nmi | spring 2024

lecture 22: hyperplane separation

this topic bc i had this conversation with a math 685 student last spring. he said you dont need to code to practice applied math. however, you need to programming languages to get a masters in applied math from hunter college. its been that way for a while.

ive mentioned this in the passing, i am sure. so can you do applied math without ever touching code? yes. bc applied mathematics is the application of math to other stuff - like physics, medicine, etc. but. you should always know more than you want to know or think you need to know. dream on if you want math unsullied by the lesser universe. at the very least you need to know enough code to know when someone who works for you or with you is jerking you around. bc computers arent going away, becoming less. more positively, coding is another tool to let you do more math and/or more with math.

reality lols

am i going to record a video? no, i am full of lies. instead, the below.

1. minkowskis separating hyperplane theorem, links, item 4 below. not quite two pages, entirely terms you would have known before the first day of class. it proves the hyperplane by coming at it from both sides.
2. the text after this list.
3. statistical learning 9.1, links, item 7 below. easily digestible video, with graphics. if youre in a hurry, maybe play it at 2x speed to scrape what you need.

the theorem

the usual hyperplane separation theorem states for every pair of convex, non-empty sets $X, Y \subseteq \mathbb{R}^n$ with disjoint interiors, there exists a non-null vector $z \in \mathbb{R}^{n-1} \backslash \{0\}$ where z is the separating hyperplane, much surprise.

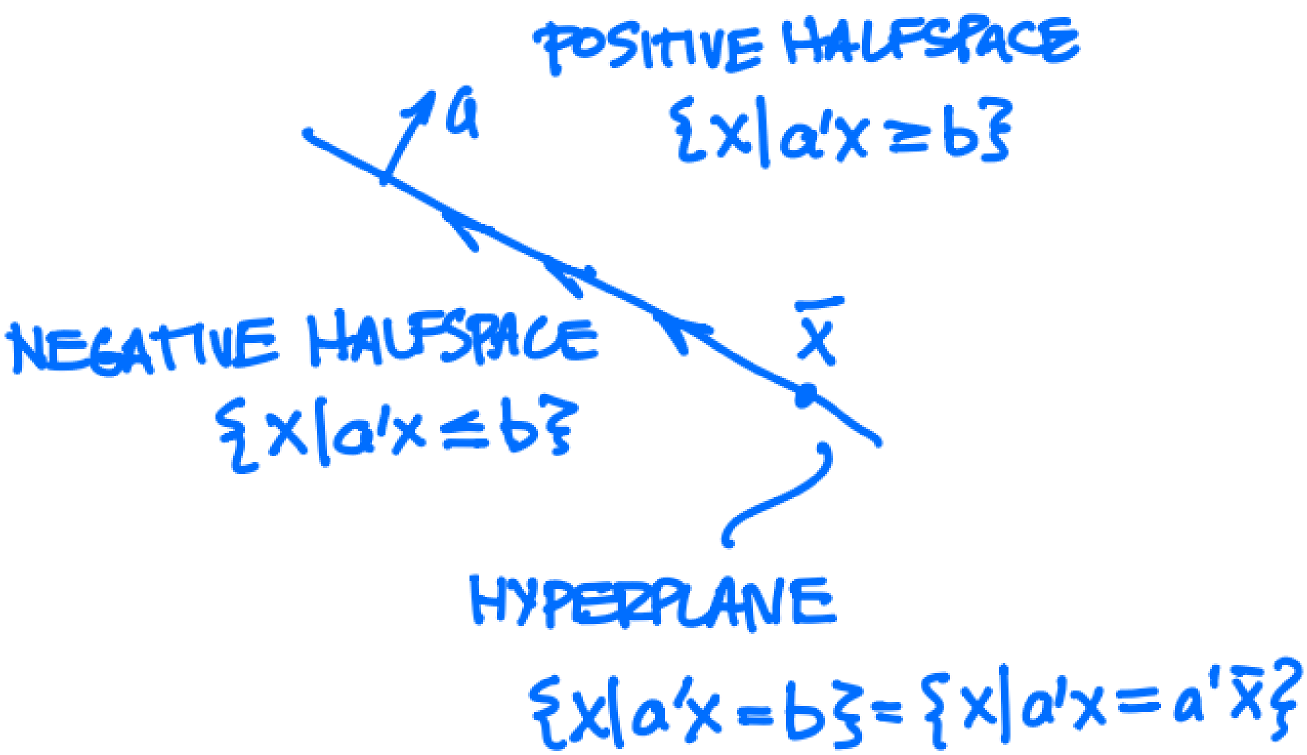
eg, for two convex sets $\in \mathbb{R}^2$, you can draw a $\mathbb{R}^1$ line between them.

refer to link 4 for a reasonably short proof. the one i meant to step through in lecture, alas.

