Convex Analysis and Optimization

Chapter 1 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 March 24, 2004

CHAPTER 1: SOLUTION MANUAL

1.1

Let C be a nonempty subset of $\Re^n$, and let λ_1 and λ_2 be positive scalars. Show that if C is convex, then $(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C$ [cf. Prop. 1.2.1(c)]. Show by example that this need not be true when C is not convex.

Solution: We always have $(\lambda_1 + \lambda_2)C \subset \lambda_1 C + \lambda_2 C$, even if C is not convex. To show the reverse inclusion assuming C is convex, note that a vector x in $\lambda_1 C + \lambda_2 C$ is of the form $x = \lambda_1 x_1 + \lambda_2 x_2$, where $x_1, x_2 \in C$. By convexity of C, we have

$$\frac{\lambda_1}{\lambda_1 + \lambda_2} x_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} x_2 \in C,$$

and it follows that

$$x = \lambda_1 x_1 + \lambda_2 x_2 \in (\lambda_1 + \lambda_2)C.$$

Hence $\lambda_1 C + \lambda_2 C \subset (\lambda_1 + \lambda_2)C$.

For a counterexample when C is not convex, let C be a set in $\Re^n$ consisting of two vectors, 0 and $x \neq 0$, and let $\lambda_1 = \lambda_2 = 1$. Then C is not convex, and $(\lambda_1 + \lambda_2)C = 2C = \{0, 2x\}$, while $\lambda_1 C + \lambda_2 C = C + C = \{0, x, 2x\}$, showing that $(\lambda_1 + \lambda_2)C \neq \lambda_1 C + \lambda_2 C$.

1.2 (Properties of Cones)

Show that:

(a) The intersection $\cap_{i \in I} C_i$ of a collection $\{C_i \mid i \in I\}$ of cones is a cone.

(b) The Cartesian product $C_1 \times C_2$ of two cones C_1 and C_2 is a cone.

(c) The vector sum $C_1 + C_2$ of two cones C_1 and C_2 is a cone.

(d) The closure of a cone is a cone.

(e) The image and the inverse image of a cone under a linear transformation is a cone.

Solution: (a) Let $x \in \cap_{i \in I} C_i$ and let α be a positive scalar. Since $x \in C_i$ for all $i \in I$ and each C_i is a cone, the vector αx belongs to C_i for all $i \in I$. Hence, $\alpha x \in \cap_{i \in I} C_i$, showing that $\cap_{i \in I} C_i$ is a cone.

(b) Let $x \in C_1 \times C_2$ and let α be a positive scalar. Then $x = (x_1, x_2)$ for some $x_1 \in C_1$ and $x_2 \in C_2$, and since C_1 and C_2 are cones, it follows that $\alpha x_1 \in C_1$ and $\alpha x_2 \in C_2$. Hence, $\alpha x = (\alpha x_1, \alpha x_2) \in C_1 \times C_2$, showing that $C_1 \times C_2$ is a cone.

(c) Let $x \in C_1 + C_2$ and let α be a positive scalar. Then, $x = x_1 + x_2$ for some $x_1 \in C_1$ and $x_2 \in C_2$, and since C_1 and C_2 are cones, $\alpha x_1 \in C_1$ and $\alpha x_2 \in C_2$. Hence, $\alpha x = \alpha x_1 + \alpha x_2 \in C_1 + C_2$, showing that $C_1 + C_2$ is a cone.

(d) Let $x \in \mathrm{cl}(C)$ and let α be a positive scalar. Then, there exists a sequence $\{x_k\} \subset C$ such that $x_k \to x$, and since C is a cone, $\alpha x_k \in C$ for all k. Furthermore, $\alpha x_k \to \alpha x$, implying that $\alpha x \in \mathrm{cl}(C)$. Hence, $\mathrm{cl}(C)$ is a cone.

(e) First we prove that $A \cdot C$ is a cone, where A is a linear transformation and $A \cdot C$ is the image of C under A. Let $z \in A \cdot C$ and let α be a positive scalar. Then, $Ax = z$ for some $x \in C$, and since C is a cone, $\alpha x \in C$. Because $A(\alpha x) = \alpha z$, the vector αz is in $A \cdot C$, showing that $A \cdot C$ is a cone.

Next we prove that the inverse image $A^{-1} \cdot C$ of C under A is a cone. Let $x \in A^{-1} \cdot C$ and let α be a positive scalar. Then $Ax \in C$, and since C is a cone, $\alpha A x \in C$. Thus, the vector $A(\alpha x) \in C$, implying that $\alpha x \in A^{-1} \cdot C$, and showing that $A^{-1} \cdot C$ is a cone.

1.3 (Lower Semicontinuity under Composition)

(a) Let $f : \Re^n \mapsto \Re^m$ be a continuous function and $g : \Re^m \mapsto \Re$ be a lower semicontinuous function. Show that the function h defined by $h(x) = g\big(f(x)\big)$ is lower semicontinuous.

(b) Let $f : \Re^n \mapsto \Re$ be a lower semicontinuous function, and $g : \Re \mapsto \Re$ be a lower semicontinuous and monotonically nondecreasing function. Show that the function h defined by $h(x) = g\big(f(x)\big)$ is lower semicontinuous. Give an example showing that the monotonic nondecrease assumption is essential.

Solution: (a) Let $\{x_k\} \subset \Re^n$ be a sequence of vectors converging to some $x \in \Re^n$. By continuity of f, it follows that $\big\{f(x_k)\big\} \subset \Re^m$ converges to $f(x) \in \Re^m$, so that by lower semicontinuity of g, we have

$$\liminf_{k \to \infty} g\big(f(x_k)\big) \geq g\big(f(x)\big).$$

Hence, h is lower semicontinuous.

(b) Assume, to arrive at a contradiction, that h is not lower semicontinuous at some $x \in \Re^n$. Then, there exists a sequence $\{x_k\} \subset \Re^n$ converging to x such that

$$\liminf_{k \to \infty} g\big(f(x_k)\big) < g\big(f(x)\big).$$

Let $\{x_k\}_\mathcal{K}$ be a subsequence attaining the above limit inferior, i.e.,

$$\lim_{k \to \infty,\ k \in \mathcal{K}} g\big(f(x_k)\big) = \liminf_{k \to \infty} g\big(f(x_k)\big) < g\big(f(x)\big). \tag{1.1}$$

Without loss of generality, we assume that

$$g\big(f(x_k)\big) < g\big(f(x)\big), \qquad \forall\ k \in \mathcal{K}.$$

Since g is monotonically nondecreasing, it follows that

$$f(x_k) < f(x), \qquad \forall\, k \in \mathcal{K},$$

which together with the fact $\{x_k\}_\mathcal{K} \to x$ and the lower semicontinuity of f, yields

$$f(x) \leq \liminf_{k\to\infty,\ k\in\mathcal{K}} f(x_k) \leq \limsup_{k\to\infty,\ k\in\mathcal{K}} f(x_k) \leq f(x),$$

showing that $\big\{f(x_k)\big\}_\mathcal{K} \to f(x)$. By our choice of the sequence $\{x_k\}_\mathcal{K}$ and the lower semicontinuity of g, it follows that

$$\lim_{k\to\infty,\ k\in\mathcal{K}} g\big(f(x_k)\big) = \liminf_{k\to\infty,\ k\in\mathcal{K}} g\big(f(x_k)\big) \geq g\big(f(x)\big),$$

contradicting Eq. (1.1). Hence, h is lower semicontinuous.

As an example showing that the assumption that g is monotonically nondecreasing is essential, consider the functions

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ 1 & \text{if } x > 0, \end{cases}$$

and $g(x) = -x$. Then

$$g\big(f(x)\big) = \begin{cases} 0 & \text{if } x \leq 0, \\ -1 & \text{if } x > 0, \end{cases}$$

which is not lower semicontinuous at 0.

1.4 (Convexity under Composition)

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

(a) Let $f : C \mapsto \Re$ be a convex function, and $g : \Re \mapsto \Re$ be a function that is convex and monotonically nondecreasing over a convex set that contains the set of values that f can take, $\big\{f(x) \mid x \in C\big\}$. Show that the function h defined by $h(x) = g\big(f(x)\big)$ is convex over C. In addition, if g is monotonically increasing and f is strictly convex, then h is strictly convex.

(b) Let $f = (f_1, \ldots, f_m)$, where each $f_i : C \mapsto \Re$ is a convex function, and let $g : \Re^m \mapsto \Re$ be a function that is convex and monotonically nondecreasing over a convex set that contains the set $\big\{f(x) \mid x \in C\big\}$, in the sense that for all $u, \overline{u}$ in this set such that $u \leq \overline{u}$, we have $g(u) \leq g(\overline{u})$. Show that the function h defined by $h(x) = g\big(f(x)\big)$ is convex over $C \times \cdots \times C$.

Solution: Let $x, y \in \Re^n$ and let $\alpha \in [0, 1]$. By the definitions of h and f, we have

$$\begin{aligned}
h\big(\alpha x + (1-\alpha)y\big) &= g\Big(f\big(\alpha x + (1-\alpha)y\big)\Big) \\
&= g\Big(f_1\big(\alpha x + (1-\alpha)y\big), \ldots, f_m\big(\alpha x + (1-\alpha)y\big)\Big) \\
&\leq g\big(\alpha f_1(x) + (1-\alpha) f_1(y), \ldots, \alpha f_m(x) + (1-\alpha) f_m(y)\big) \\
&= g\Big(\alpha\big(f_1(x), \ldots, f_m(x)\big) + (1-\alpha)\big(f_1(y), \ldots, f_m(y)\big)\Big) \\
&\leq \alpha g\big(f_1(x), \ldots, f_m(x)\big) + (1-\alpha) g\big(f_1(y), \ldots, f_m(y)\big) \\
&= \alpha g\big(f(x)\big) + (1-\alpha) g\big(f(y)\big) \\
&= \alpha h(x) + (1-\alpha) h(y),
\end{aligned}$$

where the first inequality follows by convexity of each f_i and monotonicity of g, while the second inequality follows by convexity of g.

If $m = 1$, g is monotonically increasing, and f is strictly convex, then the first inequality is strict whenever $x \neq y$ and $\alpha \in (0,1)$, showing that h is strictly convex.

1.5 (Examples of Convex Functions)

Show that the following functions from $\Re^n$ to $(-\infty, \infty]$ are convex:

(a)

$$f_1(x_1, \ldots, x_n) = \begin{cases} -(x_1 x_2 \cdots x_n)^{\frac{1}{n}} & \text{if } x_1 > 0, \ldots, x_n > 0, \\ \infty & \text{otherwise.} \end{cases}$$

(b) $f_2(x) = \ln\big(e^{x_1} + \cdots + e^{x_n}\big)$.

(c) $f_3(x) = \|x\|^p$ with $p \geq 1$.

(d) $f_4(x) = \frac{1}{f(x)}$, where f is concave and $f(x)$ is a positive number for all x.

(e) $f_5(x) = \alpha f(x) + \beta$, where $f : \Re^n \mapsto \Re$ is a convex function, and α and β are scalars such that $\alpha \geq 0$.

(f) $f_6(x) = e^{\beta x' A x}$, where A is a positive semidefinite symmetric $n \times n$ matrix and β is a positive scalar.

(g) $f_7(x) = f(Ax + b)$, where $f : \Re^m \mapsto \Re$ is a convex function, A is an $m \times n$ matrix, and b is a vector in $\Re^m$.

Solution: (a) Denote $X = \text{dom}(f_1)$. It can be seen that f_1 is twice continuously differentiable over X and its Hessian matrix is given by

$$\nabla^2 f_1(x) = \frac{f_1(x)}{n^2} \begin{bmatrix} \frac{1-n}{x_1^2} & \frac{1}{x_1 x_2} & \cdots & \frac{1}{x_1 x_n} \\ \frac{1}{x_2 x_1} & \frac{1-n}{x_2^2} & \cdots & \frac{1}{x_2 x_n} \\ & & \vdots & \\ \frac{1}{x_n x_1} & \frac{1}{x_1 x_2} & \cdots & \frac{1-n}{x_n^2} \end{bmatrix}$$

for all $x = (x_1, \ldots, x_n) \in X$. From this, direct computation shows that for all $z = (z_1, \ldots, z_n) \in \Re^n$ and $x = (x_1, \ldots, x_n) \in X$, we have

$$z' \nabla^2 f_1(x) z = \frac{f_1(x)}{n^2} \left(\left(\sum_{i=1}^n \frac{z_i}{x_i} \right)^2 - n \sum_{i=1}^n \left(\frac{z_i}{x_i} \right)^2 \right).$$

Note that this quadratic form is nonnegative for all $z \in \Re^n$ and $x \in X$, since $f_1(x) < 0$, and for any real numbers $\alpha_1, \ldots, \alpha_n$, we have

$$(\alpha_1 + \cdots + \alpha_n)^2 \leq n(\alpha_1^2 + \cdots + \alpha_n^2),$$

in view of the fact that $2\alpha_j \alpha_k \leq \alpha_j^2 + \alpha_k^2$. Hence, $\nabla^2 f_1(x)$ is positive semidefinite for all $x \in X$, and it follows from Prop. 1.2.6(a) that f_1 is convex.

(b) We show that the Hessian of f_2 is positive semidefinite at all $x \in \Re^n$. Let $\beta(x) = e^{x_1} + \cdots + e^{x_n}$. Then a straightforward calculation yields

$$z'\nabla^2 f_2(x)z = \frac{1}{\beta(x)^2} \sum_{i=1}^{n} \sum_{j=1}^{n} e^{(x_i + x_j)} (z_i - z_j)^2 \geq 0, \qquad \forall\ z \in \Re^n.$$

Hence by Prop. 1.2.6(a), f_2 is convex.

(c) The function $f_3(x) = \|x\|^p$ can be viewed as a composition $g\big(f(x)\big)$ of the scalar function $g(t) = t^p$ with $p \geq 1$ and the function $f(x) = \|x\|$. In this case, g is convex and monotonically increasing over the nonnegative axis, the set of values that f can take, while f is convex over $\Re^n$ (since any vector norm is convex, see the discussion preceding Prop. 1.2.4). Using Exercise 1.4, it follows that the function $f_3(x) = \|x\|^p$ is convex over $\Re^n$.

(d) The function $f_4(x) = \frac{1}{f(x)}$ can be viewed as a composition $g\big(h(x)\big)$ of the function $g(t) = -\frac{1}{t}$ for $t < 0$ and the function $h(x) = -f(x)$ for $x \in \Re^n$. In this case, the g is convex and monotonically increasing in the set $\{t \mid t < 0\}$, while h is convex over $\Re^n$. Using Exercise 1.4, it follows that the function $f_4(x) = \frac{1}{f(x)}$ is convex over $\Re^n$.

(e) The function $f_5(x) = \alpha f(x) + \beta$ can be viewed as a composition $g\big(f(x)\big)$ of the function $g(t) = \alpha t + \beta$, where $t \in \Re$, and the function $f(x)$ for $x \in \Re^n$. In this case, g is convex and monotonically increasing over $\Re$ (since $\alpha \geq 0$), while f is convex over $\Re^n$. Using Exercise 1.4, it follows that f_5 is convex over $\Re^n$.

(f) The function $f_6(x) = e^{\beta x' Ax}$ can be viewed as a composition $g\big(f(x)\big)$ of the function $g(t) = e^{\beta t}$ for $t \in \Re$ and the function $f(x) = x'Ax$ for $x \in \Re^n$. In this case, g is convex and monotonically increasing over $\Re$, while f is convex over $\Re^n$ (since A is positive semidefinite). Using Exercise 1.4, it follows that f_6 is convex over $\Re^n$.

(g) This part is straightforward using the definition of a convex function.

1.6 (Ascent/Descent Behavior of a Convex Function)

Let $f : \Re \mapsto \Re$ be a convex function.

(a) *(Monotropic Property)* Use the definition of convexity to show that f is "turning upwards" in the sense that if x_1, x_2, x_3 are three scalars such that $x_1 < x_2 < x_3$, then

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

(b) Use part (a) to show that there are four possibilities as x increases to ∞: (1) $f(x)$ decreases monotonically to $-\infty$, (2) $f(x)$ decreases monotonically to a finite value, (3) $f(x)$ reaches some value and stays at that value, (4) $f(x)$ increases monotonically to ∞ when $x \geq \overline{x}$ for some $\overline{x} \in \Re$.

Solution: (a) Let x_1, x_2, x_3 be three scalars such that $x_1 < x_2 < x_3$. Then we can write x_2 as a convex combination of x_1 and x_3 as follows

$$x_2 = \frac{x_3 - x_2}{x_3 - x_1} x_1 + \frac{x_2 - x_1}{x_3 - x_1} x_3,$$

so that by convexity of f, we obtain

$$f(x_2) \leq \frac{x_3 - x_2}{x_3 - x_1} f(x_1) + \frac{x_2 - x_1}{x_3 - x_1} f(x_3).$$

This relation and the fact

$$f(x_2) = \frac{x_3 - x_2}{x_3 - x_1} f(x_2) + \frac{x_2 - x_1}{x_3 - x_1} f(x_2),$$

imply that

$$\frac{x_3 - x_2}{x_3 - x_1}\big(f(x_2) - f(x_1)\big) \leq \frac{x_2 - x_1}{x_3 - x_1}\big(f(x_3) - f(x_2)\big).$$

By multiplying the preceding relation with $x_3 - x_1$ and by dividing it with $(x_3 - x_2)(x_2 - x_1)$, we obtain

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

(b) Let $\{x_k\}$ be an increasing scalar sequence, i.e., $x_1 < x_2 < x_3 < \cdots$. Then according to part (a), we have for all k

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2} \leq \cdots \leq \frac{f(x_{k+1}) - f(x_k)}{x_{k+1} - x_k}. \tag{1.2}$$

Since $\big(f(x_k) - f(x_{k-1})\big)/(x_k - x_{k-1})$ is monotonically nondecreasing, we have

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \to \gamma, \tag{1.3}$$

where γ is either a real number or ∞. Furthermore,

$$\frac{f(x_{k+1}) - f(x_k)}{x_{k+1} - x_k} \leq \gamma, \qquad \forall\, k. \tag{1.4}$$

We now show that γ is independent of the sequence $\{x_k\}$. Let $\{y_j\}$ be any increasing scalar sequence. For each j, choose x_{k_j} such that $y_j < x_{k_j}$ and $x_{k_1} < x_{k_2} < \cdots < x_{k_j}$, so that we have $y_j < y_{j+1} < x_{k_{j+1}} < x_{k_{j+2}}$. By part (a), it follows that

$$\frac{f(y_{j+1}) - f(y_j)}{y_{j+1} - y_j} \leq \frac{f(x_{k_{j+2}}) - f(x_{k_{j+1}})}{x_{k_{j+2}} - x_{k_{j+1}}},$$

and letting $j \to \infty$ yields

$$\lim_{j \to \infty} \frac{f(y_{j+1}) - f(y_j)}{y_{j+1} - y_j} \le \gamma.$$

Similarly, by exchanging the roles of $\{x_k\}$ and $\{y_j\}$, we can show that

$$\lim_{j \to \infty} \frac{f(y_{j+1}) - f(y_j)}{y_{j+1} - y_j} \ge \gamma.$$

Thus the limit in Eq. (1.3) is independent of the choice for $\{x_k\}$, and Eqs. (1.2) and (1.4) hold for any increasing scalar sequence $\{x_k\}$.

