A Course in Convexity

Complete Reading Notes

Alexander Barvinok

Organized by the Book's Eight Chapters

2026

Contents

1 Chapter I: Convex Sets at Large **1**

1.1 I.1 Convex Sets: Definitions, Examples, and Operations . . . 1
1.2 I.2 Convex Hulls and Carathéodory's Theorem . . . 3
1.3 I.3 Application: Positive Homogeneous Polynomials . . . 5
1.4 I.4 Radon's and Helly's Theorems . . . 7
1.5 I.5 Applications of Helly's Theorem in Combinatorial Geometry . . . 8
1.6 I.6 Application: Uniform Approximation . . . 10
1.7 I.7 Euler Characteristic and Valuations on Convex Sets . . . 11
1.8 I.8 Linear Transformations, Pushforward Valuations, and Intrinsic Volumes . . . 12
1.9 I.9 Polyhedra and Linear Transformations: Fourier–Motzkin Elimination . . . 14

2 Chapter II: Faces and Extreme Points **15**

2.1 II.1 Separation . . . 16
2.2 II.2 Convex Sets in Euclidean Space . . . 17
2.3 II.3 Extreme Points and the Finite-Dimensional Krein–Milman Theorem . . . 18
2.4 II.4 Extreme Points of Polyhedra . . . 19
2.5 II.5 The Birkhoff Polytope . . . 20
2.6 II.6 Permutation Polytopes and the Schur–Horn Theorem . . . 21
2.7 II.7 Transportation Polyhedra . . . 22
2.8 II.8 Convex Cones . . . 23
2.9 II.9 The Moment Curve and the Moment Cone . . . 24
2.10 II.10 Application: Quadrature of High Algebraic Degree . . . 25
2.11 II.11 The Cone of Nonnegative Polynomials . . . 26
2.12 II.12 The Positive Semidefinite Cone . . . 27
2.13 II.13 Linear Equations in Positive Semidefinite Matrices . . . 28
2.14 II.14 Application: Convexity Theorems for Quadratic Maps . . . 30
2.15 II.15 Application: Graph Realization . . . 31
2.16 II.16 Closed Convex Sets . . . 32

3 Chapter III: Convex Sets in Topological Vector Spaces **34**

3.1 III.1 Separation in Euclidean and General Vector Spaces . . . 34
3.2 III.2 Topological Vector Spaces, Convex Sets, and Hyperplanes . . . 35

3.3 III.3 Separation in Topological Vector Spaces . . . 38
3.4 III.4 The Krein–Milman Theorem . . . 39
3.5 III.5 Polyhedra in L^∞ . . . 40
3.6 III.6 Linear Optimal Control . . . 41
3.7 III.7 Lyapunov's Convexity Theorem . . . 43
3.8 III.8 The Simplex of Probability Measures . . . 44
3.9 III.9 Extreme Points of Intersections and Their Applications . . . 45

4 Chapter IV: Polarity, Duality, and Linear Programming 46
4.1 IV.1 Polarity in Euclidean Space . . . 46
4.2 IV.2 Membership in the Moment Cone . . . 48
4.3 IV.3 Duality of Vector Spaces . . . 50
4.4 IV.4 Duality of Topological Vector Spaces . . . 51
4.5 IV.5 Orders Defined by Convex Cones . . . 52
4.6 IV.6 Linear Programming Problems . . . 53
4.7 IV.7 Zero Duality Gap . . . 54
4.8 IV.8 Polyhedral Linear Programming . . . 56
4.9 IV.9 Transportation Problems . . . 57
4.10 IV.10 Semidefinite Programming . . . 58
4.11 IV.11 Clique Number and Chromatic Number . . . 61
4.12 IV.12 Linear Programming in L^∞ . . . 62
4.13 IV.13 Linear Programming Form of Uniform Approximation . . . 63
4.14 IV.14 Optimal Transport . . . 65

5 Chapter V: Convex Bodies and Ellipsoids 66
5.1 V.1 Ellipsoids . . . 66
5.2 V.2 Ellipsoids of Maximum Volume in Convex Bodies . . . 68
5.3 V.3 Norms and Their Approximation . . . 70
5.4 V.4 The Ellipsoid Method . . . 72
5.5 V.5 Gaussian Measure on Euclidean Space . . . 74
5.6 V.6 Low-Rank Matrix Approximation . . . 76
5.7 V.7 Measure and Metric on the Unit Sphere . . . 77

6 Chapter VI: Faces of Polytopes 80
6.1 VI.1 Polytopes and Polarity . . . 80

6.2 VI.2 Faces of Permutation Polytopes . . . 82
6.3 VI.3 The Euler–Poincaré Formula . . . 84
6.4 VI.4 Polytopes with Many Faces: Cyclic Polytopes . . . 85
6.5 VI.5 Simple Polytopes . . . 86
6.6 VI.6 h-Vectors and the Dehn–Sommerville Equations . . . 87
6.7 VI.7 The Upper Bound Theorem . . . 89
6.8 VI.8 Centrally Symmetric Polytopes . . . 91

7 Chapter VII: Lattices and Convex Bodies 92
7.1 Lattices: Definitions, Examples, and Bases . . . 92
7.2 Determinant, Density, Sublattices, and Index . . . 93
7.3 Blichfeldt and Minkowski's Convex Body Theorem . . . 95
7.4 Applications: Four Squares and Diophantine Approximation . . . 96
7.5 Sphere Packings . . . 97
7.6 Minkowski–Hlawka: Random Lattices and Existence . . . 98
7.7 Dual Lattices, Covering Radius, and Transference . . . 99
7.8 The Flatness Theorem . . . 100
7.9 Short Vectors and LLL-Reduced Bases . . . 102

8 Chapter VIII: Lattice Points and Polyhedra 104
8.1 Generating Functions of Simplicial Rational Cones . . . 105
8.2 Rational Cones: Bases, Triangulations, and Rationality . . . 107
8.3 Rational Polyhedra and the Generating-Function Valuation . . . 108
8.4 Brion's Theorem: From a Polyhedron to Its Vertex Cones . . . 109
8.5 Ehrhart Polynomials and Reciprocity . . . 111
8.6 Totally Unimodular Polytopes and Dual Decomposition . . . 112

1 Chapter I: Convex Sets at Large

This chapter studies the global properties of finite-dimensional convex sets. Its main themes are finite descriptions of convex hulls, intersection properties of families of convex sets, valuations on indicator functions, and linear projections of polyhedra. Carathéodory's theorem bounds the number of generators needed to represent a point. Radon's theorem turns affine dependence into an intersection of two convex hulls. Helly's theorem passes from intersections of small subfamilies to an intersection of the entire finite family. These results then lead to separation, centerpoints, uniform approximation, and polyhedral elimination.

1.1 I.1 Convex Sets: Definitions, Examples, and Operations

In $\mathbb{R}^d$ we use the inner product $\langle x, y\rangle$, the Euclidean norm $\|x\|$, and the distance $\|x-y\|$. Most results in this chapter depend only on the real vector-space structure. The additional structure of $\mathbb{R}^d$ is needed only when topology, compactness, or volume enters.

Definition 1.1 (Convex combinations, segments, and convex sets). Given $x_1, \dots, x_m \in V$, a point of the form

$$x = \sum_{i=1}^{m} \alpha_i x_i, \qquad \alpha_i \geq 0, \qquad \sum_{i=1}^{m} \alpha_i = 1,$$

is a *convex combination* of these points. The segment joining x and y is

$$[x, y] = \{\lambda x + (1-\lambda)y : 0 \leq \lambda \leq 1\}.$$

A set $C \subset V$ is *convex* if $[x, y] \subset C$ for every $x, y \in C$. By convention, the empty set is convex.

Definition 1.2 (Convex hull). The *convex hull* of a set $S \subset V$, denoted by $\operatorname{conv}(S)$, is the set of all finite convex combinations of points of S.

Example 1.3 (Three standard models). The following sets represent three basic descriptions in convex geometry:

1. a finite intersection of half-spaces, $\{x : \langle a_i, x\rangle \leq b_i,\ i = 1, \dots, m\}$;
2. the convex hull of finitely many points, $\operatorname{conv}\{v_1, \dots, v_N\}$;
3. pointwise constraints in a function space, for example $\{f \in C[0,1] : 0 \leq f(t) \leq 1 \text{ for every } t\}$.

The third example shows that the definition of convexity does not require finite dimension.

Definition 1.4 (Minkowski sum and dilation). For $A, B \subset V$, define

$$A + B = \{a + b : a \in A, b \in B\}, \qquad \lambda A = \{\lambda a : a \in A\}.$$

The set $A + u$ is a translate of A, and λA is a positive dilation when $\lambda > 0$.

Proposition 1.5 (Basic stability of convexity). An arbitrary intersection of convex sets is convex. Linear images and inverse images preserve convexity. If A and B are convex, then $A + B$ is convex. Moreover, if A is convex and $\alpha, \beta \geq 0$, then

$$(\alpha + \beta)A = \alpha A + \beta A.$$

Proof. The first three assertions follow by substituting a segment into the relevant definition. For the nontrivial inclusion in the last identity, assume $\alpha + \beta > 0$. Then

$$\alpha a + \beta b = (\alpha + \beta)\left(\frac{\alpha}{\alpha + \beta}a + \frac{\beta}{\alpha + \beta}b\right) \in (\alpha + \beta)A.$$

The expression in parentheses is a convex combination of two points of A. The reverse inclusion follows by taking $a = b$. ⊛

Remark 1.6 (Geometric interpretation). A linear subspace permits coefficients of either sign with no restriction on their sum. An affine subspace requires the coefficients to sum to 1. A convex set imposes the additional requirement that the coefficients be nonnegative. Convexity may therefore be viewed as affine linear algebra with a sign restriction.

Remark 1.7 (Operations that do not preserve convexity). The union $A \cup B$ is generally not convex, and convexity of $A+B$ does not imply that A and B are convex separately. If negative dilation factors are allowed, the identity $(\alpha+\beta)A = \alpha A+\beta A$ may also fail. Convexity concerns the entire segment between every pair of points, so the shape of the boundary alone does not determine it.

1.2 I.2 Convex Hulls and Carathéodory's Theorem

Theorem 1.8 (Minimality of the convex hull). For every $S \subset V$, the set $\operatorname{conv}(S)$ is the smallest convex set containing S. It is convex, and if $S \subset C$ with C convex, then $\operatorname{conv}(S) \subset C$.

Proof. A convex combination of two convex combinations is again a convex combination of the original points, so $\operatorname{conv}(S)$ is convex. For the other assertion, induct on the number of points in a combination. Write $\sum_{i=1}^{m} \alpha_i x_i$ as a point on the segment between x_m and a convex combination of the first $m-1$ points. Every convex set containing S must contain this point. ⊛

Definition 1.9 (Polytopes and polyhedra). The convex hull of a finite set is a *polytope*. A finite intersection of closed half-spaces,

$$P = \{x \in \mathbb{R}^d : \langle a_i, x \rangle \le b_i,\ i = 1, \dots, m\},$$

is a *polyhedron*. A polytope is bounded, whereas a polyhedron may be unbounded.

Example 1.10 (Simplex, cube, and cross-polytope). The standard d-simplex, cube, and cross-polytope can be written as

$$\begin{aligned} \Delta_d &= \{x \in \mathbb{R}^{d+1} : x_i \ge 0,\ \sum_i x_i = 1\}, \\ [0,1]^d &= \{x \in \mathbb{R}^d : 0 \le x_i \le 1\}, \\ O_d &= \{x \in \mathbb{R}^d : \sum_i |x_i| \le 1\} = \operatorname{conv}\{\pm e_1, \dots, \pm e_d\}. \end{aligned}$$

They emphasize, respectively, affine constraints, coordinate box constraints, and ℓ_1 geometry.

Theorem 1.11 (Carathéodory's theorem). If $S \subset \mathbb{R}^d$ and $x \in \operatorname{conv}(S)$, then there are at most $d+1$ points $v_1, \dots, v_{d+1} \in S$ and nonnegative coefficients λ_i such that

$$x = \sum_{i=1}^{d+1} \lambda_i v_i, \qquad \sum_{i=1}^{d+1} \lambda_i = 1.$$

More precisely, if S lies in an affine subspace of dimension r, then $r+1$ points suffice.

Proof. Begin with any finite representation $x = \sum_{i=1}^{m} \alpha_i v_i$. Suppose that $m > d+1$ and every $\alpha_i > 0$. The augmented vectors $(v_i, 1) \in \mathbb{R}^{d+1}$ are linearly dependent, so there is a nonzero vector $\gamma = (\gamma_i)$ satisfying

$$\sum_i \gamma_i v_i = 0, \qquad \sum_i \gamma_i = 0.$$

Set $t = \min_{\gamma_i > 0} \alpha_i/\gamma_i$ and $\alpha_i' = \alpha_i - t\gamma_i$. The new coefficients remain nonnegative, still sum to 1, and still represent x, while at least one coefficient becomes 0. Repeating the elimination leaves at most $d+1$ terms. ⊛

Remark 1.12 (Eliminating coefficients along a feasible direction). The proof first finds a direction γ that preserves all equality constraints. It then moves along this direction until one nonnegative coefficient reaches 0. The sparsity of basic feasible solutions in linear programming follows from the same argument.

Corollary 1.13 (The convex hull of a compact set is compact). If $S \subset \mathbb{R}^d$ is compact, then $\operatorname{conv}(S)$ is compact.

Proof. Let Δ_d be the standard simplex. By Theorem 1.11, $\operatorname{conv}(S)$ is the image of the compact set $S^{d+1} \times \Delta_d$ under the continuous map

$$(v_1, \dots, v_{d+1}; \lambda) \longmapsto \sum_{i=1}^{d+1} \lambda_i v_i.$$

It is therefore compact. ⊛

1.3 I.3 Application: Positive Homogeneous Polynomials

Let $\mathcal{H}_{2k,n}$ be the space of real homogeneous polynomials of degree $2k$ in n variables. Its dimension is

$$N = \binom{n+2k-1}{2k}.$$

An orthogonal transformation U acts by substitution: $(Up)(x) = p(U^{-1}x)$.

Lemma 1.14 (Orthogonally invariant polynomials). If $p \in \mathcal{H}_{2k,n}$ satisfies $Up = p$ for every $U \in O(n)$, then there is $\gamma \in \mathbb{R}$ such that

$$p(x) = \gamma \|x\|^{2k}.$$

Proof. The group $O(n)$ acts transitively on the unit sphere: any two unit vectors are related by an orthogonal transformation. Thus p is constant, say equal to γ, on the unit sphere. Homogeneity then gives $p(x) = \|x\|^{2k} p(x/\|x\|)$. ⊛

Proposition 1.15 (Finite decomposition into even powers). There are finitely many vectors $c_1, \dots, c_m \in \mathbb{R}^n$ such that, for every $x \in \mathbb{R}^n$,

$$\|x\|^{2k} = \sum_{i=1}^{m} \langle c_i, x\rangle^{2k}.$$

Idea of the proof. For a unit vector c, define $p_c(x) = \langle c, x\rangle^{2k}$. Average over the sphere with respect to the rotation-invariant probability measure:

$$p(x) = \int_{S^{n-1}} \langle c, x\rangle^{2k}\, dc.$$

The preceding lemma gives $p(x) = \gamma \|x\|^{2k}$ with $\gamma > 0$. The set $\{p_c : c \in S^{n-1}\}$ is compact, and its convex hull is compact by the preceding corollary. Hence the integral, as a limit of Riemann sums, lies in that convex hull. Carathéodory's theorem reduces the average to a finite convex combination, and the weights can be absorbed into the lengths of the vectors c_i. ⊛

Definition 1.16 (Positive and nonnegative homogeneous polynomials). A polynomial $p \in \mathcal{H}_{2k,n}$ is *positive* if $p(x) > 0$ for every $x \neq 0$, and *nonnegative* if $p(x) \geq 0$ for every x. The positive polynomials form an open convex cone; the nonnegative polynomials form a closed convex cone.

Proposition 1.17 (Representation of a positive homogeneous polynomial). If $p \in \mathcal{H}_{2k,n}$ is positive, then there are an integer $s \geq k$ and vectors $c_1, \dots, c_m$ such that

$$\|x\|^{2s-2k} p(x) = \sum_{i=1}^{m} \langle c_i, x\rangle^{2s}.$$

Proof. Formally replace $f \in \mathcal{H}_{2k,n}$ by the constant-coefficient differential operator $f(\partial)$, and apply it to the higher-degree identity $\|x\|^{2r} = \sum_i \langle u_i, x\rangle^{2r}$. On one hand,

$$f(\partial) \langle u, x\rangle^{2r} = \frac{(2r)!}{(2r-2k)!} f(u) \langle u, x\rangle^{2r-2k}.$$

For $r \geq 2k$, repeated differentiation also gives

$$f(\partial) \|x\|^{2r} = C_{r,k} \|x\|^{2r-4k} \Phi_r(f)(x), \qquad C_{r,k} = \frac{2^{2k} r!}{(r-2k)!},$$

where Φ_r is a linear map on $\mathcal{H}_{2k,n}$. Checking monomials shows that $\Phi_r(f) = f + O(r^{-1})$ in this finite-dimensional space. Thus $\Phi_r \to I$, and for sufficiently large r the polynomial $q = \Phi_r^{-1}(p)$ is still positive. Applying $q(\partial)$ to the identity for $\|x\|^{2r}$ yields

$$C_{r,k} \|x\|^{2r-4k} p(x) = \frac{(2r)!}{(2r-2k)!} \sum_i q(u_i) \langle u_i, x\rangle^{2r-2k}.$$

Every coefficient on the right is positive, so it can be absorbed into the corresponding linear form. Set $s = r - k$ to obtain the stated representation. ⊛

Remark 1.18 (Strict positivity and the boundary case). Strict positivity places p in the interior of the closed cone, so small perturbations remain positive. The proof uses precisely this interior stability. If one assumes only $p \geq 0$, the polynomial may lie on the boundary of the cone, and a small perturbation in $\Phi_r^{-1}(p)$ may produce negative values. The same perturbation argument then fails.

1.4 I.4 Radon's and Helly's Theorems

Theorem 1.19 (Radon's theorem). Among any $d+2$ or more points in $\mathbb{R}^d$, there are two disjoint nonempty subsets R and B such that

$$\operatorname{conv}(R) \cap \operatorname{conv}(B) \neq \varnothing.$$

Proof. For points $v_1, \dots, v_m$, the augmented vectors $(v_i, 1) \in \mathbb{R}^{d+1}$ are linearly dependent. Choose a nonzero γ satisfying $\sum_i \gamma_i v_i = 0$ and $\sum_i \gamma_i = 0$. Divide the indices according to $\gamma_i > 0$ and $\gamma_i < 0$, and set $\beta = \sum_{\gamma_i>0} \gamma_i = -\sum_{\gamma_i<0} \gamma_i$. Then

$$\sum_{\gamma_i>0} \frac{\gamma_i}{\beta} v_i = \sum_{\gamma_i<0} \frac{-\gamma_i}{\beta} v_i$$

is a convex combination of each group of points. ⊛

Theorem 1.20 (Helly's theorem). Let $A_1, \dots, A_m \subset \mathbb{R}^d$ be a finite family of convex sets with $m \geq d+1$. If every $d+1$ of the sets have nonempty intersection, then

$$\bigcap_{i=1}^{m} A_i \neq \varnothing.$$

Equivalently, if the full system is infeasible, then some subsystem of at most $d+1$ constraints is already infeasible.

Proof. Induct on m. For each i, the induction hypothesis gives a point $p_i \in \bigcap_{j\neq i} A_j$. If two of the p_i coincide, the conclusion follows at once. Otherwise apply Radon's theorem to these points. It gives disjoint index sets I, J and a point

$$p \in \operatorname{conv}\{p_i : i \in I\} \cap \operatorname{conv}\{p_j : j \in J\}.$$

If $k \notin I$, every p_i with $i \in I$ belongs to A_k, so $p \in A_k$. The other group gives the same conclusion when $k \notin J$. Since $I \cap J = \varnothing$, every k is absent from at least one of the two groups. Hence p belongs to every A_k. ⊛

Corollary 1.21 (Infinite Helly theorem for compact convex sets). Let $\{A_i : i \in I\}$ be a family of compact convex subsets of $\mathbb{R}^d$ with $|I| \geq d+1$. If every $d+1$ sets with distinct indices have nonempty intersection, then the entire family has nonempty intersection.

Proof. The finite Helly theorem gives the finite intersection property, and compactness then gives a point in the total intersection. Compactness prevents the intersection points from escaping to infinity. For example, every finite subfamily of the intervals $[n, \infty)$ has a common point, but the total intersection is empty. ⊛

Remark 1.22 (A common linear-algebraic core). Carathéodory's theorem uses affine dependence to remove redundant generators. Radon's theorem uses the signs in the same dependence to partition the point set. Helly's theorem then turns the Radon partition into a local-to-global principle for systems of constraints. This chain explains the constants $d+1$, $d+2$, and $d+1$.

Remark 1.23 (The Helly hypothesis). Helly's theorem requires every $d+1$ sets to have nonempty intersection, not merely every pair. In $\mathbb{R}^2$, three pairwise intersecting convex sets may still have empty total intersection. For an infinite family, compactness allows the finite intersection property to produce a common point.

1.5 I.5 Applications of Helly's Theorem in Combinatorial Geometry

Proposition 1.24 (Kirchberger's separation theorem). Let $R, B \subset \mathbb{R}^d$ be finite sets of red and blue points. Suppose that in every subconfiguration of at most $d+2$ points, the red points and blue points can be strictly separated by a hyperplane. Then one hyperplane strictly separates all of R from all of B.

Proof. Encode the hyperplane $\langle c, x\rangle = \alpha$ by a parameter point $(c, \alpha) \in \mathbb{R}^{d+1}$. For a red point r and a blue point b, define the open half-spaces

$$A_r = \{(c, \alpha) : \langle c, r\rangle < \alpha\}, \qquad A_b = \{(c, \alpha) : \langle c, b\rangle > \alpha\}.$$

A parameter point belongs to all these sets exactly when its hyperplane strictly separates all the points. The hypothesis says that every $d + 2$ of these sets in the parameter space have a common point. Apply Helly's theorem in $\mathbb{R}^{d+1}$. Since the intersection is open, a small perturbation avoids the degenerate parameter $c = 0$. ⊛

Theorem 1.25 (Centerpoint theorem). For every Borel probability measure μ on $\mathbb{R}^d$, there is a point $y \in \mathbb{R}^d$ such that every closed half-space H containing y satisfies

$$\mu(H) \geq \frac{1}{d+1}.$$

If μ is the empirical distribution of n points, then no cut through y can leave more than $dn/(d+1)$ points entirely on the opposite side.

Idea of the proof. Consider all closed half-spaces G whose complements have measure less than $1/(d+1)$. The open complements of any $d+1$ such half-spaces have total measure less than 1, so they cannot cover the whole space. Thus the corresponding G have a common point. First use sufficiently large half-spaces to confine the intersection to a compact set, and then combine Helly's theorem with the finite intersection property for compact sets. This gives a point y belonging to every G. If a half-space containing y had measure less than $1/(d+1)$, its opposite side would contradict the construction. Continuity of measure passes from open to closed half-spaces. ⊛

Remark 1.26 (Half-space depth). In one dimension a centerpoint is a median, so each closed half-line contains at least half the mass. In higher dimensions the bound is $1/(d+1)$. It controls half-space depth in every direction, not the medians of separate coordinates.

1.6 I.6 Application: Uniform Approximation

Given functions $f_1, \dots, f_m : T \to \mathbb{R}$ and a target function $g : T \to \mathbb{R}$, we seek $x = (\xi_1, \dots, \xi_m)$ such that

$$f_x(t) = \sum_{i=1}^{m} \xi_i f_i(t), \qquad |g(t) - f_x(t)| \le \varepsilon \quad (t \in T).$$

Proposition 1.27 (A local test on a finite set). Suppose that T is finite and that, on every subset of T with at most $m+1$ points, there is a single x attaining error at most ε. Then there is an x attaining that error on all of T.

Proof. For every $t \in T$, define the closed convex strip in coefficient space

$$A(t) = \left\{ x \in \mathbb{R}^m : \left| g(t) - \sum_i \xi_i f_i(t) \right| \le \varepsilon \right\}.$$

It is the intersection of two closed half-spaces. The hypothesis states precisely that every subfamily of at most $m+1$ of the sets $A(t)$ has nonempty intersection. Apply Helly's theorem in $\mathbb{R}^m$. ⊛

Proposition 1.28 (Uniform approximation on an infinite set). Suppose there are finitely many sample points $\sigma_1, \dots, \sigma_n \in T$ such that

$$f_x(\sigma_1) = \cdots = f_x(\sigma_n) = 0 \Longrightarrow x = 0.$$

If error at most ε can be attained on every subset with at most $m+1$ points, then a single x attains that error on all of T.

Proof. Let $A = \bigcap_j A(\sigma_j)$. The function

$$N(x) = \max_j |f_x(\sigma_j)|$$

is a norm on $\mathbb{R}^m$ by hypothesis, so it has a positive minimum on the Euclidean unit sphere. Hence A is closed and bounded, and therefore compact. Replace each $A(t)$ by $A(t) \cap A$. The finite Helly theorem gives the finite intersection property, and compactness gives a point in the total intersection. ⊛

Remark 1.29 (The condition on the sample points). For infinite T, solutions of finite subsystems may become unbounded as the subsystem grows. The sample-point condition bounds the coefficient set and places all solutions in one compact set. The finite intersection property then yields a common point.

1.7 I.7 Euler Characteristic and Valuations on Convex Sets

Definition 1.30 (Indicator function). The indicator function of $A \subset \mathbb{R}^d$ is denoted by $[A]$. It equals 1 on A and 0 elsewhere. It satisfies $[A][B] = [A \cap B]$.

Lemma 1.31 (Inclusion-exclusion). For any finite family $A_1, \dots, A_m$,

$$[\cup_{i=1}^m A_i] = \sum_{\varnothing \neq I \subseteq [m]} (-1)^{|I|-1} [\cap_{i \in I} A_i].$$

Proof. Verify the identity $[\cup_i A_i] = 1 - \prod_i (1 - [A_i])$ pointwise and expand the product. ⊛

Definition 1.32 (Algebras of convex sets and valuations). Let $\mathcal{K}(\mathbb{R}^d)$ be the real linear span of indicator functions of compact convex sets, and let $\mathcal{C}(\mathbb{R}^d)$ be the real linear span of indicator functions of closed convex sets. Since intersections of convex sets remain convex, both spaces are algebras under pointwise multiplication. A linear map from either algebra to a real vector space is called a *valuation*.

Theorem 1.33 (Euler characteristic). There is a unique valuation

$$\chi : \mathcal{C}(\mathbb{R}^d) \to \mathbb{R}$$

such that $\chi([A]) = 1$ for every nonempty closed convex set A and $\chi(\varnothing) = 0$.

Main idea of the construction. Uniqueness follows from linearity: if $f = \sum_i \alpha_i [A_i]$, then necessarily $\chi(f) = \sum_{A_i \neq \varnothing} \alpha_i$. The point is to prove that this value is independent of the representation.

For the algebra of compact convex sets, proceed by induction on the dimension. Slice with the height function $\ell(x) = x_d$, and write $H_t = \{x : x_d = t\}$ and $f_t = f|_{H_t}$. Define the jump on entering level t from below by

$$J_f(t) = \chi_{d-1}(f_t) - \lim_{\epsilon \downarrow 0} \chi_{d-1}(f_{t-\epsilon}).$$

For a finite linear combination of compact convex sets, a jump can occur only at the lowest point of one of the sets. Thus $\chi_d(f) = \sum_t J_f(t)$ is a finite sum. Linearity follows from the linearity of slicing, limits, and summation. A nonempty compact convex set has exactly one jump from 0 to 1, so its value is 1.

For closed sets that may be unbounded, intersect with a large ball B_R and let $R \to \infty$. For each fixed finite representation, the result eventually stabilizes and defines a valuation on $\mathcal{C}(\mathbb{R}^d)$. ⊛

Corollary 1.34 (Inclusion-exclusion for Euler characteristic). When the relevant indicator functions belong to the preceding algebra,

$$\chi\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{\varnothing \neq I \subseteq [m]} (-1)^{|I|-1} \chi\left(\bigcap_{i \in I} A_i\right).$$

In particular, $\chi(A \cup B) = \chi(A) + \chi(B) - \chi(A \cap B)$.

