IE 511: Integer Programming, Spring 2021	**2 Feb, 2021**

Lecture 3: Fundamental theorem of linear inequalities, Cones

Lecturer: Karthik Chandrasekaran *Scribe: Karthik*

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.,

$$IP = \max\{c^T x : Ax = b, x \geq 0, x \in \mathbb{Z}^n\} \ . \tag{3.1}$$

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^*_{LP}. If the solution x^*_{LP} is integral then x^*_{LP} 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 IP-optimal objective value.

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 serves as an important technique in the design of 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 recap 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 δ.

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

Question 0: How to verify if $Ax = b, x \geq 0, x \in \mathbb{Z}^n$ have a feasible solution?

We will first focus on a simpler question:

> Question 1: How to verify if $Ax = b, x \geq 0$ have a feasible solution?

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

> Question 2: How to verify if $Ax = b$ have a feasible solution?

We already know how to answer Question 2 (via 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. *There exists x satisfying $Ax = b$ if and only if rank$(A) =$ rank$[A\ b]$.*

The goal of this lecture is to answer Question 1. It is helpful to reformulate Question 1: equivalently, it is asking whether b can be written as a *non-negative* linear combination of columns of A.

For example, consider Figure 3.1: let a_1, a_2, a_3 be the columns of A. In Figure 3.1(i), we can write b as a non-negative linear combination of a_is but in Figure 3.1(ii), the point b cannot be written as a non-negative linear combination of a_is. This is because we have a hyperplane that separates b from all the a_is.

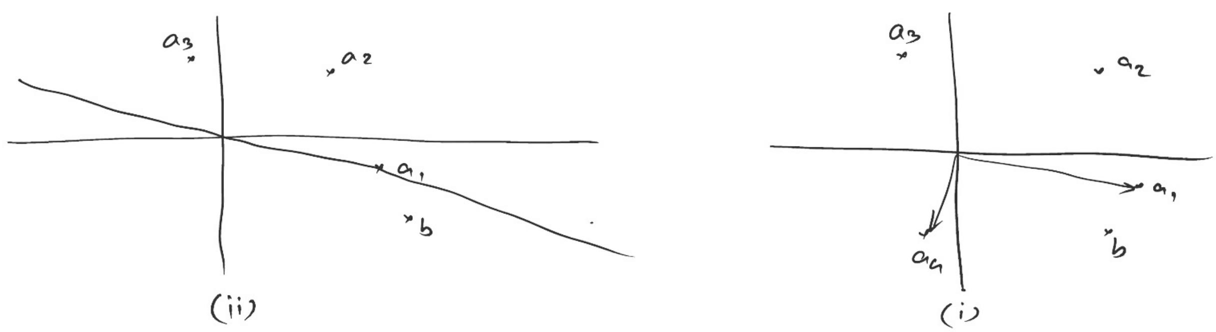

Figure 3.1: Example of feasibility

3.2 Fundamental theorem of linear inequalities

The fundamental theorem of linear inequalities allows us to answer Question 1 by formalizing the two scenarios mentioned in Figure 3.1. It shows that exactly one of the two scenarios can occur.

Theorem 2 (Fundamental theorem of linear inequalities)**.** *Let $a_1, \ldots, a_n, b \in \mathbb{R}^m$, $t =$ 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 the handout posted in compass2g or Schrijver's book titled "Theory of linear and integer programming"). □

The theory of linear programming is built on the fundamental theorem of linear inequalities (i.e., Theorem 2). The fundamental theorem allows us to answer Question 1 using the following lemma:

Lemma 3 (Farkas Lemma)**.** *$Ax = b, x \geq 0$ has a feasible 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 the fundamental theorem to $a_1, \ldots, a_n, b$, we know that exactly one of two scenarios among (i) and (ii) can happen.

- If (i) happens, then it gives a point x satisfying $Ax = b, x \geq 0$.
- If (ii) happens, then it gives a direction/vector c for which $c^T b < 0, c^T a_1 \geq 0, c^T a_2 \geq 0, \ldots, c_n^T \geq 0$. Set $y = c$. Then, we observe that $y^T A \geq 0, y^T b < 0$.

□

Fundamental theorem has several important consequences. These consequences are the tools to understand if the "lucky scenarios" happen for IPs, so we will learn these consequences. These consequences are collectively termed as *Theory of linear inequalities/Polyhedral theory.*

3.3 Polyhedral Theory

Recall that a polyhedron is a set of points satisfying a finite set of linear inequalities. Hence, the term "polyhedral theory" is synonymous with "theory of linear inequalities". The theory of linear inequalities is foundational to the theory of integer linear programming.

3.3.1 Cones

We begin our discussion of polyhedral theory with 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 it is closed under non-negative linear combinations. I.e., C is a *cone* if

$$\lambda x + \mu y \in C \ \forall x, y \in C, \text{ and } \forall \lambda, \mu \geq 0 \ .$$

Example 0: Unbounded ice cream cone with its apex at the origin.

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. See Figure 3.3 for an example.

