Notes for Graph Theory Course

Shuang Zhao

ACM Class: F0303026
Student ID: 5030309961
Shanghai Jiao Tong University
E-mail: shuangz@acm.org

November 15, 2004

Contents

Symbols

1. To any vector $x \in \mathbb{R}^n$, define $||x||$ by the Euclidean length of vector x, namely

$$||x|| := \sqrt{x_1^2 + \ldots + x_n^2}$$

2. To any vectors $x, y \in \mathbb{R}^n$, define $x \leq y$ by

$$x \leq y \iff x_1 \leq y_1, x_2 \leq y_2, \ldots, x_n \leq y_n$$

3. Let $A \in \mathbb{R}^{m \times n}$. Define the the i-th row of A by vector $a_i^\top \in \mathbb{R}^m$, the j-th column of A by vector $a_j' \in \mathbb{R}^n$, and the j-th element on the i-th row by $a_{ij} \in \mathbb{R}$, where $1 \leq i \leq m$ and $1 \leq j \leq n$.

4. Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$. Define

$$\begin{aligned} P(A,b) &:= \{x \in \mathbb{R}^n \mid Ax \leq b\} \\ P^=(A,b) &:= \{x \in \mathbb{R}^n \mid Ax = b\} \\ r_F^=(A,b) &:= \{i \in \{1,2,\ldots,m\} \mid a_i^\top x = b_i,\ \forall x \in F\} \\ r_z^=(A,b) &:= r_{\{z\}}^=(A,b) \\ r^=(A,b) &:= r_{P(A,b)}^=(A,b) \end{aligned}$$

$$\begin{aligned} A_{F,b} &:= \text{a submatrix of } A \text{ containing } a_i^\top \text{ for all } i \in r_F^=(A,b) \\ A_{z,b} &:= \text{a submatrix of } A \text{ containing } a_i^\top \text{ for all } i \in r_z^=(A,b) \\ A_b &:= \text{a submatrix of } A \text{ containing } a_i^\top \text{ for all } i \in r^=(A,b) \end{aligned}$$

Chapter 1

Polytopes, polyhedra, Farakas' lemma, and linear programming

1.1 Convex sets

First there are some base concepts within this course:

Definition 1.1. *A subset C of $\mathbb{R}^n$ is called* `convex` *if for all $x, y \in C$ and any $0 \leq \lambda \leq 1$ also $\lambda x + (1-\lambda)y \in C$. So C is convex if with any two points in C, the whole line segment connecting x and y belongs to C.*

Definition 1.2. *Let set $X \subseteq \mathbb{R}^n$,*

$$\text{Conv.hull}(X) := \Big\{ x \mid \exists t \in \mathbb{N}, \exists x_1, \ldots, x_t \in X, \exists \lambda_1, \ldots, \lambda_t \geq 0 : \\ x = \sum_{i=1}^{t} \lambda_i x_i, \sum_{i=1}^{t} \lambda_i = 1 \Big\}$$

Definition 1.3. *We call a subset H of $\mathbb{R}^n$ a* `hyperplane` *if there exist a vector $c \in \mathbb{R}^n$ with $c \neq 0$ and a $\delta \in \mathbb{R}$ such that: $H = \{x \in \mathbb{R}^n \mid c^\top x = \delta\}$.*

Definition 1.4. *We call a subset H of $\mathbb{R}^n$ a* `halfspace` *(or an* `affine halfspace`*) if there exist a vector $c \in \mathbb{R}^n$ with $c \neq 0$ and a $\delta \in \mathbb{R}$ such that*

$$H = \{x \in \mathbb{R}^n \mid c^\top x \leq \delta\}$$

And following is the first important theorem in this chapter:

Theorem 1.1 (Separating hyperplane theorem). *Let C be a closed convex set in $\mathbb{R}^n$ and let $z \notin C$. Then there exists a hyperplane separating z and C.*

Basic idea. To prove this theorem, firstly we choose a point $y \in C$ such that y is nearest to z among all points in C. Since C is closed, such y always exists. Then we pick a hyperplane separates z and y with normal vector $z - y$. Without loss of generality, the hyperplane passes $(z+y)/2$. That is, we pick the hyperplane

$$H = \{x \mid (z-y)^\top x = \frac{(z-y)^\top (z+y)}{2} = \frac{||z||^2 - ||y||^2}{2}\}$$

Next we want to show H indeed separates z and C. A fast way to do this is to find a contradiction. Assume there exists an $x \in C$ lying on the same side of H with z. Then one can easily verify the existence of another point w on the line segment connecting x and y, such that $||w - z|| < ||y - z||$. This contradicts the fact that y is nearest to z. So the theorem follows.

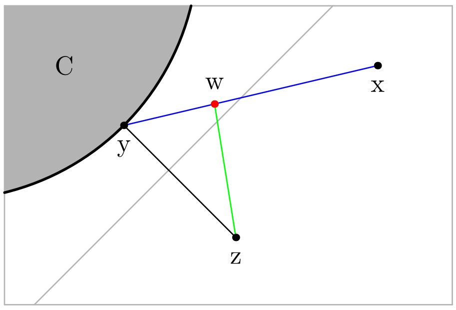

Figure: To prove separating hyperplane theorem

Proof. Since C is closed, there exists a $y \in C$ such that for each $x \in C$, $||y-z|| \leq ||x-z||$. Let $c = z - y$, $\delta = (||z||^2 - ||y||^2)/2$ and $H = \{x \mid c^\top x = \delta\}$.

From $c^\top z + c^\top y = 2\delta$ and $c^\top (z - y) = ||c||^2 > 0$, one has $c^\top z > \delta$ and $c^\top y < \delta$. Assume there exists an $x \in C$, such that $c^\top x \geq \delta$. Next we show there is a point $w = \lambda x + (1 - \lambda)y$ satisfying $0 \leq \lambda \leq 1$ and $||w - z|| < ||y - z||$.

$$\begin{aligned} ||w - z||^2 &= ||\lambda x + (1 - \lambda)y - z||^2 = ||\lambda(x - y) + (y - z)||^2 \\ &= ||\lambda(x - y) + c||^2 = \lambda^2 ||x - y||^2 - 2\lambda(x - y)^\top c + ||c||^2 \end{aligned}$$

Let

$$0 < \lambda < \min\left\{\frac{2c^\top (x - y)}{||x - y||^2}, 1\right\}$$

Then $||w - z||^2 < ||y - z||^2 = ||c||^2$. This contradicts the fact that for each $x \in C$, $||y - z|| \leq ||x - z||$. Thus for each $x \in C$, $c^\top x < \delta$. Therefore H is a hyperplane separating z and C. □

Separating hyperplane theorem is a fundamental but very useful theorem. It is widely used in the proof of many propositions. We can obtain **Farakas' Lemma** by applying this theorem on a convex cone.

As a direct consequence of separating hyperplane theorem, we have

Proposition 1.2 (Exercise [1] 2.1). *Each closed convex set is the intersection of a collection of halfspaces, possibly infinite many of them.*

Proof. Let C be a closed convex set in $\mathbb{R}^n$. Then

$$\begin{aligned} C &= \mathbb{R}^n \setminus \bigcup_{z \notin C} \{z\} \\ &= \mathbb{R}^n \setminus \bigcup_{z \notin C} \{x \mid c_z^\top x > \delta_z\} \\ &= \bigcap_{z \notin C} \{x \mid c_z^\top x \leq \delta_z\} \end{aligned}$$

The second equality follows the separating hyperplane theorem. □

1.2 Polytopes, polyhedra and cones

Polytope and polyhedra are special cases of convex sets:

Definition 1.5. *A subset P of $\mathbb{R}^n$ is a* `polyhedron` *iff there exists an $m \times n$ matrix A and a vector $b \in \mathbb{R}^m$ such that*

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

A polyhedron is the intersection of a finite number of halfspaces.

Definition 1.6. *A subset P of $\mathbb{R}^n$ is called a* `polytope` *iff P is the convex hull of a finite number of vectors. That is, there exists vectors $x_1, \dots, x_t \in \mathbb{R}^n$ such that*

$$P = \text{Conv.hull}\{x_1, \dots, x_t\}$$

Besides given the definitions of polytopes and polyhedra, we also show the relationship between them:

Theorem 1.3. *P is a polytope iff P is a bounded polyhedron.*

To prove Theorem 1.3, first we provide the definition of *extreme point*:

Definition 1.7. *Let P be a convex set. A point $z \in P$ is called a* `extreme point` *of P if z is not on the line segment connecting any two other points in P. That is, there do not exist points x, y in P such that $x \neq z$, $y \neq z$ and $z = (x + y)/2$.*

Extreme points have an important property:

Theorem 1.4. *Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and P be the polyhedron $P(A, b)$. For each $z \in P(A, b)$, z is an extreme point of P, iff $\text{rank}(A_{z,b}) = n$.*

Basic idea. A simple method to prove this proposition is by contradiction.

First we suppose z is an extreme point of $P(A, b)$ with $\text{rank}(A_{z,b}) < n$. By some simple calculations we can find $x \neq z$ and $y \neq z$ such that $z = (x + y)/2$.

Then we assume z is not an extreme point but with $\text{rank}(A_{z,b}) = n$. Let $z = (x+y)/2$, and we show $x = y$. Thus $x = y = z$.

Proof. Necessity. Suppose z is an extreme point of P with $\text{rank}(A_{z,b}) < n$. Then there exists a vector $0 \neq c \in \mathbb{R}^n$ satisfying $(A_{z,b})c = 0$.

Then for each $i \notin r_z^=(A, b)$, it holds $a_i^\top z < b_i$. So there exists a small real $\delta_i > 0$, such that $a_i^\top (z \pm \delta_i c) < b_i$.

Now let $\delta = \min\{\delta_i\}$, and $x = z - \delta c$, $y = z + \delta c$, then $x, y \in P$. Since $c \neq 0, \delta > 0$, we have $x \neq z$, $y \neq z$ and $z = (x + y)/2$. This contradicts the fact that z is an extreme point of P.

