Convex Analysis and Optimization

Chapter 2 Solutions

Dimitri P. Bertsekas

with

Angelia Nedić and Asuman E. Ozdaglar

Massachusetts Institute of Technology

Athena Scientific, Belmont, Massachusetts
http://www.athenasc.com

LAST UPDATE October 8, 2005

CHAPTER 2: SOLUTION MANUAL

2.1

Let $f : \Re^n \mapsto \Re$ be a given function.

(a) Consider a vector x^* such that f 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 $d \in \Re^n$, the function $g : \Re \mapsto \Re$, defined by $g(\alpha) = f(x^* + \alpha d)$, has $\alpha^* = 0$ as its local minimum].

(b) Assume that f is not convex. Show that a vector x^* need not be a local minimum of f if it is a local minimum of f along every line passing through x^*. *Hint*: Use the function $f : \Re^2 \mapsto \Re$ 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 $p < m < q$, while $f(0, 0) = 0$.

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 $f(\overline{x}) < f(x^*)$ for some $\overline{x}$ in the sphere centered at x^* within which f is assumed convex. Then, by convexity, for all $\alpha \in (0, 1)$,

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

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

(b) Consider the function $f(x_1, x_2) = (x_2 - px_1^2)(x_2 - qx_1^2)$, where $0 < p < q$ and let $x^* = (0, 0)$.

We first show that $g(\alpha) = f(x^* + \alpha d)$ is minimized at $\alpha = 0$ for all $d \in \Re^2$. 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,

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

Thus $g''(0) = 2d_2^2$, which is greater than 0 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 $(0, 0)$.

Let's 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.

2.2 (Lipschitz Continuity of Convex Functions)

Let $f : \Re^n \mapsto \Re$ be a convex function and X be a bounded set in $\Re^n$. Show that f is Lipschitz continuous over X, i.e., there exists a positive scalar L such that

$$\big|f(x) - f(y)\big| \leq L\|x - y\|, \qquad \forall\ x, y \in X.$$

Solution: Let ϵ be a positive scalar and let C_ϵ be the set given by

$$C_\epsilon = \big\{z \mid \|z - x\| \leq \epsilon,\ \text{for some } x \in \text{cl}(X)\big\}.$$

We claim that the set C_ϵ is compact. Indeed, since X is bounded, so is its closure, which implies that $\|z\| \leq \max_{x\in \text{cl}(X)} \|x\| + \epsilon$ for all $z \in C_\epsilon$, showing that C_ϵ is bounded. To show the closedness of C_ϵ, let $\{z_k\}$ be a sequence in C_ϵ converging to some z. By the definition of C_ϵ, there is a corresponding sequence $\{x_k\}$ in $\text{cl}(X)$ such that

$$\|z_k - x_k\| \leq \epsilon, \qquad \forall\ k. \tag{2.1}$$

Because $\text{cl}(X)$ is compact, $\{x_k\}$ has a subsequence converging to some $x \in \text{cl}(X)$. Without loss of generality, we may assume that $\{x_k\}$ converges to $x \in \text{cl}(X)$. By taking the limit in Eq. (2.1) as $k \to \infty$, we obtain $\|z - x\| \leq \epsilon$ with $x \in \text{cl}(X)$, showing that $z \in C_\epsilon$. Hence, C_ϵ is closed.

We now show that f has the Lipschitz property over X. Let x and y be two distinct points in X. Then, by the definition of C_ϵ, the point

$$z = y + \frac{\epsilon}{\|y - x\|}(y - x)$$

is in C_ϵ. Thus

$$y = \frac{\|y - x\|}{\|y - x\| + \epsilon} z + \frac{\epsilon}{\|y - x\| + \epsilon} x,$$

showing that y is a convex combination of $z \in C_\epsilon$ and $x \in C_\epsilon$. By convexity of f, we have

$$f(y) \le \frac{\|y-x\|}{\|y-x\|+\epsilon} f(z) + \frac{\epsilon}{\|y-x\|+\epsilon} f(x),$$

implying that

$$f(y) - f(x) \le \frac{\|y-x\|}{\|y-x\|+\epsilon}\big(f(z)-f(x)\big) \le \frac{\|y-x\|}{\epsilon}\left(\max_{u\in C_\epsilon} f(u) - \min_{v\in C_\epsilon} f(v)\right),$$

where in the last inequality we use Weierstrass' theorem (f is continuous over $\Re^n$ by Prop. 1.4.6 and C_ϵ is compact). By switching the roles of x and y, we similarly obtain

$$f(x) - f(y) \le \frac{\|x-y\|}{\epsilon}\left(\max_{u\in C_\epsilon} f(u) - \min_{v\in C_\epsilon} f(v)\right),$$

which combined with the preceding relation yields $\big|f(x) - f(y)\big| \le L\|x-y\|$, where $L = \big(\max_{u\in C_\epsilon} f(u) - \min_{v\in C_\epsilon} f(v)\big)/\epsilon$.

2.3 (Exact Penalty Functions)

Let $f : Y \mapsto \Re$ be a function defined on a subset Y of $\Re^n$. Assume that f is Lipschitz continuous with constant L, i.e.,

$$\big|f(x) - f(y)\big| \le L\|x-y\|, \qquad \forall\, x, y \in Y.$$

Let also X be a nonempty closed subset of Y, and c be a scalar with $c > L$.

(a) Show that if x^* minimizes f over X, then x^* minimizes

$$F_c(x) = f(x) + c \inf_{y\in X} \|y-x\|$$

over Y.

(b) Show that if x^* minimizes $F_c(x)$ over Y, then $x^* \in X$, so that x^* minimizes f over X.

Solution: We note that by Weierstrass' Theorem, the minimum of $\|y-x\|$ over $y \in X$ is attained, so we can write $\min_{y\in X} \|y-x\|$ in place of $\inf_{y\in X} \|y-x\|$.

(a) By assumption, x^* minimizes f over X, so that $x^* \in X$, and we have for all $c > L$, $y \in X$, and $x \in Y$,

$$F_c(x^*) = f(x^*) \le f(y) \le f(x) + L\|y-x\| \le f(x) + c\|y-x\|,$$

where we use the Lipschitz continuity of f to get the second inequality. Taking the minimum over all $y \in X$, we obtain

$$F_c(x^*) \le f(x) + c\min_{y\in X} \|y-x\| = F_c(x), \qquad \forall\, x \in Y.$$

Hence, x^* minimizes $F_c(x)$ over Y for all $c > L$.

(b) It will suffice to show that $x^* \in X$. Suppose, to arrive at a contradiction, that x^* minimizes $F_c(x)$ over Y, but $x^* \notin X$.

We have that $F_c(x^*) = f(x^*) + c \min_{y \in X} \|y - x^*\|$. Let $\tilde{x} \in X$ be a point where the minimum of $\|y - x\|$ over $y \in X$ is attained. Then $\tilde{x} \neq x^*$, and we have

$$\begin{aligned} F_c(x^*) &= f(x^*) + c\|\tilde{x} - x^*\| \\ &> f(x^*) + L\|\tilde{x} - x^*\| \\ &\geq f(\tilde{x}) \\ &= F_c(\tilde{x}), \end{aligned}$$

which contradicts the fact that x^* minimizes $F_c(x)$ over Y. (Here, the first inequality follows from $c > L$ and $\tilde{x} \neq x^*$, and the second inequality follows from the Lipschitz continuity of f.)

2.4 (Ekeland's Variational Principle [Eke74])

This exercise shows how ϵ-optimal solutions of optimization problems can be approximated by (exactly) optimal solutions of some other slightly perturbed problems. Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper function, and let $\overline{x} \in \Re^n$ be a vector such that

$$f(\overline{x}) \leq \inf_{x \in \Re^n} f(x) + \epsilon,$$

where $\epsilon > 0$. Then, for any $\delta > 0$, there exists a vector $\tilde{x} \in \Re^n$ such that

$$\|\overline{x} - \tilde{x}\| \leq \frac{\epsilon}{\delta}, \qquad f(\tilde{x}) \leq f(\overline{x}),$$

and $\tilde{x}$ is the unique optimal solution of the perturbed problem of minimizing $f(x) + \delta\|x - \tilde{x}\|$ over $\Re^n$.

Solution: For some $\delta > 0$, define the function $F : \Re^n \mapsto (-\infty, \infty]$ by

$$F(x) = f(x) + \delta\|x - \overline{x}\|.$$

The function F is closed in view of the assumption that f is closed. Hence, by Prop. 1.2.2(b), it follows that all the level sets of F are closed. The level sets are also bounded, since for all $\gamma > f^*$, we have

$$\big\{x \mid F(x) \leq \gamma\big\} \subset \big\{x \mid f^* + \delta\|x - \overline{x}\| \leq \gamma\big\} = B\left(\overline{x}, \frac{\gamma - f^*}{\delta}\right), \tag{2.2}$$

where $B\big(\overline{x}, (\gamma - f^*)/\delta\big)$ denotes the closed ball centered at $\overline{x}$ with radius $(\gamma - f^*)/\delta$. Hence, it follows by Weierstrass' Theorem that F attains a minimum over $\Re^n$, i.e., the set $\arg\min_{x \in \Re^n} F(x)$ is nonempty and compact.

Consider now minimizing f over the set $\arg\min_{x \in \Re^n} F(x)$. Since f is closed by assumption, we conclude by using Weierstrass' Theorem that f attains a minimum at some $\tilde{x}$ over the set $\arg\min_{x \in \Re^n} F(x)$. Hence, we have

$$f(\tilde{x}) \leq f(x), \qquad \forall\ x \in \arg\min_{x \in \Re^n} F(x). \tag{2.3}$$

Since $\tilde{x} \in \arg\min_{x \in \Re^n} F(x)$, it follows that $F(\tilde{x}) \leq F(x)$, for all $x \in \Re^n$, and

$$F(\tilde{x}) < F(x), \qquad \forall\, x \notin \arg\min_{x \in \Re^n} F(x),$$

which by using the triangle inequality implies that

$$\begin{aligned} f(\tilde{x}) &< f(x) + \delta\|x - \overline{x}\| - \delta\|\tilde{x} - \overline{x}\| \\ &\leq f(x) + \delta\|x - \tilde{x}\|, \qquad \forall\, x \notin \arg\min_{x \in \Re^n} F(x). \end{aligned} \tag{2.4}$$

Using Eqs. (2.3) and (2.4), we see that

$$f(\tilde{x}) < f(x) + \delta\|x - \tilde{x}\|, \qquad \forall\, x \neq \tilde{x},$$

thereby implying that $\tilde{x}$ is the unique optimal solution of the problem of minimizing $f(x) + \delta\|x - \tilde{x}\|$ over $\Re^n$.

Moreover, since $F(\tilde{x}) \leq F(x)$ for all $x \in \Re^n$, we have $F(\tilde{x}) \leq F(\overline{x})$, which implies that

$$f(\tilde{x}) \leq f(\overline{x}) - \delta\|\tilde{x} - \overline{x}\| \leq f(\overline{x}),$$

and also

$$F(\tilde{x}) \leq F(\overline{x}) = f(\overline{x}) \leq f^* + \epsilon.$$

Using Eq. (2.2), it follows that $\tilde{x} \in B(\overline{x}, \epsilon/\delta)$, proving the desired result.

2.5 (Approximate Minima of Convex Functions)

Let X be a closed convex subset of $\Re^n$, and let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed convex function such that $X \cap \mathrm{dom}(f) \neq \emptyset$. Assume that f and X have no common nonzero direction of recession. Let X^* be the set of global minima of f over X (which is nonempty and compact by Prop. 2.3.2), 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 a $\epsilon > 0$ 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 $\overline{x} \in X$, which using also the lower semicontinuity of f, satisfies

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

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 $\overline{x}$ 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

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

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

2.6 (Directions Along Which a Function is Flat)

The purpose of the exercise is to provide refinements of results relating to set intersections and existence of optimal solutions (cf. Props. 1.5.6 and 2.3.3). Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper convex function, and let F_f be the set of all directions y such that for every $x \in \text{dom}(f)$, the limit $\lim_{\alpha\to\infty} f(x + \alpha y)$ exists. We refer to F_f as the set of *directions along which f is flat*. Note that

$$L_f \subset F_f \subset R_f,$$

where L_f and R_f are the constancy space and recession cone of f, respectively. Let X be a subset of $\Re^n$ specified by linear inequality constraints, i.e.,

$$X = \{x \mid a_j'x \leq b_j,\ j = 1, \ldots, r\},$$

where a_j are vectors in $\Re^n$ and b_j are scalars. Assume that

$$R_X \cap F_f \subset L_f,$$

where R_X is the recession cone of X.

(a) Let

$$C_k = \big\{x \mid f(x) \leq w_k\big\},$$

where $\{w_k\}$ is a monotonically decreasing and convergent scalar sequence, and assume that $X \cap C_k \neq \emptyset$ for all k. Show that

$$X \cap \left(\cap_{k=0}^{\infty} C_k\right) \neq \emptyset.$$

(b) Show that if $\inf_{x \in X} f(x)$ is finite, the function f attains a minimum over the set X.

