ORIE 6300 Mathematical Programming I September 6, 2016

Lecture 5

Lecturer: Damek Davis *Scribe: Rundong Wu*

1 Review

A while back, we defined polyhedrons and polytopes as follows.

Definition 1 *A* Polyhedron *is* $P = \{x \in \mathbb{R}^n : Ax \leq b\}$

Definition 2 *A* Polytope *is given by* $Q = conv(v_1, v_2, ..., v_k)$*, where the* v_i *are the vertices of the polytope, for* k *finite.*

Also recall the equivalence of extreme points, vertices and basic feasible solutions, and recall the definition of a bounded polyhedron.

Definition 3 *A polyhedron* P *is* bounded *iff* $\exists M > 0$ *such that* $||x|| \leq M, \ \forall x \in P$.

We showed bounded polyhedra were polytopes by taking the extreme points and seeing that they were the vertices for P as a polytope.

Recall also the Separating Hyperplane Theorem from a previous lecture.

Theorem 1 *(Separating Hyperplane) Let* $C \subseteq \mathbb{R}^n$ *be a closed, nonempty and convex set. Let* $y \in \mathbb{R}^n \backslash C$ *and let*

$$x^* = P_C(y) := \operatorname{argmin}_{x \in C} \frac{1}{2}\|x - y\|^2.$$

Then there exists a number $b \in \mathbb{R}$*, such that with* $a = y - x^*$*, we have*

$$(\forall x \in C) \qquad a^T x \leq a^T x^* < b < a^T y.$$

2 The polar of a set

Now we want to prove that polytopes are bounded polyhedra. To do this, we need to introduce one more concept.

Definition 4 *Let* $S \subseteq \mathbb{R}^n$*. We call the set*

$$S^\circ = \{z \in \mathbb{R}^n : z^T x \leq 1, \forall x \in S\},$$

the polar set of S*.*

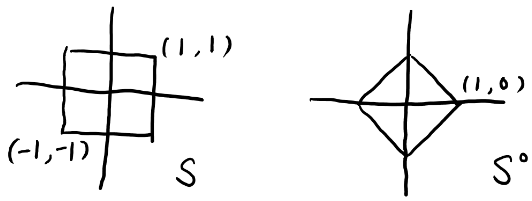

Figure 1: The polar of the ℓ_∞ ball is the ℓ_1 norm ball.

Example 1 (Polars of the ℓ_∞ and ℓ_1 balls.) *Consider $S = \{(x_1, x_2) : -1 \le x_1 \le 1,\ -1 \le x_2 \le 1\}$, the region is shown in the left of Figure 1.*

By definition, $x \in S^\circ$ if, and only if, $|x_1| + |x_2| = \sup_{(z_1,z_2)\in S}\{x_1z_1 + x_2z_2\} \le 1$. Thus, $S^\circ = \{x \in \mathbb{R}^n : \ |x_1| + |x_2| \le 1\}$, which is shown on the right hand side of Figure 1.

Now let's consider $S^{\circ\circ}$. By definition, $x \in (S^\circ)^\circ$ if, and only if, $\max\{|z_1|,\ |z_2|\} = \sup_{|x_1|+|x_2|\le 1}\{x_1z_1 + x_2z_2\} \le 1$. *Thus, $(S^\circ)^\circ = \{(x_1, x_2) : -1 \le x_1 \le 1,\ -1 \le x_2 \le 1\} = S$.*

Lemma 2 *If C is a closed convex subset of $\mathbb{R}^n$ with $0 \in C$, then $C^{\circ\circ} := (C^\circ)^\circ = C$.*

Proof:

- ($\supseteq$) Suppose that $x \in C$. Then $x \in C^{\circ\circ}$ if, and only if, $z^T x \le 1$ for all $z \in C^\circ$. This is clearly true because $z \in C^\circ$ implies that $z^T x \le 1$.
- ($\subseteq$) We will show that if $x \notin C$, then $x \notin C^{\circ\circ}$. First note that C is closed and convex with at least $z = 0 \in C$. If $x \notin C$, then by the Separating Hyperplane Theorem, there exists $0 \neq a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ with $a^T x > b > a^T z$ for all $z \in C$. Since $0 \in C$, we have $b > 0$. Let $\tilde{a} = a/b \neq 0$. Therefore $\tilde{a}^T x > 1 > \tilde{a}^T z$, for all $z \in C$. This implies $\tilde{a} \in C^\circ$. But $\tilde{a}^T x > 1$, so $x \notin C^{\circ\circ}$.

Therefore $C^{\circ\circ} = C$. □

3 Polytopes are Bounded Polyhedra

Now we can prove our result, at least sort of. We'll assume that 0 is in the interior of the polytope. We claim that this can be done without loss of generality; this is because we can translate the polytope to have $0 \in P$, apply the following proof and then translate back if needed.

