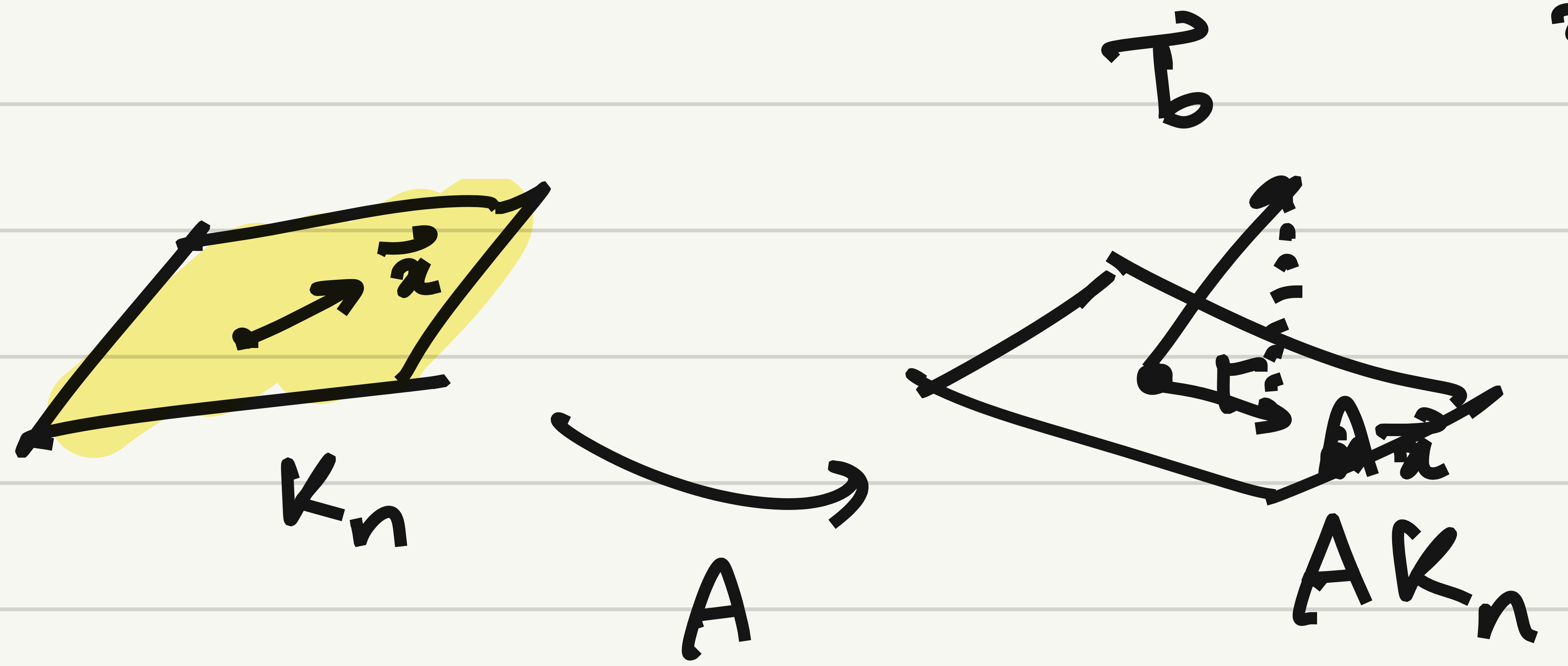
COL726 Lecture 12: Conjugate Gradient method

1. wrap-up GMRES ←
2. Minor exam logistics
3. CG

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

$K_n = \langle \vec{b}, A\vec{b}, A^2\vec{b}, \dots, A^{n-1}\vec{b} \rangle$

$$\vec{x}_n = \operatorname*{argmin}_{\vec{x} \in K_n} \|\vec{b} - A\vec{x}\|_2$$



$\vec{r}_n = \vec{b} - A\vec{x}_n$

Thm: At nth step of GMRES,

$$\frac{\|\vec{r}_n\|}{\|\vec{b}\|} \leq \inf_{p_n \in P_n} \|p_n(A)\| \leq \kappa(V) \inf_{p_n \in P_n} \|p_n\|_{\Lambda(A)}$$

