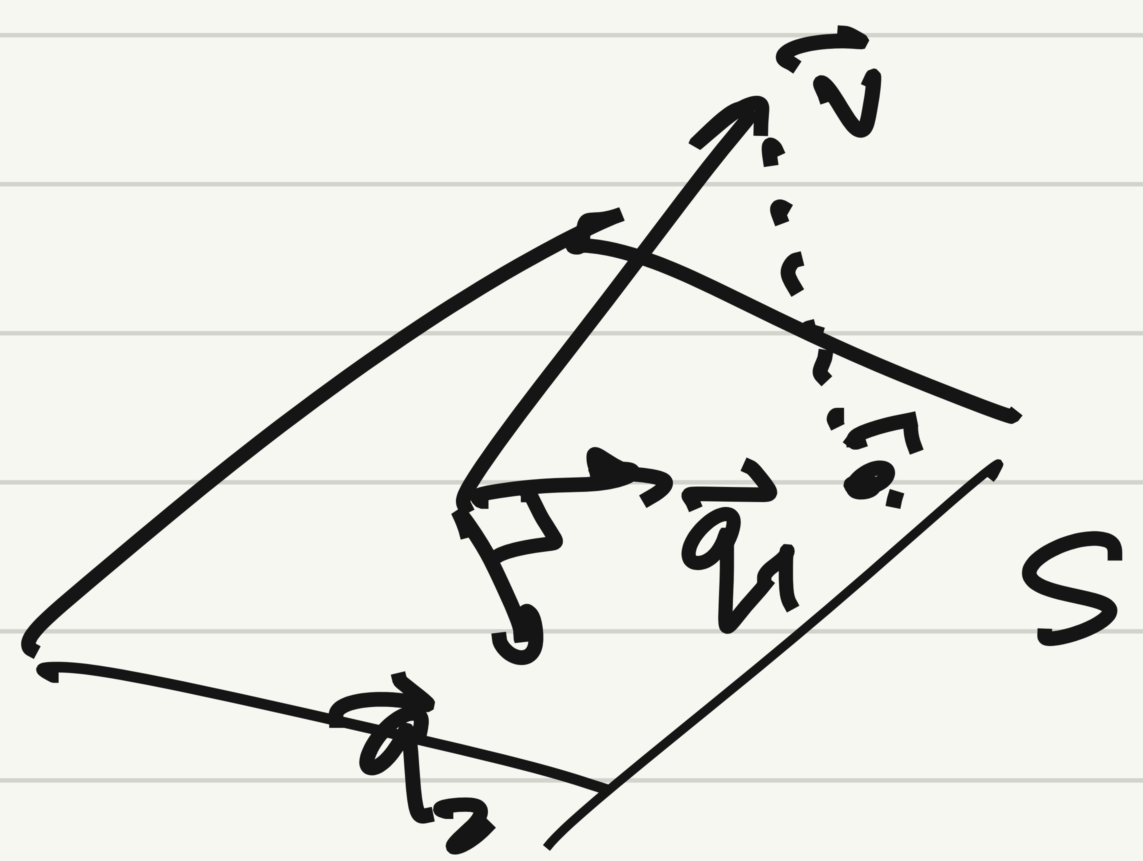
COL 726: QR factorization



$$\overbrace{\langle \vec{q}_1, \ldots, \vec{q}_n \rangle}$$

$$P\vec{v} = \underbrace{QQ^*}_{P}\vec{v}$$

$$Q = \begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_n \end{bmatrix}$$

$$S = \text{range}(A) = \langle \vec{a}_1, \ldots, \vec{a}_n \rangle \xrightarrow{?} \langle \vec{q}_1, \ldots, \vec{q}_n \rangle$$

