Convex Optimization Theory

Chapter 2

Exercises and Solutions: Extended Version

Dimitri P. Bertsekas

Massachusetts Institute of Technology

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

CHAPTER 2: EXERCISES AND SOLUTIONS†

SECTION 2.1: Extreme Points

2.1

Show by example that the set of extreme points of a nonempty compact set need not be closed. *Hint*: Consider a line segment $C_1 = \big\{(x_1, x_2, x_3) \mid x_1 = 0, x_2 = 0, -1 \le x_3 \le 1\big\}$ and a circular disk $C_2 = \big\{(x_1, x_2, x_3) \mid (x_1 - 1)^2 + x_2^2 \le 1,\ x_3 = 0\big\}$, and verify that the set $\text{conv}(C_1 \cup C_2)$ is compact, while its set of extreme points is not closed.

Solution: For the sets C_1 and C_2 as given in the hint, the set $C_1 \cup C_2$ is compact, and its convex hull is also compact by Prop. 1.2.2. The set of extreme points of $\text{conv}(C_1 \cup C_2)$ is not closed, since it consists of the two end points of the line segment C_1, namely $(0, 0, -1)$ and $(0, 0, 1)$, and all the points $x = (x_1, x_2, x_3)$ such that

$$x \ne 0, \qquad (x_1 - 1)^2 + x_2^2 = 1, \qquad x_3 = 0.$$

2.2 (Krein-Milman Theorem)

Show that a convex and compact subset of $\Re^n$ is equal to the convex hull of its extreme points.

Solution: By convexity, C contains the convex hull of its extreme points. To show the reverse inclusion, we use induction on the dimension of the space. On the real line, a compact convex set C is a line segment whose endpoints are the extreme points of C, so every point in C is a convex combination of the two endpoints. Suppose now that every vector in a compact and convex subset of $\Re^{n-1}$ can be represented as a convex combination of extreme points of the set. We will show that the same is true for compact and convex subsets of $\Re^n$.

---

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

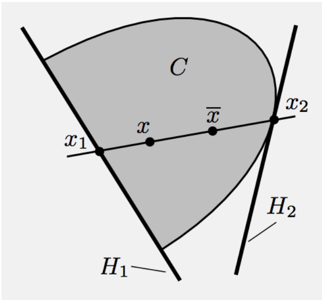

Figure 2.1. Construction used in the induction proof of the Krein-Milman Theorem (Exercise 2.2): any vector x of a convex and compact set C can be represented as a convex combination of extreme points of C. If $\overline{x}$ is another point in C, the points x_1 and x_2 shown can be represented as convex combinations of extreme points of the lower dimensional convex and compact sets $C \cap H_1$ and $C \cap H_2$, which are also extreme points of C by Prop. 2.1.1.

Let C be a compact and convex subset of $\Re^n$, and choose any $x \in C$. If x is the only point in C, it is an extreme point and we are done, so assume that $\overline{x}$ is another point in C, and consider the line that passes through x and $\overline{x}$. Since C is compact, the intersection of this line and C is a compact line segment whose endpoints, say x_1 and x_2, belong to the relative boundary of C. Let H_1 be a hyperplane that passes through x_1 and contains C in one of its closed halfspaces. Similarly, let H_2 be a hyperplane that passes through x_2 and contains C in one of its closed halfspaces (see Fig. 2.1). The intersections $C \cap H_1$ and $C \cap H_2$ are compact convex sets that lie in the hyperplanes H_1 and H_2, respectively. By viewing H_1 and H_2 as $(n-1)$-dimensional spaces, and by using the induction hypothesis, we see that each of the sets $C \cap H_1$ and $C \cap H_2$ is the convex hull of its extreme points. Hence, x_1 is a convex combination of some extreme points of $C \cap H_1$, and x_2 is a convex combination of some extreme points of $C \cap H_2$. By Prop. 2.1.1, all the extreme points of $C \cap H_1$ and all the extreme points of $C \cap H_2$ are also extreme points of C, so both x_1 and x_2 are convex combinations of some extreme points of C. Since x lies in the line segment connecting x_1 and x_2, it follows that x is a convex combination of some extreme points of C, showing that C is contained in the convex hull of the extreme points of C.

2.3

Let C be a nonempty convex subset of $\Re^n$, and let A be an $m \times n$ matrix with linearly independent columns. Show that a vector $x \in C$ is an extreme point of C if and only if Ax is an extreme point of the image AC. Show by example that if the columns of A are linearly dependent, then Ax can be an extreme point of AC, for some non-extreme point x of C.

Solution: Suppose that x is not an extreme point of C. Then $x = \alpha x_1 + (1-\alpha)x_2$ for some $x_1, x_2 \in C$ with $x_1 \neq x$ and $x_2 \neq x$, and a scalar $\alpha \in (0,1)$, so that $Ax = \alpha Ax_1 + (1-\alpha)Ax_2$. Since the columns of A are linearly independent, we have $Ay_1 = Ay_2$ if and only if $y_1 = y_2$. Therefore, $Ax_1 \neq Ax$ and $Ax_2 \neq Ax$, implying that Ax is a convex combination of two distinct points in AC, i.e., Ax is not an extreme point of AC.

Suppose now that Ax is not an extreme point of AC, so that $Ax = \alpha Ax_1 + (1-\alpha)Ax_2$ for some $x_1, x_2 \in C$ with $Ax_1 \neq Ax$ and $Ax_2 \neq Ax$, and a scalar $\alpha \in (0,1)$. Then, $A\big(x - \alpha x_1 - (1-\alpha)x_2\big) = 0$ and since the columns of A are linearly independent, it follows that $x = \alpha x_1 - (1-\alpha)x_2$. Furthermore, because $Ax_1 \neq Ax$ and $Ax_2 \neq Ax$, we must have $x_1 \neq x$ and $x_2 \neq x$, implying that x is not an extreme point of C.

As an example showing that if the columns of A are linearly dependent, then Ax can be an extreme point of AC, for some non-extreme point x of C, consider the 1×2 matrix $A = [1\ 0]$, whose columns are linearly dependent. The polyhedral set C given by

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

has two extreme points, (0,0) and (0,1). Its image $AC \subset \Re$ is given by

$$AC = \{x_1 \mid x_1 \geq 0\},$$

whose unique extreme point is $x_1 = 0$. The point $x = (0, 1/2) \in C$ is not an extreme point of C, while its image $Ax = 0$ is an extreme point of AC. Actually, all the points in C on the line segment connecting (0,0) and (0,1), except for (0,0) and (0,1), are non-extreme points of C that are mapped under A into the extreme point 0 of AC.

2.4

Let C be a nonempty closed convex subset of $\Re^n$. Show that the following are equivalent.

(i) All boundary points of C are extreme points of C.

(ii) Every hyperplane that supports C at some point intersects C only at that point.

(iii) Every line intersects the boundary of C at no more than two points.

Solution: The result is clearly true if C consists of a single point, so assume that C consists of more than one point.

We first show that (i) implies (ii). Assume that all boundary points of C are extreme points. If there is a hyperplane that supports C and intersects C at two distinct points, the entire line segment connecting the two points would lie on the boundary of C, but the midpoint of this line segment would not be an extreme point - a contradiction.

Next we show that (ii) implies (iii). Assume that every hyperplane that supports C at some point intersects C only at that point. Suppose that there is a line that intersects the boundary of C at three distinct boundary points x_1, x_2, x_3, with x_2 being the midpoint. Consider a hyperplane H that supports C at x_2, i.e., a vector $a \neq 0$ such that

$$a'x \geq a'x_2, \qquad \forall\ x \in C.$$

Then since by the hypothesis, H intersects C only at x_2, we must have $a'x_1 > a'x_2$ and $a'x_3 > a'x_2$, which is a contradiction since x_2 lies strictly between x_1 and x_3.

Finally, we show that (iii) implies (i). Assume that every line intersects the boundary of C at no more than two points. If there is a boundary point x_2 that is not extreme and therefore lies strictly between two points $x_1, x_3 \in C$, then either x_1 or x_3 must be an interior point, for otherwise the line that passes through x_1, x_2, x_3 would contain more than two boundary points. Thus, by the Line Segment Principle (Prop. 1.3.1), every point that lies strictly between x_1 and x_3, including x_2, is an interior point of C. This contradicts the hypothesis that x_2 is a boundary point of C.

2.5 (Matrix Inequalities)

Let A be a symmetric $n \times n$ matrix with components denoted a_{ij} and eigenvalues denoted $\lambda_1, \ldots, \lambda_n$, and let Λ_A be the set of all vectors of $\Re^n$ obtained by permutations of these eigenvalues.

(a) Let C be a convex set that contains Λ_A, and let $f : C \mapsto \Re^n$ be a convex function. Show that for any orthonormal set of vectors $v_1, \ldots, v_n$ in $\Re^n$, we have

$$f(v_1' A v_1, \ldots, v_n' A v_n) \leq \max_{(\xi_1, \ldots, \xi_n) \in \Lambda_A} f(\xi_1, \ldots, \xi_n).$$

Hint: Let S be the doubly stochastic matrix with components $s_{ij} = (v_i' u_j)^2$, where $u_1, \ldots, u_n$ are orthonormal eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$. Show that $v = S\lambda$, where

$$v = (v_1' A v_1, \ldots, v_n' A v_n), \qquad \lambda = (\lambda_1, \ldots, \lambda_n),$$

and use the Birkhoff-von Neumann Theorem.

(b) Let A be positive semidefinite. Show that for any orthonormal set of vectors $v_1, \ldots, v_n$ in $\Re^n$, we have

$$\det A = \lambda_1 \cdots \lambda_n \leq v_1' A v_1 \cdots v_n' A v_n.$$

Furthermore, the inequality is sharp in the sense that it is satisfied as an equality for some orthonormal set of vectors. *Hint*: Use part (a) with $f(x_1, \ldots, x_n) = -(x_1 \cdots x_n)^{1/2}$, and C equal to the nonnegative orthant.

(c) *(Hadamard's Determinant Inequality)* We have

$$(\det A)^2 \leq (a_{11}^2 + \cdots + a_{n1}^2) \cdots (a_{1n}^2 + \cdots + a_{nn}^2).$$

Furthermore, if in addition A is positive semidefinite, we have

$$\det A \leq a_{11} \cdots a_{nn}.$$

Solution: (a) Let $u_1, \ldots, u_n$ be orthonormal eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$. The orthogonality of $u_1, \ldots, u_n$ implies that

$$v_i = (v_i' u_1) u_1 + \cdots + (v_i' u_n) u_n, \qquad i = 1, \ldots, n.$$

Using this relation, it is straightforward to verify that

$$v = S\lambda,$$

where

$$v = (v_1' A v_1, \ldots, v_n' A v_n), \qquad \lambda = (\lambda_1, \ldots, \lambda_n),$$

and S is the $n \times n$ matrix with components $s_{ij} = (v_i' u_j)^2$ for all i and j. We now note that S is a doubly stochastic matrix. The reason is that we have for each i,

$$\|v_i\|^2 = \left\|(v_i' u_1) u_1 + \cdots + (v_i' u_n) u_n\right\|^2,$$

so that by using the orthonormality of $u_1, \ldots, u_n$, we have

$$\|v_i\|^2 = (v_i' u_1)^2 + \cdots + (v_i' u_n)^2.$$

This implies that S is doubly stochastic, since $\|v_i\| = 1$ by assumption, and the ith row of the matrix S is $\big((v_i' u_1)^2, \ldots, (v_i' u_n)^2\big)$.

The Birkhoff-von Neumann Theorem asserts that S can be expressed as a convex combination of permutation matrices, i.e., there exist $\mu_j \geq 0$, $j = 1, \ldots, m$, with $\sum_{j=1}^m \mu_j = 1$, and such that

$$S = \mu_1 P_1 + \cdots + \mu_m P_m,$$

where $P_1, \ldots, P_m$ are permutation matrices. Hence,

$$v = S\lambda = \mu_1 (P_1 \lambda) + \cdots + \mu_m (P_m \lambda).$$

Since the vectors $P_j \lambda$, $j = 1, \ldots, m$, belong to Λ_A, they also belong to C. Since v is a convex combination of $P_j \lambda$, $j = 1, \ldots, m$, it follows that $v \in C$. Thus, using the convexity of f, we have

$$f(v) \leq \mu_1 f(P_1 \lambda) + \cdots + \mu_m f(P_m \lambda) \leq \max_{(\xi_1, \ldots, \xi_n) \in \Lambda_A} f(\xi_1, \ldots, \xi_n).$$

(b) The inequality follows from part (a) and the hint. The inequality is satisfied as an equality if the vectors $v_1, \ldots, v_n$ are normalized eigenvectors corresponding to $\lambda_1, \ldots, \lambda_n$.

(c) Let $B = A'A$. We apply part (b) to B with the orthonormal vectors being the unit vectors $e_1, \ldots, e_n$ of $\Re^n$. We obtain

$$\det B \leq e_1' B e_1 \cdots e_n' B e_n = (a_{11}^2 + \cdots + a_{n1}^2) \cdots (a_{1n}^2 + \cdots + a_{nn}^2),$$

where the last equality can be verified by straightforward calculation. Since $\det B = (\det A)^2$, the desired inequality follows.

If A positive semidefinite, we apply part (b) to A with the orthonormal vectors being the unit vectors $e_1, \ldots, e_n$ of $\Re^n$, to obtain

$$\det A \leq a_{11} \cdots a_{nn}.$$

2.6 (Faces)

Let P be a polyhedral set. For any hyperplane H that passes through a boundary point of P and contains P in one of its halfspaces, we say that the set $F = P \cap H$ is a *face* of P. Show the following:

(a) Each face is a polyhedral set.

(b) Each extreme point of P, viewed as a singleton set, is a face.

(c) If P is not an affine set, there is a face of P whose dimension is $\dim(P) - 1$.

(d) The number of distinct faces of P is finite.

Solution: (a) Let P be a polyhedral set in $\Re^n$, and let $F = P \cap H$ be a face of P, where H is a hyperplane passing through some boundary point $\overline{x}$ of P and containing P in one of its halfspaces. Then H is given by $H = \{x \mid a'x = a'\overline{x}\}$ for some nonzero vector $a \in \Re^n$. By replacing $a'x = a'\overline{x}$ with two inequalities $a'x \leq a'\overline{x}$ and $-a'x \leq -a'\overline{x}$, we see that H is a polyhedral set in $\Re^n$. Since the intersection of two nondisjoint polyhedral sets is a polyhedral set, the set $F = P \cap H$ is polyhedral.

(b) Let P be given by

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

for some vectors $a_j \in \Re^n$ and scalars b_j. Let v be an extreme point of P, and without loss of generality assume that the first n inequalities define v, i.e., the first n of the vectors a_j are linearly independent and such that

$$a_j'v = b_j, \qquad \forall\ j = 1, \ldots, n$$

[cf. Prop. 2.1.4(a)]. Define the vector $a \in \Re^n$, the scalar b, and the hyperplane H as follows

$$a = \frac{1}{n}\sum_{j=1}^{n} a_j, \qquad b = \frac{1}{n}\sum_{j=1}^{n} b_j, \qquad H = \big\{x \mid a'x = b\big\}.$$

Then, we have

$$a'v = b,$$

so that H passes through v. Moreover, for every $x \in P$, we have $a_j'x \leq b_j$ for all j, implying that $a'x \leq b$ for all $x \in P$. Thus, H contains P in one of its halfspaces.

We will next prove that $P \cap H = \{v\}$. We start by showing that for every $\overline{v} \in P \cap H$, we must have

$$a_j'\overline{v} = b_j, \qquad \forall\ j = 1, \ldots, n. \tag{2.1}$$

To arrive at a contradiction, assume that $a_j'\overline{v} < b_j$ for some $\overline{v} \in P \cap H$ and $j \in \{1, \ldots, n\}$. Without loss of generality, we can assume that the strict inequality holds for $j = 1$, so that

$$a_1'\overline{v} < b_1, \qquad a_j'\overline{v} \leq b_j, \qquad \forall\ j = 2, \ldots, n.$$

By multiplying each of the above inequalities with $1/n$ and by summing the obtained inequalities, we obtain

$$\frac{1}{n}\sum_{j=1}^{n} a_j'\overline{v} < \frac{1}{n}\sum_{j=1}^{n} b_j,$$

implying that $a'\overline{v} < b$, which contradicts the fact that $\overline{v} \in H$. Hence, Eq. (2.1) holds, and since the vectors $a_1, \ldots, a_n$ are linearly independent, it follows that $v = \overline{v}$, showing that $P \cap H = \{v\}$.

As discussed in Section 2.1, every extreme point of P is a relative boundary point of P. Since every relative boundary point of P is also a boundary point of P, it follows that every extreme point of P is a boundary point of P. Thus, v is a boundary point of P, and as shown earlier, H passes through v and contains P in one of its halfspaces. By definition, it follows that $P \cap H = \{v\}$ is a face of P.

