Disclaimer: These notes have not been subjected to the usual scrutiny reserved for formal publications.

Recall the definition of a polyhedron: A polyhedron is a set of points that satisfy a finite set of linear inequalities, i.e., \( P = \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \} \) for some constraint matrix \( A \) and RHS vector \( b \). Also recall that an IP formulation is given by \( \max \{ c^T x : x \in P \cap \mathbb{Z}^n \} \), i.e.,

\[
IP = \max \{ c^T x : Ax = b, x \geq 0, x \in \mathbb{Z}^n \}.
\] (3.1)

An obvious algorithmic attempt to solve the IP is to forget about the integrality constraints, solve the resulting LP relaxation of the problem and get an optimal solution \( x^* \). If the solution \( x^* \) is integral then \( x^* \) is also an optimum for the IP and hence we are done; otherwise, we could suitably round the solution (e.g., set \( x_{round}(i) := \lfloor x^*_{LP}(i) \rfloor \)). This technique has two issues:

1. The rounded solution could be infeasible for the IP.
2. The objective value of the rounded solution could be far from the objective value of the optimal solution of the IP.

Exercise: Give an IP to illustrate these issues.

However, sometimes rounding the optimal solution may give an IP-feasible solution which is approximately optimal. This is a technique for designing approximation algorithms. Sometimes we could get even lucky: the optimum solution \( x^*_{LP} \) to the LP relaxation could be integral. For several discrete optimization problems, we do have IP-formulations for which this lucky scenario happens. In order to understand and have tools to recognize such scenarios, it is important to learn the fundamentals of linear programs. In the next few lectures, we will focus on the fundamentals of linear programs.

### 3.1 Optimization and Feasibility

Optimization problems can be solved (efficiently) provided a corresponding feasibility problem can be solved (efficiently). If we can solve the feasibility version of IP (3.1), then we can solve the optimization version as follows:

- Verify feasibility of \( Ax = b, x \geq 0, c^T x \geq \delta, x \in \mathbb{Z}^n \)
- Do binary search over the choice of \( \delta \).

So, we will focus on the feasibility version. We would like to understand

**Question 0:** When is \( Ax = b, x \geq 0, x \in \mathbb{Z}^n \) feasible?
We will first focus on a simpler question:

**Question 1:** When is $Ax = b, x \geq 0$ feasible?

In fact, let us refresh by answering an even simpler question:

**Question 2:** When is $Ax = b$ feasible?

We already know how to answer Question 2. From linear algebra, recall that the system $Ax = b$ is feasible if and only if the vector $b$ lies in the column span of $A$, i.e., $b$ can be written as a linear combination of the columns of $A$. Formally, we have the following proposition:

**Proposition 1.** $Ax = b$ has a solution if and only if rank $(A) = \text{rank}[A \ b]$.

The goal of this lecture is to answer Question 1. It is asking whether $b$ can be written as a non-negative linear combination of columns of $A$. For example, let $a_1, a_2, a_3$ be the columns of $A$ (consider Figure 3.1). In Figure 3.1(i), we can write $b$ as a non-negative linear combination of $a_i$s but in Figure 3.1(ii), the point $b$ cannot be written as a non-negative linear combination of $a_i$s. This is because we have a hyperplane that separates $b$ from all the $a_i$s.

![Figure 3.1: Example of feasibility](image)

**3.2 Fundamental theorem of linear inequalities**

The fundamental theorem of linear inequalities allows us to answer Question 1 by formalizing the above-mentioned two cases. It shows that exactly one of the two cases can occur.

**Theorem 2.** Let $a_1,\ldots,a_n \in \mathbb{R}^m, \ t = \text{rank}[a_1 \ a_2 \ \ldots \ a_n \ b]$. Then exactly one of the following statements hold:

(i) $b$ is a non-negative linear combination of linearly independent vectors from $a_1,\ldots,a_n$.

