Assignment 4 due today

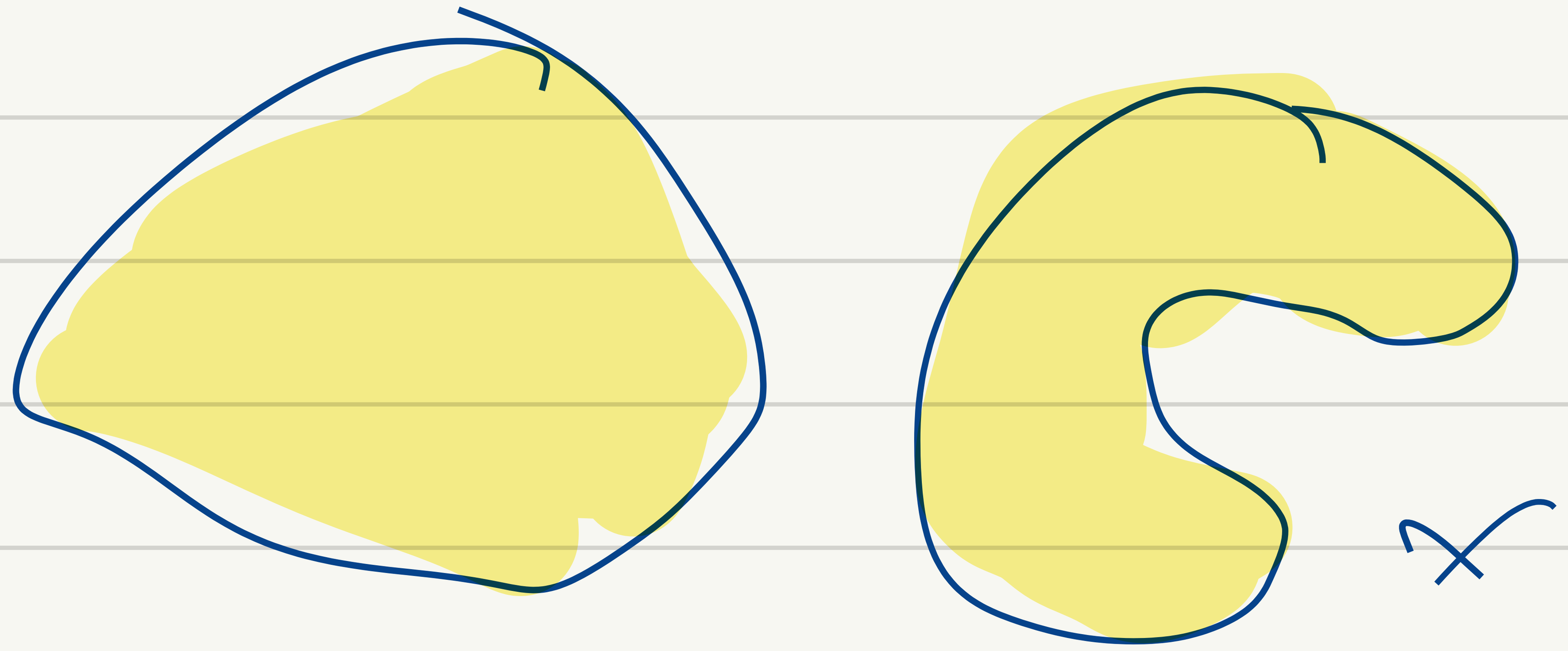
Assignment 5 out

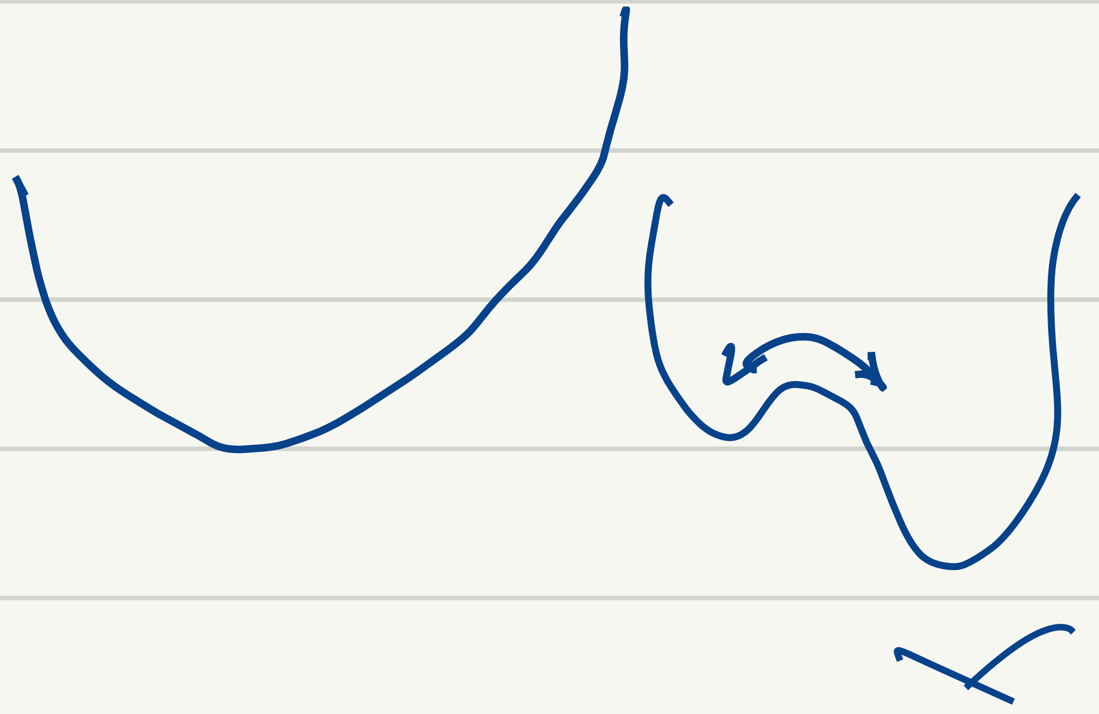
Optimization

→ continuous

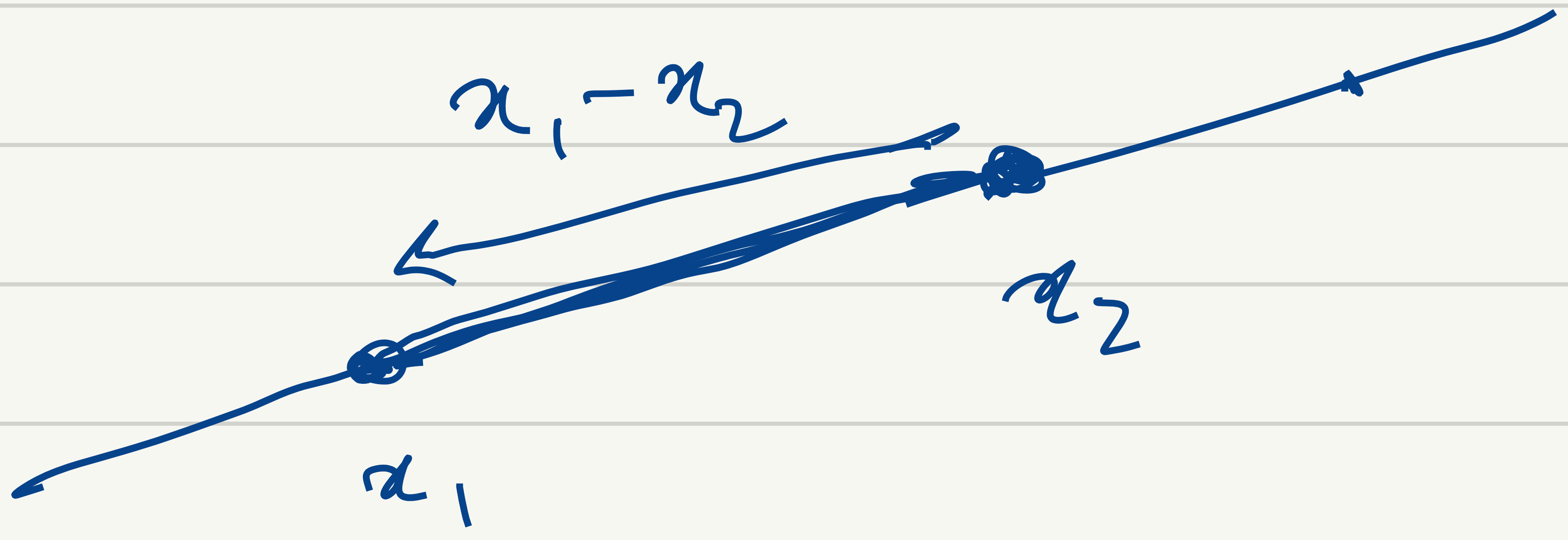
→ unconstrained

→ convex vs. nonconvex





Convex sets



$\vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$

$$\vec{x}_2 + \theta(\vec{x}_1 - \vec{x}_2) = \theta\vec{x}_1 + (1-\theta)\vec{x}_2 \quad \forall\, \theta \in \mathbb{R}$$

line joining $\vec{x}_1, \vec{x}_2$

$y: \theta\vec{x}_1 + (1-\theta)\vec{x}_2\,, \quad 0 \le \theta \le 1$: line segment between $\vec{x}_1, \vec{x}_2$

$\underbrace{y: \theta_1\vec{x}_1 + \theta_2\vec{x}_2}\,, \quad \theta_1, \theta_2 \ge 0$ $\quad \vec{y}$ is a $\boxed{\text{convex combination}}$ of $\vec{x}_1$ and $\vec{x}_2$

