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 \( C \subseteq \mathbb{R}^n \) be a closed, nonempty and bounded set, and let \( f : C \to \mathbb{R} \) be continuous on \( C \). Then \( f \) attains a minimum on \( C \).

\[ \text{Theorem 2 (Separating Hyperplane)} \quad \text{Let } C \subseteq \mathbb{R}^n \text{ be a closed, nonempty and convex set. Let } y \in \mathbb{R}^n, y \notin C. \text{ Then there exists } 0 \neq a \in \mathbb{R}^n, b \in \mathbb{R} \text{ such that } a^T y > b \text{ and } a^T x < b \text{ for all } x \in C. \]

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 \( x \in C - \hat{C} \), \( f(z) \leq f(x) \), 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,

\[ 0 < a^T a = (y - z)^T (y - z) = a^T y - a^T z \]

so then

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

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.

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

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 \( a^Tz < a^Ty \),
\[ b = \frac{1}{2}(a^Ty + a^Tz) \geq \frac{1}{2}(2a^Tz) = a^Tz > a^Tx. \]

\[ \square \]

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 \( S \subseteq \mathbb{R}^n \), then its polar is \( S^o = \{ z \in \mathbb{R}^n : z^Tx \leq 1, \forall x \in S \} \).

**Theorem 3** If \( C \) is a closed convex subset of \( \mathbb{R}^n \) with \( 0 \in C \), then \( C^{o0} := (C^0)^0 = C \).

**Proof:**

- (\( \supseteq \)) If \( x \in C \), we want to show that \( x \in C^{o0} \), i.e., that \( z^Tx \leq 1 \) for all \( z \in C^o \). But \( z \in C^o \) implies \( z^Tx \leq 1 \), so this holds.

- (\( \subseteq \)) We will show that if \( x \notin C \), then \( x \notin C^{o0} \). 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 \( a^Tx > b > a^Tz \) 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 \( \tilde{a} \in C^o \). But \( \tilde{a}^Tx > 1 \), so \( x \notin C^{o0} \).

Therefore \( C^{o0} = C. \)

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

**Proof:** Let \( P = Q^o \). Then we know that \( P^o = Q^{o0} = Q \). Since \( Q \) is a polytope, \( Q = \text{conv}\{v_1, \ldots, v_k\} \) for some \( k \) finite vectors \( v_1, \ldots, v_k \in \mathbb{R}^n \). Now \( P = Q^o = \{ z \in \mathbb{R}^n : z^Tv \leq 1, \forall v \in Q \} \), 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 \( \lambda_i \geq 0, \sum \lambda_i = 1 \). Therefore
\[ z^Tz = z^T(\sum_{i=1}^k \lambda_i v_i) = \sum_{i=1}^k \lambda_i(z^Tv_i) \leq \sum_{i=1}^k \lambda_i = 1. \]

Therefore
\[ P = \{ z \in \mathbb{R}^n : v_i^Tz \leq 1, i = 1, \ldots, k \}, \]
so \( P \) is a polyhedron. \( Q \) has 0 in its interior, so for some \( \epsilon > 0 \), all \( x \in \mathbb{R}^n \) 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 \( P = Q^o \),

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

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