**ORIE 6300 Mathematical Programming I** September 4, 2014

Lecture 4

*Lecturer: David P. Williamson* *Scribe: Paul Upchurch*

# 1 Introduction

Last time we talked about polyhedra and polytopes. This time we will define bounded polyhedra and discuss their relationship with polytopes. Recall from the last lecture the following definitions.

A polyhedron is $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, $A \in \mathbb{R}^{m \times n}$, $m \geq n$.
A polytope is $Q = \text{conv}(v_1, \ \ldots, \ v_k)$ for finite $k$.
$x \in P$ is a vertex if $\exists c \in \mathbb{R}^n$ such that $c^T x < c^T y$ for all $y \in P$, $y \neq x$.
$x \in P$ is an extreme point if $\nexists y, z \in P \ \ y, z \neq x$ such that $x = \lambda y + (1-\lambda)z, \lambda \in [0,1]$.
$x \in P$ is a basic feasible solution if $x \in P$ and it is basic (i.e., the rank of $A_=$ is $n$).

Notice that the number of vertices of $P$ is finite since given the $m$ constraints in $Ax \leq b$, we can choose $n$ of them to be met with equality; thus there are at most $\binom{m}{n}$ basic solutions.

# 2 Polyhedra and Polytopes

Now we are interested in the following two questions:

- Q1: When is a polytope a polyhedron?
- A1: A polytope is always a polyhedron.

- Q2: When is a polyhedron a polytope?
- A2: A polyhedron is almost always a polytope.

We can give a counterexample to show why a polyhedron is not always but almost always a polytope: an unbounded polyhedra is not a polytope. See Figure 1.

**Definition 1** *A polyhedron $P$ is* bounded *if $\exists M > 0$, such that $\|x\| \leq M$ for all $x \in P$.*

What we can show is this: every bounded polyhedron is a polytope, and vice versa. In this lecture, we will show one side of the proof in one direction; we will show the other direction in the next lecture. To start with, we need the following lemma.

**Lemma 1** *Any polyhedron $P = \{x \in \Re^n : Ax \leq b\}$ is convex.*

**Proof:** If $x, y \in P$, then $Ax \leq b$ and $Ay \leq b$. Therefore,

$$A(\lambda x + (1-\lambda)y) = \lambda Ax + (1-\lambda)Ay \leq \lambda b + (1-\lambda)b = b.$$

Thus $x + (1-\lambda)y \in P$. □

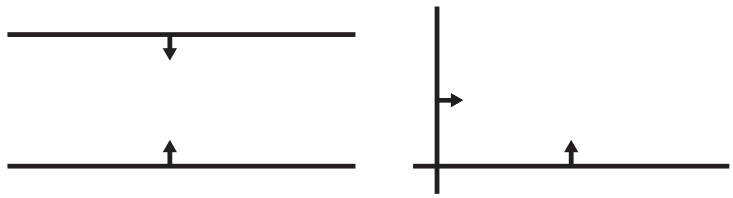

Figure 1: Examples of unbounded polyhedra that are not polytopes. (left) No extreme points, (right) one extreme point.

## 3 Representation of Bounded Polyhedra

We can now show the following theorem.

**Theorem 2** *(**Representation of Bounded Polyhedra**) A bounded polyhedron $P$ is the set of all convex combinations of its vertices, and is therefore a polytope.*

**Proof:** Let $v_1, v_2, \ldots, v_k$ be the vertices of $P$. Since $v_i \in P$ and $P$ is convex (by previous lemma), then any convex combination $\sum_{i=1}^k \lambda_i v_i \in P$. So it only remains to show that any $x \in P$ can be written as $x = \sum_{i=1}^k \lambda_i v_i$, with $\lambda_i \geq 0$ and $\sum_{i=1}^k \lambda_i = 1$.

Let $A_=$ be all the constraints that $x$ meets with equality (all rows $a_i$ such that $a_i x = b_i$). Let $ra(x)$ be the rank of the corresponding $A_=$. Recall from last time that $ra(x) = n$ if and only if $x$ is a vertex of $P$. Now we prove the theorem by induction on $n - ra(x)$.

*Base case:* Let $n - ra(x) = 0$. Then $ra(x) = n$ and since $x \in P$, $x$ is a basic feasible solution, and therefore a vertex of $P$.