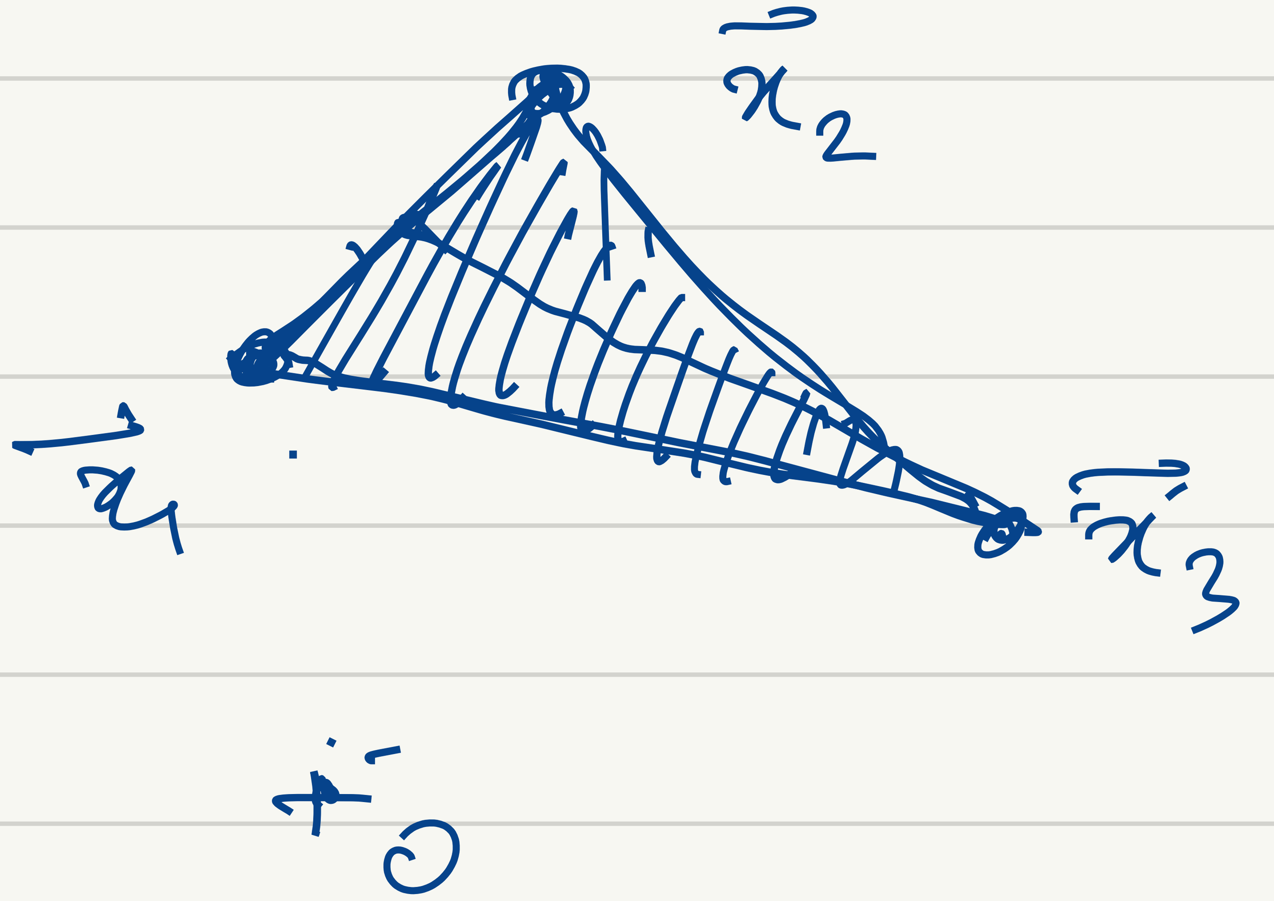
$\theta_1 + \theta_2 = 1$

n points $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$: convex comb.: $\vec{y} = \sum \theta_i \vec{x}_i$ s.t. $\theta_i \geq 0$

$$\sum \theta_i = 1$$

$$\equiv \theta_1 \vec{x}_1 + (1-\theta_1)\vec{z}$$ where $\vec{z}$ is c.c. of $\vec{x}_2, \ldots, \vec{x}_n$

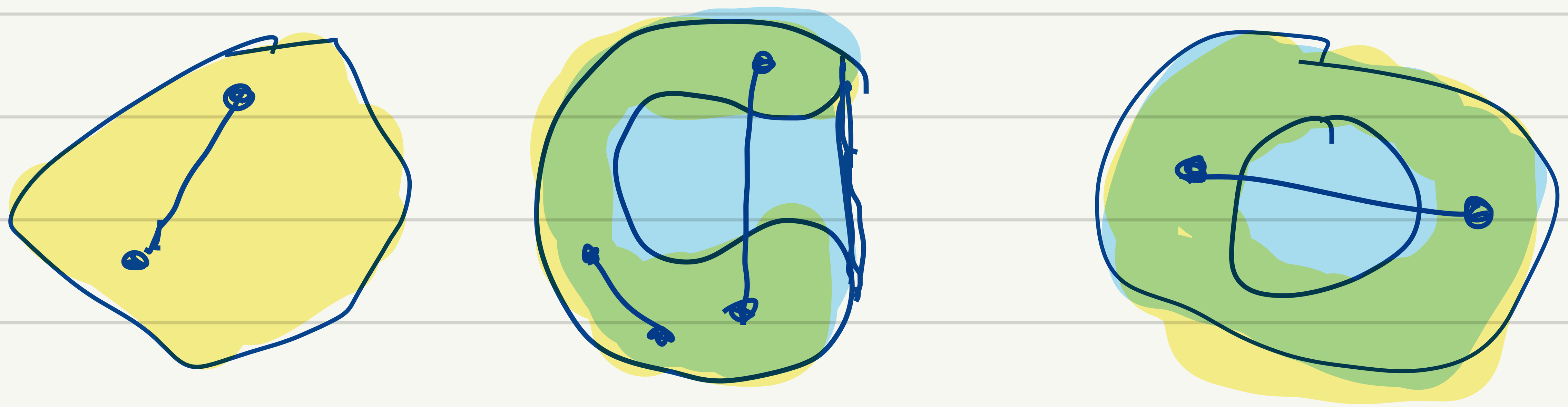
e.g. $\theta_i = \frac{1}{n} \Rightarrow \vec{y} = \dfrac{\vec{x}_1 + \cdots + \vec{x}_n}{n}$



weighted average: $\vec{y} = \dfrac{w_1\vec{x}_1 + \cdots + w_n\vec{x}_n}{w_1 + \cdots + w_n}$

$\boxed{\text{Convex set}}$: set that contains <u>all convex combinations</u> of its elements

$$\vec{x}, \vec{y} \in S \Rightarrow \theta\vec{x} + (1-\theta)\vec{y} \in S$$



$\boxed{\text{Convex hull}}$: set of all c.c. of points in S

conv(S)

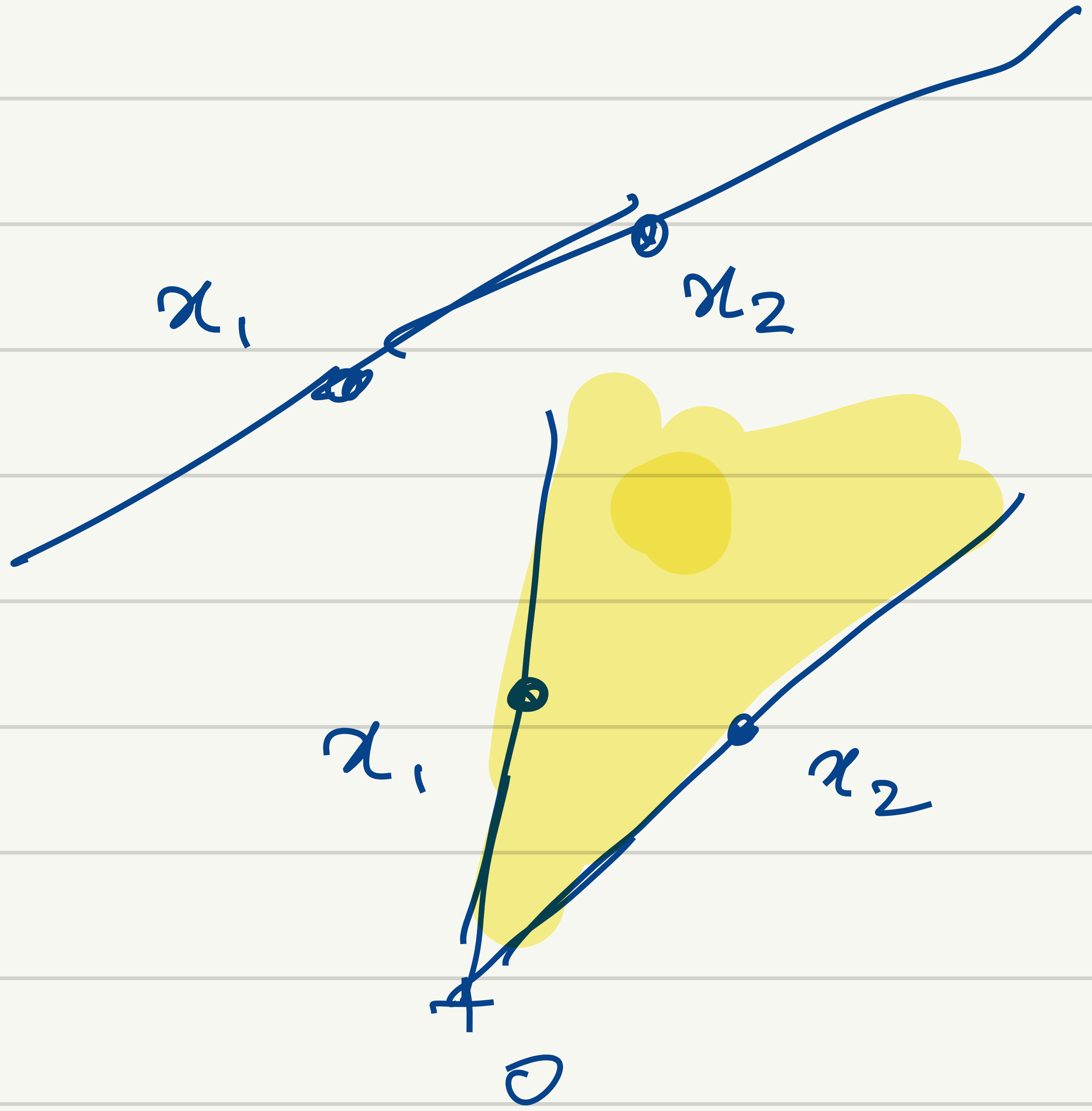
Convex hull = smallest convex set $\supseteq S$

