Mathematical Programming	Lecture 7
OR 630 Fall 2005	September 15, 2005
Notes by Dan Sheldon	

1 Separating Hyperplane Theorem

Recall the statements of Weierstrass's Theorem (without proof) and the Separating Hyperplane Theorem from the previous lecture.

Theorem 1 *(Weierstrass) A continuous real-valued function defined on a nonempty compact set attains its minimum and maximum.*

Theorem 2 *(Separating Hyperplane) Let $C \subseteq \mathbf{R}^n$ be closed and convex. Let $x \in \mathbf{R}^n, x \notin C$. Then there exists $0 \neq a \in \mathbf{R}^n, \beta \in \mathbf{R}$ such that $a^T x > \beta$ and $a^T z < \beta$ for all $z \in C$.*

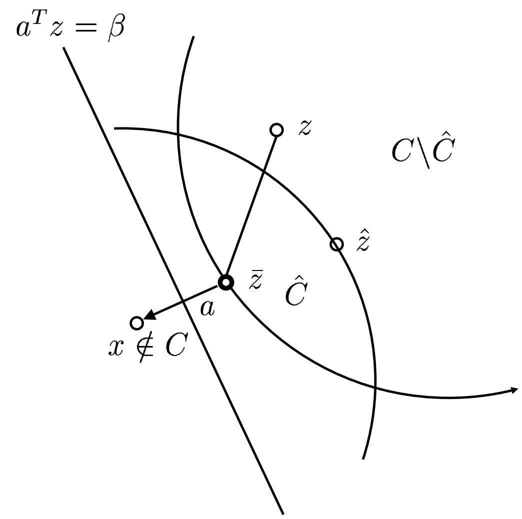

Figure 1: Illustration of proof technique

Figure 1 illustrates the proof technique. Using Weierstrass's Theorem, we will find the point $\bar{z}$ that minimizes the distance from x over C. However, C is not necessarily compact (it may be unbounded) so instead we choose an arbitrary point $\hat{z} \in C$, and restrict ourselves to the set $\hat{C}$ of all points in C no farther from x than $\hat{z}$. This set is compact, so we can apply Weierstrass's Theorem to find $\bar{z}$. We let $a = x - \bar{z}$, and choose β half way between $a^T x$ and $a^T \bar{z}$ to get the illustrated hyperplane. Finally, we show that any point $z \in C$ is farther from x than $\bar{z}$, so it is on the correct side of the hyperplane.

Proof: If C is empty, any a and appropriate β work, so assume C is nonempty. Choose $\hat{z} \in C$, and let

$$\hat{C} = \{z \in C : ||z - x|| \leq ||\hat{z} - x||\}.$$

This is the intersection of two closed sets, so it is closed. It is also bounded: if $z \in \hat{C}$, then by the triangle inequality, $||z|| \leq ||x|| + ||z - x|| \leq ||x|| + ||\hat{z} - x||$. So $\hat{C}$ is compact. Let $\bar{z}$ be the minimizer on $\hat{C}$ of

$$f(z) = \tfrac{1}{2}||z - x||^2 = \tfrac{1}{2}(z - x)^T(z - x).$$

1. Note that $\bar{z}$ minimizes f over all of C. Indeed, if $z \in C \backslash \hat{C}$, then $f(\bar{z}) \leq f(\hat{z}) < f(z)$.

2. Let $a := x - \bar{z}$. Then $a \neq 0$, since $\bar{z} \in C, x \notin C$.

3. Let $\beta := \frac{1}{2}(a^T x + a^T \bar{z})$. Then $a^T x > \beta$. Indeed,

$$0 < a^T a = a^T(x - \bar{z}) = a^T x - a^T \bar{z}$$

$$\Rightarrow \quad a^T x > a^T \bar{z} \quad \Rightarrow \quad a^T x > \frac{a^T x + a^T \bar{z}}{2} = \beta.$$

4. It remains to show that $a^T z < \beta$ for all $z \in C$. Consider any $z \in C$ and $z_\lambda := (1 - \lambda)\bar{z} + \lambda z \in C$ for $0 < \lambda \leq 1$. Since $\bar{z}$ minimizes f over C, $f(z_\lambda) \geq f(\bar{z})$, that is