P_S : ortho. proj. onto S

$A \in \mathbb{C}^{m \times n}$ full rank

$\boxed{\text{Gram-Schmidt}}$

$$\vec{a}_1 \longrightarrow \vec{a}_1 / \|\vec{a}_1\|_2 = \vec{q}_1$$

$$\vec{a}_2 \longrightarrow P_{\langle q_1 \rangle^\perp} \vec{a}_2 = \vec{a}_2 - \underbrace{\vec{q}_1 \vec{q}_1^*}\, \vec{a}_2 = \vec{v}_2 \longrightarrow \frac{\vec{v}_2}{\|\vec{v}_2\|_2} = q_2$$

$$\vec{a}_3 \to P_{\langle \vec{q}_1, \vec{q}_2 \rangle^\perp} \vec{a}_3 = \vec{a}_3 - q_1 q_1^* a_3 - q_2 q_2^* a_3 = \vec{v}_2 \to \,:$$

$\vec{a}_j \longrightarrow \vec{v}_j = \vec{a}_j - \sum_{i<j} \vec{q}_i \vec{q}_i^{\,*} \vec{a}_j \longrightarrow \vec{q}_j = \vec{v}_j / \|\vec{v}_j\|_2$

for each $j = 1, \dots, n$

$\quad \vec{v}_j = \vec{a}_j$

$\quad$ for each $i = 1, \dots, j-1$

$\qquad \vec{v}_j \mathrel{-}= \vec{q}_i \vec{q}_i^{\,*} \vec{a}_j$

$\quad \vec{q}_j = \vec{v}_j / \|\vec{v}_j\|$

$\boxed{\text{classical Gram-Schmidt orthogonalization}}$

$\vec{v}_3 = P_{\langle q_2 \rangle^\perp} P_{\langle q_1 \rangle^\perp} \vec{a}_3$

$\boxed{\text{Modified Gram-Schmidt}}$

$\vec{v}_j = P_{\langle \vec{q}_1, \dots, \vec{q}_{j-1} \rangle^\perp} \vec{a}_j = P_{\langle q_{j-1} \rangle^\perp} \cdots P_{\langle q_2 \rangle^\perp} P_{\langle q_1 \rangle^\perp} \vec{a}_j$

for each $i = 1, \dots, j-1$

$\quad \vec{v}_j \mathrel{-}= \vec{q}_i \vec{q}_i^{\,*} \vec{v}_j$

Mathematically equivalent,
Numerically more stable!

$$\vec a_1 \longrightarrow \vec v_1 = \vec a_1 \longrightarrow \vec q_1 = \vec v_1 / \|\vec v_1\|$$

$$\vec a_2 \longrightarrow P_{\langle q_1\rangle^\perp} a_1 = \vec v_2 \longrightarrow \vec q_2 = \vec v_2 / \|\vec v_2\|$$

$$\vec a_3 \longrightarrow P_{\langle q_1\rangle^\perp} \vec a_3 \longrightarrow P_{\langle \vec q_2\rangle^\perp} P_{\langle q_1\rangle^\perp} \vec a_3 = \vec v_3 \longrightarrow \vec q_3$$

$$\vdots$$

$$\vec a_j \longrightarrow P_{\langle q_1\rangle^\perp} a_j \longrightarrow P_{\langle q_2\rangle^\perp} P_{\langle q_1\rangle^\perp} a_j \longrightarrow \cdots \longrightarrow P_{\langle q_{j-1}\rangle^\perp} \cdots P_{\langle q_1\rangle^\perp} \vec a_j = \vec v_j$$

