Iterative methods

$\underline{A}\vec{x} = \underline{\vec{b}}$

$\vec{x}_0 \to \vec{x}_1 \to \vec{x}_2 \to \cdots \quad \vec{x}_* = A^{-1}\vec{b}$

$\vec{x}_n \neq$

$\|x_n - x_*\| \qquad \|\vec{b} - A\vec{x}_n\|$

"Direct methods"

$O(m)$ iters

$O(m^2)$ flops per iter

$\to O(m^3)$ flops

structure

- $\boxed{\text{Sparse}}$ → non-sparse = $\boxed{\text{dense}}$

  $\#\,\text{nonzeros}(A) \ll m^2$

  If $O(m)$ nonzeros : can compute $A\vec{x}$ in $O(m)$ flops

- $\boxed{\text{low rank}}$

  $A = \begin{bmatrix} T \end{bmatrix} \begin{bmatrix} W \end{bmatrix}$ $\quad \text{rank}(A) = r \ll m$

  $O(rm)$ flops

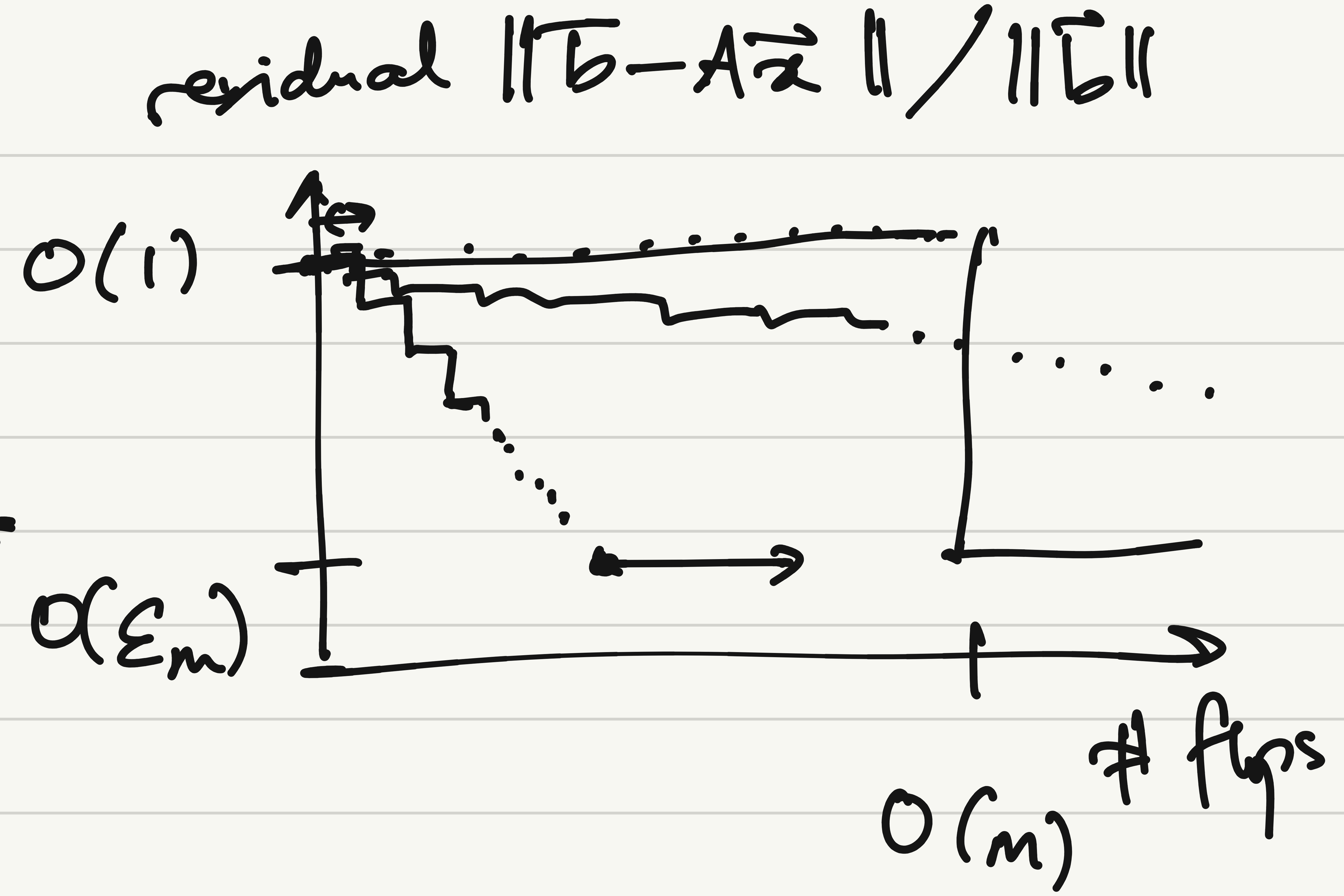
- sparse + low rank

$$\vec{x} \to \boxed{A} \to A\vec{x}$$

$$\vec{x} \to \boxed{A^*} \to A^*\vec{x}$$

$f: \mathbb{R}^m \to \mathbb{R}^m$, linear

$$\underbrace{\underbrace{\vec{x}_0 \to \vec{x}_1}_{} \to \vec{x}_2 \to \cdots \quad \vec{x}_*}_{}$$