lin. comb. $\sum \theta_i \vec{x}_i$, $\theta_i \in \mathbb{R}$

convex comb. $\sum \theta_i \vec{x}_i$, $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_n = 1$

affine comb. $\sum \theta_i \vec{x}_i$ s.t. $\theta_1 + \cdots + \theta_n = 1$

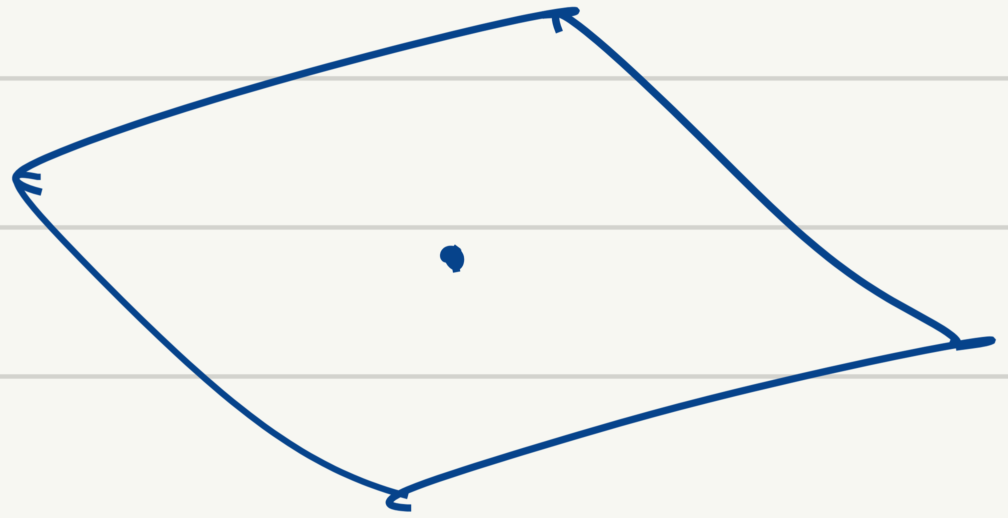
conic comb. $\sum \theta_i \vec{x}_i$ s.t. $\theta_i \geq 0$

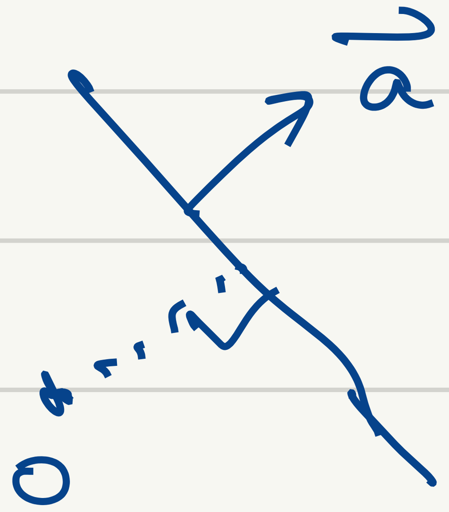


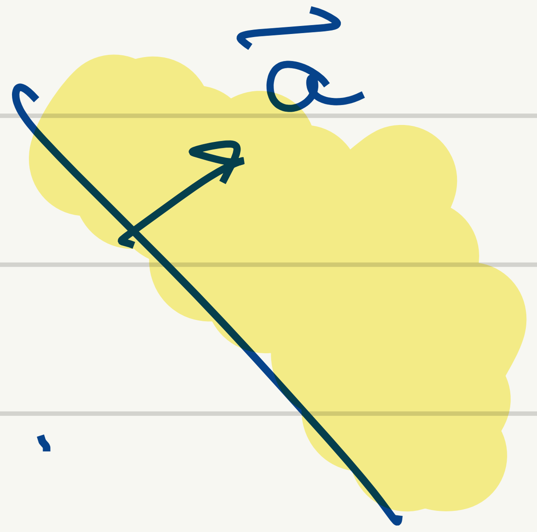
Examples and operations

hyperplane $\{\vec{x} : \vec{a}^T \vec{x} = b\}$

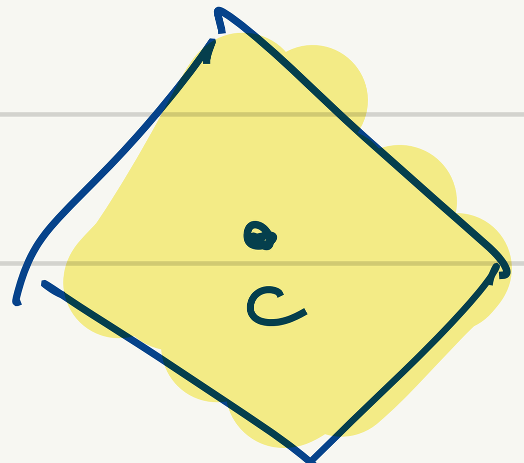
half-space $\{\vec{x} : \vec{a}^T \vec{x} \geq b\}$





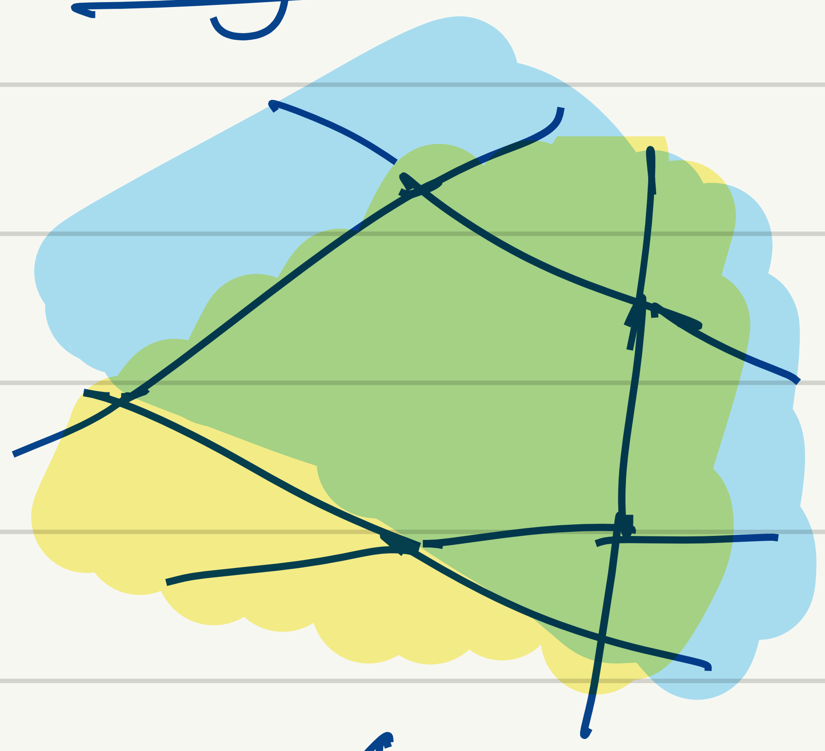


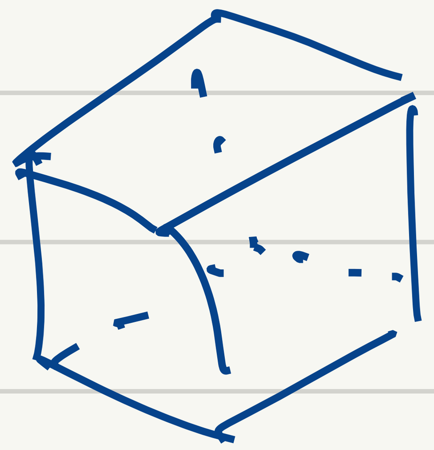
Norm ball $\{\vec{x} : \|\vec{x} - \vec{c}\| \leq r\}$