Sufficiency. Assume z is not an extreme point of P but holds $\text{rank}(A_{z,b}) = n$. Then there exists $x \neq z, y \neq z$, such that $z = (x + y)/2$.

For each $i \in r_z^=(A, b)$, we have

$$a_i^\top x \leq b_i = a_i^\top z \quad \Longrightarrow \quad a_i^\top (x - z) = \frac{a_i^\top (x - y)}{2} \leq 0$$
$$a_i^\top y \leq b_i = a_i^\top z \quad \Longrightarrow \quad a_i^\top (y - z) = \frac{a_i^\top (y - x)}{2} \leq 0$$
$$\therefore a_i^\top (x - y) = 0$$

So $A_{z,b}(x-y) = 0$. Since $\text{rank}(A_{z,b}) = n$, we have $x-y = 0$, $x = y$. Therefore $x = y = z$. It contradicts $x \neq z$ and $y \neq z$. □

And from this theorem, we can easily derive:

Corollary 1.4.1. *The total number of extreme points of a polyhedron is always finite.*

Proof. Let $P = P(A, b)$ be a nonempty polyhedron where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^n$. Then there are at most 2^m extreme points of P. □

Following Theorem 1.4, we can show something more:

Theorem 1.5. *Let $P = P(A, b)$ be a bounded polyhedron, with extreme points $x_1, \dots, x_t$. Then*

$$P = \text{Conv.hull}(x_1, \dots, x_t)$$

Basic idea. To prove the equality, we need only to show

$$\text{Conv.hull}(x_1, \dots, x_t) \subseteq P$$

and

$$P \subseteq \text{Conv.hull}(x_1, \dots, x_t)$$

According to the convexity of P, we can easily obtain the former inclusion. Next we need only to show the latter one. That is, for each $z \in P$, we need to show $z \in \text{Conv.hull}(x_1, \dots, x_t)$. By the property of extreme points, if $\text{rank}(A_{z,b}) = n$, z itself is an extreme point of P. So z is in the convex hull generated by all extreme points of P. If $\text{rank}(A_{z,b}) = k < n$, we use mathematical induction to show the inclusion. Namely, we are going to find two vectors y and z where $\text{rank}(A_{x,b}) > k$ and $\text{rank}(A_{y,b}) > k$. By induction hypothesis, $x, y \in \text{Conv.hull}(x_1, \dots, x_t)$. Thus $z \in \text{Conv.hull}(x_1, \dots, x_t)$.

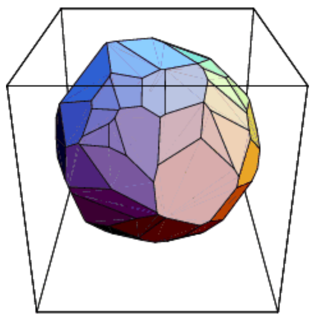

Figure: A bounded polyhedron, a polytope

Proof. Let $C = \text{Conv.hull}(x_1, \dots, x_t)$.

By the convexity of P, we know $C \subseteq P$. Next we need only to show $P \subseteq C$.

For each $z \in P$, if $\text{rank}(A_{z,b}) = n$, z itself is an extreme point. Then $z \in \{x_1, \dots, x_t\} \subseteq C$. If $\text{rank}(A_{z,b}) < n$, there exists a vector $c \neq 0$, such that $(A_{z,b})c = 0$.

Let

$$\mu_0 = \max\{\mu \mid z + \mu c \in P\}$$
$$\nu_0 = \max\{\nu \mid z - \nu c \in P\}$$

And $x = z + \mu_0 c$, $y = z - \nu_0 c$. Since P is bounded and closed, μ_0 and ν_0 exist and are both finite.

Let $|r_z^=(A, b)| = t$. Without loss of generality, $r_z^=(A, b) = \{1, 2, \dots, t\}$. Following the definition of μ_0, we know

$$a_1^\top x = b_1, \ \dots, \ a_t^\top x = b_t$$

and

$$a_{t+1}^\top x \le b_{t+1}, \ \ldots, \ a_m^\top x \le b_m$$

Since μ_0 attains maximum, we obtain there exists a $t' > t$, such that $a_{t'}^\top x = b_{t'}$. Thus $A_{z,b}$ is a submatrix of $A_{x,b}$, and $A_{x,b}$ has at least one more line than $A_{z,b}$. As $a_{t'}^\top z < b_{t'}$ and $a_{t'}^\top x = a_{t'}^\top(z+\mu_0 c) = b_{t'}$, $a_{t'}^\top c \neq 0$. So $(A_{x,b})c \neq 0$. This implies $\text{rank}(A_{x,b}) > \text{rank}(A_{z,b})$.

Similarly we can show $\text{rank}(A_{y,b}) > \text{rank}(A_{z,b})$. By our induction hypothesis, $x, y \in C$. So z, a convex combination of x and y, is a member of C. □

As a direct consequence, we have

Corollary 1.5.1. *Each bounded polyhedron is a polytope.*

To prove Theorem 1.3, we need only to prove backward:

Theorem 1.6. *Each polytope is a bounded polyhedron.*

Proof. Let P be a polytope in $\mathbb{R}^n$, say $P = \text{Conv.hull}(x_1, \ldots, x_t)$. And we prove this theorem by induction on $\dim(P)$.

If P is contained in some hyperplane, namely $\dim(P) < n$, the theorem follows from the induction hypothesis.

If $\dim(P) = n$, this implies $x_2 - x_1, \ldots, x_t - x_1$ span $\mathbb{R}^n$. Thus there exists a $x_0 \in P$ and a real number $r > 0$, such that $B(x_0, r) := \{y \ : \ ||y - x_0|| \le r\}$ is contained in P.

Without loss of generality, $x_0 = 0$. Define P^* by

$$P^* := \{y \in \mathbb{R}^n \mid x^\top y \le 1 \quad \forall x \in P\}$$

For each $y \in P^*$, it holds $x_j^\top y \le 1$ for $j = 1, \ldots, t$. At the same time, for each y satisfying $x_j^\top y \le 1$ for $j = 1, \ldots, t$, we have

$$x^\top y = \sum_{j=1}^{t} \lambda_j x_j^\top y \le \sum_{j=1}^{t} \lambda_j = 1$$

for each $x \in P$. Therefore,

$$P^* = \{y \in \mathbb{R}^n \mid x_j^\top y \le 1 \quad j = 1, \ldots, t\}$$

Moreover, P^* is bounded. As $B(x_0, r) = B(0, r) \subseteq P$, for each $0 \neq y \in P^*$, let

$$x' = r \cdot ||y||^{-1} y$$

Then $x' \in B(0, r) \subseteq P$, hence $x^\top y = r \cdot ||y|| \le 1$. So $||y|| < 1/r$, namely $P^* \subseteq B(0, 1/r)$.

By Corollary 1.5.1, we know P^* is a polytope. Thus

$$P^* = \text{Conv.hull}(y_1, \ldots, y_s)$$

Next we show $P = (P^*)^*$, and this implies

$$P = \{x \in \mathbb{R}^n \mid y_j^\top x \le 1 \quad \forall j = 1, \ldots, s\}$$

By the definition of P^*, we know $P \subseteq (P^*)^*$. And for each $z \notin P$, by the separating hyperplane theorem, there exists a hyperplane $H = \{x \mid c^\top x = \delta\}$ such that $c^\top x < \delta$ for all $x \in P$, and $c^\top z > \delta$. As $0 \in P$, we have $c^\top 0 = 0 < \delta$. Without loss of generality, $\delta = 1$. So $c \in P^*$, and $z \notin (P^*)^*$. Therefore $P = (P^*)^*$, namely P is a polyhedron. □

Comment. In our proof of Theorem 1.6, we used a key concept P^*. P^* is call the *dual* of P. We will focus on the concept of Polarity and Duality in section 1.5.

Theorem 1.3 follows from Corollary 1.5.1 and Theorem 1.6.

Convex cone is another key concept in this course:

Definition 1.8. *A subset C of $\mathbb{R}^n$ is called a* ***convex cone*** *if for any $x, y \in C$ and any $\lambda, \mu \geq 0$ one has $\lambda x + \mu y \in C$.*

Definition 1.9. *For any $X \subseteq \mathbb{R}^n$, $\text{Cone}(X)$ is the smallest cone containning X. That is:*

$$\text{Cone}(X) := \{\lambda_1 x_1 + \ldots + \lambda_t x_t \mid x_1, \ldots, x_t \in X; \lambda_1, \ldots, \lambda_t \geq 0\}$$

A cone C is called ***finitely generated*** *if $C = \text{Cone}(X)$ for some finite set X.*

And convex cone is kind of polyhedra:

Proposition 1.7 (Exercise [1] 2.7). *Let $C \subseteq \mathbb{R}^n$. Then C is a convex cone, iff C is the intersection of a collection of linear halfspaces.*

Proof. Let $z \in \mathbb{R}^n$ and $z \notin C$. Then there exists a hyperplane

$$H = \{x \in \mathbb{R}^n \mid c^\top x = \delta\}$$

separating z and C. That is, $c^\top z > \delta$ and for any $x \in C$, $c^\top x < \delta$. Since C is a convex cone, $0 \in C$. So $c^\top 0 = 0 < \delta$.
If there exist an $x' \in C$ such that $c^\top x' = \theta > 0$. Without loss of generality, $\theta = 1$. Then $(2\delta)x' \in C$. But

$$c^\top (2\delta x') = 2\delta(c^\top x') = 2\delta > \delta$$

This contradicts the fact that for all $x \in C$, $c^\top x < \delta$. So for any $x \in C$, $c^\top x < 0$, and for $z \notin C$, $c^\top z > 0$. Therefore,

$$\begin{aligned} C &= \mathbb{R}^n \backslash \bigcup_{z \notin C} \{x \mid c^\top x > 0\} \\ &= \bigcap_{z \notin C} \{x \mid c^\top x \leq 0\} \end{aligned}$$

