Lecture 2

Last time

- Syllabus
- Motivation
- Overview

Today

- The setting
- convex sets
- Separation

The setting

We will work on a real Euclidean space $\mathbb{E}$. We denote its inner product by $\langle \cdot, \cdot \rangle$.

(Finite dimensional Hilbert space)

Examples

- Standard space: $\mathbb{R}^d$ with $x^T y$.
- Symmetric matrices:
  $S^n = \{$ "symmetric $n \times n$ matrices" $\}$
  with $\langle X, Y \rangle = tr(X^T Y)$.

We consider problems of the form

$$\begin{cases} \min & f(x) \\ \text{s.t.} & x \in C \end{cases} \quad \text{or} \quad \begin{cases} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0 \quad \forall i \in [m] \end{cases}$$

$[m] := \{1, \ldots, m\}$.

for some functions $f, g_i : E \to \mathbb{R} \cup \{\pm\infty\}$ and $C \subseteq E$.

They can take infinite values.

In the next few lectures we will focus on potential assumptions for C and f, g_i.

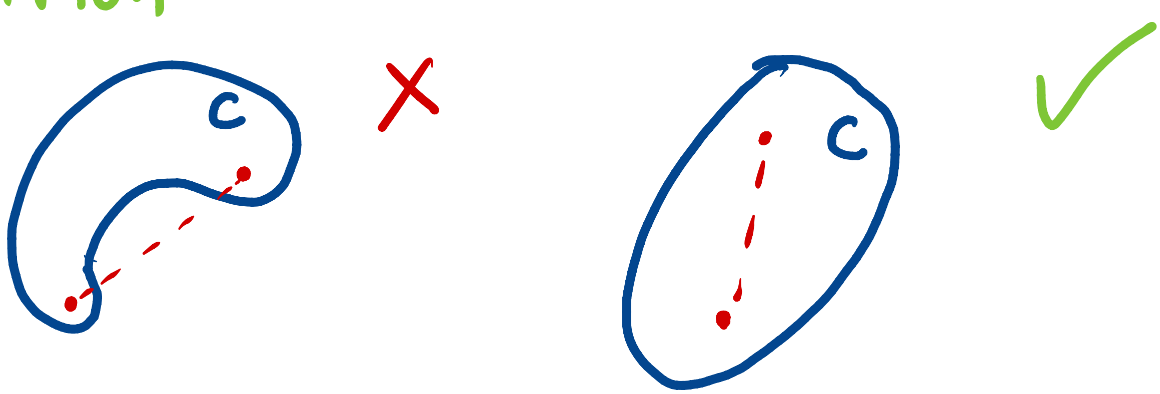
Convex sets

Def: A set $C \subseteq E$ is **convex** if for all $x, y \in C$ and $\lambda \in [0,1]$ we have

$$\lambda x + (1-\lambda) \in C.$$

+

Intuition



Examples

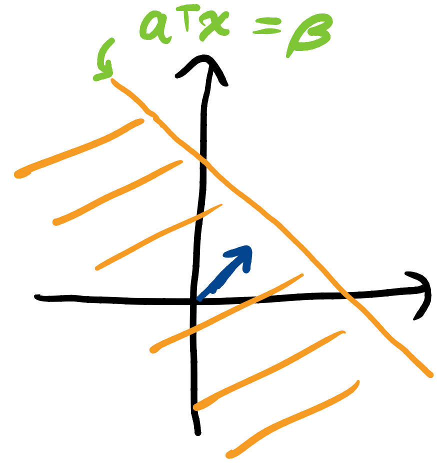
Two simple examples

▷ Half-spaces

$$H := \{ x \mid a^T x \leq \beta \}$$

for some $a \in E^*$ and $\beta \in \mathbb{R}$.

$E \setminus \{0\}$



▷ Unit ball

Define $\|x\| := \sqrt{\langle x, x \rangle}$, and

$$B = \{ x \in E \mid \|x\| \leq 1 \}.$$

The next result gives an easy way to identify convex sets.

Proposition: Arbitrary intersections of convex sets are convex.

Proof: Exercise. $\square$

Example

▷ Polyhedra

Any set of the form

$$P = \{ x \in E \mid \langle a_i, x \rangle \leq \beta_i \quad \forall i \in [m] \}$$

For some $a_i \in E$ and $\beta_i \in \mathbb{R}$.

