

18.9 Polytopes

There are two ways to pin down a convex set with finitely many numbers. One lists finitely many points and takes their convex hull; the other lists finitely many linear inequalities and takes their common solutions. A triangle can be given either way, by its three corners or by its three sides. This section proves that for bounded sets the two descriptions always describe the same sets, which is the Minkowski–Weyl theorem. On the way it locates the optima of Section 6’s linear programs over bounded regions: they occur at corners, and a corner of $\{\mathbf{x} : \mathbf{Ax} \leq \mathbf{b}\}$ can be recognized by a rank condition on the constraints that hold with equality there. The other direction of the theorem comes from Fourier–Motzkin elimination, an algorithm for projecting a system of inequalities that works the way Gaussian elimination does for equations.

Throughout, the field is $\mathbb{R}$ and the space is $\mathbb{R}^n$ with $n \geq 1$ and the dot product. An inequality between vectors is meant **entrywise**: $\mathbf{Ax} \leq \mathbf{b}$ says that $(\mathbf{Ax})_i \leq b_i$ for every i . For $\mathbf{A} \in M_{m \times n}(\mathbb{R})$ we write $\mathbf{a}_1^\top, \dots, \mathbf{a}_m^\top$ for its rows, so that $\mathbf{Ax} \leq \mathbf{b}$ is the list of m inequalities $\mathbf{a}_i^\top \mathbf{x} \leq b_i$.

18.9.1 Two finite descriptions

The first description already appeared in Section 8, just after Proposition 18.8.2.

A polytope is the convex hull of finitely many points; a polyhedron is the solution set of finitely many linear inequalities.

Definition 18.9.1. Polytope

A **polytope** in $\mathbb{R}^n$ is a set of the form $\text{conv}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ for **finitely many** points $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$, $k \geq 0$.

Definition 18.9.2. Polyhedron

A **polyhedron** in $\mathbb{R}^n$ is a set of the form

$$P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} \leq \mathbf{b}\}$$

for some $m \geq 0$, $\mathbf{A} \in M_{m \times n}(\mathbb{R})$, and $\mathbf{b} \in \mathbb{R}^m$; that is, the intersection of **finitely many** sets $\{\mathbf{x} : \mathbf{a}_i^\top \mathbf{x} \leq b_i\}$. An extreme point of a polyhedron or of a polytope (Definition 18.8.1) is called a **vertex**.

In words: a polytope is built from the inside, by mixing finitely many given points; a polyhedron is carved from the outside, by finitely many cuts. When $\mathbf{a}_i \neq \mathbf{0}$, the set $\{\mathbf{x} : \mathbf{a}_i^\top \mathbf{x} \leq b_i\}$ is a closed half-space, bounded by the hyperplane $\mathbf{a}_i^\top \mathbf{x} = b_i$ (Definition 5.5.2). A zero row $\mathbf{a}_i = \mathbf{0}$ is allowed and harmless: the condition $0 \leq b_i$ either holds for every $\mathbf{x}$ or for none. An **equation** $\mathbf{a}^\top \mathbf{x} = \beta$ is allowed too, as the two inequalities $\mathbf{a}^\top \mathbf{x} \leq \beta$ and $-\mathbf{a}^\top \mathbf{x} \leq -\beta$. By Theorem 18.1.3, the polytope $\text{conv}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is the set of all $\sum_j t_j \mathbf{v}_j$ with every $t_j \geq 0$ and $\sum_j t_j = 1$.

Some examples, simplest first.

1. **The empty set and a point** are both. $\emptyset = \text{conv} \emptyset$, and $\emptyset = \{\mathbf{x} : 0 \leq -1\}$ with the single zero row. A point $\mathbf{v} = \text{conv}\{\mathbf{v}\}$ is the solution set of the $2n$ inequalities $x_i \leq v_i$ and $-x_i \leq -v_i$.
2. **All of $\mathbb{R}^n$** is a polyhedron, with $m = 0$ (or with the one row $0 \leq 0$). So is a closed half-space. Neither is a polytope, because a polytope is bounded: if $\|\mathbf{v}_j\| \leq R$ for every j ,

- then $\|\sum_j t_j \mathbf{v}_j\| \leq \sum_j t_j \|\mathbf{v}_j\| \leq R$.
3. **The triangle** $T \subseteq \mathbb{R}^2$ with corners $(0,0), (3,0), (0,3)$ is $\text{conv}\{(0,0), (3,0), (0,3)\}$, and it is the polyhedron $\{x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 3\}$. That the two descriptions agree is not a definition but a fact; the worked example in the section on Fourier–Motzkin elimination below derives each from the other.
 4. **The cube** $B_\infty = [-1, 1]^n$ is the polyhedron given by the $2n$ inequalities $\pm x_i \leq 1$, and by Section 8 it is the hull of its 2^n sign vectors (Example (The diamond, the cube, and the sphere)) and Theorem 18.8.5). So it is both.

A nonexample by minimal change. Replace the cube’s square $B_\infty \subseteq \mathbb{R}^2$ by the disk B_2 . It is still compact and convex, still the unit ball of a norm, and still the solution set of inequalities $\mathbf{a}^\top \mathbf{x} \leq 1$, one for **every** unit vector $\mathbf{a}$: by Theorem 11.1.5, $\mathbf{a}^\top \mathbf{x} \leq \|\mathbf{x}\| \leq 1$ on the disk, and a point with $\|\mathbf{x}\| > 1$ violates the inequality for $\mathbf{a} = \mathbf{x}/\|\mathbf{x}\|$. What fails is the word **finitely**. The disk is not a polytope: by Proposition 18.8.2 a polytope has finitely many extreme points, while the disk has a whole circle of them (Example (The diamond, the cube, and the sphere)). We will see after the Minkowski–Weyl theorem that it is not a polyhedron either.

Why these definitions. Both allow only finite data, and that is the point: finitely many inequalities can be checked, finitely many points can be listed, and the whole theory of this section is about moving between two finite lists. An infinite family of inequalities describes every closed convex set, by the separation theorem (Theorem 18.3.2), so dropping “finitely many” would make the second definition say nothing new.

Section 1 already showed that every polyhedron is convex and closed; we record it for citation.

Proposition 18.9.1. Polyhedra Are Closed and Convex

Every polyhedron $P = \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ is convex and closed.

Proof. With the linear functional $\varphi_i(\mathbf{x}) = \mathbf{a}_i^\top \mathbf{x}$, each set $\{\mathbf{x} : \mathbf{a}_i^\top \mathbf{x} \leq b_i\} = \{\varphi_i \leq b_i\}$ is closed by Proposition 18.1.6 (c), and convex by Proposition 18.1.1 (b), as the preimage of the interval $(-\infty, b_i]$ under φ_i . So their intersection P is closed by Proposition 18.1.6 (b) and convex by Proposition 18.1.1 (a). ■

Warning. A polyhedron need not have any vertices, even when it is not empty. The strip $\{\mathbf{x} \in \mathbb{R}^2 : -1 \leq x_1 + x_2 \leq 1\}$ is a closed convex set with no extreme points at all: every point $\mathbf{x}$ of it is the midpoint of $\mathbf{x} \pm (1, -1)$, which lie in it. Minkowski’s theorem (Theorem 18.8.5) does not apply because the strip is unbounded, and in the language of the next theorem, its constraint rows $\pm(1, 1)$ have rank $1 < 2$.