Theorem 3 *If $Q \subseteq \mathbb{R}^n$ is a polytope with 0 in the interior of Q, then Q is a (bounded) polyhedron.*

Proof: Our proof strategy is as follows. (1) We will first show that the polar of a polytope is a polyhedron. (2) We then show that that since the polytope has 0 in its interior, then the polar of the polytope is bounded. So then $P = Q^\circ$ is a bounded polyhedron. (3) We know from a previous lecture that any bounded polyhedron is a polytope, so $P = Q^\circ$ is a polytope. (4) But then applying the proof that the polar of a polytope is a polyhedron, we get that $P^\circ = Q^{\circ\circ} = Q$ (by the lemma above) is a polyhedron. It is easy to prove that a polytope is bounded. Thus, we must prove (1) and (2).

(1) We first prove that the polar of Q is a polyhedron. Let $P = Q^\circ$. Then we know that $P^\circ = Q^{\circ\circ} = Q$. Since Q is a polytope, $Q = \text{conv}\{v_1, \dots, v_k\}$ for some k finite vectors $v_1, \dots, v_k \in \mathbb{R}^n$. We claim that

$$P = \{z \in \mathbb{R}^n : v_i^T z \leq 1,\ i = 1, \dots, k\}.$$

One the one hand, $P = Q^\circ = \{z \in \mathbb{R}^n : x^T z \leq 1,\ \forall x \in Q\}$, so $v_i^T z = z^T v_i \leq 1$ for $i = 1, 2, \dots, k$, which implies that $P \subseteq \{z \in \mathbb{R}^n : v_i^T z \leq 1,\ i = 1, \dots, k\}$. On the other hand, if $z^T v_i \leq 1$ for $i = 1, \dots, k$, then for any $x \in Q$, with $x = \sum_{i=1}^k \lambda_i v_i$ for some $\lambda_i \geq 0, \sum_i \lambda_i = 1$, we have

$$z^T x = z^T \sum_{is=1}^{k} \lambda_i v_i = \sum_{i=1}^{k} \lambda_i (z^T v_i) \leq \sum_{i=1}^{k} \lambda_i = 1,$$

which proves the claim. Thus P is a polyhedron.

(2) Because $0 \in int(Q)$, there exists some $\epsilon > 0$ such that all $x \in \mathbb{R}^n$ with $\|x\| < \epsilon$ lie in Q. If $z \in P$, $z \neq 0$, then, because $\|x\| < \epsilon$, we have

$$x = \frac{\epsilon}{2} \frac{z}{||z||} \in Q.$$

Because $P = Q^\circ$,

$$x^T z \leq 1 \quad \Rightarrow \quad \frac{\epsilon z^T z}{2||z||} \leq 1 \quad \Rightarrow \quad ||z|| \leq \frac{2}{\epsilon}.$$

Hence P is a bounded polyhedron. $\square$

4 Normal Cone

Modern optimization theory crucially relies on a concept called the *normal cone.*

Definition 5 *Let $S \subset \mathbb{R}^n$ be a closed, convex set. The* normal cone *of S is the set-valued mapping $N_S:\ \mathbb{R}^n \to 2^{\mathbb{R}^n}$, given by*

$$N_S(x) = \begin{cases} \{g \in \mathbb{R}^n | \, (\forall z \in S)\ g^T(z - x) \leq 0\} & if\ x \in S \\ \emptyset & if\ x \notin S \end{cases}$$

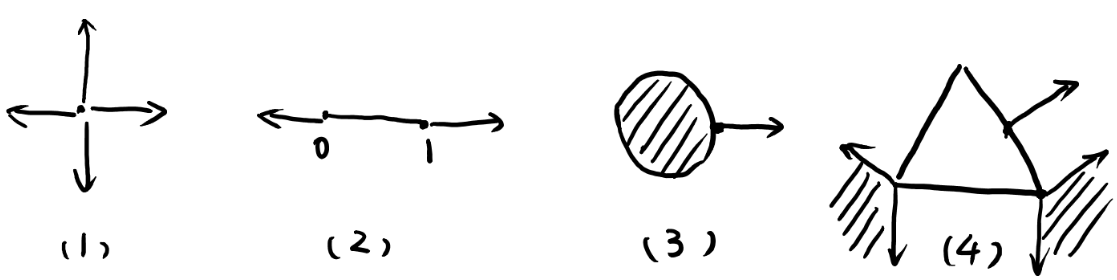

Figure 2: Normal cones of several convex sets.