We consider separately each of the three possibilities $\gamma < 0, \gamma = 0$, and $\gamma > 0$. First, suppose that $\gamma < 0$, and let $\{x_k\}$ be any increasing sequence. By using Eq. (1.4), we obtain

$$\begin{aligned} f(x_k) &= \sum_{j=1}^{k-1} \frac{f(x_{j+1}) - f(x_j)}{x_{j+1} - x_j}(x_{j+1} - x_j) + f(x_1) \\ &\le \sum_{j=1}^{k-1} \gamma(x_{j+1} - x_j) + f(x_1) \\ &= \gamma(x_k - x_1) + f(x_1), \end{aligned}$$

and since $\gamma < 0$ and $x_k \to \infty$, it follows that $f(x_k) \to -\infty$. To show that f decreases monotonically, pick any x and y with $x < y$, and consider the sequence $x_1 = x$, $x_2 = y$, and $x_k = y + k$ for all $k \ge 3$. By using Eq. (1.4) with $k = 1$, we have

$$\frac{f(y) - f(x)}{y - x} \le \gamma < 0,$$

so that $f(y) - f(x) < 0$. Hence f decreases monotonically to $-\infty$, corresponding to case (1).

Suppose now that $\gamma = 0$, and let $\{x_k\}$ be any increasing sequence. Then, by Eq. (1.4), we have $f(x_{k+1}) - f(x_k) \le 0$ for all k. If $f(x_{k+1}) - f(x_k) < 0$ for all k, then f decreases monotonically. To show this, pick any x and y with $x < y$, and consider a new sequence given by $y_1 = x$, $y_2 = y$, and $y_k = x_{K+k-3}$ for all $k \ge 3$, where K is large enough so that $y < x_K$. By using Eqs. (1.2) and (1.4) with $\{y_k\}$, we have

$$\frac{f(y) - f(x)}{y - x} \le \frac{f(x_{K+1}) - f(x_K)}{x_{K+1} - x_K} < 0,$$

implying that $f(y) - f(x) < 0$. Hence f decreases monotonically, and it may decrease to $-\infty$ or to a finite value, corresponding to cases (1) or (2), respectively.

If for some K we have $f(x_{K+1}) - f(x_K) = 0$, then by Eqs. (1.2) and (1.4) where $\gamma = 0$, we obtain $f(x_k) = f(x_K)$ for all $k \ge K$. To show that f stays at the value $f(x_K)$ for all $x \ge x_K$, choose any x such that $x > x_K$, and define $\{y_k\}$

as $y_1 = x_K$, $y_2 = x$, and $y_k = x_{N+k-3}$ for all $k \geq 3$, where N is large enough so that $x < x_N$. By using Eqs. (1.2) and (1.4) with $\{y_k\}$, we have

$$\frac{f(x) - f(x_K)}{x - x_K} \leq \frac{f(x_N) - f(x)}{x_N - x} \leq 0,$$

so that $f(x) \leq f(x_K)$ and $f(x_N) \leq f(x)$. Since $f(x_K) = f(x_N)$, we have $f(x) = f(x_K)$. Hence $f(x) = f(x_K)$ for all $x \geq x_K$, corresponding to case (3).

Finally, suppose that $\gamma > 0$, and let $\{x_k\}$ be any increasing sequence. Since $\big(f(x_k) - f(x_{k-1})\big)/(x_k - x_{k-1})$ is nondecreasing and tends to γ [cf. Eqs. (1.3) and (1.4)], there is a positive integer K and a positive scalar ϵ with $\epsilon < \gamma$ such that

$$\epsilon \leq \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}, \qquad \forall\, k \geq K. \tag{1.5}$$

Therefore, for all $k > K$

$$f(x_k) = \sum_{j=K}^{k-1} \frac{f(x_{j+1}) - f(x_j)}{x_{j+1} - x_j}(x_{j+1} - x_j) + f(x_K) \geq \epsilon(x_k - x_K) + f(x_K),$$

implying that $f(x_k) \to \infty$. To show that $f(x)$ increases monotonically to ∞ for all $x \geq x_K$, pick any $x < y$ satisfying $x_K < x < y$, and consider a sequence given by $y_1 = x_K$, $y_2 = x$, $y_3 = y$, and $y_k = x_{N+k-4}$ for $k \geq 4$, where N is large enough so that $y < x_N$. By using Eq. (1.5) with $\{y_k\}$, we have

$$\epsilon \leq \frac{f(y) - f(x)}{y - x}.$$

Thus $f(x)$ increases monotonically to ∞ for all $x \geq x_K$, corresponding to case (4) with $\overline{x} = x_K$.

1.7 (Characterization of Differentiable Convex Functions)

Let $f : \Re^n \mapsto \Re$ be a differentiable function. Show that f is convex over a nonempty convex set C if and only if

$$\big(\nabla f(x) - \nabla f(y)\big)'(x - y) \geq 0, \qquad \forall\, x, y \in C.$$

Note: The condition above says that the function f, restricted to the line segment connecting x and y, has monotonically nondecreasing gradient.

Solution: If f is convex, then by Prop. 1.2.5(a), we have

$$f(y) \geq f(x) + \nabla f(x)'(y - x), \qquad \forall\, x, y \in C.$$

By exchanging the roles of x and y in this relation, we obtain

$$f(x) \geq f(y) + \nabla f(y)'(x - y), \qquad \forall\, x, y \in C,$$

and by adding the preceding two inequalities, it follows that

$$\big(\nabla f(y) - \nabla f(x)\big)'(x - y) \geq 0. \tag{1.6}$$

Conversely, let Eq. (1.6) hold, and let x and y be two points in C. Define the function $h : \Re \mapsto \Re$ by

$$h(t) = f\big(x + t(y - x)\big).$$

Consider some $t, t' \in [0, 1]$ such that $t < t'$. By convexity of C, we have that $x + t(y - x)$ and $x + t'(y - x)$ belong to C. Using the chain rule and Eq. (1.6), we have

$$\begin{aligned} &\left(\frac{dh(t')}{dt} - \frac{dh(t)}{dt}\right)(t' - t) \\ &\qquad = \Big(\nabla f\big(x + t'(y - x)\big) - \nabla f\big(x + t(y - x)\big)\Big)'(y - x)(t' - t) \\ &\qquad \geq 0. \end{aligned}$$

Thus, dh/dt is nondecreasing on $[0, 1]$ and for any $t \in (0, 1)$, we have

$$\frac{h(t) - h(0)}{t} = \frac{1}{t}\int_0^t \frac{dh(\tau)}{d\tau}\,d\tau \leq h(t) \leq \frac{1}{1 - t}\int_t^1 \frac{dh(\tau)}{d\tau}\,d\tau = \frac{h(1) - h(t)}{1 - t}.$$

Equivalently,

$$th(1) + (1 - t)h(0) \geq h(t),$$

and from the definition of h, we obtain

$$tf(y) + (1 - t)f(x) \geq f\big(ty + (1 - t)x\big).$$

Since this inequality has been proved for arbitrary $t \in [0, 1]$ and $x, y \in C$, we conclude that f is convex.

1.8 (Characterization of Twice Continuously Differentiable Convex Functions)

Let C be a nonempty convex subset of $\Re^n$ and let $f : \Re^n \mapsto \Re$ be twice continuously differentiable over $\Re^n$. Let S be the subspace that is parallel to the affine hull of C. Show that f is convex over C if and only if $y'\nabla^2 f(x)y \geq 0$ for all $x \in C$ and $y \in S$. [In particular, when C has nonempty interior, f is convex over C if and only if $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$.]

Solution: Suppose that $f : \Re^n \mapsto \Re$ is convex over C. We first show that for all $x \in \text{ri}(C)$ and $y \in S$, we have $y'\nabla^2 f(x)y \geq 0$. Assume to arrive at a contradiction, that there exists some $\overline{x} \in \text{ri}(C)$ such that for some $y \in S$, we have

$$y'\nabla^2 f(\overline{x})y < 0.$$

Without loss of generality, we may assume that $\|y\| = 1$. Using the continuity of $\nabla^2 f$, we see that there is an open ball $B(\overline{x}, \epsilon)$ centered at $\bar{x}$ with radius ϵ such that $B(\overline{x}, \epsilon) \cap \mathrm{aff}(C) \subset C$ [since $\overline{x} \in \mathrm{ri}(C)$], and

$$y'\nabla^2 f(x)y < 0, \qquad \forall\ x \in B(\overline{x}, \epsilon). \tag{1.7}$$

By Prop. 1.1.13(a), for all positive scalars α with $\alpha < \epsilon$, we have

$$f(\bar{x} + \alpha y) = f(\bar{x}) + \alpha \nabla f(\bar{x})'y + \frac{1}{2} y'\nabla^2 f(\bar{x} + \bar{\alpha} y)y,$$

for some $\bar{\alpha} \in [0, \alpha]$. Furthermore, $\|(\overline{x} + \overline{\alpha}y) - \overline{x}\| \leq \epsilon$ [since $\|y\| = 1$ and $\bar{\alpha} < \epsilon$]. Hence, from Eq. (1.7), it follows that

$$f(\bar{x} + \alpha y) < f(\bar{x}) + \alpha \nabla f(\bar{x})'y, \qquad \forall\ \alpha \in [0, \epsilon).$$

On the other hand, by the choice of ϵ and the assumption that $y \in S$, the vectors $\bar{x} + \alpha y$ are in C for all α with $\alpha \in [0, \epsilon)$, which is a contradiction in view of the convexity of f over C. Hence, we have $y'\nabla^2 f(x)y \geq 0$ for all $y \in S$ and all $x \in \mathrm{ri}(C)$.

Next, let $\overline{x}$ be a point in C that is not in the relative interior of C. Then, by the Line Segment Principle, there is a sequence $\{x_k\} \subset \mathrm{ri}(C)$ such that $x_k \to \overline{x}$. As seen above, $y'\nabla^2 f(x_k)y \geq 0$ for all $y \in S$ and all k, which together with the continuity of $\nabla^2 f$ implies that

$$y'\nabla^2 f(\overline{x})y = \lim_{k \to \infty} y'\nabla^2 f(x_k)y \geq 0, \qquad \forall\ y \in S.$$

It follows that $y'\nabla^2 f(x)y \geq 0$ for all $x \in C$ and $y \in S$.

Conversely, assume that $y'\nabla^2 f(x)y \geq 0$ for all $x \in C$ and $y \in S$. By Prop. 1.1.13(a), for all $x, z \in C$ we have

$$f(z) = f(x) + (z - x)'\nabla f(x) + \tfrac{1}{2}(z - x)'\nabla^2 f\big(x + \alpha(z - x)\big)(z - x)$$

for some $\alpha \in [0, 1]$. Since $x, z \in C$, we have that $(z - x) \in S$, and using the convexity of C and our assumption, it follows that

$$f(z) \geq f(x) + (z - x)'\nabla f(x), \qquad \forall\ x, z \in C.$$

From Prop. 1.2.5(a), we conclude that f is convex over C.

1.9 (Strong Convexity)

Let $f : \Re^n \mapsto \Re$ be a differentiable function. We say that f *is strongly convex with coefficient* α if

$$\big(\nabla f(x) - \nabla f(y)\big)'(x - y) \geq \alpha \|x - y\|^2, \qquad \forall\ x, y \in \Re^n, \tag{1.8}$$

where α is some positive scalar.

(a) Show that if f is strongly convex with coefficient α, then f is strictly convex.

(b) Assume that f is twice continuously differentiable. Show that strong convexity of f with coefficient α is equivalent to the positive semidefiniteness of $\nabla^2 f(x) - \alpha I$ for every $x \in \Re^n$, where I is the identity matrix.

Solution: (a) Fix some $x, y \in \Re^n$ such that $x \neq y$, and define the function $h : \Re \mapsto \Re$ by $h(t) = f\big(x + t(y - x)\big)$. Consider scalars t and s such that $t < s$. Using the chain rule and the equation

$$\big(\nabla f(x) - \nabla f(y)\big)'(x - y) \geq \alpha \|x - y\|^2, \qquad \forall\, x, y \in \Re^n, \tag{1.9}$$

for some $\alpha > 0$, we have

$$\begin{aligned}
&\left(\frac{dh(s)}{dt} - \frac{dh(t)}{dt}\right)(s - t) \\
&\quad = \Big(\nabla f\big(x + s(y - x)\big) - \nabla f\big(x + t(y - x)\big)\Big)'(y - x)(s - t) \\
&\quad \geq \alpha (s - t)^2 \|x - y\|^2 > 0.
\end{aligned}$$

Thus, dh/dt is strictly increasing and for any $t \in (0, 1)$, we have

$$\frac{h(t) - h(0)}{t} = \frac{1}{t} \int_0^t \frac{dh(\tau)}{d\tau}\, d\tau < \frac{1}{1 - t} \int_t^1 \frac{dh(\tau)}{d\tau}\, d\tau = \frac{h(1) - h(t)}{1 - t}.$$

Equivalently, $th(1) + (1 - t)h(0) > h(t)$. The definition of h yields $tf(y) + (1 - t)f(x) > f\big(ty + (1 - t)x\big)$. Since this inequality has been proved for arbitrary $t \in (0, 1)$ and $x \neq y$, we conclude that f is strictly convex.

(b) Suppose now that f is twice continuously differentiable and Eq. (1.9) holds. Let c be a scalar. We use Prop. 1.1.13(b) twice to obtain

$$f(x + cy) = f(x) + cy'\nabla f(x) + \frac{c^2}{2} y'\nabla^2 f(x + tcy)y,$$

and

$$f(x) = f(x + cy) - cy'\nabla f(x + cy) + \frac{c^2}{2} y'\nabla^2 f(x + scy)y,$$

for some t and s belonging to $[0, 1]$. Adding these two equations and using Eq. (1.9), we obtain

$$\frac{c^2}{2} y'\big(\nabla^2 f(x + scy) + \nabla^2 f(x + tcy)\big)y = \big(\nabla f(x + cy) - \nabla f(x)\big)'(cy) \geq \alpha c^2 \|y\|^2.$$

We divide both sides by c^2 and then take the limit as $c \to 0$ to conclude that $y'\nabla^2 f(x)y \geq \alpha \|y\|^2$. Since this inequality is valid for every $y \in \Re^n$, it follows that $\nabla^2 f(x) - \alpha I$ is positive semidefinite.

For the converse, assume that $\nabla^2 f(x) - \alpha I$ is positive semidefinite for all $x \in \Re^n$. Consider the function $g : \Re \mapsto \Re$ defined by

$$g(t) = \nabla f\big(tx + (1 - t)y\big)'(x - y).$$

Using the Mean Value Theorem (Prop. 1.1.12), we have

$$\big(\nabla f(x) - \nabla f(y)\big)'(x - y) = g(1) - g(0) = \frac{dg(t)}{dt}$$

for some $t \in [0, 1]$. On the other hand,

$$\frac{dg(t)}{dt} = (x - y)'\nabla^2 f\big(tx + (1 - t)y\big)(x - y) \geq \alpha\|x - y\|^2,$$

where the last inequality holds because $\nabla^2 f\big(tx+(1-t)y\big) - \alpha I$ is positive semidefinite. Combining the last two relations, it follows that f is strongly convex with coefficient α.

1.10 (Posynomials)

A *posynomial* is a function of positive scalar variables $y_1, \ldots, y_n$ of the form

$$g(y_1, \ldots, y_n) = \sum_{i=1}^{m} \beta_i y_1^{a_{i1}} \cdots y_n^{a_{in}},$$

where a_{ij} and β_i are scalars, such that $\beta_i > 0$ for all i. Show the following:

(a) A posynomial need not be convex.

(b) By a logarithmic change of variables, where we set

$$f(x) = \ln\big(g(y_1, \ldots, y_n)\big), \qquad b_i = \ln \beta_i,\ \forall\ i, \qquad x_j = \ln y_j,\ \forall\ j,$$

we obtain a convex function

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

where $\exp(z) = e^{z_1} + \cdots + e^{z_m}$ for all $z \in \Re^m$, A is an $m \times n$ matrix with entries a_{ij}, and $b \in \Re^m$ is a vector with components b_i.

(c) Every function $g : \Re^n \mapsto \Re$ of the form

$$g(y) = g_1(y)^{\gamma_1} \cdots g_r(y)^{\gamma_r},$$

where g_k is a posynomial and $\gamma_k > 0$ for all k, can be transformed by a logarithmic change of variables into a convex function f given by

$$f(x) = \sum_{k=1}^{r} \gamma_k \ln \exp(A_k x + b_k),$$

with the matrix A_k and the vector b_k being associated with the posynomial g_k for each k.

Solution: (a) Consider the following posynomial for which we have $n = m = 1$ and $\beta = \frac{1}{2}$,

$$g(y) = y^{\frac{1}{2}}, \qquad \forall\, y > 0.$$

This function is not convex.

(b) Consider the following change of variables, where we set

$$f(x) = \ln\big(g(y_1, \ldots, y_n)\big), \qquad b_i = \ln \beta_i,\ \forall\, i, \qquad x_j = \ln y_j,\ \forall\, j.$$

With this change of variables, $f(x)$ can be written as

$$f(x) = \ln\left(\sum_{i=1}^{m} e^{b_i + a_{i1}x_1 + \cdots + a_{in}x_n}\right).$$

Note that $f(x)$ can also be represented as

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

where $\ln \exp(z) = \ln\big(e^{z_1} + \cdots + e^{z_m}\big)$ for all $z \in \Re^m$, A is an $m \times n$ matrix with entries a_{ij}, and $b \in \Re^m$ is a vector with components b_i. Let $f_2(z) = \ln(e^{z_1} + \cdots + e^{z_m})$. This function is convex by Exercise 1.5(b). With this identification, $f(x)$ can be viewed as the composition $f(x) = f_2(Ax + b)$, which is convex by Exercise 1.5(g).

(c) Consider a function $g : \Re^n \mapsto \Re$ of the form

$$g(y) = g_1(y)^{\gamma_1} \cdots g_r(y)^{\gamma_r},$$

where g_k is a posynomial and $\gamma_k > 0$ for all k. Using a change of variables similar to part (b), we see that we can represent the function $f(x) = \ln g(y)$ as

$$f(x) = \sum_{k=1}^{r} \gamma_k \ln \exp(A_k x + b_k),$$

with the matrix A_k and the vector b_k being associated with the posynomial g_k for each k. Since $f(x)$ is a linear combination of convex functions with nonnegative coefficients [part (b)], it follows from Prop. 1.2.4(a) that $f(x)$ is convex.

1.11 (Arithmetic-Geometric Mean Inequality)

Show that if $\alpha_1, \ldots, \alpha_n$ are positive scalars with $\sum_{i=1}^{n} \alpha_i = 1$, then for every set of positive scalars $x_1, \ldots, x_n$, we have

$$x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \leq \alpha_1 x_1 + a_2 x_2 + \cdots + \alpha_n x_n,$$

with equality if and only if $x_1 = x_2 = \cdots = x_n$. *Hint*: Show that $-\ln x$ is a strictly convex function on $(0, \infty)$.