18.9.2 Vertices are basic feasible points

Section 8 showed that a linear function on a nonempty compact convex set attains its maximum at an extreme point (Corollary 18.8.6). For a linear program over a bounded region $\{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$, that shrinks the search to the vertices. To make it a finite search we need a way to recognize a vertex from the data $\mathbf{A}, \mathbf{b}$, and a proof that there are finitely many.

Stand at a point $\mathbf{x}$ of P and look at the inequalities that are tight there. Each tight inequality $\mathbf{a}_i^\top \mathbf{x} = b_i$ confines $\mathbf{x}$ to a hyperplane. If the tight rows are rich enough to pin $\mathbf{x}$ down, there is no room to move in any direction, and $\mathbf{x}$ should be a corner. If not, there should be a direction along which one can move both ways without leaving P .

Definition 18.9.3. Active Set, Basic Feasible Point

Let $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ and $\mathbf{x} \in P$. The **active set** of $\mathbf{x}$ is $I(\mathbf{x}) = \{i : \mathbf{a}_i^\top \mathbf{x} = b_i\}$, the indices of the inequalities that hold with equality at $\mathbf{x}$. The point $\mathbf{x}$ is a **basic feasible point** of the system if among the active rows $\{\mathbf{a}_i : i \in I(\mathbf{x})\}$ there are n **linearly independent** ones.

Equivalently, the matrix $\mathbf{A}_{I(\mathbf{x})}$ formed by the active rows has a row space of dimension n , that is, rank n (Theorem 3.5.5); by Theorem 3.5.7 this says that its null space is $\{\mathbf{0}\}$, which means that $\mathbf{x}$ is the **only** solution of the equations $\mathbf{a}_i^\top \mathbf{y} = b_i, i \in I(\mathbf{x})$. The word “feasible” records that $\mathbf{x}$ lies in P ; a solution of n independent equations $\mathbf{a}_i^\top \mathbf{y} = b_i$ that violates some other inequality is a **basic point** but not a basic feasible one. Note that the definition depends on the system $\mathbf{A}\mathbf{x} \leq \mathbf{b}$, not only on the set P ; the theorem shows that the resulting notion depends only on P .

Theorem 18.9.2. Vertices Are Basic Feasible Points

Let $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ and $\mathbf{x} \in P$. Then $\mathbf{x}$ is a vertex of P if and only if $\mathbf{x}$ is a basic feasible point.

Idea. Both directions are about the directions $\mathbf{u}$ with $\mathbf{a}_i^\top \mathbf{u} = 0$ for every active i , the directions along which every tight constraint stays tight. If the active rows have rank n , the only such direction is $\mathbf{0}$, and an averaging argument pins down any segment through $\mathbf{x}$. If their rank is less than n , a nonzero such $\mathbf{u}$ exists by rank-nullity, and a short step $\pm \varepsilon \mathbf{u}$ keeps the tight constraints tight and, because there are finitely many, keeps the slack ones slack.

Proof. Write $I = I(\mathbf{x})$ and let $\mathbf{A}_I$ be the matrix of active rows (empty if $I = \emptyset$).

($\Leftarrow$) Suppose $\mathbf{A}_I$ has rank n , and let $\mathbf{x} = \frac{1}{2}(\mathbf{y} + \mathbf{z})$ with $\mathbf{y}, \mathbf{z} \in P$. For $i \in I$,

$$b_i = \mathbf{a}_i^\top \mathbf{x} = \frac{1}{2} \mathbf{a}_i^\top \mathbf{y} + \frac{1}{2} \mathbf{a}_i^\top \mathbf{z},$$

with $\mathbf{a}_i^\top \mathbf{y} \leq b_i$ and $\mathbf{a}_i^\top \mathbf{z} \leq b_i$; if either were strict, the average would be less than b_i . So $\mathbf{a}_i^\top \mathbf{y} = b_i = \mathbf{a}_i^\top \mathbf{x}$, that is, $\mathbf{A}_I(\mathbf{y} - \mathbf{x}) = \mathbf{0}$. Since $\mathbf{A}_I$ has n columns and rank n , its null space is $\{\mathbf{0}\}$ by Theorem 3.5.7, so $\mathbf{y} = \mathbf{x}$, and then $\mathbf{z} = 2\mathbf{x} - \mathbf{y} = \mathbf{x}$. By Lemma 18.8.1, $\mathbf{x}$ is a vertex.

($\Rightarrow$) Suppose $\mathbf{A}_I$ has rank less than n (this includes $I = \emptyset$). By Theorem 3.5.7 its null space is nonzero; choose $\mathbf{u} \neq \mathbf{0}$ with $\mathbf{a}_i^\top \mathbf{u} = 0$ for every $i \in I$. For $i \notin I$, the slack $b_i - \mathbf{a}_i^\top \mathbf{x}$ is positive. Let $\varepsilon > 0$ be smaller than $(b_i - \mathbf{a}_i^\top \mathbf{x})/|\mathbf{a}_i^\top \mathbf{u}|$ for every $i \notin I$ with $\mathbf{a}_i^\top \mathbf{u} \neq 0$; there are finitely many such i , so such an ε exists, and if there are none, take $\varepsilon = 1$. Then for $i \in I$, $\mathbf{a}_i^\top (\mathbf{x} \pm \varepsilon \mathbf{u}) = b_i$, and for $i \notin I$,

$$\mathbf{a}_i^\top (\mathbf{x} \pm \varepsilon \mathbf{u}) \leq \mathbf{a}_i^\top \mathbf{x} + \varepsilon |\mathbf{a}_i^\top \mathbf{u}| \leq b_i.$$

So $\mathbf{x} \pm \varepsilon \mathbf{u} \in P$, they differ because $\mathbf{u} \neq \mathbf{0}$, and $\mathbf{x}$ is their midpoint. Hence $\mathbf{x}$ is not a vertex. This proves the theorem. ■

The payoff is finiteness.

Corollary 18.9.3. A Polyhedron Has Finitely Many Vertices

Let $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ with $\mathbf{A} \in M_{m \times n}(\mathbb{R})$. Every vertex of P is the unique solution of $\mathbf{A}_J \mathbf{x} = \mathbf{b}_J$ for some set J of n row indices with $\mathbf{A}_J$ invertible. In particular P has at most $\binom{m}{n}$ vertices.

Idea. By the theorem, n independent active rows pin a vertex down, so a vertex is determined by which n rows those are.

Proof. Let $\mathbf{x}$ be a vertex. By Theorem 18.9.2, some n active rows, with index set $J \subseteq I(\mathbf{x})$, are linearly independent. The $n \times n$ matrix $\mathbf{A}_J$ then has independent rows, so its rank is n (Theorem 3.5.5), its null space is $\{\mathbf{0}\}$ (Theorem 3.5.7), and it is invertible (Theorem 3.4.4 (b)). As $J \subseteq I(\mathbf{x})$, $\mathbf{A}_J \mathbf{x} = \mathbf{b}_J$, so $\mathbf{x} = \mathbf{A}_J^{-1} \mathbf{b}_J$ is determined by J . Different vertices therefore come from different sets J , and there are $\binom{m}{n}$ sets of n indices. ■

This gives the finite recipe for a vertex list: for each choice of n rows with $\mathbf{A}_J$ invertible, solve $\mathbf{A}_J \mathbf{x} = \mathbf{b}_J$ and keep the solution if it satisfies the remaining inequalities.