(c) Since P is not an affine set, it cannot consist of a single point, so we must have $\dim(P) > 0$. Let P be given by

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

for some vectors $a_j \in \Re^n$ and scalars b_j. Also, let A be the matrix with rows a_j' and b be the vector with components b_j, so that

$$P = \{x \mid Ax \le b\}.$$

An inequality $a_j'x \le b_j$ of the system $Ax \le b$ is *redundant* if it is implied by the remaining inequalities in the system. If the system $Ax \le b$ has no redundant inequalities, we say that the system is *nonredundant*. An inequality $a_j'x \le b_j$ of the system $Ax \le b$ is an *implicit equality* if $a_j'x = b_j$ for all x satisfying $Ax \le b$.

By removing the redundant inequalities if necessary, we may assume that the system $Ax \le b$ defining P is nonredundant. Since P is not an affine set, there exists an inequality $a_{j_0}'x \le b_{j_0}$ that is not an implicit equality of the system $Ax \le b$. Consider the set

$$F = \left\{x \in P \mid a_{j_0}'x = b_{j_0}\right\}.$$

Note that $F \ne \emptyset$, since otherwise $a_{j_0}'x \le b_{j_0}$ would be a redundant inequality of the system $Ax \le b$, contradicting our earlier assumption that the system is nonredundant. Note also that every point of F is a boundary point of P. Thus, F is the intersection of P and the hyperplane $\left\{x \mid a_{j_0}'x = b_{j_0}\right\}$ that passes through a boundary point of P and contains P in one of its halfspaces, i.e., F is a face of P. Since $a_{j_0}'x \le b_{j_0}$ is not an implicit equality of the system $Ax \le b$, the dimension of F is $\dim(P) - 1$.

(d) Let P be a polyhedral set given by

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

with $a_j \in \Re^n$ and $b_j \in \Re$, or equivalently

$$P = \{x \mid Ax \le b\},$$

where A is an $r \times n$ matrix and $b \in \Re^r$. We will show that F is a face of P if and only if F is nonempty and

$$F = \{x \in P \mid a_j'x = b_j,\ j \in J\},$$

where $J \subset \{1, \ldots, r\}$. From this it will follow that the number of distinct faces of P is finite.

By removing the redundant inequalities if necessary, we may assume that the system $Ax \leq b$ defining P is nonredundant. Let F be a face of P, so that $F = P \cap H$, where H is a hyperplane that passes through a boundary point of P and contains P in one of its halfspaces. Let $H = \{x \mid\ c'x = c\overline{x}\}$ for a nonzero vector $c \in \Re^n$ and a boundary point $\overline{x}$ of P, so that

$$F = \{x \in P \mid c'x = c\overline{x}\}$$

and

$$c'x \leq c\overline{x}, \qquad \forall\ x \in P.$$

These relations imply that the set of points x such that $Ax \leq b$ and $c'x \leq c\overline{x}$ coincides with P, and since the system $Ax \leq b$ is nonredundant, it follows that $c'x \leq c\overline{x}$ is a redundant inequality of the system $Ax \leq b$ and $c'x \leq c\overline{x}$. Therefore, the inequality $c'x \leq c\overline{x}$ is implied by the inequalities of $Ax \leq b$, so that there exists some $\mu \in \Re^r$ with $\mu \geq 0$ such that

$$\sum_{j=1}^{r} \mu_j a_j = c, \qquad \sum_{j=1}^{r} \mu_j b_j = c'\overline{x}.$$

Let $J = \{j \mid \mu_j > 0\}$. Then, for every $x \in P$, we have

$$c'x = c\overline{x} \quad \Longleftrightarrow \quad \sum_{j \in J} \mu_j a_j'x = \sum_{j \in J} \mu_j b_j \quad \Longleftrightarrow \quad a_j'x = b_j,\ j \in J, \tag{2.2}$$

implying that

$$F = \{x \in P \mid a_j'x = b_j, j \in J\}.$$

Conversely, let F be a nonempty set given by

$$F = \{x \in P \mid a_j'x = b_j, j \in J\},$$

for some $J \subset \{1, \ldots, r\}$. Define

$$c = \sum_{j \in J} a_j, \qquad \beta = \sum_{j \in J} b_j.$$

Then, we have

$$\{x \in P \mid a_j'x = b_j, j \in J\} = \{x \in P \mid c'x = \beta\},$$

[cf. Eq. (2.2) where $\mu_j = 1$ for all $j \in J$]. Let $H = \{x \mid c'x = \beta\}$, so that in view of the preceding relation, we have that $F = P \cap H$. Since every point of F is a boundary point of P, it follows that H passes through a boundary point of P. Furthermore, for every $x \in P$, we have $a_j'x \leq b_j$ for all $j \in J$, implying that $c'x \leq \beta$ for every $x \in P$. Thus, H contains P in one of its halfspaces. Hence, F is a face.

2.7 (Isomorphic Polyhedral Sets)

Let P and Q be polyhedral sets in $\Re^n$ and $\Re^m$, respectively. We say that P and Q are *isomorphic* if there exist affine functions $f : P \mapsto Q$ and $g : Q \mapsto P$ such that

$$x = g\big(f(x)\big), \quad \forall\, x \in P, \qquad y = f\big(g(y)\big), \quad \forall\, y \in Q.$$

(a) Show that if P and Q are isomorphic, then their extreme points are in one-to-one correspondence.

(b) Let A be an $r \times n$ matrix and b be a vector in $\Re^r$, and let

$$P = \{x \in \Re^n \mid Ax \le b,\ x \ge 0\},$$

$$Q = \big\{(x, z) \in \Re^{n+r} \mid Ax + z = b,\ x \ge 0,\ z \ge 0\big\}.$$

Show that P and Q are isomorphic.

Solution: (a) Let P and Q be isomorhic polyhedral sets, and let $f : P \mapsto Q$ and $g : Q \mapsto P$ be affine functions such that

$$x = g\big(f(x)\big), \quad \forall\, x \in P, \qquad y = f\big(g(y)\big), \quad \forall\, y \in Q.$$

Assume that x^* is an extreme point of P and let $y^* = f(x^*)$. We will show that y^* is an extreme point of Q. Since x^* is an extreme point of P, by Exercise 2.6(b), it is also a face of P, and therefore, there exists a vector $c \in \Re^n$ such that

$$c'x < c'x^*, \qquad \forall\, x \in P,\ x \ne x^*.$$

For any $y \in Q$ with $y \ne y^*$, we have

$$f\big(g(y)\big) = y \ne y^* = f(x^*),$$

implying that

$$g(y) \ne g(y^*) = x^*, \qquad \text{with } \ g(y) \in P.$$

Hence,

$$c'g(y) < c'g(y^*), \qquad \forall\, y \in Q,\ y \ne y^*.$$

Let the affine function g be given by $g(y) = By + d$ for some $n \times m$ matrix B and vector $d \in \Re^n$. Then, we have

$$c'(By + d) < c'(By^* + d), \qquad \forall\, y \in Q,\ y \ne y^*,$$

implying that

$$(B'c)'y < (B'c)'y^*, \qquad \forall\, y \in Q,\ y \ne y^*.$$

If y^* were not an extreme point of Q, then we would have $y^* = \alpha y_1 + (1 - \alpha)y_2$ for some distinct points $y_1, y_2 \in Q$, $y_1 \ne y^*$, $y_2 \ne y^*$, and $\alpha \in (0, 1)$, so that

$$(B'c)'y^* = \alpha(B'c)'y_1 + (1 - \alpha)(B'c)'y_2 < (B'c)'y^*,$$

which is a contradiction. Hence, y^* is an extreme point of Q.

Conversely, if y^* is an extreme point of Q, then by using a symmetrical argument, we can show that x^* is an extreme point of P.

(b) For the sets

$$P = \{x \in \Re^n \mid Ax \le b,\ x \ge 0\},$$

$$Q = \big\{(x, z) \in \Re^{n+r} \mid Ax + z = b,\ x \ge 0,\ z \ge 0\big\},$$

let f and g be given by

$$f(x) = (x, b - Ax), \qquad \forall\ x \in P,$$

$$g(x, z) = x, \qquad \forall\ (x, z) \in Q.$$

Evidently, f and g are affine functions. Furthermore, clearly

$$f(x) \in Q, \qquad g\big(f(x)\big) = x, \qquad \forall\ x \in P,$$

$$g(x, z) \in P, \qquad f\big(g(x, z)\big) = x, \qquad \forall\ (x, z) \in Q.$$

Hence, P and Q are isomorphic.

SECTION 2.2: Polar Cones

2.8 (Cone Decomposition Theorem)

Let C be a nonempty closed convex cone in $\Re^n$ and let x be a vector in $\Re^n$. Show that:

(a) $\hat{x}$ is the projection of x on C if and only if

$$\hat{x} \in C, \qquad (x - \hat{x})'\hat{x} = 0, \qquad x - \hat{x} \in C^*.$$

(b) The following two statements are equivalent:

(i) x_1 and x_2 are the projections of x on C and C^*, respectively.

(ii) $x = x_1 + x_2$ with $x_1 \in C$, $x_2 \in C^*$, and $x_1' x_2 = 0$.

Solution: (a) Let $\hat{x}$ be the projection of x on C, which exists and is unique since C is closed and convex. By the Projection Theorem (Prop. 1.1.9), we have

$$(x - \hat{x})'(y - \hat{x}) \le 0, \qquad \forall\ y \in C.$$

Since C is a cone, we have $(1/2)\hat{x} \in C$ and $2\hat{x} \in C$, and by taking $y = (1/2)\hat{x}$ and $y = 2\hat{x}$ in the preceding relation, it follows that

$$(x - \hat{x})'\hat{x} = 0.$$

By combining the preceding two relations, we obtain

$$(x - \hat{x})'y \le 0, \qquad \forall\, y \in C,$$

implying that $x - \hat{x} \in C^*$.

Conversely, if $\hat{x} \in C$, $(x - \hat{x})'\hat{x} = 0$, and $x - \hat{x} \in C^*$, then it follows that

$$(x - \hat{x})'(y - \hat{x}) \le 0, \qquad \forall\, y \in C,$$

and by the Projection Theorem, $\hat{x}$ is the projection of x on C.

(b) Suppose that property (i) holds, i.e., x_1 and x_2 are the projections of x on C and C^*, respectively. Then, by part (a), we have

$$x_1 \in C, \qquad (x - x_1)'x_1 = 0, \qquad x - x_1 \in C^*.$$

Let $y = x - x_1$, so that the preceding relation can equivalently be written as

$$x - y \in C = (C^*)^*, \qquad y'(x - y) = 0, \qquad y \in C^*.$$

By using part (a), we conclude that y is the projection of x on C^*. Since by the Projection Theorem, the projection of a vector on a closed convex set is unique, it follows that $y = x_2$. Thus, we have $x = x_1 + x_2$ and in view of the preceding two relations, we also have $x_1 \in C$, $x_2 \in C^*$, and $x_1'x_2 = 0$. Hence, property (ii) holds.

Conversely, suppose that property (ii) holds, i.e., $x = x_1 + x_2$ with $x_1 \in C$, $x_2 \in C^*$, and $x_1'x_2 = 0$. Then, evidently the relations

$$x_1 \in C, \qquad (x - x_1)'x_1 = 0, \qquad x - x_1 \in C^*,$$

$$x_2 \in C^*, \qquad (x - x_2)'x_2 = 0, \qquad x - x_2 \in C$$

are satisfied, so that by part (a), x_1 and x_2 are the projections of x on C and C^*, respectively. Hence, property (i) holds.

2.9

Let C be a nonempty closed convex cone in $\Re^n$ and let a be a vector in $\Re^n$. Show that for any scalars $\beta > 0$ and $\gamma \ge 0$, we have

$$\max_{\|x\| \le \beta,\, x \in C} a'x \le \gamma \qquad \text{if and only if} \qquad a \in C^* + \big\{x \mid \|x\| \le \gamma/\beta\big\}.$$

(This may be viewed as an "approximate" version of the Polar Cone Theorem, which is obtained for $\gamma = 0$.)

Solution: If $a \in C^* + \big\{x \mid \|x\| \le \gamma/\beta\big\}$, then

$$a = \hat{a} + \overline{a} \qquad \text{with} \quad \hat{a} \in C^* \quad \text{and} \quad \|\overline{a}\| \le \gamma/\beta.$$

Since C is a closed convex cone, by the Polar Cone Theorem (Prop. 2.2.1), we have $(C^*)^* = C$, implying that for all x in C with $\|x\| \le \beta$,

$$\hat{a}'x \le 0 \qquad \text{and} \qquad \overline{a}'x \le \|\overline{a}\| \cdot \|x\| \le \gamma.$$

Hence,

$$a'x = (\hat{a} + \overline{a})'x \le \gamma, \qquad \forall\ x \in C \ \text{ with } \ \|x\| \le \beta,$$

thus implying that

$$\max_{\|x\| \le \beta,\ x \in C} a'x \le \gamma.$$

Conversely, assume that $a'x \le \gamma$ for all $x \in C$ with $\|x\| \le \beta$. Let $\hat{a}$ and $\overline{a}$ be the projections of a on C^* and C, respectively. By the Cone Decomposition Theorem (cf. Exercise 2.8), we have $a = \hat{a} + \overline{a}$ with $\hat{a} \in C^*$, $\overline{a} \in C$, and $\hat{a}'\overline{a} = 0$. Since $a'x \le \gamma$ for all $x \in C$ with $\|x\| \le \beta$ and $\overline{a} \in C$, we obtain

$$a' \frac{\overline{a}}{\|\overline{a}\|} \beta = (\hat{a} + \overline{a})' \frac{\overline{a}}{\|\overline{a}\|} \beta = \|\overline{a}\| \beta \le \gamma,$$

implying that $\|\overline{a}\| \le \gamma/\beta$, and showing that $a \in C^* + \big\{x \mid \|x\| \le \gamma/\beta\big\}$.

2.10 (Dimension and Lineality Space of Polar Cones)

Show that for any nonempty cone C in $\Re^n$, we have

$$L_{C^*} = \big(\mathrm{aff}(C)\big)^{\perp},$$

$$\dim(C) + \dim(L_{C^*}) = n,$$

$$\dim(C^*) + \dim\big(L_{\mathrm{conv}(C)}\big) \le \dim(C^*) + \dim\big(L_{cl(\mathrm{conv}(C))}\big) = n,$$

where L_X denotes the lineality space of a convex set X.

Solution: Note that $\mathrm{aff}(C)$ is a subspace of $\Re^n$ because C is a cone in $\Re^n$. We first show that

$$L_{C^*} = \big(\mathrm{aff}(C)\big)^{\perp}.$$

Let $y \in L_{C^*}$. Then, by the definition of the lineality space (see Chapter 1), both vectors y and $-y$ belong to the recession cone R_{C^*}. Since $0 \in C^*$, it follows that $0 + y$ and $0 - y$ belong to C^*. Therefore,

$$y'x \le 0, \qquad (-y)'x \le 0, \qquad \forall\ x \in C,$$

implying that

$$y'x = 0, \qquad \forall\ x \in C. \tag{2.3}$$

Let the dimension of the subspace $\mathrm{aff}(C)$ be m. By Prop. 1.3.2, there exist vectors $x_0, x_1, \ldots, x_m$ in $\mathrm{ri}(C)$ such that $x_1 - x_0, \ldots, x_m - x_0$ span $\mathrm{aff}(C)$. Thus, for any $z \in \mathrm{aff}(C)$, there exist scalars $\beta_1, \ldots, \beta_m$ such that

$$z = \sum_{i=1}^{m} \beta_i (x_i - x_0).$$

By using this relation and Eq. (2.3), for any $z \in \mathrm{aff}(C)$, we obtain

$$y'z = \sum_{i=1}^{m} \beta_i y'(x_i - x_0) = 0,$$

implying that $y \in \big(\mathrm{aff}(C)\big)^{\perp}$. Hence, $L_{C^*} \subset \big(\mathrm{aff}(C)\big)^{\perp}$.

Conversely, let $y \in \big(\mathrm{aff}(C)\big)^{\perp}$, so that in particular, we have

$$y'x = 0, \qquad (-y)'x = 0, \qquad \forall\ x \in C.$$

Therefore, $0+\alpha y \in C^*$ and $0+\alpha(-y) \in C^*$ for all $\alpha \geq 0$, and since C^* is a closed convex set, by the Recession Cone Theorem [Prop. 1.4.1(b)], it follows that y and $-y$ belong to the recession cone R_{C^*}. Hence, y belongs to the lineality space of C^*, showing that $\big(\mathrm{aff}(C)\big)^{\perp} \subset L_{C^*}$ and completing the proof of the equality $L_{C^*} = \big(\mathrm{aff}(C)\big)^{\perp}$.

