MATH 5590 - Convex Geometry Lecture Notes

Libao Jin (ljin1@uwyo.edu)

May 23, 2019

Contents

1 Introduction **1**
1.1 A simple idea with powerful applications . . . 1
1.1.1 Areas . . . 1

2 Affine Spaces and Convex Sets **2**
2.1 Affine Space . . . 2
2.2 Affine Subspace Analogies . . . 2
2.3 Affine hulls . . . 3
2.4 The affine case . . . 3
2.5 Cones . . . 4
2.6 Positive Hulls . . . 4
2.7 Convex Sets . . . 5
2.8 The add-a-dimension trick . . . 6
2.9 Pointed cones . . . 7
2.10 Topology of convexity . . . 8
2.10.1 Supporting hyperplanes and half spaces . . . 9
2.10.2 The Nearest-Point Map . . . 10
2.10.3 Separating hyperplane . . . 10
2.10.4 Positive Polynomials (Application of Carathéodory Theorem) . . . 12
2.11 Helly's Theorem and Applications . . . 16
2.11.1 Geometric Combinatorics. . . . 17
2.11.2 Approximating functions: . . . 17
2.11.3 Polyhedra . . . 22
2.11.4 Important Corollary of Fourier-Motzkin Elimination . . . 24
2.12 Visualizing Polars . . . 28
2.12.1 Cone (convex, linear) . . . 28
2.13 Linear Programming . . . 29
2.13.1 Cones and Sections . . . 30
2.14 Polarity . . . 30
2.14.1 Linear Programming Duality . . . 32

3 Ellipsoids and cones over ellipsoids **37**
3.1 Characterizing ellipsoidal cones . . . 37
3.2 Application – Chronogeometry . . . 37

4 Approximating convex bodies with ellipsoids **38**
4.1 Face lattices of polytopes . . . 43
4.1.1 Euler - Poincaré Relations . . . 45
4.2 h-vectors of simple polytopes . . . 48

5 Convexity and the integer lattice **49**
5.1 Minkowski's Geometry of Numbers 49
5.2 Basic Theory of Lattices in $\mathbb{R}^n$ 50
5.3 Lattice-point enumeratoins in rational polytopes 53

Disclaimer: there might be some typo, use this notes with caution. Also please feel free to let me know if there is anything wrong.

1 Introduction

Definition 1.1. $S \subseteq \mathbb{R}^n$ is convex if $\forall x, y \in S$, the line segment $[x, y]$ is contained in S.

Example 1.1.

1. Line segments.
2. Points.
3. Empty set $\emptyset$.
4. Triangles, Polytopes.
5. Cube.
6. Ellipsoid.

1.1 A simple idea with powerful applications

1.1.1 Areas

- Convexity
  - **Combinatorics (graph theory)**
    - Study the traveling-salesman problem (TSP).
    - The TSP-polytope encodes cycles in graph - optimize with Integer programming.
  - **Operations research (optimization)**
    - Linear/Integer programming
    - Economics: Preference-orderings encoded by convex cones (*Cones and Duality*)
  - **Algebra, Algebraic Geometry and Commutative Algebra**
    - Berstein's Theorem: Given n polynomials $f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n)$ in variables, can count the number of solutions to

$$\begin{cases} f_1(x) = 0 \\ \vdots \\ f_n(x) = 0 \end{cases}$$

      by computing the "mixed volume" of the "Newton polytopes" of $f_1, \ldots, f_n$.
  - (Functional) analysis (*Convex Theory and Its Applications in Functional Analysis*)

Definition 1.2. $K \subseteq \mathbb{R}^n$ is convex iff. $\forall x, y \in K$,

$$x+\lambda(y-x) \in K, \forall 0 \leq \lambda \leq 1 \iff (1-\lambda)x+\lambda y \in K, \forall 0 \leq \lambda \leq 1 \iff \mu x+\lambda y, \forall \lambda, \mu \in [0,1], \lambda+\mu = 1.$$

Definition 1.3. Given $x, y \in \mathbb{R}^n$, call $\lambda x + \mu y$ with $\lambda, \mu \in [0,1], \lambda + \mu = 1$ a convex combination of x and y. In general, given $x_1, x_2, \ldots, x_r \in \mathbb{R}^n$ call $\lambda_1 x_1 + \cdots + \lambda_r x_r$ with $0 \leq \lambda_1, \ldots, \lambda_r \leq 1, \lambda_1 + \cdots + \lambda_r = 1$ a *convex combination*. If drop $\leq$ constraints, this is an *affine combination*.

2 Affine Spaces and Convex Sets

2.1 Affine Space

Recall: $L \subseteq \mathbb{R}^n (L \neq \emptyset)$ is a linear subspace if $\lambda x + \mu y \in L$ for all $x, y \in L$, $\lambda, \mu \in \mathbb{R}$.

Definition 2.1. $A \subseteq \mathbb{R}^n (A \neq \emptyset)$ is an affine subspace if $\lambda x + \mu y \in A$ for all $x, y \in A, \lambda, \mu \in \mathbb{R}, \lambda + \mu = 1$.

Proposition 2.1. $A \subseteq \mathbb{R}^n$ is an affine subspace iff. there exists a linear subspace $L \subseteq \mathbb{R}^n$ and $w \in \mathbb{R}^n$ such that $A = L + w$.

Proposition 2.2. Affine subspaces are closed under *arbitrary affine* combinations: If $x_1, \ldots, x_r \in A$ and $\lambda_1 + \cdots + \lambda_r = 1$, then $\sum_{i=1}^{r} \lambda_i x_i \in A$.

Proof. By induction on r. Have $r = 2$ case by definition of "affine".

Without loss of generality. Some λ_i is not 1, say $\lambda_r \neq 1 \implies \sum_{i=1}^{r-1} \lambda_i \neq 0$. By induction,

$$\frac{\lambda_1}{\sum_{i=1}^{r-1} \lambda_i} x_1 + \cdots + \frac{\lambda_{r-1}}{\sum_{i=1}^{r-1} \lambda_i} x_{r-1} \in A$$

is an affine combination of $r-1$ points in A. OTOH, $\lambda_r = 1 - \sum_{i=1}^{r-1} \lambda_i$, so,

$$\sum_{i=1}^{r-1} \lambda_i x_i + \lambda_r x_r$$

is an affine combination of two points in A, so in A. □

2.2 Affine Subspace Analogies

Table 1: Affine Subspace Analogies

	Linear spaces	Affine spaces
Defining closure property	Closure under linear combinations of pairs	Closure under affine combinations of pairs
The "hull" of $S \subseteq \mathbb{R}^n$ (from the inside)	Set of *all* linear combinations of (arbitrary many) elements of S	Set of *all* affine combinations of (arbitrary many) elements of S
The "hull" of $S \subseteq \mathbb{R}^n$ (from the outside)	The intersection of all linear systems containing S	The intersection of all affine spaces containing S
The "positive" version	Cones: closed under positive under positive (i.e., non-negative) linear combinations	Convex Sets: closed under positive affine combinations

2.3 Affine hulls

Recall: The *linear hull* (aka, linear span) of $S \subseteq \mathbb{R}^n$ is ("from the inside")

$$\text{lin}(S) := \text{span}(S) := \{\lambda_1 x_1 + \cdots + \lambda_r x_r : r \in \mathbb{Z}_{\geq 1}, x_i \in S, \lambda_i \in \mathbb{R}, \forall i\}, \tag{2.1}$$

and ("from the outside")

$$\text{lin}(S) := \bigcap_{L \subseteq \mathbb{R}^n, \text{linear, such that} L \supset S} L. \tag{2.2}$$

Proposition 2.3. Right-hand side of (2.1) is a linear subspace.

Proof. "Obviously" closed under linear combinations of pairs. □

Proposition 2.4. Intersection $\bigcap_{\alpha \in I} L_\alpha$ of linear subspaces $L_\alpha \subseteq \mathbb{R}^n$ is a linear subspace.

Proof. Fix $\lambda, \mu \in \mathbb{R}, x, y \in \bigcap_{\alpha \in I} L_\alpha \implies \forall \alpha \in I, x, y \in L_\alpha \implies \forall \alpha \in I, \lambda x + \mu y \in L_\alpha \implies \lambda x + \mu y \in \bigcap_{\alpha \in I} L_\alpha$. □

Proposition 2.5. RHS of (2.1) = RHS of (2.2).

Proof. $[\supset], [A \cap B \subseteq A]$ RHS of (2.1) is among the sets being intersected on the RHS of (2.2). $[\subset]$, S is contained in the RHS of (2.2) and RHS of (2.2) is closed under arbitrary linear combinations (because it's a linear subspace). □

2.4 The affine case

Definition 2.2. The affine hull (aka, affine span) of $S \subseteq \mathbb{R}^n$ is ("from the inside")

$$\text{aff}(S) := \{\lambda_1 x_1 + \cdots + \lambda_r x_r : r \in \mathbb{Z}_{\geq 1}, x_i \in S, \lambda_i \in \mathbb{R}, \forall i, \sum_{i=1}^{r} \lambda_i = 1\}. \tag{2.3}$$

and ("from the outside")

$$\text{aff}(S) := \bigcap_{A \subseteq \mathbb{R}^n, \text{affine}, A \supset S} A. \tag{2.4}$$

Why the same?

Proposition 2.6. The RHS of (2.3) is an affine subspace.

Proof. Let $x, y \in$ RHS of (2.3), we have

$$x = \lambda_1 x_1 + \cdots + \lambda_r x_r : r \in \mathbb{Z}_{\geq 1}, x_i \in S, \lambda_i \in \mathbb{R}, \forall i, \sum_{i=1}^{r} \lambda_i = 1,$$

$$y = \mu_1 y_1 + \cdots + \mu_s y_s : s \in \mathbb{Z}_{\geq 1}, y_i \in S, \mu_i \in \mathbb{R}, \forall i, \sum_{i=1}^{s} \mu_i = 1.$$

Bookkeeping trick:

$$(z_1, \ldots, z_t) := (x_1, \ldots, x_r, y_1, \ldots, y_s),$$

where $t = r + s$.

$$(\sigma_1, \dots, \sigma_t) := (\lambda_1, \dots, \lambda_r, 0, \dots, 0),$$
$$(\tau_1, \dots, \tau_t) := (0, \dots, 0, \mu_1, \dots, \mu_s),$$

Then we have

$$x = \sigma_1 z_1 + \cdots + \sigma_t z_t,$$
$$y = \tau_1 z_1 + \cdots + \tau_t z_t$$

Given $\lambda, \mu \in \mathbb{R}$ such that $\lambda + \mu = 1$, have $\lambda x + \mu y = (\lambda\sigma_1 + \mu\tau_1)z_1 + \cdots + (\lambda\sigma_t + \mu\tau_t)z_t$. With coeeficients summing as

$$\begin{aligned}(\lambda\sigma_1 + \mu\tau_1) + \cdots + (\lambda\sigma_t + \mu\tau_t) &= \lambda \sum_{i=1}^{t} \sigma_i + \mu \sum_{i=1}^{t} \tau_i \\ &= \lambda \sum_{i=1}^{r} \lambda_i + \mu \sum_{i=1}^{s} \mu_i \\ &= \lambda + \mu \\ &= 1.\end{aligned}$$

□

Proposition 2.7. An intersection $\bigcap_{\alpha \in I} A_\alpha$ of affine subspaces $A_\alpha \subseteq \mathbb{R}^n$ is an affine subspace.

Proof. Same as linear space. □

Proposition 2.8. RHS of (2.3) = RHS of (2.4).

Proof. Same as linear space. □

2.5 Cones

Definition 2.3. A subset $C \subseteq \mathbb{R}^n$ is a *cone* if $\forall x, y \in C$ and $\lambda, \mu \geq 0$, have

$$\lambda x + \mu y \in C.$$

(i.e., cones are closed under positive linear combinations)
Note: C is a cone $\iff$ C is closed under arbitrary positive linear combinations.

2.6 Positive Hulls

Definition 2.4. The positive hull (aka, positive span) of $S \subseteq \mathbb{R}^n$ is

$$\text{pos}(S) := \bigcap_{C \subseteq \mathbb{R}^n, \text{cone}, C \supset S} C \quad \text{(outside)};$$
$$\text{pos}(S) := \{\lambda_1 x_1 + \cdots + \lambda_r x_r : r \in \mathbb{Z}_{\geq 1}, x_i \in S, \lambda_i \geq 0, \forall i\} \quad \text{(inside)}.$$

Proposition 2.9. 1. Intersections of cones are cones.

2. Both RHS's in Definition 2.4 are cones.

3. Both RHS's in Definition 2.4 are equal.

Proof. Same as in linear case. □

Note: C is a cone $\iff$ C is closed under arbitrary positive combinations.

2.7 Convex Sets

Definition 2.5. Let $x, y \in K, \lambda, \mu \geq 0, \lambda + \mu = 1 \implies \lambda x + \mu y \in K$, then K is a *convex set*.

Definition 2.6. The *convex hull* of $S \subseteq \mathbb{R}^n$ is

$$\text{conv}(S) := \{\lambda_1 x_1 + \cdots + \lambda_r x_r : r \in \mathbb{Z}_{\geq 1}, x_i \in S, \lambda_i \geq 0, \forall i, \sum_{i=1}^{r} \lambda_i = 1\} \quad \text{(inside)};$$

$$\text{conv}(S) := \bigcap_{K \subseteq \mathbb{R}^n, \text{convex}, K \supset S} K \quad \text{(outside)}.$$

where $\lambda_1 x_1 + \cdots + \lambda_r x_r$ are all convex combinations of elements of S.

Proposition 2.10.

1. Intersections of convex sets are convex.
2. Both RHS's in Definition 2.6 are convex.
3. Both RHS's in Definition 2.6 are equal.

Proof. Same as in affine case. □

Proposition 2.11. K is a convex $\iff$ K is closed under arbitrary convex combinations.

Proof. Same as in affine case. □

Example 2.1. $S = \{e_1, e_2, e_3\} \subseteq \mathbb{R}^3$ (standard basis vectors).

- $\text{lin}(S) = \mathbb{R}^3$.
- $\text{aff}(S) =$ the plane containing (the heads of) e_1, e_2, e_3.
- $\text{pos}(S) =$ (closed)1^{st} octant.
- $\text{conv}(S) =$ the triangle contains the heads of e_1, e_2, e_3 as the vertices.

Remark 2.1. $\text{conv}(S)$ is all combinations of elements that are both affine and positive. *Question*: $\text{conv}(S) = \text{pos}(S) \cap \text{aff}(S)$. *Ans*: Yes, in previous example. Yes, in general, if $0 \notin \text{aff}(S)$. Maybe false, if $0 \in \text{aff}(S)$.

Lemma 2.1. Suppose $0 \notin \text{aff}(S)$. If $x \in S$ and $\gamma x \in S (\gamma \in \mathbb{R})$, then $\gamma = 1$.

Proof. Suppose, TGAC, have $x, \gamma x \in S$, with $\gamma \neq 1$. WTS (want to see), $0 \in \text{aff}(S)$. So WTS 0 is an affine combination of elements of S.

$$1(\gamma x) - \gamma x = 0,$$

which is linear but affine, because the coefficients sum to $1 - \gamma$, so rescale. Hence,

$$0 = \frac{1}{1-\gamma}\gamma x + \frac{-\gamma}{1-\gamma} x \in \text{aff}(S).$$

□

2.8 The add-a-dimension trick

$\mathbb{R}^n \to \mathbb{R}^{n+1} : \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \mapsto \begin{bmatrix} x_1 & \cdots & x_n & 1 \end{bmatrix}$. Given $S \subseteq \mathbb{R}^n$, can "lift to height 1" in $\mathbb{R}^{n+1}$. $S' := \{\begin{bmatrix} x \\ 1 \end{bmatrix} \in \mathbb{R}^{n+1} : x \in S\}$.

The moral: Many *affine* (e.g., convexity) properties of S can be proved more easily by looking at *linear properties* of $\text{pos}(S')$ (avoid that "bookkeeping trick" needed for linear)

Recall: Defining closure properties involve only two elements. But hulls require considering arbitrary combinations.

Theorem 2.1 (Carathéodory's Theorem for Cones)**.** Let $S \subseteq \mathbb{R}^n$ and $v \in \text{pos}(S)$. Then $\exists T \subseteq S, |T| = n$ such that $v \in \text{pos}(T)$.

Proof. Let r be minimum such that $\exists x_1, \ldots, x_r \in S$ such that $v \in \text{pos}(T)$, where $T = \{x_1, \ldots, x_r\}$. [WTS $r \leq n$] $\implies$ $\alpha_1, \ldots, \alpha_r \geq 0$ such that $v = \alpha_1 x_1 + \cdots \alpha_r x_r$. Suppose TGAL, $r \geq n+1$, that implies there exists linear dependent $\exists \beta_1, \ldots, \beta_r$ such that $\beta_1 x_1 + \cdots + \beta_r x_r = 0$, for some $\beta_i \neq 0$. We have $\forall \lambda \in \mathbb{R}$, $\lambda\beta_1 x_1 + \cdots + \lambda\beta_r x_r = 0$, then

$$v = (\alpha_1 + \lambda\beta_1)x_1 + \cdots + (\alpha_r + \lambda\beta_r)x_r.$$

Apply Lemma 2.2 to

$$a = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, b = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

to find $\lambda \in \mathbb{R}$ such that $a + \lambda b \in \mathbb{R}^r_{\geq 0}$ with some coordinate zero. Get v expressed as a positive combination of elements that are less than r of S. □

Lemma 2.2. $\forall x \in \mathbb{R}^n_{\geq 0}, \forall v \in \mathbb{R}^n \backslash \{0\}, \exists \lambda \in \mathbb{R}$ such that some coordinate of $x + \lambda v$ is 0, but $x + \lambda v \in \mathbb{R}^n_{\geq 0}$.

Proof. Exercise. □

Theorem 2.2 (Carathéodory's Theorem for Convex Sets)**.** Let $S \subseteq \mathbb{R}^n$ and $v \in \text{conv}(S)$. Then $\exists T \subseteq S, |T| = n+1$ such that $v \in \text{conv}(T)$.

Proof. Use the add-a-dimension trick. Let $v' = \begin{bmatrix} v \\ 1 \end{bmatrix}$, $S' = \{\begin{bmatrix} x \\ 1 \end{bmatrix} : x \in S\} \subseteq \mathbb{R}^{n+1}$. Then $v \in \text{conv}(S) \implies v' \in \text{pos}(S')$. Then by Theorem 2.1, we have $T' \subseteq S'$, $|T'| = n+1$ such that $v' \in \text{pos}(T')$. Let $T' = \{x'_1, \ldots, x'_{n+1}\}$, where $x'_i = \begin{bmatrix} x_i \\ 1 \end{bmatrix}$, we have

$$v' = \alpha_1 x'_1 + \cdots + \alpha_{n+1} x'_{n+1}, \quad \alpha_i \geq 0.$$

By last coordinates, done. Or If had $\alpha_1 + \cdots + \alpha_n = 1$, done. Else, $\exists \gamma \in \mathbb{R}$ such that $\gamma\alpha_1 + \cdots + \gamma\alpha_{n+1} = 1, \gamma \neq 1 \implies x', \gamma x' \in \text{aff}(S')$, but $0 \notin \text{aff}(S')$. By Lemma from 2.2, we have $\gamma = 1$. □

2.9 Pointed cones

Remark 2.2. Define *dimension* of an affine space $A \subseteq \mathbb{R}^n$ is $\dim(L)$, where $A = L + w$ for some $w \in \mathbb{R}^n$.

Definition 2.7. A *hyperplane* $H \subseteq \mathbb{R}^n$ is an $(n-1)$-dimensional affine subspace of $\mathbb{R}^n$. Equivalently, H is a *fiber* of some linear functional: $H := \{u \in \mathbb{R}^n : c(u) = \alpha\} =: H_{c,\alpha}$ for some $c \in (\mathbb{R}^\kappa)^*$, i.e., $c : \mathbb{R}^n \to \mathbb{R}$ and $\alpha \in \mathbb{R}$. Get half-spaces,

$$H^+_{c,\alpha} = \{u \in \mathbb{R}^n : c(u) \geq \alpha\},$$
$$H^-_{c,\alpha} = \{u \in \mathbb{R}^n : c(u) \leq \alpha\}.$$

Definition 2.8. A cone C is *pointed* if $\exists c \in (\mathbb{R}^n)^*$ such that

$$C \subseteq H^+_{c,\alpha}, \quad \alpha \in \mathbb{R}, \text{ and } C \cap H_{c,\alpha} = \{v\}, \quad v \in C,$$

where v is the *apex* of C.

Note: for linear cones (only ones so far), the apex is 0.

Theorem 2.3 (Radon's Theorem for Cones)**.** Suppose $S \subseteq \mathbb{R}^n \backslash \{0\}$, $|S| \geq n+1$, $\operatorname{pos}(S)$ is pointed. Then

(a) There exists a partition $\{S_1, S_2\}$ of S such that $\operatorname{pos}(S_1) \cap \operatorname{pos}(S_2) \neq \{0\}$, where S_1 and S_2 satisfy the following conditions $S_1, S_2 \neq \emptyset, S_1 \cap S_2 = \emptyset, S_1 \cup S_2 = S$.

(b) This partition is unique if and only if $|S| = n+1$ and every n-subseteq is linearly independent.

Proof.

(a) $|S| \geq \dim \mathbb{R}^n$, that implies we have a linear dependency, i.e.,

$$\sum_{x \in S} \alpha_x x = 0, \quad \alpha_x \in \mathbb{R} \text{ not all zero.} \tag{2.5}$$

Let $P := \{x \in S : \alpha_x > 0\}, N := \{x \in S : \alpha_x < 0\}$. Choose $Z_1, Z_2 \subseteq S$ such that $Z_1 \cup Z_2 = \{x \in S : \alpha_x = 0\}, Z_1 \cap Z_2 = \emptyset$. Let $S_1 := P \cup Z_1$, $S_2 := N \cup Z_2$. Certainly, we have $S_1 \cap S_2 = \emptyset, S_1 \cup S_2 = S$ and S_1 and S_2 not both empty.

Claim: S_1 and S_2 are both not empty. Proof: S is pointed, so $\exists$ linear $c : \mathbb{R}^n \to \mathbb{R}$ such that $c(x) > 0, \forall x \in S$ (because $0 \notin S$). By (2.5), we have

$$0 = c(0) = \sum_{x \in S} \alpha_x c(x) = \underbrace{\sum_{x \in P} \alpha_x c(x)}_{\leq 0} + \underbrace{\sum_{x \in N} \alpha_x c(x)}_{\geq 0},$$

and not both 0, so neither is.

(b) Note: $|S| \geq n+1$, so $|S| = n+1$ and every n-subseteq is linear independent if and only if every proper subseteq is linearly independent. $\Longrightarrow$: (By contrapositive) Suppose that we have a proper linearly dependent subseteq of S. $\exists$ a linear dependency (2.5) with some $\alpha_x = 0$. This means, Z_1, Z_2 aren't both empty. That implies we can move between Z_1 and Z_2 to

get different partitions S_1 and S_2. $\Longleftarrow$: Suppose every proper subseteq of S is linearly independent, then $|S|= n+1$. Let $S = \{x_1, \dots, x_{n+1}\} \subseteq \mathbb{R}^n$. We have a linear operator $A : \mathbb{R}^{n+1} \to \mathbb{R}^n, \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n+1} \end{bmatrix} \mapsto \sum_{i=1}^{n+1} \beta_i x_i$.

