Convex Optimization Theory

Chapter 3

Exercises and Solutions: Extended Version

Dimitri P. Bertsekas

Massachusetts Institute of Technology

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

CHAPTER 3: EXERCISES AND SOLUTIONS†

SECTION 3.1: Constrained Optimization

3.1 (Local Minima Along Lines)

(a) Consider a vector x^* such that a given function $f : \Re^n \mapsto \Re$ 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) Consider the nonconvex 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 $p < m < q$, so x^* is not a local minimum of f even though it is a local minimum along every line passing through x^*.

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

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

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

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^*.

---

† This set of exercises will be periodically updated as new exercises are added. Many of the exercises and solutions given here were developed as part of my earlier convex optimization book [BNO03] (coauthored with Angelia Nedić and Asuman Ozdaglar), and are posted on the internet of that book's web site. The contribution of my coauthors in the development of these exercises and their solutions is gratefully acknowledged. Since some of the exercises and/or their solutions have been modified and also new exercises have been added, all errors are my sole responsibility.

(b) We first show that the function $g : \Re \mapsto \Re$ defined by $g(\alpha) = f(x^* + \alpha d)$ has a local minimum 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,

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

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

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

3.2 (Equality-Constrained Quadratic Programming)

(a) Consider the quadratic program

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}\|x\|^2 + c'x \\ &\text{subject to} \quad Ax = 0, \end{aligned} \tag{3.1}$$

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

$$x^* = -\big(I - A'(AA')^{-1}A\big)c \tag{3.2}$$

is the unique solution.

(b) Consider the more general quadratic program

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}(x - \overline{x})'Q(x - \overline{x}) + c'(x - \overline{x}) \\ &\text{subject to} \quad Ax = b, \end{aligned} \tag{3.3}$$

where c and A are as before, Q is a symmetric positive definite matrix, $b \in \Re^m$, and $\overline{x}$ is a vector in $\Re^n$, which is feasible, i.e., satisfies $A\overline{x} = b$.

Use the transformation $y = Q^{1/2}(x - \overline{x})$ to write this problem in the form of part (a) and show that the optimal solution is

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

where λ is given by

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

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

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

and show that the optimal solution is

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

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

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

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

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

It can be seen that the vector

$$x^* = -\big(I - A'(AA')^{-1}A\big)c \tag{3.4}$$

satisfies this condition and is thus the unique solution of the quadratic programming problem (3.1). (The matrix AA' is invertible because A has rank m.)

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

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

Using Eq. (3.4) we see that the solution of this problem is

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

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

$$x^* = \overline{x} - Q^{-1}(c - A'\lambda), \tag{3.5}$$

where λ is given by

$$\lambda = \left(AQ^{-1}A'\right)^{-1} AQ^{-1}c. \tag{3.6}$$

The quadratic program (3.3) contains as a special case the program

$$\begin{aligned} &\text{minimize} \quad \tfrac{1}{2}x'Qx + c'x \\ &\text{subject to} \quad Ax = b. \end{aligned} \tag{3.7}$$

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

$$\overline{x} = Q^{-1}A'(AQ^{-1}A')^{-1}b. \tag{3.8}$$

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

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

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

$$\tfrac{1}{2}(x-\overline{x})'Q(x-\overline{x}) + c'(x-\overline{x}) = \tfrac{1}{2}x'Qx + c'x + \left(\tfrac{1}{2}\overline{x}'Q\overline{x} - c'\overline{x} - b'(AQ^{-1}A')^{-1}b\right).$$

The last term in parentheses on the right-hand side above is constant, thus establishing that the programs (3.3) and (3.7) have the same optimal solution when $\overline{x}$ is given by Eq. (3.8). By combining Eqs. (3.5) and (3.8), we obtain the optimal solution of program (3.7):

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

where λ is given by Eq. (3.6).

3.3 (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 $X \cap \text{dom}(f) \neq \emptyset$. Assume that f and X have no common nonzero direction of recession. Let X^* be the set of minima of f over X (which is nonempty and compact by Prop. 3.2.2), 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 $f(x) \le f^* + \delta$ satisfies $\min_{x^* \in X^*} \|x - x^*\| \le \epsilon$.

(b) If f is real-valued, for every $\delta > 0$ there exists an $\epsilon > 0$ such that every vector $x \in X$ with $\min_{x^* \in X^*} \|x - x^*\| \le \epsilon$ satisfies $f(x) \le f^* + \delta$.

(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

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

It follows that, for all k, x_k belongs to the set $\big\{x \in X \mid f(x) \le f^* + \delta_0\big\}$, 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

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

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

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

(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

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

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., $\overline{x} \in X^*$.

SECTION 3.2: Existence of Optimal Solutions

3.4 (Minimization of Quasiconvex Functions)

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

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

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

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

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

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

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

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

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

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

$$f\big(\alpha\overline{x} + (1-\alpha)x^*\big) \leq \gamma^* = f(x^*), \qquad \forall\ \alpha \in [0,1]. \tag{3.9}$$

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

$$f\big(\alpha_k\overline{x} + (1-\alpha_k)x^*\big) < f(x^*), \qquad \text{for some } \alpha_k \in (0,1/k]; \tag{3.10}$$

otherwise, in view of Eq. (3.9), 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 (3.10) 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) and (2), we will show that the set of minima of f over X, which is given by

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

is nonempty.

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

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

and

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

while by assumption $R_X \cap R_f = L_X \cap L_f$. Then it follows by Prop. 3.2.4 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 a polyhedral set. By Prop. 3.2.4, it follows that X^* is nonempty.

3.5 (Properties of Quasiconvex Functions)

Show the following properties of quasiconvex functions (cf. Exercise 3.4):

(a) A function $f : \Re^n \mapsto (-\infty, \infty]$ is quasiconvex if and only if

$$f\big(\alpha x + (1-\alpha)y\big) \le \max\big\{f(x),\, f(y)\big\}, \qquad \forall\, x, y \in \Re^n,\ \alpha \in [0,1]. \quad (3.11)$$

(b) A differentiable function $f : \Re^n \mapsto (-\infty, \infty)$ is quasiconvex if and only if

$$f(y) \le f(x) \quad \Rightarrow \quad \nabla f(x)'(y-x) \le 0, \qquad \forall\, x, y \in \Re^n.$$

(c) If $f : \Re^n \mapsto [-\infty, \infty]$ is a quasiconvex function, A is a matrix and b is a vector, the function

$$g(x) = f(Ax+b), \qquad x \in \Re^n,$$

is quasiconvex.

