

Chapter 2

Convex Polyhedra

We begin the chapter by introducing basic concepts of convex sets and linear functions in a Euclidean space. We review some of fundamental facts about convex polyhedral sets determined by systems of linear equations and inequalities, including Farkas' theorem of the alternative which is considered a keystone of the theory of mathematical programming.

2.1 The Space \mathbb{R}^n

Throughout this book, \mathbb{R}^n denotes the n -dimensional Euclidean space of real column n -vectors. The norm of a vector x with components x_1, \dots, x_n is given by

$$\|x\| = \left[\sum_{i=1}^n (x_i)^2 \right]^{1/2} .$$

The inner product of two vectors x and y in \mathbb{R}^n is expressed as

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i .$$

The closed unit ball, the open unit ball and the unit sphere of \mathbb{R}^n are respectively defined by

$$\begin{aligned} B_n &:= \{x \in \mathbb{R}^n : \|x\| \leq 1\}, \\ \text{int}(B_n) &:= \{x \in \mathbb{R}^n : \|x\| < 1\}, \\ S_n &:= \{x \in \mathbb{R}^n : \|x\| = 1\}. \end{aligned}$$

Given a nonempty set $Q \subseteq \mathbb{R}^n$, we denote the *closure* of Q by $\text{cl}(Q)$ and its *interior* by $\text{int}(Q)$. The *conic hull*, the *positive hull* and the *affine hull* of Q are respectively given by

$$\text{cone}(Q) := \{ta : a \in Q, t \in \mathbb{R}, t \geq 0\},$$

$$\text{pos}(Q) := \left\{ \sum_{i=1}^k t_i a^i : a^i \in Q, t_i \in \mathbb{R}, t_i \geq 0, i = 1, \dots, k \text{ with } k \in \mathbb{N} \right\},$$

$$\text{aff}(Q) := \left\{ \sum_{i=1}^k t_i a^i : a^i \in Q, t_i \in \mathbb{R}, i = 1, \dots, k \text{ and } \sum_{i=1}^k t_i = 1 \text{ with } k \in \mathbb{N} \right\},$$

where \mathbb{N} denotes the set of natural numbers (Figs. 2.1, 2.2 and 2.3).

Fig. 2.1 Conic hull (with $Q = Q_1 \cup Q_2$)

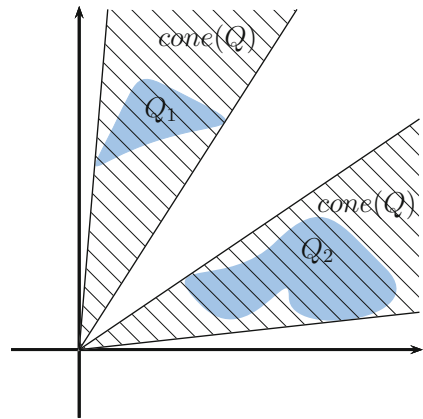


Fig. 2.2 Positive hull (with $Q = Q_1 \cup Q_2$)

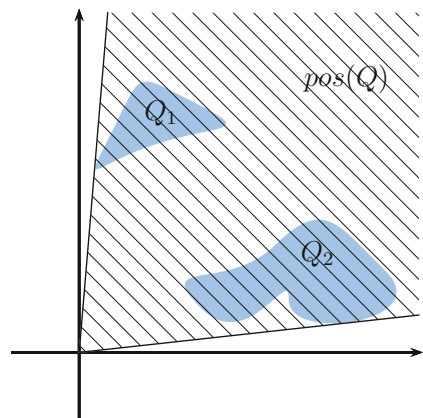
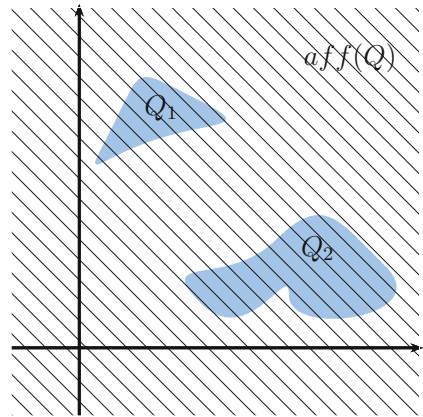


Fig. 2.3 Affine hull (with $Q = Q_1 \cup Q_2$)



Among the sets described above $\text{cone}(Q)$ and $\text{pos}(Q)$ are cones, that is, they are invariant under multiplication by positive numbers; $\text{pos}(Q)$ is also invariant under addition of its elements; and $\text{aff}(Q)$ is an affine subspace of \mathbb{R}^n . For two vectors x and y of \mathbb{R}^n , inequalities $x > y$ and $x \geq y$ mean respectively $x_i > y_i$ and $x_i \geq y_i$ for all $i = 1, \dots, n$. When $x \geq y$ and $x \neq y$, we write $x \geq y$. So a vector x is positive, that is $x \geq 0$, if its components are non-negative; and it is strictly positive if its components are all strictly positive. The set of all positive vectors of \mathbb{R}^n is the positive orthant \mathbb{R}_+^n . Sometimes row vectors are also considered. They are transposes of column vectors. Operations on row vectors are performed in the same manner as on column vectors. Thus, for two row n -vectors c and d , their inner product is expressed by

$$\langle c, d \rangle = \langle c^T, d^T \rangle = \sum_{i=1}^n c_i d_i,$$

where the upper index T denotes the transpose. On the other hand, if c is a row vector and x is a column vector, then the product cx is understood as a matrix product which is equal to the inner product $\langle c^T, x \rangle$.

Convex sets

We call a subset Q of \mathbb{R}^n *convex* if the segment joining any two points of Q lies entirely in Q , which means that for every $x, y \in Q$ and for every real number $\lambda \in [0, 1]$, one has $\lambda x + (1 - \lambda)y \in Q$ (Figs. 2.4, 2.5). It follows directly from the definition that the intersection of convex sets, the Cartesian product of convex sets, the image and inverse image of a convex set under a linear transformation, the interior and the closure of a convex set are convex. In particular, the sum $Q_1 + Q_2 := \{x + y : x \in Q_1, y \in Q_2\}$ of two convex sets Q_1 and Q_2 is convex; the conic hull of a convex set is convex. The positive hull and the affine hull of any set are convex.

The *convex hull* of Q , denoted $\text{co}(Q)$ (Fig. 2.6), consists of all convex combinations of elements of Q , that is,

Fig. 2.4 Convex set

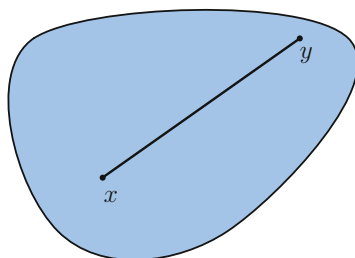
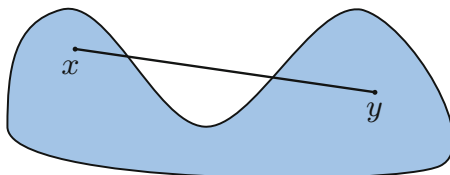


Fig. 2.5 Nonconvex set



$$\text{co}(Q) := \left\{ \sum_{i=1}^k \lambda_i x^i : x^i \in Q, \lambda_i \geq 0, i = 1, \dots, k \text{ and } \sum_{i=1}^k \lambda_i = 1 \text{ with } k \in \mathbb{N} \right\}.$$

It is the intersection of all convex sets containing Q . The closure of the convex hull of Q will be denoted by $\overline{\text{co}}(Q)$, which is exactly the intersection of all closed convex sets containing Q . The positive hull of a set is the conic hull of its convex hull. A convex combination $\sum_{i=1}^k \lambda_i x^i$ is strict if all coefficients λ_i are strictly positive.

Given a nonempty convex subset Q of \mathbb{R}^n , the *relative interior* of Q , denoted $\text{ri}(Q)$, is its interior relative to its affine hull, that is,

$$\text{ri}(Q) := \{x \in Q : (x + \varepsilon B_n) \cap \text{aff}(Q) \subseteq Q \text{ for some } \varepsilon > 0\}.$$

Equivalently, a point x in Q is a relative interior point if and only if for any point y in Q there is a positive number δ such that the segment joining the points $x - \delta(x - y)$ and $x + \delta(x - y)$ entirely lies in Q . As a consequence, any strict convex combination of a finite collection $\{x^1, \dots, x^k\}$ belongs to the relative interior of its convex hull (see also Lemma 6.4.8). It is important to note also that every nonempty convex set in \mathbb{R}^n has a nonempty relative interior. Moreover, if two convex sets Q_1 and Q_2 have at least one relative interior point in common, then $\text{ri}(Q_1 \cap Q_2) = \text{ri}(Q_1) \cap \text{ri}(Q_2)$.

Fig. 2.6 Convex hull of Q

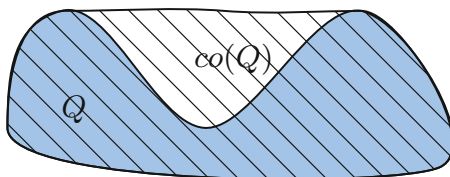
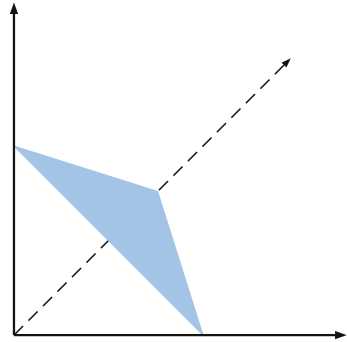


Fig. 2.7 The standard simplex in \mathbb{R}^3



Example 2.1.1 (Standard simplex) Let e^i be the i th coordinate unit vector of \mathbb{R}^n , that is its components are all zero except for the i th component equal to one. Let Δ denote the convex hull of e^1, \dots, e^n . Then a vector x with components x_1, \dots, x_n is an element of Δ if and only if $x_i \geq 0, i = 1, \dots, n$ and $\sum_{i=1}^n x_i = 1$. This set has no interior point. However, its relative interior consists of x with $x_i > 0, i = 1, \dots, n$ and $\sum_{i=1}^n x_i = 1$. The set Δ is called the *standard simplex* of \mathbb{R}^n (Fig. 2.7).

Caratheodory's theorem

It turns out that the convex hull of a set Q in the space \mathbb{R}^n can be obtained by convex combinations of at most $n + 1$ elements of Q . First we see this for positive hull.

Theorem 2.1.2 *Let $\{a^1, \dots, a^k\}$ be a collection of vectors in \mathbb{R}^n . Then for every nonzero vector x from the positive hull $\text{pos}\{a^1, \dots, a^k\}$ there exists an index set $I \subseteq \{1, \dots, k\}$ such that*

- (i) *the vectors $a^i, i \in I$ are linearly independent;*
- (ii) *x belongs to the positive hull $\text{pos}\{a^i, i \in I\}$.*

Proof Since the collection $\{a^1, \dots, a^k\}$ is finite, we may choose an index set I of minimum cardinality such that $x \in \text{pos}\{a^i, i \in I\}$. It is evident that there are strictly positive numbers $t_i, i \in I$ such that $x = \sum_{i \in I} t_i a^i$. We prove that (i) holds for this I . Indeed, if not, one can find an index $j \in I$ and real numbers s_i such that

$$a^j - \sum_{i \in I \setminus \{j\}} s_i a^i = 0.$$

Set

$$\varepsilon = \min \left\{ t_j \text{ and } -\frac{t_i}{s_i} : i \in I \text{ with } s_i < 0 \right\}$$

and express

$$\begin{aligned}
x &= \sum_{i \in I} t_i a^i - \varepsilon \left(a^j - \sum_{i \in I \setminus \{j\}} s_i a^i \right) \\
&= (t_j - \varepsilon) a^j + \sum_{i \in I \setminus \{j\}} (t_i + \varepsilon s_i) a^i.
\end{aligned}$$

It is clear that in the latter sum those coefficients corresponding to the indices that realize the minimum in the definition of ε are equal to zero. By this, x lies in the positive hull of less than $|I|$ vectors of the collection. This contradiction completes the proof. \square

A collection of vectors $\{a^1, \dots, a^k\}$ in \mathbb{R}^n is said to be *affinely independent* if the dimension of the subspace $\text{aff}\{a^1, \dots, a^k\}$ is equal to $k - 1$. By convention a set consisting of a solitary vector is affinely independent. The next result is a version of Caratheodory's theorem and well-known in convex analysis.