$$\downarrow$$

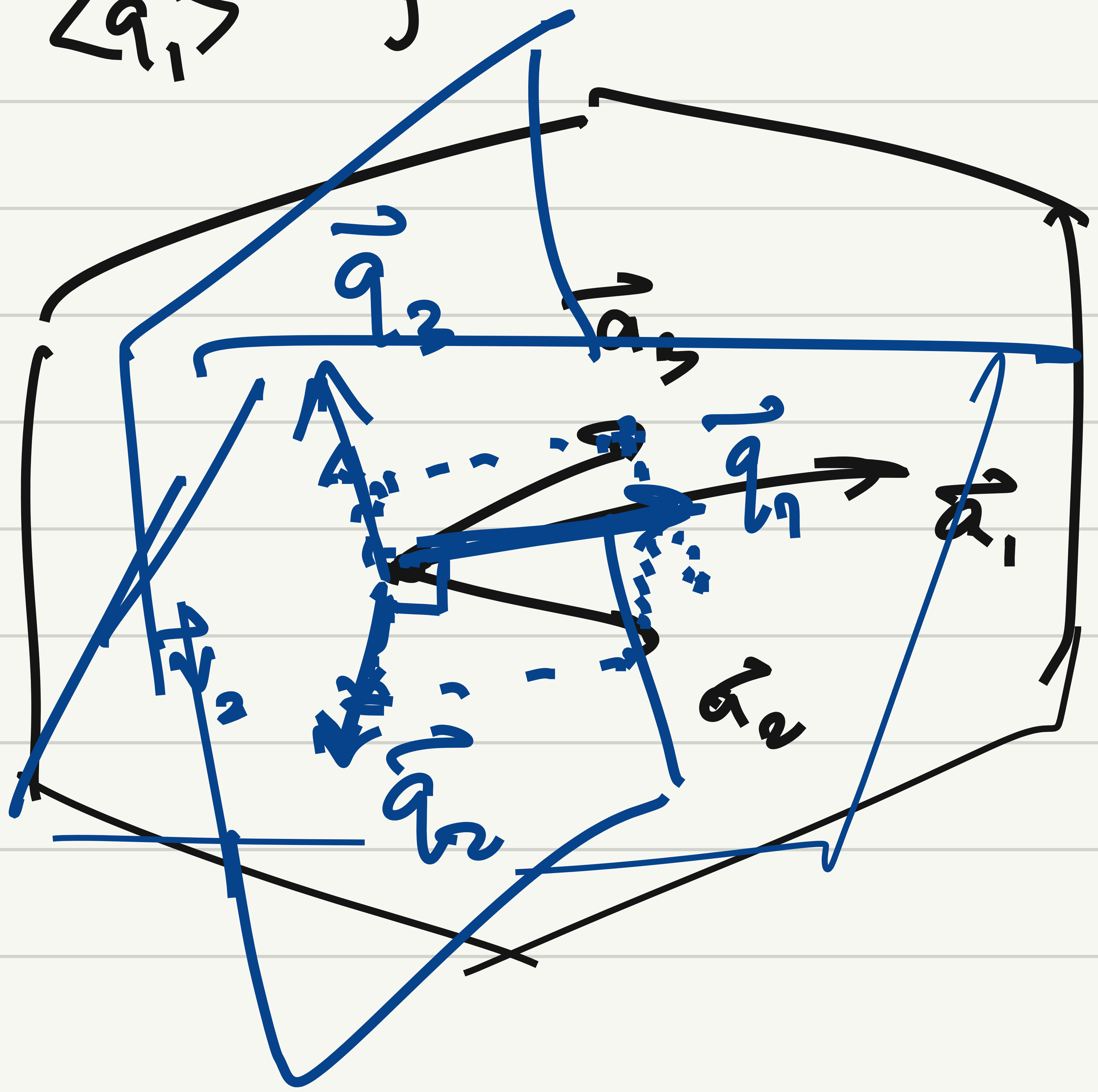
$$\vec q_j$$

for $i = 1, \dots, n$

  $\vec q_i = \vec v_i / \|\vec v_i\|$

  for $j = i+1, \dots n$

    $\vec v_j \mathrel{-}= \vec q_i \vec q_i^{\,*} \vec v_j$



$$\underbrace{\vec{a}_1, \dots, \vec{a}_n}_{} \longrightarrow \underbrace{\vec{q}_1, \dots, \vec{q}_n}_{} \text{ orthonormal}$$

$$A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix} \qquad \hat{Q} = \begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_n \end{bmatrix}$$

