

18.5 Cones and Farkas's Lemma

Chapter 3 answered one question about a linear system completely: $\mathbf{Ax} = \mathbf{b}$ has a solution exactly when $\mathbf{b}$ lies in the column space of $\mathbf{A}$ (Theorem 3.1.1). Its applications section then hit a wall. Traffic flows cannot be negative, and once the unknowns are required to satisfy $\mathbf{x} \geq \mathbf{0}$, the column space is the wrong set to look in. This section finds the right set, which is a **cone**, and answers the new question with the separation theorems of §03: either $\mathbf{Ax} = \mathbf{b}$ has a nonnegative solution, or a single vector $\mathbf{y}$ certifies that it has none. That is Farkas's lemma, and the next two sections run on it.

One step is harder than it looks. Separation needs a **closed** set, and nothing proved so far says that the relevant cone is closed. Most of the work goes into that step, and the last part of the section shows, with an explicit cone, that it can really fail.

Throughout, the field is $\mathbb{R}$. Vectors in $\mathbb{R}^n$ are columns, $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^\top \mathbf{x}$ is the standard inner product, and $\|\cdot\|$ is the Euclidean norm it induces, the norm that Theorem 18.3.1 uses. Convergence $\mathbf{y}_k \rightarrow \mathbf{y}$ means $\|\mathbf{y}_k - \mathbf{y}\| \rightarrow 0$, and “closed” is meant in the sequential sense of Definition 18.1.6: a set S is closed when every limit of a convergent sequence of points of S again lies in S . Convergence in this norm is entrywise convergence, by the elementary bound

$$|y_i| \leq \|\mathbf{y}\| \leq |y_1| + \cdots + |y_n| \quad \text{for } \mathbf{y} \in \mathbb{R}^n \text{ and } 1 \leq i \leq n. \quad (18.5.1)$$

The left inequality holds because $y_i^2 \leq y_1^2 + \cdots + y_n^2$, and the right one is the triangle inequality for $\mathbf{y} = y_1 \mathbf{e}_1 + \cdots + y_n \mathbf{e}_n$. Applied to $\mathbf{y}_k - \mathbf{y}$, with the algebra of limits of Chapter 16's introduction on the right, it shows that $\|\mathbf{y}_k - \mathbf{y}\| \rightarrow 0$ exactly when every entry of $\mathbf{y}_k - \mathbf{y}$ tends to 0.

18.5.1 Convex cones

For a vector with real entries we write $\mathbf{x} \geq \mathbf{0}$ when **every entry** of $\mathbf{x}$ is ≥ 0 , and $\mathbf{x} > \mathbf{0}$ when every entry is > 0 . The set

$$\mathbb{R}_{\geq 0}^m = \{\mathbf{x} \in \mathbb{R}^m : \mathbf{x} \geq \mathbf{0}\}$$

is the **nonnegative orthant**. The relation $\mathbf{x} \geq \mathbf{0}$ is an entrywise statement about a vector, and it has nothing to do with the Loewner order $\mathbf{A} \succeq \mathbf{O}$ of Chapter 13, which is a statement about a matrix.

Here is the question. Let $\mathbf{A} \in M_{m \times n}(\mathbb{R})$ have columns $\mathbf{a}_1, \dots, \mathbf{a}_n$. By Theorem 0.9.3, $\mathbf{Ax} = x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n$, so

$$\{\mathbf{b} \in \mathbb{R}^m : \mathbf{Ax} = \mathbf{b} \text{ for some } \mathbf{x} \geq \mathbf{0}\} = \{x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n : \text{every } x_j \geq 0\}.$$

Without the sign condition this set would be the span of the columns, a subspace. With it, the set is closed under addition and under multiplication by **nonnegative** scalars, but not under multiplication by -1 . Such sets turn up whenever quantities cannot be negative, so they get a name.

A cone is a set you may stretch and add in, but never reflect.

Definition 18.5.1. Convex Cone

Let V be a vector space over $\mathbb{R}$. A subset $K \subseteq V$ is a **convex cone** if K is **nonempty** and

$$au + bv \in K \quad \text{for all } u, v \in K \text{ and all real } a, b \geq 0.$$

In words: a convex cone contains every **nonnegative** combination of two of its members, and hence, by induction on the number of terms, every nonnegative combination of finitely many of its members. Three special cases unpack the single condition. Taking $a = b = 0$ shows $0 \in K$, so every convex cone contains the origin. Taking $b = 0$ shows that K is closed under multiplication by nonnegative scalars. Taking $a = b = 1$ shows that it is closed under addition, and conversely a nonempty set with these two closure properties is a convex cone.

A convex cone is a convex set in the sense of Definition 18.1.1: take $a = 1 - t$ and $b = t$ with $0 \leq t \leq 1$. The converse fails, and the next examples show how.

Examples. Each is checked against the one condition of the definition.

- **The orthant.** If $\mathbf{u}, \mathbf{v} \geq \mathbf{0}$ and $a, b \geq 0$, then every entry of $a\mathbf{u} + b\mathbf{v}$ is a sum of products of nonnegative numbers, so $a\mathbf{u} + b\mathbf{v} \geq \mathbf{0}$. Hence $\mathbb{R}_{\geq 0}^m$ is a convex cone.
- **A half-space through the origin.** Fix $\mathbf{a} \in \mathbb{R}^n$ and let $H = \{\mathbf{x} : \langle \mathbf{x}, \mathbf{a} \rangle \geq 0\}$. If $\mathbf{u}, \mathbf{v} \in H$ and $a, b \geq 0$, then $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{a} \rangle = a\langle \mathbf{u}, \mathbf{a} \rangle + b\langle \mathbf{v}, \mathbf{a} \rangle \geq 0$. So H is a convex cone. When $\mathbf{a} \neq \mathbf{0}$ it is one of the two closed half-spaces bounded by the hyperplane $\mathbf{a}^\perp$.
- **The ice-cream cone.** In $\mathbb{R}^{n+1}$, write vectors as $(\mathbf{x}, t)$ with $\mathbf{x} \in \mathbb{R}^n$ and $t \in \mathbb{R}$, and let $L = \{(\mathbf{x}, t) : \|\mathbf{x}\|_2 \leq t\}$. If $(\mathbf{x}, t), (\mathbf{y}, s) \in L$ and $a, b \geq 0$, the triangle inequality and homogeneity of the norm (Definition 16.1.1) give $\|a\mathbf{x} + b\mathbf{y}\|_2 \leq a\|\mathbf{x}\|_2 + b\|\mathbf{y}\|_2 \leq at + bs$. So L is a convex cone. In $\mathbb{R}^3$ it is the round solid cone that gives the family its name.
- **Positive semidefinite matrices.** In the real vector space of real symmetric $n \times n$ matrices, the set of $\mathbf{A} \succeq \mathbf{0}$ is a convex cone. This is part (a) of Exercise (C2: The cone of positive matrices). Chapter 13 used the phrase there, and this is the definition it was using.
- **Degenerate cases.** $\{0\}$ and V are convex cones. More generally, **every subspace is a convex cone**, since it is closed under all linear combinations and in particular the nonnegative ones. A subspace is the extreme case of a cone that happens to be closed under reflection as well.

Nonexamples by minimal change. In $\mathbb{R}^2$, take the union of two opposite quadrants, $Q = \{\mathbf{x} : x_1x_2 \geq 0\}$. It is closed under nonnegative scaling, since $(ax_1)(ax_2) = a^2x_1x_2$. The clause that fails is addition: $(1, 0)$ and $(0, -1)$ lie in Q , but their sum $(1, -1)$ does not. Or keep convexity and lose the scaling: the closed unit disc is convex, but $2\mathbf{e}_1$ is not in it. Or translate the orthant to $\mathbf{1} + \mathbb{R}_{\geq 0}^2$: it is convex and closed under addition, but it does not contain $\mathbf{0}$.