$\vec{r}_0 = \vec{b} - A\cancel{\vec{x}_0}\ 0$

$A = V \Lambda V^{-1}$

$\max_{\lambda_j} |p(\lambda_j)|$

$\vec{x} \in K_n$

$\updownarrow$

$[c_0, c_1, c_2, \dots, c_{n-1}]$

$\updownarrow$

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

$\updownarrow$

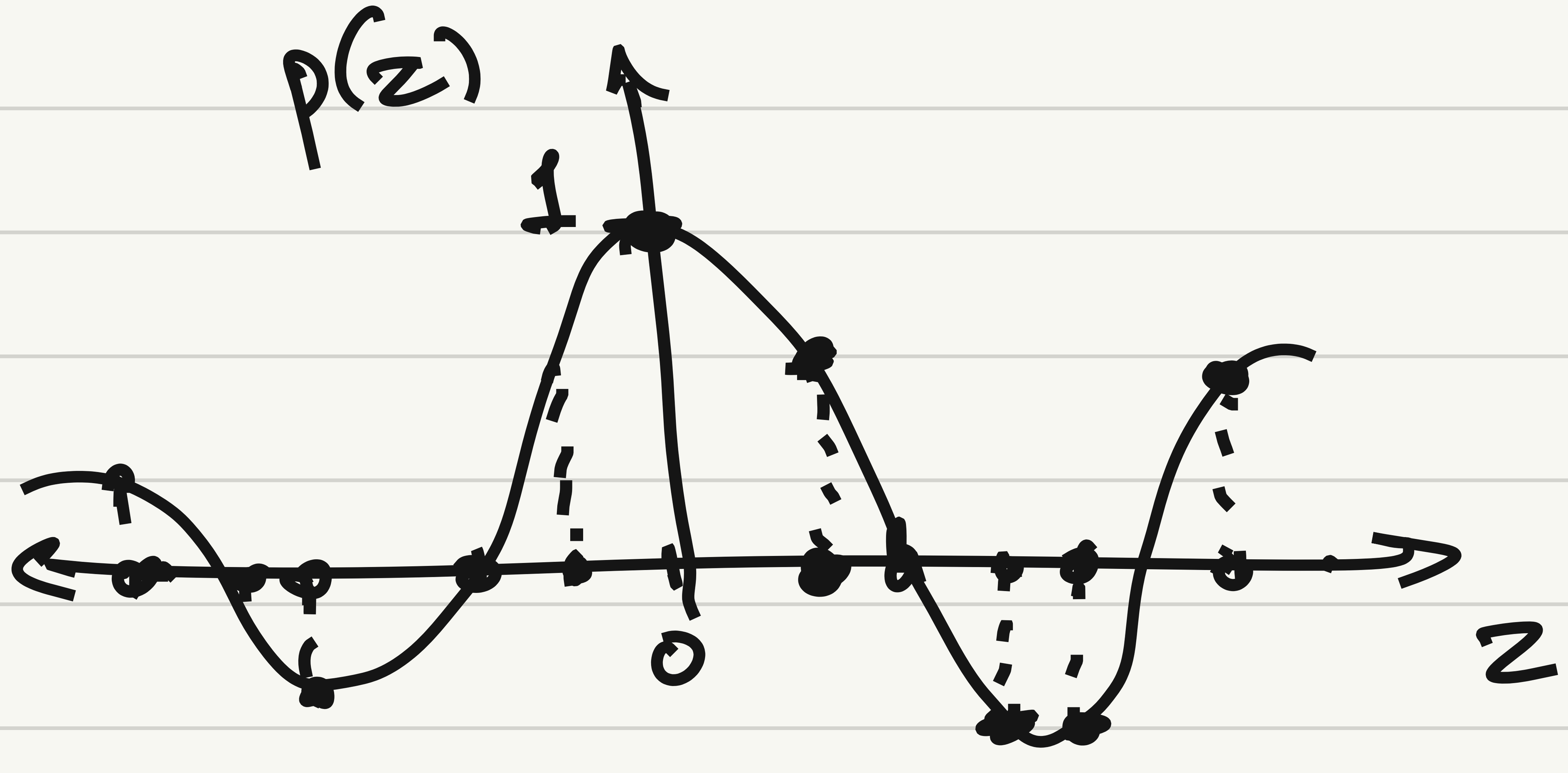
$\vec{x} = q(A)\vec{b}$

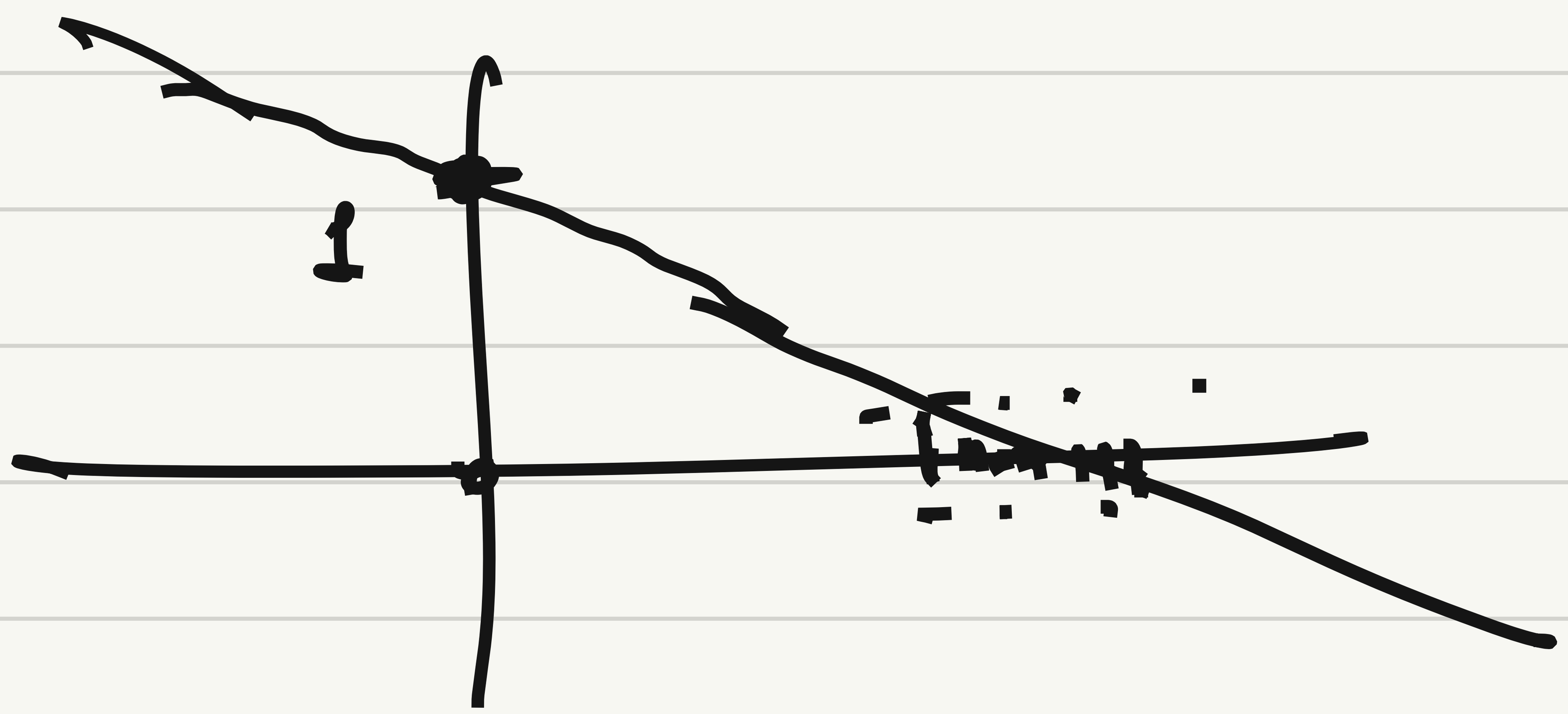
$\vec{r} = p(A)\vec{b}$