So C is the interesection of linear halfspaces. □

In fact, convex cones are polyhedra in one higher dimension. And we define another two useful symbols:

$$\begin{aligned} \text{lift}(P) &:= \left\{ \begin{pmatrix} x \\ 1 \end{pmatrix} \middle| \, x \in P \right\} \\ \text{slice}(P) &:= \left\{ x \, \middle| \, \begin{pmatrix} x \\ 1 \end{pmatrix} \in P \right\} \end{aligned}$$

Proposition 1.8.

$$\text{Conv.hull}(S) = \text{slice}(\text{Cone}(\text{lift}(S)))$$

Proof. Let $Y = \text{lift}(S)$. Then

$$\begin{aligned}
&\forall x \in \text{slice}(\text{Cone}(Y)) \\
\iff & \begin{pmatrix} x \\ 1 \end{pmatrix} \in \text{Cone}(Y) \\
\iff & \begin{pmatrix} x \\ 1 \end{pmatrix} = \sum_{p \in Y} \lambda_p p \qquad \text{where } \lambda_p \geq 0 \text{ and } p = \begin{pmatrix} p' \\ 1 \end{pmatrix} \\
\iff & x = \sum_{p \in Y} \lambda_p p' \qquad \text{where } \lambda_p \geq 0 \text{ and } \sum_{p \in Y} \lambda_p = 1 \\
\iff & x \in \text{Conv.hull}(S)
\end{aligned}$$

□

Now we know that statements about polytopes and polyhedra can be translated into statements about cones in one higher dimension. And next we use this **homogenization** technique to prove the first part of Caratheodory's theorem.

Theorem 1.9 (Caratheodory's theorem).

(i) Given a set S, for any point p in Conv.hull(S) *there is a subset T with p in* Conv.hull(T)*, with $|T| = \dim(S) + 1$, and the points of T are affinely independent*[1]*.*

(ii) Given a set S, for any point p in Cone(S) *there is a subset T with p in* Cone(T)*, with $|T| = \dim(S)$, and the points of T are linearly independent.*

Proof. (ii). Suppose there exists a subset T' with p in Cone(T') and the size T' is minimal, but $|T'| > \dim(S)$. Then

$$P = \sum_{s \in T'} c_s s \qquad \text{where } c_s > 0$$

Since $|T'| > \dim(S)$, the vectors in T' are linearly dependent. Thus

$$\sum_{s \in T'} d_s c_s s = 0 \qquad \text{where } d_s \neq 0$$

Pick the element s_0 with the largest $d_s = d_{s_0}$. We have

$$c_{s_0} s_0 = \sum_{s \in T' \setminus \{s_0\}} -\frac{d_s}{d_{s_0}} c_s s$$

And we use this sum to express $c_{s_0} s_0$. This eliminates the appearance of s_0 in the sum, and keep all the other coefficients nonnegative. This contradicts the choice of T'. □

Proof. (i). Let $S' = \text{lift}(S)$. Then by part (ii) of this theorem we know there exists a $T' \subseteq S'$ such that

$$\begin{pmatrix} p \\ 1 \end{pmatrix} \in \text{Cone}(T')$$

and $|T'| = \dim(S') = \dim(S) + 1$ with all vectors in T' are linearly independent.

Let $T = \text{slice}(T')$. Then $T \subseteq S$, $p \in \text{slice}(\text{Cone}(T')) = \text{Conv.hull}(T)$ and all vectors in T are affinely independent. Moreover, $|T| = |T'| = \dim(S') = \dim(S) + 1$. □

[1]We say n vectors $x_1, \ldots, x_t$ are affinely independent, iff $\begin{pmatrix} x_1 \\ 1 \end{pmatrix} \ldots \begin{pmatrix} x_n \\ 1 \end{pmatrix}$ are linearly independent.

Theorem 1.10 (Exercise [1] 2.12, Polyhedron decomposition theorem). *Let P be a subset of $\mathbb{R}^n$. Show that P is a polyhedron, iff $P = Q + C$ for some polytope Q and some finitely generated convex cone C.*

Proof. Necessity. (Minkowski's Theorem)

$$\begin{aligned}
& P \text{ is a polyhedron} \\
\Longrightarrow\ & P = \{x \mid x \in \mathbb{R}^n, b \in \mathbb{R}^m, A \in \mathbb{R}^{m\times n},\ Ax \le b\} \\
\Longrightarrow\ & \text{Let } X = \text{lift}(P) \\
& \text{Then Cone}(X) = \left\{ x \;\middle|\; x \in \mathbb{R}^{n+1}, \begin{pmatrix} -b & A \\ -1 & 0 \end{pmatrix} x \le 0 \right\} \\
\Longrightarrow\ & \text{cone}(X) \text{ is finitely generated} \\
\Longrightarrow\ & \text{Cone}(X) = \{\lambda_1 x_1 + \ldots \lambda_t x_t \mid \lambda_1, \ldots, \lambda_t \ge 0\} \\
& \text{Let } x_i = \begin{pmatrix} x'_i \\ a_i \end{pmatrix}, \text{and } a_1, \ldots, a_s > 0, a_{s+1}, \ldots, a_t = 0 \\
\Longrightarrow\ & \forall x \in P \\
& \begin{pmatrix} x \\ 1 \end{pmatrix} = \lambda_1 \begin{pmatrix} x'_1 \\ a_1 \end{pmatrix} + \ldots + \lambda_t \begin{pmatrix} x'_t \\ a_t \end{pmatrix} \\
& \qquad = \sum_{i=0}^{s} \lambda_i a_i \begin{pmatrix} \frac{x'_i}{a_i} \\ 1 \end{pmatrix} + \sum_{i=s+1}^{t} \lambda_i \begin{pmatrix} x'_i \\ 0 \end{pmatrix} \quad \text{where } \sum_{i=0}^{s} \lambda_i a_i = 1 \\
\Longrightarrow\ & x = (\lambda_1 a_1 \frac{x'_1}{a_1} + \ldots + \lambda_s a_s \frac{x'_s}{a_s}) + (\lambda_{s+1} x'_{s+1} + \ldots + \lambda_t x'_t)
\end{aligned}$$

Let

$$Q = \text{Conv.hull}(\frac{x'_1}{a_1}, \ldots, \frac{x'_s}{a_s}) \quad \text{and} \quad C = \text{Cone}(x'_{s+1}, \ldots, x'_t)$$

then we have

$$P = Q + C$$

Sufficiency. (Weyl's Theorem) If $P = Q + C$ where

$$Q = \{a_1 x_1 + \ldots a_s x_s \mid a_1, \ldots, a_s \ge 0, a_1 + \ldots + a_s = 1\}$$

is a polytope and

$$C = \{b_1 x'_1 + \ldots b_t x'_t \mid b_1, \ldots, b_t \ge 0\}$$

is a finitely generated convex cone.

Let $X = \text{lift}(P)$, then

$$\forall \begin{pmatrix} x \\ 1 \end{pmatrix} \in X, \quad \begin{pmatrix} x \\ 1 \end{pmatrix} = a_1 \begin{pmatrix} x_1 \\ 1 \end{pmatrix} + \ldots + a_s \begin{pmatrix} x_s \\ 1 \end{pmatrix} + b_1 \begin{pmatrix} x'_1 \\ 0 \end{pmatrix} + \ldots + b_t \begin{pmatrix} x'_t \\ 0 \end{pmatrix}$$

$$\text{Cone}(X) = \text{Cone}\left\{ \begin{pmatrix} x_1 \\ 1 \end{pmatrix}, \ldots, \begin{pmatrix} x_s \\ 1 \end{pmatrix}, \begin{pmatrix} x'_1 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} x'_t \\ 0 \end{pmatrix} \right\}$$

So Cone(X) is finitely generated, and therefore it is the intersection of a finite number of linear halfspaces. This implies P is the intersection of a finite number of halfscapes in one lower dimension. □

1.3 Farkas' Lemma

By applying the separating hyperplane theorem to a convex cone, we have:

Theorem 1.11. *Let C be a convex cone in $\mathbb{R}^n$. Then to any vector $z \in \mathbb{R}^n$, either*

$$z \in C$$

or

$$\exists c \in \mathbb{R}^n \text{ such that for each } x \in C,\ c^\top x \leq 0 \text{ and } c^\top z > 0$$

but not both.

Proof. For each $z \notin C$, by the separating hyperplane theorem, there exists a hyperplane $H = \{x \in \mathbb{R}^n \mid c^\top x = \delta\}$ satisfying that for all $x \in C$, $c^\top x < \delta$ and $c^\top z > \delta$. If there is an $x' \in C$ such that $c^\top x' = \varepsilon > 0$, we obtain the contradiction

$$\delta > c^\top \left(\frac{2\delta}{\varepsilon} x'\right) = \frac{2\delta}{\varepsilon} \cdot \varepsilon = 2\delta > 0$$

So for all $x \in C$, $c^\top x \leq 0 < \delta$ and $c^\top z > \delta > 0$. □

With Theorem 1.11 in hand, we can prove Farkas' Lemma by some translation:

Theorem 1.12 (Farkas' Lemma). *Let A be an $m \times n$ matrix and $b \in \mathbb{R}^m$. Then either*

$$Ax = b \text{ has a nonnegative solution } x_0$$

or

$$\exists y_0 \in \mathbb{R}^m \text{ such that } y_0^\top A \leq 0 \text{ and } y_0^\top b > 0.$$

but not both.

Proof. To an $m \times n$ matrix A, define Cone(A) by[2]

$$\text{Cone}(A) := \text{Cone}(a'_1, \ldots, a'_m)$$

$$\begin{aligned}
&Ax = b \text{ does not have nonnegative solution} \\
\iff & b \notin \text{Cone}(A) \\
\iff & \exists y \quad \text{such that } \forall x \in \text{Cone}(A), y^\top x \leq 0 \text{ and } y^\top b > 0 \\
\iff & \exists y \quad \text{such that } y^\top A \leq 0 \text{ and } y^\top b > 0
\end{aligned}$$

The second step follows Theorem 1.11. □

There are several variants of Farkas' Lemma, that can be easily derived from Theorem 1.12.

