Announcements

- Minor : Monday 14 th Feb

  Syllabus = everything (including next week)

- Assignment 2 : due ~~10 Feb~~ 14 Feb ?

Stability of LU factorization

$$\tilde{L}\tilde{U} = A + \underbrace{\delta A}$$

$$\frac{\|\delta A\|}{\|L\|\|U\|} = O(\varepsilon_m)$$

exact LU factorization of A

If $\|L\|\|U\| = O(\|A\|)$ then backward stable

$\|L\| = O(1)$ because $\underbrace{|l_{ij}| \leq 1}$

QR factorization:

$$\tilde{Q}\tilde{R} = A + \delta A$$

$$\frac{\|\delta A\|}{\|A\|} = O(\varepsilon_m)$$

Can $\|U\|$ be $\gg \|A\|$?

$$A = \begin{bmatrix} 1 & & & & 1 \\ -1 & 1 & & & 1 \\ -1 & -1 & 1 & & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -1 & -1 & \cdots & -1 & 1 \end{bmatrix} \longrightarrow U = \begin{bmatrix} 1 & & & & 1 \\ & 1 & & & 2 \\ & & 1 & & 4 \\ & & & \ddots & \vdots \\ & & & & 2^{m-1} \end{bmatrix}$$

$\|U\| = O(2^m)$

Mathematically: stable

Practically: useless!

Can fix by smarter pivoting strategies

or switch to QR $\sim 4/3\, m^3$ flops

Cholesky factorization

$A = \underbrace{LL^*}_{} = \underbrace{R^*R}_{}$: L lower tri, R upper tri

only exists for Hermitian positive definite matrices

What is a <u>Hermitian matrix</u>? $A^* = A$

$\boxed{\text{Quadratic form}}$ $q : \mathbb{R}^m \to \mathbb{R}$

$$q(\vec{x}) = \sum_i a_{ii} x_i^2 + \sum\sum_{i<j} 2 a_{ij} x_i x_j$$

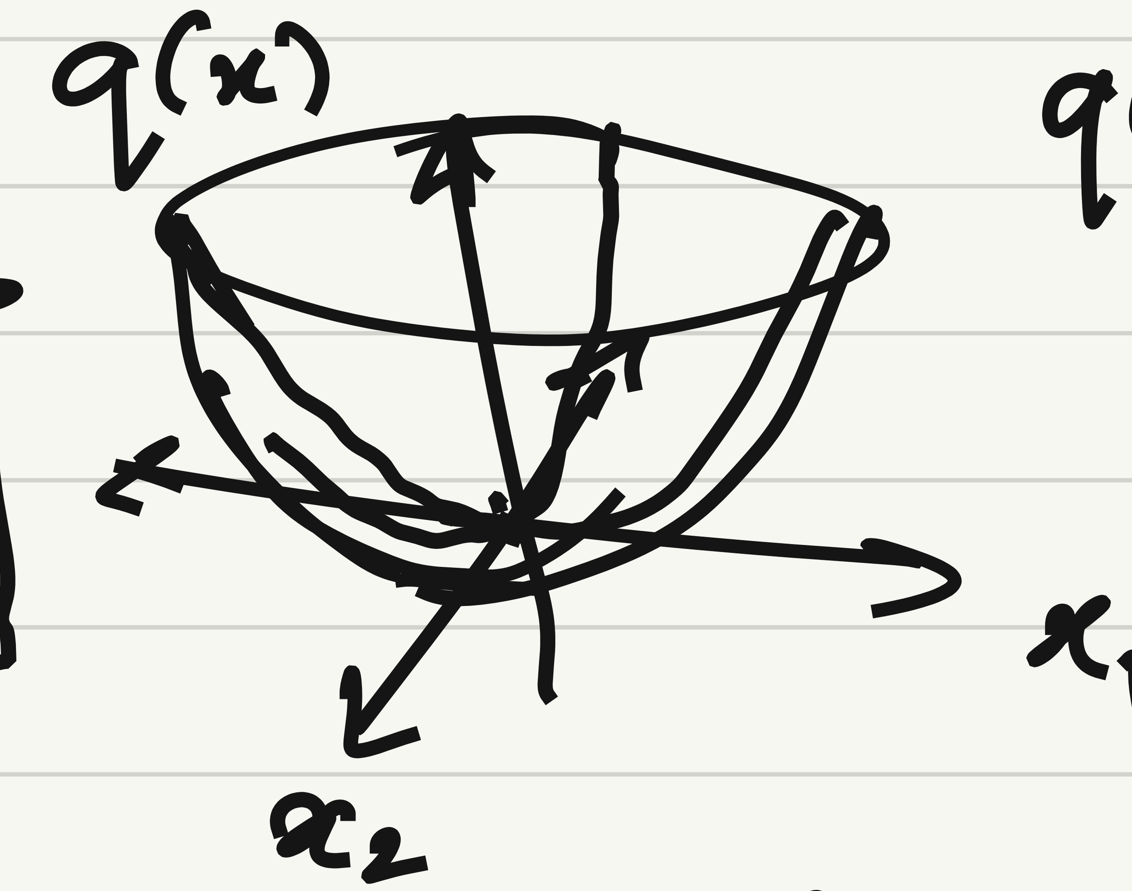
$$q(\vec{x}) = \vec{x}^* A \vec{x} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{12} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}$$