Why this definition. The scalars must be nonnegative. With all real scalars allowed, the definition becomes the definition of a subspace, and the sign information that started this section would be lost. The requirement $0 \in K$ is a convention that some authors drop: they call $\{\mathbf{x} : \mathbf{x} > \mathbf{0}\}$ a cone as well. Including the origin costs nothing here and keeps the statements below free of exceptions.

Warning. A convex cone need not be closed. In $\mathbb{R}^2$ let

$$C = \{(x, y) : x > 0\} \cup \{(0, 0)\}.$$

It is a convex cone. Let $\mathbf{u}, \mathbf{v} \in C$ and $a, b \geq 0$. The first entry of $a\mathbf{u} + b\mathbf{v}$ is $au_1 + bv_1 \geq 0$. If it is positive, the vector lies in C . If it is zero, then $au_1 = 0$ and $bv_1 = 0$; since the only member of C with first entry 0 is $\mathbf{0}$, this forces $a\mathbf{u} = \mathbf{0}$ and $b\mathbf{v} = \mathbf{0}$, and the sum is $\mathbf{0} \in C$.

But $(1/k, 1) \in C$ for every $k \geq 1$, while the limit $(0, 1)$ is not in C . Separating a point from a cone needs the cone to be closed, so closedness must be proved. This particular cone comes back at the end of the section as the image of a closed cone.

The cones in the question at the start of this section are the simplest kind.

Definition 18.5.2. Finitely Generated Cone

Let $a_1, \dots, a_m$ be vectors in a real vector space V . The **cone generated by** them is

$$\text{cone}(a_1, \dots, a_m) = \left\{ \sum_{i=1}^m x_i a_i : x_1, \dots, x_m \geq 0 \right\}.$$

A cone of this form is **finitely generated**. For the empty list we set $\text{cone}() = \{0\}$, the value of the empty sum. When $V = \mathbb{R}^n$ and $\mathbf{A}$ is the matrix with columns $\mathbf{a}_1, \dots, \mathbf{a}_m$, this is $\{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{R}_{\geq 0}^m\}$, and we also write it $\text{cone}(\mathbf{A})$.

It is a convex cone: a nonnegative combination of two nonnegative combinations of the a_i is again one, with coefficients $ax_i + bx'_i \geq 0$. It is also the **smallest** convex cone containing $a_1, \dots, a_m$, since any convex cone containing the a_i contains all their nonnegative combinations, by the remark after Definition 18.5.1. The orthant is $\text{cone}(\mathbf{e}_1, \dots, \mathbf{e}_m)$. A subspace with basis $v_1, \dots, v_k$ is $\text{cone}(v_1, -v_1, \dots, v_k, -v_k)$, since cv_i is cv_i or $|c|(-v_i)$ according to the sign of c . The ice-cream cone is **not** finitely generated, a fact we shall neither need nor prove.

Quick check. Is $\text{cone}((1, 0), (1, 1), (0, 1))$ the same as $\text{cone}((1, 0), (0, 1))$? Is $\text{cone}((1, 0), (-1, 0))$ a line or a half-line?

Solution. Both cones are the quadrant $\mathbb{R}_{\geq 0}^2$. The second list's cone is the quadrant because $x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 = (x_1, x_2)$. The first contains the second, and every generator of the first lies in the quadrant, which is a convex cone, so the first is contained in the quadrant too. A redundant generator changes nothing. The cone of $(1, 0)$ and $(-1, 0)$ is the whole x -axis $\{(t, 0) : t \in \mathbb{R}\}$, a line: for $t \geq 0$ use $t(1, 0)$, and for $t < 0$ use $|t|(-1, 0)$.

18.5.2 The dual cone

A certificate that $\mathbf{A}\mathbf{x} = \mathbf{b}$ has no nonnegative solution will be a vector $\mathbf{y}$ that makes a non-negative inner product with every column of $\mathbf{A}$ but a negative one with $\mathbf{b}$. Every nonnegative combination of the columns then also has a nonnegative inner product with $\mathbf{y}$, so $\mathbf{b}$ cannot be such a combination. The set of vectors $\mathbf{y}$ that pass the first test deserves a name.

The dual cone of K is the set of directions that make an angle of at most 90° with everything in K .

Definition 18.5.3. Dual Cone

Let V be a finite-dimensional real inner product space and let $S \subseteq V$ be **any** subset. The **dual cone** of S is

$$S^* = \{y \in V : \langle x, y \rangle \geq 0 \text{ for every } x \in S\}.$$

We write S^{**} for $(S^*)^*$.

In words: $y \in S^*$ when the linear functional $\langle \cdot, y \rangle$ is nonnegative on the whole of S , that is, when S lies in the closed half-space $\{x : \langle x, y \rangle \geq 0\}$. The inner product is what places S^* inside V rather than in the dual space of Chapter 5 (Theorem 5.1.1 is the case $V = \mathbb{R}^n$).

Warning. S^* is **not** the dual space. The star here produces a subset of V itself, not a space of functionals, and it is defined for any subset, not only for a subspace. For a subspace U the dual cone is the orthogonal complement: if $\langle u, y \rangle \geq 0$ for every $u \in U$, then applying this to $-u \in U$ gives $\langle u, y \rangle \leq 0$ as well, so $\langle u, y \rangle = 0$, and the converse is immediate. So for a subspace we always write $U^\perp$, never U^* .

The basic properties take one line each, and they explain why every dual cone is a good set to separate from.

Proposition 18.5.1. Basic Properties of the Dual Cone

Let V be a finite-dimensional real inner product space and let $S, T \subseteq V$.

- (a) S^* is a closed convex cone.
- (b) If $S \subseteq T$, then $T^* \subseteq S^*$.
- (c) $S \subseteq S^{**}$.
- (d) $\text{cone}(a_1, \dots, a_m)^* = \{y : \langle a_i, y \rangle \geq 0 \text{ for } i = 1, \dots, m\}$. In $\mathbb{R}^n$, with $\mathbf{A}$ the matrix of columns $\mathbf{a}_i$, this is $\{\mathbf{y} : \mathbf{A}^\top \mathbf{y} \geq \mathbf{0}\}$.

Idea. S^* is the intersection of the sets $\{y : \langle x, y \rangle \geq 0\}$, one for each $x \in S$. Parts (a), (b) and (d) read off this description (more members of S mean more conditions, and for a cone the generators suffice), and (c) is the symmetry of the inner product.

Proof.

- (a) $0 \in S^*$, so S^* is nonempty. If $y, z \in S^*$ and $a, b \geq 0$, then $\langle x, ay + bz \rangle = a\langle x, y \rangle + b\langle x, z \rangle \geq 0$ for every $x \in S$, using linearity in the second slot, which holds because the field is $\mathbb{R}$. Finally $S^* = \bigcap_{x \in S} \{y : \langle x, y \rangle \geq 0\}$ is an intersection of closed half-spaces (all of V when $x = 0$), closed by Proposition 18.1.6 (c) and (b), since $\langle x, \cdot \rangle$ is a linear functional; if $S = \emptyset$, then $S^* = V$, which is closed.
- (b) A vector that is nonnegative against every member of T is in particular nonnegative against every member of S .
- (c) If $x \in S$, then $\langle x, y \rangle \geq 0$ for every $y \in S^*$, by the definition of S^* . Since the inner product is symmetric over $\mathbb{R}$, this says $\langle y, x \rangle \geq 0$ for every $y \in S^*$, which is $x \in S^{**}$.
- (d) The generators lie in the cone, so (b) gives $\subseteq$. Conversely, if $\langle a_i, y \rangle \geq 0$ for every i , then $\langle \sum x_i a_i, y \rangle = \sum x_i \langle a_i, y \rangle \geq 0$ whenever every $x_i \geq 0$. The matrix form follows because the i -th entry of $\mathbf{A}^\top \mathbf{y}$ is $\mathbf{a}_i^\top \mathbf{y} = \langle \mathbf{a}_i, \mathbf{y} \rangle$.

