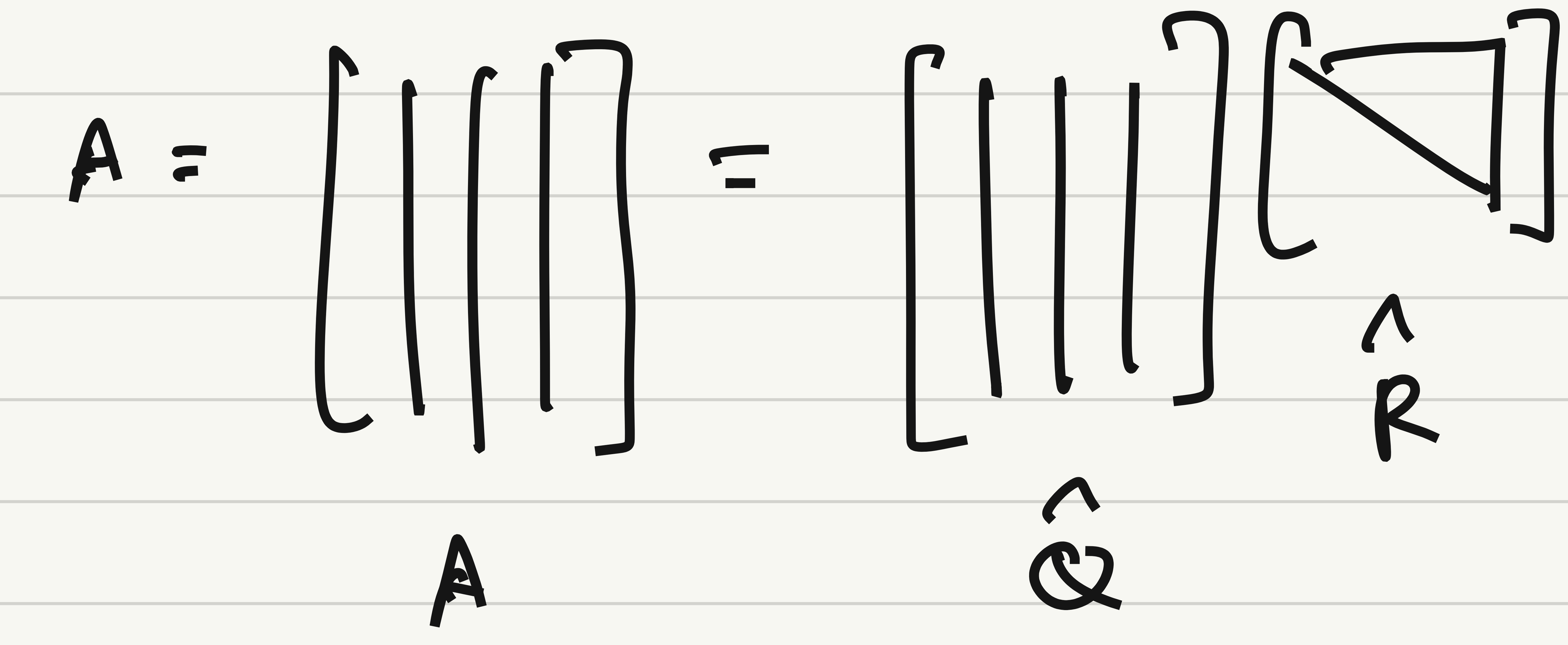
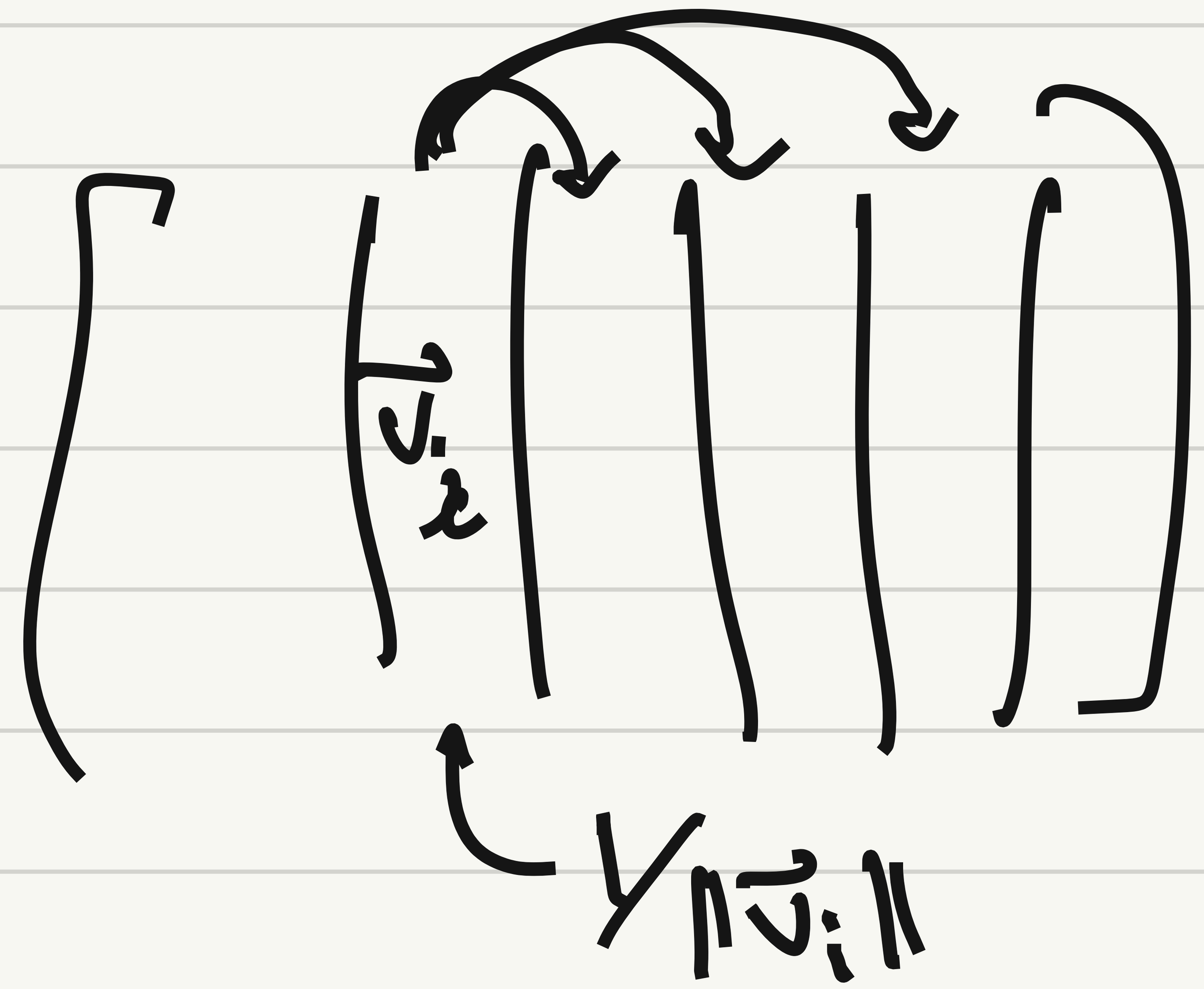
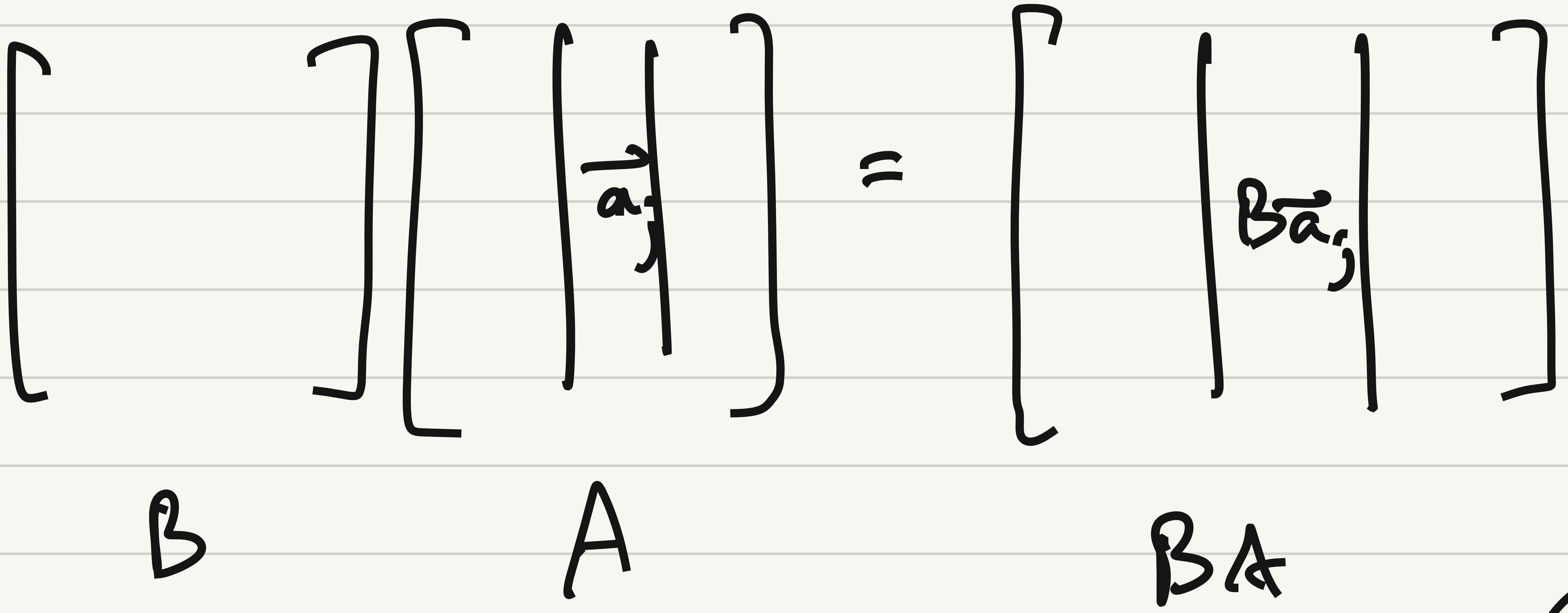
QR factorization