Solution: (a) Let f be quasiconvex. If either $f(x) = \infty$ or $f(y) = \infty$, then Eq. (3.11) holds, so assume that

$$\max\big\{f(x), f(y)\big\} < \infty.$$

Consider the level set

$$L = \Big\{z \ \Big|\ f(z) \le \max\big\{f(x), f(y)\big\}\Big\}.$$

Since x and y belong to L, and by quasiconvexity, L is convex, we have $\alpha x + (1-\alpha)y \in L$ for all $\alpha \in [0,1]$, so Eq. (3.11) holds.

Conversely, let Eq. (3.11) hold, and consider two points x and y in a level set $\big\{x \mid f(x) \le \gamma\big\}$. For any $\alpha \in [0,1]$, we have

$$f\big(\alpha x + (1-\alpha)y\big) \le \max\big\{f(x),\, f(y)\big\} \le \gamma,$$

so the level set $\big\{x \mid f(x) \le \gamma\big\}$ is convex, implying that f is quasiconvex.

(b) Let f be quasiconvex, and let x and y be such that $f(y) \le f(x)$. Assume, to arrive at a contradiction, that $\nabla f(x)'(y-x) > 0$. Then by Taylor's Theorem, we have $f\big(x + \epsilon(y-x)\big) > f(x)$ for all sufficiently small $\epsilon > 0$. This contradicts Eq. (3.11), which holds by quasiconvexity of f, as shown in part (a).

Conversely, assume that

$$f(y) \le f(x) \quad \Rightarrow \quad \nabla f(x)'(y-x) \le 0, \qquad \forall\, x, y \in \Re^n. \qquad (3.12)$$

Note that in one dimension, this relation implies that if we have

$$f(z) > \max\big\{f(x), f(y)\big\}$$

for some point z in the line segment connecting x and y, then the slope of f at z is zero. This shows that we cannot have $f(z) > \max\big\{f(x), f(y)\big\}$ for any z in the line segment connecting x and y, and by part (a) proves that f is quasiconvex. This is the idea underlying the following argument.

Assume, to arrive at a contradiction, that f is not quasiconvex, so that, by part (a), there exist x and y, and $\overline{\alpha} \in (0,1)$ such that $f(y) \leq f(x)$ and

$$f(z) > \max\big\{f(x), f(y)\big\} = f(x),$$

where

$$z = \overline{\alpha}x + (1-\overline{\alpha})y.$$

By the continuity of f, there exists a scalar $\beta \in (\overline{\alpha}, 1]$ such that

$$f\big(\alpha x + (1-\alpha)y\big) > f(x), \qquad \forall\ \alpha \in (\overline{\alpha}, \beta),$$

and

$$f(w) = f(x),$$

where

$$w = \beta x + (1-\beta)y.$$

Using the Mean Value Theorem, it follows that there exists $\gamma \in (\overline{\alpha}, \beta)$ such that

$$\nabla f(\overline{w})'(z-w) = f(z) - f(w) > 0,$$

where

$$\overline{w} = \gamma x + (1-\gamma)y.$$

Since the vector $z - w$ is colinear with the vector $y - \overline{w}$, follows that

$$\nabla f(\overline{w})'(y-\overline{w}) > 0,$$

which contradicts Eq. (3.12), since by construction we have

$$f(y) \leq f(x) = f(w) \leq f(\overline{w}).$$

(c) To show that g is quasiconvex, we must show that for all $\gamma \in \Re$, the set

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

is convex. We have

$$V_\gamma = \{x \mid Ax + b = y,\ y \in L_\gamma\},$$

where

$$L_\gamma = \big\{y \mid f(y) \leq \gamma\big\}.$$

Since f is convex, it follows that L_γ is convex, which implies that V_γ is convex.

3.6 (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. 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.,

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

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

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

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

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

(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 that $X \cap \text{dom}(f) \ne \emptyset$.

Solution: (a) 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

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

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. 3.2.4 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 $L_X \cap L_f \ne R_X \cap R_f$. 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, $x + \alpha y \in X \cap \big(\cap_{k=0}^{\infty} C_k\big)$ 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

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

Using Prop. 1.4.2(c), 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

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

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

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

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

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

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 $\mathrm{aff}(X)$, 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 $\overline{X} \cap \left(\cap_{k=0}^{\infty} C_k\right)$ is nonempty. Since $\overline{X} \subset X$, it follows that $X \cap \left(\cap_{k=0}^{\infty} C_k\right)$ is nonempty.

(b) 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

$$C_k = \left\{x \in \Re^n \mid f(x) \leq w_k\right\}.$$

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 $X \cap \left(\cap_{k=0}^{\infty} C_k\right)$, the set of minimizers of f over X, is nonempty.

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

$$F_f = L_f = \left\{y \mid y = 0\right\},$$

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.

3.7 (Bidirectionally Flat Functions - Intersections of Sets Defined by Quadratic Functions)

The purpose of the exercise is to provide refinements of various results relating to set intersections, closedness under linear transformations, existence of optimal solutions, and closedness under partial minimization. Important special cases arise when the sets involved are defined by convex quadratic functions.

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 3.6). 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 $\cap_{j=1}^r \text{dom}(g_j)$, 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.

Solution: (a) As a first step, we will show that either $\cap_{k=1}^\infty C_k \ne \emptyset$ 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 $\cap_{k=1}^\infty C_k \ne \emptyset$. 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 $\cap_{k=1}^{\infty} C_k \neq \emptyset$. This completes the proof that

$$\cap_{k=1}^{\infty} C_k = \emptyset \quad \Rightarrow \quad \text{there exists } \overline{j} \in \{1, \ldots, r\} \text{ and } y \in \cap_{j=0}^{r} R_j \text{ with } y \notin F_{g_{\overline{j}}}. \tag{3.13}$$

We now use induction on r. For $r = 0$, the preceding proof shows that $\cap_{k=1}^{\infty} C_k \neq \emptyset$. Assume that $\cap_{k=1}^{\infty} C_k \neq \emptyset$ for all cases where $r < \overline{r}$. We will show that $\cap_{k=1}^{\infty} C_k \neq \emptyset$ for $r = \overline{r}$. Assume the contrary. Then, by Eq. (3.13), 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

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

Since these sets are nonempty, by the induction hypothesis, $\cap_{k=1}^{\infty} \overline{C}_k \neq \emptyset$. 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 $\text{dom}(\phi) = (0, 1)$ [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 $\text{dom}(g_1)$, which is equal to

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

(in fact the two sets are disjoint).