Example 1.35 (Finite additivity of Euler characteristic). If a set assembled from convex sets is partitioned into finitely many disjoint pieces, χ adds their values. Overlaps are corrected by inclusion-exclusion. Unlike volume, χ assigns the value 1 to every nonempty closed convex piece, so it records topological and combinatorial structure rather than scale.

Remark 1.36 (The domain of Euler characteristic). Here χ is defined on the algebra spanned by indicator functions of closed convex sets. Values on nonconvex unions are determined by linear relations or inclusion-exclusion. The Euler characteristic of an open interval or open box changes sign with dimension and is not always 1.

1.8 I.8 Linear Transformations, Pushforward Valuations, and Intrinsic Volumes

Theorem 1.37 (Linear pushforward on the indicator-function algebra). If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, there is a unique linear map $T_\# : \mathcal{K}(\mathbb{R}^n) \to \mathcal{K}(\mathbb{R}^m)$ satisfying

$$T_\#[A] = [T(A)]$$

on the generators.

Proof. Defining the map term by term in different representations need not give the same result. For $f \in \mathcal{K}(\mathbb{R}^n)$ and $y \in \mathbb{R}^m$, take Euler characteristic along the fiber:

$$(T_{\#}f)(y) = \chi\big(f \cdot [T^{-1}(y)]\big).$$

If $f = \sum_i \alpha_i [A_i]$, then

$$(T_{\#}f)(y) = \sum_{i : y \in T(A_i)} \alpha_i = \sum_i \alpha_i [T(A_i)](y).$$

This proves at once that the result belongs to $\mathcal{K}(\mathbb{R}^m)$, that the definition is independent of the representation, and that the map is linear. ⊛

Corollary 1.38 (Linear relations are preserved under projection). If compact convex sets satisfy $\sum_i \alpha_i [A_i] = 0$, then

$$\sum_i \alpha_i [T(A_i)] = 0.$$

The assertion can fail for general nonconvex or noncompact sets.

Definition 1.39 (Intrinsic volumes). For $0 \le k \le d$, orthogonally project a compact convex set A onto a k-dimensional subspace L, take its k-dimensional volume, and average over the rotation-invariant probability measure on the Grassmannian $G_k(\mathbb{R}^d)$:

$$w_k(A) = \int_{G_k(\mathbb{R}^d)} \mathrm{vol}_k(P_L A)\, dL.$$

Set $w_0(A) = \chi(A)$ and $w_d(A) = \mathrm{vol}_d(A)$. Each w_k is a valuation invariant under rigid motions, homogeneous of degree k: $w_k(\lambda A) = \lambda^k w_k(A)$ for $\lambda > 0$.

Example 1.40 (Support function). For fixed $c \in \mathbb{R}^d$, the support function of a nonempty compact convex set is

$$h_A(c) = \max_{x \in A} \langle c, x \rangle .$$

It satisfies $h_{A+B}(c) = h_A(c) + h_B(c)$ and, with value zero assigned to the empty set, extends to a valuation on $\mathcal{K}(\mathbb{R}^d)$. The support function linearizes Minkowski addition and describes supporting hyperplanes in polarity and duality.

Remark 1.41 (Euler characteristic along fibers). The map $T_\#$ computes Euler characteristic on every fiber. The resulting function is exactly the indicator function of the image: a nonempty fiber contributes 1, and an empty fiber contributes 0. Thus each generator $[A]$ is sent to the indicator function of its image.

1.9 I.9 Polyhedra and Linear Transformations: Fourier–Motzkin Elimination

Lemma 1.42 (Projection by one coordinate). If $P \subset \mathbb{R}^d$ is a polyhedron, then its image under the coordinate projection

$$\pi(x_1, \dots, x_d) = (x_1, \dots, x_{d-1})$$

is also a polyhedron.

Proof. Partition the inequalities defining P according to whether the coefficient of x_d is positive, negative, or zero, with index sets I_+, I_-, and I_0. Inequalities in I_0 remain unchanged. Those in I_+ give upper bounds for x_d, and those in I_- give lower bounds. A value of x_d satisfies every constraint exactly when every lower bound is no greater than every upper bound. Pairing each lower bound with each upper bound produces a finite system of linear inequalities in $x_1, \dots, x_{d-1}$ that describes $\pi(P)$. ⊛

Theorem 1.43 (Linear images of polyhedra). If $P \subset \mathbb{R}^n$ is a polyhedron and $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then $T(P) \subset \mathbb{R}^m$ is a polyhedron.

Proof. If T is injective, then $T : P \to \operatorname{im} T$ is a change of coordinates, so linear inequalities remain linear inequalities. In general, embed P into the graph of T:

$$\widetilde{T}(x) = (T(x), x) \in \mathbb{R}^{m+n}.$$

The map $\widetilde{T}$ is injective, so $\widetilde{T}(P)$ is a polyhedron. Eliminate the last n coordinates one at a time using the preceding projection lemma. The final image is $T(P)$. ⊛

Definition 1.44 (Fourier–Motzkin elimination). The procedure that eliminates one variable by pairing every lower bound with every upper bound, and then repeats this step variable by variable, is called Fourier–Motzkin elimination.

Remark 1.45 (The number of constraints after elimination). If one step has p upper bounds and q lower bounds, it may create pq new inequalities. Repeated elimination can therefore increase the number of constraints rapidly, and many new constraints may be redundant. The theorem proves the existence of a finite inequality description for every linear image, but naive elimination can be expensive. Dominated or redundant inequalities may be removed after each step.

2 Chapter II: Faces and Extreme Points

This chapter studies the local boundary structure of finite-dimensional closed convex sets. Separation places a linear functional between a point and a set. Supporting hyperplanes cut out faces, and faces reduce a problem to a lower-dimensional affine space. Extreme points are the zero-dimensional instances of this structure. The second half of the chapter applies these ideas to polyhedra, transportation flows, moment cones, cones of nonnegative polynomials, and the positive semidefinite cone. Low-rank positive semidefinite matrices are then used to study convex images of quadratic maps. We write $\operatorname{aff} A$, $\operatorname{conv} A$, $\operatorname{cone} A$, $\operatorname{ext} A$, and $\operatorname{rec} A$ for the affine hull, convex hull, conical hull, set of extreme points, and recession cone.

2.1 II.1 Separation

Definition 2.1 (Affine objects and half-spaces). If $L \leq V$ is a linear subspace, then $u + L$ is an affine subspace. A set of the form

$$H = \{x \in V : f(x) = \alpha\}, \qquad f \neq 0,$$

is an affine hyperplane. The corresponding open half-spaces are $f(x) < \alpha$ and $f(x) > \alpha$; replacing the strict inequalities by weak ones gives the closed half-spaces. If A lies in one of the closed half-spaces determined by H, then H *separates* A. The separation is strict if A lies in an open half-space.

Two sets A and B are separated by H if they lie in opposite closed half-spaces. They are strictly separated if they lie in opposite open half-spaces. Equivalently, separation means that there are f and α such that $f(a) \leq \alpha \leq f(b)$ for all $a \in A$ and $b \in B$.

Definition 2.2 (Algebraically open convex set). A convex set $A \subset V$ is *algebraically open* if its intersection with every line is an open interval, an open ray, the whole line, or the empty set. In $\mathbb{R}^d$, a convex set is algebraically open exactly when it is open in the Euclidean topology. The two notions differ in infinite-dimensional vector spaces.

Theorem 2.3 (Separation theorem). Let V be a real vector space, let $A \subset V$ be a nonempty algebraically open convex set, and suppose $u \notin A$. Then an affine hyperplane through u strictly separates A.

Proof. Translate so that $u = 0$. The one-dimensional case is immediate, so assume $\dim V \geq 2$.

Step 1: the two-dimensional case. Radially project A onto the unit circle. Convexity makes the image a connected arc, and algebraic openness makes it an open arc. If its length exceeded π, it would contain antipodal points, and convexity would force $0 \in A$. A line through the origin in an endpoint direction of the arc strictly separates A.

Step 2: a line disjoint from A. Apply the first step in any two-dimensional plane containing 0 to obtain a line through 0 that does not meet A.

Step 3: maximality and a quotient. Choose a maximal linear subspace H disjoint from A, by dimension or Zorn's lemma. If $\operatorname{codim} H \geq 2$, the quotient map $\pi : V \to V/H$ sends A to an algebraically open convex set. Step 2 gives a line disjoint from $\pi(A)$ whose inverse image contradicts maximality. Hence H has codimension one and strictly separates A. ⊛

Remark 2.4 (Why algebraic openness is needed). This theorem is an algebraic form of the Hahn–Banach theorem. The counterexample in the book using finitely supported sequences shows that, in infinite-dimensional spaces, a convex set without algebraic openness need not admit a separating hyperplane through an exterior point.

2.2 II.2 Convex Sets in Euclidean Space

Lemma 2.5 (Moving inward enters the interior). If $A \subset \mathbb{R}^d$ is convex, $u_0 \in \operatorname{int} A$, and $u_1 \in A$, then

$$u_t = (1-t)u_0 + tu_1 \in \operatorname{int} A, \qquad 0 \le t < 1.$$

More precisely, if $B(u_0, \varepsilon) \subset A$, then $B(u_t, (1-t)\varepsilon) \subset A$.

It follows that $\operatorname{int} A$ is convex. If a convex subset of $\mathbb{R}^d$ has empty interior, it cannot contain $d+1$ affinely independent points and therefore lies in a proper affine subspace. This is why the relative interior should be taken inside $\operatorname{aff} A$.

Definition 2.6 (Faces and exposed points). Let $K \subset \mathbb{R}^d$ be closed and convex. If a supporting hyperplane H satisfies $F = K \cap H$, then F is a *face* of K. If $F = \{u\}$, then u is an *exposed point*. A nonempty face different from K is a proper face.

Remark 2.7 (Terminology). The faces in the book are what are often called *exposed faces*. Some sources also call a subset F a face when an interior point of an open segment lies in F only if both endpoints lie in F. Such a face need not be cut out by a single supporting hyperplane. These notes follow the book's convention.

Theorem 2.8 (Supporting hyperplane at a boundary point). If a convex set $K \subset \mathbb{R}^d$ has nonempty interior and $u \in \partial K$, then a hyperplane through u supports K.

Proof. The set $\operatorname{int} K$ is open and convex and does not contain u. Apply Theorem 2.3 to obtain a hyperplane through u that strictly separates the interior. Closure and convexity extend the separation to K. ⊛

Corollary 2.9 (Boundary points lie on proper faces). If $K \subset \mathbb{R}^d$ is closed, convex, and has nonempty interior, every $u \in \partial K$ belongs to a proper face.

Proof. Closedness gives $u \in K$. If H is a supporting hyperplane through u, then $K \cap H$ is nonempty and contains no interior point. It is therefore a proper face containing u. ⊛

Theorem 2.10 (Separation of a point and a convex set). If $A \subset \mathbb{R}^d$ is nonempty and convex and $u \notin A$, then an affine hyperplane through u separates A.

Idea of the proof. Work in the smallest affine space $L = \operatorname{aff} A$, where A has nonempty relative interior. If $u \notin L$, choose a hyperplane through u that separates L. If $u \in L$, apply the separation theorem in L to the algebraically open convex set $\operatorname{relint}_L A$. Since $A \subset \operatorname{cl}(\operatorname{relint}_L A)$, the separation extends to A, and the hyperplane can then be extended to $\mathbb{R}^d$. ⊛

Example 2.11 (Ball, cube, and simplex). Every proper face of a Euclidean ball is a point. Faces of a cube arise by fixing selected coordinates to 0 or 1. Faces of a d-simplex are exactly the convex hulls of subsets of its vertices. The ball has a smooth boundary, while the cube shows that one boundary point may admit many supporting hyperplanes.

2.3 II.3 Extreme Points and the Finite-Dimensional Krein–Milman Theorem

Definition 2.12 (Extreme point). A point $a \in A$ is an *extreme point* of A if

$$a = \frac{b+c}{2}, \quad b, c \in A \quad \Longrightarrow \quad b = c = a.$$

The set of extreme points is denoted by $\operatorname{ext} A$. Equivalently, a does not lie in the relative interior of any nondegenerate segment contained in A.

Proposition 2.13 (Linear functionals select extreme points). Let $f : V \to \mathbb{R}$ be linear. If f has a unique maximum or minimum on A, that optimizer is an extreme point. More generally, if the optimal set $F = \{x \in A : f(x) = \max_A f\}$ has an extreme point u, then $u \in \operatorname{ext} A$.

Proof. If $u = (a+b)/2$, then $f(u) = (f(a) + f(b))/2$. Since neither endpoint exceeds the optimum, equality forces a and b to lie in the optimal set. Extremality of u within that set gives $a = b = u$. ⊛

Theorem 2.14 (Finite-dimensional Krein–Milman and Minkowski theorem). If $K \subset \mathbb{R}^d$ is compact and convex, then

$$K = \operatorname{conv}(\operatorname{ext} K).$$

Induction on dimension. If the affine dimension of K is less than d, apply induction in $\operatorname{aff} K$. For a boundary point u, Theorem 2.8 gives a proper face F containing u. This face has lower dimension and $\operatorname{ext} F \subset \operatorname{ext} K$, so $u \in \operatorname{conv}(\operatorname{ext} K)$. If u is an interior point, intersect K with any line through u. The intersection is a closed interval $[a, b]$ whose endpoints lie on the boundary. Since u is a convex combination of a and b, it too is generated by extreme points. ⊛

Corollary 2.15 (Linear optimization at an extreme point). If K is nonempty, compact, and convex and f is linear, then f attains its maximum at an extreme point of K.

Proof. The maximizer set is a nonempty compact face. Apply Theorem 2.14 to this face and use the fact that an extreme point of a face is an extreme point of K. ⊛

Compactness is essential for this generation result. The book also proves that every closed convex set containing no line has at least one extreme point. This fact is used in Section II.16 for unbounded sets.

2.4 II.4 Extreme Points of Polyhedra

Let

$$P = \{x \in \mathbb{R}^d : \langle c_i, x\rangle \le \beta_i,\ i = 1, \dots, m\}, \qquad I(u) = \{i : \langle c_i, u\rangle = \beta_i\}.$$

Theorem 2.16 (Vertex criterion from active constraints). A point $u \in P$ is a vertex if and only if $\{c_i : i \in I(u)\}$ spans $\mathbb{R}^d$. In particular, at least d linearly independent constraints are active at every vertex.

Proof. If the active normals do not span $\mathbb{R}^d$, choose nonzero y orthogonal to all of them. The inactive constraints have strict slack at u, so $u \pm \varepsilon y \in P$ for sufficiently small $\varepsilon > 0$. Thus u is the midpoint of two distinct points of P. Conversely, if the active normals span the space and $u = (v + w)/2$, every active inequality must be an equality at both v and w. The equations $\langle c_i, x\rangle = \beta_i$ for $i \in I(u)$ have at most one solution, so $v = w = u$. ⊛

Corollary 2.17 (One direction of the Weyl–Minkowski theorem). Every bounded polyhedron is a polytope.

Proof. Each vertex is uniquely determined by some set of d independent active constraints, so there are only finitely many vertices. Apply Theorem 2.14. ⊛

2.5 II.5 The Birkhoff Polytope

Definition 2.18 (Doubly stochastic matrix). A matrix $X = (x_{ij}) \in \mathbb{R}^{n\times n}$ is *doubly stochastic* if

$$x_{ij} \geq 0, \qquad \sum_i x_{ij} = 1 \quad (j = 1, \dots, n), \qquad \sum_j x_{ij} = 1 \quad (i = 1, \dots, n).$$

The doubly stochastic matrices form the Birkhoff polytope B_n. A permutation $\sigma \in S_n$ corresponds to the permutation matrix P_σ.

Theorem 2.19 (Birkhoff–von Neumann theorem). The vertices of B_n are exactly the $n!$ permutation matrices. Equivalently, every doubly stochastic matrix is a convex combination of permutation matrices.

Dimension-counting proof. The affine space of matrices with all row and column sums equal to 1 has dimension $(n-1)^2$. If X is a vertex, the vertex criterion requires at least $(n-1)^2$ active nonnegativity constraints, hence at least that many zero entries. If every row had at least two positive entries, there would be at most $n(n-2) < (n-1)^2$ zeros, a contradiction. Some row therefore has only one positive entry. Its row sum makes that entry equal to 1, and the column sum makes all other entries in the same column zero. Deleting that row and column leaves an extreme point of B_{n-1}; otherwise a decomposition of the smaller matrix would extend to one of X. Induction completes the proof. ⊛

Example 2.20 (Assignment problem). Given a payoff matrix $C = (c_{ij})$, maximize $\sum_{ij} c_{ij}x_{ij}$ over B_n. A linear objective has an optimum at a vertex, and every vertex is a permutation matrix. Thus the linear programming relaxation has an integral assignment as an optimum.

2.6 II.6 Permutation Polytopes and the Schur–Horn Theorem

For $a = (a_1, \dots, a_n)$, define

$$P(a) = \operatorname{conv}\{\sigma(a) : \sigma \in S_n\}.$$

This polytope is the image of B_n under $X \mapsto Xa$. If the coordinates of a are not all equal, then $\dim P(a) = n - 1$.

Theorem 2.21 (Schur–Horn theorem). Given $\lambda = (\lambda_1, \dots, \lambda_n)$:

1. if a real symmetric or complex Hermitian matrix A has eigenvalues λ_i, then $\operatorname{diag} A \in P(\lambda)$;
2. conversely, every $a \in P(\lambda)$ is the diagonal of a real symmetric matrix with eigenvalues λ_i.

The Schur direction. For a real symmetric matrix, write $A = U \operatorname{diag}(\lambda) U^T$ with U orthogonal. For a complex Hermitian matrix, write $A = U \operatorname{diag}(\lambda) U^*$ with U unitary. In either case,

$$a_k = A_{kk} = \sum_i |u_{ki}|^2 \lambda_i.$$

The matrix $D = (|u_{ki}|^2)$ is doubly stochastic, so $D = \sum_\sigma \theta_\sigma P_\sigma$. Hence $a = D\lambda = \sum_\sigma \theta_\sigma \sigma(\lambda) \in P(\lambda)$. Horn's converse construction is deeper and is quoted without proof in the book. ⊛

Remark 2.22 (Convexification of spectral data). Eigenvalues depend nonlinearly on a matrix, but once the spectrum is fixed, the set of all possible diagonals is exactly a polytope.

2.7 II.7 Transportation Polyhedra

Let $G = (V, E)$ be a directed graph with net demand b_i at vertex i. A feasible flow $x = (x_{ij})$ satisfies

$$\sum_{j:j\to i} x_{ji} - \sum_{j:i\to j} x_{ij} = b_i, \qquad x_{ij} \geq 0.$$

The feasible flows form the transportation polyhedron $T(G, b)$.

Proposition 2.23 (The support of an extreme flow is acyclic). If x is an extreme point of $T(G, b)$, the undirected support graph formed by its positive-flow edges contains no cycle and is therefore a forest.

Proof. If the positive support contains a cycle, orient the cycle arbitrarily. Add ε along forward edges and subtract it along backward edges to obtain y; use the opposite perturbation to obtain z. Each vertex has one entering and one leaving cycle edge, so flow balance is unchanged. Taking ε smaller than the least positive flow on the cycle preserves nonnegativity. Then $x = (y + z)/2$ with $y \neq z$, contradicting extremality. ⊛

Corollary 2.24 (Integral supply and demand give integral extreme flows). If every $b_i \in \mathbb{Z}$, then every extreme point of $T(G, b)$ is an integral flow.

Leaf-removal proof. The positive support is a forest, so it has a vertex of degree at most one. The balance equation at that vertex uniquely determines the incident flow using only addition and subtraction of integers. Delete the leaf, update the neighboring demand, and continue recursively. ⊛

Thus, when the feasible region is nonempty and bounded, the minimum-cost transportation problem has an integral optimum supported on a forest. The Birkhoff polytope is the special case of unit supply and demand on a bipartite graph.

2.8 II.8 Convex Cones

Definition 2.25 (Cone, extreme ray, and base). A convex cone $K \subset V$ satisfies $\alpha x + \beta y \in K$ for all $x, y \in K$ and $\alpha, \beta \geq 0$. A ray $\operatorname{cone}(u)$ is an *extreme ray* if $u = (x+y)/2$ with $x, y \in K$ implies $x, y \in \operatorname{cone}(u)$. A set $B \subset K \setminus \{0\}$ is a *base* of K if every nonzero $u \in K$ has a unique representation $u = \lambda b$ with $\lambda > 0$ and $b \in B$.

If a hyperplane H supports K and $H \cap K \neq \varnothing$, then $0 \in H$. Otherwise scaling an intersection point along its ray would cross H and violate support. Every nonempty face of a cone is therefore a cone.

Lemma 2.26 (Extreme rays and extreme points of a base). If B is a convex base and $u = \lambda b \neq 0$, then u spans an extreme ray of K if and only if $b \in \operatorname{ext} B$.

Proof. Multiplying a convex decomposition of b by λ gives a decomposition of u in the cone. Conversely, normalize the two terms in a decomposition of u by the unique base representation. Extremality of b forces the normalized directions to agree. ⊛

Corollary 2.27 (Krein–Milman theorem for cones). If $K \subset \mathbb{R}^d$ has a compact convex base, every $u \in K$ is a conical combination of generators of extreme rays. Moreover, K is closed.

Proof. Apply Theorem 2.14 to the compact base and restore the radial scale. Since the base is compact and bounded away from the origin, a bounded neighborhood meets only a bounded interval of scales. Its intersection with K is therefore a continuous image of a compact set, which proves closedness. ⊛

Remark 2.28 (Extreme rays). Every nonzero $u \in K$ satisfies $u = (0 + 2u)/2$, so nonzero cone generators are not extreme points. The corresponding extremal objects are rays. Whether the origin is extreme depends on whether the cone is pointed. After taking a compact base, extreme rays correspond exactly to ordinary extreme points of the base.

2.9 II.9 The Moment Curve and the Moment Cone

Fix $\alpha < \beta$ and let

$$g(t) = (1, t, t^2, \dots, t^d) \in \mathbb{R}^{d+1}, \qquad M_{d+1}[\alpha, \beta] = \operatorname{cone}\{g(t) : t \in [\alpha, \beta]\}.$$

The set $\operatorname{conv} g([\alpha, \beta])$ is a compact convex base, so the moment cone is closed and full-dimensional. The restriction of a linear functional c to the curve, $\langle c, g(t)\rangle$, is a univariate polynomial of degree at most d.

Proposition 2.29 (Number of atoms in a boundary moment). Suppose $u \in \partial M_{d+1}$ and

$$u = \sum_{j=1}^{m} \lambda_j g(t_j), \quad \lambda_j > 0, \quad \alpha \le t_1 < \cdots < t_m \le \beta.$$

Then $m \le (d+2)/2$. Equality can occur only when d is even and $t_1 = \alpha$, $t_m = \beta$.

Proof. Choose a supporting hyperplane $\langle c, x\rangle = 0$ to the cone at u, and set $p(t) = \langle c, g(t)\rangle \ge 0$. Since $0 = \langle c, u\rangle = \sum_j \lambda_j p(t_j)$ with $\lambda_j > 0$, every t_j is a root of p. Interior roots have even multiplicity; only endpoint roots may be simple. The weakest degree bound is therefore $2(m-2) + 2 \le d$. This gives $m \le (d+2)/2$, and equality requires both endpoints. ⊛

Corollary 2.30 (Moment representations with few atoms). Every $u \in M_{d+1}[\alpha, \beta]$ is a conical combination of at most

$$m = \begin{cases} (d+1)/2, & d \text{ odd}, \\ (d+2)/2, & d \text{ even} \end{cases} = \left\lfloor \frac{d}{2} \right\rfloor + 1$$

points of the moment curve.

Moving an interior point to the boundary. Boundary points are covered by Proposition 2.29. If u is interior, continuously shrink the parameter interval to $[\alpha + t, \beta - t]$ until u first reaches the boundary of the corresponding moment cone. Apply the boundary result there. If u remains in the cone when the interval collapses to one point, it is already a multiple of one generator. ⊛

The conical version of Carathéodory's theorem gives only $d + 1$ generators. The moment curve nearly halves this number because every interior root of a nonnegative polynomial has even multiplicity.

2.10 II.10 Application: Quadrature of High Algebraic Degree

Theorem 2.31 (Positive quadrature with m nodes). Let $\rho \geq 0$ be continuous on $[\alpha, \beta]$. For every positive integer m, there are nodes $\alpha \leq t_1 < \cdots < t_m \leq \beta$ and weights $\lambda_i \geq 0$ such that

$$\int_\alpha^\beta f(t)\rho(t)\,dt = \sum_{i=1}^m \lambda_i f(t_i)$$

for every polynomial f of degree at most $2m-1$.

Proof. Set $d = 2m-1$ and form the moment vector

$$u = \left(\int \rho, \int t\rho, \ldots, \int t^d\rho\right).$$

Riemann sums are conical combinations of points on the moment curve. Since the moment cone is closed, $u \in M_{d+1}$. The preceding sparse representation gives $u = \sum_{i=1}^m \lambda_i g(t_i)$. Pairing this identity with the coefficient vector of $f(t) = \langle c, g(t)\rangle$ gives the formula. ⊛

Example 2.32 (Low-order formulas for uniform measure). On $[0,1]$, one node gives the midpoint rule

$$\int_0^1 f(t)\,dt = f(1/2),$$

which is exact for linear polynomials. A two-node formula exact through degree three is

$$\int_0^1 f(t)\,dt = \frac{1}{2}f\left(\frac{1}{2} - \frac{1}{2\sqrt{3}}\right) + \frac{1}{2}f\left(\frac{1}{2} + \frac{1}{2\sqrt{3}}\right).$$

In general, m nodes cannot integrate every polynomial of degree $2m$ exactly. A nonnegative polynomial with a double root at every node gives a contradiction.

Remark 2.33 (Why the algebraic degree increases). The increased degree comes from root multiplicities of nonnegative polynomials on the boundary of the moment cone. The integral functional is represented by finitely many moments, and a measure with few atoms matches those moments.

2.11 II.11 The Cone of Nonnegative Polynomials

Identify $\mathbb{R}^{d+1}$ with the polynomials of degree at most d, and define

$$K_+[\alpha,\beta] = \{p : p(t) \geq 0 \text{ for every } t \in [\alpha,\beta]\}.$$

This is a closed, full-dimensional convex cone. Its boundary consists exactly of the nonnegative polynomials that vanish somewhere on the interval.

Proposition 2.34 (Extreme rays of the nonnegative polynomial cone). The cone $K_+[\alpha,\beta]$ has a compact base. If $d = 2k$, the generators of its extreme rays are exactly

$$\lambda \prod_{i=1}^{k} (t-t_i)^2, \qquad \lambda(t-\alpha)(\beta-t)\prod_{i=1}^{k-1}(t-t_i)^2.$$

If $d = 2k+1$, they are exactly

$$\lambda(t-\alpha)\prod_{i=1}^{k}(t-t_i)^2, \qquad \lambda(\beta-t)\prod_{i=1}^{k}(t-t_i)^2,$$

where $\lambda > 0$, the points $t_i \in [\alpha,\beta]$, and repetitions are allowed.

Saturation by roots. Intersect the cone with $\int_\alpha^\beta p = 1$ to obtain a closed base. In finite dimension, $\int |p|$ is equivalent to the Euclidean norm of the coefficient vector, so the base is bounded and compact. If an extreme polynomial p had degree less than d, then $p \pm \varepsilon(t-\alpha)p$ would remain nonnegative for small ε and give a nonproportional decomposition. Hence $\deg p = d$. If $p = qr$ and r has no root on the interval, then $r \pm \varepsilon$ remain positive and again give a nonproportional decomposition. Thus all d roots must lie in the interval. Interior roots have even multiplicity and endpoints may be simple, leaving exactly the four displayed forms. Conversely, if $p = (p_1+p_2)/2$ with $p_1, p_2 \geq 0$, every root of p, with its required multiplicity, must occur in both p_1 and p_2. These d roots use the entire degree, so both summands are proportional to p. ⊛

Corollary 2.35 (Sum-of-squares representation on an interval). If $p \geq 0$ on $[\alpha, \beta]$, then

$$\begin{aligned} d = 2k: \quad & p(t) = \sum_j r_j(t)^2 + (t-\alpha)(\beta - t) \sum_i q_i(t)^2, \\ & \deg r_j \leq k, \qquad \deg q_i \leq k-1; \\ d = 2k+1: \quad & p(t) = (t-\alpha) \sum_i q_i(t)^2 + (\beta - t) \sum_j r_j(t)^2, \\ & \deg q_i, \deg r_j \leq k. \end{aligned}$$

Proof. Apply the conical Krein–Milman theorem to express p as a conical combination of the generators in Proposition 2.34, then absorb each product into a squared polynomial. ⊛

Remark 2.36 (The distinction between one and several variables). Univariate nonnegative polynomials have the preceding sum-of-squares representation. The multivariate case is different. The Motzkin and Choi–Lam polynomials show that a nonnegative polynomial need not be a sum of polynomial squares, so a corresponding semidefinite relaxation may have a gap.