Example. The vertices of a quadrilateral

Find the vertices of

$$P = \{\mathbf{x} \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, x_1 + 2x_2 \leq 4, 3x_1 + x_2 \leq 6\},$$

and the maximum of $\varphi(\mathbf{x}) = x_1 + x_2$ over P .

Solution. Number the constraints (1) $-x_1 \leq 0$, (2) $-x_2 \leq 0$, (3) $x_1 + 2x_2 \leq 4$, (4) $3x_1 + x_2 \leq 6$. Any two of the four rows are linearly independent, so each of the $\binom{4}{2} = 6$ pairs gives one basic point. We solve each pair as equations and test the other two inequalities.

pair	basic point	other constraints
(1), (2)	(0, 0)	$0 \leq 4, 0 \leq 6$: feasible
(1), (3)	(0, 2)	$-2 \leq 0, 2 \leq 6$: feasible
(1), (4)	(0, 6)	$12 \leq 4$ fails
(2), (3)	(4, 0)	$12 \leq 6$ fails
(2), (4)	(2, 0)	$-2 \leq 0, 2 \leq 4$: feasible
(3), (4)	$(\frac{8}{5}, \frac{6}{5})$	$-\frac{8}{5} \leq 0, -\frac{6}{5} \leq 0$: feasible

For the last pair, subtracting (3) from twice (4) gives $5x_1 = 8$, so $x_1 = \frac{8}{5}$ and $x_2 = 6 - 3x_1 = \frac{6}{5}$; check $\frac{8}{5} + \frac{12}{5} = 4$. By [Theorem 18.9.2](#) the vertices are the four feasible basic points $(0, 0)$, $(0, 2)$, $(2, 0)$, and $(\frac{8}{5}, \frac{6}{5})$, and P is a quadrilateral. The two infeasible basic points are where two of the lines cross outside P .

P is bounded, since $0 \leq x_1 \leq 2$ and $0 \leq x_2 \leq 2$ on it, so it is compact ([Proposition 18.9.1](#), Corollary 16.2.5) and nonempty. By Corollary 18.8.6 the maximum of φ is attained at a vertex. The values are $0, 2, 2, \frac{14}{5}$, so the maximum is $\frac{14}{5}$, at $(\frac{8}{5}, \frac{6}{5})$.

Quick check. A polyhedron in $\mathbb{R}^2$ is given by $m = 5$ inequalities. What is the largest number of vertices the corollary allows? Could a polyhedron in $\mathbb{R}^3$ given by $m = 2$ inequalities have a vertex?

Solution. At most $\binom{5}{2} = 10$. In $\mathbb{R}^3$ a vertex needs 3 linearly independent active rows, and there are only 2 rows in all, so no: the rank of the active rows is at most $2 < 3$, and by [Theorem 18.9.2](#) such a polyhedron has no vertices. Geometrically, the intersection of two half-spaces in $\mathbb{R}^3$ always contains a whole line through each of its points.

The recipe is honest but slow: $\binom{m}{n}$ grows very fast. The **simplex method**, the classical algorithm for linear programs, uses the same theorem more cleverly. It moves from a vertex to a neighboring vertex, one sharing $n - 1$ of its independent active constraints, along an edge on which the objective improves, and stops at a vertex where no edge improves it.

18.9.3 Fourier–Motzkin elimination

Now the other direction. Given a polytope $\text{conv}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, how do we find inequalities for it? The hull is the set of $\mathbf{x}$ for which **there exist** weights $\mathbf{t} \geq \mathbf{0}$ with $\sum_j t_j = 1$ and $\mathbf{x} = \sum_j t_j \mathbf{v}_j$. The conditions on the pair $(\mathbf{x}, \mathbf{t})$ are linear inequalities, so the set of pairs is a polyhedron in $\mathbb{R}^{n+k}$, and the polytope is its image under the projection that forgets $\mathbf{t}$. So we need to know that a projection of a polyhedron is a polyhedron, and a way to compute it. For equations, eliminating a variable is Gaussian elimination. For inequalities the device is to pair every upper bound on the variable with every lower bound.

Write a point of $\mathbb{R}^{n+1}$ as $(\mathbf{x}, s)$ with $\mathbf{x} \in \mathbb{R}^n$ and $s \in \mathbb{R}$, and let $\pi(\mathbf{x}, s) = \mathbf{x}$.

Lemma 18.9.4. Fourier–Motzkin Elimination

Let $P = \{(\mathbf{x}, s) \in \mathbb{R}^n \times \mathbb{R} : \mathbf{a}_i^\top \mathbf{x} + c_i s \leq b_i \text{ for } i = 1, \dots, m\}$. Split the indices by the sign of the coefficient of s : $I_+ = \{i : c_i > 0\}$, $I_- = \{i : c_i < 0\}$, $I_0 = \{i : c_i = 0\}$. Then $\pi(P)$ is the set of $\mathbf{x} \in \mathbb{R}^n$ satisfying

$$\begin{aligned} \mathbf{a}_i^\top \mathbf{x} &\leq b_i && \text{for } i \in I_0, \\ (c_i \mathbf{a}_j - c_j \mathbf{a}_i)^\top \mathbf{x} &\leq c_i b_j - c_j b_i && \text{for } i \in I_+, j \in I_-. \end{aligned}$$

In particular $\pi(P)$ is a polyhedron, and if every $b_i = 0$, every right-hand side above is 0.

Idea. For fixed $\mathbf{x}$, each row with $c_i > 0$ is an upper bound on s , each row with $c_i < 0$ is a lower bound, and a row with $c_i = 0$ does not mention s . A number s satisfying all of them exists exactly when every lower bound is at most every upper bound, and those comparisons are the new rows. Each new row is the combination $c_i \cdot (\text{row } j) + (-c_j) \cdot (\text{row } i)$, with both multipliers positive and chosen to cancel s .

Proof. Let Q be the set described by the new rows. For $i \in I_+$ and $j \in I_-$, put

$$U_i(\mathbf{x}) = \frac{b_i - \mathbf{a}_i^\top \mathbf{x}}{c_i}, \quad L_j(\mathbf{x}) = \frac{b_j - \mathbf{a}_j^\top \mathbf{x}}{c_j}.$$

Dividing row i by $c_i > 0$ shows that it says $s \leq U_i(\mathbf{x})$, and dividing row j by $c_j < 0$, which reverses the inequality, shows that it says $s \geq L_j(\mathbf{x})$. Multiplying $L_j(\mathbf{x}) \leq U_i(\mathbf{x})$ by the positive number $-c_i c_j$ and rearranging gives exactly the new row for the pair (i, j) :

$$L_j(\mathbf{x}) \leq U_i(\mathbf{x}) \iff (c_i \mathbf{a}_j - c_j \mathbf{a}_i)^\top \mathbf{x} \leq c_i b_j - c_j b_i.$$

($\subseteq$) If $(\mathbf{x}, s) \in P$, then $L_j(\mathbf{x}) \leq s \leq U_i(\mathbf{x})$ for every pair, and the rows in I_0 hold since they do not involve s . So $\mathbf{x} \in Q$.

