Chapter 2: The Geometry of Linear Programming

March 18, 2022

Scribe: Eric Chan

1 Basic Concepts

A formal definition polyhedron can be described as below.

Definition 1.1. A **polyhedron** is a set that can be described in the form $\{x \in \mathbb{R}^n \mid Ax \geq b\}$, where A is an $m \times n$ matrix and b is a vector in $\mathbb{R}^m$.

Remark. *The definition given above is always referred to as* **general form** *of a polyhedron. Also, we write* **standard form** *of a polyhedron as*

$$\{x \in \mathbb{R}^n \mid Ax = b,\ x > 0\}.$$

The following definitions are essential to LP.

Definition 1.2. Let P be a polyhedron. A vector $x \in P$ is an **extreme point** of P if we cannot find two vectors $y, z \in P$, both different and a scalar $\lambda \in [0, 1]$, such that $x = \lambda y + (1 - \lambda)z$.

Definition 1.3. Let x^* be an element of $\mathbb{R}^n$ and define index set $I(x^*)$ as an **active** set at x^*, in which $I(x^*) = \{i \mid a_i^T x^* = b_i\}$.

Definition 1.4. Consider a polyhedron P defined by linear equality and inequality constraints, and let x^* be an element of $\mathbb{R}^n$.

(a) The vector x^* is a **basic solution** if all equality constraints are active and out of the constraints that are active at x^*, there are n of them that are linearly independent.

(b) If x^* is a basic solution and a feasible solution, we say that it is a **basic feasible solution**.

With definitions above, we can establish an equality between extreme point and basic feasible solution, which is very important in simplex method.

Theorem 1.5. *Let P be a nonempty polyhedron and let $x^* \in P$. Then we can say that x^* is an extreme point is equivalent to x^* is a basic feasible solution.*

Proof.

Let's first assume that $x^* \in P$ is an extreme point.

Suppose that x^* is not a basic feasible solution. Define $I = \{i \mid a_i^T x^* = b_i\}$. Since x^* is not a basic feasible solution, then there doesn't exist n linearly independent vectors a_i, $i \in I$, so a_I cannot span the whole $\mathbb{R}^n$ space. Then there must exist some $d \in a_I^\perp$, such that $a_i^T d = 0$.

Let $\varepsilon > 0$ be very small, consider $y = x^* + \varepsilon d$, $z = x^* - \varepsilon d$. Notice that $a_i^T y = a_i^T z = a_i^T x^* = b_i$ for $i \in I$. Furthermore, for $i \notin I$, we have $a_i^T x^* > b_i$ as x^* is a feasible point. By choosing a proper small ε satisfying $\varepsilon |a_i^T d| < a_i^T x^* - b_i$, we can guarantee that y is also a feasible point of P. The same is true for z. So we finally have $x^* = (y+z)/2$ which implies that x^* is not an extreme point, contradicting with our assumption.

Let's then assume that x^* is a basic feasible solution.

Suppose that x^* is not an extreme point. Then from definition, we can find some y, $z \in P$ and $\lambda \in (0,1)$ such that $x^* = \lambda y + (1 - \lambda z)$. Define $I = \{i \mid a_i^T x^* = b_i\}$. As x^* is a basic feasible solution, we have $A_I x^* = b_I$ and $\mathrm{rank}(A_I) = n$. As the matrix is full rank, we can conclude that x^* is the only solution to the system above. Also, as y and z are feasible solutions, $a_i^T y \geq b_i$, $a_i^T z \geq b_i$, then with

$$b_i = a_i^T x^* = a_i^T (\lambda y + (1-\lambda) z) \geq \lambda b_i + (1-\lambda) b_i = b_i,$$

we can say $a_i^T y = a_i^T z = b_i$ for $i \in I$. Thus y and z are also solutions to the linear system $A_I x = b_I$, which makes a contradiction.

We have thus completed the proof. □

Remark. *Notice that the theorem above is very important since it has established the equivalence between two definitions, which may be very useful to many proofs presented later. Also, notice that the second part of proof above is very inspiring, using an inequality to get an equality.*

2 Construct Basic Solutions

The theorem presented below provides us with a way to construct basic solutions.

Theorem 2.1. *Consider a polyhedron P in standard form and $A \in \mathbb{R}^{m \times n}$ has full row rank. Then $x \in \mathbb{R}^n$ is a basic solution iff we have $Ax = b$, and there exists indices $B(1), B(2), \ldots, B(m)$ such that:*

(a) The columns $A_{B(1)}, A_{B(2)}, \ldots, A_{B(m)}$ are linearly independent;

(b) If $i \neq B(1), B(2), \ldots, B(m)$, then $x_i = 0$.