- A is onto, because, e.g., $x_1, \dots, x_n$ span $\mathbb{R}^n$, that implies $\dim(\ker A) = 1$. Then $\ker A$ is spanned by some nonzero vector

$$a := \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_{n+1} \end{bmatrix} \quad \text{with} \sum_{i=1}^{n+1} \alpha_i x_i = 0.$$

- Every n-subseteq is linearly independent, $\alpha_i \neq 0, \forall i$. That implies $S_1 := \{x_i, \alpha_i > 0\}$ and $S_2 := \{x_i : \alpha_i < 0\}$ is the only partition. [A different partition would given an element of $\ker A$ not in span$\{a\}$, because it would have its $+/-$ coordinates n different entries.]

□

Theorem 2.4 (Radon's Theorem for Convex Sets)**.** Suppose $S \subseteq \mathbb{R}^n$, $|S|\geq n+2$, then

(a) There exists a strict partition $\{S_1, S_2\}$ of S with $\text{conv}(S_1) \cap \text{conv}(S_2) \neq \emptyset$.

(b) This partition is unique if and only if $|S|= n+2$ and every $n+1$-subseteq is affinely independent.

Proof. Apply Theorem 2.3 to $S' = \left\{ \begin{bmatrix} x \\ 1 \end{bmatrix} : x \in S \right\} \subseteq \mathbb{R}^{n+1}$. □

Definition 2.9. $S \subseteq \mathbb{R}^n$ is *affinely independent* if $\forall x \in S, x \notin \text{aff}(S \backslash \{x\})$.

Proposition 2.12. Suppose $S \subseteq \mathbb{R}^n$, then the following are equivalent:

(a) S is affinely independent.

(b) If $S' := \left\{ \begin{bmatrix} x \\ 1 \end{bmatrix} \in \mathbb{R}^{n+1} : x \in S \right\}$, then S' is linearly independent.

(c) $\lambda_1 x_1 + \cdots + \lambda_r x_r = 0, x_i \in S, \lambda_1 + \cdots + \lambda_r = 0 \implies \lambda_i = 0, \forall i$.

2.10 Topology of convexity

$\mathbb{R}^n$ as a Euclidean Space, we have inner production $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. And we have properties as follows:

- bilinear
- positive definite ($\langle x, x \rangle \geq 0$ with $\langle x, x \rangle = 0 \iff x = 0$)
- symmetric ($\langle x, y \rangle = \langle y, x \rangle$)
- "Law of cosines": $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle$.
- Cauchy-Schwarz: $|\langle x, y \rangle| \leq \|x\| \|y\|$ with $<$ unless x and y are parallel.

- Triangle inequality: $\|x+y\| \leq \|x\|+\|y\|$ with $<$ unless x and y are parallel.

Any basis determines an inner product such that the basis is orthogonal, i.e.,

$$\langle b_i, b_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$

Then with $\langle \cdot, \cdot \rangle$, can identify each $x \in \mathbb{R}^n$ with $x^* \in (\mathbb{R}^n)^*$ by $x^*(y) := \langle x, y \rangle$. Fact: every linear functional $c \in (\mathbb{R}^n)^*$ is x^* for some $x \in \mathbb{R}^n$ (Riesz Representation Theorem). In this case, can compute "dot product",

$$\left\langle \sum_i \alpha_i b_i, \sum_j \beta_j b_j \right\rangle = \sum_i \alpha_i \beta_i.$$

Then we have the norm

$$\|x\| = \langle x, x \rangle^{1/2}.$$

Norm gives a topology on $\mathbb{R}^n$. Recall

- $S \subseteq \mathbb{R}^n$ is *open* if and only if $\forall x \in S, \exists, \varepsilon > 0$ such that $B_\varepsilon(x) \subseteq S$, where $B_\varepsilon(x) := \{y \in \mathbb{R} : \|x-y\| < \varepsilon\}$ is the *open* ball of radius ε centered at x.
- $S \subseteq \mathbb{R}^n$ closed if and only if $\mathbb{R}^n \backslash S$ is open, i.e., if S has points arbitarily close to x, then $x \in S$.
- The *closure* of $S \subseteq \mathbb{R}^n$

$$\overline{S} := \bigcap_{C \subseteq \mathbb{R}^n \text{ closed}, \, C \supseteq S} C.$$

- The boundary ∂S of $S \subseteq \mathbb{R}$ is a set of points $x \in \mathbb{R}^n$ such that $\forall \varepsilon > 0$, $B_\varepsilon(x) \cap S \neq \emptyset$ and $B_\varepsilon(x) \cap S^{\complement} \neq \emptyset$, where $S^{\complement} = \mathbb{R} \backslash S$.
- The interior S° of S is $S^\circ := \{x \in S : \exists \varepsilon > 0, B_\varepsilon(x) \subseteq S\}$.

Every affine subspace $A \subseteq \mathbb{R}^n$ inherits the topology whose open sets are $B_\varepsilon(x) \cap A$ for $x \in A$. Given $S \subseteq \mathbb{R}^n$, have *relative* closure, boundary, interior (etc.) defined by looking at closure (etc.) with respect to $\text{aff}(S)$.

2.10.1 Supporting hyperplanes and half spaces

Let $K \subseteq \mathbb{R}^n$ be a closed convex subset. We write $H_{x,\alpha}, H^+_{x,\alpha}, H^-_{x,\alpha}$ for $H_{x^*,\alpha}, H^+_{x^*,\alpha}, H^-_{x^*,\alpha}$

Definition 2.10. $H_{x,\alpha}$ is a *supporting hyperplane* of K if and only if

$$K \cap H_{x,\alpha} \neq \emptyset, \quad K \subseteq H^-_{x,\alpha}.$$

The halfspace containing K is a supporting halfspace.

Recall: If $K \subseteq \mathbb{R}^n$ is convex if and only if

$$K = \bigcap_{L \subseteq \mathbb{R}^n \text{ convex}, L \supseteq K} L.$$

Theorem 2.5. Suppose $K \subseteq \mathbb{R}^n$ is closed and convex, then

$$K = \bigcap_{H^- \subseteq \mathbb{R}^n \text{ a supporting halfspace of } K} H^-.$$

2.10.2 The Nearest-Point Map

Recall:

- $S \subseteq \mathbb{R}^n$ is compact if and only if S is closed and bounded;
- If S is compact and $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous, the $f(S)$ is compact.
- If S is compact and $f : S \to \mathbb{R}$ is continuous, then f attains its minimum and maximum on S (Extreme Value Theorem).

Theorem 2.6. Suppose $K \subseteq \mathbb{R}^n$ is closed and convex, suppose $x \in \mathbb{R}^n \backslash K$, then $\exists! \, x' \in K$ such that

$$\|x - x'\| = \inf_{y \in K} \|x - y\|.$$

Proof. Existence: Intersect K with sufficient big $\overline{B_\varepsilon(x)}$, $K \cap B_\varepsilon(x)$ is compact. So by Extreme Value THeorem, since $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is continuous, a closest point exists.
Uniqueness: Use isoceles triangle. Suppose we have $x' \neq x''$ in K, both with $\|x - x'\| = \|x - x''\|$ minimal. Take a look at the midpoint of x' and x'', we have $m := (x' + x'')/2$. By convexity, midpoint $m \in K$. Also $x - x'$, $x - x''$ are not parellel. By strict triangle inquality,

$$\|x - m\| = \|x - \frac{1}{2}(x' + x'')\| = \frac{1}{2}\|(x - x') + (x - x'')\| < \frac{1}{2}(\|x - x'\| + \|x - x''\|) = \|x - x'\|.$$

Contradicts the minimality. □

Use to define the nearest-point map $p_K : \mathbb{R}^n \to K$.

2.10.3 Separating hyperplane

Theorem 2.7. Suppose $K \subset \mathbb{R}^n$ is closed, convex, $x \in \mathbb{R}^n \backslash K$. Then hyperplane perpendicular to $x - p_K(x) =: x'$m then x' separates K from x: K does not intersect the open halfspace containing x.

Proof. Let $\alpha := \langle x - x', x' \rangle$, $H := H_{x-x',\alpha} \ni x'$, $H^+ := H^+_{x-x',\alpha} \backslash H \ni x$, WTS: $K \cap H^+ = \emptyset$ (i.e., H supports). Suppose, TGAC, $\exists y \in K \cap H^+$. Then (claim), $\exists z \in [x', y]$, with $z \neq x'$, but $z \in B_{\|x-x'\|}(x)$, which follows from Lemma 2.3. □

Lemma 2.3. If $\langle x, y \rangle > 0$, then $\delta \in (0, 1)$ such that $\|x - \delta y\| < \|x\|$.

Proof. WTS: $\|x - \delta y\|^2 < \|x\|^2$:

$$\|x\|^2 + \delta^2 \|y\|^2 - 2\delta \langle x, y \rangle < \|x\|^2,$$

i.e., WTS $0 < 2\delta \langle x, y \rangle - \delta^2 \|y\|^2$. That is, WTS

$$f(\delta) > 0 \text{ for some } \delta \in (0, 1),$$

where $f(t) := -\|y\|^2 t^2 + 2\langle x, y \rangle t = t(2\langle x, y \rangle - \|y\|^2 t)$. □

Theorem 2.8. Suppose $K \subseteq \mathbb{R}^n$ is closed, convex. Then

$$K = \bigcap_{H^- \subseteq \mathbb{R}^n \text{ a supporting hyperplane of } K} H^-.$$

Proof. $\subseteq$: $K \subseteq \bigcap_{K' \text{ convex}, K' \supseteq K} K' \subseteq \bigcap_{H^- \subseteq \mathbb{R}^n} H^-$.
$\supseteq$: Given $x \in \mathbb{R}^n \backslash K$, we have $H^- := H_{x-x',\alpha}$, (as previous proof) with $x \notin H^- \supseteq K \implies x \notin$ RHS intersection. □

Rem: Often don't need *all* supporting hyperplanes. For triangle, need only 3.

Proposition 2.13. Suppose $K \subset \mathbb{R}^n$ closed and convex, then $p_K : \mathbb{R}^n \to K$ is continuous.

Proof. For $x, y \in \mathbb{R}^n \backslash K$, we have nearest points $x' := p_K(x), y' := p_K(y)$ and supporting hyperplanes H_x and H_y with K on opposite side from x with respect to y. That implies

$$\langle x - x', y' - x' \rangle \leq 0,$$
$$\langle y - y', x' - y' \rangle \leq 0.$$

Use bilinearity to add:

$$\begin{aligned} \langle (x-y) - (x'-y'), y' - x' \rangle \leq 0 &\implies \langle x - y, y' - x' \rangle - \langle x' - y', y' - x' \rangle \leq 0 \\ &\implies \langle x' - y', x' - y' \rangle \leq \langle x - y, x' - y' \rangle \\ &\implies \|x' - y'\|^2 \leq \langle x - y, x' - y' \rangle \leq \|x - y\| \|x' - y'\|. \end{aligned}$$

So, either $p_K(x) = p_K(y)$ or $\|x' - y'\| \leq \|x - y\|$. Hence, as $x \to y$, we have $p_K(x) = x' \to y' = p_K(y)$, hence prove the continuity. □

Question: Suppose $S \subseteq \mathbb{R}^n$, does S closed imply $\text{conv}(S)$ closed? No.

Example 2.2. Let $S = \{(x, y) : y \geq \frac{1}{x^2}, x \neq 0\}$, which is closed. But the $\text{conv}(S) = \{(x, y) : y > 0\}$ is not closed.

But "unboundedness" is the only problem here.

Theorem 2.9. If $S \subseteq \mathbb{R}^n$ is compact, then $\text{conv}(S)$ is compact.

Proof. By Theorem 2.2,

$$\text{conv}(S) = \left\{ \sum_{i=1}^{n+1} \lambda_i s_i : s_i \in S, 0 \leq \lambda_i \leq 1, \sum_{i=1}^{n+1} \lambda_i = 1, i = 1, \ldots, n+1 \right\}.$$

Let

$$T := \left\{ (\lambda_1, \ldots, \lambda_{n+1}) \in [0,1]^{n+1} : \sum_{i=1}^{n+1} \lambda_i = 1 \right\} \implies T = \text{conv}\{e_1, \ldots, e_{n+1}\} \subseteq \mathbb{R}^{n+1},$$

which is *the standard n-simplex* and compact. Define $\Phi : T \times S^{n+1} \to \text{conv}(S)$, $\Phi : (\lambda_1, \ldots, \lambda_{n+1}; s_1, \ldots, s_{n+1}) \mapsto \sum_{i=1}^{n+1} \lambda_i s_i$, which is an onto map (by Theorem 2.2). Both T and S are compact, which implies $T \times S^{n+1}$ is compact, and Φ is continuous, then $\Phi(T \times S^{n+1})$ is continuous. □

2.10.4 Positive Polynomials (Application of Carathéodory Theorem)

Fix n to be the number of variables, $\mathbb{R}[x_1, \ldots, x_n] = \mathbb{R}[\mathbf{x}]$ — Algebra of Polynomials. Notation:

$$f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}] \implies \sum_{a=(\alpha_1,\ldots,\alpha_n)\in\mathbb{Z}^n_{\geq 0}} \lambda_a x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

where $\mathbf{x}^a = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ with only finitely many $\lambda_a \neq 0$. For $b = (\beta_1, \ldots, \beta_n) \in \mathbb{R}^n$.

$$f(b) = \sum_{a\in\mathbb{Z}^n_{\geq 0}} \lambda_a \beta_1^{\alpha_1} \cdots \beta_n^{\alpha_n} \in \mathbb{R}.$$

Definition 2.11. $f(x) \in \mathbb{R}[\mathbf{x}]$ is *positive* (resp. nonnegative) if $f(a) > 0$ (resp. $f(a) \geq 0$) for $a \neq 0$.

How to tell that $f(\mathbf{x})$ is nonnegative? Clearly,

$$f(\mathbf{x}) = \sum_{i=1}^{m} (g_i(\mathbf{x}))^2, g_i(\mathbf{x}) \in \mathbb{R}[\mathbf{x}] \implies f(\mathbf{x}) \text{ is nonnegative.}$$

But the converse is not true.

Example 2.3. $f(x, y, z) = z^6 + x^4y^2 + x^2y^4 - 3x^2y^2z^2$ is nonnegative, but not a sum of squares.

Hilbert's 17^{th} problem:

$$f(\mathbf{x}) \text{ is nonnegative} \iff f(\mathbf{x}) = \sum_{i=1}^{m} (r_i(\mathbf{x}))^2$$

for some $r_i \in \mathbb{R}[x_1, \ldots, x_n]$, i.e., rational functions $p(\mathbf{x})/q(\mathbf{x}), p(\mathbf{x}), q(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$. Artin (1927): Yes! Can "certify" nonnegativity this way.
Another approach to "certifying" positivity. Note: Every $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ is a sum of homogeneous polynomials.
Further reading: Hilbert's seventeenth problem

Definition 2.12. $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ is homogeneous of degree d if

$$f(\lambda a) = \lambda^d f(a) \forall a \in \mathbb{R}^d, \lambda \in \mathbb{R}.$$

Write $H_{d,n}$ for the vector space of such.

Example 2.4.

(a) $z^6 + x^4y^2 + x^2y^4 - 3x^2y^2z^2 \in H_{6,3}$.

(b) $\|\mathbf{x}\|^{2k} = \left(\sum_{i=1}^{n} x_i^2\right)^k \in H_{2k,n}$.

(c) $2x + 3y - 6z \in H_{1,3}$, where $H_{1,3} = (\mathbb{R}^3)^*$.

(d) $\sum_{a\in\mathbb{Z}^n_{\geq 0}} \lambda_a \mathbf{x}^a \in H_{d,n}\backslash\{0\} \iff \lambda_a \neq 0 \implies \sum_{i=1}^{n} \alpha_i = d$. $\{\mathbf{x}^a : a \in \geq 0^n \text{ and } \sum_i \alpha_i = d\} =: B$ is a basis of $H_{d,n} \implies \dim H_{d,n} = \binom{n+d-1}{d}$.

(e) Get Euclidean inner product on $H_{d,n}$ by declaring $\mathbf{x}^a$ to be O.N., then get topology, hence "closed", "boundary", "compact", etc.

Fix $d = 2k, k \in \mathbb{Z}_{\geq 0}$, then work in $H_{d,n} = H_{2k,n}$.

Theorem 2.10. Suppose $p(\mathbf{x}) \in H_{2k,n}$. Then $p(\mathbf{x})$ is positive if and only if can write

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

where $c_1(\mathbf{x}), \ldots, c_m(\mathbf{x}) \in H_{1,n}, s \in \mathbb{Z}_{\geq 1}, m \leq \dim H_{2k,n} + 1$.

Proof. Given $f(\mathbf{x}) \in H_{2k,n}$, define formal differential operator $f(\partial)$ as follows: if $f(\mathbf{x}) = \sum_{a \in \mathbb{Z}^n_{\geq 0}} \lambda_a x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, then

$$f(\partial) := \sum_{a \in \mathbb{Z}^n_{\geq 0}} \lambda_a \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}.$$

Apply to both sides

$$\|\mathbf{x}\|^{2k} = \sum_{i=1}^{m} (c_i(\mathbf{x}))^{2s}, c_i(\mathbf{x}) = \sum_{j=1}^{n} \gamma_{ij} x_j.$$

On RHS $\forall f(\mathbf{x}) \in H_{2k,n}$,

$$\begin{aligned}
f(\partial)(c_i(\mathbf{x}))^{2s} &= \sum_{a \in \mathbb{Z}^n_{\geq 0}} \lambda_a \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \left(\sum_{j=1}^{n} \gamma_{ij} x_j \right)^{2s} \\
&= \sum_{a \in \mathbb{Z}^n_{\geq 0}} \lambda_a \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n - 1}}{\partial x_n^{\alpha_n - 1}} 2s \left(\sum_{j=1}^{n} \gamma_{ij} x_j \right)^{2s-1} \gamma_{in} \\
&\vdots \\
&= \sum_{a \in \mathbb{Z}^n_{\geq 0}} \lambda_a \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_{n-1}}}{\partial x_n^{\alpha_{n-1}}} \frac{(2s)!}{(2s - \alpha_n)!} \left(\sum_{j=1}^{n} \gamma_{ij} x_j \right)^{2s-\alpha_n} \gamma_{in}^{\alpha_n} \\
&\vdots \\
&= \sum_{a \in \mathbb{Z}^n_{\geq 0}} \lambda_a \frac{(2s)!}{(2s-2k)!} \left(\sum_{j=1}^{n} \gamma_{ij} x_j \right)^{2s-2k} \gamma_{i1}^{\alpha_1} \cdots \gamma_{in}^{\alpha_n} \\
&= \frac{(2s)!}{(2s-2k)!} \left(\sum_{j=1}^{n} \gamma_{ij} x_j \right)^{2s-2k} \sum_{a \in \mathbb{Z}^n_{\geq 0}} \lambda_a \gamma_{i1}^{\alpha_1} \cdots \gamma_{in}^{\alpha_n} \\
&= \frac{(2s)!}{(2s-2k)!} (c_i(\mathbf{x}))^{2s-2k} f(\gamma_{i1}, \ldots, \gamma_{in}).
\end{aligned}$$

On the LHS of

$$\begin{aligned}
f(\partial)\|\mathbf{x}\|^{2s} &= \sum_{a\in\mathbb{Z}^n_{\geq 0}} \lambda_a \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} (x_1^2 + \cdots x_n^2)^s \\
&= \sum_{a\in\mathbb{Z}^n_{\geq 0}} \lambda_a \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n - 1}}{\partial x_n^{\alpha_n - 1}} s(x_1^2 + \cdots x_n^2)^{s-1} 2x_n \\
&= \sum_{a\in\mathbb{Z}^n_{\geq 0}} \lambda_a \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n - 2}}{\partial x_n^{\alpha_n - 2}} [s(s-1)(x_1^2 + \cdots x_n^2)^{s-2} 4 * x_n^2 + s(x_1^2 + \cdots x_n^2)^{s-1} 2] \\
&\vdots \\
&= \sum_{a\in\mathbb{Z}^n_{\geq 0}} \lambda_a \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_{n-1}}}{\partial x_n^{\alpha_{n-1}}} [\frac{s!}{(s-\alpha_n)!} (x_1^2 + \cdots x_n^2)^{s-\alpha_n} 2^{\alpha_n} * x_n^{\alpha_n} + s(x_1^2 + \cdots x_n^2)^{s-1} 2] \\
&\vdots \\
&= \sum_{a\in\mathbb{Z}^n_{\geq 0}} \lambda_a \frac{2^k s!}{(s-2k)!} (x_1^2 + \cdots x_n^2)^{s-k} \Phi_s(f)(\mathbf{x}),
\end{aligned}$$

where $\Phi_s : H_{2k,n} \to H_{2k,n}, f(\mathbf{x}) \mapsto f(\mathbf{x}) + O(\frac{1}{s})$. In particular $\Phi_s \to I$ in $\mathcal{L}(H_{2k,n})$, which is the space of linear operators. In particular, Φ_s is invertible for $s \gg 0$. Note: positive polynomials are an open set in $H_{2k,n}$. If let $q(\mathbf{x}) := \Phi_s^{-1}(p(\mathbf{x}))$, then, for $s \gg 0$, $q(\mathbf{x})$ is positive, and, putting $q(\mathbf{x})$ for $f(\mathbf{x})$, get $\frac{2^{2k}s!}{(s-2k)!}\|\mathbf{x}\|^{2s-2k}$, and

$$p(\mathbf{x}) = \frac{(2s)!}{(2s-2k)!} \sum_{i=1}^{m} q(\gamma_{i1}, \ldots, \gamma_{in}) c_i(\mathbf{x})^{2s-2k}.$$

Push positive constants "under" the even exponent on $c_i(\mathbf{x})$.

□

Let $\mathcal{O}_n$ be the orthogonal group (w.r.t. standard inner product on $\mathbb{R}^n$).

$$\{U\ M_n(\mathcal{R}) : U^T U + I\}.$$

"isometries": $\|Ua\| = \|a\|, \forall a \in \mathbb{R}^n, \langle Ua, Ub\rangle = \langle a, b\rangle$.

- $\mathcal{O}_n$ acts on $\mathbb{R}^n \implies \mathcal{O}_n$ acts on functions $\varphi : \mathbb{R}^n \to \mathbb{R}$ via $(U\varphi)(Ua) = \varphi(a), \forall a \in \mathbb{R}^n$. Then $(U\varphi)a) = \varphi(U^{-1}a)$.
- Action restricts to a group representation of $\mathcal{O}_n$ over $H_{2k,n}$, i.e., $H_{2k,n}$ is invariant under the action of $\mathcal{O}_n$: $f \in H_{2k,n} \implies f(\lambda a) = \lambda^{2k} f(a) \implies (Uf)(\lambda a) = f(U^{-1}\lambda a) = \lambda^{2k} f(U^{-1}a) = \lambda^{2k}(Uf)(a) \implies Uf \in H_{2k,n}$.
- $\|\mathbf{x}\|^{2k} \in \mathrm{Fix}(\mathcal{O}_n \curvearrowright H_{2k,n}) := \{f \in H_{2k,n} : Uf = f, \forall U \in \mathcal{O}_n\}$.

Proposition 2.14. $\mathrm{Fix}(\mathcal{O}_n \curvearrowright H_{2k,n}) = \mathrm{span}_{\mathbb{R}}\{\|\mathbf{x}\|^{2k}\}$.

