Lecture 3

Dual sets and dual functions

- separation theorems
- polar sets
- dual norms

First separation theorem

Let $S \subseteq \mathbb{R}^n$ be a non-empty closed convex set and $v \not\in S$. Then there exists a $y \in \mathbb{R}^n$ s.t. $\langle y, v \rangle < \langle y, x \rangle$ for all $x \in S$.

Proof:

It suffices to prove that for all $y \in \mathbb{R}^n$:

$\langle y, v - x \rangle > 0 \quad \forall x \in S$
want to prove that \( \langle y, v - x \rangle > 0 \) for some \( y \in \mathbb{R}^n \)

there may be multiple hyperplanes that separate the convex set from the point, but a good candidate is the tangent at \( x^* \) closest \( x \in S \) to \( v \)

\[ \langle v - x^*, v - x \rangle > 0 \]

\[ \langle v - x^*, v - x - (x - x^*) \rangle > 0 \]

\[ \Rightarrow \|v - x^*\|^2 - \langle v - x^*, x - x^* \rangle > 0 \]

suffices to prove that \( \langle v - x^*, x - x^* \rangle \leq 0 \) for \( x \in S \) does this hold pictorially?

\[ \text{Yes, the tangent is } \perp \text{ to } v - x^* \]

and because of convexity the angle between \( x - x^* \) and \( v - x^* \) is 90°
\( \mathbf{x}^* \) minimizes \( \| x - \mathbf{v} \|^2 \) over \( S \)

Then \( \langle x - x^*, v - x^* \rangle \leq 0 \) \( \forall x \in S \)

**Proof**

Take \( z = (1 - \epsilon) x^* + \epsilon x \), \( x \in S \), \( \epsilon > 0 \)

\[
\| z - v \|^2 = \| (1 - \epsilon) x^* + \epsilon x - v \|^2 = \\
\| x^* - v \| - \epsilon (\epsilon \| x^* - x \|) = \| x^* - v \| + \epsilon^2 \| x^* - x \| \\
- 2 \epsilon \langle x^* - v, x^* - x \rangle \\
\Rightarrow \| z - v \|^2 - \| x^* - v \| = 2 \epsilon \| x^* - x \| - 2 \epsilon \langle v - x^*, x^* - x^* \rangle \\
\text{because } x^* \text{ minimizes RHS} \geq 0 \]

\[
\Rightarrow \epsilon^2 \| x^* - x \| - 2 \epsilon (\epsilon \| x^* - x^* \|) > 0 \]

\( \epsilon > 0 \)

\[
\Rightarrow (v - x^*)^T (x - x^*) \leq \frac{\epsilon}{2} \| x^* - x \| + \epsilon > 0 \\
\forall x \in S \\
\epsilon \rightarrow 0^+ \Rightarrow (v - x^*)^T (x - x^*) = 0
\]

**Technical detail**

\[
\| v - x^* \|^2 - \langle v - x^*, x - x^* \rangle > 0
\]

sufficient to prove that \( \langle v - x^*, x - x^* \rangle \leq 0 \)

Here we used that \( S \) is closed and then \( \| v - x^* \|^2 > 0 \)
Second separation theorem

Let $S_1$ and $S_2$ be two disjoint closed convex sets. If one of these sets is also bounded (say $S_2$) then $\forall y \in \mathbb{R}^n$

\[ s + \sup_{x \in S_1} (y \cdot x) < \min_{z \in S_2} (y \cdot z) \]

Proof:

Intuition: I want to prove that $\exists y \in \mathbb{R}^n$

\[ s + f_{x \in S_1, z \in S_2} \]

\[ y \cdot (x - z) < 0 = y \cdot 0 \]

with the zero vector

an element of the Minkowski difference $S_1 - S_2 = \{ x - z : x \in S_1, z \in S_2 \}$

we are "lucky" and we can use the first separation theorem, because:

1. $S_1$ convex $\Rightarrow S_1 - S_2$ convex

2. $S_1$ closed $\Rightarrow S_1 - S_2$ closed

3. $0 \in S_1 - S_2$ ($S_1, S_2$ are disjoint)

\[ \sup_{w \in S_2} y \cdot w < 0 \Rightarrow \sup_{x \in S_1} (x \cdot w) < 0 \Rightarrow \]

\[ \sup_{x \in S_1} x \cdot w < \inf_{z \in S_2} z \cdot w = \min_{z \in S_2} z \cdot w \]
Let \( W \) be a sequence in \( S_1 - S_2 \) and \( W_n \to W \). We want to prove that \( W \in S_1 - S_2 \).

\[ W_n = x_n - z_n, \quad x_n \in S_1, \quad z_n \in S_2 \]

exists a subsequence \( n_i \) such that \( z_{n_i} \to z \in S_2 \)

\[ x_{n_i} = \frac{W_{n_i} + z_{n_i}}{2} \to W + z \]

since the limit exists and \( S_1 \) is closed, we have that \( W + z \in S_1 \).

\[ W = (W + z) - z \in S_1 \]

\[ z \in S_2 \]

\[ \Rightarrow W \in S_1 - S_2 \]

\( S_1 - S_2 \) is closed

The assumption of our set to be bounded cannot be removed if we want strict separation.
Theorem (more general separating hyperplane)

Let $S_1, S_2$ two disjoint convex sets. Then

\[ \forall y \in \mathbb{R}^n \quad \sup_{x \in S_1} y^T x \leq \inf_{y \in S_2} y^T x \]

Definition - Supporting hyperplane

Given a set $S \subseteq \mathbb{R}^n$ and a point $x_0$ at the boundary of $S$.

A hyperplane $\{x \mid g^T x = g^T x_0\}$ is called a supporting hyperplane to $S$ at point $x_0$

if

\[ g^T x \leq g^T x_0 \quad \forall x \in S \]

Supporting hyperplane theorem

If $C \subset \mathbb{R}^n$ is convex then at any boundary point there exists a separating hyperplane.
Polar sets

Alternative representation of a convex set $C$.

**Support function**

$$S_c(y) = \sup \{ y^T x \mid x \in C \}$$

which shows why polar sets are important and intuitive.

**Lemma**

$C_1, C_2$ two closed convex sets

$$C = C_2 \iff S_{C_1}(y) = S_{C_2}(y) \forall y \in \mathbb{R}^m$$

**Proof intuition**

since the sets are closed and convex it is enough to leave the information about the boundary

**Proof sketch**

1. Prove that the boundaries are the same
2. $x \in \text{bound}(S_1) \Leftrightarrow x \in S_2$ and use the separating hyperplane theorem
knowing that \( S_c(y) = \sup \{ y^T x \mid x \in C \} \) is an important definition we can define the dual object of a set

**polar set of \( C \)**

\[ C^* = \{ y \in \mathbb{R}^n \mid y \cdot x \leq 1, \forall x \in C \} \]

**Observations**

1. \( C^* \) is convex (no matter what \( C \) is)

   \[ y_1, y_2 \in C^* \quad (\theta y_1 + (1-\theta)y_2) \cdot x = \theta y_1 \cdot x + (1-\theta)y_2 \cdot x \leq \theta \cdot 1 + (1-\theta) \cdot 1 = 1 \rightarrow \theta y_1 + (1-\theta)y_2 \in C^* \]

2. **Question:** Assume \( C \) is a closed convex set. When does \( C^* \) contain all the information about \( C \)?

   **Answer:** because in \( \emptyset \) there is no information stored. \( 0 \in C^* \) and \( S_C(0) = 0 \) no matter the convex set \( C \)

let \( y \in \mathbb{R}^n \) \( \theta \cdot 0 \) and \( \mu = \sup y \cdot x \mid x \in C \) > 0

\[ \mu = \sup \{ y \cdot x \mid x \in C \} > 0 \text{ iff } y \cdot x \leq \mu, \forall x \in C \]

\[ y \cdot x = \mu \text{ iff } x = \frac{1}{\mu} y \in C \]

