimum at $\alpha = 0$ for all $d \in \mathbf{R}^2$. We have

$$g(\alpha) = f(x^* + \alpha d) = (\alpha d_2 - p\alpha^2 d_1^2)(\alpha d_2 - q\alpha^2 d_1^2) = \alpha^2 (d_2 - p\alpha d_1^2)(d_2 - q\alpha d_1^2).$$

Also,

$$g'(\alpha) = 2\alpha(d_2 - p\alpha d_1^2)(d_2 - q\alpha d_1^2) + \alpha^2(-pd_1^2)(d_2 - q\alpha d_1^2) + \alpha^2(d_2 - p\alpha d_1^2)(-qd_1^2).$$

Thus $g'(0) = 0$. Furthermore,

$$\begin{aligned} g''(\alpha) = {} & 2(d_2 - p\alpha d_1^2)(d_2 - q\alpha d_1^2) + 2\alpha(-pd_1^2)(d_2 - q\alpha d_1^2) \\ & + 2\alpha(d_2 - p\alpha d_1^2)(-qd_1^2) + 2\alpha(-pd_1^2)(d_2 - q\alpha d_1^2) + \alpha^2(-pd_1^2)(-qd_1^2) \\ & + 2\alpha(d_2 - p\alpha d_1^2)(-qd_1^2) + \alpha^2(-pd_1^2)(-qd_1^2). \end{aligned}$$

Thus $g''(0) = 2d_2^2$, which is positive if $d_2 \neq 0$. If $d_2 = 0$, $g(\alpha) = pq\alpha^4 d_1^4$, which is clearly minimized at $\alpha = 0$. Therefore, $(0, 0)$ is a local minimum of f along every line that passes through $(0, 0)$.

We now show that if $p < m < q$, $f(y, my^2) < 0$ if $y \neq 0$ and that $f(y, my^2) = 0$ otherwise. Consider a point of the form (y, my^2). We have $f(y, my^2) = y^4(m - p)(m - q)$. Clearly, $f(y, my^2) < 0$ if and only if $p < m < q$ and $y \neq 0$. In any $\epsilon-$neighborhood of $(0, 0)$, there exists a $y \neq 0$ such that for some $m \in (p, q)$, (y, my^2) also belongs to the neighborhood. Since $f(0, 0) = 0$, we see that $(0, 0)$ is not a local minimum.

Problem 4

(a) Consider the quadratic program

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & 1/2\ |x|^2 + c'x \\ \text{subject to} \quad & Ax = 0 \end{aligned} \tag{1}$$

where $c \in \mathbf{R}^n$ and A is an $m \times n$ matrix of rank m. Use the Projection Theorem to show that

$$x^* = -(I - A'(AA')^{-1}A)c$$

is the unique solution.
(b) Consider the more general quadratic program

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & 1/2\ (x-\bar{x})'Q(x-\bar{x}) + c'(x-\bar{x}) \\ \text{subject to} \quad & Ax = b \end{aligned} \tag{2}$$

where c and A are as before, Q is a symmetric positive definite matrix, $b \in \mathbf{R}^m$, and $\bar{x}$ is a vector in $\mathbf{R}^n$, which is feasible, i.e., satisfies $A\bar{x} = b$. Use the transformation $y = Q^{1/2}(x - \bar{x})$ to write this problem in the form of part (a) and show that the optimal solution is

$$x^* = \bar{x} - Q^{-1}(c - A'\lambda),$$

where λ is given by

$$\lambda = (AQ^{-1}A')^{-1}AQ^{-1}c.$$

(c) Apply the result of part (b) to the program

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & 1/2\ x'Qx + c'x) \\ \text{subject to} \quad & Ax = b \end{aligned} \tag{3}$$

and show that the optimal solution is

$$x^* = -Q^{-1}(c - A'\lambda - A'(AQ^{-1}A')^{-1}b).$$

Solution.
(a) By adding the constant term $1/2\|c\|^2$ to the cost function, we can equivalently write this problem as

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & 1/2\|c + x\|^2 \\ \text{subject to} \quad & Ax = 0 \end{aligned}$$

which is the problem of projecting the vector $-c$ on the subspace $X = \{x \mid Ax = 0\}$. By the optimality condition for projection, a vector x^* such that $Ax^* = 0$ is the unique projection if and only if

$$(c + x^*)'x = 0, \qquad \forall\ x \text{ with } Ax = 0.$$

It can be seen that the vector

$$x^* = -(I - A'(AA')^{-1}A)c$$

satisfies this condition and is thus the unique solution of the quadratic programming problem in (a). (The matrix AA' is invertible because A has rank m.)

