Chapter 2

Convex sets

2.1 Affine and convex sets

2.1.1 Lines and line segments

Suppose $x_1 \neq x_2$ are two points in $\mathbf{R}^n$. Points of the form

$$y = \theta x_1 + (1 - \theta)x_2,$$

where $\theta \in \mathbf{R}$, form the *line* passing through x_1 and x_2. The parameter value $\theta = 0$ corresponds to $y = x_2$, and the parameter value $\theta = 1$ corresponds to $y = x_1$. Values of the parameter θ between 0 and 1 correspond to the (closed) *line segment* between x_1 and x_2.

Expressing y in the form

$$y = x_2 + \theta(x_1 - x_2)$$

gives another interpretation: y is the sum of the *base point* x_2 (corresponding to $\theta = 0$) and the *direction* $x_1 - x_2$ (which points from x_2 to x_1) scaled by the parameter θ. Thus, θ gives the fraction of the way from x_2 to x_1 where y lies. As θ increases from 0 to 1, the point y moves from x_2 to x_1; for $\theta > 1$, the point y lies on the line beyond x_1. This is illustrated in figure 2.1.

2.1.2 Affine sets

A set $C \subseteq \mathbf{R}^n$ is *affine* if the line through any two distinct points in C lies in C, *i.e.*, if for any $x_1, x_2 \in C$ and $\theta \in \mathbf{R}$, we have $\theta x_1 + (1-\theta)x_2 \in C$. In other words, C contains the linear combination of any two points in C, provided the coefficients in the linear combination sum to one.

This idea can be generalized to more than two points. We refer to a point of the form $\theta_1 x_1 + \cdots + \theta_k x_k$, where $\theta_1 + \cdots + \theta_k = 1$, as an *affine combination* of the points $x_1, \ldots, x_k$. Using induction from the definition of affine set (*i.e.*, that it contains every affine combination of two points in it), it can be shown that

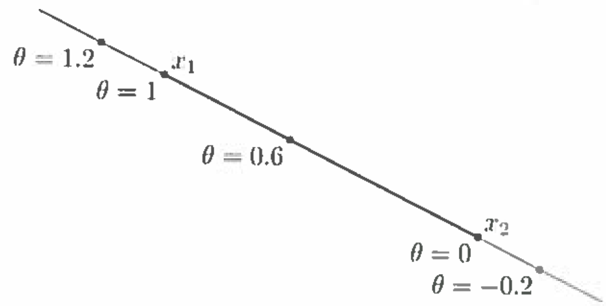

Figure 2.1 The line passing through x_1 and x_2 is described parametrically by $\theta x_1 + (1-\theta)x_2$, where θ varies over $\mathbf{R}$. The line segment between x_1 and x_2, which corresponds to θ between 0 and 1, is shown darker.

an affine set contains every affine combination of its points: If C is an affine set, $x_1, \ldots, x_k \in C$, and $\theta_1 + \cdots + \theta_k = 1$, then the point $\theta_1 x_1 + \cdots + \theta_k x_k$ also belongs to C.

If C is an affine set and $x_0 \in C$, then the set

$$V = C - x_0 = \{x - x_0 \mid x \in C\}$$

is a subspace, *i.e.*, closed under sums and scalar multiplication. To see this, suppose $v_1, v_2 \in V$ and $\alpha, \beta \in \mathbf{R}$. Then we have $v_1 + x_0 \in C$ and $v_2 + x_0 \in C$, and so

$$\alpha v_1 + \beta v_2 + x_0 = \alpha(v_1 + x_0) + \beta(v_2 + x_0) + (1 - \alpha - \beta)x_0 \in C,$$

since C is affine, and $\alpha + \beta + (1 - \alpha - \beta) = 1$. We conclude that $\alpha v_1 + \beta v_2 \in V$, since $\alpha v_1 + \beta v_2 + x_0 \in C$.

Thus, the affine set C can be expressed as

$$C = V + x_0 = \{v + x_0 \mid v \in V\},$$

i.e., as a subspace plus an offset. The subspace V associated with the affine set C does not depend on the choice of x_0, so x_0 can be chosen as any point in C. We define the *dimension* of an affine set C as the dimension of the subspace $V = C - x_0$, where x_0 is any element of C.

Example 2.1 *Solution set of linear equations.* The solution set of a system of linear equations, $C = \{x \mid Ax = b\}$, where $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$, is an affine set. To show this, suppose $x_1, x_2 \in C$, *i.e.*, $Ax_1 = b$, $Ax_2 = b$. Then for any θ, we have

$$\begin{aligned} A(\theta x_1 + (1-\theta)x_2) &= \theta A x_1 + (1-\theta) A x_2 \\ &= \theta b + (1-\theta) b \\ &= b, \end{aligned}$$

which shows that the affine combination $\theta x_1 + (1-\theta)x_2$ is also in C. The subspace associated with the affine set C is the nullspace of A.