■

Part (a) holds for **every** set S , including one that is neither convex nor closed. Part (c) is only an inclusion, and the question of when it is an equality is really the question of this section.

Examples.

- **The orthant is its own dual.** By (d), $(\mathbb{R}_{\geq 0}^m)^* = \text{cone}(\mathbf{e}_1, \dots, \mathbf{e}_m)^* = \{\mathbf{y} : \langle \mathbf{e}_i, \mathbf{y} \rangle = y_i \geq 0 \text{ for all } i\} = \mathbb{R}_{\geq 0}^m$.
- **A single generator.** $\text{cone}(\mathbf{a})^* = \{\mathbf{y} : \langle \mathbf{a}, \mathbf{y} \rangle \geq 0\}$, a closed half-space when $\mathbf{a} \neq \mathbf{0}$. A thin cone has a fat dual.
- **The degenerate cases.** The dual cone of $\{0\}$ is all of V , and the dual cone of V is $\{0\}$, because a y in it must satisfy $\langle -y, y \rangle \geq 0$, that is $\|y\|^2 \leq 0$. As subspaces, by the warning, we write these as $\{0\}^\perp = V$ and $V^\perp = \{0\}$.
- **A wedge in the plane.** For $K = \text{cone}((1, 0), (1, 1))$, part (d) gives $K^* = \{\mathbf{y} : y_1 \geq 0, y_1 + y_2 \geq 0\}$. This is the wedge between the directions $(0, 1)$ and $(1, -1)$, each of which is perpendicular to one edge of K . [Exercise \(B3: A dual cone and its dual\)](#) finishes this computation.

18.5.3 Finitely generated cones are closed

Now the key step. We want to separate a point $\mathbf{b}$ from $\text{cone}(\mathbf{A})$, and §03 separates points only from **closed** convex sets. So we must show that $\text{cone}(\mathbf{A})$ is closed.

Here is why that is not obvious. Suppose $\mathbf{A}\mathbf{x}_k \rightarrow \mathbf{b}$ with every $\mathbf{x}_k \geq \mathbf{0}$. The obvious move is to take a limit of the $\mathbf{x}_k$, but they need not converge, or even be bounded. If the columns of $\mathbf{A}$ are dependent, then $\mathbf{A}\mathbf{x}_k$ can stay put while $\mathbf{x}_k$ runs off along the null space. The cure has two parts. First, throw away the dependence, generator by generator, which is the “count, then drop a point” move of Theorem 18.2.2 adapted to cones. Second, once the columns are independent, prove that $\mathbf{x}_k$ is recovered from $\mathbf{A}\mathbf{x}_k$ by a fixed linear map, so that the $\mathbf{x}_k$ cannot escape.

Theorem 18.5.2. Carathéodory's Theorem for Cones

Let V be a real vector space and let $a_1, \dots, a_m \in V$. Every $b \in \text{cone}(a_1, \dots, a_m)$ lies in $\text{cone}(a_i : i \in S)$ for some set of indices $S \subseteq \{1, \dots, m\}$ such that the list $(a_i)_{i \in S}$ is **linearly independent**. In particular, if $\dim V = n$, then b is a nonnegative combination of at most n of the a_i .

Idea. Take a representation of b that uses as few generators as possible. If those generators were dependent, a dependence relation could be added to the representation, with a scalar multiple chosen so that one coefficient reaches zero and none goes negative. That removes a generator, which contradicts minimality. It is the move of Theorem 18.2.2 with a *linear* dependence in place of an affine one, and that difference is why the bound is n rather than $n + 1$.

Proof. By [Definition 18.5.2](#), $b = \sum_{i=1}^m x_i a_i$ with every $x_i \geq 0$. Discarding the terms with $x_i = 0$, we can write $b = \sum_{i \in S} x_i a_i$ with $x_i > 0$ for every $i \in S$, for some $S \subseteq \{1, \dots, m\}$. The empty set is allowed, when $b = 0$. Among all such representations choose one with $|S|$ as small as possible. We claim that $(a_i)_{i \in S}$ is linearly independent.

Suppose not. Then there are reals c_i , for $i \in S$, not all zero, with $\sum_{i \in S} c_i a_i = 0$. Replacing

every c_i by $-c_i$ if necessary, we may assume that $c_i > 0$ for at least one i . Put

$$t = \min \left\{ \frac{x_i}{c_i} : i \in S, c_i > 0 \right\},$$

a minimum over a nonempty finite set, so it exists. It is > 0 , and it is attained at some $i_0 \in S$. Then

$$b = \sum_{i \in S} x_i a_i - t \sum_{i \in S} c_i a_i = \sum_{i \in S} (x_i - t c_i) a_i.$$

Every new coefficient is ≥ 0 . If $c_i \leq 0$, then $x_i - t c_i \geq x_i > 0$. If $c_i > 0$, then $t \leq x_i / c_i$, so $x_i - t c_i \geq 0$. Also $x_{i_0} - t c_{i_0} = 0$. Discarding the zero coefficients leaves a representation of b with positive coefficients on a subset of $S \setminus \{i_0\}$, which contradicts the minimality of $|S|$. Hence $(a_i)_{i \in S}$ is linearly independent, and $b \in \text{cone}(a_i : i \in S)$.

If $\dim V = n$, a linearly independent list in V has length at most n by Theorem 1.6.3 (a), so $|S| \leq n$. This proves the theorem. ■

The second part is a statement about one cone with independent generators.

Lemma 18.5.3. Injective Images of the Orthant Are Closed

Let $\mathbf{B} \in M_{n \times k}(\mathbb{R})$ have **linearly independent columns**. Then $\text{cone}(\mathbf{B}) = \{\mathbf{B}\mathbf{x} : \mathbf{x} \in \mathbb{R}_{\geq 0}^k\}$ is a closed subset of $\mathbb{R}^n$.

Idea. Independent columns make $\mathbf{x} \mapsto \mathbf{B}\mathbf{x}$ injective, and an injective matrix has a left inverse $\mathbf{L}$ with $\mathbf{L}\mathbf{B} = \mathbf{I}_k$. So if $\mathbf{B}\mathbf{x}_j$ converges, then $\mathbf{x}_j = \mathbf{L}(\mathbf{B}\mathbf{x}_j)$ converges too, because $\mathbf{L}$ is a fixed matrix. The limit is still $\geq \mathbf{0}$, and it maps to the limit of the $\mathbf{B}\mathbf{x}_j$. The left inverse is what keeps the $\mathbf{x}_j$ from running away, and injectivity is exactly what provides it.

Proof. If $k = 0$, the cone is $\{\mathbf{0}\}$, which is closed. Let $k \geq 1$. Since the columns of $\mathbf{B}$ are linearly independent, Corollary 11.4.4 ((b) $\Rightarrow$ (c), with $F = \mathbb{R}$, where $\mathbf{B}^* = \mathbf{B}^\top$) shows that $\mathbf{B}^\top \mathbf{B} \in M_k(\mathbb{R})$ is invertible. Put $\mathbf{L} = (\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top \in M_{k \times n}(\mathbb{R})$, so that $\mathbf{L}\mathbf{B} = \mathbf{I}_k$.