2.12 II.12 The Positive Semidefinite Cone

Let Sym_n be the space of real symmetric matrices with inner product

$$\langle A, B \rangle = \operatorname{tr}(AB).$$

The notation $A \succeq 0$ means that $x^T A x \geq 0$ for every $x \in \mathbb{R}^n$. These matrices form the cone $\mathbb{S}^n_+$. Its interior consists of positive definite matrices, and $\{A \succeq 0 : \operatorname{tr} A = 1\}$ is a compact convex base.

Proposition 2.37 (Rank determines the minimal face). If $A \succeq 0$ and $\operatorname{rank} A = r$, then A lies in the relative interior of a face F with

$$\dim F = \frac{r(r+1)}{2},$$

and F is isometric to $\mathbb{S}^r_+$ through a linear map that preserves rank. When $r = n$, A lies in the interior of the whole cone.

Proof. If $r = 0$, then $A = 0$ and its minimal face is $\{0\}$. If $r = n$, positive definiteness places A in the interior of the whole cone. For $0 < r < n$, after orthogonal diagonalization, assume $A = \operatorname{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$ with $\lambda_i > 0$. Let $C = \operatorname{diag}(0, \dots, 0, 1, \dots, 1) \succeq 0$. The hyperplane $\langle C, X\rangle = 0$ supports $\mathbb{S}^n_+$ and contains A. If $X \succeq 0$ lies in this hyperplane, its last $n - r$ diagonal entries vanish. The inequality $|x_{ij}|^2 \le x_{ii}x_{jj}$ then makes the corresponding rows and columns zero. The face consists exactly of arbitrary positive semidefinite blocks in the upper-left $r \times r$ position. Its dimension is $r(r+1)/2$, and the block of A is positive definite, so A is a relative interior point. ⊛

Corollary 2.38 (The face lattice of the positive semidefinite cone). Every subspace $L \le \mathbb{R}^n$ determines a face

$$F_L = \{X \succeq 0 : L \subseteq \ker X\}.$$

If $\operatorname{codim} L = r$, then $\dim F_L = r(r+1)/2$. Moreover, F_{L_1} is a face of F_{L_2} exactly when $L_2 \subseteq L_1$. In particular, the extreme rays of the positive semidefinite cone are generated by the rank-one matrices xx^T.

Idea of the proof. An orthogonal change of basis sends L to the last $n-r$ coordinate directions. The block description in Proposition 2.37 gives the face, its dimension, and the inclusion relation. Taking $r = 1$ gives the extreme rays. ⊛

2.13 II.13 Linear Equations in Positive Semidefinite Matrices

Theorem 2.39 (Barvinok–Pataki rank reduction). Let r be a nonnegative integer, and let an affine subspace $\mathcal{A} \subset \operatorname{Sym}_n$ satisfy $\mathcal{A} \cap \mathbb{S}^n_+ \neq \varnothing$ and

$$\operatorname{codim} \mathcal{A} < \frac{(r+2)(r+1)}{2}.$$

Then there is $X \in \mathcal{A} \cap \mathbb{S}^n_+$ with $\operatorname{rank} X \le r$. If the section is defined by k linear equations, one may take

$$\operatorname{rank} X \le \left\lfloor \frac{\sqrt{8k+1} - 1}{2} \right\rfloor.$$

Proof. The closed convex set $K = \mathcal{A} \cap \mathbb{S}^n_+$ contains no line, so it has an extreme point X_0. Set $m = \operatorname{rank} X_0$. Proposition 2.37 places X_0 in the relative interior of a positive semidefinite face F of dimension $m(m+1)/2$. If $\dim(F \cap \mathcal{A}) > 0$, then X_0 can be perturbed in both directions inside the section and is not extreme. Hence $\operatorname{codim} \mathcal{A} \geq \dim F = m(m+1)/2$. Comparison with the hypothesis gives $m \leq r$. ⊛

Corollary 2.40 (Two quadratic equations and semidefinite lifting). Let $q_A(x) = x^T Ax$ and $q_B(x) = x^T Bx$. The equations

$$q_A(x) = \alpha, \qquad q_B(x) = \beta$$

have a real solution if and only if there is $X \succeq 0$ satisfying

$$\langle A, X \rangle = \alpha, \qquad \langle B, X \rangle = \beta.$$

Consequently, the image of $x \mapsto (q_A(x), q_B(x))$ is a convex cone in $\mathbb{R}^2$.

Proof. The forward implication uses $X = xx^T$. Conversely, apply Theorem 2.39 to the two linear equations to obtain a feasible matrix of rank at most one. When nonzero, it has the form xx^T. Thus the quadratic image is the linear image of the positive semidefinite cone under $X \mapsto (\langle A, X \rangle, \langle B, X \rangle)$ and is therefore convex. This is Dines's theorem. ⊛

The book also proves a critical improvement. Let r be a positive integer. If $\operatorname{codim} \mathcal{A} = (r+2)(r+1)/2$, $n \geq r+2$, and $\mathcal{A} \cap \mathbb{S}^n_+$ is nonempty and bounded, then a feasible matrix of rank at most r still exists. The proof studies the map sending a boundary point of an $(r+2)$-dimensional section to its kernel line in real projective space. A topological nonembedding result produces two boundary points with a common kernel; moving within their common positive semidefinite face reaches a boundary point of lower rank.

Remark 2.41 (The critical rank bound). The rank bound comes from comparing the number of constraints with the dimension of the minimal face, not from ordinary linear elimination. At equality, boundedness of the section and the stated dimension condition are also needed.

2.14 II.14 Application: Convexity Theorems for Quadratic Maps

Theorem 2.42 (Brickman's theorem). Let $n > 2$, let $q_1, q_2 : \mathbb{R}^n \to \mathbb{R}$ be quadratic forms, and let $\mathbb{S}^{n-1} = \{x : \|x\| = 1\}$. Then

$$\{(q_1(x), q_2(x)) : x \in \mathbb{S}^{n-1}\}$$

is convex in $\mathbb{R}^2$.

Proof. Let A_1, A_2 be the matrices of the quadratic forms and consider the spectrahedron

$$B = \{X \succeq 0 : \operatorname{tr} X = 1\}.$$

Its image under $X \mapsto (\langle A_1, X\rangle, \langle A_2, X\rangle)$ is convex. Fix a point in this image. Including the trace constraint gives at most three independent linear equations. The feasible section is nonempty and bounded. When the codimension is less than 3, use Theorem 2.39; at codimension 3, use the critical improvement with $r = 1$. In either case there is a rank-one feasible matrix $X = xx^T$. Since $\operatorname{tr} X = 1$, we have $\|x\| = 1$. Thus the linear image of B is exactly the quadratic image of the sphere. ⊛

Corollary 2.43 (Toeplitz–Hausdorff theorem). For a complex matrix A, its numerical range

$$W(A) = \{z^* A z : z \in \mathbb{C}^n,\ \|z\| = 1\}$$

is convex in the complex plane.

For $n = 1$ the numerical range is a point. If $n > 1$, write $z = x + iy$. Both $\operatorname{Re}(z^*Az)$ and $\operatorname{Im}(z^*Az)$ are quadratic forms in $2n$ real variables, so the result follows from Theorem 2.42.

Theorem 2.44 (Au–Yeung–Poon finite mixture theorem). Let $r \geq 1$, $k < (r+2)(r+1)/2$, and $n \geq r + 2$. If $\phi : \mathbb{R}^n \to \mathbb{R}^k$ consists of k quadratic forms, then every point of $\operatorname{conv} \phi(\mathbb{S}^{n-1})$ is a convex combination of at most r points of $\phi(\mathbb{S}^{n-1})$.

Idea of the proof. Lift a convex combination to a trace-one positive semidefinite matrix. Together with the trace equation, the codimension is at most the critical value. Apply Theorem 2.39 in the strict case and the bounded-section improvement at equality to reduce the rank to at most r. The spectral decomposition then writes the matrix as a convex combination of at most r rank-one matrices from the sphere. ⊛

Remark 2.45 (Semidefinite lifting). The point is not to verify directly that the quadratic image is closed under segments. Instead, one identifies the quadratic image of a nonconvex set with the linear image of a positive semidefinite convex set. The low-rank theorem shows that, for these low-dimensional images, the semidefinite lift introduces no additional image points.

2.15 II.15 Application: Graph Realization

Definition 2.46 (Gram matrix and graph realization). The Gram matrix of vectors $v_1, \dots, v_n \in \mathbb{R}^d$ is $X = (\langle v_i, v_j \rangle)$. It satisfies $X \succeq 0$ and $\operatorname{rank} X \le d$. Conversely, every matrix satisfying these conditions is the Gram matrix of vectors in $\mathbb{R}^d$. A weighted graph $G = (V, E; \rho)$ is *realizable in* $\mathbb{R}^d$ if its vertices can be placed at points $v_i \in \mathbb{R}^d$ such that $\|v_i - v_j\| = \rho_{ij}$ for every edge.

The distance constraints are linear equations in the Gram matrix:

$$x_{ii} - 2x_{ij} + x_{jj} = \rho_{ij}^2, \qquad (i,j) \in E.$$

Thus realizability is equivalent to positive semidefinite feasibility, while realization in $\mathbb{R}^d$ also requires rank at most d.

Proposition 2.47 (Low-dimensional realization of sparse graphs). If G has k edges and

$$k < \frac{(d+2)(d+1)}{2},$$

then G is realizable if and only if it is realizable in $\mathbb{R}^d$. In particular, every realizable graph with at most nine edges can be realized in $\mathbb{R}^3$.

Proof. Each edge gives one linear matrix equation. If the positive semidefinite system is feasible, Theorem 2.39 gives a feasible matrix of rank at most d. A Gram factorization yields a realization in $\mathbb{R}^d$. ⊛

The bound is sharp. The $d+2$ vertices of a regular $(d+1)$-simplex have all pairwise distances equal to 1. They are realizable in $\mathbb{R}^{d+1}$ but not in $\mathbb{R}^d$, and the number of edges is exactly $(d+2)(d+1)/2$.

The book also fixes one vertex at $v_n = 0$ and forms the Gram matrix of the remaining $n-1$ vectors. For a connected graph, path lengths bound every $\|v_i\|$, so the positive semidefinite feasible section is bounded. Combining this observation with the critical rank reduction shows that, at the critical number of edges, realizability is still equivalent to realizability in $\mathbb{R}^d$, except for a copy of K_{d+2} together with isolated vertices.

Remark 2.48 (Interpretation). Distance realization and Euclidean matrix completion are positive semidefinite section problems. The rank of the matrix is the required embedding dimension. With fewer edge constraints, the section is more likely to meet a low-rank face.

2.16 II.16 Closed Convex Sets

Definition 2.49 (Recession cone). For a nonempty closed convex set $A \subset \mathbb{R}^d$, define

$$\operatorname{rec} A = \{u : a + tu \in A \text{ for every } a \in A,\ t \geq 0\}.$$

If A contains no ray, set $\operatorname{rec} A = \{0\}$.

Lemma 2.50 (A ray direction extends from every point). If a closed convex set A contains the ray $a + tu$ for $t \geq 0$, then it contains the same ray starting from every $b \in A$. All such directions form the closed convex cone $\operatorname{rec} A$.

Proof. Translate so that $a = 0$. For fixed $t \geq 0$ and $0 < \varepsilon < 1$, convexity gives

$$(1-\varepsilon)b + tu = (1-\varepsilon)b + \varepsilon(\varepsilon^{-1}tu) \in A.$$

Letting $\varepsilon \downarrow 0$ and using closedness gives $b+tu \in A$. The direction set is closed under nonnegative linear combinations. If u is not a recession direction, then some $tu \notin A$; closedness excludes all nearby directions, so the recession cone is closed. ⊛

Closedness keeps the limiting endpoint of a ray in the set. For a polyhedron $P = \{x : \langle c_i, x\rangle \leq \beta_i\}$,

$$\operatorname{rec} P = \{u : \langle c_i, u\rangle \leq 0,\ i = 1, \dots, m\}.$$

Lemma 2.51 (Factoring out lineality). If a closed convex set A contains a line, there is a subspace L such that its orthogonal projection $A' \subset L^{\perp}$ is closed, convex, and contains no line, and

$$A = A' + L.$$

The space L is the lineality space of A, consisting of the directions in which one may move indefinitely in both directions from every point.

Idea of the proof. Let L consist of all directions allowing indefinite motion in both directions. Convexity makes it a subspace and gives $A+L = A$. Project onto $L^{\perp}$ to obtain A'. Maximality excludes lines from A', and closedness follows from convergent sequences together with $A+L = A$. ⊛

Theorem 2.52 (Extreme-point and recession decomposition). If a closed convex set $A \subset \mathbb{R}^d$ contains no line, every $x \in A$ can be written as

$$x = y + z, \qquad y \in \operatorname{conv}(\operatorname{ext} A), \quad z \in \operatorname{rec} A.$$

Idea of the proof. Induct on dimension and work in $\operatorname{aff} A$. Intersect A with a line through x. The intersection is either a closed interval or a closed ray. Its endpoints lie on the boundary and hence in proper faces of lower dimension; apply induction to these faces. A ray contributes a vector in $\operatorname{rec} A$, and every extreme point of a face is an extreme point of A. Combining the resulting terms gives the stated decomposition. ⊛

3 Chapter III: Convex Sets in Topological Vector Spaces

This chapter extends separation and extreme-point theory to topological vector spaces. In a locally convex space, continuous linear functionals separate a closed convex set from an exterior point. Local convexity ensures that sufficiently many such functionals exist, and the weak-star topology supplies compactness for convex sets of functions and measures. The results are applied to subsets of L^∞, linear control, vector measures, and probability measures subject to moment constraints.

3.1 III.1 Separation in Euclidean and General Vector Spaces

Definition 3.1 (Separation and strict separation). Let A, B be nonempty subsets of a real vector space V. An affine hyperplane $\{x : f(x) = \alpha\}$ *separates* A and B if there are a nonzero linear functional $f : V \to \mathbb{R}$ and $\alpha \in \mathbb{R}$ such that

$$f(a) \le \alpha \le f(b) \qquad \text{for every } a \in A, b \in B.$$

The separation is strict when both inequalities are strict.

Definition 3.2 (Minkowski difference and algebraic interior). Define

$$A - B = \{a - b : a \in A, b \in B\}.$$

A point x of a convex set $A \subset V$ belongs to its *algebraic interior* if the intersection of A with every line through x contains an open interval around x. The set A is algebraically open if all its points have this property.

Theorem 3.3 (Separation in Euclidean space). If $A, B \subset \mathbb{R}^d$ are nonempty, convex, and disjoint, then an affine hyperplane separates them.

Proof. The set $A - B$ is convex and does not contain 0. Separate $A - B$ from the origin to obtain a nonzero f such that $f(a - b) \le 0$. Thus $\sup_{a \in A} f(a) \le \inf_{b \in B} f(b)$. Choosing α between these two values gives the result. ⊛

Theorem 3.4 (Strict separation of a closed convex set and an exterior point). If $A \subset \mathbb{R}^d$ is closed and convex and $u \notin A$, then an affine hyperplane strictly separates A and u.

Proof. Choose a small ball D such that $u + D$ and $A + D$ are disjoint. They are open and convex, so ordinary separation is automatically strict. Alternatively, choose a nearest point $v \in A$ to u and use the perpendicular bisector with normal $u - v$. ⊛

Theorem 3.5 (Algebraic separation theorem). Let $A, B \subset V$ be nonempty, convex, and disjoint. If A is algebraically open, then an affine hyperplane separates A and B. The conclusion still holds when the algebraic interior of A is merely nonempty.

Remark 3.6 (Infinite-dimensional spaces). In finite dimensions, every convex set that is not full-dimensional lies in a smaller affine subspace. In infinite dimensions, a convex set may have empty interior without lying in any hyperplane. Separation therefore requires a condition such as nonempty algebraic or topological interior, or closedness together with local convexity.

3.2 III.2 Topological Vector Spaces, Convex Sets, and Hyperplanes

Definition 3.7 (Topological vector space). A real vector space V with a Hausdorff topology is a *topological vector space* if

$$(x, y) \mapsto x + y, \qquad (\lambda, x) \mapsto \lambda x$$

are continuous. Translations and nonzero scalar multiplications are then homeomorphisms.

Definition 3.8 (Interior and closure). For $A \subset V$, write $\operatorname{int}(A)$ for its interior and $\operatorname{cl}(A)$ for its closure. A point x belongs to $\operatorname{cl}(A)$ exactly when every neighborhood of x meets A.

Lemma 3.9 (Interior segment lemma). If A is convex, $x_0 \in \operatorname{int}(A)$, and $x_1 \in A$, then

$$(1 - t)x_0 + tx_1 \in \operatorname{int}(A) \qquad (0 \le t < 1).$$

Proof. Choose a neighborhood $U \subset A$ of x_0. The affine map

$$T(x) = \frac{x - x_1}{1 - t} + x_1$$

is continuous and sends $x_t = (1-t)x_0 + tx_1$ to x_0. Hence $T^{-1}(U)$ is a neighborhood of x_t. Each of its points is a convex combination of x_1 and a point of U, so it lies in A. ⊛

Theorem 3.10 (Closure and interior of a convex set). If $A \subset V$ is convex, then both $\operatorname{int}(A)$ and $\operatorname{cl}(A)$ are convex.

Proof. The assertion about the interior follows from the preceding lemma. If $x_0, x_1 \in \operatorname{cl}(A)$, continuity of $(x, y) \mapsto tx + (1-t)y$ sends sufficiently small neighborhoods of x_0, x_1 into any prescribed neighborhood of the target point. Choose points of A in those neighborhoods and take their convex combination. The target point therefore belongs to $\operatorname{cl}(A)$. ⊛

Theorem 3.11 (A hyperplane is closed or dense). An affine hyperplane H in a topological vector space satisfies exactly one of

$$H \text{ is closed,} \qquad \text{or} \qquad \operatorname{cl}(H) = V.$$

Theorem 3.12 (Closed hyperplanes and continuous linear functionals). Let $f : V \to \mathbb{R}$ be a nonzero linear functional. The level set $H_\alpha = \{x : f(x) = \alpha\}$ is closed if and only if f is continuous.

Remark 3.13. Linearity does not imply continuity in infinite dimensions. The kernel of a discontinuous linear functional is a dense hyperplane. Although it has algebraic codimension one, it cannot serve as a closed supporting hyperplane defined by a continuous functional.

Definition 3.14 (Continuous dual and weak-star topology). The continuous dual V^* is the vector space of continuous linear functionals $f : V \to \mathbb{R}$. The *weak-star topology* is the weakest topology making every evaluation map

$$f \longmapsto f(x), \qquad x \in V,$$

continuous. Its basic neighborhoods restrict only finitely many values $f(x_1), \dots, f(x_m)$.

Theorem 3.15 (Banach–Alaoglu theorem in the form used here). If U is a neighborhood of the origin in V, then the polar

$$U^{\text{polar}} = \{f \in V^* : |f(x)| \le 1 \text{ for every } x \in U\}$$

is compact in the weak-star topology.

Proof. Identify f with the coordinate family $(f(x))_{x\in U} \in [-1,1]^U$. Tikhonov's theorem makes the product space compact. The neighborhood U is absorbing: for every $x \in V$, some positive multiple of x belongs to U. Consequently, a limit of these coordinate families determines a linear functional on all of V by rescaling, and the linear identities pass to the limit. The resulting functional is bounded by 1 on U. A linear functional bounded on a neighborhood of the origin is continuous, so the limit still belongs to U^{polar}. Thus the image of the polar is closed in the product. Evaluation at any point outside U is a rescaled evaluation at a point of U, so the product topology on this image agrees with the weak-star topology. ⊛

Lemma 3.16 (Closed cone generated by a compact base). If $C \subset V$ is compact and convex and $0 \notin C$, then

$$\text{cone}(C) = \{\lambda x : \lambda \ge 0, x \in C\}$$

is a closed convex cone.

Proof. Let u lie outside the cone. Since C is closed and does not contain the origin, there is a neighborhood W of the origin disjoint from C. Continuity of scalar multiplication gives a neighborhood U_1 of u and a $\delta > 0$ such that $\alpha U_1 \subset W$ whenever $|\alpha| < \delta$. Hence U_1 misses λC for every $\lambda > \delta^{-1}$. The set $\{\lambda x : 0 \le \lambda \le \delta^{-1},\ x \in C\}$ is a continuous image of the compact set $[0, \delta^{-1}] \times C$, and is therefore closed. A neighborhood U_2 of u misses this set. The intersection $U_1 \cap U_2$ misses the entire cone. ⊛

3.3 III.3 Separation in Topological Vector Spaces

Lemma 3.17 (Topological openness implies algebraic openness). The section of an open set by any affine line in a topological vector space is open. Hence every open set is algebraically open.

Theorem 3.18 (Separation when the interior is nonempty). If $A, B \subset V$ are convex and disjoint and $\operatorname{int}(A) \neq \varnothing$, then a closed affine hyperplane separates them. Equivalently, there is a nonzero continuous linear functional f such that

$$f(a) \leq f(b) \qquad (a \in A, b \in B).$$

Proof. Algebraic separation first gives a hyperplane H. If H were not closed, it would be dense in V and would meet the interior of A. A hyperplane through an interior point cannot place all of A on one side. Thus H is closed and its defining functional is continuous. ⊛

Definition 3.19 (Locally convex space). A topological vector space V is *locally convex* if every neighborhood of every point contains a convex neighborhood of that point. Every normed space is locally convex.

Theorem 3.20 (Strict separation in a locally convex space). Let V be locally convex, let $A \subset V$ be closed and convex, and let $u \notin A$. Then a continuous linear functional f strictly separates A and u:

$$\sup_{a \in A} f(a) < f(u).$$

Proof. Closedness provides a neighborhood of u disjoint from A. Local convexity gives convex neighborhoods W_1, W_2 of the origin such that

$$(A + W_2) \cap (u + W_1) = \varnothing.$$

These sets are open and convex. The preceding theorem supplies a closed separating hyperplane, and separation of two open sets is strict. ⊛

Remark 3.21 (The role of local convexity). Local convexity ensures the existence of enough continuous linear functionals. Without it, a space may have only the zero continuous functional, making continuous support and separation impossible.

3.4 III.4 The Krein–Milman Theorem

Theorem 3.22 (Krein–Milman theorem). Let V be a locally convex topological vector space and let $K \subset V$ be nonempty, compact, and convex. Then

$$K = \operatorname{cl}\big(\operatorname{conv}(\operatorname{ext}(K))\big).$$

Proof. First use Zorn's lemma to choose a minimal member F among the nonempty compact convex extreme subsets of K. If F contained two distinct points, strict separation would give a continuous functional distinguishing them. The minimizer set of that functional on F would be a smaller nonempty compact convex extreme subset, a contradiction. Thus F is a point and K has an extreme point.

Now set $A = \operatorname{cl}(\operatorname{conv}(\operatorname{ext}(K)))$. If $u \in K \setminus A$, choose a continuous linear functional f with $f(u) < f(a)$ for every $a \in A$. Minimizing f on K gives a compact convex face. The first part supplies an extreme point of K in that face, while separation forces this point to lie outside A, a contradiction. ⊛

Corollary 3.23 (A linear functional attains an extremum at an extreme point). If $f : V \to \mathbb{R}$ is continuous and linear, some $u \in \operatorname{ext}(K)$ satisfies

$$f(u) = \max_{x \in K} f(x).$$

Remark 3.24 (Finite and infinite dimensions). A finite-dimensional compact convex set equals the convex hull of its extreme points. In infinite dimensions one generally needs the closure of that convex hull. For subsets of L^∞, finite convex combinations of extreme points need not cover the whole set.

Remark 3.25 (Extreme points and linear functionals). Extreme points are the elements of a compact convex set that admit no nontrivial convex decomposition. Every continuous linear functional attains its optimum at some extreme point.

3.5 III.5 Polyhedra in L^∞

Definition 3.26 (Dual pairing of L^1 and L^∞). Give $L^1[0,1]$ its norm topology and regard $L^\infty[0,1]$ as its continuous dual with the weak-star topology. The pairing is

$$\langle f, u\rangle = \int_0^1 f(t)u(t)\,dt.$$

Functions are identified when they agree almost everywhere.

Proposition 3.27 (Weak-star compactness of the infinite-dimensional interval). The set

$$B = \{u \in L^\infty[0,1] : 0 \le u(t) \le 1 \text{ almost everywhere}\}$$

is weak-star compact.

Proof. The set $\{u : |u| \le 1\}$ is the polar of the unit ball of L^1, hence compact by Banach–Alaoglu. The set B is a closed subset described by inequalities of continuous pairings. ⊛

Proposition 3.28 (An L^∞ polyhedron and its extreme points). Given $f_1, \dots, f_m \in L^1[0,1]$ and $\beta \in \mathbb{R}^m$, let

$$P = \left\{u \in L^\infty : 0 \le u \le 1 \text{ almost everywhere}, \int_0^1 f_i(t)u(t)\,dt = \beta_i, i = 1, \dots, m\right\}.$$

Then P is weak-star compact and convex. If $u \in \operatorname{ext}(P)$, then

$$u(t) \in \{0,1\} \qquad \text{almost everywhere.}$$

Proof. If u lies strictly between 0 and 1 on a set of positive measure, then

$$\{0 < u < 1\} = \bigcup_{n\ge 2}\{1/n \le u \le 1 - 1/n\} \quad \text{up to a null set.}$$

There is therefore an $\varepsilon > 0$ such that $E = \{\varepsilon \le u \le 1-\varepsilon\}$ has positive measure. Choose $m+1$ pairwise disjoint positive-measure subsets $E_j \subset E$. In

$$v = \sum_{j=1}^{m+1} \varepsilon_j \mathbf{1}_{E_j},$$

choose coefficients, not all zero, that preserve all m moment constraints. This is a homogeneous system of m equations in $m+1$ unknowns. Scale the solution so that $\max_j |\varepsilon_j| < \varepsilon$. Then

$u \pm v$ remain in $[0,1]$ and satisfy every moment constraint, so u is not extreme. ⊛

Corollary 3.29 (Controls with values in a polytope). Fix a polytope $Q \subset \mathbb{R}^r$, functions $f_1, \dots, f_m \in L^1([0,1], \mathbb{R}^r)$, and scalars β_i. The set

$$P_Q = \left\{ \begin{array}{l} u \in L^\infty([0,1], \mathbb{R}^r) : u(t) \in Q \text{ almost everywhere}, \\ \displaystyle\int_0^1 \langle f_i(t), u(t) \rangle \, dt = \beta_i, \qquad i = 1, \dots, m \end{array} \right\}$$

is weak-star compact and convex. Every extreme point of P_Q takes values among the vertices of Q almost everywhere.

Proof. A finite inequality description of Q makes the pointwise constraint weak-star closed, and boundedness of Q places the feasible set in a weak-star compact ball. Suppose $u(t)$ is not a vertex on a set of positive measure. The finitely many faces of Q allow us to choose a fixed nonzero direction e, a number $\varepsilon > 0$, and a positive-measure set E on which $u(t) \pm \varepsilon e \in Q$. Partition E into $m+1$ positive-measure sets and use the preceding homogeneous-system argument with perturbations in direction e. This gives two distinct feasible controls whose average is u. ⊛

Remark 3.30. Under finite-dimensional box constraints, basic feasible solutions lie mostly at endpoints. The almost-everywhere endpoint property is the infinite-dimensional form of the same phenomenon and explains the endpoint-valued controls below.