*Inductive Step:* Suppose we have shown that for any $y \in P$ such that $n - ra(y) < \ell$ for some $\ell > 0$, $y$ can be written as a convex combination of $v_1, v_2, \ldots, v_k$. Consider $x \in P$ with $ra(x) = n - \ell < n$. Then the rank of $A_= < n$, and thus there exists $z$ such that $A_= z = 0$. Since $P$ is bounded, there exist constants $\overline{\alpha} > 0$ and $\underline{\alpha} < 0$ such that $x + \alpha z \in P$ if and only if $\underline{\alpha} \leq \alpha \leq \overline{\alpha}$. Geometrically, this is equivalent to moving from $x$ in the direction $\alpha z$ until we run into a constraint.

Then we can express $x$ as

$$x = \frac{\overline{\alpha}}{\overline{\alpha} - \underline{\alpha}}(x + \underline{\alpha} z) + \frac{-\underline{\alpha}}{\overline{\alpha} - \underline{\alpha}}(x + \overline{\alpha} z).$$

Therefore, $x$ is a convex combinations of two points in $P$. Now all we need to show is that $x + \underline{\alpha} z$ and $x + \overline{\alpha} z$ are convex combinations of vertices. Since $x + \overline{\alpha} z \in P$, but $x + \alpha z \notin P$ for $\alpha > \overline{\alpha}$, there exists some constraint $a_j$ such that $a_j x < b_j$, but $a_j(x + \overline{\alpha} z) = b_j$. This implies that $ra(x + \overline{\alpha} z) > ra(x)$, so then $n - ra(x + \overline{\alpha} z) < n - ra(x) = \ell$. Therefore, $x + \overline{\alpha} z$ can be expressed as a convex combination of vertices $v_1, v_2, \ldots, v_k$ by induction; we suppose $x + \overline{\alpha} z = \sum_{i=1}^k \alpha_i v_i$, where $\alpha_i \geq 0$ and $\sum_{i=1}^k \alpha_i = 1$. Similarly, it must be the case that $x + \underline{\alpha} z$ is a convex combination of the vertices, and we can write $x + \underline{\alpha} z = \sum_{i=1}^k \beta_i v_i$, where $\beta_i \geq 0$ and $\sum_{i=1}^k \beta_i = 1$.

Therefore, we have

$$\begin{aligned} x &= \frac{\overline{\alpha}}{\overline{\alpha}-\underline{\alpha}}(x+\underline{\alpha}z)+\frac{-\underline{\alpha}}{\overline{\alpha}-\underline{\alpha}}(x+\overline{\alpha}z) \\ &= \frac{\overline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\sum_{i=1}^{k}\alpha_i v_i+\frac{-\underline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\sum_{i=1}^{k}\beta_i v_i \\ &= \sum_{i=1}^{k}\left(\frac{\overline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\alpha_i+\frac{-\underline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\beta_i\right)v_i \\ &= \sum_{i=1}^{k}\delta_i v_i, \end{aligned}$$

where $\delta_i = \frac{\overline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\alpha_i + \frac{-\underline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\beta_i \geq 0$ and

$$\begin{aligned} \sum_{i=1}^{k}\delta_i &= \sum_{i=1}^{k}\left(\frac{\overline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\alpha_i+\frac{-\underline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\beta_i\right) \\ &= \frac{\overline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\sum_{i=1}^{k}\alpha_i+\frac{-\underline{\alpha}}{\overline{\alpha}-\underline{\alpha}}\sum_{i=1}^{k}\beta_i \\ &= \frac{\overline{\alpha}}{\overline{\alpha}-\underline{\alpha}}+\frac{-\underline{\alpha}}{\overline{\alpha}-\underline{\alpha}}=1. \end{aligned}$$

Thus $x$ is a convex combination of the vertices. □

## 4 Separating Hyperplane Theorem

To begin showing the proof in the opposite direction (that is, showing that every polytope is a bounded polyhedron), we will need a theorem called the *separating hyperplane theorem*. To prove the theorem, we will use the following theorem from analysis, which we give without proof.