By definition, we have $\dim(C) = \dim\big(\mathrm{aff}(C)\big)$ and since $L_{C^*} = \big(\mathrm{aff}(C)\big)^{\perp}$, we have $\dim(L_{C^*}) = \dim\Big(\big(\mathrm{aff}(C)\big)^{\perp}\Big)$. This implies that

$$\dim(C) + \dim(L_{C^*}) = n.$$

By replacing C with C^* in the preceding relation, and by using the Polar Cone Theorem (Prop. 2.2.1), we obtain

$$\dim(C^*) + \dim\big(L_{(C^*)^*}\big) = \dim(C^*) + \dim\big(L_{\mathrm{cl}(\mathrm{conv}(C))}\big) = n.$$

Furthermore, since

$$L_{\mathrm{conv}(C)} \subset L_{\mathrm{cl}(\mathrm{conv}(C))},$$

it follows that

$$\dim(C^*) + \dim\big(L_{\mathrm{conv}(C)}\big) \leq \dim(C^*) + \dim\big(L_{\mathrm{cl}(\mathrm{conv}(C))}\big) = n.$$

2.11 (Polar Cone Operations)

Show the following:

(a) For any nonempty cones $C_i \subset \Re^{n_i}$, $i = 1, \ldots, m$, we have

$$(C_1 \times \cdots \times C_m)^* = C_1^* \times \cdots \times C_m^*.$$

(b) For any collection of nonempty cones $\{C_i \mid i \in I\}$, we have

$$\big(\cup_{i\in I} C_i\big)^* = \cap_{i\in I} C_i^*.$$

(c) For any two nonempty cones C_1 and C_2, we have

$$(C_1 + C_2)^* = C_1^* \cap C_2^*.$$

(d) For any two nonempty closed convex cones C_1 and C_2, we have

$$(C_1 \cap C_2)^* = \text{cl}(C_1^* + C_2^*).$$

Furthermore, if $\text{ri}(C_1) \cap \text{ri}(C_2) \neq \emptyset$, then the cone $C_1^* + C_2^*$ is closed and the closure operation in the preceding relation can be omitted.

(e) Consider the following cones in $\Re^3$

$$C_1 = \big\{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 \leq x_3^2,\ x_3 \leq 0\big\},$$

$$C_2 = \big\{(x_1, x_2, x_3) \mid x_2 = -x_3\big\}.$$

Verify that $\text{ri}(C_1) \cap \text{ri}(C_2) = \emptyset$, $(1, 1, 1) \in (C_1 \cap C_2)^*$, and $(1, 1, 1) \notin C_1^* + C_2^*$, thus showing that the closure operation in the relation of part (c) may not be omitted when $\text{ri}(C_1) \cap \text{ri}(C_2) = \emptyset$.

Solution: (a) It suffices to consider the case where $m = 2$. Let $(y_1, y_2) \in (C_1 \times C_2)^*$. Then, we have $(y_1, y_2)'(x_1, x_2) \leq 0$ for all $(x_1, x_2) \in C_1 \times C_2$, or equivalently

$$y_1' x_1 + y_2' x_2 \leq 0, \qquad \forall\ x_1 \in C_1, \quad \forall\ x_2 \in C_2.$$

Since C_2 is a cone, 0 belongs to its closure, so by letting $x_2 \to 0$ in the preceding relation, we obtain $y_1' x_1 \leq 0$ for all $x_1 \in C_1$, showing that $y_1 \in C_1^*$. Similarly, we obtain $y_2 \in C_2^*$, and therefore $(y_1, y_2) \in C_1^* \times C_2^*$, implying that $(C_1 \times C_2)^* \subset C_1^* \times C_2^*$.

Conversely, let $y_1 \in C_1^*$ and $y_2 \in C_2^*$. Then, we have

$$(y_1, y_2)'(x_1, x_2) = y_1' x_1 + y_2' x_2 \leq 0, \qquad \forall\ x_1 \in C_1, \quad \forall\ x_2 \in C_2,$$

implying that $(y_1, y_2) \in (C_1 \times C_2)^*$, and showing that $C_1^* \times C_2^* \subset (C_1 \times C_2)^*$.

(b) A vector y belongs to the polar cone of $\cup_{i \in I} C_i$ if and only if $y' x \leq 0$ for all $x \in C_i$ and all $i \in I$, which is equivalent to having $y \in C_i^*$ for every $i \in I$. Hence, y belongs to $\big(\cup_{i \in I} C_i\big)^*$ if and only if y belongs to $\cap_{i \in I} C_i^*$.

(c) Let $y \in (C_1 + C_2)^*$, so that

$$y'(x_1 + x_2) \leq 0, \qquad \forall\ x_1 \in C_1, \qquad \forall\ x_2 \in C_2. \tag{2.4}$$

Since the zero vector is in the closures of C_1 and C_2, by letting $x_2 \to 0$ with $x_2 \in C_2$ in Eq. (2.4), we obtain

$$y' x_1 \leq 0, \qquad \forall\ x_1 \in C_1,$$

and similarly, by letting $x_1 \to 0$ with $x_1 \in C_1$ in Eq. (2.4), we obtain

$$y' x_2 \leq 0, \qquad \forall\ x_2 \in C_2.$$

Thus, $y \in C_1^* \cap C_2^*$, showing that $(C_1 + C_2)^* \subset C_1^* \cap C_2^*$.

Conversely, let $y \in C_1^* \cap C_2^*$. Then, we have

$$y'x_1 \leq 0, \qquad \forall\, x_1 \in C_1,$$

$$y'x_2 \leq 0, \qquad \forall\, x_2 \in C_2,$$

implying that

$$y'(x_1 + x_2) \leq 0, \qquad \forall\, x_1 \in C_1, \quad \forall\, x_2 \in C_2.$$

Hence $y \in (C_1 + C_2)^*$, showing that $C_1^* \cap C_2^* \subset (C_1 + C_2)^*$.

(d) Since C_1 and C_2 are closed convex cones, by the Polar Cone Theorem (Prop. 2.2.1) and by part (b), it follows that

$$C_1 \cap C_2 = (C_1^*)^* \cap (C_2^*)^* = (C_1^* + C_2^*)^*.$$

By taking the polars and by using the Polar Cone Theorem, we obtain

$$(C_1 \cap C_2)^* = \big((C_1^* + C_2^*)^*\big)^* = \mathrm{cl}\big(\mathrm{conv}(C_1^* + C_2^*)\big).$$

The cone $C_1^* + C_2^*$ is convex, so that

$$(C_1 \cap C_2)^* = \mathrm{cl}(C_1^* + C_2^*).$$

Suppose now that $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) \neq \emptyset$. We will show that $C_1^* + C_2^*$ is closed by using Prop. 1.4.14. According to this proposition, if for any nonempty closed convex sets $\overline{C}_1$ and $\overline{C}_2$ in $\Re^n$, the equality $y_1 + y_2 = 0$ with $y_1 \in R_{\overline{C}_1}$ and $y_2 \in R_{\overline{C}_2}$ implies that y_1 and y_2 belong to the lineality spaces of $\overline{C}_1$ and $\overline{C}_2$, respectively, then the vector sum $\overline{C}_1 + \overline{C}_2$ is closed.

Let $y_1 + y_2 = 0$ with $y_1 \in R_{C_1^*}$ and $y_2 \in R_{C_2^*}$. Because C_1^* and C_2^* are closed convex cones, we have $R_{C_1^*} = C_1^*$ and $R_{C_2^*} = C_2^*$, so that $y_1 \in C_1^*$ and $y_2 \in C_2^*$. The lineality space of a cone is the set of vectors y such that y and $-y$ belong to the cone, so that in view of the preceding discussion, to show that $C_1^* + C_2^*$ is closed, it suffices to prove that $-y_1 \in C_1^*$ and $-y_2 \in C_2^*$.

Since $y_1 = -y_2$ and $y_1 \in C_1^*$, it follows that

$$y_2'x \geq 0, \qquad \forall\, x \in C_1, \tag{2.5}$$

and because $y_2 \in C_2^*$, we have

$$y_2'x \leq 0, \qquad \forall\, x \in C_2,$$

which combined with the preceding relation yields

$$y_2'x = 0, \qquad \forall\, x \in C_1 \cap C_2. \tag{2.6}$$

In view of the fact $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) \neq \emptyset$, and Eqs. (2.5) and (2.6), it follows that the linear function $y_2'x$ attains its minimum over the convex set C_1 at a point in the relative interior of C_1, implying that $y_2'x = 0$ for all $x \in C_1$ (cf. Prop. 1.3.4). Therefore, $y_2 \in C_1^*$ and since $y_2 = -y_1$, we have $-y_1 \in C_1^*$. By exchanging the roles of y_1 and y_2 in the preceding analysis, we similarly show that $-y_2 \in C_2^*$, completing the proof.

(e) By drawing the cones C_1 and C_2, it can be seen that $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) = \emptyset$ and

$$C_1 \cap C_2 = \big\{(x_1, x_2, x_3) \mid x_1 = 0,\ x_2 = -x_3,\ x_3 \le 0\big\},$$

$$C_1^* = \big\{(y_1, y_2, y_3) \mid y_1^2 + y_2^2 \le y_3^2,\ y_3 \ge 0\big\},$$

$$C_2^* = \big\{(z_1, z_2, z_3) \mid z_1 = 0,\ z_2 = z_3\big\}.$$

Clearly, $x_1 + x_2 + x_3 = 0$ for all $x \in C_1 \cap C_2$, implying that $(1, 1, 1) \in (C_1 \cap C_2)^*$. Suppose that $(1, 1, 1) \in C_1^* + C_2^*$, so that $(1, 1, 1) = (y_1, y_2, y_3) + (z_1, z_2, z_3)$ for some $(y_1, y_2, y_3) \in C_1^*$ and $(z_1, z_2, z_3) \in C_2^*$, implying that $y_1 = 1$, $y_2 = 1 - z_2$, $y_3 = 1 - z_2$ for some $z_2 \in \Re$. However, this point does not belong to C_1^*, which is a contradiction. Therefore, $(1, 1, 1)$ is not in $C_1^* + C_2^*$. Hence, when $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) = \emptyset$, the relation

$$(C_1 \cap C_2)^* = C_1^* + C_2^*$$

may fail.

2.12 (Linear Transformations and Polar Cones)

Let C be a nonempty cone in $\Re^n$, K be a nonempty closed convex cone in $\Re^m$, and A be a linear transformation from $\Re^n$ to $\Re^m$. Show that

$$(AC)^* = (A')^{-1} \cdot C^*, \qquad \big(A^{-1} \cdot K\big)^* = \mathrm{cl}(A'K^*).$$

Show also that if $\mathrm{ri}(K) \cap R(A) \neq \emptyset$, then the cone $A'K^*$ is closed and $(A')^{-1}$ and the closure operation in the above relation can be omitted.

Solution: We have $y \in (AC)^*$ if and only if $y'Ax \le 0$ for all $x \in C$, which is equivalent to $(A'y)'x \le 0$ for all $x \in C$. This is in turn equivalent to $A'y \in C^*$. Hence, $y \in (AC)^*$ if and only if $y \in (A')^{-1} \cdot C^*$, showing that

$$(AC)^* = (A')^{-1} \cdot C^*. \tag{2.7}$$

We next show that for a closed convex cone $K \subset \Re^m$, we have

$$\big(A^{-1} \cdot K\big)^* = \mathrm{cl}(A'K^*).$$

Let $y \in \big(A^{-1} \cdot K\big)^*$ and to arrive at a contradiction, assume that $y \notin \mathrm{cl}(A'K^*)$. By the Strict Separation Theorem (Prop. 1.5.3), the closed convex cone $\mathrm{cl}(A'K^*)$ and the vector y can be strictly separated, i.e., there exist a vector $a \in \Re^n$ and a scalar b such that

$$a'x < b < a'y, \qquad \forall\ x \in \mathrm{cl}(A'K^*).$$

If $a'x > 0$ for some $x \in \text{cl}(A'K^*)$, then since $\text{cl}(A'K^*)$ is a cone, we would have $\lambda x \in \text{cl}(A'K^*)$ for all $\lambda > 0$, implying that $a'(\lambda x) \to \infty$ when $\lambda \to \infty$, which contradicts the preceding relation. Thus, we must have $a'x \leq 0$ for all $x \in \text{cl}(A'K^*)$, and since $0 \in \text{cl}(A'K^*)$, it follows that

$$\sup_{x \in \text{cl}(A'K^*)} a'x = 0 \leq b < a'y. \tag{2.8}$$

Therefore, $a \in \big(\text{cl}(A'K^*)\big)^*$, and since $\big(\text{cl}(A'K^*)\big)^* \subset (A'K^*)^*$, it follows that $a \in (A'K^*)^*$. In view of Eq. (2.7) and the Polar Cone Theorem (Prop. 2.2.1), we have

$$(A'K^*)^* = A^{-1}(K^*)^* = A^{-1} \cdot K,$$

implying that $a \in A^{-1} \cdot K$. Because $y \in \big(A^{-1} \cdot K\big)^*$, it follows that $y'a \leq 0$, contradicting Eq. (2.8). Hence, we must have $y \in \text{cl}(A'K^*)$, showing that

$$\big(A^{-1} \cdot K\big)^* \subset \text{cl}(A'K^*).$$

To show the reverse inclusion, let $y \in A'K^*$ and assume, to arrive at a contradiction, that $y \notin (A^{-1} \cdot K)^*$. By the Strict Separation Theorem (Prop. 1.5.3), the closed convex cone $(A^{-1} \cdot K)^*$ and the vector y can be strictly separated, i.e., there exist a vector $\overline{a} \in \Re^n$ and a scalar $\overline{b}$ such that

$$\overline{a}'x < \overline{b} < \overline{a}'y, \qquad \forall\, x \in (A^{-1} \cdot K)^*.$$

Similar to the preceding analysis, since $(A^{-1} \cdot K)^*$ is a cone, it can be seen that

$$\sup_{x \in (A^{-1} \cdot K)^*} \overline{a}'x = 0 \leq \overline{b} < \overline{a}'y, \tag{2.9}$$

implying that $\overline{a} \in \big((A^{-1} \cdot K)^*\big)^*$. Since K is a closed convex cone and A is a linear (and therefore continuous) transformation, the set $A^{-1} \cdot K$ is a closed convex cone. Furthermore, by the Polar Cone Theorem, we have that $\big((A^{-1} \cdot K)^*\big)^* = A^{-1} \cdot K$. Therefore, $\overline{a} \in A^{-1} \cdot K$, implying that $A\overline{a} \in K$. Since $y \in A'K^*$, we have $y = A'v$ for some $v \in K^*$, and it follows that

$$y'\overline{a} = (A'v)'\overline{a} = v'A\overline{a} \leq 0,$$

contradicting Eq. (2.9). Hence, we must have $y \in (A^{-1} \cdot K)^*$, implying that

$$A'K^* \subset (A^{-1} \cdot K)^*.$$

Taking the closure of both sides of this relation, we obtain

$$\text{cl}(A'K^*) \subset (A^{-1} \cdot K)^*,$$

completing the proof.

Suppose that $\text{ri}(K^*) \cap R(A) \neq \emptyset$. We will show that the cone $A'K^*$ is closed by using Prop. 1.4.13. According to this proposition, if $R_{K^*} \cap N(A')$ is a subspace of the lineality space L_{K^*} of K^*, then

$$\text{cl}(A'K^*) = A'K^*.$$

Thus, it suffices to verify that $R_{K^*} \cap N(A')$ is a subspace of L_{K^*}. Indeed, we will show that $R_{K^*} \cap N(A') = L_{K^*} \cap N(A')$.

Let $y \in K^* \cap N(A')$. Because $y \in K^*$, we obtain

$$(-y)'x \geq 0, \qquad \forall\, x \in K. \tag{2.10}$$

For $y \in N(A')$, we have $-y \in N(A')$ and since $N(A') = R(A)^\perp$, it follows that

$$(-y)'z = 0, \qquad \forall\, z \in R(A). \tag{2.11}$$

In view of the relation $\text{ri}(K) \cap R(A) \neq \emptyset$, and Eqs. (2.10) and (2.11), the linear function $(-y)'x$ attains its minimum over the convex set K at a point in the relative interior of K, implying that $(-y)'x = 0$ for all $x \in K$ (cf. Prop. 1.3.4). Hence $(-y) \in K^*$, so that $y \in L_{K^*}$ and because $y \in N(A')$, we see that $y \in L_{K^*} \cap N(A')$. The reverse inclusion follows directly from the relation $L_{K^*} \subset R_{K^*}$, thus completing the proof.

2.13 (Pointed Cones and Bases)

Let C be a closed convex cone in $\Re^n$. We say that C is a *pointed cone* if $C \cap (-C) = \{0\}$. A convex set $D \subset \Re^n$ is said to be a *base* for C if $C = \text{cone}(D)$ and $0 \notin \text{cl}(D)$. Show that the following properties are equivalent:

(a) C is a pointed cone.

(b) $\text{cl}(C^* - C^*) = \Re^n$.

(c) $C^* - C^* = \Re^n$.

(d) C^* has nonempty interior.

(e) There exist a nonzero vector $\hat{x} \in \Re^n$ and a positive scalar δ such that $\hat{x}'x \geq \delta \|x\|$ for all $x \in C$.

(f) C has a bounded base.

Hint: Use Exercise 2.11 to show the implications (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (f) $\Rightarrow$ (a).

Solution: (a) $\Rightarrow$ (b) Since C is a pointed cone, $C \cap (-C) = \{0\}$, so that

$$\big(C \cap (-C)\big)^* = \Re^n.$$

On the other hand, by Exercise 2.11, it follows that

$$\big(C \cap (-C)\big)^* = \text{cl}(C^* - C^*),$$

which when combined with the preceding relation yields $\mathrm{cl}(C^* - C^*) = \Re^n$.