Proof. Suppose $p(x) \in \mathrm{Fix}(\mathcal{O} \curvearrowright H_{2k,n}), a \in \mathbb{R}^n, \|a\| = 1$. Define $S^{n-1} := \partial B_1(0) = \{b \in \mathbb{R}^n : \|b\| = 1\}, \gamma := p(a), q(\mathbf{x}) = p(x) - \gamma\|\mathbf{x}\|^{2k}$. WTS: $q(\mathbf{x}) = 0$. Have $q(\mathbf{x}) \in \mathrm{Fix}(\mathcal{O}_n \curvearrowright H_{2k,n})$. $q(a) = \gamma - \gamma = 0$. Let $b \in S^{n-1}$. WTS $q(b) = 0$. Have $U_b \in \mathcal{O}_n$ such that $U_b^{-1} b = a$. $q(b) = q(U_b^{-1}a) = (U_b q)(a) = q(a) = 0$. $q \in H_{2k,n} \implies q(\lambda b) = \lambda^d q(b) = 0$, where $b \in S^{n-1}, \lambda \in \mathbb{R} \implies p(\mathbf{x}) = \lambda\|\mathbf{x}\|^{2k}$. □

Lemma 2.4. We can write

$$\|\mathbf{x}\|^{2k} = \sum_{i=1}^{m} (c_i(\mathbf{x}))^{2k},$$

for some $c_1(\mathbf{x}), \dots, c_m(\mathbf{x}) \in H_{1,n} = (\mathbb{R}^n)^*$ with $m \le \dim H_{2k,n} + 1$.

Proof. Let $S^{n-1} := \{c \in (\mathbb{R}^n)^* : \|c\| = 1\}$ be the unit sphere. $\forall c \in S^{n-1}$, let $p_c(\mathbf{x}) := (c(\mathbf{x}))^{2k}$, $p_-(\mathbf{x}) : S^{n-1} \to H_{2k,n}$,

$$\mathcal{K} := \operatorname{conv}\{p_c(\mathbf{x}) : c \in S^{n-1}\}.$$

$c \mapsto p_c(\mathbf{x})$ is a continuous map of compact set S^{n-1} into $H_{2k,n}$, continuous images of compact set are compact, convex hulls of compact set are compact. That implies $\mathcal{K}$ is compact.
Digression: Recall Riemann integration.

(a) Integrating $\varphi : [0,1] \to \mathbb{R}$,

$$\int_{[0,1]} \varphi(a)\, da := \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \varphi(a_{Nj}),$$

where $a_{N1}, \dots, a_{NN}$ are "mesh points" chosen "uniformly".

(b) Integrating $\varphi : D \to \mathbb{R}$, where D is a domain of measure 1.

$$\int_{D} \varphi(a)\, da = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \varphi(a_{Nj}),$$

where mesh points a_{Nj} are "uniform".

Claim 2.1. $\mathcal{K}$ contains $\lambda\|\mathbf{x}\|^{2k}$ for some $\lambda > 0$.

Proof. Let

$$p(\mathbf{x}) := \int_{S^{n-1}} P_c(\mathbf{x})\, dc = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} P_{c_{Nj}}(\mathbf{x}),$$

where $c_{N1}(\mathbf{x}), \dots, c_{NN}(\mathbf{x})$ are "mesh points" on the unit sphere S^{n-1} chosen "unitformly" and hence rotationally symmetric in the limit $N \to \infty$. Therefore, $p(\mathbf{x}) \in \operatorname{Fix}(\mathcal{O}_n \curvearrowright H_{2k,n}) \implies p(\mathbf{x}) = \lambda\|\mathbf{x}\|^{2k}$ for some $\lambda \in \mathbb{R}$. □

Clean-up:

(a) $p(\mathbf{x}) \in \mathcal{K}$: the Riemann sums are convex combinations of elements of $\mathcal{K}$, hence the sume (for each N) is in $\mathcal{K}$. And $\mathcal{K}$ is compact (hence closed), then limit of $p(\mathbf{x})$ is in $\mathcal{K}$.

(b) $p_c(\mathbf{x})$ are even powers, hence nonnegative. $p(\mathbf{x})$ is a limit of convex combination of such, that gives $p(\mathbf{x})$ is nonnegative, hence $\lambda \ge 0$.

(c) $\lambda \ne 0$, because $p(\mathbf{x}) = 0 \implies c(a) = 0$ for almost all $a \in \mathbb{R}^n$. Hence $\lambda > 0$.

Finishing-up: WTS $\|\mathbf{x}\|^{2k} = \sum_{i=1}^m c_i(\mathbf{x})^{2k}$. Saw: $\exists \lambda > 0$ such that

$$\lambda \|x\|^{2k} \in \operatorname{conv}\{p_c(\mathbf{x}) : c \in S^{n-1}\} \subset H_{2k,n}, p_c(\mathbf{x}) = c(\mathbf{x})^{2k}.$$

Then apply Carathéodory's theorem: $\exists \bar{c}_1(\mathbf{x}), \cdots, \bar{c}_m(\mathbf{x})$, $m = \dim H_{2k,n} + 1$, such that $\|\mathbf{x}\|^{2k} = \sum_{i=1}^m \frac{\alpha_i}{\lambda} \bar{c}_i(\mathbf{x})^{2k}$. Set

$$c_i(\mathbf{x}) := (\sqrt[2k]{\alpha_i/\lambda}\bar{c}_i(\mathbf{x}))^{2k}.$$

□

Exercise 2.1. (a) $n = 2, k = 2$, we have

$$(x^2 + y^2)^2 = \frac{1}{6}(x+y)^4 + \frac{1}{6}(x-y)^4 + \frac{2}{3}x^4 + \frac{2}{3}y^4.$$

(b) $n = 3, k = 2$, we have

$$(x^2 + y^2 + z^2)^2 = \sum_{\pm} \frac{1}{12}(x \pm y \pm z)^4 + \frac{2}{3}x^4 + \frac{2}{3}y^4 + \frac{2}{3}z^4.$$

(c) Note: Can be a vector-valued integral.

2.11 Helly's Theorem and Applications

Recall Radon's Theorem 2.4.

Theorem 2.11 (Helly)**.** Suppose $A_1, \ldots, A_m \subset \mathbb{R}^d$, convex, $m \geq d+1$ and every $d+1$ of the A_i's has a nonempty intersection. Then

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

Theorem 2.12. Suppose $\mathcal{K}$ is a finite family of convex subsets $K \subset \mathbb{R}^d, d+1 \leq |\mathcal{K}| < \infty$. Suppose $\forall \mathcal{J} \subset \mathcal{K}$ with $|\mathcal{J}| \leq d+1$,

$$\bigcap_{K \in \mathcal{J}} K \neq \emptyset.$$

Then

$$\bigcap_{K \in \mathcal{K}} K \neq \emptyset.$$

Proof. Trivial if $|\mathcal{K}| = d+1$. Proceed by induction on $|\mathcal{K} \geq d+2$, then $\forall K \in \mathcal{K}, \exists P_K \in \mathbb{R}^d$ such that by induction applied $\mathcal{K} \backslash \{K\}$

$$P_K \in \bigcap_{J \in \mathcal{K}, J \neq K} J \neq \emptyset.$$

Have Radon partition of $\{P_K : K \in \mathcal{K}\}$ $(d+2$ points):

$$\emptyset \neq \mathcal{R}, \mathcal{B} \subset \mathcal{K}, \mathcal{R} \cup \mathcal{B} = \mathcal{K}, \mathcal{R} \cap \mathcal{B} = \emptyset,$$

such that

$$\exists P \in \operatorname{conv}\{P_R : R \in \mathcal{R}\} \cap \operatorname{conv}\{P_B : B \in \mathcal{B}\}.$$

□

Claim 2.2. $P \in \bigcup_{K\in\mathcal{K}} K$.

Proof. $\forall R \in \mathcal{R}$, we have $P_R \in \bigcup_{J\in\mathcal{K},J\neq R} J \subset \bigcup_{B\in\mathcal{B}} B \implies \text{conv}\{P_R : R \in \mathcal{R}\} \subset \bigcap_{B\in\mathcal{B}} B$. Likewise $\text{conv}\{P_B : B \in \mathcal{B}\} \subset \bigcap_{R\in\mathcal{R}} R \implies P \in \left(\bigcap_{R\in\mathcal{R}} R\right) \cap \left(\bigcap_{B\in\mathcal{B}} B\right) = \left(\bigcap_{K\in\mathcal{K}} K\right)$. □

Proposition 2.15. Need $|\mathcal{K}| < \infty$. But works for $|\mathcal{K} = \infty$ if all $K \in \mathcal{K}$ are compact.

2.11.1 Geometric Combinatorics.

Definition 2.13. A hyperplane $H_{c,\alpha} = \{x \in \mathbb{R}^d : \langle c, x\rangle = \alpha\}$ separates $R, B \subset \mathbb{R}^d$ if $R \subset H^-_{c,\alpha} := \{x \in \mathbb{R}^d : \langle c, x\rangle < \alpha\}$, $B \subset H^-_{c,\alpha} := \{x \in \mathbb{R}^d : \langle c, x\rangle > \alpha\}$.

Question: When can $R, B \subset \mathbb{R}^d$ be separated?

Theorem 2.13. Suppose $R, B \subset \mathbb{R}^d$, $R \cap B = \emptyset$, $|R \cup B| < \infty$. Suppose $\forall S \subset R \cup B$, $|S| \leq d + 2$. There exists a hyperplane separating $S \cap R$, $S \cap B$, then there exists a hyperplane separating R, B.

Proof. Seek hyperplane $H_{c,\alpha}$ separating every $r \in R$ from every $b \in B$.

$$\forall r \in R, K_r := \{(c, \alpha) \in \mathbb{R}^{d+1} : r \in H^-_{c,\alpha}\}, \forall b \in B, K_b := \{(c, \alpha) \in \mathbb{R}^{d+1} : b \in H^+_{c,\alpha}\}.$$

Seek $(c, \alpha) \in \left(\bigcap_{r\in R} K_r\right) \cap \left(\bigcap_{b\in B} K_b\right)$, then $H_{c,\alpha}$ does the job. Have

- $r \in R, b \in B$, K_r and K_b are convex: $\forall (c_1, \alpha_1), (c_2, \alpha_2) \in K_r, 0 \leq t_1, t_2 \leq 1, t_1 + t_2 = 1$. Have $\langle t_1c_1 + t_2c_2, r\rangle = t_1\langle c_1, r\rangle + t_2\langle c_2, r\rangle < t_1\alpha_1 + t_2\alpha_2$.
- Every $d + 2$ subfamily of $\{K_r : r \in R\} \cup \{K_b : b \in B\}$ has nonempty intersection. So Helly yields the claim. (Have to patch the possibility that the common point has form $c = 0$)

□

Example 2.5. $d = 2$: if can't separate R from B in $\mathbb{R}^2$, then there exists 4-point obstruction.

2.11.2 Approximating functions:

Fix $\varepsilon > 0$, X a set. Given $g : X \to \mathbb{R}$ and functions $f_1, \ldots, f_m : X \to \mathbb{R}$, seek to approximate g with some

$$f := \sum_{i=1}^{m} \alpha_i f_i, \alpha_i \in \mathbb{R},$$

such that $|g(x) - f(x)| \leq \varepsilon, \forall x \in X$.

Theorem 2.14. Suppose $g : X \to \mathbb{R}$, $0 \leq |x| < \infty$, $\mathcal{F} := \{f_1, \ldots, f_m\}$, $f_i : X \to \mathbb{R}$. Suppose $\forall S \subset X$ such that $|S| \leq m + 1$, $\exists f \in \text{lin}\,\mathcal{F}$ ($\mathbb{R}^k$ for some $k \leq m$) such that $|g(x) - f(x)| \leq \varepsilon$, $\forall x \in S$. Then $\exists f \in \text{lin}\,\mathcal{F}$ such that

$$|g(x) - f(x)| \leq \varepsilon, \forall x \in X.$$

Proof. Let $K_x := \{f \in \text{lin}\,\mathcal{F} : |g(x) - f(x)| \leq \varepsilon\}, \forall x \in X$.

Remark 2.3.

- $K_x \subset \mathbb{R}^k$ for $k \leq m$.
- $\forall S \subset X$, $|S| \leq m + 1$, $\bigcap_{x\in S} K_x \neq \emptyset$.

K_x is convex, suppose for fixed x, $|g(x) - f_1(x)| \leq \varepsilon$, $|g(x) - f_2(x)| \leq \varepsilon$. $\alpha f_1(x) + (1-\alpha) f_2(x)$ can approach to $g(x)$, where $0 \leq \alpha \leq 1$. Then Helly gives $f \in \bigcap_{x \in X} K_x$, as sought.

□

Generalizing to $|X| = \infty$. Need "compact" version of Helly's:

Theorem 2.15. Suppose $\mathcal{K}$ is a family of compact convex subsets $K \subset \mathbb{R}^d$. Suppose $\forall \mathcal{J} \subset \mathcal{K}$ with $|\mathcal{J}| \leq d+1$, $\bigcap_{K \in \mathcal{J}} K \neq \emptyset$, then

$$\bigcap_{K \in \mathcal{K}} K \neq \emptyset.$$

Proof. If $\bigcap_{K \in \mathcal{K}} K = \emptyset$, then there exists finite subfamily $\mathcal{K}' \subset \mathcal{K}$ such that

$$\bigcap_{K \in \mathcal{K}'} K = \emptyset.$$

Apply finite Helly. □

Theorem 2.16. Suppose $g : X \to \mathbb{R}$, $|X| > 0$ (maybe infinite), $\mathcal{F} := \{f_1, \ldots, f_m\}$, $f_i : X \to \mathbb{R}$. Suppose "regularity": $\exists T \subset X, |T| < \infty$ such that

$$\forall f \in \operatorname{lin} \mathcal{F}, f|_T \equiv 0 \implies f \equiv 0.$$

I.e., there exists a projection $P_T : \mathbb{R}^X \to \mathbb{R}^T$, $f \mapsto f|_T$ such that $\operatorname{lin} \mathcal{F} \cap \ker P_T = \{0\}$, where $\ker P_T$ is a finite co-dimensional subspace of $\mathbb{R}^X$. That is, not only is $\operatorname{lin} \mathcal{F} \subset \mathbb{R}^X$ finite dimensional, but also intersects finite co-dimensional "coordinate subspace" trivially. Suppose $\forall S \subset X$ such that $|S| \leq m+1$, $\exists f \in \operatorname{lin} \mathcal{F}$ such that $|g(x) - f(x)| \leq \varepsilon, \forall x \in S$. Then

$$\exists f \in \operatorname{lin} \mathcal{F}, |g(x) - f(x)| \leq \varepsilon, \forall x \in X.$$

Proof. (sketch) With regularity, $\bigcap_{x \in T} K_x$ are closed and also bounded subsets of the finite dimensional $\operatorname{lin} \mathcal{F} \subset \mathbb{R}^X \rightsquigarrow K_x$ are compact, by Helly's. To show bounded, note have "∞-norm" on $\mathbb{R}^T$, hence on $\operatorname{lin} \mathcal{F}$:

$$\|f\|_\infty := \max\{|f(t)| : t \in T\}.$$

By regularity, $\|f\|_\infty > 0$ for $f \neq 0.$, homogeneity: $\|\lambda f\|_\infty = |\lambda| \|f\|_\infty$. Get $\|f\|_\infty \leq R/\delta, \forall f \in K$, where $R := \varepsilon + \max\{|g(t) : t \in T\}$ and $\delta := \min\{\|f\|_\infty : \|f\|_2 = 1\}$, where $f := \sum_{i=1}^m \alpha_i f_i$. □

Towards an algorithmic theory of lattice-point enumeration. The algebra generated by compact (resp. closed) convex sets and the Euler characteristic.

Definition 2.14. Given $S \subset \mathbb{R}^d$, the *indicator function* of S is

$$[S] : \mathbb{R}^d \to \mathbb{R}, x \mapsto \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}.$$

Definition 2.15. The algebra generated by closed (resp. compact) subsets of $\mathbb{R}^d$ is

$$\mathcal{C}(\mathbb{R}^d) := \operatorname{span}_{\mathbb{R}}\{[C] : C \subset \mathbb{R}^d \text{ closed, convex}\}$$
$$\mathcal{K}(\mathbb{R}^d) := \operatorname{span}_{\mathbb{R}}\{[K] : K \subset \mathbb{R}^d \text{ compact, convex}\}$$

where $\operatorname{span}_{\mathbb{R}}$ means finite linear combinations.

Remark 2.4.

- $[S] \in \mathbb{R}^{(\mathbb{R}^d)} := \{f : \mathbb{R}^d \to \mathbb{R}\}$ — algebra (vector space with multiplication), $\forall f, g \in [S], \lambda, \mu \in \mathbb{R}, a \in \mathbb{R}^d$,

$$(\lambda f + \mu g)(a) := \lambda f(a) + \mu g(a),$$
$$(fg)(a) := f(a)g(a).$$

- $0 := [\emptyset]$, $1 = [\mathbb{R}^d]$, $[S^{\complement}] = 1 - [S]$.
- $[S][T] = [S \cap T]$.
- $[S \cup T] = [S] + [T] - [S \cap T]$.

Caution: $\{[C] : C$ closed, convex$\}$ spans $\mathcal{C}(\mathbb{R}^d)$, but is not a basis:

$$[C] + [D] - [S \cap T] - [S \cup T] = 0$$

is a *linear dependency*.

Proposition 2.16 (Inclusion/Exclusion)**.** Suppose $A_1, \ldots, A_n \subset \mathbb{R}^d$ are convex and closed (resp. compact). Then $[A_1 \cup \cdots \cup A_m] \in \mathcal{C}(\mathbb{R}^d)$ (resp. $\mathcal{K}(\mathbb{R}^d)$) via (letting $I := \{1, \ldots, m\}$)

$$\left[\bigcup_{i \in I} A_i\right] = \sum_{J \subset I, J \neq \emptyset} (-1)^{|J|-1} \left[\bigcap_{j \in J} A_j\right].$$

Proof. By DeMorgan's law,

$$\left[\bigcup_{i \in I} A_i\right] = \left[\left(\bigcap_{i \in I} A_i^{\complement}\right)^{\complement}\right] = 1 - \prod_{i \in I}(1 - [A_i]) = 1 - \sum_{J \subset I}(-1)^{|J|} \prod_{j \in J}[A_i].$$

By convention, $\prod_{j \in \emptyset}[A_i] = 1$. □

"Recall" Euler characteristic from topology, let HB be a solid handle body, then

$$\chi(HB) = 1 - (\text{number of handles}).$$

If HB is a solid ball, then $\chi(HB) = 1$, where $\chi : \mathcal{C}(\mathbb{R}^d) - \mathbb{R}$ is linear.

Definition 2.16. A linear functional $v : \mathcal{C}(\mathbb{R}^d) \to \mathbb{R}$ is called a valuation. (Note: $\mathcal{K}(\mathbb{R}^d) \subset \mathcal{C}(\mathbb{R}^d)$), call $v : \mathcal{K}(\mathbb{R}^d) \to \mathbb{R}$ a valuation, too). Note: valuations satisfies inclusion/exclusion

$$v\left(\left[\bigcup_{i \in I} A_i\right]\right) = \sum_{J \subset I, J \neq \emptyset} (-1)^{|J|-1} v\left(\left[\bigcap_{j \in J} A_j\right]\right).$$

Theorem 2.17 (Euler characteristic)**.** There exists a unique valuation

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

such that $\chi([C]) = 1, \forall C \subset \mathbb{R}^d$ nonempty, closed, convex.

Proof. The hard part: proving that $\chi([C]) = 1$ doesn't overdetermine χ (a possibility because the $[C]$ are linearly dependent).
The easy part: Uniqueness. Suppose $f \in \mathcal{C}(\mathbb{R}^d) \rightsquigarrow f = \sum_{i=1}^m \lambda_i [C_i]$ for some $\lambda \in \mathbb{R}, C_i \subset \mathbb{R}^d$ nonempty, closed, convex. By linearity,

$$\chi(f) = \sum_{i=1}^m \lambda_i \chi([C_i]) = \sum_{i=1 \text{ such that } C_i \neq \emptyset}^m \lambda_i.$$

Back to hard part: Existence. Trivial if $d = 0 : \mathbb{R}^d = \{0\}$. Then $\{[\{0\}]\}$ really is a basis. $\mathcal{K}(\mathbb{R}^0) = \mathcal{C}(\mathbb{R}^0) \cong \mathbb{R}$. For $d \geq 1$, proceed by induction. First construct $\chi : \mathcal{K}(\mathbb{R}^d) \to \mathbb{R}$, then will extend to $\mathcal{C}(\mathbb{R}^d)$. Note: suppose $K \subset \mathbb{R}^d$ is compact, convex. Then $\mathcal{K}$ has a unique "minimal height", i.e., $\forall \tau \in \mathbb{R}$, let $H_\tau := \{a \in (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d : \alpha_d = \tau\}$, there exists a unique $\tau \in \mathbb{R}$ such that $H_\tau \cap \mathcal{K} \neq \emptyset$, but $H_{\tau - \varphi} \cap K = \emptyset$ for all sufficient $\varepsilon > 0$. Need convexity, closedness. Indeed, need compactness.
$H_\tau \cong \mathbb{R}^{d-1}$, so by induction, have Euler characteristic

$$\chi_\tau : \mathcal{K}(H_\tau) \to \mathbb{R}.$$

For $f \in \mathcal{K}(\mathbb{R}^d)$, let $f_\tau : \sum_{i=1}^m \lambda_i [K_i \cap H_\tau]$, where $f =: \sum_{i=1}^m \lambda_i [K_i]$.

$$f_\tau \in \mathcal{K}(H_\tau) \rightsquigarrow \chi_\tau(f_\tau) = \sum_{i : K_i \cap \neq \emptyset} \lambda_i.$$

$$\lim_{\varepsilon \to 0^+} \chi_{\tau - \varepsilon}(f_{\tau - \varepsilon}) = \sum_{i : K_i \cap H_{\tau - \varepsilon} \neq \emptyset} \lambda_i.$$

Consider

$$\begin{aligned}
\chi_\tau(f) - \lim_{\varepsilon \to 0^+} \chi_{\tau-\varepsilon}(f_{\tau-\varepsilon}) &= \sum_{i : K_i \cap H_\tau \neq \emptyset} \lambda_i - \lim_{\varepsilon \to 0^+} \sum_{i : K_i \cap H_{\tau-\varepsilon} \neq \emptyset} \lambda_i \\
&= \lim_{\varepsilon \to 0^+} \left(\sum_{i : K_i \cap H_\tau \neq \emptyset, K_i \cap H_{\tau-\varepsilon} = \emptyset} \lambda_i - \sum_{i : K_i \cap H_\tau = \emptyset, K_i \cap H_{\tau-\varepsilon} \neq \emptyset} \lambda_i \right) \\
&= \lim_{\varepsilon \to 0^+} \sum_{i : K_i \cap H_\tau \neq \emptyset, K_i \cap H_{\tau-\varepsilon} = \emptyset} \lambda_i \\
&= \sum_{i : \tau \text{ is the unique minimal height among points in } K_i} \lambda_i.
\end{aligned}$$

Note: this expression is independent of how f is reprensented as a linear combination of the $[K]$. Because K_i is *compact*: $\forall \varepsilon > 0$ sufficient small, if $K_i \cap H_\tau = \emptyset$, then $K_i \cap H_{\tau - \varepsilon} = \emptyset$, so the second term is 0. Because every nonempty K_i appearing in $f = \sum \lambda_i [K_i]$ has a unique minimal height, can define

$$\chi(f) := \sum_{\tau \in \mathbb{R}} [\chi_\tau(f) - \lim_{\varepsilon \to 0^+} \chi_{\tau - \varepsilon}(f_{\tau - \varepsilon})].$$

Note: only finitely many nonzero terms in this sum, so well-defined, and

$$\chi(f) = \sum_{i : K_i \neq \emptyset} \lambda_i.$$

Now extend χ to $\mathcal{C}(\mathbb{R}^d)$

$$\chi(f) = \lim_{\rho\to\infty} \chi\left(\sum_{i=1}^{n} \lambda_i [K_i \cap B_\rho(0)]\right).$$

Limit is well-defined because value stabilizes as soon as ρ is big enough so that $B_\rho(0)$ meets every nonempty K_i. □

Definition 2.17. Call this χ Euler characteristic.