(ii) There exists a hyperplane $\{x : c^T x = 0\}$ containing $t - 1$ linearly independent vectors from $a_1,\ldots,a_n$ s.t $c^T b < 0$ and $c^T a_1 \geq 0,\ldots,c^T a_n \geq 0$. If $a_1,\ldots,a_n,b$ are rational, then $c$ is rational.

**Proof.** Constructive proof by Simplex Algorithm (see Schrijver’s book on theory of linear and integer programming or the handout posted in compass2g).
The theory of linear programming is built on Theorem 2. It allows us to answer Question 1 using the following lemma:

**Lemma 3 (Farkas Lemma).** $Ax = b, x \geq 0$ has a solution if and only if there does not exist $y$ such that $y^T A \geq 0, y^T b < 0$.

**Proof.** Let $a_1, \ldots, a_n$ be columns of $A$. Applying fundamental theorem to $a_1, \ldots, a_n, b$, we know that exactly one of two situations (i) and (ii) can happen.

- (i) gives a solution $x$ for $Ax = b, x \geq 0$.
- (ii) gives $c$ for which $c^T b < 0, c^T a_1 \geq 0, c^T a_2 \geq 0, \ldots, c^T a_n \geq 0$. Set $y = c$. Then $y^T A \geq 0, y^T b < 0$.

Fundamental theorem has several important consequences. These consequences are the tools to understand ideal formulations for IPs, so we will learn these consequences.

### 3.3 Polyhedral Theory: Cones

Cones are fundamental geometric structures in the theory of linear inequalities.

**Definition 4.** A set $C \subseteq \mathbb{R}^n$ of points is a cone if

$$\lambda x + \mu y \in C, \forall x, y \in C, \forall \lambda, \mu \geq 0.$$ 

In other words, a set $C$ is a cone if it is closed under non-negative linear combinations.

**Example 1:** Let $a^1, \ldots, a^m \in \mathbb{R}^n$. Consider

$$C := \left\{ \sum_{i=1}^{m} \lambda_i a^i : \lambda_1, \ldots, \lambda_m \geq 0 \right\}.$$ 

Note that $C$ is closed under non-negative linear combinations and hence, it is a cone.

![Figure 3.2: Example of a cone generated by vectors $a_1, a_2, a_3$ in 2-dimensions](image)

Cones like Example 1 are said to be finitely generated.
**Definition 5.** A cone \( C \subseteq \mathbb{R}^n \) is **finitely generated** if there exist vectors \( a^1, \ldots, a^m \in \mathbb{R}^n \) such that
\[
C = \left\{ \sum_{i=1}^{m} \lambda_i a^i : \lambda_1, \ldots, \lambda_m \geq 0 \right\}.
\]

Denote \( \text{Cone}\{a^1, \ldots, a^m\} := \left\{ \sum_{i=1}^{m} \lambda_i a^i : \lambda_1, \ldots, \lambda_m \geq 0 \right\} \).

Note that by definition, if a point \( X \) is in \( \text{Cone}\{X^1, \ldots, X^m\} \) then \( X \) is a non-negative linear combination of vectors from \( \{X^1, \ldots, X^m\} \). The first important consequence of the fundamental theorem is Caratheodary’s first theorem.

**Theorem 6.** (Caratheodary’s first theorem) Let \( X^1, \ldots, X^m \in \mathbb{R}^n \) and suppose \( X \in \text{Cone}\{X^1, \ldots, X^m\} \). Then \( X \) can be written as a non-negative linear combination of linearly independent vectors from \( \{X^1, \ldots, X^m\} \).

**Proof.** From fundamental theorem. \( \square \)

Here is another example of a cone.

**Example 2:** Let \( A \in \mathbb{R}^{m \times n} \). Consider
\[
C = \{ x \in \mathbb{R} : Ax \leq 0 \}.
\]

Again note that \( C \) is closed under non-negative linear combinations and hence, it is a cone. Cones like Example 2 are said to be polyhedral.