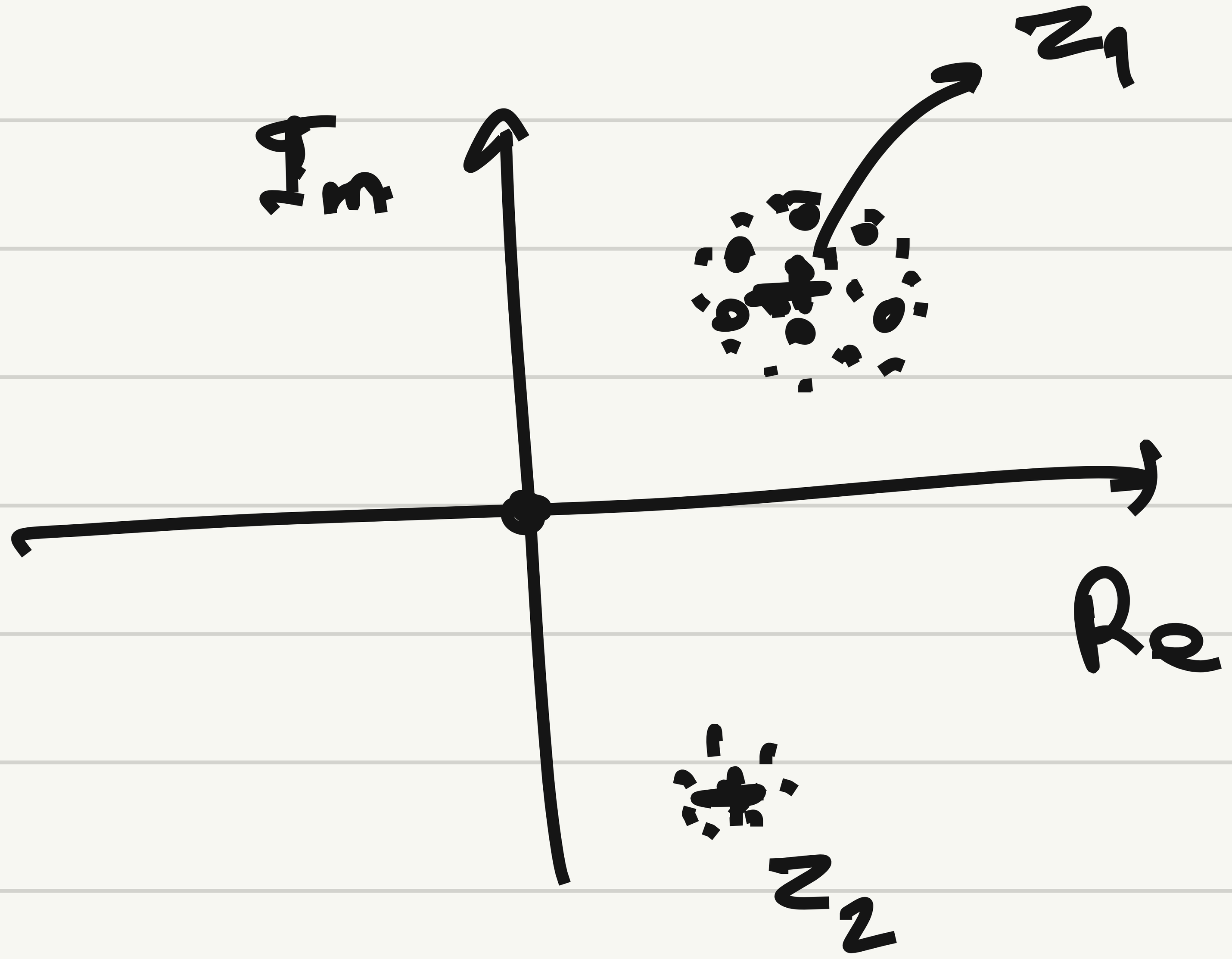
$$\frac{\|\vec{r}_n\|}{\|\vec{r}_0\|} \le \kappa(V) \inf_{p_n \in P_n} \underbrace{\|p_n\|_{\Lambda(A)}}_{= \max_{\lambda_j} |p(\lambda_j)|}$$

$\Rightarrow$ Residual at nth iter is small if:

- $\kappa(V)$ is small (eigenvectors of A are close to ortho.)
- and $\exists\, p_n$ of degree n, $p_n(0) = 1$ which is small on all λ_j