We also have a converse: every affine set can be expressed as the solution set of a system of linear equations.

The set of all affine combinations of points in some set $C \subseteq \mathbf{R}^n$ is called the *affine hull* of C, and denoted $\mathbf{aff}\, C$:

$$\mathbf{aff}\, C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_1, \ldots, x_k \in C,\ \theta_1 + \cdots + \theta_k = 1\}.$$

The affine hull is the smallest affine set that contains C, in the following sense: if S is any affine set with $C \subseteq S$, then $\mathbf{aff}\, C \subseteq S$.

2.1.3 Affine dimension and relative interior

We define the *affine dimension* of a set C as the dimension of its affine hull. Affine dimension is useful in the context of convex analysis and optimization, but is not always consistent with other definitions of dimension. As an example consider the unit circle in $\mathbf{R}^2$, *i.e.*, $\{x \in \mathbf{R}^2 \mid x_1^2 + x_2^2 = 1\}$. Its affine hull is all of $\mathbf{R}^2$, so its affine dimension is two. By most definitions of dimension, however, the unit circle in $\mathbf{R}^2$ has dimension one.

If the affine dimension of a set $C \subseteq \mathbf{R}^n$ is less than n, then the set lies in the affine set $\mathbf{aff}\, C \neq \mathbf{R}^n$. We define the *relative interior* of the set C, denoted $\mathbf{relint}\, C$, as its interior relative to $\mathbf{aff}\, C$:

$$\mathbf{relint}\, C = \{x \in C \mid B(x, r) \cap \mathbf{aff}\, C \subseteq C \text{ for some } r > 0\},$$

where $B(x, r) = \{y \mid \|y - x\| \leq r\}$, the ball of radius r and center x in the norm $\|\cdot\|$. (Here $\|\cdot\|$ is any norm; all norms define the same relative interior.) We can then define the *relative boundary* of a set C as $\mathbf{cl}\, C \setminus \mathbf{relint}\, C$, where $\mathbf{cl}\, C$ is the closure of C.

Example 2.2 Consider a square in the (x_1, x_2)-plane in $\mathbf{R}^3$, defined as

$$C = \{x \in \mathbf{R}^3 \mid -1 \leq x_1 \leq 1,\ -1 \leq x_2 \leq 1,\ x_3 = 0\}.$$

Its affine hull is the (x_1, x_2)-plane, *i.e.*, $\mathbf{aff}\, C = \{x \in \mathbf{R}^3 \mid x_3 = 0\}$. The interior of C is empty, but the relative interior is

$$\mathbf{relint}\, C = \{x \in \mathbf{R}^3 \mid -1 < x_1 < 1,\ -1 < x_2 < 1,\ x_3 = 0\}.$$

Its boundary (in $\mathbf{R}^3$) is itself; its relative boundary is the wire-frame outline,

$$\{x \in \mathbf{R}^3 \mid \max\{|x_1|, |x_2|\} = 1,\ x_3 = 0\}.$$

2.1.4 Convex sets

A set C is *convex* if the line segment between any two points in C lies in C, *i.e.*, if for any $x_1, x_2 \in C$ and any θ with $0 \leq \theta \leq 1$, we have

$$\theta x_1 + (1 - \theta) x_2 \in C.$$

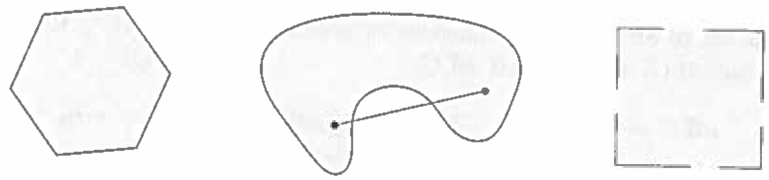

Figure 2.2 Some simple convex and nonconvex sets. *Left.* The hexagon, which includes its boundary (shown darker), is convex. *Middle.* The kidney shaped set is not convex, since the line segment between the two points in the set shown as dots is not contained in the set. *Right.* The square contains some boundary points but not others, and is not convex.

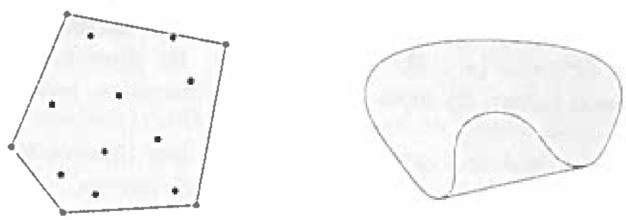

Figure 2.3 The convex hulls of two sets in $\mathbf{R}^2$. *Left.* The convex hull of a set of fifteen points (shown as dots) is the pentagon (shown shaded). *Right.* The convex hull of the kidney shaped set in figure 2.2 is the shaded set.

