The Fundamental Theorem of Polytopes

Genova, May 2025

Abstract

"Proofs should only be communicated by consenting adults in private". (Victor Klee)

0 Definitions

Definition 1. A subset $A \subseteq \mathbb{R}^d$ is *convex* if for any two points $\mathbf{x}, \mathbf{y}$ of A, the entire segment

$$[\mathbf{x}, \mathbf{y}] \stackrel{\text{def}}{=} \{t\mathbf{x} + (1-t)\mathbf{y} \; ; \; 0 \leq t \leq 1\}$$

is contained in A. Given finitely many points $\mathbf{x_1}, \ldots, \mathbf{x_n}$ in A, a *convex combination* is a point $\mathbf{x} = \sum_{i=1}^{n} \lambda_i \mathbf{x_i}$, where $\sum_{i=1}^{n} \lambda_i = 1$ and $\lambda_i \geq 0$ for all i. The *convex hull of* A, denoted by $\operatorname{conv}(A)$, is the set of all convex combinations of points from A. In other words,

$$\operatorname{conv}(A) \stackrel{\text{def}}{=} \{\sum_{i=1}^{n} \lambda_i \mathbf{a_i} \quad \text{such that } n \in \mathbb{N} \setminus \{0\}, \sum_{i=1}^{n} \lambda_i = 1,\ \lambda_i \geq 0 \text{ and } \mathbf{a_i} \in A \text{ for all } i\}.$$

Lemma 2. $\operatorname{conv}(A)$ *is the smallest convex subset containing* A.

Proof. Easy! Start by showing that $\operatorname{conv}(A)$ is convex... □

Definition 3. A *polytope* is the convex hull of finitely many points in $\mathbb{R}^d$.

Definition 4. Let P be a polytope in $\mathbb{R}^d$. A *face of* P is any subset $F \subset \mathbb{R}^d$ of the form

$$F = \{\mathbf{x} \in P \text{ such that } \mathbf{c} \cdot \mathbf{x} = c_0\},$$

where $\mathbf{c} \cdot \mathbf{x} \leq c_0$ is an inequality satisfied by all $\mathbf{x}$ in P.
We also say that the linear inequality $\mathbf{c} \cdot \mathbf{x} \leq c_0$ *supports the face* F *of* P.

In other words, "faces" are where linear functions are maximized within P. Faces may have different dimensions: For example, P is a face of itself, by taking $\mathbf{c} = \mathbf{0}$ and $c_0 = 0$. But also the empty set is a face of any polytope P, by taking $\mathbf{c} = \mathbf{0}$ and $c_0 = 1$.

Definition 5. Faces of dimension 0, 1 and $\dim P - 1$ are called *vertices*, *edges*, and *facets*.

Proposition 6. *Let P be a polytope in $\mathbb{R}^d$. Let $\mathbf{v} \in P$. The following are equivalent:*
(i) $\mathbf{v}$ *is a vertex of* P*;*
(ii) $\mathbf{v}$ *cannot be written as a convex combination of points in* $P \setminus \{v\}$*;*
(iii) *the only vector* $\mathbf{w}$ *for which both* $\mathbf{v} + \mathbf{w}$ *and* $\mathbf{v} - \mathbf{w}$ *are in* P*, is* $\mathbf{w} = \mathbf{0}$.
(iv) *there are d constraints $\mathbf{a}_j^\top \mathbf{x} \leq b_j$ valid for P which are tight at $\mathbf{v}$ (that is, $\mathbf{a}_j^\top \mathbf{v} = b_j$ for all $1 \leq j \leq d$) and the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_d$ are linearly independent.*

Proof. (i) ⇒ (ii): By contradiction, $\mathbf{v} = \sum \lambda_i \mathbf{x}_i$, with $\mathbf{x}_i \in P$ different from $\mathbf{v}$, with $\lambda_i \geq 0$ for all i, and with $\sum_i \lambda_i = 1$. By the assumption, $\{\mathbf{v}\} = H \cap P$, where H is some hyperplane of the form $\mathbf{c}^\top \mathbf{x} = c_0$, and such that $\mathbf{c}^\top \mathbf{x}_i < c_0$ for all i, because $\mathbf{v}$ is the only vertex in $H \cap P$. But then we reach the contradiction

$$c_0 = \mathbf{c}^\top \mathbf{v} = \mathbf{c}^\top (\sum \lambda_i \mathbf{x}_i) = \sum \lambda_i \mathbf{c}^\top \mathbf{x}_i < \sum \lambda_i c_0 = c_0.$$

(ii) ⇒ (iii): Write $\mathbf{v} = \frac{1}{2}(\mathbf{v} + \mathbf{w}) + \frac{1}{2}(\mathbf{v} - \mathbf{w})$.
(iii) ⇒ (iv): By contradiction, let $\mathbf{a}_j^\top \mathbf{x} \leq b_j$ be any constraint valid for P that is *not* tight at v. That means that for some small vector $\mathbf{w}$, both $\mathbf{a}_j^\top (\mathbf{v} + \mathbf{w}) \leq b_j$ and $\mathbf{a}_j^\top (\mathbf{v} - \mathbf{w}) \leq b_j$. Hence, both $\mathbf{v} + \mathbf{w}$ and $\mathbf{v} - \mathbf{w}$ are in P.
(iv) ⇒ (i): Define a hyperplane by $H \overset{\text{def}}{=} \{\mathbf{x} \text{ such that } \sum_{j=1}^d \mathbf{a}_j^\top \mathbf{x} = \sum_{j=1}^d b_j\}$. Then every $\mathbf{x}$ in P satisfies $\mathbf{a}_j^\top \mathbf{x} \leq b_j$ for all j, so in particular it satisfies $\sum_{j=1}^d \mathbf{a}_j^\top \mathbf{x} \leq \sum_{j=1}^d b_j$. Moreover, if $\mathbf{x}$ is any element in $P \cap H$, then all d constraints are tight at $\mathbf{x}$; but since a rank-d system of linear equations in $\mathbb{R}^d$ has only one solution, we conclude that $\mathbf{x} = \mathbf{v}$. □

