Mathematical Tools from Linear and Convex Optimization (MTLCO)

Ref : [2012][J.A.D. Loera, etal]
Algebraic and geometric ideas in the theory of discrete optimization. - Chap 1

1.1 Convex sets and polyhedra. 2/5/2019

We only consider subsets of Euclidean sp. $\mathbb{R}^n$.

Def

$S \subseteq \mathbb{R}^n$ is $\boxed{\text{convex}}$ if

$$\lambda x_1 + (1-\lambda) x_2 \in S, \ \forall \ \begin{matrix} \lambda \in [0,1] \\ x_1, x_2 \in S \end{matrix}$$

Convention: The empty set $\emptyset$ is convex.

Def

Given $A = \{x_1, \cdots, x_m\} \subseteq \mathbb{R}^n$. The linear combination $\sum_{i=1}^{m} r_i x_i$ is called

(i) an $\boxed{\text{affine combination}}$ if $\sum_i r_i = 1$;

(ii) a $\boxed{\text{conic combination}}$ if $r_i \geq 0, \ \forall i$;

(iii) a $\boxed{\text{convex combination}}$ if both affine and conic.

Rmk :

① Any intersection of convex sets is again convex. finite or infinite.

Def

For $A \subseteq \mathbb{R}^n$, the $\boxed{\text{convex hull}}$ of A is

$\boxed{\text{conv}(A)}$:= the intersection of all convex sets containing A.

= smallest convex set containing A.

② Linear transformations map convex sets to convex sets.

Def Given $c \neq 0$ in $\mathbb{R}^n$.

Given $\alpha \in \mathbb{R}$, and $c \in \mathbb{R}^n$.

Define $f_c : \mathbb{R}^n \to \mathbb{R}, \ x \longmapsto c^T x$.

$$\boxed{H_\alpha} := \{x \in \mathbb{R}^n : f_c(x) = \alpha\}.$$

$$\boxed{H_\alpha^+} := \{x \in \mathbb{R}^n : f_c(x) \geq \alpha\}.$$

$$\boxed{H_\alpha^-} := \{x \in \mathbb{R}^n : f_c(x) \leq \alpha\}.$$

H_α is called an $\boxed{\text{affine hyperplane}}$ or simply $\boxed{\text{hyperplane}}$.

H_α^+ / H_α^- are called $\boxed{\text{halfspaces}}$.

Rmk :

H_α, H_α^+, and H_α^- are convex sets.

Def

A set of the type

$$P = \{x \in \mathbb{R}^n : c_i^T x \leq \beta_i, \ i = 1, \cdots, m\}$$

i.e. intersection of half-spaces.

is called a $\boxed{\text{polyhedron}}$, where $c_i \in \mathbb{R}^n, \beta_i \in \mathbb{R}$.

Rmk :

Putting $C = \begin{bmatrix} - c_1 - \\ \vdots \\ - c_m - \end{bmatrix}$ and $\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$, we may write $P = \{x \in \mathbb{R}^n \mid Cx \leq \beta\}$, where "$\leq$" is understood componentwise.

Lem 1.1.3.

Every polyhedron is a convex set.

<Pf>

Half-spaces are convex and a polyhedron is the intersection of some half-spaces. #

Def

A $\boxed{\text{polytope}}$ is the convex hull of a finite set in $\mathbb{R}^n$.

Rmk : Near-future goal is Weyl-Minkowski thm

Polytopes $\Longleftrightarrow$ bounded polyhedra.

Lem 1.1.5

For $A \subseteq \mathbb{R}^n$,

$$\mathrm{conv}(A) = \left\{ \sum_{i=1}^{m} \gamma_i x_i \;\middle|\; \begin{array}{l} x_1, \cdots, x_m \in A, \text{ for some } m \\ \gamma_i \geq 0, \forall i, \sum_{i=1}^{m} \gamma_i = 1 \end{array} \right\}$$

$= \{$ all finite convex combinations among elements of $A \}$.

⟨Pf⟩

Denote $B := \mathrm{RHS}$ (right hand side).

"$\subseteq$"

It's easy to check B is convex. (by def).

Moreover, taking one element $x_1 \in A$ and $\gamma_1 = 1$, $\gamma_1 x_1 = x_1$ is a convex combination of $\{x_1\}$.

Thus, $A \subseteq B$. $\Rightarrow \mathrm{conv}(A) \subseteq \mathrm{conv}(B) \underset{B:\text{ convex}}{=} B$. #

"$\supseteq$"

We need to prove every finite convex combinatn among elements of A lies in $\mathrm{conv}(A)$.

Induction on number of nonzero terms in $\sum_i \gamma_i x_i$.

Abusing notation, denote it by m.

For $m=1$, $\gamma_1 = 1$, the result is obvious.

Assuming induction hypothesis, given $\sum_{i=1}^{m+1} \gamma_i x_i$.

$\gamma := \sum_{i=1}^{m} \gamma_i$. $(= 1 - \gamma_{m+1})$. $(0 < \gamma < 1)$.