Roughly speaking, a set is convex if every point in the set can be seen by every other point, along an unobstructed straight path between them, where unobstructed means lying in the set. Every affine set is also convex, since it contains the entire line between any two distinct points in it, and therefore also the line segment between the points. Figure 2.2 illustrates some simple convex and nonconvex sets in $\mathbf{R}^2$.

We call a point of the form $\theta_1 x_1 + \cdots + \theta_k x_k$, where $\theta_1 + \cdots + \theta_k = 1$ and $\theta_i \geq 0$, $i = 1, \ldots, k$, a *convex combination* of the points $x_1, \ldots, x_k$. As with affine sets, it can be shown that a set is convex if and only if it contains every convex combination of its points. A convex combination of points can be thought of as a *mixture* or *weighted average* of the points, with θ_i the fraction of x_i in the mixture.

The *convex hull* of a set C, denoted $\mathbf{conv}\, C$, is the set of all convex combinations of points in C:

$$\mathbf{conv}\, C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C,\ \theta_i \geq 0,\ i = 1, \ldots, k,\ \theta_1 + \cdots + \theta_k = 1\}.$$

As the name suggests, the convex hull $\mathbf{conv}\, C$ is always convex. It is the smallest convex set that contains C: If B is any convex set that contains C, then $\mathbf{conv}\, C \subseteq B$. Figure 2.3 illustrates the definition of convex hull.

The idea of a convex combination can be generalized to include infinite sums, integrals, and, in the most general form, probability distributions. Suppose $\theta_1, \theta_2, \ldots$

satisfy

$$\theta_i \geq 0, \quad i = 1, 2, \ldots, \qquad \sum_{i=1}^{\infty} \theta_i = 1,$$

and $x_1, x_2, \ldots \in C$, where $C \subseteq \mathbf{R}^n$ is convex. Then

$$\sum_{i=1}^{\infty} \theta_i x_i \in C,$$

if the series converges. More generally, suppose $p : \mathbf{R}^n \to \mathbf{R}$ satisfies $p(x) \geq 0$ for all $x \in C$ and $\int_C p(x)\, dx = 1$, where $C \subseteq \mathbf{R}^n$ is convex. Then

$$\int_C p(x) x \, dx \in C,$$

if the integral exists.

In the most general form, suppose $C \subseteq \mathbf{R}^n$ is convex and x is a random vector with $x \in C$ with probability one. Then $\mathbf{E}\, x \in C$. Indeed, this form includes all the others as special cases. For example, suppose the random variable x only takes on the two values x_1 and x_2, with $\mathbf{prob}(x = x_1) = \theta$ and $\mathbf{prob}(x = x_2) = 1 - \theta$, where $0 \leq \theta \leq 1$. Then $\mathbf{E}\, x = \theta x_1 + (1 - \theta) x_2$, and we are back to a simple convex combination of two points.

2.1.5 Cones

A set C is called a *cone*, or *nonnegative homogeneous*, if for every $x \in C$ and $\theta \geq 0$ we have $\theta x \in C$. A set C is a *convex cone* if it is convex and a cone, which means that for any x_1, $x_2 \in C$ and θ_1, $\theta_2 \geq 0$, we have

$$\theta_1 x_1 + \theta_2 x_2 \in C.$$

Points of this form can be described geometrically as forming the two-dimensional pie slice with apex 0 and edges passing through x_1 and x_2. (See figure 2.4.)

A point of the form $\theta_1 x_1 + \cdots + \theta_k x_k$ with $\theta_1, \ldots, \theta_k \geq 0$ is called a *conic combination* (or a *nonnegative linear combination*) of $x_1, \ldots, x_k$. If x_i are in a convex cone C, then every conic combination of x_i is in C. Conversely, a set C is a convex cone if and only if it contains all conic combinations of its elements. Like convex (or affine) combinations, the idea of conic combination can be generalized to infinite sums and integrals.

The *conic hull* of a set C is the set of all conic combinations of points in C, *i.e.*,

$$\{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \ \theta_i \geq 0, \ i = 1, \ldots, k\},$$

which is also the smallest convex cone that contains C (see figure 2.5).

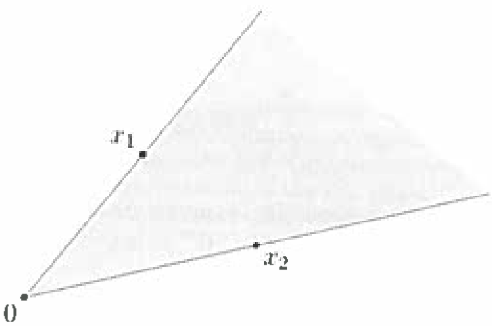

Figure 2.4 The pie slice shows all points of the form $\theta_1 x_1 + \theta_2 x_2$, where $\theta_1,\ \theta_2 \geq 0$. The apex of the slice (which corresponds to $\theta_1 = \theta_2 = 0$) is at 0; its edges (which correspond to $\theta_1 = 0$ or $\theta_2 = 0$) pass through the points x_1 and x_2.