Convex

Polyhedron = intersection of hyperplanes & half-spaces



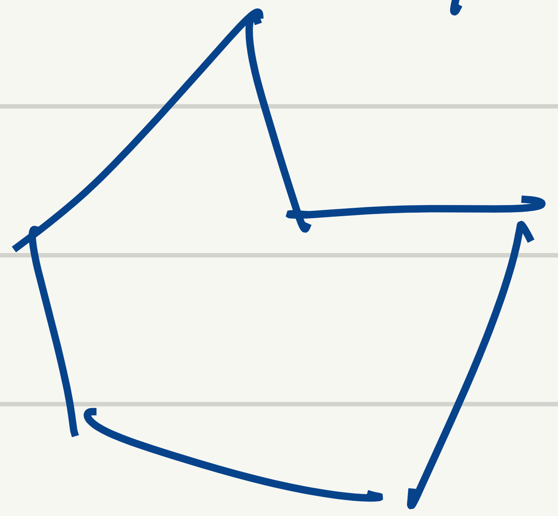


Set of all symmetric $n \times n$ matrices S^n

symm. pos. def. matrices S^n_{++}

" " semidef. " S^n_{+}

convex

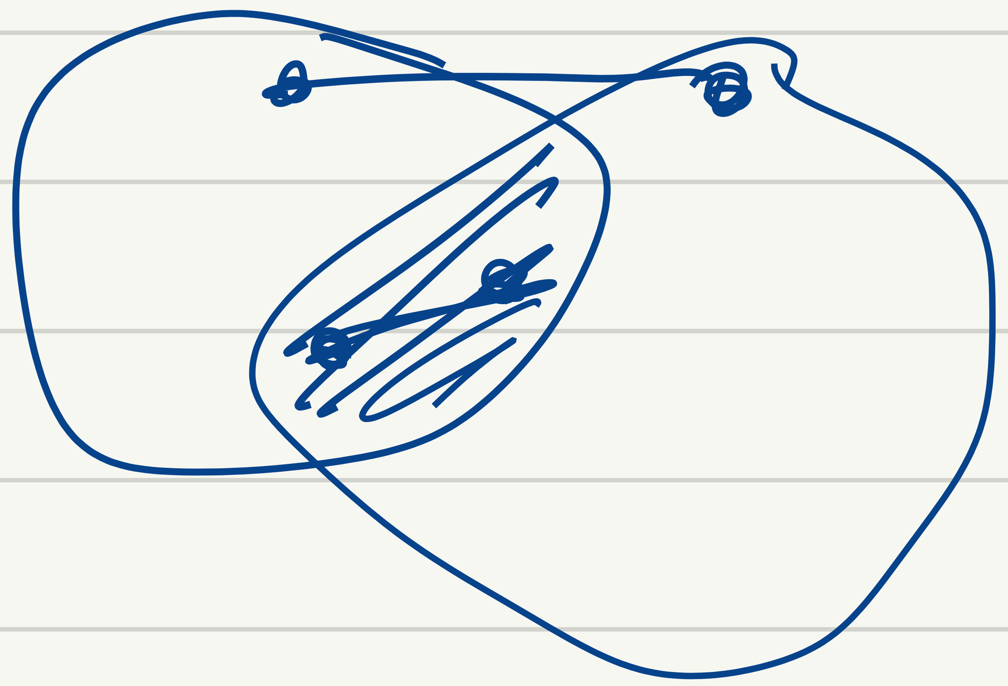


×

A, B s.p.d $\Rightarrow \theta_1 A + \theta_2 B$ s.p.d if $\theta_1, \theta_2 \geq 0$

C_1, C_2 convex sets

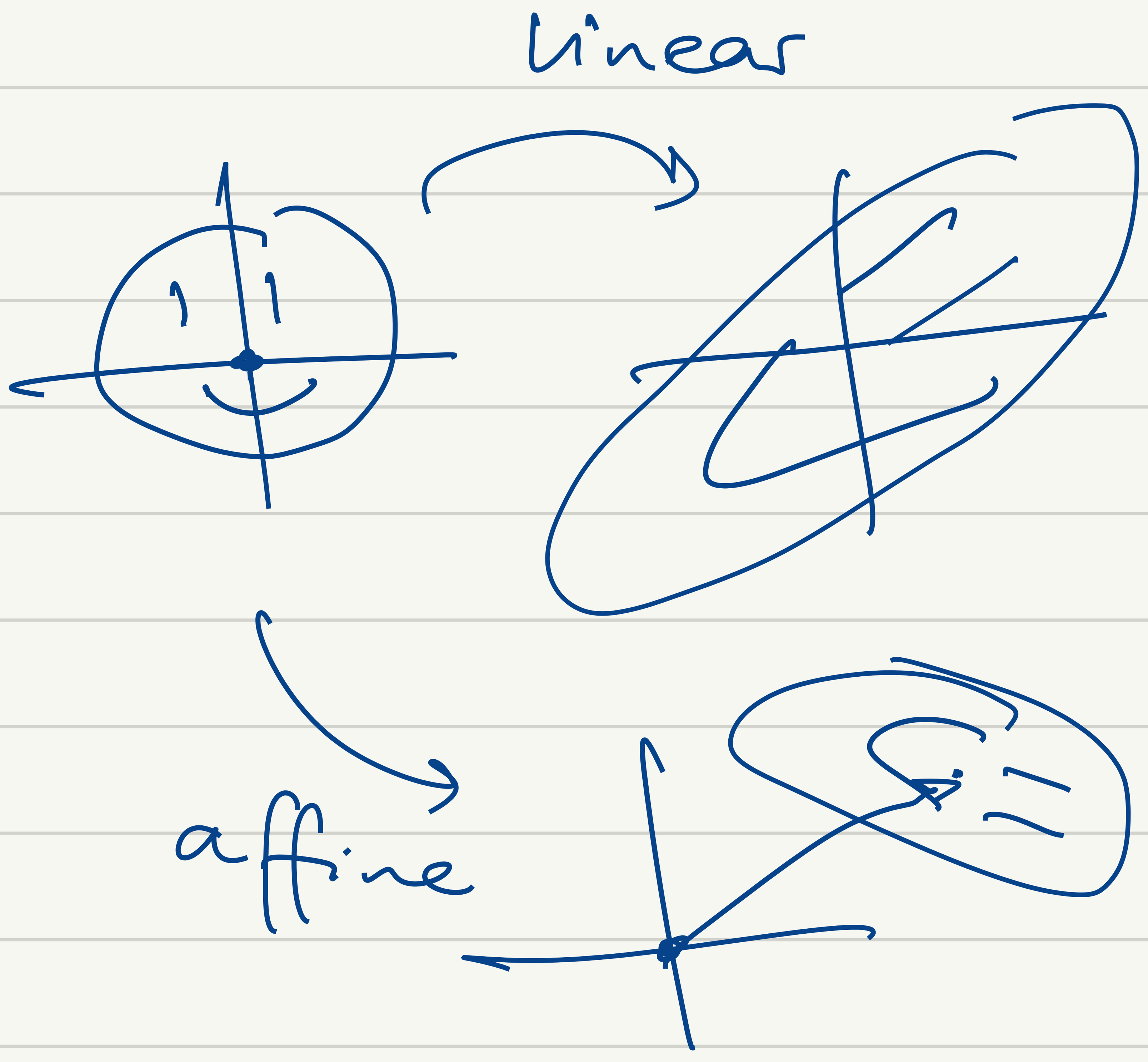
- $C_1 \cap C_2$ convex
- $C_1 \cup C_2$ not convex



Affine function : $f(\vec{x}) = A\vec{x} + \vec{b}$

$\boxed{\text{image}}$

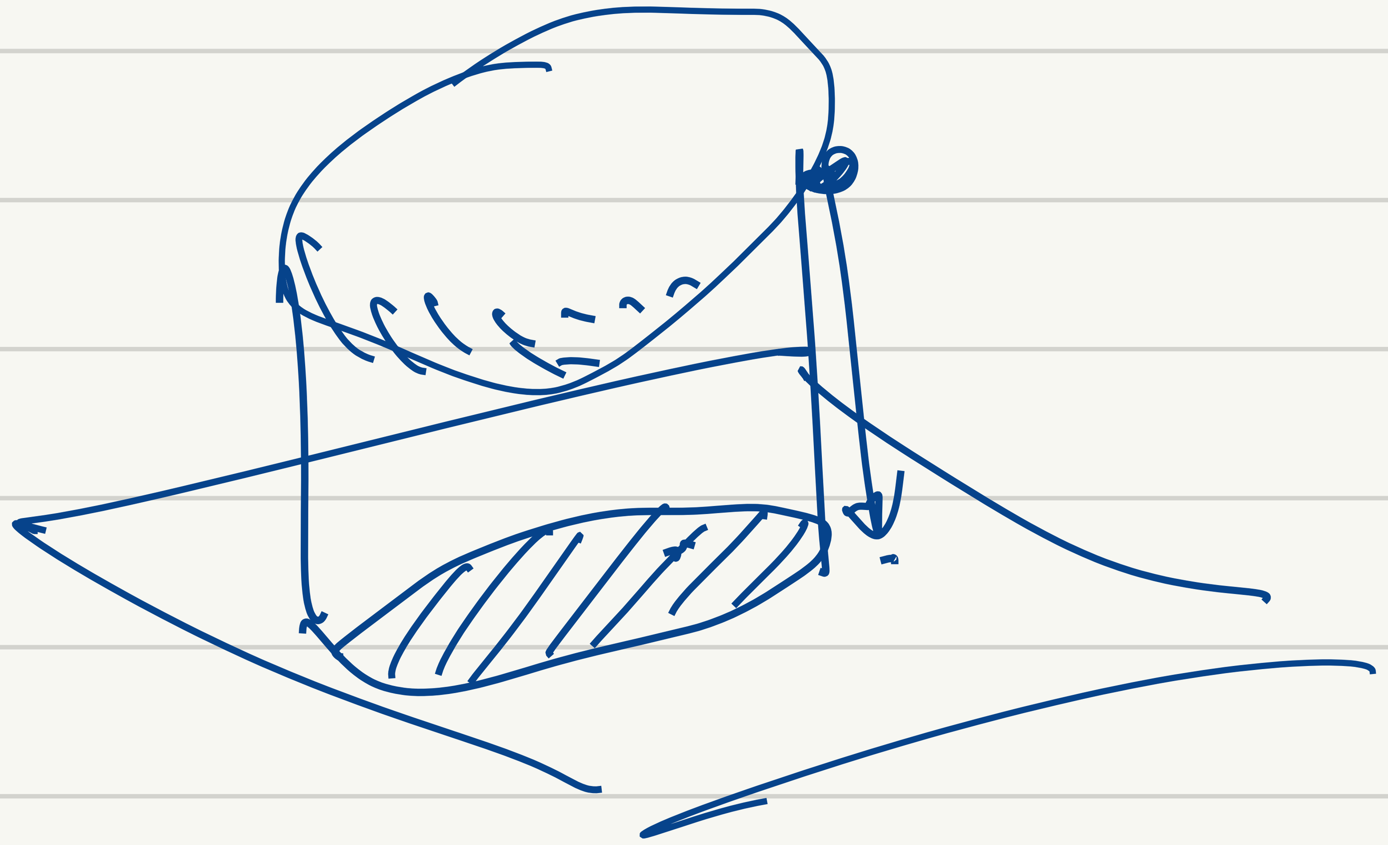
C convex, f affine $\Rightarrow$ $\{ f(\vec{x}) : \vec{x} \in C \}$ convex