(b) ⇒ (c) Since C is a closed convex cone, by the polar cone operations of Exercise 2.11, it follows that

$$\big(C \cap (-C)\big)^* = \mathrm{cl}(C^* - C^*) = \Re^n.$$

By taking the polars and using the Polar Cone Theorem (Prop. 2.2.1), we obtain

$$\Big(\big(C \cap (-C)\big)^*\Big)^* = C \cap (-C) = \{0\}. \tag{2.12}$$

Now, to arrive at a contradiction assume that there is a vector $\hat{x} \in \Re^n$ such that $\hat{x} \notin C^* - C^*$. Then, by the Separating Hyperplane Theorem (Prop. 1.5.2), there exists a nonzero vector $a \in \Re^n$ such that

$$a'\hat{x} \geq a'x, \qquad \forall\, x \in C^* - C^*.$$

If $a'x > 0$ for some $x \in C^* - C^*$, then since $C^* - C^*$ is a cone, the right hand-side of the preceding relation can be arbitrarily large, a contradiction. Thus, we have $a'x \leq 0$ for all $x \in C^* - C^*$, implying that $a \in (C^* - C^*)^*$. By the polar cone operations of Exercise 2.11(b) and the Polar Cone Theorem, it follows that

$$(C^* - C^*)^* = (C^*)^* \cap (-C^*)^* = C \cap (-C).$$

Thus, $a \in C \cap (-C)$ with $a \neq 0$, contradicting Eq. (2.12). Hence, we must have $C^* - C^* = \Re^n$.

(c) ⇒ (d) Because $C^* \subset \mathrm{aff}(C^*)$ and $-C^* \subset \mathrm{aff}(C^*)$, we have $C^* - C^* \subset \mathrm{aff}(C^*)$ and since $C^* - C^* = \Re^n$, it follows that $\mathrm{aff}(C^*) = \Re^n$, showing that C^* has nonempty interior.

(d) ⇒ (e) Let v be a vector in the interior of C^*. Then, there exists a positive scalar δ such that the vector $v + \delta \frac{y}{\|y\|}$ is in C^* for all $y \in \Re^n$ with $y \neq 0$, i.e.,

$$\left(v + \delta \frac{y}{\|y\|}\right)' x \leq 0, \qquad \forall\, x \in C, \quad \forall\, y \in \Re^n,\ y \neq 0.$$

By taking $y = x$, it follows that

$$\left(v + \delta \frac{x}{\|x\|}\right)' x \leq 0, \qquad \forall\, x \in C,\ x \neq 0,$$

implying that

$$v'x + \delta\|x\| \leq 0, \qquad \forall\, x \in C,\ x \neq 0.$$

Clearly, this relation holds for $x = 0$, so that

$$v'x \leq -\delta\|x\|, \qquad \forall\, x \in C.$$

Multiplying the preceding relation with -1 and letting $\hat{x} = -v$, we obtain

$$\hat{x}'x \geq \delta \|x\|, \qquad \forall\, x \in C.$$

(e) $\Rightarrow$ (f) Let

$$D = \big\{ y \in C \mid \hat{x}'y = 1 \big\}.$$

Then, D is a closed convex set since it is the intersection of the closed convex cone C and the closed convex set $\{y \mid \hat{x}'y = 1\}$. Obviously, $0 \notin D$. Thus, to show that D is a base for C, it remains to prove that $C = \text{cone}(D)$. Take any $x \in C$. If $x = 0$, then $x \in \text{cone}(D)$ and we are done, so assume that $x \neq 0$. We have by hypothesis

$$\hat{x}'x \geq \delta \|x\| > 0, \qquad \forall\, x \in C,\ x \neq 0,$$

so we may define $\hat{y} = \frac{x}{\hat{x}'x}$. Clearly, $\hat{y} \in D$ and $x = (\hat{x}'x)\hat{y}$ with $\hat{x}'x > 0$, showing that $x \in \text{cone}(D)$ and that $C \subset \text{cone}(D)$. Since $D \subset C$, the inclusion $\text{cone}(D) \subset C$ is obvious. Thus, $C = \text{cone}(D)$ and D is a base for C. Furthermore, for every y in D, since y is also in C, we have

$$1 = \hat{x}'y \geq \delta \|y\|,$$

showing that D is bounded and completing the proof.

(f) $\Rightarrow$ (a) Since C has a bounded base, $C = \text{cone}(D)$ for some bounded convex set D with $0 \notin \text{cl}(D)$. To arrive at a contradiction, we assume that the cone C is not pointed, so that there exists a nonzero vector $d \in C \cap (-C)$, implying that d and $-d$ are in C. Let $\{\lambda_k\}$ be a sequence of positive scalars. Since $\lambda_k d \in C$ for all k and D is a base for C, there exist a sequence $\{\mu_k\}$ of positive scalars and a sequence $\{y_k\}$ of vectors in D such that

$$\lambda_k d = \mu_k y_k, \qquad \forall\, k.$$

Therefore, $y_k = \frac{\lambda_k}{\mu_k} d \in D$ for all k and because D is bounded, the sequence $\big\{y_k\big\}$ has a subsequence converging to some $y \in \text{cl}(D)$. Without loss of generality, we may assume that $y_k \to y$, which in view of $y_k = \frac{\lambda_k}{\mu_k} d$ for all k, implies that $y = \alpha d$ and $\alpha d \in \text{cl}(D)$ for some $\alpha \geq 0$. Furthermore, by the definition of base, we have $0 \notin \text{cl}(D)$, so that $\alpha > 0$. Similar to the preceding, by replacing d with $-d$, we can show that $\tilde{\alpha}(-d) \in \text{cl}(D)$ for some positive scalar $\tilde{\alpha}$. Therefore, $\alpha d \in \text{cl}(D)$ and $\tilde{\alpha}(-d) \in \text{cl}(D)$ with $\alpha > 0$ and $\tilde{\alpha} > 0$. Since D is convex, its closure $\text{cl}(D)$ is also convex, implying that $0 \in \text{cl}(D)$, contradicting the definition of a base. Hence, the cone C must be pointed.

SECTION 2.3: Polyhedral Sets and Functions

2.14

Show that a closed convex cone is polyhedral if and only if its polar cone is polyhedral.

Solution: Let the closed convex cone C be polyhedral, and of the form

$$C = \big\{x \mid a_j'x \leq 0,\ j = 1, \ldots, r\big\},$$

for some vectors a_j in $\Re^n$. By Farkas' Lemma, we have

$$C^* = \text{cone}\big(\{a_1, \ldots, a_r\}\big),$$

so the polar cone of a polyhedral cone is finitely generated. Conversely, using the Polar Cone Theorem, we have

$$\text{cone}\big(\{a_1, \ldots, a_r\}\big)^* = \big\{x \mid a_j'x \leq 0,\ j = 1, \ldots, r\big\},$$

so the polar of a finitely generated cone is polyhedral. Thus, a closed convex cone is polyhedral if and only if its polar cone is finitely generated. By the Minkowski-Weyl Theorem (Prop. 2.3.2), a cone is finitely generated if and only if it is polyhedral. Therefore, a closed convex cone is polyhedral if and only if its polar cone is polyhedral.

2.15 (Closedness of Finitely Generated Cones)

This exercise proves that a finitely generated cone is closed without invoking Prop. 1.4.13. Let $a_1, \ldots, a_r$ be vectors in $\Re^n$ and let A be the $n \times r$ matrix that has as columns these vectors. Consider the cone generated by $a_1, \ldots, a_r$:

$$\text{cone}\big(\{a_1, \ldots, a_r\}\big) = \{A\mu \mid \mu \geq 0\}.$$

(a) Show that if $a_1, \ldots, a_r$ are linearly independent, then $\text{cone}\big(\{a_1, \ldots, a_r\}\big)$ is closed. *Hint*: Show that if $y_k = \{A\mu_k\}$ and $y_k \to y$, then $y = A\mu$ with

$$\mu = \lim_{k \to \infty} \mu_k = \lim_{k \to \infty} (A'A)^{-1}A'y_k = (A'A)^{-1}A'y.$$

(b) Show that $\text{cone}\big(\{a_1, \ldots, a_r\}\big)$ is closed without the linear independence assumption of part (a). *Hint*: Use Caratheodory's Theorem to show that $\text{cone}\big(\{a_1, \ldots, a_r\}\big)$ is equal to the union of a finite number of cones generated by linearly independent vectors.

Solution: (a) Consider a sequence $\{y_k\} \subset \text{cone}\big(\{a_1, \ldots, a_r\}\big)$ with $y_k \to y$. We will show that $y \in \text{cone}\big(\{a_1, \ldots, a_r\}\big)$. For each k, we have $y_k = A\mu_k$ for some $\mu_k \geq 0$, from which we obtain,

$$A'y_k = A'A\mu_k.$$

Since $a_1, \ldots, a_r$ are assumed linearly independent, the matrix $A'A$ is invertible, and we have

$$\mu_k = (A'A)^{-1}A'y_k.$$

It follows that

$$\mu_k \to \mu,$$

where

$$\mu = (A'A)^{-1}A'y.$$

Furthermore, since $\mu_k \geq 0$, we have $\mu \geq 0$. Taking the limit in the relation $y_k = A\mu_k$, we obtain $y = A\mu$ with $\mu \geq 0$, so $y \in \text{cone}\big(\{a_1, \ldots, a_r\}\big)$.

(b) By Caratheodory's Theorem, every vector in $\text{cone}\big(\{a_1, \ldots, a_r\}\big)$ is a positive combination of linearly independent vectors. Thus, $\text{cone}\big(\{a_1, \ldots, a_r\}\big)$ is the union of $\text{cone}\big(\{a_j \mid j \in J\}\big)$ as J ranges over all subsets of $\{1, \ldots, r\}$ such that the set $\{a_j \mid j \in J\}$ is linearly independent. Each of these cones is closed by part (a), so their union is also closed.

2.16

Let P be a polyhedral set in $\Re^n$, with a Minkowski-Weyl Representation

$$P = \left\{ x \;\middle|\; x = \sum_{j=1}^{m} \mu_j v_j + y,\ \sum_{j=1}^{m} \mu_j = 1,\ \mu_j \geq 0,\ j = 1, \ldots, m,\ y \in C \right\},$$

where $v_1, \ldots, v_m$ are some vectors in $\Re^n$ and C is a finitely generated cone in $\Re^n$ (cf. Prop. 2.3.3). Show that:

(a) The recession cone of P is equal to C.

(b) Each extreme point of P is equal to some vector v_i that cannot be represented as a convex combination of the vectors v_j with $v_j \neq v_i$.

Solution: (a) We first show that C is a subset of R_P, the recession cone of P. Let $\overline{y} \in C$, and choose any $\alpha \geq 0$ and $x \in P$ of the form $x = \sum_{j=1}^{m} \mu_j v_j$. Since C is a cone, $\alpha\overline{y} \in C$, so that $x + \alpha\overline{y} \in P$ for all $\alpha \geq 0$. It follows that $\overline{y} \in R_P$. Hence $C \subset R_P$. Conversely, to show that $R_P \subset C$, let $\overline{y} \in R_P$ and take any $x \in P$. Then $x + k\overline{y} \in P$ for all $k \geq 1$. Since $P = V + C$, where $V = \text{conv}\big(\{v_1, \ldots, v_m\}\big)$, it follows that

$$x + k\overline{y} = v^k + y^k, \qquad \forall\, k \geq 1,$$

with $v^k \in V$ and $y^k \in C$ for all $k \geq 1$. Because V is compact, the sequence $\{v^k\}$ has a limit point $v \in V$, and without loss of generality, we may assume that $v^k \to v$. Then

$$\lim_{k\to\infty} \|k\overline{y} - y^k\| = \lim_{k\to\infty} \|v^k - x\| = \|v - x\|,$$

implying that

$$\lim_{k\to\infty} \big\|\overline{y} - (1/k)y^k\big\| = 0.$$

Therefore, the sequence $\{(1/k)y^k\}$ converges to $\overline{y}$. Since $y^k \in C$ for all $k \geq 1$, the sequence $\{(1/k)y^k\}$ is in C, and by the closedness of C, it follows that $\overline{y} \in C$. Hence, $R_P \subset C$.

(b) Any point in P has the form $v + y$ with $v \in \mathrm{conv}(\{v_1, \ldots, v_m\})$ and $y \in C$, or equivalently

$$v + y = \frac{1}{2}v + \frac{1}{2}(v + 2y),$$

with v and $v + 2y$ being two distinct points in P if $y \neq 0$. Therefore, none of the points $v + y$, with $v \in \mathrm{conv}(\{v_1, \ldots, v_m\})$ and $y \in C$, is an extreme point of P if $y \neq 0$. Hence, an extreme point of P must be in the set $\{v_1, \ldots, v_m\}$. Since by definition, an extreme point of P is not a convex combination of points in P, an extreme point of P must be equal to some v_i that cannot be expressed as a convex combination of the remaining vectors v_j, $j \neq i$.

2.17 (Compact Polyhedral Sets)

Show that a nonempty compact convex set is polyhedral if and only if it has a finite number of extreme points. Show by example that the compactness assumption is essential.

Solution: By the Minkowski-Weyl Representation Theorem (Prop. 2.3.3), a polyhedral set has a finite number of extreme points. Conversely, let P be a compact convex set having a finite number of extreme points $\{v_1, \ldots, v_m\}$. By the Krein-Milman Theorem (Exercise 2.2), a compact convex set is equal to the convex hull of its extreme points, so that $P = \mathrm{conv}(\{v_1, \ldots, v_m\})$, which is a polyhedral set by Minkowski-Weyl Representation Theorem.

As an example showing that the assertion fails if compactness of the set is replaced by a weaker assumption that the set is closed and contains no lines, consider the set $D \subset \Re^3$ given by

$$D = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 \leq 1,\ x_3 = 1\}.$$

Let $C = \mathrm{cone}(D)$. It can seen that C is not a polyhedral set. On the other hand, C is closed, convex, does not contain a line, and has a unique extreme point at the origin.

[For a more formal argument, note that if C were polyhedral, then the set

$$D = C \cap \{(x_1, x_2, x_3) \mid x_3 = 1\}$$

would also be polyhedral by Prop. 2.3.4, since both C and $\{(x_1, x_2, x_3) \mid x_3 = 1\}$ are polyhedral sets. Thus, by Prop. 2.3.3, it would follow that D has a finite number of extreme points. But this is a contradiction because the set of extreme points of D coincides with $\{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 = 1,\ x_3 = 1\}$, which contains an infinite number of points. Thus, C is not a polyhedral cone, and therefore not a polyhedral set, while C is closed, convex, does not contain a line, and has a unique extreme point at the origin.]

2.18 (Polyhedral Set Decomposition)

Show that a polyhedral set can be written as the vector sum of a subspace and a polyhedral set that contains at least one extreme point. *Hint*: Use the decomposition result of Prop. 1.4.4.

Solution: By the remarks following Prop. 1.4.4, a convex set C can be written as

$$C = L_C + (C \cap L_C^\perp),$$

where L_C is the lineality space of C. Furthermore, the set $C \cap L_C^\perp$ contains no lines. If C is polyhedral, then $C \cap L_C^\perp$ is also polyhedral and since it contains no lines, by Prop. 2.1.2, it must contain an extreme point.

2.19 (Cones Generated by Polyhedral Sets)

Show that if P is a polyhedral set in $\Re^n$ containing the origin, then cone(P) is a polyhedral cone. Give an example showing that if P does not contain the origin, then cone(P) may not be a polyhedral cone.

Solution: We give two proofs. The first is based on the Minkowski-Weyl Representation of a polyhedral set P (cf. Prop. 2.3.3), while the second is based on a representation of P by a system of linear inequalities.