(c) Show by example that f need not attain a minimum over X if we just assume (as in Prop. 2.3.3) that $X \cap \mathrm{dom}(f) \neq \emptyset$.

Solution: (a) We will follow the line of proof of Prop. 1.5.6, with a modification to use the condition $R_X \cap F_f \subset L_f$ in place of the condition $R_X \cap R_f \subset L_f$.

We use induction on the dimension of the set X. Suppose that the dimension of X is 0. Then X consists of a single point. This point belongs to $X \cap C_k$ for all k, and hence belongs to the intersection $X \cap \left(\cap_{k=0}^{\infty} C_k\right)$.

Assume that, for some $l < n$, the intersection $\overline{X} \cap \left(\cap_{k=0}^{\infty} C_k\right)$ is nonempty for every set $\overline{X}$ of dimension less than or equal to l that is specified by linear inequality constraints, and is such that $\overline{X} \cap C_k$ is nonempty for all k and $R_{\overline{X}} \cap F_f \subset L_f$. Let X be of the form

$$X = \{x \mid a_j' x \leq b_j,\ j = 1, \ldots, r\},$$

and be such that $X \cap C_k$ is nonempty for all k, satisfy $R_X \cap F_f \subset L_f$, and have dimension $l+1$. We will show that the intersection $X \cap \left(\cap_{k=0}^{\infty} C_k\right)$ is nonempty.

If $L_X \cap L_f = R_X \cap R_f$, then by Prop. 1.5.5 applied to the sets $X \cap C_k$, we have that $X \cap \left(\cap_{k=0}^{\infty} C_k\right)$ is nonempty, and we are done. We may thus assume that $L_X \cap L_f \neq R_X \cap R_f$. Let $\overline{y} \in R_X \cap R_f$ with $-\overline{y} \notin R_X \cap R_f$.

If $\overline{y} \notin F_f$, then, since $\overline{y} \in R_X \cap R_f$, for all $x \in X \cap \mathrm{dom}(f)$ we have $\lim_{\alpha \to \infty} f(x + \alpha y) = -\infty$ and $x + \alpha y \in X$ for all $\alpha \geq 0$. Therefore, $x + \alpha y \in X \cap \left(\cap_{k=0}^{\infty} C_k\right)$ for sufficiently large α, and we are done.

We may thus assume that $\overline{y} \in F_f$, so that $\overline{y} \in R_X \cap F_f$ and therefore also $\overline{y} \subset L_f$, in view of the hypothesis $R_X \cap F_f \subset L_f$. Since $-\overline{y} \notin R_X \cap R_f$, it follows that $-\overline{y} \notin R_X$. Thus, we have

$$\overline{y} \in R_X, \qquad -\overline{y} \notin R_X, \qquad \overline{y} \in L_f.$$

From this point onward, the proof that $X \cap \left(\cap_{k=0}^{\infty} C_k\right) \neq \emptyset$ is nearly identical to the corresponding part of the proof of Prop. 1.5.6.

Using Prop. 1.5.1(e), it is seen that the recession cone of X is

$$R_X = \{y \mid a_j' y \leq 0,\ j = 1, \ldots, r\},$$

so the fact $\overline{y} \in R_X$ implies that

$$a_j' \overline{y} \leq 0, \qquad \forall\ j = 1, \ldots, r,$$

while the fact $-\overline{y} \notin R_X$ implies that the index set

$$J = \{j \mid a_j' \overline{y} < 0\}$$

is nonempty.

Consider a sequence $\{x_k\}$ such that

$$x_k \in X \cap C_k, \quad \forall\, k.$$

We then have

$$a_j' x_k \le b_j, \qquad \forall\, j = 1, \ldots, r, \quad \forall\, k.$$

We may assume that

$$a_j' x_k < b_j, \qquad \forall\, j \in J, \quad \forall\, k;$$

otherwise we can replace x_k with $x_k + \overline{y}$, which belongs to $X \cap C_k$ (since $\overline{y} \in R_X$ and $\overline{y} \in L_f$).

Suppose that for each k, we start at x_k and move along $-\overline{y}$ as far as possible without leaving the set X, up to the point where we encounter the vector

$$\overline{x}_k = x_k - \beta_k \overline{y},$$

where β_k is the positive scalar given by

$$\beta_k = \min_{j \in J} \frac{a_j' x_k - b_j}{a_j' \overline{y}}.$$

Since $a_j' \overline{y} = 0$ for all $j \notin J$, we have $a_j' \overline{x}_k = a_j' x_k$ for all $j \notin J$, so the number of linear inequalities of X that are satisfied by $\overline{x}_k$ as equalities is strictly larger than the number of those satisfied by x_k. Thus, there exists $j_0 \in J$ such that $a_{j_0}' \overline{x}_k = b_{j_0}$ for all k in an infinite index set $\mathcal{K} \subset \{0, 1, \ldots\}$. By reordering the linear inequalities if necessary, we can assume that $j_0 = 1$, i.e.,

$$a_1' \overline{x}_k = b_1, \qquad a_1' x_k < b_1, \qquad \forall\, k \in \mathcal{K}.$$

To apply the induction hypothesis, consider the set

$$\overline{X} = \{x \mid a_1' x = b_1,\ a_j' x \le b_j,\ j = 2, \ldots, r\},$$

and note that $\{\overline{x}_k\}_{\mathcal{K}} \subset \overline{X}$. Since $\overline{x}_k = x_k - \beta_k \overline{y}$ with $x_k \in C_k$ and $\overline{y} \in L_f$, we have $\overline{x}_k \in C_k$ for all k, implying that $\overline{x}_k \in \overline{X} \cap C_k$ for all $k \in \mathcal{K}$. Thus, $\overline{X} \cap C_k \neq \emptyset$ for all k. Because the sets C_k are nested, so are the sets $\overline{X} \cap C_k$. Furthermore, the recession cone of $\overline{X}$ is

$$R_{\overline{X}} = \{y \mid a_1' y = 0,\ a_j' y \le 0,\ j = 2, \ldots, r\},$$

which is contained in R_X, so that

$$R_{\overline{X}} \cap F_f \subset R_X \cap F_f \subset L_f.$$

Finally, to show that the dimension of $\overline{X}$ is smaller than the dimension of X, note that the set $\{x \mid a_1' x = b_1\}$ contains $\overline{X}$, so that a_1 is orthogonal to the subspace $S_{\overline{X}}$ that is parallel to $\mathrm{aff}(\overline{X})$. Since $a_1' \overline{y} < 0$, it follows that $\overline{y} \notin S_{\overline{X}}$. On the other hand, $\overline{y}$ belongs to S_X, the subspace that is parallel to aff(X), since for all k, we have $x_k \in X$ and $x_k - \beta_k \overline{y} \in X$.

Based on the preceding, we can use the induction hypothesis to assert that the intersection $\overline{X} \cap \big(\cap_{k=0}^{\infty} C_k\big)$ is nonempty. Since $\overline{X} \subset X$, it follows that $X \cap \big(\cap_{k=0}^{\infty} C_k\big)$ is nonempty.

(b) We will use part (a) and the line of proof of Prop. 2.3.3 [condition (2)]. Denote

$$f^* = \inf_{x \in X} f(x),$$

and assume without loss of generality that $f^* = 0$ [otherwise, we replace $f(x)$ by $f(x) - f^*$]. We choose a scalar sequence $\{w_k\}$ such that $w_k \downarrow f^*$, and we consider the (nonempty) level sets

$$C_k = \big\{x \in \Re^n \mid f(x) \le w_k\big\}.$$

The set $X \cap C_k$ is nonempty for all k. Furthermore, by assumption, $R_X \cap F_f \subset L_f$ and X is specified by linear inequality constraints. By part (a), it follows that $X \cap \big(\cap_{k=0}^{\infty} C_k\big)$, the set of minimizers of f over X, is nonempty.

(c) Let $X = \Re$ and $f(x) = x$. Then

$$F_f = L_f = \big\{y \mid y = 0\big\},$$

so the condition $R_X \cap F_f \subset L_f$ is satisfied. However, we have $\inf_{x \in X} f(x) = -\infty$ and f does not attain a minimum over X. Note that Prop. 2.3.3 [under condition (2)] does not apply here, because the relation $R_X \cap R_f \subset L_f$ is not satisfied.

2.7 (Bidirectionally Flat Functions)

The purpose of the exercise is to provide refinements of the results involving convex quadratic functions and relating to set intersections, closedness under linear transformations, existence of optimal solutions, and closedness under partial minimization [cf. Props. 1.5.7, 1.5.8(c), 1.5.9, 2.3.3, and 2.3.9].

Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper convex function, and let F_f be the set of directions along which f is flat (cf. Exercise 2.6). We say that f is *bidirectionally flat* if $L_f = F_f$ (i.e., if it is flat in some direction it must be flat, and hence constant, in the opposite direction). Note that every convex quadratic function is bidirectionally flat. More generally, a function of the form

$$f(x) = h(Ax) + c'x,$$

where A is an $m \times n$ matrix and $h : \Re^m \mapsto (-\infty, \infty]$ is a coercive closed proper convex function, is bidirectionally flat. In this case, we have

$$L_f = F_f = \{y \mid Ay = 0,\, c'y = 0\}.$$

Let $g_j : \Re^n \mapsto (-\infty, \infty]$, $j = 0, 1, \ldots, r$, be closed proper convex functions that are bidirectionally flat.

(a) Assume that each vector x such that $g_0(x) \le 0$ belongs to $\cap_{j=1}^r \mathrm{dom}(g_j)$, and that for some scalar sequence $\{w_k\}$ with $w_k \downarrow 0$, the set

$$C_k = \big\{x \mid g_0(x) \le w_k,\ g_j(x) \le 0,\ j = 1, \ldots, r\big\}$$

is nonempty for each k. Show that the intersection $\cap_{k=0}^\infty C_k$ is nonempty.

(b) Assume that each g_j, $j = 1, \ldots, r$, is real-valued and the set

$$C = \big\{x \mid g_j(x) \le 0,\ j = 1, \ldots, r\big\}$$

is nonempty. Show that for any $m \times n$ matrix A, the set AC is closed.

(c) Show that a closed proper convex function $f : \Re^n \mapsto (-\infty, \infty]$ that is bidirectionally flat attains a minimum over the set C of part (b), provided that $\inf_{x \in C} f(x)$ is finite.

(d) Let $F : \Re^{n+m} \mapsto (-\infty, \infty]$ be a function of the form

$$F(x, z) = \begin{cases} \overline{F}(x, z) & \text{if } (x, z) \in C, \\ \infty & \text{otherwise,} \end{cases}$$

where $\overline{F}$ is a bidirectionally flat real-valued convex function on $\Re^{n+m}$ and C is a subset of $\Re^{n+m}$ that is specified by r convex inequalities involving bidirectionally flat real-valued convex functions on $\Re^{n+m}$ [cf. part (b)]. Consider the function

$$p(x) = \inf_{z \in \Re^m} F(x, z).$$

Assume that there exists a vector $\overline{x} \in \Re^n$ such that $-\infty < p(\overline{x}) < \infty$. Show that p is convex, closed, and proper. Furthermore, for each $x \in \mathrm{dom}(p)$, the set of points that attain the infimum of $F(x, \cdot)$ over $\Re^m$ is nonempty.

Solution: (a) As a first step, we will show that either $\cap_{k=1}^\infty C_k \ne \varnothing$ or else

$$\text{there exists } \overline{j} \in \{1, \ldots, r\} \text{ and } y \in \cap_{j=0}^r R_{g_j} \text{ with } y \notin F_{g_{\overline{j}}}.$$

Let $\overline{x}$ be a vector in C_0, and for each $k \ge 1$, let x_k be the projection of $\overline{x}$ on C_k. If $\{x_k\}$ is bounded, then since the g_j are closed, any limit point $\tilde{x}$ of $\{x_k\}$ satisfies

$$g_j(\tilde{x}) \le \liminf_{k \to \infty} g_j(x_k) \le 0,$$

so $\tilde{x} \in \cap_{k=1}^\infty C_k$, and $\cap_{k=1}^\infty C_k \ne \varnothing$. If $\{x_k\}$ is unbounded, let y be a limit point of the sequence $\big\{(x_k - \overline{x})/\|x_k - \overline{x}\| \mid x_k \ne \overline{x}\big\}$, and without loss of generality, assume that

$$\frac{x_k - \overline{x}}{\|x_k - \overline{x}\|} \to y.$$

We claim that

$$y \in \cap_{j=0}^r R_{g_j}.$$

Indeed, if for some j, we have $y \notin R_{g_j}$, then there exists $\alpha > 0$ such that $g_j(\overline{x} + \alpha y) > w_0$. Let

$$z_k = \overline{x} + \alpha \frac{x_k - \overline{x}}{\|x_k - \overline{x}\|},$$

and note that for sufficiently large k, z_k lies in the line segment connecting $\overline{x}$ and x_k, so that $g_1(z_k) \le w_0$. On the other hand, we have $z_k \to \overline{x} + \alpha y$, so using the closedness of g_j, we must have

$$g_j(\overline{x} + \alpha y) \le \liminf_{k \to \infty} g_1(z_k) \le w_0,$$

which contradicts the choice of α to satisfy $g_j(\overline{x} + \alpha y) > w_0$.