($\supseteq$) Let $\mathbf{x} \in Q$. If $I_- \neq \emptyset$, let $s = \max_{j \in I_-} L_j(\mathbf{x})$; if $I_- = \emptyset$ and $I_+ \neq \emptyset$, let $s = \min_{i \in I_+} U_i(\mathbf{x})$; if both are empty, let $s = 0$. In each case $L_j(\mathbf{x}) \leq s \leq U_i(\mathbf{x})$ for all $j \in I_-$ and $i \in I_+$: in the first case because every $L_j(\mathbf{x}) \leq s$ by the choice of s , and $s = L_{j_0}(\mathbf{x}) \leq U_i(\mathbf{x})$ for every i since $\mathbf{x} \in Q$; the second case is the same with the roles exchanged. The rows in I_0 hold because $\mathbf{x} \in Q$. So $(\mathbf{x}, s) \in P$ and $\mathbf{x} \in \pi(P)$.

The last sentence is visible: the new right-hand sides are b_i and $c_i b_j - c_j b_i$.

■

As an algorithm:

Algorithm. Fourier–Motzkin elimination of one variable. Input: a system of m inequalities $\mathbf{a}_i^\top \mathbf{x} + c_i s \leq b_i$ in the unknowns $\mathbf{x} \in \mathbb{R}^n$ and $s \in \mathbb{R}$.

1. Sort the rows into I_+ ($c_i > 0$), I_- ($c_i < 0$) and I_0 ($c_i = 0$).
2. Copy every row of I_0 , without its s term.
3. For every pair $i \in I_+$, $j \in I_-$, write down the row $c_i \cdot (\text{row } j) + (-c_j) \cdot (\text{row } i)$, which is $(c_i \mathbf{a}_j - c_j \mathbf{a}_i)^\top \mathbf{x} \leq c_i b_j - c_j b_i$.
4. (Optional.) Delete every new row of the form $0 \leq \beta$ with $\beta \geq 0$, which every $\mathbf{x}$ satisfies.
A row $0 \leq \beta$ with $\beta < 0$ means the projection is empty.

Output: $|I_0| + |I_+||I_-|$ inequalities in $\mathbf{x}$ alone, at most, whose solution set is the projection of the input's solution set. To eliminate several variables, repeat.

When the variable to be eliminated appears in an **equation** with a nonzero coefficient, there is a shortcut: solve the equation for it and substitute into the other rows. The projection is unchanged, since the equation determines the eliminated variable from the others, and substitution is exactly Gaussian elimination.

Example. A triangle, both ways

Let $Q = \text{conv}\{(0,0), (3,0), (0,3), (1,1)\} \subseteq \mathbb{R}^2$. Find a system of inequalities describing Q , and then recover the vertices of Q from it.

Solution. From points to inequalities. By Theorem 18.1.3, $\mathbf{x} \in Q$ if and only if there are $t_1, \dots, t_4 \geq 0$ with $\sum_j t_j = 1$ and

$$\mathbf{x} = t_1(0,0) + t_2(3,0) + t_3(0,3) + t_4(1,1),$$

that is,

$$x_1 = 3t_2 + t_4, \quad x_2 = 3t_3 + t_4.$$

Three equations involve the weights, so we substitute three of them away and keep $s = t_4$: $t_2 = \frac{1}{3}(x_1 - s)$, $t_3 = \frac{1}{3}(x_2 - s)$, and

$$t_1 = 1 - t_2 - t_3 - s = 1 - \frac{1}{3}(x_1 + x_2 + s).$$

The four conditions $t_j \geq 0$, each multiplied by 3 where needed, become a system in (x_1, x_2, s) :

$$\begin{array}{ll} (1) & x_1 + x_2 + s \leq 3, \\ (2) & -x_1 + s \leq 0, \\ (3) & -x_2 + s \leq 0, \\ (4) & -s \leq 0. \end{array}$$

Run the algorithm on s . The coefficients of s are 1, 1, 1, -1, so $I_+ = \{1, 2, 3\}$, $I_- = \{4\}$, and $I_0 = \emptyset$. Each pair $(i, 4)$ has $c_i = 1$ and $-c_4 = 1$, so the new row is row i plus row 4, which just deletes the s :

$$x_1 + x_2 \leq 3, \quad -x_1 \leq 0, \quad -x_2 \leq 0.$$

By Lemma 18.9.4, $Q = \{x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 3\}$, the triangle T of the examples above. The generator $(1,1)$ left no trace, as it should: it lies inside the triangle.

From inequalities to vertices. The three rows $(-1, 0)$, $(0, -1)$, $(1, 1)$ are pairwise independent, so each of the $\binom{3}{2} = 3$ pairs gives a basic point: $(0,0)$, $(0,3)$, and $(3,0)$. Each satisfies the third inequality, so all three are vertices by Theorem 18.9.2. At $(1,1)$ no constraint is active, so $I = \emptyset$ and it is not a vertex, which matches Proposition 18.8.2: the extreme points of the hull are among the four generators, and $(1,1)$ is not one of them.

Warning. Fourier–Motzkin does not return a minimal description. Eliminating s from $-x_1 + s \leq 0$, $s \leq 1$, $-s \leq 0$, and $x_1 - s \leq 1$ pairs two upper bounds with two lower bounds, and gives $-x_1 \leq 0$, $x_1 \leq 2$, and twice the useless row $0 \leq 1$. Even without trivial rows, the output can contain inequalities implied by the others. A single step can turn m rows into as many as $(m/2)^2$, and eliminating many variables compounds this, so the method is a proof device and a hand tool for small systems, not an efficient algorithm.

18.9.4 The Minkowski–Weyl theorem

Cones get the same treatment. Section 5 wrote $\text{cone}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ for the set of all nonnegative combinations $\sum_j c_j \mathbf{v}_j$, $c_j \geq 0$, a convex cone (Definition 18.5.1), and called such cones **finitely generated**; we allow $k = 0$, with $\text{cone}() = \{\mathbf{0}\}$. The inequality side is a homogeneous system.

Definition 18.9.4. Polyhedral Cone

A **polyhedral cone** in $\mathbb{R}^n$ is a set of the form $C = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{0}\}$ for some $m \geq 0$ and $\mathbf{A} \in M_{m \times n}(\mathbb{R})$.

A polyhedral cone is a polyhedron, and it is a convex cone: it contains $\mathbf{0}$, and if $\mathbf{A}\mathbf{u} \leq \mathbf{0}$, $\mathbf{A}\mathbf{v} \leq \mathbf{0}$, and $a, b \geq 0$, then $\mathbf{A}(a\mathbf{u} + b\mathbf{v}) = a\mathbf{A}\mathbf{u} + b\mathbf{A}\mathbf{v} \leq \mathbf{0}$. By [Proposition 18.9.1](#) it is closed, while for finitely generated cones closedness took the work of Theorem 18.5.4.

Theorem 18.9.5. Minkowski–Weyl Theorem

- (a) A subset of $\mathbb{R}^n$ is a polytope if and only if it is a bounded polyhedron.
- (b) A subset of $\mathbb{R}^n$ is a finitely generated cone if and only if it is a polyhedral cone.