Corollary 1.12.1. *The system $Ax \leq b$ has a solution x, iff there is no vector y satisfying $y \geq 0$, $y^\top A = 0$ and $y^\top b < 0$.*

[2]Please notice this symbol, as it will be used in further definitions and proofs.

Proof. Let A' be the matrix

$$A' = \begin{pmatrix} A & -A & I \end{pmatrix}$$

Then

$$\begin{aligned} & \text{The system } Ax \le b \text{ has a solution.} \\ \iff & \text{There exists } x_1, x_2, y \ge 0 \text{ such that } A(x_1 - x_2) + y = b \\ \iff & \begin{pmatrix} A & -A & I \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ y \end{pmatrix} = b \qquad \text{where } \begin{pmatrix} x_1 \\ x_2 \\ y \end{pmatrix} \ge 0 \\ \iff & \text{The system } A'x' = b \text{ has a nonnegative solution.} \end{aligned}$$

Applying Theorem 1.12 to the system $A'x' = b$ gives the corollary. □

Corollary 1.12.2. *Suppose the system $Ax \le b$ has at least one solution. Then for every solution x of $Ax \le b$ one has $c^\top x \le \delta$, iff there exists a vector $y \ge 0$ such that $y^\top A = c^\top$ and $y^\top b \le \delta$.*

1.4 Linear programming

One of the standard forms of a linear programming (LP) problem is:

$$\begin{aligned} \text{maximize} \quad & c^\top x \\ \text{subject to} \quad & Ax \le b \end{aligned}$$

So LP can be considered as maximizing a 'linear function' $c^\top x$ over a polyhedron $P = P(A, b)$. Geometrically, this can be seen as shifting a hyperplane to its 'highest' level, under the condition that it intersects P.

Clearly, the minimization problem can be translated to the maximization problem:

$$\min\{c^\top x \mid x \in P(A, b)\} = -\max\{-c^\top x \mid x \in P(A, b)\}$$

One says the x is a *feasible solution* if $x \in P(A, b)$, namely $Ax \le b$. If x attains the maximum, it is called an *optimum solution.*

The main theorem of this section is the Duality theorem of LP, due to von Neumann. The theorem states that if

$$\max\{c^\top x \mid Ax \le b\}$$

is finite, then its duality

$$\min\{y^\top b \mid y \ge 0, y^\top A = c^\top\}$$

is finite, and the value of the maximum is equal to the value of another. In order to show this, we first prove:

Lemma 1.1. *Let P be a nonempty polyhedron in $\mathbb{R}^n$ and let $c \in \mathbb{R}^n$. If $\sup\{c^\top x \mid x \in P\}$ is finite, then $\max\{c^\top x \mid x \in P\}$ is attained.*

Proof. Let $P = P(A, b)$, and $\delta = \sup\{c^\top x \mid x \in P\}$. Next we show there exists an $x \in \mathbb{R}^n$ such that $c^\top x \ge \delta$.

$$\begin{aligned} & \text{There exists some } x \in P, \text{ such that } c^\top x \ge \delta \\ \iff & \text{The system } \begin{pmatrix} A \\ -c^\top \end{pmatrix} x \le \begin{pmatrix} b \\ -\delta \end{pmatrix} \text{ has a solution} \end{aligned}$$

By Farkas' Lemma in the form of Corollary 1.12.1, we have either

$$\begin{pmatrix} A \\ -c^\top \end{pmatrix} x \leq \begin{pmatrix} b \\ -\delta \end{pmatrix} \quad \text{has a solution}$$

or

$$\exists y' \in \mathbb{R}^{n+1},\ y' \geq 0 \quad \text{such that} \quad y'^\top \begin{pmatrix} A \\ -c^\top \end{pmatrix} = 0 \quad \text{and} \quad y'^\top \begin{pmatrix} b \\ -\delta \end{pmatrix} < 0$$

but not both. Next we show the latter case never occur.

Suppose such y' exists. Let

$$y' = \begin{pmatrix} y \\ \lambda \end{pmatrix} \quad \text{where } y \geq 0 \text{ and } \lambda \geq 0$$

If $\lambda = 0$, it follows $y^\top A = 0$ and $y^\top b < 0$. This contradicts the fact that P is nonempty. So $\lambda > 0$. Without loss of generality, $\lambda = 1$. We have $y^\top A = c^\top$ and $y^\top b < \delta$ for all $x \in P$. And it follows a contradiction that for each $x \in P$,

$$\delta = c^\top x = y^\top Ax \leq y^\top b < \delta$$

□

From this we derive:

Theorem 1.13 (Duality theorem of LP). *Let A be an $m \times n$ matrix $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Then*

$$\max\{c^\top x \mid Ax \leq b\} = \min\{y^\top b \mid y \geq 0, y^\top A = c^\top\}$$

provided that both sets are nonempty.

Proof. First note that

$$\sup\{c^\top x \mid Ax \leq b\} \leq \inf\{y^\top b \mid y \geq 0, y^\top A = c^\top\}$$

because for each $x \in P(A, b)$ and $y \in \mathbb{R}^m_+$ satisfying $y^\top A = c^\top$, it holds

$$c^\top x = y^\top Ax = y^\top (Ax) \leq y^\top b$$

As both sets are nonempty, the supremum and infimum are finite. By Lemma 1.1, there exists an $x_0 \in P(A, c)$ and a real δ such that

$$c^\top x_0 = \max\{c^\top x \mid Ax \leq b\} = \delta$$

Next we want to find a vector $y_0 \in \mathbb{R}^m_+$, $y_0^\top A = c^\top$ such that $y_0^\top b = \delta$.

Let $k = r^=_{x_0}(A, b)$. Without loss of generality, we have

$$\begin{cases} a_1^\top x_0 = b_1 \\ \cdots\cdots \\ a_k^\top x_0 = b_k \\ a_{k+1}^\top x_0 < b_{k+1} \\ \cdots\cdots \\ a_m^\top x_0 < b_m \end{cases}$$

where $0 \leq k \leq m$. Then by Theorem 1.11, either

$$c \in \text{Cone}(a_1, \dots, a_k)$$

or

$$\exists y \in \mathbb{R}^m \quad \text{such that } y^\top a_i \leq 0 \text{ and } y^\top c > 0 \text{ for all } i = 1, \dots, k$$

Now we show that the latter case will never occur. Otherwise, if such y exists, we can find a small enough positive ε such that $A(x_0 + \varepsilon y) \leq b$ and $c^\top(x_0 + \varepsilon y) > c^\top x_0$. This contradicts the choice of x_0.

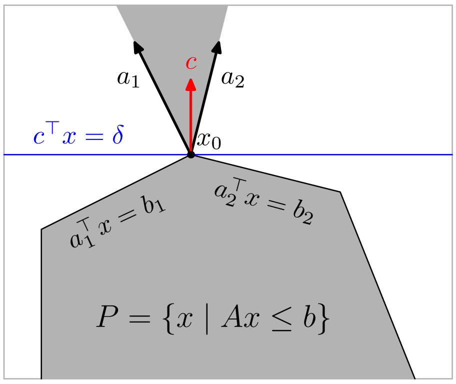

Figure: c always lies in $\text{Cone}(a_1, \dots, a_k)$

Therefore, we have $k \geq 1$ and $c \in \text{Cone}(a_1, \dots, a_k)$, say

$$c = \lambda_1 a_1 + \ldots + \lambda_k a_k \quad \text{where } \lambda_1, \dots, \lambda_k \geq 0$$

Let

$$y_0 = \begin{pmatrix} \lambda_1 & \dots & \lambda_k & 0 & \dots & 0 \end{pmatrix}^\top \in \mathbb{R}^m_+$$

Then it is clear that $y_0^\top A = c^\top$, and

$$\begin{aligned} \delta =& c^\top x_0 \\ =& (\lambda_1 a_1^\top + \ldots + \lambda_k a_k^\top) x_0 \\ =& \lambda_1 b_1 + \ldots + \lambda_k b_k \\ =& y_0^\top b \end{aligned}$$

□

There are some variants of the Duality theorem:

Corollary 1.13.1. *Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Then both*

$$\max\{c^\top x \mid x \geq 0, Ax \leq b\}$$

and

$$\min\{y^\top b \mid y \geq 0, y^\top A \geq c^\top\}$$

exist and are equal, provided both sets are nonempty.

Proof.

$$
\begin{aligned}
&\max\{c^\top x \mid x \geq 0, Ax \leq b\}\\
=&\max\left\{c^\top x \;\middle|\; \begin{pmatrix} A \\ -I \end{pmatrix} x \leq \begin{pmatrix} b \\ 0 \end{pmatrix}\right\}\\
=&\min\left\{ \begin{pmatrix} y_1^\top & y_2^\top \end{pmatrix} \begin{pmatrix} b \\ 0 \end{pmatrix} \;\middle|\; \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \geq 0, \begin{pmatrix} y_1^\top & y_2^\top \end{pmatrix} \begin{pmatrix} A \\ -I \end{pmatrix} = c^\top \right\}\\
=&\min\{y_1^\top b \mid y_1, y_2 \geq 0,\ y_1^\top A - y_2^\top = c^\top\}\\
=&\min\{y^\top b \mid y \geq 0, y^\top A \geq c^\top\}
\end{aligned}
$$

Here the second equality follows the Duality theorem. □

Corollary 1.13.2. *Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Then both*

$$\max\{c^\top x \mid Ax \geq b\}$$

and

$$\min\{y^\top b \mid y \leq 0, y^\top A = c^\top\}$$

exist and are equal, provided both sets are nonempty.