Eden: For vector space V, W and a *basis* $B = \{b_i : i \in I\}$, can extend any function $\varphi : B \to W$ to a linear transformation $\Phi : V \to W$ such that

$$\Phi(v) =: \Phi\left(\sum_{i\in I} \lambda_i b_i\right) := \sum_{i\in I} \lambda_i \varphi(b_i). \tag{2.6}$$

The key: the representation $v = \sum_{i\in I} \lambda_i b_i$ of $v \in V$ as a linear combination of B is unique, so (2.6) is well-defined.
Original sin: $\mathcal{K}(\mathbb{R}^n) := \text{span}\{[K] : K \subset \mathbb{R}^n \text{ compact convex}\}$ but *not* as a basis ($[A] + [B] = [A \cup B] - [A \cap B]$). But often want to construct a well-defined valuation $\mathcal{K}(\mathbb{R}^n) \to W$ from just an assignment $\varphi([K]) \in W$ on the $[K]$.
Salvation:

Theorem 2.18. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Then there exists an "extension"

$$\mathcal{T} : \mathcal{K}(\mathbb{R}^n) \to \mathcal{K}(\mathbb{R}^m), [K] \mapsto [T(K)], \forall K \subset \mathbb{R}^n \text{ compact convex},$$

such that $\mathcal{T}([K]) = [T(K)]$.

Proof. (working backwards).
Want. Given $f = \sum_i \lambda_i [K_i] \in \mathcal{K}(\mathbb{R}^n)$ to define

$$\mathcal{T}(f) = \sum_i \lambda_i [T(K_i)] \tag{2.7}$$

Problem is representation is not unique, so we must prove that (2.7) defines a function.
Observe.

$$\begin{aligned}
(2.7) &\iff \forall y \in \mathbb{R}^m, \mathcal{T}(f)(y) = \sum_i \lambda_i [T(K_i)](y) \\
&\iff \forall y \in \mathbb{R}^m, \mathcal{T}(f)(y) = \sum_{i: y\in T(K_i)} \lambda_i \\
&\iff \forall y \in \mathbb{R}^m, \mathcal{T}(f)(y) = \sum_{i: K_i \cap T^{-1}(y) \neq \emptyset} \lambda_i \\
&\iff \forall y \in \mathbb{R}^m, \mathcal{T}(f)(y) = \chi\left(\sum_i \lambda_i [K_i \cap T^{-1}(y)]\right) \\
&\iff \forall y \in \mathbb{R}^m, \mathcal{T}(f)(y) = \chi\left(\sum_i \lambda_i [K_i][T^{-1}(y)]\right) \\
&\iff \forall y \in \mathbb{R}^m, \mathcal{T}(f)(y) = \chi\left([T^{-1}(y)] f\right).
\end{aligned}$$

where $[T(K_i)](y) = \begin{cases} 1 & y \in T(K_i) \\ 0 & \text{otherwise} \end{cases}$, $[K_i \cap \in T^{-1}(y)] \in \mathcal{K}(\mathcal{R}^n)$. The point: RHS is independent of the representation. So the function I is well-defined. Get linearity "for free" from (2.7). □

Remark 2.5. χ itself is the extension of the unique linear transformation $\mathbb{R}^n \to \mathbb{R}^0$.

Question: Why not for $\mathcal{C}$ in place of $\mathcal{K}$ (Have x on $\mathcal{C}(\mathbb{R}^n)$). Image of *compact, convex* K under T is compact, convex. But image of merely closed (convex) C is not necesarrily closed. But can extend reesult to certain unbounded closed convex sets.

Example 2.6. $C := \{(x, y) \in \mathbb{R}^2 : y \geq 1/x, x > 0\}$, $T : \mathbb{R}^2 > \mathbb{R}^1, (x, y) \mapsto (x)$, $T(C) = (0, \infty)$.

2.11.3 Polyhedra

Recall that $\forall C \subset \mathbb{R}^n$ closed convex,

$$C = \bigcap_{\overline{H^-} \subset \mathbb{R}^n, \text{ closed half-space}, \overline{H^-} \supset C} \overline{H^-},$$

where $\overline{H^-} := \{x \in \mathbb{R}^n : \langle a, x \rangle \leq \beta, a \in \mathbb{R}^n, \beta \in \mathbb{R}\}$.

Definition 2.18. A *polyhedron* is the intersection *finitely many* closed half-spaces. For $P \subset \mathbb{R}^n$ a polyhedron, can write

$$P = \{x \in \mathbb{R}^n : \langle a_i, x \rangle \leq \beta_i, a_i \in \mathbb{R}^n, \beta_i \in \mathbb{R}, i = 1, \ldots, m\}.$$

or

$$\{x \in \mathbb{R}^n : Ax \leq b\}, \text{ where } A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} \in M_{m \times n}(\mathbb{R}), b = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \in \mathbb{R}^m.$$

Remark 2.6.

1. Polyhedra are closed, convex, but may be unbounded.
2. Polyhedra can be bounded.

Example 2.7.

(a) $P = \left\{ x \in \mathbb{R}^2 : \begin{bmatrix} -1 & -3 \\ -1 & 0 \\ -1 & 1 \end{bmatrix} x \leq \begin{bmatrix} -9 \\ -3 \\ 3 \end{bmatrix} \right\}$.

(b) $P = \left\{ x \in \mathbb{R}^2 : \begin{bmatrix} -1 & -3 \\ -1 & 0 \\ -1 & 1 \\ 1 & 1 \\ 2 & -1 \end{bmatrix} x \leq \begin{bmatrix} -9 \\ -3 \\ 3 \\ 15 \\ 18 \end{bmatrix} \right\}$.

Recall: For $C \subset \mathbb{R}^n$ closed convex, $T : \mathbb{R}^n \to \mathbb{R}^m$ linear, $T(C) \subset \mathbb{R}^m$ may *not* be closed. But

Theorem 2.19. For $P \subset \mathbb{R}^n$ a polyhedron, $T : \mathbb{R}^n \to \mathbb{R}^m$, *do* have that $T(P)$ is a polyhedron.

Proof. First, the case in which $T : \mathbb{T}^n \to \mathbb{R}^m$ is 1-1. Let $P := \{x \in \mathbb{R}^n : Ax \leq b\}$. Then $T(P) = \{y \in \mathbb{R}^m : \exists x \in \mathbb{R}^m \text{ such that } Tx = y \text{ and } Ax \leq b\}$. T is 1-1, so have $T^{-1} : \operatorname{im} T \to \mathbb{R}^n$ and

$$\begin{aligned} T(P) &= \{y \in \mathbb{R}^m : AT^{-1}y \leq b \text{ and } y \in \operatorname{im} T\} \\ &= \{y \in \mathbb{R}^m : AT^{-1}y \leq b \text{ and } \langle c, y\rangle = 0, \forall c \in B\}, \end{aligned}$$

where B is a basis of $(\operatorname{im} T)^{\perp} = \operatorname{kernel} T^*$. For T not (necessarily) 1-1, need a Lemma 2.5. □

Lemma 2.5. Let $\mathrm{pr}_n : \mathbb{R}^n \to \mathbb{R}^{n-1}$ be the projection map $(\xi_1, \ldots, \xi_{n-1}, \xi_n) \mapsto (\xi_1, \ldots, \xi_{n-1})$. Then $\mathrm{pr}_n(P)$ is a polyhedron.

Proof. Uses the Fourier-Motzkin Elimination algorithm. Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, where

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}, a_i = \begin{bmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \vdots \\ \alpha_{in} \end{bmatrix}, \alpha_{ij} \in \mathbb{R}, b = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}, \beta_i \in \mathbb{R}.$$

Then $y \in \mathrm{pr}_n(P)$ if and only if the 1-dim fiber over y meets P. It suffices to show that y is in every horizontal-facing half-space and y meets *every* downward-facing half-space *before* meeting *any* upward-facing half-space. So let $I_0 := \{i : \alpha_{in} = 0\}$ (horizontal-facing), $I_- := \{i : \alpha_{in} < 0\}$ (downward-facing), and $I_+ := \{i : \alpha_{in} > 0\}$ (upward-facing). Then

$$\begin{aligned} \mathrm{pr}(P) &= \left\{(\xi_1, \ldots, \xi_{n-1}) \in \mathbb{R}^{n-1} : \exists \xi_n \in \mathbb{R}, \sum_{j=1}^{n} \alpha_{ij}\xi_j \leq \beta_i, 1 \leq i \leq m\right\} \\ &= \left\{(\xi_1, \ldots, \xi_{n-1}) \in \mathbb{R}^{n-1} : \exists \xi_n \in \mathbb{R}, \sum_{j=1}^{n-1} \alpha_{ij}\xi_j \leq \beta_i, i \in I_0; \sum_{j=1}^{n-1} \alpha_{ij}\xi_j + \alpha_{in}\xi_n \leq \beta_i, i \in I_- \cup I_+,\right\} \end{aligned}$$

The goal is to eliminate ξ_n.

$$\xi_n \geq \frac{1}{\alpha_{in}}\left(\beta_i - \sum_{j=1}^{n-1} \alpha_{ij}\xi_j\right), \forall i \in I_-, \xi_n \geq \frac{1}{\alpha_{kn}}\left(\beta_k - \sum_{j=1}^{n-1} \alpha_{kj}\xi_j\right), \forall k \in I_+.$$

Note: $\exists \lambda$ such that $\mu \leq \lambda \leq \nu$ if and only if $\mu \leq \nu$. This leads to

$$\begin{aligned} \mathrm{pr}_n(P) = \Bigg\{&(\xi_1, \ldots, \xi_{n-1}) \in \mathbb{R}^{n-1} : \sum_{j=1}^{n} \alpha_{ij}\xi_j \leq \beta_i, \forall i \in I_0, \\ &\frac{1}{\alpha_{in}}\left(\beta_i - \sum_{j=1}^{n-1} \alpha_{ij}\xi_j\right) \leq \frac{1}{\alpha_{kn}}\left(\beta_k - \sum_{j=1}^{n-1} \alpha_{kj}\xi_j\right), \forall i \in I_-, \forall k \in I_+\Bigg\}. \end{aligned}$$

□

Theorem 2.20. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then $T(P)$ is a polyhedron.

Proof. Let $\hat{T} : \mathbb{R}^n \to \mathbb{R}^{m+n}, x \mapsto (T(x), x)$. $\hat{T}$ is linear and $1-1$. Last time, we've shown that $\hat{T}(P)$ is a polyhedron. Then

$$T(x) = \mathrm{pr}_{n+1} \circ \mathrm{pr}_{n+2} \circ \cdots \circ \mathrm{pr}_{n+m}(\hat{T}(x)) \implies T(P) = \mathrm{pr}_{n+1} \circ \mathrm{pr}_{n+2} \circ \cdots \circ \mathrm{pr}_{n+m}(\hat{T}(P)).$$

Then $T(P)$ is a polyhedron. □

Definition 2.19. $\mathcal{P}(\mathbb{R}^n) := \mathrm{span}\{[P] : P \subset \mathbb{R}^n \text{ a polyhedron}\}$.

Corollary 2.1. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear. Then $\exists \mathcal{T} : \mathcal{P}(\mathbb{R}^n) \to \mathcal{P}(\mathbb{R}^m)$ such that $\mathcal{T}([P]) = [T(P)]$.

Proof. Same as $\mathcal{K}(\mathbb{R}^n)$. □

Remark 2.7. Sets of the form

$$\left\{ x \in \mathbb{R}^n : \begin{cases} \langle a_1, x\rangle & \le \beta_1 \\ \vdots & \\ \langle a_r, x\rangle & \le \beta_r \\ \langle a_1', x\rangle & = \gamma_1' \\ \vdots & \\ \langle a_s', x\rangle & = \gamma_s' \end{cases} \right\}$$

are polyhedra, because

$$\langle a', x\rangle = \gamma \iff \begin{cases} \langle a', x\rangle \le \gamma \\ \langle a', x\rangle \ge \gamma \end{cases} \iff \begin{cases} \langle a', x\rangle \le \gamma \\ \langle -a', x\rangle \le -\gamma \end{cases}$$

2.11.4 Important Corollary of Fourier-Motzkin Elimination

Definition 2.20. $P \subset \mathbb{R}^n$ is a polytope if $P = \mathrm{conv}(V)$ for some finite $V \subset \mathbb{R}^n$.

Theorem 2.21. The polytopes are precisely the bounded polyhedra.

In particular:

Proposition 2.17. Every polytope is a bounded polyhedron.

Proof. Suppose $P = \mathrm{conv}\{v_1, \ldots, v_m\} \subset \mathbb{R}^n$ is a polytope, which is clearly bounded. Then

$$P = \{\sum_{i=1}^m \lambda_i v_i : \lambda_i \ge 0, \forall i, \sum_{i=1}^m \lambda_i = 1\}$$

Note: $\forall x := (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$, $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$,

$$x = \sum_{i=1}^m \lambda_i v_i \iff x - \sum_{i=1}^m \lambda_i v_i = 0. \iff \begin{bmatrix} I_n & v_1 & \cdots & v_m \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \\ \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix} = 0.$$

Then

$$P = \mathrm{pr}\left(\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \\ \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix}\right) = \mathrm{pr}(\mathbf{x}),$$

where $\mathrm{pr} = \mathrm{pr}_{n+1} \circ \mathrm{pr}_{n+2} \circ \cdots \circ \mathrm{pr}_{n+m}$, and $\mathbf{x} \in \mathbb{R}^{n+m}$. Then P is the projected image of a polyhedron, thus P is a polyhedron. □

Proposition 2.18. Every bounded polyhedron is a polytope.

More generally,

Proposition 2.19. If P is a polyhedron, then there exists a polytope Q and a linear cone C such that $P = Q + C$, where $Q + C$ is the Minkowski sum.

Goal:

(a) The bounded polyhedron are the polytopes.

(b) Indeed the polyhedra in $\mathbb{R}^n$ are the subsets $P \subset \mathbb{R}^n$ such that

$$P = \mathrm{conv}(V) + \mathrm{pos}(R), \exists V, R \subset \mathbb{R}^n,$$

where V, R finite.

Theorem 2.22. Suppose $P \subset \mathbb{R}^n$ is a polyhedron. Then P is a "V-polyhedron", i.e., of the form

$$P = \mathrm{conv}(V) + \mathrm{pos}(R), \exists V, R \subset \mathbb{R}^n.$$

Proof. Note that the projection of a V-polyhedron is ("easily") a V-polyhedron. Projection is linear, so respects convex / positive combinations and Minkowski sums. Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, where A is $m \times n$, $b \in \mathbb{R}^m$. Note $Ax \leq b \iff Ax \leq y$ and $y = b$. Then

$$P = \mathrm{pr}\left(\left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{m+n} : Ax \leq y\right\} \cup \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{m+n} : y = b\right\}\right)$$

, where $K = \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{m+n} : Ax \leq y\right\}$ is a polyhedron, and $H = \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{m+n} : y = b\right\}$ is an affine subspace. Then

$$K = \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{m+n} : \begin{bmatrix} A & -I_m \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq O\right\}.$$

□

Claim 2.3.

$$K = \mathrm{pos}\left(\left\{\begin{bmatrix} e_i^n \\ Ae_i^n \end{bmatrix} : 1 \leq i \leq n\right\} \cup \left\{-\begin{bmatrix} e_i^n \\ Ae_i^n \end{bmatrix} : 1 \leq i \leq n\right\} \cup \left\{\begin{bmatrix} 0 \\ e_j^m \end{bmatrix} : 1 \leq j \leq m\right\}\right),$$

where $e_1^n, \ldots, e_n^n$ and $e_1^m, \ldots, e_m^m$ are standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

Proof. Note:

- $Ax \leq y \iff Ax + s = y$ for some "slack" vector

$$s = \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_m \end{bmatrix}, \sigma \geq 0, \forall j.$$

- Also,

$$\forall x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{R}^n,$$

can write $\xi_i = \lambda_i^+ - \lambda_i^-$, with $\lambda_i^+, \lambda_i^- \geq 0$. E.g., if $\xi \geq 0$ then set $\lambda_i^+ := \xi_i, \lambda_i^- = 0$. If $\xi_i \leq 0, \lambda_i^+ = 0, \lambda_i^- = -\xi_i$. Thus, given $\begin{bmatrix} x \\ y \end{bmatrix} \in K$, have a positive combination

$$\sum_{i=1}^{n} \lambda_i^+ \begin{bmatrix} e_i^n \\ Ae_i^n \end{bmatrix} + \sum_{i=1}^{n} \lambda_i^- \left(- \begin{bmatrix} e_i^n \\ Ae_i^n \end{bmatrix} \right) + \sum_{j=1}^{m} \sigma_j \begin{bmatrix} 0 \\ e_j^n \end{bmatrix}. = \begin{bmatrix} l \\ Al \end{bmatrix} + \begin{bmatrix} 0 \\ s \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix},$$

where $l = \begin{bmatrix} \lambda_1^+ - \lambda_1^- \\ \vdots \\ \lambda_n^+ - \lambda_n^- \end{bmatrix}$. OTOH, if have a positive combination, get

$$\begin{bmatrix} A & -I_m \end{bmatrix} = O + \begin{bmatrix} 0 \\ -s \end{bmatrix} \leq O \rightsquigarrow \begin{bmatrix} x \\ y \end{bmatrix} \in K.$$

Thus, $P = \text{pr}(C \cap H)$, where C is a finitely generated cone and H is an affine subspace.

□

Remains: If Q is a V-polyhedron, and H is affine, then $Q \cap H$ is a V-polyhedron.
Warm-up: Suppose $Q \subset \mathbb{R}^n$ is a bounded V-polyhedron (aka. polytope) (not out situation). Let $H := \left\{ x = \begin{bmatrix} \xi_1 \\ \vdots \xi_n \end{bmatrix} \in \mathbb{R}^n : \xi_n = 0 \right\}$. Want $Q \cap H$ as a polytope. Dual to projection. Use a dual to Fourier-Motzkin. Intuition: every vertex of $Q \cap H$ is either in H or is on edge joining two vertices of Q, $a, b \in \text{vert}(Q), w = ta + (1-t)b, 0 \leq t \leq 1$. In last coordinate,

$$0 = t\alpha_n + (1-t)\beta_n \rightsquigarrow t(\alpha_n - \beta_n) = -\beta_n \rightsquigarrow t = \frac{-\beta_n}{\alpha_n - \beta_n}, 1 - t = \frac{\alpha_n}{\alpha_n + \beta_n} \rightsquigarrow w = \frac{-\beta n a + \alpha n b}{\alpha_n - \beta_n}.$$

So, intuitively,

$$Q \cap H = \text{conv}\left(\{a \in \text{vert}(Q) : \alpha = 0\} \cup \left\{ \frac{-\beta_n a + \alpha_n b}{\alpha_n - \beta_n} : a, b \in \text{vert}(Q), \alpha_n > 0, \beta_n < 0 \right\} \right)$$

"Clearly", RHS $\subset Q \cap H$. Next time, the converse containment.

Lemma 2.6. Suppose $C = \{x \in \mathbb{R}^n : Ax \leq 0\}$. Then $C = \text{pos}(R)$ for some $R \subset \mathbb{R}^n$ finite.

Proof. Recall

$$C = \mathrm{pr}(D \cap L) = \mathrm{pr}\left(\left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{m+n} : Ax \le y\right\} \cup \left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{m+n} : y = 0 \in \mathbb{R}^m\right\}\right)$$

It is known that for some (explicit) $S \subset \mathbb{R}^{n+m}$ finite

$$\left\{\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{m+n} : Ax \le y\right\} = \mathrm{pos}(S).$$

Remains to prove that $D \cap L$ is a V-polyhedron — in fact a cone. Set coordinates equal to zero 1-by-1:

$$H = \{(\xi_1, \dots, \xi_{n+m}) \in \mathbb{R}^{n+1}, \xi_{n+m} = 0\}.$$

Suppose $a \in D \cap L$. Let $S = \{b_1, \dots, b_l\}$. Then we have

$$a = \sum_{i=1}^{l} \lambda_i b_i \text{ and } 0 = \sum_{i=1}^{l} \lambda_i \beta_{i,n+m}.$$

This leads to

$$0 = \sum_{i:\beta_{i,n+m}>0} \lambda_i \beta_{i,n+m} + \sum_{i:\beta_{i,n+m}<0} \lambda_i \beta_{i,n+m} \rightsquigarrow \sum_{i:\beta_{i,n+m}>0} \lambda_i \beta_{i,n+m} = \sum_{i:\beta_{i,n+m}<0} \lambda_i(-\beta_{i,n+m}).$$

Let $\Lambda = \sum_{i:\beta_{i,n+m}>0} \lambda_i \beta_{i,n+m}$, $\Lambda \ge 0$. Then

$$\begin{aligned} a &= \sum_{i:\beta_{i,n+m}=0} \lambda_i b_i + \sum_{i:\beta_{i,n+m}>0} \lambda_i b_i + \sum_{i:\beta_{i,n+m}<0} \lambda_i b_i. \\ &= \sum_{i:\beta_{i,n+m}=0} \lambda_i b_i + \frac{1}{\Lambda} \sum_{i:\beta_{i,n+m}<0} \left(\sum_{j:\beta_{j,n+m}>0} \lambda_j \beta_{j,n+m}\right) \lambda_i b_i + \frac{1}{\Lambda} \sum_{j:\beta_{j,n+m}>0} \left(\sum_{i:\beta_{i,n+m}<0} \lambda_i(-\beta_{i,n+m})\right) \lambda_j b_j \\ &= \sum_{i:\beta_{i,n+m}=0} \lambda_i b_i + \frac{1}{\Lambda} \sum_{(i,j):\beta_{i,n+m}<0,\beta_{j,n+m}>0} \lambda_i \lambda_j (\beta_{j,n+m} b_i - \beta_{i,n+m} b_j). \end{aligned}$$

Observe that

$$\sum_{i:\beta_{i,n+m}=0} \lambda_i b_i \in \mathrm{pos}\{b_i \in S : \beta_{i,n+m} = 0\}.$$

For $\beta_{i,n+m} < 0$ and $\beta_{j,n+m} > 0$, have $\beta_{j,n+m} b_i - \beta_{i,n+m} b_j \in D = \mathrm{pos}(S)$ and

$$\sum_{(i,j):\beta_{i,n+m}<0,\beta_{j,n+m}>0} \frac{\lambda_i \lambda_j}{\Lambda} = 1,$$

hence D is a V-polyhedron. We've proved Weyl-Minkowski Theorem: The H-polyhedra ($\{x : Ax \le b\}$) are precisely the V-polyhedra ($\mathrm{conv}(V) + \mathrm{pos}(R)$).

□

Remark 2.8. Goal follows from Lemma.

Proof. Suppose $P = \{x \in \mathbb{R}^n : Ax \le b\}$. Homogenize:

$$C(P) = \left\{ \begin{bmatrix} x \\ \gamma \end{bmatrix} \in \mathbb{R}^{n+1} : \begin{bmatrix} A & -b \\ \mathbf{0}^T & -1 \end{bmatrix} \begin{bmatrix} x \\ \gamma \end{bmatrix} \le 0 \right\}.$$

Then

$$\begin{aligned} P &= \left\{ x \in \mathbb{R}^n : \begin{bmatrix} x \\ 1 \end{bmatrix} \in C(P) \right\} \\ &= \mathrm{pr}_{n+1}(C(P) \cap H), \end{aligned}$$

where $H := \{(\xi_1, \dots, \xi_{n+1}) \in \mathbb{R}^{n+1} : \xi_{n+1} = 1\}$. By Lemma 2.6, we have $C(P) = \mathrm{pos}(S)$ for some $S \subset \mathbb{R}^{n+1}$, finite. Note: $\begin{bmatrix} x \\ \gamma \end{bmatrix} \in C(P) \implies \gamma \ge 0$. Without loss of generality, $S = S_0 \cup S_1$, where

$$\begin{aligned} S_0 &= \{(\rho_1, \dots, \rho_{n+1}) \in S : \rho_{n+1} = 0\}, \\ S_1 &= \{(\rho_1, \dots, \rho_{n+1}) \in S : \rho_{n+1} = 1\}. \end{aligned}$$

(rescaling $r \in S$ if necessary). □

Claim 2.4. $C(P) \cap H$ is a V-polyhedron: $C(P) \cap H = \mathrm{pos}(S_0) + \mathrm{conv}(S_1)$.