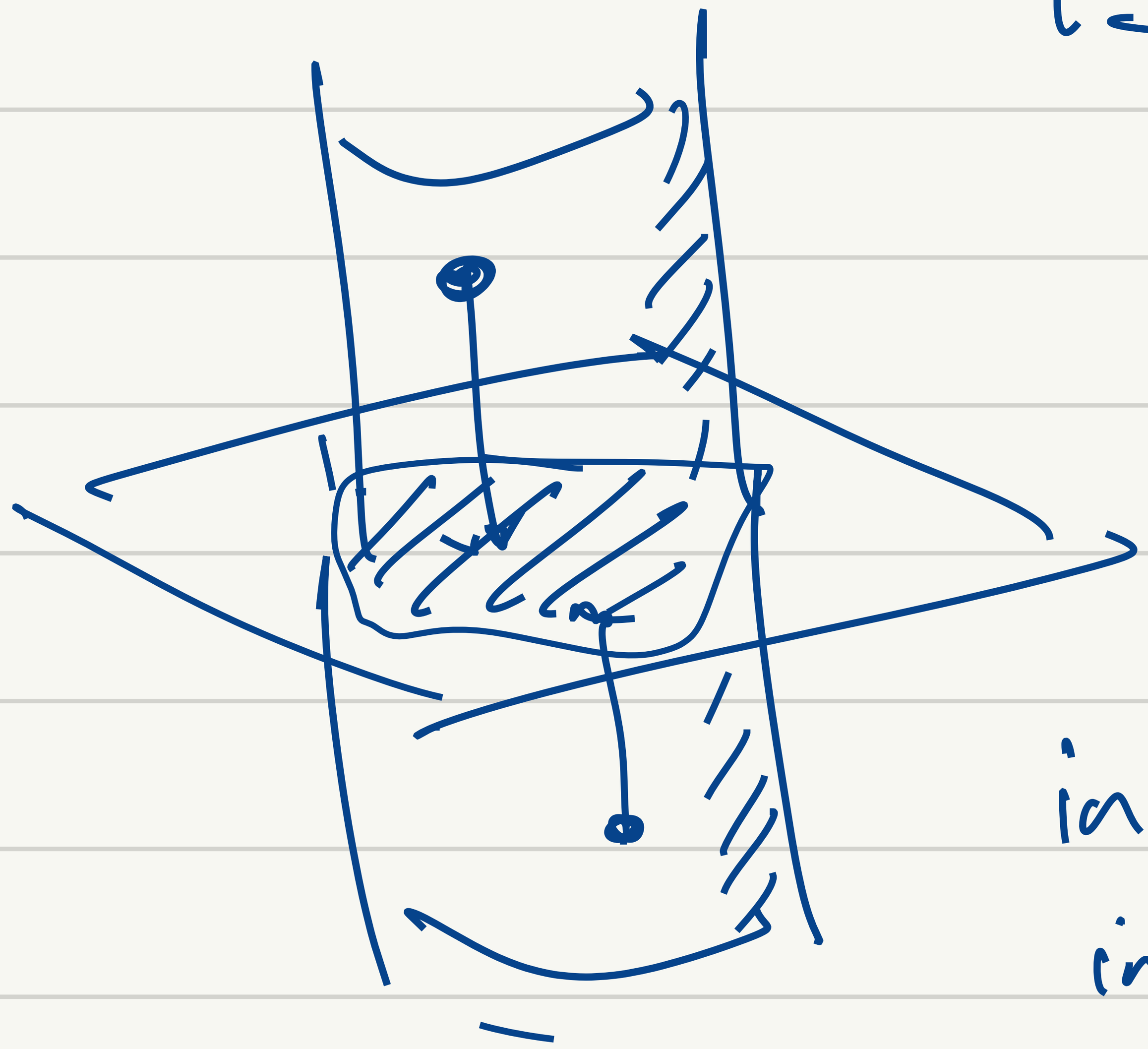
$\{ \vec{x} : f(\vec{x}) \in C \}$ convex

$\boxed{\text{inverse image}}$

$f(x,y,z) = (x,y,0)$



image



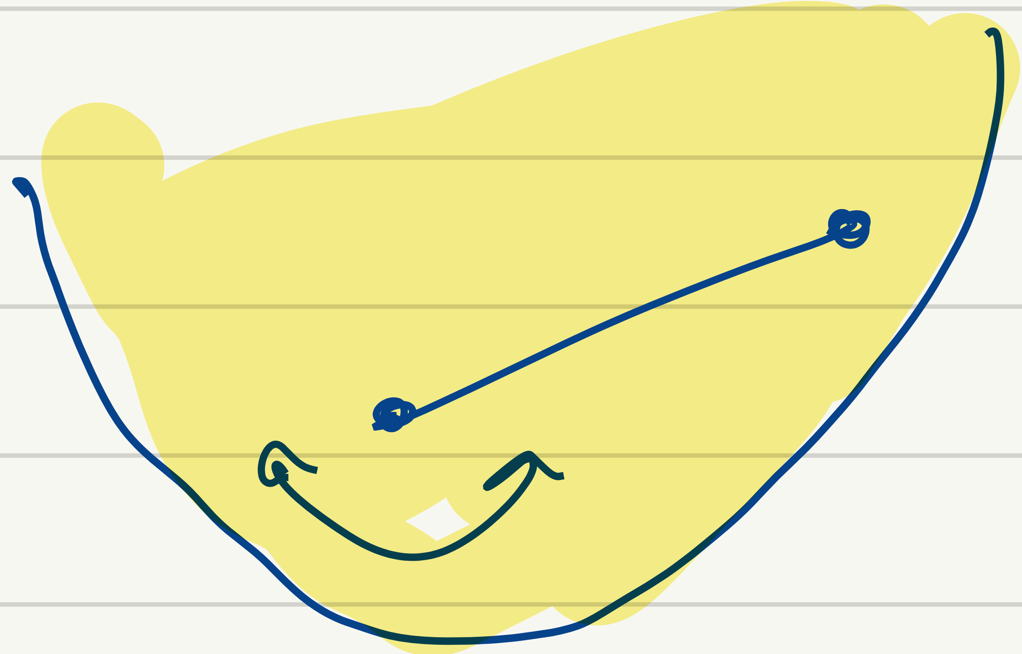
inverse image

$S_1 = \{ (\vec{x},\vec{y}) : \vec{x},\vec{y} \in \mathbb{R}^n, \|\vec{x}-\vec{y}\| \le r \}$

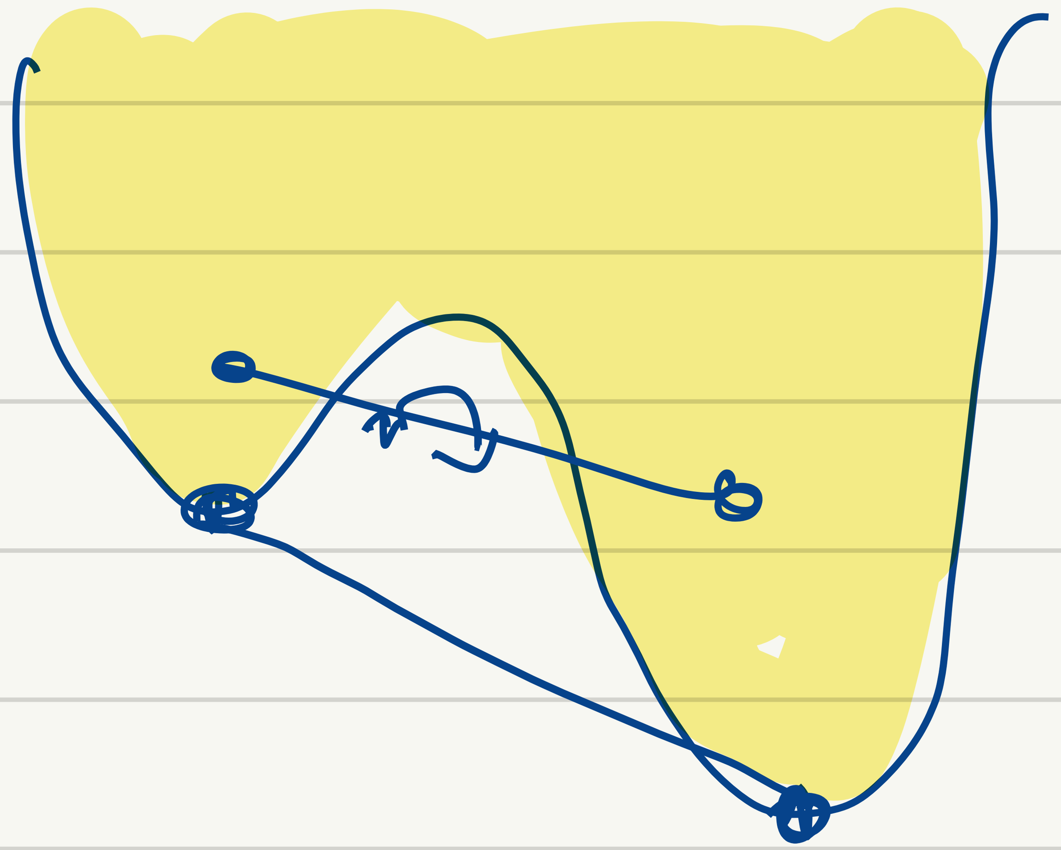
$\vec{z} = \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix}$ $\quad A\vec{z} = \begin{bmatrix} I & -I \end{bmatrix} \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} = \vec{x} - \vec{y}$

$S_2 = \{ (\vec{x},\vec{y}) : \|\vec{x}-\vec{y}\| = r \}$ convex?