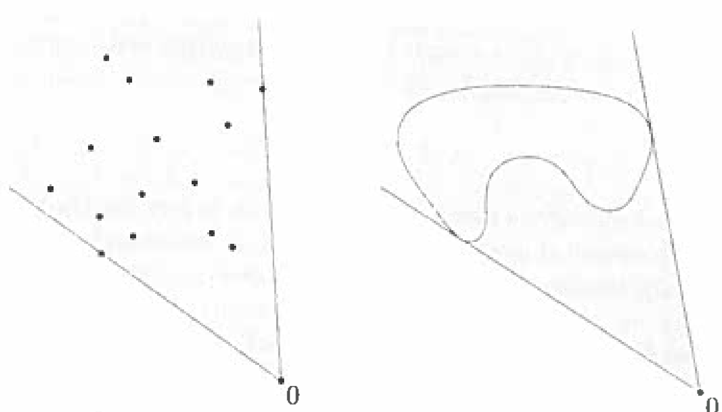

Figure 2.5 The conic hulls (shown shaded) of the two sets of figure 2.3.

2.2 Some important examples

In this section we describe some important examples of convex sets which we will encounter throughout the rest of the book. We start with some simple examples.

- The empty set $\emptyset$, any single point (*i.e.*, singleton) $\{x_0\}$, and the whole space $\mathbf{R}^n$ are affine (hence, convex) subsets of $\mathbf{R}^n$.
- Any line is affine. If it passes through zero, it is a subspace, hence also a convex cone.
- A line segment is convex, but not affine (unless it reduces to a point).
- A *ray*, which has the form $\{x_0 + \theta v \mid \theta \geq 0\}$, where $v \neq 0$, is convex, but not affine. It is a convex cone if its base x_0 is 0.
- Any subspace is affine, and a convex cone (hence convex).

2.2.1 Hyperplanes and halfspaces

A *hyperplane* is a set of the form

$$\{x \mid a^T x = b\},$$

where $a \in \mathbf{R}^n$, $a \neq 0$, and $b \in \mathbf{R}$. Analytically it is the solution set of a nontrivial linear equation among the components of x (and hence an affine set). Geometrically, the hyperplane $\{x \mid a^T x = b\}$ can be interpreted as the set of points with a constant inner product to a given vector a, or as a hyperplane with *normal vector* a; the constant $b \in \mathbf{R}$ determines the offset of the hyperplane from the origin. This geometric interpretation can be understood by expressing the hyperplane in the form

$$\{x \mid a^T(x - x_0) = 0\},$$

where x_0 is any point in the hyperplane (*i.e.*, any point that satisfies $a^T x_0 = b$). This representation can in turn be expressed as

$$\{x \mid a^T(x - x_0) = 0\} = x_0 + a^\perp,$$

where $a^\perp$ denotes the orthogonal complement of a, *i.e.*, the set of all vectors orthogonal to it:

$$a^\perp = \{v \mid a^T v = 0\}.$$

This shows that the hyperplane consists of an offset x_0, plus all vectors orthogonal to the (normal) vector a. These geometric interpretations are illustrated in figure 2.6.

A hyperplane divides $\mathbf{R}^n$ into two *halfspaces*. A (closed) halfspace is a set of the form

$$\{x \mid a^T x \leq b\}, \tag{2.1}$$

where $a \neq 0$, *i.e.*, the solution set of one (nontrivial) linear inequality. Halfspaces are convex, but not affine. This is illustrated in figure 2.7.

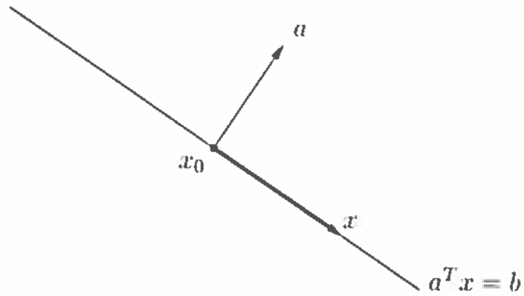

Figure 2.6 Hyperplane in $\mathbf{R}^2$, with normal vector a and a point x_0 in the hyperplane. For any point x in the hyperplane, $x - x_0$ (shown as the darker arrow) is orthogonal to a.

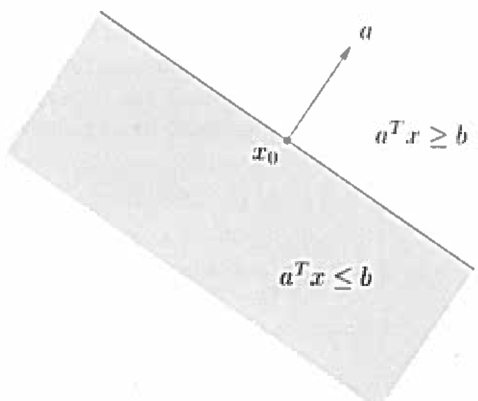

Figure 2.7 A hyperplane defined by $a^T x = b$ in $\mathbf{R}^2$ determines two halfspaces. The halfspace determined by $a^T x \geq b$ (not shaded) is the halfspace extending in the direction a. The halfspace determined by $a^T x \leq b$ (which is shown shaded) extends in the direction $-a$. The vector a is the outward normal of this halfspace.