Idea. Four implications, two tools.

① *Bounded polyhedron $\Rightarrow$ polytope.* This is Section 8 plus the vertex theorem: a bounded polyhedron is compact, so it is the hull of its vertices by Minkowski's theorem, and it has finitely many vertices.

② *Polyhedral cone $\Rightarrow$ finitely generated.* Cut the cone with the cube $[-1, 1]^n$. The result is a bounded polyhedron, hence by ① the hull of finitely many points, and scaling those points generates the cone.

③ *Polytope $\Rightarrow$ polyhedron,* and ④ *finitely generated $\Rightarrow$ polyhedral.* Write the set as the projection of a polyhedron in a bigger space, with the weights as extra coordinates, and eliminate the weights one at a time by Fourier–Motzkin. For a cone all right-hand sides are 0 and stay 0.

Proof. ① **A bounded polyhedron is a polytope.** Let $P = \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ be bounded. If $P = \emptyset$, then $P = \text{conv } \emptyset$. Otherwise P is closed and convex by [Proposition 18.9.1](#), and bounded, so it is compact by Corollary 16.2.5. By Theorem 18.8.5, $P = \text{conv}(\text{ext } P)$, and $\text{ext } P$ is finite by [Corollary 18.9.3](#). So P is a polytope.

② **A polyhedral cone is finitely generated.** Let $C = \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{0}\}$, and let $P = C \cap [-1, 1]^n$, the polyhedron given by the rows of $\mathbf{A}$ together with the $2n$ rows $\pm x_i \leq 1$. It is bounded and contains $\mathbf{0}$, so by ① $P = \text{conv}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, where $\mathbf{v}_1, \dots, \mathbf{v}_k$ are its vertices; in particular every $\mathbf{v}_j \in P \subseteq C$. We claim $C = \text{cone}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. If $\mathbf{c} \geq \mathbf{0}$, then $\mathbf{A} \sum_j c_j \mathbf{v}_j = \sum_j c_j \mathbf{A}\mathbf{v}_j \leq \mathbf{0}$, since each $\mathbf{A}\mathbf{v}_j \leq \mathbf{0}$ and each $c_j \geq 0$; so every such combination lies in C . Conversely let $\mathbf{x} \in C$. If $\mathbf{x} = \mathbf{0}$, take all $c_j = 0$. Otherwise put $r = \|\mathbf{x}\|_\infty > 0$. Then $\mathbf{A}(\mathbf{x}/r) = \frac{1}{r} \mathbf{A}\mathbf{x} \leq \mathbf{0}$ and every entry of $\mathbf{x}/r$ lies in $[-1, 1]$, so $\mathbf{x}/r \in P$, say $\mathbf{x}/r = \sum_j t_j \mathbf{v}_j$ with $t_j \geq 0$. Then $\mathbf{x} = \sum_j (rt_j) \mathbf{v}_j$ with $rt_j \geq 0$.

③ **A polytope is a bounded polyhedron.** Let $Q = \text{conv}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, and let $\mathbf{V}$ have columns $\mathbf{v}_j$. It is bounded, as in example 2 after [Definition 18.9.2](#). By Theorem 18.1.3, Q is the image, under the map $(\mathbf{x}, \mathbf{t}) \mapsto \mathbf{x}$ that forgets the last k coordinates, of

$$\begin{aligned} \hat{P} = \{(\mathbf{x}, \mathbf{t}) \in \mathbb{R}^n \times \mathbb{R}^k : & \mathbf{x} - \mathbf{V}\mathbf{t} \leq \mathbf{0}, \quad -\mathbf{x} + \mathbf{V}\mathbf{t} \leq \mathbf{0}, \\ & -\mathbf{t} \leq \mathbf{0}, \quad \mathbf{1}^\top \mathbf{t} \leq 1, \quad -\mathbf{1}^\top \mathbf{t} \leq -1\}, \end{aligned}$$

a polyhedron in $\mathbb{R}^{n+k}$. Forgetting the last k coordinates is forgetting one coordinate k times in a row, and by [Lemma 18.9.4](#) each step takes a polyhedron to a polyhedron. So Q is a polyhedron.

④ **A finitely generated cone is polyhedral.** Let $C = \text{cone}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \{\mathbf{V}\mathbf{c} : \mathbf{c} \geq \mathbf{0}\}$, where $\mathbf{V}$ has columns $\mathbf{v}_j$. It is the image under the same map of $\{(\mathbf{x}, \mathbf{c}) : \mathbf{x} - \mathbf{V}\mathbf{c} \leq \mathbf{0}, -\mathbf{x} + \mathbf{V}\mathbf{c} \leq \mathbf{0}, -\mathbf{c} \leq \mathbf{0}\}$, whose right-hand sides are all 0. By the last sentence of [Lemma 18.9.4](#), eliminating the k coordinates of $\mathbf{c}$ one at a time keeps every right-hand side 0, and the rows copied from I_0 keep theirs as well. So $C = \{\mathbf{x} : \mathbf{B}\mathbf{x} \leq \mathbf{0}\}$ for the matrix $\mathbf{B}$ of the final rows (if no rows remain, $C = \mathbb{R}^n = \{\mathbf{x} : \mathbf{0} \leq \mathbf{0}\}$). This completes the proof. ■

Two consequences show what the theorem is for. First, **the intersection of two polytopes is a polytope**. By (a) each is a bounded polyhedron; putting the two systems of inequalities together describes the intersection, which is therefore a bounded polyhedron, hence a polytope by (a) again. From the point description alone this is not obvious at all: the corners of the intersection are, in general, new points found nowhere in the two lists. Second, (b) gives a new proof of Theorem 18.5.4, since every polyhedral cone is closed. In that proof the difficulty of Section 5 has moved into Fourier–Motzkin elimination. The same elimination shows that a linear image $\mathbf{M}(C)$ of a polyhedral cone $C = \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{0}\}$ is polyhedral: it is the image, under the map forgetting $\mathbf{x}$, of $\{(\mathbf{y}, \mathbf{x}) : \mathbf{y} - \mathbf{M}\mathbf{x} \leq \mathbf{0}, -\mathbf{y} + \mathbf{M}\mathbf{x} \leq \mathbf{0}, \mathbf{A}\mathbf{x} \leq \mathbf{0}\}$, and the argument of ④ applies. So a linear image of a polyhedral cone is closed, and a closed convex cone with a nonclosed linear image, like the one in Section 5, can never be polyhedral.

The disk B_2 is not a polyhedron. If it were, it would be a bounded polyhedron, hence a polytope by (a), and it is not one.

18.9.5 The cube and the cross-polytope

The two unit balls with corners from Section 8 are the standard example of the theorem, and they illustrate a duality between the two descriptions.

Example. The cube and the cross-polytope

For $n \geq 1$, describe B_∞ and B_1 both as polytopes and as polyhedra, and find their vertices from the inequalities by [Theorem 18.9.2](#).

Solution. The cube. B_∞ is the polyhedron given by the $2n$ rows $\mathbf{e}_i^\top \mathbf{x} \leq 1$ and $-\mathbf{e}_i^\top \mathbf{x} \leq 1$. At a point $\mathbf{x}$ of it, the active rows are $\mathbf{e}_i$ for each i with $x_i = 1$ and $-\mathbf{e}_i$ for each i with $x_i = -1$; they are distinct standard vectors up to sign, so they are linearly independent, and their number is the rank. The rank is n exactly when every $|x_i| = 1$. So the vertices are the 2^n sign vectors, as Example (The diamond, the cube, and the sphere) found directly, and B_∞ is their hull.