Let P be a polyhedral set with Minkowski-Weyl representation

$$P = \left\{ x \;\middle|\; x = \sum_{j=1}^m \mu_j v_j + y,\ \sum_{j=1}^m \mu_j = 1,\ \mu_j \geq 0,\ j = 1, \ldots, m,\ y \in C \right\},$$

where $v_1, \ldots, v_m$ are some vectors in $\Re^n$ and C is a finitely generated cone in $\Re^n$. Let C be given by

$$C = \left\{ y \;\middle|\; y = \sum_{i=1}^r \lambda_i a_i,\ \lambda_i \geq 0,\ i = 1, \ldots, r \right\},$$

where $a_1, \ldots, a_r$ are some vectors in $\Re^n$, so that

$$P = \left\{ x \;\middle|\; x = \sum_{j=1}^m \mu_j v_j + \sum_{i=1}^r \lambda_i a_i,\ \sum_{j=1}^m \mu_j = 1,\ \mu_j \geq 0,\ \forall\, j,\ \lambda_i \geq 0,\ \forall\, i \right\}.$$

We claim that

$$\text{cone}(P) = \text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big).$$

Since $P \subset \text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big)$, it follows that

$$\text{cone}(P) \subset \text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big).$$

Conversely, let $y \in \text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big)$. Then, we have

$$y = \sum_{j=1}^{m} \overline{\mu}_j v_j + \sum_{i=1}^{r} \overline{\lambda}_i a_i,$$

with $\overline{\mu}_j \geq 0$ and $\overline{\lambda}_i \geq 0$ for all i and j. If $\overline{\mu}_j = 0$ for all j, then $y = \sum_{i=1}^{r} \overline{\lambda}_i a_i \in C$, and since $C = R_P$ (cf. Exercise 2.16), it follows that $y \in R_P$. Because the origin belongs to P and $y \in R_P$, we have $0 + y \in P$, implying that $y \in P$, and consequently $y \in \text{cone}(P)$. If $\overline{\mu}_j > 0$ for some j, then by setting $\overline{\mu} = \sum_{j=1}^{m} \overline{\mu}_j$, $\mu_j = \overline{\mu}_j / \overline{\mu}$ for all j, and $\lambda_i = \overline{\lambda}_i / \overline{\mu}$ for all i, we obtain

$$y = \overline{\mu} \left(\sum_{j=1}^{m} \mu_j v_j + \sum_{i=1}^{r} \lambda_i a_i \right),$$

where $\overline{\mu} > 0$, $\mu_j \geq 0$ with $\sum_{j=1}^{m} \mu_j = 1$, and $\lambda_i \geq 0$. Therefore $y = \overline{\mu}\, \overline{x}$ with $\overline{x} \in P$ and $\overline{\mu} > 0$, implying that $y \in \text{cone}(P)$ and showing that

$$\text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big) \subset \text{cone}(P).$$

We now give an alternative proof using the representation of P by a system of linear inequalities. Let P be given by

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

where $a_1, \ldots, a_r$ are vectors in $\Re^n$ and $b_1, \ldots, b_r$ are scalars. Since P contains the origin, it follows that $b_j \geq 0$ for all j. Define the index set J as follows

$$J = \{j \mid b_j = 0\}.$$

We consider separately the two cases where $J \neq \emptyset$ and $J = \emptyset$. If $J \neq \emptyset$, then we will show that

$$\text{cone}(P) = \big\{x \mid a_j' x \leq 0,\ j \in J\big\}.$$

To see this, note that since $P \subset \big\{x \mid a_j' x \leq 0,\ j \in J\big\}$, we have

$$\text{cone}(P) \subset \big\{x \mid a_j' x \leq 0,\ j \in J\big\}.$$

Conversely, let $\overline{x} \in \big\{x \mid a_j' x \leq 0,\ j \in J\big\}$. We will show that $\overline{x} \in \text{cone}(P)$. If $\overline{x} \in P$, then $\overline{x} \in \text{cone}(P)$ and we are done, so assume that $\overline{x} \notin P$, implying that the set

$$\overline{J} = \{j \notin J \mid a_j' \overline{x} > b_j\} \tag{2.13}$$

is nonempty. By the definition of J, we have $b_j > 0$ for all $j \notin J$, so let

$$\mu = \min_{j \in \overline{J}} \frac{b_j}{a_j' \overline{x}},$$

and note that $0 < \mu < 1$. We have

$$a_j'(\mu\overline{x}) \leq 0, \qquad \forall\, j \in J,$$

$$a_j'(\mu\overline{x}) \leq b_j, \qquad \forall\, j \in \overline{J}.$$

For $j \notin \overline{J} \cup J$ and $a_j'\overline{x} \leq 0 < b_j$, since $\mu > 0$, we still have $a_j'(\mu\overline{x}) \leq 0 < b_j$. For $j \notin \overline{J} \cup J$ and $0 < a_j'\overline{x} \leq b_j$, since $\mu < 1$, we have $0 < a_j'(\mu\overline{x}) < b_j$. Therefore, $\mu\overline{x} \in P$, implying that $\overline{x} = \frac{1}{\mu}(\mu\overline{x}) \in \text{cone}(P)$. It follows that

$$\big\{x \mid a_j'x \leq 0,\ j \in J\big\} \subset \text{cone}(P),$$

and hence, $\text{cone}(P) = \big\{x \mid a_j'x \leq 0,\ j \in J\big\}$.

If $J = \emptyset$, then we will show that $\text{cone}(P) = \Re^n$. To see this, take any $\overline{x} \in \Re^n$. If $\overline{x} \in P$, then clearly $\overline{x} \in \text{cone}(P)$, so assume that $\overline{x} \notin P$, implying that the set $\overline{J}$ as defined in Eq. (2.13) is nonempty. Note that $b_j > 0$ for all j, since J is empty. The rest of the proof is similar to the preceding case.

As an example, where $\text{cone}(P)$ is not polyhedral when P does not contain the origin, consider the polyhedral set $P \subset \Re^2$ given by

$$P = \big\{(x_1, x_2) \mid x_1 \geq 0,\ x_2 = 1\big\}.$$

Then, we have

$$\text{cone}(P) = \big\{(x_1, x_2) \mid x_1 > 0,\ x_2 > 0\big\} \cup \big\{(x_1, x_2) \mid x_1 = 0, x_2 \geq 0\big\},$$

which is not closed and therefore not polyhedral.

2.20

Show that if P is a polyhedral set in $\Re^n$ containing the origin, then $\text{cone}(P)$ is a polyhedral cone. Give an example showing that if P does not contain the origin, then $\text{cone}(P)$ may not be a polyhedral cone.

Solution: We give two proofs. The first is based on the Minkowski-Weyl Representation of a polyhedral set P (cf. Prop. 2.3.3), while the second is based on a representation of P by a system of linear inequalities.

Let P be a polyhedral set with Minkowski-Weyl representation

$$P = \left\{ x \;\middle|\; x = \sum_{j=1}^m \mu_j v_j + y,\ \sum_{j=1}^m \mu_j = 1,\ \mu_j \geq 0,\ j = 1, \ldots, m,\ y \in C \right\},$$

where $v_1, \ldots, v_m$ are some vectors in $\Re^n$ and C is a finitely generated cone in $\Re^n$. Let C be given by

$$C = \left\{ y \;\middle|\; y = \sum_{i=1}^r \lambda_i a_i,\ \lambda_i \geq 0,\ i = 1, \ldots, r \right\},$$

where $a_1, \ldots, a_r$ are some vectors in $\Re^n$, so that

$$P = \left\{ x \;\Big|\; x = \sum_{j=1}^{m} \mu_j v_j + \sum_{i=1}^{r} \lambda_i a_i, \; \sum_{j=1}^{m} \mu_j = 1, \; \mu_j \geq 0, \; \forall\, j, \; \lambda_i \geq 0, \; \forall\, i \right\}.$$

We claim that

$$\text{cone}(P) = \text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big).$$

Since $P \subset \text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big)$, it follows that

$$\text{cone}(P) \subset \text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big).$$

Conversely, let $y \in \text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big)$. Then, we have

$$y = \sum_{j=1}^{m} \overline{\mu}_j v_j + \sum_{i=1}^{r} \overline{\lambda}_i a_i,$$

with $\overline{\mu}_j \geq 0$ and $\overline{\lambda}_i \geq 0$ for all i and j. If $\overline{\mu}_j = 0$ for all j, then $y = \sum_{i=1}^{r} \overline{\lambda}_i a_i \in C$, and since $C = R_P$ (cf. Exercise 2.16), it follows that $y \in R_P$. Because the origin belongs to P and $y \in R_P$, we have $0 + y \in P$, implying that $y \in P$, and consequently $y \in \text{cone}(P)$. If $\overline{\mu}_j > 0$ for some j, then by setting $\overline{\mu} = \sum_{j=1}^{m} \overline{\mu}_j$, $\mu_j = \overline{\mu}_j / \overline{\mu}$ for all j, and $\lambda_i = \overline{\lambda}_i / \overline{\mu}$ for all i, we obtain

$$y = \overline{\mu} \left(\sum_{j=1}^{m} \mu_j v_j + \sum_{i=1}^{r} \lambda_i a_i \right),$$

where $\overline{\mu} > 0$, $\mu_j \geq 0$ with $\sum_{j=1}^{m} \mu_j = 1$, and $\lambda_i \geq 0$. Therefore $y = \overline{\mu}\, \overline{x}$ with $\overline{x} \in P$ and $\overline{\mu} > 0$, implying that $y \in \text{cone}(P)$ and showing that

$$\text{cone}\big(\{v_1, \ldots, v_m, a_1, \ldots, a_r\}\big) \subset \text{cone}(P).$$

We now give an alternative proof using the representation of P by a system of linear inequalities. Let P be given by

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

where $a_1, \ldots, a_r$ are vectors in $\Re^n$ and $b_1, \ldots, b_r$ are scalars. Since P contains the origin, it follows that $b_j \geq 0$ for all j. Define the index set J as follows

$$J = \{ j \mid b_j = 0 \}.$$

We consider separately the two cases where $J \neq \emptyset$ and $J = \emptyset$. If $J \neq \emptyset$, then we will show that

$$\text{cone}(P) = \big\{ x \mid a_j' x \leq 0, \; j \in J \big\}.$$

To see this, note that since $P \subset \big\{x \mid a_j'x \le 0,\ j \in J\big\}$, we have

$$\text{cone}(P) \subset \big\{x \mid a_j'x \le 0,\ j \in J\big\}.$$

Conversely, let $\overline{x} \in \big\{x \mid a_j'x \le 0,\ j \in J\big\}$. We will show that $\overline{x} \in \text{cone}(P)$. If $\overline{x} \in P$, then $\overline{x} \in \text{cone}(P)$ and we are done, so assume that $\overline{x} \notin P$, implying that the set

$$\overline{J} = \{j \notin J \mid a_j'\overline{x} > b_j\} \tag{2.14}$$

is nonempty. By the definition of J, we have $b_j > 0$ for all $j \notin J$, so let

$$\mu = \min_{j \in \overline{J}} \frac{b_j}{a_j'\overline{x}},$$

and note that $0 < \mu < 1$. We have

$$a_j'(\mu\overline{x}) \le 0, \qquad \forall\ j \in J,$$

$$a_j'(\mu\overline{x}) \le b_j, \qquad \forall\ j \in \overline{J}.$$

For $j \notin \overline{J} \cup J$ and $a_j'\overline{x} \le 0 < b_j$, since $\mu > 0$, we still have $a_j'(\mu\overline{x}) \le 0 < b_j$. For $j \notin \overline{J} \cup J$ and $0 < a_j'\overline{x} \le b_j$, since $\mu < 1$, we have $0 < a_j'(\mu\overline{x}) < b_j$. Therefore, $\mu\overline{x} \in P$, implying that $\overline{x} = \frac{1}{\mu}(\mu\overline{x}) \in \text{cone}(P)$. It follows that

$$\big\{x \mid a_j'x \le 0,\ j \in J\big\} \subset \text{cone}(P),$$

and hence, $\text{cone}(P) = \big\{x \mid a_j'x \le 0,\ j \in J\big\}$.

If $J = \emptyset$, then we will show that $\text{cone}(P) = \Re^n$. To see this, take any $\overline{x} \in \Re^n$. If $\overline{x} \in P$, then clearly $\overline{x} \in \text{cone}(P)$, so assume that $\overline{x} \notin P$, implying that the set $\overline{J}$ as defined in Eq. (2.14) is nonempty. Note that $b_j > 0$ for all j, since J is empty. The rest of the proof is similar to the preceding case.

As an example, where $\text{cone}(P)$ is not polyhedral when P does not contain the origin, consider the polyhedral set $P \subset \Re^2$ given by

$$P = \big\{(x_1, x_2) \mid x_1 \ge 0,\ x_2 = 1\big\}.$$

Then, we have

$$\text{cone}(P) = \big\{(x_1, x_2) \mid x_1 > 0,\ x_2 > 0\big\} \cup \big\{(x_1, x_2) \mid x_1 = 0, x_2 \ge 0\big\},$$

which is not closed and therefore not polyhedral.

2.21 (Support Function of a Polyhedral Set)

Show that the support function of a polyhedral set is a polyhedral function.

Solution: Let X be a polyhedral set with Minkowski-Weyl representation

$$X = \text{conv}\big(\{v_1, \ldots, v_m\}\big) + \text{cone}\big(\{d_1, \ldots, d_r\}\big)$$

for some vectors $v_1, \ldots, v_m, d_1, \ldots, d_r$ (cf. Prop. 2.3.3). The support function of X takes the form

$$\begin{aligned}\sigma_X(y) &= \sup_{x \in X} y'x \\ &= \sup_{\substack{\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_r \geq 0 \\ \sum_{i=1}^m \alpha_i = 1}} \left\{ \sum_{i=1}^m \alpha_i v_i' y + \sum_{j=1}^r \beta_j d_j' y \right\} \\ &= \begin{cases} \max_{i=1,\ldots,m} v_i' y & \text{if } d_j' y \leq 0,\ j = 1, \ldots, r, \\ \infty & \text{otherwise.} \end{cases}\end{aligned}$$

Thus the support function is polyhedral.

2.22 (Conjugate of a Polyhedral Function)

(a) Show that the conjugate of a function can be specified in terms of the support function of its epigraph with the formula

$$f^\star(y) = \sigma_{\text{epi}(f)}(y, -1).$$

(b) Use part (a) to show that the conjugate of a polyhedral function is polyhedral.

Solution: (a) We have

$$f^\star(y) = \sup_{x \in \Re^n} \big\{ x'y - f(x) \big\},$$

which can equivalently be written as

$$f^\star(y) = \sup_{(x,w) \in \text{epi}(f)} \{x'y - w\}.$$

Since the expression in braces in the right-hand side is the inner product of the vectors (x, w) and $(y, -1)$, the supremum above is the value of the support function of epi(f) at $(y, -1)$:

$$f^\star(y) = \sigma_{\text{epi}(f)}(y, -1), \qquad \forall\, y \in \Re^n.$$

(See Fig. 2.2.)

(b) Let us apply the result of part (a) to the case where f is a polyhedral function, so that epi(f) is a polyhedral set. From Exercise 2.21, the support function $\sigma_{\text{epi}(f)}$ is a polyhedral function, and it can be seen that $\sigma_{\text{epi}(f)}(y, -1)$, viewed as a function of y, is polyhedral.

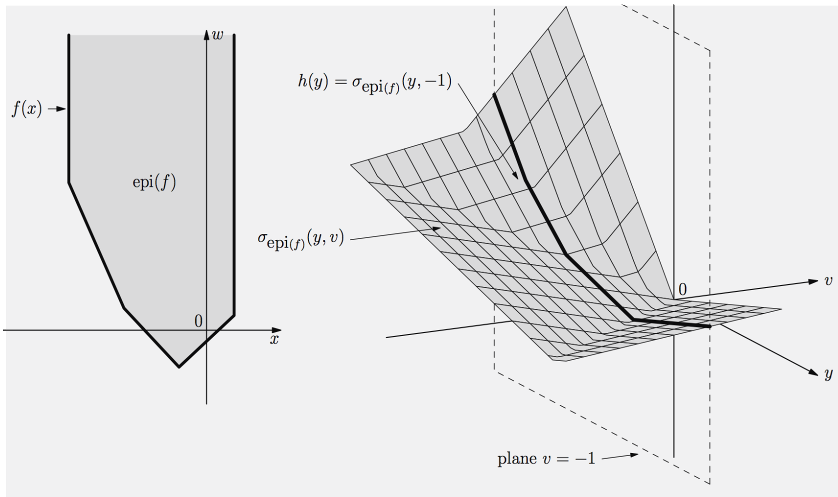

Figure 2.2. Construction of the conjugate of a function f from the support function $\sigma_{\mathrm{epi}(f)}(y,v)$ of $\mathrm{epi}(f)$ (cf. Exercise 2.22). The conjugate is obtained by setting $v=-1$:

$$h(y) = \sigma_{\mathrm{epi}(f)}(y,-1), \qquad \forall\, y \in \Re^n.$$

If f is polyhedral as in the figure, then $\mathrm{epi}(f)$ and $\sigma_{\mathrm{epi}(f)}(y,v)$ are polyhedral, so the conjugate is also polyhedral.

2.23 (Polar Sets)

This exercise introduces a notion of polar set that generalizes the notion of polar cone. Polar sets originated in Euclidean geometry, where they can be used to provide elegant proofs to many classical theorems. Given a nonempty set $C \subset \Re^n$, the *polar set of* C is defined as

$$C^\circ = \{y \mid y'x \le 1, \forall\, x \in C\}.$$

Thus the polar set C° is the level set $\big\{y \mid \sigma_C(y) \le 1\big\}$ of the support function σ_C of C. Since a single level set is sufficient to characterize all level sets of a support function (in view of positive homogeneity), it follows from the Conjugacy Theorem (Prop. 1.6.1), that any set is fully characterized by its polar up to convex closure, i.e., two sets with the same polar set have the same convex closure.

(a) Show that C° is a closed convex set. Furthermore, C° is bounded if and only if the origin in an interior point of $\mathrm{conv}(C)$.

(b) Show that the polar set of a cone is equal to its polar cone.

(c) Consider the subset $\hat{C}$ of $\Re^{n+1}$ obtained from C via the lifting procedure,

$$\hat{C} = \big\{(x,1) \mid x \in C\big\}.$$

Show that C° is obtained from the polar of the cone generated by $\hat{C}$, by "slicing" at the level -1:

$$C^\circ = \Big\{y \mid (y,-1) \in \big(\mathrm{cone}(\hat{C})\big)^*\Big\}.$$

(d) Show that if C is a finite set, then C° is a polyhedral set.