Corollary 2.1.3 *Let $\{a^1, \dots, a^k\}$ be a collection of vectors in \mathbb{R}^n . Then for every $x \in \text{co}\{a^1, \dots, a^k\}$ there exists an index set $I \subseteq \{1, \dots, k\}$ such that*

- (i) *the vectors $a^i, i \in I$ are affinely independent*
- (ii) *x belongs to the convex hull of $a^i, i \in I$.*

Proof We consider the collection of vectors $v^i = (a^i, 1), i = 1, \dots, k$ in the space $\mathbb{R}^n \times \mathbb{R}$. It is easy to verify that x belongs to the convex hull $\text{co}\{a^1, \dots, a^k\}$ if and only if the vector $(x, 1)$ belongs to the positive hull $\text{pos}\{v^1, \dots, v^k\}$. Applying Theorem 2.1.2 to the latter positive hull we deduce the existence of an index set $I \subseteq \{1, \dots, k\}$ such that the vector $(x, 1)$ belongs to the positive hull $\text{pos}\{v^i, i \in I\}$ and the collection $\{v^i, i \in I\}$ is linearly independent. Then x belongs to the convex hull $\text{co}\{a^i, i \in I\}$ and the collection $\{a^i, i \in I\}$ is affinely independent. \square

Linear operators and matrices

A mapping $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^k$ is called a linear operator between \mathbb{R}^n and \mathbb{R}^k if

- (i) $\phi(x + y) = \phi(x) + \phi(y)$,
- (ii) $\phi(tx) = t\phi(x)$

for every $x, y \in \mathbb{R}^n$ and $t \in \mathbb{R}$. The kernel and the image of ϕ are the sets

$$\begin{aligned}
\text{Ker}\phi &= \{x \in \mathbb{R}^n : \phi(x) = 0\}, \\
\text{Im}\phi &= \{y \in \mathbb{R}^k : y = \phi(x) \text{ for some } x \in \mathbb{R}^n\}.
\end{aligned}$$

These sets are linear subspaces of \mathbb{R}^n and \mathbb{R}^k respectively.

We denote the $k \times n$ -matrix whose columns are c_1, \dots, c_n by C , where c_i is the vector image of the i th coordinate unit vector e^i by ϕ . Then for every vector x of \mathbb{R}^n one has

$$\phi(x) = Cx.$$

The mapping $x \mapsto Cx$ is clearly a linear operator from \mathbb{R}^n to \mathbb{R}^k . This explains why one can identify a linear operator with a matrix. The space of $k \times n$ matrices is denoted by $L(\mathbb{R}^n, \mathbb{R}^k)$. The *transpose* of a matrix C is denoted by C^T . The norm and the inner product in the space of matrices are given by

$$\|C\| = \left(\sum_{i=1, \dots, n} \sum_{j=1, \dots, n} |c_{ij}|^2 \right)^{1/2},$$

$$\langle C, B \rangle = \sum_{i=1, \dots, n} \sum_{j=1, \dots, n} c_{ij} b_{ij}.$$

The norm $\|C\|$ is called also the *Frobenius norm*.

The inner product $\langle C, B \rangle$ is nothing but the trace of the matrix CB^T . Sometimes the space $L(\mathbb{R}^n, \mathbb{R}^k)$ is identified with the $n \times k$ -dimensional Euclidean space $\mathbb{R}^{n \times k}$.

Linear functionals

A particular case of linear operators is when the value space is one-dimensional. This is the space of linear functionals on \mathbb{R}^n and often identified with the space \mathbb{R}^n itself. Thus, each linear functional ϕ is given by a vector d_ϕ by the formula

$$\phi(x) = \langle d_\phi, x \rangle.$$

When $d_\phi \neq 0$, the kernel of ϕ is called a *hyperplane*; the vector d_ϕ is a *normal vector* to this hyperplane. Geometrically, d_ϕ is orthogonal to the hyperplane $\text{Ker}\phi$. The sets

$$\{x \in \mathbb{R}^n : \langle d_\phi, x \rangle \geq 0\},$$

$$\{x \in \mathbb{R}^n : \langle d_\phi, x \rangle \leq 0\}$$

are closed halfspaces and the sets

$$\{x \in \mathbb{R}^n : \langle d_\phi, x \rangle > 0\},$$

$$\{x \in \mathbb{R}^n : \langle d_\phi, x \rangle < 0\}$$

are open halfspaces bounded by the hyperplane $\text{Ker}\phi$. Given a real number α and a nonzero vector d of \mathbb{R}^n , one also understands a hyperplane of type

$$H(d, \alpha) = \{x \in \mathbb{R}^n : \langle d, x \rangle = \alpha\}.$$

The sets

$$H_+(d, \alpha) = \{x \in \mathbb{R}^n : \langle d, x \rangle \geq \alpha\},$$

$$H_-(d, \alpha) = \{x \in \mathbb{R}^n : \langle d, x \rangle \leq \alpha\}$$

are positive and negative halfspaces and the sets

$$\begin{aligned}\text{int}(H_+(d, \alpha)) &= \{x \in \mathbb{R}^n : \langle d, x \rangle > \alpha\}, \\ \text{int}(H_-(d, \alpha)) &= \{x \in \mathbb{R}^n : \langle d, x \rangle < \alpha\}\end{aligned}$$

are positive and negative open halfspaces.

Theorem 2.1.4 *Let Q be a nonempty convex set in \mathbb{R}^n and let $\langle d, \cdot \rangle$ be a positive functional on Q , that is $\langle d, x \rangle \geq 0$ for every $x \in Q$. If $\langle d, x \rangle = 0$ for some relative interior point x of Q , then $\langle d, \cdot \rangle$ is zero on Q .*

Proof Let y be any point in Q . Since x is a relative interior point, there exists a positive number δ such that $x + t(y - x) \in Q$ for $|t| \leq \delta$. Applying $\langle d, \cdot \rangle$ to this point we obtain

$$\langle d, x + t(y - x) \rangle = t \langle d, y \rangle \geq 0$$

for all $t \in [-\delta, \delta]$. This implies that $\langle d, y \rangle = 0$ as requested. \square

2.2 System of Linear Inequalities

We shall mainly deal with two kinds of systems of linear equations and inequalities. The first system consists of k inequalities

$$\langle a^i, x \rangle \leq b_i, \quad i = 1, \dots, k, \quad (2.1)$$

where a^1, \dots, a^k are n -dimensional column vectors and b_1, \dots, b_k are real numbers; and the second system consists of k equations which involves positive vectors only

$$\begin{aligned}\langle a^i, x \rangle &= b_i, \quad i = 1, \dots, k \\ x &\geq 0.\end{aligned} \quad (2.2)$$

Denoting by A the $k \times n$ -matrix whose rows are the transposes of a^1, \dots, a^k and by b the column k -vector of components b_1, \dots, b_k , we can write the systems (2.1) and (2.2) in matrix form

$$Ax \leq b \quad (2.3)$$

and

$$\begin{aligned}Ax &= b \\ x &\geq 0.\end{aligned} \quad (2.4)$$

Notice that any system of linear equations and inequalities can be converted to the two matrix forms described above. To this end it suffices to perform three operations:

- (a) Express each variable x_i as difference of two non-negative variables $x_i = x_i^+ - x_i^-$ where

$$\begin{aligned}x_i^+ &= \max\{x_i; 0\}, \\x_i^- &= \max\{-x_i; 0\}.\end{aligned}$$

- (b) Introduce a non-negative *slack variable* y_i in order to obtain equivalence between inequality $\langle a^i, x \rangle \leq b_i$ and equality $\langle a^i, x \rangle + y_i = b_i$. Similarly, with a non-negative *surplus variable* z_i one may express inequality $\langle a^i, x \rangle \geq b_i$ as equality $\langle a^i, x \rangle - z_i = b_i$.
- (c) Express equality $\langle a^i, x \rangle = b_i$ by two inequalities $\langle a^i, x \rangle \leq b_i$ and $\langle a^i, x \rangle \geq b_i$.

Example 2.2.1 Consider the following system

$$\begin{aligned}x_1 + 2x_2 &= 1, \\-x_1 - x_2 &\geq 0.\end{aligned}$$

It is written in form (2.3) as

$$\begin{pmatrix} 1 & 2 \\ -1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

and in form (2.4) with a surplus variable y as

$$\begin{pmatrix} 1 & -1 & 2 & -2 & 0 \\ -1 & 1 & -1 & 1 & -1 \end{pmatrix} (x_1^+, x_1^-, x_2^+, x_2^-, y)^T = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\(x_1^+, x_1^-, x_2^+, x_2^-, y)^T \geq 0.$$

Redundant equation

Given a system (2.4) we say it is *redundant* if at least one of the equations (called *redundant equation*) can be expressed as a linear combination of the others. In other words, it is redundant if there is a nonzero k -dimensional vector λ such that

$$\begin{aligned}A^T \lambda &= 0, \\ \langle b, \lambda \rangle &= 0.\end{aligned}$$

Moreover, redundant equations can be dropped from the system without changing its solution set. Similarly, an inequation of (2.1) is called redundant if its removal from the system does not change the solution set.

Proposition 2.2.2 *Assume that $k \leq n$ and that the system (2.4) is consistent. Then it is not redundant if and only if the matrix A has full rank.*

Proof If one of equations, say $\langle a^1, x \rangle = b_1$, is redundant, then a^1 is a linear combination of a^2, \dots, a^k . Hence the rank of A is not maximal, it is less than k . Conversely, when the rank of A is maximal (equal to k), no row of A is a linear combination of the others. Hence no equation of the system can be expressed as a linear combination of the others. \square

Farkas' theorem

One of the theorems of the alternative that are pillars of the theory of linear and nonlinear programming is Farkas' theorem or Farkas' lemma. There are a variety of ways to prove it, the one we present here is elementary.

Theorem 2.2.3 (Farkas' theorem) *Exactly one of the following systems has a solution:*

- (i) $Ax = b$ and $x \geq 0$;
- (ii) $A^T y \geq 0$ and $\langle b, y \rangle < 0$.

Proof If the first system has a solution x , then for every y with $A^T y \geq 0$ one has

$$\langle b, y \rangle = \langle Ax, y \rangle = \langle x, A^T y \rangle \geq 0,$$

which shows that the second system has no solution.

Now suppose the first system has no solution. Then either the system

$$Ax = b$$

has no solution, or it does have a solution, but every solution of it is not positive. In the first case, choose m linearly independent columns of A , say a_1, \dots, a_m , where m is the rank of A . Then the vectors a_1, \dots, a_m, b are linearly independent too (because b does not lie in the space spanned by a_1, \dots, a_m). Consequently, the system

$$\begin{aligned} \langle a_i, y \rangle &= 0, \quad i = 1, \dots, m, \\ \langle b, y \rangle &= -1 \end{aligned}$$

admits a solution. This implies that the system (ii) has solutions too. It remains to prove the solvability of (ii) when $Ax = b$ has solutions and they are all non-positive. We do it by induction on the dimension of x . Assume $n = 1$. If the system $a_{i1}x_1 = b_i, i = 1, \dots, k$ has a negative solution x_1 , then $y = -(b_1, \dots, b_k)^T$ is a solution of (ii) because $A^T y = -(a_{11}^2 + \dots + a_{k1}^2)x_1 > 0$ and $\langle b, y \rangle = -(b_1^2 + \dots + b_k^2) < 0$. Now assume $n > 1$ and that the result is true for the case of dimension $n - 1$. Given an n -vector x , denote by \bar{x} the $(n - 1)$ -vector consisting of the first $(n - 1)$ components of x . Let \bar{A} be the matrix composed of the first $(n - 1)$ columns of A . It is clear that the system

$$\bar{A}\bar{x} = b \text{ and } \bar{x} \geq 0$$

has no solution. By induction there is some \bar{y} such that

$$\begin{aligned} \bar{A}^T \bar{y} &\geq 0, \\ \langle b, \bar{y} \rangle &< 0. \end{aligned}$$

If $\langle a_n, \bar{y} \rangle \geq 0$, we are done. If $\langle a_n, \bar{y} \rangle < 0$, define new vectors

$$\begin{aligned} \hat{a}_i &= \langle a_i, \bar{y} \rangle a_n - \langle a_n, \bar{y} \rangle a_i, \quad i = 1, \dots, n-1, \\ \hat{b} &= \langle b, \bar{y} \rangle a_n - \langle a_n, \bar{y} \rangle b \end{aligned}$$

and consider a new system

$$\hat{a}_1 \xi_1 + \dots + \hat{a}_{n-1} \xi_{n-1} = \hat{b}. \quad (2.5)$$

We claim that this system of k equations has no positive solution. Indeed, if not, say ξ_1, \dots, ξ_{n-1} were non-negative solutions, then the vector x with

$$\begin{aligned} x_i &= \xi_i, \quad i = 1, \dots, n-1, \\ x_n &= -\frac{1}{\langle a_n, \bar{y} \rangle} \left(\langle a_1 \xi_1 + \dots + a_{n-1} \xi_{n-1}, \bar{y} \rangle - \langle b, \bar{y} \rangle \right) \end{aligned}$$

should be a positive solution of (i) because $-\langle b, \bar{y} \rangle > 0$ and $\langle \bar{A}x, \bar{y} \rangle = \langle \xi, \bar{A}^T \bar{y} \rangle \geq 0$ for $\xi = (\xi_1, \dots, \xi_{n-1})^T \geq 0$, implying $x_n \geq 0$. Applying the induction hypothesis to (2.5) we deduce the existence of a k -vector \hat{y} with

$$\begin{aligned} \langle \hat{a}_i, \hat{y} \rangle &\geq 0, \quad i = 1, \dots, n-1, \\ \langle \hat{b}, \hat{y} \rangle &< 0. \end{aligned}$$

Then the vector $y = \langle a_n, \hat{y} \rangle \bar{y} - \langle a_n, \bar{y} \rangle \hat{y}$ satisfies the system (ii). The proof is complete. \square

A number of consequences can be derived from Farkas' theorem which are useful in the study of linear systems and linear programming problems.

