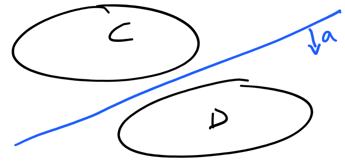
Separating and Supporting Hyperplanes

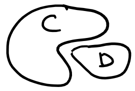
Sunday, January 24, 2021 10:23 PM

Thm Separating Hyperplane* Let C, D be convex non-intersecting sets in $\mathbb{R}^n$, then $\exists$ $a \in \mathbb{R}^n$ $(a \neq 0)$ and $\mu \in \mathbb{R}$ s.t.

$$a^T x \le \mu \quad \forall x \in C$$
$$a^T x \ge \mu \quad \forall x \in D$$



Clearly not true if we drop convexity



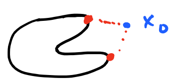
* Also true in more general Hilbert spaces but needs Hahn-Banach thm and you have to modify slightly (compactness, open)

Many variants (depending on open/closed, one set a singleton, strict inequalities)

First, some tools

Def A set S is a Chebyshev set if $\forall x_0$, $\exists!$ $x \in S$ s.t. $x = \underset{y \in S}{\operatorname{argmin}} \|y - x_0\|$
(ie., a unique best approximation pt.)

Ex: open unit ball isn't Chebyshev

Ex:  isn't Chebyshev

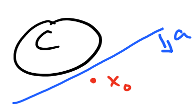
Fact Any nonempty, closed, convex set in a Hilbert space is Chebyshev
(proof in APPM 5440 (Hunter + Nachtergaele), exercise 6.1; counterexample in Banach space exer. 6.2)

For fun: are all Chebyshev sets convex? (cf. Frank Deutsch, "Best Approx. in Hilbert Space")

1) Classical: Motzkin's Thm. In $\mathbb{R}^2$ w/ Eucl. norm, yes, Chebyshev $\Rightarrow$ convex
   in fact, true in $(\mathbb{R}^n, \|\cdot\|_2)$
   in fact, true in any finite dim. space w/ "smooth" metric
2) not true in some weird non-Hilbert spaces
3) in general Hilbert space ... open problem!

Thm: Supporting hyperplanes, simplified variant

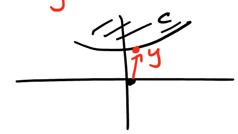
(i) If C is convex, closed and $D = \{x_0\}$, $x_0 \notin C$ then
$\exists$ $a \in \mathbb{R}^n$ s.t. $a^T x < a^T x_0$ $\forall x \in C$



(ii) Same but C need not be closed, $x_0 \notin \overline{C}$
(similar proof: just close C)

(iii) as in (ii) but allow $x_0 \in \overline{C} \setminus C$
(needs compactness argument)

Proof of (i): wlog let $x_0 = 0$. C is chebyshev so let y be the unique closest point to 0, and define $a = -y$



WTS $a^T x < a^T x_0 = 0 \quad \forall x \in C$

i.e., $y^T x > 0 \quad \forall x \in C$

let $\langle x, y \rangle := x^T y$

So let $x \in C$ be arbitrary,

$$\cancel{\|y\|^2} \le \|\underbrace{y + \varepsilon(x-y)}_{\substack{z \in C \text{ by convexity} \\ \|y\| \le \|z\| \text{ by best approx.}}}\|^2 = \cancel{\|y\|^2} + 2\varepsilon\langle y, x-y\rangle + \varepsilon^2 \|x-y\|^2$$

$$0 \le 2\cancel{\varepsilon}\langle y, x\rangle - 2\cancel{\varepsilon}\langle y, y\rangle + \varepsilon^{\cancel{2}}\|x-y\|^2$$

i.e., $2\|y\|^2 \le 2\langle y, x\rangle + \varepsilon\|x-y\|^2$

or $\langle y, x\rangle \geq \|y\|^2 - \varepsilon/2 \, \|x-y\|^2$ ($y \neq 0$ so $\|y\| \neq 0$)

$\Rightarrow \langle y, x\rangle > 0$ by choosing ε sufficiently small. $\square$

Related to Theorems of Alternatives

↳ generically: either A is true, B is false (or non-empty / or empty)

or A is false, B is true

"one or the other, not both"

"either-or"

Ex: Fredholm Alternative (finite-dim. version)

Either $\{x : Ax = b\}$ is nonempty

OR

$\{\lambda : A^T\lambda = 0, \ \lambda^T b \neq 0\}$ is non-empty (and NOT BOTH).

Why care? Professor asks you to prove there's a solution to $Ax=b$.

Simple: find a solution x. This is a "certificate"

But if professor asks you to prove there isn't a solution to $Ax=b$...

First task of duality → (e.g. A is singular, but you might get "lucky", e.g., $b=0$).

Solution: find a "certificate" λ.

Ex: Farkas Lemma → changes to Fredholm Alternative in red

Either $\{Ax = b, \ x \geq 0\}$ is non-empty

or $\{\lambda : A^T\lambda \geq 0, \ \lambda^T b < 0\}$ is non-empty

(and not both).

We'll prove something similar:

Ex: Theorem of Alternatives (for strict linear inequalities)

The set $\{x : Ax < b\}$ is empty (I)

iff the sets $C = \{b - Ax : x \in \mathbb{R}^n\}$ and $D = \mathbb{R}^m_{++}$ do not intersect iff the hyperplane separation theorem and its converse hold, i.e.,

$\exists\ \lambda \geq 0\ (\lambda \neq 0)$ s.t. $A^T\lambda = 0,\ \lambda^T b \leq 0.$ II

proof

II $\Rightarrow$ I (Converse of hyperplane separation)

Suppose such λ exists, and for contradiction, assume $\exists\ x$ s.t. $Ax < b$

then $\underbrace{\lambda^T A}_{A^T\lambda = 0} x < \underbrace{\lambda^T b}_{\lambda^T b \leq 0}$ since $\lambda \geq 0$

so $0 < \lambda^T b \leq 0$, contradiction.

I $\Rightarrow$ II By the separation thm., we know $\exists\ a$ ← call it λ now

s.t. $\lambda^T(b - Ax) \leq \mu,\quad x \in \mathbb{R}^n$ i.e., $\lambda^T z \leq \mu\quad \forall z \in C$

$\lambda^T y \geq \mu,\quad y \in \mathbb{R}^n_{++}$ $\lambda^T y \geq \mu\quad \forall y \in D$

$\Rightarrow\ \lambda^T A x = 0$ otherwise take $s \cdot x$ for $s \to \pm\infty$ to get contradiction

and $\lambda^T A x = 0\quad \forall x\ \Rightarrow\ \lambda^T A = 0$, i.e., $A^T\lambda = 0$.

$\Rightarrow\ \lambda \geq 0$ otherwise if $\lambda_i < 0$ pick $y_i = s,\ s \to +\infty$ to get contradiction

Also, need $\mu \leq 0$ since if $\mu > 0$, take $y_i \to 0^+,\ \lambda^T y \to 0^+$

Combine:

Have $\lambda \geq 0,\ \lambda \neq 0$, $A^T\lambda = 0$, and $\underbrace{\lambda^T(b - Ax)}_{= \lambda^T b} \leq \mu \leq 0$

$= \lambda^T b \leq 0$