Proof. [$\supset$] "easy".
[$\subset$] Suppose $x \in C(P) \cap H$. Then we have

$$x = (\xi_1, \dots, \xi_{n+1}) = \sum_{r \in S} \lambda_r r, \lambda_r \ge 0 \text{ and } \xi_{n+1} = 1.$$

Then

$$1 = \xi_{n+1} = \sum_{s=(\sigma_1,\dots,\sigma_{d+1}) \in S_0} \mu_s \sigma_{d+1} + \sum_{t=(\tau_1,\dots,\tau_{d+1}) \in S_1} \nu_t \tau_{d+1} = \sum_{t \in S_1} \nu_t,$$

where $\mu_s, \nu_t \ge 0$. Then

$$x = \underbrace{\sum_{s \in S_0} \mu_s s}_{\in \mathrm{pos}(S_0)} + \underbrace{\sum_{t \in S_1} \nu_t t}_{\in \mathrm{conv}(S_1)}.$$

That implies P is the projection of a V-polyhedron, hence P is one, too. □

2.12 Visualizing Polars

2.12.1 Cone (convex, linear)

Suppose C is a cone. $C^0 = \{a \in \mathbb{R}^n : \langle a, x \rangle \le 1, \forall x \in C\}$.

Proposition 2.20. $C^0 = \{a \in \mathbb{R}^n : \langle a, x \rangle \le 0, \forall x \in C\}$.

Proof. $\{a \in \mathbb{R}^n : \langle a, x \rangle \le 0, \forall x \in C\} \subset C^0$. Suppose, TGAC, $\exists a \in C^0 \backslash RHS \implies \exists x \in C$ such that $0 < \langle a, x \rangle =: \lambda$. C is cone, which implies $\frac{2}{\lambda} x \in C$ but $\langle a, \frac{2}{\lambda} x \rangle = 2$. □

Corollary 2.2. For $L \subset \mathbb{R}^n$ a linear subspace, $L^0 = L^\perp$.

Visualizing polar of K - a closed convex subset. Add a dimension.

Proposition 2.21. Let $K \subset \mathbb{R}^n$ be closed convex. Let $C(K) = \left\{\lambda \begin{bmatrix} x \\ 1 \end{bmatrix} : x \in K, \lambda \geq 0\right\}$. Then $K^0 = \left\{a \in \mathbb{R}^n : \begin{bmatrix} a \\ -1 \end{bmatrix} \in C(K)^0\right\}$.

Proof.

$$\begin{aligned} \begin{bmatrix} a \\ -1 \end{bmatrix} \in C(K)^0 &\iff \left\langle \begin{bmatrix} a \\ -1 \end{bmatrix}, \lambda \begin{bmatrix} x \\ 1 \end{bmatrix} \right\rangle \leq 0, \forall x \in K, \lambda \geq 0 \\ &\iff \lambda a x - \lambda \leq 0, \forall x \in K, \lambda \geq 0 \\ &\iff \lambda a x - \lambda \leq 0, \forall x \in K, \lambda > 0 \\ &\iff ax \leq 1, \forall x \in K \\ &\iff a \in K^0. \end{aligned}$$

□

Recall the polar map $\mathcal{D} : \mathcal{C}(\mathbb{R}^d) \to \mathcal{C}(\mathbb{R}^d), \mathcal{D}([C]) = [C^0]$ which is a linear map. Dualizing via $\mathcal{D}$. The dual of a cone $C \subset \mathbb{R}^n$: $C^* = \{a \in \mathbb{R}^n : \langle a, x\rangle \geq 0, \forall x \in C\}$.

Definition 2.21. A *section* of a cone $C \subset \mathbb{R}^n$ is $S : (L + x) \cap C$, where L is a linear subspace of $\mathbb{R}^n$, and $x \in C$.

Given a section $S \subset C$, $a \in C^*$, get the perpendicular section Σ through a:

$$\Sigma = [(\text{aff}(S) - x)^\perp + a] \cap C^*,$$

where $x \in S$.

2.13 Linear Programming

Example 2.8. Suppose we have n ingredients (e.g., milk, eggs) and m nutrients (e.g., protein, carbs), α_{ij} is the amount of nutrients i in ingredient j. Have nutritional need of β_i of nutrient $i, \forall i$.

$$A = (\alpha_{ij})_{ij}, b = (\beta_1, \ldots, \beta m).$$

Ingredient j costs γ_j (per unit). Making a dish meets the nutritional requirements but costs as little as possible.

$$Dish = x = (\xi_i, \ldots, \xi_n) \in \mathbb{R}^n,$$

where ξ_j is the amount of ingredient j in dish. That is, seek to minimize

$$\sum_{j=1}^{n} \gamma_j \xi_j = cx \quad \text{subject to } x \in \mathbb{R}^n_{\geq 0}, Ax = b.$$

IOW: Seek

$$\min\{\langle c, x\rangle : Ax = b, x \in \mathbb{R}^n_{\geq 0}\}.$$

Writing $y = (\eta_1, \ldots, \eta_m)$.

- Primal problem: Pick (nonnegative) ingredient quantities to minimize the cost of preparing the dish.
- Dual problem: Pick nutrient prices to maximize cost to customer of buying a bundle β_i nutrient i for $1 \leq i \leq m$ subject to real prices (maybe negative!) and I am selling the nutrition content of ingredient j for at most what the cook had to pay.

2.13.1 Cones and Sections

Proposition 2.22. Suppose C is a closed cone, $S \subset C$ a section of C. Let $S^* \subset C^*$ be the perpendicular section of C^* through $c_0 \in C^*$. Let $x_0 \in S$. Assume $\min\langle c, x_0\rangle$ and $c \in S^*$, $\min_{x\in S}\langle c_0, x\rangle$ exist. Then $\min_{x\in S, c\in S^*}\langle c, x\rangle = 0$.

Proof. Note, $\forall x \in C, c \in C^*, \langle c, x\rangle \geq 0$, then $\exists x' \in S, \langle c_0, x\rangle = \min_{x\in S}\langle c_0, x'\rangle \geq 0$. WTS, $c' \in S^*$ such that $\langle c', x'\rangle = 0$. □

Claim 2.5. $c \in S^* \iff c|_S c_0|_S$ is a constant.

Recall: S is a section: $S^* = [(\text{lin}(S) - x_0)^\perp + c_0] \cap C^*$, then we have

$$\begin{aligned} c \in S^* &\iff c - c_0 \perp \text{lin}(S) - x_0 \text{ and } \langle c, x\rangle \geq 0, \forall x \in C \\ &\iff \langle c - c_0, x - x_0\rangle = 0, \forall x \in S \\ &\iff \langle c - c_0, x\rangle = \langle c - c_0, x_0\rangle, \forall x \in S \end{aligned}$$

2.14 Polarity

Recall that if $C \subset \mathbb{R}^n$ is a closed, convex set then

$$C = \bigcap_{a\in\mathbb{R}^n, \beta\in\mathbb{R}: C\subset H^-_{a,\beta}} H^-_{a,\beta},$$

where $H^-_{a,\beta} = \{x \in \mathbb{R}^n : \langle a, x\rangle \leq \beta\}$. Assume $0 \in C \rightsquigarrow \beta \geq 0$. For $\beta \neq 0$, can rescale a to $a' := \frac{1}{\beta}$ so that $H^-_{a,\beta} = H^-_{a',1}$. Then

$$C = \bigcap_{a\in\mathbb{R}^n: H^-_{a,1}\supset C} H^-_{a,1} \cap \bigcap_{a\in\mathbb{R}^n: H^-_{a,0}\supset C} H^-_{a,0}.$$

$C^0 = \{a \in \mathbb{R}^n : H^-_{a,1} \supset C\}$ is the set of all such $a \in \mathbb{R}^n$, and $C^0 = \{a \in \mathbb{R}^n : \langle a, x\rangle \leq 1, \forall x \in C\}$.

Definition 2.22. Given $S \subset \mathbb{R}^n$, the *polar dual* of S is

$$S^0 = \{a \in \mathbb{R}^n : \langle a, x\rangle \leq 1, \forall x \in S\}.$$

Example 2.9.

$$C := \{(x, y) : -1 \leq x \leq 1, -1 \leq y \leq 1\} \implies C^0 = \{(x, y) : |x|+|y|\leq 1\}.$$

Remark 2.9.

(a) $S^0 = \bigcap_{x\in S} H^-_{x,1}$, then S^0 is closed, convex.

(b) $\{0\}^0 = \mathbb{R}^d$, $(\mathbb{R}^d)^0 = \{0\}$.

(c) $S \subset T \implies T^0 \subset S^0$.

(d) $\left(\bigcup_{\alpha\in I} S_\alpha\right)^0 = \bigcap_{\alpha\in I} S^0_\alpha$.

(e) $S \subset (S^0)^0$.

Theorem 2.23. Suppose $C \subset \mathbb{R}^d$ is closed, convex, $0 \in C$. Then $C = (C^0)^0$.

Proof. Suppose $x \in (C^0)^0$ and TGAC, $x \notin C$. Then there exists $H^-_{a,1}$ such that $C \subset H^-_{a,1}$ but $x \notin H^-_{a,1}$ hyperplanes. Then $x \notin (C^0)^0$. □

Theorem 2.24. Polarity extends to a valuation on $\mathcal{C}(\mathbb{R}^d)$, $\exists \mathcal{D} : \mathcal{C}(\mathbb{R}^d) \to \mathcal{C}(\mathbb{R}^d)$ such that $\mathcal{D}([C]) = [C^0], \forall C \subset \mathbb{R}^d$ closed, convex.

Proof. Want, given

$$g = \sum_{i=1}^{n} \lambda_i [C_i] \in \mathcal{C}(\mathbb{R}^d) \tag{2.8}$$

to define

$$\mathcal{D}(g) = \sum_{i=1}^{n} \lambda_i [C_i^0].$$

Problem is that (2.8) is not unique. Have

$$\begin{aligned}
\sum_{i=1}^{n} \lambda_i [C_i] &= \sum_{i=1}^{n} \lambda_i [\{a \in \mathbb{R}^n : \langle a, x \rangle \leq 1, \forall x \in C_i\}] \\
&= \sum_{i=1}^{n} \lambda_i \left[\bigcap_{x \in C_i} H^-_{x,1} \right] \\
&= \sum_{i=1}^{n} \lambda_i \left[1 - \left(\bigcap_{x \in C_i} H^-_{x,1} \right)^{\complement} \right] \\
&= \sum_{i=1}^{n} \lambda_i - \sum_{i=1}^{n} \lambda_i \left[\bigcup_{x \in C_i} \left(H^-_{x,1} \right)^{\complement} \right] \\
&= \chi(g) - \sum_{i=1}^{n} \lambda_i \left[\bigcup_{x \in C_i} \left(H^-_{x,1} \right)^{\complement} \right].
\end{aligned}$$

Regain closedness,

$$a \in (H^-_{x,1}) \iff \langle a, x \rangle > 1 \iff \langle a, x \rangle \geq 1 + \varepsilon, \forall \varepsilon > 0. \iff a \in H^+_{x,1+\varepsilon}, \forall \varepsilon > 0.$$

$$\sum_{i=1}^{n} \lambda_i [C_i] = \chi(g) - \sum_{i=1}^{n} \lambda_i \lim_{\varepsilon \to 0^+} \left[\bigcup_{x \in C_i} H^+_{x,1+\varepsilon} \right].$$

Now for $y \in \mathbb{R}^d$,

$$\begin{aligned}
\sum_{i=1}^{n} \lambda_i \left[\bigcup_{x \in C_i} H^+_{x,1+\varepsilon} \right](y) &= \sum_{i=1}^{n} \lambda_i \left[\{a \in \mathbb{R}^n : \langle a, x \rangle \geq 1 + \varepsilon \text{ for some } x \in C_i\} \right](y) \\
&= \sum_{i : C_i \cap H^+_{y,1+\varepsilon} \neq \emptyset} \lambda_i \\
&= \chi \left(\sum_{i=1}^{n} \lambda_i [C_i \cap H^+_{y,1+\varepsilon}] \right) \\
&= \chi \left(\sum_{i=1}^{n} \lambda_i [C_i][H^+_{y,1+\varepsilon}] \right) \\
&= \chi(g \cdot [H^+_{y,1+\varepsilon}]) \\
&= \chi(f_\varepsilon(y)).
\end{aligned}$$

Note that

$$\left[\{a \in \mathbb{R}^n : \langle a, x \rangle \geq 1 + \varepsilon \text{ for some } x \in C_i\} \right](y) = \begin{cases} 1 & \text{iff } \exists x \in C_i, x \in H^+_{y,1+\varepsilon}, \\ 0 & \text{otherwise.} \end{cases}$$

□

Definition 2.23. $D(g)(y) := \chi(g) = \lim_{\varepsilon \to 0^+} \chi(f_\varepsilon(y))$.

Independent of the convex combination. Follows from construction that $\mathcal{D} : \mathcal{C}(\mathbb{R}^d) \to \mathcal{C}(\mathbb{R}^d) \implies \mathcal{D}$ is a valuation.

Remark 2.10. If $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\{(\xi_1, \ldots, \xi_d) : \sqrt[p]{\sum_i (\xi_i)^p} = 1\}, \{(\xi_1, \ldots, \xi_d) : \sqrt[q]{\sum_i (\xi_i)^q} = 1\}.$$

2.14.1 Linear Programming Duality

Given $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$. We have the primal linear programming problem and the dual problem

$$\min\{\langle c, x \rangle : Ax = b, x \in \mathbb{R}^n_{\geq 0}\} = \max\{\langle y, b \rangle : A^* y \leq c, y \in \mathbb{R}^m\}, \tag{2.9}$$

where A^* is the adjoint $\langle A^* y, x \rangle = \langle y, Ax \rangle$, provided that both sets are not empty (in particular, the optima exist).
Other formulations exist. E.g.,

$$\min\{\langle y, b \rangle : A^* y \geq c, y \in \mathbb{R}^m\} = \max\{\langle c, x \rangle : Ax = b, x \in \mathbb{R}^n_{\geq 0}\}.$$

All reduces to the same problem, as follows.
Symmetrize: Since $\min\{\langle c, x \rangle : Ax = b, x \in \mathbb{R}^n_{\geq 0}\}$ is well-defined, $\exists x_0 \geq 0$ such that $Ax_0 = b$. $\max\{\langle y, b \rangle : A^* y \leq c, y \in \mathbb{R}^m\} = \max\{\langle y, Ax_0 \rangle : A^* y \leq c, y \in \mathbb{R}^m\}$. Recall that

$$\langle y, Ax_0 \rangle = \langle A^* y, x_0 \rangle.$$

Recall that $\operatorname{Im} A^* = (\ker A)^\perp$.

$$\begin{aligned}\max\{\langle y, Ax_0\rangle : A^*y \le c, y \in \mathbb{R}^m\} &= \max\{\langle A^*y, x_0\rangle : A^*y \le c, y \in \mathbb{R}^m\}\\ &= \max\{\langle z, x_0\rangle : z \le c, z \in \operatorname{Im} A^*\}\\ &= \max\{\langle z, x_0\rangle : z \le c, z \in (\ker A)^\perp\}\\ &= \max\{\langle z - c, x_0\rangle : z - c \le 0, z \in \ker A^\perp\} + \langle c, x\rangle.\end{aligned}$$

Note: Fox $x \subset \mathbb{R}$, $\max(X) = -\min(-X)$ (if $\max(X)$ exists). Then RHS of (2.9) is

$$-\min\{\langle c - z, x_0\rangle : c - z \ge 0, z \in \ker A^\perp\} + \langle c, x\rangle = -\min\{\langle d, x_0\rangle : d \ge 0, c - d \in \ker A^\perp\} + \langle c, x\rangle.$$

While LHS of (2.9) is

$$\min\{\langle c, x'\rangle : x' \ge 0, x - x' \in \ker A\}.$$

Let $L := \ker A$, $C = \ge 0^n$ ($\rightsquigarrow C^* = \mathbb{R}^n_{\ge 0}$). Then

$$\min\{\langle c, x'\rangle : x' \ge 0, x - x' \in \ker A\} = \min\{\langle c, x'\rangle : x' \in C, b = Ax' = Ax\} = \min\{\langle c, x'\rangle : x' \in C, x' \in L + x\}$$

Thus,

$$LPD \iff \langle c_0, x_0\rangle = \min_{x \in S}\langle c_0, x\rangle + \min_{c \in S^\perp}\langle c, x_0\rangle, \tag{2.10}$$

where $S, S^\perp$ are the perpendicular sections.

$$S := (L + x) \cap C \text{ and } S^\perp := (L^\perp + c) \cap C^*,$$

provided that S, $S^\perp$ are not empty.

Theorem 2.25 (Linear programming duality for cones, version 1.0)**.** Suppose $C \subset \mathbb{R}^n$ is *polyhedral* (linear) cone, $L \subset \mathbb{R}^n$ is a linear subspace. Suppose $x_0, c_0 \in \mathbb{R}^n$ are such that perpendicular sections

$$S = (L + x) \cap C \text{ and } S^\perp = (L^\perp + c) \cap C^*$$

are nonempty. Then (2.10) holds.

Reducing (2.10) *further*: Let $\min_{x \in S}\langle c, x'\rangle$ be attained at $\hat{x} \in S \subset L + x$, $\min_{d \in S^\perp}\langle d, x'\rangle$ be attained at $\hat{d} \in S^\perp \subset L^\perp + c$. That implies

$$\exists v \in L, x = \hat{x} + v, \quad \exists w \in L^\perp, c = \hat{c} + w.$$

Then

$$\begin{aligned}\langle c, x\rangle = \langle c, \hat{x}\rangle + \langle \hat{c}, x\rangle &\iff \langle \hat{c} + w, \hat{x} + v\rangle = \langle \hat{c} + w, \hat{x}\rangle + \langle \hat{c}, \hat{x} + v\rangle\\ &\iff \langle \hat{c} + w, v\rangle = \langle \hat{c}, \hat{x} + v\rangle\\ &\iff \langle w, v\rangle = \langle \hat{c}, \hat{x}\rangle.\end{aligned}$$

Theorem 2.26 (Linear programming duality for cones version 2.0)**.** Suppose $C, L, x, c, S, S^\perp$ as in version 1.0, then

$$\min_{x' \in S, c' \in S^\perp}\langle c', x'\rangle = 0.$$

(In particular, all these minima exist.)

Let $C, L, x, c, S, S^\perp$ be as in hypothesis of LPD for cones. Assume *nontrivially*: $x \neq 0$, $c \neq 0$, $L \neq \{0\}$, $L^\perp \neq \{0\}$. (Else, LPD is trivially true).

Lemma 2.7. $L^\perp + c = \{c' \in (\mathbb{R}^n)^* : c'\Big|_{L+x} = c\Big|_{L+x} + \alpha, \exists \alpha \in \mathbb{R} \text{ depending on } c\}$.

Proof. [$\subset$]. Suppose $c' \in L^\perp + c \rightsquigarrow \exists w \in L^\perp, c' = w + c$. For all $x' \in L + x$, have $x' = v + x$ for some $v \in L$. Then

$$\langle c', x' \rangle - \langle c' - c, x' \rangle = \langle w, x' \rangle = \langle w, v + x \rangle = \langle w, x \rangle.$$

□

[$\supset$] Fix $c' \in (\mathbb{R})^*$ such that $c'\Big|_{L+x'} - c\Big|_{L+x}$ is constant, say $\alpha \in \mathbb{R}$, pick $w = c' - c$. $\forall v \in L, \alpha = \langle w, v + x \rangle = \langle w, x \rangle + \langle w, x \rangle$. Set $v = 0$, then $\alpha = \langle w, x \rangle$. Then $\langle w, v \rangle = 0 \rightsquigarrow w \in L^\perp \rightsquigarrow c' = w + c \in L^\perp + c$.

Lemma 2.8. The minima

(i) $\min_{x \in S} \langle c, x' \rangle$,

(ii) $\min_{c \in S^\perp} \langle c', x \rangle$,

(iii) $\min_{x \in S, c \in S^\perp} \langle c', x' \rangle$,

are attained.

Proof. For (i) and (ii), computing minima of images of polyhedral sets (S or $S^\perp$) under linear maps

$$\langle c, \cdot \rangle : \mathbb{R}^n \to \mathbb{R}, \langle \cdot, x \rangle : (\mathbb{R}^n)^* \to \mathbb{R}.$$

By Fourier-Motzkin, these images are closed. Since $S \subset C$ and $S^\perp \subset C^*$, the images in $\mathbb{R}$ are bounded below by 0, if $c \in S^\perp$, $x \in S$. Otherwise, if not, fix $c \in S^\perp \subset (L^\perp + c) \cap C^*$, then we have $c = c' + w$ for some $w \in L^\perp$. Therefore,

$$\begin{aligned} \langle c, x' \rangle &= \langle c' + w, x' \rangle \\ &= \langle c' + w, v + x \rangle \\ &= \underbrace{\langle c', x' \rangle}_{\geq 0} + \underbrace{\langle w, x \rangle}_{\text{constant in } x, \text{say} alpha} . \end{aligned}$$

Dualize the argument to show that (ii) is attained. For (iii) get lower bound,

$$\{\langle c', x' \rangle : c' \in S^\perp \subset C^*, x \in S \subset C\} \geq 0.$$

We will now show that attain this lower bound. *The Topological Heart* — A *positive* Hahn-Banach Theorem. Let C be a cone in $\mathbb{R}^n$. Let $U \subset \mathbb{R}^n$ be a linear subspace. Assume that $U \cap C \neq U$. □

Definition 2.24. Given cone C and $U \subset \mathbb{R}^n$ (a linear subspace), call $C' := C \cap U$, a sectional subcone.

C is not a linear subspace, then $0 \in \partial C$. Recall: Have a supporting hyperplane at 0. I.e., $\exists c \in (\mathbb{R})^*$ such that $\langle c, x \rangle \geq 0, \forall x \in C$, i.e., $C^* \subset (\mathbb{R}^n)^*$ is nonempty. In particular, C' has a supporting hyperplane in U^*: $\exists \tilde{c} : U^* \to \mathbb{R}$ such that $\tilde{c}(x) \geq 0, \forall x \in C'$.