By hypothesis, $\sum_{i=1}^{m} \frac{\gamma_i}{\gamma} x_i \in \mathrm{conv}(A)$.

$\because x_{m+1} \in \mathrm{conv}(A)$ and $\mathrm{conv}(A)$ is convex

$\therefore \gamma \cdot \left(\sum_{i=1}^{m} \frac{\gamma_i}{\gamma} x_i \right) + (1-\gamma) x_{m+1} \in \mathrm{conv}(A)$

i.e. $\sum_{i=1}^{m} \gamma_i x_i + \gamma_{m+1} x_{m+1} \in \mathrm{conv}(A)$. #

Def

$C \subseteq \mathbb{R}^n$ is called a **cone** if

$\lambda x + \mu y \in C$, $\forall\, x, y \in C$, $\lambda, \mu \in \mathbb{R}_+ := [0, \infty)$.

① A cone C is called **polyhedral** if

$C = \{ x \mid Ax \geq 0 \}$, for some A.

check: necessarily a cone.

② A cone C is **finitely generated** if

$$C = \left\{ \sum_{i=1}^{m} \gamma_i x_i \;\middle|\; \gamma_i \geq 0, \forall i, x_1, \cdots, x_m \in S \right\},$$

where $S \subseteq \mathbb{R}^n$ is a finite set.

Rmk

Cones defined here are sometimes called convex cones in other literatures.

Thm 1.1.7 (Caratheodory theorem).

$S \subseteq \mathbb{R}^n$, $x \in \mathrm{conv}(S)$.

Then x is the convex combination of at most $(n+1)$ points in S.

⟨Pf⟩ (Prove by contradiction).

Take $\bar{x} \in \mathrm{conv}(S)$.

By Lem 1.1.5,

write $\bar{x} = \sum_{i=1}^{m} \lambda_i x_i$ w/ $x_i \in S$, $\lambda_i > 0$, $\sum_{i=1}^{m} \lambda_i = 1$ and m smallest possible

Suppose $m \geq n+2$.

claim: $\exists\, \gamma_i \in \mathbb{R}$ s.t. $\sum_{i=1}^{m} \gamma_i x_i = 0$, $\sum_{i=1}^{m} \gamma_i = 0$ w/ at least one $\gamma_i > 0$

⟨Pf of claim⟩

$\because m-1 \geq n+1$

$\therefore x_2 - x_1, \cdots, x_m - x_1$ is linearly dependent.

$\Rightarrow \exists\, \gamma_2, \cdots, \gamma_m$ not all zero s.t. $\sum_{i=2}^{m} \gamma_i (x_i - x_1) = 0$

Define $\gamma_1 = -(\gamma_2 + \cdots + \gamma_m)$.

Then $\sum_{i=1}^{m} \gamma_i = 0$, $\sum_{i=1}^{m} \gamma_i x_i = 0$.

$\because \gamma_2, \cdots, \gamma_m$ are not all zero and $\sum_{i=1}^{m} \gamma_i = 0$

$\therefore$ at least one $\gamma_i > 0$. # of claim

Let's summarize what we have so far:

$x_1 \quad x_2 \quad \cdots \quad x_m$

$\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_m \to \bar{x} = \sum \lambda_i x_i, \ \sum \lambda_i = 1, \ \lambda_i \geq 0$

$\gamma_1 \quad \gamma_2 \quad \cdots \quad \gamma_m \to \sum \gamma_i = 0$, $\sum \gamma_i x_i = 0$, $\gamma_{i_0} > 0$, for some i_0.

If we can find $\alpha > 0$ s.t.

$\lambda_1 - \alpha\gamma_1 \geq 0, \ \ldots, \ \lambda_m - \alpha\gamma_m \geq 0$ $(*_1)$, and

$\lambda_k - \alpha\gamma_k = 0$, for some k, $(*_2)$

then

$$\bar{x} = \sum_{i=1}^{m} (\lambda_i - \alpha\gamma_i) x_i \quad \text{and} \quad \sum_{i=1}^{m} (\lambda_i - \alpha\gamma_i) = 1$$

i.e. $\bar{x}$ can be expressed as a convex comb. of $(m-1)$ elements of A, contradicting w/ minimality of m and we will be done.

To guarantee $\lambda_i - \alpha\gamma_i \geq 0$,

case1 $\gamma_i \leq 0$, then it's automatically true no matter what $\alpha > 0$ is.

case2 $\gamma_i > 0$, then it's equivalent to $\alpha \leq \frac{\lambda_i}{\gamma_i}$

Therefore, choosing $\alpha = \min\{\lambda_i/\gamma_i \mid \gamma_i > 0\}$ will guarantee $\lambda_i - \alpha\gamma_i \geq 0, \ \forall i$. i.e. $(*_1)$

Moreover, for $k = \underset{i}{\operatorname{argmin}} \{\lambda_i/\gamma_i \mid \gamma_i > 0\}$,

$$\lambda_k - \alpha\gamma_k = \lambda_k - \frac{\lambda_k}{\gamma_k} \cdot \gamma_k = 0. \quad \text{i.e. } (*_2)$$

Thus, the desired α does exist and we reach a contradiction and complete the proof. #

Lem 1.1.8 (Radon's theorem).