Proof.

$$
\begin{aligned}
&\max\{c^\top x \mid Ax \geq b\}\\
=&\max\{c^\top x \mid (-A)x \leq (-b)\}\\
=&\min\{y^\top(-b) \mid y \geq 0, y^\top(-A) = c^\top\}\\
=&\min\{(-y)^\top b \mid (-y) \leq 0, (-y)^\top A = c^\top\}\\
=&\min\{y^\top b \mid y \leq 0, y^\top A = c^\top\}
\end{aligned}
$$

Here the second equality follows the Duality theorem. □

Theorem 1.14 (Exercise [1] 2.25). *Let a matrix, a column vector, and a row vector be given:*

$$\begin{pmatrix} A & B & C \\ D & E & F \\ G & H & K \end{pmatrix} \qquad \begin{pmatrix} a \\ b \\ c \end{pmatrix} \qquad \begin{pmatrix} d & e & f \end{pmatrix}$$

where $A, B, C, D, E, F, G, H, K$ are matrices, a, b, c are column vectors, and d, e, f are row vectors (of appropriate dimensions). Then

$$
\begin{aligned}
\max\{\ dx + ey + fz :\ & x \geq 0, z \leq 0\\
& Ax + By + Cz \leq a\\
& Dx + Ey + Fz = b\\
& Gx + Hy + Kz \geq c\ \}\\
= \min\{\ ua + vb + wc :\ & u \geq 0, w \leq 0\\
& uA + vD + wG \geq d\\
& uB + vE + wH = e\\
& uC + vF + wK \leq f\ \}
\end{aligned}
$$

Proof.

$$\begin{aligned}
&\max\{\ dx+ey+fz && : \ x\geq 0, z\leq 0\\
& && \quad Ax+By+Cz\leq a\\
& && \quad Dx+Ey+Fz=b\\
& && \quad Gx+Hy+Kz\geq c\ \}\\
&=\max\left\{\ \begin{pmatrix} d & e & -e & -f\end{pmatrix}\begin{pmatrix}x\\ y_1\\ y_2\\ -z\end{pmatrix}\right. && : \ \begin{pmatrix}x\\ y_1\\ y_2\\ -z\end{pmatrix}\geq 0\\
& && \quad \left.\begin{pmatrix} A & B & -B & -C\\ D & E & -E & -F\\ -D & -E & E & F\\ -G & -H & H & K\end{pmatrix}\begin{pmatrix}x\\ y_1\\ y_2\\ -z\end{pmatrix}\leq\begin{pmatrix}a\\ b\\ -b\\ -c\end{pmatrix}\right\}\\
&=\min\left\{\ \begin{pmatrix} u & v_1 & v_2 & -w\end{pmatrix}\begin{pmatrix}a\\ b\\ -b\\ -c\end{pmatrix}\right. && : \ \begin{pmatrix}u^\top\\ v_1^\top\\ v_2^\top\\ -w^\top\end{pmatrix}\geq 0\\
& && \quad \left.\begin{pmatrix} A^\top & D^\top & -D^\top & -G^\top\\ B^\top & E^\top & -E^\top & -H^\top\\ -B^\top & -E^\top & E^\top & H^\top\\ -C^\top & -F^\top & F^\top & K^\top\end{pmatrix}\begin{pmatrix}u^\top\\ v_1^\top\\ v_2^\top\\ -w^\top\end{pmatrix}\geq\begin{pmatrix}d^\top\\ e^\top\\ -e^\top\\ -f^\top\end{pmatrix}\right\}\\
&=\min\{\ ua+vb+wc && : \ u\geq 0, w\leq 0\\
& && \quad uA+vD+wG\geq d\\
& && \quad uB+vE+wH=e\\
& && \quad uC+vF+wK\leq f\ \}
\end{aligned}$$

Here, the first and last equality are doing translations and the middle one follows the Duality theorem of the form 1.13.1. □

Theorem 1.15 (The Von Neumann's Minimax Theorem on two-person zero-sum game). *Let* $P_n=\{x\in\mathbb{R}^n_+ \mid x_1+\ldots+x_n=1\}$. *Then for every* $A\in\mathbb{R}^{m\times n}$,

$$\max_{x\in P_m}\min_{y\in P_n} x^\top Ay=\min_{y\in P_n}\max_{x\in P_m} x^\top Ay$$

Proof. Let

$$v_1=\max_{x\in P_m}\min_{y\in P_n} x^\top Ay$$

and

$$v_2=\min_{y\in P_n}\max_{x\in P_m} x^\top Ay$$

Next we show $v_1=v_2$ by linear programming.

Without loss of generality, $a_{ij}>0$ for all i,j.[3] For any fixed x, $\min_{y\in P_n} x^\top Ay$ is

[3]Let B be an $m\times n$ matrix containing only 1's. Then for all $x\in P_m$ and $y\in P_n$, $x^\top By=1$. So if the minimal element in A is smaller than or equal to 0, say t, let $A'=A+(1-t)B$. Then

$$\max_x\min_y x^\top A'y=\max_x\min_y x^\top Ay+(1-t)=v_1+(1-t)$$
$$\min_y\max_x x^\top A'y=\min_y\max_x x^\top Ay+(1-t)=v_2+(1-t)$$

Hence to prove $v_1=v_2$, we need only to show $\max_x\min_y x^\top A'y=\min_y\max_x x^\top A'y$, where A' is an $m\times n$ matrix with all its elements greater than 0.

attained at an extreme point of the polytope $\{Ay \mid y \in P_n\}$.[4] That is,

$$v_1 = \max_{x\in P_m} \min_{y\in P_n} x^\top Ay = \max_{x\in P_m} \min\{x^\top a'_1, \dots, x^\top a'_n\}$$

This follows v_1 is the maximum of the LP problem:

$$\max_{v\in\mathbb{R}} v$$

subject to

$$\exists x \in P_m,\ x^\top a'_i \geq v,\ i = 1, 2, \dots, n$$

It is obvious that this problem has a feasible solution yielding $v > 0$. So by defining $x'_i = x_i/v$ and $x' = \begin{pmatrix} x'_1 & \dots & x'_n \end{pmatrix}^\top$, we have

$$v_1 = \max \frac{1}{\sum_{i=1}^n x_i/v} = \max \frac{1}{\sum_{i=1}^n x'_i}$$

Hence $1/v_1$ is the minimum of the LP problem:

$$\min \sum_{i=1}^n x'_i = x'^\top \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

subject to

$$x'^\top A \geq \begin{pmatrix} 1 & \dots & 1 \end{pmatrix} \quad \text{and} \quad x' \geq 0$$

Similarly, v_2 is the minimum of the LP problem:

$$\min_{v\in\mathbb{R}} v$$

subject to

$$\exists y \in P_n,\ a_i^\top y \leq v,\ i = 1, 2, \dots, m$$

Define y'_i by y_i/v and $y' = \begin{pmatrix} y'_1 & \dots & y'_m \end{pmatrix}^\top$, we know $1/v_2$ is the maximum of the LP problem:

$$\max \sum_{i=1}^m y'_i = \begin{pmatrix} 1 & \dots & 1 \end{pmatrix} y'$$

subject to

$$Ay' \leq \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad \text{and} \quad y' \geq 0$$

According to the Duality Theorem of the form 1.13.1, we obtain $1/v_1 = 1/v_2$. This completes the proof. □

[4] We will give a strictly proof to this property in Section 1.6 (Corollary 1.24.1).

1.5 Polarity and duality

Definition 1.10. *Let C be a convex cone in $\mathbb{R}^n$, define $C^{\maltese}$ by*

$$C^{\maltese} := \{y \in \mathbb{R}^n \mid x^\top y \leq 0 \quad \forall x \in C\}$$

And $C^{\maltese}$ is called the `polar` *of C.*

Theorem 1.16 (Exercise [1] 2.13). *For any subset C of $\mathbb{R}^n$,*

(i) For each convex cone C, $C^{\maltese}$ is a closed convex cone.

(ii) For each closed convex cone C, $C^{\maltese\maltese} = C$.

Proof. (i). For any $y_1, y_2 \in C^{\maltese}$, it holds $x^\top y_1 \leq 0, x^\top y_2 \leq 0$ for all $x \in C$. And for any $\lambda, \mu \in \mathbb{R}_+$, we have

$$x^\top(\lambda y_1 + \mu y_2) = \lambda(x^\top y_1) + \mu(x^\top y_2) \leq 0$$

Hence $\lambda y_1 + \mu y_2 \in C^{\maltese}$. So $C^{\maltese}$ is a closed convex cone. □

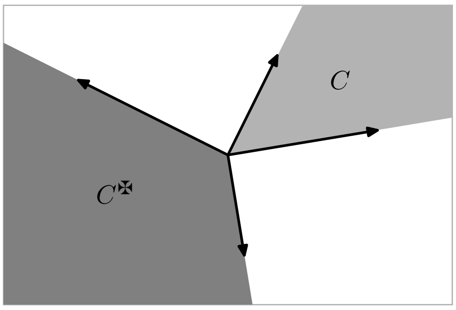

Figure: C and $C^{\maltese}$

Proof. (ii). For any $x \in C$, by the definition of $C^{\maltese}$, we have $y^\top x \leq 0$ for all $y \in C^{\maltese}$. So $C \subseteq C^{\maltese\maltese}$.

And for any $x' \notin C$, by Theorem 1.11, there exists a hyperplane $c^\top x = 0$, such that for all $x \in C$, $c^\top x \leq 0$ and $c^\top x' > 0$. From the definition of $C^{\maltese}$, we know $c \in C^{\maltese}$. And since $c^\top x' > 0$, $x' \notin C^{\maltese\maltese}$. So $C = C^{\maltese\maltese}$. □

Theorem 1.17. *Let $A \in \mathbb{R}^{m\times n}$, and C be the cone $P(A, 0)$. Then*

$$C = \text{Cone}(A^\top)^{\maltese}$$

Proof. For each $x \in C$, that is, $Ax \leq 0$. And for all $y \in \text{Cone}(A^\top)$, that is, $y = \lambda_1 a_1 + \ldots + \lambda_m a_m$ where $\lambda_1, \ldots, \lambda_m \geq 0$, we have

$$y^\top x = \sum_{i=1}^m (\lambda_i a_i^\top) x = \sum_{i=1}^m \lambda_i (a_i^\top x) \leq 0$$

So $x \in \text{Cone}(A^\top)^{\maltese}$. Thus $C \subseteq \text{Cone}(A^\top)^{\maltese}$.