**Definition 7.** A cone \( C \subseteq \mathbb{R}^n \) is a **polyhedral** if there exists \( A \in \mathbb{R}^{m \times n} \) such that
\[
C = \{ x \in \mathbb{R}^n : Ax \leq 0 \}.
\]

**Note:** An ice cream cone is not a polyhedral cone. This is because it does not have an inequality description with finite number of inequalities. It needs infinitely many inequalities.

Our first main result about cones is the equivalence between polyhedral cones and finitely generated cones.

**Theorem 8** (Farkas-Weyl-Minkowski). Let \( C \) be a cone. Then, \( C \) is polyhedral if and only if \( C \) is finitely generated.

**Proof.** \( \iff \): Let \( C = \text{Cone}\{X^1, \ldots, X^m\} \) for \( X^1, \ldots, X^m \in \mathbb{R}^n \). We may assume that \( \text{span}\{X^1, \ldots, X^m\} = \mathbb{R}^n \) (if not, redo the proof in the subspace spanned by \( X^1, \ldots, X^m \)).

By the fundamental theorem of linear inequalities, \( y \in \text{Cone}\{X^1, \ldots, X^m\} \) if and only if for all hyperplanes \( \{ x : c^T x = 0 \} \) containing \( n - 1 \) linearly independent vectors from \( X^1, \ldots, X^m \) with \( c^T X^i \geq 0 \forall i \in [m] \), we have \( c^T y \geq 0 \). \( \text{(I)} \)

Consider all hyperplanes \( \{ x : c^T x = 0 \} \) that are spanned by \( (n - 1) \)-linearly independent vectors from \( X^1, \ldots, X^m \) such that \( c^T X^i \geq 0 \forall i \in [m] \). The number of such hyperplanes is at most \( \binom{m}{n-1} \) and is hence finite. Let them be defined by \( c_1, \ldots, c_d \). We will show that
\[
\text{Cone}\{X^1, \ldots, X^m\} = \{ x : c_i^T x \geq 0 \forall i \in [d] \}.
\]

There are two cases:
1. If \( \bar{x} \in \text{Cone}\{X^1, \ldots, X^m\} \), then (I) implies that \( c_i^T \bar{x} \geq 0 \) \( \forall i \in [d] \) and hence \( \bar{x} \) is in the RHS set.

2. If \( \bar{x} \notin \text{Cone}\{X^1, \ldots, X^m\} \), then (I) implies that there exists \( i \in [d] \) with \( c_i^T \bar{x} < 0 \). Hence, \( \bar{x} \) is not in the RHS set.

Hence if \( C \) is a finitely generated cone, then \( C \) is a polyhedral cone.

\( \Leftarrow \):
Let \( C = \{ x : a_i^T x \leq 0 \) \( \forall i \in [m] \} \). By the previous part, we have

\[
\text{Cone}\{a_1, \ldots, a_m\} = \{ x : b_i^T x \leq 0 \) \( \forall i \in [d] \} \quad \text{for some } b_1, b_2, \ldots, b_d \in \mathbb{R}^n.
\]

We will show that \( C \) is a finitely generated cone that is generated by \( b_1, \ldots, b_d \), i.e.,

\[
C = \text{Cone}\{b_1, \ldots, b_d\}.
\]

To prove this equality we prove \( \text{Cone}\{b_1, \ldots, b_d\} \subseteq C \) and \( C \subseteq \text{Cone}\{b_1, \ldots, b_d\} \).

1. \( \text{Cone}\{b_1, \ldots, b_d\} \subseteq C \): We have \( b_i \in C \) for all \( i \in [d] \) because from Equation (3.2) we have

\[
b_i^T a_1 \leq 0, \ldots, b_i^T a_m \leq 0
\]

which implies that \( a_i^T b_1 \leq 0, \ldots, a_i^T b_d \leq 0 \). If the generators \( b_1, \ldots, b_d \) are in a cone \( C \), then it follows that their non-negative linear combinations are also in \( C \), i.e., \( \text{Cone}\{b_1, \ldots, b_d\} \subseteq C \).