3.6 III.6 Linear Optimal Control

Definition 3.31 (Linear control system). Let $A : [0,1] \to \mathbb{R}^{n \times n}$ and $B : [0,1] \to \mathbb{R}^{n \times r}$ be continuous. Consider

$$x'(t) = A(t)x(t) + B(t)u(t), \qquad x(0) = x_0,$$

where $u \in L^\infty([0,1], \mathbb{R}^r)$. The solution is absolutely continuous and satisfies the equation almost everywhere. If $X'(t) = A(t)X(t)$ and $X(0) = I$, then $X(t)$ is invertible and

$$x(t) = X(t) \left(x_0 + \int_0^t X(s)^{-1} B(s) u(s) \, ds \right).$$

Proposition 3.32 (The L^∞ linear programming form). Fixing the endpoint $x(1) = x_1$ gives finitely many integral equations

$$\int_0^1 \langle F_i(t), u(t)\rangle \, dt = \beta_i, \qquad F_i \in L^1([0,1], \mathbb{R}^r).$$

A running cost linear in the state and control, with continuous coefficients, has the form

$$\int_0^1 \langle g(t), u(t)\rangle \, dt + \text{constant}, \qquad g \in L^1([0,1], \mathbb{R}^r).$$

If the control is pointwise constrained to a polytope, the problem is a linear program in L^∞.

Example 3.33 (Reduction of a second-order control problem). Suppose the coefficients in

$$x''(t) = a_0(t)x(t) + a_1(t)x'(t) + b(t)u(t)$$

are smooth, the initial and terminal values of x, x' are fixed, and $0 \le u \le 1$. Writing $y = (x, x')^{\mathsf T}$ gives

$$y' = Ay + Bu, \qquad A = \begin{pmatrix} 0 & 1 \\ a_0 & a_1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 \\ b \end{pmatrix}.$$

If $X' = AX$ and $X(0) = I$, then the terminal condition is

$$\int_0^1 X(t)^{-1} B(t) u(t)\, dt = X(1)^{-1} y_1 - y_0.$$

These are two scalar moment equations. For a cost $\int_0^1 (c_0 x + c_1 x' + du)\, dt$, put $c = (c_0, c_1)^{\mathsf T}$. Substitution of the solution formula and interchange of the two integrals gives a constant plus $\int_0^1 g(s)u(s)\, ds$, where

$$g(s) = d(s) + \int_s^1 c(t)^{\mathsf T} X(t) X(s)^{-1} B(s)\, dt.$$

Thus both the endpoint conditions and the objective depend linearly on the scalar control.

Corollary 3.34 (Endpoint-valued optimal control). If a scalar control problem is feasible and $0 \le u \le 1$, then an optimal control u^* exists with

$$u^*(t) \in \{0, 1\} \qquad \text{almost everywhere.}$$

If the optimum is unique, the unique optimizer has this property.

Proof. The feasible set is weak-star compact and the linear objective is weak-star continuous. An optimum therefore exists. Choose an extreme point of the optimal face and use the extreme-point description from Section III.5. A unique optimizer is itself extreme. ⊛

3.7 III.7 Lyapunov's Convexity Theorem

Theorem 3.35 (Lyapunov convexity theorem for nonatomic measures). Given $f_1, \dots, f_m \in L^1[0,1]$, define for measurable E

$$\Phi(E) = \left(\int_E f_1(t)\,dt, \dots, \int_E f_m(t)\,dt \right).$$

Then the range

$$\{\Phi(E) : E \subset [0,1] \text{ measurable}\}$$

is a compact convex subset of $\mathbb{R}^m$.

Proof. Replace indicators of sets by all functions $0 \le u \le 1$, and define

$$T(u) = \left(\int_0^1 f_1 u, \dots, \int_0^1 f_m u \right).$$

The box $B = \{0 \le u \le 1\}$ is weak-star compact and convex, so $T(B)$ is compact and convex. Conversely, for $a \in T(B)$, the fiber $B_a = \{u \in B : T(u) = a\}$ is nonempty, compact, and convex. Krein–Milman gives an extreme point of the fiber, and Section III.5 shows that this point is almost everywhere an indicator $\mathbf{1}_E$. Hence $a = \Phi(E)$, so the indicators and the convex box have the same image. ⊛

Corollary 3.36 (Convexity of an integral image). Let $S \subset \mathbb{R}$ be Lebesgue measurable and let $F_i : S \times [0,1] \to \mathbb{R}$ be given functions. Let Γ consist of the measurable selections $x : [0,1] \to S$ for which each composition $t \mapsto F_i(x(t), t)$ is measurable and integrable. Then the set

$$\left\{ \left(\int_0^1 F_1(x(t),t)\,dt, \dots, \int_0^1 F_m(x(t),t)\,dt \right) : x \in \Gamma \right\}$$

is convex.

Proof. Choose $x, y \in \Gamma$ and $0 \le \theta \le 1$. Apply Lyapunov's theorem to the integrable functions $F_i(x(t), t) - F_i(y(t), t)$. Their integral over some measurable set E equals θ times their integral over $[0, 1]$. The selection that equals x on E and y outside E belongs to Γ and realizes the desired convex combination of the two integral vectors. ⊛

Remark 3.37 (Splitting under a nonatomic measure). Choosing between two selections on complementary measurable sets realizes any convex combination of their finite vectors of integrals. The equality concerns the integrals, not the pointwise values of the selections. With atoms, the prescribed integrals may not be attainable by such a partition.

3.8 III.8 The Simplex of Probability Measures

Definition 3.38 (Simplex of measures). Give $C[0, 1]$ the uniform norm and let $V[0, 1] = C[0, 1]^*$ carry the weak-star topology. Define

$$\mathcal{P} = \left\{ \mu \in V[0,1] : \int f \, d\mu \ge 0 \text{ for every } f \ge 0, \quad \int 1 \, d\mu = 1 \right\}.$$

This is the set of Borel probability measures on $[0, 1]$. The point t corresponds to the Dirac measure δ_t, for which $\int f \, d\delta_t = f(t)$.

Proposition 3.39 (Compactness of the probability simplex). The set $\mathcal{P}$ is weak-star compact and convex.

Proposition 3.40 (Extreme points of the probability simplex).

$$\operatorname{ext}(\mathcal{P}) = \{\delta_t : t \in [0, 1]\}.$$

Proof. If δ_t is written as the average of two probability measures, testing against continuous functions with a strict maximum at t forces both measures to equal δ_t.

Conversely, let μ be extreme. For $0 < h < 1$, weighting μ by h and normalizing writes μ as a convex combination of two probability measures. Extremality gives

$$\int fh \, d\mu = \left(\int f \, d\mu \right) \left(\int h \, d\mu \right).$$

By linear extension this holds for all continuous f, h. The kernel of μ is therefore a multiplicatively closed hyperplane. Compactness and the finite intersection property show that it is the kernel of evaluation at some point t, so $\mu = \delta_t$. ⊛

3.9 III.9 Extreme Points of Intersections and Their Applications

Lemma 3.41 (One hyperplane constraint). Suppose every intersection of a convex set K with a line is a closed bounded interval, a point, or empty. If H is an affine hyperplane, then every extreme point of $K \cap H$ is a convex combination of at most two extreme points of K.

Theorem 3.42 (Dual Carathéodory theorem). Under the preceding condition, if an affine subspace A has codimension m, then every extreme point of $K \cap A$ is a convex combination of at most $m+1$ extreme points of K.

Proof. Intersect successively with hyperplanes to see first that the target point is a convex combination of finitely many extreme points of K. Choose a representation using the fewest points:

$$u = \sum_{i=1}^{r} \lambda_i u_i.$$

The points u_i are affinely independent, and u lies in the relative interior of their simplex. If $r > m+1$, the codimension-m section contains a nondegenerate segment through u inside this simplex, contradicting extremality in the section. ⊛

Proposition 3.43 (Atomic optimizers under moment constraints). Given continuous functions $f_1, \dots, f_m, g$ and numbers α_i, let

$$B = \left\{ \mu \in \mathcal{P} : \int f_i \, d\mu = \alpha_i,\ i = 1, \dots, m \right\}.$$

If $B \neq \varnothing$, the maximum and minimum of $\int g \, d\mu$ over B are attained by probability measures supported on at most $m+1$ points. Every value between this minimum and maximum is also attained.

Proof. The set B is the intersection of $\mathcal{P}$ with m weak-star closed affine hyperplanes, so it is compact and convex. A continuous linear objective attains an extremum at an extreme point of B. The dual Carathéodory theorem writes this point as a convex combination of at most $m+1$ extreme points of $\mathcal{P}$, and Section III.8 identifies those points as Dirac measures. The image of the compact convex set B under $\mu \mapsto \int g \, d\mu$ is a compact interval, which proves the final assertion. ⊛

Remark 3.44 (Finite support). This gives the basic convex-geometric form of the finite-support principle: each additional independent affine constraint increases the required support of an extreme feasible measure by at most one point.

4 Chapter IV: Polarity, Duality, and Linear Programming

This chapter begins with polar sets in Euclidean space and develops duality for vector spaces, topological vector spaces, and convex cones. Weak duality follows from nonnegative pairings with a dual cone; strong duality is tied to closedness of a suitable linear image. The later sections treat polyhedral linear programming, semidefinite programming, L^∞ programming, uniform approximation, and optimal transport.

4.1 IV.1 Polarity in Euclidean Space

Definition 4.1 (Polar set). For nonempty $A \subset \mathbb{R}^d$, define

$$A^{\text{polar}} = \{y \in \mathbb{R}^d : \langle y, x\rangle \le 1 \text{ for every } x \in A\}.$$

The polar is always a closed convex set containing the origin.

Proposition 4.2 (Basic operations with polar sets). For nonempty $A, B \subset \mathbb{R}^d$ and $\lambda > 0$,

$$A \subset B \Longrightarrow B^{\text{polar}} \subset A^{\text{polar}}, \qquad (\lambda A)^{\text{polar}} = \lambda^{-1} A^{\text{polar}},$$

$$\Big(\bigcup_i A_i\Big)^{\text{polar}} = \bigcap_i A_i^{\text{polar}}, \qquad A \subset (A^{\text{polar}})^{\text{polar}}.$$

If L is a linear subspace, then $L^{\text{polar}} = L^\perp$. If $A = \operatorname{conv}\{v_1, \dots, v_m\}$, then

$$A^{\text{polar}} = \{y : \langle y, v_i\rangle \le 1, i = 1, \dots, m\}.$$

Theorem 4.3 (Bipolar theorem). If $A \subset \mathbb{R}^d$ is closed and convex and $0 \in A$, then

$$(A^{\text{polar}})^{\text{polar}} = A.$$

For arbitrary nonempty A,

$$(A^{\text{polar}})^{\text{polar}} = \operatorname{cl}\big(\operatorname{conv}(A \cup \{0\})\big).$$

Proof. The inclusion $A \subset (A^{\text{polar}})^{\text{polar}}$ is immediate. If $u \notin A$, strict separation gives $c \neq 0$ and $\alpha > 0$ such that

$$\langle c, x\rangle < \alpha < \langle c, u\rangle \qquad (x \in A).$$

For $y = c/\alpha$, we have $y \in A^{\text{polar}}$ but $\langle y, u\rangle > 1$, so $u \notin (A^{\text{polar}})^{\text{polar}}$. ⊛

Corollary 4.4 (The converse direction of Weyl–Minkowski). Every polytope is a polyhedron.

Proof. After translating and restricting to $\operatorname{aff}(P)$, assume $0 \in \operatorname{int}(P)$. If P is the convex hull of finitely many points, then P^{polar} is given by finitely many inequalities. It is bounded because 0 is interior to P, hence it is a polytope. Expressing P^{polar} as a finite convex hull and taking the polar again turns its vertices into a finite half-space description of P. ⊛

Lemma 4.5 (Polar of a convex cone). If $K \subset \mathbb{R}^d$ is a convex cone, then

$$K^{\text{polar}} = \{y : \langle y, x\rangle \leq 0 \text{ for every } x \in K\}.$$

Thus K^{polar} is a convex cone and

$$(K_1 + K_2)^{\text{polar}} = K_1^{\text{polar}} \cap K_2^{\text{polar}}.$$

Remark 4.6 (Polarity and faces). Polarity reverses inclusion and exchanges generators with constraint normals. A supporting hyperplane $\langle y, x\rangle = 1$ cuts out a face in the original space, while y is a point describing that face in the dual space. Vertex and inequality descriptions are therefore exchanged by polarity.

Theorem 4.7 (Lawrence's polarity transform for valuations). Let $\mathcal{C}(\mathbb{R}^d)$ be the vector space generated by indicators of nonempty closed convex sets. There is a linear operator $D : \mathcal{C}(\mathbb{R}^d) \to \mathcal{C}(\mathbb{R}^d)$ such that

$$D(\mathbf{1}_A) = \mathbf{1}_{A^{\text{polar}}}$$

for every nonempty closed convex set A. Thus every linear identity among indicators of closed convex sets remains valid with the same coefficients after taking polars.

Proof. For $\varepsilon > 0$, let $H_{y,\varepsilon} = \{x : \langle x, y\rangle \geq 1 + \varepsilon\}$. Define

$$D_\varepsilon(g)(y) = \chi(g) - \chi(g\mathbf{1}_{H_{y,\varepsilon}}).$$

Multiplication by a fixed halfspace indicator and application of the Euler characteristic are linear operations. For $g = \mathbf{1}_A$, the expression is 1 precisely when $A \cap H_{y,\varepsilon}$ is empty. As $\varepsilon \downarrow 0$, it tends to 1 exactly when $\langle x, y\rangle \leq 1$ for all $x \in A$. Taking this limit on a finite linear combination of indicators therefore defines D and proves its linearity. ⊛

Example 4.8 (Dual valuations). If ν is a valuation on indicators of closed convex sets, then $\nu^*(g) = \nu(Dg)$ is also a valuation. For convex bodies containing the origin in their interiors, this gives the dual volume $\mathrm{vol}^*(A) = \mathrm{vol}(A^{\mathrm{polar}})$. For a closed convex cone K, the normalized spherical angle is the measure of K on the unit sphere under the rotation invariant probability measure. The angle of K^{polar} is its exterior spherical angle, and the polarity transform shows that this too extends to a valuation on cone indicators.

4.2 IV.2 Membership in the Moment Cone

Definition 4.9 (Moment cone). Let

$$g(t) = (1, t, t^2, \ldots, t^d), \qquad 0 \leq t \leq 1,$$

and define

$$M_{d+1} = \mathrm{cone}\{g(t) : 0 \leq t \leq 1\} \subset \mathbb{R}^{d+1}.$$

A vector $a = (a_0, \ldots, a_d)$ belongs to M_{d+1} exactly when there is a finite nonnegative Borel measure μ such that

$$a_k = \int_0^1 t^k \, d\mu(t), \qquad k = 0, \ldots, d.$$

Definition 4.10 (Cone of nonnegative polynomials). Let

$$K_d = \left\{ c = (c_0, \ldots, c_d) : \sum_{k=0}^{d} c_k t^k \geq 0 \text{ on } [0,1] \right\}.$$

Lemma 4.11 (Dual membership test).

$$a \in M_{d+1} \quad \Longleftrightarrow \quad \langle a, c\rangle \geq 0 \text{ for every } c \in K_d.$$

Proof. Membership $c \in K_d$ means $\langle c, g(t)\rangle \geq 0$ for every t, so $-K_d = M_{d+1}^{\text{polar}}$. Since M_{d+1} is a closed convex cone, the bipolar theorem gives $M_{d+1} = (-K_d)^{\text{polar}}$, which is the displayed family of inequalities. ⊛

Theorem 4.12 (Hankel criterion for truncated moments). Let $a = (a_0, \dots, a_d)$.

1. If $d = 2m$, then $a \in M_{d+1}$ if and only if

$$[a_{i+j}]_{i,j=0}^{m} \succeq 0, \qquad [a_{i+j+1} - a_{i+j+2}]_{i,j=0}^{m-1} \succeq 0.$$

2. If $d = 2m+1$, then $a \in M_{d+1}$ if and only if

$$[a_{i+j+1}]_{i,j=0}^{m} \succeq 0, \qquad [a_{i+j} - a_{i+j+1}]_{i,j=0}^{m} \succeq 0.$$

Proof. On $[0,1]$, nonnegative polynomials of degree at most $2m$ are sums of squares of polynomials of degree at most m and of $t(1-t)$ times squares of degree at most $m-1$. In degree at most $2m+1$, they are sums of t times squares and $(1-t)$ times squares of degree at most m. Apply the dual membership test to these generators. Nonnegativity of the resulting quadratic forms in the coefficient vectors is exactly positive semidefiniteness of the displayed Hankel matrices. ⊛

Remark 4.13 (Finitely many semidefinite conditions). The infinitely many polynomial tests are equivalent to finitely many linear matrix inequalities. The moment problem is therefore a membership problem for positive semidefinite cones.

4.3 IV.3 Duality of Vector Spaces

Definition 4.14 (Dual pair). A nondegenerate bilinear form

$$\langle \cdot, \cdot \rangle : E \times F \to \mathbb{R}$$

on real vector spaces E, F is a *duality*. Nondegeneracy means that a vector pairing to zero with every vector on the other side must itself be zero.

Example 4.15 (Four basic dual pairs).

1. $E = F = \mathbb{R}^d$ with the standard inner product;
2. symmetric matrices with $\langle X, Y \rangle = \operatorname{tr}(XY)$;
3. $L^1[0,1]$ and $L^\infty[0,1]$ with the integral pairing;
4. $C[0,1]$ and signed Borel measures with $\langle f, \mu \rangle = \int f \, d\mu$.

Definition 4.16 (Adjoint map). Let $A : E_1 \to E_2$ be linear. A linear map $A^* : F_2 \to F_1$ is its *adjoint* if

$$\langle Ae, f \rangle_2 = \langle e, A^* f \rangle_1 \qquad (e \in E_1, f \in F_2).$$

Nondegeneracy makes the adjoint unique whenever it exists.

Example 4.17 (Constraint map). If

$$A(e) = \big(\langle e, f_1 \rangle, \dots, \langle e, f_m \rangle\big),$$

then

$$A^*(y) = \sum_{i=1}^{m} y_i f_i.$$

Thus multiplying constraints by multipliers and adding them is exactly the adjoint operation.

Definition 4.18 (Polarity in a dual pair). For nonempty $A \subset E$, define

$$A^{\text{polar}} = \{ f \in F : \langle e, f \rangle \le 1 \text{ for every } e \in A \}.$$

The polar of a subset of F is defined symmetrically in E.

4.4 IV.4 Duality of Topological Vector Spaces

Definition 4.19 (Weak topology of a dual pair). For a dual pair (E, F), the weak topology on E is the weakest topology making every map $e \mapsto \langle e, f\rangle$, for $f \in F$, continuous. The topology on F is defined symmetrically. A basic open set restricts only finitely many pairing values.

Theorem 4.20 (Continuous dual under the weak topology). Under these weak topologies:

1. every $f \in F$ defines a continuous linear functional on E;
2. every continuous linear functional on E has the unique form $e \mapsto \langle e, f\rangle$ for some $f \in F$;
3. the symmetric assertions hold for F.

Proof. If $\phi : E \to \mathbb{R}$ is continuous, some neighborhood

$$U = \{e : |\langle e, f_i\rangle| < \varepsilon_i, i = 1, \dots, m\}$$

is mapped by ϕ to a bounded interval. Hence the common kernel $L = \{e : \langle e, f_i\rangle = 0, i = 1, \dots, m\}$ lies in $\ker \phi$. Linear algebra makes ϕ a linear combination of these coordinate functionals, so $\phi(e) = \langle e, \sum_i \lambda_i f_i\rangle$. Uniqueness follows from nondegeneracy. ⊛

Corollary 4.21 (Topological bipolar theorem). If E, F form the continuous dual pair of locally convex topological vector spaces, and $A \subset E$ is closed, convex, and contains the origin, then

$$(A^{\text{polar}})^{\text{polar}} = A.$$

Remark 4.22. The topology determines which linear functionals are continuous. Changing it changes the continuous dual and may also change the feasible set and the duality gap of a dual problem.

4.5 IV.5 Orders Defined by Convex Cones

Definition 4.23 (Cone order). A convex cone $K \subset V$ defines

$$x \preceq_K y \quad \Longleftrightarrow \quad y - x \in K.$$

This preorder is a partial order if $K \cap (-K) = \{0\}$. If K is closed, the order is stable under limits.

Example 4.24 (Basic cone orders).

1. $K = \mathbb{R}^d_+$: coordinatewise order;
2. $K = \mathbb{S}^n_+$: the Loewner order $X \preceq Y$;
3. $K = L^1_+$ or L^∞_+: almost-everywhere pointwise order;
4. $K = C_+[0,1]$: pointwise order everywhere;
5. the cone of positive measures: comparison against all nonnegative continuous functions.

Definition 4.25 (Dual cone). For a dual pair (E, F), the dual cone of $K \subset E$ is

$$K^* = \{f \in F : \langle x, f\rangle \geq 0 \text{ for every } x \in K\}.$$

If K is closed and convex, then in a locally convex dual setting $(K^*)^* = K$.

Proposition 4.26 (Self-dual and paired cones). Under the natural pairings,

$$(\mathbb{R}^d_+)^* = \mathbb{R}^d_+, \qquad (\mathbb{S}^n_+)^* = \mathbb{S}^n_+,$$

$$(L^\infty_+)^* = L^1_+, \qquad (C_+[0,1])^* = V_+[0,1].$$

4.6 IV.6 Linear Programming Problems

Definition 4.27 (Primal and dual conic programs). Let (E_i, F_i) be dual pairs, let $K_i \subset E_i$ be convex cones, and let $A : E_1 \to E_2$ have adjoint $A^* : F_2 \to F_1$. Given $c \in F_1$ and $b \in E_2$, define the primal problem

$$\begin{aligned} \gamma = \inf_x \quad & \langle x, c\rangle_1 \\ \text{subject to} \quad & Ax \succeq_{K_2} b, \qquad x \succeq_{K_1} 0, \end{aligned}$$

and the dual problem

$$\begin{aligned} \beta = \sup_\ell \quad & \langle b, \ell\rangle_2 \\ \text{subject to} \quad & A^*\ell \preceq_{K_1^*} c, \qquad \ell \succeq_{K_2^*} 0. \end{aligned}$$

Theorem 4.28 (Weak duality and complementary slackness). For every primal feasible x and dual feasible ℓ,

$$\langle x, c\rangle_1 \geq \langle b, \ell\rangle_2, \qquad \text{hence} \qquad \gamma \geq \beta.$$

If

$$\langle x, c - A^*\ell\rangle_1 = 0, \qquad \langle Ax - b, \ell\rangle_2 = 0,$$

then both points are optimal and $\gamma = \beta$. Conversely, if optimizers exist and $\gamma = \beta$, the two complementary conditions hold.

Proof. The cone and dual-cone conditions give $\langle Ax - b, \ell\rangle_2 \geq 0$ and $\langle x, c - A^*\ell\rangle_1 \geq 0$. The adjoint identity yields

$$\langle x, c\rangle_1 \geq \langle x, A^*\ell\rangle_1 = \langle Ax, \ell\rangle_2 \geq \langle b, \ell\rangle_2.$$

Equality holds exactly when both nonnegative slack pairings vanish. ⊛

Remark 4.29 (Sign convention). The primal problem is a minimization and the dual is a maximization, so the primal value is the upper bound. Reversing the primal convention requires reversing the corresponding inequalities.

4.7 IV.7 Zero Duality Gap

Definition 4.30 (Standard form and lifted image cone). With $K_2 = \{0\}$, the primal problem becomes

$$\gamma = \inf\{\langle x, c\rangle : Ax = b, x \in K\},$$

and the dual is

$$\beta = \sup\{\langle b, \ell\rangle : A^*\ell \preceq_{K^*} c\}.$$

Define

$$\widehat{A}(K) = \{(Ax, \langle x, c\rangle) : x \in K\} \subset E_2 \times \mathbb{R}.$$

Theorem 4.31 (Closed-image criterion for zero gap). Let (E_1, F_1) and (E_2, F_2) be continuous dual pairs of locally convex topological vector spaces. Let $K \subset E_1$ be a closed convex cone, let $A : E_1 \to E_2$ be continuous and linear, and suppose its adjoint is continuous. If $\widehat{A}(K)$ is closed in the product topology and the primal is feasible, then

$$\gamma = \beta.$$

If also $\gamma > -\infty$, a primal optimizer exists.

Proof. If $\gamma = -\infty$, weak duality rules out every dual feasible point, so $\beta = -\infty$. Now assume $\gamma > -\infty$. The intersection of the line $\{(b,t) : t \in \mathbb{R}\}$ with the closed cone $\widehat{A}(K)$ records all primal objective values; closedness makes its lower endpoint attainable. Let x_0 be a primal optimizer. For $\varepsilon > 0$, the point $(b, \gamma - \varepsilon)$ lies outside the image cone. Strictly separate it from the cone. Conic scaling shows that the separator can be written (ℓ, σ) with

$$\langle Ax, \ell\rangle + \sigma\langle x, c\rangle \le 0 \quad (x \in K), \qquad \langle b, \ell\rangle + \sigma(\gamma - \varepsilon) > 0.$$

Evaluation at x_0 gives $\langle b, \ell\rangle + \sigma\gamma \le 0$, hence $\sigma < 0$. Normalize to $\sigma = -1$. Then $c - A^*\ell \in K^*$, so ℓ is dual feasible, and

$$\langle b, \ell\rangle > \gamma - \varepsilon.$$

Let $\varepsilon \downarrow 0$ and combine with weak duality. ⊛

Lemma 4.32 (Closed image from a compact base). Let $K \subset V$ have a compact convex base B, let $T : V \to W$ be continuous and linear, and suppose

$$\ker T \cap K = \{0\}.$$

Then $T(K)$ is a closed convex cone.

Proof. The set $T(B)$ is compact and convex and does not contain the origin, while $T(K) = \operatorname{cone}(T(B))$. Apply the closed-cone lemma from Section III.2. ⊛

Remark 4.33 (Sufficient conditions for a closed image). Useful sufficient conditions include a polyhedral cone, whose linear image is closed; a compact base together with $\ker T \cap K = \{0\}$; and the positive definite combination used in Section IV.10. Interior conditions instead give a local stability property at the prescribed right-hand side. This local property is enough for zero gap without requiring the entire image cone to be closed.

Proposition 4.34 (Stability and an interior condition). Suppose the primal is feasible and $\gamma > -\infty$. If

$$(b, \gamma - \varepsilon) \notin \operatorname{cl}(\widehat{A}(K)) \qquad \text{for every } \varepsilon > 0,$$

then $\gamma = \beta$. In finite dimensions this condition holds if A is onto and there is a feasible $x_0 \in \operatorname{int}(K)$.

Proof. For the first assertion, separate $(b, \gamma - \varepsilon)$ from the closure of the image cone. The proof of the closed-image theorem still applies: evaluate the separating functional on feasible points whose objective values tend to γ to obtain $\sigma < 0$, then normalize it.

For the interior condition, choose a balanced neighborhood W of the origin with $x_0 + W \subset K$ and $|\langle x, c\rangle| \le s$ on $x_0 + W$ for some $s > 0$. The set AW is a neighborhood of the origin because a surjective linear map between finite-dimensional spaces is open. Given $\varepsilon > 0$, choose a feasible x_ε with cost below $\gamma + \varepsilon/4$ and $0 < \delta < 1$ such that $\delta(s + |\gamma|) < \varepsilon/4$. For $u = \delta A w \in \delta AW$, the two points

$$x_\pm = (1 - \delta)x_\varepsilon + \delta(x_0 \pm w)$$

are feasible for $b \pm u$, and their costs are below $\gamma + \varepsilon/2$. Let $\gamma_\pm$ be the corresponding perturbed infima. Averaging a feasible point for $b + u$ with one for $b - u$ gives a feasible point for b. Thus $2\gamma \le \gamma_+ + \gamma_-$, and the upper bounds on $\gamma_\pm$ imply $\gamma_\pm \ge \gamma - \varepsilon/2$. The neighborhood $(b + \delta AW) \times (\gamma - 5\varepsilon/4, \gamma - 3\varepsilon/4)$ therefore misses the image cone. ⊛

4.8 IV.8 Polyhedral Linear Programming

Definition 4.35 (Finite-dimensional standard form). For $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$, the classical primal and dual pair is

$$\gamma = \inf\{c^{\mathsf T} x : Ax = b, x \ge 0\}, \qquad \beta = \sup\{b^{\mathsf T} y : A^{\mathsf T} y \le c\}.$$

Theorem 4.36 (Strong duality for polyhedral linear programs). If the primal is feasible, then $\gamma = \beta$. If $\gamma > -\infty$, both primal and dual optimizers exist.