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

   Simplifying,

$$\lambda(\bar{z} - x)^T(z - \bar{z}) + \tfrac{1}{2}\lambda^2||z - \bar{z}||^2 \geq 0$$
$$\Rightarrow \quad a^T(\bar{z} - z) \geq -\tfrac{1}{2}\lambda||z - \bar{z}||^2.$$

   But this is true for all $0 < \lambda \leq 1$, so we take the limit as λ approaches 0, and see that $a^T(\bar{z} - z) \geq 0$, or

$$a^T z \leq a^T \bar{z} < \frac{a^T x + a^T \bar{z}}{2} = \beta.$$

□

2 The polar of a set

Definition 1 *If $S \subseteq \mathbf{R}^n$, then its* polar *is $S^\circ = \{z \in \mathbf{R}^n : z^T x \leq 1, \forall x \in S\}$.*

Remark 1 *As an intersection of halfspaces containing 0, the polar of a set is always a closed convex set containing $0 \in \mathbf{R}^n$.*

Theorem 3 *(Bipolar) If C is a closed convex subset of $\mathbf{R}^n$ with $0 \in C$, then $C^{\circ\circ} := (C^\circ)^\circ = C$.*

Proof:

- ($\supseteq$) If $x \in C$, we want to show that $x \in C^{\circ\circ}$, i.e., that $z^T x \leq 1$ for all $z \in C^\circ$. But $z \in C^\circ$ implies $z^T x \leq 1$, so this holds.

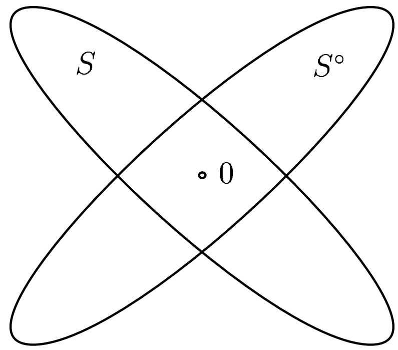

Figure 2: Illustration of the polar of a set (containing the origin)

- ($\subseteq$) We show this by proving that if $\bar{x} \notin C$ then $\bar{x} \notin C^{\circ\circ}$. By the Separating Hyperplane Theorem, there exists $0 \neq a \in \mathbf{R}^n$ and $\beta \in \mathbf{R}$ with $a^T\bar{x} > \beta > a^T x$ for all $x \in C$. Since $0 \in C$, it must be the case that $\beta > 0$. We'll scale a to get an inner product of 1, i.e., look at $z = a/\beta \neq 0$. Then $z^T\bar{x} > 1 > z^T x$ for all $x \in C$. So $x \in C^\circ$. But $z^T\bar{x} > 1$, so $\bar{x} \notin C^{\circ\circ}$.

$\square$

Now we're ready to prove that polytopes are polyhedra. We'll restrict our proof to polytopes Q with $0 \in int(Q)$, but it is true in general. If the interior of Q is nonempty, we can just translate Q so it contains the origin. The translated polytope is a polyhedron, which we can translate back to see that Q is a polyhedron. It is slightly harder if the interior of Q is empty. In this case Q lies completely in some (proper) affine subspace of $\mathbf{R}^n$. We can imagine translating and rotating the affine subspace so it contains the origin and lines up with the coordinate axes. Then we can "lop off" the axes corresponding to the orthogonal complement of the affine subspace. With respect to this lower dimensional subspace, the interior of Q will be nonempty, so our theorem will apply.

Theorem 4 *If $Q \subseteq \mathbf{R}^n$ is a polytope with $0 \in int(Q)$, then Q is a (bounded) polyhedron.*

Proof: We know that $Q = Q^{\circ\circ}$, that is $Q = P^\circ$ for $P = Q^\circ$. Note that if $Q = \text{conv}\{v_1, \ldots, v_n\}$, then

$$\begin{aligned} P &= \{z \in \mathbf{R}^n : x^T z \leq 1, \forall x \in Q\} \\ &= \{z \in \mathbf{R}^n : v_i^T z \leq 1, i = 1, \ldots, k\}. \end{aligned}$$

so P is a polyhedron. Q has 0 in its interior, so for some $\epsilon > 0$, all $x \in \mathbf{R}^n$ with $||x|| \leq \epsilon$ lie in Q. If $z \in P$, $z \neq 0$, then

$$x = \epsilon \frac{z}{||z||} \in Q.$$

Then

$$x^T z \leq 1 \quad \Rightarrow \quad \frac{\epsilon z^T z}{||z||} \leq 1 \quad \Rightarrow \quad ||z|| \leq \frac{1}{\epsilon}.$$