(e) Show that

$$(C^\circ)^\circ = \mathrm{cl}\big(\mathrm{conv}\big(\{0\} \cup C\big)\big),$$

so if C is a closed convex set containing the origin, we have $(C^\circ)^\circ = C$.

(f) Consider a bounded polyhedral set P. For each extreme point v of P, consider the halfspace $H_v = \{y \mid y'v \leq 1\}$. Show that the polar set P° is the intersection of the halfspaces H_v, where v ranges over the extreme points of P.

(g) Consider a circle in the plane that is centered at the origin, and a convex polygon that is inscribed in the circle and contains the origin in its interior. Show that the polar set is a polygon that can be circumscribed around some circle centered at the origin.

Solution: (a) Clearly, we have

$$0 \in \mathrm{int}(\mathrm{conv}(C)) \qquad \Longleftrightarrow \qquad \sigma_C(y) > 0, \qquad \forall\, y \neq 0.$$

Since σ_C is positively homogeneous, it is equal to its recession function, so

$$0 \in \mathrm{int}(\mathrm{conv}(C)) \qquad \Longleftrightarrow \qquad R_{\sigma_C} = \{0\}.$$

Since $R_{\sigma_C} = \{0\}$ if and only if the nonempty level sets of σ_C are compact, and C° is a level set, we have

$$0 \in \mathrm{int}(\mathrm{conv}(C)) \qquad \Longleftrightarrow \qquad \big\{y \mid \sigma_C(y) \leq 1\big\} = C^\circ \text{ is compact.}$$

(b) If C is a cone, by Example 5.2.2, σ_C is the indicator function of the polar set C^*. Since C° is a nonempty level set of σ_C, it follows that $C^\circ = C^\star$.

(c) Using the definition of cone:

$$\mathrm{cone}(\hat{C}) = \big\{(\lambda x, \lambda) \mid x \in C,\ \lambda > 0\big\}.$$

Using the definition of polar cone:

$$(\mathrm{cone}(\hat{C}))^\star = \big\{(y, w) \mid y'\lambda x + w'\lambda \leq 0, x \in C, \lambda > 0\big\}$$

Therefore

$$\begin{aligned}\{y \mid (y, -1) \in (\mathrm{cone}(\hat{C}))^\star\} &= \{y \mid y'\lambda x - \lambda \leq 0, x \in C, \lambda > 0\} \\ &= \{y \mid y'x \leq 1\} \\ &= C^\circ\end{aligned}$$

(d) If C is a finite set, C° is the intersection of a finite number of halfspaces. Furthermore, C° is nonempty since it contains the origin, so it is polyhedral.

(e) We first show that

$$C^\circ = (\mathrm{cl}(C)^\circ) = (\mathrm{conv}(C)^\circ) = (\{0\} \cap C)^\circ.$$

The first two equations hold because C, $\mathrm{cl}(C)$, and $\mathrm{conv}(C)$ have the same support function. The third equation is true by the definition of polar set.

Assume that C is closed, convex, and contains the origin. To show that $(C^\circ)^\circ = C$, note that

$$\begin{aligned}(C^\circ)^\circ &= \{x \mid x'y \le 1,\ \forall\ y \in C^\circ\} \\ &= \big\{x \mid (x,1)'(y,-1) \le 0,\ \forall\ (y,-1) \in D^*\big\},\end{aligned}$$

where $D = \mathrm{cone}(\hat{C})$ and the second equation follows from part (c). Since D is a cone, its polar set is a cone by part (b). We write the above equation as

$$(C^\circ)^\circ = \big\{x \mid (x,1)'(\lambda y, -\lambda) \le 0,\ \forall\ (y,-1) \in D^*,\ \forall\ \lambda > 0)\big\},$$

or equivalently,

$$(C^\circ)^\circ = \big\{x \mid (x,1)'(\bar{y}) \le 0,\ \forall\ \bar{y} \in D^\star\big\},$$

and note that

$$\big\{(\lambda x, \lambda) \mid (x,1)'\bar{y} \le 0,\ \forall\ \bar{y} \in D^\star\big\} = (D^*)^*$$

and $(D^*)^* = D$ because $\hat{C}$ is closed and convex. Now it follows that

$$(C^\circ)^\circ = \big\{x \mid (x,1) \in (D^\star)^\star = D\big\}$$

where D is the cone of "lifted" C. Therefore $(C^\circ)^\circ = C$. For an arbitrary set C without any assumption,

$$\mathrm{cl}(\mathrm{conv}(\{0\} \cap C)) = \big(\mathrm{cl}(\mathrm{conv}(\{0\} \cap C)^\circ)\big)^\circ = (C^\circ)^\circ.$$

(f) By the Minkowski-Weyl representation, a bounded polyhedral set P is the convex hull of its extreme points. Thus

$$P\circ = \{y \mid y'x \le 1,\ \forall\ x \in \mathrm{conv}(\{v_1, v_2, ..., v_r\})\}.$$

We have

$$y'x \le 1, \qquad \forall\ x \in \mathrm{conv}(\{v_1, ..., v_r\}) \qquad \Longleftrightarrow \qquad y'v_i \le 1,\ \forall\ i = 1, ..., r.$$

Therefore

$$P\circ = \{y \mid y'x \le 1,\ \forall\ x \in \mathrm{conv}(\{v_1, v_2, ..., v_r\})\} = H_{v_1} \cap ... \cap H_{v_r}.$$

(g) Using part (f), the polar set of the convex polygon P is the intersection of $H_v = \{y \mid y'v \le 1\}$, where v ranges over the extreme points of P. Furthermore, P° is bounded because 0 is an interior point of P, and it is polyhedral because it is the intersection of a finite number of halfspaces.

If P is inscribed in the circle $\{x \mid \|x\| = r\}$, all extreme points v satisfy $\|v\| = r$, and H_v corresponds to a tangent hyperplane of the circle centered at origin with radius $1/r$. Thus, the intersection of H_v can be circumscribed around the circle $\{x \mid \|x\| = 1/r\}$.

SECTION 2.4: Polyhedral Aspects of Optimization

2.24 (Gordan's Theorem of the Alternative [Gor1873])

Let $a_1, \ldots, a_r$ be vectors in $\Re^n$.

(a) Show that exactly one of the following two conditions holds:

(i) There exists a vector $x \in \Re^n$ such that

$$a_1' x < 0, \ldots, a_r' x < 0.$$

(ii) There exists a vector $\mu \in \Re^r$ such that $\mu \neq 0$, $\mu \geq 0$, and

$$\mu_1 a_1 + \cdots + \mu_r a_r = 0.$$

(b) Show that an equivalent statement of part (a) is the following: a polyhedral cone has nonempty interior if and only if its polar cone does not contain a line, i.e., a set of the form $\{x + \alpha z \mid \alpha \in \Re\}$, where x lies in the polar cone and z is a nonzero vector.

Note: This result is also given with an alternative proof in Section 5.6.

Solution: (a) Assume that there exist $\hat{x} \in \Re^n$ and $\mu \in \Re^r$ such that both conditions (i) and (ii) hold, i.e.,

$$a_j' \hat{x} < 0, \qquad \forall\, j = 1, \ldots, r, \tag{2.15}$$

$$\mu \neq 0, \qquad \mu \geq 0, \qquad \sum_{j=1}^{r} \mu_j a_j = 0. \tag{2.16}$$

By premultiplying Eq. (2.15) with $\mu_j \geq 0$ and summing the obtained inequalities over j, we have

$$\sum_{j=1}^{r} \mu_j a_j' \hat{x} < 0.$$

On the other hand, from Eq. (2.16), we obtain

$$\sum_{j=1}^{r} \mu_j a_j' \hat{x} = 0,$$

which is a contradiction. Hence, both conditions (i) and (ii) cannot hold simultaneously.

The proof will be complete if we show that the conditions (i) and (ii) cannot *fail* to hold simultaneously. Assume that condition (i) fails to hold, and consider the sets given by

$$C_1 = \left\{ w \in \Re^r \mid a_j' x \leq w_j,\ j = 1, \ldots, r,\ x \in \Re^n \right\},$$

$$C_2 = \{ \xi \in \Re^r \mid \xi_j < 0,\ j = 1, \ldots, r \}.$$

It can be seen that both C_1 and C_2 are convex. Furthermore, because the condition (i) does not hold, C_1 and C_2 are disjoint sets. Therefore, by the Separating Hyperplane Theorem (Prop. 1.5.2), C_1 and C_2 can be separated, i.e., there exists a nonzero vector $\mu \in \Re^r$ such that

$$\mu' w \geq \mu' \xi, \qquad \forall\, w \in C_1, \qquad \forall\, \xi \in C_2,$$

implying that

$$\inf_{w \in C_1} \mu' w \geq \mu' \xi, \qquad \forall\, \xi \in C_2.$$

Since each component ξ_j of $\xi \in C_2$ can be any negative scalar, for the preceding relation to hold, μ_j must be nonnegative for all j. Furthermore, by letting $\xi \to 0$, in the preceding relation, it follows that

$$\inf_{w \in C_1} \mu' w \geq 0,$$

implying that

$$\mu_1 w_1 + \cdots + \mu_r w_r \geq 0, \qquad \forall\, w \in C_1.$$

By setting $w_j = a_j' x$ for all j, we obtain

$$(\mu_1 a_1 + \cdots + \mu_r a_r)' x \geq 0, \qquad \forall\, x \in \Re^n,$$

and because this relation holds for all $x \in \Re^n$, we must have

$$\mu_1 a_1 + \cdots + \mu_r a_r = 0.$$

Hence, the condition (ii) holds, showing that the conditions (i) and (ii) cannot fail to hold simultaneously.

Alternative proof: We will show the equivalent statement of part (b), i.e., that a polyhedral cone contains an interior point if and only if the polar C^* does not contain a line.

Let

$$C = \big\{ x \mid a_j' x \leq 0, j = 1, \ldots, r \big\},$$

where $a_j \neq 0$ for all j. Assume that C contains an interior point, and to arrive at a contradiction, assume that C^* contains a line. Then there exists a $d \neq 0$ such that d and $-d$ belong to C^*, i.e., $d'x \leq 0$ and $-d'x \leq 0$ for all $x \in C$, so that $d'x = 0$ for all $x \in C$. Thus for the interior point $\overline{x} \in C$, we have $d'\overline{x} = 0$, and since $d \in C^*$ and $d = \sum_{j=1}^r \mu_j a_j$ for some $\mu_j \geq 0$, we have

$$\sum_{j=1}^r \mu_j a_j' \overline{x} = 0.$$

This is a contradiction, since $\overline{x}$ is an interior point of C, and we have $a_j' \overline{x} < 0$ for all j.

Conversely, assume that C^* does not contain a line. Then C^* has an extreme point, and since the origin is the only possible extreme point of a cone, it follows that the origin is an extreme point of C^*, which is the cone generated by $\{a_1, \ldots, a_r\}$. Therefore $0 \notin \mathrm{conv}\big(\{a_1, \ldots, a_r\}\big)$, and there exists a hyperplane that strictly separates the origin from $\mathrm{conv}\big(\{a_1, \ldots, a_r\}\big)$. Thus, there exists a vector x such that $y'x < 0$ for all $y \in \mathrm{conv}\big(\{a_1, \ldots, a_r\}\big)$, so in particular,

$$a_j'x < 0, \qquad \forall\ j = 1, \ldots, r,$$

and x is an interior point of C.

(b) Let C be a polyhedral cone given by

$$C = \big\{x \mid a_j'x \le 0, j = 1, \ldots, r\big\},$$

where $a_j \neq 0$ for all j. The interior of C is given by

$$\mathrm{int}(C) = \big\{x \mid a_j'x < 0,\ j = 1, \ldots, r\big\},$$

so that C has nonempty interior if and only if the condition (i) of part (a) holds.

By Farkas' Lemma, the polar cone of C is given by

$$C^* = \left\{x \;\middle|\; x = \sum_{j=1}^{r} \mu_j a_j, \mu_j \ge 0,\ j = 1, \ldots, r\right\}.$$

We now show that C^* contains a line if and only if there is a $\mu \in \Re^r$ such that $\mu \neq 0$, $\mu \ge 0$, and $\sum_{j=1}^{r} \mu_j a_j = 0$ [condition (ii) of part (a) holds]. Suppose that C^* contains a line, i.e., a set of the form $\{x + \alpha z \mid \alpha \in \Re\}$, where $x \in C^*$ and z is a nonzero vector. Since C^* is a closed convex cone, by the Recession Cone Theorem (Prop. 1.4.1), it follows that z and $-z$ belong to R_{C^*}. This, implies that $0 + z = z \in C^*$ and $0 - z = -z \in C^*$, and therefore z and $-z$ can be represented as

$$z = \sum_{j=1}^{r} \mu_j a_j, \qquad \forall\ j,\ \mu_j \ge 0,\ \mu_j \neq 0 \text{ for some } j,$$

$$-z = \sum_{j=1}^{r} \overline{\mu}_j a_j, \qquad \forall\ j,\ \overline{\mu}_j \ge 0,\ \overline{\mu}_j \neq 0 \text{ for some } j.$$

Thus, $\sum_{j=1}^{r} (\mu_j + \overline{\mu}_j) a_j = 0$, where $(\mu_j + \overline{\mu}_j) \ge 0$ for all j and $(\mu_j + \overline{\mu}_j) \neq 0$ for at least one j, showing that the condition (ii) of part (a) holds.

Conversely, suppose that $\sum_{j=1}^{r} \mu_j a_j = 0$ with $\mu_j \ge 0$ for all j and $\mu_j \neq 0$ for some j. Assume without loss of generality that $\mu_1 > 0$, so that

$$-a_1 = \sum_{j \neq 1} \frac{\mu_j}{\mu_1} a_j,$$

with $\mu_j/\mu_1 \ge 0$ for all j, which implies that $-a_1 \in C^*$. Since $a_1 \in C^*$, $-a_1 \in C^*$, and $a_1 \neq 0$, it follows that C^* contains a line, completing the proof.

2.25 (Linear System Alternatives)

Let $a_1, \ldots, a_r$ be vectors in $\Re^n$ and let $b_1, \ldots b_r$ be scalars. Show that exactly one of the following two conditions holds:

(i) There exists a vector $x \in \Re^n$ such that

$$a_1'x \le b_1, \ldots, a_r'x \le b_r.$$

(ii) There exists a vector $\mu \in \Re^r$ such that $\mu \ge 0$ and

$$\mu_1 a_1 + \cdots + \mu_r a_r = 0, \qquad \mu_1 b_1 + \cdots + \mu_r b_r < 0.$$

Note: This result is a special case of Motzkin's Transposition Theorem, given with an alternative proof in Section 5.6.

Solution: Assume that there exist $\hat{x} \in \Re^n$ and $\mu \in \Re^r$ such that both conditions (i) and (ii) hold, i.e.,

$$a_j'\hat{x} \le b_j, \qquad \forall\, j = 1, \ldots, r, \tag{2.17}$$

$$\mu \ge 0, \qquad \sum_{j=1}^{r} \mu_j a_j = 0, \qquad \sum_{j=1}^{r} \mu_j b_j < 0. \tag{2.18}$$

By premultiplying Eq. (2.17) with $\mu_j \ge 0$ and summing the obtained inequalities over j, we have

$$\sum_{j=1}^{r} \mu_j a_j'\hat{x} \le \sum_{j=1}^{r} \mu_j b_j.$$

On the other hand, by using Eq. (2.18), we obtain

$$\sum_{j=1}^{r} \mu_j a_j'\hat{x} = 0 > \sum_{j=1}^{r} \mu_j b_j,$$

which is a contradiction. Hence, both conditions (i) and (ii) cannot hold simultaneously.

The proof will be complete if we show that conditions (i) and (ii) cannot *fail* to hold simultaneously. Assume that condition (i) fails to hold, and consider the sets given by

$$P_1 = \{\xi \in \Re^r \mid \xi_j \le 0,\ j = 1, \ldots, r\},$$

$$P_2 = \big\{w \in \Re^r \mid a_j'x - b_j = w_j,\ j = 1, \ldots, r,\ x \in \Re^n\big\}.$$

Clearly, P_1 is a polyhedral set. For the set P_2, we have

$$P_2 = \{w \in \Re^r \mid Ax - b = w,\ x \in \Re^n\} = R(A) - b,$$

where A is the matrix with rows a_j' and b is the vector with components b_j. Thus, P_2 is an affine set and is therefore polyhedral. Furthermore, because the condition (i) does not hold, P_1 and P_2 are disjoint polyhedral sets, and they can be strictly separated [Prop. 1.5.3 under condition (3)]. Hence, there exists a vector $\mu \in \Re^r$ such that

$$\sup_{\xi \in P_1} \mu'\xi < \inf_{w \in P_2} \mu' w.$$

Since each component ξ_j of $\xi \in P_1$ can be any negative scalar, for the preceding relation to hold, μ_j must be nonnegative for all j. Furthermore, since $0 \in P_1$, it follows that

$$0 < \inf_{w \in P_2} \mu' w,$$

implying that

$$0 < \mu_1 w_1 + \cdots + \mu_r w_r, \qquad \forall\ w \in P_2.$$

By setting $w_j = a_j'x - b_j$ for all j, we obtain

$$\mu_1 b_1 + \cdots + \mu_r b_r < (\mu_1 a_1 + \cdots + \mu_r a_r)' x, \qquad \forall\ x \in \Re^n.$$

Since this relation holds for all $x \in \Re^n$, we must have

$$\mu_1 a_1 + \cdots + \mu_r a_r = 0,$$

implying that

$$\mu_1 b_1 + \cdots + \mu_r b_r < 0.$$

Hence, the condition (ii) holds, showing that the conditions (i) and (ii) cannot fail to hold simultaneously.