Proof. The cone $\mathbb{R}^n_+$ is polyhedral, so its image under $x \mapsto (Ax, c^{\mathsf T} x)$ is again a polyhedral cone and is closed. Section IV.7 gives zero gap and primal attainment. Put the dual in standard form using nonnegative and slack variables and apply the same argument to obtain dual attainment. ⊛

Corollary 4.37 (Coordinatewise complementary slackness). Primal and dual feasible points x, y are both optimal if and only if

$$x_j\big(c_j - (A^{\mathsf{T}}y)_j\big) = 0, \qquad j = 1, \dots, n.$$

Example 4.38 (The dual of the diet problem). In the primal program, interpret x_j as the quantity of ingredient j, c_j as its unit cost, and A_{ij} as its content of nutrient i. The equations $Ax = b$ specify the required nutrient totals. A dual variable y_i assigns a unit price to nutrient i. The inequality $\sum_i A_{ij}y_i \le c_j$ says that the nutrients in an ingredient cost no more than that ingredient. At an optimum, every ingredient used in a positive quantity has equality in this price constraint. The dual variables are unrestricted in sign because the nutrient requirements are equations.

4.9 IV.9 Transportation Problems

Definition 4.39 (Flow and potential formulations). On a finite directed graph $G = (V, E)$, let b_i be the net demand at vertex i, and let $c_{ij} \ge 0$ be the cost of edge (i, j). The primal and dual are

$$\min_{x \ge 0} \sum_{(i,j)\in E} c_{ij}x_{ij} \quad \text{subject to} \quad \sum_{j:(j,i)\in E} x_{ji} - \sum_{j:(i,j)\in E} x_{ij} = b_i,$$

$$\max_{\lambda} \sum_i b_i\lambda_i \quad \text{subject to} \quad \lambda_j - \lambda_i \le c_{ij} \quad ((i,j) \in E).$$

Corollary 4.40 (Complementarity for optimal flows). If a feasible flow exists, primal and dual optimizers exist. Feasible x, λ are optimal if and only if

$$x_{ij} > 0 \Longrightarrow \lambda_j - \lambda_i = c_{ij}.$$

Proof. Nonnegative costs bound the primal objective below by 0, while $\lambda = 0$ is dual feasible. Thus both values are finite whenever the primal is feasible, and polyhedral strong duality gives attainment on both sides. ⊛

Remark 4.41 (Network simplex method). The positive support of an extreme feasible flow is a forest. When it is a spanning tree, first set potentials so that every positive tree edge satisfies equality. If a nontree edge violates its dual constraint, add it to create a cycle and adjust the flow around the cycle until an old edge becomes zero. This pivot lowers the objective and produces a new tree-supported feasible flow.

An initial flow can be constructed by adding an artificial vertex 0. Assume $\sum_i b_i = 0$. For each supply vertex, add the edge $i \to 0$ and send $-b_i$ along it; for each demand vertex, add $0 \to i$ and send b_i along it. Set every original edge flow to zero. Minimizing the total flow on these artificial edges gives a feasibility test: the original network has a feasible flow exactly when this minimum is zero. The original transportation cost can then be minimized over flows with all artificial edge flows equal to zero.

4.10 IV.10 Semidefinite Programming

Definition 4.42 (Primal and dual semidefinite programs). Under the pairing $\langle X, Y\rangle = \operatorname{tr}(XY)$ and the self-dual cone $\mathbb{S}^n_+$, consider

$$\gamma = \inf\{\langle C, X\rangle : \langle A_i, X\rangle = b_i, X \succeq 0\},$$

$$\beta = \sup\left\{\sum_i b_i y_i : \sum_i y_i A_i \preceq C\right\}.$$

Example 4.43 (A strictly feasible program without a primal optimizer). Let

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad b_1 = 1.$$

The primal problem minimizes X_{11} subject to $X \succeq 0$ and $2X_{12} = 1$. For every $a > 0$, the matrix

$$X_a = \begin{pmatrix} a & 1/2 \\ 1/2 & 1/(4a) \end{pmatrix}$$

is feasible, so the infimum is 0. A feasible matrix cannot have $X_{11} = 0$, since positive semidefiniteness would then force $X_{12} = 0$. Thus the infimum is not attained. There are strictly feasible matrices, such as $\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} \succ 0$. The dual constraint $C - yA_1 \succeq 0$ forces $y = 0$, so the dual optimum is attained with value 0.

Proposition 4.44 (A sufficient closedness condition). Suppose there are $\alpha_1, \dots, \alpha_m, \rho \in \mathbb{R}$ such that

$$B = \sum_i \alpha_i A_i + \rho C \succ 0.$$

If the primal is feasible, then $\gamma = \beta$; if $\gamma > -\infty$, a primal optimizer exists.

Proof. For $T(X) = (\langle A_1, X\rangle, \dots, \langle A_m, X\rangle, \langle C, X\rangle)$, if $X \succeq 0$ and $T(X) = 0$, then $\langle B, X\rangle = 0$. Since $B \succ 0$, this forces $X = 0$. The positive semidefinite cone has a compact base, so the compact-base lemma makes $T(\mathbb{S}^n_+)$ closed. ⊛

Proposition 4.45 (Strict feasibility). Suppose the constraint matrices $A_1, \dots, A_m$ are linearly independent. If there is a primal feasible $X_0 \succ 0$ and $\gamma > -\infty$, then $\gamma = \beta$ and a dual optimizer exists. If instead the primal is feasible and there is a dual feasible y_0 with

$$C - \sum_i (y_0)_i A_i \succ 0,$$

then $\gamma = \beta$ and a primal optimizer exists.

Proof. For the first assertion, linear independence makes the constraint map onto, so the interior criterion in Section IV.7 gives zero gap. Feasibility and the finite lower bound make $\gamma = \beta$ finite. Choose dual feasible $y^{(k)}$ with $b^{\mathsf{T}} y^{(k)} \to \beta$. If a subsequence had $\|y^{(k)}\| \to \infty$, compactness of the unit sphere would give, after a further subsequence,

$$z = \lim_k \frac{y^{(k)}}{\|y^{(k)}\|}, \qquad \|z\| = 1, \qquad -\sum_i z_i A_i \succeq 0, \qquad b^{\mathsf{T}} z = 0.$$

The last two conclusions follow by dividing the dual slack and objective by $\|y^{(k)}\|$. The positive semidefinite matrix $-\sum_i z_i A_i$ pairs to zero with $X_0 \succ 0$, so it is zero. Linear independence then forces $z = 0$, a contradiction. The optimizing sequence is therefore bounded. A convergent subsequence has a dual feasible limit because the positive semidefinite cone is closed, and its objective is β.

For the second assertion, the strictly positive definite slack is the positive definite combination in the preceding proposition. Weak duality bounds the primal value below by the finite value at y_0, and feasibility bounds it above. The closed-image result gives zero gap and primal attainment. ⊛

Corollary 4.46 (Semidefinite complementary slackness and rank). If X and y are optimal with zero gap and $S = C - \sum_i y_i A_i \succeq 0$, then

$$\langle X, S \rangle = 0, \qquad \operatorname{rank}(X) + \operatorname{rank}(S) \leq n.$$

Remark 4.47. Unlike polyhedral linear programs, semidefinite programs do not always attain their optimal values. A closed lifted image gives primal attainment when the primal infimum is finite. Primal strict feasibility gives dual attainment under the preceding assumptions, whereas dual strict feasibility gives primal attainment.

4.11 IV.11 Clique Number and Chromatic Number

Definition 4.48 (A semidefinite program for the Lovász theta number). For an undirected graph $G = (V, E)$ on at least one vertex, let J be the all-ones matrix and define

$$\vartheta(G) = \max\{\langle J, X\rangle : \operatorname{tr}(X) = 1, X \succeq 0,\ X_{ij} = 0 \text{ if } i \neq j \text{ and } (i,j) \notin E\}.$$

Theorem 4.49 (Sandwich theorem). If $\omega(G)$ is the clique number and $\chi(G)$ the chromatic number, then

$$\omega(G) \leq \vartheta(G) \leq \chi(G).$$

Proof. The indicator vector z of a clique of size k gives the feasible rank-one matrix $X = zz^{\mathsf{T}}/k$ with objective value k. Conversely, write a feasible X as a Gram matrix $X_{ij} = \langle v_i, v_j\rangle$. Vertices of the same color are nonadjacent, so their vectors are orthogonal. Sum by color and apply Cauchy–Schwarz to get $\|\sum_i v_i\|^2 \leq \chi(G)$. ⊛

Proposition 4.50 (Dual spectral form). For $i < j$, let A_{ij} be the symmetric matrix with 1 in positions (i,j) and (j,i) and zero elsewhere. The value $\vartheta(G)$ is also the least t for which

$$tI - J - \sum_{\substack{i<j \\ (i,j)\notin E}} y_{ij} A_{ij} \succeq 0.$$

The identity matrix is positive definite and lies in the span of the constraints, so there is no duality gap. The primal point $X = I/|V|$ is positive definite and feasible, which also gives dual attainment. Equivalently, minimize the largest eigenvalue of a symmetric matrix whose diagonal entries and entries corresponding to edges of G are 1, with its other entries free.

Remark 4.51. This semidefinite program places a computable convex quantity between two integer graph parameters. It replaces rank-one matrices generated by clique indicators with general Gram matrices.

4.12 IV.12 Linear Programming in L^∞

Definition 4.52 (Primal and dual problems). Given $f_i, g \in L^1[0,1]$, consider

$$\gamma = \inf\left\{ \int_0^1 gu : \int_0^1 f_i u = b_i,\ 0 \le u \le 1 \text{ almost everywhere} \right\}.$$

Here $u \in L^\infty[0,1]$. Its dual, with $\lambda \in \mathbb{R}^m$ and $p \in L^1[0,1]$, can be written as

$$\beta = \sup_{\lambda, p}\left\{ \sum_i \lambda_i b_i - \int_0^1 p : p \ge 0, -p + \sum_i \lambda_i f_i \le g \text{ almost everywhere} \right\}.$$

Lemma 4.53 (The cone used to homogenize the box). In $L^\infty[0,1] \oplus L^\infty[0,1]$, let

$$K = \{(u,v) : u, v \ge 0,\ u+v \text{ is constant almost everywhere}\}.$$

Under the integral pairing with $L^1[0,1] \oplus L^1[0,1]$, its dual cone is

$$K^* = \left\{ (a,h) : \int_0^1 \min\{a(t), h(t)\}\, dt \ge 0 \right\}.$$

Proof. If $u + v = \delta \ge 0$, then $au + hv \ge \delta \min\{a,h\}$ almost everywhere. Thus the displayed integral condition makes every pairing nonnegative. Conversely, choose $u = \mathbf{1}_{\{a \le h\}}$ and $v = 1 - u$. Their pairing with (a,h) equals $\int_0^1 \min\{a,h\}$, so nonnegativity of all pairings implies the condition. ⊛

To put the primal in canonical form, define

$$A(u,v) = \left(\int_0^1 f_1 u, \dots, \int_0^1 f_m u, \int_0^1 (u+v) \right), \qquad c = (g, 0).$$

The right-hand side is $(b_1, \dots, b_m, 1)$, and

$$A^*(\lambda, \rho) = \left(\sum_i \lambda_i f_i + \rho, \rho \right).$$

The dual-cone lemma turns $c - A^*(\lambda,\rho) \in K^*$ into

$$\rho \le \int_0^1 \min\left\{ g - \sum_i \lambda_i f_i, 0 \right\}.$$

For fixed λ, the largest possible ρ is the right-hand side. Writing its negative integrand as $p = \max\{0, \sum_i \lambda_i f_i - g\}$ gives the dual formulation above.

Proposition 4.54 (Zero gap and primal attainment). If the primal is feasible, then $\gamma = \beta$, and a primal optimizer exists.

Proof. The section $u + v = 1$ is a weak-star compact base of K. A nonnegative pair in the kernel of the lifted constraint map satisfies $\int_0^1 (u + v) = 0$, hence $u = v = 0$ almost everywhere. The compact-base lemma makes the lifted image closed. The closed-image theorem gives zero gap, and the bound $\gamma \geq -\int_0^1 |g| > -\infty$ gives primal attainment. ⊛

Proposition 4.55 (Maximum principle and endpoint values). For fixed λ, the best slack variable is

$$p(t) = \max\left\{0, -g(t) + \sum_i \lambda_i f_i(t)\right\}.$$

If the primal and dual solutions are optimal and the zero set of $g - \sum_i \lambda_i f_i$ has measure zero, then almost everywhere

$$u(t) = 1 \text{ on } \{p > 0\}, \qquad u(t) = 0 \text{ on } \{p = 0\}.$$

Remark 4.56. Section III.6 used extreme points to prove the existence of an endpoint-valued solution. Duality and complementary slackness now determine where the control takes each endpoint value.

4.13 IV.13 Linear Programming Form of Uniform Approximation

Definition 4.57. Given $g, f_1, \ldots, f_m \in C[0,1]$, let $L = \operatorname{lin}\{f_1, \ldots, f_m\}$. The primal problem is

$$\gamma = \min_{a \in \mathbb{R}^m} \left\| g - \sum_i a_i f_i \right\|_\infty,$$

equivalently, minimize an error radius r subject to two linear inequalities at every t. For a signed measure μ, write $\|\mu\| = |\mu|([0,1])$ for its total variation.

Theorem 4.58 (Measure dual of uniform approximation).

$$\gamma = \max\left\{\int g\, d\mu : \|\mu\| \leq 1,\ \int f_i\, d\mu = 0, i = 1, \ldots, m\right\},$$

and both primal and dual optimizers exist.

Proof. Represent the dual variables of the two error inequalities by positive measures μ_+, μ_-, and let $\mu = \mu_+ - \mu_-$. The total-mass constraint becomes the total-variation bound $\|\mu\| \le 1$, while the coefficient variables give the annihilation conditions. The positive-measure cone has a weak-star compact base, so the closed-image theorem gives zero gap. For primal attainment, any $f \in L$ with $\|g - f\|_\infty \le \|g\|_\infty$ satisfies $\|f\|_\infty \le 2\,\|g\|_\infty$. This closed ball in the finite-dimensional space L is compact, and the continuous error function attains its minimum there. ⊛

Corollary 4.59 (A finite contact criterion). For $f \in L$, put $r = \|g - f\|_\infty$. The function f is a best uniform approximation if and only if there are $k \le m+1$ distinct points t_j and nonzero coefficients s_j such that

$$\sum_j |s_j| = 1.$$

They must satisfy

$$\sum_j s_j f_i(t_j) = 0 \quad (i = 1, \dots, m), \qquad s_j\big(g(t_j) - f(t_j)\big) = |s_j|\, r \quad (j = 1, \dots, k).$$

Thus a positive coefficient is supported where the error is r, and a negative coefficient where the error is $-r$. The signed measure $\mu = \sum_j s_j \delta_{t_j}$ is a dual optimizer.

Proof. The extreme points of the total-variation unit ball are δ_t and $-\delta_t$. For $r > 0$, choose an extreme point of the dual optimal face. Section III.9 expresses it as a combination of at most $m+1$ signed Dirac measures. Equality in the dual bound $\int (g - f)\, d\mu \le r\,\|\mu\|$ gives the stated signs and contact conditions. If $r = 0$, choose $m+1$ distinct points and a nonzero solution of the m homogeneous annihilation equations, discard zero coefficients, and normalize the remaining coefficients. Conversely, the displayed conditions make μ dual feasible with $\int g\, d\mu = \int (g - f)\, d\mu = r$, so weak duality proves that f is optimal. ⊛

4.14 IV.14 Optimal Transport

Definition 4.60 (Kantorovich primal and dual problems). Let $X = [0,1]$; the same argument applies to compact metric spaces. Let μ_1, μ_2 be finite positive Borel measures of equal mass, and let $c \in C(X \times X)$. A coupling $\pi \geq 0$ is a positive measure on $X \times X$ with marginals μ_1, μ_2. The primal and dual problems are

$$\gamma = \min_{\pi} \left\{ \int c(x,y)\, d\pi : P_1\pi = \mu_1, P_2\pi = \mu_2 \right\},$$

$$\beta = \sup_{\phi,\psi} \left\{ \int \phi\, d\mu_1 + \int \psi\, d\mu_2 : \phi(x) + \psi(y) \leq c(x,y) \right\}.$$

The dual potentials ϕ, ψ belong to $C(X)$. For every $f \in C(X)$, the marginal maps satisfy

$$\int f\, d(P_1\pi) = \int f(x)\, d\pi(x,y), \qquad \int f\, d(P_2\pi) = \int f(y)\, d\pi(x,y).$$

Proposition 4.61 (Kantorovich duality and existence of an optimal coupling). For $X = [0,1]$, or any compact metric space, with continuous cost, feasibility implies $\gamma = \beta$, and an optimal coupling exists.

Proof. Equal total mass ensures feasibility: if this mass is $M > 0$, then $\mu_1 \otimes \mu_2 / M$ is a coupling; if $M = 0$, use the zero measure. The cone of positive measures has the weak-star compact base of probability measures. If a nonnegative measure has zero first marginal, it is zero, so the kernel of the lifted map meets the cone only at zero. Section IV.7 gives a closed image, zero gap, and primal attainment. ⊛

Remark 4.62 (Finite and continuous transportation models). With finite support this is the transportation linear program of Section IV.9, where ϕ and ψ are vertex potentials on the two sides. The continuous model replaces the finite transportation matrix by a measure on the product space.

5 Chapter V: Convex Bodies and Ellipsoids

This chapter studies metric properties of convex bodies. A maximum-volume inscribed ellipsoid gives a quadratic approximation to a convex body and leads to norm approximation and the ellipsoid method. Concentration of Gaussian and spherical measure is then used for low-rank positive semidefinite approximation, random projection, and sections of convex bodies.

5.1 V.1 Ellipsoids

Definition 5.1 (Ellipsoid). Let

$$B_2^d = \{x \in \mathbb{R}^d : \|x\|_2 \le 1\}.$$

If $T : \mathbb{R}^d \to \mathbb{R}^d$ is invertible and $a \in \mathbb{R}^d$, then

$$E = T(B_2^d) + a$$

is an *ellipsoid* with center a. Equivalently, for some $Q \succ 0$,

$$E = \{x : \langle Q(x-a), x-a\rangle \le 1\}, \qquad Q = (TT^*)^{-1}.$$

Consequently,

$$\operatorname{vol}(E) = |\det T| \operatorname{vol}(B_2^d) = \frac{\operatorname{vol}(B_2^d)}{\sqrt{\det Q}}.$$

Lemma 5.2 (Positive definite representation; Barvinok V.1.2). Every ellipsoid has a representation

$$E = S(B_2^d) + a, \qquad S \succ 0.$$

For a fixed center, the positive definite matrix S is unique.

Proof. Take the polar decomposition $T = SU$, where $S = (TT^*)^{1/2} \succ 0$ and U is orthogonal. Since $U(B_2^d) = B_2^d$, we have $T(B_2^d) = S(B_2^d)$. Uniqueness follows from uniqueness of the positive definite square root. ⊛

Lemma 5.3 (Determinant inequality; Barvinok V.1.3). If $X, Y \succ 0$, then

$$\det\left(\frac{X+Y}{2}\right) \geq \sqrt{\det X \det Y},$$

with equality exactly when $X = Y$. More generally, $X \mapsto \log\det X$ is strictly concave on the positive definite cone.

Main argument. A congruence transformation reduces X to I, after which Y can be orthogonally diagonalized. The assertion becomes the coordinatewise inequality

$$(1+\lambda_i)/2 \geq \sqrt{\lambda_i}.$$

Equality throughout forces every $\lambda_i = 1$. ⊛

Proposition 5.4 (Polarity and volume). If

$$E = \{x : \langle Qx, x\rangle \leq 1\}, \qquad Q \succ 0,$$

then

$$E^\circ = \{y : \langle Q^{-1}y, y\rangle \leq 1\}, \qquad \operatorname{vol}(E)\operatorname{vol}(E^\circ) = \operatorname{vol}(B_2^d)^2.$$

Remark 5.5 (Principal axes and volume). A Euclidean ball depends on the chosen inner product, whereas the family of ellipsoids is preserved by invertible affine transformations. The volume formula shows that maximizing ellipsoid volume is equivalent to maximizing $\log\det S$. This also explains the natural role of $-\log\det X$ in the geometry of the positive definite cone.

5.2 V.2 Ellipsoids of Maximum Volume in Convex Bodies

Definition 5.6 (Convex body). A compact convex set $K \subset \mathbb{R}^d$ with nonempty interior is a *convex body*. It is centrally symmetric about the origin when $K = -K$.

Theorem 5.7 (Existence and uniqueness of the John ellipsoid; Barvinok V.2.2). Every convex body $K \subset \mathbb{R}^d$ contains a unique ellipsoid E_K of maximum volume.

Main argument. Consider

$$\mathcal{A} = \{(S, a) : S \succeq 0,\ S(B_2^d) + a \subset K\}.$$

Compactness of K makes $\mathcal{A}$ closed and bounded, so $\det S$ attains its maximum. Nonempty interior makes the maximum positive. If (S_1, a_1) and (S_2, a_2) are both optimal, convexity makes their midpoint feasible and strict concavity of $\log\det$ gives $S_1 = S_2$. If the centers differed, normalize the two ellipsoids to unit balls centered at $\pm a$. Their convex hull contains an ellipsoid elongated in the a-direction with larger volume, a contradiction. ⊛

Theorem 5.8 (John's inclusion for a general convex body; Barvinok V.2.4). Let $E \subset K \subset \mathbb{R}^d$ be a maximum-volume inscribed ellipsoid and translate its center to the origin. Then

$$E \subset K \subset dE.$$

The constant d is attained by a simplex and is best possible.

Proof. Normalize $E = B_2^d$. If $x = \rho e_1 \in K$ with $\rho > d$, then $C = \operatorname{conv}(B_2^d \cup \{x\}) \subset K$. Inside C, consider

$$E_\tau = \left\{ \xi : \frac{(\xi_1 - \tau)^2}{(1+\tau)^2} + \frac{\xi_2^2 + \cdots + \xi_d^2}{\beta_\tau^2} \le 1 \right\},$$

where

$$\beta_\tau^2 = \frac{(\rho - \tau)^2 - (1+\tau)^2}{\rho^2 - 1}.$$

A tangent-line calculation in two-dimensional sections gives $E_\tau \subset C$, while

$$\log \frac{\operatorname{vol} E_\tau}{\operatorname{vol} B_2^d} = \left(1 - \frac{d-1}{\rho - 1}\right)\tau + O(\tau^2) > 0$$

for sufficiently small $\tau > 0$, contradicting maximality. ⊛

Theorem 5.9 (John's inclusion for centrally symmetric bodies; Barvinok V.2.5). If $K = -K$, its maximum-volume inscribed ellipsoid E is centered at the origin and

$$E \subset K \subset \sqrt{d}\, E.$$

The cross-polytope attains the constant $\sqrt{d}$.

Main argument. Uniqueness and the reflection $x \mapsto -x$ force the center to be zero. Normalize $E = B_2^d$. If $\pm\rho e_1 \in K$, place in $\operatorname{conv}(B_2^d \cup \{\pm\rho e_1\})$ a centered ellipsoid with semiaxes $\alpha, \beta, \ldots, \beta$. Tangency gives

$$\alpha^2 = \rho^2 - (\rho^2 - 1)\beta^2.$$

Writing $\beta = 1 - \delta$, the first-order change in log volume is

$$(\rho^2 - d)\delta + O(\delta^2).$$

If $\rho > \sqrt{d}$, the volume increases, contradicting maximality. ⊛

Proposition 5.10 (Contact points of the John ellipsoid). Under the normalization $B_2^d \subset K$, the unit ball is the maximum-volume inscribed ellipsoid if and only if there are contact points $u_i \in \partial K \cap \mathbb{S}^{d-1}$ and $\lambda_i > 0$ such that

$$\sum_i \lambda_i = 1, \qquad \sum_i \lambda_i u_i = 0, \qquad \sum_i \lambda_i (u_i \otimes u_i) = \frac{1}{d} I.$$

The number of contact points can be chosen to satisfy $m \le (d^2 + 3d)/2$.

Example 5.11 (Cube, cross-polytope, and simplex). For the cube $K = [-1,1]^d$, the John ellipsoid is B_2^d. For the cross-polytope $K = \{x : \|x\|_1 \le 1\}$, it is $d^{-1/2}B_2^d$. The John ball of the standard d-simplex has radius $1/\sqrt{d(d+1)}$, measured in its affine hull.

Remark 5.12 (Inscribed and circumscribed ellipsoids). The John ellipsoid is the maximum-volume inscribed ellipsoid; the Loewner ellipsoid is the minimum-volume circumscribed ellipsoid. Polarity relates them. When the center is not the origin, translate before taking the polar.

5.3 V.3 Norms and Their Approximation

Definition 5.13 (Minkowski functional). If $K \subset \mathbb{R}^d$ is a convex body with $K = -K$, so that the origin lies in its interior, define

$$p_K(x) = \inf\{\lambda > 0 : x \in \lambda K\}.$$

This is the *Minkowski functional* of K.

Lemma 5.14 (Correspondence between norms and convex bodies; Barvinok V.3.2). If p is a norm on $\mathbb{R}^d$, then $K_p = \{x : p(x) \le 1\}$ is a centrally symmetric convex body. Conversely, p_K is a norm and $K = K_{p_K}$.

Proof. The triangle inequality makes K_p convex. In finite dimensions, norms are continuous and equivalent to the Euclidean norm, so K_p is compact with nonempty interior. Conversely, homogeneity and positive definiteness are immediate. If $x \in \lambda K$ and $y \in \mu K$, then convexity gives $x + y \in (\lambda + \mu)K$. Taking infima proves the triangle inequality. ⊛

Corollary 5.15 (Quadratic approximation). For every norm p, some $Q \succ 0$ satisfies

$$\sqrt{\langle Qx, x\rangle} \le p(x) \le \sqrt{d}\,\sqrt{\langle Qx, x\rangle} \qquad (x \in \mathbb{R}^d).$$

Theorem 5.16 (Polynomial approximation; Barvinok V.3.4). Given a norm $p : \mathbb{R}^d \to \mathbb{R}$ and an integer $n \ge 1$, there is a homogeneous polynomial q of degree $2n$ such that

$$q = \sum_j q_j^2,$$

where each q_j is homogeneous of degree n, and

$$q(x)^{1/(2n)} \le p(x) \le \binom{n+d-1}{n}^{1/(2n)} q(x)^{1/(2n)}.$$

For $n = d$, the approximation factor is at most 2.

Proof. The bipolar theorem gives

$$p(x) = \max_{y \in K_p^\circ} |\langle x, y \rangle| .$$

After lifting $x \mapsto x^{\otimes n}$, the quantity $p(x)^n$ is the support function of

$$A = \operatorname{conv}\{\pm y^{\otimes n} : y \in K_p^\circ\}.$$

This body lies in a symmetric tensor space of dimension at most $D = \binom{n+d-1}{n}$. Apply the centrally symmetric John theorem in the affine hull of A to obtain $E \subset A \subset \sqrt{D}\, E$. The square of the support function of an ellipsoid is a quadratic form in $x^{\otimes n}$, hence a sum of squares of degree $2n$. The inclusions give the bounds. ⊛

Example 5.17 (Explicit approximation of ℓ_∞). Let

$$p(x) = \max_i |x_i| , \qquad q(x) = \sum_{i=1}^{d} x_i^{2n}.$$

Then

$$p(x) \le q(x)^{1/(2n)} \le d^{1/(2n)} p(x).$$

As the degree increases, the sum of even powers approximates the largest coordinate magnitude with a factor tending to 1.

5.4 V.4 The Ellipsoid Method

Lemma 5.18 (Ellipsoid covering of a half-ball; Barvinok V.4.2). Let $B^+ = \{x \in B_2^d : x_d \ge 0\}$. There is an ellipsoid $E \supset B^+$ such that

$$\frac{\operatorname{vol}(E)}{\operatorname{vol}(B_2^d)} \le \exp\left(-\frac{1}{2(d+1)}\right).$$

By affine transformation, the same conclusion holds for a half-ellipsoid cut through its center.