$S \subseteq \mathbb{R}^n$ w/ $\#(S) = n+2$.

Then $\exists A \neq \emptyset \neq B$ w/ $S = A \amalg B$ s.t. $\operatorname{conv}(A) \cap \operatorname{conv}(B) \neq \emptyset$.

disjoint union

a partition of S

<Pf> (Prove by construction).

By claim in the Pf of Thm 1.1.7,

$\exists \gamma_i \in \mathbb{R}$ s.t. $\sum_{i=1}^{n+2} \gamma_i x_i = 0$, $\sum_{i=1}^{n+2} \gamma_i = 0$ w/ at least one $\gamma_i > 0$

$P_1 := \{i \mid \gamma_i > 0\}$. $A := \{x_i \mid i \in P_1\}$.

$P_2 := \{j \mid \gamma_j < 0\}$. $B := \{x_j \mid j \in P_2\}$.

$A \neq \emptyset$ by condition. $B \neq \emptyset$ since $\sum_{i=1}^{n+2} \gamma_i = 0$.

$\gamma := \sum_{i \in P_1} \gamma_i$. (>0)

claim: $\operatorname{conv}(A) \cap \operatorname{conv}(B) \neq \emptyset$.

$$\bar{x} = \sum_{i \in P_1} \left(\frac{\gamma_i}{\gamma}\right) \cdot x_i \in \operatorname{conv}(A).$$

$\because \sum_{i=1}^{n+2} \gamma_i x_i = 0$

$$\therefore \bar{x} = \sum_{j \in P_2} \left(\frac{-\gamma_j}{\gamma}\right) \cdot x_j. \quad (*)$$

$\because \sum_{i=1}^{n+2} \gamma_i = 0 \quad \therefore \sum_{i \in P_1} \gamma_i = \sum_{j \in P_2} (-\gamma_j)$.

$\Rightarrow \sum_{j \in P_2} \left(\frac{-\gamma_j}{\gamma}\right) = 1$ w/ $-\gamma_j/\gamma > 0, \ \forall j \in P_2$.

By $(*)$, $\bar{x} \in \operatorname{conv}(B)$.

$\Rightarrow \operatorname{conv}(A) \cap \operatorname{conv}(B) \supseteq \{\bar{x}\} \neq \emptyset$. #

1.2 Farka's lemma and feasibility of polyhedra

1.2.1 Solvability of a system of lineer eq'ns.

Thm 1.2.1 (Fredholm's alternative theorem).

The system of lineer eq'ns $Ax = b$ has a sol'n iff $\forall y$ w/ $y^T A = 0^T$, we have $y^T b = 0$.

Before proving it, we recall the Gaussian elimination procedure:

$$(A|b) \to (D|d) = \left(\begin{array}{ccc|c|c} 1 & & 0 & * & d_1 \\ & \ddots & & \vdots & \vdots \\ 0 & & 1 & * & d_r \\ \hline & & & 0 & d_{r+1} \\ & 0 & & \vdots & \vdots \\ & & & 0 & d_m \end{array}\right)$$

where $\exists$ invertible $Q \in \mathbb{R}^{m \times m}$ s.t. $QA = D$, $Qb = d$.

Lem 1.2.2

$Ax=b$ is solvable iff $d_{r+1}=\cdots=d_m=0$.

⟨Pf⟩

Use the fact that Gaussian elimination procedure preserves the sol'n of the system. #

⟨Pf of Thm 1.2.1⟩

($\Rightarrow$)

Assume x_0 is a sol'n. Then $Ax_0=b$.

$\because y^TA=0^T \quad \therefore y^TAx_0=0$.

$\Rightarrow y^Tb=0$. #

($\Leftarrow$)

Suppose $Ax=b$ has no sol'n.

By Lem 1.2.2, $\exists\ j\in\{r+1,\cdots,m\}$ s.t. $d_j\neq 0$.

Denote $Q=\begin{pmatrix}- q_1 - \\ \vdots \\ - q_m -\end{pmatrix}$, where $QA=D$, $Qb=d$.

Then $q_j^TA=0$ while $q_jb=d_j\neq 0$.

Taking $y=q_j$. ※. Thus, $Ax=b$ has a sol'n. #

1.2.2 Solvability of a system of linear inequalities. $\boxed{2/9/2019}$

We will introduce the Fourier-Motzkin elimination algorithm. (by Fourier in 1800s, Motzkin in 1930s)

Algorithm 1.1 Fourier-Motzkin elimination

Input

$Ax\leq b$, where $A: m\times n$, $b: m\times 1$, $x\in\mathbb{R}^n$.

a column index j, $1\leq j\leq n$. $A=\begin{bmatrix}- a_1 - \\ \vdots \\ - a_m -\end{bmatrix}=[a_{ij}]$.