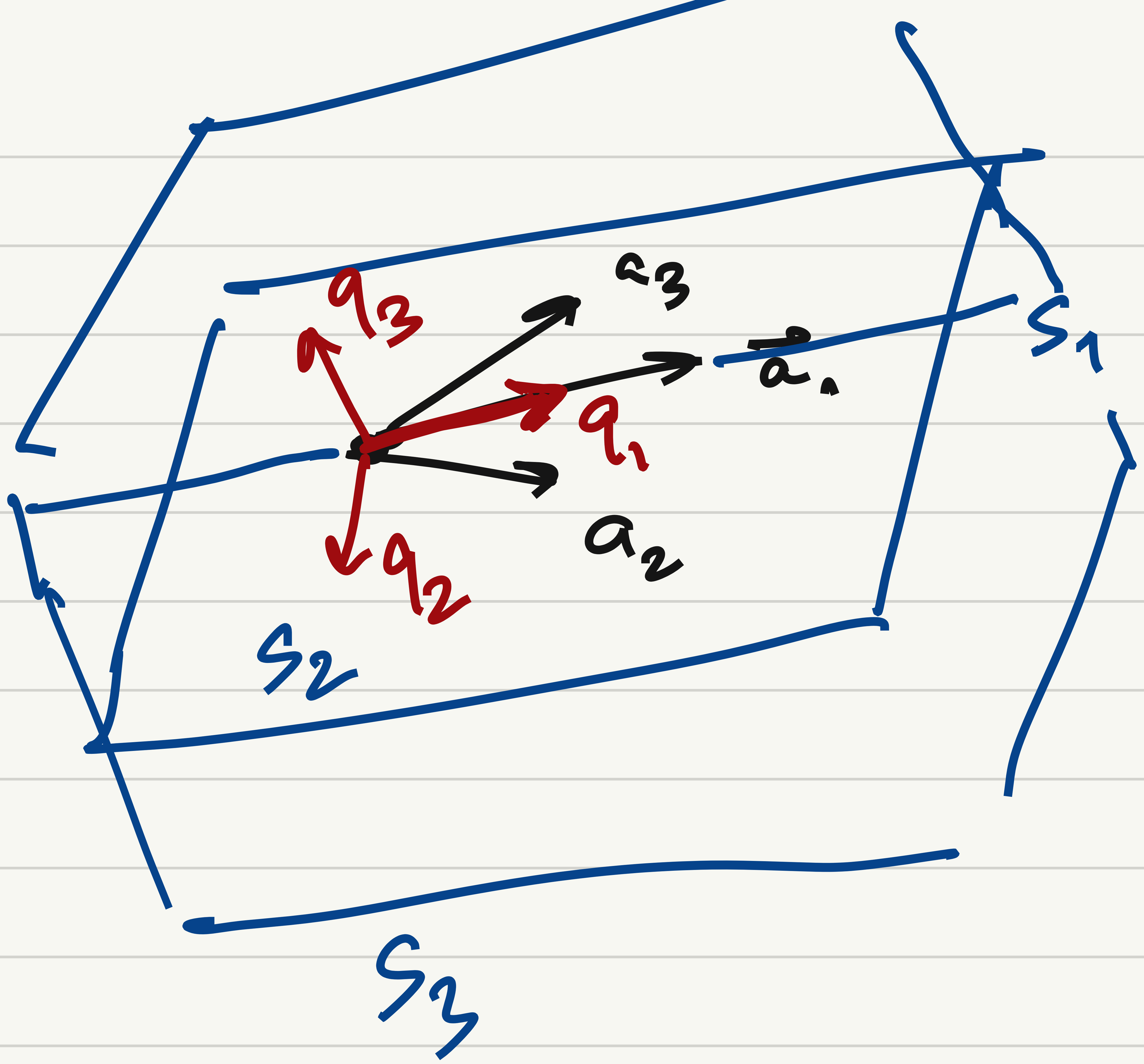
$A \in \mathbb{C}^{m \times n}$, $\vec{b} \in \mathbb{C}^m$, $\underbrace{\vec{b} \overset{?}{\in} \text{range}(A)}$. what is $\vec{x}$ s.t. $A\vec{x} \simeq \vec{b}$?