Convex functions



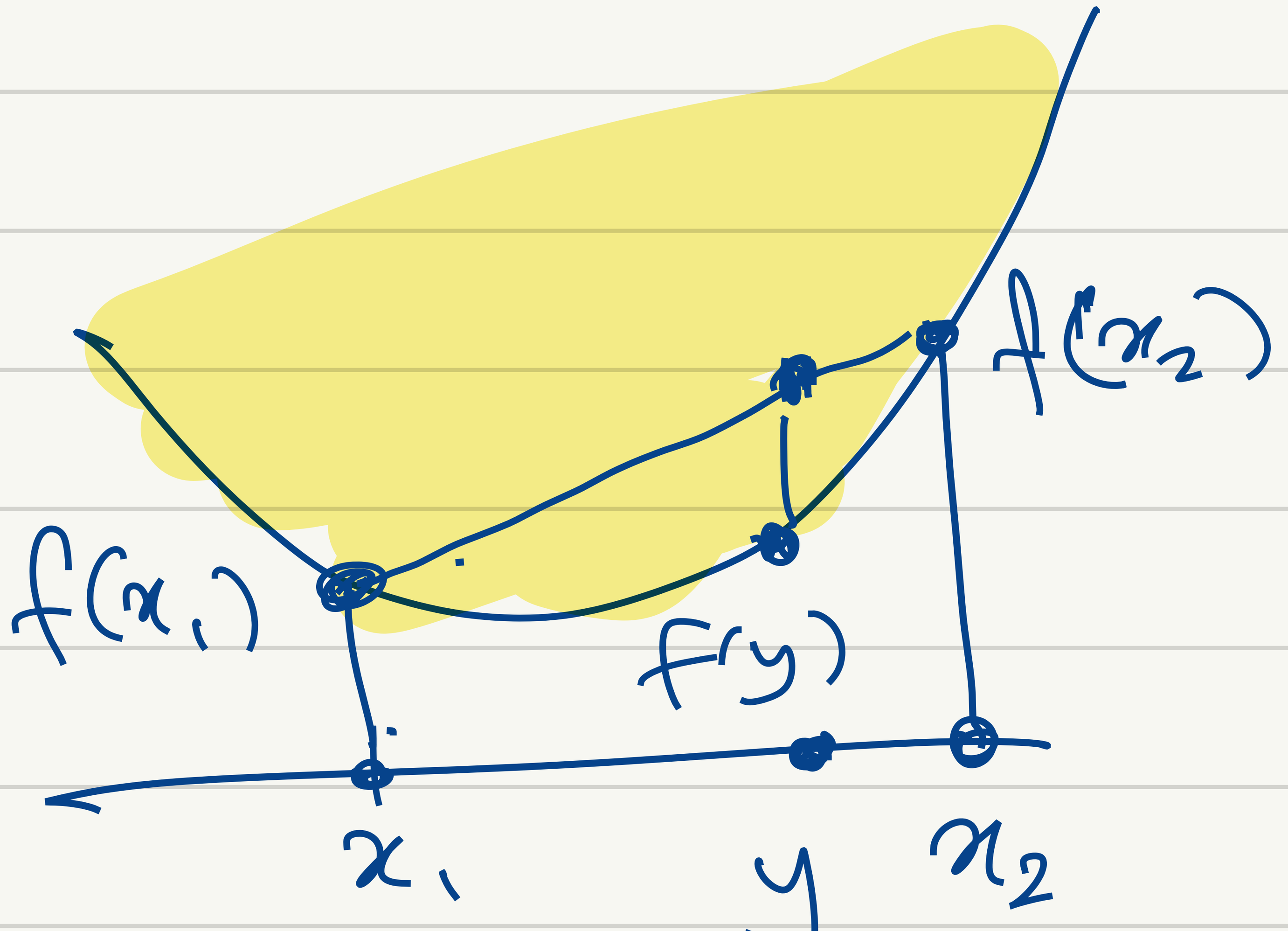
$f''(x) \geq 0$?



$\boxed{\text{Epigraph}}$ of function :

$$\text{epi}(f) = \{(\vec{x}, y) : y \geq f(\vec{x})\}$$

$f : \mathbb{R}^n \to \mathbb{R}$ then $\text{epi}(f) \subseteq \mathbb{R}^{n+1}$



f is a $\boxed{\text{convex function}}$ $\iff$ $\text{epi}(f)$ is convex set.

$\iff$ line segment between $(x_1, f(x_1))$, $(x_2, f(x_2))$ is in $\text{epi}(f)$

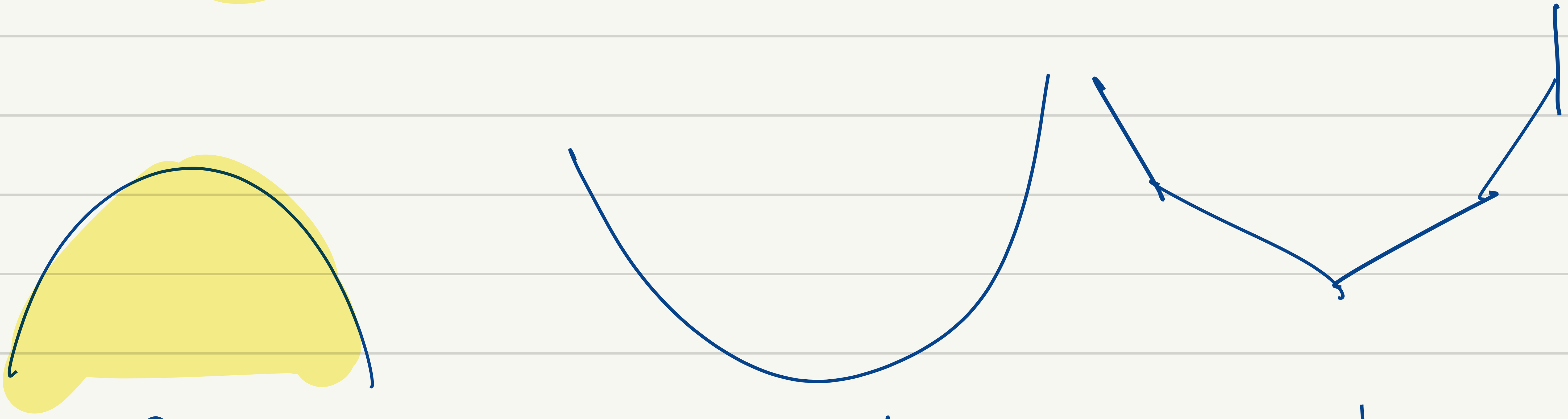
$$\iff \underbrace{f(\theta x_1 + (1-\theta)x_2)}_{\vec{y}} \leq \underbrace{\theta f(x_1) + (1-\theta) f(x_2)}$$

$$f(\theta_1 \vec{x}_1 + \cdots + \theta_n \vec{x}_n) \le \theta_1 f(\vec{x}_1) + \cdots + \theta_n f(\vec{x}_n)$$: **Jensen's inequality**

$$f(E[\vec{x}]) \le E[f(\vec{x})]$$

Strict convexity : $f(\theta x_1 + (1-\theta) x_2) < \theta f(x_1) + (1-\theta) x_2 \quad \forall \; 0 < \theta < 1$