(b) 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) 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.

SECTION 3.3: Partial Minimization of Convex Functions

3.8 (Asymptotic Slopes of Functions of Two Vectors)

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

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

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

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

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

Solution: By Prop. 1.4.5(a), the recession cone of F has the form

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

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

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

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

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

(b) This is a consequence of the fact that the recession function of $F(x, \cdot)$ is the same for all $x \in X$, and the comments following Prop. 1.4.7.

3.9 (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 [this shows Prop. 3.3.1(a) by a somewhat different argument, which does not rely on the assumption $F(x, z) > -\infty$ for all (x, z)].

Solution: (a) We clearly have $E_f \subset \text{epi}(f)$. Let $(x, w) \in \text{epi}(f)$, so that $f(x) \leq w$. Let $\{w_k\}$ be a decreasing sequence such that $w_k \to w$, so that $f(x) < w_k$ and

$(x, w_k) \in E_f$ for all k. Since $(x, w_k) \to (x, w)$, it follows that $(x, w) \in \mathrm{cl}(E_f)$. Hence, $\mathrm{epi}(f) \subset \mathrm{cl}(E_f)$.

Let f be convex, and let $(x, w), (y, v) \in E_f$. Then for any $\alpha \in [0, 1]$, we have

$$f\big(\alpha x + (1-\alpha)y\big) \le \alpha f(x) + (1-\alpha)f(y) < \alpha w + (1-\alpha)v.$$

It follows that $\big(\alpha x + (1-\alpha)y, \alpha w + (1-\alpha)v\big) \in E_f$, so that E_f is convex.

Let E_f be convex, and for any $x, y \in \mathrm{dom}(f)$, let $\big\{(x, w_k)\big\}$, $\big\{(y, v_k)\big\}$ be sequences in E_f such that $w_k \downarrow f(x)$ and $v_k \downarrow f(y)$, respectively [we allow the possibility that $f(x)$ and/or $f(y)$ are equal to $-\infty$]. Since E_f is convex, we have

$$f\big(\alpha x + (1-\alpha)y\big) < \alpha w_k + (1-\alpha)v_k, \qquad \forall\ \alpha \in [0, 1],\ k = 0, 1, \ldots.$$

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

$$f\big(\alpha x + (1-\alpha)y\big) \le \alpha f(x) + (1-\alpha)f(y), \qquad \forall\ \alpha \in [0, 1],$$

so f is convex.

(b) We have $f(x) < w$ if and only if there exists $z \in \Re^m$ such that $F(x, z) < w$:

$$E_f = \big\{(x, w) \mid F(x, z) < w \text{ for some } z \in \Re^m\big\}.$$

Thus 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) If F is convex, by part (a) the set $\big\{(x, z, w) \mid F(x, z) < w\big\}$ is convex, and by part (b), E_f is the image of this set under a linear transformation. Therefore, E_f is convex, so by part (a), f is convex.

3.10 (Closures of Partially Minimized Functions)

Consider a function $F : \Re^{n+m} \mapsto (-\infty, \infty]$ and the function $f : \Re^n \mapsto [-\infty, \infty]$ defined by

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

(a) Show that

$$P\Big(\mathrm{cl}\big(\mathrm{epi}(F)\big)\Big) \subset \mathrm{cl}\big(\mathrm{epi}(f)\big),$$

where $P(\cdot)$ denotes projection on the space of (x, w).

(b) Let $\overline{f}$ be defined by

$$\overline{f}(x) = \inf_{z \in \Re^m} (\mathrm{cl}\, F)(x, z),$$

where $\mathrm{cl}\, F$ is the closure of F. Show that the closures of f and $\overline{f}$ coincide.

Solution: (a) We first note that from Prop. 3.3.1, we have

$$P\big(\mathrm{epi}(F)\big) \subset \mathrm{epi}(f) \subset \mathrm{cl}\Big(P\big(\mathrm{epi}(F)\big)\Big), \tag{3.14}$$

where $P(\cdot)$ denotes projection on the space of (x, w). To show the equation

$$P\big(\mathrm{cl}\big(\mathrm{epi}(F)\big)\big) \subset \mathrm{cl}\big(\mathrm{epi}(f)\big), \tag{3.15}$$

let $(\overline{x}, \overline{w})$ belong to $P\big(\mathrm{cl}\big(\mathrm{epi}(F)\big)\big)$. Then there exists $\overline{z}$ such that $(\overline{x}, \overline{z}, \overline{w}) \in \mathrm{cl}\big(\mathrm{epi}(F)\big)$, and hence there is a sequence $(x_k, z_k, w_k) \in \mathrm{epi}(F)$ such that $x_k \to \overline{x}$, $z_k \to \overline{z}$, and $w_k \to \overline{w}$. Thus we have $f(x_k) \leq F(x_k, z_k) \leq w_k$, implying that $(x_k, w_k) \in \mathrm{epi}(f)$ for all k. It follows that $(\overline{x}, \overline{w}) \in \mathrm{cl}\big(\mathrm{epi}(f)\big)$.