Let $\mathbf{y}_j \in \text{cone}(\mathbf{B})$ with $\mathbf{y}_j \rightarrow \mathbf{y}$. Write $\mathbf{y}_j = \mathbf{B}\mathbf{x}_j$ with $\mathbf{x}_j \geq \mathbf{0}$. Then $\mathbf{x}_j = \mathbf{L}\mathbf{B}\mathbf{x}_j = \mathbf{L}\mathbf{y}_j$. By Eq. 18.5.1, $\mathbf{y}_j \rightarrow \mathbf{y}$ entrywise. Each entry of $\mathbf{L}\mathbf{y}_j$ is a fixed linear combination of the entries of $\mathbf{y}_j$, so by the algebra of limits $\mathbf{x}_j = \mathbf{L}\mathbf{y}_j \rightarrow \mathbf{L}\mathbf{y}$ entrywise. Put $\mathbf{x} = \mathbf{L}\mathbf{y}$.

Each entry of $\mathbf{x}_j$ is ≥ 0 , and a nonstrict inequality survives a limit, so $\mathbf{x} \geq \mathbf{0}$. By the same reasoning with the fixed matrix $\mathbf{B}$, we have $\mathbf{y}_j = \mathbf{B}\mathbf{x}_j \rightarrow \mathbf{B}\mathbf{x}$ entrywise. A convergent sequence of real numbers has only one limit, so $\mathbf{B}\mathbf{x} = \mathbf{y}$. Hence $\mathbf{y} = \mathbf{B}\mathbf{x}$ with $\mathbf{x} \geq \mathbf{0}$, that is, $\mathbf{y} \in \text{cone}(\mathbf{B})$. This proves that the cone is closed. ■

The two pieces now fit together.

Theorem 18.5.4. Finitely Generated Cones Are Closed

Let $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. Then $\text{cone}(\mathbf{a}_1, \dots, \mathbf{a}_m)$ is a closed convex cone.

Idea. By [Theorem 18.5.2](#) the cone is the union of the cones of its independent sublists. There are finitely many of them, and each is closed by [Lemma 18.5.3](#). A finite union of closed sets is closed, by Proposition 18.1.6 (b).

Proof. Write $K = \text{cone}(\mathbf{a}_1, \dots, \mathbf{a}_m)$, a convex cone by the remark after [Definition 18.5.2](#). Let $\mathcal{S}$ be the set of those $S \subseteq \{1, \dots, m\}$ for which $(\mathbf{a}_i)_{i \in S}$ is linearly independent, and for $S \in \mathcal{S}$ put $K_S = \text{cone}(\mathbf{a}_i : i \in S)$. The empty set is in $\mathcal{S}$, and $\mathcal{S}$ is finite, with at most 2^m members.

First, $K = \bigcup_{S \in \mathcal{S}} K_S$. Each $K_S \subseteq K$, because a nonnegative combination of some of the $\mathbf{a}_i$ is a nonnegative combination of all of them with the remaining coefficients zero. And $K \subseteq \bigcup_S K_S$ is [Theorem 18.5.2](#).

Next, each K_S is closed. If $S = \emptyset$, then $K_S = \{\mathbf{0}\}$. Otherwise $K_S = \text{cone}(\mathbf{B}_S)$, where $\mathbf{B}_S$ is the matrix whose columns are the $\mathbf{a}_i$ with $i \in S$. Those columns are linearly independent, so [Lemma 18.5.3](#) applies.

Hence K is a finite union of closed sets, closed by Proposition 18.1.6 (b). ■

The proof uses no analysis beyond the algebra of limits of Chapter 16's introduction, through [Eq. 18.5.1](#) and the subsequence argument of Proposition 18.1.6 (b). In particular it never uses compactness, fact (A3) (the only compactness in the background would be the equivalence of norms, which fixing the Euclidean norm avoids); a finitely generated cone other than $\{\mathbf{0}\}$ is unbounded, so compactness had nothing to offer anyway. What the proof uses is **finite generation**, twice: for finitely many pieces, and for writing each piece as the image of an orthant under a matrix with a left inverse. The last part of the section shows that without it the conclusion can fail. In Farkas's lemma below, compactness enters only through the existence half of Theorem 18.3.1.

18.5.4 Farkas's lemma

We can now separate. Theorem 18.3.2 would give some hyperplane between $\mathbf{b}$ and the cone, but for a cone we want one **through the origin**, with a normal vector whose sign on the cone we control. Both come most cleanly from the nearest point of Theorem 18.3.1.

Lemma 18.5.5. Separating a Point from a Closed Cone

Let V be a finite-dimensional real inner product space, let $K \subseteq V$ be a **closed** convex cone, and let $b \in V$ with $b \notin K$. Then there is $y \in K^*$ with $\langle b, y \rangle < 0$.

Idea. Let p be the point of K nearest to b , and use $y = p - b$, the vector from b back to the cone. The inequality of Theorem 18.3.1 says that every point of K is seen from p at an angle of at least 90° from b . Because K is a cone, we may test it against 0 and $2p$, which shows

that the separating hyperplane passes through the origin.

Proof. K is nonempty, closed, and convex, so by Theorem 18.3.1 (a) there is a point $p \in K$ nearest to b , and by part (b) of the same theorem

$$\langle b - p, z - p \rangle \leq 0 \quad \text{for every } z \in K. \quad (18.5.2)$$

Put $y = p - b$, which is nonzero since $p \in K$ and $b \notin K$. The points $z = 0$ and $z = 2p$ lie in K , since K is a cone. Substituting them into Eq. 18.5.2 gives $\langle b - p, p \rangle \geq 0$ and $\langle b - p, p \rangle \leq 0$, so $\langle y, p \rangle = 0$. Now Eq. 18.5.2 reads $\langle b - p, z \rangle \leq 0$, that is $\langle z, y \rangle \geq 0$, for every $z \in K$. So $y \in K^*$. Finally,

$$\langle b, y \rangle = \langle p - y, y \rangle = \langle p, y \rangle - \langle y, y \rangle = -\|y\|^2 < 0,$$

since $y \neq 0$. This proves the lemma. ■

With $K = \text{cone}(\mathbf{A})$, the lemma becomes the theorem this section is named after. It is stated for a matrix, since that is how it will be used.

Theorem 18.5.6. Farkas's Lemma

Let $\mathbf{A} \in M_{m \times n}(\mathbb{R})$ and $\mathbf{b} \in \mathbb{R}^m$. **Exactly one** of the following has a solution:

- (F1) $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \geq \mathbf{0}$;
- (F2) $\mathbf{A}^\top \mathbf{y} \geq \mathbf{0}$ and $\mathbf{b}^\top \mathbf{y} < 0$, with $\mathbf{y} \in \mathbb{R}^m$.

Idea. The two systems cannot both be solvable, which is one line of algebra: pair $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{y}$. That at least one is solvable is geometry. If (F1) fails, then $\mathbf{b}$ lies outside the cone of the columns, the cone is closed by Theorem 18.5.4, and Lemma 18.5.5 produces $\mathbf{y}$.