$q : \mathbb{C}^m \to \mathbb{R}$ <u>real quadratic form</u>

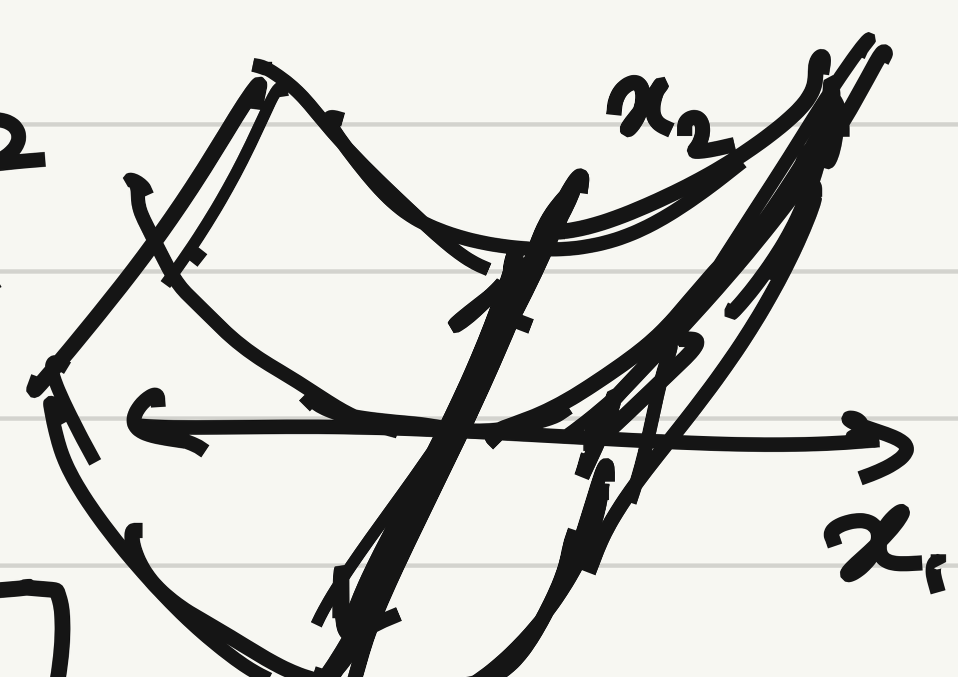
$q(\vec{x}) = \underbrace{\vec{x}^* A \vec{x}}$ where A is Hermitian

