Lecture 3 Dual sets and Dual functions

- separation theorems $\rightarrow$ point from closed convex set; $\rightarrow$ two disjoint closed convex sets
- polar sets
- dual norms

First separation theorem

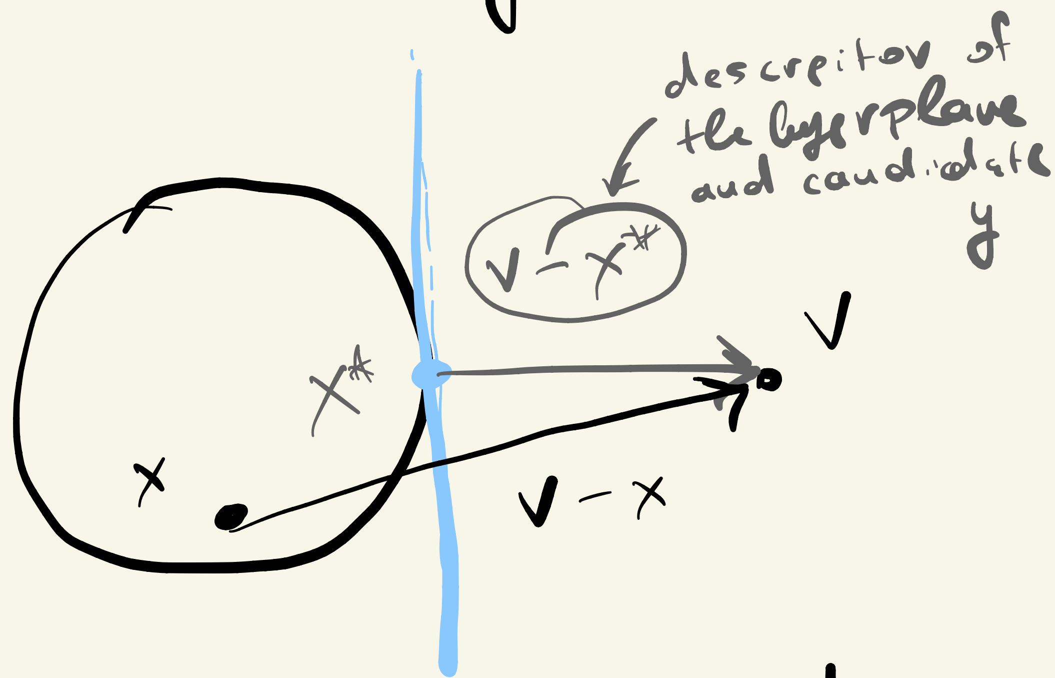
Let $S \subseteq \mathbb{R}^n$ be a non-empty closed convex set and $v \notin S$. Then there exists a $y \in \mathbb{R}^n$ s.t. $\langle y, v\rangle > \langle y, x\rangle \; \forall x \in S$

proof

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

$$\langle y, v - x\rangle > 0 \quad \forall x \in S$$

want to prove that $\langle y, v - x \rangle > 0 \ \forall x \in S$ 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 $v\}$