Corollary 7. *A polytope is the convex hull of its vertices (and of no proper subset thereof).*

Proof. Let A be finite. Let V be the list of vertices of $P \overset{\text{def}}{=} \text{conv}(A)$. Let $V' = V \cup A$. Since $A \subseteq V' \subseteq P$, clearly $P = \text{conv}(A) \subseteq \text{conv}(V') \subseteq \text{conv}(P) = P$, which implies $P = \text{conv}(V')$. But all points of V' not in V can be written as combination of points in V, so $P = \text{conv}(V)$. □

Example 8 (Simplices)**.** We define the *standard d-dimensional simplex* as

$$\Delta^d \overset{\text{def}}{=} \text{conv}\{\mathbf{e}_1, \ldots, \mathbf{e}_{d+1}\} \subseteq \mathbb{R}^{d+1}.$$

This can also be described in terms of facets as follows:

$$\Delta^d = \{\mathbf{x} \in \mathbb{R}^{d+1} \text{ such that } \mathbf{x}_i \geq 0 \text{ for all } i \text{ and } \sum_{i=1}^d \mathbf{x}_i = 1\}.$$

By construction, the d-simplex has $d+1$ vertices and $d+1$ facets. For example: Δ^1 is a segment of length $\sqrt{2}$ in $\mathbb{R}^2$; Δ^2 is an equilateral triangle in $\mathbb{R}^3$; Δ^3 is a regular tetrahedron in $\mathbb{R}^4$.

Example 9 (Cubes)**.** We define the *standard d-dimensional cube* as

$$C^d \overset{\text{def}}{=} \text{conv}\{\pm 1, \pm 1, \ldots, \pm 1\} \subseteq \mathbb{R}^d.$$

This can also be described in terms of facets as follows:

$$C^d = \{\mathbf{x} \in \mathbb{R}^d \text{ such that } -1 \leq \mathbf{x}_i \leq +1 \text{ for all } i\ \}.$$

By construction, the d-cube has 2^d vertices and $2d$ facets. For example, C^1 is the segment $[-1, 1] \subseteq \mathbb{R}$, C^2 is a square, C^3 a cube.

Example 10 (Crosspolytopes)**.** We define the *standard crosspolytope* as

$$C^{d*} \overset{\text{def}}{=} \text{conv}\{\mathbf{e}_1, -\mathbf{e}_1, \ldots, \mathbf{e}_d, -\mathbf{e}_d\} \subseteq \mathbb{R}^d.$$

This can also be described in terms of facets as follows:

$$C^{d*} = \{\mathbf{x} \in \mathbb{R}^d \text{ such that } \sum_{i=1}^d |\mathbf{x}_i| \leq 1\}.$$

By construction, the cube has $2d$ vertices and 2^d facets. For example, C^{1*} is the segment $[-1, 1] \subseteq \mathbb{R}$, same as C^1; C^{2*} is the square C^2 rotated of 45 degrees; C^{3*} is the octahedron.

Definition 11 (Affine/Projective Equivalence)**.** Two polytopes $P \subseteq \mathbb{R}^d$ and $Q \subseteq \mathbb{R}^e$ are *affinely equivalent* (resp. *projectively equivalent*) if there is an affine (resp. projective) map $f : \mathbb{R}^d \to \mathbb{R}^e$ that yields a bijection between P and Q when restricted to P.

Definition 12 (Combinatorial Equivalence)**.** Two polytopes $P \subseteq \mathbb{R}^d$ and $Q \subseteq \mathbb{R}^e$ are *combinatorially equivalent* if there is a bijection between the sets of their faces that preserves the inclusion relation. (Equivalently: if they have the same face poset.)

Proposition 13. *Affinely $\Rightarrow$ projectively $\Rightarrow$ combinatorially equivalent. Converses are false.*

Proof. The implications are obvious. Affine maps preserve parallelism, so a rectangle, a square and a parallelogram are all affinely equivalent, but a trapezoid is not affinely equivalent to them; however, all quadrilaterals are projectively equivalent. In general, any projective equivalence in d-space is determined by the image of $d + 2$ points; so if you perturb the position of just one of the vertices of a regular pentagon, say, you get a polygon with five edges no longer projectively equivalent to the regular one; but they would still be combinatorially equivalent. □

Typically, we have "combinatorial equivalence" in mind. So the convex hull of $d + 1$ affinely independent points in $\mathbb{R}^d$ is called "$\underline{\text{a}}$ simplex".

Definition 14. A *polyhedron* is any subset of $\mathbb{R}^d$ of the form $P = \{\mathbf{x} \text{ such that } A\mathbf{x} \leq \mathbf{b}\}$, for some $\mathbf{b} \in \mathbb{R}^m$ and some matrix $A \in \mathbb{R}^{m \times d}$. A *polyhedral-cone* is any subset of $\mathbb{R}^d$ of the form $P = \{\mathbf{x} \text{ such that } A\mathbf{x} \leq \mathbf{0}\}$. In other words, polyhedra are intersections of finitely many closed half-spaces, and polyhedral-cones are intersection of finitely many closed linear halfspaces.

Definition 15 (Bounded sets)**.** A set $B \subset \mathbb{R}^d$ is *bounded* if it is contained in a large cube; that is, if there exists an integer $n \in \mathbb{N}$ such that $B \subseteq [-n, n]^d$.

Definition 16. A set U in $\mathbb{R}^d$ *contains a ray* if there exist $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^d$ such that the whole set $\{\mathbf{x} + u\mathbf{y} : u \geq 0\}$ (called *infinite ray from* $\mathbf{x}$ *in direction* $\mathbf{y}$) is contained in U.