Corollary 2.2.4 *Exactly one of the following systems has a solution:*

- (i) $Ax = 0, \langle c, x \rangle = 1$ and $x \geq 0$;
- (ii) $A^T y \geq c$.

Proof If (ii) has a solution y , then for a positive vector x with $Ax = 0$ one has

$$0 = \langle y, Ax \rangle = \langle A^T y, x \rangle \geq \langle c, x \rangle.$$

So (i) is not solvable. Conversely, if (i) has no solution, then applying Farkas' theorem to the inconsistent system

$$\begin{pmatrix} A \\ c^T \end{pmatrix} x = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ and } x \geq 0$$

yields the existence of a vector y and of a real number t such that

$$\begin{pmatrix} A^T c \end{pmatrix} \begin{pmatrix} y \\ t \end{pmatrix} \geq 0 \text{ and } \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} y \\ t \end{pmatrix} \right\rangle < 0.$$

Hence $t < 0$ and $-y/t$ is a solution of (ii). □

Corollary 2.2.5 *Exactly one of the following systems has a solution:*

- (i) $Ax \geq 0$ and $x \geq 0$;
- (ii) $A^T y \leq 0$ and $y > 0$.

Proof By introducing a surplus variable $z \in \mathbb{R}^k$ we convert (i) to an equivalent system

$$\begin{aligned} Ax - Iz &= 0, \\ \begin{pmatrix} x \\ z \end{pmatrix} &\geq 0, \\ \left\langle c, \begin{pmatrix} x \\ z \end{pmatrix} \right\rangle &= 1, \end{aligned}$$

where c is an $(n+k)$ -vector whose n first components are all zero and the remaining components are one. According to Corollary 2.2.4 it has no solution if and only if the following system has a solution

$$\begin{pmatrix} A^T \\ -I \end{pmatrix} y \geq c.$$

It is clear that the latter system is equivalent to (ii). □

The next corollary is known as Motzkin's theorem of the alternative.

Corollary 2.2.6 (Motzkin's theorem) *Let A and B be two matrices having the same number of columns. Exactly one of the following systems has a solution:*

- (i) $Ax > 0$ and $Bx \geq 0$;
- (ii) $A^T y + B^T z = 0$, $y \geq 0$ and $z \geq 0$.

Proof The system (ii) is evidently equivalent to the following one

$$\begin{pmatrix} A^T & B^T \\ e^T & 0 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} y \\ z \end{pmatrix} \geq 0.$$

By Farkas' theorem it is compatible (has a solution) if and only if the following system is incompatible:

$$\begin{pmatrix} A & e \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left\langle \begin{pmatrix} x \\ t \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle < 0.$$

The latter system is evidently equivalent to the system of (i). □

Some classical theorems of alternatives are immediate from Corollary 2.2.6.

- *Gordan's theorem* (B is the zero matrix):
Exactly one of the following systems has a solution
 - (1) $Ax > 0$;
 - (2) $A^T y = 0$ and $y \geq 0$.
- *Ville's theorem* (B is the identity matrix):
Exactly one of the following systems has a solution
 - (3) $Ax > 0$ and $x \geq 0$;
 - (4) $A^T y \leq 0$ and $y \geq 0$.
- *Stiemke's theorem* (A is the identity matrix and B is replaced by $\begin{pmatrix} B \\ -B \end{pmatrix}$):
Exactly one of the following systems has a solution
 - (5) $Bx = 0$ and $x > 0$;
 - (6) $B^T y \geq 0$.

2.3 Convex Polyhedra

A set that can be expressed as the intersection of a finite number of closed half-spaces is called a *convex polyhedron*. A convex bounded polyhedron is called a *polytope*. According to the definition of closed half-spaces, a convex polyhedron is the solution set to a finite system of inequalities

$$\langle a^i, x \rangle \leq b_i, \quad i = 1, \dots, k \tag{2.6}$$

where a^1, \dots, a^k are n -dimensional column vectors and b_1, \dots, b_k are real numbers. When $b_i = 0, i = 1, \dots, k$, the solution set to (2.6) is a cone and called a *convex polyhedral cone*. We assume throughout this section that the system is not redundant and solvable.

Supporting hyperplanes and faces

Let P be a convex polyhedron and let

$$H = \{x \in \mathbb{R}^n : \langle v, x \rangle = \alpha\}$$

be a hyperplane with v nonzero. We say H is a *supporting hyperplane* of P at a point $x \in P$ if the intersection of H with P contains x and P is contained in one of the closed half-spaces bounded by H (Fig. 2.8). In this case, the nonempty set $H \cap P$ is called a *face* of P . Thus, a nonempty subset F of P is a face if there is a nonzero vector $v \in \mathbb{R}^n$ such that

$$\langle v, y \rangle \leq \langle v, x \rangle \quad \text{for all } x \in F, y \in P.$$

When a face is zero-dimensional, it is called a *vertex*. A nonempty polyhedron may have no vertex. By convention P is a face of itself; other faces are called *proper faces*. One-dimensional faces are called *edges*. Two vertices are said to be adjacent if they are end-points of an edge.

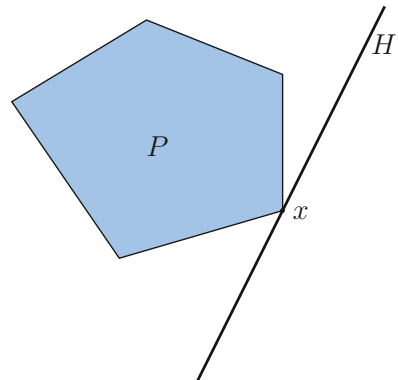
Example 2.3.1 Consider a system of three inequalities in \mathbb{R}^2 :

$$x_1 + x_2 \leq 1 \tag{2.7}$$

$$-x_1 - x_2 \leq 0 \tag{2.8}$$

$$-x_1 \leq 0. \tag{2.9}$$

Fig. 2.8 Supporting hyperplane



The polyhedron defined by (2.7) and (2.8) has no vertex. It has two one-dimensional faces determined respectively by $x_1 + x_2 = 1$ and $x_1 + x_2 = 0$, and one two-dimensional face, the polyhedron itself. The polyhedron defined by (2.7)–(2.9) has two vertices (zero-dimensional faces) determined respectively by

$$\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases} \text{ and } \begin{cases} x_1 = 0 \\ x_2 = 1 \end{cases},$$

three one-dimensional faces given by

$$\begin{cases} x_1 + x_2 \leq 1 \\ -x_1 - x_2 \leq 0 \\ x_1 = 0 \end{cases}, \begin{cases} x_1 + x_2 = 1 \\ -x_1 \leq 0 \end{cases} \text{ and } \begin{cases} -x_1 - x_2 = 0 \\ -x_1 \leq 0 \end{cases},$$

and one two-dimensional face, the polyhedron itself.

Proposition 2.3.2 *Let P be a convex polyhedron. The following properties hold.*

- (i) *The intersection of any two faces is a face if it is nonempty.*
- (ii) *Two different faces have no relative interior point in common.*

Proof We prove (i) first. Assume F_1 and F_2 are two faces with nonempty intersection. If they coincide, there is nothing to prove. If not, let H_1 and H_2 be two supporting hyperplanes that generate these faces, say

$$\begin{aligned} H_1 &= \{x \in \mathbb{R}^n : \langle v^1, x \rangle = \alpha_1\}, \\ H_2 &= \{x \in \mathbb{R}^n : \langle v^2, x \rangle = \alpha_2\}. \end{aligned}$$

Since these hyperplanes contain the intersection of distinct faces F_1 and F_2 , the vector $v = v^1 + v^2$ is not zero. Consider the hyperplane

$$H = \{x \in \mathbb{R}^n : \langle v, x \rangle = \alpha_1 + \alpha_2\}.$$

It is a supporting hyperplane of P because it evidently contains the intersection of the faces F_1 and F_2 , and for every point x in P , one has

$$\langle v, x \rangle = \langle v^1, x \rangle + \langle v^2, x \rangle \leq \alpha_1 + \alpha_2. \quad (2.10)$$

It remains to show that the intersection of H and P coincides with the intersection $F_1 \cap F_2$. The inclusion

$$F_1 \cap F_2 \subseteq H \cap P$$

being clear, we show the converse. Let x be in $H \cap P$. Then (2.10) becomes equality for this x . But $\langle v^1, x \rangle \leq \alpha_1$ and $\langle v^2, x \rangle \leq \alpha_2$, so that equality of (2.10) is possible only when the two latter inequalities are equalities. This proves that x belongs to both F_1 and F_2 .

For the second assertion notice that if F_1 and F_2 have a relative interior point in common, then in view of Theorem 2.1.4, the functional $\langle v^1, \cdot \rangle$ is constant on F_2 . It follows that $F_2 \subseteq H_1 \cap P \subseteq F_1$. Similarly, one has $F_1 \subseteq F_2$, and hence equality holds. \square

Let x be a solution of the system (2.6). Define the *active index set* at x to be the set

$$I(x) = \left\{ i \in \{1, \dots, k\} : \langle a^i, x \rangle = b_i \right\}.$$

The remaining indices are called *inactive indices*.

Theorem 2.3.3 *Assume that P is a convex polyhedron given by (2.6). A nonempty proper convex subset F of P is a face if and only if there is a nonempty maximal index set $I \subseteq \{1, \dots, k\}$ such that F is the solution set to the system*

$$\langle a^i, x \rangle = b_i, \quad i \in I \tag{2.11}$$

$$\langle a^j, x \rangle \leq b_j, \quad j \in \{1, \dots, k\} \setminus I, \tag{2.12}$$

in which case the dimension of F is equal to $n - \text{rank}\{a^i : i \in I\}$.

Proof Denote the solution set to the system (2.11, 2.12) by F' that we suppose nonempty. To prove that it is a face, we set

$$v = \sum_{i \in I} a^i \quad \text{and} \quad \alpha = \sum_{i \in I} b_i.$$

Notice that v is nonzero because F' is not empty and the system (2.6) is not redundant. It is clear that the negative half-space $H_-(v, \alpha)$ contains P . Moreover, if x is a solution to the system, then, of course, x belongs to P and to H at the same time, which implies $F' \subseteq H \cap P$. Conversely, any point x of the latter intersection satisfies

$$\begin{aligned} \langle a^i, x \rangle &\leq b_i, \quad i = 1, \dots, k, \\ \sum_{i \in I} \langle a^i, x \rangle &= \sum_{i \in I} b_i. \end{aligned}$$

The latter equality is possible only when those inequalities with indices from I are equalities. In other words, x belongs to F' .

Now, let F be a proper face of P . Pick a relative interior point \bar{x} of F and consider the system (2.11, 2.12) with $I = I(\bar{x})$ the active index set of \bar{x} . Being a proper face of P , F has no interior point, and so the set I is nonempty. As before, F' is the solution set to that system. By the first part, it is a face. We wish to show that it coincides with F . For this, in view of Proposition 2.3.2 it suffices to show that \bar{x} is also a relative interior point of F' . Let x be another point in F' . We have to prove that there is a positive number δ such that the segment $[\bar{x}, \bar{x} + \delta(x - \bar{x})]$ lies in F' . Indeed, note

that for indices j outside the set I , inequalities $\langle a^j, \bar{x} \rangle \leq b_j$ are strict. Therefore, there is $\delta > 0$ such that

$$\langle a^j, \bar{x} \rangle + \delta \langle a^j, x - \bar{x} \rangle \leq b_j$$

for all $j \notin I$. Moreover, being a linear combination of \bar{x} and x , the endpoint $\bar{x} + \delta(x - \bar{x})$ satisfies the equalities (2.11) too. Consequently, this point belongs to F' , and hence so does the whole segment. Since F and F' are two faces with a relative interior point in common, they must be the same. \square

In general, for a given face F of P , there may exist several index sets I for which F is the solution set to the system (2.11, 2.12). We shall, however, understand that no inequality can be equality without changing the solution set when saying that the system (2.11, 2.12) determines the face F . So, if two inequalities combined yields equality, their indices will be counted in I .