Solution: Consider the function $f(x) = -\ln(x)$. Since $\nabla^2 f(x) = 1/x^2 > 0$ for all $x > 0$, the function $-\ln(x)$ is strictly convex over $(0, \infty)$. Therefore, for all positive scalars $x_1, \ldots, x_n$ and $\alpha_1, \ldots \alpha_n$ with $\sum_{i=1}^n \alpha_i = 1$, we have

$$-\ln(\alpha_1 x_1 + \cdots + \alpha_n x_n) \le -\alpha_1 \ln(x_1) - \cdots - \alpha_n \ln(x_n),$$

which is equivalent to

$$e^{\ln(\alpha_1 x_1 + \cdots + \alpha_n x_n)} \ge e^{\alpha_1 \ln(x_1) + \cdots + \alpha_n \ln(x_n)} = e^{\alpha_1 \ln(x_1)} \cdots e^{\alpha_n \ln(x_n)},$$

or

$$\alpha_1 x_1 + \cdots + \alpha_n x_n \ge x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

as desired. Since $-\ln(x)$ is strictly convex, the above inequality is satisfied with equality if and only if the scalars $x_1, \ldots, x_n$ are all equal.

1.12 (Young and Holder Inequalities)

Use the result of Exercise 1.11 to verify Young's inequality

$$xy \le \frac{x^p}{p} + \frac{y^q}{q}, \qquad \forall\, x \ge 0,\ \forall\, y \ge 0,$$

where $p > 0$, $q > 0$, and

$$1/p + 1/q = 1.$$

Then, use Young's inequality to verify Holder's inequality

$$\sum_{i=1}^n |x_i y_i| \le \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \left(\sum_{i=1}^n |y_i|^q\right)^{1/q}.$$

Solution: According to Exercise 1.11, we have

$$u^{\frac{1}{p}} v^{\frac{1}{q}} \le \frac{u}{p} + \frac{v}{q}, \qquad \forall\, u > 0, \quad \forall\, v > 0,$$

where $1/p + 1/q = 1$, $p > 0$, and $q > 0$. The above relation also holds if $u = 0$ or $v = 0$. By setting $u = x^p$ and $v = y^q$, we obtain Young's inequality

$$xy \le \frac{x^p}{p} + \frac{y^q}{q}, \qquad \forall\, x \ge 0, \quad \forall\, y \ge 0.$$

To show Holder's inequality, note that it holds if $x_1 = \cdots = x_n = 0$ or $y_1 = \cdots = y_n = 0$. If $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ are such that $(x_1, \ldots, x_n) \ne 0$ and $(y_1, \ldots, y_n) \ne 0$, then by using

$$x = \frac{|x_i|}{\left(\sum_{j=1}^n |x_j|^p\right)^{1/p}} \quad \text{and} \quad y = \frac{|y_i|}{\left(\sum_{j=1}^n |y_j|^q\right)^{1/q}}$$

in Young's inequality, we have for all $i = 1, \ldots, n$,

$$\frac{|x_i|}{\left(\sum_{j=1}^n |x_j|^p\right)^{1/p}} \frac{|y_i|}{\left(\sum_{j=1}^n |y_j|^q\right)^{1/q}} \leq \frac{|x_i|^p}{p\left(\sum_{j=1}^n |x_j|^p\right)} + \frac{|y_i|^q}{q\left(\sum_{j=1}^n |y_j|^q\right)}.$$

By adding these inequalities over $i = 1, \ldots, n$, we obtain

$$\frac{\sum_{i=1}^n |x_i| \cdot |y_i|}{\left(\sum_{j=1}^n |x_j|^p\right)^{1/p} \left(\sum_{j=1}^n |y_j|^q\right)^{1/q}} \leq \frac{1}{p} + \frac{1}{q} = 1,$$

which implies Holder's inequality.

1.13

Let C be a nonempty convex set in $\Re^{n+1}$, and let $f : \Re^n \mapsto [-\infty, \infty]$ be the function defined by

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

Show that f is convex.

Solution: Let (x, w) and (y, v) be two vectors in $\text{epi}(f)$. Then $f(x) \leq w$ and $f(y) \leq v$, implying that there exist sequences $\big\{(x, \overline{w}_k)\big\} \subset C$ and $\big\{(y, \overline{v}_k)\big\} \subset C$ such that for all k,

$$\overline{w}_k \leq w + \frac{1}{k}, \qquad \overline{v}_k \leq v + \frac{1}{k}.$$

By the convexity of C, we have for all $\alpha \in [0, 1]$ and all k,

$$\big(\alpha x + (1 - \alpha y),\ \alpha\overline{w}_k + (1 - \alpha)\overline{v}_k\big) \in C,$$

so that for all k,

$$f\big(\alpha x + (1 - \alpha)y\big) \leq \alpha\overline{w}_k + (1 - \alpha)\overline{v}_k \leq \alpha w + (1 - \alpha)v + \frac{1}{k}.$$

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

$$f\big(\alpha x + (1 - \alpha)y\big) \leq \alpha w + (1 - \alpha)v,$$

so that $\alpha(x, w) + (1 - \alpha)(y, v) \in \text{epi}(f)$. Hence, $\text{epi}(f)$ is convex, implying that f is convex.

1.14

Show that the convex hull of a nonempty set coincides with the set of all convex combinations of its elements.

Solution: The elements of X belong to $\text{conv}(X)$, so all their convex combinations belong to $\text{conv}(X)$ since $\text{conv}(X)$ is a convex set. On the other hand, consider any two convex combinations of elements of X, $x = \lambda_1 x_1 + \cdots + \lambda_m x_m$ and $y = \mu_1 y_1 + \cdots + \mu_r y_r$, where $x_i \in X$ and $y_j \in X$. The vector

$$(1-\alpha)x + \alpha y = (1-\alpha)\left(\lambda_1 x_1 + \cdots + \lambda_m x_m\right) + \alpha\left(\mu_1 y_1 + \cdots + \mu_r y_r\right),$$

where $0 \leq \alpha \leq 1$, is another convex combination of elements of X.

Thus, the set of convex combinations of elements of X is itself a convex set, which contains X, and is contained in $\text{conv}(X)$. Hence it must coincide with $\text{conv}(X)$, which by definition is the intersection of all convex sets containing X.

1.15

Let C be a nonempty convex subset of $\Re^n$. Show that

$$\text{cone}(C) = \cup_{x \in C}\{\gamma x \mid \gamma \geq 0\}.$$

Solution: Let $y \in \text{cone}(C)$. If $y = 0$, then $y \in \cup_{x \in C}\{\gamma x \mid \gamma \geq 0\}$. If $y \neq 0$, then by definition of $\text{cone}(C)$, we have

$$y = \sum_{i=1}^{m} \lambda_i x_i,$$

for some positive integer m, nonnegative scalars λ_i, and vectors $x_i \in C$. Since $y \neq 0$, we cannot have all λ_i equal to zero, implying that $\sum_{i=1}^{m} \lambda_i > 0$. Because $x_i \in C$ for all i and C is convex, the vector

$$x = \sum_{i=1}^{m} \frac{\lambda_i}{\sum_{j=1}^{m} \lambda_j} x_i$$

belongs to C. For this vector, we have

$$y = \left(\sum_{i=1}^{m} \lambda_i\right) x,$$

with $\sum_{i=1}^{m} \lambda_i > 0$, implying that $y \in \cup_{x \in C}\big\{\gamma x \mid \gamma \geq 0\big\}$ and showing that

$$\text{cone}(C) \subset \cup_{x \in C}\{\gamma x \mid \gamma \geq 0\}.$$

The reverse inclusion follows from the definition of $\text{cone}(C)$.

1.16 (Convex Cones)

Show that:

(a) For any collection of vectors $\{a_i \mid i \in I\}$, the set $C = \{x \mid a_i'x \leq 0,\, i \in I\}$ is a closed convex cone.

(b) A cone C is convex if and only if $C + C \subset C$.

(c) For any two convex cones C_1 and C_2 containing the origin, we have

$$C_1 + C_2 = \text{conv}(C_1 \cup C_2),$$

$$C_1 \cap C_2 = \bigcup_{\alpha \in [0,1]} \big(\alpha C_1 \cap (1-\alpha) C_2\big).$$

Solution: (a) Let $x \in C$ and let λ be a positive scalar. Then

$$a_i'(\lambda x) = \lambda a_i'x \leq 0, \qquad \forall\, i \in I,$$

showing that $\lambda x \in C$ and that C is a cone. Let $x, y \in C$ and let $\lambda \in [0,1]$. Then

$$a_i'\big(\lambda x + (1-\lambda)y\big) = \lambda a_i'x + (1-\lambda)a_i'y \leq 0, \qquad \forall\, i \in I,$$

showing that $\big(\lambda x + (1-\lambda)y\big) \in C$ and that C is convex. Let a sequence $\{x_k\} \subset C$ converge to some $\bar{x} \in \Re^n$. Then

$$a_i'\bar{x} = \lim_{k \to \infty} a_i'x_k \leq 0, \qquad \forall\, i \in I,$$

showing that $\bar{x} \in C$ and that C is closed.

(b) Let C be a cone such that $C + C \subset C$, and let $x, y \in C$ and $\alpha \in [0,1]$. Then since C is a cone, $\alpha x \in C$ and $(1-\alpha)y \in C$, so that $\alpha x + (1-\alpha)y \in C + C \subset C$, showing that C is convex. Conversely, let C be a convex cone and let $x, y \in C$. Then, since C is a cone, $2x \in C$ and $2y \in C$, so that by the convexity of C, $x + y = \frac{1}{2}(2x + 2y) \in C$, showing that $C + C \subset C$.

(c) First we prove that $C_1 + C_2 \subset \text{conv}(C_1 \cup C_2)$. Choose any $x \in C_1 + C_2$. Since $C_1 + C_2$ is a cone [see Exercise 1.2(c)], the vector $2x$ is in $C_1 + C_2$, so that $2x = x_1 + x_2$ for some $x_1 \in C_1$ and $x_2 \in C_2$. Therefore,

$$x = \frac{1}{2}x_1 + \frac{1}{2}x_2,$$

showing that $x \in \text{conv}(C_1 \cup C_2)$.

Next, we show that $\text{conv}(C_1 \cup C_2) \subset C_1 + C_2$. Since $0 \in C_1$ and $0 \in C_2$, it follows that

$$C_i = C_i + 0 \subset C_1 + C_2, \qquad i = 1, 2,$$

implying that

$$C_1 \cup C_2 \subset C_1 + C_2.$$

By taking the convex hull of both sides in the above inclusion and by using the convexity of $C_1 + C_2$, we obtain

$$\mathrm{conv}(C_1 \cup C_2) \subset \mathrm{conv}(C_1 + C_2) = C_1 + C_2.$$

We finally show that

$$C_1 \cap C_2 = \bigcup_{\alpha \in [0,1]} \big(\alpha C_1 \cap (1-\alpha) C_2\big).$$

We claim that for all α with $0 < \alpha < 1$, we have

$$\alpha C_1 \cap (1-\alpha) C_2 = C_1 \cap C_2.$$

Indeed, if $x \in C_1 \cap C_2$, it follows that $x \in C_1$ and $x \in C_2$. Since C_1 and C_2 are cones and $0 < \alpha < 1$, we have $x \in \alpha C_1$ and $x \in (1-\alpha)C_2$. Conversely, if $x \in \alpha C_1 \cap (1-\alpha)C_2$, we have

$$\frac{x}{\alpha} \in C_1,$$

and

$$\frac{x}{(1-\alpha)} \in C_2.$$

Since C_1 and C_2 are cones, it follows that $x \in C_1$ and $x \in C_2$, so that $x \in C_1 \cap C_2$.

If $\alpha = 0$ or $\alpha = 1$, we obtain

$$\alpha C_1 \cap (1-\alpha) C_2 = \{0\} \subset C_1 \cap C_2,$$

since C_1 and C_2 contain the origin. Thus, the result follows.

1.17

Let $\{C_i \mid i \in I\}$ be an arbitrary collection of convex sets in $\Re^n$, and let C be the convex hull of the union of the collection. Show that

$$C = \bigcup_{\overline{I} \subset I,\ \overline{I}:\ \text{finite set}} \left\{ \sum_{i \in \overline{I}} \alpha_i C_i \;\Big|\; \sum_{i \in \overline{I}} \alpha_i = 1,\ \alpha_i \geq 0,\ \forall\, i \in \overline{I} \right\},$$

i.e., the convex hull of the union of the C_i is equal to the set of all convex combinations of vectors from the C_i.

Solution: By Exercise 1.14, C is the set of all convex combinations $x = \alpha_1 y_1 + \cdots + \alpha_m y_m$, where m is a positive integer, and the vectors $y_1, \ldots, y_m$ belong to the union of the sets C_i. Actually, we can get C just by taking those combinations in which the vectors are taken from different sets C_i. Indeed, if two of the vectors, y_1 and y_2 belong to the same C_i, then the term $\alpha_1 y_1 + \alpha_2 y_2$ can be replaced by αy, where $\alpha = \alpha_1 + \alpha_2$ and

$$y = (\alpha_1/\alpha) y_1 + (\alpha_2/\alpha) y_2 \in C_i.$$

Thus, C is the union of the vector sums of the form

$$\alpha_1 C_{i_1} + \cdots + \alpha_m C_{i_m},$$

with

$$\alpha_i \geq 0, \ \forall \, i = 1, \ldots, m, \quad \sum_{i=1}^{m} \alpha_i = 1,$$

and the indices $i_1, \ldots, i_m$ are all different, proving our claim.

1.18 (Convex Hulls, Affine Hulls, and Generated Cones)

Let X be a nonempty set. Show that:

(a) X, $\mathrm{conv}(X)$, and $\mathrm{cl}(X)$ have the same affine hull.

(b) $\mathrm{cone}(X) = \mathrm{cone}\big(\mathrm{conv}(X)\big)$.

(c) $\mathrm{aff}\big(\mathrm{conv}(X)\big) \subset \mathrm{aff}\big(\mathrm{cone}(X)\big)$. Give an example where the inclusion is strict, i.e., $\mathrm{aff}\big(\mathrm{conv}(X)\big)$ is a strict subset of $\mathrm{aff}\big(\mathrm{cone}(X)\big)$.

(d) If the origin belongs to $\mathrm{conv}(X)$, then $\mathrm{aff}\big(\mathrm{conv}(X)\big) = \mathrm{aff}\big(\mathrm{cone}(X)\big)$.

Solution: (a) We first show that X and $\mathrm{cl}(X)$ have the same affine hull. Since $X \subset \mathrm{cl}(X)$, there holds

$$\mathrm{aff}(X) \subset \mathrm{aff}\big(\mathrm{cl}(X)\big).$$

Conversely, because $X \subset \mathrm{aff}(X)$ and $\mathrm{aff}(X)$ is closed, we have $\mathrm{cl}(X) \subset \mathrm{aff}(X)$, implying that

$$\mathrm{aff}\big(\mathrm{cl}(X)\big) \subset \mathrm{aff}(X).$$

We now show that X and $\mathrm{conv}(X)$ have the same affine hull. By using a translation argument if necessary, we assume without loss of generality that X contains the origin, so that both $\mathrm{aff}(X)$ and $\mathrm{aff}\big(\mathrm{conv}(X)\big)$ are subspaces. Since $X \subset \mathrm{conv}(X)$, evidently $\mathrm{aff}(X) \subset \mathrm{aff}\big(\mathrm{conv}(X)\big)$. To show the reverse inclusion, let the dimension of $\mathrm{aff}\big(\mathrm{conv}(X)\big)$ be m, and let $x_1, \ldots, x_m$ be linearly independent vectors in $\mathrm{conv}(X)$ that span $\mathrm{aff}\big(\mathrm{conv}(X)\big)$. Then every $x \in \mathrm{aff}\big(\mathrm{conv}(X)\big)$ is a linear combination of the vectors $x_1, \ldots, x_m$, i.e., there exist scalars $\beta_1, \ldots, \beta_m$ such that

$$x = \sum_{i=1}^{m} \beta_i x_i.$$

By the definition of convex hull, each x_i is a convex combination of vectors in X, so that x is a linear combination of vectors in X, implying that $x \in \mathrm{aff}(X)$. Hence, $\mathrm{aff}\big(\mathrm{conv}(X)\big) \subset \mathrm{aff}(X)$.

(b) Since $X \subset \mathrm{conv}(X)$, clearly $\mathrm{cone}(X) \subset \mathrm{cone}\big(\mathrm{conv}(X)\big)$. Conversely, let $x \in \mathrm{cone}\big(\mathrm{conv}(X)\big)$. Then x is a nonnegative combination of some vectors in

$\text{conv}(X)$, i.e., for some positive integer p, vectors $x_1, \ldots, x_p \in \text{conv}(X)$, and nonnegative scalars $\alpha_1, \ldots, \alpha_p$, we have

$$x = \sum_{i=1}^{p} \alpha_i x_i.$$

Each x_i is a convex combination of some vectors in X, so that x is a nonnegative combination of some vectors in X, implying that $x \in \text{cone}(X)$. Hence $\text{cone}\big(\text{conv}(X)\big) \subset \text{cone}(X)$.

(c) Since $\text{conv}(X)$ is the set of all convex combinations of vectors in X, and $\text{cone}(X)$ is the set of all nonnegative combinations of vectors in X, it follows that $\text{conv}(X) \subset \text{cone}(X)$. Therefore

$$\text{aff}\big(\text{conv}(X)\big) \subset \text{aff}\big(\text{cone}(X)\big).$$

As an example showing that the above inclusion can be strict, consider the set $X = \big\{(1,1)\big\}$ in $\Re^2$. Then $\text{conv}(X) = X$, so that

$$\text{aff}\big(\text{conv}(X)\big) = X = \big\{(1,1)\big\},$$

and the dimension of $\text{conv}(X)$ is zero. On the other hand, $\text{cone}(X) = \big\{(\alpha,\alpha) \mid \alpha \geq 0\big\}$, so that

$$\text{aff}\big(\text{cone}(X)\big) = \big\{(x_1,x_2) \mid x_1 = x_2\big\},$$

and the dimension of $\text{cone}(X)$ is one.

(d) In view of parts (a) and (c), it suffices to show that

$$\text{aff}\big(\text{cone}(X)\big) \subset \text{aff}\big(\text{conv}(X)\big) = \text{aff}(X).$$

It is always true that $0 \in \text{cone}(X)$, so $\text{aff}\big(\text{cone}(X)\big)$ is a subspace. Let the dimension of $\text{aff}\big(\text{cone}(X)\big)$ be m, and let $x_1, \ldots, x_m$ be linearly independent vectors in $\text{cone}(X)$ that span $\text{aff}\big(\text{cone}(X)\big)$. Since every vector in $\text{aff}\big(\text{cone}(X)\big)$ is a linear combination of $x_1, \ldots, x_m$, and since each x_i is a nonnegative combination of some vectors in X, it follows that every vector in $\text{aff}\big(\text{cone}(X)\big)$ is a linear combination of some vectors in X. In view of the assumption that $0 \in \text{conv}(X)$, the affine hull of $\text{conv}(X)$ is a subspace, which implies by part (a) that the affine hull of X is a subspace. Hence, every vector in $\text{aff}\big(\text{cone}(X)\big)$ belongs to $\text{aff}(X)$, showing that $\text{aff}\big(\text{cone}(X)\big) \subset \text{aff}(X)$.