Output

$Dx\leq d$, where $D: M\times n$, $d: M\times 1$, $x\in\mathbb{R}^n$ w/ jth column of D being zero.

Steps:

① $I:=\{1,2,\cdots,m\}$. Partition I into

$N:=\{i\in I \mid a_{ij}<0\}$, negative

$Z:=\{i\in I \mid a_{ij}=0\}$, and zero

$P:=\{i\in I \mid a_{ij}>0\}$. positive

② Create D and d by adding one row/entry at a time:

(i) for $i\in Z$, do — this makes the jth column of D zero.

Put row $\vec{a}_i$ in D and b_i in d.

(ii) for each pair (s,t) w/ $s\in N$, $t\in P$, do

Create new row of D: $a_{tj}\vec{a}_s-a_{sj}\vec{a}_t$.

Create new entry of d: $a_{tj}b_s-a_{sj}b_t$.

③ Return $M\times n$ matrix D and $M\times 1$ vector d.

$(M=|Z|+|P|\cdot|N|)$.

Rmk:

① F-M elimination is NOT very efficient in practice, but it can be used to prove theorems in the theory of linear programming.

② By-product of F-M elimination:

$U_j\in\mathbb{R}_+^{M\times m}$ s.t. $D=U_jA$ and $d=U_jb$.

$U_j=\begin{bmatrix}- \vec{e}_i - \\ a_{tj}\vec{e}_s-a_{sj}\vec{e}_t\end{bmatrix}\begin{matrix}\}\ i\in Z \\ \}\ s\in N, t\in P.\end{matrix}$

Thm 1.2.3

Algorithm 1.1 produces D and d s.t.

$Ax\leq b$ is solvable iff $Dy\leq d$ is solvable.

⟨Pf⟩

($\Rightarrow$) Given a sol'n x of $Ax\leq b$.

i.e. $\vec{a}_i^Tx\leq b_i$, $\forall\, i=1,\cdots,m$.

For $i\in Z$, $\vec{a}_i$ and b_i are copied to D and d.

Thus $Dx \le d$ is satisfied on the corresponding row.

For $s \in N$ and $t \in P$,

$\because a_{sj} < 0$ and $a_{tj} > 0$

$\therefore -a_{sj} \vec{a_t}^T x \le -a_{sj} b_t$ and

$a_{tj} \vec{a_s}^T x \le a_{tj} b_s$.

$\Rightarrow (a_{tj}\vec{a_s} - a_{sj}\vec{a_t})^T x \le a_{tj} b_s - a_{sj} b_t$.

i.e. on the row of D and entry of d w.r.t. (s,t),

x satisfies $Dx \le d$. #

($\Leftarrow$)

$\because$ jth column of D is zero

$\therefore$ any sol'n of $Dx \le d$ can be made a sol'n x w/ $x_j = 0$. (Simply replace the jth entry by 0 will achieve this.)

Let x be a sol'n of $Dx \le d$ w/ $x_j = 0$.

claim: $\exists L, U$ w/ $L \le U$ s.t. $y = x + \lambda e_j$ is a sol'n of $Ay \le b$, $\forall \lambda \in [L, U]$. L: lower. U: upper.

Suffice to prove this.

<Pf of claim>.

1° Define

$y_i := \frac{1}{a_{ij}}(b_i - \vec{a_i}^T x)$, $i \in N \cup P$.

$U := \begin{cases} \min\{y_t \mid t \in P\}, & \text{if } P \ne \emptyset. \\ \infty, & \text{if } P = \emptyset. \end{cases}$

$L := \begin{cases} \max\{y_s \mid s \in N\}, & \text{if } N \ne \emptyset. \\ -\infty, & \text{if } N = \emptyset. \end{cases}$

2° claim: $L \le U$.

For $t \in P$ and $s \in N$, since $Dx \le d$,

$(a_{tj}\vec{a_s}^T - a_{sj}\vec{a_t}^T) x \le a_{tj} b_s - a_{sj} b_t$.

$\Rightarrow \frac{b_t - \vec{a_t}^T x}{a_{tj}} \ge \frac{b_s - \vec{a_s}^T x}{a_{sj}}$ (since $a_{sj} < 0$, $a_{tj} > 0$).

i.e. $y_t \ge y_s$.

$\because t \in P$ and $s \in N$ are arbitrary.

$\therefore \min\{y_t \mid t \in P\} \ge \max\{y_s \mid s \in N\}$.

i.e. $U \ge L$. # of claim 2°

3° For $\lambda \in [L, U]$, denote $x^\lambda = x + \lambda e_j$.

For $i \in \{1, 2, \cdots, m\}$,

$b_i - \vec{a_i}^T x^\lambda = b_i - \vec{a_i}^T x - \lambda \underbrace{\vec{a_i}^T e_j}_{a_{ij}}$

case 1 $i \in Z$.