Corollary 2.3.4 *If an m -dimensional convex polyhedron has a vertex, then it has faces of any dimension less than m .*

Proof The corollary is evident for a zero-dimensional polyhedron. Suppose P is a polyhedron of dimension $m > 0$. By Theorem 2.3.3 without loss of generality we may assume that P is given by the system (2.11, 2.12) with $|I| = n - m$ and that the family $\{a^i, i \in I\}$ is linearly independent. Since P has a vertex, there is some $i_0 \in \{1, \dots, k\} \setminus I$ such that the vectors $a^i, i \in I \cup \{i_0\}$ are linearly independent. Then the system

$$\begin{aligned} \langle a^i, x \rangle &= b_i, \quad i \in I \cup \{i_0\}, \\ \langle a^j, x \rangle &\leq b_j, \quad j \in \{1, \dots, k\} \setminus (I \cup \{i_0\}) \end{aligned}$$

generates an $(m - 1)$ -dimensional face of P . Notice that this system has a solution because P is generated by the non-redundant system (2.11, 2.12). Continuing the above process we are able to construct a face of any dimension less than m . \square

Corollary 2.3.5 *Let F be a face of the polyhedron P determined by the system (2.11, 2.12). Then for every $x \in F$ one has*

$$I(x) \supseteq I.$$

Equality holds if and only if x is a relative interior point of F .

Proof The inclusion $I \subseteq I(x)$ is evident because $x \in F$. For the second part, we first assume $I(x) = I$, that is

$$\begin{aligned} \langle a^i, x \rangle &= b_i, \quad i \in I, \\ \langle a^j, x \rangle &< b_j, \quad j \in \{1, \dots, k\} \setminus I. \end{aligned}$$

It is clear that if $y \in \text{aff}(F)$, then $\langle a^i, y \rangle = b_i, i \in I$, and if $y \in x + \varepsilon B_k$ with $\varepsilon > 0$ sufficiently small, then $\langle a^j, y \rangle < b_j, j \in \{1, \dots, k\} \setminus I$. We deduce that

$\text{aff}(F) \cap (x + \varepsilon B_k) \subseteq F$, which shows that x is a relative interior point of F . Conversely, let x be a relative interior point of F . Using the argument in the proof of Theorem 2.3.3 we know that F is also a solution set to the system

$$\begin{aligned} \langle a^i, y \rangle &= b_i, \quad i \in I(x), \\ \langle a^j, y \rangle &\leq b_j, \quad j \in \{1, \dots, k\} \setminus I(x). \end{aligned}$$

Since the system (2.11, 2.12) determines F , we have $I(x) \subseteq I$, and hence equality follows. \square

Corollary 2.3.6 *Let F be a face of the polyhedron P determined by the system (2.11, 2.12). Then a point $v \in F$ is a vertex of F if and only if it is a vertex of P .*

Proof It is clear that every vertex of P is a vertex of F if it belongs to F . To prove the converse, let us deduce a system of inequalities from (2.11, 2.12) by expressing equalities $\langle a^i, x \rangle = b_i$ as two inequalities $\langle a^i, x \rangle \leq b_i$ and $\langle -a^i, x \rangle \leq -b_i$. If v is a vertex of F , then the active constraints at v consists of the vectors $a^i, -a^i, i \in I$ and some $a^j, j \in J \subseteq \{1, \dots, k\} \setminus I$, so that the rank of the family $\{a^i, -a^i, a^j : i \in I, j \in J\}$ is equal to n . It follows that the family $\{a^i, a^j : i \in I, j \in J\}$ has rank equal to n too. In view of Theorem 2.3.3 the point v is a vertex of P . \square

Given a face F of a polyhedron, according to the preceding corollary the active index set $I(x)$ is constant for every relative interior point x of F . Therefore, we call it active index set of F and denote it by I_F .

A collection of subsets of a polyhedron is said to be a *partition* of it if the elements of the collection are disjoint and their union contains the entire polyhedron.

Corollary 2.3.7 *The collection of all relative interiors of faces of a polyhedron forms a partition of the polyhedron.*

Proof It is clear from Proposition 2.3.2(ii) that relative interiors of different faces are disjoint. Moreover, given a point x in P , consider the active index set $I(x)$. If it is empty, then the point belongs to the interior of P and we are done. If it is not empty, by Corollary 2.3.5, the face determined by system (2.11, 2.12) with $I = I(x)$ contains x in its relative interior. The proof is complete. \square

We now deduce a first result on representation of elements of polyhedra by vertices.

Corollary 2.3.8 *A convex polytope is the convex hull of its vertices.*

Proof The corollary is evident when the dimension of a polytope is less or equal to one in which case it is a point or a segment with two end-points. We make the induction hypothesis that the corollary is true when a polytope has dimension less than m with $1 < m < n$ and prove it for the case when P is a polytope determined by the system (2.6) and has dimension equal to m . Since P is a convex set, the convex hull of its vertices is included in P itself. Conversely, let x be a point of P . In view of

Corollary 2.3.7 it is a relative interior point of some face F of P . If F is a proper face of P , its dimension is less than m , and so we are done. It remains to treat the case where x is a relative interior point of P . Pick any point $y \neq x$ in P and consider the line passing through x and y . Since P is bounded, the intersection of this line with P is a segment, say with end-points c and d . Let F_c and F_d be faces of P that contain c and d in their relative interiors. As c and d are not relative interior points of P , the faces F_c and F_d are proper faces of P , and hence they have dimension strictly less than m . By induction c and d belong to the convex hulls of the vertices of F_c and F_d respectively. By Corollary 2.3.6 they belong to the convex hull of the vertices of P , and hence so does x because x belongs to the convex hull of c and d . \square

A similar result is true for polyhedral cones. It explains why one-dimensional faces of a polyhedral cone are called extreme rays.

Corollary 2.3.9 *A nontrivial polyhedral cone with vertex is the convex hull of its one-dimensional faces.*

Proof By definition a polyhedral cone P is defined by a homogeneous system

$$\langle a^i, x \rangle \leq 0, \quad i = 1, \dots, k. \quad (2.13)$$

Choose any nonzero point y in P and consider the hyperplane H given by

$$\langle a^1 + \dots + a^k, x - y \rangle = 0. \quad (2.14)$$

We claim that the vector $a^1 + \dots + a^k$ is nonzero. Indeed, if not, the inequalities (2.13) would become equalities for all $x \in P$, and P would be either a trivial cone, or a cone without vertex. Moreover, $P \cap H$ is a bounded polyhedron, because otherwise one should find a nonzero vector u satisfying $\langle a^i, u \rangle = 0, i = 1, \dots, k$ and P could not have vertices. In view of Corollary 2.3.8, $P \cap H$ is the convex hull of its vertices. To complete the proof it remains to show that a vertex v of $P \cap H$ is the intersection of a one-dimensional face of P with H . Indeed, the polytope $P \cap H$ being determined by the system (2.13) and (2.14), there is a set $J \subset \{1, \dots, k\}$ with $|J| = n - 1$ such that the vectors $a^j, j \in J$ and $a^1 + \dots + a^k$ are linearly independent and v is given by system

$$\langle a^j, x \rangle = 0, \quad j \in J \quad (2.15)$$

$$\begin{aligned} \langle a^1 + \dots + a^k, x \rangle &= \langle a^1 + \dots + a^k, y \rangle, \\ \langle a^i, x \rangle &\leq 0, \quad i \in \{1, \dots, k\} \setminus J. \end{aligned} \quad (2.16)$$

It is clear that (2.15) and (2.16) determine a one-dimensional face of P whose intersection with H is v . \square

Separation of convex polyhedra

Given two convex polyhedra P and Q in \mathbb{R}^n , we say that a nonzero vector v separates them if

$$\langle v, x \rangle \geq \langle v, y \rangle \text{ for all vectors } x \in P, y \in Q$$

and strict inequality is true for some of them (Fig. 2.9). The following result can be considered as a version of Farkas' theorem or Gordan's theorem.

Theorem 2.3.10 *If P and Q are convex polyhedra without relative interior points in common, then there is a nonzero vector separating them.*

Proof We provide a proof for the case where both P and Q have interior points only. Without loss of generality we may assume that P is determined by the system (2.6) and Q is determined by the system

$$\langle d^j, x \rangle \leq c_j, \quad j = 1, \dots, m.$$

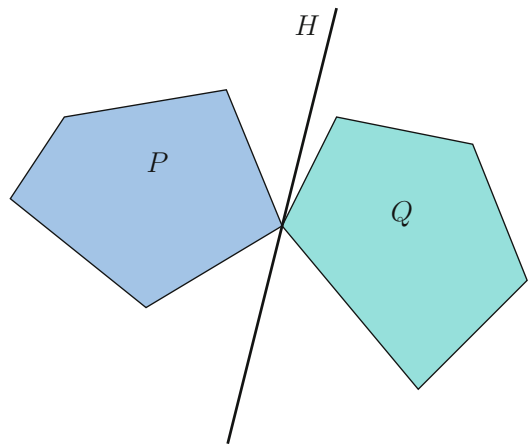
Thus, the following system:

$$\begin{aligned} \langle a^i, x \rangle &< b_i, \quad i = 1, \dots, k \\ \langle d^j, x \rangle &< c_j, \quad j = 1, \dots, m \end{aligned}$$

has no solution because the first k inequalities determine the interior of P and the last m inequalities determine the interior of Q . This system is equivalent to the following one:

$$\begin{pmatrix} -A & b \\ -D & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

Fig. 2.9 Separation



where A is the $k \times n$ -matrix whose rows are transposes of a^1, \dots, a^k , D is the $m \times n$ -matrix whose rows are transposes of d^1, \dots, d^k , b is the k -vector with the components b_1, \dots, b_k and c is the m -vector with the components c_1, \dots, c_m . According to Gordan's theorem, there exist positive vectors $\lambda \in \mathbb{R}^k$ and $\mu \in \mathbb{R}^m$ and a real number $s \geq 0$, not all zero, such that

$$\begin{aligned} A^T \lambda + D^T \mu &= 0, \\ \langle b, \lambda \rangle + \langle c, \mu \rangle + s &= 0. \end{aligned}$$

It follows from the latter equality that (λ, μ) is nonzero. We may assume without loss of generality that $\lambda \neq 0$. We claim that $A^T \lambda \neq 0$. Indeed, if not, choose x an interior point of P and y an interior point of Q . Then $D^T \mu = 0$ and hence

$$\langle b, \lambda \rangle > \langle Ax, \lambda \rangle = \langle x, A^T \lambda \rangle = 0$$

and

$$\langle c, \mu \rangle \geq \langle Dy, \mu \rangle = \langle y, D^T \mu \rangle = 0,$$

which is in contradiction with the aforesaid equality. Defining v to be the nonzero vector $-A^T \lambda$, we deduce for every $x \in P$ and $y \in Q$ that

$$\langle v, x \rangle = \langle -A^T \lambda, x \rangle = \langle \lambda, -Ax \rangle \geq \langle \lambda, -b \rangle \geq \langle \mu, c \rangle \geq \langle \mu, Dy \rangle = \langle v, y \rangle.$$

Of course inequality is strict when x and y are interior points. By this v separates P and Q as requested. \square

Asymptotic cones

Given a nonempty convex and closed subset C of \mathbb{R}^n , we say that a vector v is an *asymptotic* or a *recession direction* of C if

$$x + tx \in C \text{ for all } x \in C, t \geq 0.$$

The set of all asymptotic directions of C is denoted by C_∞ (Fig. 2.10). It is a convex cone. It can be seen that a closed convex set is bounded if and only if its *asymptotic cone* is trivial. The set $C_\infty \cap (-C_\infty)$ is a linear subspace and called the *lineality space* of C .

An equivalent definition of asymptotic directions is given next.