Proof. For $d = 1$, the interval $[0, 1]$ covers the upper half of $[-1, 1]$ and has half its length. Assume now that $d \geq 2$. Translate the center by $1/(d+1)$ along e_d, and scale the squared transverse and longitudinal semiaxes by

$$\frac{d^2}{d^2-1}, \qquad \frac{d^2}{(d+1)^2}.$$

Substitution into $x_1^2 + \cdots + x_{d-1}^2 \leq 1 - x_d^2$ verifies containment. A determinant calculation and $1 + t \leq e^t$ give the volume bound. ⊛

Proposition 5.19 (Ellipsoid feasibility algorithm). Let the feasible set of strict linear inequalities

$$P = \{x : \langle a_i, x\rangle < \beta_i, i = 1, \ldots, m\}$$

have the following property for known numbers $R > r > 0$: if it is nonempty, some x_0 satisfies $B(x_0, r) \subset P \cap B(0, R)$. An inequality with $a_i = 0$ is either satisfied everywhere or proves infeasibility immediately; discard the former before making cuts. Start with $E_0 = B(0, R)$. If the center is infeasible, retain the half-ellipsoid containing $P \cap E_k$, as determined by a violated inequality, and replace it by the covering ellipsoid above. After at most

$$1 + \left\lfloor 2d(d+1) \log \frac{R}{r} \right\rfloor$$

cuts, the algorithm finds a feasible point or proves $P = \varnothing$.

Proof. Every step preserves $B(x_0, r) \subset E_k$, while

$$\operatorname{vol} E_k \leq e^{-k/(2(d+1))} \operatorname{vol} E_0.$$

Once the right side is smaller than $\operatorname{vol} B(x_0, r)$, nonemptiness is impossible. Comparing the two ball volumes and solving for k gives the bound. ⊛

Remark 5.20 (Iteration count). The ellipsoid method needs only a separation procedure that returns either feasibility or a separating hyperplane. The containing ellipsoid loses a fixed fraction of volume at each step, so the full constraint description need not be explicitly enumerated.

5.5 V.5 Gaussian Measure on Euclidean Space

Definition 5.21 (Standard Gaussian measure). For a Borel set $A \subset \mathbb{R}^n$, define

$$\gamma_n(A) = (2\pi)^{-n/2} \int_A e^{-\|x\|_2^2/2}\,\mathrm{d}x.$$

If $g \sim N(0, I_n)$, then $\gamma_n(A) = \mathbb{P}\{g \in A\}$.

Lemma 5.22 (One-dimensional Gaussian tail; Barvinok V.5.2). If $Z \sim N(0,1)$ and $t \geq 0$, then

$$\mathbb{P}(Z \geq t) = \mathbb{P}(Z \leq -t) \leq e^{-t^2/2}.$$

Proof. For $\lambda \geq 0$, Markov's inequality and $\mathbb{E}e^{\lambda Z} = e^{\lambda^2/2}$ give $\mathbb{P}(Z \geq t) \leq e^{\lambda^2/2 - \lambda t}$. Take $\lambda = t$. ⊛

Proposition 5.23 (Concentration of the Gaussian radius; Barvinok V.5.4). If $g \sim N(0, I_n)$, then for $\delta \geq 0$,

$$\mathbb{P}\{\|g\|_2^2 \geq n + \delta\} \leq \left(\frac{n}{n+\delta}\right)^{-n/2} e^{-\delta/2}.$$

For $0 < \delta < n$,

$$\mathbb{P}\{\|g\|_2^2 \leq n - \delta\} \leq \left(\frac{n}{n-\delta}\right)^{-n/2} e^{\delta/2}.$$

At $\delta = n$, the lower-tail probability and the limiting bound are both zero.

Chernoff method. Use

$$\mathbb{E}e^{\lambda\|g\|^2/2} = (1-\lambda)^{-n/2}, \qquad 0 \leq \lambda < 1,$$

and

$$\mathbb{E}e^{-\lambda\|g\|^2/2} = (1+\lambda)^{-n/2}.$$

Optimize with $\lambda = \delta/(n+\delta)$ for the upper tail and $\lambda = \delta/(n-\delta)$ for the lower tail. ⊛

Corollary 5.24 (Relative deviations; Barvinok V.5.5). For $0 < \varepsilon < 1$,

$$\mathbb{P}\left\{\|g\|_2^2 \geq \frac{n}{1-\varepsilon}\right\} \leq e^{-\varepsilon^2 n/4}, \qquad \mathbb{P}\{\|g\|_2^2 \leq (1-\varepsilon)n\} \leq e^{-\varepsilon^2 n/4}.$$

Proof. In the upper-tail bound take $\delta = \varepsilon n/(1-\varepsilon)$. Its logarithm is $-\frac{n}{2}[\log(1-\varepsilon)+\varepsilon/(1-\varepsilon)]$, and the expression in brackets is at least $\varepsilon^2/2$ by its power-series expansion. For the lower tail take $\delta = \varepsilon n$ and use $\log(1-\varepsilon)+\varepsilon \le -\varepsilon^2/2$. ⊛

Proposition 5.25 (Concentration of positive semidefinite quadratic forms; Barvinok V.5.6). Let $Q \succeq 0$, let $q(x) = \langle Qx, x\rangle$, and $\|Q\|_F = \sqrt{\operatorname{tr}(Q^2)}$. For $g \sim N(0, I_n)$ and every $\tau \ge 0$:

1. $$\mathbb{P}\{q(g) < \operatorname{tr} Q - \tau \|Q\|_F\} \le e^{-\tau^2/4};$$

2. if $\tau\lambda_{\max}(Q) \le \|Q\|_F$, then
$$\mathbb{P}\{q(g) > \operatorname{tr} Q + \tau \|Q\|_F\} \le e^{-\tau^2/8}.$$

Proof. If $Q = 0$, both events are empty. Otherwise, let $F = \|Q\|_F > 0$ and diagonalize Q. For $s \ge 0$,
$$\mathbb{E}e^{-sq(g)/2} = \det(I + sQ)^{-1/2}.$$
Since $\log(1+t) \ge t - t^2/2$ for $t \ge 0$, Markov's inequality gives
$$\mathbb{P}\{q(g) < \operatorname{tr} Q - \tau F\} \le \exp\left(\frac{s^2F^2}{4} - \frac{s\tau F}{2}\right).$$
Take $s = \tau/F$ to obtain the lower-tail bound.

For $s\lambda_{\max}(Q) < 1$,
$$\mathbb{E}e^{sq(g)/2} = \det(I - sQ)^{-1/2}.$$
The inequality $\log(1-t) \ge -t - t^2$ holds for $0 \le t \le 1/2$. Take $s = \tau/(2F)$; the assumed condition gives $s\lambda_{\max}(Q) \le 1/2$. Another application of Markov's inequality therefore yields
$$\mathbb{P}\{q(g) > \operatorname{tr} Q + \tau F\} \le \exp\left(\frac{s^2F^2}{2} - \frac{s\tau F}{2}\right) = e^{-\tau^2/8}.$$
⊛

Remark 5.26 (Spectrum of a quadratic form). After diagonalization, a positive semidefinite quadratic form is a weighted sum of independent squares of Gaussian variables. Exponential moments reduce the weights to the three scales $\operatorname{tr} Q$, $\|Q\|_F$, and $\lambda_{\max}(Q)$.

5.6 V.6 Low-Rank Matrix Approximation

Theorem 5.27 (Low-rank approximate positive semidefinite solution; Barvinok V.6.1). Let $k \geq 1$, let $A_1, \dots, A_k \succeq 0$, let $\alpha_i \geq 0$, and suppose some $X \succeq 0$ satisfies $\langle A_i, X \rangle = \alpha_i$ for all i. Given $0 < \varepsilon < 1$, if an integer m satisfies

$$m \geq \frac{8}{\varepsilon^2} \log(4k),$$

there is $X_0 \succeq 0$ with $\operatorname{rank} X_0 \leq m$ and

$$(1-\varepsilon)\alpha_i \leq \langle A_i, X_0 \rangle \leq (1+\varepsilon)\alpha_i \qquad (i = 1, \dots, k).$$

Gaussian sampling. Write $X = TT^*$ and set $B_i = T^* A_i T$, reducing the reference solution to I. Take independent $g_1, \dots, g_m \sim N(0, I)$ and let

$$Y_0 = \frac{1}{m} \sum_{j=1}^{m} g_j \otimes g_j, \qquad X_0 = T Y_0 T^*.$$

Then $\operatorname{rank} X_0 \leq m$ and $\operatorname{tr} B_i = \alpha_i$. Let Q_i consist of m diagonal blocks equal to B_i, and concatenate the g_j into a single standard Gaussian vector g. Thus

$$m \langle A_i, X_0 \rangle = \langle Q_i g, g \rangle, \qquad \operatorname{tr} Q_i = m\alpha_i, \qquad \|Q_i\|_F = \sqrt{m}\, \|B_i\|_F .$$

If $\alpha_i = 0$, positive semidefiniteness gives $B_i = 0$, so this constraint holds exactly. Otherwise, $\|B_i\|_F \leq \operatorname{tr} B_i = \alpha_i$ and $\lambda_{\max}(B_i) \leq \|B_i\|_F$. Apply the quadratic-form bounds with $\tau = \varepsilon\sqrt{m}$. The spectral condition holds, and a deviation by more than $\varepsilon m \alpha_i$ is a deviation by more than $\tau \|Q_i\|_F$. The total failure probability is therefore at most

$$e^{-\varepsilon^2 m/4} + e^{-\varepsilon^2 m/8} \leq 2e^{-\varepsilon^2 m/8} \leq \frac{1}{2k}.$$

The union bound gives simultaneous success with probability at least $1/2$. ⊛

Corollary 5.28 (Graph realization with distortion; Barvinok V.6.3). If a weighted graph G has $k \geq 1$ edges and is realizable in some Euclidean space, then for $0 < \varepsilon < 1$ and any integer $m \geq 8\varepsilon^{-2} \log(4k)$ its vertices can be placed in $\mathbb{R}^m$ so that every edge satisfies

$$(1-\varepsilon)\rho_{ij}^2 \leq \|v_i - v_j\|_2^2 \leq (1+\varepsilon)\rho_{ij}^2.$$

Proof. Use the Gram matrix X of the original realization. For an edge $\{i, j\}$,

$$\rho_{ij}^2 = x_{ii} - 2x_{ij} + x_{jj} = \langle A_{ij}, X \rangle, \qquad A_{ij} = (e_i - e_j)(e_i - e_j)^* \succeq 0.$$

Apply the theorem to these k measurements. The bound $\operatorname{rank} X_0 \le m$ is equivalent to a Gram realization by vectors in $\mathbb{R}^m$. ⊛

Remark 5.29 (Gram matrices). A configuration of vectors determines a positive semidefinite Gram matrix. Gaussian sampling replaces a high-rank Gram matrix with a low-rank one. Rank is the dimension of the realization, and the selected linear measurements record the geometry that must be preserved.

5.7 V.7 Measure and Metric on the Unit Sphere

Definition 5.30 (Spherical Haar measure). Let $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$, and let ν_{n-1} be the rotation-invariant probability measure. For a Borel set $A \subset \mathbb{S}^{n-1}$, let $\widehat{A} = \{rx : r \ge 0, x \in A\}$. Then

$$\nu_{n-1}(A) = \gamma_n(\widehat{A}).$$

Lemma 5.31 (Spherical cap bound; Barvinok V.7.1). For $y \in \mathbb{S}^{n-1}$ and $\varepsilon > 0$,

$$\nu_{n-1}\{x : \langle x, y \rangle \ge \varepsilon\} \le 2e^{-\varepsilon^2 n/16}.$$

Thus the inner product of two independent uniform spherical vectors typically has order $n^{-1/2}$.

Proof. By rotation invariance take $y = e_1$. Lifting the event to its Gaussian cone gives $g_1 \ge \varepsilon \|g\|$. Apply the one-dimensional upper-tail bound to g_1, the Gaussian-radius lower-tail bound to $\|g\|^2$, and then the union bound. ⊛

Proposition 5.32 (Scale of a random projection). Let $1 \le k \le n$ and $0 < \varepsilon < 1$. If $L \subset \mathbb{R}^n$ is a fixed k-dimensional subspace and x is uniform on $\mathbb{S}^{n-1}$, then

$$\mathbb{P}\left\{(1-\varepsilon)\sqrt{\frac{k}{n}} \le \|\operatorname{proj}_L x\|_2 \le (1-\varepsilon)^{-1}\sqrt{\frac{k}{n}}\right\} \ge 1 - 4e^{-\varepsilon^2 k/4}.$$

The same bound holds for fixed $x \in \mathbb{S}^{n-1}$ and a uniformly random k-dimensional subspace. Thus the relative error is small with high probability when k is large.

Proof. Write $x = g/\|g\|_2$ with $g \sim N(0, I_n)$, and put $U = \|\mathrm{proj}_L\, g\|_2^2$ and $V = \|g\|_2^2$. By rotation invariance, U is the sum of k independent squared standard Gaussian variables, while V is the corresponding sum of n variables. The bounds for relative deviations and a union bound show that

$$(1-\varepsilon)k \le U \le \frac{k}{1-\varepsilon}, \qquad (1-\varepsilon)n \le V \le \frac{n}{1-\varepsilon}$$

hold together with probability at least $1 - 2e^{-\varepsilon^2 k/4} - 2e^{-\varepsilon^2 n/4} \ge 1 - 4e^{-\varepsilon^2 k/4}$. Since $\|\mathrm{proj}_L\, x\|_2^2 = U/V$, taking square roots proves the claim. Rotation invariance exchanges a fixed subspace and a random vector with a random subspace and a fixed vector. ⊛

Definition 5.33 (Spherical distance and neighborhoods). For $n \ge 2$ and $x, y \in \mathbb{S}^{n-1}$, define

$$\mathrm{dist}(x, y) = \arccos \langle x, y \rangle .$$

This is the length of the shorter great-circle arc joining the points. For a closed set $A \subset \mathbb{S}^{n-1}$ and $\rho \ge 0$, set

$$A_\rho = \{x : \mathrm{dist}(x, y) \le \rho \text{ for some } y \in A\}.$$

The spherical cap with center y and radius ρ is $C(y, \rho) = \{x : \mathrm{dist}(x, y) \le \rho\}$; it is the whole sphere when $\rho \ge \pi$.

Theorem 5.34 (Spherical isoperimetric inequality; Barvinok V.7.1, Problem 4). Let $n \ge 2$, let $A \subset \mathbb{S}^{n-1}$ be closed, and let $C(y, r)$ be a spherical cap with $0 < \nu_{n-1}(A) = \nu_{n-1}(C(y, r)) < 1$. Then for every $\rho > 0$,

$$\nu_{n-1}(A_\rho) \ge \nu_{n-1}(C(y, r + \rho)).$$

Thus spherical caps have the smallest neighborhoods among sets of the same measure.

Corollary 5.35 (Concentration around a set of measure one half). If $A \subset \mathbb{S}^{n-1}$ is closed and $\nu_{n-1}(A) \ge 1/2$, then for $\rho > 0$,

$$\nu_{n-1}(A_\rho) \ge 1 - 2\exp\left(-\frac{n\rho^2}{4\pi^2}\right).$$

Proof. If $\nu_{n-1}(A) = 1$, the claim is immediate. Otherwise, a cap with the measure of A has radius at least $\pi/2$. For $0 < \rho \le \pi/2$, the complement of its ρ-neighborhood is contained in a cap defined by $\langle x, -y\rangle \ge \sin\rho$. The spherical cap bound and $\sin\rho \ge 2\rho/\pi$ therefore bound the complementary measure by $2e^{-n\rho^2/(4\pi^2)}$. For $\rho > \pi/2$, the enlarged cap is the whole sphere. Apply the isoperimetric inequality in both cases. ⊛

Corollary 5.36 (Concentration of Lipschitz functions; Barvinok V.7.1, Problem 7). Let $n \ge 2$ and let $f : \mathbb{S}^{n-1} \to \mathbb{R}$ satisfy

$$|f(x) - f(y)| \le \operatorname{dist}(x, y) \qquad (x, y \in \mathbb{S}^{n-1}).$$

If α is a median of f, so that each of $\{f \le \alpha\}$ and $\{f \ge \alpha\}$ has measure at least $1/2$, then for every $\rho > 0$,

$$\nu_{n-1}\{x : |f(x) - \alpha| > \rho\} \le 4\exp\left(-\frac{n\rho^2}{4\pi^2}\right).$$

Proof. The Lipschitz condition makes both median sets closed. The ρ-neighborhood of $\{f \le \alpha\}$ lies in $\{f \le \alpha+\rho\}$, while that of $\{f \ge \alpha\}$ lies in $\{f \ge \alpha-\rho\}$. Apply the preceding concentration bound to the two median sets and add the resulting bounds for the upper and lower tails. ⊛

Theorem 5.37 (Qualitative Dvoretzky theorem). For every positive integer k and $0 < \varepsilon < 1$, there is $m = m(k, \varepsilon)$ such that every convex body $K \subset \mathbb{R}^n$ with $K = -K$ and $n \ge m$ has a k-dimensional subspace L and a centered ellipsoid $E \subset L$ satisfying

$$(1 - \varepsilon)E \subset K \cap L \subset (1 + \varepsilon)E.$$

One may take m of order

$$\exp(O(\varepsilon^{-2}k)).$$

Remark 5.38 (John and Dvoretzky). John's theorem approximates an entire d-dimensional centrally symmetric body by an ellipsoid within factor $\sqrt{d}$. Dvoretzky's theorem finds a fixed-dimensional section of a high-dimensional body with arbitrarily small relative error from an ellipsoid. Concentration of Lipschitz functions on the sphere supplies the measure estimate used in this existence result.

6 Chapter VI: Faces of Polytopes

This chapter studies face lattices and face numbers of polytopes. Polarity pairs vertices with facets, Euler characteristic gives alternating relations among face numbers, and cyclic polytopes attain the largest possible face numbers for a fixed number of vertices. The duality between simple and simplicial polytopes leads to f-vectors, h-vectors, and the Dehn–Sommerville equations.

6.1 VI.1 Polytopes and Polarity

Lemma 6.1 (A face is generated by its vertices; Barvinok VI.1.1). If $P = \operatorname{conv}\{v_1, \dots, v_m\}$ and F is a face of P, then

$$F = \operatorname{conv}\{v_i : v_i \in F\}.$$

Thus every face of a polytope is a polytope, and a polytope has finitely many faces.

Proof. Choose c, α with $\langle c, x\rangle \le \alpha$ on P and $F = \{x \in P : \langle c, x\rangle = \alpha\}$. If $x = \sum_i \lambda_i v_i \in F$, every vertex satisfying $\langle c, v_i\rangle < \alpha$ must have $\lambda_i = 0$; otherwise the weighted sum would be strictly less than α. ⊛

Corollary 6.2 (A face of a face). If G is a face of F and F is a face of a polytope P, then G is a face of P.

Proof. Let f expose F in P and let g expose G in F. For sufficiently small $\varepsilon > 0$, the functional $f + \varepsilon g$ first separates F from the other finitely many vertices of P, then separates G within F. ⊛

Theorem 6.3 (Face correspondence under polarity; Barvinok VI.1.3). Let $P \subset \mathbb{R}^d$ be a d-polytope containing 0 in its interior and set $Q = P^\circ$. For a face F of P, define

$$\widehat{F} = \{x \in Q : \langle x, y\rangle = 1 \text{ for every } y \in F\},$$

with $\widehat{\varnothing} = Q$, $\widehat{P} = \varnothing$, and $\dim \varnothing = -1$. Then:

1. $\widehat{F}$ is a face of Q, and $\dim F + \dim \widehat{F} = d - 1$;
2. $F \subset G$ implies $\widehat{G} \subset \widehat{F}$;
3. every face H of Q has a unique representation $H = \widehat{F}$.

Thus polarity gives a bijection between the faces of P and P° that reverses inclusion.

Proof. The empty and whole faces satisfy the assertions by the stated conventions. Let F be a nonempty proper face, and average its vertices to obtain $v \in \operatorname{relint} F$. Since every vertex u of F satisfies $\langle x, u\rangle \le 1$ for $x \in Q$, equality $\langle x, v\rangle = 1$ holds exactly when all these vertex inequalities are equalities. Hence $\widehat{F} = \{x \in Q : \langle x, v\rangle = 1\}$ is a face of Q.

Choose a functional exposing F, and normalize its vector c so that $\langle c, u\rangle = 1$ for the vertices in F and $\langle c, u\rangle < 1$ for all other vertices of P. This normalization is possible because $0 \in \operatorname{int} P$. If $\dim F = k$, then $\dim \operatorname{span} F = k + 1$, and the equality constraints defining $\widehat{F}$ specify the affine space $c + (\operatorname{span} F)^{\perp}$, of dimension $d - k - 1$. All remaining vertex inequalities are strict at c. A sufficiently small neighborhood of c within this affine space therefore lies in $\widehat{F}$, proving the dimension formula.

If $F \subset G$, the equalities for G include those for F, so $\widehat{G} \subset \widehat{F}$. Applying polarity twice returns F: any $y \in P$ that pairs to 1 with every $x \in \widehat{F}$ must pair to 1 with the exposing vector c, and hence lies in F. The converse follows from the definition. The same argument for Q, together with $Q^{\circ} = P$, proves that the correspondence is a bijection. ⊛

Definition 6.4 (Vertices, edges, facets, and ridges). A zero-dimensional face is a vertex and a one-dimensional face is an edge. In a d-polytope, a $(d-1)$-face is a facet and a $(d-2)$-face is a ridge. Two vertices are adjacent when they span an edge.

Corollary 6.5 (Diamond property). If $G \subset F$ are faces with $\dim F - \dim G = 2$, there are exactly two intermediate faces H_1, H_2 satisfying $G \subsetneq H_i \subsetneq F$. In particular, every ridge lies in exactly two facets.

Lemma 6.6 (Irredundant facet description; Barvinok VI.1.5). Every d-polytope can be written as

$$P = \{x \in \mathbb{R}^d : f_i(x) \le \alpha_i, i = 1, \dots, m\},$$

where the sets $F_i = \{x \in P : f_i(x) = \alpha_i\}$ are exactly its facets.

Remark 6.7 (Faces in this chapter). Every nonempty proper face here is exposed by a supporting hyperplane. General convex sets can have nonexposed extreme points, but every face of a finitely generated polytope is exposed by a linear functional.

6.2 VI.2 Faces of Permutation Polytopes

Definition 6.8 (Permutation polytope). For $a = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, define

$$P(a) = \operatorname{conv}\{\sigma(a) : \sigma \in S_n\}.$$

When the coordinates are distinct, reorder them so that $\alpha_1 > \alpha_2 > \cdots > \alpha_n$.

Lemma 6.9 (Rearrangement exchange; Barvinok VI.2.1). If $x_i > x_j$ and $y_i < y_j$, and y' is obtained by exchanging y_i, y_j, then

$$\langle x, y' \rangle - \langle x, y \rangle = (x_i - x_j)(y_j - y_i) > 0.$$

Proposition 6.10 (Classification of nonempty faces; Barvinok VI.2.2). Assume $\alpha_1 > \cdots > \alpha_n$. Take an ordered partition $\mathcal{S} = (S_1, \dots, S_k)$ of $\{1, \dots, n\}$, with $s_j = |S_j|$. Place the largest s_1 values α_i in coordinates S_1, the next s_2 in S_2, and so on, allowing arbitrary permutations within each block. Let $F_{\mathcal{S}}$ be the convex hull of these vertices. Then

$$F_{\mathcal{S}} \text{ is a face of } P(a), \qquad \dim F_{\mathcal{S}} = n - k,$$

and every nonempty face of $P(a)$ arises in this way.

Proof. First classify the faces containing a. If $c = (\gamma_1, \dots, \gamma_n)$ is maximized at a, the exchange lemma gives $\gamma_1 \geq \cdots \geq \gamma_n$. The maximal constant blocks of the γ_i allow exactly the permutations within the corresponding ordered partition. Moreover,

$$F_{\mathcal{S}} \cong P(a^{(1)}) \times \cdots \times P(a^{(k)}), \qquad \dim F_{\mathcal{S}} = \sum_{j=1}^{k} (s_j - 1) = n - k.$$

The action of S_n moves a general face to one containing a. ⊛

Corollary 6.11 (Edges and facets). Two vertices are adjacent exactly when one is obtained from the other by exchanging two values α_r, α_{r+1} adjacent in numerical order. Facets correspond to ordered partitions into two blocks.

Theorem 6.12 (Rado's theorem; Barvinok VI.2.3). Let $\alpha_1 \geq \cdots \geq \alpha_n$, and arrange the coordinates of b as $\beta_1 \geq \cdots \geq \beta_n$. Then $b \in P(a)$ if and only if

$$\sum_{i=1}^{k} \beta_i \leq \sum_{i=1}^{k} \alpha_i \quad (k = 1, \dots, n-1),$$

and

$$\sum_{i=1}^{n} \beta_i = \sum_{i=1}^{n} \alpha_i.$$

Equivalently, b is majorized by a.

Proof. The facet description from the classification is

$$\sum_{i \in S} x_i \leq \sum_{i=1}^{|S|} \alpha_i$$

for every nonempty proper subset S, together with equality of total sums. For $|S| = k$, the left side is largest on the k greatest coordinates of b, giving the stated partial sums. If a has repeated coordinates, approximate it by strictly decreasing vectors and pass to the limit. ⊛

Remark 6.13 (Majorization). Together with Birkhoff–von Neumann, Rado's theorem gives three equivalent descriptions: partial-sum inequalities, convex combinations of permutations, and $b = Da$ for a doubly stochastic matrix D.

6.3 VI.3 The Euler–Poincaré Formula

Lemma 6.14 (Euler characteristic of boundary and interior; Barvinok VI.3.1). If $P \subset \mathbb{R}^d$ is a d-polytope, then

$$\chi(\partial P) = 1 + (-1)^{d-1}, \qquad \chi(\operatorname{int} P) = (-1)^d.$$

Proof. Slice by a generic height function and induct on d. An interior boundary slice is homeomorphic to a $(d-2)$-sphere, while the top and bottom each contribute a convex face. Summing the jumps of Euler characteristic proves the first formula. The second follows from $[P] = [\operatorname{int} P] + [\partial P]$ and $\chi(P) = 1$. ⊛

Theorem 6.15 (Euler–Poincaré formula; Barvinok VI.3.2). Let $f_i(P)$ be the number of i-faces of P. Then

$$\sum_{i=0}^{d-1} (-1)^i f_i(P) = 1 + (-1)^{d-1}.$$

Set $f_{-1}(P) = f_d(P) = 1$.

Proof. The relative interiors of the nonempty faces partition P:

$$[P] = \sum_{\varnothing \neq F \preceq P} [\operatorname{relint} F].$$

Apply Euler characteristic and use $\chi(\operatorname{relint} F) = (-1)^{\dim F}$. ⊛

Definition 6.16 (f-vector). The vector

$$f(P) = (f_0(P), f_1(P), \ldots, f_{d-1}(P))$$

is the f-vector of P.

Example 6.17 (Dimension three). For a three-dimensional polytope, $f_0 - f_1 + f_2 = 2$. A tetrahedron gives $4 - 6 + 4 = 2$, and a cube gives $8 - 12 + 6 = 2$.

Remark 6.18 (What Euler's relation determines). The formula supplies one universal linear relation but does not characterize all possible integer f-vectors. The Dehn–Sommerville equations add relations for simple and simplicial polytopes; nonlinear realizability conditions are deeper.

6.4 VI.4 Polytopes with Many Faces: Cyclic Polytopes

Definition 6.19 (Cyclic polytope). For the moment curve $q(t) = (t, t^2, \dots, t^d)$ and parameters $t_1 < \cdots < t_n$, with $n \geq d+1$, define

$$C(d,n) = \operatorname{conv}\{q(t_1), \dots, q(t_n)\}.$$

Proposition 6.20 (Neighborliness; Barvinok VI.4.2). If $I \subset \{1, \dots, n\}$ and $|I| \leq d/2$, then

$$\operatorname{conv}\{q(t_i) : i \in I\}$$

is a face of $C(d,n)$. Thus every set of at most $\lfloor d/2 \rfloor$ vertices spans a face.

Proof. The assertion is immediate for $I = \varnothing$. Otherwise, write $I = \{i_1, \dots, i_r\}$ and take

$$p(t) = -\prod_{j=1}^{r} (t - t_{i_j})^2.$$

This polynomial has degree $2r \leq d$, vanishes at the selected parameters, and is strictly negative at all the others. Writing $p(t) = c_1 t + \cdots + c_d t^d - \alpha$, with unused coefficients equal to zero, gives a linear functional on the moment curve whose supporting hyperplane exposes exactly these vertices. ⊛

Theorem 6.21 (Gale evenness criterion). Let $I \subset \{1, \dots, n\}$ with $|I| = d$. The simplex $\operatorname{conv}\{q(t_i) : i \in I\}$ is a facet of $C(d,n)$ if and only if every two indices outside I have an even number of indices from I between them.