$$\Rightarrow \vec{b} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots$$

$$\langle \vec{a}_1, \dots, \vec{a}_n \rangle = \langle \vec{q}_1, \dots, \vec{q}_n \rangle$$

$$\vec{a}_j \in \langle \vec{q}_1, \dots, \vec{q}_n \rangle \Rightarrow \vec{a}_j = r_{1j}\vec{q}_1 + \cdots + r_{nj}\vec{q}_n = \hat{Q}\vec{r}_j \qquad \vec{r}_j = \begin{bmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{nj} \end{bmatrix}$$

$$\vec{r}_j = Q^* \vec{a}_j$$



$$\begin{array}{ccccccc}
\langle \vec{a}_1 \rangle & \subseteq & \langle \vec{a}_1, \vec{a}_2 \rangle & \subseteq & \langle \vec{a}_1, \dots, \vec{a}_3 \rangle & \subseteq & \cdots \\
\parallel & & \parallel & & \parallel & & \\
\langle \vec{q}_1 \rangle & \subseteq & \langle \vec{q}_1, \vec{q}_2 \rangle & \subseteq & \langle \vec{q}_1, \dots, \vec{q}_3 \rangle & \subseteq &
\end{array}$$

$$\langle \vec{q}_1, \dots, \vec{q}_j \rangle = \langle \vec{a}_1, \dots, \vec{a}_j \rangle$$

$$\vec{a}_j = r_{1j}\vec{q}_1 + \cdots + r_{jj}\vec{q}_j = \hat{Q} \begin{bmatrix} r_{1j} \\ \vdots \\ r_{jj} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

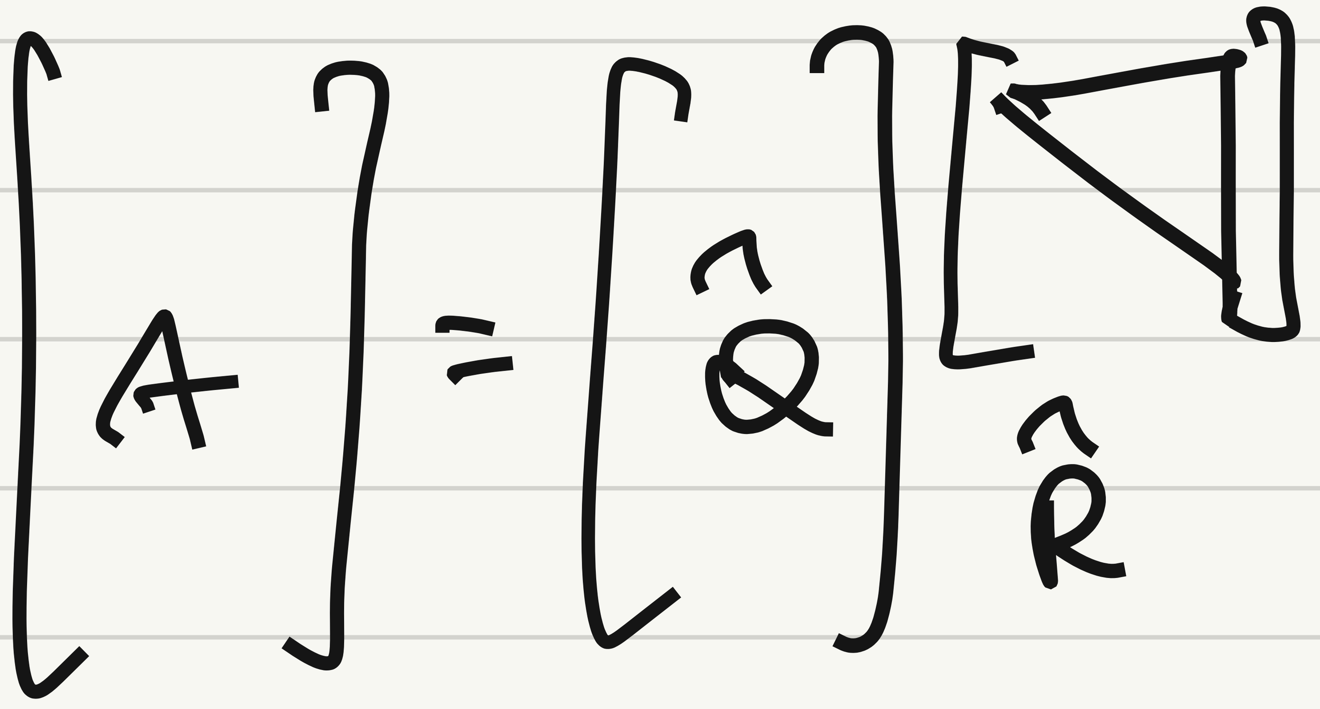
$$A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_j & \cdots & \vec{a}_n \end{bmatrix} = \hat{Q} \underbrace{\begin{bmatrix} r_{11} & \cdots & r_{1j} & \cdots & r_{1n} \\ & \ddots & \vdots & & \vdots \\ & & r_{jj} & & \vdots \\ & & & \ddots & \\ & & & & r_{nn} \end{bmatrix}}_{\hat{R}}$$

