Today: Convexity → any local minimum is a global minimum

- Convex sets
- seperating hyperplane theorem
- convex functions

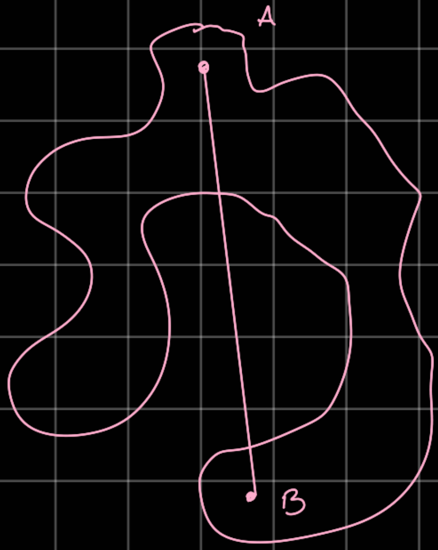
Convex combination:

$\sum_{i=1}^{n} \lambda_i \vec{x}_i$ is a convex combination of $\vec{x}_i$ if $\lambda_i \geq 0$ and $\sum_{i=1}^{n} \lambda_i = 1$

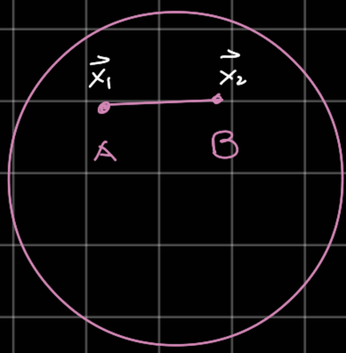
↳ can think about this as different weights or as probabilities (sum to 1 all ≥ 0)

Convex set

- A set C is convex if the line segment joining any two points in the set is contained in the set.



NOT CONVEX



CONVEX

let's mathematize this idea:

$\vec{x}_3 = \theta \vec{x}_1 + (1-\theta)\vec{x}_2 \in C \quad \forall \theta$

if $\vec{x}_1, \vec{x}_2 \in C$; $\theta \in [0,1]$

eg Hyperplane

$C = \{ \vec{x} \mid \vec{a}^T \vec{x} = b \}$ → equivalent to:

$\vec{a}^T(\vec{x} - \vec{x}_0) = 0$

$\vec{a}^T \vec{x} - \vec{a}^T \vec{x}_0 = 0$

$\vec{a}^T \vec{x} = \vec{a}^T \vec{x}_0$

Ayah attempt at proving/disproving convexity

For hyperplanes

$$C = \{\vec{x} \mid \vec{a}^T \vec{x} = b\}$$

$$= \{\vec{x} \mid \vec{a}^T(\vec{x} - \vec{x}_0) = 0\}$$

Know: $\vec{x}_3 = \theta \vec{x}_1 + (1-\theta)\vec{x}_2 \in C \quad \forall \theta$

if $\vec{x}_1, \vec{x}_2 \in C$; $\theta \in [0,1]$

$\vec{x}_1 \doteq \vec{a}^T \vec{x}_1 = b \Rightarrow \vec{a}^T \vec{x}_1 = a^T \vec{x}_0$

$\vec{x}_2 \doteq \vec{a}^T \vec{x}_2 = b \qquad \vec{a}^T \vec{x}_2 = a^T \vec{x}_0$

$\vec{x}_3 = \theta \vec{x}_1 + (1-\theta)\vec{x}_2$

$a^T \vec{x}_3 = \vec{a}^T(\theta \vec{x}_1 + (1-\theta)\vec{x}_2)$

$= \theta \vec{a}^T \vec{x}_1 + (1-\theta)\vec{a}^T \vec{x}_2$

$= \theta b + (1-\theta) b$

$= \theta b + b - \theta b$

$= b \Rightarrow$ it is convex.

Ranade: → she did the same thing as I did so yay!

$\vec{x}_1 \in C, \ \vec{x}_2 \in C$

$\vec{x}_3 = \theta \vec{x}_1 + (1-\theta)\vec{x}_2$

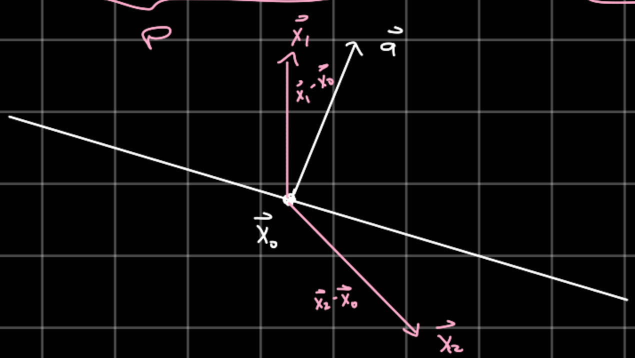
⋮

Hyperplanes & the seperating hyperplane thm

- hyperplanes divide space into 2 half-spaces

$\{\vec{x} \mid \vec{a}^T \vec{x} \geq b\}$ $\qquad$ $\{\vec{x} \mid \vec{a}^T \vec{x} \leq b\}$

$$\underbrace{\{\vec{x} \mid \vec{a}^T(\vec{x} - \vec{x}_0) \geq 0\}}_{P} \qquad \underbrace{\{\vec{x} \mid \vec{a}^T(\vec{x} - \vec{x}_0) \leq 0\}}_{N}$$



Consider :

$$P = \{ A \mid A \in S^n \text{ ie } A \in \mathbb{R}^{n\times n} \text{ symm}, A \text{ PSD} \}$$

PSD ↳ $\vec{x}^T A \vec{x} \geq 0 \quad \forall \vec{x} \in \mathbb{R}^n$

Q/ is P convex?

