Examples of Convex Sets and Separating Hyperplane Theorem

Orestis Plevrakis

In this lecture, we will see some more examples of convex sets and we will prove the separating hyperplane theorem.

1 Examples of Convex Sets

1.1 Balls

Let $x_0 \in \mathbb{R}^n$ and $r > 0$. We define the open ball with center x_0 and radius r as

$$B(x_0, r) := \{x \in \mathbb{R}^n \mid \|x - x_0\| < r\}$$

The closed ball is defined as $\overline{B(x_0, r)} := \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq r\}$.

Proposition 1. *Let $x_0 \in \mathbb{R}^n$ and $r > 0$. Then, both $B(x_0, r)$, $\overline{B(x_0, r)}$ are convex.*

Proof. I give the proof for $K := \overline{B(x_0, r)}$ (the proof for open balls is identical). Let $x, y \in K$ and $\theta \in [0, 1]$. I will show that $(1 - \theta)x + \theta y \in K$.

$$\begin{aligned}
\|(1-\theta)x + \theta y - x_0\| &= \|(1-\theta)x + \theta y - ((1-\theta)x_0 + \theta x_0)\| \\
&= \|(1-\theta)(x - x_0) + \theta(y - x_0)\| \\
&\leq (1-\theta)\|x - x_0\| + \theta\|y - x_0\| \leq (1-\theta)r + \theta r = r
\end{aligned}$$

□

1.2 Ellipsoids

Last time we saw that ellipsoids are sets of the form $\{x \in \mathbb{R}^n \mid (x - x_0)^\top A(x - x_0) \leq 1\}$, where $x_0 \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, $A \succ 0$.

Proposition 2. *Ellipsoids are convex.*

Proof. We will use the following characterization of positive semi-definite matrices (PSD): a matrix $A \in \mathbb{R}^{n\times n}$ is PSD iff there exists a matrix $Q \in \mathbb{R}^{n\times n}$ such that $A = Q^\top Q$ (you will prove this at the second homework). Now, fix an $x_0 \in \mathbb{R}^n$, an $A \in \mathbb{R}^{n\times n}$, $A \succ 0$, a matrix Q as above, and consider the corresponding ellipsoid E. Then, $x \in E \leftrightarrow \|Q(x - x_0)\| \leq 1$. Let $x, y \in E$ and $\theta \in [0, 1]$. The proof now proceeds as the previous one:

$$\begin{aligned}
\|Q((1-\theta)x + \theta y - x_0)\| &= \|Q((1-\theta)(x - x_0) + \theta(y - x_0))\| \\
&\leq (1-\theta)\|Q(x - x_0)\| + \theta\|Q(y - x_0)\| \leq 1 - \theta + \theta = 1
\end{aligned}$$

□

1.3 Polyhedra

A polyhedron is defined as a set of the form $\{x \in \mathbb{R}^n \mid a_i \cdot x \leq b_i,\ i = 1, \ldots, m\}$, for some $a_i \in \mathbb{R}^n$, $a_i \neq 0$, $b_i \in \mathbb{R}$. In other words, a polyhedron is a finite intersection of halfspaces: $\cap_{i=1}^m \{x \in \mathbb{R}^n \mid a_i \cdot x \leq b_i\}$. Using the following proposition, we get that polyhedra are convex.

Proposition 3. *Let K_1, K_2 convex sets in $\mathbb{R}^n$. Then $K_1 \cap K_2$ is convex.*

Proof. If $x, y \in K_1 \cap K_2$, then $[x, y] \subseteq K_1, K_2$, and thus $[x, y] \subseteq K_1 \cap K_2$. □

In two dimensions, bounded polyhedra are convex polygons.

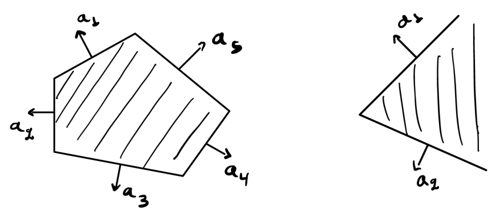
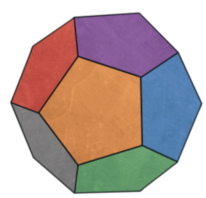


Figure 1: A 2d bounded polyhedron (convex polygon) on the left, and an unbounded on the right.