If $y \in \cap_{j=0}^r F_{g_j}$, since all the g_j are bidirectionally flat, we have $y \in \cap_{j=0}^r L_{g_j}$. If the vectors $\overline{x}$ and x_k, $k \ge 1$, all lie in the same line [which must be the line $\{\overline{x} + \alpha y \mid \alpha \in \Re\}$], we would have $g_j(\overline{x}) = g_j(x_k)$ for all k and j. Then it follows that $\overline{x}$ and x_k all belong to $\cap_{k=1}^\infty C_k$. Otherwise, there must be some x_k, with k large enough, and such that, by the Projection Theorem, the vector $x_k - \alpha y$ makes an angle greater than $\pi/2$ with $x_k - \overline{x}$. Since the g_j are constant on the line $\{x_k - \alpha y \mid \alpha \in \Re\}$, all vectors on the line belong to C_k, which contradicts the fact that x_k is the projection of $\overline{x}$ on C_k.

Finally, if $y \in R_{g_0}$ but $y \notin F_{g_0}$, we have $g_0(x + \alpha y) \to -\infty$ as $\alpha \to \infty$, so that $\cap_{k=1}^\infty C_k \ne \emptyset$. This completes the proof that

$$\cap_{k=1}^\infty C_k = \emptyset \;\Rightarrow\; \text{there exists } \overline{j} \in \{1, \ldots, r\} \text{ and } y \in \cap_{j=0}^r R_j \text{ with } y \notin F_{g_{\overline{j}}}. \quad (1)$$

We now use induction on r. For $r = 0$, the preceding proof shows that $\cap_{k=1}^\infty C_k \ne \emptyset$. Assume that $\cap_{k=1}^\infty C_k \ne \emptyset$ for all cases where $r < \overline{r}$. We will show that $\cap_{k=1}^\infty C_k \ne \emptyset$ for $r = \overline{r}$. Assume the contrary. Then, by Eq. (1), there exists $\overline{j} \in \{1, \ldots, r\}$ and $y \in \cap_{j=0}^r R_j$ with $y \notin F_{g_{\overline{j}}}$. Let us consider the sets

$$\overline{C}_k = \big\{x \mid g_0(x) \le w_k,\, g_j(x) \le 0,\, j = 1, \ldots, r,\, j \ne \overline{j}\big\}.$$

Since these sets are nonempty, by the induction hypothesis, $\cap_{k=1}^\infty \overline{C}_k \ne \emptyset$. For any $\tilde{x} \in \cap_{k=1}^\infty \overline{C}_k$, the vector $\tilde{x} + \alpha y$ belongs to $\cap_{k=1}^\infty \overline{C}_k$ for all $\alpha > 0$, since $y \in \cap_{j=0}^r R_j$. Since $g_0(\tilde{x}) \le 0$, we have $\tilde{x} \in \mathrm{dom}(g_{\overline{j}})$, by the hypothesis regarding the domains of the g_j. Since $y \in \cap_{j=0}^r R_j$ with $y \notin F_{g_{\overline{j}}}$, it follows that $g_{\overline{j}}(\tilde{x} + \alpha y) \to -\infty$ as $\alpha \to \infty$. Hence, for sufficiently large α, we have $g_{\overline{j}}(\tilde{x} + \alpha y) \le 0$, so $\tilde{x} + \alpha y$ belongs to $\cap_{k=1}^\infty C_k$.

Note: To see that the assumption

$$\big\{x \mid g_0(x) \le 0\big\} \subset \cap_{j=1}^r \mathrm{dom}(g_j)$$

is essential for the result to hold, consider an example in $\Re^2$. Let

$$g_0(x_1, x_2) = x_1, \qquad g_1(x_1, x_2) = \phi(x_1) - x_2,$$

where the function $\phi : \Re \mapsto (-\infty, \infty]$ is convex, closed, and coercive with $\mathrm{dom}(\phi) = (0, 1)$ [for example, $\phi(t) = -\ln t - \ln(1 - t)$ for $0 < t < 1$]. Then it can be verified that $C_k \neq \emptyset$ for every k and sequence $\{w_k\} \subset (0,1)$ with $w_k \downarrow 0$ [take $x_1 \downarrow 0$ and $x_2 \geq \phi(x_1)$]. On the other hand, we have $\cap_{k=0}^{\infty} C_k = \emptyset$. The difficulty here is that the set $\{x \mid g_0(x) \leq 0\}$, which is equal to

$$\{x \mid x_1 \leq 0,\, x_2 \in \Re\},$$

is not contained in $\mathrm{dom}(g_1)$, which is equal to

$$\{x \mid 0 < x_1 < 1,\, x_2 \in \Re\}$$

(in fact the two sets are disjoint).

(b) We will use part (a) and the line of proof of Prop. 1.5.8(c). In particular, let $\{y_k\}$ be a sequence in $A\,C$ converging to some $\overline{y} \in \Re^n$. We will show that $\overline{y} \in A\,C$. We let

$$g_0(x) = \|Ax - \overline{y}\|^2, \qquad w_k = \|y_k - \overline{y}\|^2,$$

and

$$C_k = \big\{x \mid g_0(x) \leq w_k,\, g_j(x) \leq 0,\, j = 1, \ldots, r\big\}.$$

The functions involved in the definition of C_k are bidirectionally flat, and each C_k is nonempty by construction. By applying part (a), we see that the intersection $\cap_{k=0}^{\infty} C_k$ is nonempty. For any x in this intersection, we have $Ax = \overline{y}$ (since $y_k \to \overline{y}$), showing that $\overline{y} \in A\,C$.

(c) We will use part (a) and the line of proof of Prop. 2.3.3 [condition (3)]. Denote

$$f^* = \inf_{x \in C} f(x),$$

and assume without loss of generality that $f^* = 0$ [otherwise, we replace $f(x)$ by $f(x) - f^*$]. We choose a scalar sequence $\{w_k\}$ such that $w_k \downarrow f^*$, and we consider the (nonempty) sets

$$C_k = \big\{x \in \Re^n \mid f(x) \leq w_k,\, g_j(x) \leq 0,\, j = 1, \ldots, r\big\}.$$

By part (a), it follows that $\cap_{k=0}^{\infty} C_k$, the set of minimizers of f over C, is nonempty.

(d) Use the line of proof of Prop. 2.3.9.

2.8 (Minimization of Quasiconvex Functions)

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

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

are convex. Let X be a convex subset of $\Re^n$, let f be a quasiconvex function such that $X \cap \mathrm{dom}(f) \neq \emptyset$, 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. 2.3.3 and Exercise 2.7 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 the set X is of the form

$$X = \{x \mid a_j' x \le b_j,\, j = 1, \ldots, r\},$$

where a_j are vectors in $\Re^n$ and b_j are scalars.

(3) $f^* > -\infty$, the set X is of the form

$$X = \Big\{x \mid x' Q_j x + a_j' x + b_j \le 0,\ j = 1, \ldots, r\Big\},$$

where Q_j are symmetric positive semidefinite $n \times n$ matrices, a_j are vectors in $\Re^n$, and b_j are scalars, and for some $\overline{\gamma} \in \Gamma$ and all $\gamma \in \Gamma$ with $\gamma \le \overline{\gamma}$, the level sets V_γ are of the form

$$V_\gamma = \Big\{x \mid x' Q x + c' x + b(\gamma) \le 0\Big\},$$

where Q is a symmetric positive semidefinite $n \times n$ matrix, c is a vector in $\Re^n$, and $b(\gamma)$ is a monotonically nondecreasing function of γ, such that the set $\big\{b(\gamma) \mid f^* < \gamma \le \overline{\gamma}\big\}$ is bounded.

(4) $f^* > -\infty$, the set X is of the form

$$X = \Big\{x \mid g_j(x) \le 0,\ j = 1, \ldots, r\Big\},$$

and for some $\overline{\gamma} \in \Gamma$ and all $\gamma \in \Gamma$ with $\gamma \le \overline{\gamma}$, the level sets V_γ are of the form

$$V_\gamma = \Big\{x \mid g_0(x) + b(\gamma) \le 0\Big\},$$

where g_j, $j = 0, 1, \ldots, r$, are real-valued, convex, and bidirectionally flat functions (cf. Exercise 2.7), and $b(\gamma)$ is a monotonically nondecreasing function of γ, such that the set $\big\{b(\gamma) \mid f^* < \gamma \le \overline{\gamma}\big\}$ is bounded.

Solution: (a) Let x^* be a local minimum of f over X and assume, to arrive at a contradiction, that there exists a vector $\overline{x} \in X$ such that $f(\overline{x}) < f(x^*)$. Then, $\overline{x}$ 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 $\overline{x}$ belongs to the set, implying that

$$f\big(\alpha \overline{x} + (1 - \alpha) x^*\big) \le f(x^*), \qquad \forall\, \alpha \in [0, 1]. \tag{1}$$

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

$$f\big(\alpha_k \overline{x} + (1-\alpha_k)x^*\big) < f(x^*), \qquad \text{for some } \alpha_k \in (0, 1/k]; \tag{2}$$

otherwise, in view of Eq. (1), we would have that $f(x)$ is constant for x on the line segment connecting x^* and $(1/k)\overline{x} + \big(1-(1/k)\big)x^*$. Equation (2) contradicts the local optimality of x^*.

(b) We consider the level sets

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

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 $C \subset D$, we have $R_C \subset R_D$, 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)-(4), we 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 by Prop. 1.5.5 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 $R_X \cap R_f \subset L_f$ and X is specified by linear inequalities. By Prop. 1.5.6, it follows that X^* is nonempty.

Let condition (3) hold. The sets V_{γ^k} have the form

$$V_{\gamma^k} = \{x \in \Re^n \mid x'Qx + c'x + b(\gamma^k) \leq 0\}.$$

In view of the assumption that $b(\gamma)$ is bounded for $\gamma \in (f^*, \overline{\gamma}]$, we can consider a subsequence $\{b(\gamma^k)\}_\mathcal{K}$ that converges to a scalar. Furthermore, X is specified by convex quadratic inequalities, and the intersection $X \cap V_{\gamma^k}$ is nonempty for all $k \in \mathcal{K}$. By Prop. 1.5.7, it follows that X^* is nonempty.

Similarly, under condition (4), the result follows using Exercise 2.7(a).

2.9 (Partial Minimization)

(a) Let $f : \Re^n \mapsto [-\infty, \infty]$ be a function and consider the subset of $\Re^{n+1}$ given by

$$E_f = \big\{(x, w) \mid f(x) < w\big\}.$$

Show that E_f is related to the epigraph of f as follows:

$$E_f \subset \text{epi}(f) \subset \text{cl}(E_f).$$

Show also that f is convex if and only if E_f is convex.

(b) Let $F : \Re^{m+n} \mapsto [-\infty, \infty]$ be a function and let

$$f(x) = \inf_{z \in \Re^m} F(x, z), \qquad x \in \Re^n.$$

Show that E_f is the projection of the set $\big\{(x, z, w) \mid F(x, z) < w\big\}$ on the space of (x, w).

(c) Use parts (a) and (b) to show that convexity of F implies convexity of f (cf. Prop. 2.3.5).

Solution: (a) The epigraph of f is given by

$$\text{epi}(f) = \big\{(x, w) \mid f(x) \le w\big\}.$$

If $(x, w) \in E_f$, then it follows that $(x, w) \in \text{epi}(f)$, showing that $E_f \subset \text{epi}(f)$. Next, assume that $(x, w) \in \text{epi}(f)$, i.e., $f(x) \le w$. Let $\{w_k\}$ be a sequence with $w_k > w$ for all k , and $w_k \to w$. Then we have, $f(x) < w_k$ for all k, implying that $(x, w_k) \in E_f$ for all k, and that the limit $(x, w) \in \text{cl}(E_f)$. Thus we have the desired relations,

$$E_f \subset \text{epi}(f) \subset \text{cl}(E_f). \tag{2.5}$$

We next show that f is convex if and only if E_f is convex. By definition, f is convex if and only if epi(f) is convex. Assume that epi(f) is convex. Suppose, to arrive at a contradiction, that E_f is not convex. This implies the existence of vectors $(x_1, w_1) \in E_f$, $(x_2, w_2) \in E_f$, and a scalar $\alpha \in (0, 1)$ such that $\alpha(x_1, w_1)+ (1 - \alpha)(x_2, w_2) \notin E_f$, from which we get

$$\begin{aligned} f\big(\alpha x_1 + (1-\alpha)x_2\big) &\ge \alpha w_1 + (1-\alpha)w_2 \\ &> \alpha f(x_1) + (1-\alpha) f(x_2), \end{aligned} \tag{2.6}$$

where the second inequality follows from the fact that (x_1, w_1) and (x_2, w_2) belong to E_f. We have $\big(x_1, f(x_1)\big) \in \text{epi}(f)$ and $\big(x_2, f(x_2)\big) \in \text{epi}(f)$. In view of the convexity assumption of epi(f), this yields $\alpha\big(x_1, f(x_1)\big) + (1-\alpha)\big(x_2, f(x_2)\big) \in$ epi(f) and therefore,

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2).$$

Combined with Eq. (2.6), the preceding relation yields a contradiction, thus showing that E_f is convex.