(b) By taking closure in Eq. (3.14), we see that

$$\mathrm{cl}\big(\mathrm{epi}(f)\big) = \mathrm{cl}\Big(P\big(\mathrm{epi}(F)\big)\Big). \tag{3.16}$$

Denoting $\overline{F} = \mathrm{cl}\, F$ and replacing F with $\overline{F}$, we also have

$$\mathrm{cl}\big(\mathrm{epi}(\overline{f})\big) = \mathrm{cl}\Big(P\big(\mathrm{epi}(\overline{F})\big)\Big). \tag{3.17}$$

On the other hand, by taking closure in Eq. (3.15), we have

$$\mathrm{cl}\Big(P\big(\mathrm{epi}(\overline{F})\big)\Big) \subset \mathrm{cl}\Big(P\big(\mathrm{epi}(F)\big)\Big),$$

which, in view of $\mathrm{epi}(\overline{F}) \supset \mathrm{epi}(F)$, implies that

$$\mathrm{cl}\Big(P\big(\mathrm{epi}(\overline{F})\big)\Big) = \mathrm{cl}\Big(P\big(\mathrm{epi}(F)\big)\Big). \tag{3.18}$$

By combining Eqs. (3.16)-(3.18), we see that

$$\mathrm{cl}\big(\mathrm{epi}(f)\big) = \mathrm{cl}\big(\mathrm{epi}(\overline{f})\big).$$

3.11 (Counterexample for Partially Minimized Functions)

Consider the function of $x \in \Re$ and $z \in \Re$ given by

$$F(x, z) = \begin{cases} e^{-\sqrt{xz}} & \text{if } x \geq 0,\ z \geq 0, \\ \infty & \text{otherwise.} \end{cases}$$

Verify that F convex and closed, but that the function $f(x) = \inf_{z \in \Re} F(x, z)$ is convex but not closed.

Solution: To prove that F is convex, we note that it is the composition of the monotonically increasing exponential function, and the function $-\sqrt{xz}$ which can be shown to convex over $\big\{(x, z) \mid x \geq 0,\ z \geq 0\big\}$ (one way to do this is to use Prop. 1.1.7). It is straightforward to verify that f is the nonclosed function

$$f(x) = \inf_{z \in \Re} F(x, z) = \begin{cases} 0 & \text{if } x > 0, \\ 1 & \text{if } x = 0, \\ \infty & \text{if } x < 0. \end{cases}$$

In particular, if $x > 0$ and $\epsilon > 0$, we have $0 < e^{-\sqrt{xz}} < \epsilon$, provided z is sufficiently large.

3.12 (Image Operation)

Let $h : \Re^m \mapsto (-\infty, \infty]$ be a function and A be an $m \times n$ matrix. Consider the function $A \circ h : \Re^n \mapsto [-\infty, \infty]$ defined by

$$(A \circ h)(x) = \inf_{Az=x} h(z), \qquad x \in \Re^n. \tag{3.19}$$

We refer to $A \circ h$ as the *image of* h *under* A. (The terminology comes from the special case where h is the indicator function of a set; then $A \circ h$ is the indicator function of the image of the set under A.)

(a) Show that if h is convex, then $A \circ h$ is convex.

(b) Assume that h is closed proper convex and that $R(A) \cap \text{dom}(h) \neq \emptyset$, where $R(A)$ is the range of A. Assume further that every direction of recession of h that belongs to the nullspace of A is a direction along which h is constant. Then $A \circ h$ is closed proper convex, and for every $x \in \text{dom}(f)$, the infimum in the definition of $(A \circ h)(x)$ is attained. *Hint*: Use Prop. 3.3.4.

Solution: Consider the function $F : \Re^{n+m} \mapsto (-\infty, \infty]$ given by

$$F(x, z) = \begin{cases} h(z) & \text{if } Az = x, \\ \infty & \text{otherwise.} \end{cases}$$

We have

$$(A \circ h)(x) = \inf_{z \in \Re^m} F(x, z), \qquad x \in \Re^n,$$

so $A \circ h$ is obtained from by partial minimization, and its properties follow by applying the theory of Section 3.3.

(a) By Prop. 3.3.1(a), $A \circ h$ is convex.

(b) We will use Prop. 3.3.4. Let $\overline{x}$ be a point in $R(A) \cap \text{dom}(h)$. Choose $\overline{\gamma}$ so that the set $\{z \mid h(z) \leq \overline{\gamma},\ Az = \overline{x}\}$ is nonempty, and note that

$$\{z \mid h(z) \leq \overline{\gamma},\ Az = \overline{x}\} = \{z \mid F(\overline{x}, z) \leq \overline{\gamma}\}.$$

From this we see that our assumption that every direction of recession of h that belongs to the nullspace of A is a direction along which h is constant is equivalent to the recession cone of $\{z \mid F(\overline{x}, z) \leq \overline{\gamma}\}$ being equal to its lineality space. Thus Prop. 3.3.4 applies, and by the conclusion of the proposition, $A \circ h$ is closed proper convex, and for every $x \in \text{dom}(A \circ h)$, the infimum in the definition of $(A \circ h)(x)$ is attained.

3.13 (Infimal Convolution Operation)

Consider the function $f : \Re^n \mapsto [-\infty, \infty]$ defined by

$$f(x) = \inf_{x_1+\cdots+x_m=x} \{f_1(x_1) + \cdots + f_m(x_m)\}, \qquad x \in \Re^n,$$

where for $i = 1, \ldots, m$, $f_i : \Re^n \mapsto (-\infty, \infty]$ is a closed proper convex function. The function f is called the *infimal convolution of* $f_1, \ldots, f_m$. The terminology comes from the resemblance with the classical formula of integral convolution in the case of two functions when $f(x)$ takes the form

$$\inf_{z \in \Re^n} \big\{ f_1(x - z) + f_2(z) \big\}.$$

Assume that $\cap_{i=1}^m \mathrm{dom}(f_i) \neq \emptyset$, and furthermore that every $d = (d_1, \ldots, d_m)$ that is a direction of recession of $f_1 + \cdots + f_m$ and satisfies

$$d_1 + \cdots + d_m = 0,$$

is a direction along which $f_1 + \cdots + f_m$ is constant. Show that f is closed proper convex, and that the infimum in the definition of $f(x)$ is attained for every $x \in \mathrm{dom}(f)$. *Hint*: Show that infimal convolution is a special case of the image operation of Exercise 3.12, and apply the result of that exercise.

Solution: We can write f as the result of an image operation, $(A \circ h)$ [cf. Eq. (3.19)], where $h : \Re^{mn} \mapsto (-\infty, \infty]$ is the function given by

$$h(x_1, \ldots, x_m) = f_1(x_1) + \cdots + f_m(x_m),$$

and A is the $n \times (mn)$ matrix defined by

$$A(x_1, \ldots, x_m) = x_1 + \cdots + x_m.$$

Our assumption on directions of recession is the same as the assumption of Exercise 3.12(b), specialized to the infimal convolution context.