2.26 (Integer Programming and Unimodular Matrices)

Integer programming problems are optimization problems, which as part of their constraints include the requirement that the optimization variables take integer values, such as 0 or 1. An important method for solving such problems relies on the solution of a continuous optimization problem, called the *relaxed problem*, which is derived from the original by neglecting the integer constraints while maintaining all the other constraints. If the relaxed problem happens to have integer components, it will then solve optimally not just the relaxed problem, but also the original integer programming problem. Thus, polyhedral sets whose extreme points have integer components are of special significance. We will characterize an important class of such sets.

Let us say that a square matrix with integer components is *unimodular* if its determinant is 0, 1, or -1, and let us say that a rectangular matrix with integer components is *totally unimodular* if each of its square submatrices is unimodular. If A is an invertible matrix, by Cramer's rule, its inverse A^{-1} has components of the form

$$\left[A^{-1}\right]_{ij} = \frac{\text{polynomial in the components of } A}{\text{determinant of } A}.$$

It follows that if A is an invertible matrix with integer components that is unimodular, its inverse has integer components. Furthermore, for any vector b with integer components, the unique solution $A^{-1}b$ of the system

$$Ax = b$$

has integer components.

Let P be a polyhedral set of the form

$$P = \{x \mid Ax = b,\, c \leq x \leq d\},$$

where A is an $m \times n$ matrix, b is a vector in $\Re^m$, and c and d are vectors in $\Re^n$. Assume that all the components of A, b, c, and d are integer, and that A is totally unimodular. Show that all the extreme points of P have integer components.

Solution: Let v be an extreme point of P. Consider the subset of indices

$$I = \{i \mid c_i < v_i < d_i\},$$

and without loss of generality, assume that

$$I = \{1, \ldots, \overline{m}\}$$

for some integer $\overline{m}$. Let $\overline{A}$ be the matrix consisting of the first $\overline{m}$ columns of A and let $\overline{v}$ be the vector consisting of the first $\overline{m}$ components of v. Note that each of the last $n - \overline{m}$ components of v is equal to either the corresponding component of c or to the corresponding component of d, which are integer. Thus the extreme point v has integer components if and only if the subvector $\overline{v}$ has integer components.

By Prop. 2.1.4, $\overline{A}$ has linearly independent columns, so $\overline{v}$ is the unique solution of the system of equations

$$\overline{A}y = \overline{b},$$

where $\overline{b}$ is equal to b minus the last $n - \overline{m}$ columns of A multiplied with the corresponding components of v (each of which is equal to either the corresponding component of c or the corresponding component of d, so that $\overline{b}$ has integer components). Equivalently, there exists an invertible $\overline{m} \times \overline{m}$ submatrix $\tilde{A}$ of $\overline{A}$ and a subvector $\tilde{b}$ of $\overline{b}$ with $\overline{m}$ components such that

$$\overline{v} = (\tilde{A})^{-1}\tilde{b}.$$

Since by hypothesis, A is totally unimodular, the invertible submatrix $\tilde{A}$ is unimodular, and it follows that $\overline{v}$ (and hence also the extreme point v) has integer components.

2.27 (Unimodularity I)

Let A be an $n \times n$ invertible matrix with integer entries. Show that A is unimodular if and only if the solution of the system $Ax = b$ has integer components for every vector $b \in \Re^n$ with integer components. *Hint*: To prove that A is unimodular when the given property holds, use the system $Ax = u_i$, where u_i is the ith unit vector, to show that A^{-1} has integer components, and then use the equality $\det(A) \cdot \det(A^{-1}) = 1$. To prove the converse, use Cramer's rule.

Solution: Suppose that the system $Ax = b$ has integer components for every vector $b \in \Re^n$ with integer components. Since A is invertible, it follows that the vector $A^{-1}b$ has integer components for every $b \in \Re^n$ with integer components. For $i = 1, \ldots, n$, let e_i be the vector with ith component equal to 1 and all other components equal to 0. Then, for $b = e_i$, the vectors $A^{-1}e_i$, $i = 1, \ldots, n$, have integer components, implying that the columns of A^{-1} are vectors with integer components, so that A^{-1} has integer entries. Therefore, $\det(A^{-1})$ is integer, and since $\det(A)$ is also integer and $\det(A) \cdot \det(A^{-1}) = 1$, it follows that either $\det(A) = 1$ or $\det(A) = -1$, showing that A is unimodular.

Suppose now that A is unimodular. Take any vector $b \in \Re^n$ with integer components, and for each $i \in \{1, \ldots, n\}$, let A_i be the matrix obtained from A by replacing the ith column of A with b. Then, according to Cramer's rule, the components of the solution $\hat{x}$ of the system $Ax = b$ are given by

$$\hat{x}_i = \frac{\det(A_i)}{\det(A)}, \qquad i = 1, \ldots, n.$$

Since each matrix A_i has integer entries, it follows that $\det(A_i)$ is integer for all $i = 1, \ldots, n$. Furthermore, because A is invertible and unimodular, we have either $\det(A) = 1$ or $\det(A) = -1$, implying that the vector $\hat{x}$ has integer components.

2.28 (Unimodularity II)

Let A be an $m \times n$ matrix.

(a) Show that A is totally unimodular if and only if its transpose A' is totally unimodular.

(b) Show that A is totally unimodular if and only if every subset J of $\{1, \ldots, n\}$ can be partitioned into two subsets J_1 and J_2 such that

$$\left| \sum_{j \in J_1} a_{ij} - \sum_{j \in J_2} a_{ij} \right| \leq 1, \qquad \forall\ i = 1, \ldots, m.$$

Solution: (a) The proof is straightforward from the definition of the totally unimodular matrix and the fact that B is a submatrix of A if and only if B' is a submatrix of A'.

(b) Suppose that A is totally unimodular. Let J be a subset of $\{1, \ldots, n\}$. Define z by $z_j = 1$ if $j \in J$, and $z_j = 0$ otherwise. Also let $w = Az$, $c_i = d_i = \frac{1}{2}w_i$ if w_i is even, and $c_i = \frac{1}{2}(w_i - 1)$ and $d_i = \frac{1}{2}(w_i + 1)$ if w_i is odd. Consider the polyhedral set

$$P = \{x \mid c \leq Ax \leq d,\ 0 \leq x \leq z\},$$

and note that $P \neq \emptyset$ because $\frac{1}{2}z \in P$. Since A is totally unimodular, the polyhedron P has integer extreme points. Let $\hat{x} \in P$ be one of them. Because $0 \leq \hat{x} \leq z$ and $\hat{x}$ has integer components, it follows that $\hat{x}_j = 0$ for $j \notin J$ and

$\hat{x}_j \in \{0,1\}$ for $j \in J$. Therefore, $z_j - 2\hat{x}_j = \pm 1$ for $j \in J$. Define $J_1 = \{j \in J \mid z_j - 2\hat{x}_j = 1\}$ and $J_2 = \{j \in J \mid z_j - 2\hat{x}_j = -1\}$. We have

$$\begin{aligned}\sum_{j\in J_1} a_{ij} - \sum_{j\in J_2} a_{ij} &= \sum_{j\in J} a_{ij}(z_j - 2\hat{x}_j)\\ &= \sum_{j=1}^{n} a_{ij}(z_j - 2\hat{x}_j)\\ &= [Az]_i - 2[A\hat{x}]_i\\ &= w_i - 2[A\hat{x}]_i,\end{aligned}$$

where $[Ax]_i$ denotes the ith component of the vector Ax. If w_i is even, then since $c_i \leq [A\hat{x}]_i \leq d_i$ and $c_i = d_i = \frac{1}{2}w_i$, it follows that $[A\hat{x}]_i = w_i$, so that

$$w_i - 2[A\hat{x}]_i = 0, \qquad \text{when } w_i \text{ is even.}$$

If w_i is odd, then since $c_i \leq [A\hat{x}]_i \leq d_i$, $c_i = \frac{1}{2}(w_i - 1)$, and $d_i = \frac{1}{2}(w_i + 1)$, it follows that

$$\frac{1}{2}(w_i - 1) \leq [A\hat{x}]_i \leq \frac{1}{2}(w_i + 1),$$

implying that

$$-1 \leq w_i - 2[A\hat{x}]_i \leq 1.$$

Because $w_i - 2[A\hat{x}]_i$ is integer, we conclude that

$$w_i - 2[A\hat{x}]_i \in \{-1, 0, 1\}, \qquad \text{when } w_i \text{ is odd.}$$

Therefore,

$$\left| \sum_{j\in J_1} a_{ij} - \sum_{j\in J_2} a_{ij} \right| \leq 1, \quad \forall\ i = 1, \ldots, m. \tag{2.19}$$

Suppose now that the matrix A is such that any $J \subset \{1, \ldots, n\}$ can be partitioned into two subsets so that Eq. (2.19) holds. We prove that A is totally unimodular, by showing that each of its square submatrices is unimodular, i.e., the determinant of every square submatrix of A is -1, 0, or 1. We use induction on the size of the square submatrices of A.

To start the induction, note that for $J \subset \{1, \ldots, n\}$ with J consisting of a single element, from Eq. (2.19) we obtain $a_{ij} \in \{-1, 0, 1\}$ for all i and j. Assume now that the determinant of every $(k-1) \times (k-1)$ submatrix of A is -1, 0, or 1. Let B be a $k \times k$ submatrix of A. If $\det(B) = 0$, then we are done, so assume that B is invertible. Our objective is to prove that $|\det B| = 1$. By Cramer's rule and the induction hypothesis, we have $B^{-1} = \frac{B^*}{\det(B)}$, where $b^*_{ij} \in \{-1, 0, 1\}$. By the definition of B^*, we have $Bb^*_1 = \det(B)e_1$, where b^*_1 is the first column of B^* and $e_1 = (1, 0, \ldots 0)'$.

Let $J = \{j \mid b^*_{j1} \neq 0\}$ and note that $J \neq \emptyset$ since B is invertible. Let $\overline{J}_1 = \{j \in J \mid b^*_{j1} = 1\}$ and $\overline{J}_2 = \{j \in J \mid j \notin \overline{J}_1\}$. Then, since $[Bb^*_1]_i = 0$ for $i = 2, \ldots, k$, we have

$$[Bb^*_1]_i = \sum_{j=1}^{k} b_{ij}b^*_{j1} = \sum_{j\in\overline{J}_1} b_{ij} - \sum_{j\in\overline{J}_2} b_{ij} = 0, \qquad \forall\ i = 2, \ldots, k.$$

Thus, the cardinality of the set J is even, so that for any partition $(\tilde{J}_1, \tilde{J}_2)$ of J, it follows that $\sum_{j\in\tilde{J}_1} b_{ij} - \sum_{j\in\tilde{J}_2} b_{ij}$ is even for all $i = 2, \ldots, k$. By assumption, there is a partition (J_1, J_2) of J such that

$$\left| \sum_{j\in J_1} b_{ij} - \sum_{j\in J_2} b_{ij} \right| \leq 1 \qquad \forall\ i = 1, \ldots, k, \tag{2.20}$$

implying that

$$\sum_{j\in J_1} b_{ij} - \sum_{j\in J_2} b_{ij} = 0, \qquad \forall\ i = 2, \ldots, k. \tag{2.21}$$

Consider now the value $\alpha = \left|\sum_{j\in J_1} b_{1j} - \sum_{j\in J_2} b_{1j}\right|$, for which in view of Eq. (2.20), we have either $\alpha = 0$ or $\alpha = 1$. Define $y \in \Re^k$ by $y_i = 1$ for $i \in J_1$, $y_i = -1$ for $i \in J_2$, and $y_i = 0$ otherwise. Then, we have $\big|[By]_1\big| = \alpha$ and by Eq. (2.21), $[By]_i = 0$ for all $i = 2, \ldots, k$. If $\alpha = 0$, then $By = 0$ and since B is invertible, it follows that $y = 0$, implying that $J = \emptyset$, which is a contradiction. Hence, we must have $\alpha = 1$ so that $By = \pm e_1$. Without loss of generality assume that $By = e_1$ (if $By = -e_1$, we can replace y by $-y$). Then, since $Bb_1^* = \det(B)e_1$, we see that $B\big(b_1^* - \det(B)y\big) = 0$ and since B is invertible, we must have $b_1^* = \det(B)y$. Because y and b_1^* are vectors with components -1, 0, or 1, it follows that $b_1^* = \pm y$ and $\big|\det(B)\big| = 1$, completing the induction and showing that A is totally unimodular.

2.29 (Unimodularity III)

Show that a matrix A is totally unimodular if one of the following holds:

(a) The entries of A are -1, 0, or 1, and there are exactly one 1 and exactly one -1 in each of its columns.

(b) The entries of A are 0 or 1, and in each of its columns, the entries that are equal to 1 appear consecutively.

Solution: (a) We show that the determinant of any square submatrix of A is -1, 0, or 1. We prove this by induction on the size of the square submatrices of A. In particular, the 1×1 submatrices of A are the entries of A, which are -1, 0, or 1. Suppose that the determinant of each $(k-1) \times (k-1)$ submatrix of A is -1, 0, or 1, and consider a $k \times k$ submatrix B of A. If B has a zero column, then $\det(B) = 0$ and we are done. If B has a column with a single nonzero component (1 or -1), then by expanding its determinant along that column and by using the induction hypothesis, we see that $\det(B) = 1$ or $\det(B) = -1$. Finally, if each column of B has exactly two nonzero components (one 1 and one -1), the sum of its rows is zero, so that B is singular and $\det(B) = 0$, completing the proof and showing that A is totally unimodular.

(b) The proof is based on induction as in part (a). The 1×1 submatrices of A are the entries of A, which are 0 or 1. Suppose now that the determinant of each $(k-1)\times(k-1)$ submatrix of A is -1, 0, or 1, and consider a $k\times k$ submatrix B of A.

Since in each column of A, the entries that are equal to 1 appear consecutively, the same is true for the matrix B. Take the first column b_1 of B. If $b_1 = 0$, then B is singular and $\det(B) = 0$. If b_1 has a single nonzero component, then by expanding the determinant of B along b_1 and by using the induction hypothesis, we see that $\det(B) = 1$ or $\det(B) = -1$. Finally, let b_1 have more than one nonzero component (its nonzero entries are 1 and appear consecutively). Let l and p be rows of B such that $b_{i1} = 0$ for all $i < l$ and $i > p$, and $b_{i1} = 1$ for all $l \leq i \leq p$. By multiplying the lth row of B with (-1) and by adding it to the $l+1$st, $l+2$nd, ..., kth row of B, we obtain a matrix $\overline{B}$ such that $\det(B) = \det(\overline{B})$ and the first column $\overline{b}_1$ of $\overline{B}$ has a single nonzero component. Furthermore, the determinant of every square submatrix of $\overline{B}$ is -1, 0, or 1 (this follows from the fact that the determinant of a square matrix is unaffected by adding a scalar multiple of a row of the matrix to some of its other rows, and from the induction hypothesis). Since $\overline{b}_1$ has a single nonzero component, by expanding the determinant of $\overline{B}$ along $\overline{b}_1$, it follows that $\det(\overline{B}) = 1$ or $\det(\overline{B}) = -1$, implying that $\det(B) = 1$ or $\det(B) = -1$, completing the induction and showing that A is totally unimodular.

2.30 (Unimodularity IV)

Let A be a matrix with entries -1, 0, or 1, and exactly two nonzero entries in each of its columns. Show that A is totally unimodular if and only if the rows of A can be divided into two subsets such that for each column the following hold: if the two nonzero entries in the column have the same sign, their rows are in different subsets, and if they have the opposite sign, their rows are in the same subset.

Solution: If A is totally unimodular, then by Exercise 2.28(a), its transpose A' is also totally unimodular, and by Exercise 2.28(b), the set $I = \{1, \ldots, m\}$ can be partitioned into two subsets I_1 and I_2 such that

$$\left| \sum_{i \in I_1} a_{ij} - \sum_{i \in I_2} a_{ij} \right| \leq 1, \qquad \forall\, j = 1, \ldots, n.$$

Since $a_{ij} \in \{-1, 0, 1\}$ and exactly two of $a_{1j}, \ldots, a_{mj}$ are nonzero for each j, it follows that

$$\sum_{i \in I_1} a_{ij} - \sum_{i \in I_2} a_{ij} = 0, \qquad \forall\, j = 1, \ldots, n.$$

Take any $j \in \{1, \ldots, n\}$, and let l and p be such that $a_{ij} = 0$ for all $i \neq l$ and $i \neq p$, so that in view of the preceding relation and the fact $a_{ij} \in \{-1, 0, 1\}$, we see that: if $a_{lj} = -a_{pj}$, then both l and p are in the same subset (I_1 or I_2); if $a_{lj} = a_{pj}$, then l and p are not in the same subset.