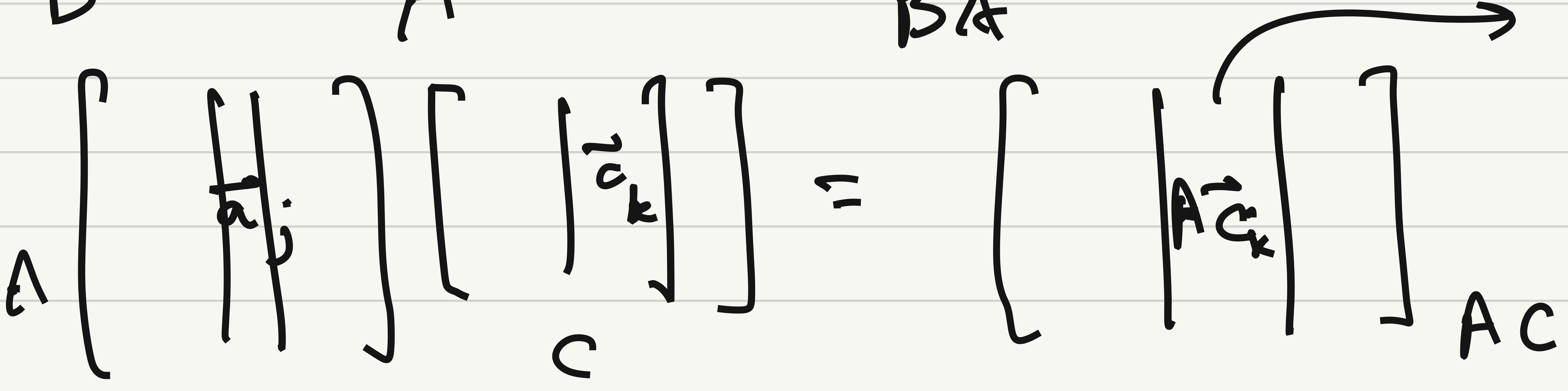


$A \to AR_1 \to AR_1R_2 \to \cdots \to AR_1\cdots R_n = \hat{Q}$





: each column gets transformed <u>independently</u>



$\sum \vec{a}_j c_{jk}$: columns of A get <u>mixed together</u>

R_i : $1/r_{ii}$, $-r_{ij}/r_{ii}$

$$A \xrightarrow{R_1} AR_1 \xrightarrow{R_2} \cdots \longrightarrow AR_1 \cdots R_n = \hat{Q}$$

$$\underbrace{A \xrightarrow{Q_1} Q_1 A \xrightarrow{Q_2} \cdots \longrightarrow Q_n \cdots Q_1 A = R}_{\boxed{\text{Householder triangularization}}}$$

$A = QR$

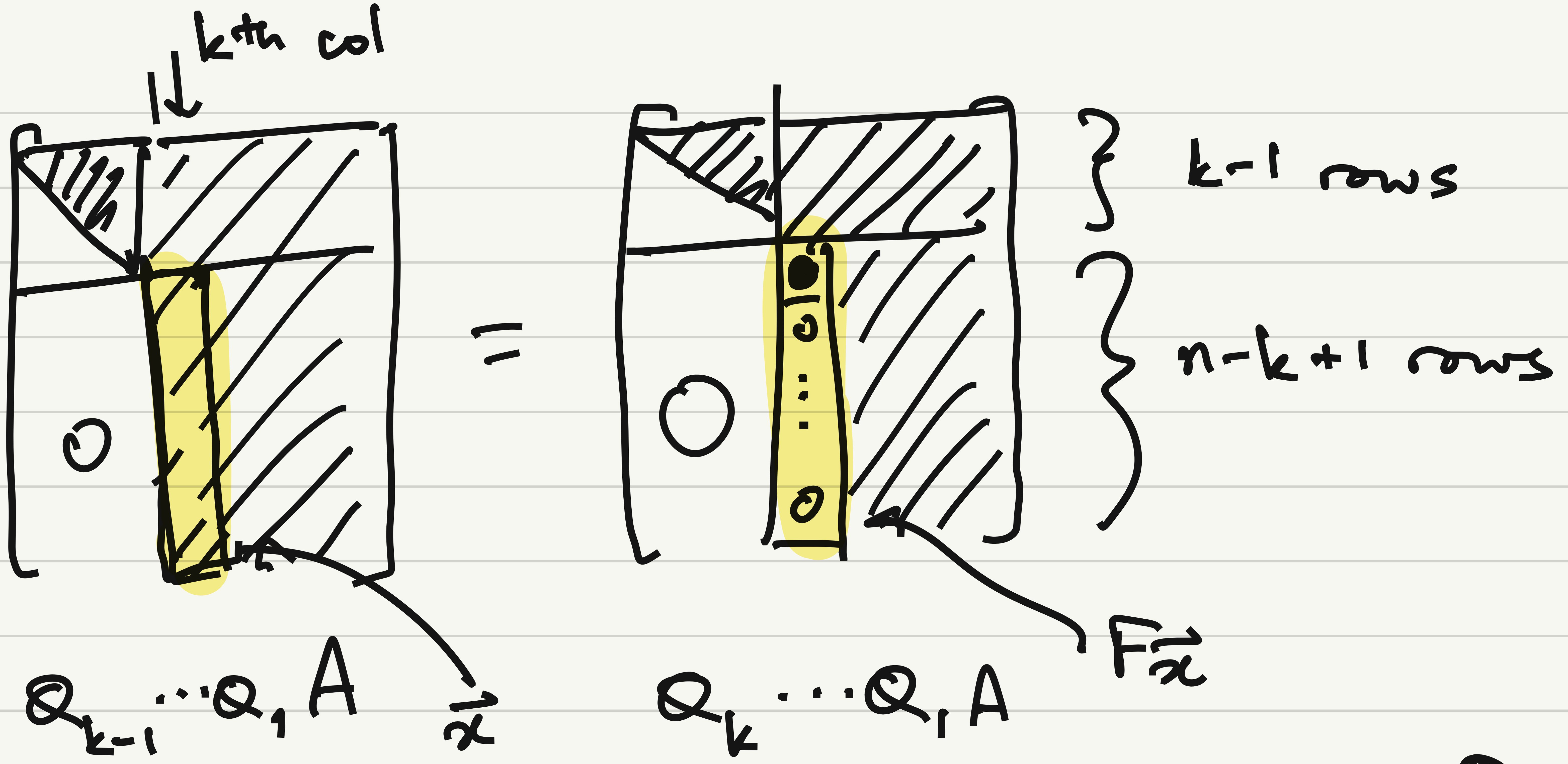
Best condition number for any matrix is 1

