Mathematical Programming	**Lecture 4**
OR 630 Fall 2005	**September 6, 2005**
Instructor: Mike Todd	**Scribe: Chris Provan**

The material we cover here may be found in Chapter 2 of Bertsimas and Tsitsiklis or, for a more condensed reading, in Chapter 7 of Schrijver (available online at http://encompass.library.cornell.edu/cgi-bin/scripts/ebooks.cgi?bookid=17885).

We closed the last lecture with a brief discussion of the relationship between polytopes and polyhedra.

Recall that a ***polytope*** is the set of all convex combinations of a finite set of points $v_1, v_2, \ldots, v_k$. We can also think of a polytope with k generating points as a linear transformation of a $(k-1)$-*dimensional simplex*. A ***polyhedron*** is the intersection of a finite number of half-spaces, and if bounded, can also be thought of as an inverse linear image of a *simplex*.

This lecture will focus on making the relationship between polytopes and polyhedra more explicit.

Definition 1 *Let Q be a convex set in $\mathbb{R}^n$. Then $x \in Q$ is an* **extreme point** *of Q if x cannot be written as $(1-\lambda)y + \lambda z$ for $y, z \in Q$, $y \neq z$, $0 < \lambda < 1$. $x \in Q$ is a* **vertex** *of Q if $\exists f \in \mathbb{R}^n$ with argmax$\{f^T z : z \in Q\} = \{x\}$ (i.e. x is the* unique *optimal solution for some objective coefficient vector f).*

It is interesting to note that because these definitions are generalized for all convex sets - not just polyhedra - a point could possibly be extreme but not be a vertex. One set of examples are the points on an oval where the line segments of the sides meet the curves of the ends.

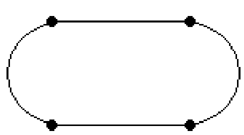

Figure 1: *Four extreme points in a two-dimensional convex set that are not vertices.*

For the purposes of the subsequent theorem and definitions, let us define the following notation for a polyhedron:

$$\begin{aligned} Q^* \; &:= \{y \in \mathbb{R}^m : \; A_x^T y \leq c_x; A_w^T y = c_w\} \\ &=: \{y \in \mathbb{R}^m : \; a_j^T y \leq c_j, j \in N_x; a_j^T y = c_j, j \in N_w\} \end{aligned}$$

Definition 2 $\mathbf{I(y)} := \{j \in N_x \cup N_w : a_j^T y = c_j\}$ *and* $\mathbf{ra(y)} := rank(\{a_j : j \in I(y)\})$

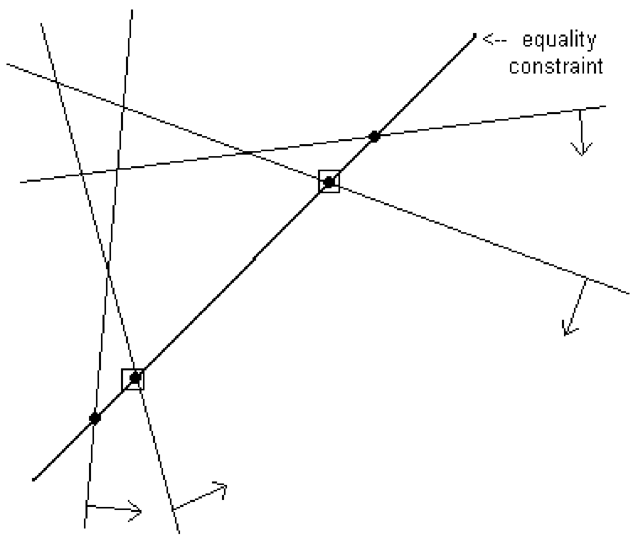

Figure 2: *A geometric representation of the four basic solutions (dots) and two basic feasible solutions (boxed) of an LP problem with one equality and four inequality constraints on two variables.*

Definition 3 *Call $y \in \mathbf{R}^m$ a* **basic solution** *of Q^* if $N_w \subseteq I(y)$ and $ra(y) = m$. y is a* **basic feasible solution** *of Q^* if it also lies inside Q^*. (See Figure 2).*

Since there are only a finite number of constraints defining Q^*, there are only a finite number of ways to choose $I(y)$, and if $ra(y) = m$ then y is uniquely determined by $I(y)$. So there are at most $\binom{|N_x \cup N_w|}{m}$ basic solutions.

Theorem 1 *(**Characterization of Vertices**). Let Q^* be defined as above. The following are equivalent:*
(a) y is a vertex of Q^*.
(b) y is an extreme point of Q^*.
(c) y is a basic feasible solution of Q^*.

Proof: We first prove that (a) $\Rightarrow$ (b). Let y be a vertex of Q^* and suppose by way of contraposition that y is not an extreme point of Q^*. So $\exists s, t \in Q^*$, $s \neq t$, $0 < \lambda < 1$ such that $y = (1-\lambda)s + \lambda t$. Since y is a vertex, $\exists f \in \mathbf{R}^m$ such that

$$\begin{aligned} \operatorname{argmax}\{f^T z : z \in Q\} = \{y\} \quad &\Rightarrow f^T s < f^T y, f^T t < f^T y \\ &\Rightarrow (1-\lambda) f^T s + \lambda t < f^T y. \end{aligned}$$

But we also have

$$\begin{aligned} f^T y \quad &= f^T[(1-\lambda)s + \lambda t] \\ &= (1-\lambda) f^T s + \lambda f^T t, \end{aligned}$$

which is a contradiction. So y must be an extreme point.

