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) Let $C \subseteq \Re^n$ be be a closed, nonempty and bounded set, and let $f : C \to \Re$ be continuous on C. Then f attains a minimum on C.*

Theorem 2 *(Separating Hyperplane) Let $C \subseteq \Re^n$ be a closed, nonempty and convex set. Let $y \in \Re^n, y \notin C$. Then there exists $0 \neq a \in \Re^n, b \in \Re$ such that $a^T y > b$ and $a^T x < b$ for all $x \in C$.*

Proof: Define

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

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

Last time we showed that $\hat{C}$ is a closed, bounded, and non-empty set, so that we can 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

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

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, $f(z) \leq f(x_\lambda)$, i.e.

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

Rewriting, we obtain

$$\frac{1}{2}[2(z - y)^T \lambda(x - z) + \lambda^2 (x - z)^T (x - z)] \geq 0$$

or

$$(z - y)^T (x - z) + \frac{1}{2}\lambda(x - z)^T (x - z) \geq 0$$

or

$$a^T (z - x) + \frac{1}{2}\lambda(x - z)^T (x - z) \geq 0$$

or

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

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

$$b = \frac{1}{2}(a^Ty + a^Tz) \geq \frac{1}{2}(2a^Tz) = a^Tz > a^Tx.$$

□

2 The polar of a set

To get to the proof that polytopes are bounded polyhedra, we need to introduce one more concept.

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

Theorem 3 *If C is a closed convex subset of $\Re^n$ 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 $z^Tx \leq 1$ for all $z \in C^\circ$. But $z \in C^\circ$ implies $z^Tx \leq 1$, so this holds.
- ($\subseteq$) We will show that if $x \notin C$, then $x \notin C^{\circ\circ}$. If $x \notin C$, then by the Separating Hyperplane Theorem, there exists $0 \neq a \in \Re^n$ and $b \in \Re$ with $a^Tx > b > a^Tz$ for all $z \in C$. Since $0 \in C$, then $b > 0$. Let $\tilde{a} = a/b \neq 0$. Therefore $\tilde{a}^Tx > 1 > \tilde{a}^Tz$, for all $z \in C$. This implies $\tilde{a} \in C^\circ$. But $\tilde{a}^Tx > 1$, so $x \notin C^{\circ\circ}$.

Therefore $C^{\circ\circ} = C$. □

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 (and we'll leave it to the class to show on a problem set); this is because we can translate the polytope so that this is true if needed.

Theorem 4 *If $Q \subseteq \Re^n$ is a polytope with 0 in the interior of Q, then Q is a (bounded) polyhedron.*

Proof: Let $P = Q^\circ$. Then we know that $P^\circ = Q^{\circ\circ} = Q$. Since Q is a polytope, $Q = \text{conv}\{v_1, \ldots, v_k\}$ for some k finite vectors $v_1, \ldots, v_k \in \Re^n$. Now $P = Q^\circ = \{z \in \Re^n : x^Tz \leq 1, \forall x \in Q\}$, so $v_i^Tz = z^Tv_i \leq 1$ for $i = 1, 2, \ldots, k$. For any $x \in Q, x = \sum_{i=1}^k \lambda_i v_i$ where $\lambda_i \geq 0, \sum_i \lambda_i = 1$. Therefore

$$z^Tx = z^T(\sum_{i=1}^k \lambda_i v_i) = \sum_{i=1}^k \lambda_i(z^Tv_i) \leq \sum_{i=1}^k \lambda_i = 1.$$

Therefore

$$P = \{z \in \Re^n : v_i^Tz \leq 1, \, i = 1, \ldots, k\},$$

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

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

since $||x|| = \epsilon$. Then since $P = Q^\circ$,

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

Hence P is a bounded polyhedron. By our previous result, $P = Q^\circ$ is a polytope. And from what we just proved, this implies that P° is a bounded polyhedron, which means that $(Q^\circ)^\circ = Q$ is a bounded polyhedron. $\square$