Proof. The selected d points are affinely independent by the Vandermonde determinant. A hyperplane through them corresponds on the moment curve to a nonzero polynomial of degree at most d with the selected parameters as roots. It is therefore a scalar multiple of $\prod_{i\in I}(t-t_i)$. Between any two unselected parameters, its sign changes once for each selected parameter. Its values at all unselected parameters have the same sign exactly when the evenness condition holds. In that case choose the scalar so that all those values are negative; the hyperplane then supports the polytope and exposes precisely the selected simplex. Conversely, a supporting hyperplane forces all unselected values to have the same sign and hence forces evenness. ⊛

Remark 6.22. Neighborliness maximizes the low-dimensional face numbers, and Gale evenness determines the entire face lattice. It therefore depends only on d, n, not on the specific parameters t_i. The upper bound theorem below shows that $C(d, n)$ maximizes every face number among d-polytopes with n vertices.

6.5 VI.5 Simple Polytopes

Definition 6.23 (Simple and simplicial polytopes). A d-polytope is *simple* if exactly d facets contain each vertex. It is *simplicial* if every facet is a simplex.

Proposition 6.24 (Duality). If $0 \in \operatorname{int} P$, then

$$P \text{ is simple} \quad \Longleftrightarrow \quad P^\circ \text{ is simplicial.}$$

Proof. Under polarity a vertex of P corresponds to a facet of P°, and the facets containing the vertex correspond to vertices of the dual facet. Exactly d incident facets means that this $(d-1)$-face has exactly d affinely independent vertices and is a simplex. ⊛

Proposition 6.25 (Local structure; Barvinok VI.5.2). Let $P \subset \mathbb{R}^d$ be simple and let v be a vertex. Then:

1. v has exactly d adjacent vertices;
2. any $k < d$ of them, together with v, lie in a unique k-face;
3. the intersection of any k facets containing v is a $(d-k)$-face;
4. if a linear functional is greater at v than at every adjacent vertex, then v is its unique maximizer on P;
5. every nonempty face of P is simple.

Proof. Translate v to the origin and use the d active constraints at v as coordinates, so the local feasible cone is $\mathbb{R}^d_+$. Its coordinate axes give the d edges, and setting selected coordinates to zero gives all local faces. A functional decreases along every axis exactly when all its coordinate coefficients are negative, so it is maximized at the origin on the whole orthant. Repeating the active-constraint count on each face proves the last assertion. ⊛

Example 6.26 (Cube and cross-polytope). The d-cube is simple: exactly d coordinate constraints are active at each vertex. Its polar, the cross-polytope, is simplicial. For $d > 2$, the cross-polytope is not simple.

Remark 6.27 (Linear functionals and vertices). Vertices of a simple polytope correspond to nondegenerate basic feasible solutions of a linear program. Active normals form a basis, and moving along a neighboring edge corresponds to a pivot. A degenerate vertex may have more than d active inequalities.

6.6 VI.6 h-Vectors and the Dehn–Sommerville Equations

Definition 6.28 (h-vector). For a simple d-polytope P, set $f_d(P) = 1$ and define

$$h_k(P) = \sum_{i=k}^{d} (-1)^{i-k} \binom{i}{k} f_i(P), \qquad k = 0, \dots, d.$$

The vector $(h_0, \dots, h_d)$ is the h-vector.

Lemma 6.29 (Inversion between f- and h-vectors; Barvinok VI.6.2).

$$f_i(P) = \sum_{k=i}^{d} \binom{k}{i} h_k(P), \qquad i = 0, \dots, d.$$

These equations determine the $h_k(P)$ uniquely.

Proof. For $f(t) = \sum_{i=0}^{d} f_i(P) t^i$ and $h(t) = \sum_{k=0}^{d} h_k(P) t^k$, the definition says $h(t) = f(t-1)$, while the inverse is $f(t) = h(t+1)$. ⊛

Theorem 6.30 (Index interpretation; Barvinok VI.6.3). Let P be a simple d-polytope, and choose a linear functional ℓ taking different values on adjacent vertices. The index of a vertex v is the number of adjacent vertices u with $\ell(u) < \ell(v)$. The number of vertices of index k is $h_k(P)$.

Double counting. Mark the unique ℓ-maximum on every i-face. A vertex of index k is the maximum of the unique i-face determined by any i of its k lower neighbors. Hence

$$f_i(P) = \sum_{k=i}^{d} \binom{k}{i} \#\{v : \operatorname{ind}_\ell(v) = k\}.$$

Compare this with the inversion formula. ⊛

Corollary 6.31 (Dehn–Sommerville equations; Barvinok VI.6.4). For a simple d-polytope,

$$h_k(P) = h_{d-k}(P), \qquad k = 0, \dots, d.$$

Proof. Replacing ℓ by $-\ell$ exchanges lower and higher neighbors at every vertex and changes index k to $d-k$. ⊛

Example 6.32 (Simplex and cube). For a d-simplex, $h = (1, 1, \dots, 1)$. For the d-cube, $h_k = \binom{d}{k}$.

Remark 6.33 (Genericity). The index interpretation assumes that the polytope is simple. A simplicial polytope receives the dual convention through polarity; a general polytope does not have the same local counting formula.

6.7 VI.7 The Upper Bound Theorem

Lemma 6.34 (Comparison with a facet; Barvinok VI.7.1). If P is a simple d-polytope and F is a facet, then

$$h_k(F) \le h_k(P), \qquad k = 0, \dots, d-1.$$

If every $k+1$ facets of P intersect, then equality holds at k.

Proof. Perturb a supporting functional of F so that it takes distinct values at all vertices and every vertex in F has a lower value than every vertex outside F. A vertex of F has exactly one neighbor outside F, and that neighbor has a higher value. Its index in F is therefore its index in P, giving the inequality.

Suppose that every $k+1$ facets of P intersect and that a vertex v outside F has index k. The $d-k$ upper edges at v determine a face G on which the functional is minimized at v. Local simplicity expresses G as the intersection of k facets. Thus $F \cap G \neq \varnothing$. But every point of F has a lower functional value than v, contradicting minimality on G. There are no additional vertices of index k outside F, so equality holds. ⊛

Lemma 6.35 (Double counting over facets; Barvinok VI.7.2). Let P be a simple d-polytope. For $k = 0, \dots, d-1$,

$$\sum_{F \text{ a facet of } P} h_k(F) = (d-k)h_k(P) + (k+1)h_{k+1}(P).$$

Proof. Choose a generic linear functional and count pairs (F, v) in which F is a facet and v has index k within F. Each facet through v contains all but one of its d adjacent edges. A vertex of index k in P is counted once for each of the $d-k$ facets omitting a higher edge. A vertex of index $k+1$ is counted once for each of the $k+1$ facets omitting a lower edge. Vertices of any other index contribute nothing. ⊛

Theorem 6.36 (Upper bound for h-numbers; Barvinok VI.7.3). If P is a simple d-polytope with n facets, then

$$h_k(P) \le \binom{n-d+k-1}{k}, \qquad k = 0, \dots, d.$$

If every k facets intersect, equality holds at that k.

Proof. The preceding lemmas give $nh_i(P) \geq (d-i)h_i(P) + (i+1)h_{i+1}(P)$, hence

$$h_{i+1}(P) \leq \frac{n-d+i}{i+1} h_i(P).$$

Iterate from $h_0 = 1$. ⊛

Theorem 6.37 (Upper bound theorem; Barvinok VI.7.4). Let P be any d-polytope with n facets. Let Q be a simple d-polytope with n facets such that every collection of $k \leq \lfloor d/2 \rfloor$ facets intersects in a $(d-k)$-face. Then

$$f_i(P) \leq f_i(Q), \qquad i = 0, \ldots, d-1.$$

After translating $C(d,n)$ so that the origin lies in its interior, one may take $Q = C(d,n)^\circ$. Equivalently, no d-polytope with n vertices has more i-faces than $C(d,n)$.

Remark 6.38 (Extremal face numbers). For $r = \lfloor d/2 \rfloor$,

$$h_k(Q) = \binom{n-d+k-1}{k} \quad (0 \leq k \leq r),$$

and the remaining entries follow from $h_k(Q) = h_{d-k}(Q)$. Thus $f_i(Q) = \sum_{k=i}^{d} \binom{k}{i} h_k(Q)$, and if $i > r$, $f_i(Q) = \binom{n}{d-i}$.

Proof. For simple P, the h-number bound applies up to $d/2$, and Dehn–Sommerville reflects it to all indices. The inversion formula has nonnegative coefficients, so the f-vector inequality follows. A small perturbation of the facet constants turns a general P into a simple polytope with the same number of facets and no fewer faces of any dimension. ⊛

Remark 6.39 (Order of the number of faces). For fixed d, the total number of faces of a polytope with n facets is $O(n^{\lfloor d/2 \rfloor})$. Merely listing all faces can require this many outputs.

6.8 VI.8 Centrally Symmetric Polytopes

Theorem 6.40 (Figiel–Lindenstrauss–Milman theorem; Barvinok VI.8.1). There is an absolute constant $\gamma > 0$ such that every centrally symmetric d-polytope with V vertices and F facets satisfies

$$\log V \cdot \log F \geq \gamma d.$$

Proof. Translate the center of symmetry to the origin and apply a linear transformation so that the John ellipsoid is B_2^d. Then $B_2^d \subset P \subset \sqrt{d}\, B_2^d$. Radially project the vertices to the sphere. The spherical cap bound and a union bound show that, outside a set of directions of measure at most $1/3$, $h_P(c) \lesssim \sqrt{\log V}$. Normalize the facet description as

$$P = \{x : \langle a_j, x\rangle \leq 1, j = 1, \ldots, F\}, \qquad \|a_j\| \leq 1.$$

A second cap estimate shows that, outside another exceptional set of measure at most $1/3$, $h_P(c) \gtrsim \sqrt{d/\log F}$. The two good sets intersect, so $\sqrt{d/\log F} \lesssim \sqrt{\log V}$. Squaring and rearranging proves the theorem. ⊛

Remark 6.41 (Comparison with the upper bound theorem). The upper bound theorem controls the largest possible face numbers for a fixed number of vertices or facets. Central symmetry yields a lower bound for the product of the logarithms of the numbers of vertices and facets. The proof uses John position and spherical measure in addition to the combinatorial structure of the polytope.

Remark 6.42 (Probabilistic method). Each exceptional set of directions has measure below $1/3$. The union bound then guarantees a direction satisfying both estimates. The argument proves existence without constructing such a direction explicitly.

7 Chapter VII: Lattices and Convex Bodies

This chapter studies the interaction between lattices and convex bodies. The determinant measures the volume of a fundamental domain. Blichfeldt's theorem and Minkowski's convex body theorem turn volume conditions into the existence of lattice points. Applications include short vectors, sphere packings, Diophantine approximation, dual lattices, the flatness theorem, and LLL reduction.

7.1 Lattices: Definitions, Examples, and Bases

Definition 7.1 (Lattice and lattice basis). An additive subgroup $\Lambda \subset \mathbb{R}^d$ is *discrete* if $B(0,\varepsilon) \cap \Lambda = \{0\}$ for some $\varepsilon > 0$. It is a full-rank lattice if it also spans $\mathbb{R}^d$. Linearly independent vectors $u_1, \dots, u_d \in \Lambda$ form a lattice basis when

$$\Lambda = \{\mu_1 u_1 + \cdots + \mu_d u_d : \mu_i \in \mathbb{Z}\}.$$

Every bounded set contains only finitely many lattice points, and nonzero lattice vectors have a positive minimum length. Standard examples are

$$\mathbb{Z}^d, \qquad A_d = \{x \in \mathbb{Z}^{d+1} : \sum_i x_i = 0\}, \qquad D_d = \{x \in \mathbb{Z}^d : \sum_i x_i \text{ is even}\},$$

as well as E_8, obtained by adjoining a half-integral translate to D_8. The lattices A_2, A_3, D_4, E_8 give notable sphere packings in their respective dimensions.

Lemma 7.2 (Nearest lattice layer to a rational subspace). Let $b_1, \dots, b_k \in \Lambda$ be linearly independent, $k < d$, and $L = \operatorname{lin}(b_1, \dots, b_k)$. Some $v \in \Lambda \setminus L$ attains

$$\operatorname{dist}(v, L) = \min_{w \in \Lambda \setminus L} \operatorname{dist}(w, L) > 0.$$

Proof. Let $\Pi = \{\sum_{i=1}^k \alpha_i b_i : 0 \le \alpha_i \le 1\}$. A bounded neighborhood of Π contains finitely many lattice points, so a nearest point outside L exists. Every point of L can be translated by an element of $\sum_i \mathbb{Z} b_i$ back into Π, without changing the distance. The chosen point therefore minimizes the distance to all of L. ⊛

Theorem 7.3 (Existence of a lattice basis; Barvinok VII.1.4). Every positive-rank lattice has a basis. More explicitly, let $b_1, \dots, b_d \in \Lambda$ be any linearly independent vectors. Put $L_k = \operatorname{lin}(b_1, \dots, b_k)$, $L_0 = \{0\}$, and choose $u_k \in \Lambda \cap (L_k \setminus L_{k-1})$ closest to L_{k-1}. Then $u_1, \dots, u_d$ is a basis of Λ.

Proof. Induct with $\Lambda_k = \Lambda \cap L_k$. For $v \in \Lambda_k$, compare its coefficient in the b_k-direction with that of u_k, and call the ratio μ. If $\mu \notin \mathbb{Z}$, then $v - \lfloor \mu \rfloor u_k$ is outside L_{k-1} but has distance $\{\mu\} \operatorname{dist}(u_k, L_{k-1})$, contradicting minimality. Hence $\mu \in \mathbb{Z}$, and induction applies to $v - \mu u_k$. ⊛

Definition 7.4 (Fundamental parallelepiped). For a basis $u_1, \dots, u_d$, the half-open set

$$\Pi = \left\{ \sum_{i=1}^{d} \alpha_i u_i : 0 \le \alpha_i < 1 \right\}$$

is a fundamental parallelepiped. The half-open convention gives one representative of each lattice congruence class; closed neighboring translates would overlap on their boundaries.

Remark 7.5. A discrete subgroup need not have full rank in its ambient space. For example, $\mathbb{Z}(1,0) \subset \mathbb{R}^2$ has rank one. A lower-rank lattice should be regarded inside its linear span.

7.2 Determinant, Density, Sublattices, and Index

Lemma 7.6 (Unique decomposition by a fundamental domain). Every $x \in \mathbb{R}^d$ has a unique expression $x = v + y$ with $v \in \Lambda$ and $y \in \Pi$. Thus the translates $\{\Pi + v : v \in \Lambda\}$ tile $\mathbb{R}^d$ without overlap.

Proof. Write $x = \sum_i \alpha_i u_i$, take $v = \sum_i \lfloor \alpha_i \rfloor u_i$ and $y = \sum_i \{\alpha_i\} u_i$. In two such decompositions, the difference of the two y's is a lattice vector whose coefficient in every basis direction has magnitude below one, so it is zero. ⊛

Theorem 7.7 (Determinant and asymptotic point density; Barvinok VII.2.3). The volume of a fundamental parallelepiped is independent of the basis and is denoted $\det \Lambda$. For $B(\rho) = \{x : \|x\| \le \rho\}$,

$$\lim_{\rho \to \infty} \frac{\operatorname{vol} B(\rho)}{|B(\rho) \cap \Lambda|} = \det \Lambda.$$

Thus $1/\det \Lambda$ is the asymptotic number of lattice points per unit volume.

Proof. If $\Pi \subset B(\gamma)$, let $X(\rho)$ be the union of fundamental cells centered at $B(\rho) \cap \Lambda$. Then

$$\operatorname{vol} X(\rho) = |B(\rho) \cap \Lambda| \operatorname{vol} \Pi, \qquad B(\rho - \gamma) \subset X(\rho) \subset B(\rho + \gamma).$$

Use $\operatorname{vol} B(r) = r^d \operatorname{vol} B(1)$ and let $\rho \to \infty$. The limit depends only on the lattice. For the basis matrix $U = [u_1 \ \cdots \ u_d]$, one has $\det \Lambda = |\det U|$. ⊛

Definition 7.8 (Sublattice and index). If $\Lambda_0 \subseteq \Lambda$ is also full rank, it is a sublattice. The number of cosets in Λ/Λ_0 is the index $[\Lambda : \Lambda_0]$.

Theorem 7.9 (Index and volume). For a full-rank sublattice $\Lambda_0 \subseteq \Lambda$,

$$[\Lambda : \Lambda_0] = |\Pi_0 \cap \Lambda| = \frac{\det \Lambda_0}{\det \Lambda}.$$

Thus the half-open parallelepiped spanned by independent $u_i \in \mathbb{Z}^d$ contains exactly

$$|\det[u_1 \ \cdots \ u_d]|$$

integer points.

Example 7.10 (Pick's formula). For an integral polygon $P \subset \mathbb{R}^2$,

$$\left|P \cap \mathbb{Z}^2\right| = \operatorname{area}(P) + \frac{1}{2}\left|\partial P \cap \mathbb{Z}^2\right| + 1.$$

The book combines lattice-point counting in half-open parallelograms with triangulation. Chapter VIII extends this to Ehrhart polynomials $\left|kP \cap \mathbb{Z}^d\right|$.

7.3 Blichfeldt and Minkowski's Convex Body Theorem

Theorem 7.11 (Blichfeldt's theorem). If a measurable set $X \subset \mathbb{R}^d$ satisfies $\operatorname{vol} X > \det \Lambda$, then distinct $x, y \in X$ exist with $x - y \in \Lambda$.

Proof. Fold the translates $X_u = (X - u) \cap \Pi$ into one fundamental domain. If they covered almost every point at most once, then $\sum_u \operatorname{vol} X_u \le \operatorname{vol} \Pi$, contradicting the volume assumption. Hence some $z \in \Pi$ lies in $X_u \cap X_v$ with $u \neq v$, and $z + u, z + v \in X$. ⊛

Equivalently, the periodic count $N(z) = \sum_{u \in \Lambda} [X](z + u)$ satisfies

$$\int_{\Pi} N(z)\, \mathrm{d}z = \operatorname{vol} X.$$

An average above one forces a point covered at least twice.

Theorem 7.12 (Minkowski's convex body theorem; Barvinok VII.3.2). Let $A \subset \mathbb{R}^d$ be convex and symmetric about the origin, so $A = -A$. If

$$\operatorname{vol} A > 2^d \det \Lambda,$$

then A contains a nonzero lattice point. If A is compact, the strict inequality may be replaced by $\operatorname{vol} A \ge 2^d \det \Lambda$.

Proof. Apply Blichfeldt to $X = A/2$. Distinct $x, y \in X$ have $u = x - y \in \Lambda$. Since $2x, 2y \in A$, symmetry gives $-2y \in A$, and convexity gives $u = \frac{1}{2}(2x) + \frac{1}{2}(-2y) \in A$. For equality with compact A, apply the strict version to $(1 + \varepsilon)A$ and use discreteness as $\varepsilon \downarrow 0$. ⊛

For a convex body $A = -A$, define successive minima by

$$\lambda_k(A, \Lambda) = \inf\{t > 0 : \dim \operatorname{lin}(tA \cap \Lambda) \ge k\}.$$

Minkowski's second theorem gives

$$\lambda_1 \cdots \lambda_d \operatorname{vol} A \le 2^d \det \Lambda.$$

Lemma 7.13 (Volume of the Euclidean unit ball). For $\beta_d = \operatorname{vol} B_2^d$,

$$\beta_d = \frac{\pi^{d/2}}{\Gamma(d/2 + 1)}, \qquad \Gamma(s) = \int_0^\infty t^{s-1} e^{-t}\, \mathrm{d}t \quad (s > 0).$$

Comparing $\int_{\mathbb{R}^d} e^{-\|x\|^2}\, \mathrm{d}x = \pi^{d/2}$ with polar coordinates proves the formula. Stirling's approximation gives $\beta_d^{1/d} \asymp d^{-1/2}$.

7.4 Applications: Four Squares and Diophantine Approximation

Lemma 7.14 (A lattice defined by congruences). If $c_i \in \mathbb{Z}^d$ and $\gamma_i \in \mathbb{N}$, then

$$\Lambda = \{x \in \mathbb{Z}^d : \langle c_i, x\rangle \equiv 0 \pmod{\gamma_i},\ i = 1, \dots, m\}$$

is a sublattice with $\det \Lambda \le \gamma_1 \cdots \gamma_m$.

It is the kernel of the homomorphism $x \mapsto (\langle c_i, x\rangle \bmod \gamma_i)_i$; the image has at most the product of the moduli elements, and the index formula gives the bound.

Theorem 7.15 (Lagrange's four-square theorem). Every positive integer n can be written as

$$n = x_1^2 + x_2^2 + x_3^2 + x_4^2, \qquad x_i \in \mathbb{Z}.$$

Davenport–Minkowski argument. The four-square identity reduces the problem to prime n. A pigeonhole argument modulo n gives α, β with $\alpha^2 + \beta^2 + 1 \equiv 0 \pmod{n}$. Define

$$\Lambda = \{x \in \mathbb{Z}^4 :\ x_1 \equiv \alpha x_3 + \beta x_4, \quad x_2 \equiv \beta x_3 - \alpha x_4 \pmod{n}\}.$$

Then $\det \Lambda \le n^2$, and n divides the squared norm of every lattice vector. The open four-dimensional ball of radius $\sqrt{2n}$ has volume $2\pi^2 n^2 > 2^4 \det \Lambda$. Minkowski gives nonzero $x \in \Lambda$ with $0 < \|x\|^2 < 2n$. The only positive multiple of n in this range is n. ⊛

Theorem 7.16 (Rational approximation and Hurwitz's constant). There is an absolute $0 < C < 1$ such that, for every irrational α, arbitrarily large $q \in \mathbb{N}$ and $p \in \mathbb{Z}$ satisfy

$$\left|\alpha - \frac{p}{q}\right| \le \frac{C}{q^2}.$$

One may take $C = 1/\sqrt{5}$, and this uniform constant is optimal.

Convex-geometric argument for $C = 1$. For large Q, consider

$$A_Q = \{(x, y) : |\alpha x - y| \le Q^{-1},\ |x| \le Q\}.$$

Its area is 4, and it is compact and symmetric about the origin. The equality case of Minkowski's theorem gives a nonzero integer point $(q, p) \in A_Q$. For $Q > 1$, a point with $q = 0$ would also have $p = 0$. Replacing (q, p) by its negative if necessary gives $0 < q \le Q$ and $|\alpha q - p| \le Q^{-1} \le q^{-1}$. The denominators cannot remain bounded because $Q^{-1} \to 0$ and α is irrational. ⊛

Continued fractions improve the constant to $1/\sqrt{5}$, and the golden ratio proves optimality. The same convex-body construction in dimension $n + 1$ simultaneously approximates $\alpha_1, \dots, \alpha_n$ with error of order $q^{-1-1/n}$.

7.5 Sphere Packings

Definition 7.17 (Packing radius and density). The packing radius $\rho(\Lambda)$ is the largest ρ for which the open balls $\{B(u, \rho) : u \in \Lambda\}$ are pairwise disjoint. Their asymptotic covered proportion is the packing density $\sigma(\Lambda)$.

If $\lambda_1 = \min_{u \neq 0} \|u\|$, then

$$\rho(\Lambda) = \frac{\lambda_1}{2}, \qquad \sigma(\Lambda) = \frac{\beta_d \rho(\Lambda)^d}{\det \Lambda}.$$

Example 7.18 (Square and hexagonal lattices). For $\mathbb{Z}^2$, one has $\rho = 1/2$ and $\sigma = \pi/4$. For a suitable isometric realization of A_2,

$$\sigma(A_2) = \frac{\pi}{\sqrt{12}} > \frac{\pi}{4}.$$

The hexagonal arrangement wastes less area in each fundamental cell.

Corollary 7.19 (Minkowski's short-vector bound). Every lattice satisfies

$$\rho(\Lambda) \le \frac{\Gamma(d/2 + 1)^{1/d}}{\sqrt{\pi}} \det(\Lambda)^{1/d} = \sqrt{\frac{d}{2\pi e}} \left(1 + O\left(\frac{\log d}{d}\right)\right) \det(\Lambda)^{1/d}.$$

This is equivalent to the coarse bound $\sigma(\Lambda) \le 1$ and illustrates the general principle that the volume of a fundamental domain forces a short vector.

7.6 Minkowski–Hlawka: Random Lattices and Existence

Lemma 7.20 (Average count under a random translate). If $M \subset \mathbb{R}^d$ is measurable and Π is a fundamental domain, then

$$\int_{\Pi} |M \cap (\Lambda + x)| \, \mathrm{d}x = \operatorname{vol} M.$$

If $\operatorname{vol} M < \det \Lambda$, the average count is below one, so some translate misses M.

Theorem 7.21 (Minkowski–Hlawka theorem; Barvinok VII.6.2). Let $d > 1$, let $M \subset \mathbb{R}^d$ be bounded and Jordan measurable, and let $\delta > \operatorname{vol} M$. There is a lattice Λ such that

$$\det \Lambda = \delta, \qquad M \cap (\Lambda \setminus \{0\}) = \varnothing.$$

Main argument. Scale to $\delta = 1$. Choose $\alpha > 0$ sufficiently small and slice M by $x_d = k\alpha$. Fix basis vectors $u_i = \alpha^{-1/(d-1)} e_i$ for $i < d$, and let $u_d(x) = x + \alpha e_d$, with x uniform in the transverse fundamental domain. Every resulting lattice has determinant one. Averaging over the layers makes the expected number of lattice points with $k \neq 0$ equal to $\alpha \sum_{k \neq 0} \operatorname{vol}_{d-1} M_k < 1$. Some choice has none, and choosing the transverse cell large enough also excludes nonzero points from the zero layer. ⊛

Corollary 7.22 (Existence bound for lattice packings). For every $\sigma < 2^{-d}$, a d-dimensional lattice exists with packing density at least σ. Equivalently, some determinant-one lattice has packing radius $\Omega(\sqrt{d})$.

7.7 Dual Lattices, Covering Radius, and Transference

Definition 7.23 (Dual lattice).

$$\Lambda^* = \{y \in \mathbb{R}^d : \langle x, y\rangle \in \mathbb{Z} \text{ for every } x \in \Lambda\}.$$

If U is a basis matrix, $U^{-\mathsf{T}}$ is a dual basis, and

$$\det \Lambda^* = \frac{1}{\det \Lambda}, \qquad (\Lambda^*)^* = \Lambda, \qquad (T\Lambda)^* = (T^{\mathsf{T}})^{-1}\Lambda^*.$$

Lemma 7.24 (Short vectors in a lattice and its dual).

$$\rho(\Lambda)\rho(\Lambda^*) \le \frac{d}{4}.$$

Proof. Apply Minkowski's Euclidean short-vector bound to both lattices and multiply, using reciprocal determinants. ⊛

Definition 7.25 (Covering radius).

$$\mu(\Lambda) = \sup_{x \in \mathbb{R}^d} \operatorname{dist}(x, \Lambda)$$

is the least radius for which closed balls centered at lattice points cover the space.

Theorem 7.26 (Transference between covering and packing; Barvinok VII.7.4). For a dimensional constant $c(d) \le d^{3/2}/4$,

$$\frac{1}{4} \le \mu(\Lambda)\rho(\Lambda^*) \le c(d).$$

Main argument. For the lower bound, choose $u_1, \dots, u_d$ successively shortest among lattice vectors outside the span of the preceding choices. The midpoint $u_d/2$ is at distance at least $\|u_d\|/2$ from the lattice. Indeed, a closer lattice point v would make either v or $2v - u_d$ a vector outside $\operatorname{lin}(u_1, \dots, u_{d-1})$ shorter than u_d. For a shortest nonzero $w \in \Lambda^*$, at least one integer pairing $\langle u_i, w\rangle$ is nonzero. Since $\|u_i\| \le \|u_d\|$, this gives $\|u_d\|\,\|w\| \ge 1$ and hence $\mu(\Lambda)\rho(\Lambda^*) \ge 1/4$. For the upper bound, project orthogonally to a shortest vector u. If Λ_1 is the projected lattice, then

$$\mu(\Lambda)^2 \le \mu(\Lambda_1)^2 + \rho(\Lambda)^2.$$

Use $\Lambda_1^* \subseteq \Lambda^*$, induction, and the preceding short-vector inequality. ⊛

7.8 The Flatness Theorem

For $w \neq 0$, define

$$\operatorname{width}_w(K) = \max_{x \in K} \langle w, x\rangle - \min_{x \in K} \langle w, x\rangle .$$

When $w \in \Lambda^*$, lattice points lie on the integer layers $\langle w, x\rangle = k$.