Then $a_{ij} = 0$, and $b_i - \vec{a_i}^T x^\lambda = b_i - \vec{a_i}^T x$.

$\because \vec{a_i}$ and b_i are copied to D and d, for $i \in Z$

$\therefore b_i - \vec{a_i}^T x \ge 0$. $\Rightarrow \vec{a_i}^T x^\lambda \le b_i$.

case 2 $i \in N$.

Then $b_i - \vec{a_i}^T x^\lambda = a_{ij}\left(\frac{b_i - \vec{a_i}^T x}{a_{ij}} - \lambda\right) = a_{ij}(y_i - \lambda)$

$\because i \in N$ $\therefore y_i \le \max\{y_s \mid s \in N\} = L \le \lambda$.

$\Rightarrow \overset{-}{a_{ij}}(\overset{-}{y_i - \lambda}) \ge 0$. $\Rightarrow b_i \ge \vec{a_i}^T x^\lambda$.

case 3 $i \in P$

Similar to case 2. (left as exercise).

Thus, $\forall i \in \{1, \cdots, m\}$, $\vec{a_i}^T x^\lambda \le b_i$.

$\Rightarrow Ax^\lambda \le b$. #

Rmk:

By applying $U_1, \cdots, U_n$ to A, we obtain

$U := U_n \cdots U_1$, a nonnegative matrix, s.t.

$D = UA = 0$.

Then $Ax \le b$ has a sol'n $\Leftrightarrow d := Ub \ge 0$.

In particular, fixing A,

$\{b \mid Ax \le b \text{ has a sol'n}\}$

$= \{b \mid Ub \ge 0\}$, is a polyhedral cone.

Lem 1.2.5 A: $m \times n$ matrix, $b \in \mathbb{R}^{m \times 1}$.

Then $\exists$ nonnegative $k \times m$ matrix U w/ $UA = 0$

s.t. $Ub \ge 0 \Leftrightarrow Ax \le b$ has a sol'n.

Cor 1.2.6 (Farkas' lemma)

$Ax \le b$ has a sol'n if and only if

$\begin{cases} y^TA = 0^T \\ y \geq 0 \\ y^Tb < 0 \end{cases}$ does not have a sol'n.

<Pf>

($\Rightarrow$)

Suppose the contrary. Let x be a sol'n of $Ax \le b$, and y be a sol'n of the system.

$\because y \geq 0$

$\therefore y^TAx \le y^Tb. \Rightarrow 0^Tx \le y^Tb < 0 \Rightarrow 0 < 0.$ $\ast$ #

($\Leftarrow$)

Suppose the contrary. i.e. $Ax \le b$ has no sol'n.

By Thm 1.2.3, $\exists\ U = \begin{pmatrix} -u_1- \\ \vdots \\ -u_k- \end{pmatrix}$ s.t. $UA = 0$, $U \geq 0$, and $u_i^Tb < 0$, for some i.

Taking $y = u_i$.

Then $y^TA = 0^T$. $y \geq 0$. $y^Tb < 0$. $\ast$ Thus $Ax \le b$ has a sol'n. #

Lem 1.2.7 (Farkas' lemma, version 2)

Either $Ax = b$, $x \geq 0$ is solvable

or $y^TA \geq 0^T$, $y^Tb < 0$ is solvable,

but not both.

<Pf>

$Ax = b, x \geq 0 \iff \begin{cases} Ax \le b \\ -Ax \le -b \\ -Ix \le 0. \end{cases}$ Denote $\bar{A} = \begin{pmatrix} A \\ -A \\ -I \end{pmatrix}$, $\bar{b} = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$.

Then $Ax = b$, $x \geq 0$ is solvable $\iff \bar{A}z \le \bar{b}$ is solvable.

1° "Not both".

If both the systems are solvable, then

$y^TAx = y^Tb < 0$, w/ $y^TA \geq 0^T$, $x \geq 0$. $\ast$.

2° Assume $Ax = b$, $x \geq 0$ is NOT solvable.

i.e. $\bar{A}z \le \bar{b}$ is NOT solvable.

$\boxed{P \text{ or } Q}$ $\equiv \boxed{\sim P \Rightarrow Q}$ $\equiv \boxed{\sim Q \Rightarrow P}$

By Cor 1.2.6, $\exists\ z = (z_1, z_2, w)^T$ s.t.

$\begin{cases} z^T\bar{A} = 0^T \\ z \geq 0 \\ z^T\bar{b} < 0. \end{cases}$ i.e. $\begin{cases} (z_1 - z_2)^TA = w^T \\ z_1 \geq 0, z_2 \geq 0, w \geq 0 \\ z_1^Tb - z_2^Tb < 0. \end{cases}$ Thus, $y = z_1 - z_2$ is a sol'n of the 2nd sys. #

Rmk: Geometric interpretation of Farkas' lem.: (ver. 2)

$K(A) :=$ cone generated by columns of A.