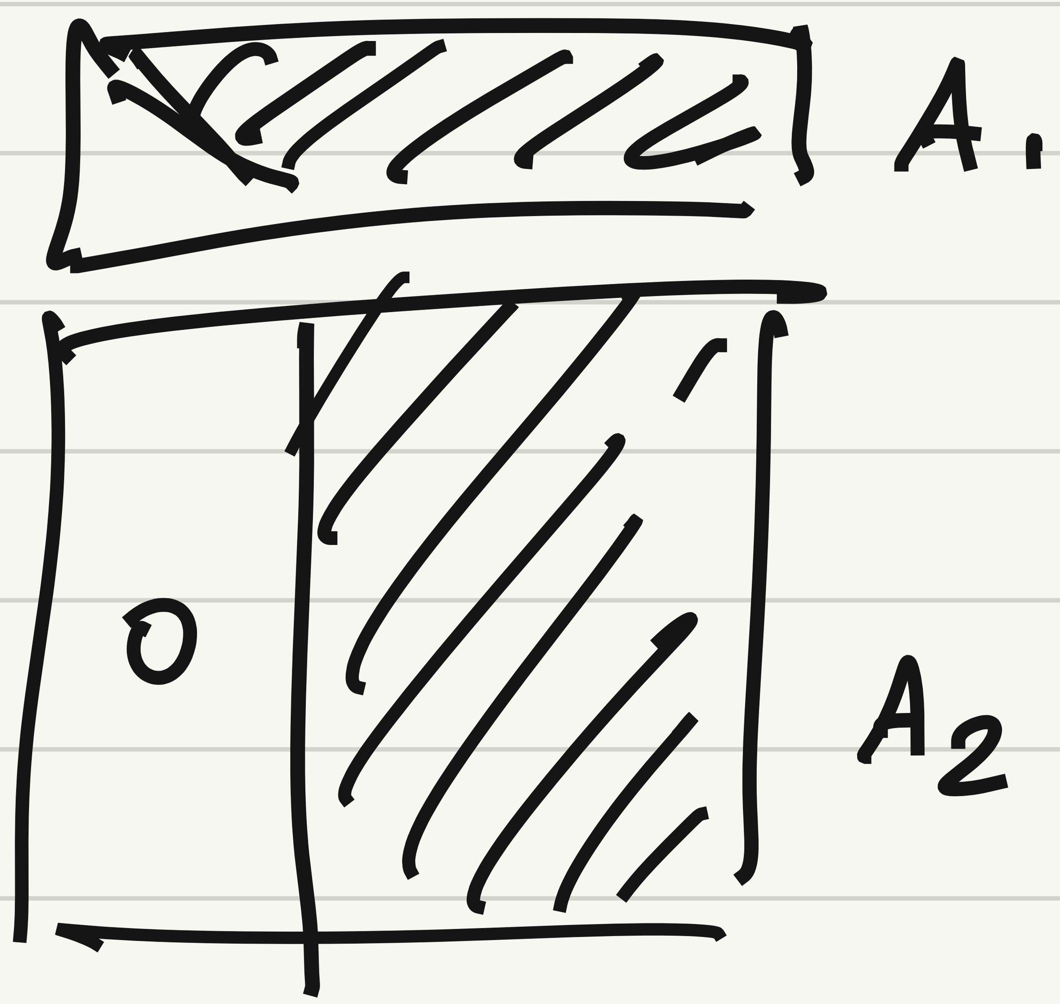
$$\kappa(A) = \sigma_1/\sigma_n \qquad \sigma_1 \geq \sigma_n$$

If Q is unitary, $\kappa(Q) = 1$

A

$$Q_k = \begin{array}{c} \begin{array}{cc} k-1 & n-k+1 \end{array} \\ \begin{bmatrix} I & 0 \\ 0 & F \end{bmatrix} \end{array}$$





$$Q_k A = \begin{bmatrix} IA_1 + \cancel{0A_2} \\ \\ \cancel{0A_1} + \underbrace{FA_2} \end{bmatrix}$$

$$x = \begin{bmatrix} \bullet \\ \bullet \\ \bullet \\ \vdots \\ \bullet \end{bmatrix} \Big\} n-k+1 \qquad F\vec{x} = \begin{bmatrix} \bullet \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

I need a unitary $\underline{F}$ s.t. $F\vec{x} =$

$$Q_k^* Q_k = \begin{bmatrix} I & \\ & \underbrace{F^*F}_{I} \end{bmatrix} = I$$

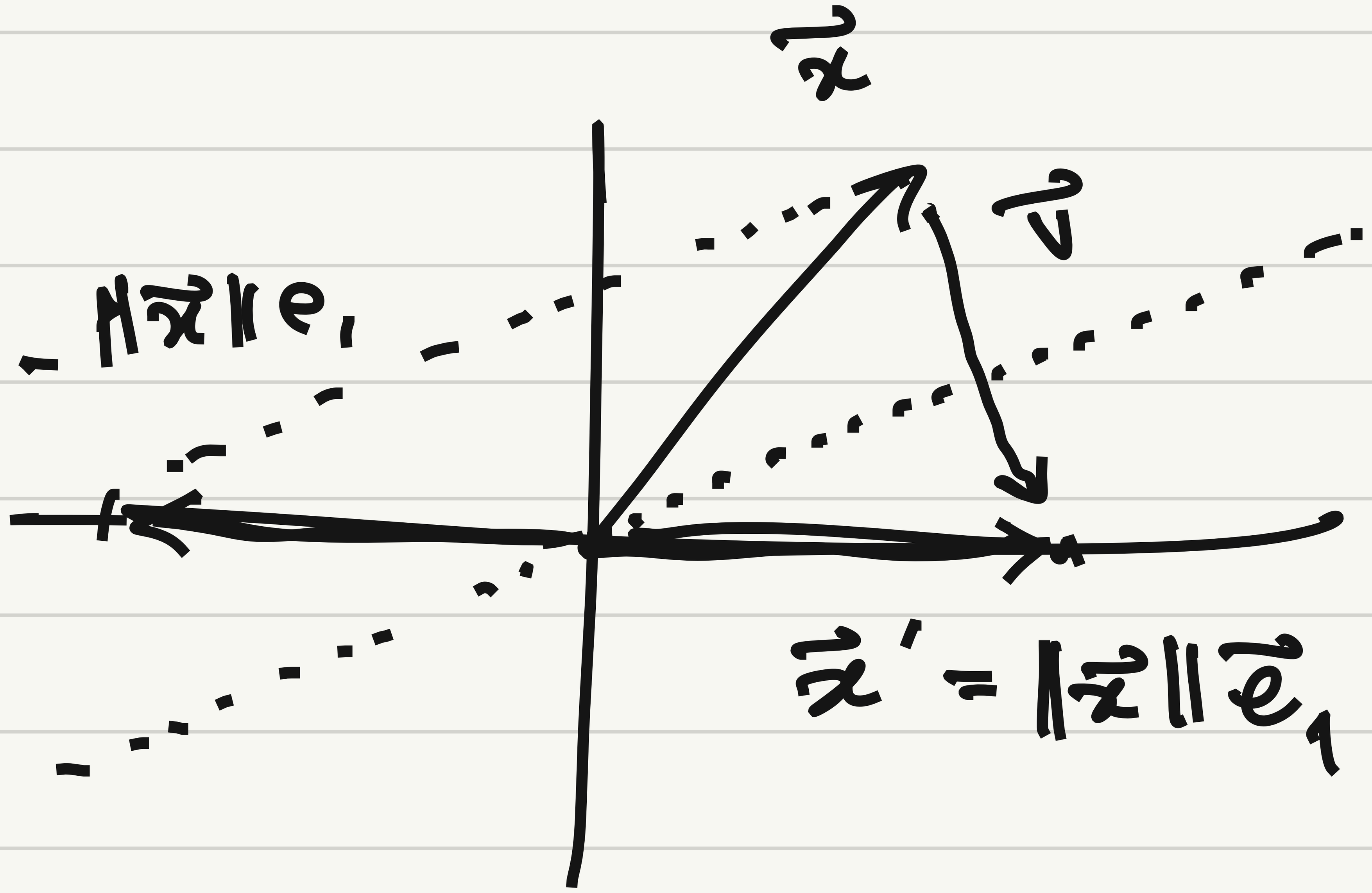
$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_\ell \end{bmatrix} \qquad \ell = n - k + 1$$