1.19

Let $\{f_i \mid i \in I\}$ be an arbitrary collection of proper convex functions $f_i : \Re^n \mapsto (-\infty, \infty]$. Define

$$f(x) = \inf\Big\{w \mid (x,w) \in \text{conv}\big(\cup_{i\in I}\text{epi}(f_i)\big)\Big\}, \qquad x \in \Re^n.$$

Show that $f(x)$ is given by

$$f(x) = \inf\left\{ \sum_{i\in\overline{I}} \alpha_i f_i(x_i) \;\middle|\; \sum_{i\in\overline{I}} \alpha_i x_i = x,\ x_i \in \Re^n,\ \sum_{i\in\overline{I}} \alpha_i = 1,\ \alpha_i \geq 0,\ \forall\, i \in \overline{I},\ \overline{I} \subset I,\ \overline{I}:\ \text{finite} \right\}.$$

Solution: By definition, $f(x)$ is the infimum of the values of w such that $(x,w) \in C$, where C is the convex hull of the union of nonempty convex sets $\text{epi}(f_i)$. We have that $(x,w) \in C$ if and only if (x,w) can be expressed as a convex combination of the form

$$(x,w) = \sum_{i\in\overline{I}} \alpha_i (x_i, w_i) = \left(\sum_{i\in\overline{I}} \alpha_i x_i, \sum_{i\in\overline{I}} \alpha_i w_i \right),$$

where $\overline{I} \subset I$ is a finite set and $(x_i, w_i) \in \text{epi}(f_i)$ for all $i \in \overline{I}$. Thus, $f(x)$ can be expressed as

$$f(x) = \inf\left\{ \sum_{i\in\overline{I}} \alpha_i w_i \;\middle|\; (x,w) = \sum_{i\in\overline{I}} \alpha_i (x_i, w_i),\ (x_i, w_i) \in \text{epi}(f_i),\ \alpha_i \geq 0,\ \forall\, i \in \overline{I},\ \sum_{i\in\overline{I}} \alpha_i = 1 \right\}.$$

Since the set $\big\{ \big(x_i, f_i(x_i)\big) \mid x_i \in \Re^n \big\}$ is contained in $\text{epi}(f_i)$, we obtain

$$f(x) \leq \inf\left\{ \sum_{i\in\overline{I}} \alpha_i f_i(x_i) \;\middle|\; x = \sum_{i\in\overline{I}} \alpha_i x_i,\ x_i \in \Re^n,\ \alpha_i \geq 0,\ \forall\, i \in \overline{I},\ \sum_{i\in\overline{I}} \alpha_i = 1 \right\}.$$

On the other hand, by the definition of $\text{epi}(f_i)$, for each $(x_i, w_i) \in \text{epi}(f_i)$ we have $w_i \geq f_i(x_i)$, implying that

$$f(x) \geq \inf\left\{ \sum_{i\in\overline{I}} \alpha_i f_i(x_i) \;\middle|\; x = \sum_{i\in\overline{I}} \alpha_i x_i,\ x_i \in \Re^n,\ \alpha_i \geq 0,\ \forall\, i \in \overline{I},\ \sum_{i\in\overline{I}} \alpha_i = 1 \right\}.$$

By combining the last two relations, we obtain

$$f(x) = \inf\left\{ \sum_{i\in\overline{I}} \alpha_i f_i(x_i) \;\middle|\; x = \sum_{i\in\overline{I}} \alpha_i x_i,\ x_i \in \Re^n,\ \alpha_i \geq 0,\ \forall\, i \in \overline{I},\ \sum_{i\in\overline{I}} \alpha_i = 1 \right\},$$

where the infimum is taken over all representations of x as a convex combination of elements x_i such that only finitely many coefficients α_i are nonzero.

1.20 (Convexification of Nonconvex Functions)

Let X be a nonempty subset of $\Re^n$ and let $f : X \mapsto \Re$ be a function that is bounded below over X. Define the function $F : \text{conv}(X) \mapsto \Re$ by

$$F(x) = \inf\bigl\{w \mid (x, w) \in \text{conv}\bigl(\text{epi}(f)\bigr)\bigr\}.$$

Show that:

(a) F is convex over $\text{conv}(X)$ and it is given by

$$F(x) = \inf\left\{\sum_i \alpha_i f(x_i) \;\middle|\; \sum_i \alpha_i x_i = x,\ x_i \in X,\ \sum_i \alpha_i = 1,\ \alpha_i \geq 0,\ \forall\ i\right\},$$

where the infimum is taken over all representations of x as a convex combination of elements of X (i.e., with finitely many nonzero coefficients α_i).

(b)

$$\inf_{x\in\text{conv}(X)} F(x) = \inf_{x\in X} f(x).$$

(c) Every $x^* \in X$ that attains the minimum of f over X, i.e., $f(x^*) = \inf_{x\in X} f(x)$, also attains the minimum of F over $\text{conv}(X)$.

Solution: (a) Since $\text{conv}\bigl(\text{epi}(f)\bigr)$ is a convex set, it follows from Exercise 1.13 that F is convex over $\text{conv}(X)$. By Caratheodory's Theorem, it can be seen that $\text{conv}\bigl(\text{epi}(f)\bigr)$ is the set of all convex combinations of elements of $\text{epi}(f)$, so that

$$F(x) = \inf\left\{\sum_i \alpha_i w_i \;\middle|\; (x, w) = \sum_i \alpha_i (x_i, w_i),\right.$$
$$\left.(x_i, w_i) \in \text{epi}(f),\ \alpha_i \geq 0,\ \sum_i \alpha_i = 1\right\},$$

where the infimum is taken over all representations of x as a convex combination of elements of X. Since the set $\bigl\{\bigl(z, f(z)\bigr) \mid z \in X\bigr\}$ is contained in $\text{epi}(f)$, we obtain

$$F(x) \leq \inf\left\{\sum_i \alpha_i f(x_i) \;\middle|\; x = \sum_i \alpha_i x_i,\ x_i \in X,\ \alpha_i \geq 0,\ \sum_i \alpha_i = 1\right\}.$$

On the other hand, by the definition of $\text{epi}(f)$, for each $(x_i, w_i) \in \text{epi}(f)$ we have $w_i \geq f(x_i)$, implying that

$$F(x) \geq \inf\left\{\sum_i \alpha_i f(x_i) \;\middle|\; (x, w) = \sum_i \alpha_i (x_i, w_i),\right.$$
$$\left.(x_i, w_i) \in \text{epi}(f),\ \alpha_i \geq 0,\ \sum_i \alpha_i = 1\right\},$$
$$= \inf\left\{\sum_i \alpha_i f(x_i) \;\middle|\; x = \sum_i \alpha_i x_i,\ x_i \in X,\ \alpha_i \geq 0,\ \sum_i \alpha_i = 1\right\},$$

which combined with the preceding inequality implies the desired relation.

(b) By using part (a), we have for every $x \in X$

$$F(x) \leq f(x),$$

since $f(x)$ corresponds to the value of the function $\sum_i \alpha_i f(x_i)$ for a particular representation of x as a finite convex combination of elements of X, namely $x = 1 \cdot x$. Therefore, we have

$$\inf_{x \in X} F(x) \leq \inf_{x \in X} f(x),$$

and since $X \subset \mathrm{conv}(X)$, it follows that

$$\inf_{x \in \mathrm{conv}(X)} F(x) \leq \inf_{x \in X} f(x).$$

Let $f^* = \inf_{x \in X} f(x)$. If $\inf_{x \in \mathrm{conv}(X)} F(x) < f^*$, then there exists $z \in \mathrm{conv}(X)$ with $F(z) < f^*$. According to part (a), there exist points $x_i \in X$ and nonnegative scalars α_i with $\sum_i \alpha_i = 1$ such that $z = \sum_i \alpha_i x_i$ and

$$F(z) \leq \sum_i \alpha_i f(x_i) < f^*,$$

implying that

$$\sum_i \alpha_i \big(f(x_i) - f^* \big) < 0.$$

Since each α_i is nonnegative, for this inequality to hold, we must have $f(x_i) - f^* < 0$ for some i, but this cannot be true because $x_i \in X$ and f^* is the optimal value of f over X. Therefore

$$\inf_{x \in \mathrm{conv}(X)} F(x) = \inf_{x \in X} f(x).$$

(c) If $x^* \in X$ is a global minimum of f over X, then x^* also belongs to $\mathrm{conv}(X)$, and by part (b)

$$\inf_{x \in \mathrm{conv}(X)} F(x) = \inf_{x \in X} f(x) = f(x^*) \geq F(x^*),$$

showing that x^* is also a global minimum of F over $\mathrm{conv}(X)$.

1.21 (Minimization of Linear Functions)

Show that minimization of a linear function over a set is equivalent to minimization over its convex hull. In particular, if $X \subset \Re^n$ and $c \in \Re^n$, then

$$\inf_{x \in \text{conv}(X)} c'x = \inf_{x \in X} c'x.$$

Furthermore, the infimum in the left-hand side above is attained if and only if the infimum in the right-hand side is attained.

Solution: Let $f : X \mapsto \Re$ be the function $f(x) = c'x$, and define

$$F(x) = \inf\bigl\{w \mid (x, w) \in \text{conv}\bigl(\text{epi}(f)\bigr)\bigr\},$$

as in Exercise 1.20. According to this exercise, we have

$$\inf_{x \in \text{conv}(X)} F(x) = \inf_{x \in X} f(x),$$

and

$$\begin{aligned} F(x) &= \inf\left\{\sum_i \alpha_i c' x_i \;\Big|\; \sum_i \alpha_i x_i = x,\, x_i \in X,\, \sum_i \alpha_i = 1,\, \alpha_i \geq 0\right\} \\ &= \inf\left\{c'\left(\sum_i \alpha_i x_i\right) \;\Big|\; \sum_i \alpha_i x_i = x,\, x_i \in X,\, \sum_i \alpha_i = 1,\, \alpha_i \geq 0\right\} \\ &= c'x, \end{aligned}$$

showing that

$$\inf_{x \in \text{conv}(X)} c'x = \inf_{x \in X} c'x.$$

According to Exercise 1.20(c), if $\inf_{x \in X} c'x$ is attained at some $x^* \in X$, then $\inf_{x \in \text{conv}(X)} c'x$ is also attained at x^*. Suppose now that $\inf_{x \in \text{conv}(X)} c'x$ is attained at some $x^* \in \text{conv}(X)$, i.e., there is $x^* \in \text{conv}(X)$ such that

$$\inf_{x \in \text{conv}(X)} c'x = c'x^*.$$

Then, by Caratheodory's Theorem, there exist vectors $x_1, \ldots, x_{n+1}$ in X and nonnegative scalars $\alpha_1, \ldots, \alpha_{n+1}$ with $\sum_{i=1}^{n+1} \alpha_i = 1$ such that $x^* = \sum_{i=1}^{n+1} \alpha_i x_i$, implying that

$$c'x^* = \sum_{i=1}^{n+1} \alpha_i c' x_i.$$

Since $x_i \in X \subset \text{conv}(X)$ for all i and $c'x \geq c'x^*$ for all $x \in \text{conv}(X)$, it follows that

$$c'x^* = \sum_{i=1}^{n+1} \alpha_i c' x_i \geq \sum_{i=1}^{n+1} \alpha_i c' x^* = c'x^*,$$

implying that $c'x_i = c'x^*$ for all i corresponding to $\alpha_i > 0$. Hence, $\inf_{x \in X} c'x$ is attained at the x_i's corresponding to $\alpha_i > 0$.

1.22 (Extension of Caratheodory's Theorem)

Let X_1 and X_2 be nonempty subsets of $\Re^n$, and let $X = \text{conv}(X_1) + \text{cone}(X_2)$. Show that every vector x in X can be represented in the form

$$x = \sum_{i=1}^{k} \alpha_i x_i + \sum_{i=k+1}^{m} \alpha_i y_i,$$

where m is a positive integer with $m \le n+1$, the vectors $x_1, \ldots, x_k$ belong to X_1, the vectors $y_{k+1}, \ldots, y_m$ belong to X_2, and the scalars $\alpha_1, \ldots, \alpha_m$ are nonnegative with $\alpha_1 + \cdots + \alpha_k = 1$. Furthermore, the vectors $x_2 - x_1, \ldots, x_k - x_1, y_{k+1}, \ldots, y_m$ are linearly independent.

Solution: The proof will be an application of Caratheodory's Theorem [part (a)] to the subset of $\Re^{n+1}$ given by

$$Y = \big\{(x, 1) \mid x \in X_1\big\} \cup \big\{(y, 0) \mid y \in X_2\big\}.$$

If $x \in X$, then

$$x = \sum_{i=1}^{k} \gamma_i x_i + \sum_{i=k+1}^{m} \gamma_i y_i,$$

where the vectors $x_1, \ldots, x_k$ belong to X_1, the vectors $y_{k+1}, \ldots, y_m$ belong to X_2, and the scalars $\gamma_1, \ldots, \gamma_m$ are nonnegative with $\gamma_1 + \cdots + \gamma_k = 1$. Equivalently, $(x, 1) \in \text{cone}(Y)$. By Caratheodory's Theorem part (a), we have that

$$(x, 1) = \sum_{i=1}^{k} \alpha_i (x_i, 1) + \sum_{i=k+1}^{m} \alpha_i (y_i, 0),$$

for some positive scalars $\alpha_1, \ldots, \alpha_m$ and vectors

$$(x_1, 1), \ldots (x_k, 1),\ (y_{k+1}, 0), \ldots,\ (y_m, 0),$$

which are linearly independent (implying that $m \le n+1$) or equivalently,

$$x = \sum_{i=1}^{k} \alpha_i x_i + \sum_{i=k+1}^{m} \alpha_i y_i, \qquad 1 = \sum_{i=1}^{k} \alpha_i.$$

Finally, to show that the vectors $x_2 - x_1, \ldots,\ x_k - x_1,\ y_{k+1}, \ldots, y_m$ are linearly independent, assume to arrive at a contradiction, that there exist $\lambda_2, \ldots, \lambda_m$, not all 0, such that

$$\sum_{i=2}^{k} \lambda_i (x_i - x_1) + \sum_{i=k+1}^{m} \lambda_i y_i = 0.$$

Equivalently, defining $\lambda_1 = -(\lambda_2 + \cdots + \lambda_m)$, we have

$$\sum_{i=1}^{k} \lambda_i (x_i, 1) + \sum_{i=k+1}^{m} \lambda_i (y_i, 0) = 0,$$

which contradicts the linear independence of the vectors

$$(x_1, 1), \ldots, (x_k, 1), (y_{k+1}, 0), \ldots, (y_m, 0).$$

1.23

Let X be a nonempty bounded subset of $\Re^n$. Show that

$$\mathrm{cl}\big(\mathrm{conv}(X)\big) = \mathrm{conv}\big(\mathrm{cl}(X)\big).$$

In particular, if X is compact, then $\mathrm{conv}(X)$ is compact (cf. Prop. 1.3.2).

Solution: The set $\mathrm{cl}(X)$ is compact since X is bounded by assumption. Hence, by Prop. 1.3.2, its convex hull, $\mathrm{conv}\big(\mathrm{cl}(X)\big)$, is compact, and it follows that

$$\mathrm{cl}\big(\mathrm{conv}(X)\big) \subset \mathrm{cl}\Big(\mathrm{conv}\big(\mathrm{cl}(X)\big)\Big) = \mathrm{conv}\big(\mathrm{cl}(X)\big).$$

It is also true in general that

$$\mathrm{conv}\big(\mathrm{cl}(X)\big) \subset \mathrm{conv}\Big(\mathrm{cl}\big(\mathrm{conv}(X)\big)\Big) = \mathrm{cl}\big(\mathrm{conv}(X)\big),$$

since by Prop. 1.2.1(d), the closure of a convex set is convex. Hence, the result follows.

1.24 (Radon's Theorem)

Let $x_1, \ldots, x_m$ be vectors in $\Re^n$, where $m \geq n+2$. Show that there exists a partition of the index set $\{1, \ldots, m\}$ into two disjoint sets I and J such that

$$\mathrm{conv}\big(\{x_i \mid i \in I\}\big) \cap \mathrm{conv}\big(\{x_j \mid j \in J\}\big) \neq \emptyset.$$

Hint: The system of $n+1$ equations in the m unknowns $\lambda_1, \ldots, \lambda_m$,

$$\sum_{i=1}^m \lambda_i x_i = 0, \qquad \sum_{i=1}^m \lambda_i = 0,$$

has a nonzero solution λ^*. Let $I = \{i \mid \lambda_i^* \geq 0\}$ and $J = \{j \mid \lambda_j^* < 0\}$.

Solution: Consider the system of $n+1$ equations in the m unknowns $\lambda_1, \ldots, \lambda_m$

$$\sum_{i=1}^m \lambda_i x_i = 0, \qquad \sum_{i=1}^m \lambda_i = 0.$$

Since $m > n+1$, there exists a nonzero solution, call it λ^*. Let

$$I = \{i \mid \lambda_i^* \geq 0\}, \qquad J = \{j \mid \lambda_j^* < 0\},$$

and note that I and J are nonempty, and that

$$\sum_{k \in I} \lambda_k^* = \sum_{k \in J} (-\lambda_k^*) > 0.$$

Consider the vector

$$x^* = \sum_{i \in I} \alpha_i x_i,$$

where

$$\alpha_i = \frac{\lambda_i^*}{\sum_{k \in I} \lambda_k^*}, \qquad i \in I.$$

In view of the equations $\sum_{i=1}^m \lambda_i^* x_i = 0$ and $\sum_{i=1}^m \lambda_i^* = 0$, we also have

$$x^* = \sum_{j \in J} \alpha_j x_j,$$

where

$$\alpha_j = \frac{-\lambda_j^*}{\sum_{k \in J} (-\lambda_k^*)}, \qquad j \in J.$$

It is seen that the α_i and α_j are nonnegative, and that

$$\sum_{i \in I} \alpha_i = \sum_{j \in J} \alpha_j = 1,$$

so x^* belongs to the intersection

$$\text{conv}\big(\{x_i \mid i \in I\}\big) \cap \text{conv}\big(\{x_j \mid j \in J\}\big).$$

Given four distinct points in the plane (i.e., $m = 4$ and $n = 2$), Radon's Theorem guarantees the existence of a partition into two subsets, the convex hulls of which intersect. Assuming, there is no subset of three points lying on the same line, there are two possibilities:

(1) Each set in the partition consists of two points, in which case the convex hulls intesect and define the diagonals of a quadrilateral.

(2) One set in the partition consists of three points and the other consists of one point, in which case the triangle formed by the three points must contain the fourth.

In the case where three of the points define a line segment on which they lie, and the fourth does not, the triangle formed by the two ends of the line segment and the point outside the line segment form a triangle that contains the fourth point. In the case where all four of the points lie on a line segment, the degenerate triangle formed by three of the points, including the two ends of the line segment, contains the fourth point.