Theorem 2.27 (Positive Hahn-Banach Theorem)**.** Let $C' \subset C \subset \mathbb{R}^n$ be a sectional subcone: $C' = C \cap U$ for some linear subspace U. Assume $C' \neq U$. Let $\tilde{c} \in U^*$ supporting C' be given. Then there exists an extension $c \in C^*$ of $\tilde{c}$:

$$c\Big|_U = \tilde{c} \text{ and } c(x) \geq 0, \forall x \in C.$$

Theorem 2.28. Suppose $C \subset V$, where $\dim V = n$, is a cone. $U \subset V$ is a linear subspace. Let $C' = U \cap C$. Suppose $\tilde{c} \in U^* = \{\text{linear functionals } U \to \mathbb{R}\}$ supports C': $\tilde{c} \geq 0, \forall x \in C'$. Then there exists an extension $c \in V^*$ of $\tilde{c}$ that supports C:

$$c(x) = \tilde{c}(x), \forall x \in U.$$

Proof. By induction on $\operatorname{codim}(U) = \dim(V) - \dim(U)$. Suppose $\operatorname{codim}(U) = 1$. Fix $y_0 \in V$ such that $V = U \oplus \operatorname{span}(y_0)$. Then $\forall y \in V, \exists! \, x \in U, \lambda \in \mathbb{R}$ such that $y = x + \lambda y_0$. Then

$$\{\text{extensions } c \in V^* \text{ of } \tilde{c}\} \iff \{\text{choice } \xi := c(y_0) \in \mathbb{R}\} \text{ via } c(y) = c(x + \lambda y_0) = \tilde{c}(x) + \lambda\xi.$$

Thus, seek $\xi \in \mathbb{R}$ such that

$$\tilde{c}(x) + \lambda\xi \geq 0, \forall x \in U, \lambda \in \mathbb{R}, x + \lambda y_0 \in C.$$

Have the case that $\lambda = 0$ because $\tilde{c}$ supports $C' = C \cap U$. In other words, seek $\xi \in \mathbb{R}$ such that

$$\begin{cases} \xi \geq \dfrac{-1}{\lambda}\tilde{c}(x), \forall x \in U, \lambda > 0, \\ \xi \leq \dfrac{-1}{\lambda}\tilde{c}(x), \forall x \in U, \lambda < 0. \end{cases}$$

So, seek $\xi \in \mathbb{R}$ such that

$$\frac{-1}{\lambda_2}\tilde{c}(x_2) \leq \xi \leq \frac{-1}{\lambda_1}\tilde{c}(x_1), \forall x_1, x_2 \in U, \lambda_1, \lambda_2 \in \mathbb{R}, \lambda_2 < 0 < \lambda_1.$$

and

$$x_1 + \lambda_1 y_0 \in C, \quad x_2 + \lambda_2 y_0 \in C.$$

Can eliminate ξ. Equivalently, need

$$\frac{-1}{\lambda_2}\tilde{c}(x_2) \leq \frac{-1}{\lambda_1}\tilde{c}(x_1), \forall x_1, x_2 \in U, \lambda_1, \lambda_2 \in \mathbb{R}, \lambda_2 < 0 < \lambda_1, \frac{-1}{\lambda_2}x_2 \leq_C y_0 \leq_C \frac{-1}{\lambda_1}x_1.$$

That is, need

$$-\lambda_1\tilde{c}(x_2) \geq -\lambda_2\tilde{x_1}, \forall x_1, x_2 \in U, \lambda_1, \lambda_2 \in \mathbb{R}, \lambda_2 < 0 < \lambda_1, -\lambda x_2 \geq_C \lambda_1\lambda_2 y_0 \geq_C -\lambda_2 x_1.$$

Indeed, have

$$\tilde{c}(u_2) \geq \tilde{c}(u_1), \forall u_1, u_2 \in U, u_2 \geq_C u_1 \iff u_2 \geq_{C'} u_1,$$

because $\tilde{c}$ supports C' in U. Precede by induction. If $\operatorname{codim}(U) > 1$, have

$$U = U_r \subset U_{r+1} \subset \cdots \subset U_n = V,$$

with $\dim(U_{i+1}) = \dim(U_i) + 1$. Apply base case to

$$C \cap U_i \subset C \cap U_{i+1}.$$

□

Helpful Observation: C induces a *partial order* on V via $x \leq_C y$ if and only if $y - x \in C$.

Remark 2.11.

- $0 \leq_C x \iff x \in C$.
- $x \leq_C y \implies \forall \begin{cases} \lambda > 0, \lambda x \leq_C \lambda y \\ \lambda < 0, \lambda x \geq_C \lambda y \end{cases}$.
- $x \leq_C y$ and $z \in V \implies x + z \leq_C y + z$.
- If $c \in V^*$ supports C, then $x \leq_C y \implies c(x) \leq c(y)$.

It is a partial order:

- *Transitivity*:

$$\begin{aligned} x \leq_C y \text{ and } y \leq_C z &\iff y - x \in C \text{ and } z - y \in C \\ &\iff (x - y) + (y - z) = x - z \in C \\ &\iff x \leq_C z. \end{aligned}$$

Back to Linear Programming Duality
$C, L, S = (L + x_0) \cap C, S^\perp = (L^\perp + c_0) \cap C^*$. Last time,

$$x' \in S, \langle c_0, x' \rangle = \min_{x \in S} \langle c_0, x \rangle c' \in S^\perp, \langle c', x_0 \rangle = \min_{c \in S^\perp} \langle c, x_0 \rangle.$$

Saw:

$$\text{LP duality} \iff \langle c', x' \rangle = 0.$$

In particular, LP duality $\implies \min_{x \in S, c \in S^\perp} \langle c, x \rangle$ exists.
Suppose, TGAC, $\langle c', x' \rangle = \beta > 0$. Define $\tilde{c} : L + x_0 \to \mathbb{R}$ via $\tilde{c}(x) = c'(x) - \beta, \forall x \in L + x_0$. Extend $\tilde{c}$ by linearity to $\text{lin}(L + x_0) = U$.
Proof (LPD for polyhedral cones) Have $x' \in S$, $c' \in S^\perp \subset L^\perp + c_0$, $\langle c_0, x' \rangle = \min_{x \in S} \langle c_0, x \rangle$, $\langle c', x_0 \rangle = \min_{c \in S^\perp} \langle c, x_0 \rangle$. WTS: $\beta = \langle c', x' \rangle = 0$. Have $\tilde{c} : L + x_0 \to \mathbb{R}$, $\tilde{c} : x_{\tilde{c}} \mapsto \langle c', x \rangle - \beta$. Extended by linearity to $\text{lin}(L + x_0) = U$. Note: Every element of $L^\perp + c_0$ attain its minimum over S at x', because, on $L + x_0$, they differ from each other by constants. $\langle \tilde{c}, x' \rangle = 0$, $\langle \tilde{c}, x \rangle \geq 0, \forall x \in S$. Then

$$\tilde{c} \text{ is nonnegative on } S \implies \tilde{c} \text{ is nonnegative on } \text{pos}(S) = C'.$$

Case 1: Suppose $\text{pos}(S) = C \cap U$, have $\tilde{c} \in (C')^* \subset U^*$. Apply previous theorem to extend $\tilde{c} \in U^*$ to $c \in (\mathbb{R}^n)^*$ supporting C. Have $c \in (L^\perp + c_0) \cap C^*$, where $L^\perp + c_0$ still differs from c' by a constant. Then $\langle c, x' \rangle = 0$ is minimum attained over S by elements of $S^\perp$.
Case 2: By previous theorem, we can assume, without loss of generality, that $U = \mathbb{R}^n$ (the whoe primal space). Then without loss of generality, $C' = C$. Only know (so far) that $\tilde{c}$ is nonnegative on $\text{pos}(S)$. WTS: $\tilde{c} \in \partial C^*$ and $x' \in \partial C$, and they support each other. I.e., $\ker \langle \tilde{c}, \cdot \rangle$ is a supporting hyperplane at x', and vice versa. Then $\langle \tilde{c}, x' \rangle = 0$.

Claim 2.6. Without loss of generality, c' is not constant on S. (If were, then $\tilde{c} = c' - \beta \equiv 0$), so $\tilde{c}$ supports C, therefore $\tilde{c} \in C^*$.

Then c' attains its minimum on ∂C. Else can go in some direction still in C attaining a smaller value than at x'. Then $\langle \tilde{c}, x' \rangle = 0$ and x' is on $\partial S \subset \partial C$, then $\tilde{c}$ supporting $\operatorname{pos}(S)$ at x'.

$$\begin{aligned} \partial \operatorname{pos}(S) \subset \partial C &\implies \tilde{c} \text{ supports } C \text{ at } x' \in \partial C \\ &\implies \tilde{c} \in C^* \implies \tilde{c} \in L^\perp + c_0 \\ &\implies \tilde{c} \in S^\perp. \end{aligned}$$

3 Ellipsoids and cones over ellipsoids

Definition 3.1. Let $B_k \subset \mathbb{R}^k$ be the closed unit ball in $\mathbb{R}^k$. Let $A : \mathbb{R}^k \to \mathbb{R}^n$ be an affine linear map, i.e., there exists linear map $L : \mathbb{R}^k \to \mathbb{R}^n$ and $b \in \mathbb{R}^n$ such that $Ax = Lx + b$. Then $A(B_k) \subset \mathbb{R}^n$ is an ellipsoid.

Definition 3.2. $S \subset \mathbb{R}^n$ is centrally symmetric if $\exists c, -(S - c) + c = S$.

Note: Ellipsoids are centrally symmetric.

Theorem 3.1. Suppose that $K \subset \mathbb{R}^n$ compact, closed and centrally symmetric, the following are equivalent:

(a) K is an ellipsoid.

(b) Blashke (1916): For all bundles of parallel chord, the midpoints all lie in a hyperplane.

(c) Myer-Reisner (1989): For all bundles of parallel hyperplanes $\{H\}$, the centroids of the section the section $H \cap K$ all lie on a lie.

(d) Blashke (1916?): Every "shadow boundaries" is contained in a hyperplane, where shadow boundary is the preimage of the boundary $P(K)$ for some projection P.

(e) Aitchison-Petty-Rogers (1973): $\exists p \in K$, which is not the centroid such that for all hyperplanes $H \ni p$, the section $K \cap H$ is centrally symmetric. (p is called a "false center") (Requires $\dim K \geq 3$)

(f) Brunn (1889): Every section of K is an ellipsoid.

3.1 Characterizing ellipsoidal cones

Definition 3.3. A cone $C \subset \mathbb{R}^n$ is ellipsoidal if some section of C is an ellipsoid of $\dim(C) - 1$.

3.2 Application – Chronogeometry

In special relativity, space-time is Minkowski spacetime:

(a) M: a 4 dimensional *real affine space*. That is, a set M together with transitive and free group action of $\mathbb{R}^4$ on M. I.e., $\forall a \in M$, $v \mapsto a + v$ is a bijection. $\mathbb{R}^4 \to M$ at each point $a \in M$, M looks like $\mathbb{R}^4$ with a as the origin.

(b) A cone $C \subset \mathbb{R}^4$ – the light cone.

(c) Points = "events".

- $a + C$: points in M that a can affect.
- $b - C$: point in M that can affect b.

Question: Why do we observe 1-dimensional and a 3-dimensional space?
Answer: Fix $a \in M$, $b \in \text{int}(a + C)$ the "time-time" is the line through a and b, the "space-summand" is the affine span of $\partial(a + C) \cap \partial(b - C)$.
Empirical fact: $\text{aff}(\partial(a + C) \cap \partial(b - M))$ is 3-dimensional, $\forall b \in \text{int}(a + C)$.

Theorem 3.2. Let $C \in \mathbb{R}^n$ be a closed convex cone. Assume that $\forall b \in \text{int}(C)$, there exists hyperplane H such that $\partial C \cap \partial(b - C) \subset H$. Then C is an ellipsoidal cone.

The reason why "space is 3-dimensional" is C is ellipsoidal. Conversely, by the theorem, *only* an ellipsoidal cone would give 3-dimensional space.

Definition 3.4 (FBI property for cones)**.** A closed cone $C \subset \mathbb{R}^n$ satisfies the *flat boundary intersections* property if $\forall b \in \text{int}(C)$, there exists hyperplane H such that $\partial C \cap \partial(b - C) \subset H$. In this case, call C FBI.

Fact: C is ellipsoidal $\implies$ C is FBI.
To prove the convese:

Definition 3.5 (CSS property for cones)**.** Closed cone $C \subset \mathbb{R}^n$ satisfies the *centrally symmetric sections* (CSS) properties if every boundary section of C is centrally symmetric.

Theorem 3.3 (Olavjanischnikoff (1941))**.** A closed cone $C \subset \mathbb{R}^3$ is CSS if and only if C is ellipsoidal.

Theorem 3.4 (Jerónimo, McAllister)**.** Closed cone $C \subset \mathbb{R}^n$ is CSS if and only if C is ellipsoidal.

Outline of argument

1. 3-dimensional CSS characterization of ellipsoidal cones $\implies$ 3-dimensional FBI characterization of ellipsoidal cones: Uses fact that convex bodies of dimension less than 5 contain inscribed parallelepiped.
2. 3-dimensional FBI characterization of ellipsoidal cones $\implies$ n-dimensional FBI characterization of ellipsoidal cones.
3. n-dimensional FBI characterization of ellipsoidal cones $\implies$ n-dimensional CSS characterization of ellipsoidal cones.

4 Approximating convex bodies with ellipsoids

Definition 4.1. A *convex body* $K \subset \mathbb{R}^n$ is a compact convex subset of $\mathbb{R}^n$ with nonempty interior.

Goal: Every convex body $K \subset \mathbb{R}^2$ contains a unique ellipsoid E of maximum volume, and, (if E is centered at 0),

$$E \subset K \subset nE.$$

Recall: $E \subset \mathbb{R}^n$ is an ellipsoid if $E = T(B) + a$ for some linear $T : \mathbb{R}^k \to \mathbb{R}^n, a \in \mathbb{R}^n$, where $B = \{x \in \mathbb{R}^k, \|x\| \leq 1\}$.

Remark 4.1.

(a) For full-dimensional ellipsoids $E \subset \mathbb{R}^n$ (i.e., E $\neq \emptyset$) suffices to consider $T : \mathbb{R}^n \to \mathbb{R}^n$ invertible.

(b) In this case: $\text{vol}(E) = |\det(T)|\text{vol}(B)$, where $B \subset \mathbb{R}^n$ is a unit ball.

(c)

$$\begin{aligned} E &= \{Ty + a : \langle y, y\rangle \leq 1\} \\ &= \{x \in \mathbb{R}^n : \langle T^{-1}(x-a), T^{-1}(x-a)\rangle \leq 1\} \\ &= \{x \in \mathbb{R}^n : \langle x-a, Q(x-a)\rangle \leq 1\}, \end{aligned}$$

where $Q = (T^{-1})^* T^{-1} = (TT^*)^{-1}$ is a positive definite operator.

Definition 4.2. $S \in \mathcal{L}(\mathbb{R}^n)$ is positive definite if S has an orthonormal basis of eigenvectors and corresponding eigenvalues are all positive.

RecalL:

- S is positive definite if and only if $S^* = S$ and $\langle x, Sx\rangle > 0$ for all $x \neq 0$.

Remark 4.2. If basis for $\mathbb{R}^n$ is the eigenvectors of Q, with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then

$$E = \{(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n : \sum_{i=1}^{n} \lambda_i (\xi_i - \alpha_i)^2 \leq 1\},$$

where $a = (\alpha_1, \ldots, \alpha_n)$ (in basis of eigenvectors).

Recall: We have *polar decomposition*: Every invertible $T \in \mathcal{L}(\mathbb{R}^n)$ is a composition $T = SU$, where U is orthogonal ($U^*U = I$) and S is positive definite.

Corollary 4.1. Every full-dimensional ellipsoid E can be written $E = S(B) + a$, where $S \in \mathcal{L}(\mathbb{R}^n)$ is positive definite and $a \in \mathbb{R}^n$.

Proof. We have $E = T(B) + a = SU(B) + a$, and $U(B) = B$. □

Technical Lemma:

Lemma 4.1. Suppose $S, T \in \mathcal{L}(\mathbb{R}^n)$ are positive definite. Then

$$\det\left(\frac{S+T}{2}\right) \geq \sqrt{\det(S)\det(T)}$$

with equality if and only if $S = T$.

Proof. X is positive definite, we have basis such that the matrix of X is

$$[X] = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

for $\lambda_i > 0$. In $\frac{X+Y}{2}$, divide row i by $\sqrt{\lambda_i}$ and divide column i by $\sqrt{\lambda_i}$.

$$\det\left(\frac{X+Y}{2}\right) = \left(\prod_{i=1}^{n} \lambda_i\right) \det\left(\frac{I+Y'}{2}\right),$$

where $Y' = X^{-1/2} Y X^{-1/2}$, where

$$[X^{-1/2}] = \begin{bmatrix} \lambda_1^{-1/2} & & & \\ & \lambda_2^{-1/2} & & \\ & & \ddots & \\ & & & \lambda_n^{-1/2} \end{bmatrix}$$

On the right-hand side,

$$\sqrt{\det(X)\det(Y)} = \det(X)\sqrt{\det(Y')} = (\prod_{i=1}^{n} \lambda_i)\det(Y').$$

Reduced problem to showing, for Y' positive definite

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

with equality if and only if $Y' = I$. Now think in basis such that

$$[Y'] = \begin{bmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & \\ & & & \mu_n \end{bmatrix},$$

where $\mu_i > 0, i = 1, \ldots, n$. So, WTS

$$\prod_{i=1}^{n} \frac{1+\mu_i}{2} \geq \sqrt{\prod_{i=1}^{n} \mu_i} = \prod_{i=1}^{n} \sqrt{\mu_i}.$$

Have the inequality ($\geq$) from AMGM with equality if and only if $\mu_i = 1 \iff Y' = I \iff X = Y$. □

Remark 4.3. CF: "AMGM": $\forall \sigma, \tau > 0$, $\frac{\sigma+\tau}{2} \geq \sqrt{\sigma\tau} \iff (\sigma - \tau)^2 \geq 0$ with equality if and only if $\sigma = \tau$.

Goal (Part 1):

Theorem 4.1. Suppose $K \subset \mathbb{R}^n$ is a convex body. Then there exists a unique ellipsoid $E \subset K$ of maximum volume.

Proof. Existence: Seek positive definite $S \in \mathcal{L}(\mathbb{R}^n)$ and $a \in \mathbb{R}^n$ such that $S(B) + a \subset K$ and $|\det(S)| = \frac{\mathrm{vol}(S(B)+a)}{\mathrm{vol}(B)}$ is maximal. Let $\mathrm{Sym}_n \subset \mathcal{L}(\mathbb{R}^n)$ be the set of positive definite matrices. Let $X \subset \mathrm{Sym}_n \times \mathbb{R}^n$ be

$$X = \{(S, a) : S(B_n) + a \subset K, a \in \mathbb{R}^n, S \text{ is positive semi-definite}\},$$

where $B_n = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$. By the Remark below, we have $|\det(S)|$ attains its maximum on X. That implies there exists an ellipsoid $E_1 \subset K$ of maximum volume $\mathrm{vol}(E_1) = |\det(S_1)|\mathrm{vol}(B_n)$. *Uniqueness*: Suppose we have

$$E_1 = S_1(B_n) + a_1 \subset K, E_2 = S_2(B_n) + a_2 \subset K,$$

where S_1, S_2 is positive definite, of maximum volume. X is convex, then

$$E_3 = \frac{S_1 + S_2}{2}(B_n) + \frac{a_1 + a_2}{2},$$

because $\frac{1}{2}(S_1, a_1) + \frac{1}{2}(S_2, a_2) \in X$. By Lemma 4.1, since S_1, S_2 have maximum volume, then

$$\det(S_1) = \det(S_2) \implies \det(S_1) \geq \det\left(\frac{S_1 + S_2}{2}\right) \geq \det(S_1).$$

Have equality, we then conclude $S_1 = S_2$. Therefore E_1, E_2 are two translations of same linear image of B_n. Applying an affine transformation, assume, without loss of generality, E_1, E_2 are unit balls and $\frac{a_1+a_2}{2} = 0$. K is convex which implies $\text{conv}\{E_1 \cup E_2\} \subset K$. If $a \neq -a$, can "stretch" E_3 to get $E_4 \subset K$ of greater volume. Hence $E_1 = E_2$.

□

Remark 4.4.

- Since K compact, X is compact in $\text{Sym}_n \times \mathbb{R}^n$.
- $\det(S)$ is a polynomial function of entries of $(S, a) \in X$.

Theorem 4.2. Suppose $K \subset \mathbb{R}^n$ is a convex body, $E \subset K$ is the max volume eillipsoid. Assume (after translation if necessary) E is centered at the origin. Then $E \subset K \subset nE$.

Proof. Apply linear operator if necessary, assume, without loss of generality, $E = B_n$. WTS: $\forall r \in K, \|r\| \leq n$. Suppose $\exists r \in K$ such that $\|r\| > n$. Without loss of generality, $r = (\rho, 0, \ldots, 0)$ for some $\rho > 0$. Since K is convex, $E \subset K$, $r \in K$, $\text{conv}(E \cup \{r\}) \subset K$.
Note:

- The equation of ellipse ∂E_1 has the form

$$\frac{(x - \tau)^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1.$$

- Given center $(\tau, 0, \ldots, 0)$ of E_1, there exists unique values of α, β such that
  - ∂E_1 passes through $(-1, 0, \ldots, 0)$.
  - ∂E_1 meets the L from r to "top half" of ∂E at exactly one point.
  - For given α, β, the volume of E_1 is given by

$$\text{vol}(E_1) = \alpha\beta^{n-1} \text{vol}(B_n),$$

because $E_1 = T(B_n) + r$, where

$$T = \begin{bmatrix} \alpha & & & \\ & \beta & & \\ & & \ddots & \\ & & & \beta \end{bmatrix} \implies \det(T) = \alpha\beta^{n-1}.$$

Strategy:

- Find α, β in terms of τ.
- Find τ making $\alpha\beta^{n-1} > 1$ hence contradicting maximality of $\mathrm{vol}(E)$.

$(-1, 0, \ldots, 0) \in \partial E_1 \implies \alpha = \tau + 1$. Thus the equation of ellipse becomes

$$\frac{(x-\tau)^2}{(\tau+1)^2} + \frac{y^2}{\beta^2} = 1,$$

and the equation of line:

$$\frac{y-0}{x-\rho} = -\frac{1}{\sqrt{\rho^2-1}} \implies y = -\frac{x-\rho}{\sqrt{\rho^2-1}}.$$

Points (x, y) of intersection satisfy

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

Want exactly one point of intersection, that is, want the discriminant of the intersection to be 0 to solve for β^2, get

$$\alpha = \tau + 1, \quad \beta^2 = 1 - \frac{2\tau}{\rho - 1}.$$

For each sufficient small $\tau > 0$, constructed $E_1 \subset K$ centered at $(\tau, 0, \ldots, 0)$ with equation of ∂E,

$$\frac{(x_1-\tau)^2}{\alpha^2} + \frac{1}{\beta^2}\sum_{i=2}^{n} x_i^2 = 1$$

with $\alpha = 1 + \tau$, $\beta^2 = 1 - \frac{2\tau}{\rho-1}$. Have $\mathrm{vol}(E_1) = \alpha\beta^{n-1}\,\mathrm{vol}(E)$. TGAC, seek $\tau > 0$ such that

$$\begin{aligned} \alpha\beta^{n-1} > 1 &\iff \ln(\alpha) + (n-1)\ln(\beta) > 0 \\ &\iff \ln(\alpha) > -\frac{n-1}{2}\ln(\beta^2) \\ &\iff \ln(1+\tau) > -\frac{n-1}{2}\ln\left(1 - \frac{2\tau}{\rho-1}\right). \end{aligned} \tag{4.1}$$

Recall MacLaurin Series $ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$ for $x \in (-1, 1)$. LHS of (4.1): $\tau + O(\tau^2)$ some small $\tau > 0$. RHS of (4.1):

$$-\frac{n-1}{2}\left(-\frac{2\tau}{\rho-1} + O(\tau^2)\right) = \frac{n-1}{\rho-1}\tau + O(\tau^2).$$

So, get (4.1) for sufficient small $\tau > 0$ so $\|r\| = \rho > n$. Then $\mathrm{vol}(E_1) > \mathrm{vol}(E)$. Contradiction. □

Theorem 4.3. Suppose $K \subset \mathbb{R}^n$ is a *centrally symmetric* convex body. Assume maximum volume ellipsoid is centered at the origin ($K = -K$). Then

$$K \subset \sqrt{n}E.$$

Proof. As before, without loss of generality, $E = B_n$. TGAC, have $r \in K$, $\|r\| > \sqrt{n}$. Centrally symmetry implies $-r \in K$. For given β, find α such that $\partial E_1 \cap (xy\text{-plane})$ with equation

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$$

meets L at single point. As before, set discriminant to be zero to find

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

Set $\beta = 1 - \varepsilon$ for $\varepsilon > 0$. Seek $\varepsilon > 0$ sufficient small such that $\alpha\beta^{n-1} > 1$, then find $\ln(\alpha) + (n-1)\ln(\beta) > 0$ holds for $\varepsilon > 0$ sufficient small. □

4.1 Face lattices of polytopes

Recall: Suppose $K \subset \mathbb{R}^n$ closed convex, $F \subset K$ is a *face* of K.

$$\langle a, F\rangle = \beta, \quad \langle a, K\backslash F\rangle < \beta,$$

for some $a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$.
Question: Is every face of a face a face? That is, suppose E is a face of F, F is a face of K. Must E be a face of K?