For each $x' \notin C$, there exists an $i \in \{1, 2, \ldots, m\}$, such that $a_i^\top x' > 0$. Let $y = a_i \in \text{Cone}(A^\top)$, then $y^\top x' > 0$. So $x' \notin \text{Cone}(A^\top)^{\maltese}$. Therefore $C = \text{Cone}(A^\top)^{\maltese}$. □

Theorem 1.18 (Exercise [1] 2.28). *Let A be an $m \times n$ matrix and let $b \in \mathbb{R}^n$. Let $P = P(A, b)$, $P \neq \emptyset$ and C be the convex cone $P(A, 0)$. Let the the set $\mathcal{C}$ consists of all vectors c for which $\max\{c^\top x \mid x \in P\}$ is finite. Then $\mathcal{C} = C^{\maltese}$.*

Proof. To prove $\mathcal{C} = C^{\maltese}$, we just show that $\mathcal{C} \subseteq C^{\maltese}$ and conversely $C^{\maltese} \subseteq \mathcal{C}$. And by Theorem 1.16 and Theorem 1.17, we know $C^{\maltese} = \text{Cone}(A^\top)$.

For all $c \in \mathcal{C}$, by the Duality theorem of LP of the form Theorem 1.13 we obtain $c^\top = y^\top A$ for some $y \geq 0$. That is, $c \in \text{Cone}(A^\top)$. So $\mathcal{C} \subseteq C^{\maltese}$.

Conversely, for each $c \in C^{\maltese}$, we have

$$c = \sum_{i=1}^{m} \lambda_i a_i \quad \text{where} \quad \lambda_1, \ldots, \lambda_m \geq 0, \ \sum_{i=1}^{m} \lambda_i = 1$$

Then for all $x \in P$,

$$c^\top x = \sum_{i=1}^{m} \lambda_i a_i^\top x \leq \sum_{i=1}^{m} \lambda_i b_i \leq \sum_{i=1}^{m} b_i$$

So $\max\{c^\top x \mid x \in P\}$ is finite, namely $C^{\maltese} \subseteq \mathcal{C}$. □

Next we extend the concept of polar to Polyhedra by define:

Definition 1.11. *Let P be a polyhedron in $\mathbb{R}^n$, define P^* by*

$$P^* := \{y \in \mathbb{R}^n \mid x^\top y \leq 1 \quad \forall x \in P\}$$

And P^* is called the `dual` *of P.*

Proposition 1.19. *Let C be a convex cone in $\mathbb{R}^n$. Then $C^{\maltese} = C^*$.*

Proof. By the definition of $C^{\maltese}$ and C^*, we know $C^{\maltese} \subseteq C^*$. For each $x \notin C^{\maltese}$, namely there exists a $c \in C$, such that $c^\top x = \delta > 0$. Without loss of generality, $\delta = 1$. Then $x \notin C^*$. So $C^{\maltese} = C^*$. □

Theorem 1.20 (Exercise [1] 2.14). *Let P be a polyhedron.*

(i) P^* is again a polyhedron.

(ii) P contains the origin, iff $(P^*)^* = P$.

(iii) The origin is an internal point of P, iff P^* is bounded.

Proof. (i). Let $X = \text{lift}(P)$, $Q = \text{Cone}(X)$. Then by Theorem 1.16, Q^* is again a convex cone.

Since

$$x^\top y \leq 1 \iff \begin{pmatrix} x & 1 \end{pmatrix} \begin{pmatrix} y \\ -1 \end{pmatrix} \leq 0$$

and it follows

$$y \in P^* \iff \begin{pmatrix} y \\ -1 \end{pmatrix} \in Q^*$$

Hence $P^* = -\text{slice}(-Q^*)$ is a polyhedron[5]. □

Proof. (ii). Necessity. It is easy to show that $\forall p \in P$, $p \in (P^*)^*$. So $P \subseteq (P^*)^*$.

For any $p' \notin P$, there exists a hyperplane $c^\top x = \delta$ such that $\forall p \in P$, $c^\top p < \delta$, and $c^\top p' > \delta$. Since P contains the origin, $c^\top 0 = 0 < \delta$. Without loss of generality, $\delta = 1$. It follows $\forall p \in P$, $c^\top p < 1$. So $c \in P^*$. As $c^\top p' > 1$, $p' \notin (P^*)^*$. So $P = (P^*)^*$.

Sufficiency. Since $0 \in (P^*)^*$ and $P = (P^*)^*$, one has $0 \in P$. □

[5] Here we use $-\mathcal{A}$ to express the set $\{-x \mid x \in \mathcal{A}\}$.

Proof. (iii). Necessity. If the origin is an internal point of P, that is there exists $r > 0$, such that $B(0, r) \subseteq P$. And for each $y \neq 0$ in P^*, let $x = r||y||^{-1}y$. Since $||x|| = r$, $x \in B(0, r) \subseteq P$. So $x^\top y \leq 1$. That is, $r||y|| \leq 1$, $||y|| \leq 1/r$. Hence $P^* \subseteq B(0, 1/r)$, P^* is bounded.

Sufficiency. If P^* is bounded, that is, P^* is a polytope. Let

$$P^* = \text{Conv.hull}(y_1, \ldots, y_t)$$

Then $(P^*)^* = \{x \in \mathbb{R}^n \mid y_j^\top x \leq 1 \quad j = 1, \ldots, t\}$.

If $0 \notin P$, there exists a hyperplane $c^\top x = \delta$, such that for each $p \in P$, $c^\top p < \delta$, and $c^\top 0 = 0 > \delta$. So $kc \in P^*$, $\forall k \in \mathbb{R}_+$. This contradicts P^* is bounded. So $0 \in P$, therefore $P = (P^*)^*$, namely

$$P = \{x \in \mathbb{R}^n \mid y_j^\top x \leq 1 \quad j = 1, \ldots, t\}$$

As $y_j^\top 0 = 0 < 1$ for all $j = 1, \ldots, t$, the origin is an internal point of P. □

Now we recall the basic idea of the proof of Theorem 1.6. In fact, the idea is quite simple.

We have a polytope P, and we want to prove it is also a bounded polyhedron. It is not so easy to do directly, so we use a indirect way to show this. First we show the polar of a polytope is a polyhedron, thus P^* is a polyhedron. And by Theorem 1.5.1, we know P^* is a polytope. This follows $(P^*)^*$ is a polyhedron. Since $P = (P^*)^*$, we obtain P is a polyhedron.

1.6 Faces, edges and vertices

Definition 1.12. *Let $P = P(A, b)$ is a polyhedron with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$. Then $x \in P(A, b)$ is called an* `inner point` *of P, iff $a_i^\top x < b_i$ for all $i \notin r^=(A, b)$. That is, $r_x^=(A, b) = r^=(A, b)$.*

Proposition 1.21. *Every nonempty polyhedron has an inner point.*

Proof. Let P be a polyhedron in $\mathbb{R}^n$, and we prove by induction on n.

If P is contained in some hyperplane, the induction hypothesis gives the proposition. So we may assume that P is not contained in any affine hyperplane. This follows there are t affinely independent vectors $x_1, x_2, \ldots, x_t \in P$, namely $x_2 - x_1, \ldots, x_t - x_1$ span $\mathbb{R}^n$. It follows that there exists a vector $x_0 \in P$ and a real $r > 0$, such that $B(x_0, r)$ is contained in P. Then x_0 is an inner point of P. □

Next, we give the definition of the dimension of a polyhedron:

Definition 1.13. *A polyhedron P is of* `dimension` *k, denoted by $\dim(P) = k$, if the maximal number of affinely independent points in P is $k + 1$.*

We say $\dim(P) = -1$, if $P = \emptyset$.

Lemma 1.2. *Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Then $\dim(P(A, b)) = n - \text{rank}(A_b)$ provided that $P(A, b)$ is nonempty.*

Proof. The solution space for $A_b x = 0$ is of dimension $n - \text{rank}(A_b)$. That is, there are $n - \text{rank}(A_b)$ linearly independent vectors $x_1, \ldots, x_{n-\text{rank}(A_b)}$ satisfying $A_b x = 0$. And this follows $0, x_1, \ldots, x_{n-\text{rank}(A_b)}$ are affinely independent. According to Proposition 1.21, we can take an inner point in $P(A, b)$, say $\tilde{x}$. This implies that there exists a small enough real $\varepsilon > 0$, satisfying $\tilde{x}, \tilde{x} + \varepsilon x_1, \ldots, \tilde{x} + \varepsilon x_{n-\mathrm{rank}(A_b)}$ are affinely independent points in $P(A, b)$. So $\dim(P(A, b)) \geq n - \mathrm{rank}(A_b)$.