vocabulary expansion pack

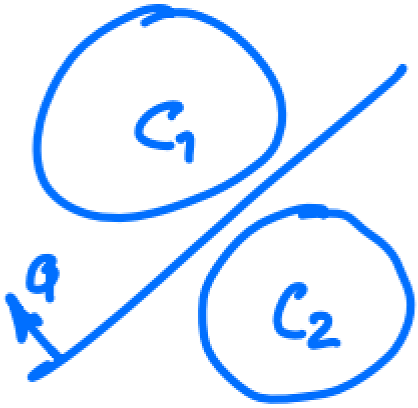
... that uses a different notation for $\mathbb{R}^{n-1}$

a **hyperplane** is a set of the form $\{x|a'x = b\}$ where $\vec{a}$ is a nonzero vector in $\mathbb{R}^n$ and b is a scalar. (ie, $a' \in \mathbb{R}^{n-1}$.)



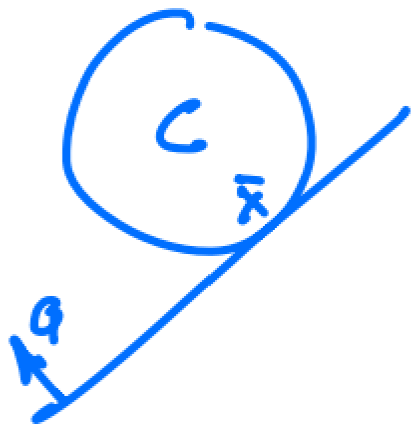
sets C_1, C_2 are **separated** by hyperplane $H = \{x|a'x = b\}$ if each lies in a different closed halfspace associated with H. as can be seen in the below underwhelming graphic of what is surely the real world.

eg, $a'x_1 \leq b \leq a'x_2$ for any $x_1 \in C_1, x_2 \in C_2$.



---

if $\bar{x}$ belongs to the closure of C, a hyperplane that separates C and the singleton $\{\bar{x}\}$ is said to be **supporting** C at $\bar{x}$.

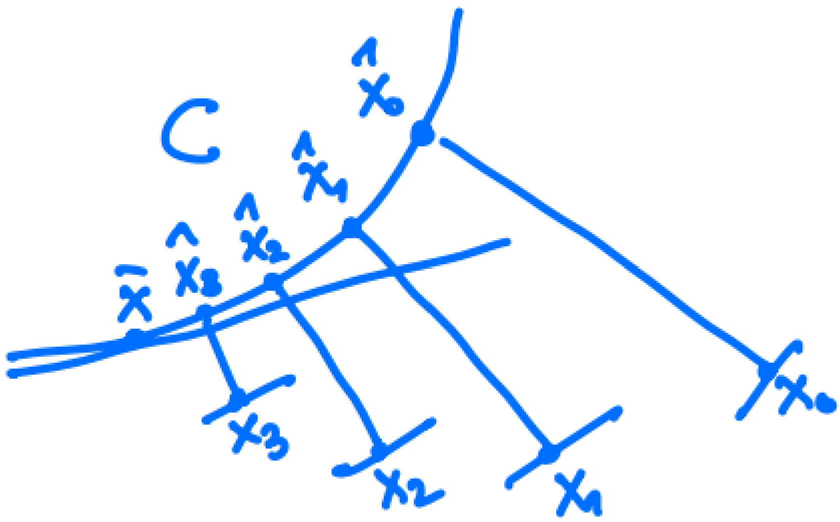


supporting hyperplane theorem. let C be convex and let $\bar{x}$ be a vector that is not an interior point of C. then there exists a hyperplane that passes through $\bar{x}$ and contains C in one of its closed halfspaces.

proof-lite. take sequence $\{x_k\}$ that does not belong to $\bar{C}$ and converges to $\bar{x}$. let $\hat{x}_k$ be the project of x_k on $\bar{C}$. for all $x \in \bar{C}$

$$a_k' x \geq a_k' x_k \quad \text{for any } x \in \bar{C}, k = 0, 1, \ldots$$

where $a_k = \frac{\hat{x}_k - x_k}{||\hat{x}_k - x_k||}$. let a be a limit point of $\{a_k\}$ and take limit as $k \to \infty$. ■