$$F\vec{x} = \begin{bmatrix} ? \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad ? \rightarrow z\|\vec{x}\|, \quad |z| = 1$$

F is unitary

$$\Rightarrow \|F\vec{x}\|_2 = \|\vec{x}\|_2$$

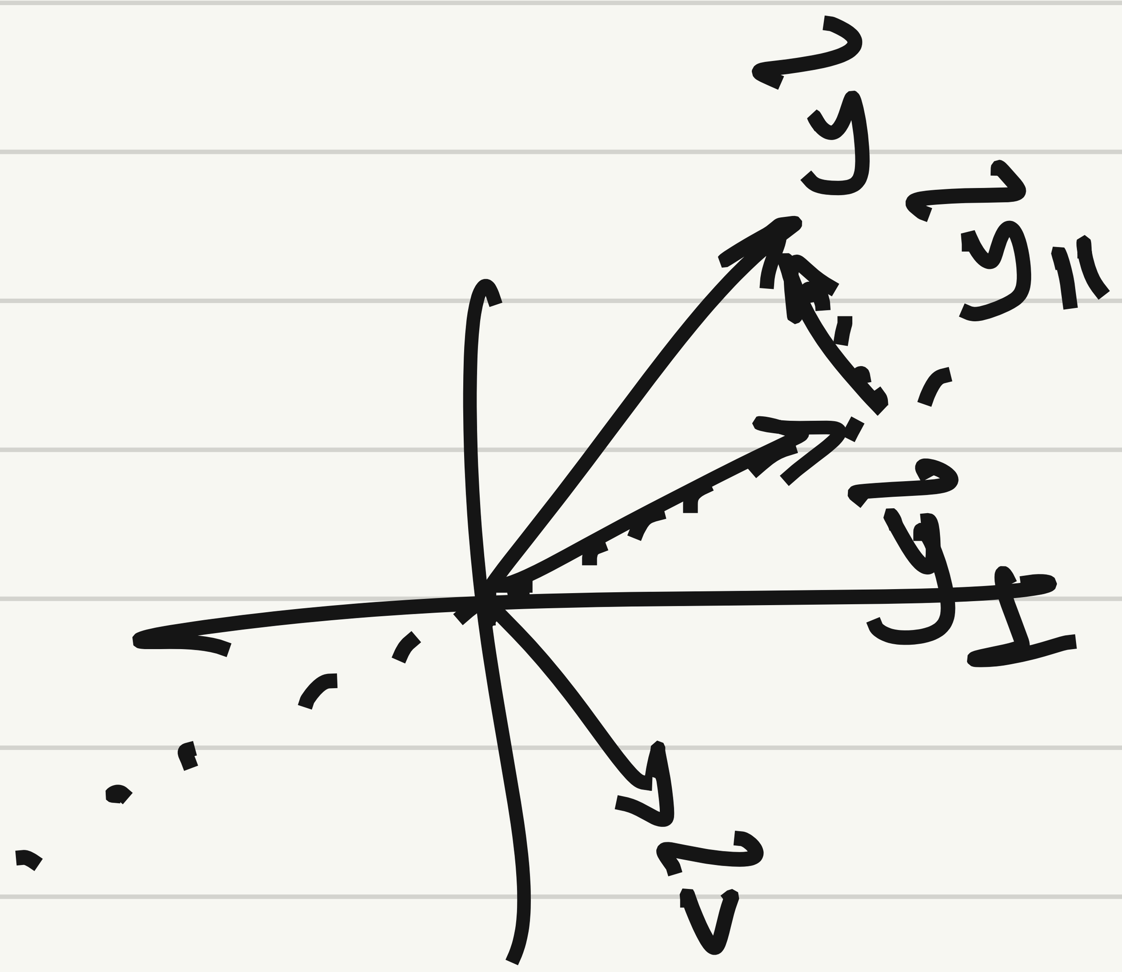
multiple of $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$



$$\vec{v} = \vec{x}' - \vec{x}$$

$$F = \underbrace{I - 2\frac{\vec{v}\vec{v}^*}{\vec{v}^*\vec{v}}}$$

To maximize $\|\vec{v}\|$, choose $\vec{x}' = -\text{sign}(x_1)\|\vec{x}\|\vec{e}_1$



$$\vec{y}_\parallel = P_v\vec{y} = \frac{\vec{v}\vec{v}^*}{\vec{v}^*\vec{v}}\vec{y}$$

$$\vec{y}_\perp = \left(I - \frac{\vec{v}\vec{v}^*}{\vec{v}^*\vec{v}}\right)\vec{y}$$

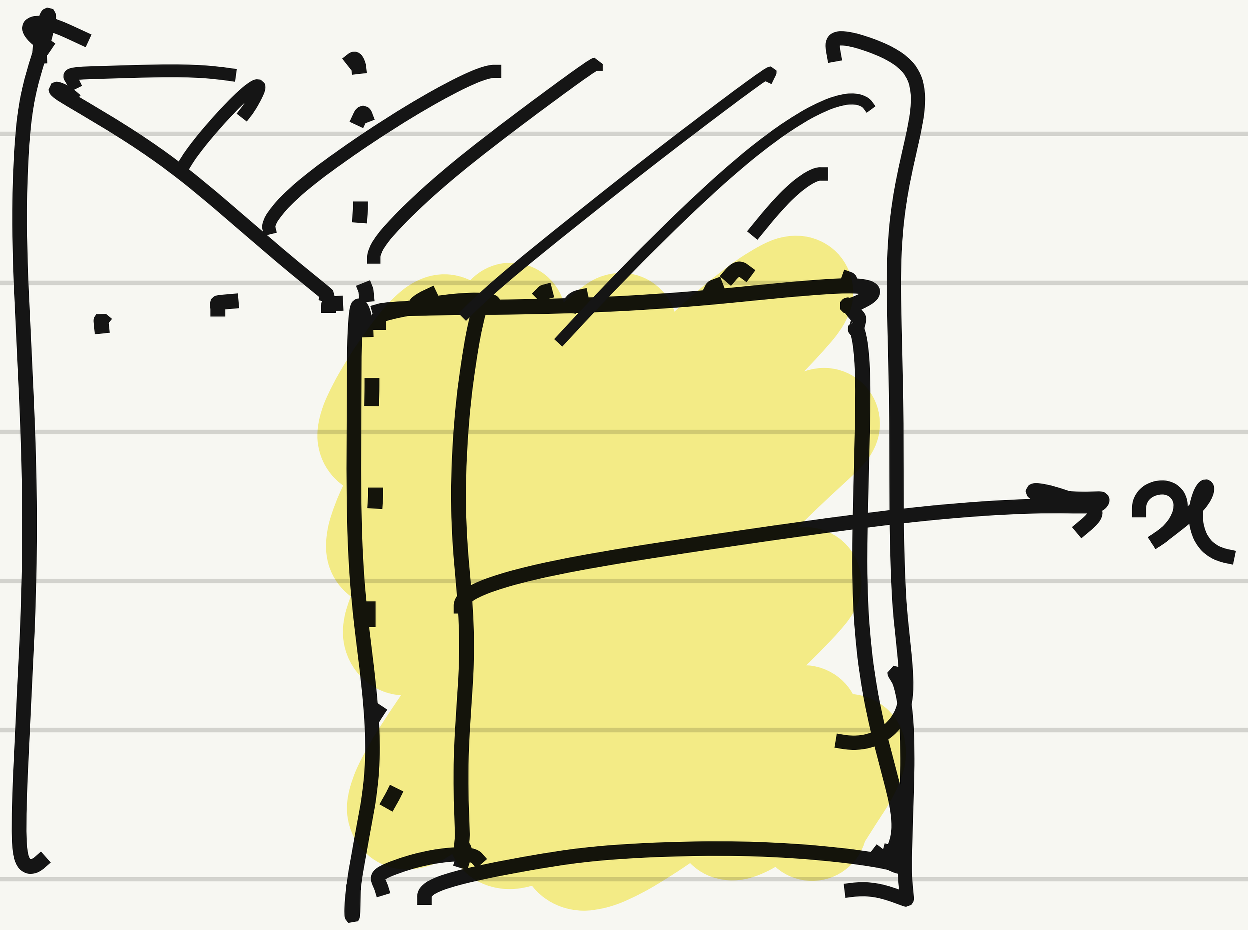
$$F\vec{y} = \vec{y}_\perp - \vec{y}_\parallel$$