(b) By introducing the transformation $y = Q^{1/2}(x - \bar{x})$, we can write the problem as

$$\begin{aligned} &\underset{y}{\text{minimize}} && 1/2\|y\|^2 + \left(Q^{-1/2}c\right)' y \\ &\text{subject to} && AQ^{-1/2}y = 0 \end{aligned}$$

Using part (a), we see that the solution of this problem is

$$y^* = -\left(I - Q^{-1/2}A'\left(AQ^{-1}A'\right)^{-1}AQ^{-1/2}\right)Q^{-1/2}c$$

and by passing to the x-coordinate system through the inverse transformation $x^* - \bar{x} = Q^{-1/2}y^*$, we obtain the optimal solution

$$x^* = \bar{x} - Q^{-1}(c - A'\lambda),$$

where λ is given by

$$\lambda = \left(AQ^{-1}A'\right)^{-1}AQ^{-1}c. \tag{4}$$

(c) The quadratic program in part (b) contains as a special case the program

$$\begin{aligned} &\underset{x}{\text{minimize}} && 1/2x'Qx + c'x \\ &\text{subject to} && Ax = b \end{aligned}$$

This special case is obtained when $\bar{x}$ is given by

$$\bar{x} = Q^{-1}A'(AQ^{-1}A')^{-1}b. \tag{5}$$

Indeed $\bar{x}$ as given above satisfies $A\bar{x} = b$ as required, and for all x with $Ax = b$, we have

$$x'Q\bar{x} = x'A'(AQ^{-1}A')^{-1}b = b'(AQ^{-1}A')^{-1}b,$$

which implies that for all x with $Ax = b$,

$$1/2(x-\bar{x})'Q(x-\bar{x}) + c'(x-\bar{x}) = 1/2x'Qx + c'x + (1/2\bar{x}'Q\bar{x} - c'\bar{x} - b'(AQ^{-1}A')^{-1}b).$$

The last term in parentheses on the right-hand side above is constant, thus establishing that the programs (2) and (3) have the same optimal solution when $\bar{x}$ is given by Eq. 5. Therefore, we obtain the optimal solution of program (3):

$$x^* = -Q^{-1}\left(c - A'\lambda - A'(AQ^{-1}A')^{-1}b\right),$$

where λ is given by Eq. 4.

Problem 5

Let X be a closed convex subset of $\mathbf{R}^n$, and let $f : \mathbf{R}^n \mapsto (-\infty, \infty]$ be a closed convex function such that $X \cap dom(f) \neq \emptyset$. Assume that f and X have no common nonzero direction of recession. Let X^* be the set of minima of f over X (which is nonempty and compact), and let $f^* = \inf_{x \in X} f(x)$. Show that:
(a) For every $\epsilon > 0$ there exists a $\delta > 0$ such that every vector $x \in X$ with $f(x) \leq f^* + \delta$ satisfies $\min_{x^* \in X^*} \|x - x^*\| \leq \epsilon$.
(b) If f is real-valued, for every $\delta > 0$ there exists an $\epsilon > 0$ such that every vector $x \in X$ with $\min_{x^* \in X^*} \|x - x^*\| \leq \epsilon$ satisfies $f(x) \leq f^* + \delta$.
(c) Every sequence $\{x_k\} \subset X$ satisfying $f(x_k) \to f^*$ is bounded and all its limit points belong to X^*.

Solution.
(a) Let $\epsilon > 0$ be given. Assume, to arrive at a contradiction, that for any sequence $\{\delta_k\}$ with $\delta_k \downarrow 0$, there exists a sequence $\{x_k\} \in X$ such that for all k

$$f^* \leq f(x_k) \leq f^* + \delta_k, \qquad \min_{x^* \in X^*} \|x_k - x^*\| \geq \epsilon.$$

It follows that, for all k, x_k belongs to the set $\{x \in X \mid f(x) \leq f^* + \delta_0\}$, which is compact since f and X are closed and have no common nonzero direction of recession. Therefore, the sequence $\{x_k\}$ has a limit point $\bar{x} \in X$, which using also the lower semicontinuity of f, satisfies

$$f(\bar{x}) \leq \liminf_{k \to \infty} f(x_k) = f^*, \qquad \|\bar{x} - x^*\| \geq \epsilon, \quad \forall\, x^* \in X^*,$$

a contradiction.

(b) Let $\delta > 0$ be given. Assume, to arrive at a contradiction, that there exist sequences $\{x_k\} \subset X$, $\{x_k^*\} \subset X^*$, and $\{\epsilon_k\}$ with $\epsilon_k \downarrow 0$ such that

$$f(x_k) > f^* + \delta, \qquad \|x_k - x_k^*\| \leq \epsilon_k, \qquad \forall\, k = 0, 1, \ldots$$

(here x_k^* is the projection of x_k on X^*). Since X^* is compact, there is a subsequence $\{x_k^*\}_{\mathcal{K}}$ that converges to some $x^* \in X^*$. It follows that $\{x_k\}_{\mathcal{K}}$ also converges to x^*. Since f is real-valued, it is continuous, so we must have $f(x_k) \to f(x^*)$, a contradiction.

(c) Let $\bar{x}$ be a limit point of the sequence $\{x_k\} \subset X$ satisfying $f(x_k) \to f^*$. By lower semicontinuity of f, we have that

$$f(\bar{x}) \leq \liminf_{k \to \infty} f(x_k) = f^*.$$

Because $\{x_k\} \in X$ and X is closed, we have $\bar{x} \in X$, which in view of the preceding relation implies that $f(\bar{x}) = f^*$, i.e., $\bar{x} \in X^*$.

6.253 Convex Analysis and Optimization
Spring 2010