3.14 (Epigraph Relations for Image and Infimal Convolution)

Consider the image operation $A \circ h$ of Exercise 3.12, and the relation

$$(A \circ h)(x) = \inf_{z \in \Re^m} F(x, z), \qquad x \in \Re^n,$$

where $F : \Re^{n+m} \mapsto (-\infty, \infty]$ is given by

$$F(x, z) = \begin{cases} h(z) & \text{if } Az = x, \\ \infty & \text{otherwise.} \end{cases}$$

(a) Show that the epigraph relation

$$P\big(\mathrm{epi}(F)\big) \subset \mathrm{epi}(f) \subset \mathrm{cl}\Big(P\big(\mathrm{epi}(F)\big)\Big)$$

where $P(\cdot)$ denotes projection on the space of (x, w) [cf. Prop. 3.3.1(b)] takes the form

$$\overline{A}\,\mathrm{epi}(h) \subset \mathrm{epi}(f) \subset \mathrm{cl}\big(\overline{A}\,\mathrm{epi}(h)\big),$$

where $\overline{A}$ is the linear transformation that maps a vector $(x, w) \in \Re^{n+1}$ to (Ax, w).

(b) Show that in the special case of the infimal convolution operation (cf. Exercise 3.13) the epigraph relation takes the form

$$\text{epi}(f_1) + \cdots + \text{epi}(f_m) \subset \text{epi}(f) \subset \text{cl}\big(\text{epi}(f_1) + \cdots + \text{epi}(f_m)\big).$$

Solution: In the case of the image operation, we have

$$\text{epi}(F) = \big\{(x, z, w) \mid Az = x,\ h(z) \le w\big\},$$

so that

$$\begin{aligned} P\big(\text{epi}(F)\big) &= \big\{(x, w) \mid \text{for some } z \text{ with } Az = x,\ h(z) \le w\big\} \\ &= \big\{(Az, w) \mid h(z) \le w\big\} \\ &= \overline{A}\,\text{epi}(h). \end{aligned}$$

We thus obtain

$$\overline{A}\,\text{epi}(h) \subset \text{epi}(f) \subset \text{cl}\big(\overline{A}\,\text{epi}(h)\big).$$

(b) Similar to part (a).

3.15 (Conjugates of Linear Composition and Image)

Consider the composition of a closed proper convex function $f : \Re^m \mapsto (-\infty, \infty]$ with a linear transformation A, an $m \times n$ matrix, i.e., the function $f \circ A : \Re^n \mapsto [-\infty, \infty]$ given by

$$(f \circ A)(x) = f(Ax).$$

Consider also the function $A' \circ f^\star$ given by

$$(A' \circ f^\star)(x) = \inf_{A'z = x} f^\star(z), \qquad x \in \Re^n,$$

which was called the image function of $f^\star$ under A' in Example 3.13. Show that the conjugate of $A' \circ f^\star$ is $f \circ A$, and that the conjugate of $f \circ A$ is the closure of $A' \circ f^\star$, provided $f \circ A$ is proper [which is true if and only if the range of A contains a point in $\text{dom}(f)$]. Give an example where the conjugate of $f \circ A$ is different from $A' \circ f^\star$.

Solution: We have for all $z \in \Re^n$,

$$\begin{aligned} f(Az) &= \sup_y \big\{z'A'y - f^\star(y)\big\} \\ &= \sup_x \sup_{A'y = x} \big\{z'x - f^\star(y)\big\} \\ &= \sup_x \Big\{z'x - \inf_{A'y = x} f^\star(y)\Big\} \\ &= \sup_x \big\{z'x - (A' \circ f^\star)(x)\big\}. \end{aligned}$$

Thus the conjugate of $A' \circ f^\star$ is $f \circ A$. By the Conjugacy Theorem [Prop. 1.6.1(c)], the conjugate of $f \circ A$ is the closure of $A' \circ f^\star$, provided $f \circ A$ is proper.

Let $f^\star : \Re^2 \mapsto (-\infty, \infty]$ be the closed proper convex function

$$f^\star(y_1, y_2) = \begin{cases} e^{-\sqrt{y_1 y_2}} & \text{if } y_1 \geq 0,\ y_2 \geq 0, \\ \infty & \text{otherwise,} \end{cases}$$

and let A' be projection on the space of y_1, i.e., $A'(y_1, y_2) = y_1$ (cf. Exercise 3.11). Then for $x \geq 0$,

$$(A' \circ f^\star)(x) = \inf_{A'y = x} f^\star(y) = \inf_{y_1 = x} e^{-\sqrt{y_1 y_2}} = \begin{cases} 0 & \text{if } x > 0, \\ 1 & \text{if } x = 0, \end{cases}$$

while for $x < 0$,

$$(A' \circ f^\star)(x) = \infty.$$

Thus $A' \circ f^\star$ is convex proper but not closed.

3.16 (Conjugates of Sum and Infimal Convolution)

Consider the function

$$(f_1 + \cdots + f_m)(x) = f_1(x) + \cdots + f_m(x),$$

where $f_i : \Re^n \mapsto (-\infty, \infty]$, $i = 1, \ldots, m$, are closed proper convex functions, with conjugates denoted by $f_i^\star$. Consider also the function $f_1^\star \oplus \cdots \oplus f_m^\star$ given by

$$(f_1^\star \oplus \cdots \oplus f_m^\star)(x) = \inf_{x_1 + \cdots + x_m = x} \big\{ f_1^\star(x_1) + \cdots + f_m^\star(x_m) \big\}, \qquad x \in \Re^n,$$

which was called the infimal convolution of $f_1^\star, \ldots, f_m^\star$ in Exercise 3.13. Show that the conjugate of $f_1^\star \oplus \cdots \oplus f_m^\star$ is $f_1 + \cdots + f_m$, and the conjugate of $f_1 + \cdots + f_m$ is the closure of $f_1^\star \oplus \cdots \oplus f_m^\star$, provided $f_1 + \cdots + f_m$ is proper [which is true if and only if $\cap_{i=1}^m \text{dom}(f_i) \neq \emptyset$].