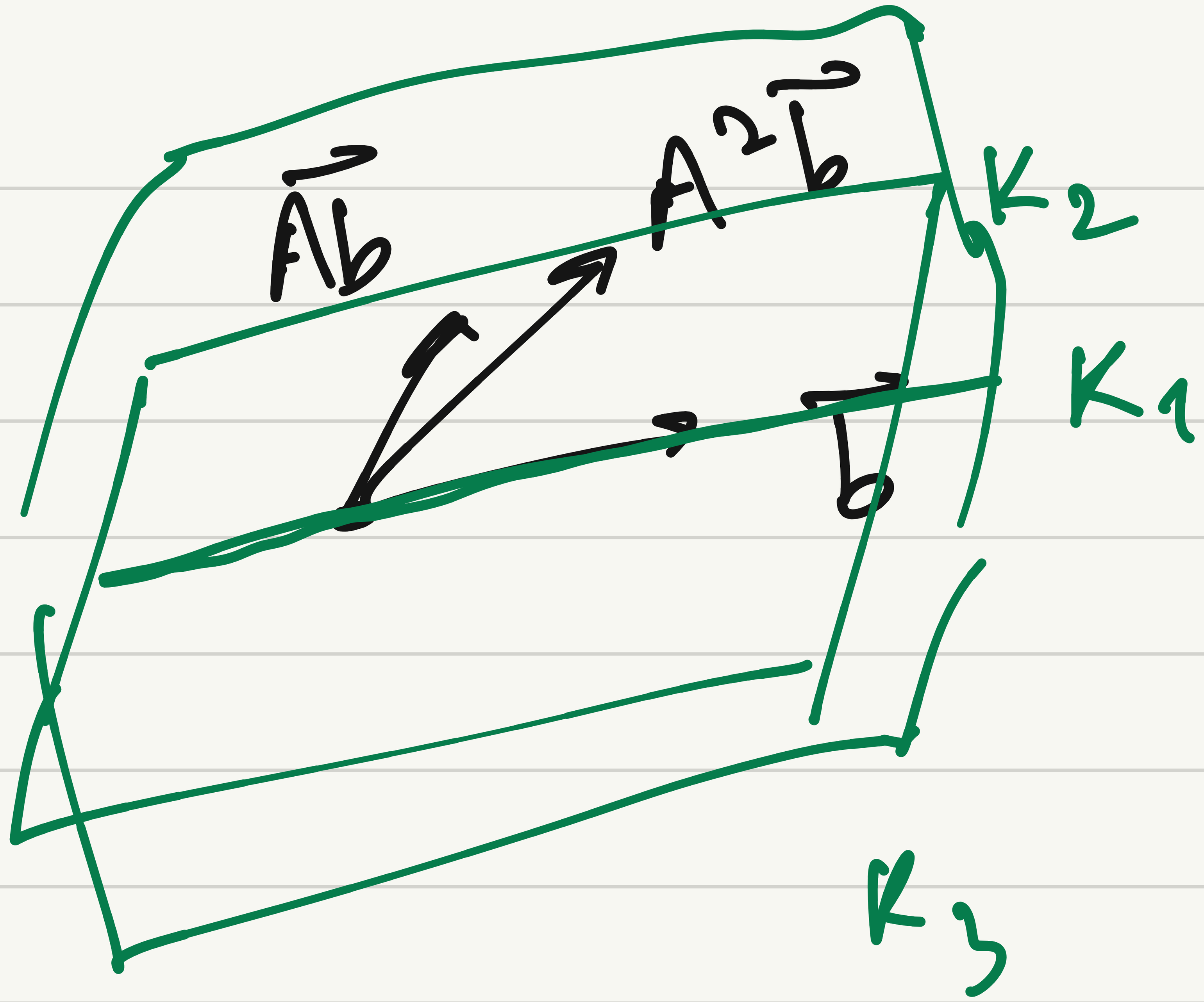


$$\to \boxed{A} \to \quad , \quad \vec{b}$$

$$\underbrace{\vec{b} \to A\vec{b} \to A^2\vec{b} \to \cdots}_{} \quad \boxed{\text{Krylov sequence}}$$