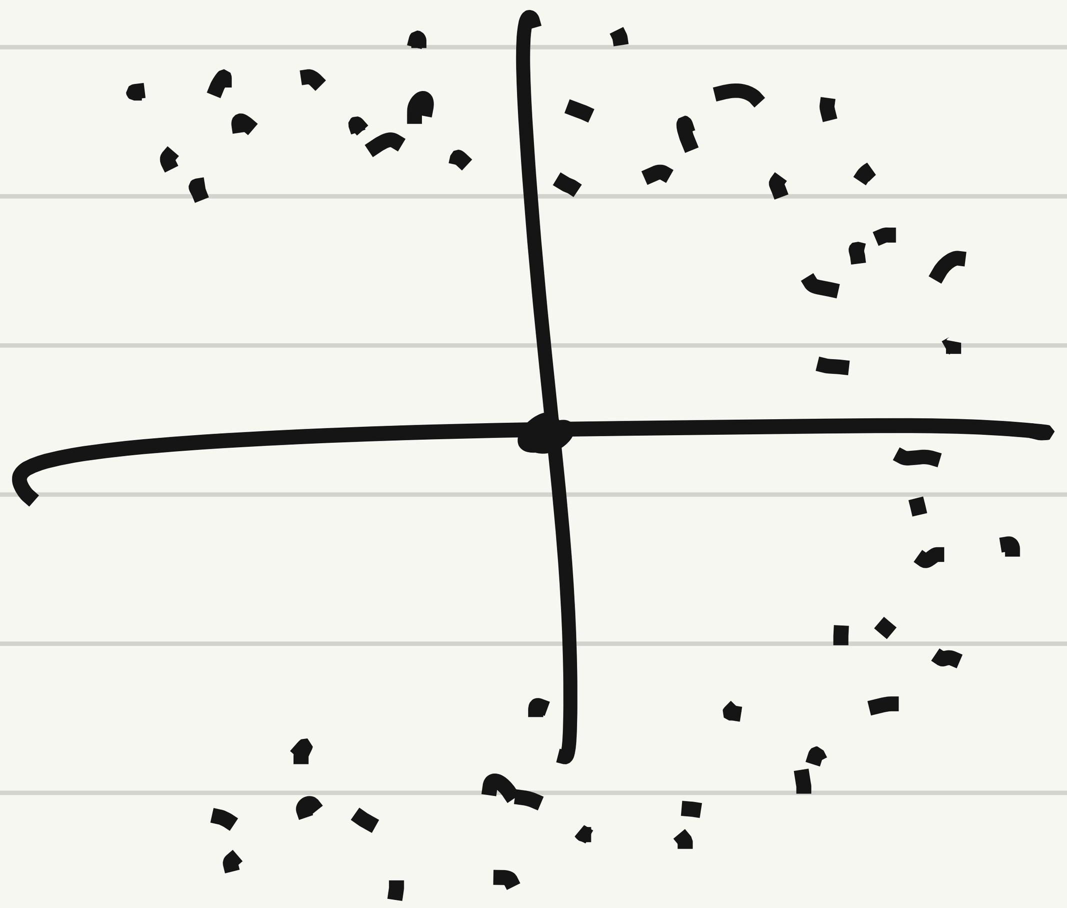




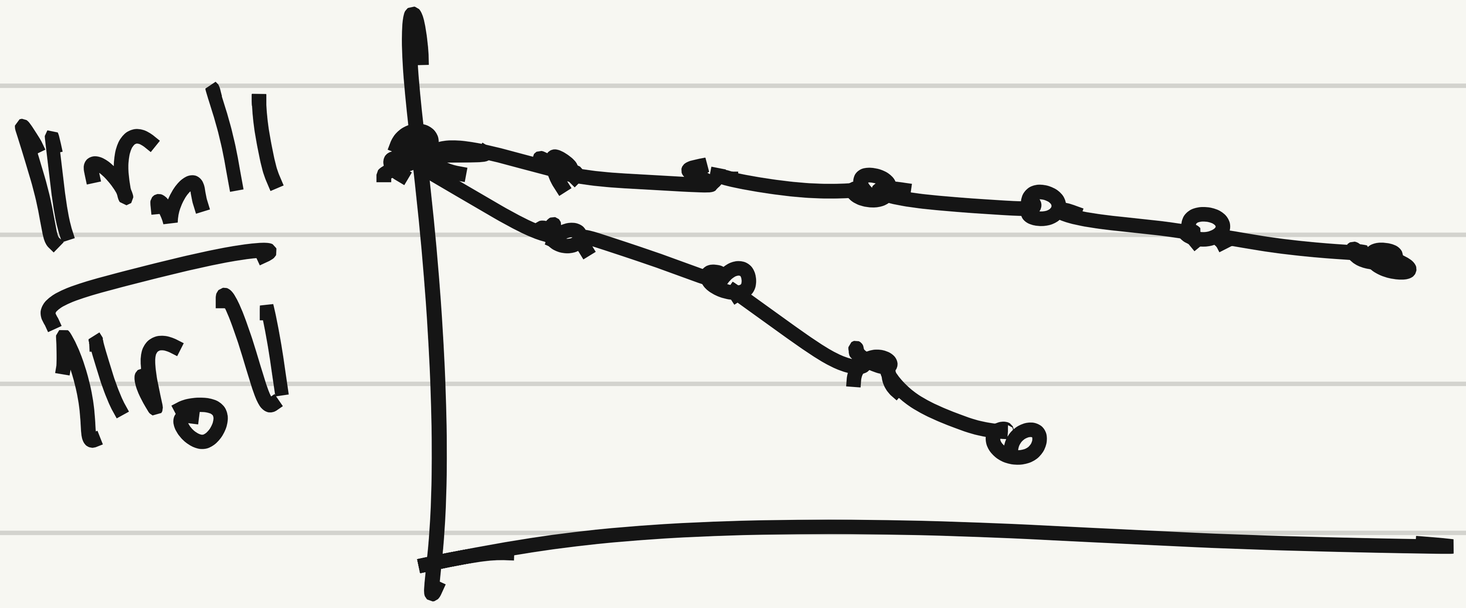
$$p_1(z) = 1 - \frac{z}{z_1}$$

$$p_2(z) = \left(1 - \frac{z}{z_1}\right)\left(1 - \frac{z}{z_2}\right)$$

$x_0 \to x_1 \to x_2 \to \cdots \to x_n \to \cdots$



GMRES will converge slowly



Termination criterion?

- max # iterations
- tolerance on $\|\vec{r}_n\|$
- ...

Minor exam:

Online only, on Gradescope

Monday 4-5pm + 15 min for scanning & uploading
Upload by 5:15pm

Syllabus = everything including today

Open book, open notes, open internet

But no collaboration or discussion : with other people
except for clarifications on Piazza

Conjugate Gradients (CG)

$A\vec{x} = \vec{b}$: only if A is Hermitian positive definite $A^* = A$

Most famous Krylov subspace method! $\vec{x}^* A \vec{x} > 0$ for all $\vec{x} \neq 0$

We'll only consider real case: $A^T = A$, $\vec{x}^T A \vec{x} > 0$ if $\vec{x} \neq 0$

$\boxed{\text{Residual}}$ $\vec{r} = \vec{b} - A\vec{x}$,

$\boxed{\text{Error}}$ $\vec{e} = \vec{x}_* - \vec{x}$ $A\vec{x}_* = \vec{b}$

$A\vec{e} = \vec{r}$

Big picture, no proofs

① define problem-dependent norm

$$\|\vec{x}\|_A = \sqrt{\vec{x}^T A \vec{x}}$$

$$\vec{x}_n = \underset{\vec{x} \in K_n}{\operatorname{argmin}} \|\vec{e}_n\|_A \qquad \|\vec{x}_* - \vec{x}_n\|_A$$

$$f(\vec{x}) = \|\vec{x}_* - \vec{x}\|_A^2 = (\vec{x}_* - \vec{x})^T A (\vec{x}_* - \vec{x})$$

$$= \underbrace{\vec{x}_*^T A \vec{x}_*}_{\text{Const.}} + \vec{x}^T A \vec{x} - 2\vec{x}^T \underbrace{A \vec{x}_*}_{\vec{b}}$$

$$= \vec{x}^T A \vec{x} - 2\vec{x}^T \vec{b} + \text{Const.}$$

CG algo:

Initialize $\vec{x}_0 = 0$, $\vec{r}_0 = \vec{b}$, $\vec{p}_0 = \vec{b}$

repeat for $n = 0, 1, 2, \ldots$:

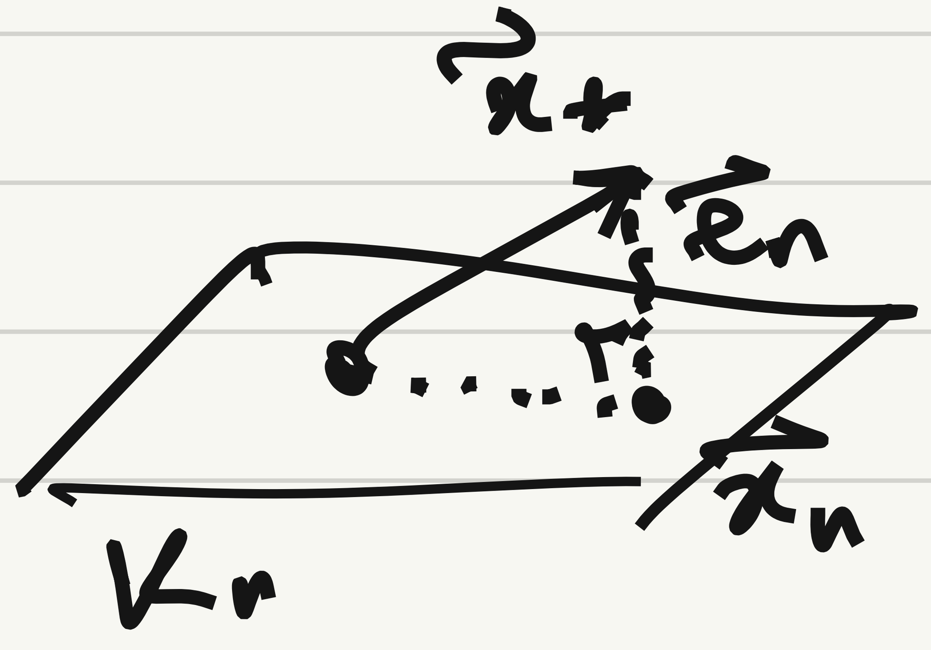
$$\alpha = \vec{r}_n^T \vec{r}_n / \vec{p}_n^T A \vec{p}_n$$

