Last week: QR factorization.

Solving least squares: $\min \|Ax-b\|_2^2$

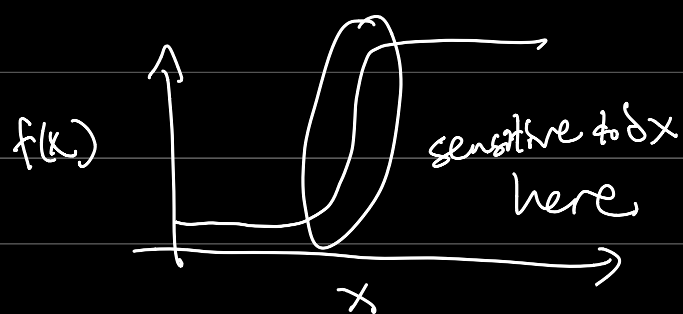
$\Rightarrow A^TAx = A^Tb$

$\Rightarrow Rx = Q^*b$

No need to construct A^TA if using QR.

Conditioning: measuring sensitivity w.r.t small perturbations.

$f(x) - f(x+\delta x) \Rightarrow$ well-conditioned = small if δx is small



Def: Let $f: X \to Y$ be our algorithm or numerical method or problem.

① A problem is well conditioned if small changes in x lead to small changes in $f(x)$.

② A problem is ill conditioned if otherwise: small changes in x can lead to large changes in $f(x)$.

③ Define $\delta f = f(x+\delta x) - f(x)$

Absolute cond. num: $\hat{\kappa} = \lim_{\delta \to 0} \sup_{\|\delta x\| \le \delta} \frac{\|\delta f\|}{\|\delta x\|}$

④ <u>Relative cond. num</u>:

$$\kappa = \lim_{\delta \to 0} \sup_{\|\delta x\| \le \delta} \frac{\|\delta f\| / \|f\|}{\|\delta x\| / \|x\|}$$

$f(A, x) = Ax \Rightarrow$ Suppose $Ax = b$.

$$(A+\delta A)(x+\delta x) = b$$

$$\underbrace{Ax}_{b} + \underbrace{\delta A x + A\delta x}_{\text{"}\delta b\text{"}} + \underbrace{\delta A \delta x}_{\text{assume} \approx 0}$$

$$\delta A x + A \delta x = 0$$

$$\Rightarrow A\delta x = -\delta A x$$

$$\delta x = -A^{-1} \delta A x$$

$$\Rightarrow \frac{\|\delta x\|}{\|x\|} \le \underbrace{\|A\| \|A^{-1}\|}_{} \frac{\|\delta A\|}{\|A\|} \Rightarrow$$

relative sensitivity of x w.r.t rel. sensitivity of A depends on $\|A^{-1}\| \|A\|$

Matrix cond. num $\kappa(A)$

<u>Ex</u>: $Ax = b \Rightarrow$ examine $A(x+\delta x) = b + \delta b$

$$\cancel{Ax} + A\delta x = \cancel{b} + \delta b$$

$$A\delta x = \delta b \Rightarrow \frac{\|\delta x\|}{\|x\|} \le \frac{\|A^{-1}\| \|\delta b\|}{\|x\|} \frac{\|b\|}{\|b\|}$$

$$\boxed{\frac{\|\delta x\|}{\|x\|}} \le \frac{\|A^{-1}\|}{\|x\|} \frac{\|\delta b\|}{\|b\|} \|Ax\| \le \underbrace{\|A\| \|A^{-1}\|}_{\kappa(A)} \boxed{\frac{\|\delta b\|}{\|b\|}}$$

$\kappa(A)$ measures how close A is to being singular. If $\kappa(A) = 1$, as far away from singular as possible. If $\kappa(A) \gg 1$
$\Rightarrow$ A can become singular under a small perturbation.

→ Vandermonde matrix → large cond. num.