Obviously any set that contains an infinite ray is not bounded. The converse is false in general, e.g. $\mathbb{Z} \subset \mathbb{R}$, but true for polyhedra, because of Bolzano–Weierstrass' theorem:

Lemma 17. *A polyhedron is bounded if and only if it contains no ray.*

Proof. Let $P = \{\mathbf{x} \text{ such that } A\mathbf{x} \leq \mathbf{b}\}$ be a polyhedron that contains no ray. By contradiction, suppose P is not bounded. Then for every $n \in \mathbb{N}$, the set U contains a point $\mathbf{x}_n$ of norm larger than n. Consider the sequence $(\mathbf{y}_n)_{n \in \mathbb{N}}$ defined by

$$\mathbf{y}_n \stackrel{\text{def}}{=} \frac{\mathbf{x}_n}{||\mathbf{x}_n||}.$$

By the Bolzano–Weierstrass theorem, being bounded, $(\mathbf{y}_n)_{n \in \mathbb{N}}$ has a convergent subsequence. Let $\mathbf{y}$ be the limit point of this subsequence. Since all points of $(\mathbf{y}_n)_{n \in \mathbb{N}}$ have norm 1, and satisfy

$$A\mathbf{y}_n = \frac{1}{||\mathbf{x}_n||} A\mathbf{x}_n \leq \frac{1}{||\mathbf{x}_n||}\mathbf{b} < \frac{1}{n}\mathbf{b},$$

passing to the limit of the subsequence we have that $||\mathbf{y}|| = 1$ and $A\mathbf{y} \leq \mathbf{0}$. But then for every $\mathbf{x}$ in P, the ray $\{\mathbf{x} + u\mathbf{y} : u \geq 0\}$ is contained in P, because

$$A(\mathbf{x} + u\mathbf{y}) = A\mathbf{x} + uA\mathbf{y} \leq A\mathbf{x} \leq \mathbf{b}.$$

□

1 Fourier-Motzkin elimination and Farkas' lemma

Theorem 18 (Fourier 1827, Motzkin 1936)**.** *Let $A\mathbf{x} \le \mathbf{b}$ be a system with $n \ge 1$ variables and m linear inequalities. There is a system $A'\mathbf{x}' \le \mathbf{b}'$ with $n-1$ variables $x_2, \dots, x_n$ and at most $\max(m, \frac{m^2}{4})$ linear inequalities, such that:*

(FM1) *$A\mathbf{x} \le \mathbf{b}$ has a solution if and only if $A'\mathbf{x}' \le \mathbf{b}'$ does, and*

(FM2) *each inequality of $A'\mathbf{x}' \le \mathbf{b}'$ is a positive linear combination of one or two of the inequalities from $A\mathbf{x} \le \mathbf{b}$.*
That is, $(0|A') = MA$ and $\mathbf{b}' = M\mathbf{b}$ for some matrix M with entries ≥ 0.

Proof by example. Consider the 3-variable system

$$\begin{aligned} 2x - 5y + 4z &\le 10 \\ 3x - 6y + 3z &\le 9 \\ 5x + 10y - z &\le 15 \\ -x + 5y - 2z &\le -7 \\ -3x + 2y + 6z &\le 12. \end{aligned}$$

Let us get rid of the variable x, say. To do this, we rewrite the system as

$$\begin{aligned} x &\le 5 + \tfrac{5}{2}y - 2z \\ x &\le 2 + 2y - z \\ x &\le 3 - 2y + \tfrac{1}{5}z \\ x &\ge 7 + 5y - 2z \\ x &\ge -4 + \tfrac{2}{3}y + 2z. \end{aligned}$$

Clearly, one such x exists if and only if the following inequality has a solution:

$$\max\left(7 + 5y - 2z, -4 + \frac{2}{3}y + 2z\right) \le \min\left(5 + \frac{5}{2}y - 2z, 2 + 2y - z, 3 - 2y + \frac{1}{5}z\right).$$

But such inequality can be equivalently rewritten as a system of inequalities in y and z only, by imposing that each of the lower bounds is indeed not larger than each of the upper bounds!

$$\begin{aligned} 7 + 5y - 2z &\le 5 + \tfrac{5}{2}y - 2z \\ 7 + 5y - 2z &\le 2 + 2y - z \\ 7 + 5y - 2z &\le 3 - 2y + \tfrac{1}{5}z \\ -4 + \tfrac{2}{3}y + 2z &\le 5 + \tfrac{5}{2}y - 2z \\ -4 + \tfrac{2}{3}y + 2z &\le 2 + 2y - z \\ -4 + \tfrac{2}{3}y + 2z &\le 3 - 2y + \tfrac{1}{5}z. \end{aligned}$$

Taking all the variables to the left and all the constants to the right hand side, we get the desired system $A'\mathbf{x}' \le \mathbf{b}'$, in which every row is a positive linear combination of exactly two inequalities from $A\mathbf{x} \le \mathbf{b}$. (If in $A\mathbf{x} \le \mathbf{b}$ we had inequalities not involving x, we just copy-paste them.) $\square$

Lemma 19 (Logic Farkas Lemma, essentially Farkas, 1894)**.** *Let $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$.*
The system $A\mathbf{x} \le \mathbf{b}$ is inconsistent $\iff$ for some $\mathbf{y} \ge \mathbf{0}$ in $\mathbb{R}^m$, $A^\top \mathbf{y} = \mathbf{0}$ and $\mathbf{b}^\top \mathbf{y} = -1$.

