ORIE 6300 Mathematical Programming I	September 11, 2008

Lecture 5

Lecturer: David P. Williamson — *Scribe: Emma Qiu Wang*

Last time we talked about polyhedra and their characteristics. This time we will define polytopes and discuss about their relationship with polyhedra.

First we need some definitions.

Definition 1 *A set $S \subseteq \Re^n$ is* convex *if $\forall x, y \in S$, $\lambda x + (1-\lambda)y \in S$, $\forall \lambda \in [0,1]$.*

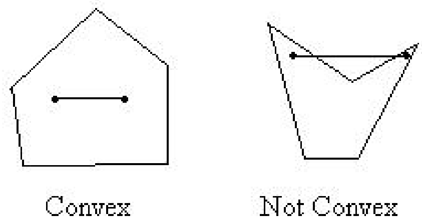

Figure 1: Examples of convex sets

Definition 2 *Given $v_1, v_2, \ldots, v_k \in \Re^n$, a* convex combination *of $v_1, v_2, \ldots, v_k$ is $v = \sum_{i=1}^k \lambda_i v_i$ for some λ_i such that $\lambda_i \geq 0$ and $\sum_{i=1}^k \lambda_i = 1$.*

Given $v_1, v_2, \ldots, v_k \in \Re^n$, let's set

$$Q = \{v \in \Re^n : v \text{ is a convex combination of } v_1, v_2, \ldots, v_k\}.$$

Lemma 1 *Q is convex.*

Proof: Pick any $x, y \in Q$. This implies that

$$\begin{array}{lll} x = \sum_{i=1}^k \alpha_i v_i & \alpha_i \geq 0 & \sum_{i=1}^k \alpha_i = 1, \\ y = \sum_{i=1}^k \beta_i v_i & \beta i \geq 0 & \sum_{i=1}^k \beta i = 1. \end{array}$$

For $\lambda \in [0,1]$, then

$$\begin{aligned} \lambda x + (1-\lambda)y &= \lambda \sum_{i=1}^k \alpha_i v_i + (1-\lambda) \sum_{i=1}^k \beta_i v_i \\ &= \sum_{i=1}^k [\lambda \alpha_i + (1-\lambda)\beta_i] v_i. \end{aligned}$$

Then we know that $\lambda\alpha_i + (1-\lambda)\beta_i \geq 0$ for all i, and that

$$\sum_{i=1}^{k}(\lambda\alpha_i + (1-\lambda)\beta_i) = \lambda\sum_{i=1}^{k}\alpha_i + (1-\lambda)\sum_{i=1}^{k}\beta_i = 1.$$

Thus

$$\lambda x + (1-\lambda)y = \sum_{i=1}^{k}\delta_i v_i,$$

where $\delta_i = \lambda\alpha_i + (1-\lambda)\beta_i$, so that $\delta_i \geq 0$ for all i, and $\sum_{i=1}^{k}\delta_i = 1$. Thus $\lambda x + (1-\lambda)y \in Q$. $\square$

For $Q = \{v \in \Re^n : v$ is a convex combination of $v_1, v_2, \ldots, v_k\}$, we say that Q is a *convex hull* of $v_1, v_2, \ldots, v_k$, and we write $Q = conv(v_1, v_2, \ldots, v_k)$.

Definition 3 *For Q the convex hull of a finite number of vectors $v_1, v_2, \ldots, v_k$, Q is a polytope.*

Observation 1 *Any extreme point of a polytope $Q = conv(v_1, v_2, \ldots, v_k)$ is v_j for some $j = 1, 2, \ldots, k$.*

Proof: Suppose we have an arbitrary point v in the polytope, and $v \neq v_1, v_2, \ldots, v_k$. Then by definition, we know that

$$\begin{aligned} v &= \sum_{i=1}^{k}\lambda_i v_i \\ &= \lambda_1 v_1 + \sum_{i=2}^{k}\lambda_i v_i \\ &= \lambda_1 v_1 + (1-\lambda_1)\sum_{i=2}^{k}\frac{\lambda_i}{1-\lambda_1}v_i \end{aligned}$$

Let $w = \sum_{i=2}^{k}\frac{\lambda_i}{1-\lambda_1}v_i$. It is easy to check that w is a convex combination of $v_2, \ldots, v_k$. Thus we can express $v = \lambda_1 v_1 + (1-\lambda_1)w$, for $v_1, w \in Q$, and thus v is is not an extreme point. $\square$

Now we are interested in the following two questions:

- Q1: When is a polytope a polyhedron?
- A1: A polytope is always a polyhedron (We will prove this in later lectures).

- 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.

Definition 4 *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 of the proof in one direction, and we'll show the other direction in the next lecture. To start with, we need the following lemma.

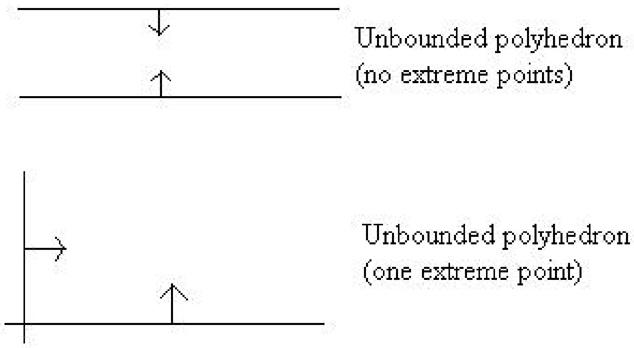

Figure 2: Examples of unbounded polyhedra

Lemma 2 *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 $x + (1-\lambda)y \in P$. $\square$

We can now show the following theorem.

Theorem 3 ***(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. (Question for the reader: Why do we have a finite number of vertices?). 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_j s.t. $a_j x = b_j$). 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 through 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 αz until we run into a constraint.

Then we can express x as

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

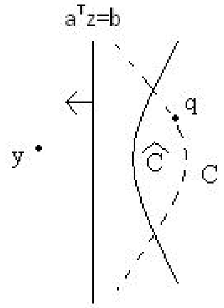

Figure 3: Separating hyperplane

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. The same thing applies to $x+\underline{\alpha}z$. Therefore x must be a convex combination of $v_1, v_2, \ldots, v_k$. $\square$

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 4 ***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.*

Theorem 5 ***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 $x \in C$.*

We first discuss the brief idea of the proof and leave the actual proof for next lecture.

Let $f(x) = \frac{1}{2} \parallel x-y \parallel$, for all $x \in C$. We'd like to apply the 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. Consider $\hat{C} = \{x \in C : \parallel q-y \parallel \geq \parallel x-y\}$. Then $\hat{C}$ is closed, non-empty and bounded; we see that $\hat{C}$ is bounded since for $x \in \hat{C}$, we have $\parallel x \parallel \leq \parallel y \parallel + \parallel y-x \parallel$ by triangle inequality and $\parallel y \parallel + \parallel y-x \parallel \leq \parallel y \parallel + \parallel q-y \parallel$ by the definition of $\hat{C}$. Then we can apply Weierstrass theorem on $\hat{C}$ to find a point z that minimizes f. We'll define the hyperplane as $a = y-z$, $b = (a^T y + a^T z)/2$. The proof that this hyperplane has the properties that we want will wait for next time.