Here is an example of a three-dimensional polyhedron:



2 Separating Hyperplane Theorem

In the last lecture we mentioned the following theorem:

Theorem 4. *Let K be a closed [1] convex set in $\mathbb{R}^n$. Let $x_0 \in \mathbb{R}^n$ such that $x_0 \notin K$. Then, there exist nonzero $a \in \mathbb{R}^n$, and $b \in \mathbb{R}$ such that*

$$a \cdot x_0 > b$$
$$a \cdot x < b, \ \forall x \in K$$

This theorem says that given a closed convex set K, and a point x_0 outside it, there exists a hyperplane that separates K from x_0. We will first prove it for the case where K is also bounded (i.e., K is compact), and after we will show how this extra assumption can be removed. Now, whenever we want to show something non-trivial that holds in $\mathbb{R}^n$, it is almost always a good idea to try to show it first in dimensions 1, 2, or even 3. In these dimensions we can visualize what is going on, and we can also use classical Euclidean geometry, making it easier to come up with a solution. It happens often that the solution for such special cases reveals insights about how to attack the general. This is what will happen here.

[1] A set is closed if it contains its boundary.

2.1 Two dimensions

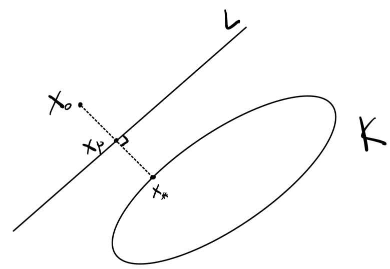
Before I go into the general case, I prove the following:

Theorem 5. *Let K be a compact convex set in $\mathbb{R}^2$. Let $x_0 \in \mathbb{R}^2$ such that $x_0 \notin K$. Then, there exists a line L that separates K from x_0.*

Proof. In Problem 1, you had to prove a very similar theorem. Mimicking the proof[2] for that one, we consider a point $x_* \in K$ that is closest to x_0, i.e.,

$$\|x_0 - x_*\| = \inf\{\|x_0 - x\| \mid x \in K\} \tag{1}$$

Such x_* exists because K is compact and the function $f : K \to \mathbb{R}$ defined as $f(x) = \|x_0 - x\|$ is continuous. Let x_μ be the midpoint of the segment $[x_0, x_*]$. We define L to be the line passing through x_μ that is perpendicular to $[x_0, x_*]$. Now, we need to show why x_0 and K lie on opposite



sides of L. There are several ways to prove this and all of them can be generalized to show the theorem for $\mathbb{R}^n$. Here is one way: Suppose there exists a point $x \in K$ such that either $x \in L$ or x is on the same side of L as x_0. This implies that the angle $\phi := \angle x_0 x_* x$ is acute (less that 90 degrees). Now that implies that the segment $[x_*, x]$ intersects the interior of the circle with center x_0 and radius $\|x_0 - x_*\|$ (see Figure 2). However, since K is convex, we have $[x_*, x] \subseteq K$, so there exists a point in K that is closer to x_0 than x_*, contradiction. $\square$

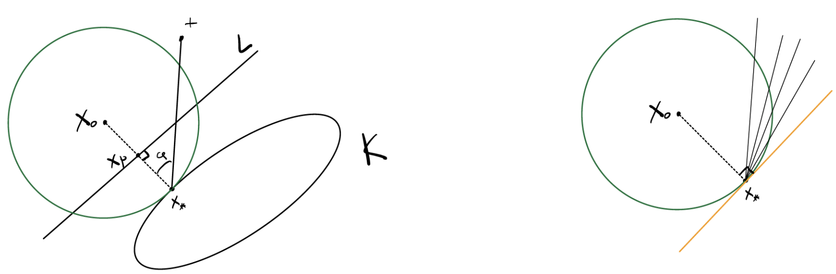


Figure 2: In the left diagram, you see an illustration of the argument. In the right diagram, the orange line is the tangent, and you can see why any halfline that starts from x_* and forms acute angle with $[x_*, x_0]$ will intersect (at least in its very beginning) the interior of the circle.