Proof. *Not both.* Suppose $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{x} \geq \mathbf{0}$, and $\mathbf{A}^\top \mathbf{y} \geq \mathbf{0}$. Then

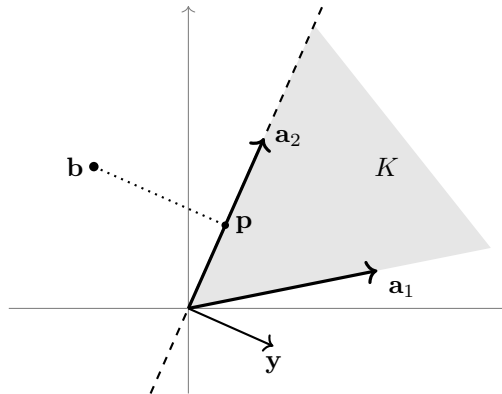
$$\mathbf{b}^\top \mathbf{y} = (\mathbf{Ax})^\top \mathbf{y} = \mathbf{x}^\top (\mathbf{A}^\top \mathbf{y}) \geq 0,$$

since it is a sum of products of the nonnegative entries of $\mathbf{x}$ and $\mathbf{A}^\top \mathbf{y}$. So $\mathbf{b}^\top \mathbf{y} < 0$ is impossible.

At least one. Suppose (F1) has no solution. Let $K = \text{cone}(\mathbf{A})$. By Definition 18.5.2, (F1) says exactly that $\mathbf{b} \in K$, so $\mathbf{b} \notin K$. By Theorem 18.5.4, K is a closed convex cone, so Lemma 18.5.5 gives $\mathbf{y} \in K^*$ with $\langle \mathbf{b}, \mathbf{y} \rangle = \mathbf{b}^\top \mathbf{y} < 0$. By Proposition 18.5.1 (d), $\mathbf{y} \in K^*$ means $\mathbf{A}^\top \mathbf{y} \geq \mathbf{0}$. So $\mathbf{y}$ solves (F2). This proves the theorem. ■

The geometric reading. Either $\mathbf{b}$ lies in the cone spanned by the columns of $\mathbf{A}$, or there is a

hyperplane through the origin, $\{\mathbf{z} : \langle \mathbf{z}, \mathbf{y} \rangle = 0\}$, with every column on one side ($\langle \mathbf{a}_j, \mathbf{y} \rangle \geq 0$) and $\mathbf{b}$ strictly on the other ($\langle \mathbf{b}, \mathbf{y} \rangle < 0$). The vector $\mathbf{y}$ is a **certificate of infeasibility**: anyone can check (F2) by multiplying out, and the “not both” computation then proves that (F1) has no solution.



$\mathbf{b}$ lies outside $K = \text{cone}(\mathbf{a}_1, \mathbf{a}_2)$, and $\mathbf{p}$ is the point of K nearest to it. The vector $\mathbf{y}$ points along $\mathbf{p} - \mathbf{b}$. The dashed line $\langle \mathbf{z}, \mathbf{y} \rangle = 0$ passes through $\mathbf{0}$, with K on the side $\langle \mathbf{z}, \mathbf{y} \rangle \geq 0$ and $\mathbf{b}$ strictly on the other

Example. Deciding a system both ways

Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \quad \mathbf{b}' = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Decide, for $\mathbf{b}$ and for $\mathbf{b}'$, which alternative of [Theorem 18.5.6](#) holds, and exhibit the solution.

Solution. The columns are $\mathbf{a}_1 = \mathbf{e}_1$, $\mathbf{a}_2 = (2, 1)$, and $\mathbf{a}_3 = \mathbf{e}_2$, all in the closed first quadrant. So $\text{cone}(\mathbf{A}) \subseteq \mathbb{R}_{\geq 0}^2$, and in fact the two are equal, since $\mathbf{e}_1$ and $\mathbf{e}_2$ are among the generators.

For $\mathbf{b} = (3, 2)$ we expect (F1). The system $x_1 + 2x_2 = 3$, $x_2 + x_3 = 2$ has, for example, the solution $\mathbf{x} = (3, 0, 2) \geq \mathbf{0}$, and also $\mathbf{x} = (1, 1, 1)$. Nonnegative solutions need not be unique.

For $\mathbf{b}' = (-1, 1)$ we expect (F2), since $\mathbf{b}'$ has a negative first entry. Try $\mathbf{y} = \mathbf{e}_1$: then $\mathbf{A}^\top \mathbf{y} = (1, 2, 0) \geq \mathbf{0}$ and $\mathbf{b}'^\top \mathbf{y} = -1 < 0$. So $\mathbf{y} = \mathbf{e}_1$ is a certificate, and by the “not both” half of [Theorem 18.5.6](#) the system $\mathbf{Ax} = \mathbf{b}'$, $\mathbf{x} \geq \mathbf{0}$ has no solution. The certificate encodes a one-line argument: the first equation reads $x_1 + 2x_2 = -1$, and its left side is ≥ 0 whenever $\mathbf{x} \geq \mathbf{0}$.

Linear programming, in the next section, meets systems of **inequalities**, $\mathbf{Ax} \leq \mathbf{b}$ with $\mathbf{x} \geq \mathbf{0}$. Farkas's lemma handles them after one standard move: add a **slack** variable to turn each inequality into an equation.

Corollary 18.5.7. Farkas's Lemma for Inequalities

Let $\mathbf{M} \in M_{m \times n}(\mathbb{R})$ and $\mathbf{d} \in \mathbb{R}^m$. Exactly one of the following has a solution:

- (i) $\mathbf{M}\mathbf{z} \leq \mathbf{d}$ with $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z} \geq \mathbf{0}$;
- (ii) $\mathbf{w} \geq \mathbf{0}$, $\mathbf{M}^\top \mathbf{w} \geq \mathbf{0}$, and $\mathbf{d}^\top \mathbf{w} < 0$, with $\mathbf{w} \in \mathbb{R}^m$.

Idea. A slack vector $\mathbf{s} = \mathbf{d} - \mathbf{M}\mathbf{z} \geq \mathbf{0}$ turns the inequalities into equations with matrix $\begin{pmatrix} \mathbf{M} & \mathbf{I}_m \end{pmatrix}$, and in [Theorem 18.5.6](#) the identity block becomes the extra condition $\mathbf{w} \geq \mathbf{0}$.

Proof. Here $\mathbf{M}\mathbf{z} \leq \mathbf{d}$ means $\mathbf{d} - \mathbf{M}\mathbf{z} \geq \mathbf{0}$. So (i) has a solution if and only if there are $\mathbf{z} \geq \mathbf{0}$ and $\mathbf{s} \geq \mathbf{0}$ in $\mathbb{R}^m$ with $\mathbf{M}\mathbf{z} + \mathbf{s} = \mathbf{d}$: given $\mathbf{z}$, take $\mathbf{s} = \mathbf{d} - \mathbf{M}\mathbf{z}$, and given $\mathbf{z}$ and $\mathbf{s}$, drop $\mathbf{s}$. In block form this reads

$$\begin{pmatrix} \mathbf{M} & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} \mathbf{z} \\ \mathbf{s} \end{pmatrix} = \mathbf{d}, \quad \begin{pmatrix} \mathbf{z} \\ \mathbf{s} \end{pmatrix} \geq \mathbf{0}.$$

Apply [Theorem 18.5.6](#) to the $m \times (n + m)$ matrix $\begin{pmatrix} \mathbf{M} & \mathbf{I}_m \end{pmatrix}$ and to $\mathbf{d}$. Exactly one of the system above and the following is solvable: $\begin{pmatrix} \mathbf{M} & \mathbf{I}_m \end{pmatrix}^\top \mathbf{w} \geq \mathbf{0}$ and $\mathbf{d}^\top \mathbf{w} < 0$. The transpose has the blocks $\mathbf{M}^\top$ and $\mathbf{I}_m$ stacked, so the first condition says $\mathbf{M}^\top \mathbf{w} \geq \mathbf{0}$ and $\mathbf{w} \geq \mathbf{0}$. That is (ii). ■

18.5.5 The dual of the dual

Farkas's lemma has a coordinate-free form, which says that nothing is lost by passing to the dual cone twice, provided the cone is closed.