f is **concave** if $-f$ is convex



Concave

strictly convex

not strictly convex

1st- and 2nd-order conditions

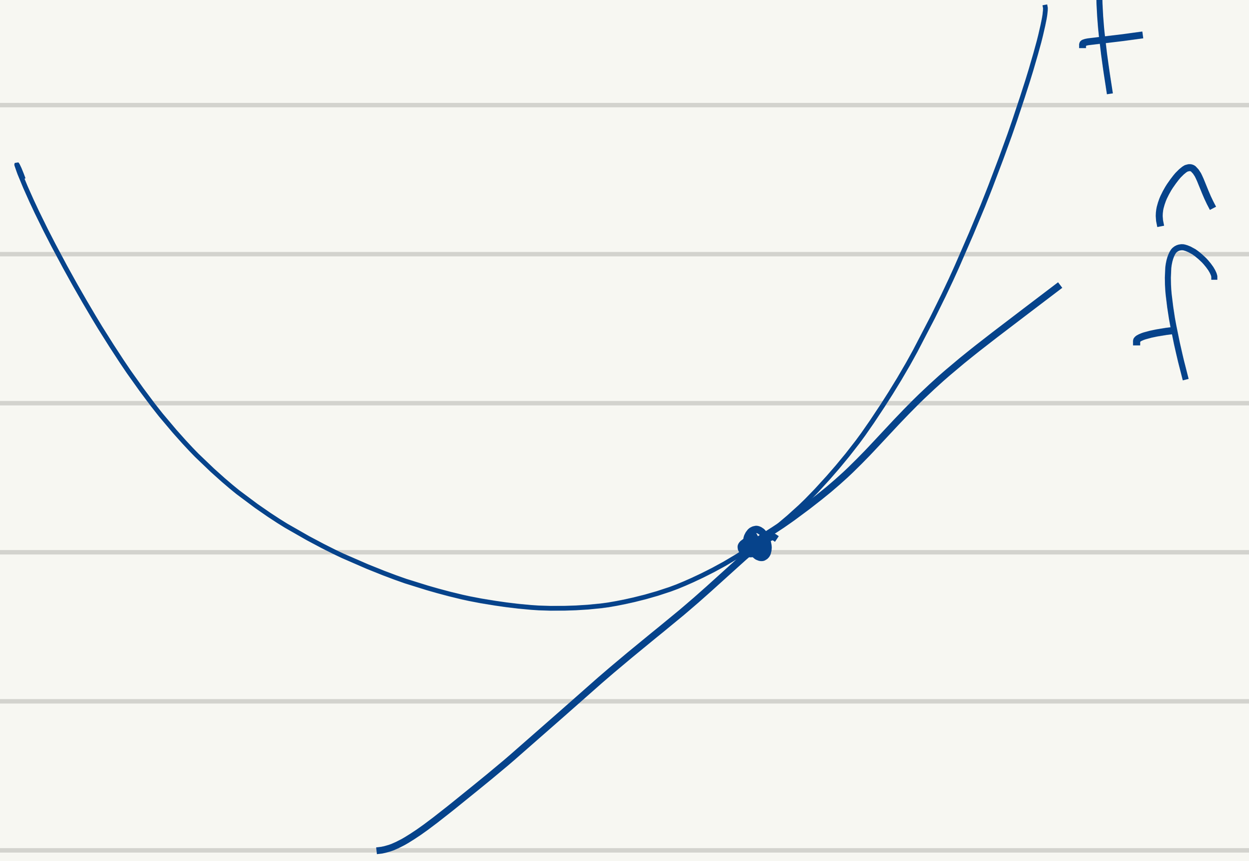
$f : \mathbb{R}^n \to \mathbb{R}$

$\nabla f : \mathbb{R}^n \to \mathbb{R}^n$

$$\nabla f(x) = \begin{bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \vdots \end{bmatrix}$$

$$f(x+h) = f(x) + \nabla f(x)^T h + o(\|h\|^2)$$

$\hat{f}(y) = f(x) + \nabla f(x)^T (y-x)$



f is convex iff $f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall\, x, y$

$\Rightarrow$ If $\nabla f(x) = 0$ then $f(y) \geq f(x) \quad \forall\, y$

$\Rightarrow$ x is global minimizer!

Suppose f is twice differentiable

$\boxed{\text{Hessian}}$ matrix $\nabla^2 f(x) = \begin{bmatrix} \partial^2 f / \partial x_1^2 & \partial^2 f / \partial x_1 \partial x_2 & \cdots \\ \partial^2 f / \partial x_1 \partial x_2 & \partial^2 f / \partial x_2^2 & \\ \vdots & & \ddots \end{bmatrix}$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

f is convex $\iff$ $\nabla^2 f(\vec{x}) \succeq 0$ for all $\vec{x}$

$\nabla^2 f(\vec{x}) \succ 0 \Rightarrow f$ is strictly convex

not $\Leftarrow$: $f(x) = x^4$

Examples & operations

All linear fns $f(\vec{x}) = A\vec{x}$

affine fns $f(\vec{x}) = A\vec{x} + \vec{b}$

$\Big\}$ convex

Quadratic fn $f(\vec{x}) = \frac{1}{2} \vec{x}^T P \vec{x} + \vec{q}^T \vec{x} + r$ (P symmetric)

is convex $\iff$ $P \succeq 0$

Notation:

$A \in \mathbb{R}^{n \times n}$

$A \succ 0$ if A is pos. def $\quad A \in S^n_{++}$

$A \succeq 0$ if pos. semidef. $\quad A \in S^n_{+}$

$A \succeq B \iff A - B \succeq 0$

$\vec{x} \in \mathbb{R}^n$

$\vec{x} \succeq 0$ if $x_i \geq 0 \ \forall i$

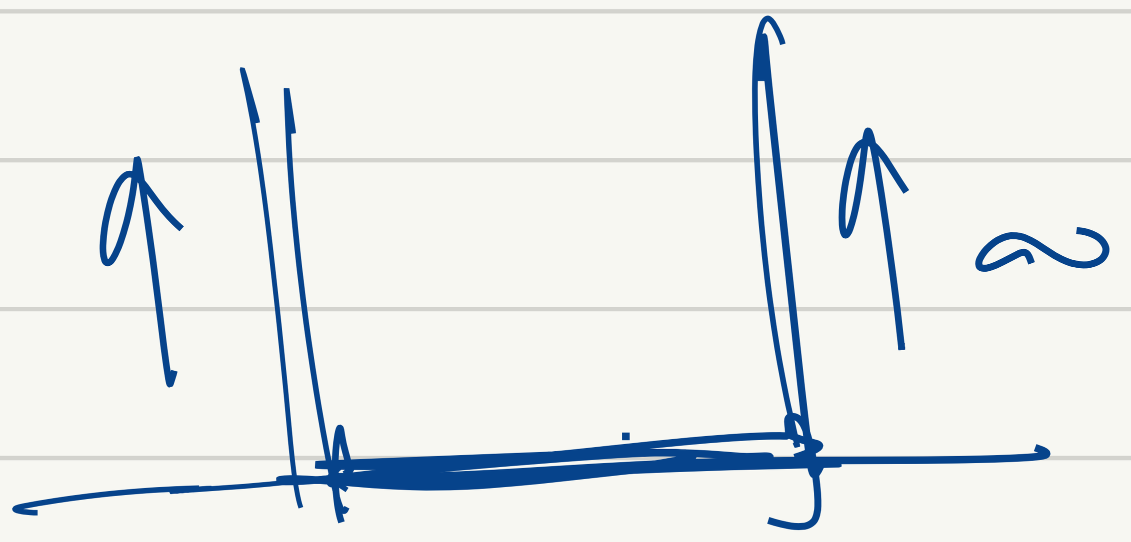
$f(\vec{x}) = \frac{1}{2}\vec{x}^T P\vec{x} + \vec{q}^T\vec{x} + r$