$$\mathcal{K}_1 = \langle \vec{b} \rangle, \quad \mathcal{K}_2 = \langle \vec{b}, A\vec{b} \rangle, \quad \mathcal{K}_3 = \langle \vec{b}, A\vec{b}, A^2\vec{b} \rangle, \quad \ldots$$

$$\mathcal{K}_n = \langle \vec{b}, A\vec{b}, \ldots, A^{n-1}\vec{b} \rangle \qquad \boxed{\text{Krylov subspaces}}$$



$$A\vec{x} = \vec{b}$$

$$\min \underbrace{\|\vec{b} - A\vec{x}\|_2}_{} \quad \text{for } \vec{x} \in K_n = \langle \vec{b}, A\vec{b}, \dots, A^{n-1}\vec{b} \rangle$$

$$\vec{x} = c_0 \underbrace{\vec{b}}_{} + c_1 \underbrace{A\vec{b}}_{} + \cdots + c_{n-1} \underbrace{A^{n-1}\vec{b}}_{}$$

$\vec{v}_1, \dots, \vec{v}_m$ eigenvectors of A $\qquad |\lambda_1| > |\lambda_2| > \cdots > |\lambda_m|$

$$\vec{b} = s_1\vec{v}_1 + \cdots + s_m\vec{v}_m \Rightarrow A^n\vec{b} = s_1 \underbrace{\lambda_1^n}_{} \vec{v}_1 + \cdots + s_m \lambda_m^n \vec{v}_m$$

$$\approx s_1 \lambda_1^n \vec{v}_1$$

Do Gram-Schmidt on $\vec{b}, A\vec{b}, A^2\vec{b}, \dots$

$\Rightarrow$ $\boxed{\text{Arnoldi iterations}}$

Arnoldi iters

start with $\vec{q}_1 = \vec{b}/\|\vec{b}\|$

for $n = 1, 2, 3, \ldots$

$\vec{v}_{n+1} = A\vec{q}_n$

project $\vec{v}_{n+1}$ onto $\langle \vec{q}_1, \ldots \vec{q}_n \rangle^\perp$

set $\vec{q}_{n+1} = \vec{v}_{n+1}/\|\vec{v}_{n+1}\|$

$A\vec{q}_n \in \langle \vec{q}_1, \ldots, \vec{q}_{n+1} \rangle$

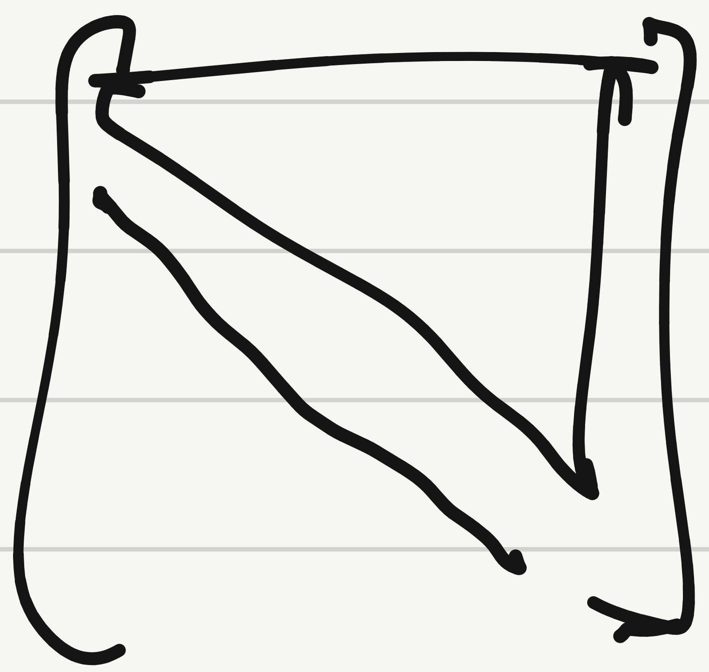
$= h_{1n}\vec{q}_1 + \cdots h_{nn}\vec{q}_n +$

$h_{n+1,n}\vec{q}_{n+1}$

$$A \begin{bmatrix} \vec{q}_n \end{bmatrix} = \begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_{n+1} \end{bmatrix} \begin{bmatrix} h_{1n} \\ \vdots \\ h_{nn} \\ h_{n+1,n} \end{bmatrix}$$

$$A \underbrace{\begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_n \end{bmatrix}}_{Q_n \in \mathbb{C}^{m \times n}} = \underbrace{\begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_n & \vec{q}_{n+1} \end{bmatrix}}_{Q_{n+1} \in \mathbb{C}^{m \times (n+1)}} \underbrace{\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ & h_{32} & \ddots & \vdots \\ & & \ddots & h_{nn} \\ & & & h_{n+1,n} \end{bmatrix}}_{\tilde{H}_n \in \mathbb{C}^{(n+1) \times n}}$$

$$\boxed{AQ_n = Q_{n+1}\tilde{H}_n}$$



$\boxed{\text{upper Hessenberg}}$

$$\min \|\vec{b} - A\vec{x}\|_2 \text{ s.t. } \vec{x} \in K_n$$
$$= \langle \vec{b}, A\vec{b}, \dots A^{n-1}\vec{b} \rangle$$
$$= \langle \vec{q}_1, \vec{q}_2, \dots, \vec{q}_n \rangle$$