want to prove that

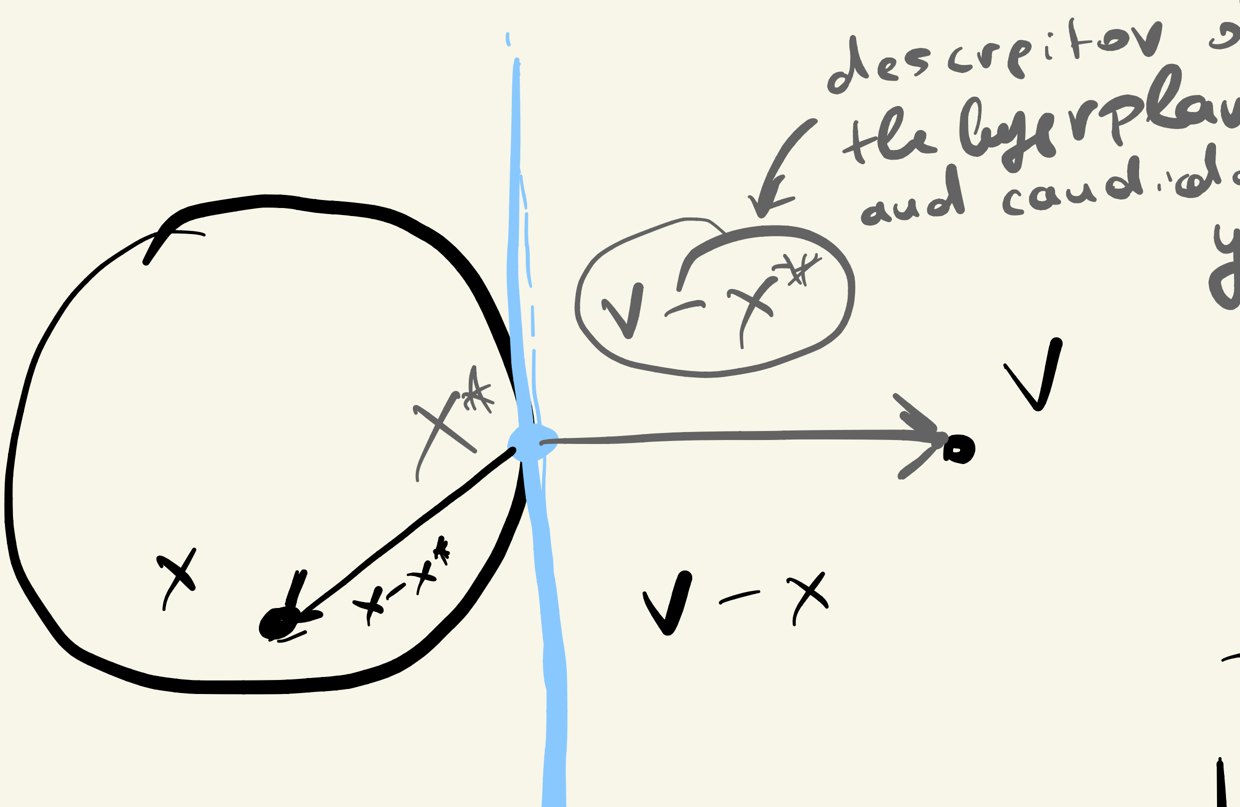
$\langle v - x^*, v - x \rangle > 0 \Rightarrow$

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

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

suffices to prove that $\langle v - x^*, x - x^* \rangle \leq 0$ $\forall x \in S$

does this hold pictorially?



yes, the tangent is $\perp$ to $v - x^*$

and because of convexity the angle between $x - x^*$ and $v - x^* \geq 90°$

x^* s.t. minimizes $\|x-v\|^2$ over S

then $\langle x-x^*, v-x^* \rangle \leq 0 \ \forall x \in S$

<u>proof</u>

take $z = (1-\varepsilon)x^* + \varepsilon x$, $x \in S$, $\varepsilon > 0$ $\left(\begin{array}{c}\text{by convexity} \\ z \in S\end{array}\right)$

$\|z-v\|^2 = \|(1-\varepsilon)x^* + \varepsilon x - v\| =$

$= \|x^* - v - \varepsilon(x^* - x)\| = \|x^* - v\| + \varepsilon^2 \|x^* - x\|$
$- 2\varepsilon(x^* - v, x^* - x)$

$\Rightarrow \|z-v\|^2 - \|x^* - v\| = \varepsilon^2 \|x^* - x\| - 2\varepsilon(v - x^*, x - x^*)$

because x^* minimizes RHS ≥ 0

$\Rightarrow \varepsilon^2 \|x^* - x\| - 2\varepsilon(v - x^*, x - x^*) \geq 0$

$\varepsilon > 0$
$\Rightarrow (v - x^*, x - x^*) \leq \frac{\varepsilon}{2} \|x^* - x\| \ \forall \varepsilon > 0$
$\forall x \in S$