Theorem 18.5.8. The Bidual of a Closed Cone

Let V be a finite-dimensional real inner product space and let $K \subseteq V$ be a **closed** convex cone. Then $K^{**} = K$.

Proof. $K \subseteq K^{**}$ is [Proposition 18.5.1](#) (c). Conversely, let $b \notin K$. By [Lemma 18.5.5](#) there is $y \in K^*$ with $\langle b, y \rangle < 0$. That single y shows $b \notin K^{**}$. Hence $K^{**} \subseteq K$, and the two are equal. ■

For $K = \text{cone}(\mathbf{A})$, which is closed by [Theorem 18.5.4](#), the theorem is Farkas's lemma. By [Proposition 18.5.1](#) (d), $K^* = \{\mathbf{y} : \mathbf{A}^\top \mathbf{y} \geq \mathbf{0}\}$, so “ $\mathbf{b} \in K^{**}$ ” says that every $\mathbf{y}$ with $\mathbf{A}^\top \mathbf{y} \geq \mathbf{0}$ has $\mathbf{b}^\top \mathbf{y} \geq 0$, that is, (F2) has no solution. And “ $\mathbf{b} \in K$ ” is (F1). This is the “dualize” move in its purest form: a question about a cone (is $\mathbf{b}$ in it?) is traded for one about its dual (does every member of the dual approve of $\mathbf{b}$?), and the trade is exact when the cone is closed.

Remark. For the wedge $K = \text{cone}((1, 0), (1, 1))$, whose dual is $\text{cone}((0, 1), (1, -1))$ by [Exercise \(B3: A dual cone and its dual\)](#), the theorem and [Proposition 18.5.1](#) (d) say that $K = \{\mathbf{x} : \langle \mathbf{x}, (0, 1) \rangle \geq 0, \langle \mathbf{x}, (1, -1) \rangle \geq 0\}$, one inequality for each generator of K^* . In general the theorem describes a closed convex cone from outside, as an intersection of half-spaces

through the origin; for a finitely generated cone finitely many suffice, which is one half of the Minkowski–Weyl theorem of Section 9.

18.5.6 When the cone is not finitely generated

The system “ $\mathbf{Ax} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$ ” asks whether $\mathbf{b}$ lies in the image of the orthant under $\mathbf{x} \mapsto \mathbf{Ax}$. Replace the orthant by a closed convex cone $K \subseteq \mathbb{R}^k$ and $\mathbf{A}$ by a matrix $\mathbf{P} \in M_{m \times k}(\mathbb{R})$, and ask whether $\mathbf{b} \in \mathbf{P}(K)$. The “not both” half of the proof of [Theorem 18.5.6](#) survives unchanged. The “at least one” half needs $\mathbf{P}(K)$ to be closed, and the next example shows that this can fail.

Example. A closed cone with a nonclosed shadow

In $\mathbb{R}^3$, with coordinates (x, y, z) , let

$$K = \{(x, y, z) : xs^2 + 2yst + zt^2 \geq 0 \text{ for all } s, t \in \mathbb{R}\},$$

the set of triples for which the symmetric matrix $\begin{pmatrix} x & y \\ y & z \end{pmatrix}$ is positive semidefinite (Definition 13.1.1). Let $\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, so that $\mathbf{P}(x, y, z) = (x, y)$ is the projection onto the first two coordinates, and write $\mathbf{P}(K) = \{\mathbf{Pk} : \mathbf{k} \in K\}$. Show that K is a closed convex cone, but that

$$\mathbf{P}(K) = \{(x, y) : x > 0\} \cup \{(0, 0)\},$$

the cone of the earlier warning, which is not closed. Then show that for $\mathbf{b} = (0, 1)$ **neither** of the two alternatives holds:

- (i) $\mathbf{Pk} = \mathbf{b}$ for some $\mathbf{k} \in K$;
- (ii) $\langle \mathbf{Pk}, \mathbf{y} \rangle \geq 0$ for every $\mathbf{k} \in K$, and $\langle \mathbf{b}, \mathbf{y} \rangle < 0$, for some $\mathbf{y} \in \mathbb{R}^2$.

Solution. K is a closed convex cone. Write $q_{\mathbf{k}}(s, t) = xs^2 + 2yst + zt^2$ for $\mathbf{k} = (x, y, z)$. It is linear in $\mathbf{k}$: $q_{a\mathbf{k}+b\mathbf{k}'} = aq_{\mathbf{k}} + bq_{\mathbf{k}'}$. So if $\mathbf{k}, \mathbf{k}' \in K$ and $a, b \geq 0$, then $q_{a\mathbf{k}+b\mathbf{k}'}(s, t) \geq 0$ for all s, t , and K is a convex cone that contains $\mathbf{0}$. If $\mathbf{k}_j \rightarrow \mathbf{k}$ with $\mathbf{k}_j \in K$, then for fixed s, t the number $q_{\mathbf{k}_j}(s, t)$ is a fixed linear combination of the entries of $\mathbf{k}_j$. So it converges to $q_{\mathbf{k}}(s, t)$, and $q_{\mathbf{k}}(s, t) \geq 0$ because a nonstrict inequality survives a limit. Hence K is closed.

Membership. By Theorem 13.1.4 (a) $\Leftrightarrow$ (e), $(x, y, z) \in K$ exactly when the principal minors x , z , and $xz - y^2$ are all ≥ 0 .

The image. $(0, 0) = \mathbf{P}(\mathbf{0}) \in \mathbf{P}(K)$. If $x > 0$ and y is arbitrary, put $z = y^2/x$. Then $x \geq 0$, $z \geq 0$, and $xz - y^2 = 0$, so $(x, y, z) \in K$ and $(x, y) \in \mathbf{P}(K)$. Conversely, if $(x, y, z) \in K$, then $x \geq 0$, and if $x = 0$ then $0 = xz \geq y^2$ forces $y = 0$. So $\mathbf{P}(K)$ is exactly the displayed set. It is not closed: $(1/k, 1) \in \mathbf{P}(K)$ converges to $(0, 1) \notin \mathbf{P}(K)$.

Neither alternative holds for $\mathbf{b} = (0, 1)$. Alternative (i) fails because $\mathbf{b} \notin \mathbf{P}(K)$. For (ii), suppose $\mathbf{y} = (y_1, y_2)$ has $\langle \mathbf{Pk}, \mathbf{y} \rangle \geq 0$ for every $\mathbf{k} \in K$. Since $(1, t) \in \mathbf{P}(K)$ for every real t , we get $y_1 + ty_2 \geq 0$ for all t . A nonconstant affine function of t takes negative values, so $y_2 = 0$. But then $\langle \mathbf{b}, \mathbf{y} \rangle = y_2 = 0$, which is not < 0 . So (ii) fails too.

It is worth seeing exactly where the proof of [Theorem 18.5.6](#) breaks. The vectors $\mathbf{k}_j = (1/j, 1, j)$ lie in K , their images $(1/j, 1)$ converge to $\mathbf{b}$, and the $\mathbf{k}_j$ themselves run off to infinity in the third coordinate. This is the escape that the left inverse in [Lemma 18.5.3](#) prevented. There is

no nearest point of $\mathbf{P}(K)$ to $\mathbf{b}$: the distance is 0, but $\mathbf{b}$ is not in the set, so [Lemma 18.5.5](#) has nothing to start from. What fails is **closedness**, which is why the finitely generated case needed its own theorem.