Next assume that E_f is convex. We show that epi(f) is convex. Let (x_1, w_1) and (x_2, w_2) be arbitrary vectors in epi(f). Consider sequences of vectors (x_1, w_1^k) and (x_2, w_2^k) such that $w_1^k > w_1$, $w_2^k > w_2$, and $w_1^k \to w_1$, $w_2^k \to w_2$. It follows that for each k, (x_1, w_1^k) and (x_2, w_2^k) belong to E_f. Since E_f is convex by assumption, this implies that for each $\alpha \in [0,1]$ and all k, the vector $\big(\alpha x_1 + (1-\alpha)x_2, \alpha w_1^k + (1-\alpha)w_2^k\big) \in E_f$, i.e., we have for each k

$$f\big(\alpha x_1 + (1-\alpha)x_2\big) < \alpha w_1^k + (1-\alpha)w_2^k.$$

Taking the limit in the preceding relation, we get

$$f\big(\alpha x_1 + (1-\alpha)x_2\big) \leq \alpha w_1 + (1-\alpha)w_2,$$

showing that $\big(\alpha x_1 + (1-\alpha)x_2, \alpha w_1 + (1-\alpha)w_2\big) \in \text{epi}(f)$. Hence epi$(f)$ is convex.

(b) Let T denote the projection of the set $\big\{(x,z,w) \mid F(x,z) < w\big\}$ on the space of (x,w). We show that $E_f = T$. Let $(x,w) \in E_f$. By definition, we have

$$\inf_{z \in \Re^m} F(x,z) < w,$$

which implies that there exists some $\overline{z} \in \Re^m$ such that

$$F(x, \overline{z}) < w,$$

showing that $(x, \overline{z}, w)$ belongs to the set $\big\{(x,z,w) \mid F(x,z) < w\big\}$, and $(x,w) \in T$. Conversely, let $(x,w) \in T$. This implies that there exists some z such that $F(x,z) < w$, from which we get

$$f(x) = \inf_{z \in \Re^m} F(x,z) < w,$$

showing that $(x,w) \in E_f$, and completing the proof.

(c) Let F be a convex function. Using part (a), the convexity of F implies that the set $\big\{(x,z,w) \mid F(x,z) < w\big\}$ is convex. Since the projection mapping is linear, and hence preserves convexity, we have, using part (b), that the set E_f is convex, which implies by part (a) that f is convex.

2.10 (Partial Minimization of Nonconvex Functions)

Let $f : \Re^n \times \Re^m \mapsto (-\infty, \infty]$ be a closed proper function. Assume that f has the following property: for all $u^* \in \Re^m$ and $\gamma \in \Re$, there exists a neighborhood N of u^* such that the set $\big\{(x,u) \mid u \in N, f(x,u) \leq \gamma\big\}$ is bounded, i.e., $f(x,u)$ is *level-bounded in x locally uniformly in u*. Let

$$p(u) = \inf_x f(x,u), \qquad P(u) = \arg\min_x f(x,u), \qquad u \in \Re^m.$$

(a) Show that the function p is closed and proper. Show also that for each $u \in \text{dom}(p)$, the set $P(u)$ is nonempty and compact.

(b) Consider the following weaker assumption: for all $u \in \Re^m$ and $\gamma \in \Re$, the set $\{x \mid f(x,u) \leq \gamma\}$ is bounded. Show that this assumption is not sufficient to guarantee the closedness of p.

(c) Let $\{u_k\}$ be a sequence such that $u_k \to u^*$ for some $u^* \in \text{dom}(p)$, and assume that $p(u_k) \to p(u^*)$ (which holds when p is continuous at u^*). Let also $x_k \in P(u_k)$ for all k. Show that the sequence $\{x_k\}$ is bounded and all its limit points lie in $P(u^*)$.

(d) Show that a sufficient condition for p to be continuous at u^* is the existence of some $x^* \in P(u^*)$ such that $f(x^*, \cdot)$ is continuous at u^*.

Solution: (a) For each $u \in \Re^m$, let $f_u(x) = f(x,u)$. There are two cases; either $f_u \equiv \infty$, or f_u is lower semicontinuous with bounded level sets. The first case, which corresponds to $p(u) = \infty$, can't hold for every u, since f is not identically equal to ∞. Therefore, $\text{dom}(p) \neq \emptyset$, and for each $u \in \text{dom}(p)$, we have by Weierstrass' Theorem that $p(u) = \inf_x f_u(x)$ is finite [i.e., $p(u) > -\infty$ for all $u \in \text{dom}(p)$] and the set $P(u) = \arg\min_x f_u(x)$ is nonempty and compact.

We now show that p is lower semicontinuous. By assumption, for all $\overline{u} \in \Re^m$ and for all $\alpha \in \Re$, there exists a neighborhood $\overline{N}$ of $\overline{u}$ such that the set $\{(x,u) \mid f(x,u) \leq \alpha\} \cap (\Re^n \times \overline{N})$ is bounded in $\Re^n \times \Re^m$. We can choose a smaller closed set N containing $\overline{u}$ such that the set $\{(x,u) \mid f(x,u) \leq \alpha\} \cap (\Re^n \times N)$ is closed (since f is lower semicontinuous) and bounded. In view of the assumption that f_u is lower semicontinuous with bounded level sets, it follows using Weierstrass' Theorem that for any scalar α,

$$p(u) \leq \alpha \text{ if and only if there exists } x \text{ such that } f(x,u) \leq \alpha.$$

Hence, the set $\{u \mid p(u) \leq \alpha\} \cap N$ is the image of the set $\{(x,u) \mid f(x,u) \leq \alpha\} \cap (\Re^n \times N)$ under the continuous mapping $(x,u) \mapsto u$. Since the image of a compact set under a continuous mapping is compact [cf. Prop. 1.1.9(d)], we see that $\{u \mid p(u) \leq \alpha\} \cap N$ is closed.

Thus, each $\overline{u} \in \Re^m$ is contained in a closed set whose intersection with $\{u \mid p(u) \leq \alpha\}$ is closed, so that the set $\{u \mid p(u) \leq \alpha\}$ itself is closed for all scalars α. It follows from Prop. 1.2.2 that p is lower semicontinuous.

(b) Consider the following example

$$f(x,u) = \begin{cases} \min\{|x - 1/u|, 1 + |x|\} & \text{if } u \neq 0,\ x \in \Re, \\ 1 + |x| & \text{if } u = 0,\ x \in \Re, \end{cases}$$

where x and u are scalars. This function is continuous in (x,u) and the level sets are bounded in x for each u, but not *locally uniformly* in u, i.e., there does not exists a neighborhood N of $u = 0$ such that the set $\{(x,u) \mid u \in N,\ f(x,u) \leq \alpha\}$ is bounded for some $\alpha > 0$.

For this function, we have

$$p(u) = \begin{cases} 0 & \text{if } u \neq 0, \\ 1 & \text{if } u = 0. \end{cases}$$

Hence, the function p is not lower semicontinuous at 0.

(c) Let $\{u_k\}$ be a sequence such that $u_k \to u^*$ for some $u^* \in \text{dom}(p)$, and also $p(u_k) \to p(u^*)$. Let α be any scalar such that $p(u^*) < \alpha$. Since $p(u_k) \to p(u^*)$, we obtain

$$f(x_k, u_k) = p(u_k) < \alpha, \tag{2.7}$$

for all k sufficiently large, where we use the fact that $x_k \in P(u_k)$ for all k. We take N to be a closed neighborhood of u^* as in part (a). Since $u_k \to u^*$, using Eq. (2.7), we see that for all k sufficiently large, the pair (x_k, u_k) lies in the compact set

$$\big\{(x,u) \mid f(x,u) \leq \alpha\big\} \cap (\Re^n \times N).$$

Hence, the sequence $\{x_k\}$ is bounded, and therefore has a limit point, call it x^*. It follows that

$$(x^*, u^*) \in \big\{(x,u) \mid f(x,u) \leq \alpha\big\}.$$

Since this is true for arbitrary $\alpha > p(u^*)$, we see that $f(x^*, u^*) \leq p(u^*)$, which, by the definition of $p(u)$, implies that $x^* \in P(u^*)$.

(d) By definition, we have $p(u) \leq f(x^*, u)$ for all u and $p(u^*) = f(x^*, u^*)$. Since $f(x^*, \cdot)$ is continuous at u^*, we have for any sequence $\{u_k\}$ converging to u^*

$$\limsup_{k\to\infty} p(u_k) \leq \limsup_{k\to\infty} f(x^*, u_k) = f(x^*, u^*) = p(u^*),$$

thereby implying that p is upper semicontinuous at u^*. Since p is also lower semicontinuous at u^* by part (a), we conclude that p is continuous at u^*.

2.11 (Projection on a Nonconvex Set)

Let C be a nonempty closed subset of $\Re^n$.

(a) Show that the distance function $d_C(x) : \Re^n \mapsto \Re$, defined by

$$d_C(x) = \inf_{w \in C} \|w - x\|,$$

is a continuous function of x.

(b) Show that the projection set $P_C(x)$, defined by

$$P_C(x) = \arg\min_{w \in C} \|w - x\|,$$

is nonempty and compact.

(c) Let $\{x_k\}$ be a sequence that converges to a vector x^* and let $\{w_k\}$ be a sequence with $w_k \in P_C(x_k)$. Show that the sequence $\{w_k\}$ is bounded and all its limit points belong to the set $P_C(x^*)$.

Solution: We define the function f by

$$f(w,x) = \begin{cases} \|w - x\| & \text{if } w \in C, \\ \infty & \text{if } w \notin C. \end{cases}$$

With this identification, we get

$$d_C(x) = \inf_w f(w,x), \qquad P_C(x) = \arg\min_w f(w,x).$$

We now show that $f(w,x)$ satisfies the assumptions of Exercise 2.10, so that we can apply the results of this exercise to this problem.

Since the set C is closed by assumption, it follows that $f(w,x)$ is lower semicontinuous. Moreover, by Weierstrass' Theorem, we see that $f(w,x) > -\infty$ for all x and w. Since the set C is nonempty by assumption, we also have that $\text{dom}(f)$ is nonempty. It is also straightforward to see that the function $\|\cdot\|$, and therefore the function f, satisfies the locally uniformly level-boundedness assumption of Exercise 2.10.

(a) Since the function $\|\cdot\|$ is lower semicontinuous and the set C is closed, it follows from Weierstrass' Theorem that for all $x^* \in \Re^n$, the infimum in $\inf_w f(w,x)$ is attained at some w^*, i.e., $P(x^*)$ is nonempty. Hence, we see that for all $x^* \in \Re^n$, there exists some $w^* \in P(x^*)$ such that $f(w^*,\cdot)$ is continuous at x^*, which follows by continuity of the function $\|\cdot\|$. Hence, the function $f(w,x)$ satisfies the sufficiency condition given in Exercise 2.10(d), and it follows that $d_C(x)$ depends continuously on x.

(b) This part follows from part (a) of Exercise 2.10.

(c) This part follows from part (c) of Exercise 2.10.

2.12 (Convergence of Penalty Methods [RoW98])

Let $f : \Re^n \mapsto (-\infty,\infty]$ be a closed proper function, let $F : \Re^n \mapsto \Re^m$ be a continuous function, and let $D \subset \Re^m$ be a nonempty closed set. Consider the problem

$$\begin{aligned} &\text{minimize} \;\; f(x) \\ &\text{subject to } F(x) \in D. \end{aligned}$$

Consider also the following approximation of this problem:

$$\begin{aligned} &\text{minimize} \;\; f(x) + \theta\big(F(x), c\big) \\ &\text{subject to } x \in \Re^n, \end{aligned} \tag{P$_c$}$$

with $c \in (0,\infty)$, where the function $\theta : \Re^m \times (0,\infty) \mapsto (-\infty,\infty]$ is lower semicontinuous, monotonically increasing in c for each $u \in \Re^m$, and satisfies

$$\lim_{c\to\infty} \theta(u,c) = \begin{cases} 0 & \text{if } u \in D, \\ \infty & \text{if } u \notin D. \end{cases}$$

Assume that for some $\overline{c} \in (0,\infty)$ sufficiently large, the level sets of the function $f(x) + \theta\big(F(x),\overline{c}\big)$ are bounded, and consider any sequence of parameter values $c_k \geq \overline{c}$ with $c_k \to \infty$. Show the following:

(a) The sequence of optimal values of the approximate problems (P_{c_k}) converges to the optimal value of the original problem.