The cross-polytope. By Example (The diamond, the cube, and the sphere) and Theorem 18.8.5, $B_1 = \text{conv}\{\pm \mathbf{e}_1, \dots, \pm \mathbf{e}_n\}$. For its inequalities, let $\mathbf{s}$ run over the 2^n sign vectors. For every $\mathbf{s}$, $\mathbf{s}^\top \mathbf{x} = \sum_i s_i x_i \leq \sum_i |x_i| = \|\mathbf{x}\|_1$, with equality when $s_i = 1$ for $x_i \geq 0$ and $s_i = -1$ for $x_i < 0$. So $\|\mathbf{x}\|_1 = \max_{\mathbf{s}} \mathbf{s}^\top \mathbf{x}$, and

$$B_1 = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{s}^\top \mathbf{x} \leq 1 \text{ for every sign vector } \mathbf{s}\},$$

a polyhedron with 2^n rows. At $\mathbf{e}_1$, the active rows are the $\mathbf{s}$ with $s_1 = 1$. They have rank n : they include $\mathbf{1}$, and for each $j \geq 2$ the vector $\mathbf{1} - 2\mathbf{e}_j$, so their span contains $\frac{1}{2}(\mathbf{1} - (\mathbf{1} - 2\mathbf{e}_j)) = \mathbf{e}_j$ for $j \geq 2$ and then $\mathbf{e}_1 = \mathbf{1} - \sum_{j \geq 2} \mathbf{e}_j$. The same holds at every $\pm\mathbf{e}_i$. Every other point $\mathbf{x}$ of B_1 has active rows of rank less than n . If $\|\mathbf{x}\|_1 < 1$, no row is active. If $\|\mathbf{x}\|_1 = 1$, then $\mathbf{x}$ has two nonzero entries x_i, x_j with $i \neq j$, since a point with one nonzero entry and $\|\mathbf{x}\|_1 = 1$ is some $\pm\mathbf{e}_i$; let $\sigma_i, \sigma_j \in \{\pm 1\}$ be their signs. An active $\mathbf{s}$ has $\sum_k s_k x_k = \sum_k |x_k|$ with each $s_k x_k \leq |x_k|$, so $s_k x_k = |x_k|$ for every k , and in particular $s_i = \sigma_i, s_j = \sigma_j$. Hence $\mathbf{u} = \sigma_i \mathbf{e}_i - \sigma_j \mathbf{e}_j \neq \mathbf{0}$ has $\mathbf{s}^\top \mathbf{u} = \sigma_i^2 - \sigma_j^2 = 0$ for every active $\mathbf{s}$, and the matrix of active rows has rank less than n by Theorem 3.5.7. By Theorem 18.9.2, the vertices are exactly the $2n$ points $\pm\mathbf{e}_i$.

Duality. The rows of the cube are $\pm\mathbf{e}_i$, which are the vertices of the cross-polytope; the rows of the cross-polytope are the sign vectors, which are the vertices of the cube. In $\mathbb{R}^3$, the cube has 8 vertices and 6 inequalities, the octahedron 6 vertices and 8 inequalities. This exchange is the polytope form of the fact that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are dual norms (Definition 18.4.4): the maximum of $\mathbf{y}^\top \mathbf{x}$ over $\mathbf{x} \in B_1$ is $\|\mathbf{y}\|_\infty$, and over $\mathbf{x} \in B_\infty$ it is $\|\mathbf{y}\|_1$.

18.9.6 Exercises

18.9.6.1 A. Check your understanding

Exercise. A1

- State the definitions of a polytope and of a polyhedron in $\mathbb{R}^n$.
- True or false: every polyhedron is a polytope. Justify your answer.
- True or false: the intersection of two polytopes in $\mathbb{R}^n$ is a polytope. Justify your answer.
- State the criterion of Theorem 18.9.2 for a point of $\{\mathbf{x} : \mathbf{Ax} \leq \mathbf{b}\}$ to be a vertex.
- True or false: every inequality produced by Fourier–Motzkin elimination (Lemma 18.9.4) is needed to describe the projection. Justify your answer.

Solution.

- A polytope is the convex hull of finitely many points of $\mathbb{R}^n$. A polyhedron is a set $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} \leq \mathbf{b}\}$ for some $\mathbf{A} \in M_{m \times n}(\mathbb{R})$ and $\mathbf{b} \in \mathbb{R}^m$, that is, an intersection of finitely many sets $\{\mathbf{x} : \mathbf{a}_i^\top \mathbf{x} \leq b_i\}$.
- False. The half-plane $\{x_2 \geq 0\} \subseteq \mathbb{R}^2$ is a polyhedron, with one row, but it is unbounded, while every polytope is bounded.
- True. By Theorem 18.9.5 (a), each is a bounded polyhedron. Listing both systems of inequalities together describes the intersection, which is therefore a polyhedron, and it is bounded because it lies inside one of the two. By Theorem 18.9.5 (a) again, it is a polytope.
- A point $\mathbf{x}$ with $\mathbf{Ax} \leq \mathbf{b}$ is a vertex if and only if the rows $\mathbf{a}_i$ with $\mathbf{a}_i^\top \mathbf{x} = b_i$ include n linearly independent ones.
- False. In the warning after Example (A triangle, both ways), two of the four output rows are $0 \leq 1$, which every point satisfies.

18.9.6.2 B. Practice

Exercise. B1: Vertices of a pentagon

Find all vertices of

$$P = \{\mathbf{x} \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 4, x_1 - x_2 \leq 2, -x_1 + 2x_2 \leq 5\},$$

and the maximum of $2x_1 + 3x_2$ over P . Hence write P as a polytope.

Solution. Number the rows (1) $-x_1 \leq 0$, (2) $-x_2 \leq 0$, (3) $x_1 + x_2 \leq 4$, (4) $x_1 - x_2 \leq 2$, (5) $-x_1 + 2x_2 \leq 5$. Any two rows are independent (no two are multiples of each other), so each of the $\binom{5}{2} = 10$ pairs gives one basic point.

pair	point	verdict
(1), (2)	(0, 0)	feasible
(1), (3)	(0, 4)	(5): $8 \leq 5$ fails
(1), (4)	(0, -2)	(2) fails
(1), (5)	(0, $\frac{5}{2}$)	feasible
(2), (3)	(4, 0)	(4): $4 \leq 2$ fails
(2), (4)	(2, 0)	feasible
(2), (5)	(-5, 0)	(1) fails
(3), (4)	(3, 1)	feasible
(3), (5)	(1, 3)	feasible
(4), (5)	(9, 7)	(3): $16 \leq 4$ fails

For instance, (3) and (4) added give $2x_1 = 6$, so $\mathbf{x} = (3, 1)$; (3) and (5) added give $3x_2 = 9$, so $\mathbf{x} = (1, 3)$; (4) and (5) added give $x_2 = 7$, then $x_1 = 9$. For the feasible points, the checks are routine: at $(0, \frac{5}{2})$, $0 + \frac{5}{2} \leq 4$ and $-\frac{5}{2} \leq 2$; at $(2, 0)$, $2 \leq 4$ and $-2 \leq 5$; at $(3, 1)$, $-3 + 2 \leq 5$; at $(1, 3)$, $1 - 3 \leq 2$. By [Theorem 18.9.2](#) the vertices are $(0, 0)$, $(2, 0)$, $(3, 1)$, $(1, 3)$, $(0, \frac{5}{2})$.