Proof by Kuhn, 1956. By induction on the number n of variables. The basis case is $n = 0$, when we have $A = \mathbf{0}$; a system of the type $\mathbf{0} \le \mathbf{b}$ is inconsistent when $\mathbf{b}$ has some negative components. Choose an i with $b_i > 0$ and set $\mathbf{y} \overset{\text{def}}{=} \frac{-1}{b_i}\mathbf{e}_i$. By construction, $\mathbf{b}^\top \mathbf{y} = \frac{-b_i}{b_i} = -1$. If instead $A\mathbf{x} \le \mathbf{b}$ has at least one variable, we perform Fourier–Motzkin elimination: The new system $A'\mathbf{x}' \le \mathbf{b}'$, where $A' \in \mathbb{R}^{m' \times (n-1)}$ for some $m' \le \max(m, m^2/2)$, $(0|A') = MA$ and

$\mathbf{b}' = M\mathbf{b}$ for some nonnegative matrix M, is still inconsistent and has one variable less. By induction, there exists a vector $\mathbf{y}' \geq \mathbf{0}$ in $\mathbb{R}^{m'}$ such that $A'^{\top}\mathbf{y} = \mathbf{0}$ and $\mathbf{b}'^{\top}\mathbf{y}' = -1$. But then if we set $\mathbf{y} \stackrel{\text{def}}{=} M^{\top}\mathbf{y}'$, by definition we get

$$A^{\top}\mathbf{y} = A^{\top}(M^{\top}\mathbf{y}') = (MA)^{\top}\mathbf{y}' = (0|A')^{\top}\mathbf{y}' = \mathbf{0}^{\top} \text{ and } \mathbf{b}^{\top}\mathbf{y} = \mathbf{b}^{\top}M\mathbf{y}' = (M\mathbf{b})^{\top}\mathbf{y}' = \mathbf{b}'^{\top}\mathbf{y}' = -1.$$

□

Definition 20. Given finitely many vectors $v_1, \ldots, v_n$ in $\mathbb{R}^d$, if V is the $d \times n$ matrix that has the v_i's as columns, we define

$$\text{cone}(V) \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbb{R}^d \text{ such that } \mathbf{x} = V\mathbf{u} \text{ for some } \mathbf{u} \in \mathbb{R}^n \text{ such that } \mathbf{u} \geq \mathbf{0}\}, \text{ and}$$

$$\text{conv}(V) \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbb{R}^d \text{ such that } \mathbf{x} = V\mathbf{t} \text{ for some } \mathbf{t} \in \mathbb{R}^n \text{ such that } \mathbf{t} \geq \mathbf{0}, \mathbf{t} \cdot \mathbf{1} = 1\}.$$

If X is of the form cone(V), we say that X is a *finitely-generated cone*, and we also sat that X is *generated as cone by the columns of* V. A priori, this notion has nothing to do with the "polyhedral-cone" notion, though today we'll see why we used the same word. (We have already a word for when X is of the form conv(V), right?)

Lemma 21 (Geometric Farkas Lemma)**.** *Let A be a real $m \times n$ matrix. Let $\mathbf{b} \in \mathbb{R}^m$. Exactly one of the following is true:*

(F1) *The point $\mathbf{b}$ lies in the finitely-generated cone C in $\mathbb{R}^m$ generated by the n columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$ of A.*

(F2) *There is a hyperplane passing through the origin, of the form $\{\mathbf{x} \in \mathbb{R}^m$ such that $\mathbf{y}^{\top}\mathbf{x} = 0\}$ for some $\mathbf{y} \in \mathbb{R}^m$, such that all vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ (and thus the cone C) lie on one side, whereas $\mathbf{b}$ lies on the other side. That is, $\mathbf{a}_i \cdot \mathbf{y} \geq 0$ for all i, whereas $\mathbf{b} \cdot \mathbf{y} < 0$.*

Proof. (F1) is the same as saying, there exists $\mathbf{x} \geq \mathbf{0}$ in $\mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$. Instead (F2) is the same as saying, there is a $\mathbf{y} \in \mathbb{R}^m$ such that $A^{\top}\mathbf{y} \geq \mathbf{0}$ in $\mathbb{R}^n$ and $\mathbf{b}^{\top}\mathbf{y} < 0$ in $\mathbb{R}$. If (F1) is true, then clearly (F2) is false, for otherwise we would have $(A\mathbf{x})^{\top}\mathbf{y} = \mathbf{b}^{\top}\mathbf{y} < 0$ and at the same time $(A\mathbf{x})^{\top}\mathbf{y} = \mathbf{x}^{\top}A^{\top}\mathbf{y} \geq \mathbf{x}^{\top}\mathbf{0} = 0$. If instead (F1) is false, then with a cheap trick $A\mathbf{x} = \mathbf{b}$ can be translated as $A'\mathbf{x} \leq \mathbf{b}'$, where

$$A' \stackrel{\text{def}}{=} \begin{pmatrix} A \\ -A \end{pmatrix} \text{ and } b' \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{b} \\ -\mathbf{b} \end{pmatrix}.$$

The fact that $A'\mathbf{x} \leq \mathbf{b}'$ has no solution $\mathbf{x} \geq 0$ means that if we define

$$A'' \stackrel{\text{def}}{=} \begin{pmatrix} A' \\ -I_n \end{pmatrix} \text{ and } b'' \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{b}' \\ \mathbf{0} \end{pmatrix},$$

then the system $A''\mathbf{x} \leq \mathbf{b}''$ has **no solution**, with A'' in $\mathbb{R}^{(2m+n)\times n}$. By Logic Farkas' Lemma,

$$\exists \mathbf{y}'' \geq \mathbf{0} \text{ in } \mathbb{R}^{2m+n} \text{ such that } A''^{\top}\mathbf{y}'' = \mathbf{0} \text{ and } \mathbf{b}''^{\top}\mathbf{y}'' = -1. \tag{1}$$