1.25 (Helly's Theorem [Hel21])

Consider a finite collection of convex subsets of $\Re^n$, and assume that the intersection of every subcollection of $n+1$ (or fewer) sets has nonempty intersection. Show that the entire collection has nonempty intersection. *Hint*: Use induction. Assume that the conclusion holds for every collection of M sets, where $M \geq n+1$, and show that the conclusion holds for every collection of $M+1$ sets. In particular, let $C_1, \ldots, C_{M+1}$ be a collection of $M+1$ convex sets, and consider the collection of $M+1$ sets $B_1, \ldots, B_{M+1}$, where

$$B_j = \cap_{\substack{i=1,\ldots,M+1,\\ i\neq j}} C_i, \qquad j = 1, \ldots, M+1.$$

Note that, by the induction hypothesis, each set B_j is the intersection of a collection of M sets that have the property that every subcollection of $n+1$ (or fewer) sets has nonempty intersection. Hence each set B_j is nonempty. Let x_j be a vector in B_j. Apply Radon's Theorem (Exercise 1.24) to the vectors $x_1, \ldots, x_{M+1}$. Show that any vector in the intersection of the corresponding convex hulls belongs to the intersection of $C_1, \ldots, C_{M+1}$.

Solution: Consider the induction argument of the hint, let B_j be defined as in the hint, and for each j, let x_j be a vector in B_j. Since $M+1 \geq n+2$, we can apply Radon's Theorem to the vectors $x_1, \ldots, x_{M+1}$. Thus, there exist nonempty and disjoint index subsets I and J such that $I \cup J = \{1, \ldots, M+1\}$, nonnegative scalars $\alpha_1, \ldots, \alpha_{M+1}$, and a vector x^* such that

$$x^* = \sum_{i\in I} \alpha_i x_i = \sum_{j\in J} \alpha_j x_j, \qquad \sum_{i\in I} \alpha_i = \sum_{j\in J} \alpha_j = 1.$$

It can be seen that for every $i \in I$, a vector in B_i belongs to the intersection $\cap_{j\in J} C_j$. Therefore, since x^* is a convex combination of vectors in B_i, $i \in I$, x^* also belongs to the intersection $\cap_{j\in J} C_j$. Similarly, by reversing the role of I and J, we see that x^* belongs to the intersection $\cap_{i\in I} C_I$. Thus, x^* belongs to the intersection of the entire collection $C_1, \ldots, C_{M+1}$.

1.26

Consider the problem of minimizing over $\Re^n$ the function

$$\max\big\{f_1(x), \ldots, f_M(x)\big\},$$

where $f_i : \Re^n \mapsto (-\infty, \infty]$, $i = 1, \ldots, M$, are convex functions, and assume that the optimal value, denoted f^*, is finite. Show that there exists a subset I of $\{1, \ldots, M\}$, containing no more than $n+1$ indices, such that

$$\inf_{x\in\Re^n} \left\{\max_{i\in I} f_i(x)\right\} = f^*.$$

Hint: Consider the convex sets $X_i = \big\{x \mid f_i(x) < f^*\big\}$, argue by contradiction, and apply Helly's Theorem (Exercise 1.25).

Solution: Assume the contrary, i.e., that for every index set $I \subset \{1, \ldots, M\}$, which contains no more than $n+1$ indices, we have

$$\inf_{x\in\Re^n} \left\{\max_{i\in I} f_i(x)\right\} < f^*.$$

This means that for every such I, the intersection $\cap_{i\in I} X_i$ is nonempty, where

$$X_i = \big\{x \mid f_i(x) < f^*\big\}.$$

From Helly's Theorem, it follows that the entire collection $\{X_i \mid i = 1,\ldots,M\}$ has nonempty intersection, thereby implying that

$$\inf_{x\in\Re^n} \left\{\max_{i=1,\ldots,M} f_i(x)\right\} < f^*.$$

This contradicts the definition of f^*. *Note*: The result of this exercise relates to the following question: what is the minimal number of functions f_i that we need to include in the cost function $\max_i f_i(x)$ in order to attain the optimal value f^*? According to the result, the number is no more than $n+1$. For applications of this result in structural design and Chebyshev approximation, see Ben Tal and Nemirovski [BeN01].

1.27

Let C be a nonempty convex subset of $\Re^n$, and let $f : \Re^n \mapsto (-\infty,\infty]$ be a convex function such that $f(x)$ is finite for all $x \in C$. Show that if for some scalar γ, we have $f(x) \geq \gamma$ for all $x \in C$, then we also have $f(x) \geq \gamma$ for all $x \in \mathrm{cl}(C)$.

Solution: Let $\overline{x}$ be an arbitrary vector in $\mathrm{cl}(C)$. If $f(\overline{x}) = \infty$, then we are done, so assume that $f(\overline{x})$ is finite. Let x be a point in the relative interior of C. By the Line Segment Principle, all the points on the line segment connecting x and $\overline{x}$, except possibly $\overline{x}$, belong to $\mathrm{ri}(C)$ and therefore, belong to C. From this, the given property of f, and the convexity of f, we obtain for all $\alpha \in (0,1]$,

$$\alpha f(x) + (1-\alpha) f(\overline{x}) \geq f\big(\alpha x + (1-\alpha)\overline{x}\big) \geq \gamma.$$

By letting $\alpha \to 0$, it follows that $f(\overline{x}) \geq \gamma$. Hence, $f(\overline{x}) \geq \gamma$ for all $x \in \mathrm{cl}(C)$.

1.28

Let C be a nonempty convex set, and let S be the subspace that is parallel to the affine hull of C. Show that

$$\mathrm{ri}(C) = \mathrm{int}(C + S^\perp) \cap C.$$

Solution: From Prop. 1.4.5(b), we have that for any vector $a \in \Re^n$, $\mathrm{ri}(C+a) = \mathrm{ri}(C) + a$. Therefore, we can assume without loss of generality that $0 \in C$, and $\mathrm{aff}(C)$ coincides with S. We need to show that

$$\mathrm{ri}(C) = \mathrm{int}(C + S^\perp) \cap C.$$

Let $x \in \text{ri}(C)$. By definition, this implies that $x \in C$ and there exists some open ball $B(x, \epsilon)$ centered at x with radius $\epsilon > 0$ such that

$$B(x, \epsilon) \cap S \subset C. \tag{1.10}$$

We now show that $B(x, \epsilon) \subset C + S^\perp$. Let z be a vector in $B(x, \epsilon)$. Then, we can express z as $z = x + \alpha y$ for some vector $y \in \Re^n$ with $\|y\| = 1$, and some $\alpha \in [0, \epsilon)$. Since S and $S^\perp$ are orthogonal subspaces, y can be uniquely decomposed as $y = y_S + y_{S^\perp}$, where $y_S \in S$ and $y_{S^\perp} \in S^\perp$. Since $\|y\| = 1$, this implies that $\|y_S\| \leq 1$ (Pythagorean Theorem), and using Eq. (1.10), we obtain

$$x + \alpha y_S \in B(x, \epsilon) \cap S \subset C,$$

from which it follows that the vector $z = x + \alpha y$ belongs to $C + S^\perp$, implying that $B(x, \epsilon) \subset C + S^\perp$. This shows that $x \in \text{int}(C + S^\perp) \cap C$.

Conversely, let $x \in \text{int}(C + S^\perp) \cap C$. We have that $x \in C$ and there exists some open ball $B(x, \epsilon)$ centered at x with radius $\epsilon > 0$ such that $B(x, \epsilon) \subset C + S^\perp$. Since C is a subset of S, it can be seen that $(C + S^\perp) \cap S = C$. Therefore,

$$B(x, \epsilon) \cap S \subset C,$$

implying that $x \in \text{ri}(C)$.

1.29

Let $x_0, \ldots, x_m$ be vectors in $\Re^n$ such that $x_1 - x_0, \ldots, x_m - x_0$ are linearly independent. The convex hull of $x_0, \ldots, x_m$ is called an *m-dimensional simplex*, and $x_0, \ldots, x_m$ are called the *vertices* of the simplex.

(a) Show that the dimension of a convex set is the maximum of the dimensions of all the simplices contained in the set.

(b) Use part (a) to show that a nonempty convex set has a nonempty relative interior.

Solution: (a) Let C be the given convex set. The convex hull of any subset of C is contained in C. Therefore, the maximum dimension of the various simplices contained in C is the largest m for which C contains $m + 1$ vectors $x_0, \ldots, x_m$ such that $x_1 - x_0, \ldots, x_m - x_0$ are linearly independent.

Let $K = \{x_0, \ldots, x_m\}$ be such a set with m maximal, and let $\text{aff}(K)$ denote the affine hull of set K. Then, we have $\dim\big(\text{aff}(K)\big) = m$, and since $K \subset C$, it follows that $\text{aff}(K) \subset \text{aff}(C)$.

We claim that $C \subset \text{aff}(K)$. To see this, assume that there exists some $x \in C$, which does not belong to $\text{aff}(K)$. This implies that the set $\{x, x_0, \ldots, x_m\}$ is a set of $m + 2$ vectors in C such that $x - x_0$, $x_1 - x_0, \ldots, x_m - x_0$ are linearly independent, contradicting the maximality of m. Hence, we have $C \subset \text{aff}(K)$, and it follows that

$$\text{aff}(K) = \text{aff}(C),$$

thereby implying that $\dim(C) = m$.

(b) We first consider the case where C is n-dimensional with $n > 0$ and show that the interior of C is not empty. By part (a), an n-dimensional convex set contains an n-dimensional simplex. We claim that such a simplex S has a nonempty interior. Indeed, applying an affine transformation if necessary, we can assume that the vertices of S are the vectors $(0, 0, \dots, 0), (1, 0, \dots, 0), \dots, (0, 0, \dots, 1)$, i.e.,

$$S = \left\{ (x_1, \dots, x_n) \;\middle|\; x_i \geq 0,\ \forall\, i = 1, \dots, n,\ \ \sum_{i=1}^{n} x_i \leq 1 \right\}.$$

The interior of the simplex S,

$$\text{int}(S) = \left\{ (x_1, \dots, x_n) \mid x_i > 0,\ \forall\, i = 1, \dots, n,\ \ \sum_{i=1}^{n} x_i < 1 \right\},$$

is nonempty, which in turn implies that $\text{int}(C)$ is nonempty.

For the case where $\dim(C) < n$, consider the n-dimensional set $C + S^\perp$, where $S^\perp$ is the orthogonal complement of the subspace parallel to $\text{aff}(C)$. Since $C + S^\perp$ is a convex set, it follows from the above argument that $\text{int}(C + S^\perp)$ is nonempty. Let $x \in \text{int}(C + S^\perp)$. We can represent x as $x = x_C + x_{S^\perp}$, where $x_C \in C$ and $x_{S^\perp} \in S^\perp$. It can be seen that $x_C \in \text{int}(C + S^\perp)$. Since

$$\text{ri}(C) = \text{int}(C + S^\perp) \cap C,$$

(cf. Exercise 1.28), it follows that $x_c \in \text{ri}(C)$, so $\text{ri}(C)$ is nonempty.

1.30

Let C_1 and C_2 be two nonempty convex sets such that $C_1 \subset C_2$.

(a) Give an example showing that $\text{ri}(C_1)$ need not be a subset of $\text{ri}(C_2)$.

(b) Assuming that the sets C_1 and C_2 have the same affine hull, show that $\text{ri}(C_1) \subset \text{ri}(C_2)$.

(c) Assuming that the sets $\text{ri}(C_1)$ and $\text{ri}(C_2)$ have nonempty intersection, show that $\text{ri}(C_1) \subset \text{ri}(C_2)$.

(d) Assuming that the sets C_1 and $\text{ri}(C_2)$ have nonempty intersection, show that the set $\text{ri}(C_1) \cap \text{ri}(C_2)$ is nonempty.

Solution: (a) Let C_1 be the segment $\big\{(x_1, x_2) \mid 0 \leq x_1 \leq 1,\ x_2 = 0\big\}$ and let C_2 be the box $\big\{(x_1, x_2) \mid 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\big\}$. We have

$$\text{ri}(C_1) = \big\{(x_1, x_2) \mid 0 < x_1 < 1,\ x_2 = 0\big\},$$

$$\text{ri}(C_2) = \big\{(x_1, x_2) \mid 0 < x_1 < 1,\ 0 < x_2 < 1\big\}.$$

Thus $C_1 \subset C_2$, while $\text{ri}(C_1) \cap \text{ri}(C_2) = \emptyset$.

(b) Let $x \in \text{ri}(C_1)$, and consider a open ball B centered at x such that $B \cap \text{aff}(C_1) \subset C_1$. Since $\text{aff}(C_1) = \text{aff}(C_2)$ and $C_1 \subset C_2$, it follows that $B \cap \text{aff}(C_2) \subset C_2$, so $x \in \text{ri}(C_2)$. Hence $\text{ri}(C_1) \subset \text{ri}(C_2)$.

(c) Because $C_1 \subset C_2$, we have

$$\text{ri}(C_1) = \text{ri}(C_1 \cap C_2).$$

Since $\text{ri}(C_1) \cap \text{ri}(C_2) \neq \emptyset$, there holds

$$\text{ri}(C_1 \cap C_2) = \text{ri}(C_1) \cap \text{ri}(C_2)$$

[Prop. 1.4.5(a)]. Combining the preceding two relations, we obtain $\text{ri}(C_1) \subset \text{ri}(C_2)$.

(d) Let x_2 be in the intersection of C_1 and $\text{ri}(C_2)$, and let x_1 be in the relative interior of C_1 [$\text{ri}(C_1)$ is nonempty by Prop. 1.4.1(b)]. If $x_1 = x_2$, then we are done, so assume that $x_1 \neq x_2$. By the Line Segment Principle, all the points on the line segment connecting x_1 and x_2, except possibly x_2, belong to the relative interior of C_1. Since $C_1 \subset C_2$, the vector x_1 is in C_2, so that by the Line Segment Principle, all the points on the line segment connecting x_1 and x_2, except possibly x_1, belong to the relative interior of C_2. Hence, all the points on the line segment connecting x_1 and x_2, except possibly x_1 and x_2, belong to the intersection $\text{ri}(C_1) \cap \text{ri}(C_2)$, showing that $\text{ri}(C_1) \cap \text{ri}(C_2)$ is nonempty.

1.31

Let C be a nonempty convex set.

(a) Show the following refinement of Prop. 1.4.1(c): $x \in \text{ri}(C)$ if and only if for every $\overline{x} \in \text{aff}(C)$, there exists a $\gamma > 1$ such that $x + (\gamma - 1)(x - \overline{x}) \in C$.

(b) Assuming that the origin lies in $\text{ri}(C)$, show that $\text{cone}(C)$ coincides with $\text{aff}(C)$.

(c) Show the following extension of part (b) to a nonconvex set: If X is a nonempty set such that the origin lies in the relative interior of $\text{conv}(X)$, then $\text{cone}(X)$ coincides with $\text{aff}(X)$.

Solution: (a) Let $x \in \text{ri}(C)$. We will show that for every $\overline{x} \in \text{aff}(C)$, there exists a $\gamma > 1$ such that $x + (\gamma - 1)(x - \overline{x}) \in C$. This is true if $\overline{x} = x$, so assume that $\overline{x} \neq x$. Since $x \in \text{ri}(C)$, there exists $\epsilon > 0$ such that

$$\big\{z \mid \|z - x\| < \epsilon\big\} \cap \text{aff}(C) \subset C.$$

Choose a point $\overline{x}_\epsilon \in C$ in the intersection of the ray $\big\{x + \alpha(\overline{x} - x) \mid \alpha \geq 0\big\}$ and the set $\big\{z \mid \|z - x\| < \epsilon\big\} \cap \text{aff}(C)$. Then, for some positive scalar α_ϵ,

$$x - \overline{x}_\epsilon = \alpha_\epsilon(x - \overline{x}).$$

Since $x \in \text{ri}(C)$ and $\overline{x}_\epsilon \in C$, by Prop. 1.4.1(c), there is $\gamma_\epsilon > 1$ such that

$$x + (\gamma_\epsilon - 1)(x - \overline{x}_\epsilon) \in C,$$

which in view of the preceding relation implies that

$$x + (\gamma_\epsilon - 1)\alpha_\epsilon(x - \overline{x}) \in C.$$

The result follows by letting $\gamma = 1 + (\gamma_\epsilon - 1)\alpha_\epsilon$ and noting that $\gamma > 1$, since $(\gamma_\epsilon - 1)\alpha_\epsilon > 0$. The converse assertion follows from the fact $C \subset \text{aff}(C)$ and Prop. 1.4.1(c).

(b) The inclusion $\text{cone}(C) \subset \text{aff}(C)$ always holds if $0 \in C$. To show the reverse inclusion, we note that by part (a) with $x = 0$, for every $\overline{x} \in \text{aff}(C)$, there exists $\gamma > 1$ such that $\tilde{x} = (\gamma - 1)(-\overline{x}) \in C$. By using part (a) again with $x = 0$, for $\tilde{x} \in C \subset \text{aff}(C)$, we see that there is $\tilde{\gamma} > 1$ such that $z = (\tilde{\gamma} - 1)(-\tilde{x}) \in C$, which combined with $\tilde{x} = (\gamma - 1)(-\overline{x})$ yields $z = (\tilde{\gamma} - 1)(\gamma - 1)\overline{x} \in C$. Hence

$$\overline{x} = \frac{1}{(\tilde{\gamma} - 1)(\gamma - 1)} z$$

with $z \in C$ and $(\tilde{\gamma} - 1)(\gamma - 1) > 0$, implying that $\overline{x} \in \text{cone}(C)$ and, showing that $\text{aff}(C) \subset \text{cone}(C)$.

(c) This follows by part (b), where $C = \text{conv}(X)$, and the fact

$$\text{cone}\big(\text{conv}(X)\big) = \text{cone}(X)$$

[Exercise 1.18(b)].

1.32

Let C be a nonempty set.

(a) If C is convex and compact, and the origin is not in the relative boundary of C, then $\text{cone}(C)$ is closed.

(b) Give examples showing that the assertion of part (a) fails if C is unbounded or the origin is in the relative boundary of C.

(c) If C is compact and the origin is not in the relative boundary of $\text{conv}(C)$, then $\text{cone}(C)$ is closed. *Hint*: Use part (a) and Exercise 1.18(b).

Solution:
(a) If $0 \in C$, then $0 \in \text{ri}(C)$ since 0 is not on the relative boundary of C. By Exercise 1.31(b), it follows that $\text{cone}(C)$ coincides with $\text{aff}(C)$, which is a closed set. If $0 \notin C$, let y be in the closure of $\text{cone}(C)$ and let $\{y_k\} \subset \text{cone}(C)$ be a sequence converging to y. By Exercise 1.15, for every y_k, there exists a nonnegative scalar α_k and a vector $x_k \in C$ such that $y_k = \alpha_k x_k$. Since $\{y_k\} \to y$, the sequence $\{y_k\}$ is bounded, implying that

$$\alpha_k \|x_k\| \le \sup_{m \ge 0} \|y_m\| < \infty, \qquad \forall\ k.$$

We have $\inf_{m \ge 0} \|x_m\| > 0$, since $\{x_k\} \subset C$ and C is a compact set not containing the origin, so that

$$0 \le \alpha_k \le \frac{\sup_{m \ge 0} \|y_m\|}{\inf_{m \ge 0} \|x_m\|} < \infty, \qquad \forall\ k.$$

Thus, the sequence $\{(\alpha_k, x_k)\}$ is bounded and has a limit point (α, x) such that $\alpha \geq 0$ and $x \in C$. By taking a subsequence of $\{(\alpha_k, x_k)\}$ that converges to (α, x), and by using the facts $y_k = \alpha_k x_k$ for all k and $\{y_k\} \to y$, we see that $y = \alpha x$ with $\alpha \geq 0$ and $x \in C$. Hence, $y \in \text{cone}(C)$, showing that $\text{cone}(C)$ is closed.

