# EECS 127 — Lecture 10.

**Today.** Convexity.

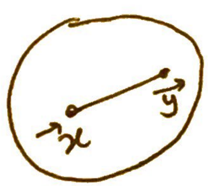
- Convex Sets
- Convex functions.

**Def$^n$:** Convex Combination.

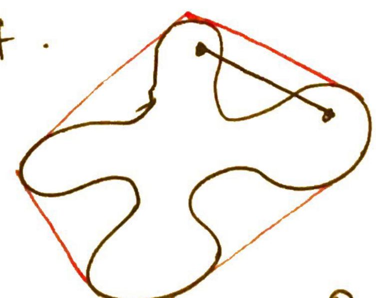
$\sum_{i=1}^{n} \lambda_i \vec{x}_i$ if $\sum_{i=1}^{n} \lambda_i = 1$ $\lambda \geq 0$.

**Def:** Convex set.

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




Convex



Not convex.

$\theta \in [0, 1]$.

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

$\theta \cdot \vec{x}_1 + (1-\theta) \cdot \vec{x}_2 \in C$

**Def:** Convex Hull.

e.g. $C = \{\vec{x} \mid \vec{a}^T \vec{x} = b\}$.     $\vec{a}^T(\vec{x} - \vec{x}_0) = 0$

Hyperplane.

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

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

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

$\Rightarrow \vec{x}_3 \in C$.   $\therefore$ C is convex.

$\{\vec{x} \mid \vec{a}^T \vec{x} \le b\}$.   Half-spaces.

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

e.g: $P = \{ A \mid A \in \mathbb{R}^{n \times n}. \ A \text{ is PSD (symmetric)} \}$.

A: PSD $\Rightarrow$ $\vec{x}^T A \cdot \vec{x} \geq 0 \quad \forall \vec{x} \in \mathbb{R}^n$.

Consider: $A_1, A_2 \in P$.

$\theta \in [0, 1]$.

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

Consider: $\vec{x}^T A_3 \vec{x} = \vec{x}^T (\theta A_1 + (1-\theta) A_2) \cdot \vec{x}$

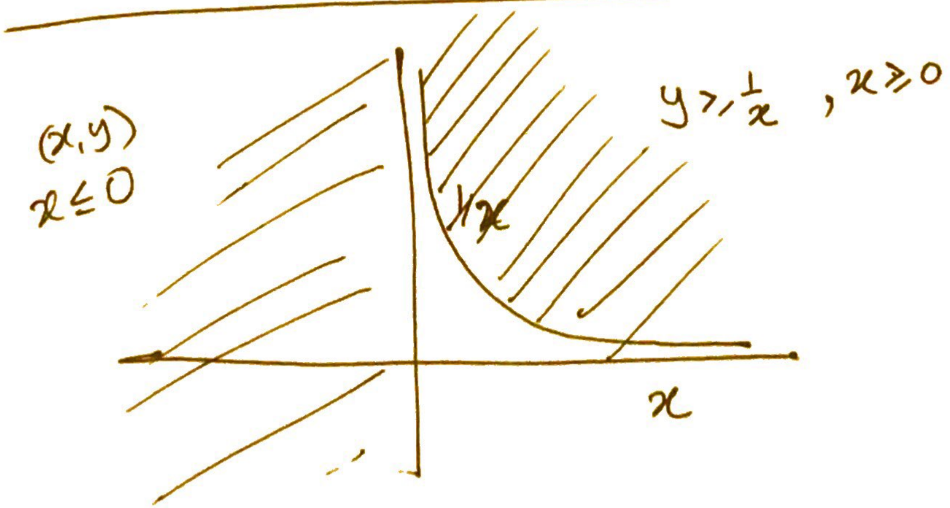
$= \theta \vec{x}^T A_1 x + (1-\theta) \vec{x}^T A_2 \cdot \vec{x}$.

$\forall \vec{x}$

$\geq 0$.

$A_3$ is also PSD. $A_3 \in P$.

---



# Separating Hyperplane Theorem.

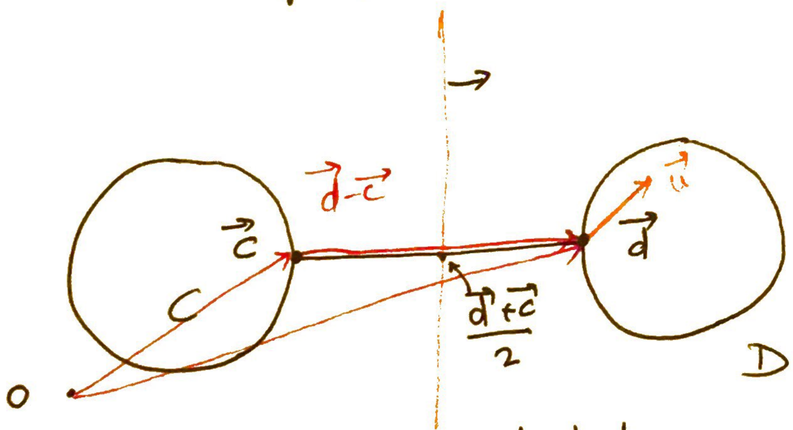
Let $C$ and $D$ be two convex sets.