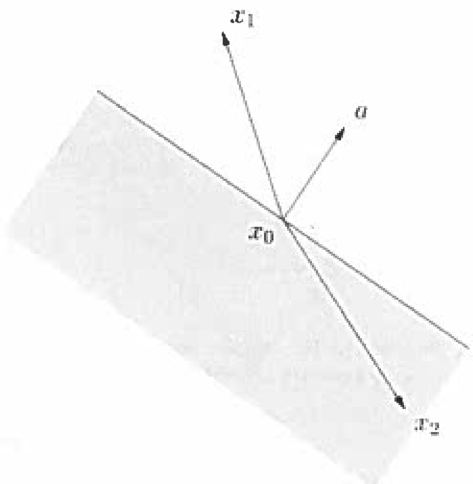

Figure 2.8 The shaded set is the halfspace determined by $a^T(x - x_0) \leq 0$. The vector $x_1 - x_0$ makes an acute angle with a, so x_1 is not in the halfspace. The vector $x_2 - x_0$ makes an obtuse angle with a, and so is in the halfspace.

The halfspace (2.1) can also be expressed as

$$\{x \mid a^T(x - x_0) \leq 0\}, \tag{2.2}$$

where x_0 is any point on the associated hyperplane, *i.e.*, satisfies $a^T x_0 = b$. The representation (2.2) suggests a simple geometric interpretation: the halfspace consists of x_0 plus any vector that makes an obtuse (or right) angle with the (outward normal) vector a. This is illustrated in figure 2.8.

The boundary of the halfspace (2.1) is the hyperplane $\{x \mid a^T x = b\}$. The set $\{x \mid a^T x < b\}$, which is the interior of the halfspace $\{x \mid a^T x \leq b\}$, is called an *open halfspace*.

2.2.2 Euclidean balls and ellipsoids

A *(Euclidean) ball* (or just ball) in $\mathbf{R}^n$ has the form

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x \mid (x - x_c)^T(x - x_c) \leq r^2\},$$

where $r > 0$, and $\|\cdot\|_2$ denotes the Euclidean norm, *i.e.*, $\|u\|_2 = (u^T u)^{1/2}$. The vector x_c is the *center* of the ball and the scalar r is its *radius*; $B(x_c, r)$ consists of all points within a distance r of the center x_c. Another common representation for the Euclidean ball is

$$B(x_c, r) = \{x_c + ru \mid \|u\|_2 \leq 1\}.$$

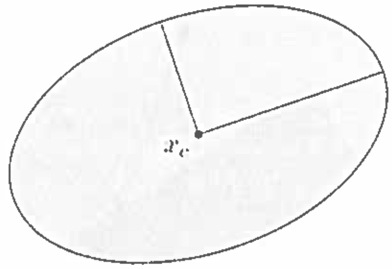

Figure 2.9 An ellipsoid in $\mathbf{R}^2$, shown shaded. The center x_c is shown as a dot, and the two semi-axes are shown as line segments.

A Euclidean ball is a convex set: if $\|x_1 - x_c\|_2 \leq r$, $\|x_2 - x_c\|_2 \leq r$, and $0 \leq \theta \leq 1$, then

$$\begin{aligned} \|\theta x_1 + (1-\theta)x_2 - x_c\|_2 &= \|\theta(x_1 - x_c) + (1-\theta)(x_2 - x_c)\|_2 \\ &\leq \theta\|x_1 - x_c\|_2 + (1-\theta)\|x_2 - x_c\|_2 \\ &\leq r. \end{aligned}$$

(Here we use the homogeneity property and triangle inequality for $\|\cdot\|_2$; see §A.1.2.)

A related family of convex sets is the *ellipsoids*, which have the form

$$\mathcal{E} = \{x \mid (x - x_c)^T P^{-1}(x - x_c) \leq 1\}, \tag{2.3}$$

where $P = P^T \succ 0$, *i.e.*, P is symmetric and positive definite. The vector $x_c \in \mathbf{R}^n$ is the *center* of the ellipsoid. The matrix P determines how far the ellipsoid extends in every direction from x_c; the lengths of the semi-axes of $\mathcal{E}$ are given by $\sqrt{\lambda_i}$, where λ_i are the eigenvalues of P. A ball is an ellipsoid with $P = r^2 I$. Figure 2.9 shows an ellipsoid in $\mathbf{R}^2$.