$$\vec{x}_{n+1} = \vec{x}_n + \alpha \vec{p}_n$$

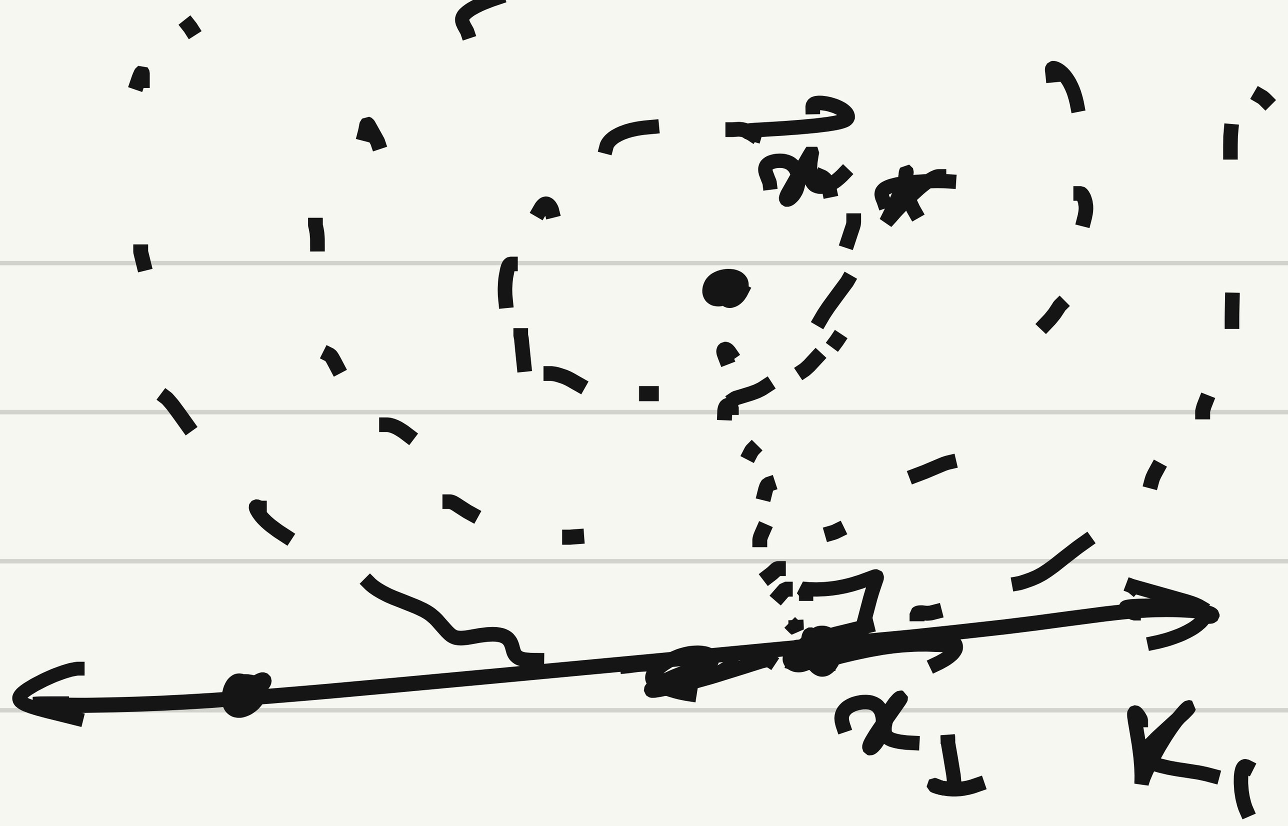
$$\vec{r}_{n+1} = \vec{r}_n - \alpha A \vec{p}_n$$

$$\beta = \vec{r}_{n+1}^T \vec{r}_{n+1} / \vec{r}_n^T \vec{r}_n$$

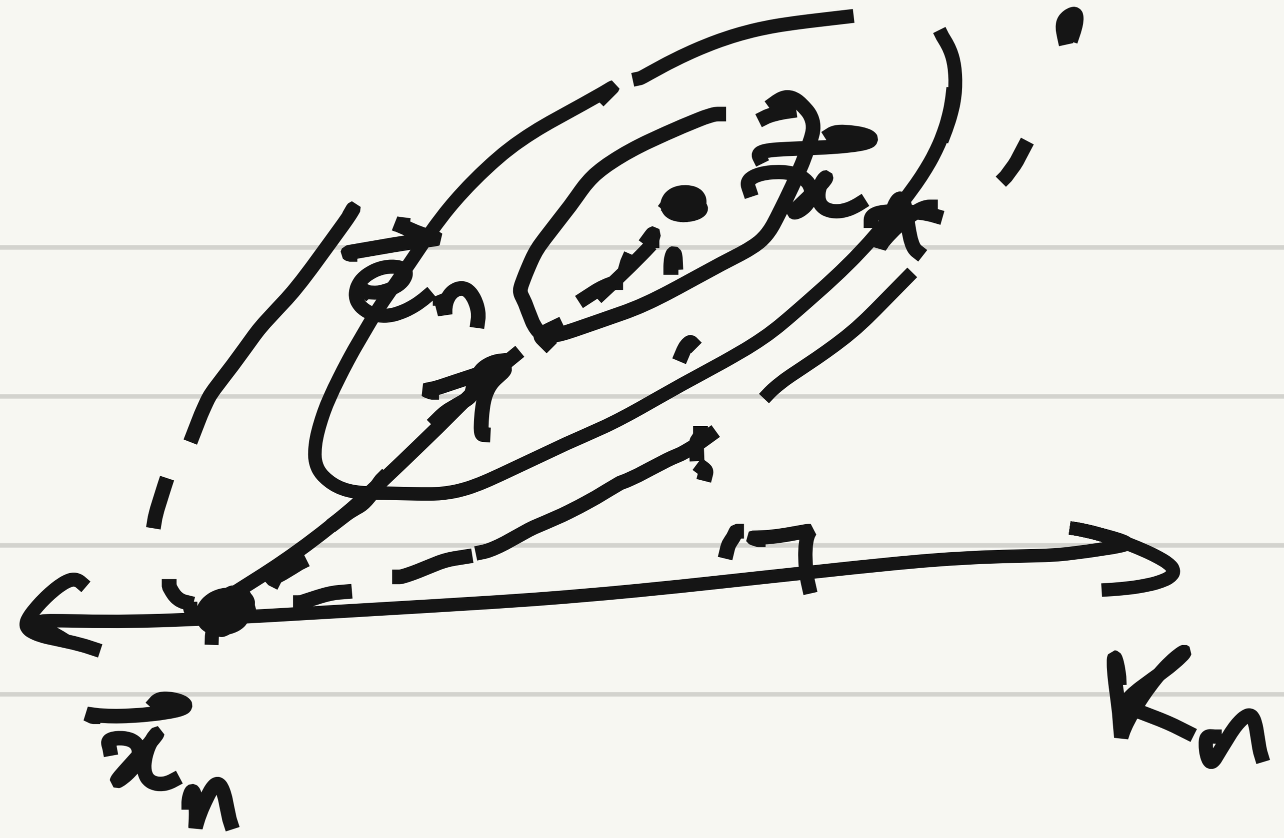
$$\vec{p}_{n+1} = \vec{r}_{n+1} + \beta \vec{p}_n$$