$\boxed{\text{positive definite}}$



$q(x) = \|x\|^2$

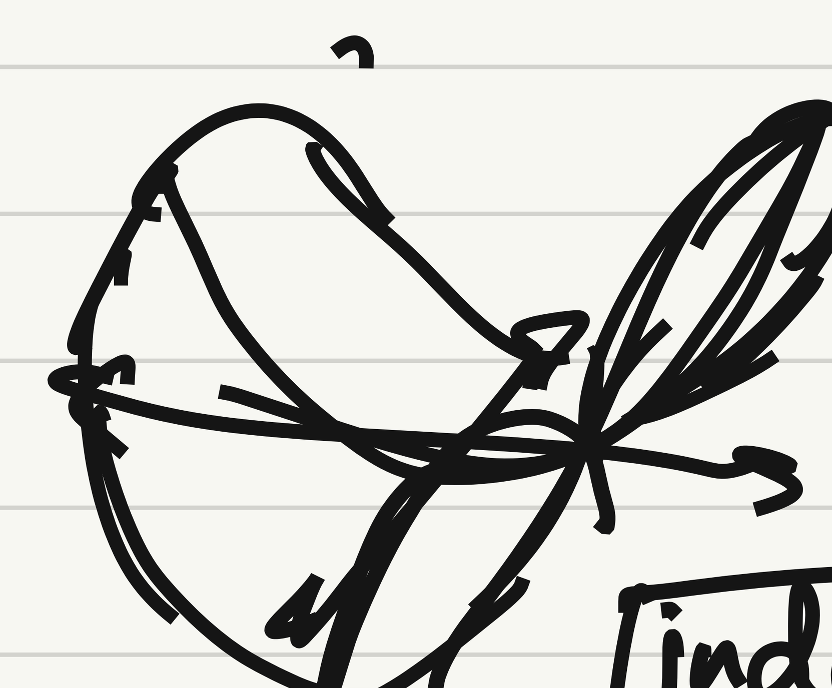
$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} : \quad q(\vec{x}) = x_1^2$$