- Yes: if K is a polytope.
- No: For general closed convex (even compact) K.

Let $K = \triangle \cup \circ$ such that the tangents to ∂K at a and b are unique. any functional supporting K at E supports all of F. But why is true for polytopes?

Theorem 4.4. Faces of polytopes are polytopes. Suppose $P = \text{conv}(V), |V| \leq \infty, V \subset \mathbb{R}^n$. Suppose $P \subset H^-_{a,\beta}, F = P \cap H_{a,\beta}$. Then $F = \text{conv}(V \cap H_{a,\beta}) = U$.

Proof.
$\supset$: "Trivial". $V \cap H_{a,\beta} \subset P \cap H_{a,\beta}$.
$\subset$: Let $x \in F$ be given. $x \in F \subset P$, then have

$$x = \sum_{v \in V} \lambda_v v,$$

with $\lambda_v \geq 0, \forall v \in V$ and

$$\sum_{v \in V} \lambda_v = 1.$$

$x \in F$, then

$$\begin{aligned}\beta &= \langle a, x\rangle \\ &= \sum_{v\in V} \lambda_r \langle a, v\rangle \\ &= \sum_{u\in U} \lambda_u \langle a, u\rangle + \sum_{v\in V\backslash\backslash U} \lambda_v \langle a, v\rangle \\ &< \beta \sum_{u\in U} \lambda_u + \beta \sum_{v\in V\backslash U} \lambda_v \\ &= \beta \left(\sum_{v\in V} \lambda_v\right) \\ &= \beta.\end{aligned}$$

If some λ_v ($v \in V\backslash U$) is nonzero. Then $\lambda_v = 0$, $\forall v \in V\backslash U$, then $v \in \text{conv}(U)$. □

Notation: Write "$F \le P$" for "F is a face of P".

Theorem 4.5. Suppose P is a polytope. Then $E \le F \le P \implies E \le P$ (i.e., $\le$ is transitive *on polytopes*).

Proof. Put $P = \text{conv}(V)$, $|V| < \infty$. By previous theorem, we have $F = \text{conv}(U)$ for some $U \subset V$.

$$F \le P \implies \langle a_{FP}, F\rangle = \beta_{FP}, \langle a_{FP}, P\backslash F\rangle < \beta_{FP},$$

and

$$E \le F \implies \langle a_{EF}, E\rangle = \beta_{EF}, \langle a_{EF}, F\backslash E\rangle < \beta_{EF},$$

for some $a_{FP}, a_{EF} \in \mathbb{R}^n$, $\beta_{FP}, \beta_{EF} \in \mathbb{R}$, $\forall \varepsilon > 0$, let $a_\varepsilon := a_{FP} + \varepsilon a_{EF}, \beta_\varepsilon = \beta_{FP} + \varepsilon\beta_{EF}$.

$$\langle a_\varepsilon, E\rangle = \langle a_{FP}, E\rangle + \varepsilon\langle a_{EF}, E\rangle = \beta_{FP} + \varepsilon\beta_{EF} = \beta_\varepsilon.$$

$$\langle a_\varepsilon, F\backslash E\rangle = \langle a_{FP}, F\backslash E\rangle + \varepsilon\langle a_{EF}, F\backslash E\rangle < \beta_{FP} + \varepsilon\beta_{EF} = \beta_\varepsilon.$$

$$\langle a_\varepsilon, P\backslash F\rangle = \langle a_{FP}, P\backslash F\rangle + \varepsilon\langle a_{EF}, P\backslash F\rangle < \beta_{FP} + \varepsilon\beta_{EF} = \beta_\varepsilon.$$

Seek $\varepsilon > 0$ such that $RHS < \beta_\varepsilon = \beta_{FP} + \varepsilon\beta_{EF}$. That is, seek $\varepsilon > 0$ such that

$$\underbrace{(\langle a_{FP}, P\backslash F\rangle - \beta_{FE})}_{<0} + \varepsilon\left(\langle a_{EF}, P\backslash F\rangle - \beta_{EF}\right) < 0. \tag{4.2}$$

The point: by previous theorem, need (4.2) at only *finitely many* point $P\backslash F$, namely, at $v \in V\backslash U$. Picking any $\varepsilon > 0$ such that

$$\max_{v\in V\backslash U} \frac{\langle a_{EF}, v\rangle - \beta_{EF}}{\beta_{EF} - \langle a_{FP}, v\rangle} < \frac{1}{\varepsilon}.$$

□

Remark 4.5.

- $P \le P$ (let $a = 0$, $\beta = 0$, then $\langle a, P\rangle = \beta$).
- $\emptyset \le P$ (let $a = 0$, $\beta = 1$, then $\langle a, P\rangle < \beta$).

- $\leq$ is relexive, anti-symmetric (because $\subset$ is anti-symmetric), transitive.

For polytope P, write $\mathcal{F}(P) = \{F \subset P : F \leq P\}$.

Corollary 4.2. $\mathcal{F}(P)$ is a partially orderred set (poset), ordered by $\leq$. Has maximal element P, minimal element $\emptyset$, ranked by dimension.

Remark 4.6.

- Define $\dim \emptyset = -1$.
- $f_i(P) = \#\{F \leq P : \dim F = i\}$ for $-1 \leq i \leq \dim(P)$.

i	-1	0	1	2	3
$f_i(P)$	1	4	6	4	1

Example 4.1.

4.1.1 Euler - Poincaré Relations

For 3-dimensional polytope P with

$$f_0(P) = V,$$
$$f_1(P) = E,$$
$$f_2(P) = F.$$

We have $V - E + F = 2$. In general,

Theorem 4.6 (Euler-Poincaré formula)**.** Suppose $P \subset \mathbb{R}^n$ is a d-dimensional polytope. Then

$$\sum_{i=0}^{d-1} (-1)^i f_i(P) = 1 + (-1)^{d-1} = \chi(\partial P).$$

Proof. Note that ∂P is *disjoint* union of the *relative* interiors of the faces $F < P$. (In particular, relative interior of a vertex is just that vertex.)

$$[\partial P] = \sum_{\emptyset < F < P} [\text{int}(F)].$$

By linearity of χ,

$$1 + (-1)^{d-1} = \chi(\partial P) = \sum_{\emptyset < F < P} \chi(\text{int}(F)) = \sum_{i=0}^{d-1} \sum_{F < P, \dim(F) = i} (-1)^i = \sum_{i=0}^{d-1} f_i(P)(-1)^i.$$

□

Follows from

Lemma 4.2. Suppose $P \subset \mathbb{R}^n$ is d-dimensional polytope. Then

$$\chi(\partial P) = 1 + (-1)^{d-1}, \quad \chi(\text{int}(P)) = (-1)^d.$$

Proof. Recall definition of χ by induction:

- Fix $a \in \mathbb{R}^n$ such that
$$H_\tau = \{x \in \text{aff}(P) : \langle a, x\rangle = \tau\}, \tau \in \mathbb{R},$$
is a family of $(d-1)$- dimensional hyperplanes in $\text{aff}(P)$. Note that $\partial P \cap H_{\tau_{\max}}$ is a point, $\partial P \cap H_\tau$ is a rectangle (frame-like), $\partial P \cap H_{\tau_{\min}}$ is a plane. Defined χ inductively:
$$\chi(\partial P) = \sum_{\tau\in\mathbb{R}} [\chi(\partial P \cap H_\tau) - \lim_{\varepsilon\to 0^+} \chi(\partial P \cap H_{\tau-\varepsilon})].$$
Note that $\chi(Q) = 1$ for all polytopes. Also, χ is linear on $\mathcal{K}(\mathbb{R}^d)$ and $\chi(\emptyset) = 0$.
  - For $\tau < \tau_{\min}$, $\chi(\partial \cap H_\tau) = 0$.
  - For $\tau = \tau_{\min}$, $\chi(\partial P \cap H_\tau) - \lim_{\varepsilon\to 0^+} \chi(\partial P \cap H_{\tau-\varepsilon}) = 1 - 0 = 1$, because $\partial P \cap H_\tau$ is a face of P, thus a polytope!
  - For $\tau_{\min} < \tau < \tau_{\max}$, $\partial P \cap H_\tau = \partial(P \cap H_\tau)$, where $P \cap H_\tau$ is $(d-1)$ - dimensional polytope. By induction, $\chi(\partial P \cap H_\tau) = 1 + (-1)^{d-2}$. Also, $\lim_{\varepsilon\to 0^+} \chi(\partial P \cap H_{\tau-\varepsilon}) = 1 + (-1)^{d-2}$.
  - For $\tau = \tau_{\max}$, $\chi(\partial P \cap H_\tau) = 1$. $\lim_{\varepsilon\to 0^+} \chi(\partial P \cap H_{\tau-\varepsilon}) = 1 + (-1)^{d-2}$.
  - For $\tau > \tau_{\max}$, summand is 0.

  Therefore, $\chi(\partial P) = 1 + \sum_\tau 0 + [1 - (1 + (-1)^{d-2})] = 1 + (-1)^{d-1}$.

$\chi(\text{int}(P))$: Note that P is a *disjoint* union of $\text{int}(P)$ and ∂P. Then $[P] = [\text{int}(P)] + [\partial P]$ (no inclusive/exclusive terms), therefore, $[\text{int}(P)] = [P] - [\partial P]$. χ is linear on $\mathcal{K}(\mathbb{R}^n)$, then
$$\chi(\text{int}(P)) = \chi(P) - \chi(\partial P) = 1 - [1 + (-1)^{d-1}] = (-1)^d.$$

□

Recall that

- χ is Euler characteristic,
- $\chi : \mathcal{K}(\mathbb{R}^n) \to \mathbb{R}$, where $\mathcal{K}$ is the algebra generated by indicator functions $[K]$ for compact convex $K \subset \mathbb{R}^n$.
- $\chi([K]) = \begin{cases} 1 & \text{if } K \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$
- For $A \subset \mathbb{R}^n$ such that $[A] \in \mathcal{K}(\mathbb{R}^n)$, write $\chi(A) = \chi([A])$.
- $\chi(\partial P) = \sum_{\tau\in\mathbb{R}} (\chi(\partial P \cap H_\tau) - \lim_{\varepsilon\to 0^+} \chi(\partial P \cap H_{\tau-\varepsilon}))$, where $\{H_\tau\}_{\tau\in\mathbb{R}}$ is a parallel family of hyperplanes in $\text{aff}(P) \subset \mathbb{R}^n$.

Remark 4.7. ∂P (resp. $\text{int}(P)$) are *relative* boundary (resp. interior) relative to $\text{aff}(P)$.

Remark 4.8.

- χ took cone of inclusion/exclusion for us.

- Including (-1)-dimensional and d-dimensional face yields

$$\sum_{i=-1}^{d}(-1)^i f_i(P) = \sum_{i=-1}^{d}(-1)^i \binom{d+1}{i+1} = 0.$$

  This is the case where P is a d-simplex.

Fact: Let $P_d = \{(f_0, \ldots, f_{d-1}) \in \mathbb{R}^d, \exists P, f_i(P) = f_i, \forall i\}$, where P is d-dimensional polytope. Then

$$\text{aff}(P_d) = \left\{(f_0, \ldots, f_{d-1}) \in \mathbb{R}^d : \sum_{i=0}^{d-1}(-1)^i f_i = 1 + (-1)^{d+1}\right\}.$$

I.e., no other (independent) affine linar transformation is satisfied by the f-vectors of *all* d-dimensional polytopes. But — can say more about f-vector if consider certain special families of polytopes. f-vectors of *simple* polytopes — Dehn-Sommervile relations (aka. duality).

Definition 4.3. A d-dimensional polytope such that every vertex is an intersection of exactly d facets is *simple*.

Example 4.2.

(a) A tetrahedron is simple.

(b) A cube is simple.

(c) A pyramid is not simple (because the vertex on the top is the intersection of 4 facets).

Generic polytopes are simple if "generic" means "bounded by generic hyperplanes".

Remark 4.9.

(a) Each vertex v_0 of P "looks locally like the origin in $\mathbb{R}^d_{\geq 0}$"

(b) Every vertex v_0 has exactly d neighbors (other vertices connected to v_0 by an edge).

(c) For every $k \leq d$ neigbors $v_1, \ldots, v_k$, $\exists!$ a k-dimensional face of P containing $v_0, v_1, \ldots, v_k$.

(d) For any polytope Q, there exists a simple polytope P such that

$$f_i(Q) \leq f_i(P), i = 0, \ldots, d.$$

  Sketch: If Q has facets supporting hyperplanes $H_j = \{x \in \mathbb{R}^d : \langle a_j, x\rangle = \beta_j\}$. perturb β_j's "generically" weakly increases the number of i-dimensional intersections. To bound above the "f-numbers" of polytopes just look at simple polytopes.

(e) P is simple, then P° is simplicial, i.e., every facet of P° is a simplex). Further, $f_i(P^\circ) = f_{d-i}(P)$. Then we can study simple polytopes to understand f-vectors of simplical spheres (at least polytopal ones).

4.2 h-vectors of simple polytopes

Definition 4.4. Let d-dimensional P be simple. Define

$$h_k(P) = \sum_{i=0}^{d} (-1)^{i-k} \binom{k}{i} f_i(P).$$

Hence, the *h-vector* $(h_0(P), \dots, h_d(P)) \subset \mathbb{R}^{d+1}$ is a linear transformation of the f-vecotr $(f_0(P), \dots, f_d(P))$. In particular,

$$f_i(P) = \sum_{k=0}^{d} \binom{k}{i} h_k(P). \tag{4.3}$$

Here,

$$\binom{k}{i} = \frac{k!}{i!\,(k-i)!} \text{ and } \binom{k}{i} = 0 \text{ if } k < i.$$

Indeed, if put

$$F(x) = \sum_{i=0}^{d} f_i(P) x^i, H(x) = \sum_{k=0}^{d} h_k(P) x^k,$$

then

$$F(x) = H(x+1), H(x) = F(x-1),$$

Thus, (4.3) *determines* the h-vector in terms of the $f_i(P), i = 0, \dots, d$.

Theorem 4.7 (Dehn-Sommerville Relations)**.** Suppose P is simple, $\dim P = d$. Then

$$h_k(P) = h_{d-k}(P), k = 0, 1, \dots, d.$$

Indeed, $k = 0$ case is Euler-Poincaré.

Proof. Recall that $h_k(P) = v_k(P, l), k = 0, \dots, d$, then we have k downstairs (lower) neighbors if and only if $d - k$ upstairs neighbors, thus have $d - k$ downstairs neighbors with respect to $-l$

$$v_k(P, l) = v_{d-k}(P, -l) = h_{d-k}(P).$$

□

To prove, look geometric/combinatorial interpretation of h-vecotr.

Theorem 4.8. Suppose $P \subset \mathbb{R}^d$, $\dim P = d$, P is simple. Fix linar functional $l : \mathbb{R}^d \to \mathbb{R}$ such that no two vertices u, v of P satisfy $l(u) = l(v)$. For vertices u, v, call u a *lower neighbor* of v if u is a neighbor of v and $l(u) < l(v)$. Let $v_k(P, l)$ be the number of vertices of P with exactly k lower neighbors. Then $h_k(P) = v_k(P, l), k = 0, \dots, d$ regardless of l.

Proof. Show that $v_k(P, l)$ satisfy (4.3) which determines $h_k(P)$. Double count

$$S = \{(F, v) : F \le P, \dim F = i, v \text{ is "max vertex" with respect to } l \text{ on } F\}.$$

On the other hand,

$$|S| = \sum_{F \subset P, \dim F = i} N_F = \sum_{F \subset P, \dim F = i} 1 = f_i(P),$$

where N_F is the number of vertices v of F with max l-value on F. On the other hand, putting $V(P)$ be the vertices of P. Then

$$|S| = \sum_{v \in V(P)} N_F.$$

Define *index* of vertex v to be the number of neighbors $u \in V$ of v such that

$$l(u) < l(v).$$

Then

$$|S| = \sum_{k=0}^{d} \sum_{v \in V(P), \text{index}(v)=k} N_F.$$

Recall that P is simple, i-dimensional faces containing v iff. i-dimensional subsets of the d neighbors of v. Then

$$|S| = \sum_{k=0}^{d} \sum_{v \in V(P), \text{index}(P)=k} \binom{k}{i} = \sum_{k=0}^{d} \binom{k}{i} v_k(P, l).$$

□

Remark 4.10.

- Euler-Poincaré for simple P, v is a special case ($k = 0$).
- All linear relations satisfied the h-vectors (equivalently f-vectors) of *all* simple polytopes are consequences of DS relations.
- f-vectors of simple polytopes are characterized.
- But f-vectors of arbitrary polytopes for dim ≥ 4 are still uncharacterized.

5 Convexity and the integer lattice

Definition 5.1. The integer lattice in $\mathbb{R}^n$ is

$$\mathbb{Z}^n = \{(\xi_1, \ldots, \xi_n) \in \mathbb{R}^n : \xi_i \in \mathbb{Z}, \forall i\}.$$

5.1 Minkowski's Geometry of Numbers

Theorem 5.1 (Minkowski's First Theorem (for $\mathbb{Z}^n$))**.** Suppose $K \subset \mathbb{R}^n$ is convex and centrally symmetric about 0. If $\text{vol}(K) > 2^n$, then

$$K \cap (\mathbb{Z}^n \backslash \{0\}) \neq \emptyset.$$

Proof. Suppose $K \subset \mathbb{R}^n$ convex, centrally sysmmetric, i.e., $K = -K$. Also have $\text{vol}(K) > 2^n$. It is equivalent that $\text{vol}(\frac{1}{2}K) > 1$. By Lemma 5.1, we have

$$\exists x, y \in \frac{1}{2}K, x - y \in \mathbb{Z}^n \backslash \{0\}.$$

Have $2x, 2y \in K$ with $0 \neq x - y \in \mathbb{Z}^n$. In fact, $x - y \in K$: centrally symmetry implies that $-2k \in K$. Convexity implies that $x - y = \frac{1}{2}(2x) + \frac{1}{2}(2y) \in K$. Therefore, $x - y \in K \cap (\mathbb{Z}^n \backslash \{0\})$. □

Remark 5.1. Can attain $\text{vol}(K) = 2^n$ with $2 \times 2 \times \cdots \times 2$ open parallelepiped.

Facts: For measurable subsets of $\mathbb{R}^n$.

- $\text{vol}(\lambda S) = |\lambda|^n \text{vol}(S)$.
- If $T \subset S$, then
$$\text{vol}(T) \leq \text{vol}(S).$$
- Let $\Pi : \{\sum_{i=1}^n \lambda_i e_i : 0 \leq \lambda_i < 1, \forall i\}$ be the fundamental parallelepiped for $\mathbb{Z}^n$, where $e_i, \forall i$ are standard basis vectors.
- Bounded convex sets have well-defined volume.

Lemma 5.1. Suppose $S \subset \mathbb{R}^n$ and $\text{vol}(S) > 1$. Then $\exists x, y \in S$ such that $x - y \in \mathbb{Z}^n \backslash \{0\}$.

Proof. Note that $\mathbb{Z}^n$ — translates of Π tile $\mathbb{R}^n$ — as a *disjoint* union. Then
$$S = \bigsqcup_{v \in \mathbb{Z}^n} S \cap (\Pi + v).$$
And $\text{vol}(S) > 1 \implies \sum_{v \in \mathbb{Z}^n} \text{vol}(S \cap (\Pi + v)) > 1$, where $S \cap (\Pi + 1)$ translate back to Π. For $v \in \mathbb{Z}^n$, let $S_v = (S - v) \cap \Pi$. Then
$$\sum_{v \in \mathbb{Z}^n} \text{vol}(S_v) > 1.$$
Then $S_v \subset \Pi (v \subset \mathbb{Z}^n)$ are *not* pairwise disjoint. $\exists u, v \in \mathbb{Z}^n (u \neq v)$ such that $S_u \cap S_v \neq \emptyset$. Let $w \in S_u \cap S_v$. Then
$$x - y = (w + u) - (w + v) = u - v \in \mathbb{Z}^n,$$
where $x = w + u$, $y = w + v$. □

5.2 Basic Theory of Lattices in $\mathbb{R}^n$

Definition 5.2. Given basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$, call
$$\Lambda = \langle v_1, \ldots, v_n \rangle_{\mathbb{Z}} = \left\{ \sum_{i=1}^n \lambda_i v_i : \lambda_i \in \mathbb{Z}, \forall i \right\}$$
a lattice (of *rank* n),
$$\Pi = \left\{ \sum_{i=1}^n \lambda_i v_i : 0 \leq \lambda_i < 1, \forall i \right\}$$
is a fundamental parallelepiped of Λ.

Example 5.1.

- $\mathbb{Z}$.
- $\langle (1,1), (1,-1) \rangle_{\mathbb{Z}}$.
- Fix $c_1, \ldots, c_m \in \mathbb{Z}^n$, $k_1, \ldots, k_m \in \mathbb{Z}$, then
$$\{a = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n : \langle c_i, a \rangle \equiv 0 \mod k_i, \forall i\}$$
is a sublattice of $\mathbb{Z}^n$.

Remark 5.2. A latice Λ of rank n is a *discrete* additive subgroup of $\mathbb{R}^n$ that spans $\mathbb{R}^n$, $\exists \varepsilon > 0$ such that $B_\varepsilon(0) \cap \Lambda = \{0\}$. In particular, given lattices $\Lambda' \subset \Lambda$ have the *index* of Λ' in Λ: $|\Lambda/\Lambda'|$.

Proposition 5.1. Suppose $\Lambda = \langle u_1, \dots, u_n\rangle_{\mathbb{Z}} = \langle v_1, \dots, v_n\rangle_{\mathbb{Z}}$. Let M be the transition matrix:

$$Mu_j = v_j, j = 1, \dots, n.$$

Then $\det(M) \in \{\pm 1\}$.

Proof. Have that entries of M and of M^{-1} are *integers*. Then $\det(M), \det(M^{-1}) \in \mathbb{Z}$. Also, $\det(M)\det(M^{-1}) = \det(I) = 1$. Therefore, $\det M = \det(M^{-1}) \in \{\pm 1\}$. □

Example 5.2. $\mathbb{Z}^n = \langle (1,0), (0,1)\rangle_{\mathbb{Z}} = \langle (1,1), (1,2)\rangle_{\mathbb{Z}}$.

Corollary 5.1. Given $\Lambda = \langle v_1, \dots, v_n\rangle_{\mathbb{Z}}$,

$$\det(\Lambda) = |\det[v_1 \cdots v_n]|$$

does *not* depend on the basis chosen. In particular, $\mathrm{vol}(\Pi)$ does not depend on the basis.