In 2-norm, $\vec{e}_n \perp K_n \iff \underbrace{\vec{v}^T \vec{e}_n = 0 \text{ for all } \vec{v} \in K_n}$



2-norm



A-norm

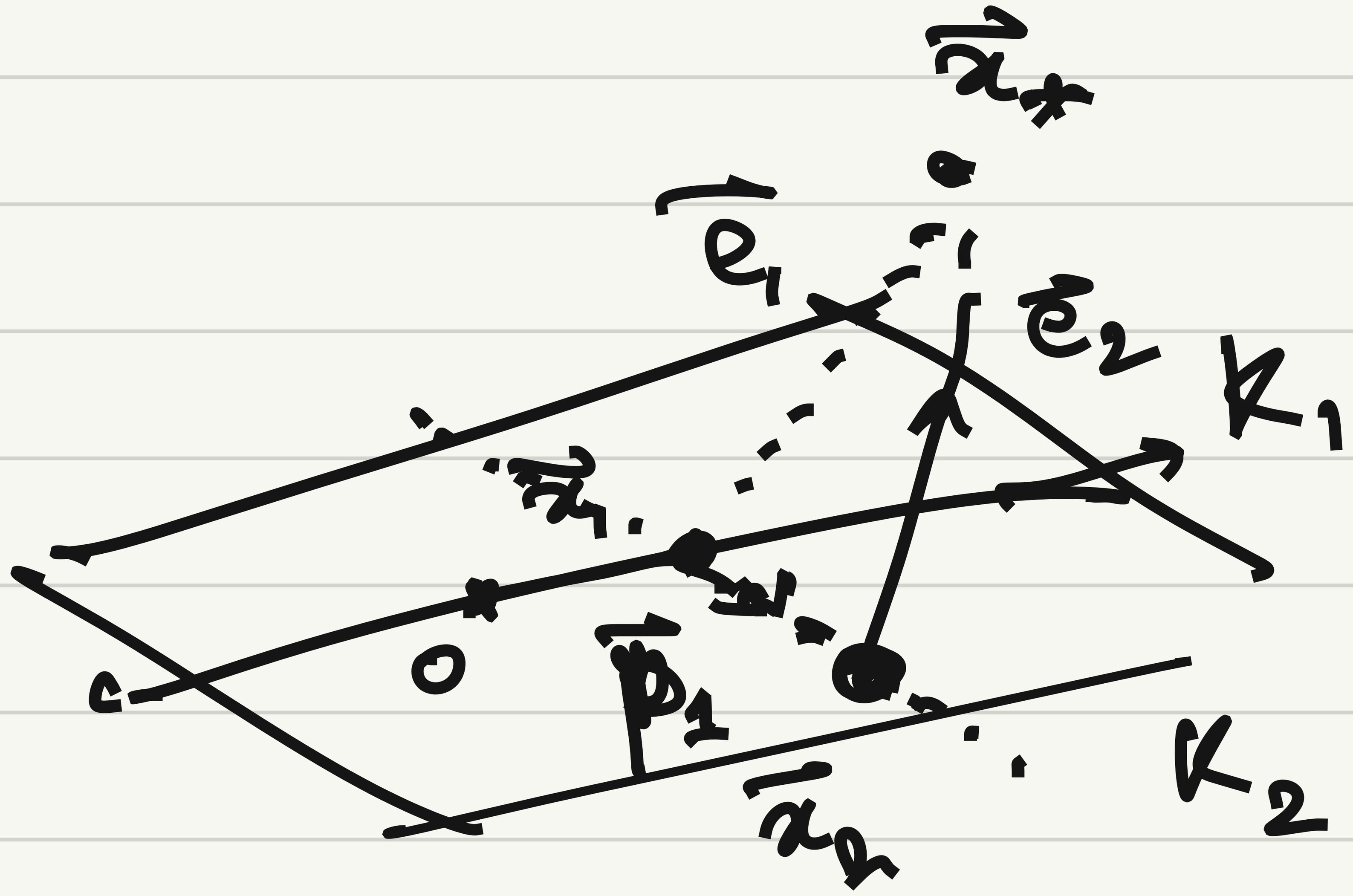
$\vec{v}$ is $\boxed{\text{A-orthogonal}}$ or $\boxed{\text{A-conjugate}}$ to $\vec{x}_* - \vec{x}_n$

$$\vec{v}^T(\vec{x}_* - \vec{x}_\perp) = 0$$

for all $\vec{v} \in K_n$

$$\vec{v}^T A(\vec{x}_* - \vec{x}_n) = 0$$

for all $\vec{v} \in K_n$



Alg.: At nth iter, $\vec{x}_n$ minimizes $f(\vec{x}_n)$ over K_n

$$\vec{r}_n = \vec{b} - A\vec{x}_n \in K_{n+1}$$

Want $\vec{x}_{n+1}$ which min $f(\vec{x}_{n+1})$ over K_{n+1}

Choose $\boxed{\text{search direction}}$ $\vec{p}_n \in K_{n+1}$, $\vec{x}_{n+1} = \vec{x}_n + \alpha \vec{p}_n$