$\boxed{\text{positive semidefinite}}$

$$A = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$$

$q(\vec{x}) = x_1^2 - x_2^2$



$\boxed{\text{indefinite}}$

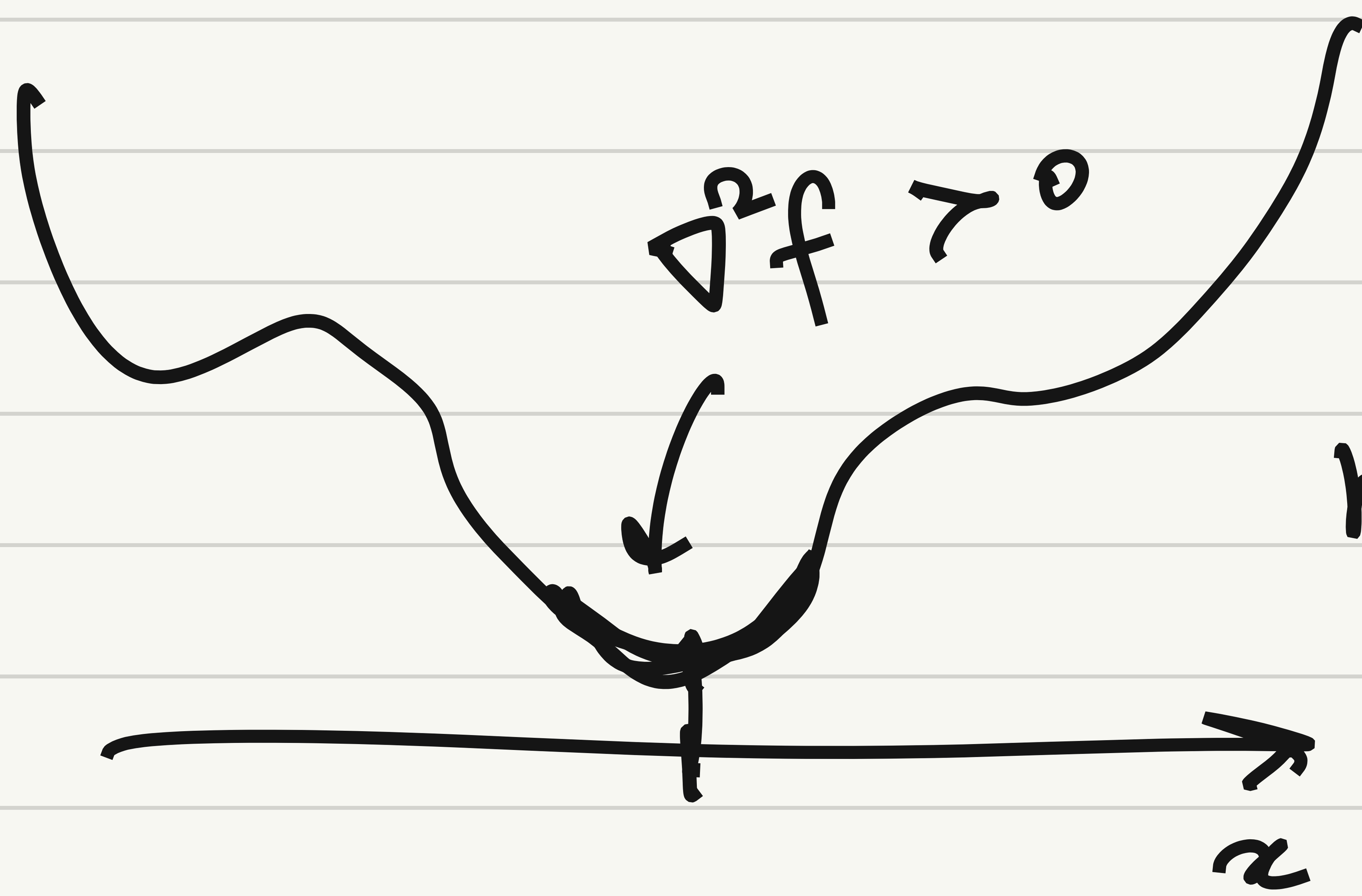
$A \in \mathbb{C}^{m \times m}$ is $\boxed{\text{Hermitian positive definite}}$ if $A^* = A$

and $\underline{\vec{x}^* A \vec{x} > 0 \;\; \forall \vec{x} \neq 0}$

$A \succ 0$

$\boxed{\text{Hermitian positive semidefinite}}$ $(A \succeq 0)$: $A^* = A$

$\vec{x}^* A \vec{x} \geq 0 \;\; \forall \vec{x} \neq 0.$



min $f(x)$

Properties: $A \succ 0$, $B \succ 0$

- $A + B \succ 0$
- $sA \succ 0$ if $s \in \mathbb{R}$, $s > 0$
- $\cancel{AB}$ is usually not Hermitian
- $X^* A X$ is Hermitian
  - $\succ 0$ if X is tall, full rank
  - $\succeq 0$ otherwise

$$\vec{x}^* X^* A X \vec{x} = \underbrace{(X\vec{x})}^* A \underbrace{(X\vec{x})} > 0$$

if $X\vec{x} \neq 0$

$A \succ 0 \Rightarrow X^* A X \succ 0$

$$X = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \qquad X^* \vec{x} = \begin{bmatrix} x_2 \\ x_5 \\ x_6 \end{bmatrix}$$

$$\underbrace{X^* A X} = \begin{bmatrix} a_{22} & a_{25} & a_{26} \\ a_{52} & a_{55} & a_{56} \\ a_{62} & a_{65} & a_{66} \end{bmatrix} \quad : \boxed{\text{principal submatrix}} \text{ of } A$$

$\Rightarrow$ every principal submatrix of A is H.P.D.