Another common representation of an ellipsoid is

$$\mathcal{E} = \{x_c + Au \mid \|u\|_2 \leq 1\}, \tag{2.4}$$

where A is square and nonsingular. In this representation we can assume without loss of generality that A is symmetric and positive definite. By taking $A = P^{1/2}$, this representation gives the ellipsoid defined in (2.3). When the matrix A in (2.4) is symmetric positive semidefinite but singular, the set in (2.4) is called a *degenerate ellipsoid*; its affine dimension is equal to the rank of A. Degenerate ellipsoids are also convex.

2.2.3 Norm balls and norm cones

Suppose $\|\cdot\|$ is any norm on $\mathbf{R}^n$ (see §A.1.2). From the general properties of norms it can be shown that a *norm ball* of radius r and center x_c, given by $\{x \mid \|x - x_c\| \leq r\}$, is convex. The *norm cone* associated with the norm $\|\cdot\|$ is the set

$$C = \{(x, t) \mid \|x\| \leq t\} \subseteq \mathbf{R}^{n+1}.$$

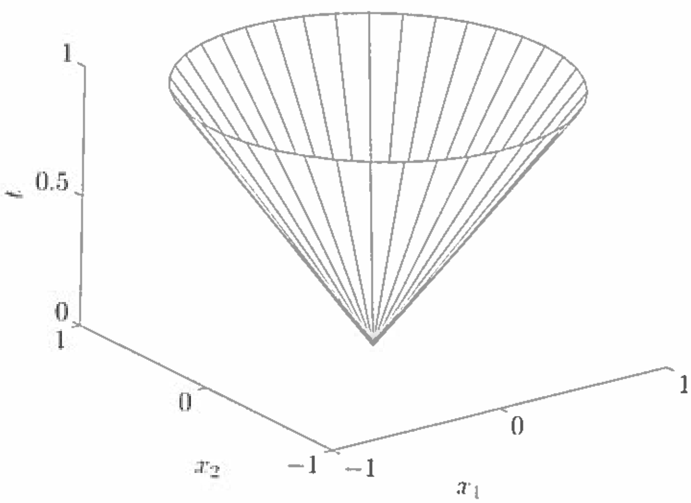

Figure 2.10 Boundary of second-order cone in $\mathbf{R}^3$, $\{(x_1, x_2, t) \mid (x_1^2+x_2^2)^{1/2} \le t\}$.

It is (as the name suggests) a convex cone.

Example 2.3 *The second-order cone* is the norm cone for the Euclidean norm, *i.e.*,

$$\begin{aligned} C &= \{(x,t) \in \mathbf{R}^{n+1} \mid \|x\|_2 \le t\} \\ &= \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \;\middle|\; \begin{bmatrix} x \\ t \end{bmatrix}^T \begin{bmatrix} I & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \le 0,\ t \ge 0 \right\}. \end{aligned}$$

The second-order cone is also known by several other names. It is called the *quadratic cone*, since it is defined by a quadratic inequality. It is also called the *Lorentz cone* or *ice-cream cone*. Figure 2.10 shows the second-order cone in $\mathbf{R}^3$.

2.2.4 Polyhedra

A *polyhedron* is defined as the solution set of a finite number of linear equalities and inequalities:

$$\mathcal{P} = \{x \mid a_j^T x \le b_j,\ j = 1, \ldots, m,\ c_j^T x = d_j,\ j = 1, \ldots, p\}. \tag{2.5}$$

A polyhedron is thus the intersection of a finite number of halfspaces and hyperplanes. Affine sets (*e.g.*, subspaces, hyperplanes, lines), rays, line segments, and halfspaces are all polyhedra. It is easily shown that polyhedra are convex sets. A bounded polyhedron is sometimes called a *polytope*, but some authors use the opposite convention (*i.e.*, polytope for any set of the form (2.5), and polyhedron

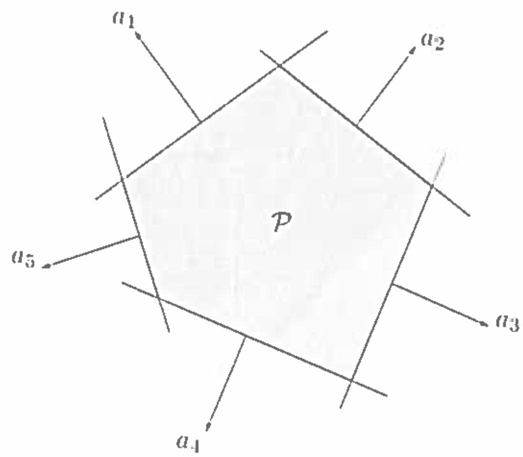

Figure 2.11 The polyhedron $\mathcal{P}$ (shown shaded) is the intersection of five halfspaces, with outward normal vectors $a_1, \dots, a_5$.

when it is bounded). Figure 2.11 shows an example of a polyhedron defined as the intersection of five halfspaces.