So if we write

$$\mathbf{y}'' = \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \\ \mathbf{u} \end{pmatrix} \text{ with } \mathbf{v}, \mathbf{w} \in \mathbb{R}^m, \mathbf{u} \in \mathbb{R}^n,$$

then we can read Equation 1 as follows:

$$\exists \mathbf{v}, \mathbf{w} \geq \mathbf{0} \text{ in } \mathbb{R}^m \text{ and } \mathbf{u} \geq \mathbf{0} \text{ in } \mathbb{R}^n \text{ such that } A^{\top}\mathbf{v} - A^{\top}\mathbf{w} - \mathbf{u} = \mathbf{0} \text{ and } \mathbf{b}^{\top}\mathbf{v} - \mathbf{b}^{\top}\mathbf{w} = -1.$$

In other words,

$$\exists \mathbf{v}, \mathbf{w} \geq \mathbf{0} \text{ in } \mathbb{R}^m \text{ such that } A^{\top}\mathbf{v} - A^{\top}\mathbf{w} \geq 0 \text{ and } \mathbf{b}^{\top}\mathbf{v} - \mathbf{b}^{\top}\mathbf{w} = -1.$$

Setting $\mathbf{y} \stackrel{\text{def}}{=} \mathbf{v} - \mathbf{w}$, we conclude that $A^{\top}\mathbf{y} \geq 0$ and $\mathbf{b}^{\top}\mathbf{y} = -1$. □

2 The Fundamental Theorem for Polyhedra

Definition 22. The Minkowski sum of two subsets $A, B \subset \mathbb{R}^d$ is

$$A + B \overset{\text{def}}{=} \{a + b \text{ such that } a \in A, b \in B\}.$$

Theorem 23 (Fundamental Theorem, Minkowski 1896, Motzkin 1936)**.** *Let $P \subseteq \mathbb{R}^d$.*

(I.) *P is a finitely-generated cone $\iff$ P is a polyhedral-cone.*

(II.) *P is a polytope $\iff$ P is a bounded polyhedron.*

(III.) *P is a Minkowski sum of a finitely-generated cone and a polytope $\iff$ P is a polyhedron.*

Proof. We follow the presentation in Schrijver, Theory of Linear and Integer Programming.

I, '$\Rightarrow$': Let $C = \text{cone}(V)$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_e$ be the columns of V. Without loss, we may assume that $\mathbf{v}_1, \ldots, \mathbf{v}_e$ span $\mathbb{R}^d$. (If not, let X be their span: then any linear halfspace H of X can be extended to a linear halfspace H' of $\mathbb{R}^d$, so that $H' \cap X = H$.) Consider all the (finitely many!) linear halfspaces $H = \{\mathbf{x} \text{ such that } \mathbf{cx} \leq 0\}$ of $\mathbb{R}^d$ such that $\mathbf{v}_1, \ldots, \mathbf{v}_e$ are in H and $\{\mathbf{x} \text{ such that } \mathbf{cx} = 0\}$ is spanned by ($d-1$ linearly independent) vectors from the list $\mathbf{v}_1, \ldots, \mathbf{v}_e$. But cone($V$) is exactly the intersection of these halfspaces. In fact, $\subseteq$ is clear by construction; $\supseteq$ is proven by contradiction, using the Geometric Farkas' Lemma: if $\mathbf{x}$ is not a nonnegative linear combination of the v_i's, then some hyperplane through the origin separates it from C, so that C and $\mathbf{x}$ would end up in opposite linear halfspaces.

I, '$\Leftarrow$': Let $C \overset{\text{def}}{=} \{\mathbf{x} \text{ such that } \mathbf{a}_1^\top \mathbf{x} \leq 0, \ldots, \mathbf{a}_m^\top \mathbf{x} \leq 0\}$. By part (I, '$\Rightarrow$') proved above, we know that the cone generated by $\mathbf{a}_1^\top, \ldots, \mathbf{a}_m^\top$ is a polyhedral-cone; in other words, there exists a matrix B such that $\text{cone}(\mathbf{a}_1^\top, \ldots, \mathbf{a}_m^\top) = \{\mathbf{x} \text{ such that } B\mathbf{x} \leq 0\}$. If we call $\mathbf{b}_1, \ldots, \mathbf{b}_t$ the rows of this matrix B, we have that

$$\text{cone}(\mathbf{a}_1^\top, \ldots, \mathbf{a}_m^\top) = \{\mathbf{x} \text{ such that } \mathbf{b}_1 \cdot \mathbf{x} \leq 0, \ldots, \mathbf{b}_t \cdot \mathbf{x} \leq 0\}. \tag{2}$$

We claim that

$$C = \text{cone}(\mathbf{b}_1^\top, \ldots, \mathbf{b}_t^\top).$$

In fact, '$\supseteq$' is because by Equation 2 we know that $\mathbf{b}_j \cdot \mathbf{a}_i \leq 0$ for all i in $\{1, \ldots, m\}$ and for all j in $\{1, \ldots, t\}$. By definition of C, this tells us that $\mathbf{b}_1^\top, \ldots, \mathbf{b}_t^\top$ are in C, so the cone they generate is also in C. It remains to show '$\subseteq$'. By contradiction, suppose that $\mathbf{y} \notin \text{cone}(\mathbf{b}_1^\top, \ldots, \mathbf{b}_t^\top)$ for some $\mathbf{y} \in C$. By the Geometric Farkas' Lemma, we know that there exists a vector $-\mathbf{w}$ such that

$$\mathbf{b}_1 \cdot (-\mathbf{w}) \geq 0, \ \ldots \ , \ \mathbf{b}_t \cdot (-\mathbf{w}) \geq 0, \text{ and } \mathbf{y} \cdot (-\mathbf{w}) < 0.$$