$x^T y \to \vec{x}^T A y$

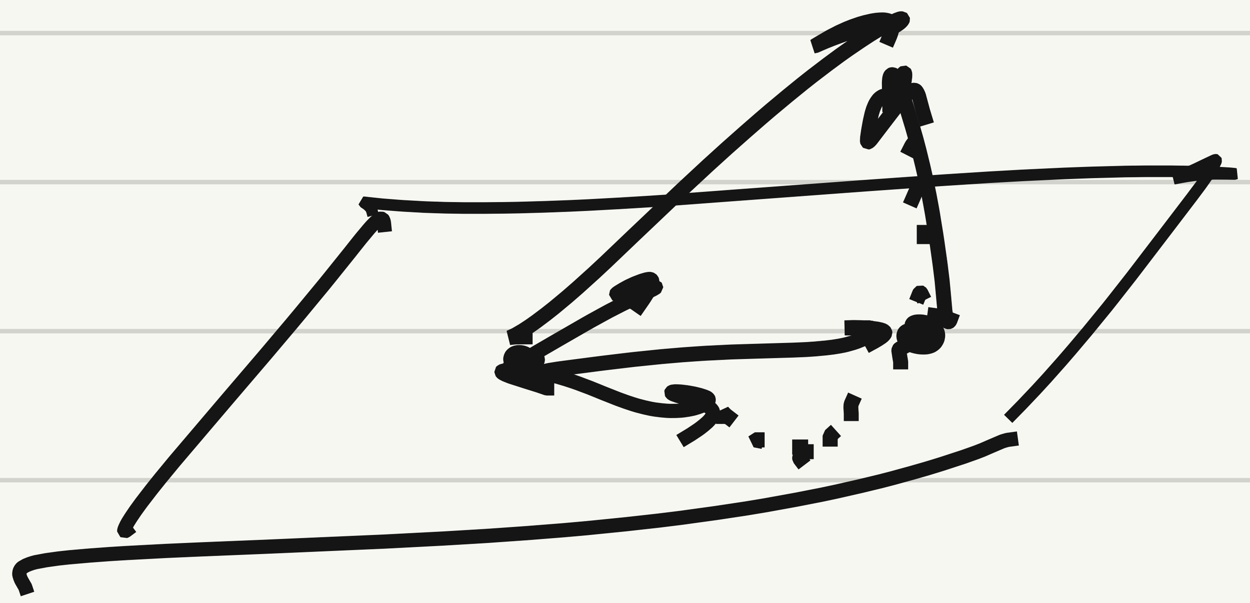
$\|x\|_2 \to \|x\|_A$

$\vec{x}_{n+1} = \vec{x}_n + \alpha \vec{p}_n$

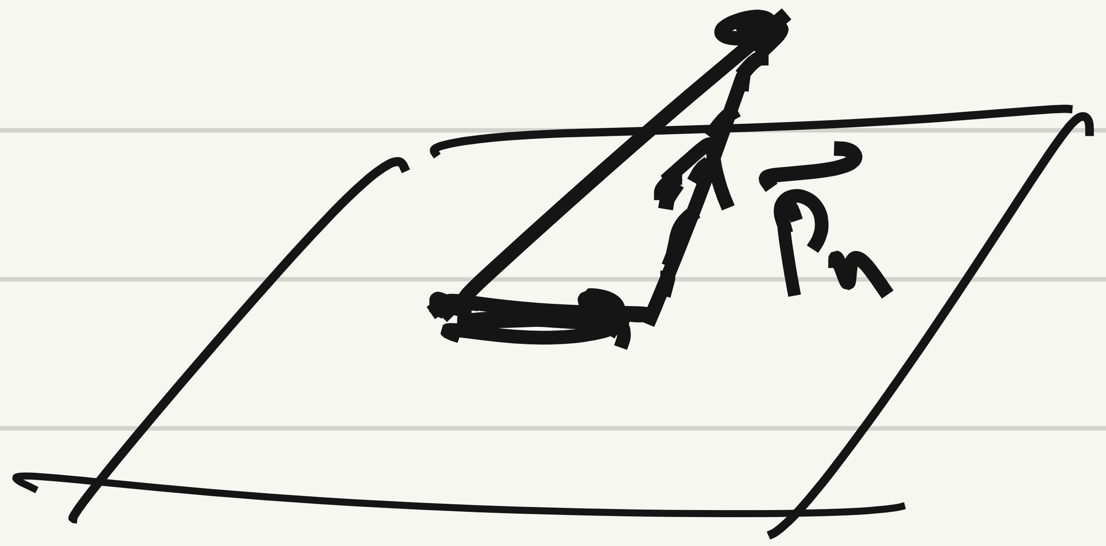
Choose α to minimize $f(\vec{x}_n + \alpha \vec{p}_n)$: quadratic

How to choose $\vec{p}_n$? ② If $\vec{p}_n$ is A-conjugate to K_n then $\vec{x}_{n+1}$ is optimal over K_{n+1}!

Pick any vector in K_{n+1}, e.g. $\vec{r}_n$, do A-conjugate decomp. with $\vec{p}_0, \dots, \vec{p}_{n-1}$



③ $\vec{r}_n$ is already A-conjugate to $\vec{p}_0, \vec{p}_1, \dots, \vec{p}_{n-2}$



K_n

$\Rightarrow \vec{p}_n = \vec{r}_n + \beta \vec{p}_{n-1}$

Shewchuk, "An Introduction to CG without the Agonizing Pain"

Convergence of CG

$\|\vec{e}_n\|_A$ decreases monotonically with n, goes to zero in $\leq m$ iters

$p(0) = 1$

CG is Krylov subspace method $\Rightarrow$ $\vec{x}_n = q(A)\cdot\vec{b}$, $\vec{e}_n = p(A)\,\vec{e}_0$

$\vec{x}_* - \vec{x}_0 = \vec{x}_*$

$$\Rightarrow \frac{\|\vec{e}_n\|_A}{\|\vec{e}_0\|_A} \leq \inf_{p_n \in P_n} \|p(A)\|_A$$

$$p(A) = V\,p(\Lambda)\,V^{-1} = V \begin{bmatrix} p(\lambda_1) & & \\ & \ddots & \\ & & p(\lambda_m) \end{bmatrix} V^{-1}$$

$A^T = A \Rightarrow$ eigenvectors are orthogonal, A-conjugate

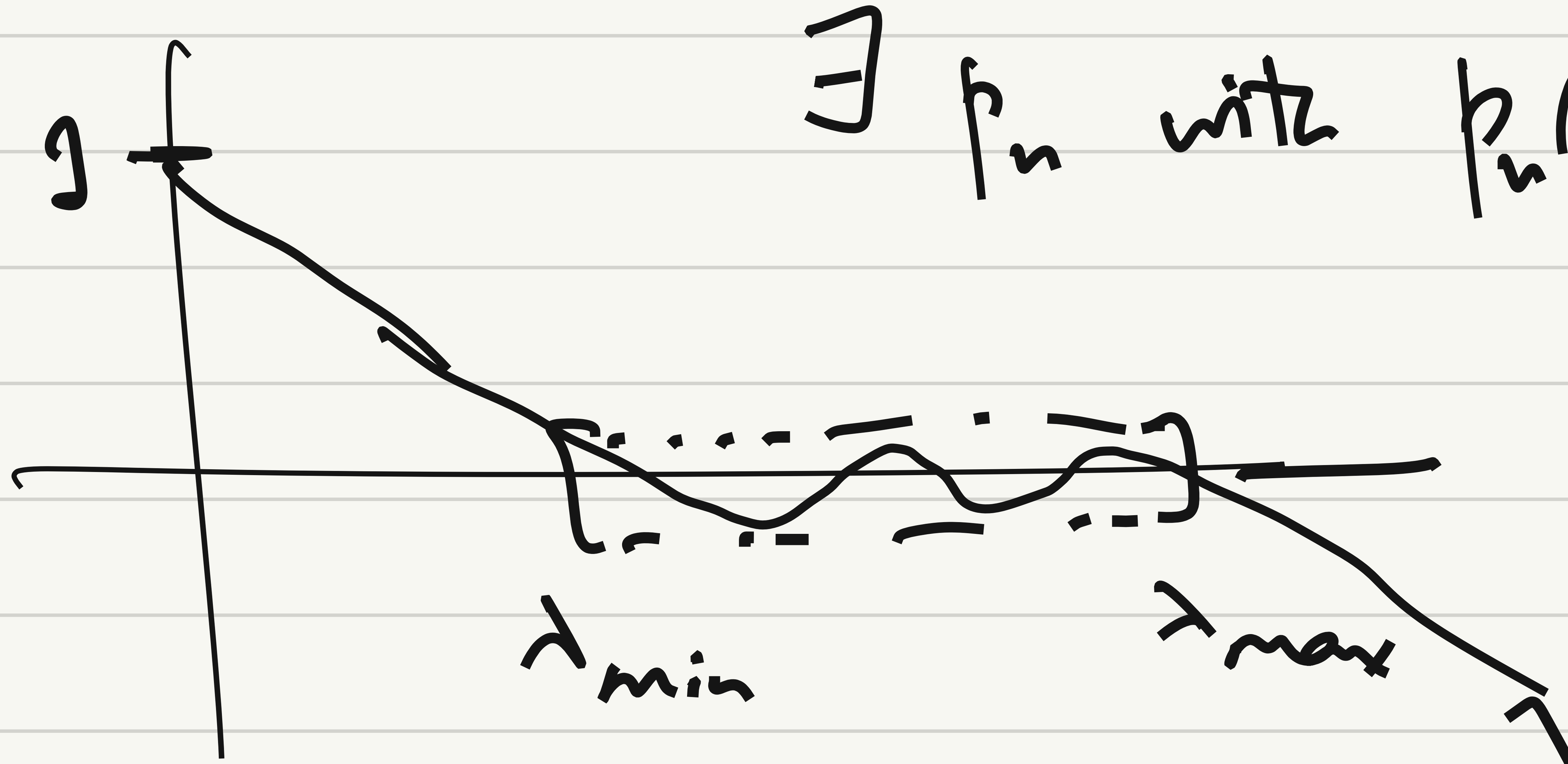
A is SPD $\Rightarrow$ eigenvalues are real, positive $\Rightarrow \|p(A)\|_A = \max_j |p(\lambda_j)|$