Given $b \in \mathbb{R}^m$. If $b \notin K(A)$, then $\exists$ hyperplane $H = \{x : a^Tx = 0\}$ separating $K(A)$ and b. i.e. $a^Tb < 0$, $a^TA \geq 0^T$.

2/10/2019

1.3 Weyl-Minkowski's representation theorem.

Thm 1.3.1 (Weyl-Minkowski theorem).

Every polytope is a polyhedron.

Every bdd polyhedron is a polytope.

The goal of this sect'n is to prove / introduce concepts surrounding this thm.

Def

A rep. of a polytope by a set of linear ineq. is called an **H-representation.**

A rep. of a polytope by the convex hull of a set of pts is called a **V-representation.**

For proving Weyl-Minkowski thm, we need a concept called "polar".

Def

$A \subseteq \mathbb{R}^n$.

The **polar** of A is

$$\boxed{A^\circ} := \{x \in \mathbb{R}^n \mid a^Tx \le 1, \forall a \in A\}$$

$$= \bigcap_{a \in A} \{x \in \mathbb{R}^n \mid a^Tx \le 1\}.$$

Ex:

① For $x \in \mathbb{R}^n$, $\{x\}^\circ$ is a closed half-sp.

② L: a line in $\mathbb{R}^2$ passing through 0.

$\Rightarrow L^\circ =$ the line perpendicular to L that passes through 0.

③ $C :=$ circle of radius 1 w/ center at 0.

$\Rightarrow C^\circ =$ closed unit ball centered at 0.

Lem 1.3.4

For $A, B \subseteq \mathbb{R}^n$, we have:

(i) The polar A° is closed, convex and contains 0.

(ii) If $A \subseteq B$, then $B^\circ \subseteq A^\circ$.

(iii) For any $S \subseteq \mathbb{R}^n$, $(\mathrm{conv}(S))^\circ = S^\circ$.

⟨Pf⟩

(i) $0 \in A^\circ$ by def.

∵ A° is the intersection of closed half-sp.

∴ A° is closed and convex. #

(ii) Clear from def.

(iii) ∵ $S \subseteq \mathrm{conv}(S)$ ∴ $(\mathrm{conv}(S))^\circ \subseteq S^\circ$.

Given $x \in S^\circ$. We need to prove $z^T x \le 1$ $\forall z \in \mathrm{conv}(S)$.

Let $z \in \mathrm{conv}(S)$.

Then $z = \sum_{i=1}^m \lambda_i x_i$, for some m, $x_1, \cdots, x_m \in S$, $\lambda_i \ge 0$, $\sum_{i=1}^m \lambda_i = 1$.

∵ $x \in S^\circ$ ∴ $x_i^T x \le 1$, $\forall i$.

$\Rightarrow z^T x = \sum_{i=1}^m \lambda_i x_i^T x \le \sum_{i=1}^m \lambda_i = 1$. Thus $x \in (\mathrm{conv}(S))^\circ$ #

Lem 1.3.5

(i) If P is a polytope w/ $0 \in P$, then $(P^\circ)^\circ = P$.

(ii) If P is a polytope, then P° is a polyhedron.

⟨Pf⟩

(i) It's easy to see $P \subseteq (P^\circ)^\circ$ by def.

Suppose $\exists\, y \in (P^\circ)^\circ$ but $y \notin P$.

Let $\{x \mid c^T x = \alpha\}$ be a separating hyperplane of y and P. i.e. $c^T y > \alpha$, $c^T p < \alpha$, $\forall p \in P$.

∵ $0 \in P$ ∴ $\alpha > 0$.

Then $(c/\alpha)^T y > 1$, $(c/\alpha)^T p < 1$, $\forall p \in P$.

$\Rightarrow c/\alpha \in P^\circ$.

∵ $y \in (P^\circ)^\circ$ ∴ $(c/\alpha)^T y = y^T (c/\alpha) \le 1$. $\times$

Thus $\nexists$ such y and $(P^\circ)^\circ \subseteq P$. #

(ii) Suffice to prove the following claim.

Claim: If $P = \mathrm{conv}(a_1, \cdots, a_m)$, then $P^\circ = \{x \in \mathbb{R}^n \mid a_i^T x \le 1, \forall i\}$.

⟨Pf of claim⟩

∵ $\{a_1, \cdots, a_m\} \subseteq P$ ∴ $P^\circ \subseteq (\{a_1, \cdots, a_m\})^\circ = \text{RHS}$.

Given $x \in \text{RHS}$. i.e. $a_i^T x \le 1$, $\forall i$.

For $p \in P$, $p = \sum_{i=1}^m \lambda_i a_i$, $\lambda_i \ge 0$, $\sum_{i=1}^m \lambda_i = 1$.

$\Rightarrow p^T x = \sum_{i=1}^m \lambda_i a_i^T x \le \sum_{i=1}^m \lambda_i = 1$. $\Rightarrow x \in P^\circ$. #

Lem 1.3.6

(i.e. we may replace 1 by 0 for cones.)

$K \subseteq \mathbb{R}^n$, a cone.