$n=1 \longrightarrow \vec{x}_1$
$n=2 \longrightarrow \vec{x}_2$
$\vdots$

$$\vec{x} = Q_n \vec{y}$$

$$\min \|\vec{b} - AQ_n\vec{y}\|_2 \quad \text{over all } \vec{y} \in \mathbb{C}^n$$

$\vec{b} \in \text{range}(Q_{n+1})$

$$= \|\vec{b} - Q_{n+1}\tilde{H}_n\vec{y}\|_2 = \|Q_{n+1}^*\vec{b} - \tilde{H}_n\vec{y}\|_2 = \|\,\|\vec{b}\|\vec{e}_1 - \tilde{H}_n\vec{y}\|_2$$

$\min \|\vec{b} - A\vec{x}\|_2$ s.t. $\vec{x} \in K_n$

$\Leftrightarrow$ $\min \| \underbrace{\tilde{H}_n}_{(n+1)\times n} \vec{y} - \underbrace{\|\vec{b}\|\vec{e}_1}\|_2$ over all $\vec{y} \in \mathbb{C}^n$

Compute QR of $\tilde{H}_n$ $\longrightarrow$ solve for $\vec{y}_n$ $\longrightarrow$ $\vec{x}_n = Q_n \vec{y}_n$

$\boxed{\text{GMRES}}$ (generalized minimum residual method)

for $n = 1, 2, 3, \ldots$ cost increases every iteration!

do Arnoldi to get $Q_n, \tilde{H}_n$ $\longleftarrow$ $O(mn)$

solve $\underbrace{\min \| \cdots \|}$ $\longleftarrow$ $O(n^3)$

using Hessenberg

$qr(\tilde{H}_n) : O(n^2)$

$qr(\tilde{H}_{n-1}) \to qr(\tilde{H}_n) : O(n)$

n iters $\longrightarrow$ $\vec{x}_n$

$\vec{x}_* = \vec{x}_n + \Delta\vec{x}$ $\Rightarrow$ solve $A \underbrace{\Delta\vec{x}} = \underbrace{\vec{b} - A\vec{x}_n}$

Convergence of GMRES

$\vec{x}_1, \vec{x}_2, \vec{x}_3, \ldots$

$\vec{r}_n = \vec{b} - A\vec{x}_n$, $\|\vec{r}_n\| \overset{?}{\longrightarrow} 0$

1. $\|\vec{r}_{n+1}\| \leq \|\vec{r}_n\|$
2. $\|\vec{r}_n\| = 0$ for some $n \leq m$

$K_m = \mathbb{C}^m$

$A\vec{q}_n \in \langle \vec{q}_1, \ldots, \vec{q}_n \rangle$: $\boxed{\text{Breakdown}}$

then $\vec{x}_* \in \langle \vec{q}_1, \ldots, \vec{q}_n \rangle$

find upper bound for $\|\vec{r}_n\|$ for any $n < m$?

Krylov subspaces $\longleftrightarrow$ polynomial approximation $\longleftrightarrow$ eigenvalues of A

min $\| \cdots \|$ over all $\vec{x} \in K_n$

Any $\vec{x} \in K_n$: $\vec{x} = c_0 \vec{b} + c_1 A\vec{b} + \cdots + c_{n-1} A^{n-1} \vec{b}$

$$= \underbrace{(c_0 I + c_1 A + \cdots + c_{n-1} A^{n-1})}\vec{b}$$

$$= q(A)\vec{b}$$

$q(z) = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1}$

min $\|\vec{r}_n\|$

$$\vec{r} = \vec{b} - A\vec{x} = \vec{b} - A\,q(A)\vec{b} = \underbrace{(I - A\,q(A))}_{p(A)}\vec{b}$$

$p(z) = 1 - z\,q(z)$

degree-n poly. with const. term 1.

$$\min \|\vec{r}_n\|_2 \Longleftrightarrow \min \underbrace{\|p(A)\vec{b}\|_2}$$

$$\|\vec{r}_n\|_2 = \inf \|p(A)\vec{b}\|_2 \leq \|\vec{b}\|_2 \inf \|p(A)\|_2$$

Q. For which matrices does GMRES converge quickly?

A. When $\exists$ poly. p s.t. $\|p(A)\|_2$ is small

<u>Next</u>: for which matrices does such poly exist??