Solution: We have for all $y \in \Re^n$,

$$\begin{aligned} f_1(y) + \cdots + f_m(y) &= \sup_{x_1} \big\{ y'x_1 - f_1^\star(x_1) \big\} + \cdots + \sup_{x_m} \big\{ y'(x_m - f_m^\star(x_m) \big\} \\ &= \sup_{x_1, \ldots, x_m} \big\{ y'(x_1 + \cdots + x_m) - f_1^\star(x_1) - \cdots - f_m^\star(x_m) \big\} \\ &= \sup_x \sup_{x_1 + \cdots + x_m = x} \big\{ y'x - f_1^\star(x_1) - \cdots - f_m^\star(x_m) \big\} \\ &= \sup_x \left\{ y'x - \inf_{x_1 + \cdots + x_m = x} \big\{ f_1^\star(x_1) + \cdots + f_m^\star(x_m) \big\} \right\} \\ &= \sup_x \big\{ y'x - (f_1^\star \oplus \cdots \oplus f_m^\star)(x) \big\}. \end{aligned}$$

Thus the conjugate of $f_1^\star \oplus \cdots \oplus f_m^\star$ is $f_1 + \cdots + f_m$. By the Conjugacy Theorem [Prop. 1.6.1(c)] and given the properness of $f_1 + \cdots + f_m$, it follows that the conjugate of $f_1 + \cdots + f_m$ is the closure of $f_1^\star \oplus \cdots \oplus f_m^\star$, provided $f_1 + \cdots + f_m$ is proper.

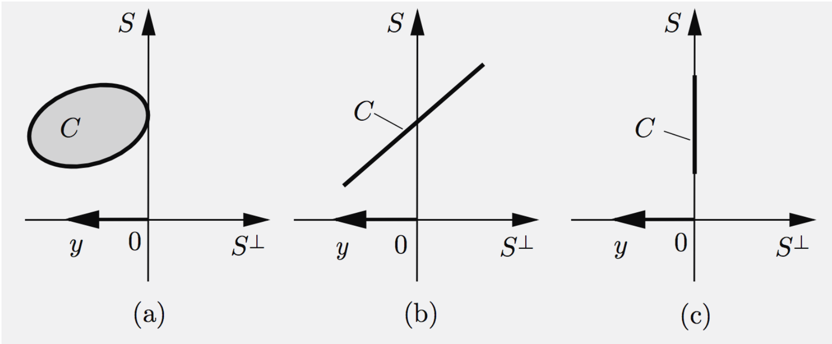

Figure 3.1. Visualization and proof of the Proper Separation Lemma of Exercise 3.17. As figure (a) illustrates, we have $S \cap \mathrm{ri}(C) = \emptyset$ if and only if S and C can be properly separated (necessarily by a hyperplane whose normal belongs to $S^{\perp}$). Thus, $S \cap \mathrm{ri}(C) = \emptyset$ if and only if there exists $y \in S^{\perp}$ such that $\sigma_C(y) > 0$ and $\sigma_C(-y) \leq 0$, as in figure (a). Equivalently, $S \cap \mathrm{ri}(C) \neq \emptyset$ if and only if for all $y \in S^{\perp}$ with $y \neq 0$, we have either

$$\sigma_C(y) > 0, \qquad \sigma_C(-y) > 0$$

as in figure (b), or

$$\sigma_C(y) = \sigma_C(-y) = 0$$

as in figure (c). This is equivalent to the statement of Lemma.

3.17 (Proper Separation Lemma)

This exercise and the next one provide machinery that will be used to develop sharper characterizations of the conjugates of the image and infimal convolution operations (see Fig. 3.1). Let S be a subspace and let C be a convex set. Then

$$S \cap \mathrm{ri}(C) \neq \emptyset$$

if and only if the support function σ_C satisfies

$$\sigma_C(y) = \sigma_C(-y) = 0, \qquad \forall\, y \in S^{\perp} \text{ with } \sigma_C(y) \leq 0.$$

Solution: Since $\mathrm{ri}(S) = S$, by the Proper Separation Theorem (Prop. 1.5.6), we have $S \cap \mathrm{ri}(C) = \emptyset$ if and only if there exists a hyperplane that properly separates S and C, i.e., a $y \in S^{\perp}$ such that

$$\sup_{x \in C} x'y > 0, \qquad \inf_{x \in C} x'y \geq 0,$$

or equivalently,

$$\sigma_C(y) > 0, \qquad \sigma_C(-y) \leq 0,$$

since the sup and inf of $x'y$ over C are the support function values $\sigma_C(y)$ and $-\sigma_C(-y)$. This shows the result, as illustrated also in Fig. 3.1.

3.18 (Relative Interior of Domain and Directions of Recession)

Let S be a subspace of $\Re^n$, let $f : \Re^n \mapsto (-\infty, \infty]$ be a proper convex function, and let $f^\star$ be its conjugate.

(a) We have

$$S \cap \mathrm{ri}\big(\mathrm{dom}(f)\big) \neq \varnothing$$

if and only if every direction of recession of $f^\star$ that belongs to $S^\perp$ is a direction along which $f^\star$ is constant.

(b) If f is closed, we have

$$S \cap \mathrm{ri}\big(\mathrm{dom}(f^\star)\big) \neq \varnothing$$

if and only if every direction of recession of f that belongs to $S^\perp$ is a direction along which f is constant.

Solution: (a) Use Exercise 3.17 with $C = \mathrm{dom}(f)$, and with σ_C equal to the recession function $r_{f^\star}$. Note that the support function of $\mathrm{dom}(f)$ is the recession function of $f^\star$, according to one of the exercises for Chapter 1.

(b) If f is closed, then by the Conjugacy Theorem [Prop. 1.6.1(c)], it is the conjugate of $f^\star$, and the result follows by applying part (a) with the roles of f and $f^\star$ interchanged.