Next we show that $\dim(P(A, b)) \leq n - \mathrm{rank}(A_b)$. Let $y_1, \ldots, y_k$ be a set of k affinely independent points in $P(A, b)$. By the definition of A_b, let $\beta \in \mathbb{R}^{|r^{=}(A,b)|}$ be a vector containing all b_i where $i \in r^{=}(A, b)$. Then to any $x \in P(A, b)$, $A_b x = \beta$. Hence $y_1, \ldots, y_k$ are affinely independent solutions to the system $A_b x = \beta$. So $k \leq n - \mathrm{rank}(A_b) + 1$. Therefore $\dim(P(A, b)) \leq n - \mathrm{rank}(A_b)$. □

Before giving the first theorem in this section, we provide some definitions:

Definition 1.14. *Let $c \in \mathbb{R}^n$, $\delta \in \mathbb{R}$ and P be a nonempty polyhedron. Then the halfspace $\{x \in \mathbb{R}^n \mid c^\top x \leq \delta\}$ is called a* **`supporting halfspace`** *of P, iff it contains P.*

If $H = \{x \mid c^\top x \leq \delta\}$ is a supporting halfspace of P, then we say $F := P \cap \{x \mid c^\top x = \delta\}$ is a **`face`** *of P, and $c^\top x \leq \delta$* **`represents`** *it. A face F is said to be* **`proper`** *if $F \neq \emptyset$ and $F \neq P$. If $F \neq \emptyset$, we say that $Ax \leq b$* **`supports`** *F.*

The face of dimension $\dim(P) - 1$ is called a **`facet`** *of P. The face of dimension 0 is called a* **`vertex`** *of P.*

The face F is supported by $c^\top x \leq \delta$ iff $\max\{c^\top x \mid x \in P\} = \delta$, and in such a case it holds F is the set of optimal solutions to the LP program $\max\{c^\top x\}$ subject to $x \in P$. By Lemma 1.1 on page 12, we know nonempty F exists if $\max\{c^\top x \mid x \in P\}$ is finite.

And by Theorem 1.4 on page 5 and Theorem 1.2, we obtain:

Theorem 1.22. *Let P be a nonempty polyhedron $P(A, b)$ where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$. Then following assertions are equivalent:*

1. *z is an extreme point of P.*
2. *z is a vertex of P.*
3. *$\mathrm{rank}(A_{z,b}) = n$.*

Theorem 1.23. *Let A be an $m \times n$ matrix, $b \in \mathbb{R}^n$, $P = P(A, b)$ and F is a nonempty face of P. Then F is a polyhedron of the form*

$$F = \{\, x \in \mathbb{R}^n \ : \begin{array}{ll} a_i^\top x = b_i & \forall i \in r_F^{=}(A, b) \\ a_i^\top x \leq b_i & \forall i \notin r_F^{=}(A, b) \,\} \end{array}$$

Proof. Let F be a face of P supported by $c^\top x \leq \delta$. Then by the definition of F, we have

$$\begin{pmatrix} A \\ c^\top \\ -c^\top \end{pmatrix} x \leq \begin{pmatrix} b \\ \delta \\ -\delta \end{pmatrix}$$

So F is a polyhedron.

Assume $r_F^{=}(A, b) = k$. Without loss of generality, $r_F^{=}(A, b) = \{1, 2, \ldots, k\}$. Then let

$$F' = \{x \in \mathbb{R}^n \mid a_1^\top x = b_1,\ \ldots,\ a_k^\top x = b_k,\ a_{k+1}^\top x \leq b_{k+1},\ \ldots,\ a_m^\top x \leq b_m\}$$

The definition of F' directly follows $F \subseteq F'$. Next we show that $F' \subseteq F$.

By Proposition 1.21, we can take an inner point of F, say x'. This follows:

$$\begin{aligned} & r^{=}_{x'}(A,b) \\ = \;& r^{=}_{x'}\left(\begin{pmatrix} A \\ c^\top \\ -c^\top \end{pmatrix}, \begin{pmatrix} b \\ \delta \\ -\delta \end{pmatrix}\right)\setminus\{m+1, m+2\} \\ = \;& r^{=}_{F}\left(\begin{pmatrix} A \\ c^\top \\ -c^\top \end{pmatrix}, \begin{pmatrix} b \\ \delta \\ -\delta \end{pmatrix}\right)\setminus\{m+1, m+2\} \\ = \;& r^{=}_{F}(A,b) \end{aligned}$$

So $a_i^\top x' = b_i$ for $i = 1, 2, \ldots, k$ and $a_i^\top x' < b_i$ for $i = k+1, \ldots, m$. This gives $c \in \text{Cone}(a_1, \ldots, a_k)$. Let $c = \lambda_1 a_1 + \ldots + \lambda_k a_k$ and $y' = \begin{pmatrix}\lambda_1 & \ldots & \lambda_k & 0 & \ldots & 0\end{pmatrix}^\top$, then $y'^\top b = \delta$ and $y'^\top A = c^\top$. For all $x \in F'$, we have $c^\top x = y'^\top Ax = y'^\top b = \delta$. So $F' \subseteq F$. $\square$

This theorem directly gives:

Corollary 1.23.1. *The number of distinct faces of a polyhedron is finite.*

Proof. To a polyhedron $P = P(A, b)$ where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^n$, the number of all its distinct faces will not exceed 2^m. $\square$

This property shows that to a polyhedron P, although there can be infinite many of vector c which satisfies $\delta = \max\{c^\top x \mid x \in P\}$ is finite, the number of distinct faces supported by $c^\top x \le \delta$ is finite.

Moreover, we have:

Corollary 1.23.2. *The intersection of two faces of P is again a face of P.*

Corollary 1.23.3. *If P is a polyhedron, F is a face of P and G is a face of F. Then G is also a face of P.*

Theorem 1.24. *Suppose $P = P(A, b) \neq \emptyset$ and $\text{rank}(A) = n - k$. Then each minimal nonempty face of P under inclusion has dimension k.*

Proof. Let F be a minimal nonempty face of P under inclusion. If $\dim(F) = 0$, according to Lemma 1.2 and Theorem 1.23, $0 = \dim(F) = n - \text{rank}(A_{F,b}) \ge n - \text{rank}(A)$, so $\text{rank}(A) = n$, $k = 0$. The theorem follows.

Next we assume $\dim(F) > 0$. By Proposition 1.21, we can take an inner point x' of F. Since $\dim(F) > 0$, there exists a $y \in F$, $y \neq x$. Consider the line L connecting x' and y:

$$L = \{x \mid x = x' + \lambda(y - x')\} \quad \text{where } \lambda \in \mathbb{R}$$

If L intersects with a hyperplane $H = \{x \mid a_i^\top x = b_i\}$, where $i \notin r^{=}_F(A, b) = r^{=}_{x'}(A, b)$. Then the nonempty set $F' = F \cap H$ does not contain x'. Thus $r^{=}_{F'}(A, b) \supseteq r^{=}_F(A, b) \cup \{i\}$. And this implies $\dim(F') < \dim(F)$, which contradicts the choice of F.

So L does not intersect with any hyperplane $H = \{x \mid a_i^\top x = b_i\}$ where $i \notin r^{=}_F(A, b)$. This follows L is contained in P, namely $Ax' + \lambda A(y - x') \le b$ for all $\lambda \in \mathbb{R}$. So $A(y - x') = 0$ for all $y \in F$. This gives $F = \{y \mid Ay = Ax'\}$. As $\text{rank}(A) = n - k$, by Theorem 1.23 we obtain $\dim(F) = k$. $\square$

As a direct consequence of this theorem, we have

Corollary 1.24.1 (Exercise [1] 2.22). *Let $P = P(A, b)$ be a nonempty polyhedron with at least one vertex. Then for all vector c such that*

$$\delta = \max\{c^\top x \mid x \in P\}$$

is finite, there exists a vertex x' of P, satisfying $c^\top x' = \delta$.

Corollary 1.24.2. *For $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}$, $P(A, b) \neq \emptyset$ has a vertex iff* $\text{rank}(A) = n$.

Since an $m\times n$ matrix A with $\text{rank}(A) = n$ has at most C_m^n invertible submatrices with the size of $n \times n$, we obtain that the polyhedron $P(A, b)$ contains at most C_m^n vertices.

Chapter 2

Integer programming, totally unimodular matrices

2.1 Integer linear programming

Let $b \in \mathbb{R}^m$, $c, d \in \mathbb{R}^n$ and $A, B \in \mathbb{R}^{m \times n}$. To the LP problem

$$\max\{c^\top x + d^\top y\}$$

subject to

$$x, y \in \mathbb{R}^n_+,\ Ax + By \leq b$$

if we add one more restriction that y is an *integer* vector, namely $y \in \mathbb{Z}^n_+$, the LP problem is now called a *mixed-integer* linear programming (MIP) problem.

Moreover, an *integer* linear programming (IP) problem is a special case of MIP problem in which there are no continuous variables:

$$\max\{c^\top x \mid x \in \mathbb{Z}^n_+, Ax \leq b\}$$

Although LP is solvable in polynomial time, the general IP problem is $\mathcal{NP}$-complete. However, in some special classes of IP problem, polynomial-time algorithms have been found.

Definition 2.1. *To a polyhedron $P = P(A, b)$, if for each vector c satisfying $\delta = \max\{c^\top x \mid x \in P\}$ is finite, there exists an integer vector $x' \in P$, such that $c^\top x' = \delta$, we say the polyhedron P is* `integral`.

By Definition 2.1, we know the IP problem over an integral polyhedron P collapses to an LP problem.

Definition 2.2. *Let P be a convex set in $\mathbb{R}^n$. Define the* `integer hull` *of P by*

$$\text{Conv.hull}(P \cap \mathbb{Z}^n)$$

which is the convex hull generated by all integer points in P.

Proposition 2.1. *A polyhedron P is integral iff the integer hull of P is P itself.*

Proof. Necessity. By Theorem 1.10 on page 10, we know $P = Q + C$ where Q is a polytope and C is a finitely generated cone. Without loss of generality, each vertex of Q is also the vertex of P. Since P is integral, Q and C can both be generated by integer points. Therefore the integer hull of P equals to P itself.

Sufficiency. Since the integer hull of P is P itself, we know each face of P which is the set of solution to some optimization problem on P, contains an integer point. □

2.2 Totally unimodular matrices

Definition 2.3. *A* **`unimodular matrix`** *is a square integer matrix whose determinant has magnitude* 1.

A **`totally unimodular matrix`** *is a matrix with all its square submatrix having determinant equals to* 0, +1 *or* −1. *That is, all nonsingular square submatrices of a totally unimodular matrix are unimodular.*

Proposition 2.2. *Let A be a totally unimodular $n \times m$ matrix. Then*

$$-A,\ A^\top,\ \begin{pmatrix} A & I_n \end{pmatrix},\ \begin{pmatrix} A \\ I_m \end{pmatrix},\ \begin{pmatrix} A & -A \end{pmatrix} \text{ and } \begin{pmatrix} A \\ -A \end{pmatrix}$$

are also totally unimodular.