$g(\vec{x}) = \frac{1}{2}\|A\vec{x} - b\|_2^2$

$\nabla f(\vec{x}) = P\vec{x} + q$

find $\nabla g(\vec{x})$, $\nabla^2 g(\vec{x})$

$\nabla^2 f(\vec{x}) = P$

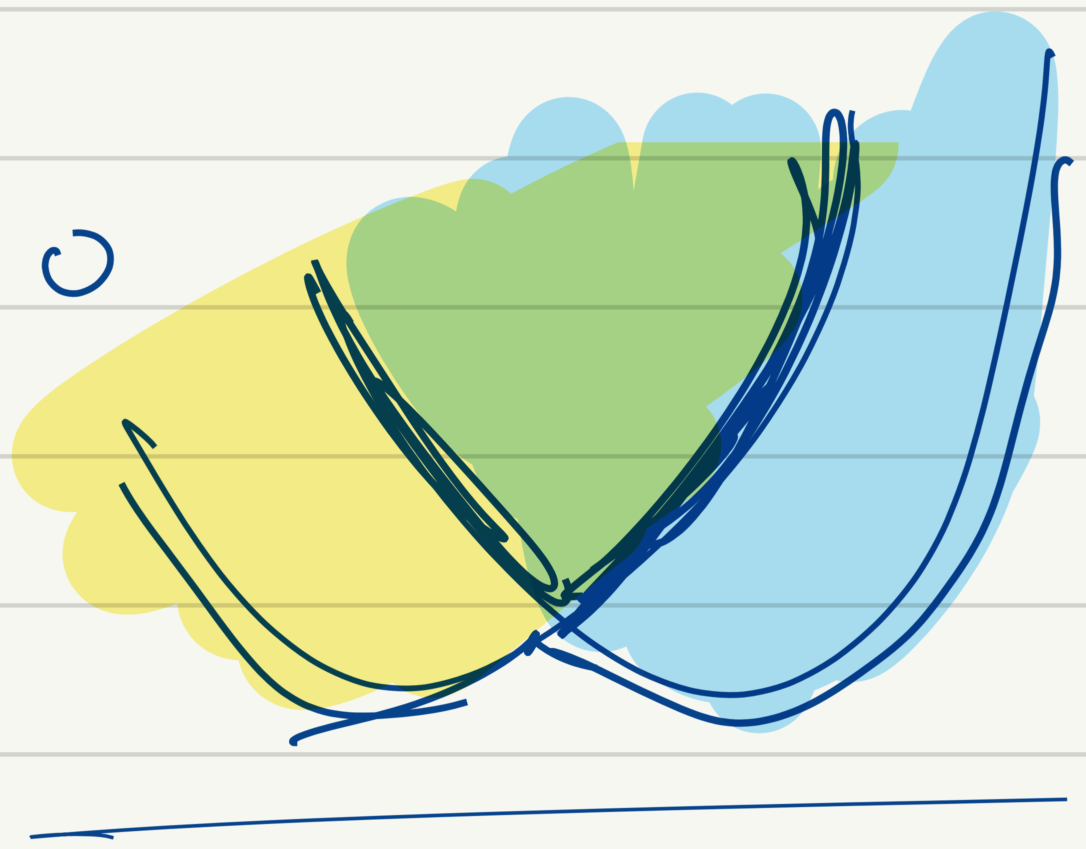


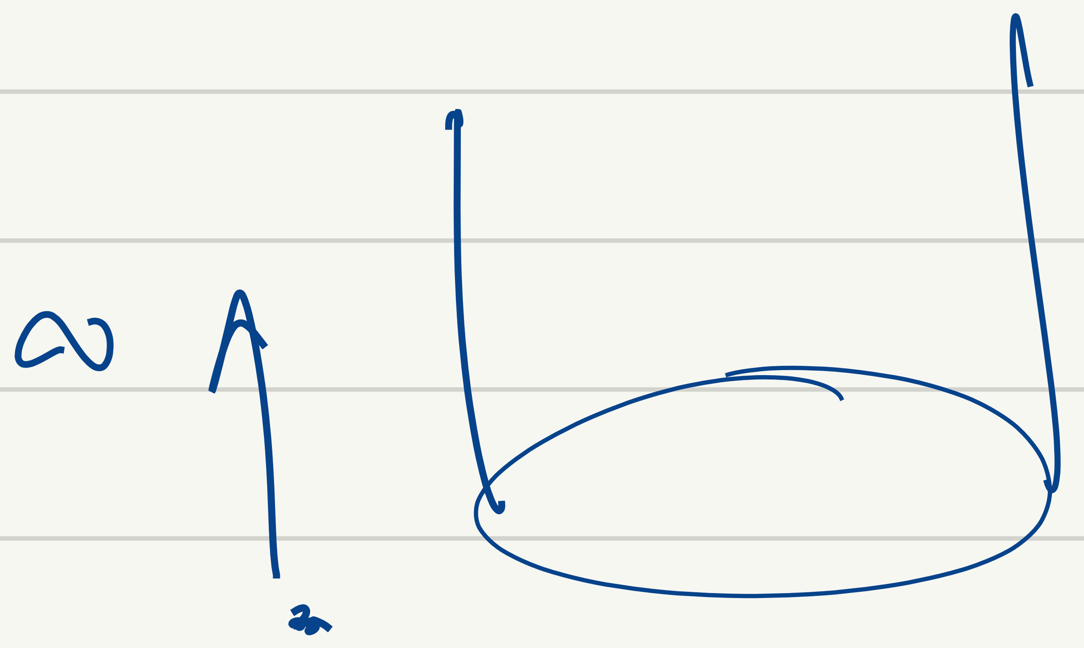
All norms are convex functions $f(\vec{x}) = \|\vec{x}\|$

If C is a convex set, $\boxed{\text{indicator function}}$ $I_C(\vec{x}) = \begin{cases} 0 & \text{if } \vec{x} \in C \\ \infty & \text{if } \vec{x} \notin C \end{cases}$

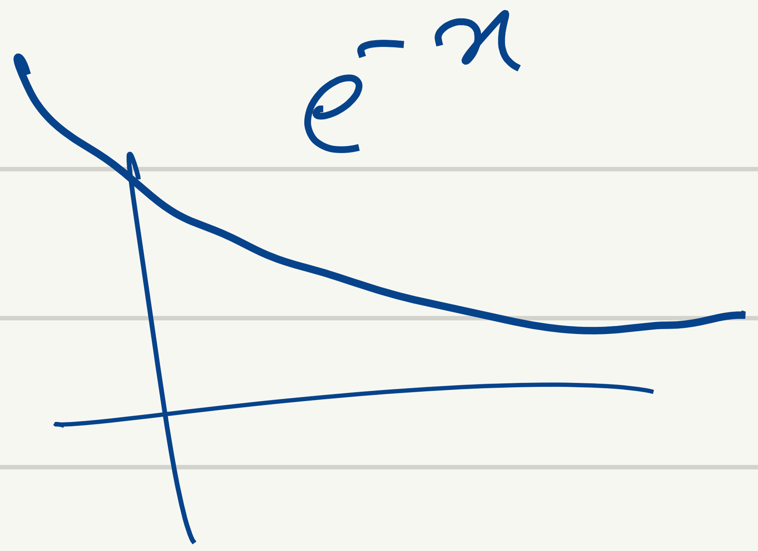
f, g are convex functions $\Rightarrow$

- $w_1 f + w_2 g$ is convex if $w_1, w_2 \geq 0$
- $\max(f(x), g(x))$ is convex
- $\min(f(x), g(x))$ is not



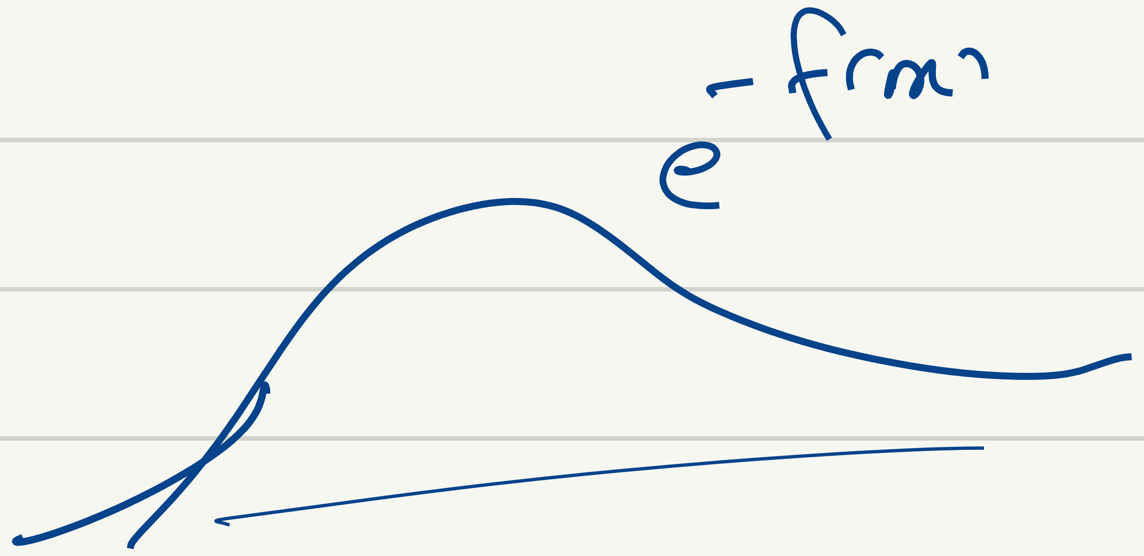


$f : \mathbb{R}^n \to \mathbb{R}$

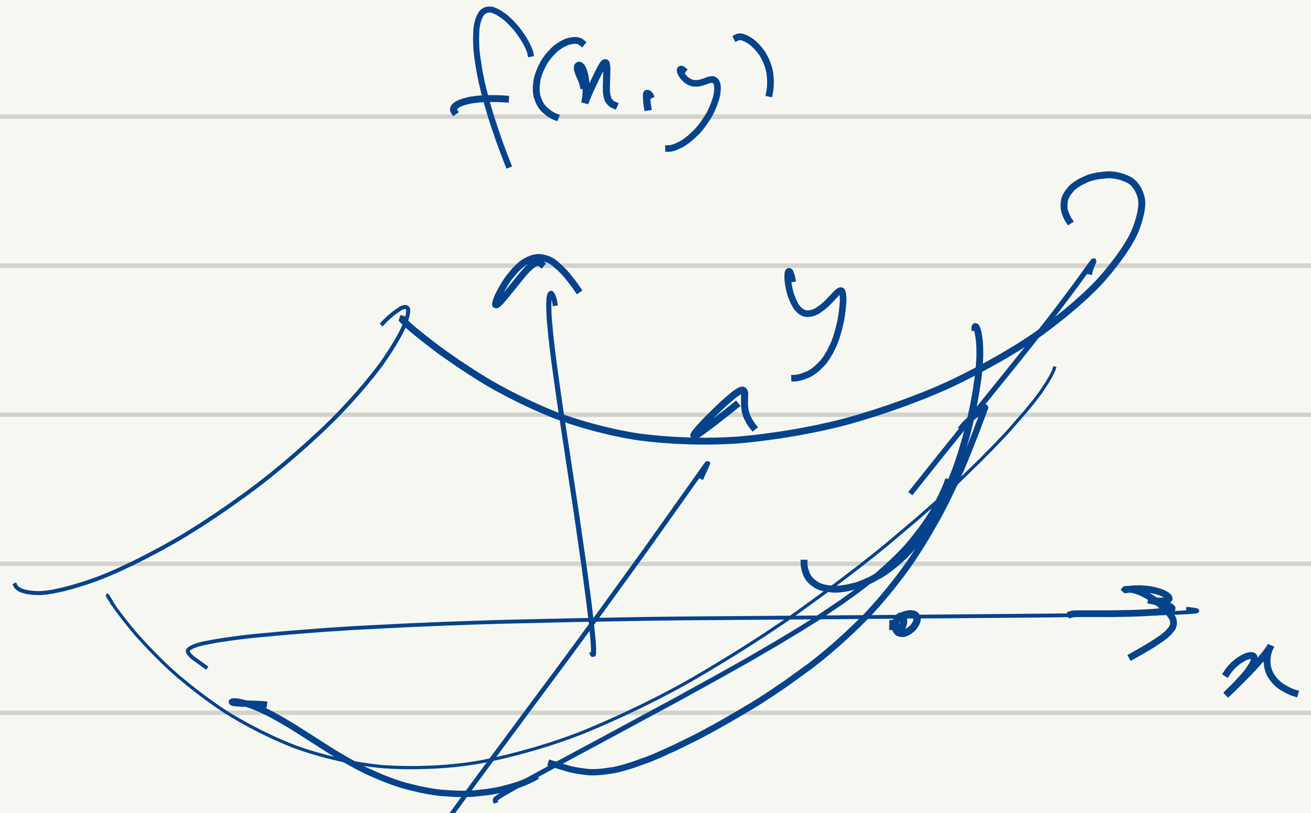


- $f(A\vec{x} + \vec{b})$ is convex
- $g(f(x))$ is convex if g is convex and nondecreasing on range(f)

eg. $e^{f(x)}$, $f(x)^2$ is convex
if $f(x) \geq 0$



Minimize over some variables



$f(x,y)$ is convex, $\quad g(x) = \inf\limits_{y \in C} f(x,y)$

C is convex set $\quad$ is also convex!

$f(x_1, x_2, \dots, x_k, \dots, x_n)$, $g(x_1, \dots, x_k) = \inf\limits_{x_k, \dots, x_n} f(x_1 \dots x_n)$