---

separating hyperplane theorem. let C_1, C_2 be two nonempty convex subsets of $\mathbb{R}^n$. if C_1, C_2 are disjoint, there exists a hyperplane that separates them. ie, there exists vector $a \neq 0$ such that

$$a'x_1 \leq a'x_2 \forall x_1 \in C_1, x_2 \in C_2.$$

proof-lite. consider convex set

$$C_1 - C_2 = \{x_2 - x_1 \big| \quad x_1 \in C_1, x_2 \in C_2\}.$$

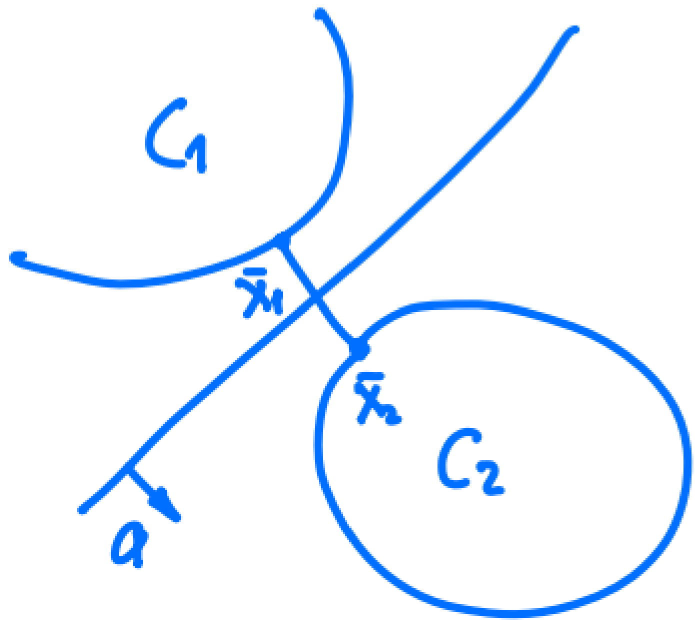
bc C_1, C_2 disjoint, the origin does not belong to $C_1 - C_2$. by the supporting hyperplane theorem (ding!), there exists vector $a \neq 0$ such that

$$a \leq a'x \quad \forall x \in C_1 - C_2. \blacksquare$$

---

a separating $\{x | a'x = b\}$ that is disjoint from C_1, C_2 is called **strictly separating**.

ie, $a'x_1 < b < a'x_2$ for any $x_1 \in C_1, x_2 \in C_2$.



strict separation theorem. let C_1, C_2 be two disjoint nonempty convex sets. if C_1 is closed and C_2 is compact, there exists a hyperplane that strictly separates them.

proof-lite. consider set $C_1 - C_2$. bc C_1 is closed and C_2 is compact, $C_1 - C_2$ is closed. bc $C_1 \cap C_2 = \emptyset \rightarrow 0 \notin C_1 - C_2$. let $\bar{x}_1 - \bar{x}_2$ be the projection of 0 onto $C_1 - C_2$. the strictly separating hyperplane is constructed.

also, thats not iif. any conditions that guarantee closedness of $C_1 - C_2$ gurantee the existence of a strictly separating hyperplane; however, a strictly separating hyperplane may exist without $C_1 - C_2$.

- well?

how does this all tie back to numerical methods? chat up your peers in the class discord if its unclear and you missed lecture.

- the links

---

in particular order:

1) math 747, functional space, spring 2023, hunter college, professor robert s thompson. its usually offered spring semester and as far as i know its thompsons class. i have decent notes but theyre cryptic and spliced with discourse as his lecture covered several chalkboards and a few days. his was specifically the strong separation wrt hilbert space with the hahn-banach separation theorem, if ive read my notes correctly. he made token mention at the end of how it differed from weak separation in a punchline kind of way. he also mentioned how the hyperplane was used in applied math / numerical analysis with regard to error, et al. and speaking of thomposons teaching chops, hes given lectures at this next resource.

2) mit ocw, an algorithmists toolkit, lectures 11 and 12. from the mother lode of all thats good online in terms of algorithms as far as im concerned. yeah, yeah, github. id originally wanted to do 3-4 special lectures, and this site alone had hundreds of applications. so maybe youre luckier than the kids next semester. *the proof sketch of lecture 11, section 3.1, was covered quickly in tuesdays lecture.*

3) robert peng, university of chicago. the hahn-banach separation theorem and some other stuff. this is that one thats 14 pages long and has 21 definitions, 16 lemmas, a corollary and 8 proofs. its pretty awesome. said the antithesis of a mathematician with a professed allergy to proofs. yes, this is more than you need for the final. by lots. but in case you were wondering if there was any math in applied math or if it was only this coding nonsense.

4) minkowskis separating hyperplane theorem, bruno salcedo, penn state. quick, easily digestible run-through of a proof.

---

in no particular order:

5) functional analysis, walter rudin, p58 and p70. if youre in math for keeps, this is a likely book to own irl. i picked mine up second-hand for a lucky 40-50u sd? (or maybe it was 100 bc some copies cost 200+.) tho i am a math tourist myself.

6) some dude at columbia had this lecture. i liked his order and organization but his conventions made me angry. this looks like hes compensating for not having a handy latex editor and i got tired of looking at it. but the examples are as clear as that 30-second guy on youtube who gives examples of hyperplanes as used in finance.

7) some dude at stanford had this paper.

8) statistical learning: 9.1 optimal separating hyperplane (from 2:14) @stanford online

and a few that i couldnt find again (there was a browser crash, lets say) just this moment.