The dual cone cannot see the difference either: $\mathbf{P}(K)^{**}$ is the closed half-plane $\{x \geq 0\}$, strictly larger than $\mathbf{P}(K)$, as [Exercise \(C3: The bidual sees only the closure\)](#) shows.

18.5.7 Exercises

18.5.7.1 A. Check your understanding

Exercise. A1

- Define a convex cone and the dual cone of a subset S of $\mathbb{R}^n$.
- State Farkas's lemma, and say in one sentence what the vector $\mathbf{y}$ in the second alternative certifies.
- Determine whether the following is correct, and justify your answer: every convex cone in $\mathbb{R}^2$ is closed.
- Determine whether the following is correct, and justify your answer: the union of two convex cones in $\mathbb{R}^2$ is a convex cone.
- In [Theorem 18.5.2](#) the bound is n generators, while Theorem 18.2.2 allows $n+1$ points. Which feature of the proof accounts for the difference?

Solution.

- A convex cone is a nonempty $K \subseteq \mathbb{R}^n$ with $a\mathbf{u} + b\mathbf{v} \in K$ for all $\mathbf{u}, \mathbf{v} \in K$ and all real $a, b \geq 0$. The dual cone of S is $S^* = \{\mathbf{y} : \langle \mathbf{x}, \mathbf{y} \rangle \geq 0 \text{ for every } \mathbf{x} \in S\}$.
- For $\mathbf{A} \in M_{m \times n}(\mathbb{R})$ and $\mathbf{b} \in \mathbb{R}^m$, exactly one of the following is solvable: $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{x} \geq \mathbf{0}$, or $\mathbf{A}^\top \mathbf{y} \geq \mathbf{0}$ with $\mathbf{b}^\top \mathbf{y} < 0$. Such a $\mathbf{y}$ certifies that the first system has no solution, since any solution $\mathbf{x}$ would give $0 > \mathbf{b}^\top \mathbf{y} = \mathbf{x}^\top \mathbf{A}^\top \mathbf{y} \geq 0$.
- Incorrect. The cone $\{(x, y) : x > 0\} \cup \{\mathbf{0}\}$ of the warning in this section is a convex cone, but $(1/k, 1) \rightarrow (0, 1)$ leaves it.
- Incorrect. The two quadrants $\mathbb{R}_{\geq 0}^2$ and $-\mathbb{R}_{\geq 0}^2$ are convex cones, and their union contains $(1, 0)$ and $(0, -1)$ but not their sum $(1, -1)$.
- For cones the dependence to be removed is a **linear** dependence $\sum c_i a_i = 0$, and a linearly independent list in an n -dimensional space has at most n members. For convex hulls it is an **affine** dependence (coefficients also summing to 0), and an affinely independent list can have $n+1$ members.

18.5.7.2 B. Practice

Exercise. B1: Recognizing cones

Determine which of the following subsets of $\mathbb{R}^2$ or $\mathbb{R}^3$ are convex cones. Justify your answers.

- $\{\mathbf{x} \in \mathbb{R}^2 : x_1 \geq |x_2|\}$.
- $\{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$.
- $\{\mathbf{x} \in \mathbb{R}^2 : x_1 \geq 1\}$.

- (d) $\{\mathbf{x} \in \mathbb{R}^2 : x_2 \geq x_1^2\}$.
 (e) $\{\mathbf{x} \in \mathbb{R}^3 : x_1 \geq 0, x_2 \geq 0, x_3 \leq x_1 + x_2\}$.

Solution.

- (a) A convex cone. It is nonempty, containing $\mathbf{0}$. If $x_1 \geq |x_2|$, $y_1 \geq |y_2|$, and $a, b \geq 0$, then by the triangle inequality for real numbers (Theorem 0.6.4)

$$|ax_2 + by_2| \leq a|x_2| + b|y_2| \leq ax_1 + by_1.$$

It is cone $((1, 1), (1, -1))$, a quarter-plane turned by 45° .

- (b) A convex cone. It is a subspace, being the null space of $\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$, and every subspace is a convex cone.
 (c) Not a convex cone: it does not contain $\mathbf{0}$, which taking $a = b = 0$ in the definition would force into it. It is convex and closed under addition; what fails is scaling by $0 \leq a < 1$, for instance $\frac{1}{2}(1, 0) = (\frac{1}{2}, 0)$ is not in it.
 (d) Not a convex cone. It is convex, but it is not closed under nonnegative scaling: $(1, 1)$ is in it, and $2(1, 1) = (2, 2)$ is not, since $2 < 4$.
 (e) A convex cone. Each of the three conditions is a homogeneous linear inequality $\langle \mathbf{x}, \mathbf{a} \rangle \geq 0$, with $\mathbf{a} = \mathbf{e}_1$, $\mathbf{e}_2$, and $(1, 1, -1)$, so the set is an intersection of three half-spaces through the origin. Each is a convex cone by the examples of this section, and an intersection of convex cones is a convex cone: it contains $\mathbf{0}$, and each member cone contains every nonnegative combination of points of the intersection.

Exercise. B2: A certificate either way

Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \end{pmatrix}$. For each of $\mathbf{b} = (0, 3)$ and $\mathbf{b} = (-1, 1)$, decide which alternative of [Theorem 18.5.6](#) holds, and exhibit a solution of that alternative.

Solution. The columns are $\mathbf{a}_1 = (1, 0)$, $\mathbf{a}_2 = (1, 1)$, and $\mathbf{a}_3 = (-1, 2)$.

For $\mathbf{b} = (0, 3)$, try $\mathbf{x} = (0, x_2, x_3)$: the equations $x_2 - x_3 = 0$ and $x_2 + 2x_3 = 3$ give $x_2 = x_3 = 1$. So $\mathbf{x} = (0, 1, 1) \geq \mathbf{0}$ solves $\mathbf{Ax} = \mathbf{b}$, and (F1) holds.

For $\mathbf{b} = (-1, 1)$, look for $\mathbf{y}$ perpendicular to the “outermost” generator $\mathbf{a}_3$ and nonnegative on the others. The vector $\mathbf{y} = (2, 1)$ gives $\mathbf{A}^\top \mathbf{y} = (2, 3, 0) \geq \mathbf{0}$ and $\mathbf{b}^\top \mathbf{y} = -2 + 1 = -1 < 0$. So (F2) holds, and by [Theorem 18.5.6](#) the system $\mathbf{Ax} = (-1, 1)$, $\mathbf{x} \geq \mathbf{0}$ has no solution. As a direct check, adding twice the first equation to the second gives $2x_1 + 3x_2 = -1$, whose left side is ≥ 0 for $\mathbf{x} \geq \mathbf{0}$.

Exercise. B3: A dual cone and its dual

Let $K = \text{cone}((1, 0), (1, 1)) \subseteq \mathbb{R}^2$.

- (a) Show that $K^* = \text{cone}((0, 1), (1, -1))$.

- (b) Compute K^{**} and check that it equals K , as [Theorem 18.5.8](#) predicts.

Solution.