for $k = 1 \ldots n$:

  let $\vec{x} = A[k:m, k]$

  construct reflection vector $\vec{v}_k$

  Apply reflection to remaining block of A



A

$$\longrightarrow A[k:m, k:n] \mathrel{-}= \frac{2\vec{v}_k \vec{v}_k^{*}}{\vec{v}_k^{*}\vec{v}_k} A[k:m, k:n]$$

Operation count $\sim 2\left(m - \frac{n}{3}\right)n^3$ flops.

$$\underbrace{\begin{bmatrix} v_1 \end{bmatrix} \begin{bmatrix} v_2 \end{bmatrix} \cdots \begin{bmatrix} v_n \end{bmatrix}}_{\text{implicit representation of } Q}, \quad \begin{bmatrix} R \end{bmatrix}$$

$$Q_n \cdots Q_1 A = R$$

$$A = \underbrace{(Q_n \cdots Q_1)^{-1}}_{Q^{-1}} R = \underbrace{Q_1^{-1} \cdots Q_n^{-1}}_{Q} R$$

Accuracy & Stability of Householder

QR using Householder is not accurate but is backward stable

Thm: If QR factorization of A is computed with Householder $\to \tilde{Q}, \tilde{R}$

then $\dfrac{\|\tilde{Q}\tilde{R} - A\|}{\|A\|} = O(\varepsilon_m)$.

$\tilde{Q}\tilde{R}$ is the QR factorization of slightly perturbed input $A + \delta A$ with $\dfrac{\|\delta A\|}{\|A\|} = O(\varepsilon_m)$

eg. solve $A\vec{x} = \vec{b}$, A is square, full rank

$QR = A \leftarrow$ Householder

$\vec{y} = Q^* \vec{b} \longleftarrow$ apply reflectors to $\vec{b}$

$\vec{x} = R^{-1}\vec{y} \leftarrow$ backsubstitution

$$\left.\begin{array}{l} \tilde{Q}, \tilde{R} = \text{house}(A) \\ \tilde{y} = \tilde{Q}^* \odot \vec{b} \\ \tilde{x} = \text{backsub}(\tilde{R}, \tilde{y}) \end{array}\right\}$$

$$\tilde{Q}, \tilde{R} = \text{house}(A)$$
$$\tilde{y} = \tilde{Q}^* \odot \vec{b}$$
$$\tilde{x} = \text{backsub}(\tilde{R}, \tilde{y})$$

$$\left[\ \tilde{Q}\tilde{R} = A + \delta A, \qquad \|\delta A\| / \|A\| = O(\varepsilon_m)\right.$$
$$\left[\ (\tilde{Q} + \delta Q)\tilde{y} = b, \qquad \|\delta Q\| = O(\varepsilon_m)\right.$$
$$\left[\ (\tilde{R} + \delta R)\hat{x} = \tilde{y}, \qquad \|\delta R\| / \|R\| = O(\varepsilon_m)\right.$$

$$\tilde{x} = (\tilde{R} + \delta R)^{-1}\tilde{y} = (\tilde{R} + \delta R)^{-1}(\tilde{Q} + \delta Q)^{-1} b$$

$$(\tilde{Q} + \delta Q)(\tilde{R} + \delta R)\tilde{x} = b$$

$$(\underbrace{\tilde{Q}\tilde{R}}_{= A + \delta A} \underbrace{+ \delta Q \tilde{R} + \tilde{Q}\delta R + \delta Q \delta R}_{\Delta A})\tilde{x} = b$$

Show that $\dfrac{\|\Delta A\|}{\|A\|} = O(\varepsilon_m)$

$$(A + \Delta A)\tilde{x} = b$$

$\iff$ [show that $\|\delta A\|$, $\|\delta Q \tilde{R}\|$, $\|\tilde{Q}\delta R\|$, $\|\delta Q \delta R\|$ are all $O(\|A\| \varepsilon_m)$

$$\|\delta Q \hat{\tilde{R}}\|$$

$$\|\delta Q\| = O(\varepsilon_m)$$

$$\leq \underbrace{\|\delta Q\|}_{O(\varepsilon_m)} \cdot \underbrace{\|\tilde{R}\|}_{\approx \|A\|} = \|A\| \, O(\varepsilon_m)$$

$$\frac{\|\tilde{R}\|}{\|A\|} = O(1)$$

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

$$= \|A\| \cdot (1 + O(\varepsilon_m))$$

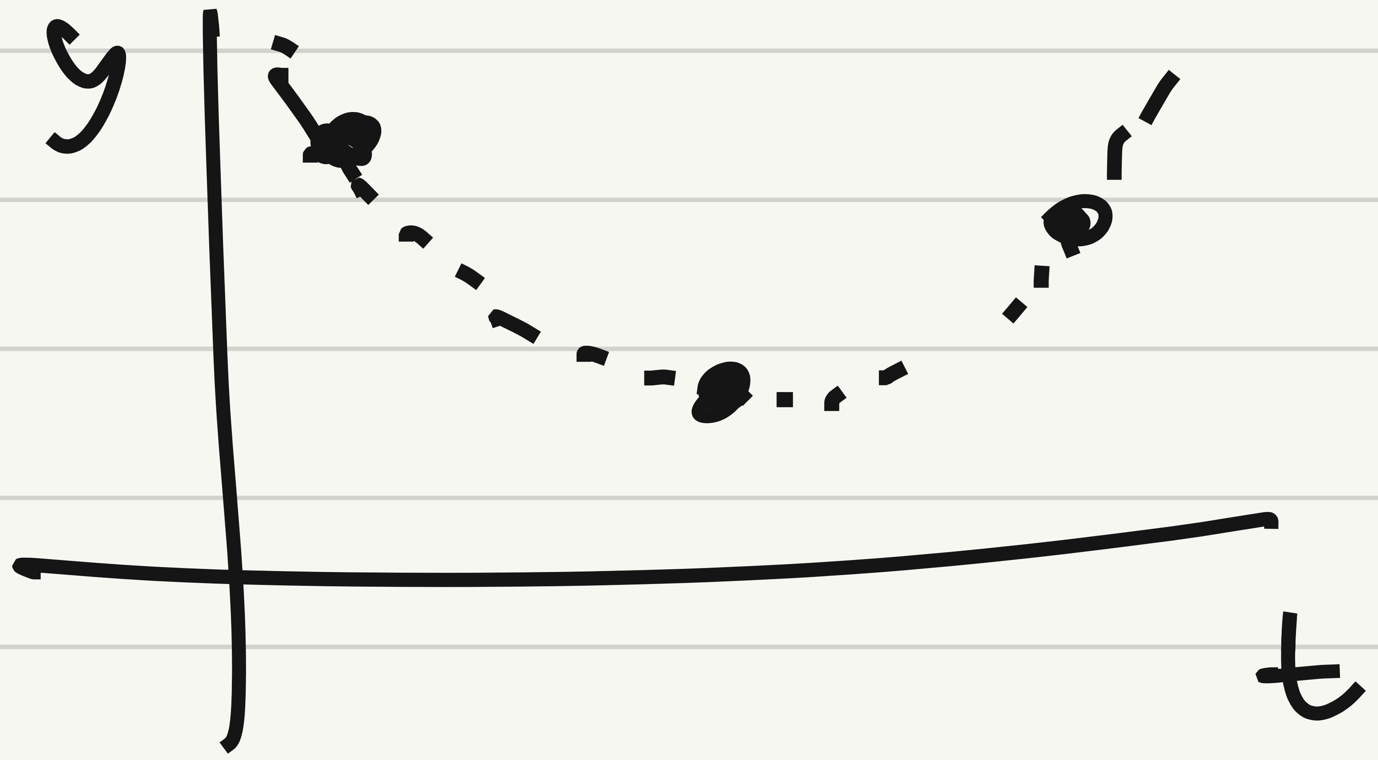
Solving $A\vec{x} = \vec{b}$ using Householder QR is backward stable.