We now prove (b) $\Rightarrow$ (c). Let y be an extreme point of Q^* but suppose by way of contraposition that y is not a basic feasible solution. For y to be an extreme point of Q^*, it must lie in Q^* and therefore $N_w \subseteq I(y)$. The only possible way y could not be a basic feasible solution is for $ra(y) < m$, and hence there is a direction vector $0 \neq d \in \mathbf{R}^m$ with $a_j^T d = 0$ for all $j \in I(y)$. So

$$a_j^T(y + \alpha d) = \{ \begin{array}{rl} a_j^T y = c_j, & if j \in I(y); \\ a_j^T y + \alpha a_j^T d < c_j + \alpha a_j^T d, & otherwise. \end{array}$$

So for some $\varepsilon > 0, \forall |\alpha| \leq \varepsilon, y + \alpha d \in Q^*$. So y can be written as a convex combination of two other points of Q^*, namely $y = \frac{y+\varepsilon d}{2} + \frac{y-\varepsilon d}{2}$, which contradicts y being an extreme point. Therefore y must also be a basic feasible solution of Q^*.

Finally, we prove (c) $\Rightarrow$ (a). Let y be a basic feasible solution of Q^*, which implies $ra(y) = m$. By definition $a_j^T y = c_j$ for $j \in I(y)$, and therefore $(\sum_{j \in I(y)} a_j)^T y = \sum_{j \in I(y)} c_j$. Because $ra(y) = m$, y is the unique point in Q^* that satisfies this equation. Let $f = \sum_{j \in I(y)} a_j$. $\forall z \in Q^*, j \in I(y) : a_j^T z \leq c_j$, which implies $f^T z \leq \sum_{j \in I(y)} c_j$. But as we have shown, $f^T y$ satisfies this constraint with equality, and therefore y maximizes $f^T z$ over Q^*. And because y uniquely satisfies with equality, it uniquely maximizes $f^T z$, and so y is by definition a vertex. $\square$

The following two corollaries are given without proof.

Corollary 1 *Any polyhedron Q has only a finite number of vertices. Specifically, if Q is defined by n constraints on m variables then it has a maximum of $\binom{n}{m}$ vertices.*

Corollary 2 *The set of basic feasible solutions does not depend on the representation of Q.*

We now prove the claim delivered without proof at the end of the last lecture – that any bounded polyhedron is a polytope. We must first give a concrete definition to the notion of boundedness.

Definition 4 *A convex set $Q \subseteq \mathbf{R}^n$ is* **bounded** *if for some $M > 0$, $\|y\| \leq M$ for all $y \in Q$. A convex set is* **pointed** *if it does not contain any line $\{y + \alpha d : \alpha \in \mathbf{R}\}$ for $0 \neq d \in \mathbf{R}^n$, $y \in \mathbf{R}^n$.*

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

Proof: Since Q is convex and contains all of its own vertices, it necessarily contains all convex combinations of its vertices. So it only remains to show that every $y \in Q$ can be written as a convex combination of vertices of Q. We will prove this result through induction on $m - ra(y)$.

Basis: Let $m - ra(y) = 0$. Then $ra(y) = m$ and since $y \in Q$, y is a basic feasible solution, and therefore a vertex, of Q.

Inductive Step: Suppose $y \in Q$ can be written as a convex combination of vertices when $m - ra(y) < k$ for some $k > 0$, and consider $\bar{y} \in Q$ with $ra(\bar{y}) = m - k < m$. Choose a direction vector $0 \neq d \in \mathbf{R}^m$ such that $a_j^T d = 0$ for $j \in I(\bar{y})$. Since Q is bounded, there exist $\underline{\alpha} < 0, \bar{\alpha} > 0$ such that $\bar{y} + \alpha d$ lies in Q whenever $\underline{\alpha} \leq \alpha \leq \bar{\alpha}$ and outside Q otherwise. Geometrically, this is equivalent to moving positively and negatively in direction d until we run into a constraint. Now consider $\bar{y} + \underline{\alpha} d$ and $\bar{y} + \bar{\alpha} d$. We can write $\bar{y} = \frac{\bar{\alpha}}{\bar{\alpha} - \underline{\alpha}}(\bar{y} + \underline{\alpha} d) + \frac{-\underline{\alpha}}{\bar{\alpha} - \underline{\alpha}}(\bar{y} + \bar{\alpha} d)$.

Therefore $\bar{y}$ is a convex combination of two points in Q. Let us look more closely now at $\bar{y} + \bar{\alpha} d$. We defined $\bar{\alpha}$ such that $\bar{y} + \bar{\alpha} d$, in addition to holding with equality to all constraints where $\bar{y}$ held with equality, it was also tight to a constraint to which $\bar{y}$ was not tight, say $a_k^T z \leq c_k$. This means that $a_k^T \bar{y} < c_k$ and $a_k^T(\bar{y} + \bar{\alpha} d) = c_k$. Since $a_k \notin span\{a_j : j \in I(\bar{y})\}$, we have

$$ra(\bar{y} + \bar{\alpha} d) = rank\{a_j : j \in I(\bar{y} + \bar{\alpha} d)\} \geq rank(\{a_j : j \in I(\bar{y})\} \cup \{a_k\}) > ra(\bar{y}) .$$

So $m - ra(\bar{y} + \bar{\alpha} d) < k$; hence $\bar{y} + \bar{\alpha} d$ is a convex combination of the vertices of Q. The same argument holds for $\bar{y} + \underline{\alpha} d$, and so $\bar{y}$ itself is a convex combination of the vertices of Q. This completes the induction step, and we have proved the theorem. $\square$