P is bounded, since $0 \leq x_1, x_2 \leq 4$ on it by rows (1) to (3), and nonempty. So by Corollary 18.8.6 the maximum of $2x_1 + 3x_2$ is attained at a vertex. The values are $0, 4, 9, 11, \frac{15}{2}$, so the maximum is 11, at $(1, 3)$. By the proof of [Theorem 18.9.5](#) (a), $P = \text{conv}\{(0, 0), (2, 0), (3, 1), (1, 3), (0, \frac{5}{2})\}$.

Exercise. B2: Inequalities for a quadrilateral

Use Fourier–Motzkin elimination to find a system of inequalities describing the quadrilateral

$$Q = \text{conv}\{(0, 0), (2, 0), (2, 1), (0, 2)\}.$$

Hence find the vertices of Q .

Hint: substitute away t_1, t_2, t_4 and keep $s = t_3$, the weight of $(2, 1)$.

Solution. $\mathbf{x} \in Q$ if and only if there are $t_1, \dots, t_4 \geq 0$ with $\sum t_j = 1$, $x_1 = 2t_2 + 2t_3$, and

$x_2 = t_3 + 2t_4$. With $s = t_3$: $t_2 = \frac{1}{2}(x_1 - 2s)$, $t_4 = \frac{1}{2}(x_2 - s)$, and

$$t_1 = 1 - t_2 - s - t_4 = 1 - \frac{1}{2}x_1 - \frac{1}{2}x_2 + \frac{1}{2}s.$$

The conditions $2t_1 \geq 0$, $2t_2 \geq 0$, $2t_4 \geq 0$, $t_3 \geq 0$ read

$$\begin{array}{ll} (1) & x_1 + x_2 - s \leq 2, \\ (2) & -x_1 + 2s \leq 0, \\ (3) & -x_2 + s \leq 0, \\ (4) & -s \leq 0. \end{array}$$

The coefficients of s are $-1, 2, 1, -1$, so $I_+ = \{2, 3\}$ and $I_- = \{1, 4\}$. The four pairs give:

$$\begin{array}{ll} (2), (1): & 2 \cdot (1) + 1 \cdot (2) = x_1 + 2x_2 \leq 4, \\ (2), (4): & 2 \cdot (4) + 1 \cdot (2) = -x_1 \leq 0, \\ (3), (1): & 1 \cdot (1) + 1 \cdot (3) = x_1 \leq 2, \\ (3), (4): & 1 \cdot (4) + 1 \cdot (3) = -x_2 \leq 0. \end{array}$$

So $Q = \{0 \leq x_1 \leq 2, x_2 \geq 0, x_1 + 2x_2 \leq 4\}$ by [Lemma 18.9.4](#). Its rows are $(-1, 0), (1, 0), (0, -1), (1, 2)$. The pair $\pm(1, 0)$ is dependent; the other five pairs give the basic points $(0, 0), (0, 2), (2, 0), (2, 1)$, and $(4, 0)$ (from $x_2 = 0, x_1 + 2x_2 = 4$), and the last violates $x_1 \leq 2$. So by [Theorem 18.9.2](#) the vertices are the four given points, and none of them was redundant.

Exercise. B3: Which of these are vertices?

Let $P = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} \geq \mathbf{0}, x_1 + x_2 + x_3 \leq 3, x_1 + x_2 \leq 2\}$. Determine which of the following points are vertices of P . Justify your answer.

- (a) $(0, 2, 1)$
- (b) $(1, 1, 1)$
- (c) $(0, 0, 3)$
- (d) $(3, 0, 0)$
- (e) $(1, 0, 0)$

Solution. We use [Theorem 18.9.2](#): check that the point lies in P , list its active rows, and compute their rank. The rows are $-\mathbf{e}_1, -\mathbf{e}_2, -\mathbf{e}_3, (1, 1, 1)$, and $(1, 1, 0)$.

- (a) $(0, 2, 1) \in P$: the entries are ≥ 0 , $0 + 2 + 1 = 3 \leq 3$, $0 + 2 \leq 2$. Active: $-\mathbf{e}_1, (1, 1, 1), (1, 1, 0)$. Their span contains $(1, 1, 1) - (1, 1, 0) = \mathbf{e}_3$, $\mathbf{e}_1$, and $(1, 1, 0) - \mathbf{e}_1 = \mathbf{e}_2$, so it is $\mathbb{R}^3$. Rank 3: **a vertex**.
- (b) $(1, 1, 1) \in P$, with active rows $(1, 1, 1)$ and $(1, 1, 0)$ only. Rank $2 < 3$: **not a vertex**. Indeed it is the midpoint of $(0, 2, 1)$ and $(2, 0, 1)$, both in P .
- (c) $(0, 0, 3) \in P$, with active rows $-\mathbf{e}_1, -\mathbf{e}_2, (1, 1, 1)$, which are independent (the third has a nonzero third entry, the first two do not). Rank 3: **a vertex**.
- (d) $(3, 0, 0) \notin P$, since $3 + 0 = 3 > 2$. It is a basic point, the solution of $x_2 = 0, x_3 = 0, x_1 + x_2 + x_3 = 3$, but not a feasible one: **not a vertex**.
- (e) $(1, 0, 0) \in P$, with active rows $-\mathbf{e}_2$ and $-\mathbf{e}_3$ only. Rank $2 < 3$: **not a vertex**. It is the midpoint of $(0, 0, 0)$ and $(2, 0, 0)$.

18.9.6.3 C. Going deeper

Exercise. C1: Unbounded polyhedra

Let $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} \leq \mathbf{b}\}$ be a **nonempty** polyhedron, and let $C = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} : \mathbf{Ax} - t\mathbf{b} \leq \mathbf{0}, -t \leq 0\}$.

- (a) Prove that $\mathbf{x} \in P$ if and only if $(\mathbf{x}, 1) \in C$.
- (b) By [Theorem 18.9.5](#) (b), C is generated by finitely many vectors $(\mathbf{w}_j, t_j)$, $j = 1, \dots, k$. Prove that every $t_j \geq 0$, and that

$$P = \text{conv}\{\mathbf{w}_j/t_j : t_j > 0\} + \left\{ \sum_{t_j=0} c_j \mathbf{w}_j : c_j \geq 0 \right\},$$

the set of sums of a point of the polytope and a point of the cone. So every nonempty polyhedron is a polytope plus a finitely generated cone.

- (c) Check the statement for $P = \{\mathbf{x} \in \mathbb{R}^2 : x_1 \geq 0, 0 \leq x_2 \leq 1\}$, with the polytope $\text{conv}\{(0, 0), (0, 1)\}$ and the cone generated by $\mathbf{e}_1$.

Solution.