Theorem 2.3.11 *A vector v is an asymptotic direction of a convex and closed set C if and only if there exist a sequence of elements $x^s \in C$ and a sequence of positive numbers t_s converging to zero such that $v = \lim_{s \rightarrow \infty} t_s x^s$.*

Proof If $v \in C_\infty$ and $x \in C$, then $x^s = x + sv \in C$ for all $s \in \mathbb{N} \setminus \{0\}$. Setting $t_s = 1/s$ we obtain $v = \lim_{s \rightarrow \infty} t_s x^s$ with $\lim_{s \rightarrow \infty} t_s = 0$. Conversely, assume that

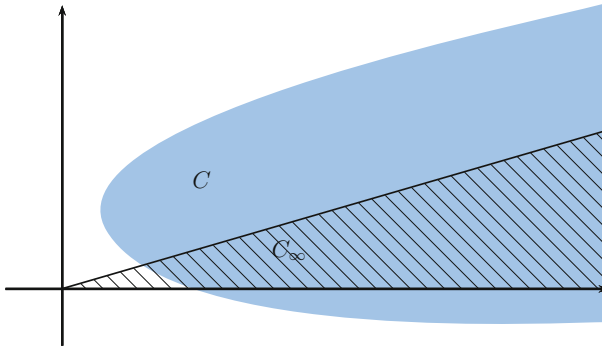


Fig. 2.10 Asymptotic cone

$v = \lim_{s \rightarrow \infty} t_s x^s$ for $x^s \in C$ and $t_s > 0$ converging to zero as s tends to ∞ . Let $x \in C$ and $t > 0$ be given. Then tt_s converges to zero as $s \rightarrow \infty$ and $0 \leqq tt_s \leqq 1$ for s sufficiently large. Hence,

$$\begin{aligned} x + tv &= \lim_{s \rightarrow \infty} (x + tt_s x^s) \\ &= \lim_{s \rightarrow \infty} \left((1 - tt_s)x + tt_s x^s + tt_s x \right) \\ &= \lim_{s \rightarrow \infty} \left((1 - tt_s)x + tt_s x^s \right). \end{aligned}$$

The set C being closed and convex, the points under the latter limit belong to the set C , and therefore their limit $x + tv$ belongs to C too. Since x and $t > 0$ were chosen arbitrarily we conclude that $v \in C_\infty$. □

Below is a formula to compute the asymptotic cone of a polyhedron.

Theorem 2.3.12 *The asymptotic cone of the polyhedron P determined by the system (2.6) is the solution set to system*

$$\langle a^i, v \rangle \leqq 0, \quad i = 1, \dots, k. \tag{2.17}$$

Proof Let v be an asymptotic direction of P . Then for every positive number t one has

$$\langle a^i, \bar{x} + tv \rangle \leqq b_i, \quad i = 1, \dots, k,$$

where \bar{x} is any point in P . By dividing both sides of the above inequalities by $t > 0$ and letting this t tend to ∞ we derive (2.17). For the converse, if v is a solution of (2.17), then for every point x in P one has

$$\langle a^i, x + tv \rangle = \langle a^i, x \rangle + t \langle a^i, v \rangle \leqq b_i, \quad i = 1, \dots, k$$

for all $t \geq 0$. Thus, the points $x + tv$ with $t \geq 0$, belong to P and v is an asymptotic direction. \square

Example 2.3.13 Consider a (nonempty) polyhedron in \mathbb{R}^3 defined by the system:

$$\begin{aligned} -x_1 - x_2 - x_3 &\leq -1, \\ x_3 &\leq 1, \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

The asymptotic cone is given by the system

$$\begin{aligned} -x_1 - x_2 - x_3 &\leq 0, \\ x_3 &\leq 0, \\ x_1, x_2, x_3 &\geq 0, \end{aligned}$$

in which the first inequality is redundant, and hence it is simply given by $x_1 \geq 0$, $x_2 \geq 0$ and $x_3 = 0$.

Using asymptotic directions we are also able to tell whether a convex polyhedron has a vertex or not. A cone is called *pointed* if it contains no straight line. When a cone C is not pointed, it contains a nontrivial linear subspace $C \cap (-C)$, called also the lineality space of C .

Corollary 2.3.14 *A convex polyhedron has vertices if and only if its asymptotic cone is pointed. Consequently, if a convex polyhedron has a vertex, then so does any of its faces.*

Proof It is easy to see that when a polyhedron has a vertex, it contains no straight line, and hence its asymptotic cone is pointed. We prove the converse by induction on the dimension of the polyhedron. The case where a polyhedron is of dimension less or equal to one is evident because a polyhedron with a pointed asymptotic cone is either a point or a segment or a ray, hence it has a vertex. Assume the induction hypothesis that the conclusion is true for all polyhedra of dimension less than m with $1 < m < n$. Let P be m -dimensional with a pointed asymptotic cone. If P has no proper face, then the inequalities (2.6) are strict, which implies that P is closed and open at the same time. This is possible only when P coincides with the space \mathbb{R}^n which contradicts the hypothesis that P_∞ is pointed. Now, let F be a proper face of P . Its asymptotic cone, being a subset of the asymptotic cone of P is pointed too. By induction, it has a vertex, which in view of Corollary 2.3.6 is also a vertex of P .

To prove the second part of the corollary it suffices to notice that if a face of P has no vertex, by the first part of the corollary, its asymptotic cone contains a straight line, hence so does the set P itself. \square

A second representation result for elements of a convex polyhedron is now formulated in a more general situation.

Corollary 2.3.15 *A convex polyhedron with vertex is the convex hull of its vertices and extreme directions.*

Proof We conduct the proof by induction on the dimension of the polyhedron. The corollary is evident when a polyhedron is zero or one-dimensional. We assume that it is true for all convex polyhedra of dimension less than m with $1 < m < n$ and prove it for an m -dimensional polyhedron P determined by system

$$\begin{aligned} \langle a^i, x \rangle &= b_i, \quad i \in I \\ \langle a^j, x \rangle &\leq b_j, \quad j \in \{1, \dots, k\} \setminus I, \end{aligned}$$

in which $|I| = n - m$ and the vectors $a^i, i \in I$ are linearly independent. Let y be an arbitrary element of P . If it belongs to a proper face of P , then by induction we can express it as a convex combination of vertices and extreme directions. If it is a relative interior point of P , then setting $a = a^1 + \dots + a^k$ that is nonzero because the asymptotic cone of P is pointed, P itself having a vertex, and considering the intersection of P with the hyperplane H determined by equality $\langle a, x \rangle = \langle a, y \rangle$ we obtain a bounded polyhedron $P \cap H$. By Corollary 2.3.8, the point y belongs to the convex hull of vertices of $P \cap H$. The vertices of $P \cap H$ belong to proper faces of P , by induction, they also belong to the convex hull of vertices and extreme directions of P , hence so does y . \square

When a polyhedron has a non-pointed asymptotic cone, it has no vertex. However, it is possible to express it as a sum of its asymptotic cone and a bounded polyhedron as well.

Corollary 2.3.16 *Every convex polyhedron is the sum of a bounded polyhedron and its asymptotic cone.*

Proof Denote the lineality space of the asymptotic cone of a polyhedron P by M . If M is trivial, we are done in view of Corollary 2.3.15 (the convex hull of all vertices of P serves as a bounded polyhedron). If M is not trivial, we decompose the space \mathbb{R}^n into the direct sum of M and its orthogonal M^\perp . Denote by P_\perp the projection of P on M^\perp . Then $P = P_\perp + M$. Indeed, let x be an element in P and let $x = x^1 + x^2$ with $x^1 \in M$ and $x^2 \in M^\perp$. Then $x^2 \in P_\perp$ and one has $x \in P_\perp + M$. Conversely, let $x^1 \in M$ and $x^2 \in P_\perp$. By definition there is some $y \in P$, say $y = y^1 + y^2$ with $y^1 \in M$ and $y^2 \in M^\perp$ such that $y^2 = x^2$. Since M is a part of the asymptotic cone of P , one deduces that

$$x = x^1 + x^2 = y^1 + (x^1 - y^1) + y^2 = y + (x^1 - y^1) \in y + M \subseteq P,$$

showing that x belongs to P . Further, we claim that the asymptotic cone of P_\perp is pointed. In fact, if not, say it contains a straight line d . Then the convexity of P implies that the space $M + d$ belongs to the asymptotic cone of P . This is a contradiction because d lies in M^\perp and M is already the biggest linear subspace contained in P . Let Q denote the convex hull of the set of all vertices of P_\perp which

is nonempty by Corollary 2.3.14. It follows from Corollary 2.3.18 below that the asymptotic cone of P is the sum of the asymptotic cone of P_{\perp} and M . We deduce $P = Q + (P_{\perp})_{\infty} + M = Q + P_{\infty}$ as requested. \square

The following calculus rule for asymptotic directions under linear transformations is useful.

Corollary 2.3.17 *Let P be the polyhedron determined by the system (2.6) and let L be a linear operator from \mathbb{R}^n to \mathbb{R}^m . Then*

$$L(P_{\infty}) = [L(P)]_{\infty}.$$

Proof The inclusion $L(P_{\infty}) \subseteq [L(P)]_{\infty}$ is true for any closed convex set. Indeed, if u is an asymptotic direction of P , then for every x in P and for every positive number t one has $x + tu \in P$. Consequently, $L(x) + tL(u)$ belongs to $L(P)$ for all $t \geq 0$. This means that $L(u)$ is an asymptotic direction of $L(P)$. For the converse inclusion, let v be a nonzero asymptotic direction of $L(P)$. By definition, for a fixed x of P , vectors $L(x) + tv$ belong to $L(P)$ for any $t \geq 0$. Thus, there are x^1, x^2, \dots in P such that $L(x^{\nu}) = L(x) + \nu v$, or equivalently $v = L(\frac{x^{\nu} - x}{\nu})$ for all $\nu = 1, 2, \dots$. Without loss of generality we may assume that the vectors $\frac{x^{\nu} - x}{\nu}$ converge to some nonzero vector u as ν tends to ∞ . Then

$$\langle a^i, u \rangle = \lim_{\nu \rightarrow \infty} \left(\langle a^i, \frac{x^{\nu}}{\nu} \rangle - \langle a^i, \frac{x}{\nu} \rangle \right) \leq 0$$

for all $i = 1, \dots, k$. In view of Theorem 2.3.12 the vector u is an asymptotic direction of P and $v = L(u) \in L(P_{\infty})$ as requested. \square

Corollary 2.3.18 *Let P, P_1 and P_2 be polyhedra in \mathbb{R}^n with $P \subseteq P_1$. Then*

$$\begin{aligned} (P)_{\infty} &\subseteq (P_1)_{\infty} \\ (P_1 \times P_2)_{\infty} &= (P_1)_{\infty} \times (P_2)_{\infty} \\ (P_1 + P_2)_{\infty} &= (P_1)_{\infty} + (P_2)_{\infty}. \end{aligned}$$

Proof The first two expressions are direct from the definition of asymptotic directions. For the third expression consider the linear transformation L from $\mathbb{R}^n \times \mathbb{R}^n$ to \mathbb{R}^n defined by $L(x, y) = x + y$, and apply Corollary 2.3.17 and the second expression to conclude. \square

Polar cones

Given a cone C in \mathbb{R}^n , the (negative) *polar cone* of C (Fig. 2.11) is the set

$$C^{\circ} := \left\{ v \in \mathbb{R}^n : \langle v, x \rangle \leq 0 \text{ for all } x \in C \right\}.$$

The polar cone of C° is called the bipolar cone of C . Here is a formula to compute the polar cone of a polyhedral cone.

Theorem 2.3.19 *The polar cone of the polyhedral cone determined by the system*

$$\langle a^i, x \rangle \leq 0, i = 1, \dots, k,$$

is the positive hull of the vectors a^1, \dots, a^k .

Proof It is clear that any positive combination of vectors a^1, \dots, a^k belongs to the polar cone of the polyhedral cone. Let v be a nonzero vector in the polar cone. Then the following system has no solution

$$\begin{aligned} \langle a^i, x \rangle &\leq 0, i = 1, \dots, k, \\ \langle v, x \rangle &> 0. \end{aligned}$$

According to Farkas' theorem the system

$$\begin{aligned} y_1 a^1 + \dots + y_k a^k &= v, \\ y_1, \dots, y_k &\geq 0 \end{aligned}$$

has a solution, which completes the proof. □

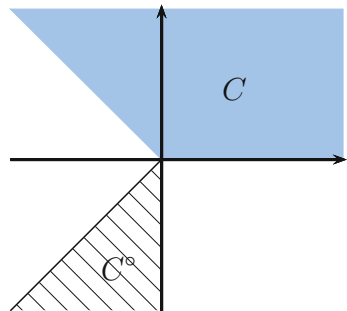
Example 2.3.20 Let C be a polyhedral cone in \mathbb{R}^3 defined by the system:

$$\begin{aligned} x_1 - x_2 &\leq 0, \\ x_3 &= 0. \end{aligned}$$

By expressing the latter equality as two inequalities $x_3 \leq 0$ and $-x_3 \leq 0$, we deduce that the polar cone of C is the positive hull of the three vectors $(1, -1, 0)^T$, $(0, 0, -1)^T$ and $(0, 0, 1)^T$. In other words, the polar cone C° consists of vectors $(t, -t, s)^T$ with $t \in \mathbb{R}_+$ and $s \in \mathbb{R}$.