A/ Consider $A_1, A_2 \in P$

$$\theta A_1 + (1-\theta) A_2 = A_3$$

↳ use the $\vec{x} A \vec{x} \geq 0$ property

$$\theta \vec{x}^T A_1 \vec{x} + (1-\theta) \vec{x}^T A_2 \vec{x} = \vec{x}^T A_3 \vec{x} \geq 0$$

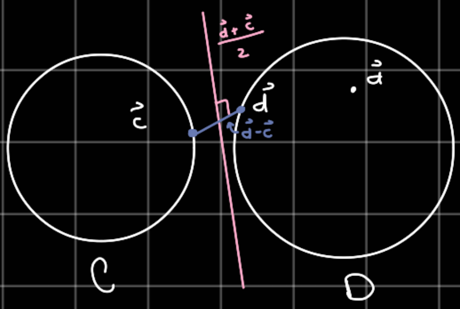
$A_3 \in P \Rightarrow P$ is convex

seperating hyperplane thm

$C \cap D = \emptyset$ ↪ not intersecting

- let C, D be convex sets, then $\exists$ hyperplane $\vec{a}^T \vec{x} = b$ then

$$\forall \vec{x} \in C \quad \vec{a}^T \vec{x} \geq b$$
$$\forall \vec{x} \in D \quad \vec{a}^T \vec{x} \leq b$$



- To prove this, we'll prove by existance

distance btwn C & D

- $\text{dist}(C, D) = \inf \{ \|\vec{c} - \vec{d}\|_2 \mid \vec{c} \in C, \vec{d} \in D \}$

↳ infimum (largest lower bound)

$y \leq f(x) \quad \forall x$

y is the largest such lower bound

- let $\vec{c}, \vec{d}$ be the points that are closest to each other

Consider

- $\vec{d} - \vec{c}$ as normal that passes through $\frac{\vec{d}+\vec{c}}{2}$

↳ the equation for this is :

$$f(\vec{x}) = (\vec{d}-\vec{c})^T \left(\vec{x} - \frac{\vec{d}+\vec{c}}{2}\right) = 0$$ } eqn for a hyperplane

$f(\vec{x}) = 0$ is a hyperplace

$$f(\vec{d}) = (\vec{d}-\vec{c})^T \left(\vec{d} - \frac{\vec{d}+\vec{c}}{2}\right) = \frac{1}{2}\|\vec{d}-\vec{c}\|^2$$

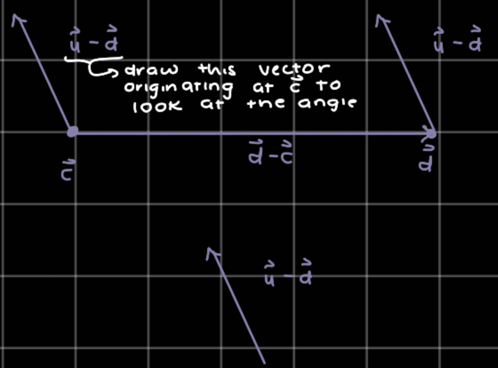
$$f(\vec{c}) = (\vec{d}-\vec{c})^T \left(\vec{c} - \frac{\vec{d}+\vec{c}}{2}\right) = -\frac{1}{2}\|\vec{d}-\vec{c}\|^2$$

↳ want to prove that $f(\vec{x}) \geq 0 \quad \forall d \in D$;

$f(\vec{x}) \leq 0 \quad \forall \vec{c} \in C$

· if possible, let $\vec{u} \in D$ s.t. $f(\vec{u}) < 0$

$$f(\vec{u}) = (\vec{d}-\vec{c})^T (\vec{u} - \tfrac{1}{2}(\vec{d}+\vec{c}))$$
$$= (\vec{d}-\vec{c})^T (\vec{u} - \tfrac{1}{2}(\vec{d}+\vec{c}) + \vec{d} - \vec{d})$$
$$= (\vec{d}-\vec{c})^T ((\vec{u}-\vec{d}) + \tfrac{1}{2}(\vec{d}-\vec{c}))$$
$$= \underbrace{\langle \vec{d}-\vec{c}, \vec{u}-\vec{d} \rangle}_{\text{if } f(\vec{u})<0 \text{, then this is } <0} + \underbrace{\tfrac{1}{2}\|\vec{d}-\vec{c}\|_2^2}_{\text{this is positive}}$$



← want to move along $\vec{d}$: $\vec{d} + t(\vec{u}-\vec{d})$

↳ since $\langle \vec{d}-\vec{c}, \vec{u}-\vec{d} \rangle < 0$, then this means that it'll be closer to $\vec{c}$.

→ weird: want to build up a contradiction from this