$\Rightarrow$ every diag. entry of A is real and > 0

$A = I \Rightarrow X^* X \succ 0$ for any tall full rank X

if X is not tall or full rank, $X^* X \succeq 0$.

<u>Cholesky factorization</u>

A is HPD,

$$A = \left[\begin{array}{c|c} 1 & \vec{b}^* \\ \hline \vec{b} & C \end{array}\right]$$

$$\left[\begin{array}{c|c} 1 & 0 \\ \hline -\vec{b} & I \end{array}\right]$$

$$L_1 A = \left[\begin{array}{c|c} 1 & \vec{b}^* \\ \hline 0 & C - \vec{b}\vec{b}^* \end{array}\right]$$

$$A = \underbrace{\left[\begin{array}{c|c} 1 & 0 \\ \hline \vec{b} & I \end{array}\right]}_{L_1} \underbrace{\left[\begin{array}{c|c} 1 & \vec{b}^* \\ \hline 0 & C - \vec{b}\vec{b}^* \end{array}\right]}$$

$$A = L_1 \left[\cdots \right]$$

$$= L_1 \underbrace{\left[\cdots \right]} L_1^*$$

$$= L_1 \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & C - \vec{b}\vec{b}^* \end{array}\right] \underbrace{\left[\begin{array}{c|c} 1 & \vec{b}^* \\ \hline 0 & I \end{array}\right]}_{L_1^*} = L_1 \underbrace{\left[\begin{array}{c|c} 1 & \\ \hline & C - \vec{b}\vec{b}^* \end{array}\right]} L_1^*$$

$$A = L_1 \underbrace{\left[\begin{array}{c|c} 1 & \\ \hline & c - \vec{b}\vec{b}^* \end{array}\right]}_{} L_1^* = L_1 L_2 \left[\begin{array}{c|c} 1 & \\ & 1 \\ \hline & D \end{array}\right] L_2^* L_1^* = \cdots$$

$$= \underbrace{(L_1 L_2 \cdots L_m)}_{L} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \underbrace{(L_m^* \cdots L_2^* L_1^*)}_{L^*}$$

① What if $a_{11} \neq 1$?

② How do I know I can recurse? $c - \vec{b}\vec{b}^* \succ 0$?

$\alpha = \sqrt{a_{11}}$ because $a_{11} > 0$

$$A = \left[\begin{array}{c|c} a_{11} & \vec{b}^* \\ \hline \vec{b} & C \end{array}\right] \qquad L_1 = \left[\begin{array}{c|c} \alpha & 0 \\ \hline \vec{b}/\alpha & I \end{array}\right] \longrightarrow A = L_1 \left[\begin{array}{c|c} 1 & \\ \hline & \underbrace{c - \frac{\vec{b}\vec{b}^*}{a_{11}}}_{} \end{array}\right] L_1^*$$

Is $c - \vec{b}\vec{b}^* / a_{11} \succ 0$?

$$A = L_1 \begin{bmatrix} 1 & \\ & c - \vec{b}\vec{b}^*/a_{11} \end{bmatrix} L_1^*$$

$$\Rightarrow \begin{bmatrix} 1 & \\ & \boxed{\dfrac{c - \vec{b}\vec{b}^*}{a_{11}}} \end{bmatrix} = \underbrace{L_1^{-1} A L_1^{-*}}_{x^* A x}$$

prin. submatrix of HPD matrix

$$A = \underbrace{(L_1 L_2 \cdots L_m)}_{\substack{L \\ = R^*}} \underbrace{(L_m^* \cdots L_2^* L_1^*)}_{\substack{L^* \\ = R}}$$

$$A = L L^* = \underline{R^* R}$$

Alg:

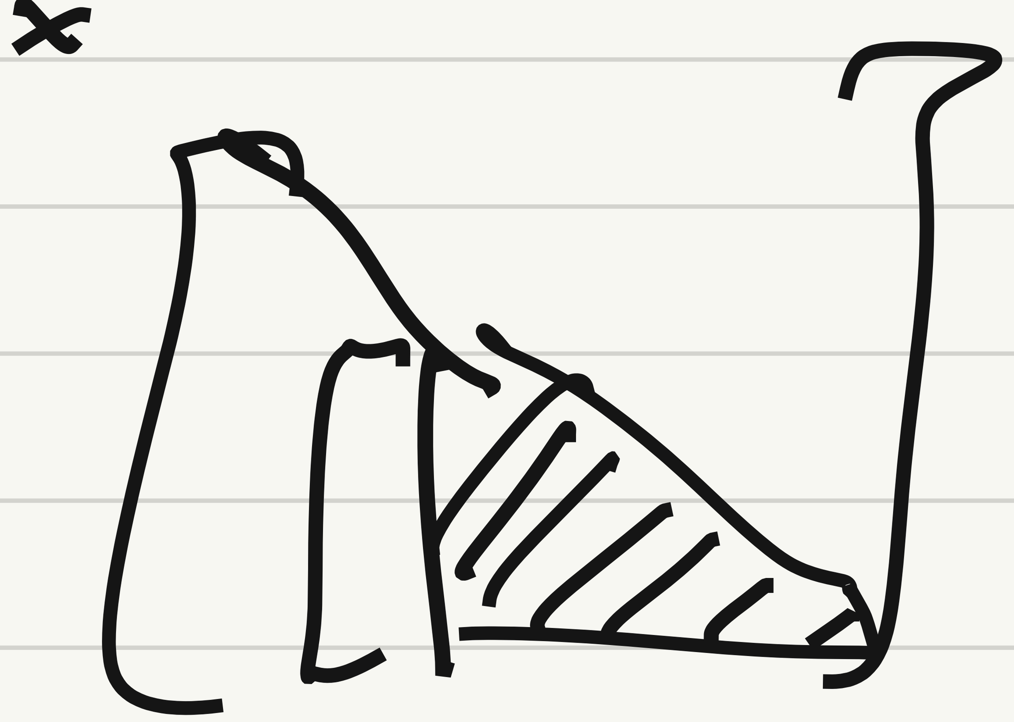
$L = A$

for each col $k = 1 \ldots m$:

$\vec{b} = L_{k+1:m,\,k}$

subtract $\vec{b}\vec{b}^* / L_{kk}$ from lower tri. of $L_{k+1:m,\,k+1:m}$

divide by $\sqrt{L_{kk}}$



k

Thm: Every HPD matrix has a unique Cholesky factorization.

Earlier: for any tall full rank X, $A = X^*X \succ 0$

Now: If $A \succ 0$ then $A = X^*X$ for some X (but not unique)

Operation count: $\sim \frac{1}{3} m^3$ flops

Stability:

$\tilde{L}\tilde{L}^* = A + \delta A$

$\frac{\|\delta A\|}{\|L\| \cdot \|L\|} = O(\varepsilon_m)$ but now $\underline{\|L\| = \|R\| = \sqrt{\|A\|}}$ (proof by SVD)

$= \frac{\|\delta A\|}{\|A\|}$

Thm: for sufficiently small ε_m; no $\tilde{L}_{kk} \leq 0$,

$\tilde{L}\tilde{L}^* = A + \delta A$ with $\|\delta A\| / \|A\| = O(\varepsilon_m)$