Backward stable $\Rightarrow$ r.f.e. $= O(\kappa \, \varepsilon_m)$

Condition number of solving $Ax = b$ $= \kappa(A)$

$\Rightarrow$ r.f.e $= O(\kappa(A) \, \varepsilon_m)$

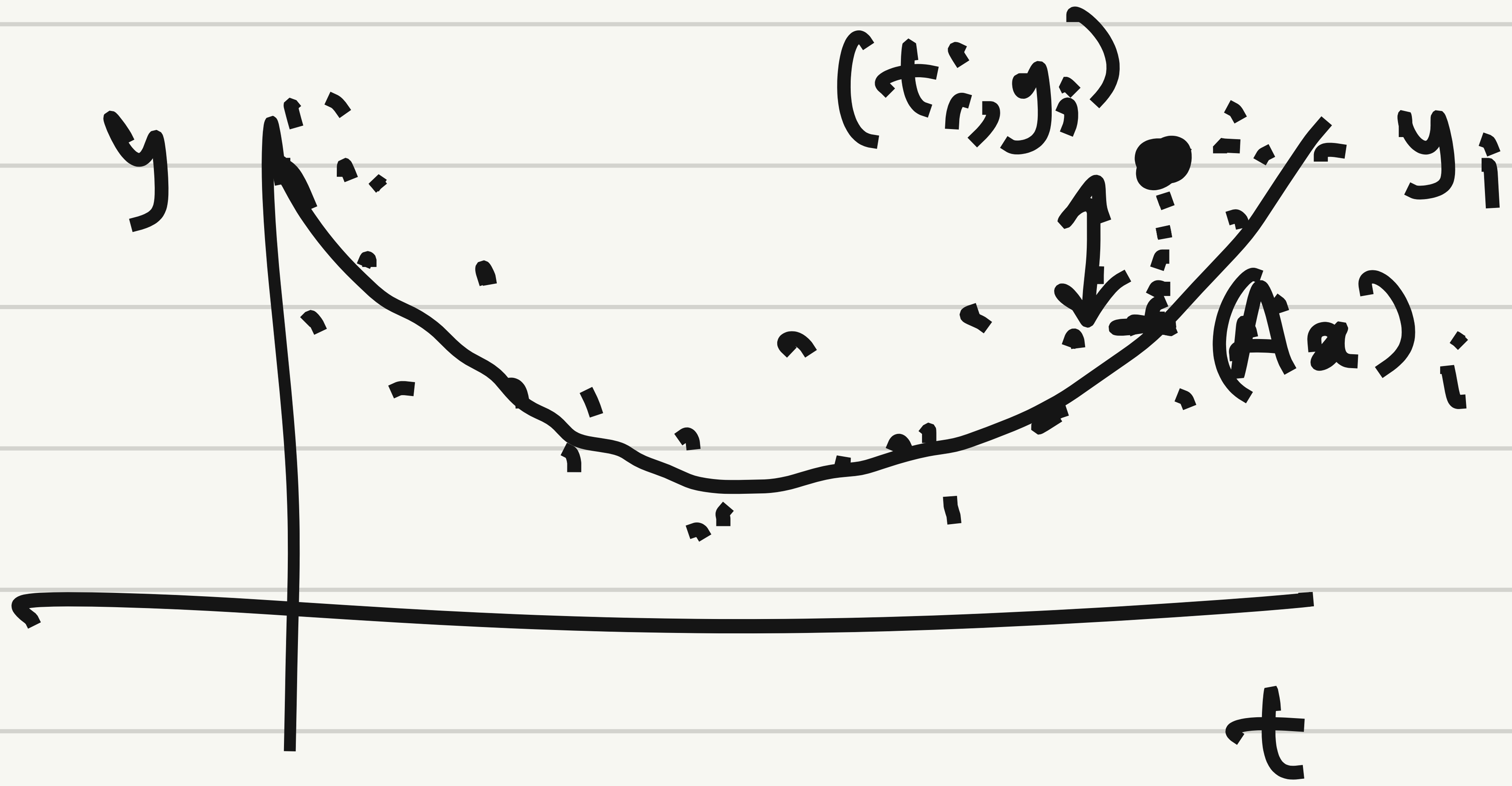
least-squares problems



$$y(t) = at^2 + bt + c$$

$$a t_1^2 + b t_1 + c = y_1$$
$$a t_2^2 + b t_2 + c = y_2$$
$$a t_3^2 + b t_3 + c = y_2$$

$$\begin{bmatrix} t_1^2 & t_1 & 1 \\ t_2^2 & t_2 & 1 \\ t_3^2 & t_3 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$



$$\underbrace{\begin{bmatrix} t_1^2 & t_1 & 1 \\ \vdots & & \\ t_m^2 & t_m & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{\vec{b}}$$

$A\vec{x} = \vec{b}$: no sol.