Corollary 2.3.21 *Let C_1 and C_2 be polyhedral cones in \mathbb{R}^n . Then the following calculus rules hold*

Fig. 2.11 Polar cone



$$\begin{aligned}(C_1 + C_2)^\circ &= C_1^\circ \cap C_2^\circ \\ (C_1 \cap C_2)^\circ &= C_1^\circ + C_2^\circ.\end{aligned}$$

Proof Let $v \in (C_1 + C_2)^\circ$. We have

$$\langle v, x + y \rangle \leq 0 \text{ for all } x \in C_1, y \in C_2.$$

By setting $y = 0$ in this inequality we deduce $v \in C_1^\circ$. Similarly, by setting $x = 0$ we obtain $v \in C_2^\circ$, and hence $v \in C_1^\circ \cap C_2^\circ$. Conversely, if v belongs to both C_1° and C_2° , then $\langle v, \cdot \rangle$ is negative on C_1 and C_2 . Consequently, it is negative on the sum $C_1 + C_2$ by linearity, which shows that $v \in (C_1 + C_2)^\circ$.

For the second equality we observe that the inclusion $C_1^\circ + C_2^\circ \subseteq (C_1 \cap C_2)^\circ$ follows from the definition. To prove the opposite inclusion we assume that C_1 is determined by the system described in Theorem 2.3.19 with $i = 1, \dots, k_1$ and C_2 is determined by that system with $i = k_1 + 1, \dots, k_1 + k_2$. Then the polyhedral cone $C_1 \cap C_2$ is determined by that system with $i = 1, \dots, k_1 + k_2$. In view of Theorem 2.3.19, the polar cone of $C_1 \cap C_2$ is the positive hull of the vectors $a^1, \dots, a^{k_1+k_2}$, which is evidently the sum of the positive hulls $\text{pos}\{a^1, \dots, a^{k_1}\}$ and $\text{pos}\{a^{k_1+1}, \dots, a^{k_1+k_2}\}$, that is the sum of the polar cones C_1° and C_2° . \square

Corollary 2.3.22 *The bipolar cone of a polyhedral cone C coincides with the cone C itself.*

Proof According to Theorem 2.3.19 a vector v belongs to the bipolar cone $C^{\circ\circ}$ if and only if

$$\left\langle v, \sum_{i=1}^k \lambda_i a^i \right\rangle \leq 0 \text{ for all } \lambda_i \geq 0, i = 1, \dots, k.$$

The latter system is equivalent to

$$\langle a^i, v \rangle \leq 0, i = 1, \dots, k,$$

which is exactly the system determining the cone C . \square

Corollary 2.3.23 *A vector v belongs to the polar cone of the asymptotic cone of a convex polyhedron if and only if the linear functional $\langle v, \cdot \rangle$ attains its maximum on the polyhedron.*

Proof It suffices to consider the case where v is nonzero. Assume v belongs to the polar cone of the asymptotic cone P_∞ . In virtue of Theorems 2.3.12 and 2.3.19, it is a positive combination of the vectors a^1, \dots, a^k . Then the linear functional $\langle v, \cdot \rangle$ is majorized by the same combination of real numbers b_1, \dots, b_k on P . Let α be its supremum on P . Our aim is to show that this value is realizable, or equivalently, the system

$$\begin{aligned} \langle a^i, x \rangle &\leq b_i, \quad i = 1, \dots, k \\ \langle v, x \rangle &\geq \alpha \end{aligned}$$

is solvable. Suppose to the contrary that the system has no solution. In view of Corollary 2.2.4, there are a positive vector y and a real number $t \geq 0$ such that

$$\begin{aligned} tv &= A^T y, \\ t\alpha &= \langle b, y \rangle + 1. \end{aligned}$$

We claim that t is strictly positive. Indeed, if $t = 0$, then $A^T y = 0$ and $\langle b, y \rangle = -1$ and for a vector x in P we would deduce

$$0 = \langle A^T y, x \rangle = \langle y, Ax \rangle \leq \langle y, b \rangle = -1,$$

a contradiction. We obtain expressions for v and α as follows

$$v = \frac{1}{t} A^T y \quad \text{and} \quad \alpha = \frac{1}{t} (\langle b, y \rangle + 1).$$

Let $\{x^r\}_{r \geq 1}$ be a maximizing sequence of the functional $\langle v, \cdot \rangle$ on P , which means $\lim_{r \rightarrow \infty} \langle v, x^r \rangle = \alpha$. Then, for every r one has

$$\begin{aligned} \langle v, x^r \rangle &= \frac{1}{t} \langle A^T y, x^r \rangle \\ &\leq \frac{1}{t} \langle y, b \rangle \\ &\leq \alpha - \frac{1}{t}, \end{aligned}$$

which is a contradiction when r is sufficiently large.

For the converse part, let \bar{x} be a point in P where the functional $\langle v, \cdot \rangle$ achieves its maximum. Then

$$\langle v, x - \bar{x} \rangle \leq 0 \quad \text{for all } x \in P.$$

In particular,

$$\langle v, u \rangle \leq 0 \quad \text{for all } u \in P_\infty,$$

and hence v belongs to the polar cone of P_∞ . □

Normal cones

Given a convex polyhedron P determined by the system (2.6) and a point x in P , we say that a vector v is a *normal vector* to P at x if

$$\langle v, y - x \rangle \leq 0 \quad \text{for all } y \in P.$$

The set of all normal vectors to P at x forms a convex cone called the *normal cone* to P at x and denoted $N_P(x)$ (Fig. 2.12). When x is an interior point of P , the normal cone at that point is zero. When x is a boundary point, the normal cone is computed by the next result.

Theorem 2.3.24 *The normal cone to the polyhedron P at a boundary point x of P is the positive hull of the vectors a^i with i being active indices at the point x .*

Proof Let \bar{x} be a boundary point in P . Then the active index set $I(\bar{x})$ is nonempty. Let v be an element of the positive hull of the vectors a^i , $i \in I(\bar{x})$, say

$$v = \sum_{i \in I(\bar{x})} \lambda_i a^i \quad \text{with } \lambda_i \geq 0, i \in I(\bar{x}).$$

Then for every point x in P and every active index $i \in I(\bar{x})$, one has

$$\langle a^i, x - \bar{x} \rangle = \langle a^i, x \rangle - b_i \leq 0,$$

which yields

$$\langle v, x - \bar{x} \rangle = \sum_{i \in I(\bar{x})} \lambda_i \langle a^i, x - \bar{x} \rangle \leq 0.$$

Hence v is normal to P at \bar{x} . For the converse, assume that v is a nonzero vector satisfying

$$\langle v, x - \bar{x} \rangle \leq 0 \quad \text{for all } x \in P. \quad (2.18)$$

We wish to establish that v is a normal vector at 0 to the polyhedron, denoted Q , that is determined by the system

$$\langle a^i, y \rangle \leq 0, \quad i \in I(\bar{x}).$$

This will certainly complete the proof because the normal cone to that polyhedron is exactly its polar cone, the formula of which was already given in Theorem 2.3.19.

Observe that normality condition (2.18) can be written as

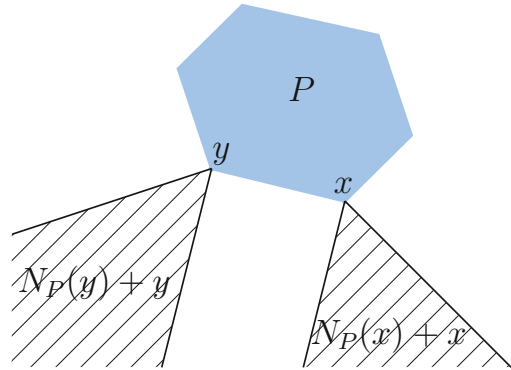
$$\langle v, y \rangle \leq 0 \quad \text{for all } y \in \text{cone}(P - \bar{x}).$$

Therefore, v will be a normal vector to Q at zero if Q coincides with $\text{cone}(P - \bar{x})$. Indeed, let y be a vector of $\text{cone}(P - \bar{x})$, say $y = t(x - \bar{x})$ for some x in P and some positive number t . Then

$$\langle a^i, y \rangle = t \langle a^i, x - \bar{x} \rangle \leq 0,$$

which yields $y \in Q$. Thus, $\text{cone}(P - \bar{x})$ is a subset of Q . For the reverse inclusion we notice that inequalities with inactive indices are strict at \bar{x} . Therefore, given a vector y in Q , one can find a small positive number t such that

Fig. 2.12 Normal cone



$$\langle a^j, \bar{x} \rangle + t \langle a^j, y \rangle \leq b_j$$

for all j inactive. Of course, when i is active, it is true that

$$\langle a^i, \bar{x} + ty \rangle = \langle a^i, \bar{x} \rangle + t \langle a^i, y \rangle \leq b_i.$$

Hence, $\bar{x} + ty$ belongs to P , or equivalently y belongs to $\text{cone}(P - \bar{x})$. This achieves the proof. \square

Example 2.3.25 Consider the polyhedron in \mathbb{R}^3 defined by the system:

$$\begin{aligned} x_1 + x_2 + x_3 &\leq 1, \\ -2x_1 - 3x_2 &\leq -1, \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

This is a convex polytope with six vertices

$$\begin{aligned} v_1 &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 1/3 \\ 0 \end{pmatrix}, \\ v_4 &= \begin{pmatrix} 1/2 \\ 0 \\ 0 \end{pmatrix}, v_5 = \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \end{pmatrix}, v_6 = \begin{pmatrix} 0 \\ 1/3 \\ 2/3 \end{pmatrix} \end{aligned}$$

and five two-dimensional faces

$$\begin{aligned} \text{co}\{v_1, v_2, v_5, v_6\}, \text{co}\{v_1, v_2, v_3, v_4\}, \text{co}\{v_1, v_4, v_5\}, \\ \text{co}\{v_3, v_4, v_5, v_6\}, \text{co}\{v_2, v_3, v_6\}. \end{aligned}$$

At the vertex v_1 there are three active constraints:

$$\begin{aligned}x_1 + x_2 + x_3 &= 1, \\x_2 &= 0, \\x_3 &= 0,\end{aligned}$$

and two non-active constraints

$$\begin{aligned}-2x_1 - 3x_2 &\leq -1, \\-x_1 &\leq 0.\end{aligned}$$

Hence the normal cone at the vertex v_1 is the positive hull of the vectors $u_1 = (1, 1, 1)^T$, $u_2 = (0, -1, 0)^T$ and $u_3 = (0, 0, -1)^T$. Notice that u_1 generates the normal cone at the point $(1/3, 1/3, 1/3)^T$ on the two-dimensional face $F_1 = \text{co}\{v_1, v_2, v_6, v_5\}$, u_2 generates the normal cone at the point $(2/3, 0, 1/4)^T$ on the two-dimensional face $F_2 = \text{co}\{v_1, v_4, v_5\}$, and the positive hull of u_1 and u_2 is the normal cone at the point $(3/4, 0, 1/4)^T$ on the one-dimensional face $[v_1, v_5]$ that is the intersection of the two-dimensional faces F_1 and F_2 .

As a direct consequence of Theorem 2.3.24, we observe that the normal cone is the same at any relative interior point of a face. We refer to this cone as the normal cone to a face. In view of Corollary 2.3.7 we obtain a collection of all normal cones of faces, whose union is called the normal cone of P and denoted by N_P . Thus, if $\mathcal{F} := \{F_1, \dots, F_q\}$ is the collection of all faces of P , then

$$N_P = \bigcup_{i=1}^q N(F_i).$$

It is to point out a distinction between this cone and the cone $N(P)$, the normal cone to P when P is considered as a face of itself. We shall see now that the collection \mathcal{N} of all normal cones $N(F_i)$, $i = 1, \dots, q$, is a nice dual object of the collection \mathcal{F} .

Theorem 2.3.26 *Assume that P is a convex polyhedron given by the system*

$$\langle a^i, x \rangle \leq b_i, \quad i = 1, \dots, k.$$

Then the following assertions hold.

- (i) *The normal cone of P is composed of all normal cones to P at its points, that is*

$$N_P = \bigcup_{x \in P} N_P(x)$$

and coincides with the polar cone of the asymptotic cone of P . In particular, it is a polyhedral cone, and it is the whole space if and only if P is a polytope (bounded polyhedron).

- (ii) In the collection \mathcal{F} , if F_i is a face of F_j , then $N(F_j)$ is a face of $N(F_i)$. Moreover, if $i \neq j$, then the normal cones $N(F_i)$ and $N(F_j)$ have no relative interior point in common.
- (iii) In the collection \mathcal{N} , if N is a face of $N(F_i)$, then there is a face F_ℓ containing the face F_i such that $N = N(F_\ell)$.