Theorem 5.2. Suppose $\Lambda \subset \mathbb{Z}^n$ is a lattice, with fundamental parallelepiped Π. Let Π_0 be fundamental parallelepiped of $\mathbb{Z}^n$. Then

$$\det(\Lambda) = |\mathbb{Z}^n/\Lambda| = \mathrm{vol}(\Pi) = |\Pi \cap \mathbb{Z}^n|.$$

Proof. (sketch/intuition) Note Π, Π_0 both (nearly) tile a large ball $B_\rho(0)$. Ratio of number of Π_0-tiles to Π-tiles is

$$\frac{\mathrm{vol}(\Pi)}{\mathrm{vol}(\Pi_0)} a = \mathrm{vol}(\Pi) = \det(\Lambda).$$

On the other hand, get *one* Π-tile for each Λ-point in $B_\rho(0)$ and likewise one Π_0-tile for each $\mathbb{Z}^n$ point, so the ratio of number of Π_0-tiles to Π-tiles is the number of $\mathbb{Z}^n$ points in each Π-tile. □

Proposition 5.2. Let $\Lambda = \{(\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n : \langle c_i, a\rangle \equiv 0 \mod k_i, \forall i\}$, where $c_i \in \mathbb{Z}^n$, $k_i \in \mathbb{Z}$. Then

$$\det(\Lambda) \le \prod_{i=1}^{m} k_i.$$

Proof. Note that $\mathrm{rank}(\Lambda) = n$, because $k_1 k_2 \cdots k_m \mathbb{Z}^n \subset \Lambda$. Then

$$\det(\Lambda) = |\mathbb{Z}^n/\Lambda| = N_{\text{cosets}},$$

where N_{cosets} is the number of cosets of Λ in $\mathbb{Z}^n$. Have injective map

$$\varphi : a + \Lambda \mapsto \oplus \mathbb{Z}/(k_i\mathbb{Z}), a \mapsto (\langle c_1, a\rangle, \langle c_2, a\rangle, \dots, \langle c_m, a\rangle).$$

Codomain has $\prod_{i=1}^m k_i$ elements, which gives upper bound on cardinality of domain $\mathbb{Z}^n/\Lambda$. □

Theorem 5.3 (Minkowski's 1st Theorem for lattice $\Lambda \subset \mathbb{R}^n$)**.** Suppose $K \subset \mathbb{R}^n$ is convex, centrally symmetric about 0. Suppose $\mathrm{vol}(K) > 2^n \det(\lambda)$.

Proof. Apply Minkowski for $\mathbb{Z}^n$ to $A^{-1}(K)$, where $A : e_i \mapsto v_i$, $\Lambda = \langle v_1, \dots, v_n\rangle_{\mathbb{Z}}$. □

Toward Lagrange's "4 squares" Theorem.

Theorem 5.4. Suppose $n \in \mathbb{Z}_{\geq 0}$, then $\exists k_1, \ldots, k_4 \in \mathbb{Z}$ such that $n = k_1^2 + k_2^2 + k_3^2 + k_4^2$.

Proof. Seek $(k_1, \ldots, k_4) \in \mathbb{Z}^4$ such that $k_1^2 + \cdots + k_4^2 = n$. So, seek $(k_1, \ldots, k_4) \in \mathbb{Z}^4$ such that $k_1^2 + k_2^2 + k_3^2 + k_4^2 \equiv 0 \mod n$ and

$$0 < k_1^2 + k_2^2 + k_3^2 + k_4^2 < 2n \iff 0 \neq (k_1, \ldots, k_4) \in B_{\sqrt{2n}}(0).$$

In other words, seek nozero point in $S_n \cap B_{\sqrt{2n}}(0)$ where $S_n = \{(k_1, \ldots, k_n) \in \mathbb{Z}^4 : k_1^2 + \cdots + k_4^2 \equiv 0 \mod n\}$. Problem: S_n is not a lattice (e.g., $(\pm 2, 1, 0, 0) \in S_5$, but $\|(0, 2, 0, 0)\|^2 = 4 \not\equiv 0 \mod 5$). Solution: Find lattice $\Lambda \subset S_n$ that Minkowski forces to place a nozero element in $B_{\sqrt{2n}}(0)$. □

Example 5.3.

- $0 = 0^2 + \cdots + 0^2$.
- $1 = 1^2 + 0^2 + \cdots + 0^2$.
- $7 = 2^2 + 1^2 + 1^2 + 1^2$.

Note that 4 squares is minimal.

Remark 5.3. Without loss of generality, $n = p$, where p is prime. Because, given

$$m = j_1^2 + j_2^2 + j_3^2 + j_4^2, n = k_1^2 + k_2^2 + k_3^2 + k_4^2,$$

then $mn = l_1^2 + l_2^2 + l_3^2 + l_4^2$, where

$$\begin{aligned} l_1 &= j_1k_1 - j_2k_2 - j_3k_3 - j_4k_4, \\ l_2 &= j_1k_2 + j_2k_1 + j_3k_4 - j_4k_3, \\ l_3 &= j_1k_3 - j_2k_4 + j_3k_1 + j_4k_2, \\ l_4 &= j_1k_4 + j_2k_3 - j_3k_2 + j_4k_1. \end{aligned}$$

Proof. Suppose $n = p$ is prime. Let $S_p = \{(k_1, k_2, k_3, k_4) \in \mathbb{Z}^4 : k_1^2 + k_2^2 + k_3^2 + k_4^2 \equiv 0 \mod p\}$. Idea: Use Minkowski's 1st Theorem to show

$$B_{\sqrt{2p}}(0) \cap (S_p \backslash \{0\}) \neq \emptyset.$$

Problem: S_p is not lattice. Need lattice $\Lambda_p \subset S_p$ such that $\mathrm{vol}(B_{\sqrt{2p}}(0)) > 2^4 \det(\Lambda)$. Idea: Impose conditions on k_1, k_2, k_3, k_4,

$$\begin{aligned} k_1^2 + k_2^2 + k_3^2 + k_4^2 \equiv 0 \mod p &\iff k_1^2 + k_2^2 + (\alpha k_1 + \beta k_2)^2 + (\beta k_1 - \alpha k_2)^2 \equiv 0 \mod p \\ &\iff (1 + \alpha^2 + \beta^2)k_1^2 + (1 + \alpha^2 + \beta^2)k_2^2 \equiv 0 \mod p. \end{aligned}$$

Then it holds for *all* $k_1, k_2 \in \mathbb{Z}$ if $\alpha, \beta \in \mathbb{Z}$ are such that $\alpha^2 + \beta^2 \equiv -1 \mod p$. □

Lemma 5.2. Such α, β exist. That is, suppose p is prime, then $\exists \alpha, \beta \in \mathbb{F}_p$ such that $\alpha^2 + \beta^2 = -1$.

Proof. For $p = 2$, trivial. For $p \geq 3$, $A = \{\alpha^2 : \alpha \in \mathbb{F}_p\}$, $B = \{-1 - \beta^2 : \beta \in \mathbb{F}_p\}$. Note that $\alpha_1^2 = \alpha_2^2 \iff (\alpha_1 - \alpha_2)(\alpha_1 + \alpha_2) = 0 \iff \alpha_1 = \pm\alpha_2$, and $\alpha_2 \neq -\alpha_2$ unless $\alpha_2 = 0$. So $\mathbb{F}_p \backslash \{0\} \to A \backslash \{0\}$ $(\alpha \mapsto \alpha^2)$ is $2 - 1$. Then

$$|A| = \frac{|\mathbb{F}_p| - 1}{2} + 1 = \frac{p + 1}{2}.$$

And

$$B = -1 - A \implies |B| = |A| = \frac{p + 1}{2}.$$

A, B each have $> \frac{1}{2}$ of $\mathbb{F}_p$, $A \cap B \neq \emptyset$. Then $\exists \alpha, \beta$ such that $\alpha^2 = 1 - \beta^2$. □

Fix $\alpha, \beta \in \mathbb{Z}$ such that (**) holds. Let

$$\Lambda_p = \{(k_1, k_2, k_3, k_4) \in \mathbb{Z}^4 : k_3 \equiv \alpha k_1 + \beta k_2 \mod p, k_4 \equiv \beta k_1 - \alpha k_2 \mod p\}.$$

I.e., $\Lambda_p = \{k \in \mathbb{Z}^4 : \langle c_i, k\rangle \equiv 0 \mod p, i = 1, 2\}$, where $c_1 = (-\alpha, -\beta, 1, 0)$, $c_2 = (-\beta, \alpha, 0, 1)$. Then $\Lambda_p \subset S_p$ is a lattice, and $\det(\Lambda_p) \leq p^2$. On the other hand, by calculus, in $\mathbb{R}^4$, $\mathrm{vol}(B_1(0)) = \frac{\pi^2}{2}$, then

$$\mathrm{vol}(B_{\sqrt{2p}}(0)) = (\sqrt{2p})^4 \frac{\pi^2}{2} = 2\pi^2 p^2 > 2^4 p^2 \geq \det(\Lambda_p).$$

Then $M : B_{\sqrt{2p}}(0) \cap (\Lambda_p \backslash \{0\}) \neq \emptyset$.

5.3 Lattice-point enumeratoins in rational polytopes

Definition 5.3. Polytope $P \subset \mathbb{R}^n$ is *rational*, if $\mathrm{vert}(P) \subset \mathbb{Q}^n$.

We seek to compute $|P \cap \mathbb{Z}^n|$.

Definition 5.4. Suppose $S \subset \mathbb{R}^n$. The *characteristic polynomial* of S is the formal Laurent polynomial

$$f_S(x_1, \ldots, x_n) = \sum_{a \in S \cap \mathbb{Z}^n} \mathbf{x}^a.$$

where $\mathbf{x}^a = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, $a = (\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Example 5.4. $f_P(x, y) = x + xy^{-1} + xy + 1 + y + y^2 + x^{-1}$.

Note that $|P \cap \mathbb{Z}^2| = f_P(1, 1)$. Bad news: $f_S(\mathbf{x})$ can be very big — even infinite. Good news: Often, "short" encodings of $f_S(\mathbf{x})$ exist (even are computable "quickly").

Example 5.5. Let $C : \mathrm{pos}\{v_1, v_2\} \subset \mathbb{R}^2$, where $v_1 = (2, 1)$, $v_2 = (0, 1)$. Let $\Pi = \{\alpha_1 v_1 + \alpha_2 v_2 : 0 \leq \alpha_1, \alpha_2 < 1\}$, then $f_C(x, y)$ has infinitely many terms.

$$f_C(x, y) = 1 + y + xy + x^2 y + y^2 + y^2 x + \cdots$$

But has "short" encoding:

$$f_C(x, y) = \frac{1 + xy}{(1 - x^2 y)(1 - y)}. \tag{5.1}$$

Why? Note: Every $a \in C \cap \mathbb{Z}^2$ can be written

$$a = b + iv_1 + jv_2, b \in \Pi.$$

in exactly *one* way with $i, j \in \mathbb{Z}_{\geq 0}$.

The right-hand side of (5.1) is

$$\frac{1 + xy}{(1 - x^2 y)(1 - y)} = (1+xy) \cdot \left(\sum_{j \in \mathbb{Z}_{\geq 0}} y^j\right) \left(\sum_{i \in \mathbb{Z}_{\geq 0}} (x^2 y)^i\right) = \left(\sum_{b \in \Pi \cap \mathbb{Z}^n} \mathbf{x}^b\right) \left(\sum_{i \in \mathbb{Z}_{\geq 0}} \mathbf{x}^{iv_1}\right) \left(\sum_{j \in \mathbb{Z}_{\geq 0}} \mathbf{x}^{jv_2}\right),$$

where $\mathbf{x}^{v_1} = x^2 y$, $\mathbf{x}^{v_2} = y$. Expanding gives

$$\sum_{b \in \Pi \cap \mathbb{Z}^n, i, j \in \mathbb{Z}_{\geq 0}} \mathbf{x}^{b + iv_1 + jv_2} = \sum_{a \in C \cap \mathbb{Z}^n} \mathbf{x}^a = f_C(\mathbf{x}).$$

Definition 5.5. A cone $C \subset \mathbb{R}^n$ is

- *rational* if $C = \text{pos}\{v_1, \ldots, v_r\}$ for some $v_i \in \mathbb{Q}^n$ (equivalently, for some $v_i \in \mathbb{Z}^n$).
- *simple* if $C = \text{pos}\{v_1, \ldots, v_r\}$ with $r = \dim(C)$. That is, $C = \text{pos}\{v_1, \ldots, v_r\}$ is simple if $v_1, \ldots, v_r$ are linearly independent.

Generalizing the example:

Theorem 5.5. Suppose $C = \text{pos}\{v_1, \ldots, v_r\}$ is a rational simple cone ($v_i \in \mathbb{Z}^n, i = 1, \ldots, r$). Let $\Pi = \{\sum_{i=1}^r \lambda_i v_i : 0 \le \lambda_i < 1\}$. Then

$$f_C(\mathbf{x}) = \frac{f_\Pi(\mathbf{x})}{\prod_{i=1}^r (1 - \mathbf{x}^{v_i})}.$$

Generalizing to non-simple cones:

Definition 5.6. A cone is *pointed* if it contains no line.

Example 5.6. In $\mathbb{R}^2$, the cone cannot be the region that $y \ge 0$, and in $\mathbb{R}^3$, it cannot be $z \ge 0$.

Theorem 5.6. Suppose $C \subset \mathbb{R}^n$ is a rational *pointed* cone. Then

$$[C] = \sum_i \varepsilon_i [C_i],$$

where C_i are rational *simple* cones, and $\varepsilon_i \in \{\pm 1\}$.

Proof. Sketch:

- C has base P that is a rational polytope.
- Triangulate P using (e.g.) *regular triangulations*.
- Fix "generic" height function $t : \text{vert}(P) \to \mathbb{R}$. Let $P' = \text{conv}\{(v, t(v)) : v \in \text{vert}(P)\}$.
- By "generic-ness", all faces of P' are simplices.
- Project "bottom" faces "down" to P to get a triangulation $\Delta = \{T_i\}$. ("Triangulation" means if T_i is a face of $T_j \in \Delta$, then $T_i \in \Delta$, and if $T_i, T_j \in \Delta$, then $T_i \cap T_j \in \Delta$). Then let $C_i = \text{pos}(T_i)$.
- By inclusion/exclusion, get $[C] = \sum_i \varepsilon_i [C_i]$.

□

Remark 5.4. $[C] = \sum_i \varepsilon_i [C_i] \implies f_C(\mathbf{x}) = \sum_i \varepsilon_i f_{C_i}(\mathbf{x})$.

Corollary 5.2. Suppose $C \subset \mathbb{R}^n$ is a rational pointed cone, then $f_C(\mathbf{x})$ is a rational function.

Proof. Let $\Delta = \{C_i\}$ be a triangulation of C into simple cones C_i. Let $C = \text{pos}(V_i), V_i \subset \mathbb{Z}_i^n$, finite and linearly independent. Then we have

$$[C] = \sum_i \varepsilon_i [C_i],$$

and so

$$f_C(\mathbf{x}) = \sum_{C_i \in \Delta} \varepsilon_i \frac{f_{\Pi_i}(\mathbf{x})}{\prod_{v \in V_i} (1 - \mathbf{x}^v)},$$

where Π_i is the fundamental parallelepiped of C_i. Multiply through by common denominators to get a rational function. □

Can use to compute $f_P(\mathbf{x})$ for a (bounded) polytope using *Brion's Theorem*.

Example 5.7. Let $P = [0,5] \subset \mathbb{R}$, $C_0 = [0,\infty)$, $C_1 = (-\infty, 5]$,

$$f_P(x) = x^0 + x^1 + \cdots + x^5,$$
$$f_{C_0}(x) = 1 + x^1 + x^2 + \cdots = \frac{1}{1-x},$$
$$f_{C_1}(x) = x^5(1 + x^{-1} + x^{-2} + \cdots) = x^5 \frac{1}{1-x^{-1}} = \frac{-x^6}{1-x}.$$

Note that

$$\tilde{f}_{C_0}(x) + \tilde{f}_{C_1}(x) = \frac{1}{1-x} + \frac{-x^6}{1-x} = \frac{1-x^6}{1-x} = \tilde{f}_P(x).$$

Take away: need to distinguish the formal series $f_C(\mathbf{x}) = \sum_{a \in C \cap \mathbb{Z}^n} \mathbf{x}^a$ and a corresponding rational function $\tilde{f}_C(\mathbf{x})$. Can't do if C *contains* a line:

Example 5.8. Let $C = \mathbb{R}$, then $f_C(x) = \sum_{i \in \mathbb{Z}} x^i$. Therefore, $x f_C(\mathbf{x}) = f_C(\mathbf{x})$. That implies $(x-1) f_C(\mathbf{x}) = 0$. So, would wnat to set $\tilde{f}_C(\mathbf{x}) = 0$. Extends to any cone C containing line through, say $a \in \mathbb{Z}^n \backslash \{0\}$:

$$(x^a - 1) f_C(\mathbf{x}) = 0.$$

Definition 5.7. Suppose $P \subset \mathbb{R}^n$ is a polytope, v is a vertex of P. Then the *vertex cone* of P at v is

$$C(P,v) = \{v + \lambda a : \lambda \in \mathbb{R}_{\geq 0}, a \in \mathbb{R}^n, \exists \varepsilon > 0, v + \varepsilon a \in P\}.$$

Theorem 5.7 (Brion, 1988)**.** Suppose $P \subset \mathbb{R}^n$ is a rational polytope. Then

$$\tilde{f}_P(\mathbf{x}) = \sum_{v \in \mathrm{vert}(P)} \tilde{f}_{C(P,v)}(\mathbf{x}).$$

Further reading: On a theorem of Brion.

Lemma 5.3. Have a "suitable" rational function $\tilde{f}_{C(P,v)}(\mathbf{x})$ for $C(P,v)$, where "suitable" means not throwing away information.

Proof. Problem is $C(P,v)$ is not a *cone* unless $v = 0$. Let $C = C(P,v)$. Add a dimension:

$$C' = \overline{\mathrm{pos}\{(a,1) : a \in C\}} \subset \mathbb{R}^{n+1}$$

is a rational polyhedral cone containing no line. So we have

$$\tilde{f}_{C'}(\mathbf{x}) = \tilde{f}_{C'}(x_1, \cdots, x_n, x_{n+1})$$

as a rational functions. Note that C is the slice at height of 1 of C'. Recover $\tilde{f}_C(x_1, \ldots, x_n)$ from $\tilde{f}_{C'}(x_1, \ldots, x_n, x_{n+1})$ using *formal* differentiation

$$\tilde{f}_C(\mathbf{x}) = \left(\frac{\partial}{\partial x_{n+1}} \tilde{f}_{C'}(x_1, \ldots, x_n, x_{n+1}) \right) \Big|_{x_{n+1}=0}.$$

To see why, think in terms of the Laurent series, $\partial/\partial x_{n+1}$ kills monomials of $f_{C'}(\mathbf{x})$ ending in $x_{n+1}^i, i = 0$ and $x_{n+1} = 0$ kills monomials of $f_{C'}(\mathbf{x})$ ending in $x_{n+1}^i, i \geq 2$. Computing $f_\Pi(\mathbf{x})$. Note that if C is *unimodular* cone, that is, $\|\det[v_1, \ldots, v_r] = 1$, in other words, Π is a fundamental parallelepiped of $\mathbb{Z}^n$, then $f_\Pi(\mathbf{x}) = 1$. □

Theorem 5.8 (Barvinok, 1993)**.** Suppose $C \subset \mathbb{R}^n$ is a rational simple cone. Then

$$[C] = \sum_i \varepsilon[C_i],$$

where C_i are *unimodular* cones. Moreover, the number of C_i is "small" - meaning is $\log_2(|\det(\text{generators of } C)|)$.

Proof. (sketch) Without loss of generality, assume that $\dim(C) = n$. Let $D = |\det[v_1 \cdots v_n] \in \mathbb{Z}$.

$$\Gamma = \left\{ \sum_{i=1}^n \lambda_i v_i : |\lambda_i| \leq D^{-1/n} \right\}.$$

I.e., Γ is "2^n copies of $\overline{\Pi}$, each copy scaled by $D^{-1/n}$". That implies $\text{vol}(\Gamma) = 2^n$. Also, Γ is centrally symmetric about the origin, thus Minkowski applies. That is, $\exists w \neq 0$ such that $w \in \Gamma \cap \mathbb{Z}^n$. Idea: Use w to "partition" C into simple cones "closer to unimodular".

(a) Case I: "inside" the cone: $w \in \Gamma \cap C$.

(b) Case II: "outside" the cone: $w \in \Gamma \backslash C$. We have $w = \sum_i \lambda_i v_i, |\lambda_i| \leq D^{-1/n}$ and

$$[v_1 v_2 v_3] = [v_1 v_2 w] + [v_1 v_3 w] - [v_1 w] - [v_2 v_3 w] + [v_2 v_3].$$

Note that new fundamental parallelepiped have volumes looking like

$$\begin{aligned} |\det[v_1 v_2 \cdots v_{n-1} w]| &= |\sum_{i=1}^n \lambda_i \det[v_1 \cdots v_{n-1} v_i] \\ &= |\lambda_n \det[v_1 \cdots v_{n-1} v_n] \\ &\leq D^{-1/n} D = D^{\frac{n-1}{n}} < D. \end{aligned}$$

At each step, $D \in \mathbb{Z}_{\geq 1}$, decreasing upper bound. The number of steps is at most m such that $(\cdots((D^{\frac{n-1}{n}})^{\frac{n-1}{n}} \cdots)^{\frac{n-1}{n}} = D^{\left(\frac{n-1}{n}\right)^m} < 2$. Therefore,

$$m < \frac{\log_2 \log_2 D}{\log_2 \left(\frac{n}{n-1}\right)}.$$

For *fixed* bound is $\log \log D$, i.e., logarithm of the "input size" of C. And generated less than n full-dim cones at each steps. Then new terms in the expression of $f_C(\mathbf{x})$ is at most n^m. □

Definition 5.8. A cone C is *unimodular* if $v_1, \ldots, v_r$ generate the lattice $\text{lin}(C) \cap \mathbb{Z}^n$.

E.g., if $\dim(C) = n$, we have $\mathbb{Z}^n = \langle v_1, \ldots, v_n \rangle_{\mathbb{Z}}$.

Remark 5.5.

- If $\dim(C) = n$ $(r = n)$,

$$\begin{aligned} C \text{ is unimodular} &\iff |\det[v_1 v_2 \cdots v_n]| = 1 \\ &\iff \text{vol}(\Pi) = 1 \\ &\iff \Pi \cap \mathbb{Z}^n = \{0\}, \end{aligned}$$

  where $\Pi = \{\sum_i \alpha_i v_i : 0 \leq \alpha_i < 1\}$.

- C is unimodular, then
$$f_C(\mathbf{x}) = \frac{1}{\prod_{i=1}^{r}(1-\mathbf{x}^{v_i})}.$$

Example 5.9. $[C_2] = [C_1 \cup C_2] - [C_1] + [C_3]$, then

$$f_C(x,y) = f_{C_1\cup C_2}(x,y) - f_{C_1}(x,y) + f_{C_3}(x,y) = \frac{1}{(1-x)(1-y)} - \frac{1}{(1-xy^5)(1-y)} + \frac{1}{1-xy^5}.$$

On the other hand,
$$f_{C_2}(x,y) = \frac{1+xy+xy^2+\cdots+xy^4}{(1-xy^5)(1-x)}.$$

Example 5.10.

$$f_C(x,y) = \frac{1+xy+\cdots+xy^4}{(1-x)(1-xy^5)} + x^2y^0\frac{1+x^{-1}y+\cdots+x^{-1}y^4}{(1-x^{-1})(1-x^{-1}y^5)} + xy^5\frac{1+y^{-1}+y^{-2}+\cdots+y^{-9}}{(1-xy^{-5})(1-x^{-1}y^{-5})}.$$

$|P\cap\mathbb{Z}^n| = f_C(1,1)$.