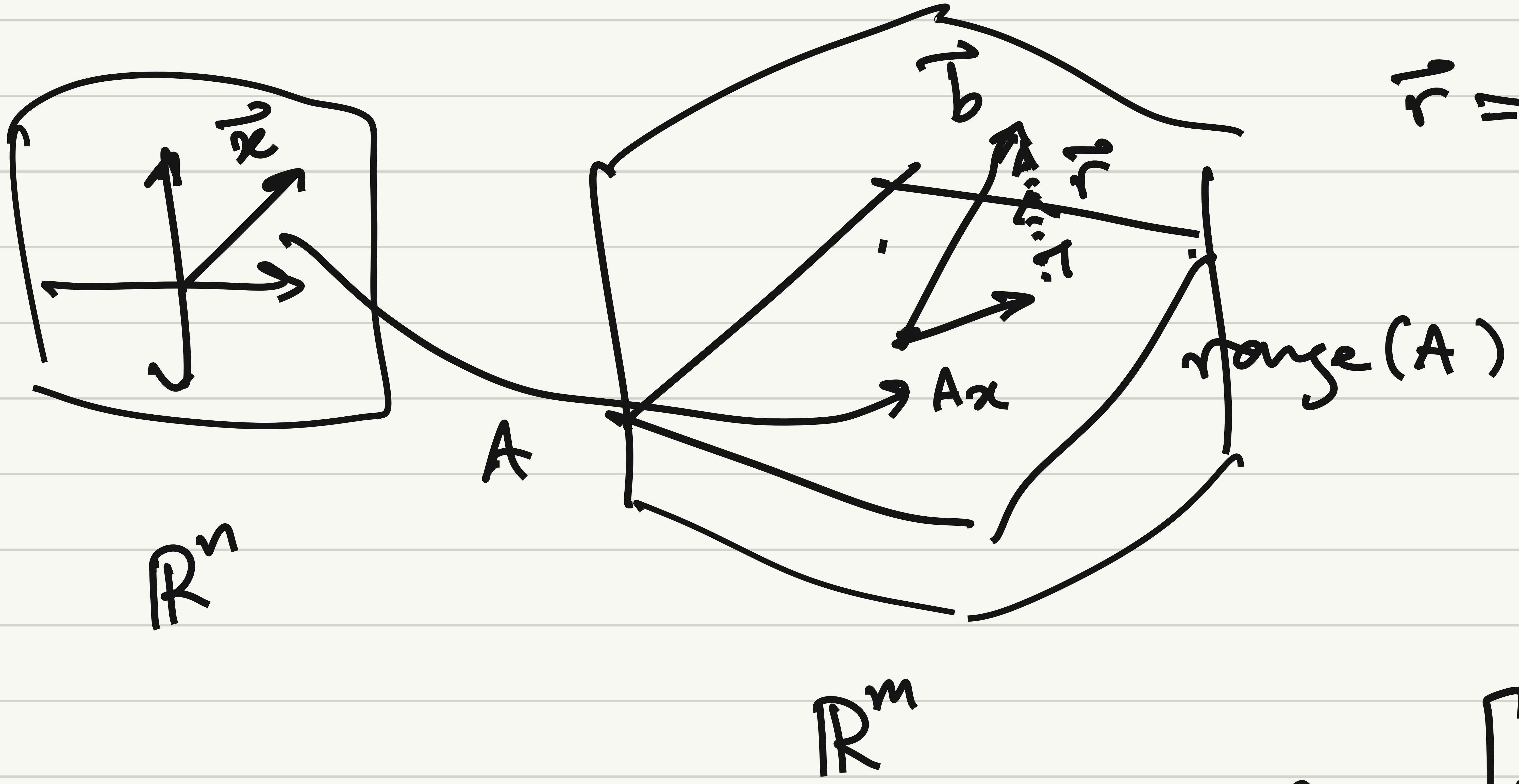
$\vec{r} = \vec{b} - A\vec{x}$

find $\vec{x}$ to minimize $\|\vec{r}\|$

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

minimize $\|\vec{r}\|_2 = \sqrt{r_1^2 + \cdots + r_m^2} \iff$ minimize $r_1^2 + \cdots + r_m^2$: $\boxed{\text{least squares problem}}$

find $\vec{x}$ to get $A\vec{x}$ as close as possible to $\vec{b}$.



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

$\vec{r} \perp \text{range}(A) = \langle \vec{a}_1, \dots, \vec{a}_n \rangle$

$\iff \vec{r} \perp \vec{a}_j \quad \forall j$

$\iff \vec{a}_j^{\,*} \vec{r} = 0 \quad \forall j$

$$\iff \underbrace{\begin{bmatrix} \vec{a}_1^{\,*} \\ \hline \vdots \\ \hline \vec{a}_n^{\,*} \end{bmatrix}}_{A^*} \vec{r} = 0 \quad \Rightarrow \quad A^* \vec{r} = 0$$

$A^*(\vec{b} - A\vec{x}) = 0 \iff \underbrace{A^*A\vec{x} = A^*\vec{b}}$ $\boxed{\text{normal equations}}$

$A \in \mathbb{C}^{m\times n}$, $m \geq n$

$\vec{x}$ solves least-sq. prob $\iff$ $\vec{x}$ solves normal eq.

A is full rank $\iff$ A^*A is invertible

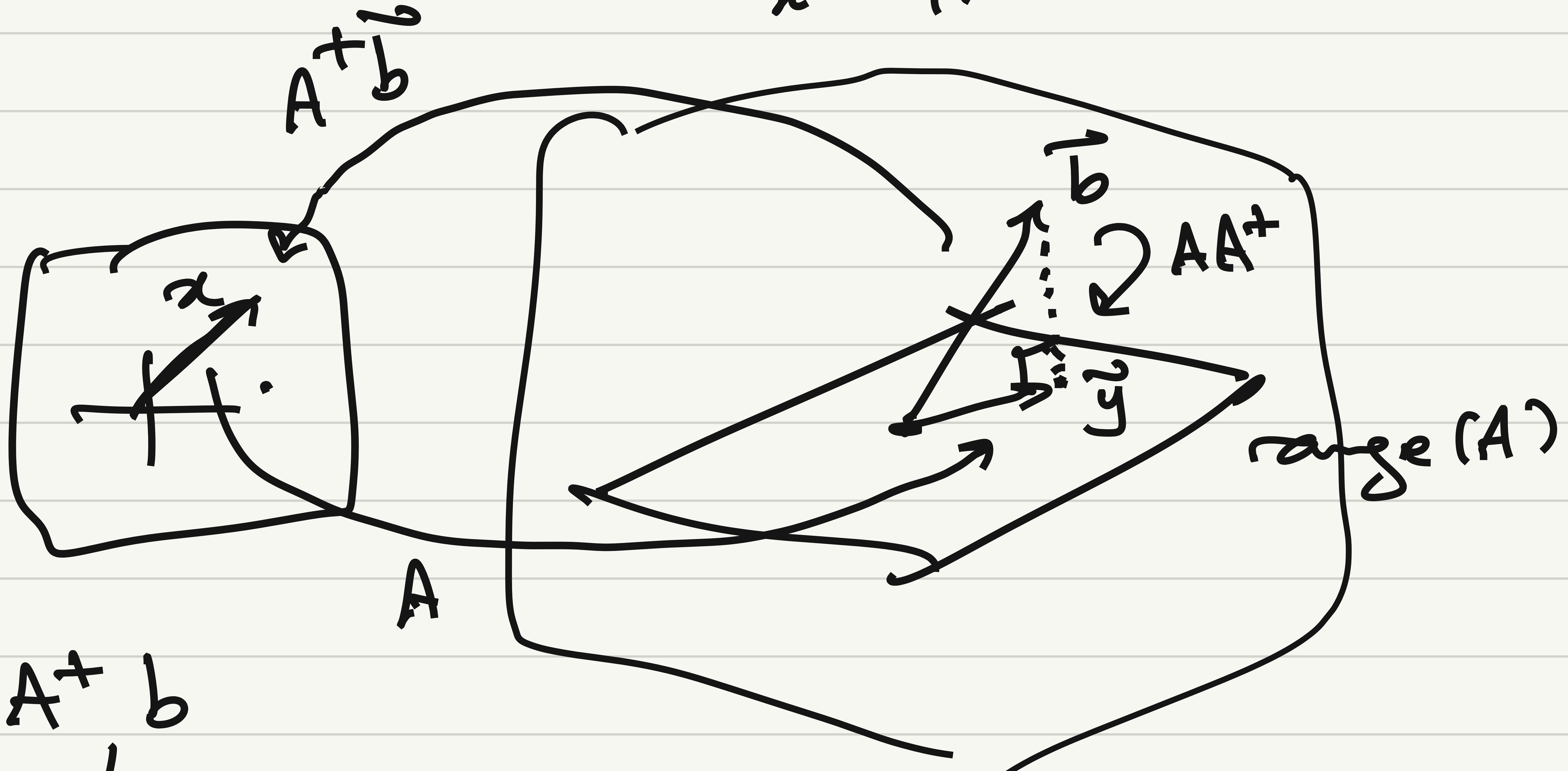
$\vec{x} = \underbrace{(A^*A)^{-1}A^*}\vec{b}$