Then $K^\circ = \{x \in \mathbb{R}^n \mid y^T x \le 0, \forall y \in K\}$.

In particular, K° is again a cone.

⟨Pf⟩

Clearly, $K^\circ \supseteq$ RHS.

Given $x \in K^\circ$.

Then $y^T x \le 1$, $\forall y \in K$.

For nonzero $y_0 \in K$. ∵ K: cone ∴ $N y_0 \in K$ $\forall N \in \mathbb{N}$

$\Rightarrow (N y_0)^T x \le 1$, $\forall N \in \mathbb{N}$.

$\Rightarrow y_0^T x \le 1/N$, $\forall N \in \mathbb{N}$. $\Rightarrow y_0^T x \le 0$.

Thus, $x \in$ RHS, and $K^\circ \subseteq$ RHS. #

Rmk

For linear subsp. L, $L^\circ = L^\perp = \{x \in \mathbb{R}^n \mid y^T x = 0 \ \forall y \in L\}$

⟨Pf⟩

Linear subsp. are special cases of cones.

$\Rightarrow L^\circ = \{x \in \mathbb{R}^n \mid y^T x \le 0, \forall y \in L\}$.

∵ $-y \in L$, $\forall y \in L$

∴ We may further write $L^\circ = L^\perp$. #

Rmk $\boxed{2/20/2019}$

If $K = \text{cone}(\{a_1, \cdots, a_m\})$, then

$$K^\circ = \{x \mid a_i^T x \le 0,\ i = 1, \cdots, m\}.$$

Ex 1.3.7

① $(\mathbb{R}^n_+)^\circ = \mathbb{R}^n_-$.

② $\{0\}^\circ = \mathbb{R}^n$, $(\mathbb{R}^n)^\circ = \{0\}$.

③ $K := \{(x_1, x_2) \mid x_1 > 0\} \cup \{(0,0)\}$.

Then $K^\circ = \text{cone}(\{(-1,0)\})$.

$\Rightarrow K^{\circ\circ} = \{(x_1, x_2) \mid x_1 \ge 0\} \supsetneq K$.

Def

Given $S_1, S_2 \subseteq \mathbb{R}^n$.

$\boxed{S_1 + S_2} := \{x + y \mid x \in S_1, y \in S_2\}$ is called the $\boxed{\text{Minkowski sum}}$ of S_1 and S_2.

Prop 1.3.8

K_1, K_2 : convex cones in $\mathbb{R}^n$. Then

① $K_1 + K_2$ is a cone.

② if $K_1 \subseteq K_2$, then $K_1^\circ \supseteq K_2^\circ$. (holds for general sets)

③ $(K_1 + K_2)^\circ = K_1^\circ \cap K_2^\circ$

④ $K_1^\circ + K_2^\circ \subseteq (K_1 \cap K_2)^\circ$.

⑤ $K \subseteq K^{\circ\circ}$. (holds for general sets).

〈Pf〉

Exercise. #

Ex 1.3.9 (In general, in ④ and ⑤ above, we do not have "=")

1° $K_1 := \{(x_1, x_2) \mid x_1 > 0\} \cup \{(0,0)\}$. $K_2 := -K_1$.

Then $K_1 \cap K_2 = \{(0,0)\}$. $\Rightarrow (K_1 \cap K_2)^\circ = \mathbb{R}^2$.

$K_1^\circ = \text{cone}(\{(-1,0)\})$ and $K_2^\circ = \text{cone}(\{(1,0)\})$.

Thus $K_1^\circ + K_2^\circ = \mathbb{R} \times \{0\} \subsetneq \mathbb{R}^2 = (K_1 \cap K_2)^\circ$. #

2° By ③ in Ex 1.3.7, $\exists$ cone K s.t. $K \subsetneq K^{\circ\circ}$. #

Thm 1.3.10 (Modified).

① If K is a finitely generated cone, then $K = K^{\circ\circ}$.

② If K is a polyhedral cone, then $K = K^{\circ\circ}$.

③ If a polyhedral cone K is written as

$$K = \{x \mid A^T x \le 0\} \text{ w/ } A^T = \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & a_m & - \end{pmatrix},$$

then $K^\circ = \text{cone}(\{a_1, \cdots, a_m\})$.

Thus, K° is a finitely generated cone.

〈Pf〉

① $K \subseteq K^{\circ\circ}$ generally holds.

Suppose $\exists\, y \in K^{\circ\circ}$ s.t. $y \notin K$.

$\because K$ is finitely generated

$\therefore \exists$ matrix $A = (a_1 \cdots a_m)$ s.t. $K = \text{cone}(\{a_1, \cdots, a_m\}) = \{Ax \mid x \ge 0\}$.

$\because y \notin K$

$\therefore Ax = y$, $x \ge 0$ has no sol'n in x.

By Farka's lemma, version 2,

$\exists\, a$ s.t. $a^T A \le 0^T$, $a^T y > 0$.

$\because a^T A \le 0 \quad \therefore a \in K^\circ$.