▷ PSD matrices

$$S^n_+ = \{ X \in S^n \mid y^T X y \geq 0 \quad \forall y \in \mathbb{R}^d \}$$

is convex since

$$0 \le y'\bar{X}y = tr(y^T X y) = tr(X yy') = \langle X, yy^T\rangle;$$

thus it an infinite intersection of half spaces. $\dashv$

Denote int C and cl C the interior and closure of C. They'll play a key role later on and interact nicely with convex sets.

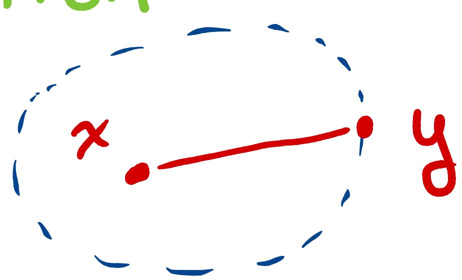
Proposition: Closures of convex sets are convex.

Proof: Exercise. $\square$

Lemma ($\Rightarrow$): Suppose C is convex. If $x \in \text{int } C$ and $y \in \text{cl } C$, then

$$(1-\lambda)x + \lambda y \in \text{int } C \quad \forall \lambda \in [0,1)$$

Intuition



The segment ●—○ belongs to int C.

Proof: Assume first $y \in S$.
There is a $\delta > 0$ s.t. $x + \delta B \subseteq S$.
By convexity, $\forall \lambda \in (0,1)$

$$(1-\lambda)(x + \delta B) + \lambda y \subseteq S$$

$$\Rightarrow (1-\lambda)x + \lambda y + (1-\lambda)\delta B \subseteq S,$$

which implies $(1-\lambda)x + \lambda y \in \operatorname{int} S$.

Now suppose $y \in \operatorname{cl} S$, then there is a sequence $\{y_k\} \subseteq S$ s.t. $y_k \to y$.
For $\lambda \in (0,1)$, we can write

$$\begin{aligned} &(1-\lambda)x + \lambda y \\ &= (1-\lambda)x + \lambda y_k + \lambda(y - y_k) \\ &= (1-\lambda)\underbrace{\left(x + \frac{\lambda}{(1-\lambda)}(y - y_k)\right)}_{z_k} + \lambda y_k. \end{aligned}$$

For large enough k, $z_k \in \operatorname{int} S$. Then, our first argument shows the conclusion.

$\square$

Corollary: The interior of convex sets is convex.

Proof: Take $x, y \in \operatorname{int} C$. Then, $y \in \operatorname{cl} C \Rightarrow (1-\lambda)x + \lambda y \in \operatorname{int} C \quad \forall \lambda \in [0,1]$ by Lemma ($\rightarrow$). $\square$

The next set of results will form the foundation of duality.

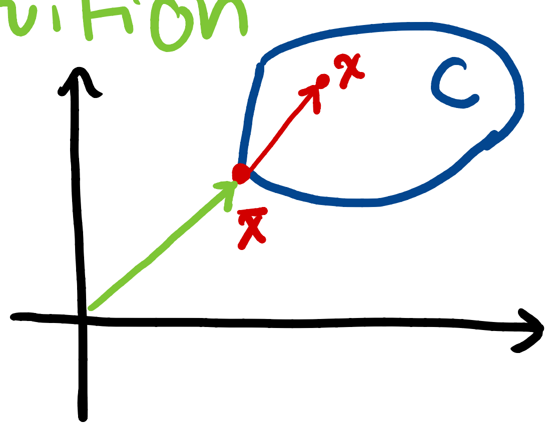
Theorem (Best approximation) Any nonempty closed convex set C has a unique shortest vector

$$\bar{x} = \underset{x \in C}{\operatorname{argmin}} \|x\|.$$

Moreover, it is characterized by

(☺) $\quad \langle \bar{x}, x - \bar{x} \rangle \geq 0 \quad \forall x \in S.$ $\dashv$

Intuition



↗ and ↗ are aligned.

Proof:

Existance. Choose any $\hat{x} \in C$, consider $C_1 = C \cap \|\hat{x}\| B$. Then

$$\underset{x \in S_1}{\min} \|x\|$$

$\|x\| \leftarrow$ continuous

$S_1 \leftarrow$ compact

achieves a minimizer x^*. Moreover, $\forall x \in C \setminus C_1$ we have $\|x^*\| \leq \|\hat{x}\| < \|x\|$.

Characterization. Let $\bar{x} \in \underset{x \in C}{\operatorname{argmin}} \|x\|$, then $\|\bar{x}\|^2 \leq \|\bar{x} + \lambda(x - \bar{x})\|^2 \quad \forall x \in C$. Expanding

$$0 \leq \lambda \|x - \bar{x}\|^2 + 2\langle \bar{x}, x - \bar{x} \rangle,$$

taking $\lambda \downarrow 0$ yields (☺). The other direction follows easily.