Proof. Assume x satisfies (a) and (b). Then we can partition A as $[B, N]$, where $B = \{A_B\}$, and N otherwise. So we can write the whole system as

$$M = \begin{bmatrix} B & N \\ 0 & I \end{bmatrix},$$

it's obvious that $\mathrm{rank}(M) = n$, thus from the definition, it is a basic solution.

Let's then assume that x is a basic solution. Define index set $I = \{i \mid x_i \neq 0\} = \{B(1), \ldots, B(k)\}$. Thus we can conclude the system

$$\begin{cases} Ax = b \\ x_i = 0, \quad i \notin I \end{cases}$$

has unique solution, equivalently, the equation

$$\sum_{i=1}^{k} A_{B(i)} x_{B(i)} = b$$

has a unique solution. It follows that $A_{B(1)}, \ldots, A_{B(k)}$ are linearly independent, which implies $k \leq m$. By the base extension theorem, we can find $m - k$ additional columns $A_{B(k+1)}, \ldots, A_{B(m)}$ so that the column set spans space $\mathbb{R}^m$. In addition, if $i \neq B(1), \ldots, B(m)$, then $i \neq B(1), \ldots, B(k)$ and $x_i = 0$. Therefore, both conditions (a) and (b) are satisfied. □

The theorem above provides us with a way to construct basic solutions, we generalize it to following steps.

1. Choose m linearly independent columns $A_{B(1)}, \ldots, A_{B(m)}$.
2. Let $x_i = 0$ for all $i \neq B(1), \ldots, B(m)$.
3. Solve the system of m equations $Ax = b$ for the unknowns $x_{B(1)}, \ldots, x_{B(m)}$.

3 Existence of Extreme Points

Definition 3.1. A polyhedron $P \subset \mathbb{R}^n$ **contains a line** if there exists a vector $x \in P$ and a nonzero vector $d \in \mathbb{R}^n$ such that $x + \lambda d \in P$ for all scalars λ.

Theorem 3.2. *Suppose that polyhedron $P = \{x \in \mathbb{R}^n \mid a_i^T x \geq b_i,\ i = 1, \ldots, m\}$ is nonempty. Then the following are equivalent:*

(a) The polyhedron P has at least one extreme point.

(b) The polyhedron P does not contain a line.

(c) There exist n vectors out of the family $a_1, \ldots, a_m$ which are linearly independent.

Proof. **(a)⇒(c)**: If P has an extreme point x, then x is also a basic feasible solution, then there exists n constraints that are active at x, with corresponding a_i being linearly independent.
(c)⇒(b): Assume that n vectors $a_1, a_2, \ldots, a_n$ are linearly independent. If P contains a line $x + \lambda d$, where d is a nonzero vector. We must have $a_i^T(x + \lambda d) \geq b_i$ for all i and all λ. As λ is an arbitrary number, we must have $a_i^T d = 0$. With the condition that a_i are linearly independent, we only have

$d = 0$, which makes a contradiction.

(b)⇒(a): Let $x \in P$ and $I = \{i \mid a_i^T x = b_i\}$. For trivial case, if n of the vectors $a_i, i \in I$ corresponding to the active constraints are linearly independent, then x is a basic feasible solution and thus an extreme point. If this is not the case, then all of the vectors $a_i, i \in I$ lie in a subspace of $\mathbb{R}^n$ and then there exists a nonzero vector $d \in \mathbb{R}^n$ such that $a_i^T d = 0,\ \ \forall i \in I$. Consider points on the line $y = x + \lambda d$. For $i \in I$, we have $a_i^T y = a_i^T x + \lambda a_i^T d = a_i^T x = b_i$. This means that these constraints remain active along the whole line. However, as the polyhedron contains no line, some constraints must be violated with λ varying. At the point where some constraint is about to be violated, a new constraint must become active, and we conclude that there exits some $\tilde{\lambda}$ and some $j \notin I$ such that $a_j^T(x + \tilde{\lambda} d) = b_j$.

The next thing we have to do is to argue a_j is not a linear combination of vectors $a_i, i \in I$. Indeed, with $a_j^T x \neq b_j$ and $a_j^T(x + \tilde{\lambda} d) = b_j$, we have $a_j^T d \neq 0$. If a_j is a linear combination of a_I, then we must have $a_j^T d = 0$, which leads a contradiction. Thus by moving from x to $x + \tilde{\lambda} d$, the number of linearly independent active constraints has been increased by at least one. By repeating this procedure, we eventually end up with a basic solution in this feasible, that is to say, a basic feasible solution. □

Remark. *It's worth noticing that this theorem points out the existence of extreme points, and the third part of the proof is valuable, since it shows us a nontrivial way to find linearly independent vectors step by step.*

A direct lemma following the theorem above is shown below.