3.19 (Closedness of Image)

Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper convex function, and let $f^\star$ be its conjugate. Consider the image function

$$(A \circ f)(x) = \inf_{Az=x} f(z), \qquad x \in \Re^n,$$

where A is an $m \times n$ matrix. Note that the convexity of $A \circ f$ was established in Exercise 3.12, and the calculation of its conjugate was given as $(f^\star \circ A')(y) = f^\star(A'y)$ in Exercise 3.15. This exercise provides criteria for closedness of $A \circ f$.

(a) If

$$R(A') \cap \mathrm{ri}\big(\mathrm{dom}(f^\star)\big) \neq \varnothing,$$

then $A \circ f$ is closed and proper, and the infimum in its definition is attained for all $x \in \mathrm{dom}(A \circ f)$. Furthermore, $A \circ f$ is the conjugate of $f^\star \circ A'$.

(b) If f is polyhedral and

$$R(A') \cap \mathrm{dom}(f^\star) \neq \varnothing,$$

then $A \circ f$ is polyhedral, and the infimum in its definition is attained for all $x \in \mathrm{dom}(A \circ f)$. Furthermore, $A \circ f$ is the conjugate of $f^\star \circ A'$.

Solution: To simplify notation, we denote $g = A \circ f$.

(a) By Exercise 3.18, every direction of recession of f that belongs to $R(A')^\perp$ [which is $N(A)$] is a direction along which f is constant. From Exercise 3.12(b), it follows that g is closed and proper, and the infimum in its definition is attained for all $x \in \mathrm{dom}(g)$. Furthermore, g is the conjugate of $f^\star \circ A'$ in view of the closedness of g and Exercise 3.15.

(b) We view g as the result of partial minimization of a polyhedral function. Therefore, using Prop. 3.3.1(b) and the fact that projection of a polyhedral set is polyhedral, the epigraph of g is polyhedral. Furthermore, the assumption $R(A') \cap \mathrm{dom}(h) \neq \emptyset$ implies that there exists $\overline{x} \in \Re^m$ such that $f^\star(A'\overline{x}) < \infty$ or

$$\sup_{y \in \Re^n} \bigl\{ y'A'\overline{x} - f(y) \bigr\} < \infty.$$

It follows that

$$y'A'\overline{x} \leq f(y), \qquad \forall\ y \in \Re^n,$$

so using the definition of g, we have $x'\overline{x} \leq g(x)$ for all $x \in \Re^m$. Therefore, $-\infty < g(x)$ for all $x \in \Re^m$, so that g is proper, and hence polyhedral by Prop. 3.3.5.

Finally, since minimizing $f(z)$ over $\{z \mid Az = x\}$ is a linear program whose optimal value is finite for all $x \in \mathrm{dom}(g)$, by Prop. 1.4.12, it follows that the infimum in the definition of g is attained for all $x \in \mathrm{dom}(g)$.

3.20 (Closedness of Infimal Convolution)

Let $f_i : \Re^n \mapsto (-\infty, \infty]$, $i = 1, \ldots, m$, be closed proper convex functions, and let $f_i^\star$ be their corresponding conjugates. Consider the infimal convolution

$$(f_1 \oplus \cdots \oplus f_m)(x) = \inf_{x_1 + \cdots + x_m = x} \bigl\{ f_1(x_1) + \cdots + f_m(x_m) \bigr\}, \quad x \in \Re^n.$$

Note that the convexity of $f_1 \oplus \cdots \oplus f_m$ was established in Exercise 3.13, and the calculation of its conjugate was given as $f_1^\star + \cdots + f_m^\star$ in Exercise 3.16. This exercise provides criteria for closedness of $f_1 \oplus \cdots \oplus f_m$. Show that if for some k, the functions $f_1, \ldots, f_k$ are polyhedral, the functions $f_{k+1}, \ldots, f_m$ are closed, and

$$\Bigl(\cap_{i=1}^k \mathrm{dom}(f_i^\star) \Bigr) \cap \Bigl(\cap_{i=k+1}^m \mathrm{ri}\bigl(\mathrm{dom}(f_i^\star)\bigr) \Bigr) \neq \emptyset, \tag{3.20}$$

then $f_1 \oplus \cdots \oplus f_m$ is closed and proper, and the infimum in its definition is attained for all $x \in \mathrm{dom}(f_1 \oplus \cdots \oplus f_m)$. In this case, $f_1 \oplus \cdots \oplus f_m$ is the conjugate of $f_1^\star + \cdots + f_m^\star$.

Solution: We first show the result in the two special cases where all the functions f_i are nonpolyhedral and polyhedral, respectively, and we then combine the two cases to show the result in its full generality. To simplify notation, we denote

$$g = f_1 \oplus \cdots \oplus f_m.$$

We first assume that $\cap_{i=1}^m \mathrm{ri}\big(\mathrm{dom}(f_i^\star)\big) \neq \emptyset$, and we show that the result follows from the corresponding result for the image operation (Exercise 3.19). Indeed, the infimal convolution operation is the special case of the image operation $A \circ f$, where f and A are given by

$$f(x) = f_1(x_1) + \cdots + f_m(x_m), \qquad A(x_1, \ldots, x_m) = x_1 + \cdots + x_m.$$

It is straightforward to verify that the condition $\cap_{i=1}^m \mathrm{ri}\big(\mathrm{dom}(f_i^\star)\big) \neq \emptyset$ is equivalent to

$$R(A') \cap \mathrm{ri}\big(\mathrm{dom}(f^\star)\big) \neq \emptyset,$$

where $f^\star$ is the conjugate of f. Hence, by Exercise 3.19(a), g is closed and proper, and the infimum in its definition is attained for all $x \in \mathrm{dom}(g)$.

A similar proof, using Exercise 3.19(b), shows the result under the assumption that $\cap_{i=1}^m \mathrm{dom}(f_i^\star) \neq \emptyset$ and that all functions f_i are polyhedral.

Finally, to show the result in the general case, we introduce the polyhedral function

$$p(y) = f_1^\star(y) + \cdots + f_k^\star(y)$$

and the closed proper convex function

$$q(y) = f_{k+1}^\star(y) + \cdots + f_m^\star(y).$$

The assumption (3.20) then amounts to $\mathrm{dom}(p) \cap \mathrm{ri}\big(\mathrm{dom}(q)\big) \neq \emptyset$, since $\mathrm{dom}(p) = \cap_{i=1}^k \mathrm{dom}(f_i^\star)$, and the condition $\cap_{i=1}^m \mathrm{ri}\big(\mathrm{dom}(f_i^\star)\big) \neq \emptyset$ implies that

$$\mathrm{ri}\big(\mathrm{dom}(q)\big) = \cap_{i=k+1}^m \mathrm{ri}\big(\mathrm{dom}(f_i^\star)\big)$$

by Prop. 1.3.8.