Proof. Firstly, it is quite clear that $-A$ and $A^\top$ are totally unimodular.

Secondly, we show $\begin{pmatrix} A & I_n \end{pmatrix}$ is totally unimodular. For each square $k \times k$ submatrix B of $\begin{pmatrix} A & I_n \end{pmatrix}$, suppose the first t columns of B come from A, and the other $k - t$ ones come from I_n. Next we induction on t.

If $t = k$, B is a submatrix of A. Since A is totally unimodular, $|\det B|$ equals to 0 or 1. If $t < k$, the last column of B contains all 0's but one 1. Let the 1 be on the r-th row of B, then $\det(B) = (-1)^{r+k}(\det B')$, where B' is a $(k-1) \times (k-1)$ matrix given by B omitting the r-th row and the k-th column. By our induction hypothesis, $|\det B'|$ equals to 0 or 1. Hence $|\det B|$ also equals 0 or 1. This follows $\begin{pmatrix} A & I_n \end{pmatrix}$ is totally unimodular.

Similarly we can show $\begin{pmatrix} A \\ I_m \end{pmatrix}$ is totally unimodular.

Next we prove $\begin{pmatrix} A & -A \end{pmatrix}$ is totally unimodular. Let B be a submatrix of $\begin{pmatrix} A & -A \end{pmatrix}$. If there exists a column i, such that the i-th columns of A and $-A$ both occur in B, we can easily derive that $\det B = 0$.

Otherwise B can be given by rearranging the columns of a submatrix of A, say C, then multiply -1 to some of its columns. This implies $|\det B| = |\det C|$, namely $|\det B|$ equals 0 or 1. □

Next we show the relationship between totally unimodular matrices and the IP problem:

Lemma 2.1. *Let A be a totally unimodular $m \times n$ matrix and let $b \in \mathbb{Z}^m$. Then each vertex of the polyhedron $P = P(A, b)$ is an integer vector.*

Proof. Let z be a vertex of P. By Theorem 1.22 on page 21, we know $\operatorname{rank}(A_{z,b}) = n$. Hence $A_{z,b}$ has a nonsingular $n \times n$ submatrix A'. Let b' be the part of b corresponding to the rows of A that occur in A'. Then, we have $A'z = b'$.

Since A is totally unimodular, $|\det A'| = 1$. Because A' is an integer matrix, $(A')^{-1}$ is also integer. Therefore $z' = b'(A')^{-1}$ is an integer vector. □

Theorem 2.3. *Let A be a totally unimodular $m \times n$ matrix and let $b \in \mathbb{Z}^m$. Then the polyhedron $P = P(A, b)$ is integral.*

Proof. Let $c \in \mathbb{R}^n$ and $x = x^*$ be an optimum solution of the LP problem

$$\max\{c^\top x \mid x \in P\}$$

Choose $d_1, d_2 \in \mathbb{Z}^n$ such that $d_1 \le x^* \le d_2$, and consider that polyhedron

$$Q := \{x \in \mathbb{R}^n \mid Ax \le b,\ d_1 \le x \le d_2\}$$

Let

$$A' := \begin{pmatrix} A \\ -I \\ I \end{pmatrix} \quad \text{and} \quad b' := \begin{pmatrix} b \\ -d_1 \\ d_2 \end{pmatrix}$$

then

$$Q = \{x \in \mathbb{R}^n \mid A'x \le b'\}$$

Since Q is of full column rank, by Theorem 1.24 on page 22, Q contains a vertex. And by Corollary 1.24.1 we obtain the maximum of LP problem $\{c^\top x \mid x \in Q\}$ is attained at a vertex of Q, say $\tilde{x}$. By Proposition 2.2, we know A' is totally unimodular. Thus according to Lemma 2.1, $\tilde{x}$ is integer.

As $x^* \in Q$, $c^\top \tilde{x} \ge c^\top x^*$. Hence $\tilde{x}$ is also an optimum solution of LP problem

$$\max\{c^\top x \mid x \in P\}$$

□

By Proposition 2.2 and Theorem 2.3, we can derive:

Corollary 2.3.1. *Let A be an $m \times n$ totally unimodular matrix. Let $b \in \mathbb{Z}^m$ and $u \in \mathbb{Z}^n$. Then each of the following polyhedra is integral:*

1. $P(A, b)$
2. $P(A, b) \cap \mathbb{R}^n_+$
3. $P^=(A, b)$
4. $P^=(A, b) \cap \{x \mid x \le u\}$

Proof. $P(A, b)$ being integral is directly given by Theorem 2.3.

Let

$$A' = \begin{pmatrix} A \\ -I \end{pmatrix} \quad \text{and} \quad b' = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

we know A' is totally unimodular and $b' \in \mathbb{Z}^{m+n}$. Thus $P(A', b') = P(A, b) \cap \mathbb{R}^n_+$ is intergal.

Similarly, let

$$A' = \begin{pmatrix} A \\ -A \end{pmatrix} \quad \text{and} \quad b' = \begin{pmatrix} b \\ -b \end{pmatrix}$$

then A' is totally unimodular and $b' \in \mathbb{Z}^{2m}$. So $P(A', b') = P^=(A, b)$ is intergal.

Finally, let

$$A' = \begin{pmatrix} A \\ -A \\ I \end{pmatrix} \quad \text{and} \quad b' = \begin{pmatrix} b \\ -b \\ u \end{pmatrix}$$

then A' is totally unimodular and $b' \in \mathbb{Z}^{2m+n}$. So $P(A', b') = P^=(A, b) \cap \{x \mid x \le u\}$ is intergal. □

In the following part of this section, we will prove the famous *Hoffman-Kruskal Theorem.* As a preparation, we prove:

Lemma 2.2. *Let $b \in \mathbb{Z}^m$, and A be an integer $m \times n$ matrix such that the polyhedron*

$$P = P(A, b) \cap \mathbb{R}_+^n = \{x \mid Ax \leq b,\ x \geq 0\}$$

is integral. And let $B = \begin{pmatrix} A & I \end{pmatrix}$. Then for any $c \in \mathbb{Z}^n$, each vertex (if any) of the polyhedron

$$Q = P^=(B, c) \cap \mathbb{R}_+^n = \{x \mid Bx = c,\ x \geq 0\}$$

is integer.

Proof. Let

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

be a vertex of Q where $z_1 \in \mathbb{R}_+^n$ and $z_2 \in \mathbb{R}_+^m$. Since $z \in Q$, namely $Bz = Az_1 + z_2 = c$, we have $Az_1 = c - z_2 \leq c$. This means $z_1 \in P$. Moreover, z_1 is an extreme point of P. If not so, namely there exists $u, v \in P$, such that $u \neq v$ and $z_1 = (u + v)/2$, we have

$$z_2 = c - Az_1 = c - \frac{A(u + v)}{2} = \frac{1}{2}(c - Au) + \frac{1}{2}(c - Av)$$

Thus

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} u \\ c - Au \end{pmatrix} + \frac{1}{2}\begin{pmatrix} v \\ c - Av \end{pmatrix}$$

This contradicts the fact that z is a vertex of Q.

Since P is integral, z_1 is integer. Therefore $z_2 = c - Az_1$ is also integer. □

With Lemma 2.2 in hand, we can now start to prove *Hoffman-Kruskal Theorem*:

Theorem 2.4 (Hoffman-Kruskal Theorem). *Let $A \in \mathbb{Z}^{m \times n}$. Then A is totally unimodular, iff for each $b \in \mathbb{Z}^m$, the polyhedron*

$$P = P(A, b) \cap \mathbb{R}_+^n = \{x \mid Ax \leq b,\ x \geq 0\}$$

is integral.

Proof. Necessity. Let $A' = \begin{pmatrix} A \\ I \end{pmatrix}$ and $b' = \begin{pmatrix} b \\ 0 \end{pmatrix}$, then A' is also totally unimodular and $P = P(A', b')$. By Theorem 2.3, we obtain P is integral.

Sufficiency. Let $B = \begin{pmatrix} A & I \end{pmatrix} \in \mathbb{Z}^{m \times (n+m)}$. Then A is totally unimodular iff each nonsingular $m \times m$ submatrix of B has determinant ± 1.

Let $C \in \mathbb{Z}^{m \times m}$ be a unsingular submatrix of B. Next we show C^{-1} is integer. To any $v \in \mathbb{Z}^m$, there exists another vector $u \in \mathbb{Z}^m$ such that

$$z = u + C^{-1}v \in \mathbb{Z}_+^m$$

Let $b = Cz$, then $b = Cz = Cu + CC^{-1}v = Cu + v$ is integer.

Without loss of generality, C contains first m columns of B, namely $B = \begin{pmatrix} C & D \end{pmatrix}$. And we raise z to $z' \in \mathbb{Z}^{n+m}$:

$$z' = \begin{pmatrix} z \\ \mathbf{0} \end{pmatrix}$$

where $\mathbf{0}$ is an all-zero vector in $\mathbb{R}^n$.

Let

$$E = \begin{pmatrix} B \\ -B \\ -I \end{pmatrix} \quad \text{and} \quad f = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$$

Then for system

$$Ez' \leq f$$

the equality holds for the first m rows and last n rows of E. Since $\text{rank}(C) = m$, and the last n rows come from $-I$, we obtain $\text{rank}(E_{z',f}) = m + n$. By Theorem 1.4 on page 5, z' is a vertex of polyhedron $P(E, f) = P^{=}(B, b) \cap \mathbb{R}^m_+$. By Lemma 2.2, z' is integer. So z is integer.

Hence for any integer vector v, $C^{-1}v = z - u$ is integer. Therefore C^{-1} is integer. This implies C is integer. □

Bibliography

[1] Alexander Schrijver: *A Course in Combinatorial Optimization*, 2004

[2] Course slides by Yaokun Wu: *Graph Theory*, 2004