Uniqueness. Suppose $\bar{x}_1, \bar{x}_2$ satisfy (☺). Then

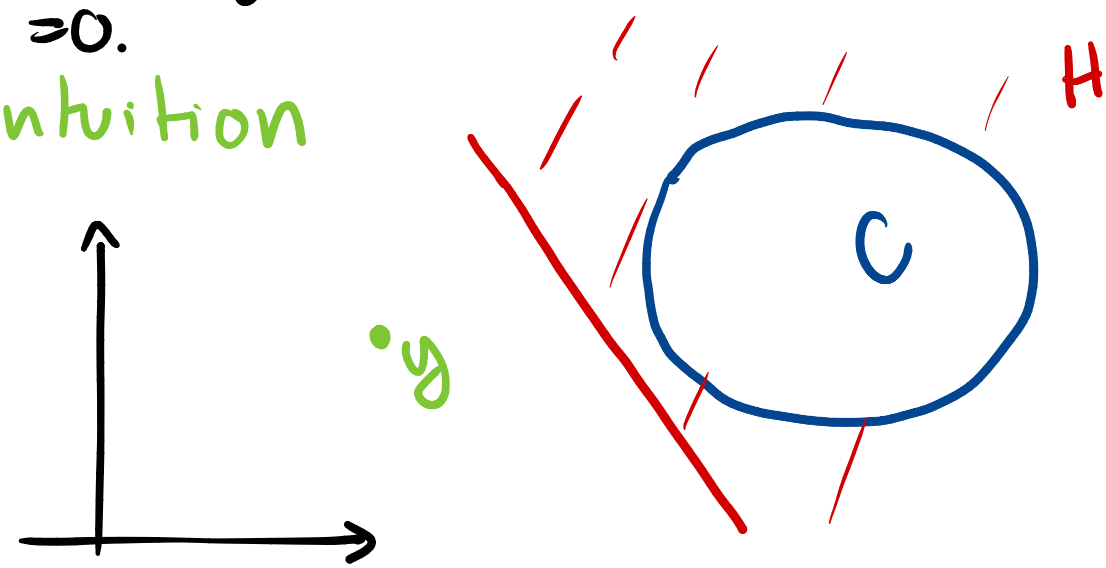
$$\begin{aligned} &\langle \bar{x}_1, \bar{x}_1 - \bar{x}_2 \rangle \leq 0 \\ + \; &\langle -\bar{x}_2, \bar{x}_1 - \bar{x}_2 \rangle \leq 0 \\ \hline &\|\bar{x}_1 - \bar{x}_2\|^2 \leq 0 \end{aligned}$$

Thus, $\bar{x}_1 = \bar{x}_2$. $\square$

Theorem (Basic separation)
Suppose C is a nonempty closed convex set and $y \notin C$. Then there exists a half space H s.t.
$C \subseteq H$ and $y \notin H$.

Proof: Apply previous result after a change of variables so that $y = 0$. $\square$

Intuition



Theorem (Hann-Banach) Suppose C is convex and $\operatorname{int} C \neq \emptyset$ and let $y \notin \operatorname{int} C$. Then, there exists a half space H containing C

$\mathrm{cl}\, C \setminus \mathrm{int}\, C$

with $y \in \mathrm{bd}\, C$.

Proof: WLOG assume $0 \in \mathrm{int}\, C$. For $n \in \mathbb{N}_+$, define

$$z_n = \left(1 + \frac{1}{n}\right) y.$$

Notice that $y = \frac{1}{n+1} 0 + \frac{n}{n+1} z_n$, then Lemma ($\to$) implies $z_n \notin \mathrm{cl}\, C$. Thus, by basic separation we obtain $\exists \{a_n\} \subseteq E^*$ s.t.

$$\langle a_n, z_n \rangle \geq \langle a_n, x \rangle \quad \forall x \in C.$$

WLOG we can take $\|a_n\| = 1$, (why?) then by the Bolzano-Weierstrass theorem there exist a subsequence a_{n_k} s.t. $a_{n_k} \to a$. Therefore

$$\langle a, y \rangle \geq \langle a, x \rangle \quad \forall x \in C.$$

The result follows by taking

$$H = \{ x \mid \langle a, x \rangle \leq \langle a, y \rangle \}. \quad \square$$