2. \( C \subseteq \text{Cone}\{b_1, \ldots, b_d\} \): For contradiction, let \( \bar{y} \in C \), with \( \bar{y} \notin \text{Cone}\{b_1, \ldots, b_d\} \). By previous part, we have \( \text{Cone}\{b_1, \ldots, b_d\} = \{ y : w_i^T y \leq 0, i = 1, \ldots, r \} \) for some \( w_1, \ldots, w_r \) which implies that there exists \( i \in [r] \) for which \( w_i^T \bar{y} > 0 \). Also, by definition of the \( w_i \)s, we have that \( w_i^T b_j \leq 0 \) for every \( j \in [d] \).

\[
w_i^T b_j \leq 0 \forall j \in [d] \implies w_i \in \text{Cone}\{a_1, \ldots, a_m\} \quad \text{(By Equation 3.2)}
\]

\[
\implies w_i = \sum_{j=1}^m \lambda_j a_j \quad \text{for some } \lambda_j \geq 0 \forall j \in [m]
\]

\[
\implies w_i^T \bar{y} = \sum_{j=1}^m \lambda_j a_j^T \bar{y} \leq 0
\]

Equation (3.3) holds since \( \bar{y} \in C \) and hence \( a_j^T \bar{y} \leq 0 \) \( \forall j \in [m] \) while \( \lambda_j \geq 0 \forall j \in [m] \). This is a contradiction to \( w_i^T \bar{y} > 0 \).

\( \square \)

Theorem \( \square \), i.e., the equivalence between polyhedral cones and finitely generated cones allows us to move between linear inequality description and non-negative linear combination description of a cone.

As mentioned earlier, the fundamental theorem of linear inequalities is a powerful result (the duality theorem for LPs can be shown using the fundamental theorem). We will see one more consequence of the fundamental theorem of linear inequalities. The purpose behind seeing these proofs is to get comfortable with viewing \( Ax = b, x \geq 0 \) as expressing \( b \) as a non-negative linear combination of the columns of \( A \).
Theorem 9 (Caratheodory’s second theorem). If
\[
\max\{c^T x : Ax \leq b\} = \min\{y^T b : y^T A = c^T, y \geq 0\}
\]
holds and both are feasible, then the minimum is attained by a \(y \geq 0\) with positive components corresponding to linearly independent rows of \(A\).

Proof. Let \(x^*\) be an optimal solution to primal maximization and let \(t := c^T x^*\). Then, \(t = \min\{y^T b : y^T A = c^T, y \geq 0\}\) which implies that \([c^T \ t]\) is a non-negative linear combination of rows of \([A \ b]\), i.e.,
\[
\begin{bmatrix}
c \\
t
\end{bmatrix} \in \text{Cone}\left\{\text{columns of } \begin{bmatrix} A^T \\ b^T \end{bmatrix}\right\}.
\]
By Caratheodory’s first theorem (Theorem 6), we have
\[
\begin{bmatrix}
c \\
t
\end{bmatrix} = \begin{bmatrix} A^T \\ b^T \end{bmatrix} y
\]
for a \(y \geq 0\) with non-zero components corresponding to linearly independent columns from \([A^T \ b^T]\).

Let \(y'\) be the positive components of \(y\) and let \([A' \ b']\) be the corresponding rows of \([A \ b]\). By Theorem 6 we have that \([A' \ b']\) has full row rank. We need to show that \(A'\) has full row rank. By complementary slackness conditions we have \(A'x^* = b'\) which implies that \(b'\) is in the column space of \(A'\). Hence, column rank\([A' \ b']\) = column rank\([A']\). We also know that row rank\([A' \ b']\) = column rank\([A' \ b']\) and column rank\([A']\) = row rank\([A']\) which implies that row rank\([A' \ b']\) = row rank\([A']\). Hence, \(A'\) has full row rank. \(\square\)