$\varepsilon \to 0^+ \Rightarrow (v - x^*, x - x^*) \leq 0$

$\square$

<u>technical detail</u>

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

suffices to prove that $\langle v - x^*, x - x^* \rangle \leq 0$

here we used that S is closed and thus $\|v - x^*\|^2 > 0$

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 $\exists y \in \mathbb{R}^n$

s.t $\sup_{x \in S_1} (y, x) < \min_{z \in S_2} (y, z)$

<u>proof</u>

<u>Intuition</u>: I want to prove that $\exists y \in \mathbb{R}^n$ s.t $\forall x \in S_1, z \in S_2$

$$y(\underbrace{x-z}_{\text{an element of the Minkowski difference } S_1 - S_2 = \{x - z : x \in S_1, z \in S_2\}}) < 0 = y \cdot \vec{0}$$

separation with the zero vector

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

① $\left.\begin{array}{l} S_1 \text{ convex} \\ S_2 \text{ convex} \end{array}\right\} \Rightarrow$ $S_1 - S_2$ convex — easy to prove

② $\left.\begin{array}{l} S_1 \text{ closed} \\ S_2 \text{ compact} \end{array}\right\} \Rightarrow$ $S_1 - S_2$ closed

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

$$\sup_{w \in S_1 - S_2} y \cdot w < 0 \Rightarrow \sup_{\substack{x \in S_1 \\ z \in S_2}} (x-z)w < 0 \Rightarrow$$

$$\sup_{x \in S_1} xw < \inf_{z \in S_2} zw = \min_{z \in S_2} zw$$

② $\left.\begin{array}{l} S_1 \text{ closed} \\ S_2 \text{ compact} \end{array}\right\} \Rightarrow S_1 - S_2$ closed

proof

Let w_n be a sequence in $S_1 - S_2$ and $w_n \to w$. We want to prove that $w \in S_1 - S_2$.

$w_n = x_n - z_n$, $x_n \in S_1$, $z_n \in S_2$

exists a subsequence n_i s.t. $z_{n_i} \to z \in S_2$

$$x_{n_i} = \underbrace{w_{n_i}}_{\text{converges}} + \underbrace{z_{n_i}}_{\text{converges}} \to w + z$$