- (a) $(\mathbf{x}, 1) \in C$ says $\mathbf{Ax} - \mathbf{b} \leq \mathbf{0}$ and $-1 \leq 0$, which is $\mathbf{x} \in P$.
- (b) Each generator lies in C , so $-t_j \leq 0$. Split the indices into $J_+ = \{j : t_j > 0\}$ and $J_0 = \{j : t_j = 0\}$.
 $(\subseteq)$ Let $\mathbf{x} \in P$. By (a), $(\mathbf{x}, 1) = \sum_j c_j(\mathbf{w}_j, t_j)$ with $c_j \geq 0$. The last coordinate gives $1 = \sum_{j \in J_+} c_j t_j$, so $J_+ \neq \emptyset$, and

$$\mathbf{x} = \sum_{j \in J_+} (c_j t_j) \frac{\mathbf{w}_j}{t_j} + \sum_{j \in J_0} c_j \mathbf{w}_j,$$

where the weights $c_j t_j$ are nonnegative and add up to 1. So $\mathbf{x}$ lies in the right-hand side.

- $(\supseteq)$ Let $\mathbf{x} = \sum_{j \in J_+} s_j \mathbf{w}_j / t_j + \sum_{j \in J_0} c_j \mathbf{w}_j$ with $s_j \geq 0$, $\sum s_j = 1$, and $c_j \geq 0$. Then

$$(\mathbf{x}, 1) = \sum_{j \in J_+} \frac{s_j}{t_j} (\mathbf{w}_j, t_j) + \sum_{j \in J_0} c_j (\mathbf{w}_j, 0),$$

a nonnegative combination of generators, so $(\mathbf{x}, 1) \in C$ and $\mathbf{x} \in P$ by (a).

- (c) A point $(0, u) + c\mathbf{e}_1$ with $0 \leq u \leq 1$ and $c \geq 0$ is (c, u) , which satisfies $x_1 \geq 0$ and $0 \leq x_2 \leq 1$. Conversely $(x_1, x_2) \in P$ is $(0, x_2) + x_1 \mathbf{e}_1$ with $(0, x_2) \in \text{conv}\{(0, 0), (0, 1)\}$ and $x_1 \geq 0$. So $P = \text{conv}\{(0, 0), (0, 1)\} + \{c\mathbf{e}_1 : c \geq 0\}$.

Exercise. C2: Linear images of polyhedra

Let $\mathbf{M} \in M_{p \times n}(\mathbb{R})$.

- (a) Prove that if $P \subseteq \mathbb{R}^n$ is a polyhedron, then $\mathbf{M}(P) = \{\mathbf{Mx} : \mathbf{x} \in P\}$ is a polyhedron in $\mathbb{R}^p$, and that if P is a polyhedral cone, so is $\mathbf{M}(P)$.
- (b) Prove that if Q is a polytope, then $\mathbf{M}(Q)$ is a polytope, directly from [Definition 18.9.1](#).
- (c) Deduce that a linear image of a polyhedral cone is closed. Why does this not contradict

the existence of a closed convex cone with a nonclosed linear image?

Solution.

- (a) Let $P = \{\mathbf{x} : \mathbf{Ax} \leq \mathbf{b}\}$. Then $\mathbf{y} \in \mathbf{M}(P)$ if and only if there is $\mathbf{x}$ with $(\mathbf{y}, \mathbf{x})$ in

$$\hat{P} = \{(\mathbf{y}, \mathbf{x}) \in \mathbb{R}^p \times \mathbb{R}^n : \mathbf{y} - \mathbf{Mx} \leq \mathbf{0}, -\mathbf{y} + \mathbf{Mx} \leq \mathbf{0}, \mathbf{Ax} \leq \mathbf{b}\},$$

a polyhedron in $\mathbb{R}^{p+n}$. So $\mathbf{M}(P)$ is the image of $\hat{P}$ under the map forgetting the last n coordinates, and n applications of [Lemma 18.9.4](#) show that it is a polyhedron. If $\mathbf{b} = \mathbf{0}$, every right-hand side in $\hat{P}$ is 0, and stays 0 under elimination, so $\mathbf{M}(P)$ is a polyhedral cone.

- (b) If $Q = \text{conv}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, then by Theorem 18.1.3 and linearity,

$$\mathbf{M}\left(\sum_j t_j \mathbf{v}_j\right) = \sum_j t_j \mathbf{Mv}_j,$$

so $\mathbf{M}(Q) = \text{conv}\{\mathbf{Mv}_1, \dots, \mathbf{Mv}_k\}$, a polytope.

- (c) By (a), the image of a polyhedral cone is a polyhedral cone, which is closed by [Proposition 18.9.1](#). So a closed convex cone whose linear image is not closed cannot be polyhedral; by [Theorem 18.9.5](#) (b), it cannot be finitely generated either. This is consistent with Section 5, where closedness was proved only for finitely generated cones.

Exercise. C3: Vertices in standard form

Let $\mathbf{A} \in M_{m \times n}(\mathbb{R})$ have columns $\mathbf{c}_1, \dots, \mathbf{c}_n$, let $\mathbf{b} \in \mathbb{R}^m$, and let $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, the feasible region of a linear program in standard form.

- (a) Prove that $\mathbf{x} \in P$ is a vertex of P if and only if the columns $\mathbf{c}_j$ with $x_j > 0$ are linearly independent.
 (b) Deduce that a vertex of P has at most $\text{rank } \mathbf{A}$ positive entries.
 (c) Find the vertices of $P = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 1, \mathbf{x} \geq \mathbf{0}\}$, the standard triangle.

Solution.

- (a) Write P as a polyhedron with the rows $\mathbf{Ax} \leq \mathbf{b}$, $-\mathbf{Ax} \leq -\mathbf{b}$, and $-\mathbf{e}_j^\top \mathbf{x} \leq 0$. At $\mathbf{x} \in P$, every row of $\pm \mathbf{A}$ is active, and $-\mathbf{e}_j$ is active exactly when $x_j = 0$. Let $S = \{j : x_j > 0\}$. By [Theorem 18.9.2](#), $\mathbf{x}$ is a vertex if and only if the active rows have rank n , that is (Theorem 3.5.7), if and only if the only $\mathbf{u}$ with $\mathbf{Au} = \mathbf{0}$ and $u_j = 0$ for all $j \notin S$ is $\mathbf{u} = \mathbf{0}$. For such $\mathbf{u}$, $\mathbf{Au} = \sum_{j \in S} u_j \mathbf{c}_j$. So the condition says that $\sum_{j \in S} u_j \mathbf{c}_j = \mathbf{0}$ forces every $u_j = 0$, which is the linear independence of $\{\mathbf{c}_j : j \in S\}$.
 (b) The columns $\mathbf{c}_j$ lie in $\text{col}(\mathbf{A})$, of dimension $\text{rank } \mathbf{A}$, and an independent list in it has at most $\text{rank } \mathbf{A}$ vectors by Theorem 1.6.3 (a). By (a), a vertex has at most that many positive entries.
 (c) Here $\mathbf{A} = (1 \ 1 \ 1)$, of rank 1. By (b) a vertex has at most one positive entry, and since the entries add up to 1, exactly one, equal to 1. Conversely $\mathbf{e}_j \in P$, and its one positive entry corresponds to the column $(1) \neq \mathbf{0}$, an independent list. So the vertices are $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.