**iff** \( \frac{y}{\mu} \cdot x \leq 1, \forall x \in C \)

\[ a \cdot \frac{y}{\mu} \cdot x > 1 \quad \forall a \in (a \neq C^* ) \text{ iff } \frac{y}{\mu} \cdot x = \frac{a}{\mu} \in C \]

if \( \mu = 0 \) all \( a, y \in C^* \) and \( S_C(a, y) = 0 \)

\( \forall a, y \).
So, if \( C \) is a closed convex set and \( S_c(y) > 0 \) for \( y \in \mathbb{R}^n \) then \( C^* \) maintains all the information about \( C \).

\[ \text{Question:} \]

When does \( S_c(y) > 0 \) for \( y \in \mathbb{R}^n \)

\[ \iff C \text{ contains the origin.} \]

\[ \text{proof} \]

using Separating hyperplane theorem

From all that discussion we are ready to formulate the

**Reconstruction theorem**

If \( C \) is a closed convex set that contains the origin then \( C^{**} = C \)
proof

We will prove that $C \subseteq C^{**}$ and $C^{**} \subseteq C$

$C^* = \{ y \mid y \cdot x \leq 1 \forall x \in C \}$

$C^{**} = \{ z \mid z \cdot y \leq 1 \forall y \in C^* \}$

$x \in C \Rightarrow y \cdot x \leq 1 \forall y \in C^* \Rightarrow x \in C^{**}$

(every element of $C^*$ lies on the ray $y \cdot x \leq 1$ with all $x \in C$

$\Rightarrow C \subseteq C^{**}$

Let $p \notin C$ (we will prove that $p \notin C^{**}$)

since $C$ is closed and convex $\Rightarrow \exists y \in \mathbb{R}^n$

$\sup_{x \in C} y \cdot x < y \cdot p$

$C$ contains the origin $\Rightarrow \sup_{x \in C} y \cdot x > 0$

If $\sup_{x \in C} y \cdot x > 0$ then we can scale up $y$

such that $y' = ay$, $a > 1 \Rightarrow y' \cdot p > 1$

$\sup_{x \in C} y' \cdot x = 1$
If \( \sup_{x \in C} y' \cdot x > 0 \) then we can scale up \( y' \) such that \( y' = a \cdot y' \), \( a > 1 \) s.t. \( y' \cdot p > 1 \) and \( \sup_{x \in C} y' \cdot x = 1 \).

If \( \sup_{x \in C} (y' \cdot x) = 0 \) then we can scale such that \( y' = a \cdot y' \), \( a > 1 \) s.t. \( y' \cdot p > 1 \) and \( \sup_{x \in C} (y' \cdot x) = 0 \).

Overall, we found \( y' \in \mathbb{R}^n \).

\[ y' \cdot p > 1 \quad \text{and} \quad \sup_{x \in C} (y' \cdot x) = 1 \]

\[ \Rightarrow y' \in \mathbb{R}^n \quad \Rightarrow \quad p \in C^{**} \quad \text{Eq.} \]
Dual norms

given a norm function \( \| \cdot \| : \mathbb{R}^n \to \mathbb{R} \)
the dual norm of \( \| \cdot \| \) is defined as
\[
\| y \|_{*} = \sup \{ y^T x \mid \| x \| \leq 1 \}
\]
or
\[
\| y \|_{*} = \sup \{ y^T x \mid x \in \mathcal{B} \}, \quad \mathcal{B} = \{ x \mid \| x \| \leq 1 \}
\]

\( \| \cdot \| \) measures how "big" is an element

\( \| \cdot \|_{*} \) measures how "big is the linear functional" associated with an element.
How much this linear functional can stretch elements

\[
\mathcal{B}^* = \left\{ y \mid \| y \|_{*} \leq 1 \right\} = \left\{ y \mid y^T x \leq 1 \text{ for } x \in \mathcal{B} \right\}
\]
polar set of \( \mathcal{B} \)

\( \mathcal{B} \) and \( \mathcal{B}^* \) are polar sets of each other. \( \Rightarrow \| x \|_{**} = \| x \| \) we only care about the unit circles.