In other words, there exists a vector $\mathbf{w}$ such that

$$\mathbf{b}_1 \cdot \mathbf{w} \leq 0, \ \ldots \ , \ \mathbf{b}_t \cdot \mathbf{w} \leq 0, \text{ and } \mathbf{y} \cdot \mathbf{w} > 0. \tag{3}$$

Via Equation 2, the first t inequalities above tell us that $\mathbf{w} \in \text{cone}(\mathbf{a}_1^\top, \ldots, \mathbf{a}_m^\top)$. So write

$$\mathbf{w} = \lambda_1 \mathbf{a}_1^\top + \ldots + \lambda_m \mathbf{a}_m^\top, \qquad \text{for some } \lambda_i \geq 0. \tag{4}$$

Since $\mathbf{y}$ in C, by definition of C we must have that $\mathbf{a}_i^\top \mathbf{y} \leq 0$ for all i. In particular, since $\lambda_i \geq 0$, we have $\lambda_i \mathbf{a}_i^\top \mathbf{y} \leq 0$ for all i. But then from Equations 3 and 4 we get the contradiction

$$0 < \mathbf{y} \cdot \mathbf{w} = \mathbf{y} \cdot (\sum_{i=1}^{m} \lambda_i \mathbf{a}_i^\top) = \sum_{i=1}^{m} \lambda_i \mathbf{a}_i^\top \mathbf{y} \leq 0.$$

III, '⇐': Let P be of the form $\{\mathbf{x} \text{ such that } A\mathbf{x} \leq \mathbf{b}\}$, for some matrix $A \in \mathbb{R}^{m\times d}$ and some vector $\mathbf{b} \in \mathbb{R}^m$. This means that P is defined by the m inequalities $\mathbf{a}_i\mathbf{x} \leq b_i$, where $\mathbf{a}_i$ is the i-th row of A. The idea is now to homogenize this system by introducing a new variable x_0, and by setting $\mathbf{a}_i\mathbf{x} - b_i x_0 \leq 0$, together with $x_0 \geq 0$. If we set $x_0 = 1$ in the new system, we recover the old one: so the solutions of the new system are simply the vectors solving of the old one, with a 1 appended. Formally, define

$$C(P) \stackrel{\text{def}}{=} \{\mathbf{x} \text{ such that } A'\mathbf{x} \leq \mathbf{0}\}, \qquad \text{where } A' = \begin{pmatrix} -1 & \mathbf{0} \\ -\mathbf{b} & A \end{pmatrix}.$$

Then

$$P = \{\mathbf{x} \in \mathbb{R}^d \text{ such that } \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \in C(P)\}. \tag{5}$$

By definition, $C(P)$ is an intersection of closed linear halfspaces. But then by part (I), '⇐', C(P) is a finitely-generated cone. Suppose that $C(P) = \text{cone}(V) \subset \mathbb{R} \times \mathbb{R}^d$. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_e$ be the column vectors of V. Without loss, we can assume that the first coordinate of each vector $\mathbf{v}_i$ is either 0 or 1. So reorder the column vectors of V as

$$\begin{pmatrix} 0 \\ \mathbf{w}_1 \end{pmatrix}, \begin{pmatrix} 0 \\ \mathbf{w}_2 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \mathbf{w}_k \end{pmatrix}, \begin{pmatrix} 1 \\ \mathbf{w}_{k+1} \end{pmatrix}, \begin{pmatrix} 1 \\ \mathbf{w}_{k+2} \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ \mathbf{w}_e \end{pmatrix}.$$

Let Q be the cone generated by $\mathbf{w}_1, \ldots, \mathbf{w}_k$ in $\mathbb{R}^d$. Let R be the convex hull of $\mathbf{w}_{k+1}, \ldots, \mathbf{w}_e$ in $\mathbb{R}^d$. Then from Equation 5 it is easy to see that

$$P = \{\mathbf{x} \in \mathbb{R}^d \text{ such that } \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \in C(P)\} = Q + R.$$

III, '⇒': Let $P \subseteq \mathbb{R}^d$ be of the form $P = \text{cone}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k) + \text{conv}(\mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \ldots, \mathbf{v}_e)$. Define

$$C(P) \stackrel{\text{def}}{=} \text{cone}\left(\begin{pmatrix} 0 \\ \mathbf{v}_1 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \mathbf{v}_k \end{pmatrix}, \begin{pmatrix} 1 \\ \mathbf{v}_{k+1} \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ \mathbf{v}_e \end{pmatrix}\right).$$

Clearly, any vector $\mathbf{x} \in \mathbb{R}^d$ belongs to P if and only if $\begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \in C(P)$. By part I, '⇒', the right hand side is a polyhedral-cone in $\mathbb{R} \times \mathbb{R}^d$. Hence it can be written (separating the first column of the matrix defining it) as

$$\{\begin{pmatrix} \lambda \\ \mathbf{x} \end{pmatrix} \text{ such that } (\mathbf{b}|A)\begin{pmatrix} \lambda \\ \mathbf{x} \end{pmatrix} \leq \mathbf{0}\},$$

for some matrix $A \in \mathbb{R}^{m\times d}$, some scalar $\lambda \in \mathbb{R}$ and some vector $\mathbf{b} \in \mathbb{R}^m$. So a vector $\mathbf{x} \in \mathbb{R}^d$ belongs to P if and only if $A\mathbf{x} \leq -\lambda\mathbf{b}$, which means that P is a polyhedral-cone.

II, '⇒': That polytopes are polyhedra follows by part (III), '⇒'. Thus it suffices to show that the convex hull of any bounded set V is bounded. This is easy: if $|\mathbf{x}| \leq M$ for all $\mathbf{x}$ in V, then $|\sum_{i=1}^n \lambda_i \mathbf{x}_i| \leq \sum_{i=1}^n \lambda_i |\mathbf{x}_i| \leq \sum_{i=1}^n \lambda_i M = M$. So $\text{conv}(V)$ is bounded.