Figure 3.2: Example of a cone generated by vectors a_1, a_2, a_3 in 2-dimensions

Cones like Example 1 are said to be finitely generated. Formally,

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

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

For a set of points/vectors $a^1, \ldots, a^m \in \mathbb{R}^n$, let $\text{Cone}\{a^1, \ldots, a^m\} := \{\sum_{i=1}^m \lambda_i a^i : \lambda_1, \ldots, \lambda_m \geq 0\}$. We will call $\text{Cone}\{a^1, \ldots, a^m\}$ as *the cone generated by the points* $a^1, \ldots, a^m$.

Remark 6. An unbounded ice cream cone is not a finitely generated cone. This is because it needs infinitely many generators.

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

Proof. By definition of finitely-generated cones and the fundamental theorem. □

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 (polyhedral - since the cone is the set of points satisfying a finite number of linear inequalities).

Definition 8. A cone $C \subseteq \mathbb{R}^n$ is a *polyhedral* if there exists $A \in \mathbb{R}^{m \times n}$ such that $C = \{x \in \mathbb{R}^n : Ax \leq 0\}$.

Remark 9. An unbounded 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.

An important result about cones is the equivalence between polyhedral cones and finitely generated cones that we prove next. This equivalence allows us to move between the linear inequality description and the non-negative linear combination description of a cone.

Theorem 10 (Farkas-Weyl-Minkowski)**.** *Let C be a cone. Then, C is polyhedral if and only if C is finitely generated.*

Proof. We prove the complete characterization by showing both implications.

$\Leftarrow$:

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

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

Consider all hyperplanes $\{x : c^T x = 0\}$ that are spanned by $(n-1)$-linearly independent vectors from $X^1, \dots, X^m$ 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 $c_1, \dots, c_d$. We will show that

$$\text{Cone}\{X^1, \dots, X^m\} = \{x : c_i^T x \geq 0 \ \forall i \in [d]\} \ .$$

This implies that the finitely generated cone C is indeed a polyhedral cone since the RHS is a polyhedral cone. We prove the equation above by considering two cases:

1. If $\bar{x} \in \text{Cone}\{X^1, \dots, X^m\}$, 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 $\bar{x} \notin \text{Cone}\{X^1, \dots, X^m\}$, 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.

$\Rightarrow$:

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

$$\text{Cone}\{a_1, \dots, a_m\} = \{x : b_i^T x \leq 0 \ \forall i \in [d]\} \ \text{ for some } \ b_1, b_2, \dots, b_d \in \mathbb{R}^n \ . \tag{3.2}$$

We will show that C is a finitely generated cone that is generated by $b_1, \dots, b_d$, i.e.,

$$C = \text{Cone}\{b_1, \dots, b_d\} \ .$$

To prove this equality we prove $\text{Cone}\{b_1, \dots, b_d\} \subseteq C$ and $C \subseteq \text{Cone}\{b_1, \dots, b_d\}$.

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

   $$b_i^T a_1 \leq 0, \dots, b_i^T a_m \leq 0$$

   (since for every $j \in [m]$, we have that $a_j \in \text{Cone}\{a_1, \dots, a_m\}$ and hence, by Equation (3.2), $b_i^T a_j \leq 0$ for every $i \in [d]$) which implies that $a_1^T b_i \leq 0, \dots, a_m^T b_i \leq 0$. If the generators $b_1, \dots, b_d$ are in a cone C, then it follows that their non-negative linear combinations are also in C, i.e., $\text{Cone}\{b_1, \dots, b_d\} \subseteq C$.

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 the forward implication of the theorem (that we have proved above), we have that $\text{Cone}\{b_1, \ldots, b_d\} = \{y : w_i^T y \leq 0, i = 1, \ldots, r\}$ for some $w_1, \ldots, w_r$. Consequently, there exists $i \in [r]$ for which $w_i^T \bar{y} > 0$. Also, by definition of the w_is, we have that $w_i^T b_j \leq 0$ for every $j \in [d]$. Hence,

$$\begin{aligned} w_i^T b_j \leq 0 \ \forall j \in [d] &\implies w_i \in \text{Cone}\{a_1, \ldots, a_m\} && \text{(By Equation 3.2)} \\ &\implies w_i = \sum_{j=1}^{m} \lambda_j a_j \text{ for some } \lambda_j \geq 0 \ \forall j \in [m] \\ &\implies w_i^T \bar{y} = \sum_{j=1}^{m} \lambda_j a_j^T \bar{y} \leq 0 && (3.3) \end{aligned}$$

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 $w_i^T \bar{y} > 0$.

□

Significance of Theorem 10. As mentioned earlier, Theorem 10, 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.

The fundamental theorem of linear inequalities is quite powerful: 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 (i.e., moving between the algebraic and the geometric viewpoints).

Theorem 11 (Caratheodary'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 point $y \geq 0$ with positive components corresponding to linearly independent rows of A.

Proof. Let x^* be an optimal solution to the primal maximization problem 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 the 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 Caratheodary's first theorem (Theorem 7), we have that

$$\begin{bmatrix} c \\ t \end{bmatrix} = \begin{bmatrix} A^T \\ b^T \end{bmatrix} y$$

for a $y \geq 0$ with non-zero components corresponding to linearly independent columns from $\begin{bmatrix} A^T \\ b^T \end{bmatrix}$.

Let y' be the positive components of y and let $[A' \ b']$ be the corresponding rows of $[A \ b]$ (corresponding to the positive components of y). By Theorem 7, 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. $\square$