Hence P is a bounded polyhedron. By our previous result, P is a polytope, say $\text{conv}\{w_1, \ldots, w_l\}$. Then

$$Q = Q^{\circ\circ} = P^{\circ} = \{x : w_j^T x \leq 1, \quad j = 1, \ldots l\},$$

i.e., Q is a polyhedron.

$\square$

3 The Farkas Lemma

We now turn back to studying Linear Programs in more detail, characterizing feasibility, infeasibility and duality. Consider the standard form problem

$$\begin{array}{rrcl} \min & c^T x & & \\ & Ax & = & b \\ & x & \geq & 0 \end{array} \tag{1}$$

and its dual

$$\begin{array}{rrcl} \max & b^T y & & \\ & A^T y & \leq & c \end{array} \tag{2}$$

We'll use the Separating Hyperplane Theorem to show that the infeasibility of (1) can be characterized using a "theorem of the alternative".

Theorem 5 *(Farkas Lemma) Exactly one of (I) and (II) below has a feasible solution.*

I. $Ax = b, \quad x \geq 0$

II. $A^T y \leq 0 \quad b^T y > 0$

In the case that (II) is feasible, we call a solution y that is feasible in (II) a short proof or certificate of the infeasibility of (I). It may not be easy to find such a certificate, but it is easy to verify that (I) is infeasible given the certificate.

Proof:

- First we show it can't be the case that both (I) and (II) have a solution. Suppose x solves (I) and y solves (II). Then

$$0 < b^T y = y^T b = y^T (Ax) = (y^T A)x = (A^T y)^T x \leq 0.$$

This is an obvious contradiction. Hence y is a short proof of the infeasibility of (I).

- Next we show that at least one has a solution. So suppose that (I) has no solution, that is

$$b \notin \underbrace{\{z = Ax : x \geq 0\}}_{(*)}$$

The set (*) is clearly convex. It is also closed, which we'll show later. So we can invoke the Separating Hyperplane Theorem, and there is some $0 \neq y \in \mathbf{R}^m, \beta \in \mathbf{R}$ such that $b^T y > \beta > z^T y$ for all $z = Ax$ for some $x \geq 0$.

If we let $x = 0$, then $z = 0$, and it must be the case that $\beta > 0$.

Now if we let $x = \alpha e_j = (0, \ldots, 0, \alpha, 0, \ldots, 0)^T$, then $z = \alpha A e_j = \alpha a_j$, where $A = [a_1, \ldots, a_n]$, for $\alpha \geq 0$. This gives us

$$\beta > a^T y = \alpha a_j^T y \quad \Rightarrow \quad a_j^T y \leq 0 \quad \text{for all } j$$

$$\Rightarrow \quad A^T y \leq 0.$$

So we have $A^T y \leq 0$, and $b^T y > \beta$ for $\beta > 0$, or $b^T y > 0$. This means y is a solution to (II).

Now we'll return to showing that $\{z = Ax \in \mathbf{R}^n : x \geq 0\}$ is closed.

Let z be a limit point of the set, i.e., $z_k \to z$, where each $z_k = Ax_k, x_k \geq 0$. Is $z = Ax$ for some $x \geq 0$? If the x_k's converge (or have any limit point), we're OK, just choose x as this limit point. For each k, $Ax = z_k, x \geq 0$ has a feasible solution, so it has a basic feasible solution, say corresponding to β_k, ν_k. There are a finite number of ways to choose β_k and ν_k, so there must be some β_k, ν_k, say, β, ν, that repeat infinitely often, and for these k we get

$$x_{k_B} = B^{-1} z_k, \quad x_{k_N} = 0.$$

So $x_{k_B} \to x_B$, say, and then $(x_B; x_N)$ gives us

$$A[x_B; x_N] = Bx_B = \lim B(B^{-1} z_k) = \lim z_k = z.$$

$\square$