(b) To see that the assertion in part (a) fails when C is unbounded, let C be the line $\big\{(x_1, x_2) \mid x_1 = 1,\ x_2 \in \Re\big\}$ in $\Re^2$ not passing through the origin. Then, $\text{cone}(C)$ is the nonclosed set $\big\{(x_1, x_2) \mid x_1 > 0,\ x_2 \in \Re\big\} \cup \big\{(0,0)\big\}$.

To see that the assertion in part (a) fails when C contains the origin on its relative boundary, let C be the closed ball $\big\{(x_1, x_2) \mid (x_1 - 1)^2 + x_2^2 \leq 1\big\}$ in $\Re^2$. Then, $\text{cone}(C)$ is the nonclosed set $\big\{(x_1, x_2) \mid x_1 > 0,\ x_2 \in \Re\big\} \cup \big\{(0,0)\big\}$ (see Fig. 1.3.2).

(c) Since C is compact, the convex hull of C is compact (cf. Prop. 1.3.2). Because $\text{conv}(C)$ does not contain the origin on its relative boundary, by part (a), the cone generated by $\text{conv}(C)$ is closed. By Exercise 1.18(b), $\text{cone}\big(\text{conv}(C)\big)$ coincides with $\text{cone}(C)$ implying that $\text{cone}(C)$ is closed.

1.33

(a) Let C be a nonempty convex cone. Show that $\text{ri}(C)$ is also a convex cone.

(b) Let $C = \text{cone}\big(\{x_1, \ldots, x_m\}\big)$. Show that

$$\text{ri}(C) = \left\{ \sum_{i=1}^{m} \alpha_i x_i \;\middle|\; \alpha_i > 0,\ i = 1, \ldots, m \right\}.$$

Solution: (a) By Prop. 1.4.1(b), the relative interior of a convex set is a convex set. We only need to show that $\text{ri}(C)$ is a cone. Let $y \in \text{ri}(C)$. Then, $y \in C$ and since C is a cone, $\alpha y \in C$ for all $\alpha > 0$. By the Line Segment Principle, all the points on the line segment connecting y and αy, except possibly αy, belong to $\text{ri}(C)$. Since this is true for every $\alpha > 0$, it follows that $\alpha y \in \text{ri}(C)$ for all $\alpha > 0$, showing that $\text{ri}(C)$ is a cone.

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

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

under the linear transformation A. Therefore, by using Prop. 1.4.3(d), we have

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

1.34

Let A be an $m \times n$ matrix and let C be a nonempty convex set in $\Re^m$. Assuming that $A^{-1} \cdot \mathrm{ri}(C)$ is nonempty, show that

$$\mathrm{ri}(A^{-1} \cdot C) = A^{-1} \cdot \mathrm{ri}(C), \qquad \mathrm{cl}(A^{-1} \cdot C) = A^{-1} \cdot \mathrm{cl}(C).$$

(Compare these relations with those of Prop. 1.4.4.)

Solution: Define the sets

$$D = \Re^n \times C, \qquad S = \big\{(x, Ax) \mid x \in \Re^n\big\}.$$

Let T be the linear transformation that maps $(x, y) \in \Re^{n+m}$ into $x \in \Re^n$. Then it can be seen that

$$A^{-1} \cdot C = T \cdot (D \cap S). \tag{1.11}$$

The relative interior of D is given by $\mathrm{ri}(D) = \Re^n \times \mathrm{ri}(C)$, and the relative interior of S is equal to S (since S is a subspace). Hence,

$$A^{-1} \cdot \mathrm{ri}(C) = T \cdot \big(\mathrm{ri}(D) \cap S\big). \tag{1.12}$$

In view of the assumption that $A^{-1} \cdot \mathrm{ri}(C)$ is nonempty, we have that the intersection $\mathrm{ri}(D) \cap S$ is nonempty. Therefore, it follows from Props. 1.4.3(d) and 1.4.5(a) that

$$\mathrm{ri}\big(T \cdot (D \cap S)\big) = T \cdot \big(\mathrm{ri}(D) \cap S\big). \tag{1.13}$$

Combining Eqs. (1.11)-(1.13), we obtain

$$\mathrm{ri}(A^{-1} \cdot C) = A^{-1} \cdot \mathrm{ri}(C).$$

Next, we show the second relation. We have

$$A^{-1} \cdot \mathrm{cl}(C) = \big\{x \mid Ax \in \mathrm{cl}(C)\big\} = T \cdot \big\{(x, Ax) \mid Ax \in \mathrm{cl}(C)\big\} = T \cdot \big(\mathrm{cl}(D) \cap S\big).$$

Since the intersection $\mathrm{ri}(D) \cap S$ is nonempty, it follows from Prop. 1.4.5(a) that $\mathrm{cl}(D) \cap S = \mathrm{cl}(D \cap S)$. Furthermore, since T is continuous, we obtain

$$A^{-1} \cdot \mathrm{cl}(C) = T \cdot \mathrm{cl}(D \cap S) \subset \mathrm{cl}\big(T \cdot (D \cap S)\big),$$

which combined with Eq. (1.11) yields

$$A^{-1} \cdot \mathrm{cl}(C) \subset \mathrm{cl}(A^{-1} \cdot C).$$

To show the reverse inclusion, $\mathrm{cl}(A^{-1} \cdot C) \subset A^{-1} \cdot \mathrm{cl}(C)$, let $\overline{x}$ be some vector in $\mathrm{cl}(A^{-1} \cdot C)$. This implies that there exists some sequence $\{x_k\}$ converging to $\overline{x}$ such that $Ax_k \in C$ for all k. Since x_k converges to $\overline{x}$, we have that Ax_k converges to $A\overline{x}$, thereby implying that $A\overline{x} \in \mathrm{cl}(C)$, or equivalently, $\overline{x} \in A^{-1} \cdot \mathrm{cl}(C)$.

1.35 (Closure of a Convex Function)

Consider a proper convex function $f : \Re^n \mapsto (-\infty, \infty]$ and the function whose epigraph is the closure of the epigraph of f. This function is called the *closure of* f and is denoted by $\mathrm{cl}\, f$. Show that:

(a) $\mathrm{cl}\, f$ is the greatest lower semicontinuous function majorized by f, i.e., if $g : \Re^n \mapsto [-\infty, \infty]$ is lower semicontinuous and satisfies $g(x) \le f(x)$ for all $x \in \Re^n$, then $g(x) \le (\mathrm{cl} f)(x)$ for all $x \in \Re^n$.

(b) $\mathrm{cl}\, f$ is a closed proper convex function and

$$(\mathrm{cl}\, f)(x) = f(x), \qquad \forall\, x \in \mathrm{ri}\big(\mathrm{dom}(f)\big).$$

(c) If $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$ and $y \in \mathrm{dom}(\mathrm{cl}\, f)$, we have

$$(\mathrm{cl}\, f)(y) = \lim_{\alpha \downarrow 0} f\big(y + \alpha(x - y)\big).$$

(d) Assume that $f = f_1 + \cdots + f_m$, where $f_i : \Re^n \mapsto (-\infty, \infty]$, $i = 1, \ldots, m$, are proper convex functions such that $\cap_{i=1}^m \mathrm{ri}\big(\mathrm{dom}(f_i)\big) \neq \varnothing$. Show that

$$(\mathrm{cl}\, f)(x) = (\mathrm{cl}\, f_1)(x) + \cdots + (\mathrm{cl}\, f_m)(x), \qquad \forall\, x \in \Re^n.$$

Solution: (a) Let $g : \Re^n \mapsto [-\infty, \infty]$ be such that $g(x) \le f(x)$ for all $x \in \Re^n$. Choose any $x \in \mathrm{dom}(\mathrm{cl}\, f)$. Since $\mathrm{epi}(\mathrm{cl}\, f) = \mathrm{cl}\big(\mathrm{epi}(f)\big)$, we can choose a sequence $\big\{(x_k, w_k)\big\} \in \mathrm{epi}(f)$ such that $x_k \to x$, $w_k \to (\mathrm{cl}\, f)(x)$. Since g is lower semicontinuous at x, we have

$$g(x) \le \liminf_{k \to \infty} g(x_k) \le \liminf_{k \to \infty} f(x_k) \le \liminf_{k \to \infty} w_k = (\mathrm{cl}\, f)(x).$$

Note also that since $\mathrm{epi}(f) \subset \mathrm{epi}(\mathrm{cl}\, f)$, we have $(\mathrm{cl}\, f)(x) \le f(x)$ for all $x \in \Re^n$.

(b) For the proof of this part and the next, we will use the easily shown fact that for any convex function f, we have

$$\mathrm{ri}\big(\mathrm{epi}(f)\big) = \big\{(x, w) \mid x \in \mathrm{ri}\big(\mathrm{dom}(f)\big),\ f(x) < w\big\}.$$

Let $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$, and consider the vertical line $L = \big\{(x, w) \mid w \in \Re\big\}$. Then there exists $\hat{w}$ such that $(x, \hat{w}) \in L \cap \mathrm{ri}\big(\mathrm{epi}(f)\big)$. Let $\overline{w}$ be such that $(x, \overline{w}) \in L \cap \mathrm{cl}\big(\mathrm{epi}(f)\big)$. Then, by Prop. 1.4.5(a), we have $L \cap \mathrm{cl}\big(\mathrm{epi}(f)\big) = \mathrm{cl}\big(L \cap \mathrm{epi}(f)\big)$, so that $(x, \overline{w}) \in \mathrm{cl}\big(L \cap \mathrm{epi}(f)\big)$. It follows from the Line Segment Principle that the vector $\big(x, \hat{w} + \alpha(\overline{w} - \hat{w})\big)$ belongs to $\mathrm{epi}(f)$ for all $\alpha \in [0, 1)$. Taking the limit as $\alpha \to 1$, we see that $f(x) \le \overline{w}$ for all $\overline{w}$ such that $(x, \overline{w}) \in L \cap \mathrm{cl}\big(\mathrm{epi}(f)\big)$, implying that $f(x) \le (\mathrm{cl}\, f)(x)$. On the other hand, since $\mathrm{epi}(f) \subset \mathrm{epi}(\mathrm{cl}\, f)$, we have $(\mathrm{cl}\, f)(x) \le f(x)$ for all $x \in \Re^n$, so $f(x) = (\mathrm{cl}\, f)(x)$.

We know that a closed convex function that is improper cannot take a finite value at any point. Since $\mathrm{cl}\, f$ is closed and convex, and takes a finite value at all points of the nonempty set $\mathrm{ri}\big(\mathrm{dom}(f)\big)$, it follows that $\mathrm{cl}\, f$ must be proper.

(c) Since the function $\mathrm{cl}\, f$ is closed and is majorized by f, we have

$$(\mathrm{cl}\, f)(y) \leq \liminf_{\alpha \downarrow 0} (\mathrm{cl}\, f)\big(y + \alpha(x-y)\big) \leq \liminf_{\alpha \downarrow 0} f\big(y + \alpha(x-y)\big).$$

To show the reverse inequality, let w be such that $f(x) < w$. Then, $(x, w) \in \mathrm{ri}\big(\mathrm{epi}(f)\big)$, while $\big(y, (\mathrm{cl}\, f)(y)\big) \in \mathrm{cl}\big(\mathrm{epi}(f)\big)$. From the Line Segment Principle, it follows that

$$\big(\alpha x + (1-\alpha)y, \alpha w + (1-\alpha)(\mathrm{cl}\, f)(y)\big) \in \mathrm{ri}\big(\mathrm{epi}(f)\big), \qquad \forall\ \alpha \in (0, 1].$$

Hence,

$$f\big(\alpha x + (1-\alpha)y\big) < \alpha w + (1-\alpha)(\mathrm{cl}\, f)(y), \qquad \forall\ \alpha \in (0, 1].$$

By taking the limit as $\alpha \to 0$, we obtain

$$\liminf_{\alpha \downarrow 0} f\big(y + \alpha(x-y)\big) \leq (\mathrm{cl}\, f)(y),$$

thus completing the proof.

(d) Let $x \in \cap_{i=1}^m \mathrm{ri}\big(\mathrm{dom}(f_i)\big)$. Since by Prop. 1.4.5(a), we have

$$\mathrm{ri}\big(\mathrm{dom}(f)\big) = \cap_{i=1}^m \mathrm{ri}\big(\mathrm{dom}(f_i)\big),$$

it follows that $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$. By using part (c), we have for every $y \in \mathrm{dom}(\mathrm{cl}\, f)$,

$$(\mathrm{cl}\, f)(y) = \lim_{\alpha \downarrow 0} f\big(y + \alpha(x-y)\big) = \sum_{i=1}^m \lim_{\alpha \downarrow 0} f_i\big(y + \alpha(x-y)\big) = \sum_{i=1}^m (\mathrm{cl}\, f_i)(y).$$

1.36

Let C be a convex set and let M be an affine set such that the intersection $C \cap M$ is nonempty and bounded. Show that for every affine set $\overline{M}$ that is parallel to M, the intersection $C \cap \overline{M}$ is bounded.

Solution: The assumption that "$C \cap M$ is bounded" must be modified to read "$\mathrm{cl}(C) \cap M$ is bounded". Assume first that C is closed. Since $C \cap M$ is bounded, by part (c) of the Recession Cone Theorem (cf. Prop. 1.5.1), $R_{C \cap M} = \{0\}$. This and the fact $R_{C \cap M} = R_C \cap R_M$, imply that $R_C \cap R_M = \{0\}$. Let S be a subspace such that $M = x + S$ for some $x \in M$. Then $R_M = S$, so that $R_C \cap S = \{0\}$. For every affine set $\overline{M}$ that is parallel to M, we have $R_{\overline{M}} = S$, so that

$$R_{C \cap \overline{M}} = R_C \cap R_{\overline{M}} = R_C \cap S = \{0\}.$$

Therefore, by part (c) of the Recession Cone Theorem, $C \cap \overline{M}$ is bounded.

In the general case where C is not closed, we replace C with $\mathrm{cl}(C)$. By what has already been proved, $\mathrm{cl}(C) \cap \overline{M}$ is bounded, implying that $C \cap \overline{M}$ is bounded.

1.37 (Properties of Cartesian Products)

Given nonempty sets $X_i \subset \Re^{n_i}$, $i = 1, \ldots, m$, let $X = X_1 \times \cdots \times X_m$ be their Cartesian product. Show that:

(a) The convex hull (closure, affine hull) of X is equal to the Cartesian product of the convex hulls (closures, affine hulls, respectively) of the X_i.

(b) If all the sets $X_1, \ldots, X_m$ contain the origin, then

$$\text{cone}(X) = \text{cone}(X_1) \times \cdots \times \text{cone}(X_m).$$

Furthermore, the result fails if one of the sets does not contain the origin.

(c) If all the sets $X_1, \ldots, X_m$ are convex, then the relative interior (recession cone) of X is equal to the Cartesian product of the relative interiors (recession cones, respectively) of the X_i.

Solution: (a) We first show that the convex hull of X is equal to the Cartesian product of the convex hulls of the sets X_i, $i = 1, \ldots, m$. Let y be a vector that belongs to conv(X). Then, by definition, for some k, we have

$$y = \sum_{i=1}^{k} \alpha_i y_i, \qquad \text{with } \alpha_i \geq 0,\ i = 1, \ldots, m, \quad \sum_{i=1}^{k} \alpha_i = 1,$$

where $y_i \in X$ for all i. Since $y_i \in X$, we have that $y_i = (x_1^i, \ldots, x_m^i)$ for all i, with $x_1^i \in X_1, \ldots, x_m^i \in X_m$. It follows that

$$y = \sum_{i=1}^{k} \alpha_i (x_1^i, \ldots, x_m^i) = \left(\sum_{i=1}^{k} \alpha_i x_1^i, \ldots, \sum_{i=1}^{k} \alpha_i x_m^i \right),$$

thereby implying that $y \in \text{conv}(X_1) \times \cdots \times \text{conv}(X_m)$.

To prove the reverse inclusion, assume that y is a vector in $\text{conv}(X_1) \times \cdots \times \text{conv}(X_m)$. Then, we can represent y as $y = (y_1, \ldots, y_m)$ with $y_i \in \text{conv}(X_i)$, i.e., for all $i = 1, \ldots, m$, we have

$$y_i = \sum_{j=1}^{k_i} \alpha_j^i x_j^i, \qquad x_j^i \in X_i,\ \forall\, j, \qquad \alpha_j^i \geq 0,\ \forall\, j, \quad \sum_{j=1}^{k_i} \alpha_j^i = 1.$$

First, consider the vectors

$$(x_1^1, x_{r_1}^2, \ldots, x_{r_{m-1}}^m),\ (x_2^1, x_{r_1}^2, \ldots, x_{r_{m-1}}^m), \ldots, (x_{k_i}^1, x_{r_1}^2, \ldots, x_{r_{m-1}}^m),$$

for all possible values of $r_1, \ldots, r_{m-1}$, i.e., we fix all components except the first one, and vary the first component over all possible x_j^1's used in the convex combination that yields y_1. Since all these vectors belong to X, their convex combination given by

$$\left(\left(\sum_{j=1}^{k_1} \alpha_j^1 x_j^1 \right), x_{r_1}^2, \ldots, x_{r_{m-1}}^m \right)$$

belongs to the convex hull of X for all possible values of $r_1, \ldots, r_{m-1}$. Now, consider the vectors

$$\left(\left(\sum_{j=1}^{k_1} \alpha_j^1 x_j^1\right), x_1^2, \ldots, x_{r_{m-1}}^m\right), \ldots, \left(\left(\sum_{j=1}^{k_1} \alpha_j^1 x_j^1\right), x_{k_2}^2, \ldots, x_{r_{m-1}}^m\right),$$

i.e., fix all components except the second one, and vary the second component over all possible x_j^2's used in the convex combination that yields y_2. Since all these vectors belong to $\mathrm{conv}(X)$, their convex combination given by

$$\left(\left(\sum_{j=1}^{k_1} \alpha_j^1 x_j^1\right), \left(\sum_{j=1}^{k_2} \alpha_j^2 x_j^2\right), \ldots, x_{r_{m-1}}^m\right)$$

belongs to the convex hull of X for all possible values of $r_2, \ldots, r_{m-1}$. Proceeding in this way, we see that the vector given by

$$\left(\left(\sum_{j=1}^{k_1} \alpha_j^1 x_j^1\right), \left(\sum_{j=1}^{k_2} \alpha_j^2 x_j^2\right), \ldots, \left(\sum_{j=1}^{k_m} \alpha_j^m x_j^m\right)\right)$$

belongs to $\mathrm{conv}(X)$, thus proving our claim.