(b) Any sequence $\{x_k\}$, where x_k is an optimal solution of the approximate problem (P_{c_k}), is bounded and each of its limit points is an optimal solution of the original problem.

Solution: (a) We set $\overline{s} = 1/\overline{c}$ and consider the function $g(x,s) : \Re^n \times \Re \mapsto (-\infty, \infty]$ defined by

$$g(x,s) = f(x) + \tilde{\theta}\big(F(x), s\big),$$

with the function $\tilde{\theta}$ given by

$$\tilde{\theta}(u,s) = \begin{cases} \theta(u, 1/s) & \text{if } s \in (0, \overline{s}], \\ \delta_D(u) & \text{if } s = 0, \\ \infty & \text{if } s < 0 \text{ or } s > \overline{s}, \end{cases}$$

where

$$\delta_D(u) = \begin{cases} 0 & \text{if } u \in D, \\ \infty & \text{if } u \notin D. \end{cases}$$

We identify the original problem with that of minimizing $g(x,0)$ in $x \in \Re^n$, and the approximate problem for parameter $s \in (0, \overline{s}]$ with that of minimizing $g(x,s)$ in $x \in \Re^n$ where $s = 1/c$. With the notation introduced in Exercise 2.10, the optimal value of the original problem is given by $p(0)$ and the optimal value of the approximate problem is given by $p(s)$. Hence, we have

$$p(s) = \inf_{x \in \Re^n} g(x,s).$$

We now show that, for the function $g(x,s)$, the assumptions of Exercise 2.10 are satisfied.

We have that $g(x,s) > -\infty$ for all (x,s), since by assumption $f(x) > -\infty$ for all x and $\theta(u,s) > -\infty$ for all (u,s). The function $\tilde{\theta}$ is such that $\tilde{\theta}(u,s) < \infty$ at least for one vector (u,s), since the set D is nonempty. Therefore, it follows that $g(x,s) < \infty$ for at least one vector (x,s), unless $g \equiv \infty$, in which case all the results of this exercise follow trivially.

We now show that the function $\tilde{\theta}$ is lower semicontinuous. This is easily seen at all points where $s \neq 0$ in view of the assumption that the function θ is lower semicontinuous on $\Re^m \times (0, \infty)$. We next consider points where $s = 0$. We claim that for any $\alpha \in \Re$,

$$\big\{u \mid \tilde{\theta}(u,0) \le \alpha\big\} = \bigcap_{s \in (0,\overline{s}]} \big\{u \mid \tilde{\theta}(u,s) \le \alpha\big\}. \tag{2.8}$$

To see this, assume that $\tilde{\theta}(u,0) \le \alpha$. Since $\tilde{\theta}(u,s) \uparrow \tilde{\theta}(u,0)$ as $s \downarrow 0$, we have $\tilde{\theta}(u,s) \le \alpha$ for all $s \in (0, \overline{s}]$. Conversely, assume that $\tilde{\theta}(u,s) \le \alpha$ for all $s \in (0, \overline{s}]$. By definition of $\tilde{\theta}$, this implies that

$$\theta(u, 1/s) \le \alpha, \qquad \forall\ s \in (0, s_0].$$

Taking the limit as $s \to 0$ in the preceding relation, we get

$$\lim_{s \to 0} \theta(u, 1/s) = \delta_D(u) = \tilde{\theta}(u,0) \le \alpha,$$

thus, proving the relation in (2.8). Note that for all $\alpha \in \Re$ and all $s \in (0, \overline{s}]$, the set

$$\big\{u \mid \tilde{\theta}(u,s) \le \alpha\big\} = \big\{u \mid \theta(u,1/s) \le \alpha\big\},$$

is closed by the lower semicontinuity of the function θ. Hence, the relation in Eq. (2.8) implies that the set $\big\{u \mid \tilde{\theta}(u,0) \le \alpha\big\}$ is closed for all $\alpha \in \Re$, thus showing that the function $\tilde{\theta}$ is lower semicontinuous everywhere (cf. Prop. 1.2.2). Together with the assumptions that f is lower semicontinuous and F is continuous, it follows that g is lower semicontinuous.

Finally, we show that g satisfies the locally uniform level boundedness property given in Exercise 2.10, i.e., for all $s^* \in \Re$ and for all $\alpha \in \Re$, there exists a neighborhood N of s^* such that the set $\big\{(x,s) \mid s \in N, g(x,s) \le \alpha\big\}$ is bounded. By assumption, we have that the level sets of the function $g(x,\overline{s}) = f(x) + \tilde{\theta}\big(F(x), 1/\overline{s}\big)$ are bounded. The definition of $\tilde{\theta}$, together with the fact that $\tilde{\theta}(u,s)$ is monotonically increasing as $s \downarrow 0$, implies that g is indeed level-bounded in x locally uniformly in s.

Therefore, all the assumptions of Exercise 2.10 are satisfied and we get that the function p is lower semicontinuous in s. Since $\tilde{\theta}(u,s)$ is monotonically increasing as $s \downarrow 0$, it follows that p is monotonically nondecreasing as $s \downarrow 0$. This implies that

$$p(s) \to p(0), \qquad \text{as } s \downarrow 0.$$

Defining $s_k = 1/c_k$ for all k, where $\{c_k\}$ is the given sequence of parameter values, we get

$$p(s_k) \to p(0),$$

thus proving that the optimal value of the approximate problem converges to the optimal value of the original problem.

(b) We have by assumption that $s_k \to 0$ with $x_k \in P_{1/s_k}$. It follows from part (a) that $p(s_k) \to p(0)$, so Exercise 2.10(b) implies that the sequence $\{x_k\}$ is bounded and all its limit points are optimal solutions of the original problem.

2.13 (Approximation by Envelope Functions [RoW98])

Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper function. For a scalar $c > 0$, define the corresponding *envelope function* $e_c f$ and the *proximal mapping* $P_c f$ by

$$e_c f(x) = \inf_w \left\{ f(w) + \frac{1}{2c}\|w - x\|^2 \right\},$$

$$P_c f(x) = \arg\min_w \left\{ f(w) + \frac{1}{2c}\|w - x\|^2 \right\}.$$

Assume that there exists some $c > 0$ with $e_c f(x) > -\infty$ for some $x \in \Re^n$. Let c_f be the supremum of the set of all such c. Show the following:

(a) For every $c \in (0, c_f)$, the set $P_c f(x)$ is nonempty and compact, while the value $e_c f(x)$ is finite and depends continuously on (x, c) with

$$e_c f(x) \uparrow f(x) \text{ for all } x, \text{ as } c \downarrow 0.$$

(b) Let $\{w_k\}$ be a sequence such that $w_k \in P_{c_k} f(x_k)$ for some sequences $\{x_k\}$ and $\{c_k\}$ such that $x_k \to x^*$ and $c_k \to c^* \in (0, c_f)$. Then $\{w_k\}$ is bounded and all its limit points belong to the set $P_{c^*} f(x^*)$.

[*Note*: The approximation $e_c f$ is an underestimate of the function f, i.e., $e_c f(x) \leq f(x)$ for all $x \in \Re^n$. Furthermore, $e_c f$ is a real-valued continuous function, whereas f itself may only be extended real-valued and lower semicontinuous.]

Solution: (a) We fix a $c_0 \in (0, c_f)$ and consider the function

$$h(w, x, c) = \begin{cases} f(w) + (\frac{1}{2c})\|w - x\|^2 & \text{if } c \in (0, c_0], \\ f(x) & \text{if } c = 0 \text{ and } w = x, \\ \infty & \text{otherwise.} \end{cases}$$

We consider the problem of minimizing $h(w, x, c)$ in w. With this identification and using the notation introduced in Exercise 2.10, for some $c \in (0, c_0)$, we obtain

$$e_c f(x) = p(x, c) = \inf_w h(w, x, c),$$

and

$$P_c f(x) = P(x, c) = \arg\min_w h(w, x, c).$$

We now show that, for the function $h(w, x, c)$, the assumptions given in Exercise 2.10 are satisfied.

We have that $h(w, x, c) > -\infty$ for all (w, x, c), since by assumption $f(x) > -\infty$ for all $x \in \Re^n$. Furthermore, $h(w, x, c) < \infty$ for at least one vector (w, x, c), since by assumption $f(x) < \infty$ for at least one vector $x \in X$.

We next show that the function h is lower semicontinuous in (w, x, c). This is easily seen at all points where $c \in (0, c_0]$ in view of the assumption that f is lower semicontinuous and the function $\|\cdot\|^2$ is lower semicontinuous. We now consider points where $c = 0$ and $w \neq x$. Let $\{(w_k, x_k, c_k)\}$ be a sequence that converges to some $(w, x, 0)$ with $w \neq x$. We can assume without loss of generality that $w_k \neq x_k$ for all k. Note that for some k, we have

$$h(w_k, x_k, c_k) = \begin{cases} \infty & \text{if } c_k = 0, \\ f(w_k) + (\frac{1}{2c_k})\|w_k - x_k\|^2 & \text{if } c_k > 0. \end{cases}$$

Taking the limit as $k \to \infty$, we have

$$\lim_{k \to \infty} h(w_k, x_k, c_k) = \infty \geq h(w, x, 0),$$

since $w \neq x$ by assumption. This shows that h is lower semicontinuous at points where $c = 0$ and $w \neq x$. We finally consider points where $c = 0$ and $w = x$. At these points, we have $h(w, x, c) = f(x)$. Let $\{(w_k, x_k, c_k)\}$ be a sequence that converges to some $(w, x, 0)$ with $w = x$. Considering all possibilities, we see that the limit inferior of the sequence $\{h(w_k, x_k, c_k)\}$ cannot be less than $f(x)$, thus showing that h is also lower semicontinuous at points where $c = 0$ and $w = x$.

Finally, we show that h satisfies the locally uniform level-boundedness property given in Exercise 2.10, i.e., for all (x^*, c^*) and for all $\alpha \in \Re$, there exists a neighborhood N of (x^*, c^*) such that the set $\big\{(w, x, c) \mid (x, c) \in N, h(w, x, c) \leq \alpha\big\}$ is bounded. Assume, to arrive at a contradiction, that there exists a sequence $\big\{(w_k, x_k, c_k)\big\}$ such that

$$h(w_k, x_k, c_k) \leq \alpha < \infty, \tag{2.9}$$

for some scalar α, with $(x_k, c_k) \to (x^*, c^*)$, and $\|w_k\| \to \infty$. Then, for sufficiently large k, we have $w_k \neq x_k$, which in view of Eq. (2.9) and the definition of the function h, implies that $c_k \in (0, c_0]$ and

$$f(w_k) + \frac{1}{2c_k}\|w_k - x_k\|^2 \leq \alpha,$$

for all sufficiently large k. In particular, since $c_k \leq c_0$, it follows from the preceding relation that

$$f(w_k) + \frac{1}{2c_0}\|w_k - x_k\|^2 \leq \alpha. \tag{2.10}$$

The choice of c_0 ensures, through the definition of c_f, the existence of some $c_1 > c_0$, some $\overline{x} \in \Re^n$, and some scalar β such that

$$f(w) \geq -\frac{1}{2c_1}\|w - \overline{x}\|^2 + \beta, \qquad \forall\, w.$$

Together with Eq. (2.10), this implies that

$$-\frac{1}{2c_1}\|w_k - \overline{x}\|^2 + \frac{1}{2c_0}\|w_k - x_k\|^2 \leq \alpha - \beta,$$

for all sufficiently large k. Dividing this relation by $\|w_k\|^2$ and taking the limit as $k \to \infty$, we get

$$-\frac{1}{2c_1} + \frac{1}{2c_0} \leq 0,$$

from which it follows that $c_1 \leq c_0$. This is a contradiction by our choice of c_1. Hence, the function $h(w, x, c)$ satisfies all the assumptions of Exercise 2.10.