Iterative methods

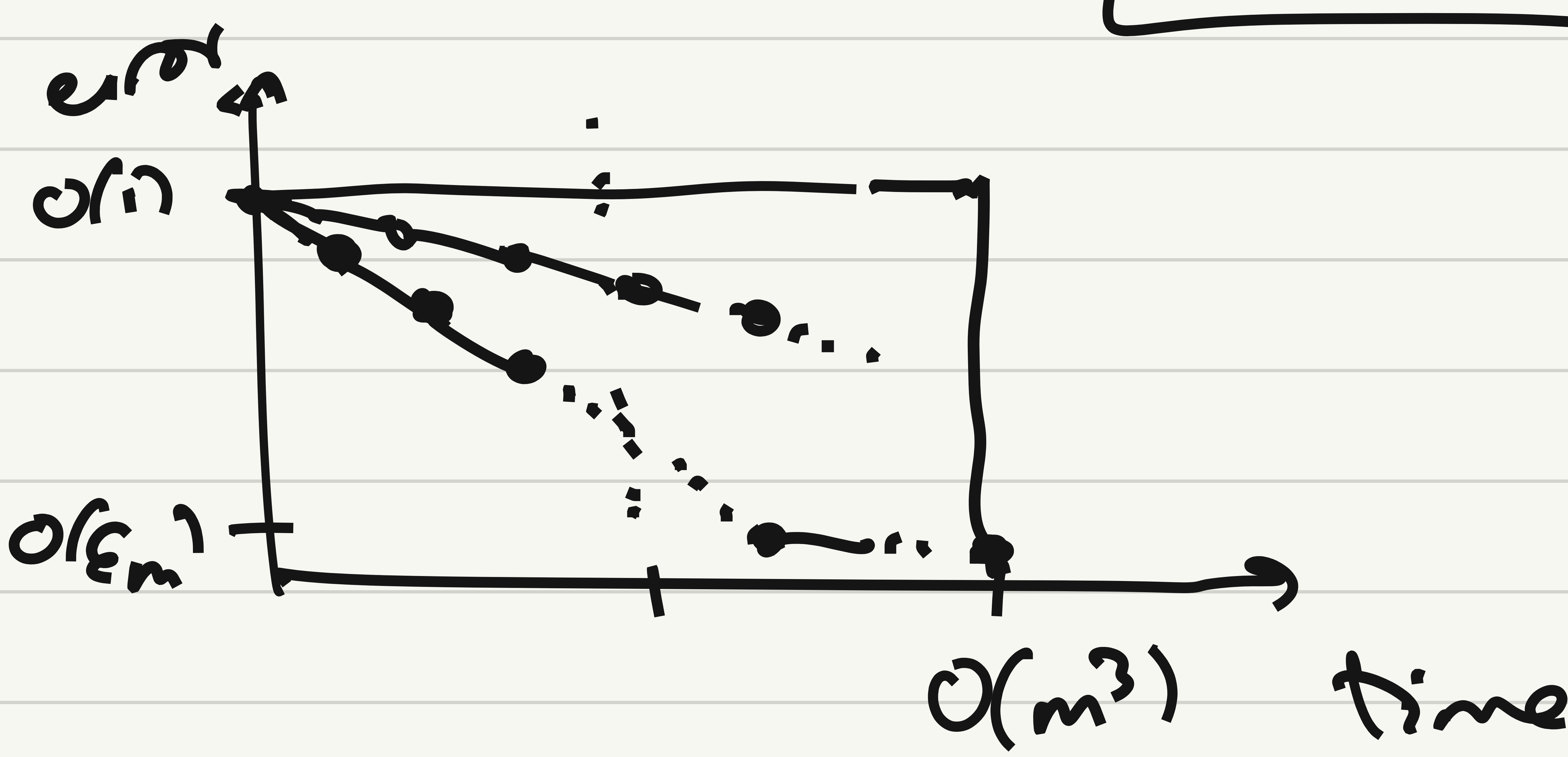
$A\vec{x} = \vec{b}$. $A \in \mathbb{C}^{m\times m}$

All algs so far take $O(m^3)$ flops. ← $\boxed{\text{direct methods}}$

$\boxed{\text{Iterative methods}}$: $\underbrace{\vec{x}^{(1)} \to \vec{x}^{(2)}} \to \cdots \to \vec{x}^* = A^{-1}\vec{b}$

Each iter has cheap cost, $\|\vec{x}^{(k)} - \vec{x}^*\|$ or $\|A\vec{x}^{(k)} - \vec{b}\|$ decreases quickly