It will be convenient to use the compact notation

$$\mathcal{P} = \{x \mid Ax \preceq b, \ Cx = d\} \tag{2.6}$$

for (2.5), where

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}, \qquad C = \begin{bmatrix} c_1^T \\ \vdots \\ c_p^T \end{bmatrix},$$

and the symbol $\preceq$ denotes *vector inequality* or *componentwise inequality* in $\mathbf{R}^m$: $u \preceq v$ means $u_i \leq v_i$ for $i = 1, \dots, m$.

Example 2.4 The *nonnegative orthant* is the set of points with nonnegative components, *i.e.*,

$$\mathbf{R}_+^n = \{x \in \mathbf{R}^n \mid x_i \geq 0, \ i = 1, \dots, n\} = \{x \in \mathbf{R}^n \mid x \succeq 0\}.$$

(Here $\mathbf{R}_+$ denotes the set of nonnegative numbers: $\mathbf{R}_+ = \{x \in \mathbf{R} \mid x \geq 0\}$.) The nonnegative orthant is a polyhedron and a cone (and therefore called a *polyhedral cone*).

Simplexes

Simplexes are another important family of polyhedra. Suppose the $k+1$ points $v_0, \dots, v_k \in \mathbf{R}^n$ are *affinely independent*, which means $v_1 - v_0, \dots, v_k - v_0$ are linearly independent. The simplex determined by them is given by

$$C = \mathbf{conv}\{v_0, \dots, v_k\} = \{\theta_0 v_0 + \cdots + \theta_k v_k \mid \theta \succeq 0, \ \mathbf{1}^T\theta = 1\}, \tag{2.7}$$

where $\mathbf{1}$ denotes the vector with all entries one. The affine dimension of this simplex is k, so it is sometimes referred to as a k-dimensional simplex in $\mathbf{R}^n$.

Example 2.5 *Some common simplexes.* A 1-dimensional simplex is a line segment; a 2-dimensional simplex is a triangle (including its interior); and a 3-dimensional simplex is a tetrahedron.

The *unit simplex* is the n-dimensional simplex determined by the zero vector and the unit vectors, *i.e.*, $0, e_1, \ldots, e_n \in \mathbf{R}^n$. It can be expressed as the set of vectors that satisfy

$$x \succeq 0, \qquad \mathbf{1}^T x \leq 1.$$

The *probability simplex* is the $(n-1)$-dimensional simplex determined by the unit vectors $e_1, \ldots, e_n \in \mathbf{R}^n$. It is the set of vectors that satisfy

$$x \succeq 0, \qquad \mathbf{1}^T x = 1.$$

Vectors in the probability simplex correspond to probability distributions on a set with n elements, with x_i interpreted as the probability of the ith element.

To describe the simplex (2.7) as a polyhedron, *i.e.*, in the form (2.6), we proceed as follows. By definition, $x \in C$ if and only if $x = \theta_0 v_0 + \theta_1 v_1 + \cdots + \theta_k v_k$ for some $\theta \succeq 0$ with $\mathbf{1}^T \theta = 1$. Equivalently, if we define $y = (\theta_1, \ldots, \theta_k)$ and

$$B = \begin{bmatrix} v_1 - v_0 & \cdots & v_k - v_0 \end{bmatrix} \in \mathbf{R}^{n \times k},$$

we can say that $x \in C$ if and only if

$$x = v_0 + By \tag{2.8}$$

for some $y \succeq 0$ with $\mathbf{1}^T y \leq 1$. Now we note that affine independence of the points $v_0, \ldots, v_k$ implies that the matrix B has rank k. Therefore there exists a nonsingular matrix $A = (A_1, A_2) \in \mathbf{R}^{n \times n}$ such that

$$AB = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} B = \begin{bmatrix} I \\ 0 \end{bmatrix}.$$

Multiplying (2.8) on the left with A, we obtain

$$A_1 x = A_1 v_0 + y, \qquad A_2 x = A_2 v_0.$$

From this we see that $x \in C$ if and only if $A_2 x = A_2 v_0$, and the vector $y = A_1 x - A_1 v_0$ satisfies $y \succeq 0$ and $\mathbf{1}^T y \leq 1$. In other words we have $x \in C$ if and only if

$$A_2 x = A_2 v_0, \qquad A_1 x \succeq A_1 v_0, \qquad \mathbf{1}^T A_1 x \leq 1 + \mathbf{1}^T A_1 v_0,$$

which is a set of linear equalities and inequalities in x, and so describes a polyhedron.

Convex hull description of polyhedra

The convex hull of the finite set $\{v_1, \ldots, v_k\}$ is

$$\mathbf{conv}\{v_1, \ldots, v_k\} = \{\theta_1 v_1 + \cdots + \theta_k v_k \mid \theta \succeq 0,\ \mathbf{1}^T\theta = 1\}.$$

This set is a polyhedron, and bounded, but (except in special cases, *e.g.*, a simplex) it is not simple to express it in the form (2.5), *i.e.*, by a set of linear equalities and inequalities.