$\boxed{\text{reduced QR factorization}}$

Reduced QR factorization: $A \in \mathbb{C}^{m\times n}$

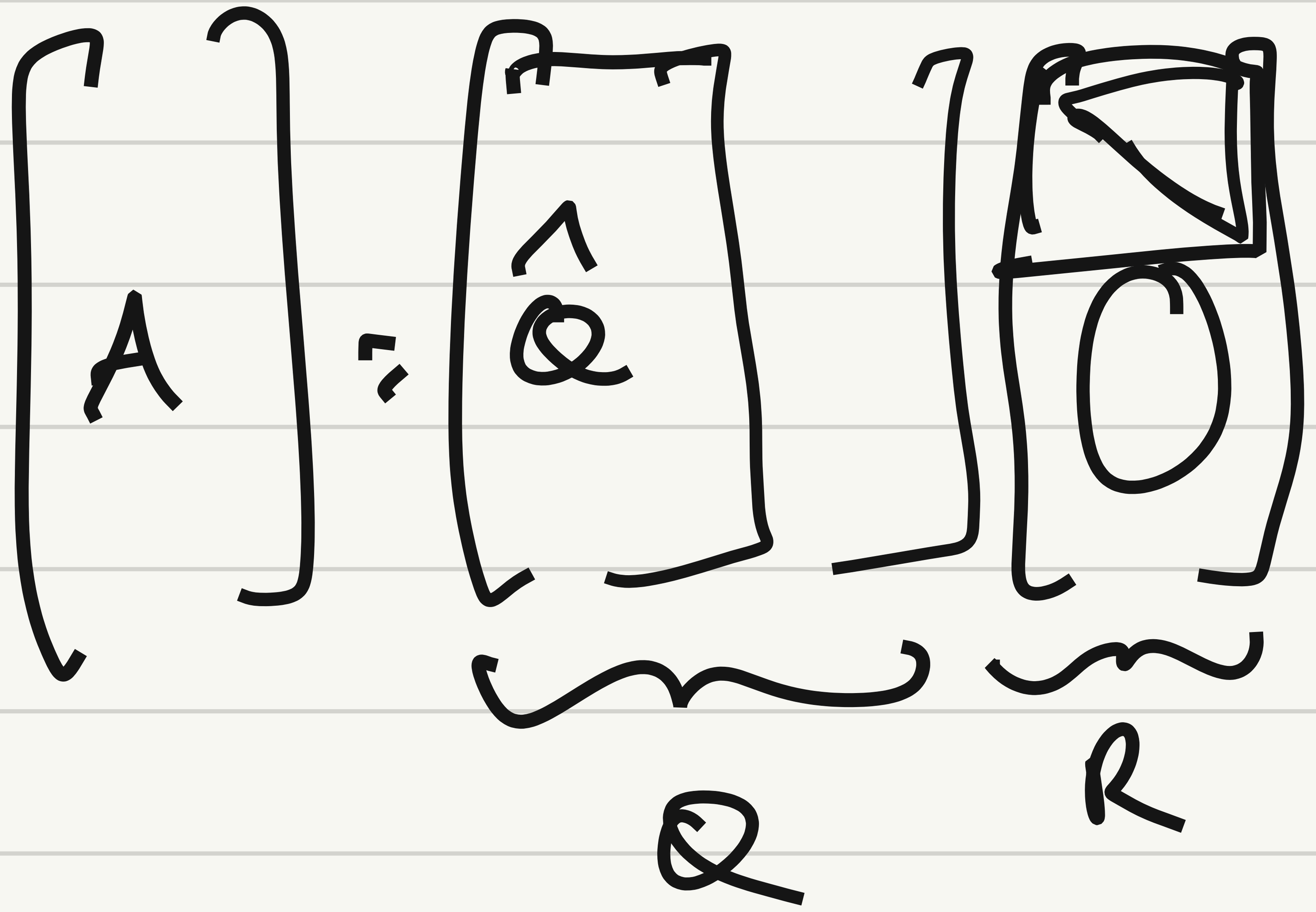
$A = \hat{Q}\hat{R}$, $\quad \hat{Q} \in \mathbb{C}^{m\times n}$, orthonormal columns