Lemma 3.3. *Every nonempty bounded polyhedron and every nonempty polyhedron in standard form has at least one basic feasible solution.*

Remark. *Although we have shown before that any linear programming in general form can be equivalently transformed into one in standard form, this lemma still does not mean every polyhedron has an extreme point. As the equivalence of two LP problems only means the equivalence of optimal solutions, not for the equivalence of feasible polyhedron.*

4 Optimality of Extreme Points

Theorem 4.1. *Consider the linear programming problem of minimizing $c^T x$ over a polyhedron P. Suppose that P* **has at least one extreme point** *and that* **there exists an optimal solution***. Then there exists an optimal solution which is an extreme point of P.*

Proof. Denote Q as the set of all optimal solutions, which is nonempty according to assumption. Let P be of the form $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ and v_{opt} be the optimal value. Then we can write Q as

$$Q = \{x \in \mathbb{R}^n \mid Ax \geq b,\ c^T x = v_{opt}\},$$

so Q is also a polyhedron. As $Q \subset P$ and P contains no line, then Q contains no line, so Q has at least one extreme point.

Let x^* be an extreme point of Q. Suppose that x^* is not an extreme point of P, then we can find $y \in P$, $z \in P$ both different from x^* and $\lambda \in (0,1)$ such that $x^* = \lambda y + (1-\lambda)z$. Then

$$v_{opt} = c^T x^* = \lambda c^T y + (1-\lambda) c^T z \geq \lambda v_{opt} + (1-\lambda) v_{opt} = v_{opt}.$$

So we must have $c^T y = c^T z = v_{opt}$, and therefore $y, z \in Q$, which contradicts with the assumption that x^* is an extreme point of Q. So we know that x^* is an extreme point of P, thus complete the proof. □

Remark. *The way we prove the theorem above is worth noticing, in which we first assume a set and from that to finish our proof. However, the assumption that* **existence of optimal solution** *is always not that easy, so we have a much stronger theorem below.*

Theorem 4.2. *Consider the linear programming problem of minimizing $c^T x$ over a polyhedron P. Suppose that P has at least one extreme point. Then, either the optimal cost is equal to $-\infty$ or there exists an extreme point which is optimal.*

Proof. We first define the terminology: an element x of P has $\operatorname{rank} k$ if we can find k but no more than k linearly independent constraints that are active at x.

First assume that the optimal cost is finite. Let $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ and consider some $x \in P$ of $\operatorname{rank} k \leq n$. Otherwise, x would be an extreme point and the proof is finished. Let $I = \{i \mid a_i^T x = b_i\}$. Since $k < n$, we can find some nonzero $d \in \mathbb{R}^n$ such that d is orthogonal to every $a_i, i \in I$. By shifting operators, we can always assume that $c^T d \leq 0$.

Suppose $c^T d < 0$. Consider half line $y = x + \lambda d$ where $\lambda > 0$. Then for $i \in I$, we have

$$a_i^T y = a_i^T x + \lambda a_i^T d = a_i^T x = b_i.$$

If the entire half-line is contained in P, then the optimal value would be $-\infty$, which we have assumed not to be the case. Therefore, the half-line must have some intersection with P, where it is about to exit. Then we have some $\tilde{\lambda} > 0$ and $j \notin I$ such that $a_j^T(x + \tilde{\lambda} d) = b_j$. Denote $y = x + \tilde{\lambda} d$ and note $c^T y \leq c^T x$. Besides, a_j is linearly independent from $a_i, i \in I$, so $\operatorname{rank} y$ is at least $k+1$.

Suppose $c^T d = 0$. Consider for arbitrary λ, $y = x + \lambda d$. As P contains no lines, we have at the boundary y such that $c^T y = c^T x$.

In either case, we have found a new point y such that $c^T y \leq c^T x$ and its rank are greater. Repeating this procedure, we finally end up with a vector w of rank n such that $c^T w \leq c^T x$.

Let $w^1, \ldots, w^r$ be the basic feasible solutions in P and let w^* be a basic feasible solution such that $c^T w^* \leq c^T w^i$ for all i. We have already shown that for every x there exists some i such that $c^T w^i \leq c^T x$. It follows that $c^T w^* \leq c^T x$ for all $x \in P$, then w^* is optimal. □

Remark. *The proof above is essentially a repetition of existence of extreme points. It's also worth noticing that the reason this proof makes sense is the lemma that number of extreme points is finite, otherwise we may not find a minimal value in the last part.*

Lemma 4.3. *Consider the linear programming problem of minimizing $c^T x$ over a nonempty polyhedron. Then either the optimal cost is equal to $-\infty$ or there exists an optimal solution.*