**Theorem 3** *(**Weierstrass**) Let $C \subseteq \Re^n$ be a closed, non-empty and bounded set. Let $f : C \to \Re$ be continuous on $C$. Then $f$ attains a maximum (and a minimum) on some point of $C$.*

Suppose $f(x) = \frac{1}{2}\|x-y\|$, for all $x \in C$. We'd like to apply Weierstrass' theorem to find the minimizer of $f$ in $C$, but $C$ may not be bounded. To get around this, we pick some $q \in C$, which we can do since $C$ is non-empty. Then, let $\hat{C} = \{x \in C : \|q-y\| \geq \|x-y\|\}$. $\hat{C}$ is closed, non-empty and bounded; we see that $\hat{C}$ is bounded since for $x \in \hat{C}$, we have $\|x\| \leq \|y\| + \|y-x\|$ by the triangle inequality and $\|y\|+\|y-x\| \leq \|y\|+\|q-y\|$ by the definition of $\hat{C}$; both $\|y\|$ and $\|q-y\|$ are constant terms. Now we can apply Weierstrass' theorem on $\hat{C}$ to find a point $z$ that minimizes $f$.

**Theorem 4** *(**Separating Hyperplane**) Let $C \subseteq \Re^n$ be closed, non-empty and convex set. Let $y \notin C$, then there exists a hyperplane $a \neq 0$, $a \in \Re^n$, $b \in \Re$, such that $a^T y > b$ and $a^T x < b$, for all $x \in C$.*

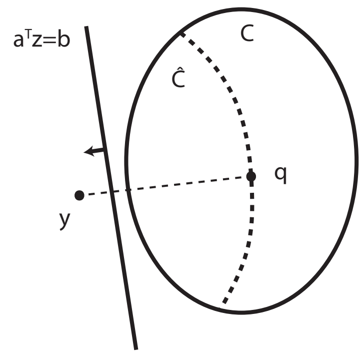


Figure 2: Separating hyperplane

**Proof:** Define

$$f(x) = \tfrac{1}{2}||x - y||^2$$

$$\hat{C} = \{x \in C : ||q - y|| \geq ||q - x||\}.$$

Apply Weierstrass' theorem. Let $z$ be the minimizer of $f$ in $\hat{C}$. Note that for any $x \in C - \hat{C}$, $f(z) \leq f(q) < f(x)$, and therefore $z$ minimizes $f$ over all of $C$, since any $x \notin \hat{C}$ must have been further away from $y$ than $q$.

Let $a = y - z$. Then $a \neq 0$, since $z \in C, y \notin C$. Let $b = \frac{1}{2}(a^T y + a^T z)$. Then,

$$0 < a^T a = a^T(y - z) = a^T y - a^T z$$

so then

$$a^T y > a^T z \quad \Rightarrow \quad 2a^T y > a^T y + a^T z \quad \Rightarrow \quad a^T y > \tfrac{1}{2}(a^T y + a^T z) = b.$$

It remains to show that $a^T x < b$ for all $x \in C$. Let $x_\lambda = (1 - \lambda)z + \lambda x \in C$ for $0 < \lambda \leq 1$. Since $z$ minimizes $f$ over $C$, $f(z) \leq f(x_\lambda)$. Thus,

$$\begin{aligned} f(x_\lambda) = \tfrac{1}{2}((1-\lambda)z + \lambda x - y)^T((1-\lambda)z + \lambda x - y) &= \tfrac{1}{2}(z - y + \lambda(x - z))^T(z - y + \lambda(x - z)) \\ &\geq \tfrac{1}{2}(z - y)^T(z - y) = f(z). \end{aligned}$$

Rewriting, we obtain

$$\begin{aligned} \tfrac{1}{2}[2(z - y)^T \lambda(x - z) + \lambda^2(x - z)^T(x - z)] &\geq 0 \\ (z - y)^T(x - z) + \tfrac{1}{2}\lambda(x - z)^T(x - z) &\geq 0 \\ a^T(z - x) + \tfrac{1}{2}\lambda(x - z)^T(x - z) &\geq 0 \end{aligned}$$

or

$$a^T(z - x) \geq -\tfrac{1}{2}\lambda(x - z)^T(x - z).$$

But we can take $\lambda \to 0$ arbitrarily small, so $a^T(z - x) \geq 0$ which implies $a^T z \geq a^T x$. Using the fact that $a^T z < a^T y$,

$$b = \tfrac{1}{2}(a^T y + a^T z) \geq \tfrac{1}{2}(2a^T z) = a^T z > a^T x.$$

□