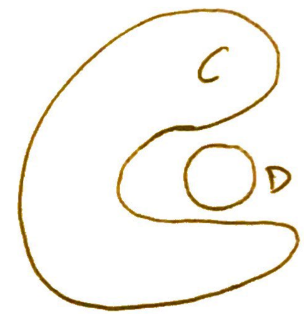
$C \cap D = \phi$.

Then, there exists a hyperplane $\vec{a}^T \vec{x} = b$.

s.t. $\forall \vec{x} \in C$, $\vec{a}^T \vec{x} \geq b$

$\forall \vec{x} \in D$, $\vec{a}^T \vec{x} \leq b$.




**Proof:** Define: "distance" between $C$, $D$.

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

Say $\vec{c}, \vec{d}$ the two points that are closest together.

Assume. $\vec{c}, \vec{d}$ exists.

Consider, the hyperplane with $\vec{d}-\vec{c}$ as a normal and that passes through the midpoint of the segment joining $\vec{d}, \vec{c}$.

midpoint $= \dfrac{\vec{d}+\vec{c}}{2}$.

$\vec{a}^T = (\vec{d}-\vec{c})^T$

$x_0 = \dfrac{\vec{d}+\vec{c}}{2}$.

$$f(\vec{x}) = \vec{a}^T (x - x_0)$$

$$= (\vec{d}-\vec{c})^T (\vec{x} - \tfrac{1}{2}(\vec{d}+\vec{c}))$$

$$f(\vec{d}) = (\vec{d}-\vec{c})^T (\vec{d} - \tfrac{1}{2}(\vec{d}+\vec{c}))$$

$$= \tfrac{1}{2} \|\vec{d}-\vec{c}\|_2^2$$

$$\geq 0$$

$$f(\vec{c}) = -\tfrac{1}{2} \|\vec{d}-\vec{c}\|^2$$

$$\leq 0.$$

---

To show: $f(\vec{x}) \geq 0 \quad \forall \vec{x} \in D$.

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

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

$$= (\vec{d} - \vec{c})^T \left(\vec{u} - \vec{d} + \vec{d} - \tfrac{1}{2}(\vec{d} + \vec{c})\right)$$

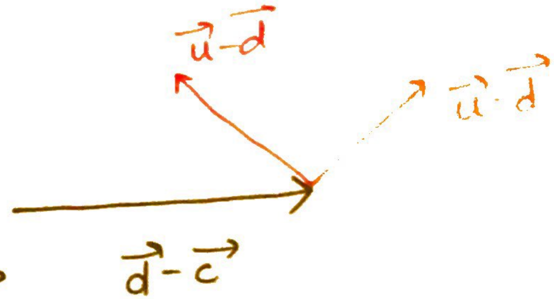
$$= (\vec{d} - \vec{c})^T \left((\vec{u} - \vec{d}) + \tfrac{1}{2}(\vec{d} - \vec{c})\right)$$

$$= (\vec{d} - \vec{c})^T (\vec{u} - \vec{d}) + \underbrace{\tfrac{1}{2}\|\vec{d} - \vec{c}\|_2^2}_{\geq 0} \qquad < 0$$

$f(\vec{u}) < 0$

$$\Longrightarrow \quad (\vec{d} - \vec{c})^T (\vec{u} - \vec{d}) < 0$$

<u>Intuition:</u> I can move along the vector $\vec{u} - \vec{d}$ to get "closer" to $\vec{c}$




<u>Consider:</u> $\vec{p} = \vec{d} + t(\vec{u} - \vec{d}) = t \cdot \vec{u} + (1 - t)\vec{d}$

<u>Consider:</u> $\|\vec{c} - \vec{p}\|_2^2 = \|\vec{c} - \vec{d} - t(\vec{u} - \vec{d})\|_2^2$

$\|\vec{c}-\vec{p}\|_2^2$

$= \|\vec{c}-\vec{d} - t(\vec{u}-\vec{d})\|_2^2$

$= \left((\vec{c}-\vec{d}) - t(\vec{u}-\vec{d})\right)^T \left((\vec{c}-\vec{d}) - t(\vec{u}-\vec{d})\right)$

$= \|\vec{c}-\vec{d}\|_2^2 \quad - 2\langle \vec{c}-\vec{d}, t(\vec{u}-\vec{d})\rangle \quad + t^2 \|\vec{u}-\vec{d}\|_2^2$ $\qquad t \in [0,1]$.