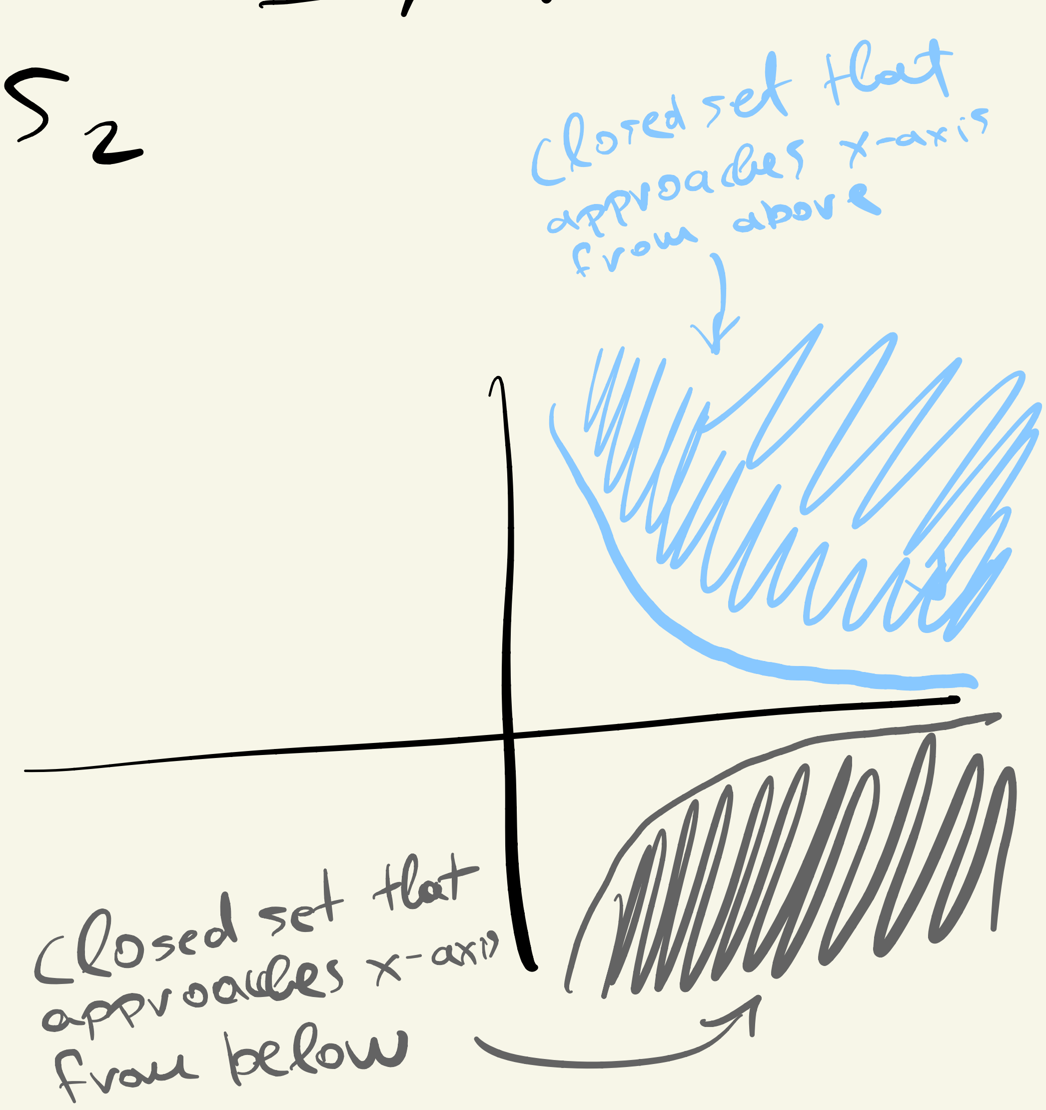
$\Rightarrow$ the limit exists and it is unique

since the limit exists and S_1 is closed we have that $w + z \in S_1$

$$w = \underbrace{(w+z)}_{\in S_1} - \underbrace{z}_{\in S_2} \Rightarrow w \in S_1 - S_2$$

$S_1 - S_2$ is closed

! the assumption of one set to be bounded cannot be removed if we want strict separation



Theorem (more general separating hyperplane theorem)

S_1, S_2 two disjoint convex sets. Then

$\exists\, y \in \mathbb{R}^n$ s.t. $\sup_{x \in S_1} y \cdot x \leq \inf_{z \in S_2} y \cdot z$

Definition – Supporting hyperplane

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

A hyperplane $\{x \mid y^T x = y^T x_0\}$ is called a supporting hyperplane to S at point x_0 if $y^T x \leq y^T x_0 \; \forall x \in S$