By assumption, we have that $f(\overline{x}) < \infty$ for some $\overline{x} \in \Re^n$. Using the definition of $e_c f(x)$, this implies that

$$\begin{aligned} e_c f(x) &= \inf_w \left\{ f(w) + \frac{1}{2c}\|w - x\|^2 \right\} \\ &\leq f(\overline{x}) + \frac{1}{2c}\|\overline{x} - x\|^2 < \infty, \qquad \forall\, x \in \Re^n, \end{aligned}$$

where the first inequality is obtained by setting $w = \overline{x}$ in $f(w) + \frac{1}{2c}\|w - x\|^2$. Together with Exercise 2.10(a), this shows that for every $c \in (0, c_0)$ and all $x \in \Re^n$, the function $e_c f(x)$ is finite, and the set $P_c f(x)$ is nonempty and compact. Furthermore, it can be seen from the definition of $h(w, x, c)$, that for all $c \in (0, c_0)$, $h(w, x, c)$ is continuous in (x, c). Therefore, it follows from Exercise 2.10(d) that for all $c \in (0, c_0)$, $e_c f(x)$ is continuous in (x, c). In particular, since $e_c f(x)$ is a monotonically decreasing function of c, it follows that

$$e_c f(x) = p(x, c) \uparrow p(x, 0) = f(x), \quad \forall\, x \text{ as } c \downarrow 0.$$

This concludes the proof for part (a).

(b) Directly follows from Exercise 2.10(c).

2.14 (Envelopes and Proximal Mappings under Convexity [RoW98])

Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper convex function. For $c > 0$, consider the envelope function $e_c f$ and the proximal mapping $P_c f$ (cf. Exercise 2.13). Show the following:

(a) The supremum of the set of all $c > 0$ such that $e_c f(x) > -\infty$ for some $x \in \Re^n$, is ∞.

(b) The proximal mapping $P_c f$ is single-valued and is continuous in the sense that $P_c f(x) \to P_{c^*} f(x^*)$ whenever $(x, c) \to (x^*, c^*)$ with $c^* > 0$.

(c) The envelope function $e_c f$ is convex and smooth, and its gradient is given by

$$\nabla e_c f(x) = \frac{1}{c}\big(x - P_c f(x)\big).$$

Note: The envelope function $e_c f$ is smooth, regardless of whether f is nonsmooth.

Solution: We consider the function g_c defined by

$$g_c(x, w) = f(w) + \frac{1}{2c}\|w - x\|^2.$$

In view of the assumption that f is lower semicontinuous, it follows that $g_c(x, w)$ is lower semicontinuous. We also have that $g_c(x, w) > -\infty$ for all (x, w) and $g_c(x, w) < \infty$ for at least one vector (x, w). Moreover, since $f(x)$ is convex by assumption, $g_c(x, w)$ is convex in (x, w), even strictly convex in w.

Note that by definition, we have

$$e_c f(x) = \inf_w g_c(x, w),$$

$$P_c f(x) = \arg\min_w g_c(x, w).$$

(a) In order to show that c_f is ∞, it suffices to show that $e_c f(0) > -\infty$ for all $c > 0$. This will follow from Weierstrass' Theorem, once we show the boundedness of the level sets of $g_c(0, \cdot)$. Assume the contrary, i.e., there exists some $\alpha \in \Re$ and a sequence $\{x_k\}$ such that $\|x_k\| \to \infty$ and

$$g_c(0, x_k) = f(x_k) + \frac{1}{2c}\|x_k\|^2 \le \alpha, \qquad \forall\ k. \tag{2.11}$$

Assume without loss of generality that $\|x_k\| > 1$ for all k. We fix an x_0 with $f(x_0) < \infty$. We define

$$\tau_k = \frac{1}{\|x_k\|} \in (0, 1),$$

and

$$\overline{x}_k = (1 - \tau_k)x_0 + \tau_k x_k.$$

Since $\|x_k\| \to \infty$, it follows that $\tau_k \to 0$. Using Eq. (2.11) and the convexity of f, we obtain

$$\begin{aligned} f(\overline{x}_k) &\le (1 - \tau_k) f(x_0) + \tau_k f(x_k) \\ &\le (1 - \tau_k) f(x_0) + \tau_k \alpha - \frac{\tau_k}{2c}\|x_k\|^2. \end{aligned}$$

Taking the limit as $k \to \infty$ in the above equation, we see that $f(\overline{x}_k) \to -\infty$. It follows from the definitions of τ_k and $\overline{x}_k$ that

$$\begin{aligned}\|\overline{x}_k\| &\leq \|1-\tau_k\|\|x_0\| + \|\tau_k\|\|x_k\| \\ &\leq \|x_0\| + 1.\end{aligned}$$

Therefore, the sequence $\{\overline{x}_k\}$ is bounded. Since f is lower semicontinuous, Weierstrass' Theorem suggests that f is bounded from below on every bounded subset of $\Re^n$. Since the sequence $\{\overline{x}_k\}$ is bounded, this implies that the sequence $f(\overline{x}_k)$ is bounded from below, which contradicts the fact that $f(\overline{x}_k) \to \infty$. This proves that the level sets of the function $g_c(0, \cdot)$ are bounded. Therefore, using Weierstrass' Theorem, we have that the infimum in $e_c f(0) = \inf_w g_c(0, w)$ is attained, and $e_c f(0) > -\infty$ for every $c > 0$. This shows that the supremum c_f of all $c > 0$, such that $e_c f(x) > -\infty$ for some $x \in \Re^n$, is ∞.

(b) Since the value c_f is equal to ∞ by part (a), it follows that $e_c f$ and $P_c f$ have all the properties given in Exercise 2.13 for all $c > 0$: The set $P_c f(x)$ is nonempty and compact, and the function $e_c f(x)$ is finite for all x, and is continuous in (x, c). Consider a sequence $\{w_k\}$ with $w_k \in P_{c_k} f(x_k)$ for some sequences $x_k \to x^*$ and $c_k \to c^* > 0$. Then, it follows from Exercise 2.13(b) that the sequence $\{w_k\}$ is bounded and all its limit points belong to the set $P_{c^*} f(x^*)$. Since $g_c(x, w)$ is strictly convex in w, it follows from Prop. 2.1.2 that the proximal mapping $P_c f$ is single-valued. Hence, we have that $P_c f(x) \to P_{c^*} f(x^*)$ whenever $(x, c) \to (x^*, c^*)$ with $c^* > 0$.

(c) The envelope function $e_c f$ is convex by Exercise 2.15 [since $g_c(x, w)$ is convex in (x, w)], and continuous by Exercise 2.13. We now prove that it is differentiable. Consider any point $\overline{x}$, and let $\overline{w} = P_c f(\overline{x})$. We will show that $e_c f$ is differentiable at $\overline{x}$ with

$$\nabla e_c f(\overline{x}) = \frac{(\overline{x} - \overline{w})}{c}.$$

Equivalently, we will show that the function h given by

$$h(u) = e_c f(\overline{x} + u) - e_c f(\overline{x}) - \frac{(\overline{x} - \overline{w})'}{c} u \tag{2.12}$$

is differentiable at 0 with $\nabla h(0) = 0$. Since $\overline{w} = P_c f(\overline{x})$, we have

$$e_c f(\overline{x}) = f(\overline{w}) + \frac{1}{2c}\|\overline{w} - \overline{x}\|^2,$$

whereas

$$e_c f(\overline{x} + u) \leq f(\overline{w}) + \frac{1}{2c}\|\overline{w} - (\overline{x} + u)\|^2, \qquad \forall\, u,$$

so that

$$h(u) \leq \frac{1}{2c}\|\overline{w} - (\overline{x} + u)\|^2 - \frac{1}{2c}\|\overline{w} - \overline{x}\|^2 - \frac{1}{c}(\overline{x} - \overline{w})' u = \frac{1}{2c}\|u\|^2, \ \forall\, u. \tag{2.13}$$

Since $e_c f$ is convex, it follows from Eq. (2.12) that h is convex, and therefore,

$$0 = h(0) = h\left(\frac{1}{2}u + \frac{1}{2}(-u)\right) \leq \frac{1}{2}h(u) + \frac{1}{2}h(-u),$$

which implies that $h(u) \geq -h(-u)$. From Eq. (2.13), we obtain

$$-h(-u) \geq -\frac{1}{2c}\|-u\|^2 = -\frac{1}{2c}\|u\|^2, \qquad \forall\ u,$$

which together with the preceding relation yields

$$h(u) \geq -\frac{1}{2c}\|u\|^2, \qquad \forall\ u.$$

Thus, we have

$$|h(u)| \leq \frac{1}{2c}\|u\|^2, \qquad \forall\ u,$$

which implies that h is differentiable at 0 with $\nabla h(0) = 0$. From the formula for $\nabla e_c f(\cdot)$ and the continuity of $P_c f(\cdot)$, it also follows that e_c is continuously differentiable.

2.15

(a) Let C_1 be a convex set with nonempty interior and C_2 be a nonempty convex set that does not intersect the interior of C_1. Show that there exists a hyperplane such that one of the associated closed halfspaces contains C_2, and does not intersect the interior of C_1.

(b) Show by an example that we cannot replace interior with relative interior in the statement of part (a).

Solution: (a) In view of the assumption that $\text{int}(C_1)$ and C_2 are disjoint and convex [cf Prop. 1.2.1(d)], it follows from the Separating Hyperplane Theorem that there exists a vector $a \neq 0$ such that

$$a'x_1 \leq a'x_2, \qquad \forall\ x_1 \in \text{int}(C_1), \quad \forall\ x_2 \in C_2.$$

Let $b = \inf_{x_2 \in C_2} a'x_2$. Then, from the preceding relation, we have

$$a'x \leq b, \qquad \forall\ x \in \text{int}(C_1). \tag{2.14}$$

We claim that the closed halfspace $\{x \mid a'x \geq b\}$, which contains C_2, does not intersect $\text{int}(C_1)$.

Assume to arrive at a contradiction that there exists some $\overline{x}_1 \in \text{int}(C_1)$ such that $a'\overline{x}_1 \geq b$. Since $\overline{x}_1 \in \text{int}(C_1)$, we have that there exists some $\epsilon > 0$ such that $\overline{x}_1 + \epsilon a \in \text{int}(C_1)$, and

$$a'(\overline{x}_1 + \epsilon a) \geq b + \epsilon\|a\|^2 > b.$$

This contradicts Eq. (2.14). Hence, we have

$$\text{int}(C_1) \subset \{x \mid a'x < b\}.$$

(b) Consider the sets

$$C_1 = \{(x_1, x_2) \mid x_1 = 0\},$$
$$C_2 = \{(x_1, x_2) \mid x_1 > 0,\ x_2 x_1 \geq 1\}.$$

These two sets are convex and C_2 is disjoint from $\mathrm{ri}(C_1)$, which is equal to C_1. The only separating hyperplane is the x_2 axis, which corresponds to having $a = (0, 1)$, as defined in part (a). For this example, there does not exist a closed halfspace that contains C_2 but is disjoint from $\mathrm{ri}(C_1)$.

2.16

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

Solution: If there exists a hyperplane H with the properties stated, the condition $M \cap \mathrm{ri}(C) = \emptyset$ clearly holds. Conversely, if $M \cap \mathrm{ri}(C) = \emptyset$, then M and C can be properly separated by Prop. 2.4.5. This hyperplane can be chosen to contain M since M is affine. If this hyperplane contains a point in $\mathrm{ri}(C)$, then it must contain all of C by Prop. 1.4.2. This contradicts the proper separation property, thus showing that $\mathrm{ri}(C)$ is contained in one of the open halfspaces.

2.17 (Strong Separation)

Let C_1 and C_2 be nonempty convex subsets of $\Re^n$, and let B denote the unit ball in $\Re^n$, $B = \{x \mid \|x\| \leq 1\}$. 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 $a \in \Re^n$ such that $\inf_{x \in C_1} a'x > \sup_{x \in C_2} a'x$.

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

(b) If C_1 and C_2 are disjoint, any one of the five conditions for strict separation, given in Prop. 2.4.3, 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 $a \in \Re^n$, $b \in \Re$, 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 $a \in \Re^n$ 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 $a \in \Re^n$ such that

$$\inf_{x \in C_1} a'x > \sup_{x \in C_2} a'x, \tag{2.15}$$

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 $\|y_1\| \leq \epsilon$, $\|y_2\| \leq \epsilon$,

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

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. 2.4.3 imply condition (1) of that proposition (see the discussion in the proof of Prop. 2.4.3), which states that the set $C_1 - C_2$ is closed, i.e.,

$$\mathrm{cl}(C_1 - C_2) = C_1 - C_2.$$

Hence, we have $0 \notin \mathrm{cl}(C_1 - C_2)$, 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.

2.18

Let C_1 and C_2 be nonempty convex subsets of $\Re^n$ such that C_2 is a cone.

(a) Suppose that there exists a hyperplane that separates C_1 and C_2 properly. Show that there exists a hyperplane which separates C_1 and C_2 properly and passes through the origin.

(b) Suppose that there exists a hyperplane that separates C_1 and C_2 strictly. Show that there exists a hyperplane that passes through the origin such that one of the associated closed halfspaces contains the cone C_2 and does not intersect C_1.