II, '⇐': This follows immediately from part (III), '⇐', because cones are not bounded. □

Remark 24. We already know from Proposition 6 how to pass from the facet to the vertex description of a polytope: to find the vertices of $A\mathbf{x} \leq \mathbf{b}$ in $\mathbb{R}^d$, with $A \in \mathbb{R}^{m\times d}$, we simply need to choose d out of the m inequalities, put equality signs in them, and solve! Two remarks: If the solution is not unique, then the rows were linearly dependent, so by Proposition 6 we may throw the infinitely-many solutions away. Also, the solution may be incompatible with the remaining inequalities; in this case, we also throw it away. But how to pass from a vertex to a facet description? For this we need another idea, namely, duality.

3 The Idea of Duality

Given a cone C in $\mathbb{R}^d$, its polar cone is the set $C^* \stackrel{\text{def}}{=} \{\mathbf{c} \text{ such that } \mathbf{cx} \leq 0 \text{ for all } \mathbf{x} \in C\}$. This is a well-studied set with the remarkable property (proven below) that $C^{**} = C$. For this reason, the polar cone is often called "dual cone". Following Schrijver's textbook, we develop an analogous (and more general) notion of polarity for polyhedra.

Definition 25. For any subset $X \subseteq \mathbb{R}^d$, the *polar* of X is

$$X^* \stackrel{\text{def}}{=} \{\mathbf{c} \in \mathbb{R}^d \text{ such that } \mathbf{cx} \leq 1 \text{ for all } \mathbf{x} \in X\}.$$

Lemma 26. *For any subset $X \subseteq \mathbb{R}^d$,*
(a) *X^* is a convex set containing $\mathbf{0}$,*
(b) *for any $Y \supseteq X$, one has $Y^* \subseteq X^*$;*
(c) *$X^{**} \supseteq X$.*

Proof. Straightforward from the definitions. □

Remark 27. In general, X^{**} is larger than X. In fact, one can show that X^{**} is the topological closure of $\text{conv}\,(X \cup \{\mathbf{0}\})$. Below we will prove a weaker statement: namely, that if X is a polyhedron containing the origin, then $X^{**} = X$. Since we can always shift coordinates to place the origin inside a given polyhedron, it makes sense for us to use the word "dual" instead of "polar". (Note: to make duality an internal operation within the world of polytopes, some authors like Ziegler require $\mathbf{0}$ to be in *the topological interior of* X; we will see that this stronger assumption is equivalent to the boundedness of X^*.)

Notation. From now on, if P can be written in the form $\{\mathbf{x} \in \mathbb{R}^d \text{ such that } A\mathbf{x} \leq \mathbf{1}\}$, we simply write $P = P(A, \mathbf{1})$. (Not all polytope are of this form: as Exercise, show that the polytopes that can be written this way are exactly those containing $\mathbf{0}$ in their interior.) Moreover, for any $m \times d$ matrix A and any $n \times d$ matrix B, we define

$$P(A, \mathbf{1}, B, \mathbf{0}) \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbb{R}^d \text{ such that } A\mathbf{x} \leq \mathbf{1}, B\mathbf{x} \leq \mathbf{0}\}.$$

By definition, $P(A, \mathbf{1}, B, \mathbf{0})$ is a polyhedron.

Theorem 28 (Duality for polyhedra)**.** *Let $P \subseteq \mathbb{R}^d$ be a polyhedron containing $\{0\}$. Then:*
(i) *P^* is a polyhedron again;*
(ii) *$P^{**} = P$;*
(iii) *if $P = \text{conv}(\{0\} \cup V) + \text{cone}(W)$, then $P^* = P(V^\top, \mathbf{1}, W^\top, \mathbf{0})$;*
(iv) *if $P = P(A, \mathbf{1}, B, \mathbf{0})$, then $P^* = \text{conv}(\{0\} \cup A^\top) + \text{cone}(B^\top)$.*

Proof. (i) By the Fundamental Theorem 23, we can write P as the Minkowski sum of a polytope and a finitely-generated cone, whence by part (iii) we conclude.

(ii) We only have to show $P^{**} \subseteq P$, the other inclusion being true for all sets P. By contradiction, suppose $\mathbf{z}$ is in P^{**} but not in P. By Farkas' lemma, there is some inequality $\mathbf{a} \cdot \mathbf{x} \leq \beta$ satisfied by all $\mathbf{x}$ in P, but not by $\mathbf{z}$. Since $\mathbf{0}$ is in P, $\beta \geq 0$. So there are two cases: If $\beta > 0$, then $\beta^{-1}\mathbf{a}$ is in P^*, so since $\mathbf{z} \in P^{**}$ we obtain$\beta^{-1}\mathbf{a} \cdot \mathbf{z} \leq 1$, which contradicts the fact that $\mathbf{a} \cdot \mathbf{z} > \beta$ by definition of $\mathbf{z}$. If instead $\beta = 0$, then $\lambda\mathbf{a}$ is in P^* for all $\lambda \geq 0$. But since $\mathbf{a} \cdot \mathbf{z} > \beta = 0$, we get that $\lambda\mathbf{a} \cdot \mathbf{z} > 1$ for some λ large enough. Another contradiction.