Example 2 *We compute several normal cones; see Figure 2.*

1. *Let* $S = \{z\}$.
$$N_S(x) = \begin{cases} \mathbb{R}^n & if\ x = z \\ \emptyset & otherwise \end{cases}$$

2. *Let* $S = [0, 1]$.
$$N_S(x) = \begin{cases} \mathbb{R}_{\leq 0} & if\ x = 0 \\ \mathbb{R}_{\geq 0} & if\ x = 1 \\ \{0\} & if\ x \in (0, 1) \\ \emptyset & otherwise \end{cases}$$

3. *Let* $S = \{x \mid \|x\| \leq 1,\ x \in \mathbb{R}^n\}$.
$$N_S(x) = \begin{cases} \mathbb{R}_{\geq 0}x & if\ \|x\| = 1 \\ \{0\} & if\ \|x\| < 1 \\ \emptyset & otherwise \end{cases}$$

4. *The normal cone of a triangle, computed at some but not all points, is depicted in Figure 2.*

Normal cones satisfy several useful properties.

Proposition 4 *Let* $S \subseteq \mathbb{R}^n$ *be a nonempty, closed, convex set. Then the following hold:*

1. *If* $x \in S$, *then* $N_S(x)$ *is a* convex cone, *i.e.*
$$(\forall \lambda_1 \geq 0), (\forall \lambda_2 \geq 0), (\forall g_1 \in N_S(x)), (\forall g_2 \in N_S(x))\ \lambda_1 g_1 + \lambda_2 g_2 \in N_S(x).$$

2. *Let* $y \in \mathbb{R}^n \backslash S$, *then* $P_S(y) = x \iff y - x \in N_S(x)$.

3. *If* $x \in \mathrm{int}(S)$, *then* $N_S(x) = \{0\}$.

4. *If* $x \in S$ *and* $x \notin \mathrm{int}(S)$, *then* $N_S(x) \supsetneq \{0\}$.

Proof:

1. We leave the proof of part 1 as an exercise.

2. ($\Rightarrow$): By separating hyperplane theorem, with $a = y - P_S(y) = y - x$, $\exists\, b \in \mathbb{R}$, *s.t.*
$$\begin{aligned} (\forall z \in S) \quad & a^T z \leq a^T P_S(y) \leq a^T y \\ \Rightarrow \quad & a^T(z - P_S(y)) \leq 0 \\ \Rightarrow \quad & a \in N_S(P_S(y)) \end{aligned}$$

   ($\Leftarrow$): If $y - x \in N_S(x)$, then
$$\begin{array}{ll} \forall z \in S & (y-x)^T(z-x) \leq 0 \\ \iff & (y-x)^T(y-x) + (y-x)^T(z-y) \leq 0 \\ \Rightarrow & ||y-x||^2 \leq (y-x)^T(y-z) \leq ||y-x||\ ||y-z||\ (Cauchy-Schwarz\ inequality) \\ \Rightarrow & ||y-x|| \leq ||y-z||,\ \forall z \in S \\ \Rightarrow & x = P_S(y) \end{array}$$

3. Suppose that $g \in N_S(x)$. Because $x \in \text{int}(S)$, there exists $\epsilon > 0$, such that $x + \epsilon g \in S$. Therefore, we have

$$\begin{aligned} & g^T((x+\epsilon g) - x) \leq 0 \\ \Rightarrow \quad & \epsilon g^T g \leq 0 \\ \Rightarrow \quad & g = 0. \end{aligned}$$

   Hence, $N_S(x) = \{0\}$.

4. Since $x \notin \text{int}(S)$, there exists a sequence $y^k \in \mathbb{R}^n \backslash S$, *s.t.* $y^k \to x$ as $k \to \infty$. We leave it as an exercise to prove that if $x^k = P_S(y^k)$, then $x^k \to x$, $k \to \infty$.

   Let $g^k = \frac{y^k - x^k}{||y^k - x^k||}$. Then, by part 5, we have $g^k \in N_S(x^k)$. Without loss of generality, we can assume that $g^k \to g \in \mathbb{R}^n$, with $\|g\| = 1$.

   We claim that $g \in N_S(x)$. To prove the claim, note that since $g^k \in N_S(x^k)$, we have $(g^k)^T(z - x^k) \leq 0$, and

$$\begin{aligned} g^T(z-x) &= (g-g^k)^T(z-x) + (g^k)^T(z-x) \\ &= (g-g^k)^T(z-x) + (g^k)^T(x^k-x) + (g^k)^T(z-x^k) \\ &\leq (g-g^k)^T(z-x) + (g^k)^T(x^k-x) \\ &\to 0 \quad \text{as} \quad k \to \infty \end{aligned}$$

   So for all $z \in S$, we have $g^T(z - x) \leq 0$, which means $g \in N_S(x)$. Obviously $0 \in N_S(x)$, so $\{0\} \subsetneq N_S(x)$.

$\square$

Proposition 5 shows that normal cones detect the boundary and interior of convex sets.

Corollary 5 *Let $S \subseteq \mathbb{R}^n$ be a nonempty, closed, convex set. Then*

- $N_S(x) = \{0\}$ *if, and only if,* $x \in \text{int}(S)$.
- $N_S(x) \supsetneq \{0\}$ *if, and only if,* $x \in S \backslash \text{int}(S)$.