Solution: (a) If C_1 and C_2 can be separated properly, we have from the Proper Separation Theorem that there exists a vector $a \neq 0$ such that

$$\inf_{x \in C_1} a'x \geq \sup_{x \in C_2} a'x, \tag{2.16}$$

$$\sup_{x \in C_1} a'x > \inf_{x \in C_2} a'x. \tag{2.17}$$

Let

$$b = \sup_{x \in C_2} a'x. \tag{2.18}$$

and consider the hyperplane

$$H = \{x \mid a'x = b\}.$$

Since C_2 is a cone, we have

$$\lambda a'x = a'(\lambda x) \leq b < \infty, \qquad \forall\ x \in C_2,\ \forall\ \lambda > 0.$$

This relation implies that $a'x \leq 0$, for all $x \in C_2$, since otherwise it is possible to choose λ large enough and violate the above inequality for some $x \in C_2$. Hence, it follows from Eq. (2.18) that $b \leq 0$. Also, by letting $\lambda \to 0$ in the preceding relation, we see that $b \geq 0$. Therefore, we have that $b = 0$ and the hyperplane H contains the origin.

(b) If C_1 and C_2 can be separated strictly, we have by definition that there exists a vector $a \neq 0$ and a scalar β such that

$$a'x_2 < \beta < a'x_1, \qquad \forall\ x_1 \in C_1, \quad \forall\ x_2 \in C_2. \tag{2.19}$$

We choose b to be

$$b = \sup_{x \in C_2} a'x, \tag{2.20}$$

and consider the closed halfspace

$$K = \{x \mid a'x \leq b\},$$

which contains C_2. By Eq. (2.19), we have

$$b \leq \beta < a'x, \qquad \forall\ x \in C_1,$$

so the closed halfspace K does not intersect C_1.

Since C_2 is a cone, an argument similar to the one in part (a) shows that $b = 0$, and hence the hyperplane associated with the closed halfspace K passes through the origin, and has the desired properties.

2.19 (Separation Properties of Cones)

Define a *homogeneous halfspace* to be a closed halfspace associated with a hyperplane that passes through the origin. Show that:

(a) A nonempty closed convex cone is the intersection of the homogeneous halfspaces that contain it.

(b) The closure of the convex cone generated by a nonempty set X is the intersection of all the homogeneous halfspaces containing X.

Solution: (a) C is contained in the intersection of the homogeneous closed halfspaces that contain C, so we focus on proving the reverse inclusion. Let $x \notin C$. Since C is closed and convex by assumption, by using the Strict Separation Theorem, we see that the sets C and $\{x\}$ can be separated strictly. From Exercise 2.18(c), this implies that there exists a hyperplane that passes through the origin such that one of the associated closed halfspaces contains C, but is disjoint from x. Hence, if $x \notin C$, then x cannot belong to the intersection of the homogeneous closed halfspaces containing C, proving that C contains that intersection.

(b) A homogeneous halfspace is in particular a closed convex cone containing the origin, and such a cone includes X if and only if it includes $\mathrm{cl}\big(\mathrm{cone}(X)\big)$. Hence, the intersection of all closed homogeneous halfspaces containing X and the intersection of all closed homogeneous halfspaces containing $\mathrm{cl}\big(\mathrm{cone}(X)\big)$ coincide. From what has been proved in part(a), the latter intersection is equal to $\mathrm{cl}\big(\mathrm{cone}(X)\big)$.

2.20 (Convex System Alternatives)

Let $g_j : \Re^n \mapsto (-\infty, \infty]$, $j = 1, \ldots, r$, be closed proper convex functions, and let X be a nonempty closed convex set. Assume that one of the following four conditions holds:

(1) $R_X \cap R_{g_1} \cap \cdots \cap R_{g_r} = L_X \cap L_{g_1} \cap \cdots \cap L_{g_r}$.

(2) $R_X \cap R_{g_1} \cap \cdots \cap R_{g_r} \subset L_{g_1} \cap \cdots \cap L_{g_r}$ and X is specified by linear inequality constraints.

(3) Each g_j is a convex quadratic function and X is specified by convex quadratic inequality constraints.

(4) Each g_j is a convex bidirectionally flat function (see Exercise 2.7) and X is specified by convex bidirectionally flat functions.

Show that:

(a) If there is no vector $x \in X$ such that

$$g_1(x) \leq 0, \ldots, g_r(x) \leq 0,$$

then there exist a positive scalar ϵ, and a vector $\mu \in \Re^r$ with $\mu \geq 0$, such that

$$\mu_1 g_1(x) + \cdots + \mu_r g_r(x) \geq \epsilon, \qquad \forall\, x \in X.$$

Hint: Show that under any one of the conditions (1)-(4), the set

$$C = \big\{u \mid \text{there exists an } x \in X \text{ such that } g_j(x) \leq u_j,\ j = 1, \ldots, r\big\}$$

is closed, by viewing it as the projection of the set

$$\big\{(x, u) \mid x \in X,\ g_j(x) \leq u_j,\ j = 1, \ldots, r\big\}$$

on the space of u. Furthermore, the origin does not belong to C, so it can be strictly separated from C by a hyperplane. The normal of this hyperplane provides the desired vector μ.

(b) If for every $\epsilon > 0$, there exists a vector $x \in X$ such that

$$g_1(x) < \epsilon, \ldots, g_r(x) < \epsilon,$$

then there exists a vector $x \in X$ such that

$$g_1(x) \leq 0, \ldots, g_r(x) \leq 0.$$

Hint: Argue by contradiction and use part (a).

Solution: (a) Consider the set

$$C = \big\{u \mid \text{there exists an } x \in X \text{ such that } g_j(x) \leq u_j,\ j = 1, \ldots, r\big\},$$

which may be viewed as the projection of the set

$$M = \Big\{(x, u) \mid x \in X,\ g_j(x) \leq u_j,\ j = 1, \ldots, r\Big\}$$

on the space of u. Let us denote this linear transformation by A. It can be seen that

$$R_M \cap N(A) = \Big\{(y, 0) \mid y \in R_X \cap R_{g_1} \cdots \cap R_{g_r}\Big\},$$

where R_M denotes the recession cone of set M. Similarly, we have

$$L_M \cap N(A) = \Big\{(y, 0) \mid y \in L_X \cap L_{g_1} \cdots \cap L_{g_r}\Big\},$$

where L_M denotes the lineality space of set M. Under conditions (1), (2), and (3), it follows from Prop. 1.5.8 that the set $AM = C$ is closed. Similarly, under condition (4), it follows from Exercise 2.7(b) that the set $AM = C$ is closed.

By assumption, there is no vector $x \in X$ such that

$$g_1(x) \leq 0, \ldots, g_r(x) \leq 0.$$

This implies that the origin does not belong to C. Therefore, by the Strict Separation Theorem, it follows that there exists a hyperplane that strictly separates the origin and the set C, i.e., there exists a vector μ such that

$$0 < \epsilon \leq \mu' u, \qquad \forall\ u \in C. \tag{2.21}$$

This equation implies that $\mu \geq 0$ since for each $u \in C$, we have that $(u_1, \ldots, u_j + \gamma, \ldots, u_r) \in C$ for all j and $\gamma > 0$. Since $\big(g_1(x), \ldots, g_r(x)\big) \in C$ for all $x \in X$, Eq. (2.21) yields

$$\mu_1 g_1(x) + \cdots + \mu_r g_r(x) \geq \epsilon, \qquad \forall\ x \in X. \tag{2.22}$$

(b) Assume that there is no vector $x \in X$ such that

$$g_1(x) \leq 0, \ldots, g_r(x) \leq 0.$$

This implies by part (a) that there exists a positive scalar ϵ, and a vector $\mu \in \Re^r$ with $\mu \geq 0$, such that

$$\mu_1 g_1(x) + \cdots + \mu_r g_r(x) \geq \epsilon, \qquad \forall\ x \in X.$$

Let x be an arbitrary vector in X and let $j(x)$ be the smallest index that satisfies $j(x) = \arg\max_{j=1,\ldots,r} g_j(x)$. Then Eq. (2.22) implies that for all $x \in X$

$$\epsilon \leq \sum_{j=1}^{r} \mu_j g_j(x) \leq \sum_{j=1}^{r} \mu_j g_{j(x)}(x) = g_{j(x)}(x) \sum_{j=1}^{r} \mu_j.$$

Hence, for all $x \in X$, there exists some $j(x)$ such that

$$g_{j(x)}(x) \geq \frac{\epsilon}{\sum_{j=1}^{r} \mu_j} > 0.$$

This contradicts the statement that for every $\epsilon > 0$, there exists a vector $x \in X$ such that

$$g_1(x) < \epsilon, \ldots, g_r(x) < \epsilon,$$

and concludes the proof.

2.21

Let C be a nonempty closed convex subset of $\Re^{n+1}$ that contains no vertical lines. Show that C is equal to the intersection of the closed halfspaces that contain it and correspond to nonvertical hyperplanes.

Solution: C is contained in the intersection of the closed halfspaces that contain C and correspond to nonvertical hyperplanes, so we focus on proving the reverse inclusion. Let $x \notin C$. Since by assumption C does not contain any vertical lines, we can apply Prop. 2.5.1, and we see that there exists a closed halfspace that correspond to a nonvertical hyperplane, containing C but not containing x. Hence, if $x \notin C$, then x cannot belong to the intersection of the closed halfspaces containing C and corresponding to nonvertical hyperplanes, proving that C contains that intersection.

2.22 (Min Common/Max Crossing Duality)

Consider the min common/max crossing framework, assuming that $w^* < \infty$.

(a) Assume that M is compact. Show that q^* is equal to the optimal value of the min common point problem corresponding to $\text{conv}(M)$.

(b) Assume that M is closed and convex, and does not contain a halfline of the form $\{(x, w + \alpha) \mid \alpha \leq 0\}$. Show that $\overline{M}$ is the epigraph of the function given by

$$f(x) = \inf\{w \mid (x, w) \in M\}, \qquad x \in \Re^n,$$

and that f is closed proper and convex.

(c) Assume that w^* is finite, and that $\overline{M}$ is convex and closed. Show that $q^* = w^*$.

Solution: (a) Let us denote the optimal value of the min common point problem and the max crossing point problem corresponding to $\text{conv}(M)$ by $w^*_{\text{conv}(M)}$ and $q^*_{\text{conv}(M)}$, respectively. In view of the assumption that M is compact, it follows from Prop. 1.3.2 that the set $\text{conv}(M)$ is compact. Therefore, by Weierstrass' Theorem, $w^*_{\text{conv}(M)}$, defined by

$$w^*_{\text{conv}(M)} = \inf_{(0,w)\in \text{conv}(M)} w$$

is finite. It can also be seen that the set

$$\overline{\text{conv}(M)} = \{(u, w) \mid \text{there exists } \overline{w} \text{ with } \overline{w} \leq w \text{ and } (u, \overline{w}) \in \text{conv}(M)\}$$

is convex. Indeed, we consider vectors $(u, w) \in \overline{\text{conv}(M)}$ and $(\tilde{u}, \tilde{w}) \in \overline{\text{conv}(M)}$, and we show that their convex combinations lie in $\overline{\text{conv}(M)}$. The definition of $\overline{\text{conv}(M)}$ implies that there exists some w_M and $\tilde{w}_M$ such that

$$w_M \leq w, \qquad (u, w_M) \in \text{conv}(M),$$

$$\tilde{w}_M \le \tilde{w}, \qquad (\tilde{u}, \tilde{w}_M) \in \text{conv}(M).$$

For any $\alpha \in [0, 1]$, we multiply these relations with α and $(1 - \alpha)$, respectively, and add. We obtain

$$\alpha w_M + (1 - \alpha)\tilde{w}_M \le \alpha w + (1 - \alpha)\tilde{w}.$$

In view of the convexity of $\text{conv}(M)$, we have $\alpha(u, w_M) + (1 - \alpha)(\tilde{u}, \tilde{w}_M) \in \text{conv}(M)$, so these equations imply that the convex combination of (u, w) and $(\tilde{u}, \tilde{w})$ belongs to $\overline{\text{conv}(M)}$. This proves the convexity of $\overline{\text{conv}(M)}$.