Next, we show the corresponding result for the closure of X. Assume that $y = (x_1, \ldots, x_m) \in \mathrm{cl}(X)$. This implies that there exists some sequence $\{y^k\} \subset X$ such that $y^k \to y$. Since $y^k \in X$, we have that $y^k = (x_1^k, \ldots, x_m^k)$ with $x_i^k \in X_i$ for each i and k. Since $y^k \to y$, it follows that $x_i \in \mathrm{cl}(X_i)$ for each i, and hence $y \in \mathrm{cl}(X_1) \times \cdots \times \mathrm{cl}(X_m)$. Conversely, suppose that $y = (x_1, \ldots, x_m) \in \mathrm{cl}(X_1) \times \cdots \times \mathrm{cl}(X_m)$. This implies that there exist sequences $\{x_i^k\} \subset X_i$ such that $x_i^k \to x_i$ for each $i = 1, \ldots, m$. Since $x_i^k \in X_i$ for each i and k, we have that $y^k = (x_1^k, \ldots, x_m^k) \in X$ and $\{y^k\}$ converges to $y = (x_1, \ldots, x_m)$, implying that $y \in \mathrm{cl}(X)$.

Finally, we show the corresponding result for the affine hull of X. Let's assume, by using a translation argument if necessary, that all the X_i's contain the origin, so that $\mathrm{aff}(X_1), \ldots, \mathrm{aff}(X_m)$ as well as $\mathrm{aff}(X)$ are all subspaces.

Assume that $y \in \mathrm{aff}(X)$. Let the dimension of $\mathrm{aff}(X)$ be r, and let $y^1, \ldots, y^r$ be linearly independent vectors in X that span $\mathrm{aff}(X)$. Thus, we can represent y as

$$y = \sum_{i=1}^{r} \beta^i y^i,$$

where $\beta^1, \ldots, \beta^r$ are scalars. Since $y^i \in X$, we have that $y^i = (x_1^i, \ldots, x_m^i)$ with $x_j^i \in X_j$. Thus,

$$y = \sum_{i=1}^{r} \beta^i (x_1^i, \ldots, x_m^i) = \left(\sum_{i=1}^{r} \beta^i x_1^i, \ldots, \sum_{i=1}^{r} \beta^i x_m^i\right),$$

implying that $y \in \text{aff}(X_1) \times \cdots \times \text{aff}(X_m)$. Now, assume that $y \in \text{aff}(X_1) \times \cdots \times \text{aff}(X_m)$. Let the dimension of $\text{aff}(X_i)$ be r_i, and let $x_i^1, \ldots, x_i^{r_i}$ be linearly independent vectors in X_i that span $\text{aff}(X_i)$. Thus, we can represent y as

$$y = \left(\sum_{j=1}^{r_1} \beta_1^j x_1^j, \ldots, \sum_{j=1}^{r_m} \beta_m^j x_m^j \right).$$

Since each X_i contains the origin, we have that the vectors

$$\left(\sum_{j=1}^{r_1} \beta_1^j x_1^j, 0, \ldots, 0 \right), \left(0, \sum_{j=1}^{r_2} \beta_2^j x_2^j, 0, \ldots, 0 \right), \ldots, \left(0, \ldots, \sum_{j=1}^{r_m} \beta_m^j x_m^j \right),$$

belong to $\text{aff}(X)$, and so does their sum, which is the vector y. Thus, $y \in \text{aff}(X)$, concluding the proof.

(b) Assume that $y \in \text{cone}(X)$. We can represent y as

$$y = \sum_{i=1}^{r} \alpha^i y^i,$$

for some r, where $\alpha^1, \ldots, \alpha^r$ are nonnegative scalars and $y_i \in X$ for all i. Since $y^i \in X$, we have that $y^i = (x_1^i, \ldots, x_m^i)$ with $x_j^i \in X_j$. Thus,

$$y = \sum_{i=1}^{r} \alpha^i (x_1^i, \ldots, x_m^i) = \left(\sum_{i=1}^{r} \alpha^i x_1^i, \ldots, \sum_{i=1}^{r} \alpha^i x_m^i \right),$$

implying that $y \in \text{cone}(X_1) \times \cdots \times \text{cone}(X_m)$.

Conversely, assume that $y \in \text{cone}(X_1) \times \cdots \times \text{cone}(X_m)$. Then, we can represent y as

$$y = \left(\sum_{j=1}^{r_1} \alpha_1^j x_1^j, \ldots, \sum_{j=1}^{r_m} \alpha_m^j x_m^j \right),$$

where $x_i^j \in X_i$ and $\alpha_i^j \geq 0$ for each i and j. Since each X_i contains the origin, we have that the vectors

$$\left(\sum_{j=1}^{r_1} \alpha_1^j x_1^j, 0, \ldots, 0 \right), \left(0, \sum_{j=1}^{r_2} \alpha_2^j x_2^j, 0, \ldots, 0 \right) \ldots, \left(0, \ldots, \sum_{j=1}^{r_m} \alpha_m^j x_m^j \right),$$

belong to the $\text{cone}(X)$, and so does their sum, which is the vector y. Thus, $y \in \text{cone}(X)$, concluding the proof.

Finally, consider the example where

$$X_1 = \{0, 1\} \subset \Re, \qquad X_2 = \{1\} \subset \Re.$$

For this example, $\text{cone}(X_1) \times \text{cone}(X_2)$ is given by the nonnegative quadrant, whereas $\text{cone}(X)$ is given by the two halflines $\alpha(0,1)$ and $\alpha(1,1)$ for $\alpha \geq 0$ and the region that lies between them.

(c) We first show that

$$\text{ri}(X) = \text{ri}(X_1) \times \cdots \times \text{ri}(X_m).$$

Let $x = (x_1, \ldots, x_m) \in \text{ri}(X)$. Then, by Prop. 1.4.1 (c), we have that for all $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_m) \in X$, there exists some $\gamma > 1$ such that

$$x + (\gamma - 1)(x - \overline{x}) \in X.$$

Therefore, for all $\overline{x}_i \in X_i$, there exists some $\gamma > 1$ such that

$$x_i + (\gamma - 1)(x_i - \overline{x}_i) \in X_i,$$

which, by Prop. 1.4.1(c), implies that $x_i \in \text{ri}(X_i)$, i.e., $x \in \text{ri}(X_1) \times \cdots \times \text{ri}(X_m)$. Conversely, let $x = (x_1, \ldots, x_m) \in \text{ri}(X_1) \times \cdots \times \text{ri}(X_m)$. The above argument can be reversed through the use of Prop. 1.4.1(c), to show that $x \in \text{ri}(X)$. Hence, the result follows.

Finally, let us show that

$$R_X = R_{X_1} \times \cdots \times R_{X_m}.$$

Let $y = (y_1, \ldots, y_m) \in R_X$. By definition, this implies that for all $x \in X$ and $\alpha \geq 0$, we have $x + \alpha y \in X$. From this, it follows that for all $x_i \in X_i$ and $\alpha \geq 0$, $x_i + \alpha y_i \in X_i$, so that $y_i \in R_{X_i}$, implying that $y \in R_{X_1} \times \cdots \times R_{X_m}$. Conversely, let $y = (y_1, \ldots, y_m) \in R_{X_1} \times \cdots \times R_{X_m}$. By definition, for all $x_i \in X_i$ and $\alpha \geq 0$, we have $x_i + \alpha y_i \in X_i$. From this, we get for all $x \in X$ and $\alpha \geq 0$, $x + \alpha y \in X$, thus showing that $y \in R_X$.

1.38 (Recession Cones of Nonclosed Sets)

Let C be a nonempty convex set.

(a) Show that

$$R_C \subset R_{\text{cl}(C)}, \qquad \text{cl}(R_C) \subset R_{\text{cl}(C)}.$$

Give an example where the inclusion $\text{cl}(R_C) \subset R_{\text{cl}(C)}$ is strict.

(b) Let $\overline{C}$ be a closed convex set such that $C \subset \overline{C}$. Show that $R_C \subset R_{\overline{C}}$. Give an example showing that the inclusion can fail if $\overline{C}$ is not closed.

Solution:
(a) Let $y \in R_C$. Then, by the definition of R_C, $x + \alpha y \in C$ for every $x \in C$ and every $\alpha \geq 0$. Since $C \subset \text{cl}(C)$, it follows that $x + \alpha y \in \text{cl}(C)$ for some $x \in \text{cl}(C)$ and every $\alpha \geq 0$, which, in view of part (b) of the Recession Cone Theorem (cf. Prop. 1.5.1), implies that $y \in R_{\text{cl}(C)}$. Hence

$$R_C \subset R_{\text{cl}(C)}.$$

By taking closures in this relation and by using the fact that $R_{\mathrm{cl}(C)}$ is closed [part (a) of the Recession Cone Theorem], we obtain $\mathrm{cl}(R_C) \subset R_{\mathrm{cl}(C)}$.

To see that the inclusion $\mathrm{cl}(R_C) \subset R_{\mathrm{cl}(C)}$ can be strict, consider the set

$$C = \big\{(x_1, x_2) \mid 0 \le x_1,\ 0 \le x_2 < 1\big\} \cup \big\{(0, 1)\big\},$$

whose closure is

$$\mathrm{cl}(C) = \{(x_1, x_2) \mid 0 \le x_1,\ 0 \le x_2 \le 1\}.$$

The recession cones of C and its closure are

$$R_C = \big\{(0, 0)\big\}, \qquad R_{\mathrm{cl}(C)} = \big\{(x_1, x_2) \mid 0 \le x_1,\ x_2 = 0\big\}.$$

Thus, $\mathrm{cl}(R_C) = \big\{(0, 0)\big\}$, and $\mathrm{cl}(R_C)$ is a strict subset of $R_{\mathrm{cl}(C)}$.

(b) Let $y \in R_C$ and let x be a vector in C. Then we have $x + \alpha y \in C$ for all $\alpha \ge 0$. Thus for the vector x, which belongs to $\overline{C}$, we have $x + \alpha y \in \overline{C}$ for all $\alpha \ge 0$, and it follows from part (b) of the Recession Cone Theorem (cf. Prop. 1.5.1) that $y \in R_{\overline{C}}$. Hence, $R_C \subset R_{\overline{C}}$.

To see that the inclusion $R_C \subset R_{\overline{C}}$ can fail when $\overline{C}$ is not closed, consider the sets

$$C = \big\{(x_1, x_2) \mid x_1 \ge 0,\ x_2 = 0\big\}, \qquad \overline{C} = \big\{(x_1, x_2) \mid x_1 \ge 0,\ 0 \le x_2 < 1\big\}.$$

Their recession cones are

$$R_C = C = \big\{(x_1, x_2) \mid x_1 \ge 0,\ x_2 = 0\big\}, \qquad R_{\overline{C}} = \big\{(0, 0)\big\},$$

showing that R_C is not a subset of $R_{\overline{C}}$.

1.39 (Recession Cones of Relative Interiors)

Let C be a nonempty convex set.

(a) Show that $R_{\mathrm{ri}(C)} = R_{\mathrm{cl}(C)}$.

(b) Show that a vector y belongs to $R_{\mathrm{ri}(C)}$ if and only if there exists a vector $x \in \mathrm{ri}(C)$ such that $x + \alpha y \in \mathrm{ri}(C)$ for every $\alpha \ge 0$.

(c) Let $\overline{C}$ be a convex set such that $\overline{C} = \mathrm{ri}(\overline{C})$ and $C \subset \overline{C}$. Show that $R_C \subset R_{\overline{C}}$. Give an example showing that the inclusion can fail if $\overline{C} \ne \mathrm{ri}(\overline{C})$.

Solution: (a) The inclusion $R_{\mathrm{ri}(C)} \subset R_{\mathrm{cl}(C)}$ follows from Exercise 1.38(b).

Conversely, let $y \in R_{\mathrm{cl}(C)}$, so that by the definition of $R_{\mathrm{cl}(C)}$, $x + \alpha y \in \mathrm{cl}(C)$ for every $x \in \mathrm{cl}(C)$ and every $\alpha \ge 0$. In particular, $x + \alpha y \in \mathrm{cl}(C)$ for every $x \in \mathrm{ri}(C)$ and every $\alpha \ge 0$. By the Line Segment Principle, all points on the line segment connecting x and $x + \alpha y$, except possibly $x + \alpha y$, belong to $\mathrm{ri}(C)$, implying that $x + \alpha y \in \mathrm{ri}(C)$ for every $x \in \mathrm{ri}(C)$ and every $\alpha \ge 0$. Hence, $y \in R_{\mathrm{ri}(C)}$, showing that $R_{\mathrm{cl}(C)} \subset R_{\mathrm{ri}(C)}$.

(b) If $y \in R_{\mathrm{ri}(C)}$, then by the definition of $R_{\mathrm{ri}(C)}$ for every vector $x \in \mathrm{ri}(C)$ and $\alpha \geq 0$, the vector $x + \alpha y$ is in $\mathrm{ri}(C)$, which holds in particular for some $x \in \mathrm{ri}(C)$ [note that $\mathrm{ri}(C)$ is nonempty by Prop. 1.4.1(b)].

Conversely, let y be such that there exists a vector $x \in \mathrm{ri}(C)$ with $x + \alpha y \in \mathrm{ri}(C)$ for all $\alpha \geq 0$. Hence, there exists a vector $x \in \mathrm{cl}(C)$ with $x + \alpha y \in \mathrm{cl}(C)$ for all $\alpha \geq 0$, which, by part (b) of the Recession Cone Theorem (cf. Prop. 1.5.1), implies that $y \in R_{\mathrm{cl}(C)}$. Using part (a), it follows that $y \in R_{\mathrm{ri}(C)}$, completing the proof.

(c) Using Exercise 1.38(c) and the assumption that $C \subset \overline{C}$ [which implies that $C \subset \mathrm{cl}(\overline{C})$], we have

$$R_C \subset R_{\mathrm{cl}(\overline{C})} = R_{\mathrm{ri}(\overline{C})} = R_{\overline{C}},$$

where the equalities follow from part (a) and the assumption that $\overline{C} = \mathrm{ri}(\overline{C})$.

To see that the inclusion $R_C \subset R_{\overline{C}}$ can fail when $\overline{C} \neq \mathrm{ri}(\overline{C})$, consider the sets

$$C = \big\{(x_1, x_2) \mid x_1 \geq 0,\ 0 < x_2 < 1\big\}, \qquad \overline{C} = \big\{(x_1, x_2) \mid x_1 \geq 0,\ 0 \leq x_2 < 1\big\},$$

for which we have $C \subset \overline{C}$ and

$$R_C = \big\{(x_1, x_2) \mid x_1 \geq 0,\ x_2 = 0\big\}, \qquad R_{\overline{C}} = \big\{(0, 0)\big\},$$

showing that R_C is not a subset of $R_{\overline{C}}$.

1.40

This exercise is a refinement of Prop. 1.5.6. Let $\{X_k\}$ and $\{C_k\}$ be sequences of closed convex subsets of $\Re^n$, such that the intersection

$$X = \cap_{k=0}^{\infty} X_k$$

is specified by linear inequality constraints as in Prop. 1.5.6. Assume that:

(1) $X_{k+1} \subset X_k$ and $C_{k+1} \subset C_k$ for all k.

(2) $X_k \cap C_k$ is nonempty for all k.

(3) We have

$$R_X = L_X, \qquad R_X \cap R_C \subset L_C,$$

where

$$R_X = \cap_{k=0}^{\infty} R_{X_k}, \qquad L_X = \cap_{k=0}^{\infty} L_{X_k},$$

$$R_C = \cap_{k=0}^{\infty} R_{C_k}, \qquad L_C = \cap_{k=0}^{\infty} L_{C_k}.$$

Then the intersection $\cap_{k=0}^{\infty}(X_k \cap C_k)$ is nonempty. *Hint*: Consider the sets $\overline{C}_k = X_k \cap C_k$ and the intersection $X \cap (\cap_{k=0}^{\infty} \overline{C}_k)$. Apply Prop. 1.5.6.

Solution: For each k, consider the set $\overline{C}_k = X_k \cap C_k$. Note that $\{\overline{C}_k\}$ is a sequence of nonempty closed convex sets and X is specified by linear inequality constraints. We will show that, under the assumptions given in this exercise, the assumptions of Prop. 1.5.6 are satisfied, thus showing that the intersection $X \cap (\cap_{k=0}^{\infty} \overline{C}_k)$ [which is equal to the intersection $\cap_{k=0}^{\infty}(X_k \cap C_k)$] is nonempty.

Since $X_{k+1} \subset X_k$ and $C_{k+1} \subset C_k$ for all k, it follows that

$$\overline{C}_{k+1} \subset \overline{C}_k, \qquad \forall\, k,$$

showing that assumption (1) of Prop. 1.5.6 is satisfied. Similarly, since by assumption $X_k \cap C_k$ is nonempty for all k, we have that, for all k, the set

$$X \cap \overline{C}_k = X \cap X_k \cap C_k = X_k \cap C_k,$$

is nonempty, showing that assumption (2) is satisfied. Finally, let R denote the set $R = \cap_{k=0}^{\infty} R_{\overline{C}_k}$. Since by assumption $\overline{C}_k$ is nonempty for all k, we have, by part (e) of the Recession Cone Theorem, that $R_{\overline{C}_k} = R_{X_k} \cap R_{C_k}$ implying that

$$\begin{aligned} R &= \cap_{k=0}^{\infty} R_{\overline{C}_k} \\ &= \cap_{k=0}^{\infty}(R_{X_k} \cap R_{C_k}) \\ &= \left(\cap_{k=0}^{\infty} R_{X_k}\right) \cap \left(\cap_{k=0}^{\infty} R_{C_k}\right) \\ &= R_X \cap R_C. \end{aligned}$$

Similarly, letting L denote the set $L = \cap_{k=0}^{\infty} L_{\overline{C}_k}$, it can be seen that $L = L_X \cap L_C$. Since, by assumption $R_X \cap R_C \subset L_C$, it follows that

$$R_X \cap R = R_X \cap R_C \subset L_C,$$

which, in view of the assumption that $R_X = L_X$, implies that

$$R_X \cap R \subset L_C \cap L_X = L,$$

showing that assumption (3) of Prop. 1.5.6 is satisfied, and thus proving that the intersection $X \cap (\cap_{k=0}^{\infty} \overline{C}_k)$ is nonempty.

1.41

Let C be a nonempty convex subset of $\Re^n$ and let A be an $m \times n$ matrix. Show that if $R_{\mathrm{cl}(C)} \cap N(A) = \{0\}$, then

$$\mathrm{cl}(A \cdot C) = A \cdot \mathrm{cl}(C), \qquad A \cdot R_{\mathrm{cl}(C)} = R_{A \cdot \mathrm{cl}(C)}.$$

Give an example showing that $A \cdot R_{\mathrm{cl}(C)}$ and $R_{A \cdot \mathrm{cl}(C)}$ may differ when $R_{\mathrm{cl}(C)} \cap N(A) \neq \{0\}$.