2.2 Proof of Theorem 4

I first prove the theorem assuming that K is bounded. The proof will simply be a direct translation of the above geometric argument, to the language of linear algebra. Let $x_* \in K$ such that

[2] You can see Problem 1 and its solution in the course website.

$\|x_0 - x_*\| = \inf\{\|x_0 - x\| \mid x \in K\}$ (observe that the existence of such x_* will be the only place we used compactness). We define

$$\begin{aligned} x_\mu &:= \frac{x_0 + x_*}{2} \\ a &:= x_0 - x_* \\ b &:= a \cdot x_\mu \end{aligned}$$

First, we show that $a \cdot x_0 > b$:

$$a \cdot x_0 - b = \; a \cdot (x_0 - x_\mu) = \frac{\|x_0 - x_*\|^2}{2} > 0$$

Now, suppose that there exists an $x \in K$, such that

$$a \cdot x \geq b \tag{2}$$

Motivated by the two-dimensional proof, we want to use 2 to show that the angle between the vectors $x_0 - x_*$ and $x - x_*$ is acute, i.e., $A := (x_0 - x_*) \cdot (x - x_*) > 0$. Let's show it:

$$a \cdot x - b = a \cdot (x - x_\mu) = (x_0 - x_*) \cdot (x - x_* + x_* - x_\mu) = A + (x_0 - x_*) \cdot (x_* - x_\mu)$$

By construction, $x_0 - x_*$, $x_* - x_\mu$ have opposite directions. Here is the algebraic proof: $(x_0 - x_*) \cdot (x_* - x_\mu) = -\frac{\|x_0 - x_*\|^2}{2} < 0$. Thus, $A > 0$.

Now, by convexity, we have $[x_*, x] \subseteq K$. Again motivated by the two-dimensional proof, the plan is to show that for *small enough* $\theta \in (0, 1)$, we have that $\|x_0 - ((1-\theta)x_* + \theta x)\| < \|x_0 - x_*\|$. [3]Here is the proof: for any $\theta \in [0, 1]$,

$$\|x_0 - ((1-\theta)x_* + \theta x)\|^2 = \|(x_0 - x_*) - \theta(x - x_*)\|^2 = \|x_0 - x_*\|^2 - 2\theta A + \theta^2\|x - x_*\|^2$$

Now, $-2\theta A + \theta^2\|x - x_*\|^2 = \theta(-2A + \theta\|x - x_*\|^2)$, and for small enough $\theta > 0$, we have $-2A + \theta\|x - x_*\|^2 < 0$. [4]

2.2.1 Removing the assumption that K is bounded

As you can observe, the only place where we used that K is bounded was to argue that

$$\exists x_* \in K, \; \|x_0 - x_*\| = \inf\{\|x_0 - x\| \mid x \in K\} \tag{3}$$

I now show that 3 holds, even when K is unbounded. Consider an arbitrary point $y \in K$, and let $R := \|x_0 - y\|$. Then, $\inf\{\|x_0 - x\| \mid x \in K\} = \inf\{\|x_0 - x\| \mid x \in K \cap \overline{B(x_0, R)}\}$. But now, $K \cap \overline{B(x_0, R)}$ is bounded, and is also closed (intersection of two closed sets is closed). Thus, since continuous functions over compact sets attain their minimum value, we get 3. $\square$

[3]If it is unclear to you why we are looking at small values of $\theta \in (0, 1)$, revisit Section 3.1 in lecture 2.

[4]Intuitively, for small $\theta > 0$, we have $\theta^2 << \theta$ which makes $-2\theta A$ dominate over $\theta^2\|x - x_*\|^2$.