(iii) Write $P = \text{conv}(0, \mathbf{v}_1, \dots \mathbf{v}_m) + \text{cone}(\mathbf{w}_1, \dots \mathbf{w}_n)$. Then we claim that

$$P^* = \{\mathbf{c} \in \mathbb{R}^d \text{ such that } \mathbf{c} \cdot \mathbf{v}_i \leq 1 \text{ for all } i = 1, \dots, m, \ \mathbf{c} \cdot \mathbf{w}_j \leq 0 \text{ for all } j = 1, \dots, n.\} \quad (6)$$

The inclusion '$\supseteq$' is clear; for the converse, write any $\mathbf{x}$ in P as a convex combination of the $\mathbf{v}_i$ plus a conic combination of the $\mathbf{w}_j$, and compute $\mathbf{cx}$. Since $\mathbf{c}\cdot\mathbf{v}_i \leq 1$ ($i = 1, \ldots, m$) and $\mathbf{c}\cdot\mathbf{w}_j \leq 0$ ($j = 1, \ldots, n$), plugging in you will get $\mathbf{c}\cdot\mathbf{x} \leq 1$. So the claim is proven. Since V and W are the matrices obtaining juxtaposing the column vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ and $\mathbf{w}_1, \ldots, \mathbf{w}_n$, respectively, we can conveniently rewrite Equation 6 as

$$P^* = \{\mathbf{c} \in \mathbb{R}^d \text{ such that } V^\top \mathbf{c} \leq 1,\ W^\top \mathbf{c} \leq 0\}.$$

(iv) We have to show

$$P(A, \mathbf{1}, B, \mathbf{0})^* = \operatorname{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{0}) + \operatorname{cone}(\mathbf{w}_1, \ldots, \mathbf{w}_n),$$

where $\mathbf{v}_i$ is the transpose of the i-th *row* of A and $\mathbf{w}_i$ is the transpose of the i-th *row* of B. Rather than performing long calculations, we define $Q \overset{\text{def}}{=} \operatorname{conv}(\{0\} \cup A^\top) + \operatorname{cone}(B^\top)$. Then by part (iii) $Q^* = P$. So by part (ii) $P^* = Q^{**} = Q$. □

Definition 29. The *relative interior* of a polytope $P \subseteq \mathbb{R}^d$ (not necessarily full-dimensional) is the set of $\mathbf{y}$ of P not contained in any proper face of P. More generally, for a polyhedron P,

$$\operatorname{relint} P \overset{\text{def}}{=} \{\mathbf{y} \in \mathbb{R}^d \text{ such that } (\mathbf{ay} = a_0 \text{ and } \forall \mathbf{x} \in P,\ \mathbf{ax} \leq a_0) \Longrightarrow \forall \mathbf{x} \in P,\ \mathbf{ax} = a_0\}.$$

Lemma 30. *For any polyhedron P containing $\{0\}$, P^* is bounded if and only if $\mathbf{0} \in \operatorname{relint} P$.*

Proof. Left as exercise. □

Corollary 31 (Duality for Polytopes, Ziegler's version)**.** *Let $P \subseteq \mathbb{R}^d$ be a polytope such that $\mathbf{0} \in \operatorname{relint} P$. Then:*

(i) *P^* is a polytope;*
(ii) *$P^{**} = P$;*
(iii) *if $P = \operatorname{conv}(V)$, then $P^* = P(V^\top, \mathbf{1})$, and*
(iv) *if $P = P(A, \mathbf{1})$, then $P^* = \operatorname{conv}(A^\top)$.*

So if I give you a polytope as a convex hull, i.e. $P = \operatorname{conv}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, then how do I find the hyperplanes defining it? It's easy: Pass to the polar! Since $P^* = P(V^\top, \mathbf{1})$, you have the polar described in terms of its facets; by finding the vertices of P^*, the computer will figure out the facets of P.

Example 32. Let us find the facets of the convex hull P of the five points

$$\begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 7 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

We leave it to you to verify that 0 is in the relative interior of P. Then if V is the 3×5 matrix that has the five points above as columns, $P^* = P(V^\top, \mathbf{1})$. In other words, P^* is the set of vectors $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ satisfying

$$\begin{aligned} -x &\leq 1 \\ x + 2y &\leq 1 \\ 7y + 2z &\leq 1 \\ x - z &\leq 1 \\ 2x + 3y - z &\leq 1. \end{aligned}$$

We now choose 3 out of these 5 inequalities (in any possible way) and set them to equalities. This leads to $\binom{5}{3} = 10$ different systems. Each of them has exactly one solution. However, the first system (namely, $-x = 1$ and $x + 2y = 1$ and $7y + 2z = 1$) has a solution incompatible with the remaining inequalities (namely, $x = -1$ and $y = 1$ and $z = -3$, which contradicts $x - z \leq 1$). The four systems that do *not* make use of the equation $-x = 1$ have the same solution (namely, $x = \frac{1}{3}$ and $y = \frac{1}{3}$ and $z = \frac{-2}{3}$). So out of ten systems, we actually only get six different acceptable solutions. Hence, P^* is the convex hull of the following six points:

$$\begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 5/7 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ -5 \\ 18 \end{pmatrix}, \begin{pmatrix} -1 \\ 5/3 \\ -2 \end{pmatrix}, \begin{pmatrix} 1/3 \\ 1/3 \\ -2/3 \end{pmatrix}.$$

If W is the 3×6 matrix whose columns are the six points above, we have that $P^{**} = P(W^\top, \mathbf{1})$. In other words, P^{**} is the set of vectors $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ satisfying

$$\begin{aligned} -x + y - 2z &\leq 1 \\ -x + y &\leq 1 \\ -x + \tfrac{5}{7}y - 2z &\leq 1 \\ -x - 5y - 18 &\leq 1 \\ -x + \tfrac{5}{3}y - 2z &\leq 1 \\ \tfrac{1}{3}x + \tfrac{1}{3}y - \tfrac{2}{3}z &\leq 1. \end{aligned}$$

But $P^{**} = P$. So P has the six facets above. The last facet, corresponding to (the intersection of P with) the hyperplane $\frac{1}{3}x + \frac{1}{3}y - \frac{2}{3}z = 1$, contains four out of the five points of P. So in case you are wondering, P is a pyramid over a pentagon. P^* is also a pyramid over a pentagon: Five of its six vertices lie on the plane $x = -1$.