- (a) By [Proposition 18.5.1](#) (d), $K^* = \{\mathbf{y} : y_1 \geq 0, y_1 + y_2 \geq 0\}$. Given such a $\mathbf{y}$, write $\mathbf{y} = (y_1 + y_2)(0, 1) + y_1(1, -1)$: indeed the right side is $(y_1, y_1 + y_2 - y_1) = (y_1, y_2)$, and both coefficients are ≥ 0 . So $\mathbf{y} \in \text{cone}((0, 1), (1, -1))$. Conversely, both generators satisfy the two inequalities, and the set defined by the inequalities is a convex cone, so it contains the cone they generate.
- (b) Applying [Proposition 18.5.1](#) (d) to the generators found in (a), $K^{**} = \{\mathbf{x} : x_2 \geq 0, x_1 - x_2 \geq 0\}$. The same decomposition argument, now with $\mathbf{x} = (x_1 - x_2)(1, 0) + x_2(1, 1)$, shows this set is $\text{cone}((1, 0), (1, 1)) = K$.

18.5.7.3 C. Going deeper

Exercise. C1: Gordan's alternative

Let $\mathbf{A} \in M_{m \times n}(\mathbb{R})$.

- (a) Prove that exactly one of the following has a solution: $\mathbf{A}\mathbf{x} = \mathbf{0}$ with $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{x} \neq \mathbf{0}$; or $\mathbf{A}^\top \mathbf{y} > \mathbf{0}$ with $\mathbf{y} \in \mathbb{R}^m$.
- (b) Deduce that if the columns of $\mathbf{A}$ are $(1, 0)$, $(0, 1)$, and $(-1, -1)$, then some nontrivial nonnegative combination of them is $\mathbf{0}$, and find one.

Hint for (a): scale $\mathbf{x}$ so that its entries sum to 1, and apply [Theorem 18.5.6](#) to $\mathbf{A}$ with an extra row of ones.

Solution.

- (a) *Not both.* If $\mathbf{A}\mathbf{x} = \mathbf{0}$ with $\mathbf{x} \geq \mathbf{0}$, $\mathbf{x} \neq \mathbf{0}$, and $\mathbf{A}^\top \mathbf{y} > \mathbf{0}$, then $0 = \mathbf{y}^\top \mathbf{A}\mathbf{x} = \sum_j x_j (\mathbf{A}^\top \mathbf{y})_j$. Every term is ≥ 0 , and the term for an index with $x_j > 0$ is > 0 . That is a contradiction.
- At least one.* Suppose the first system has no solution. If $\mathbf{A}\mathbf{x} = \mathbf{0}$, $\mathbf{x} \geq \mathbf{0}$, and $\mathbf{1}^\top \mathbf{x} = 1$, then $\mathbf{x} \neq \mathbf{0}$ would solve it. So the system

$$\begin{pmatrix} \mathbf{A} \\ \mathbf{1}^\top \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}, \quad \mathbf{x} \geq \mathbf{0},$$

has no solution. By [Theorem 18.5.6](#) there is $(\mathbf{y}, s) \in \mathbb{R}^m \times \mathbb{R}$ with $\mathbf{A}^\top \mathbf{y} + s\mathbf{1} \geq \mathbf{0}$ and $\mathbf{0}^\top \mathbf{y} + 1 \cdot s < 0$, that is, $s < 0$. Then $\mathbf{A}^\top \mathbf{y} \geq -s\mathbf{1} > \mathbf{0}$, so $\mathbf{y}$ solves the second system.

- (b) Here $\mathbf{A} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$. The second system is impossible: $\mathbf{A}^\top \mathbf{y} = (y_1, y_2, -y_1 - y_2)$, and its entries sum to 0, so they cannot all be positive. By (a) the first system is solvable. Indeed $\mathbf{x} = (1, 1, 1)$ gives $\mathbf{A}\mathbf{x} = (1 - 1, 1 - 1) = \mathbf{0}$.

Exercise. C2: The ice-cream cone is its own dual

Let $L = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} : \|\mathbf{x}\|_2 \leq t\}$, with the inner product $\langle (\mathbf{x}, t), (\mathbf{y}, s) \rangle = \mathbf{y}^\top \mathbf{x} + ts$ on $\mathbb{R}^{n+1}$.

- (a) Prove that $L \subseteq L^*$.
- (b) Prove that $L^* \subseteq L$.
- (c) Deduce that L is closed.

Hint for (b): test $(\mathbf{y}, s) \in L^*$ against the member $(-\mathbf{y}, \|\mathbf{y}\|_2)$ of L .

Solution.

- (a) Let $(\mathbf{x}, t), (\mathbf{y}, s) \in L$. By Theorem 11.1.5, $|\mathbf{y}^\top \mathbf{x}| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \leq ts$, where the last step multiplies the inequalities $0 \leq \|\mathbf{x}\|_2 \leq t$ and $0 \leq \|\mathbf{y}\|_2 \leq s$. Hence $\mathbf{y}^\top \mathbf{x} + ts \geq ts - |\mathbf{y}^\top \mathbf{x}| \geq 0$. Fix $(\mathbf{y}, s) \in L$: we have shown that its inner product with every $(\mathbf{x}, t) \in L$ is ≥ 0 , so $(\mathbf{y}, s) \in L^*$. Hence $L \subseteq L^*$.
- (b) Let $(\mathbf{y}, s) \in L^*$. The point $(-\mathbf{y}, \|\mathbf{y}\|_2)$ lies in L , so

$$0 \leq \langle (-\mathbf{y}, \|\mathbf{y}\|_2), (\mathbf{y}, s) \rangle = -\|\mathbf{y}\|_2^2 + s\|\mathbf{y}\|_2 = \|\mathbf{y}\|_2(s - \|\mathbf{y}\|_2).$$

If $\mathbf{y} \neq \mathbf{0}$, dividing by $\|\mathbf{y}\|_2 > 0$ gives $s \geq \|\mathbf{y}\|_2$. If $\mathbf{y} = \mathbf{0}$, test against $(\mathbf{0}, 1) \in L$ instead: $s \geq 0 = \|\mathbf{y}\|_2$. Either way $(\mathbf{y}, s) \in L$.

- (c) By (a) and (b), $L = L^*$, and every dual cone is closed by [Proposition 18.5.1](#) (a).

Exercise. C3: The bidual sees only the closure

Let $C = \{(x, y) : x > 0\} \cup \{(0, 0)\} \subseteq \mathbb{R}^2$, the image cone of [Example \(A closed cone with a nonclosed shadow\)](#).

- (a) Prove that $C^* = \{(y_1, 0) : y_1 \geq 0\}$.
- (b) Prove that $C^{**} = \{(x, y) : x \geq 0\}$.
- (c) Explain why this does not contradict [Theorem 18.5.8](#), and show that C^{**} is the set of all limits of convergent sequences of points of C .

Solution.

- (a) If $(y_1, y_2) \in C^*$, then testing against $(1, t) \in C$ gives $y_1 + ty_2 \geq 0$ for every real t . This forces $y_2 = 0$, as in the example, and then $t = 0$ gives $y_1 \geq 0$. Conversely, if $y_2 = 0$ and $y_1 \geq 0$, then $\langle (x, y), (y_1, 0) \rangle = xy_1 \geq 0$ for every $(x, y) \in C$, since $x \geq 0$ on C .
- (b) By (a), $C^* = \text{cone}((1, 0))$, so by [Proposition 18.5.1](#) (d),

$$C^{**} = \{(x, y) : \langle (x, y), (1, 0) \rangle \geq 0\} = \{(x, y) : x \geq 0\}.$$

- (c) [Theorem 18.5.8](#) assumes the cone is closed, and C is not. Every point of C^{**} is such a limit: if $x > 0$ or $(x, y) = \mathbf{0}$, use the constant sequence; if $x = 0$, use $(1/k, y) \in C$, which tends to $(0, y)$. Conversely, a limit of points of $C \subseteq C^{**}$ lies in C^{**} , because C^{**} is closed by [Proposition 18.5.1](#) (a). So C^{**} is exactly the set of limits of sequences in C .