Using the compactness of $\text{conv}(M)$, it can be shown that for every sequence $\big\{(u_k, w_k)\big\} \subset \text{conv}(M)$ with $u_k \to 0$, there holds $w^*_{\text{conv}(M)} \le \liminf_{k\to\infty} w_k$. Let $\big\{(u_k, w_k)\big\} \subset \text{conv}(M)$ be a sequence with $u_k \to 0$. Since $\text{conv}(M)$ is compact, the sequence $\big\{(u_k, w_k)\big\}$ has a subsequence that converges to some $(0, \overline{w}) \in \text{conv}(M)$. Assume without loss of generality that $\big\{(u_k, w_k)\big\}$ converges to $(0, \overline{w})$. Since $(0, \overline{w}) \in \text{conv}(M)$, we get

$$w^*_{\text{conv}(M)} \le \overline{w} = \liminf_{k\to\infty} w_k.$$

Therefore, by Min Common/Max Crossing Theorem I, we have

$$w^*_{\text{conv}(M)} = q^*_{\text{conv}(M)}. \tag{2.23}$$

Let q^* be the optimal value of the max crossing point problem corresponding to M, i.e.,

$$q^* = \sup_{\mu\in\Re^n} q(\mu),$$

where for all $\mu \in \Re^n$

$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu' u\}.$$

We will show that $q^* = w^*_{\text{conv}(M)}$. For every $\mu \in \Re^n$, $q(\mu)$ can be expressed as $q(\mu) = \inf_{x\in M} c'x$, where $c = (\mu, 1)$ and $x = (u, w)$. From Exercise 2.23, it follows that minimization of a linear function over a set is equivalent to minimization over its convex hull. In particular, we have

$$q(\mu) = \inf_{x\in X} c'x = \inf_{x\in\text{conv}(X)} c'x,$$

from which using Eq. (2.23), we get

$$q^* = q^*_{\text{conv}(M)} = w^*_{\text{conv}(M)},$$

proving the desired claim.

(b) The function f is convex by the result of Exercise 2.23. Furthermore, for all $x \in \text{dom}(f)$, the infimum in the definition of $f(x)$ is attained. The reason is that, for $x \in \text{dom}(f)$, the set $\big\{w \mid (x, w) \in M\big\}$ is closed and bounded below, since M is closed and does not contain a halfline of the form $\big\{(x, w + \alpha) \mid \alpha \le 0\big\}$. Thus, we have $f(x) > -\infty$ for all $x \in \text{dom}(f)$, while $\text{dom}(f)$ is nonempty, since

M is nonempty in the min common/max crossing framework. It follows that f is proper. Furthermore, by its definition, $\overline{M}$ is the epigraph of f. Finally, to show that f is closed, we argue by contradiction. If f is not closed, there exists a vector x and a sequence $\{x_k\}$ that converges to x and is such that

$$f(x) > \lim_{k\to\infty} f(x_k).$$

We claim that $\lim_{k\to\infty} f(x_k)$ is finite, i.e., that $\lim_{k\to\infty} f(x_k) > -\infty$. Indeed, by Prop. 2.5.1, the epigraph of f is contained in the upper halfspace of a nonvertical hyperplane of $\Re^{n+1}$. Since $\{x_k\}$ converges to x, the limit of $\big\{f(x_k)\big\}$ cannot be equal to $-\infty$. Thus the sequence $\big(x_k, f(x_k)\big)$, which belongs to M, converges to $\big(x, \lim_{k\to\infty} f(x_k)\big)$ Therefore, since M is closed, $\big(x, \lim_{k\to\infty} f(x_k)\big) \in M$. By the definition of f, this implies that $f(x) \leq \lim_{k\to\infty} f(x_k)$, contradicting our earlier hypothesis.

(c) We prove this result by showing that all the assumptions of Min Common/Max Crossing Theorem I are satisfied. By assumption, $w^* < \infty$ and the set $\overline{M}$ is convex. Therefore, we only need to show that for every sequence $\{u_k, w_k\} \subset M$ with $u_k \to 0$, there holds $w^* \leq \liminf_{k\to\infty} w_k$.

Consider a sequence $\{u_k, w_k\} \subset M$ with $u_k \to 0$. If $\liminf_{k\to\infty} w_k = \infty$, then we are done, so assume that $\liminf_{k\to\infty} w_k = \tilde{w}$ for some scalar $\tilde{w}$. Since $M \subset \overline{M}$ and $\overline{M}$ is closed by assumption, it follows that $(0, \tilde{w}) \in \overline{M}$. By the definition of the set $\overline{M}$, this implies that there exists some $\overline{w}$ with $\overline{w} \leq \tilde{w}$ and $(0, \overline{w}) \in M$. Hence we have

$$w^* = \inf_{(0,w)\in M} w \leq \overline{w} \leq \tilde{w} = \liminf_{k\to\infty} w_k,$$

proving the desired result, and thus showing that $q^* = w^*$.

2.23 (An Example of Lagrangian Duality)

Consider the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in X, \qquad e_i'x = d_i, \quad i = 1, \ldots, m, \end{aligned}$$

where $f : \Re^n \mapsto \Re$ is a convex function, X is a convex set, and e_i and d_i are given vectors and scalars, respectively. Consider the min common/max crossing framework where M is the subset of $\Re^{m+1}$ given by

$$M = \Big\{ \big(e_1'x - d_1, \ldots, e_m'x - d_m, f(x)\big) \mid x \in X \Big\}.$$

(a) Derive the corresponding max crossing problem.

(b) Show that the corresponding set $\overline{M}$ is convex.

(c) Show that if $w^* < \infty$ and X is compact, then $q^* = w^*$.

(d) Show that if $w^* < \infty$ and there exists a vector $\overline{x} \in \text{ri}(X)$ such that $e_i'\overline{x} = d_i$ for all $i = 1, \ldots, m$, then $q^* = w^*$ and the max crossing problem has an optimal solution.

Solution: (a) The corresponding max crossing problem is given by

$$q^* = \sup_{\mu \in \Re^m} q(\mu),$$

where $q(\mu)$ is given by

$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu' u\} = \inf_{x\in X} \left\{ f(x) + \sum_{i=1}^{m} \mu_i (e_i' x - d_i) \right\}.$$

(b) Consider the set

$$\overline{M} = \Big\{ \big(u_1, \ldots, u_m, w\big) \mid \exists\, x \in X \text{ such that } e_i'x - d_i = u_i,\ \forall\, i,\ f(x) \le w \Big\}.$$

We show that $\overline{M}$ is convex. To this end, we consider vectors $(u, w) \in \overline{M}$ and $(\tilde{u}, \tilde{w}) \in \overline{M}$, and we show that their convex combinations lie in $\overline{M}$. The definition of $\overline{M}$ implies that for some $x \in X$ and $\tilde{x} \in X$, we have

$$f(x) \le w, \qquad e_i'x - d_i = u_i, \quad i = 1, \ldots, m,$$
$$f(\tilde{x}) \le \tilde{w}, \qquad e_i'\tilde{x} - d_i = \tilde{u}_i, \quad i = 1, \ldots, m.$$

For any $\alpha \in [0, 1]$, we multiply these relations with α and 1-α, respectively, and add. By using the convexity of f, we obtain

$$f\big(\alpha x + (1-\alpha)\tilde{x}\big) \le \alpha f(x) + (1-\alpha) f(\tilde{x}) \le \alpha w + (1-\alpha)\tilde{w},$$

$$e_i'\big(\alpha x + (1-\alpha)\tilde{x}\big) - d_i = \alpha u_i + (1-\alpha)\tilde{u}_i, \quad i = 1, \ldots, m.$$

In view of the convexity of X, we have $\alpha x + (1-\alpha)\tilde{x} \in X$, so these equations imply that the convex combination of (u, w) and $(\tilde{u}, \tilde{w})$ belongs to $\overline{M}$, thus proving that $\overline{M}$ is convex.

(c) We prove this result by showing that all the assumptions of Min Common/Max Crossing Theorem I are satisfied. By assumption, w^* is finite. It follows from part (b) that the set $\overline{M}$ is convex. Therefore, we only need to show that for every sequence $\big\{(u_k, w_k)\big\} \subset M$ with $u_k \to 0$, there holds $w^* \le \liminf_{k\to\infty} w_k$.

Consider a sequence $\big\{(u_k, w_k)\big\} \subset M$ with $u_k \to 0$. Since X is compact and f is convex by assumption (which implies that f is continuous by Prop. 1.4.6), it follows from Prop. 1.1.9(c) that set M is compact. Hence, the sequence $\big\{(u_k, w_k)\big\}$ has a subsequence that converges to some $(0, \overline{w}) \in M$. Assume without loss of generality that $\big\{(u_k, w_k)\big\}$ converges to $(0, \overline{w})$. Since $(0, \overline{w}) \in M$, we get

$$w^* = \inf_{(0,w)\in M} w \le \overline{w} = \liminf_{k\to\infty} w_k,$$

proving the desired result, and thus showing that $q^* = w^*$.

(d) We prove this result by showing that all the assumptions of Min Common/Max Crossing Theorem II are satisfied. By assumption, w^* is finite. It follows from part (b) that the set $\overline{M}$ is convex. Therefore, we only need to show that the set

$$D = \big\{(e_1'x - d_1, \ldots, e_m'x - d_m) \mid x \in X\big\}$$

contains the origin in its relative interior. The set D can equivalently be written as

$$D = E \cdot X - d,$$

where E is a matrix, whose rows are the vectors e_i', $i = 1, \ldots, m$, and d is a vector with entries equal to d_i, $i = 1, \ldots, m$. By Prop. 1.4.4 and Prop. 1.4.5(b), it follows that

$$\mathrm{ri}(D) = E \cdot \mathrm{ri}(X) - d.$$

Hence the assumption that there exists a vector $\overline{x} \in \mathrm{ri}(X)$ such that $E\overline{x} - d = 0$ implies that 0 belongs to the relative interior of D, thus showing that $q^* = w^*$ and that the max crossing problem has an optimal solution.

2.24 (Saddle Points in Two Dimensions)

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

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

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 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

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

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

$$x^* = \hat{x}(z^*), \qquad z^* = \hat{z}(x^*). \tag{2.24}$$

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

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

or equivalently, if $\overline{x} = \hat{x}(\overline{z})$ and $\overline{z} = \hat{z}(\overline{x})$. Therefore, from Eq. (2.24), 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.

2.25 (Saddle Points of Quadratic Functions)

Consider a quadratic function $\phi : X \times Z \mapsto \Re$ of the form

$$\phi(x, z) = x'Qx + x'Dz - z'Rz,$$

where Q and R are symmetric positive semidefinite $n \times n$ and $m \times m$ matrices, respectively, D is some $n \times m$ matrix, and X and Z are subsets of $\Re^n$ and $\Re^m$, respectively. Derive conditions under which ϕ has at least one saddle point.

Solution: Let X and Z be closed and convex sets. Then, for each $z \in Z$, the function $t_z : \Re^n \mapsto (-\infty, \infty]$ defined by

$$t_z(x) = \begin{cases} \phi(x, z) & \text{if } x \in X, \\ \infty & \text{otherwise,} \end{cases}$$

is closed and convex in view of the assumption that Q is a positive semidefinite symmetric matrix. Similarly, for each $x \in X$, the function $r_x : \Re^m \mapsto (-\infty, \infty]$ defined by

$$r_x(z) = \begin{cases} -\phi(x, z) & \text{if } z \in Z, \\ \infty & \text{otherwise,} \end{cases}$$

is closed and convex in view of the assumption that R is a positive semidefinite symmetric matrix. Hence, Assumption 2.6.1 is satisfied. Let also Assumptions 2.6.2 and 2.6.3 hold, i.e,

$$\inf_{x \in X} \sup_{z \in Z} \phi(x, z) < \infty,$$

and

$$-\infty < \sup_{z \in Z} \inf_{x \in X} \phi(x, z).$$

By the positive semidefiniteness of Q, it can be seen that, for each $z \in Z$, the recession cone of the function t_z is given by

$$R_{t_z} = R_X \cap N(Q) \cap \{y \mid y' Dz \leq 0\},$$

where R_X is the recession cone of the convex set X and $N(Q)$ is the null space of the matrix Q. Similarly, for each $z \in Z$, the constancy space of the function t_z is given by

$$L_{t_z} = L_X \cap N(Q) \cap \{y \mid y' Dz = 0\},$$

where L_X is the lineality space of the set X. By the positive semidefiniteness of R, for each $x \in X$, it can be seen that the recession cone of the function r_x is given by

$$R_{r_x} = R_Z \cap N(R) \cap \{y \mid x' Dy \geq 0\},$$

where R_Z is the recession cone of the convex set Z and $N(R)$ is the null space of the matrix R. Similarly, for each $x \in X$, the constancy space of the function r_x is given by

$$L_{r_x} = L_Z \cap N(R) \cap \{y \mid x' Dy = 0\},$$

where L_Z is the lineality space of the set Z.

If

$$\bigcap_{z \in Z} R_{t_z} = \{0\}, \quad \text{and} \quad \bigcap_{x \in X} R_{r_x} = \{0\}, \tag{2.25}$$

then it follows from the Saddle Point Theorem part (a), that the set of saddle points of ϕ is nonempty and compact. [In particular, the condition given in Eq. (2.25) holds when Q and R are positive definite matrices, or if X and Z are compact.]

Similarly, if

$$\bigcap_{z \in Z} R_{t_z} = \bigcap_{z \in Z} L_{t_z}, \quad \text{and} \quad \bigcap_{x \in X} R_{r_x} = \bigcap_{x \in X} L_{r_x},$$

then it follows from the Saddle Point Theorem part (b), that the set of saddle points of ϕ is nonempty.