Suppose now that the rows of A can be divided into two subsets such that for each column the following property holds: if the two nonzero entries in the column have the same sign, they are in different subsets, and if they have the opposite sign, they are in the same subset. By multiplying all the rows in one of the subsets by -1, we obtain the matrix $\overline{A}$ with entries $\overline{a}_{ij} \in \{-1, 0, 1\}$, and exactly one 1 and exactly one -1 in each of its columns. Therefore, by

Exercise 2.29(a), $\overline{A}$ is totally unimodular, so that every square submatrix of $\overline{A}$ has determinant -1, 0, or 1. Since the determinant of a square submatrix of $\overline{A}$ and the determinant of the corresponding submatrix of A differ only in sign, it follows that every square submatrix of A has determinant -1, 0, or 1, showing that A is totally unimodular.

2.31 (Elementary Vectors [Roc69])

Given a vector $z = (z_1, \ldots, z_n)$ in $\Re^n$, the *support* of z is the set of indices $\{j \mid z_j \neq 0\}$. We say that a nonzero vector z of a subspace S of $\Re^n$ is *elementary* if there is no vector $\overline{z} \neq 0$ in S that has smaller support than z, i.e., for all nonzero $\overline{z} \in S$, $\{j \mid \overline{z}_j \neq 0\}$ is not a strict subset of $\{j \mid z_j \neq 0\}$. Show that:

(a) Two elementary vectors with the same support are scalar multiples of each other.

(b) For every nonzero vector y, there exists an elementary vector with support contained in the support of y.

(c) (*Conformal Realization Theorem*) We say that a vector x is in *harmony* with a vector z if

$$x_j z_j \geq 0, \qquad \forall\, j = 1, \ldots, n.$$

Show that every nonzero vector x of a subspace S can be written in the form

$$x = z^1 + \ldots + z^m,$$

where $z^1, \ldots, z^m$ are elementary vectors of S, and each of them is in harmony with x and has support contained in the support of x. *Note*: Among other subjects, this result finds significant application in network optimization algorithms (see Rockafellar [Roc69] and Bertsekas [Ber98]).

Solution: (a) If two elementary vectors z and $\overline{z}$ had the same support, the vector $z - \gamma\overline{z}$ would be nonzero and have smaller support than z and $\overline{z}$ for a suitable scalar γ. If z and $\overline{z}$ are not scalar multiples of each other, then $z - \gamma\overline{z} \neq 0$, which contradicts the definition of an elementary vector.

(b) We note that either y is elementary or else there exists a nonzero vector $\overline{z}$ with support strictly contained in the support of y. Repeating this argument for at most $n - 1$ times, we must obtain an elementary vector.

(c) We first show that every nonzero vector $y \in S$ has the property that there exists an elementary vector of S that is in harmony with y and has support that is contained in the support of y.

We show this by induction on the number of nonzero components of y. Let V_k be the subset of nonzero vectors in S that have k or less nonzero components, and let $\overline{k}$ be the smallest k for which V_k is nonempty. Then, by part (b), every vector $y \in V_{\overline{k}}$ must be elementary, so it has the desired property. Assume that all vectors in V_k have the desired property for some $k \geq \overline{k}$. We let y be a vector in V_{k+1} and we show that it also has the desired property. Let z be an elementary vector whose support is contained in the support of y. By using the negative of

z if necessary, we can assume that $y_j z_j > 0$ for at least one index j. Then there exists a largest value of γ, call it $\overline{\gamma}$, such that

$$y_j - \gamma z_j \geq 0, \qquad \forall \; j \text{ with } y_j > 0,$$

$$y_j - \gamma z_j \leq 0, \qquad \forall \; j \text{ with } y_j < 0.$$

The vector $y - \overline{\gamma} z$ is in harmony with y and has support that is strictly contained in the support of y. Thus either $y - \overline{\gamma} z = 0$, in which case the elementary vector z is in harmony with y and has support equal to the support of y, or else $y - \overline{\gamma} z$ is nonzero. In the latter case, we have $y - \overline{\gamma} z \in V_k$, and by the induction hypothesis, there exists an elementary vector $\overline{z}$ that is in harmony with $y - \overline{\gamma} z$ and has support that is contained in the support of $y - \overline{\gamma} z$. The vector $\overline{z}$ is also in harmony with y and has support that is contained in the support of y. The induction is complete.

Consider now the given nonzero vector $x \in S$, and choose any elementary vector $\overline{z}^1$ of S that is in harmony with x and has support that is contained in the support of x (such a vector exists by the property just shown). By using the negative of $\overline{z}^1$ if necessary, we can assume that $x_j \overline{z}_j^1 > 0$ for at least one index j. Let $\overline{\gamma}$ be the largest value of γ such that

$$x_j - \gamma \overline{z}_j^1 \geq 0, \qquad \forall \; j \text{ with } x_j > 0,$$

$$x_j - \gamma \overline{z}_j^1 \leq 0, \qquad \forall \; j \text{ with } x_j < 0.$$

The vector $x - z^1$, where

$$z^1 = \overline{\gamma} \, \overline{z}^1,$$

is in harmony with x and has support that is strictly contained in the support of x. There are two cases: (1) $x = z^1$, in which case we are done, or (2) $x \neq z^1$, in which case we replace x by $x - z^1$ and we repeat the process. Eventually, after m steps where $m \leq n$ (since each step reduces the number of nonzero components by at least one), we will end up with the desired decomposition $x = z^1 + \cdots + z^m$.

2.32 (Combinatorial Separation Theorem [Cam68], [Roc69])

Let S be a subspace of $\Re^n$. Consider a set B that is a Cartesian product of n nonempty intervals, and is such that $B \cap S^\perp = \emptyset$ (by an interval, we mean a convex set of scalars, which may be open, closed, or neither open nor closed.) Show that there exists an elementary vector z of S (cf. Exercise 2.31) such that

$$t'z < 0, \qquad \forall \; t \in B,$$

i.e., a hyperplane that separates B and $S^\perp$, and does not contain any point of B. *Note*: There are two points here: (1) The set B need not be closed, as required for application of the Strict Separation Theorem (cf. Prop. 1.5.3), and (2) the hyperplane normal can be one of the elementary vectors of S (not just any vector of S). For application of this result in duality theory for network optimization and monotropic programming, see Rockafellar [Roc84] and Bertsekas [Ber98].

Solution: For simplicity, assume that B is the Cartesian product of bounded open intervals, so that B has the form

$$B = \{t \mid \underline{b}_j < t_j < \overline{b}_j,\ j = 1, \ldots, n\},$$

where $\underline{b}_j$ and $\overline{b}_j$ are some scalars. The proof is easily modified for the case where B has a different form.

Since $B \cap S^\perp = \emptyset$, there exists a hyperplane that separates B and $S^\perp$. The normal of this hyperplane is a nonero vector $d \in S$ such that

$$t'd \leq 0, \qquad \forall\ t \in B.$$

Since B is open, this inequality implies that actually

$$t'd < 0, \qquad \forall\ t \in B.$$

Equivalently, we have

$$\sum_{\{j \mid d_j > 0\}} (\overline{b}_j - \epsilon) d_j + \sum_{\{j \mid d_j < 0\}} (\underline{b}_j + \epsilon) d_j < 0, \tag{2.22}$$

for all $\epsilon > 0$ such that $\underline{b}_j + \epsilon < \overline{b}_j - \epsilon$. Let

$$d = z^1 + \cdots + z^m,$$

be a decomposition of d, where $z^1, \ldots, z^m$ are elementary vectors of S that are in harmony with x, and have supports that are contained in the support of d [cf. part (c) of the Exercise 2.31]. Then the condition (2.22) is equivalently written as

$$\begin{aligned} 0 &> \sum_{\{j \mid d_j > 0\}} (\overline{b}_j - \epsilon) d_j + \sum_{\{j \mid d_j < 0\}} (\underline{b}_j + \epsilon) d_j \\ &= \sum_{\{j \mid d_j > 0\}} (\overline{b}_j - \epsilon) \left(\sum_{i=1}^m z_j^i \right) + \sum_{\{j \mid d_j < 0\}} (\underline{b}_j + \epsilon) \left(\sum_{i=1}^m z_j^i \right) \\ &= \sum_{i=1}^m \left(\sum_{\{j \mid z_j^i > 0\}} (\overline{b}_j - \epsilon) z_j^i + \sum_{\{j \mid z_j^i < 0\}} (\underline{b}_j + \epsilon) z_j^i \right), \end{aligned}$$

where the last equality holds because the vectors z^i are in harmony with d and their supports are contained in the support of d. From the preceding relation, we see that for at least one elementary vector z^i, we must have

$$0 > \sum_{\{j \mid z_j^i > 0\}} (\overline{b}_j - \epsilon) z_j^i + \sum_{\{j \mid z_j^i < 0\}} (\underline{b}_j + \epsilon) z_j^i,$$

for all $\epsilon > 0$ that are sufficiently small and are such that $\underline{b}_j + \epsilon < \overline{b}_j - \epsilon$, or equivalently

$$0 > t'z^i, \qquad \forall\ t \in B.$$

2.33 (Tucker's Complementarity Theorem)

(a) Let S be a subspace of $\Re^n$. Show that there exist disjoint index sets I and $\overline{I}$ with $I \cup \overline{I} = \{1, \ldots, n\}$, and vectors $x \in S$ and $y \in S^\perp$ such that

$$x_i > 0, \quad \forall\, i \in I, \qquad x_i = 0, \quad \forall\, i \in \overline{I},$$

$$y_i = 0, \quad \forall\, i \in I, \qquad y_i > 0, \quad \forall\, i \in \overline{I}.$$

Furthermore, the index sets I and $\overline{I}$ with this property are unique. In addition, we have

$$x_i = 0, \qquad \forall\, i \in \overline{I}, \quad \forall\, x \in S \text{ with } x \geq 0,$$

$$y_i = 0, \qquad \forall\, i \in I, \quad \forall\, y \in S^\perp \text{ with } y \geq 0.$$

Hint: Use a hyperplane separation argument based on Exercise 2.32.

(b) Let A be an $m \times n$ matrix and let b be a vector in $\Re^n$. Assume that the set $F = \{x \mid Ax = b,\ x \geq 0\}$ is nonempty. Apply part (a) to the subspace

$$S = \big\{(x, w) \mid Ax - bw = 0,\ x \in \Re^n,\ w \in \Re\big\},$$

and show that there exist disjoint index sets I and $\overline{I}$ with $I \cup \overline{I} = \{1, \ldots, n\}$, and vectors $x \in F$ and $z \in \Re^m$ such that $b'z = 0$ and

$$x_i > 0, \quad \forall\, i \in I, \qquad x_i = 0, \quad \forall\, i \in \overline{I},$$

$$y_i = 0, \quad \forall\, i \in I, \qquad y_i > 0, \quad \forall\, i \in \overline{I},$$

where $y = A'z$. *Note*: A special choice of A and b yields an important result, which relates optimal primal and dual solutions in linear programming: the Goldman-Tucker Complementarity Theorem [GoT56] (see the exercises of Chapter 5).

Solution: (a) Fix an index k and consider the following two assertions:

(1) There exists a vector $x \in S$ with $x_i \geq 0$ for all i, and $x_k > 0$.

(2) There exists a vector $y \in S^\perp$ with $y_i \geq 0$ for all i, and $y_k > 0$.

We claim that one and only one of the two assertions holds. Clearly, assertions (1) and (2) cannot hold simultaneously, since then we would have $x'y > 0$, while $x \in S$ and $y \in S^\perp$. We will show that they cannot fail simultaneously. Indeed, if (1) does not hold, the Cartesian product $B = \Pi_{i=1}^n B_i$ of the intervals

$$B_i = \begin{cases} (0, \infty) & \text{if } i = k, \\ [0, \infty) & \text{if } i \neq k, \end{cases}$$

does not intersect the subspace S, so by the result of Exercise 2.32, there exists a vector z of $S^\perp$ such that $x'z < 0$ for all $x \in B$. For this to hold, we must have $z \in B^*$ or equivalently $z \leq 0$, while by choosing $x = (0, \ldots, 0, 1, 0, \ldots, 0) \in B$, with the 1 in the kth position, the inequality $x'z < 0$ yields $z_k < 0$. Thus assertion (2) holds with $y = -z$. Similarly, we show that if (2) does not hold, then (1) must hold.

Let now I be the set of indices k such that (1) holds, and for each $k \in I$, let $x(k)$ be a vector in S such that $x(k) \geq 0$ and $x_k(k) > 0$ (note that we do not exclude the possibility that one of the sets I and $\overline{I}$ is empty). Let $\overline{I}$ be the set of indices such that (2) holds, and for each $k \in \overline{I}$, let $y(k)$ be a vector in $S^\perp$ such that $y(k) \geq 0$ and $y_k(k) > 0$. From what has already been shown, I and $\overline{I}$ are disjoint, $I \cup \overline{I} = \{1, \ldots, n\}$, and the vectors

$$x = \sum_{k \in I} x(k), \qquad y = \sum_{k \in \overline{I}} y(k),$$

satisfy

$$x_i > 0, \quad \forall\, i \in I, \qquad x_i = 0, \quad \forall\, i \in \overline{I},$$

$$y_i = 0, \quad \forall\, i \in I, \qquad y_i > 0, \quad \forall\, i \in \overline{I}.$$

The uniqueness of I and $\overline{I}$ follows from their construction and the preceding arguments. In particular, if for some $k \in \overline{I}$, there existed a vector $x \in S$ with $x \geq 0$ and $x_k > 0$, then since for the vector $y(k)$ of $S^\perp$ we have $y(k) \geq 0$ and $y_k(k) > 0$, assertions (a) and (b) must hold simultaneously, which is a contradiction.

The last assertion follows from the fact that for each k, exactly one of the assertions (1) and (2) holds.

(b) Consider the subspace

$$S = \big\{(x, w) \mid Ax - bw = 0,\, x \in \Re^n,\, w \in \Re\big\}.$$

Its orthogonal complement is the range of the transpose of the matrix $[A \ \ -b]$, so it has the form

$$S^\perp = \big\{(A'z, -b'z) \mid z \in \Re^m\big\}.$$

By applying the result of part (a) to the subspace S, we obtain a partition of the index set $\{1, \ldots, n+1\}$ into two subsets. There are two possible cases:

(1) The index $n+1$ belongs to the first subset.

(2) The index $n+1$ belongs to the second subset.

In case (2), the two subsets are of the form I and $\overline{I} \cup \{n+1\}$ with $I \cup \overline{I} = \{1, \ldots, n\}$, and by the last assertion of part (a), we have $w = 0$ for all (x, w) such that $x \geq 0$, $w \geq 0$ and $Ax - bw = 0$. This, however, contradicts the fact that the set $F = \{x \mid Ax = b,\, x \geq 0\}$ is nonempty. Therefore, case (1) holds, i.e., the index $n+1$ belongs to the first index subset. In particular, we have that there exist disjoint index sets I and $\overline{I}$ with $I \cup \overline{I} = \{1, \ldots, n\}$, and vectors (x, w) with $Ax - bw = 0$, and $z \in \Re^m$ such that

$$w > 0, \qquad b'z = 0,$$

$$x_i > 0, \quad \forall\, i \in I, \qquad x_i = 0, \quad \forall\, i \in \overline{I},$$

$$y_i = 0, \quad \forall\, i \in I, \qquad y_i > 0, \quad \forall\, i \in \overline{I},$$

where $y = A'z$. By dividing (x, w) with w if needed, we may assume that $w = 1$ so that $Ax - b = 0$, and the result follows.

REFERENCES

[Ber98] Bertsekas, D. P., 1998. Network Optimization: Continuous and Discrete Models, Athena Scientific, Belmont, MA.

[Cam68] Camion, P., 1968. "Modules Unimodulaires," J. Comb. Theory, Vol. 4, pp. 301-362.

[GoT56] Goldman, A. J., and Tucker, A. W., 1956. "Theory of Linear Programming," in Linear Inequalities and Related Systems, H. W. Kuhn and A. W. Tucker, eds., Princeton University Press, Princeton, N.J., pp. 53-97.

[Gor1873] Gordan, P., 1873. "Uber die Auflosung Linearer Gleichungen mit Reelen Coefficienten," Mathematische Annalen, Vol. 6, pp. 23-28.

[Roc69] Rockafellar, R. T., 1969. "The Elementary Vectors of a Subspace of R^N," in Combinatorial Mathematics and its Applications, by Bose, R. C., and Dowling, T. A., (Eds.), University of North Carolina Press, pp. 104-127.

[Roc84] Rockafellar, R. T., 1984. Network Flows and Monotropic Optimization, Wiley, N. Y.; republished by Athena Scientific, Belmont, MA, 1998.