→ Hilbert matrix $\begin{array}{c} \\ i \end{array}\begin{bmatrix} \frac{1}{i+j+1} & \rightarrow \\ \downarrow & \searrow \end{bmatrix}$ → worst cond. num of many matrices
$\qquad\qquad\qquad\quad j$

→ A^TA normal matrix. If A square

$$\kappa(A) = \|A\| \|A^{-1}\|$$

$$\begin{aligned}\kappa(A^TA) &= \|A^TA\| \, \|(A^TA)^{-1}\| \le \|A\|^2 \, \|A^{-T}\| \, \|A^{-1}\| \\ &\le \|A\|^2 \|A^{-1}\|^2 \\ &= \kappa(A)^2\end{aligned}$$

→ How to generalize to rectangular matrices?
↳ singular value decomposition (SVD)

SVD: ① generalize cond. numbers

② generalize solutions of least squares to singular A

③ "Rolls-Royce" of matrix factorizations

$\Rightarrow$ FTLA, norms, inverses

Given A:

Derive based on eigenvalues / vectors of

A^*A and $AA^* \Rightarrow$ Hermitian & positive semi-definite.

$$\underbrace{x^* A^*Ax}_{(Ax)^*(Ax) = \|Ax\|^2} \geq 0$$

Recall: $H \in \mathbb{C}^{n\times n}$

If $H = H^*$ & $x^*Hx \geq 0 \ \forall x \in \mathbb{C}^n$

Let $Hv_j = \lambda_j v_j \Rightarrow (\lambda_j, v_j) =$ eigenvalues & eigvecs

① $\{v_j\}$ form an orthogonal (orthonormal) basis for $\mathbb{C}^n$

② $H = H^* \Rightarrow \lambda_j \in \mathbb{R}$

③ $x^*Hx \geq 0 \Rightarrow \lambda_j \geq 0$

Pf: ① For λ_j, λ_k distinct $\quad Hv_j = \lambda_j v_j$

$Hv_k = \lambda_k v_k$

$$v_k^* H v_j = \lambda_j v_j^* v_k$$
$$\underbrace{v_j^* H v_k}_{\to\, v_k^* H v_j \text{ by } H = H^*} = \lambda_k v_j^* v_k$$

$$\Rightarrow \underbrace{(\lambda_j - \lambda_k)}_{\text{distinct} \Rightarrow \neq 0}(v_j^* v_k) = 0$$

$\Rightarrow v_j^* v_k = 0$

② $H = H^*$, $\lambda_j \in \mathbb{R}$, $v_j^* H v_j = v_j^* \lambda_j v_j$

① $v_j^*(H v_j) = \lambda_j (v_j^* v_j)$
② $(H v_j)^* v_j = \overline{\lambda_j} (v_j^* v_j)$

$① - ② \Rightarrow 0 = (\lambda_j - \overline{\lambda_j}) \underbrace{(v_j^* v_j)}_{\neq 0}$

$\Rightarrow \lambda_j \in \mathbb{R}$.

③ $\underbrace{v_j^* H v_j}_{\geq 0} = \lambda_j \underbrace{(v_j^* v_j)}_{\geq 0} \Rightarrow \lambda_j \geq 0$

A^*A & AA^* sym., pos. semi-def.

<u>SVD</u>: $A = U \Sigma V^*$, U, V related to AA^*, A^*A

U, V = unitary, Σ = real & diagonal w/ pos. entries.

<u>SVD derivation</u>: Let $A \in \mathbb{C}^{m \times n}$ full rank, suppose $m \geq n$.

① Let $v_1, \ldots, v_n$ are eigenvectors of A^*A.

$\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n \geq 0$.

② Let $\sigma_j = \| A v_j \| = \sqrt{(A v_j)^* (A v_j)}$

$= \sqrt{v_j^* A^* A v_j} = \sqrt{\lambda_j}$

③ $\underbrace{u_j = \frac{1}{\sigma_j} A v_j}$. Note $\boxed{u_j^* u_k} = \frac{1}{\sigma_j \sigma_k} (Av_j)^* (Av_k)$

$= \frac{1}{\sigma_j \sigma_k} v_j^* A^* A v_k = \boxed{0}$ if $j \neq k$.

$$u_j = \frac{1}{\sigma_j} A v_j \quad \text{for } j = 1, \dots, n$$

$$\underbrace{\begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix}}_{U} = \begin{bmatrix} | & | & & | \\ \frac{1}{\sigma_1} A v_1 & \frac{1}{\sigma_2} A v_2 & \cdots & \frac{1}{\sigma_n} A v_n \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & & | \\ Av_1 & \cdots & Av_n \\ | & & | \end{bmatrix} \underbrace{\begin{bmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_n} \end{bmatrix}}_{\Sigma^{-1}}$$

$$= A \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \Sigma^{-1}$$

$$U = A V \Sigma^{-1}$$

$\Rightarrow$ $U \Sigma V^* = A$ for $A \in \mathbb{R}^{m \times n}$ $m \geq n$ & full rank.

"reduced" SVD