$\boxed{\text{pseudoinverse}}$ A^+

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

$\vec{x} = A^{-1}\vec{b}$

$A^+A = \underbrace{(A^*A)^{-1}}\underbrace{A^*A} = I$

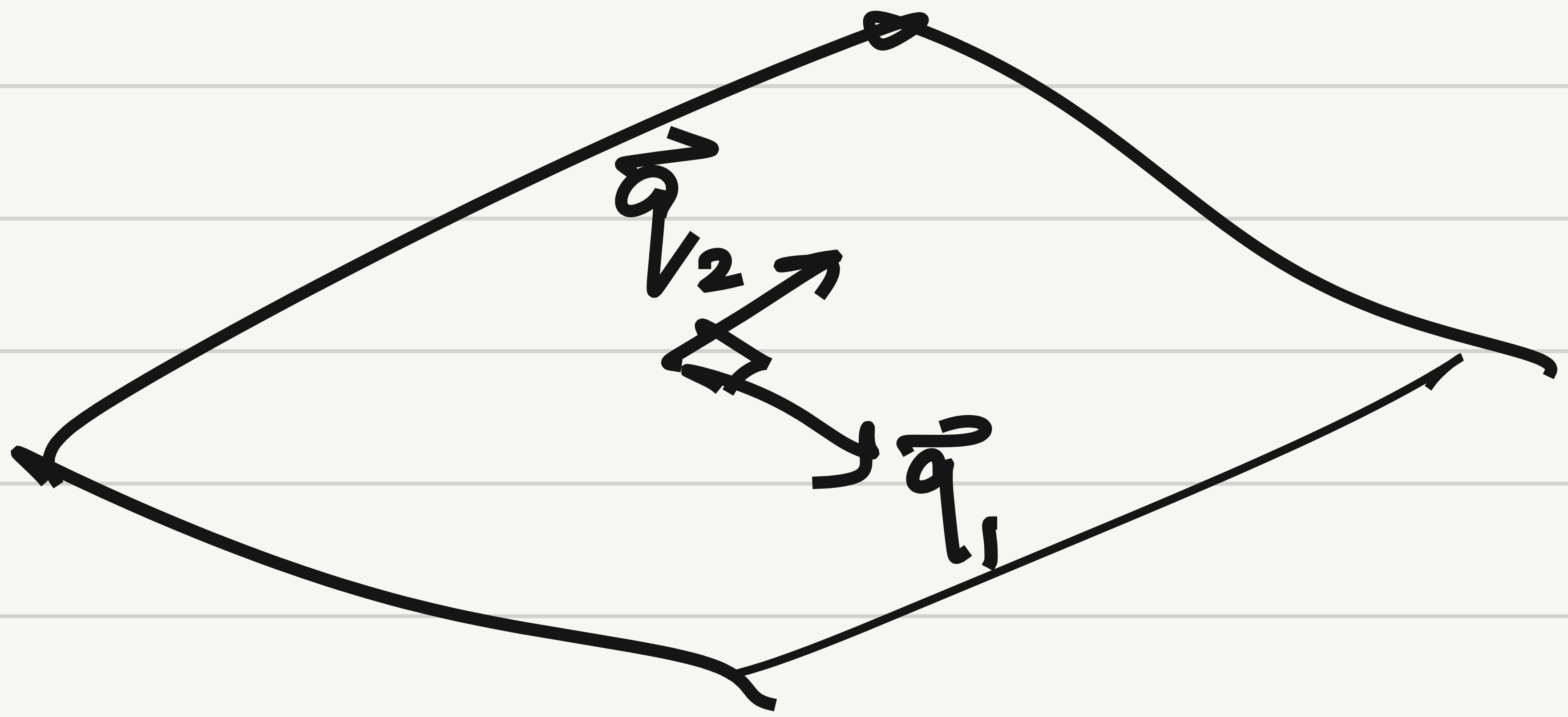


$AA^+ \neq I$ but is useful!

$\vec{y}$ = closest point to $\vec{b}$ = $A\vec{x}$ = $\underbrace{AA^+}b$

AA^+ is the orthogonal projector onto range(A)

$$\underbrace{A(A^*A)^{-1}A^*}$$



orthogonal proj $= QQ^*$ $\quad \left(= Q(\underbrace{Q^*Q}_{I})^{-1}Q \right)$

Solving l.sq. problems

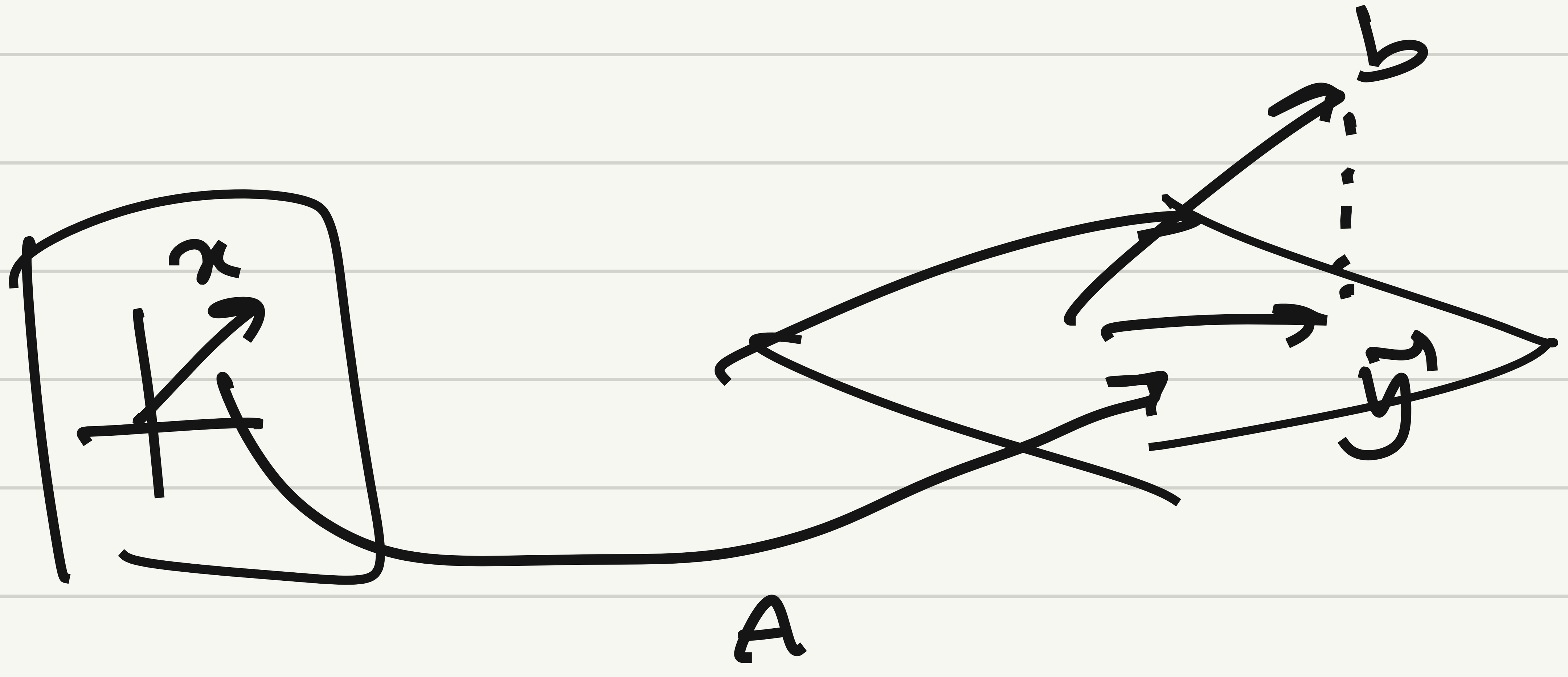
$$\kappa(A^*A) = (\kappa(A))^2$$

① solve normal eqs. $\underbrace{A^*A}\vec{x} = \underbrace{A^*b} \longrightarrow$ unstable!

② QR factorization: $A = \hat{Q}\hat{R}$

range($\hat{Q}$) = range(A)

Projection $\vec{y} = \hat{Q}\hat{Q}^* \vec{b}$



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

$\cancel{\hat{Q}}\hat{R}\vec{x} = \vec{y} = \cancel{\hat{Q}}\hat{Q}^* \vec{b}$

$\hat{R}\vec{x} = \hat{Q}^* \vec{b}$

exactly same as solving $A\vec{x} = \vec{b}$

Backward stable!

(2) SVD: $A = \hat{U}\hat{\Sigma}\hat{V}^*$