It is also defined by the direction

Supporting hyperplane theorem

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

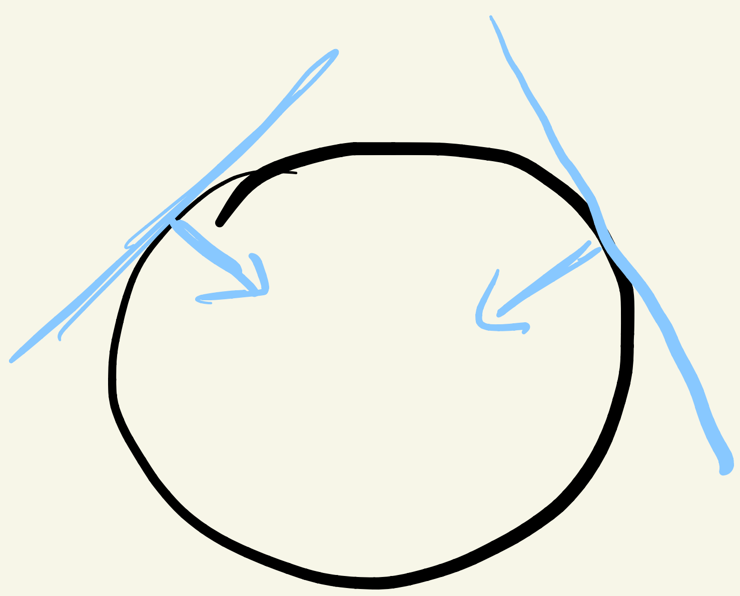
Polar sets

Alternative representation of a convex set C.

Support function

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

measures how far in a certain direction does C goes.



which shows why polar sets are important and intuitive

Lemma

C_1, C_2 two closed convex sets

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

Proof intuition $\Rightarrow$ since the sets are closed and convex it is enough to have the information about the boundary

Proof sketch: ① prove that the boundaries are the same
② $x_1 \in \text{bound}(S_1)$ $x_1 \notin S_2$ and use the separating hyperplane theorem

with strict separation

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

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

observations

① C^* is convex (no matter what C is)

$y_1 \in C^*$ $\quad (\theta y_1 + (1-\theta) y_2) x = \theta y_1 x + (1-\theta) y_2 x \leq$

$y_2 \in C^*$

$\theta \in [0,1]$ $\quad \leq \theta \cdot 1 + (1-\theta) \cdot 1 = 1 \Rightarrow$

$\Rightarrow \theta y_1 + (1-\theta) y_2 \in C^*$

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

Answer:

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

Let $y \in \mathbb{R}^n - \{0\}$ and $\mu = \sup\{y \cdot x \mid x \in C\} > 0$

$\mu = \sup\{y \cdot x \mid x \in C\} > 0$ iff $\quad y \cdot x \leq \mu, \forall x \in C$

$y \cdot x_0 = \mu, x_0 \in C$

iff $\frac{y}{\mu} \cdot x \leq 1 \quad \forall x \in C$

$a \cdot \frac{y}{\mu} \cdot x_0 > 1 \quad \forall a > 1 \quad (a \frac{y}{\mu} \notin C^*)$

if $\mu = 0$

all $a \cdot y \in C^*$

and $S_C(a \cdot y) = 0$ $\forall y$.

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

<u>Question</u> :

When does $S_C(y) \geq 0 \ \forall y \in \mathbb{R}^n$?

$\Leftrightarrow$ C contains the origin.

<u>proof</u>
using Separating hyperplane theorem

From all that discussion we are ready to formulate the

<u>Reconstruction theorem</u>

If C is a closed convex set that contains the origin then $C^{**} = C$

$\Longrightarrow$

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^* has to have $y \cdot x' \leq 1$ 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 $\Rightarrow \sup_{x \in C} y \cdot x \geq 0$

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

$\sup_{x \in C} y' x = 1$

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

$$\sup_{x \in C} y' x = 1$$

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

$$\sup_{x\in C} (y' \cdot x) = 0$$

Overall we found $y' \in \mathbb{R}^n$:

$$y'p > 1 \text{ and } \sup_{x\in C}(y'x) \leq 1$$

$\Rightarrow y' \in C^* \Longrightarrow p \notin C^{**}$ ▨

$\Longrightarrow$

Dual norms

given a norm function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$
the dual norm of $\|\cdot\|$ is defined as

$$\|y\|_* = \sup \{ y^T \cdot x \mid \|x\| \leq 1 \}$$

or

$$\|y\|_* = \sup \{ y^T x \mid x \in B \}, \quad B = \{ x : \|x\| \leq 1 \}$$

$\|\cdot\|$ ← 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

$$B^* = \{ y \mid \|y\|_* \leq 1 \} = \{ y \mid y^T x \leq 1 \ \forall x \in B \} =$$

polar set of B

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