Lemma 7.27 (A narrow direction for a lattice-free ellipsoid). If an ellipsoid E contains no lattice point, some nonzero $w \in \Lambda^*$ satisfies

$$\operatorname{width}_w(E) \le d^{3/2}.$$

For a ball, use $2r\,\|w\|$, the bound $r \le \mu(\Lambda)$, and transference. For an ellipsoid, write $E = T(B)$ with T invertible and linear; the ball B need not be centered at the origin. Apply the ball result to $\Lambda_1 = T^{-1}\Lambda$. If $v \in \Lambda_1^*$ is the resulting direction, then $w = (T^{\mathsf{T}})^{-1}v \in \Lambda^*$ and $\operatorname{width}_w(E) = \operatorname{width}_v(B)$.

Theorem 7.28 (Flatness theorem; Barvinok VII.8.3). If a convex body $K \subset \mathbb{R}^d$ contains no point of Λ, then some nonzero $w \in \Lambda^*$ satisfies

$$\operatorname{width}_w(K) \le d^{5/2}.$$

Proof. Let E be the John ellipsoid of K, centered at a. Then $E \subseteq K \subseteq a + d(E - a)$. Since E is lattice-free, choose the narrow dual direction from the lemma. Width is translation invariant and scales linearly, so $\operatorname{width}_w(K) \le d \operatorname{width}_w(E) \le d^{5/2}$. ⊛

In fixed-dimensional integer feasibility, a lattice-free convex region crosses only polynomially many integer layers in this direction. Branching over those layers reduces the dimension and forms the geometric basis of Lenstra's algorithm.

Lemma 7.29 (Parity bound for an integral polytope; Barvinok VII.8.3, Problem 2). If an integral polytope in $\mathbb{R}^d$ has more than 2^d vertices, it contains an integer point other than its vertices.

Proof. There are 2^d coordinate parity classes in $\mathbb{Z}^d$. Two distinct vertices in the same class have an integral midpoint. This midpoint belongs to the polytope and is not a vertex, since it is a strict convex combination of two distinct points. ⊛

Theorem 7.30 (Doignon's integer Helly theorem). Let $A_1, \dots, A_m \subset \mathbb{R}^d$ be convex sets, with $m \ge 2^d$. If every subfamily of 2^d sets has an integer point in its intersection, then

$$\bigcap_{i=1}^{m} A_i \cap \mathbb{Z}^d \neq \varnothing.$$

This is Problem 3 in Section VII.8.3 of the book. The number 2^d is the integer analogue of the $d + 1$ in Helly's theorem. It is sharp: for each $v \in \{0, 1\}^d$, take the halfspace

$$\sum_{j : v_j = 0} x_j + \sum_{j : v_j = 1} (1 - x_j) \ge 1,$$

which excludes v and contains every other cube vertex. Any subfamily of $2^d - 1$ halfspaces therefore contains an integer point. Their full intersection contains none: given $x \in \mathbb{Z}^d$, choose $v_j = 0$ when $x_j \le 0$ and $v_j = 1$ when $x_j \ge 1$. The left side of the corresponding inequality is at most zero.

7.9 Short Vectors and LLL-Reduced Bases

For a lattice basis $u_1, \dots, u_d$, let $w_1, \dots, w_d$ be the unnormalized Gram–Schmidt vectors:

$$u_k = w_k + \sum_{i<k} \alpha_{ki} w_i, \qquad w_k \perp \operatorname{lin}(u_1, \dots, u_{k-1}).$$

Definition 7.31 (LLL-reduced basis). With $\tau = 4/3$, a basis is LLL-reduced if

$$|\alpha_{ki}| \le \frac{1}{2} \quad (i < k),$$

and

$$\operatorname{dist}(u_k, L_{k-1})^2 \le \tau \operatorname{dist}(u_{k+1}, L_{k-1})^2, \qquad k = 1, \dots, d-1.$$

Theorem 7.32 (Approximation by the first basis vector; Barvinok VII.9.2). For an LLL-reduced basis and every nonzero $v \in \Lambda$,

$$\|u_1\| \le 2^{(d-1)/2} \|v\|.$$

Proof. Size reduction and the Lovász condition give $\|w_{k+1}\|^2 \geq \frac{1}{2}\|w_k\|^2$. Thus $\|w_k\| \geq 2^{-(k-1)/2}\|u_1\|$. A nonzero lattice vector has a nonzero integer coefficient in some last basis direction, so its distance from the preceding subspace is at least $\|w_k\|$. The worst case is $k = d$. ⊛

Algorithm 7.33 (LLL basis reduction). Starting from any basis, repeat:

1. compute the Gram–Schmidt coefficients and reduce each u_k in the order $i = k-1, k-2, \dots, 1$. When $|\alpha_{ki}| > 1/2$, replace $u_k \leftarrow u_k - [\alpha_{ki}]u_i$, where brackets denote the nearest integer;
2. after all size conditions hold, check the Lovász conditions. If an adjacent pair violates one, exchange the vectors and return to the size-reduction stage;
3. stop when all Lovász conditions hold.

All steps are unimodular integer basis changes and preserve Λ.

Theorem 7.34 (Termination potential and number of reduction stages; Barvinok VII.9.4). Let $\Lambda_k = \Lambda \cap L_k$ and

$$D(u_1, \dots, u_d) = \prod_{k=1}^{d-1} \det \Lambda_k.$$

Size reduction leaves D unchanged; every necessary exchange multiplies it by at most $\tau^{-1/2} < 1$. Put $\lambda = \min_{v \in \Lambda \setminus \{0\}} \|v\|$ and let D_0 be the potential of the initial basis. For every basis,

$$D \geq B_\lambda, \qquad B_\lambda = \lambda^{d(d-1)/2} \prod_{k=1}^{d-1} k^{-k/2}.$$

If a positive integer m satisfies

$$\tau^{-m/2} D_0 < B_\lambda,$$

then the algorithm uses at most $m+1$ complete size-reduction stages and at most $m+1$ checks of the Lovász conditions. In particular, every lattice has an LLL-reduced basis.

Proof. The first k basis vectors form a basis of Λ_k, and $\det \Lambda_k = \prod_{i=1}^{k} \|w_i\|$. Size reduction preserves the subspaces L_k. Exchanging u_k and u_{k+1} changes only L_k and replaces the factor $\|w_k\|$ by $\operatorname{dist}(u_{k+1}, L_{k-1}) < \tau^{-1/2} \|w_k\|$.

The shortest nonzero vector in Λ_k has length at least λ. Applying Minkowski's theorem in L_k to the cube inscribed in a Euclidean ball gives $\lambda \le \sqrt{k}(\det \Lambda_k)^{1/k}$, hence $\det \Lambda_k \ge \lambda^k k^{-k/2}$. Multiplying these inequalities gives $D \ge B_\lambda$. After sufficiently many exchanges the geometric decrease would contradict this lower bound. Each exchange is followed by one size-reduction stage, and the last check terminates the algorithm. ⊛

A polynomial-time analysis also controls input bit length and growth of intermediate numbers. Strict decrease of the potential alone proves only finite termination. Gram–Schmidt uses arbitrary real orthogonal transformations to obtain orthogonal coordinates; LLL permits only unimodular integer basis changes and keeps the same lattice while improving the geometry of its basis.

8 Chapter VIII: Lattice Points and Polyhedra

For a rational polyhedron $P \subset \mathbb{R}^d$, this chapter studies $P \cap \mathbb{Z}^d$ through multivariate generating functions. A lattice point $m = (m_1, \dots, m_d)$ corresponds to the monomial $x^m = x_1^{m_1} \cdots x_d^{m_d}$, and

$$f(P, x) = \sum_{m \in P \cap \mathbb{Z}^d} x^m.$$

Geometric series give the function of a simplicial rational cone, triangulation extends it to general rational cones, Brion's theorem expresses a polytope as a sum of vertex-cone functions, and Ehrhart theory counts lattice points in dilates.

8.1 Generating Functions of Simplicial Rational Cones

Definition 8.1 (Multivariate lattice-point generating function). For $m \in \mathbb{Z}^d$, write $x^m = x_1^{m_1} \cdots x_d^{m_d}$, and for $P \subset \mathbb{R}^d$ set

$$f(P, x) = \sum_{m \in P \cap \mathbb{Z}^d} x^m.$$

For a rational polyhedron P containing no line, the series converges absolutely on a nonempty open set and has a rational continuation. A polyhedron containing a line generally has no such convergence region; its generating-function valuation will be defined to be zero.

The one-dimensional model is $\sum_{n \geq 0} x^n = (1 - x)^{-1}$ for $|x| < 1$. Every factor $1 - x^u$ below comes from a geometric series along an integral ray u.

Definition 8.2 (Simplicial rational cone). For linearly independent $u_1, \dots, u_k \in \mathbb{Z}^d$, let

$$K = \operatorname{cone}(u_1, \dots, u_k) = \left\{ \sum_{i=1}^{k} \lambda_i u_i : \lambda_i \geq 0 \right\},$$

with half-open fundamental parallelepiped

$$\Pi = \left\{ \sum_{i=1}^{k} \alpha_i u_i : 0 \leq \alpha_i < 1 \right\}.$$

Lemma 8.3 (Generating function of a simplicial cone; Barvinok VIII.1.2). On

$$U = \{x \in (\mathbb{C} \setminus \{0\})^d : |x^{u_i}| < 1, i = 1, \dots, k\},$$

the series converges absolutely and

$$f(K, x) = \left(\sum_{a \in \Pi \cap \mathbb{Z}^d} x^a \right) \prod_{i=1}^{k} \frac{1}{1 - x^{u_i}}.$$

Proof. Every $m \in K \cap \mathbb{Z}^d$ has a unique expression $m = a + \sum_i \nu_i u_i$, where $a \in \Pi \cap \mathbb{Z}^d$ and $\nu_i \in \mathbb{Z}_{\geq 0}$. Split the cone coordinates into integer and fractional parts. The series factors into a finite numerator and k geometric series. ⊛

Example 8.4 (Unimodular orthant). For $K = \mathbb{R}^d_+ = \operatorname{cone}(e_1, \dots, e_d)$, $\Pi \cap \mathbb{Z}^d = \{0\}$, so

$$f(K, x) = \prod_{i=1}^{d} \frac{1}{1 - x_i}.$$

If the ray generators span a proper sublattice of $\mathbb{Z}^d \cap \operatorname{lin} K$, the numerator records the additional coset representatives in Π.

Proposition 8.5 (Cone reciprocity). For a k-dimensional simplicial rational cone,

$$f(\operatorname{relint} K, x^{-1}) = (-1)^k f(K, x).$$

Let $\Pi_- = \{\sum_i \alpha_i u_i : 0 \leq \alpha_i < 1\}$ and $\Pi_+ = \{\sum_i \alpha_i u_i : 0 < \alpha_i \leq 1\}$. The map $a \mapsto \sum_i u_i - a$ pairs their lattice points. The closed cone uses representatives in Π_-, and its relative interior uses those in Π_+. Combining this pairing with $(1 - x^{-u})^{-1} = -x^u (1 - x^u)^{-1}$ proves the identity between rational functions.

Remark 8.6 (Formal series and rational functions). The same rational function has different Laurent expansions in different regions:

$$\frac{1}{1 - x} = \sum_{n \geq 0} x^n \quad (|x| < 1), \qquad \frac{1}{1 - x} = -\sum_{n \geq 1} x^{-n} \quad (|x| > 1).$$

Brion's theorem adds rational functions; the vertex-cone series need not share a convergence region.

8.2 Rational Cones: Bases, Triangulations, and Rationality

Definition 8.7 (Rational cone and rational polyhedron). A cone $K = \{x : \langle c_i, x\rangle \le 0\}$ with $c_i \in \mathbb{Z}^d$ is rational. A polyhedron $P = \{x : \langle c_i, x\rangle \le \alpha_i\}$ with $\alpha_i \in \mathbb{Z}$ is rational. A polytope is rational or integral when its vertices lie in $\mathbb{Q}^d$ or $\mathbb{Z}^d$, respectively.

For polytopes, rational inequality and vertex descriptions are equivalent. Vertices solve integer linear systems and hence have rational coordinates; conversely, a polytope with rational vertices has rational supporting hyperplanes, whose denominators can be cleared. For an unbounded polyhedron, a finite generator description must also account for the recession cone. A rational polyhedron containing no line has rational vertices and a recession cone generated by integral rays; rational vertices alone do not determine rationality.

Lemma 8.8 (An integral polytope base for a pointed rational cone). If a nonzero rational cone K contains no line, it has an integral polytope base Q: every $x \in K \setminus \{0\}$ has a unique expression $x = \lambda y$, where $\lambda > 0$ and $y \in Q$.

Proof. Write $K = \{x : \langle c_i, x\rangle \le 0\}$ and set $c = \sum_i c_i$. Pointedness gives $\langle c, x\rangle < 0$ for nonzero $x \in K$. The section $P = K \cap \{x : \langle c, x\rangle = -1\}$ is a bounded rational polyhedron. Clearing a common denominator gives an integral polytope $Q = \delta P$ with the same radial property. ⊛

Lemma 8.9 (Triangulation). Every polytope can be triangulated into full-dimensional simplices meeting along common faces, using only its original vertices.

A generic lifting $v_i \mapsto (v_i, \tau_i)$ and projection of the lower simplicial faces gives such a triangulation.

Theorem 8.10 (Rationality for a pointed rational cone). If a rational cone $K \subset \mathbb{R}^d$ contains no line, then $f(K, x)$ converges absolutely on a nonempty open set and extends to

$$f(K, x) = \sum_{j=1}^{N} \frac{p_j(x)}{(1 - x^{u_{j1}}) \cdots (1 - x^{u_{jd}})},$$

where the p_j are Laurent polynomials and $u_{ji} \in \mathbb{Z}^d$.

Proof. Triangulate an integral base and cone over the simplices. Inclusion-exclusion writes $[K]$ as a linear combination of the simplicial cones and their common faces. Apply the preceding formula to each term. A common convergence region is $\{e^{a+ib} : a \in \operatorname{int} K^\circ, b \in \mathbb{R}^d\}$, which is nonempty because K is pointed. ⊛

8.3 Rational Polyhedra and the Generating-Function Valuation

Lemma 8.11 (Homogenization). If a rational polyhedron $P \subset \mathbb{R}^d$ contains no line, then $f(P, x)$ converges on a nonempty open set and is rational.

Proof. For $P = \{x : \langle c_i, x\rangle \le \alpha_i\}$, form

$$\widehat{K} = \{(x, t) : \langle c_i, x\rangle - \alpha_i t \le 0, t \ge 0\}.$$

The section $t = 1$ is P. Extracting the terms whose last coordinate is one from the rational cone series gives $f(P, x)$. ⊛

Let $\mathcal{P}(\mathbb{Q}^d)$ be the real span of indicators of rational polyhedra.

Theorem 8.12 (Lattice-point generating-function valuation; Barvinok VIII.3.3). There is a unique linear map

$$F : \mathcal{P}(\mathbb{Q}^d) \longrightarrow \mathbb{C}(x_1, \dots, x_d)$$

such that:

1. if P contains no line, $F([P]) = f(P, x)$ by rational continuation;
2. $F([P + m]) = x^m F([P])$ for $m \in \mathbb{Z}^d$;
3. if P contains a line, then $F([P]) = 0$.

Main argument. Polyhedra without lines span the algebra. To show that the extension is well defined, start from a relation $\sum_i \alpha_i [P_i] = 0$, multiply it by indicators of finite intersections P_I, and use inclusion-exclusion. For each fixed I, all relevant intersections lie in P_I and hence share a convergence region; pointwise lattice summation proves the corresponding rational identity. Substitution back into inclusion-exclusion yields $\sum_i \alpha_i f(P_i, x) = 0$. If P contains a line, some nonzero $m \in \mathbb{Z}^d$ satisfies $P + m = P$; translation covariance gives $F([P]) = x^m F([P])$, forcing $F([P]) = 0$. ⊛

Example 8.13 (A polyhedron containing a line). For $P_+ = [0, \infty)$ and $P_- = (-\infty, 0]$,

$$F(P_+) = \frac{1}{1-x}, \qquad F(P_-) = \frac{1}{1-x^{-1}}, \qquad F(\mathbb{R}) = 0,$$

and

$$0 = \frac{1}{1-x} + \frac{1}{1-x^{-1}} - 1.$$

The two series have no common convergence region, but their rational continuations satisfy the indicator identity $[\mathbb{R}] = [P_+] + [P_-] - [\{0\}]$.

8.4 Brion's Theorem: From a Polyhedron to Its Vertex Cones

Definition 8.14 (Supporting cone). For $v \in P$, define

$$\operatorname{cone}(P, v) = \{x : \varepsilon x + (1-\varepsilon)v \in P \text{ for some } 0 < \varepsilon < 1\}.$$

If P is given by inequalities, retaining only those active at v gives this translated cone. Points in the relative interior of the same face have the same supporting cone.

Let $\mathcal{P}_0(\mathbb{Q}^d)$ be the span of indicators of rational polyhedra containing lines. The valuation F vanishes on this subspace.

Lemma 8.15 (Vertex-cone decomposition modulo lines). For every rational polyhedron P,

$$[P] - \sum_{v \in \operatorname{ext} P} [\operatorname{cone}(P, v)] \in \mathcal{P}_0(\mathbb{Q}^d).$$

Main argument. Verify the assertion for a standard simplex by inclusion-exclusion; all remaining terms contain lines. Pass to arbitrary polytopes through linear images. If P contains a line, it has no vertices and $[P] \in \mathcal{P}_0(\mathbb{Q}^d)$, so the assertion holds directly. Otherwise, for an unbounded polyhedron use $P = \operatorname{conv}(\operatorname{ext} P) + \operatorname{rec}(P)$, apply the bounded decomposition, take the product with the recession cone, and push forward by addition. Terms containing lines remain in $\mathcal{P}_0(\mathbb{Q}^d)$. ⊛

For a polytope P, the more precise Brianchon–Gram identity is

$$[P] = \sum_{G \neq \varnothing} (-1)^{\dim G} [\operatorname{cone}(P, G)].$$

The sum runs over all nonempty faces, including P itself. Modulo $\mathcal{P}_0$, all positive-dimensional face cones vanish. For an unbounded polyhedron, the corresponding Brianchon–Gram formula sums only over its bounded faces.

Theorem 8.16 (Brion's theorem; Barvinok VIII.4.6). For every rational polyhedron P,

$$F([P]) = \sum_{v \in \operatorname{ext} P} F([\operatorname{cone}(P, v)]).$$

Proof. Apply F to the preceding decomposition; it kills the error term in $\mathcal{P}_0$. ⊛

Example 8.17 (Standard integral triangle). For $P = \operatorname{conv}\{(0,0), (1,0), (0,1)\}$, the three vertex cones give

$$F_0 = \frac{1}{(1-x_1)(1-x_2)}, \quad F_1 = \frac{x_1}{(1-x_1^{-1})(1-x_1^{-1}x_2)}, \quad F_2 = \frac{x_2}{(1-x_2^{-1})(1-x_1x_2^{-1})}.$$

After cancellation,

$$F_0 + F_1 + F_2 = 1 + x_1 + x_2 = f(P, x).$$

Each cone has infinitely many lattice points, but the sum of their rational continuations is a finite polynomial.

8.5 Ehrhart Polynomials and Reciprocity

For an integral polytope P, set $L_P(k) = \left|kP \cap \mathbb{Z}^d\right|$ for $k \in \mathbb{Z}_{>0}$.

Theorem 8.18 (Ehrhart's theorem; Barvinok VIII.5.1). If $P \subset \mathbb{R}^d$ is integral, a polynomial $L_P(t)$ of degree $\dim P$ satisfies

$$L_P(k) = \left|kP \cap \mathbb{Z}^d\right|$$

for every positive integer k.

From Brion's theorem. Put $r = \dim P$. Translate by an integral vertex and choose a basis of the lattice $\mathbb{Z}^d \cap (\mathrm{aff}(P) - v_1)$. This identifies the counting problem with that of a full-dimensional integral polytope $Q \subset \mathbb{R}^r$; the translation of kP by $-kv_1$ preserves its lattice-point count. Let $q_i \in \mathbb{Z}^r$ be the vertices of Q and $K_i = \mathrm{cone}(Q, q_i) - q_i$. Since dilation moves the vertex but not its direction cone,

$$f(kQ, x) = \sum_i x^{kq_i} F(K_i).$$

Substitute $x = e^{\tau c}$ for generic c. Each term has the form $\tau^{-r} e^{\tau k \langle c, q_i \rangle} A_i(\tau)$, where A_i is analytic near zero and independent of k. The constant term at $\tau = 0$ is $|kQ \cap \mathbb{Z}^r| = \left|kP \cap \mathbb{Z}^d\right|$. Expanding the exponential shows that only powers k^j, with $0 \le j \le r$, occur. Comparing fundamental lattice cells with the dilates kP, or equivalently using Riemann sums, gives

$$\lim_{k \to \infty} \frac{\left|kP \cap \mathbb{Z}^d\right|}{k^r} = \mathrm{vol}_r(P) > 0,$$

where the relative volume is normalized so that a fundamental cell has volume one. Thus the polynomial has degree exactly r and leading coefficient $\mathrm{vol}_r(P)$. ⊛

Example 8.19 (Standard simplex). For $\Delta_d = \{x \in \mathbb{R}^d_+ : \sum_i x_i \le 1\}$, adding a slack variable gives

$$L_{\Delta_d}(k) = \binom{k+d}{d}.$$

Its leading coefficient is $1/d! = \mathrm{vol}\, \Delta_d$.

Theorem 8.20 (Ehrhart–Macdonald reciprocity). If $r = \dim P$, then for every positive integer k,

$$L_P(-k) = (-1)^r \left|\mathrm{relint}(kP) \cap \mathbb{Z}^d\right|.$$

This follows from cone reciprocity and Brion localization. The substitution of negative integers

algebraically records lattice points in the relative interior. Moreover,

$$L_P(k) = \operatorname{vol}_r(P)k^r + \text{lower-order terms}, \qquad L_P(0) = 1,$$

with volume normalized by $\mathbb{Z}^d \cap (\operatorname{aff}(P) - v)$ for any integral vertex v. In dimension two,

$$L_P(k) = \operatorname{area}(P)k^2 + \frac{1}{2}\left|\partial P \cap \mathbb{Z}^2\right| k + 1.$$

For rational vertices, $L_P(k)$ is generally a quasipolynomial, polynomial on each residue class modulo a common denominator.

Remark 8.21 (Transportation polytopes). The number of nonnegative integral $d \times d$ matrices whose row and column sums all equal k is the lattice-point count in k times the Birkhoff polytope. It is therefore an Ehrhart polynomial of degree $(d-1)^2$.

8.6 Totally Unimodular Polytopes and Dual Decomposition

Definition 8.22 (Unimodular cones and polytopes). A cone $K = \operatorname{cone}(u_1, \dots, u_k)$ is unimodular if its integral generators form a basis of $\mathbb{Z}^d \cap \operatorname{lin} K$. An integral polytope is *totally unimodular*, in the terminology of this book, when every vertex cone translated to the origin is unimodular.

For a unimodular cone, the fundamental parallelepiped contains only the origin in its relative lattice, so

$$F(K) = \prod_{j=1}^{k} \frac{1}{1 - x^{u_j}}.$$

Corollary 8.23 (Brion formula for a totally unimodular polytope). Let P be an integral polytope with vertices $v_1, \dots, v_n$. If $\operatorname{cone}(P, v_i) = v_i + \operatorname{cone}(u_{i1}, \dots, u_{ik})$ with the generators forming a relative lattice basis, then

$$\sum_{m \in P \cap \mathbb{Z}^d} x^m = \sum_{i=1}^{n} \frac{x^{v_i}}{(1 - x^{u_{i1}}) \cdots (1 - x^{u_{ik}})}.$$

The standard simplex is totally unimodular, as are transportation polytopes with integral margins that have dimension $(m-1)(n-1)$ and are simple. The simplicity condition ensures that their vertex cones are simplicial.

Proposition 8.24 (Rational translate of a unimodular cone; Barvinok VIII.6.2, Problem 1). Let $u_1, \dots, u_d$ be a basis of $\mathbb{Z}^d$, let $K = \operatorname{cone}(u_1, \dots, u_d)$, and let $u_1^*, \dots, u_d^*$ be its dual basis, so $\left\langle u_i, u_j^* \right\rangle = \delta_{ij}$. For $v \in \mathbb{Q}^d$, put

$$w = \sum_{i=1}^{d} \lceil \langle v, u_i^* \rangle \rceil \, u_i.$$

Then

$$F([v + K]) = \frac{x^w}{(1 - x^{u_1}) \cdots (1 - x^{u_d})}.$$

Proof. For an integer point m, the coordinates $\langle m, u_i^* \rangle$ are integers. The condition $m \in v + K$ is equivalent to $\langle m, u_i^* \rangle \geq \langle v, u_i^* \rangle$ for every i. Hence the integer points of $v + K$ are exactly $w + \sum_i \nu_i u_i$ with $\nu_i \in \mathbb{Z}_{\geq 0}$, and the geometric-series formula applies. ⊛

For a full-dimensional rational polytope whose vertex direction cones are unimodular, apply this formula at each vertex and then sum by Brion's theorem. The numerator at a rational vertex v_i is x^{w_i}, with w_i obtained by taking ceilings in that cone's dual coordinates; it is generally not x^{v_i}.

Proposition 8.25 (Dissection into unimodular cones; Barvinok VIII.6.1, Problem 3). Every full-dimensional pointed rational cone admits a finite dissection into full-dimensional unimodular cones that meet along common faces.

Subdivision argument. First triangulate the cone into simplicial rational cones, choosing the same integral generator for a ray wherever it occurs. For a simplicial cone generated by integral $u_1, \dots, u_d$, let $\Delta = |\det[u_1 \ \cdots \ u_d]|$. If $\Delta > 1$, its half-open fundamental parallelepiped contains a nonzero integer point $a = \sum_i \alpha_i u_i$ with $0 \leq \alpha_i < 1$. Subdivide using the ray through a. Each new full-dimensional cone replaces one generator u_i with a, where $\alpha_i > 0$, and has determinant $\alpha_i \Delta$, a positive integer strictly smaller than Δ. Perform the same subdivision in every cone sharing the affected face so that the dissections remain compatible. Repeating this construction reduces the integer determinants until all are one. ⊛

Lemma 8.26 (Polar of a full-dimensional unimodular cone; Barvinok VIII.6.1, Problem 4). If $K = \operatorname{cone}(u_1, \dots, u_d)$ is unimodular, then $K^\circ = \operatorname{cone}(-u_1^*, \dots, -u_d^*)$ is unimodular.

Proof. The dual vectors form a basis of $\mathbb{Z}^d$, and the inequalities $\langle y, u_i \rangle \leq 0$ say precisely that y is a nonnegative combination of the vectors $-u_i^*$. ⊛

Remark 8.27 (Dual decomposition). Work in the linear span of a vertex direction cone K, with its relative lattice, and take the polar in that span. Choose lattice coordinates there. If K is not unimodular, dissect its polar into full-dimensional unimodular cones K_i:

$$[K^\circ] = \sum_i [K_i] + \text{lower-dimensional terms.}$$

After applying polarity, the polars of lower-dimensional cones contain lines and are killed by F, leaving

$$F(K) = \sum_i F(K_i^\circ).$$

Each K_i° is unimodular by the preceding lemma, so its generating function has the product formula above. The polarity valuation of Chapter IV preserves the indicator relation, and F vanishes on the cones containing lines.

Remark 8.28 (Integrality and unimodularity). In linear programming, a totally unimodular constraint matrix explains when a relaxation has an integral optimum for integral right-hand sides. The condition on a polytope used here also requires each vertex cone to be simplicial and generated by a lattice basis, which makes its generating function a monomial divided by geometric-series factors. A polytope with a totally unimodular constraint matrix can fail this cone condition when it is not simple. The polytope of $n \times n$ doubly stochastic matrices, for $n \geq 3$, provides such examples.