Proof For the first property it is evident that N_P is contained in the union of the right hand side. Let $x \in P$. There exists an index $i \in \{1, \dots, k\}$ such that $x \in \text{ri}(F_i)$. Then $N_P(x) = N(F_i)$ and equality of (i) is satisfied. To prove that N_P coincides with $(P_\infty)^\circ$, let v be a vector of the normal cone $N(F_i)$ for some i . Choose a relative interior point x_0 of the face F_i . Then, by definition,

$$\langle v, x - x_0 \rangle \leq 0 \text{ for all } x \in P.$$

By Corollary 2.3.23 the vector v belongs to the polar cone of the cone P_∞ . Conversely, let v be in $(P_\infty)^\circ$. In view of the same corollary, the linear functional $\langle v, \cdot \rangle$ attains its maximum on P at some point \bar{x} , which means that

$$\langle v, x - \bar{x} \rangle \leq 0 \text{ for all } x \in P.$$

By definition, v is a normal vector to P at \bar{x} .

For (ii), assume that F_i is a face of F_j with $i \neq j$, which implies that the active index set I_{F_i} of F_i contains the active index set I_{F_j} of F_j . Let x_j be a relative interior point of F_j . Then one has

$$N(F_j) = N_P(x_j) \subset N(F_i).$$

Suppose that $N(F_j)$ is not a face of $N(F_i)$. There exists a face

$$N_0 = \text{pos}\{a^\ell : \ell \in I_0\} \subseteq N(F_i)$$

for some $I_0 \subseteq I_{F_i}$, which contains $N(F_j)$ as a proper subset and such that its relative interior meets $N(F_j)$ at some point, say v_0 . Let F_0 be the solution set to the system

$$\begin{aligned} \langle a^\ell, x \rangle &= b_\ell, \ell \in I_0, \\ \langle a^\ell, x \rangle &\leq b_\ell, \ell \in \{1, \dots, p\} \setminus I_0. \end{aligned}$$

We see that $I_{F_j} \subseteq I_0 \subseteq I_{F_i}$, hence $F_i \subseteq F_0 \subseteq F_j$. In particular $F_0 \neq \emptyset$, hence it is a face of P . Let x_0 be a relative interior point of F_0 . We claim that

$$\langle v, x_j - x_0 \rangle = 0 \text{ for all } v \in N_0.$$

Indeed, consider the linear functional $v \mapsto \langle v, x_j - x_0 \rangle$ on N_0 . On the one hand, $\langle v, x_j - x_0 \rangle \leq 0$ for all $v \in N_0$ because $x_0 \in \text{ri}(F_0)$. On the other hand, for $v_0 \in \text{ri}(N_0) \cap N(F_j)$ above, one has $\langle v_0, x_0 - x_j \rangle \leq 0$, hence $\langle v_0, x_j - x_0 \rangle = 0$.

Consequently, $\langle v, x_j - x_0 \rangle = 0$ on N_0 . Using this fact we derive for every $v \in N_0$ that

$$\langle v, x - x_j \rangle = \langle v, x - x_0 \rangle + \langle v, x_0 - x_j \rangle \leq 0,$$

for all $x \in M$, which implies $v \in N(F_j)$ and arrive at the contradiction $N(F_j) = N_0$. To prove the second part of assertion (ii), suppose to the contrary that the normal cones $N(F_i)$ and $N(F_j)$ have a relative interior point v in common. Then for each $x \in F_i$ and $y \in F_j$ one has

$$\langle v, x - y \rangle = 0.$$

Since $\langle u, y - x \rangle \leq 0$ for all $u \in N(F_i)$ and v is a relative interior point of $N(F_i)$, one deduces

$$\langle u, x - y \rangle = 0 \text{ for all } u \in N(F_i).$$

Consequently, for $u \in N(F_i)$ it is true that

$$\langle u, z - y \rangle = \langle u, z - x \rangle + \langle u, x - y \rangle \leq 0 \text{ for all } z \in P,$$

which shows $u \in N(F_j)$. In other words $N(F_i) \subseteq N(F_j)$. The same argument with i and j interchanging the roles, leads to equality $N(F_i) = N(F_j)$. In view of the first part we arrive at the contradiction $F_i = F_j$.

We proceed to (iii). Let N be a face of $N(F_i)$ for some $i : 1 \leq i \leq k$. The case $N = N(F_i)$ being trivial, we may assume $N \neq N(F_i)$. Let $I \subseteq I_{F_i}$ be a subset of indices such that

$$N = \text{cone}\{a^\ell : \ell \in I\} \subseteq N(F_i) = \text{cone}\{a^\ell : \ell \in I_{F_i}\}.$$

Let F be the solution set to the system

$$\begin{aligned} \langle a^\ell, x \rangle &= b_\ell, \quad \ell \in I, \\ \langle a^\ell, x \rangle &\leq b_\ell, \quad \ell \in \{1, \dots, p\} \setminus I. \end{aligned}$$

Since $I \subseteq I_{F_i}$, we have $F_i \subseteq F$. In particular $F \neq \emptyset$ and F is a face of P . Now we show that $N(F) = N$ and F_i is a proper face of F . Indeed, as N is a proper face of $N(F_i)$, I is a proper subset of I_{F_i} and there is a nonzero vector $u \in R^n$ such that

$$\begin{aligned} \langle a^\ell, u \rangle &= 0 \quad \text{for } \ell \in I, \\ \langle a^\ell, u \rangle &< 0 \quad \text{for } \ell \in I_{F_i} \setminus I. \end{aligned}$$

Take $x \in \text{ri}(F_i)$ and consider the point $x + tu$ with $t > 0$. One obtains

$$\begin{aligned} \langle a^\ell, x + tu \rangle &= b_\ell, \quad \ell \in I, \\ \langle a^\ell, x + tu \rangle &= \langle a^\ell, x \rangle + t \langle a^\ell, u \rangle < b_\ell, \quad \ell \in I_{F_i} \setminus I. \end{aligned}$$

Moreover, since $\langle a^\ell, x \rangle < b_\ell$ for $\ell \in \{1, \dots, p\} \setminus I_{F_i}$, when t is sufficiently small, one also has

$$\langle a^\ell, x + tu \rangle < b_\ell, \ell \in \{1, \dots, p\} \setminus I_{F_i}.$$

Consequently,

$$N(F) = N_P(x + tu) = \text{pos}\{a^\ell : \ell \in I\} = N$$

when t is sufficiently small. It is evident that $F \neq F_i$. The proof is complete. \square

Example 2.3.27 Consider the polyhedron P in \mathbb{R}^2 defined by the system:

$$\begin{aligned} -x_1 - x_2 &\leq -1, \\ -x_1 + x_2 &\leq 1, \\ -x_2 &\leq 0. \end{aligned}$$

It has two vertices F_1 and F_2 determined respectively by

$$\begin{cases} -x_1 - x_2 = -1 \\ -x_1 + x_2 \leq 1 \\ -x_2 = 0 \end{cases} \text{ and } \begin{cases} -x_1 - x_2 = -1 \\ -x_1 + x_2 = 1 \\ -x_2 \leq 0 \end{cases}$$

three one-dimensional faces F_3 , F_4 and F_5 determined respectively by

$$\begin{cases} -x_1 - x_2 \leq -1 \\ -x_1 + x_2 \leq 1 \\ -x_2 = 0 \end{cases}, \begin{cases} -x_1 - x_2 = -1 \\ -x_1 + x_2 \leq 1 \\ -x_2 \leq 0 \end{cases} \text{ and } \begin{cases} -x_1 - x_2 \leq -1 \\ -x_1 + x_2 = 1 \\ -x_2 \leq 0 \end{cases}$$

and P itself is the unique two-dimensional face. Denote by $v_1 = (-1, -1)^T$, $v_2 = (-1, 1)^T$ and $v_3 = (0, -1)^T$. Then the normal cones of the faces F_1, \dots, F_5 are respectively the positive hulls of the families $\{v_1, v_3\}$, $\{v_1, v_2\}$, $\{v_3\}$, $\{v_1\}$ and $\{v_2\}$. The normal cone of P is zero. Moreover, the union N_P of these normal cones is the positive hull of the vectors v_2 and v_3 . It is the polar cone of the asymptotic cone of P which is defined by the system

$$\begin{aligned} -x_1 - x_2 &\leq 0, \\ -x_1 + x_2 &\leq 0, \\ -x_2 &\leq 0, \end{aligned}$$

in which the first inequality is redundant and hence it is reduced to $x_1 \geq x_2 \geq 0$.

Next we prove that the normal cone of a face is obtained from the normal cones of its vertices.

Corollary 2.3.28 *Assume that a face F of P is the convex hull of its vertices v^1, \dots, v^q . Then*

$$N(F) = \bigcap_{i=1}^q N_P(v^i).$$

Proof The inclusion $N(F) \subseteq \bigcap_{i=1}^q N_P(v^i)$ is clear from (ii) of Theorem 2.3.26. We prove the converse inclusion. Let u be a nonzero vector of the intersection $\bigcap_{i=1}^q N_P(v^i)$. Let x be a relative interior point of F . Then x is a convex combination of the vertices v^1, \dots, v^q :

$$x = \sum_{i=1}^q \lambda_i v^i$$

with $\lambda_i \geq 0$, $i = 1, \dots, q$ and $\lambda_1 + \dots + \lambda_q = 1$. We have then

$$\langle u, x' - v^i \rangle \leq 0 \text{ for all } x' \in P, i = 1, \dots, q.$$

This implies

$$\begin{aligned} \langle u, x' - x \rangle &= \langle u, \sum_{i=1}^q \lambda_i x' - \sum_{i=1}^q \lambda_i v^i \rangle \\ &= \sum_{i=1}^q \lambda_i \langle u, x' - v^i \rangle \leq 0 \end{aligned}$$

By this, u is a normal vector to P at x , and $u \in N(F)$. □

Combining this corollary with Corollary 2.3.8 we conclude that the normal cone of a bounded face is the intersection of the normal cones of all proper faces of that bounded face. This is not true for unbounded faces, for instance when a face has no proper face.

2.4 Basis and Vertices

In this section we consider a polyhedron P given by the system

$$\begin{aligned} Ax &= b \\ x &\geq 0. \end{aligned} \tag{2.19}$$

We assume throughout that the matrix A has n columns denoted a_1, \dots, a_n and k rows that are transposes of a^1, \dots, a^k and linearly independent, and that the components b_1, \dots, b_k of the vector b are non-negative numbers. A point x in P is said to be an *extreme point* of P if it cannot be expressed as a convex combination $x = ta + (1-t)a'$ for some $0 < t < 1$ and $a, a' \in P$ with $a \neq a'$. It can be seen that extreme points correspond to vertices we have defined in the previous section. Certain results we have obtained for polyhedra given in a general form (by inequalities) will be recaptured here, but our emphasis will be laid on computing issues which are much simplified under equality form (2.19).

A $k \times k$ -submatrix B composed of columns of A is said to be a *basis* if it is invertible.

Let B be a basis. By using a permutation one may assume that B is composed of the first k columns of A , and the remaining columns form a $k \times (n - k)$ -submatrix N , called a *non-basic part* of A . Let x be a vector with components x_B and x_N , where x_B is a k -dimensional vector and x_N is an $(n - k)$ -dimensional vector satisfying

$$\begin{aligned} Bx_B &= b, \\ x_N &= 0. \end{aligned}$$

If x_B is a positive vector, then x is a solution to (2.19) and called a *feasible basic solution* (associated with the basis B). If in addition x_B has no zero component, it is called *non-degenerate*; otherwise it is *degenerate*.

Example 2.4.1 Consider the polyhedron in \mathbb{R}^3 defined by the system:

$$\begin{aligned} x_1 + x_2 + x_3 &= 1, \\ 3x_1 + 2x_2 &= 1, \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

The vectors $a_1 = (1, 1, 1)^T$ and $a_2 = (3, 2, 0)^T$ are linear independent. There are three bases

$$B_1 = \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix}, B_2 = \begin{pmatrix} 1 & 1 \\ 3 & 0 \end{pmatrix} \text{ and } B_3 = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}.$$

The basic solutions corresponding to B_1, B_2 and B_3 are respectively $(-1, 2, 0)^T$, $(1/3, 0, 2/3)^T$ and $(0, 1/2, 1/2)^T$. The first solution is unfeasible, while the two last ones are feasible and non-degenerate.

Given a vector $x \in \mathbb{R}^n$, its *support*, denoted $\text{supp}(x)$, consists of indices i for which the component x_i is nonzero. The support of a nonzero vector is always nonempty.

Theorem 2.4.2 *A vector x is a vertex of the polyhedron P if and only if it is a feasible basic solution of the system (2.19).*

Proof Let x be a feasible basic solution. Assume that it is a convex combination of two solutions y and z of the system (2.19), say $x = ty + (1 - t)z$ with $t \in (0, 1)$. Then for any nonbasic index j , the component x_j is zero, so that $ty_j + (1 - t)z_j = 0$. Remembering that y and z are positive vectors, we derive $y_j = z_j = 0$. Moreover, the basic components of solutions to (2.19) satisfy equation

$$Bx_B = b$$

with B nonsingular. Therefore, they are unique, that is $x_B = y_B = z_B$. Consequently, the three solutions x , y and z are the same.