A generalization of this convex hull description is

$$\{\theta_1 v_1 + \cdots + \theta_k v_k \mid \theta_1 + \cdots + \theta_m = 1,\ \theta_i \geq 0,\ i = 1, \ldots, k\}, \tag{2.9}$$

where $m \leq k$. Here we consider nonnegative linear combinations of v_i, but only the first m coefficients are required to sum to one. Alternatively, we can interpret (2.9) as the convex hull of the points $v_1, \ldots, v_m$, plus the conic hull of the points $v_{m+1}, \ldots, v_k$. The set (2.9) defines a polyhedron, and conversely, every polyhedron can be represented in this form (although we will not show this).

The question of how a polyhedron is represented is subtle, and has very important practical consequences. As a simple example consider the unit ball in the ℓ_∞-norm in $\mathbf{R}^n$,

$$C = \{x \mid |x_i| \leq 1,\ i = 1, \ldots, n\}.$$

The set C can be described in the form (2.5) with $2n$ linear inequalities $\pm e_i^T x \leq 1$, where e_i is the ith unit vector. To describe it in the convex hull form (2.9) requires at least 2^n points:

$$C = \mathbf{conv}\{v_1, \ldots, v_{2^n}\},$$

where $v_1, \ldots, v_{2^n}$ are the 2^n vectors all of whose components are 1 or -1. Thus the size of the two descriptions differs greatly, for large n.

2.2.5 The positive semidefinite cone

We use the notation $\mathbf{S}^n$ to denote the set of symmetric $n \times n$ matrices,

$$\mathbf{S}^n = \{X \in \mathbf{R}^{n \times n} \mid X = X^T\},$$

which is a vector space with dimension $n(n+1)/2$. We use the notation $\mathbf{S}^n_+$ to denote the set of symmetric positive semidefinite matrices:

$$\mathbf{S}^n_+ = \{X \in \mathbf{S}^n \mid X \succeq 0\},$$

and the notation $\mathbf{S}^n_{++}$ to denote the set of symmetric positive definite matrices:

$$\mathbf{S}^n_{++} = \{X \in \mathbf{S}^n \mid X \succ 0\}.$$

(This notation is meant to be analogous to $\mathbf{R}_+$, which denotes the nonnegative reals, and $\mathbf{R}_{++}$, which denotes the positive reals.)

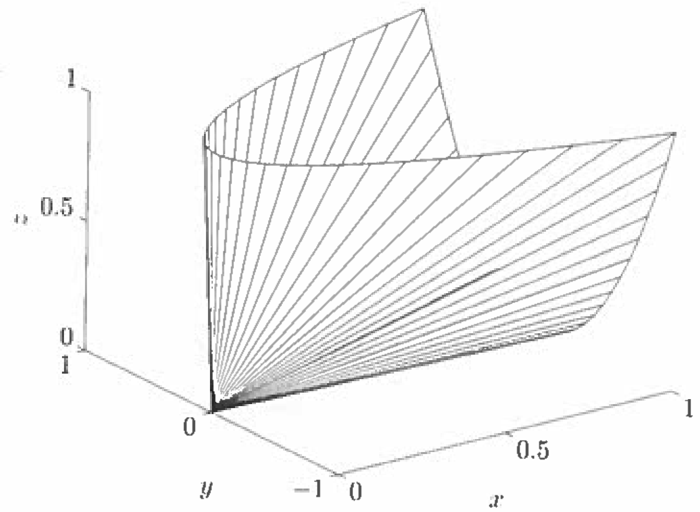

Figure 2.12 Boundary of positive semidefinite cone in $\mathbf{S}^2$.

The set $\mathbf{S}_+^n$ is a convex cone: if $\theta_1, \theta_2 \geq 0$ and A, $B \in \mathbf{S}_+^n$, then $\theta_1 A + \theta_2 B \in \mathbf{S}_+^n$. This can be seen directly from the definition of positive semidefiniteness: for any $x \in \mathbf{R}^n$, we have

$$x^T(\theta_1 A + \theta_2 B)x = \theta_1 x^T A x + \theta_2 x^T B x \geq 0,$$

if $A \succeq 0$, $B \succeq 0$ and θ_1, $\theta_2 \geq 0$.

Example 2.6 *Positive semidefinite cone in* $\mathbf{S}^2$. We have

$$X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbf{S}_+^2 \iff x \geq 0, \quad z \geq 0, \quad xz \geq y^2.$$

The boundary of this cone is shown in figure 2.12, plotted in $\mathbf{R}^3$ as (x, y, z).

2.3 Operations that preserve convexity

In this section we describe some operations that preserve convexity of sets, or allow us to construct convex sets from others. These operations, together with the simple examples described in §2.2, form a calculus of convex sets that is useful for determining or establishing convexity of sets.