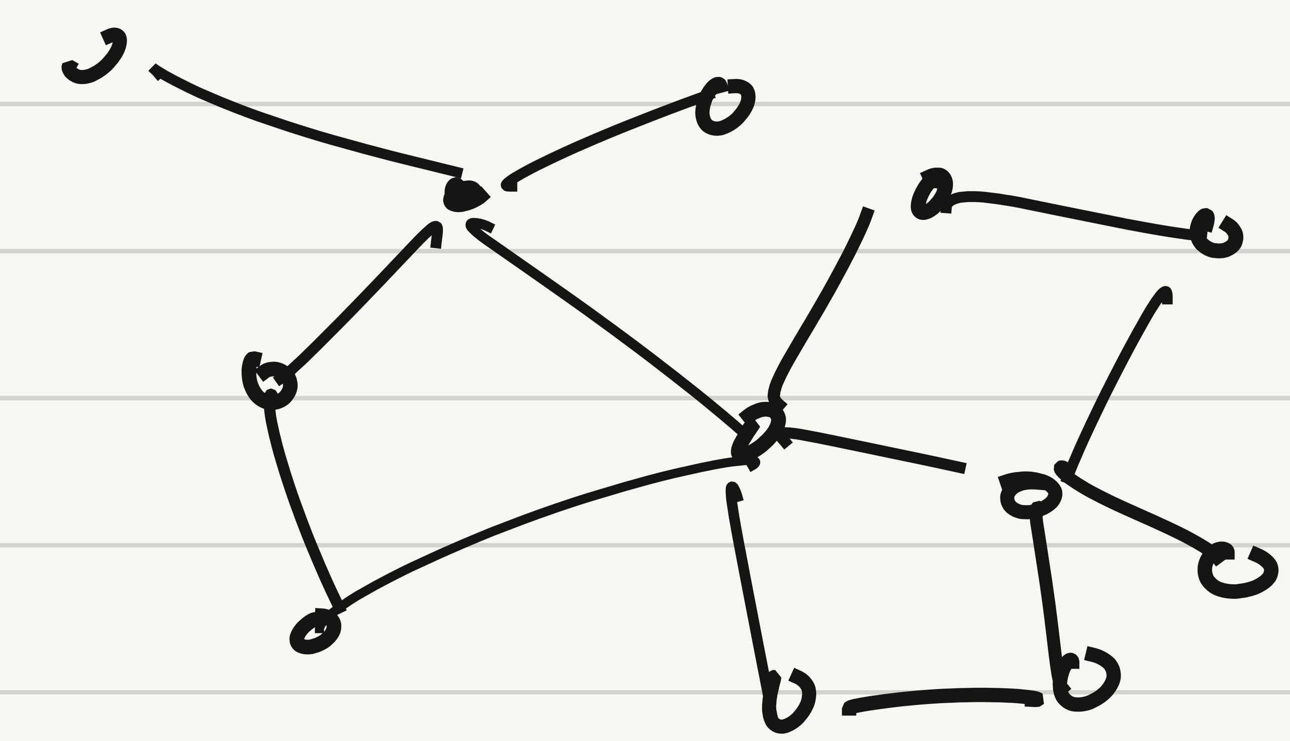
Direct methods: $\underbrace{\sim m \text{ iters}}$, each iter has $\underbrace{O(m^2) \text{ cost}}$



Consider matrices s.t. you can compute

$\vec{x} \mapsto A\vec{x}$ in $\ll m^2$ flops

eg. (i) A is $\boxed{\text{sparse}}$: very few nonzero entries eg. each row & col has $\ll m$ nonzeros



② low rank matrices $A = TW$, $T \in \mathbb{C}^{m \times r}$,

$W \in \mathbb{C}^{r \times m}$, $r \ll m$

③ sparse + low rank

⋮

for such matrices, $\vec{x} \mapsto A\vec{x}$ is cheap!

Plan: treat A as black box, compute only $A\vec{x}$ at each iter.