$\hat{R} \in \mathbb{C}^{n\times n}$, upper triangular



$\boxed{\text{full QR}}$ factorization: $A = QR$, $\quad Q \in \mathbb{C}^{m\times m}$ unitary

$R \in \mathbb{C}^{m\times n}$ upper tri.



$$Q = \left[\begin{array}{ccc|ccc} \vec{q}_1 & \cdots & \vec{q}_n & \vec{q}_{n+1} & \cdots & \vec{q}_m \\ & & & & & \end{array}\right]$$

$\vec{q}_{n+1}, \dots, \vec{q}_m \perp \text{range}(A)$.

If A is full rank,

$\{\vec{q}_1, \dots, \vec{q}_n\}$ are orthonormal basis for range(A).

$\{\vec{q}_{n+1} \dots \vec{q}_m\}$ are orthonormal basis for range$(A)^\perp$

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

$$\underbrace{A = QR} \Rightarrow QR\vec{x} = \vec{b} \Rightarrow \underbrace{R\vec{x}} = \underbrace{Q^*\vec{b}}$$

$$\begin{bmatrix} \diagdown & \\ 0 & \diagdown \end{bmatrix}\begin{bmatrix} \vec{x} \end{bmatrix} = \begin{bmatrix} Q^*\vec{b} \end{bmatrix}$$

backsubstitution
$O(n^2)$ time

What are r_{ij}?

$$\vec{v}_j = \vec{a}_j - \vec{q}_1\vec{q}_1^{\,*}\vec{a}_j - \cdots - \vec{q}_{j-1}\vec{q}_{j-1}^{\,*}\vec{a}_j, \quad \vec{q}_j = \frac{\vec{v}_j}{\|\vec{v}_j\|}$$

$$\vec{a}_j = r_{1j} q_1 + \cdots + r_{j-1,j}\vec{q}_{j-1} + r_{jj}\vec{q}_j$$

$$\Rightarrow \vec{q}_j = \frac{\vec{a}_j - r_{1j}\vec{q}_1 - \cdots - r_{j-1,j}\vec{q}_{j-1}}{r_{jj}}$$

$$\Rightarrow r_{jj} = \|\vec{v}_j\|$$

$$r_{1j} = \vec{q}_1^{\,*}\vec{a}_j$$
$$\vdots$$
$$r_{ij} = \vec{q}_i^{\,*}\vec{a}_j$$

Operation count : flops (floating-point operations) : add, sub, mul, div, sqrt

Both CGS and MGS require $\sim 2mn^2$ flops

Def: $f(x) \sim g(x)$ if $\displaystyle\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$

$$2mn^2 + 10mn + n^2 + 500 \sim 2mn^2$$

$$3mn^2 \not\sim 2mn^2$$

MGS:

for $i = 1 \ldots n$

$\quad r_{ii} = \|\vec{v}_i\| \longleftarrow \sim 2m$

$\quad \vec{q}_i = \vec{v}_i / r_{ii}$

$\quad$ for $j = i+1 \ldots n$

$$\left.\begin{array}{l} r_{ij} = \vec{q}_i^{\,*} \vec{v}_j \\ \vec{v}_j \mathrel{-}= r_{ij}\vec{q}_i \end{array}\right\} \begin{array}{l} \sim m \text{ mult,} \\ \sim m \text{ add} \\ \sim m \text{ mult} \\ \sim m \text{ sub} \end{array} \quad \sim 4m \text{ flops}$$

$$\sum_{i=1}^{n} \sum_{j=i+1}^{n} 4m \sim 2mn^2$$

Hilbert matrix = $\begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & \cdots \\ 1/2 & 1/3 & 1/4 & & \\ 1/3 & 1/4 & & & \\ 1/4 & & & & \\ \vdots & & & & \end{bmatrix}$