$$\|\vec{e}_n\|_A / \|\vec{e}_0\|_A \leq \max_j |p(\lambda_j)|$$

Two consequences:

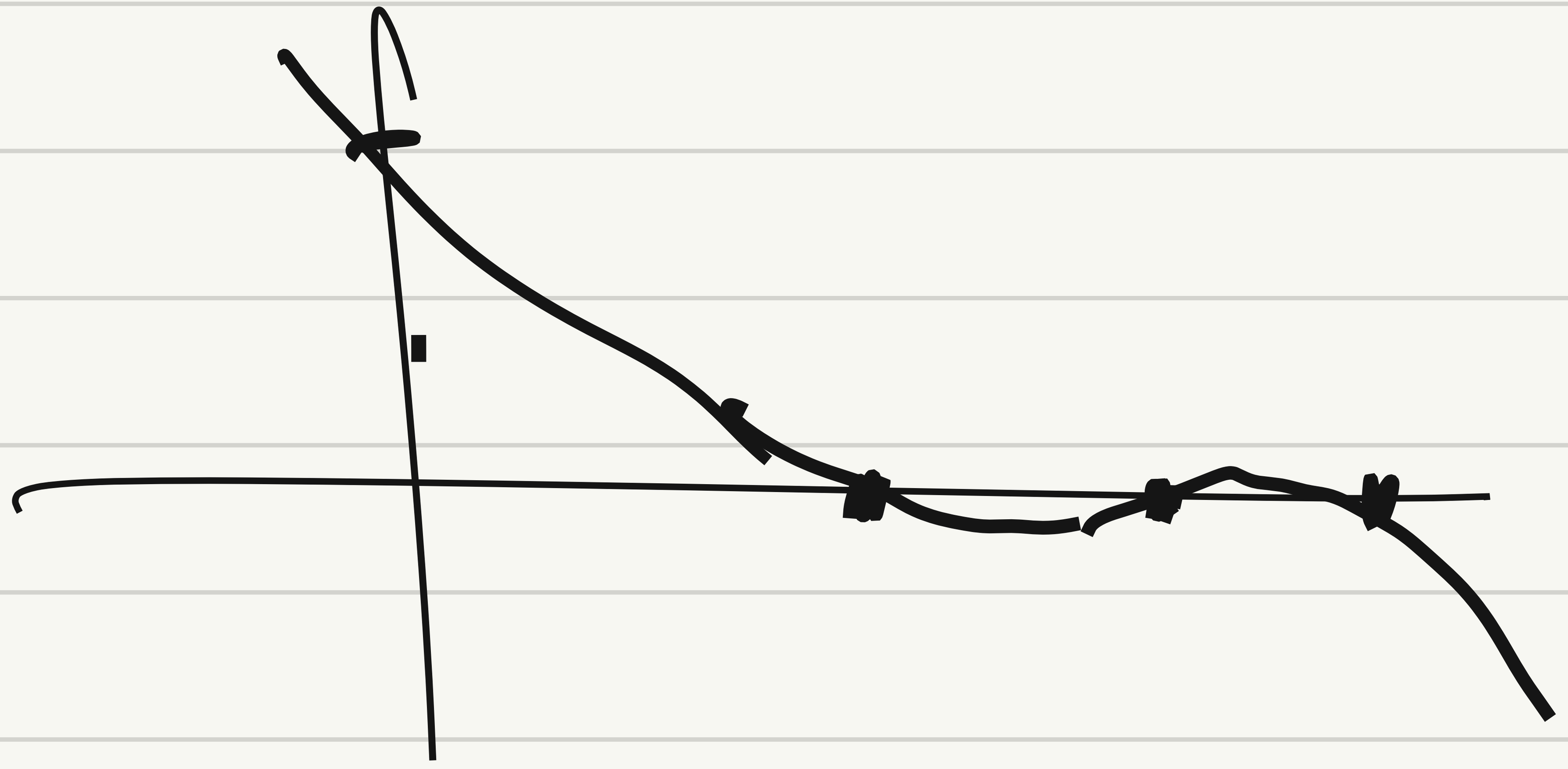
1. If only n distinct eigenvalue $\Rightarrow$ CG converges in $\le n$ iterations

2. If $\lambda_i \in [\lambda_{min}, \lambda_{max}]$ then $\kappa = \dfrac{\lambda_{max}}{\lambda_{min}}$



$\exists\ p_n$ with $p_n(0) = 1$,

$$p_n(x) \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n$$



$$\Rightarrow \frac{\|\vec{e}_n\|}{\|\vec{e}_0\|} \le 2\underbrace{\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)}_{\approx\, 1 - \frac{2}{\sqrt{\kappa}} \text{ for large } \kappa}^n$$

At each iter, error decreases by $\approx \frac{2}{\sqrt{\kappa}}$

Convergence to specified tolerance in $O(\sqrt{\kappa})$ iters