<u>WANT:</u> $2t\langle \vec{d}-\vec{c}, (\vec{u}-\vec{d})\rangle + t^2 \|\vec{u}-\vec{d}\|_2^2 < 0$

$(*)\ 2\langle \vec{d}-\vec{c}, \vec{u}-\vec{d}\rangle + t \cdot \|\vec{u}-\vec{d}\|_2^2 < 0$

Since $\langle \vec{d}-\vec{c}, \vec{u}-\vec{d}\rangle$ is negative, ) We can choose $t$

if we choose $t$, small enough, $(*) \leq 0$.

$\Rightarrow \|\vec{c}-\vec{p}\|_2^2 < \|\vec{c}-\vec{d}\|_2^2$ This contradicts the minimal distance between $C$, $D$.

$\vec{p} \in D \qquad \Rightarrow$ Our assumption was wrong

$\Rightarrow f(\vec{x})$ is a separating hyperplane. $\boxtimes$

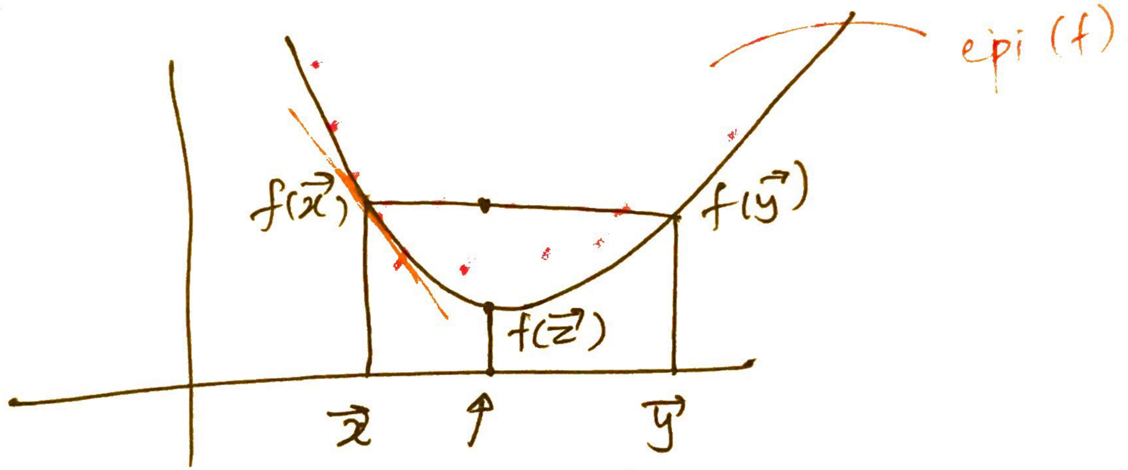
## Convex functions:

$f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if

domain $f$ is a convex set.

$$f(\theta \vec{x} + (1-\theta)\vec{y}) \leq \theta \cdot f(\vec{x}) + (1-\theta) f(\vec{y})$$

Jensens Inequality




$$\theta\vec{x} + (1-\theta)\vec{y} = \vec{z}$$

- **Epigraph:** $\quad \text{Epi } f = \{(x,t) \mid x \in \text{dom } f \; ; \; f(x) \leq t\}$

**Prop 1:** $f$ is a convex function $\Longleftrightarrow$ Epi $f$ is a convex set.

- First-order conditions.

$f$: differentiable. $f: \mathbb{R}^n \to \mathbb{R}$.

Then. $f$ is convex, iff domain$f$ is convex.

and. $$f(\vec{y}) \geq \underbrace{f(\vec{x}) + \nabla f(\vec{x})^T (\vec{y} - \vec{x})}_{\text{Taylor approximation.}} \quad \forall \vec{x}, \vec{y} \in \text{dom}.$$

Implication: If $\nabla f(\vec{x}_*) = 0$

$$\vec{f}(y) \geq f(\vec{x}_*) + 0$$

$\Rightarrow f(\vec{x}_*)$ is a global minimum.

- Second-order condition $f$: twice differentiable.

$f$ is convex iff. dom$f$ is convex.

$$\nabla^2 f(x) \succeq 0$$

Jensen in probability:

$f$ : convex :

$$f(E[x]) \leq E[f(x)]$$