If you have to do Gram-Schmidt : don't use CGS

If you just want QR : don't use GS!

$A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \\ & & \end{bmatrix}$ $\quad V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \\ & & \end{bmatrix}$ initially $V = A$

$V^{(0)} \to V^{(1)} \to V^{(2)} \to \cdots \to \hat{Q}$

$V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \\ & & \end{bmatrix}$

then $\vec{v}_1 \to \vec{v}_1 / \|\vec{v}_1\| = \vec{v}_1 / r_{11}$

$\vec{v}_2 \to \vec{v}_2 - \vec{q}_1^{\,*} \vec{v}_2 = \vec{v}_2 - \dfrac{\vec{v}_1^{\,*} \vec{v}_2}{\|\vec{v}_1\|} = \vec{v}_2 - \dfrac{r_{12}}{r_{11}} \vec{v}_1$

$\vec{v}_j \to \vec{v}_j - \dfrac{r_{1j}}{r_{11}} \vec{v}_1$

$$V^{(0)} = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \longrightarrow \underbrace{\begin{bmatrix} \frac{\vec{v}_1}{r_{11}} & \vec{v}_2 - \frac{r_{12}}{r_{11}} v_1 & \cdots & \vec{v}_j - \frac{r_{1j}}{r_{11}} \vec{v}_1 & \cdots \end{bmatrix}}_{V^{(1)}} = V^{(0)} R_1$$

$$= \underbrace{\begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}}_{V^{(0)}} \underbrace{\begin{bmatrix} \frac{1}{r_{11}} & -r_{12}/r_{11} & \cdots & -r_{1j}/r_{11} & \cdots \\ 0 & 1 & & 0 & \\ \vdots & & \ddots & & \\ \vdots & & & 1 & \\ & & & & \ddots \\ 0 & & & 0 & 1 \end{bmatrix}}_{R_1} = \begin{bmatrix} \vec{q}_1 & \underbrace{\vec{v}_2 & \cdots & \vec{v}_n}_{\text{orthogonal to } \vec{q}_1} \end{bmatrix}$$

$$V^{(0)} = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix} \longrightarrow \begin{bmatrix} \vec{q}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} \longrightarrow \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \vec{v}_3 & \cdots & \vec{v}_n \end{bmatrix} \longrightarrow \cdots$$

$$V^{(1)} = V^{(0)} R_1 \qquad V^{(2)} = V^{(1)} R_2$$

$$V^{(n)} = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \end{bmatrix}$$

$$R_2 = \begin{bmatrix} 1 & & & & \\ & 1/r_{22} & -\frac{r_{23}}{r_{22}} & \cdots & -r_{2n}/r_{22} \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$

(Exercise)

$$V^{(0)} = A$$

$$V^{(1)} = A R_1$$

$$V^{(2)} = A R_1 R_2$$

$$\vdots$$

$$\hat{Q} = V^{(n)} = A \underbrace{R_1 R_2 \cdots R_n}_{\text{upper triangular}}$$

$$\hat{Q} = A(R_1 R_2 \cdots R_n)$$

$$A = \hat{Q} \underbrace{(R_1 R_2 \cdots R_n)^{-1}}_{\hat{R}}$$

GS = "<u>triangular orthogonalization</u>"

At each step, coeffs are $\underbrace{\frac{1}{r_{jj}}}$, $\underbrace{\frac{-r_{ij}}{r_{jj}}}$, ...

Dividing by small number produced by subtraction is dangerous.

Next class: <u>Householder algorithm</u> for QR

"orthogonal triangularization"