$\because y \in K^{\circ\circ} \quad \therefore a^T y \le 0$. $\times$

Thus $\nexists$ such y and $K = K^{\circ\circ}$. #

②/③

Use notations in ③.

Let $\hat{K} = \text{cone}(\{a_1, \cdots, a_m\})$.

By def of "$\circ$", $\hat{K}^\circ = K$. $(*_1)$

By ①, since $\hat{K}$ is finitely generated, $\hat{K} = \hat{K}^{\circ\circ}$. $(*_2)$

Thus $\hat{K} \overset{(*_2)}{=} \hat{K}^{\circ\circ} = (\hat{K}^\circ)^\circ \overset{(*_1)}{=} K^\circ$, proving ③

Moreover, $K^{\circ\circ} = (K^\circ)^\circ \overset{③}{=} \hat{K}^\circ \overset{(*_1)}{=} K$, proving ②. #

Cor 1.3.11 (Cor of Lem 1.2.5).

Given a matrix A.

Then $\exists$ matrix B s.t. $BA \geq 0$ and

$\{z \mid Ax = z, x \geq 0$ has a sol'n in $x\}$ ← cone generated by columns of A

$= \{z \mid Bz \geq 0\}$.

<Pf>

$Ax = z, x \geq 0$ has a sol'n iff $\tilde{A}x = \begin{pmatrix} A \\ -A \\ -I \end{pmatrix} x \leq \begin{pmatrix} z \\ -z \\ 0 \end{pmatrix}$ has a sol'n.

By Lem 1.2.5, $\exists\ \tilde{B} = (B_1 \mid B_2 \mid B_3) \geq 0$ w/ $\tilde{B}\tilde{A} = 0$, s.t.

a sol'n exists iff $\tilde{B}\begin{pmatrix} z \\ -z \\ 0 \end{pmatrix} = B_1 z - B_2 z \geq 0$.

$\tilde{B}\tilde{A} = B_1 A - B_2 A - B_3 = 0. \Rightarrow (B_1 - B_2)A = B_3 \geq 0$.

Letting $B = B_1 - B_2$, the result follows. #

We now prove a Weyl-Minkowski thm for cones. The one for polytopes/polyhedra can be obtained via "homogenizat'n", discussed in the next sec'n.

Thm 1.3.12 (Weyl)

Every finitely generated cone is polyhedral.

<Pf>

Given $K = \text{cone}(\{a_1, \cdots, a_m\})$.

Then $K = \{z \mid Ax = z, x \geq 0$ has a sol'n in $x\}$

where $A = (a_1 \cdots a_m)$.

By Cor 1.3.11, $\exists\ B$ w/ $BA \geq 0$ s.t.

$K = \{z \mid Bz \geq 0\}$.

i.e. K is polyhedral. #

Rmk:

B in the proof of Thm 1.3.12 is NOT necessarily "minimal". (i.e. some rows of B might be redundant.)

Cor 1.3.13 (Minkowski)

Every polyhedral cone is finitely generated.

<Pf>

Given a polyhedral cone K.

By ③ in Thm 1.3.10, K° is finitely generated.

By Thm 1.3.12, K° is polyhedral.

By ③ in Thm 1.3.10, $K^{\circ\circ}$ is finitely generated.

By ② in $\cdot\cdot$ $\cdot\cdot$, $K = K^{\circ\circ}$.

Thus, K is finitely generated. #

Rmk: (Cor of Thm 1.3.12 and 1.3.13)

① K: finitely generated cone

$\Rightarrow K^\circ$: finitely generated cone.

② K_1, K_2: f.g. cone (Notation: f.g. = finitely generated)

$\Rightarrow K_1 \cap K_2$: f.g. cone.

<Pf>

① K: f.g. $\Rightarrow K^\circ$: polyhedral $\Rightarrow K^\circ$: f.g. #

② K_1, K_2: f.g. $\Rightarrow K_1, K_2$: polyhedral $\Rightarrow K_1 \cap K_2$: poly… $\Rightarrow K_1 \cap K_2$: f.g. #

Cor 1.3.14.

K_1, K_2: f.g. cone.

Then $K_1^\circ + K_2^\circ = (K_1 \cap K_2)^\circ$.

(i.e. equality in ④ of Prop 1.3.8 holds for f.g. cones).

<Pf>

We will apply ③ of Prop 1.3.8 and ①/② of Thm 1.3.10 multiple times.

$$K_1^\circ + K_2^\circ = (K_1^\circ + K_2^\circ)^{\circ\circ} \overset{\text{③ of Prop 1.3.8}}{=} (K_1^{\circ\circ} \cap K_2^{\circ\circ})^\circ = (K_1 \cap K_2)^\circ. \ \#$$

Cor 1.3.15 (Cor of standard Weyl-Minkowski)

P: an n-dim polytope in $\mathbb{R}^n$. Then

① intersect'n of P w/ a hyperplane is again a polytope, and

② image of P under a linear map is another polytope. (in particular, a projection.)