Conversely, let x be an extreme point of the polyhedron. Our aim is to show that the columns a^i , $i \in \text{supp}(x)$ are linearly independent. It is then easy to find a basis B such that x is the basic solution associated with that basis. To this end, we prove first that $\text{supp}(x)$ is minimal by inclusion among solutions of the system (2.19). In fact, if not, one can find another solution, say y , with minimal support such that $\text{supp}(y)$ is a proper subset of $\text{supp}(x)$. Choose an index j from the support of y such that

$$\frac{x_j}{y_j} = \min\left\{\frac{x_i}{y_i} : i \in \text{supp}(y)\right\}.$$

Let $t > 0$ be that quotient. Then

$$A(x - ty) = (1 - t)b \quad \text{and} \quad x - ty \geq 0.$$

If $t \geq 1$, then by setting $z = x - y$ we can express

$$x = \frac{1}{2}(y + \frac{2}{3}z) + \frac{1}{2}(y + \frac{4}{3}z),$$

a convex combination of two distinct solutions of (2.19), which is a contradiction. If $t < 1$, then take

$$z = \frac{1}{1 - t}(x - ty).$$

We see that z is a solution to (2.19) and different from x because its support is strictly contained in the support of x . It is also different from y because the component y_j is not zero while the component z_j is zero. We derive from the definition of z that x is a strict convex combination of y and z , which is again a contradiction.

Now we prove that the columns a^i , $i \in \text{supp}(x)$ are linearly independent. Suppose the contrary: there is a vector y different from x (if not take $2y$ instead) with

$$Ay = 0 \quad \text{and} \quad \text{supp}(y) \subseteq \text{supp}(x).$$

By setting

$$t = \begin{cases} -\min\{\frac{x_i}{y_i} : i \in \text{supp}(y)\} & \text{if } y \geq 0 \\ \min\{-\frac{x_i}{y_i} : i \in \text{supp}(y), y_i < 0\} & \text{else} \end{cases}$$

we obtain that $z = x + ty$ is a solution to (2.19) whose support is strictly contained in the support of x and arrive at a contradiction with the minimality of the support of x . It remains to complete the vectors a^i , $i \in \text{supp}(x)$ to a basis to see that x is indeed a basic solution. \square

Corollary 2.4.3 *The number of vertices of the polyhedron P does not exceed the binomial coefficient $\binom{n}{k}$.*

Proof This follows from Theorem 2.4.2 and the fact that the number of bases of the matrix A is at most $\binom{n}{k}$. Notice that not every basic solution has positive components. \square

We deduce again Corollary 2.3.8 about the description of polytopes in terms of extreme points (vertices), but this time for a polytope determined by the system (2.19).

Corollary 2.4.4 *If P is a polytope, then any point in it can be expressed as a convex combination of vertices.*

Proof Let x be any solution of (2.19). If the support of x is minimal, then in view of Theorem 2.4.2 that point is a vertex. If not, then there is a solution y^1 different from x , with minimal support and $\text{supp}(y^1) \subset \text{supp}(x)$. Set

$$t_1 = \min \left\{ \frac{x_j}{y_j^1} : j \in \text{supp}(y^1) \right\}.$$

This number is positive and strictly smaller than one, because otherwise the nonzero vector $x - y^1$ should be an asymptotic direction of the polyhedron and P should be unbounded. Consider the vector

$$z^1 = \frac{1}{1 - t_1}(x - t_1 y^1).$$

It is clear that this vector is a solution to (2.19) and its support is strictly smaller than the support of x . If the support of z^1 is minimal, then z^1 is a vertex and we obtain a convex combination

$$x = t_1 y^1 + (1 - t_1) z^1,$$

in which y^1 and z^1 are vertices. If not, we continue the process to find a vertex y^2 whose support is strictly contained in the support of z^1 and so on. In view of Corollary 2.4.3 after a finite number of steps one finds vertices y^1, \dots, y^p such that x is a convex combination of them. \square

Extreme rays

Extreme direction of a convex polyhedron P in \mathbb{R}^n can be defined to be a direction that cannot be expressed as a strictly positive combination of two linearly independent asymptotic vectors of P . As the case when a polyhedron is given by a system of linear inequalities (Corollary 2.3.15), we shall see that a polyhedron determined by (2.19) is completely determined by its vertices and extreme directions.

Theorem 2.4.5 *Assume that the convex polyhedron P is given by the system (2.19). Then*

- (i) *A nonzero vector v is an asymptotic direction of P if and only if it is a solution to the associated homogenous system*

$$\begin{aligned} Ax &= 0, \\ x &\geq 0. \end{aligned}$$

- (ii) *A nonzero vector v is an extreme asymptotic direction of P if and only if it is a positive multiple of a vertex of the polyhedron determined by the system*

$$\begin{aligned} Ay &= 0 & (2.20) \\ y_1 + \cdots + y_n &= 1, \\ y &\geq 0. \end{aligned}$$

Consequently P_∞ consists of all positive combinations of the vertices of this latter polyhedron.

Proof The first assertion is proven as in Theorem 2.3.12. For the second assertion, let v be a nonzero extreme direction. Then $Av = 0$ by (i) and $t := v_1 + \cdots + v_n > 0$. The vector v/t is in the polyhedron of (ii), denoted Q . Since each point of that polyhedron is an asymptotic direction of P , if v/t were a convex combination of two distinct points y^1 and y^2 in Q , then v would be a convex combination of two linearly independent asymptotic directions ty^1 and ty^2 of P , which is a contradiction. Conversely, let v be a vertex of Q . It is clear that v is nonzero. If $v = tx + (1-t)y$ for some nonzero asymptotic directions x and y of P and some $t \in (0, 1)$, then with

$$\begin{aligned} t' &= \frac{t \sum_{i=1}^n x_i}{t \sum_{i=1}^n x_i + (1-t) \sum_{i=1}^n y_i} = t \sum_{i=1}^n x_i, \\ x' &= \frac{1}{\sum_{i=1}^n x_i} x, \\ y' &= \frac{1}{\sum_{i=1}^n y_i} y, \end{aligned}$$

we express v as a convex combination $t'x' + (1 - t')y'$ of two points of Q . Note that $t' > 0$. By hypothesis, $x' = y'$ which means that x and y are linearly dependent. The proof is complete. \square

Corollary 2.4.6 *A nonzero vector is an extreme asymptotic direction of P if and only if it is a basic feasible solution of the system (2.20). Consequently, the number of extreme asymptotic directions of P does not exceed the binomial coefficient $\binom{n}{k+1}$.*

Proof This is obtained from Theorems 2.4.2 and 2.4.5. \square

Example 2.4.7 Consider the polyhedron in \mathbb{R}^3 defined by the system:

$$\begin{aligned} x_1 - x_2 &= 1 \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

The asymptotic cone of this polyhedron is the solution set to the system

$$\begin{aligned} x_1 - x_2 &= 0 \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

Any vector $(t, t, s)^T$ with $t \geq 0$ and $s \geq 0$ is an asymptotic direction. To obtain extreme asymptotic directions we solve the system

$$\begin{aligned} y_1 - y_2 &= 0 \\ y_1 + y_2 + y_3 &= 1 \\ y_1, y_2, y_3 &\geq 0. \end{aligned}$$

There are three bases corresponding to basic variables $\{y_1, y_2\}$, $\{y_1, y_3\}$ and $\{y_2, y_3\}$:

$$B_1 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad B_3 = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}.$$

The basic solution $y = (1/2, 1/2, 0)^T$ is associated with B_1 and the basic solution $y = (0, 0, 1)^T$ is associated with B_2 and B_3 . Both of them are feasible, and hence they are extreme asymptotic directions.

In the following we describe a practical way to compute extreme rays of the polyhedron P .

Corollary 2.4.8 *Assume that B is a basis of the matrix A and a_s is a non-basic column of A such that the system*

$$By = -a_s$$

has a positive solution $\bar{y} \geq 0$. Then the vector \bar{x} whose basic components are equal to \bar{y} , the s th component is equal to 1 and the other non-basic components are all zero, is an extreme ray of the polyhedron P .

Proof It is easy to check that the submatrix corresponding to the variables of \bar{y} and the variable y_s is a feasible basis of the system (2.20). It remains to apply Corollary 2.4.6 to conclude. \square

In Example 2.4.7 we have $A = (1, -1, 0)$. For the basis $B = (1)$ corresponding to the basic variable x_1 and the second non-basic column, the system $By = -a_s$ takes the form $y = 1$ and has a positive solution $y = 1$. In view of Corollary 2.4.8 the vector $(1, 1, 0)^T$ is an extreme asymptotic direction. Note that using the same basis B and the non-basic column $a_3 = (0)$ we obtain the system $y = 0$ which has a positive (null) solution. Hence the vector $(0, 0, 1)^T$ is also an extreme asymptotic direction.

Representation of Elements of a Polyhedron

A finitely generated convex set is defined to be a set which is the convex hull of a finite set of points and directions, that is, each element of it is the sum of a convex combination of a finite set of points and a positive combination of a finite set of directions. The next theorem states that convex polyhedra are finitely generated, which is Corollary 2.3.15 for a polyhedron determined by the system (2.19).

Theorem 2.4.9 *Every point of a convex polyhedron given by the system (2.19) can be expressed as a convex combination of its vertices, possibly added to a positive combination of the extreme asymptotic directions.*

Proof Let x be any point in P . If its support is minimal, then, according to the proof of Theorem 2.4.2 that point is a vertex. If not, there is a vertex v^1 whose support is minimal and strictly contained in the support of x . Set

$$t = \min \left\{ \frac{x_j}{v_j^1} : j \in \text{supp}(v^1) \right\}$$

and consider the vector $x - tv^1$. If $t \geq 1$, then the vector $z = x - v^1$ is an asymptotic direction of the polyhedron and then x is the sum of the vertex v^1 and an asymptotic direction. The direction z , in its turn, is expressed as a convex combination of extreme asymptotic directions. So the corollary follows. If $t < 1$, the technique of proof of Theorem 2.4.2 can be applied. Expressly, setting $z = (x - tv^1)/(1 - t)$ we deduce that $z \geq 0$ and

$$Az = \frac{1}{1-t}b - \frac{t}{1-t}b = b.$$

Moreover, the support of z is a proper subset of the support of x because the components j of z with j realizing the value of $t = x_j/v_j^1$ are zero. Then $x = tv^1 + (1-t)z$ with strict inclusion $\text{supp}(z) \subset \text{supp}(x)$. Continuing this process we arrive at finding

a finite number of vertices v^1, \dots, v^p and an asymptotic direction z such that x is the sum of a convex combination of v^1, \dots, v^p and z . Then expressing z as a convex combination of asymptotic extreme directions we obtain the conclusion. \square

In view of Corollaries 2.4.3 and 2.4.6 the numbers of vertices and extreme asymptotic directions of a polyhedron P are finite. Denote them respectively by v^1, \dots, v^p and z^1, \dots, z^q . Then each element x of P is expressed as

$$x = \sum_{i=1}^p \lambda_i v^i + \sum_{j=1}^q \mu_j z^j$$

with

$$\sum_{i=1}^p \lambda_i = 1, \lambda_i \geq 0, i = 1, \dots, p \text{ and } \mu_j \geq 0, j = 1, \dots, q.$$

Notice that the above representation is not unique, that is, an element x of P can be written as several combinations of $v^i, i = 1, \dots, p$ and $z^j, j = 1, \dots, q$ with different coefficients λ_i and μ_j . An easy example can be observed for the center x of the square with vertices

$$v^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, v^2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, v^3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ and } v^4 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

It is clear that x can be seen as the middle point of v^1 and v^4 , and as the middle point of v^2 and v^3 too.

Another point that should be made clear is the fact that the results of this section are related to polyhedra given by the system (2.19) and they might be false under systems of different type. For instance, in view of Theorem 2.4.9 a polyhedron determined by (2.19) has at least a vertex. This is no longer true if a polyhedron is given by another system. Take a hyperplane determined by equation $\langle d, x \rangle = 0$ for some nonzero vector $d \in \mathbb{R}^2$. It is a polyhedron without vertices. An equivalent system is given in form of (2.19) as follows

$$\begin{aligned} \langle d, x^+ \rangle - \langle d, x^- \rangle &= 0, \\ x^+, x^- &\geq 0. \end{aligned}$$

The latter system generates a polyhedron in \mathbb{R}^4 that does have vertices. However, a vertex $(x^+, x^-)^T$ of this polyhedron gives an element $x = x^+ - x^-$ of the former polyhedron, but not a vertex of it.