Let M be the affine hull of $\mathrm{dom}(q)$ and assume for simplicity, and without loss of generality, that M is a subspace. It can be seen that the condition $\mathrm{dom}(p) \cap \mathrm{ri}\big(\mathrm{dom}(q)\big) \neq \emptyset$ is equivalent to

$$\mathrm{ri}\big(\mathrm{dom}(p) \cap M\big) \cap \mathrm{ri}\big(\mathrm{dom}(q)\big) \neq \emptyset$$

[see the exercises for Chapter 1; the proof is to choose $\overline{y} \in \mathrm{ri}\big(\mathrm{dom}(p) \cap M\big)$ and $y \in \mathrm{dom}(p) \cap \mathrm{ri}\big(\mathrm{dom}(q)\big)$ [which belongs to $\mathrm{dom}(p) \cap M$], consider the line segment connecting y and $\overline{y}$, and use the Line Segment Principle to conclude that points close to y belong to $\mathrm{dom}(p) \cap \mathrm{ri}\big(\mathrm{dom}(q)\big)$]. Therefore, if we replace p with the function $\hat{p}$ given by

$$\hat{p}(y) = \begin{cases} p(y) & \text{if } y \in M, \\ \infty & \text{otherwise,} \end{cases}$$

whose domain is $\mathrm{dom}(p) \cap M$, the special case of the result shown earlier for nonpolyhedral functions applies. We thus obtain that the function

$$\hat{g}(x) = \inf_{y \in \Re^n} \big\{ r(y) + s(x - y) \big\}$$

is closed and the infimum over y is attained for all $x \in \mathrm{dom}(\hat{g})$, where r and s are the conjugates of $\hat{p}$ and q, respectively. It can be seen by a straightforward calculation that $\hat{g} = g$ and the result follows.

3.21 (Preservation of Closedness Under Linear Transformation)

Let C, C_i, $i = 1, \ldots, m$, be nonempty closed convex subsets of $\Re^n$, and let D, D_i, $i = 1, \ldots, m$, denote the domains of their support functions:

$$D = \text{dom}(\sigma_C), \qquad D_i = \text{dom}(\sigma_{C_i}), \quad i = 1, \ldots, m.$$

(a) Let A be an $m \times n$ matrix. Then AC is closed if $R(A') \cap \text{ri}(D) \neq \emptyset$.

(b) The vector sum $C_1 + \cdots + C_m$ is closed if $\cap_{i=1}^m \text{ri}(D_i) \neq \emptyset$.

(c) The vector sum $C_1 + \cdots + C_m$ is closed if for some $k \geq 1$, $C_1, \ldots, C_k$ is polyhedral, and $\left(\cap_{i=1}^k D_i\right) \cap \left(\cap_{i=k+1}^m \text{ri}(D_i)\right) \neq \emptyset$.

Solution: (a) We apply Exercise 3.19(a) with f and $f^\star$ being the support and indicator functions of C, respectively.

(b), (c) It is sufficient to consider the case of two sets ($m = 2$). We apply the result of Exercise 3.20 with f_1, f_2 equal to the support functions of C_1, C_2, and $f_1^\star, f_2^\star$ equal to the indicator functions of C_1, C_2. Then the indicator function of $C_1 + C_2$ is obtained as the infimal convolution of $f_1^\star, f_2^\star$. Under the given relative interior assumptions, Exercise 3.20 shows that the infimal convolution function is closed, so $C_1 + C_2$ is closed.

3.22 (Partial Minimization of Quasiconvex Functions)

Consider a function $F : \Re^{n+m} \mapsto [-\infty, \infty]$ of vectors $x \in \Re^n$ and $z \in \Re^m$, which is quasiconvex (cf. Exercise 3.4). Show that the function $f : \Re^n \mapsto [-\infty, \infty]$ given by

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

is quasiconvex. *Hint*: The level sets of f are obtained by projection of level sets of F on $\Re^n$ as follows: for any $\gamma \in \Re$ and monotonically decreasing scalar sequence $\{\gamma_k\}$ with $\gamma_k \to \gamma$,

$$\big\{x \mid f(x) \leq \gamma\big\} = \cap_{k=1}^\infty \big\{x \mid \text{there exists } (x, z) \text{ with } F(x, z) \leq \gamma_k\big\}.$$

Solution: We follow the hint. Since intersection and projection preserve convexity, $\big\{x \mid f(x) \leq \gamma\big\}$ is convex for all $\gamma \in \Re$, so f is quasiconvex.

SECTION 3.4: Saddle Point and Minimax Theory

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

$$g(a) = f(a) - a > 0,$$
$$g(b) = f(b) - b < 0.$$

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

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

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

$$x^* = \hat{x}(z^*), \qquad z^* = \hat{z}(x^*). \tag{3.21}$$

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. (3.21), 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.

3.24 (Saddle Points for Quadratic Functions)

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

$$\phi(x, z) = x'Qx + c'x + z'Mx - z'Rz - d'z,$$

where Q and R are symmetric positive semidefinite $n \times n$ and $m \times m$ matrices, respectively, M is an $n \times m$ matrix, $c \in \Re^n$, $d \in \Re^m$, and X and Z are polyhedral subsets of $\Re^n$ and $\Re^m$, respectively. Show that if either $\sup_{z\in Z} \inf_{x\in X} \phi(x, z)$ is finite or $\inf_{x\in X} \sup_{z\in Z} \phi(x, z)$ is finite, there exists a saddle point. Use the case where $\phi(x, z) = x + z$, $X = Z = \Re$, to show that the finiteness assumption is essential.

Solution: Here, the domain of the function

$$\sup_{z\in Z} \Phi(x, z)$$

is polyhedral and the function is convex quadratic within its domain. The result follows from the existence of solutions result for quadratic programs (Prop. 1.4.12).