Solution: Let y be in the closure of $A \cdot C$. We will show that $y = Ax$ for some $x \in \mathrm{cl}(C)$. For every $\epsilon > 0$, the set

$$C_\epsilon = \mathrm{cl}(C) \cap \left\{x \mid \|y - Ax\| \leq \epsilon\right\}$$

is closed. Since $A\cdot C \subset A\cdot \mathrm{cl}(C)$ and $y \in \mathrm{cl}(A\cdot C)$, it follows that y is in the closure of $A \cdot \mathrm{cl}(C)$, so that C_ϵ is nonempty for every $\epsilon > 0$. Furthermore, the recession cone of the set $\big\{x \mid \|Ax - y\| \leq \epsilon\big\}$ coincides with the null space $N(A)$, so that $R_{C_\epsilon} = R_{\mathrm{cl}(C)} \cap N(A)$. By assumption we have $R_{\mathrm{cl}(C)} \cap N(A) = \{0\}$, and by part (c) of the Recession Cone Theorem (cf. Prop. 1.5.1), it follows that C_ϵ is bounded for every $\epsilon > 0$. Now, since the sets C_ϵ are nested nonempty compact sets, their intersection $\cap_{\epsilon>0} C_\epsilon$ is nonempty. For any x in this intersection, we have $x \in \mathrm{cl}(C)$ and $Ax - y = 0$, showing that $y \in A \cdot \mathrm{cl}(C)$. Hence, $\mathrm{cl}(A \cdot C) \subset A \cdot \mathrm{cl}(C)$. The converse $A \cdot \mathrm{cl}(C) \subset \mathrm{cl}(A \cdot C)$ is clear, since for any $x \in \mathrm{cl}(C)$ and sequence $\{x_k\} \subset C$ converging to x, we have $Ax_k \to Ax$, showing that $Ax \in \mathrm{cl}(A \cdot C)$. Therefore,

$$\mathrm{cl}(A \cdot C) = A \cdot \mathrm{cl}(C). \tag{1.14}$$

We now show that $A \cdot R_{\mathrm{cl}(C)} = R_{A\cdot\mathrm{cl}(C)}$. Let $y \in A \cdot R_{\mathrm{cl}(C)}$. Then, there exists a vector $u \in R_{\mathrm{cl}(C)}$ such that $Au = y$, and by the definition of $R_{\mathrm{cl}(C)}$, there is a vector $x \in \mathrm{cl}(C)$ such that $x + \alpha u \in \mathrm{cl}(C)$ for every $\alpha \geq 0$. Therefore, $Ax + \alpha Au \in A \cdot \mathrm{cl}(C)$ for every $\alpha \geq 0$, which, together with $Ax \in A \cdot \mathrm{cl}(C)$ and $Au = y$, implies that y is a direction of recession of the closed set $A \cdot \mathrm{cl}(C)$ [cf. Eq. (1.14)]. Hence, $A \cdot R_{\mathrm{cl}(C)} \subset R_{A\cdot\mathrm{cl}(C)}$.

Conversely, let $y \in R_{A\cdot\mathrm{cl}(C)}$. We will show that $y \in A \cdot R_{\mathrm{cl}(c)}$. This is true if $y = 0$, so assume that $y \neq 0$. By definition of direction of recession, there is a vector $z \in A\cdot\mathrm{cl}(C)$ such that $z + \alpha y \in A\cdot\mathrm{cl}(C)$ for every $\alpha \geq 0$. Let $x \in \mathrm{cl}(C)$ be such that $Ax = z$, and for every positive integer k, let $x_k \in \mathrm{cl}(C)$ be such that $Ax_k = z + ky$. Since $y \neq 0$, the sequence $\{Ax_k\}$ is unbounded, implying that $\{x_k\}$ is also unbounded (if $\{x_k\}$ were bounded, then $\{Ax_k\}$ would be bounded, a contradiction). Because $x_k \neq x$ for all k, we can define

$$u_k = \frac{x_k - x}{\|x_k - x\|}, \qquad \forall\ k.$$

Let u be a limit point of $\{u_k\}$, and note that $u \neq 0$. It can be seen that u is a direction of recession of $\mathrm{cl}(C)$ [this can be done similar to the proof of part (c) of the Recession Cone Theorem (cf. Prop. 1.5.1)]. By taking an appropriate subsequence if necessary, we may assume without loss of generality that $\lim_{k\to\infty} u_k = u$. Then, by the choices of u_k and x_k, we have

$$Au = \lim_{k\to\infty} Au_k = \lim_{k\to\infty} \frac{Ax_k - Ax}{\|x_k - x\|} = \lim_{k\to\infty} \frac{k}{\|x_k - x\|} y,$$

implying that $\lim_{k\to\infty} \frac{k}{\|x_k - x\|}$ exists. Denote this limit by λ. If $\lambda = 0$, then u is in the null space $N(A)$, implying that $u \in R_{\mathrm{cl}(C)} \cap N(A)$. By the given condition $R_{\mathrm{cl}(C)} \cap N(A) = \{0\}$, we have $u = 0$ contradicting the fact $u \neq 0$. Thus, λ is positive and $Au = \lambda y$, so that $A(u/\lambda) = y$. Since $R_{\mathrm{cl}(C)}$ is a cone [part (a) of the Recession Cone Theorem] and $u \in R_{\mathrm{cl}(C)}$, the vector u/λ is in $R_{\mathrm{cl}(C)}$, so that y belongs to $A \cdot R_{\mathrm{cl}(C)}$. Hence, $R_{A\cdot\mathrm{cl}(C)} \subset A \cdot R_{\mathrm{cl}(C)}$, completing the proof.

As an example showing that $A\cdot R_{\mathrm{cl}(C)}$ and $R_{A\cdot\mathrm{cl}(C)}$ may differ when $R_{\mathrm{cl}(C)} \cap N(A) \neq \{0\}$, consider the set

$$C = \big\{(x_1, x_2) \mid x_1 \in \Re,\ x_2 \geq x_1^2\big\},$$

and the linear transformation A that maps $(x_1, x_2) \in \Re^2$ into $x_1 \in \Re$. Then, C is closed and its recession cone is

$$R_C = \big\{(x_1, x_2) \mid x_1 = 0,\ x_2 \geq 0\big\},$$

so that $A \cdot R_C = \{0\}$, where 0 is scalar. On the other hand, $A \cdot C$ coincides with $\Re$, so that $R_{A \cdot C} = \Re \neq A \cdot R_C$.

1.42

Let C be a nonempty convex subset of $\Re^n$. Show the following refinement of Prop. 1.5.8(a) and Exercise 1.41: if A is an $m \times n$ matrix and $R_{\mathrm{cl}(C)} \cap N(A)$ is a subspace of the lineality space of $\mathrm{cl}(C)$, then

$$\mathrm{cl}(A \cdot C) = A \cdot \mathrm{cl}(C), \qquad A \cdot R_{\mathrm{cl}(C)} = R_{A \cdot \mathrm{cl}(C)}.$$

Solution: Let S be defined by

$$S = R_{\mathrm{cl}(C)} \cap N(A),$$

and note that S is a subspace of $L_{\mathrm{cl}(C)}$ by the given assumption. Then, by Lemma 1.5.4, we have

$$\mathrm{cl}(C) = \big(\mathrm{cl}(C) \cap S^\perp\big) + S,$$

so that the images of $\mathrm{cl}(C)$ and $\mathrm{cl}(C) \cap S^\perp$ under A coincide [since $S \subset N(A)$], i.e.,

$$A \cdot \mathrm{cl}(C) = A \cdot \big(\mathrm{cl}(C) \cap S^\perp\big). \tag{1.15}$$

Because $A \cdot C \subset A \cdot \mathrm{cl}(C)$, we have

$$\mathrm{cl}(A \cdot C) \subset \mathrm{cl}\big(A \cdot \mathrm{cl}(C)\big),$$

which in view of Eq. (1.15) gives

$$\mathrm{cl}(A \cdot C) \subset \mathrm{cl}\Big(A \cdot \big(\mathrm{cl}(C) \cap S^\perp\big)\Big).$$

Define

$$\overline{C} = \mathrm{cl}(C) \cap S^\perp$$

so that the preceding relation becomes

$$\mathrm{cl}(A \cdot C) \subset \mathrm{cl}(A \cdot \overline{C}). \tag{1.16}$$

The recession cone of $\overline{C}$ is given by

$$R_{\overline{C}} = R_{\mathrm{cl}(C)} \cap S^\perp, \tag{1.17}$$

[cf. part (e) of the Recession Cone Theorem, Prop. 1.5.1], for which, since $S = R_{\mathrm{cl}(C)} \cap N(A)$, we have

$$R_{\overline{C}} \cap N(A) = S \cap \ S^{\perp} = \{0\}.$$

Therefore, by Prop. 1.5.8, the set $A \cdot \overline{C}$ is closed, implying that $\mathrm{cl}(A \cdot \overline{C}) = A \cdot \overline{C}$. By the definition of $\overline{C}$, we have $A \cdot \overline{C} \subset A \cdot \mathrm{cl}(C)$, implying that $\mathrm{cl}(A \cdot \overline{C}) \subset A \cdot \mathrm{cl}(C)$ which together with Eq. (1.16) yields $\mathrm{cl}(A \cdot C) \subset A \cdot \mathrm{cl}(C)$. The converse $A \cdot \mathrm{cl}(C) \subset \mathrm{cl}(A \cdot C)$ is clear, since for any $x \in \mathrm{cl}(C)$ and sequence $\{x_k\} \subset C$ converging to x, we have $Ax_k \to Ax$, showing that $Ax \in \mathrm{cl}(A \cdot C)$. Therefore,

$$\mathrm{cl}(A \cdot C) = A \cdot \mathrm{cl}(C). \tag{1.18}$$

We next show that $A \cdot R_{\mathrm{cl}(C)} = R_{A \cdot \mathrm{cl}(C)}$. Let $y \in A \cdot R_{\mathrm{cl}(C)}$. Then, there exists a vector $u \in R_{\mathrm{cl}(C)}$ such that $Au = y$, and by the definition of $R_{\mathrm{cl}(C)}$, there is a vector $x \in \mathrm{cl}(C)$ such that $x + \alpha u \in \mathrm{cl}(C)$ for every $\alpha \geq 0$. Therefore, $Ax + \alpha Au \in A\mathrm{cl}(C)$ for some $x \in \mathrm{cl}(C)$ and for every $\alpha \geq 0$, which together with $Ax \in A \cdot \mathrm{cl}(C)$ and $Au = y$ implies that y is a recession direction of the closed set $A \cdot \mathrm{cl}(C)$ [Eq. (1.18)]. Hence, $A \cdot R_{\mathrm{cl}(C)} \subset R_{A \cdot \mathrm{cl}(C)}$.

Conversely, in view of Eq. (1.15) and the definition of $\overline{C}$, we have

$$R_{A \cdot \mathrm{cl}(C)} = R_{A \cdot \overline{C}}.$$

Since $R_{\overline{C}} \cap N(A) = \{0\}$ and $\overline{C}$ is closed, by Exercise 1.41, it follows that

$$R_{A \cdot \overline{C}} = A \cdot R_{\overline{C}},$$

which combined with Eq. (1.17) implies that

$$A \cdot R_{\overline{C}} \subset A \cdot R_{\mathrm{cl}(C)}.$$

The preceding three relations yield $R_{A \cdot \mathrm{cl}(C)} \subset A \cdot R_{\mathrm{cl}(C)}$, completing the proof.

1.43 (Recession Cones of Vector Sums)

This exercise is a refinement of Prop. 1.5.9.

(a) Let $C_1, \ldots, C_m$ be nonempty closed convex subsets of $\Re^n$ such that the equality $y_1 + \cdots + y_m = 0$ with $y_i \in R_{C_i}$ implies that each y_i belongs to the lineality space of C_i. Then, the vector sum $C_1 + \cdots + C_m$ is a closed set and

$$R_{C_1 + \cdots + C_m} = R_{C_1} + \cdots + R_{C_m}.$$

(b) Show the following extension of part (a) to nonclosed sets: Let $C_1, \ldots, C_m$ be nonempty convex subsets of $\Re^n$ such that the equality $y_1 + \cdots + y_m = 0$ with $y_i \in R_{\mathrm{cl}(C_i)}$ implies that each y_i belongs to the lineality space of $\mathrm{cl}(C_i)$. Then, we have

$$\mathrm{cl}(C_1 + \cdots + C_m) = \mathrm{cl}(C_1) + \cdots + \mathrm{cl}(C_m),$$

$$R_{\mathrm{cl}(C_1+\cdots+C_m)} = R_{\mathrm{cl}(C_1)} + \cdots + R_{\mathrm{cl}(C_m)}.$$

Solution: (a) Let C be the Cartesian product $C_1 \times \cdots \times C_m$. Then, by Exercise 1.37, C is closed, and its recession cone and lineality space are given by

$$R_C = R_{C_1} \times \cdots \times R_{C_m}, \qquad L_C = L_{C_1} \times \cdots \times L_{C_m}.$$

Let A be a linear transformation that maps $(x_1, \ldots, x_m) \in \Re^{mn}$ into $x_1 + \cdots + x_m \in \Re^n$. The null space of A is the set of all $(y_1, \ldots, y_m)$ such that $y_1 + \cdots + y_m = 0$. The intersection $R_C \cap N(A)$ consists of points $(y_1, \ldots, y_m)$ such that $y_1 + \cdots + y_m = 0$ with $y_i \in R_{C_i}$ for all i. By the given condition, every vector $(y_1, \ldots, y_m)$ in the intersection $R_C \cap N(A)$ is such that $y_i \in L_{C_i}$ for all i, implying that $(y_1, \ldots, y_m)$ belongs to the lineality space L_C. Thus, $R_C \cap N(A) \subset L_C \cap N(A)$. On the other hand by definition of the lineality space, we have $L_C \subset R_C$, so that $L_C \cap N(A) \subset R_C \cap N(A)$. Therefore, $R_C \cap N(A) = L_C \cap N(A)$, implying that $R_C \cap N(A)$ is a subspace of L_C. By Exercise 1.42, the set $A \cdot C$ is closed and $R_{A \cdot C} = A \cdot R_C$. Since $A \cdot C = C_1 + \cdots + C_m$, the assertions of part (a) follow.

(b) The proof is similar to that of part (a). Let C be the Cartesian product $C_1 \times \cdots \times C_m$. Then, by Exercise 1.37(a),

$$\mathrm{cl}(C) = \mathrm{cl}(C_1) \times \cdots \times \mathrm{cl}(C_m), \tag{1.19}$$

and its recession cone and lineality space are given by

$$R_{\mathrm{cl}(C)} = R_{\mathrm{cl}(C_1)} \times \cdots \times R_{\mathrm{cl}(C_m)}, \tag{1.20}$$

$$L_{\mathrm{cl}(C)} = L_{\mathrm{cl}(C_1)} \times \cdots \times L_{\mathrm{cl}(C_m)}.$$

Let A be a linear transformation that maps $(x_1, \ldots, x_m) \in \Re^{mn}$ into $x_1 + \cdots + x_m \in \Re^n$. Then, the intersection $R_{\mathrm{cl}}(C) \cap N(A)$ consists of points $(y_1, \ldots, y_m)$ such that $y_1 + \cdots + y_m = 0$ with $y_i \in R_{\mathrm{cl}(C_i)}$ for all i. By the given condition, every vector $(y_1, \ldots, y_m)$ in the intersection $R_{\mathrm{cl}(C)} \cap N(A)$ is such that $y_i \in L_{\mathrm{cl}(C_i)}$ for all i, implying that $(y_1, \ldots, y_m)$ belongs to the lineality space $L_{\mathrm{cl}(C)}$. Thus, $R_{\mathrm{cl}(C)} \cap N(A) \subset L_{\mathrm{cl}(C)} \cap N(A)$. On the other hand by definition of the lineality space, we have $L_{\mathrm{cl}(C)} \subset R_{\mathrm{cl}(C)}$, so that $L_{\mathrm{cl}(C)} \cap N(A) \subset R_{\mathrm{cl}(C)} \cap N(A)$. Hence, $R_{\mathrm{cl}(C)} \cap N(A) = L_{\mathrm{cl}(C)} \cap N(A)$, implying that $R_{\mathrm{cl}(C)} \cap N(A)$ is a subspace of $L_{\mathrm{cl}(C)}$. By Exercise 1.42, we have $\mathrm{cl}(A \cdot C) = A \cdot \mathrm{cl}(C)$ and $R_{A \cdot \mathrm{cl}(C)} = A \cdot R_{\mathrm{cl}(C)}$, from which by using the relation $A \cdot C = C_1 + \cdots + C_m$, and Eqs. (1.19) and (1.20), we obtain

$$\mathrm{cl}(C_1 + \cdots + C_m) = \mathrm{cl}(C_1) + \cdots + \mathrm{cl}(C_m),$$

$$R_{\mathrm{cl}(C_1+\cdots+C_m)} = R_{\mathrm{cl}(C_1)} + \cdots + R_{\mathrm{cl}(C_m)}.$$

1.44

Let $C_1, \ldots, C_m$ be nonempty subsets of $\Re^n$ that are specified by convex quadratic inequalities, i.e., for all $i = 1, \ldots, n$,

$$C_i = \bigl\{x \mid x'Q_{ij}x + a_{ij}'x + b_{ij} \le 0,\ j = 1, \ldots, r_i\bigr\},$$

where Q_{ij} are symmetric positive semidefinite $n \times n$ matrices, a_{ij} are vectors in $\Re^n$, and b_{ij} are scalars. Show that the vector sum $C_1 + \cdots + C_m$ is a closed set.

Solution: Let C be the Cartesian product $C_1 \times \cdots \times C_m$ viewed as a subset of $\Re^{mn}$, and let A be the linear transformation that maps a vector $(x_1, \ldots, x_m) \in \Re^{mn}$ into $x_1 + \cdots + x_m$. Note that set C can be written as

$$C = \bigl\{x = (x_1, \ldots, x_m) \mid x'\overline{Q}_{ij}x + \overline{a}_{ij}'x + b_{ij} \le 0,\ i = 1, \ldots, m,\ j = 1, \ldots, r_i\bigr\},$$

where the $\overline{Q}_{ij}$ are appropriately defined symmetric positive semidefinite $mn \times mn$ matrices and the $\overline{a}_{ij}$ are appropriately defined vectors in $\Re^{mn}$. Hence, the set C is specified by convex quadratic inequalities. Thus, we can use Prop. 1.5.8(c) to assert that the set $AC = C_1 + \cdots + C_m$ is closed.

1.45 (Set Intersection and Helly's Theorem)

Show that the conclusions of Props. 1.5.5 and 1.5.6 hold if the assumption that the sets C_k are nonempty and nested is replaced by the weaker assumption that any subcollection of $n + 1$ (or fewer) sets from the sequence $\{C_k\}$ has nonempty intersection. *Hint*: Consider the sets $\overline{C}_k$ given by

$$\overline{C}_k = \cap_{i=1}^k C_i, \qquad \forall\ k = 1, 2, \ldots,$$

and use Helly's Theorem (Exercise 1.25) to show that they are nonempty.

Solution: Helly's Theorem implies that the sets $\overline{C}_k$ defined in the hint are nonempty. These sets are also nested and satisfy the assumptions of Props. 1.5.5 and 1.5.6. Therefore, the intersection $\cap_{i=1}^\infty \overline{C}_i$ is nonempty. Since

$$\cap_{i=1}^\infty \overline{C}_i \subset \cap_{i=1}^\infty C_i,$$

the result follows.