M126 lec 2 SVD Conditioning Stability. ACMS @ 1pm here. · ellipse.m, WS · (→ laptop) · ① 1/10/12

Orthogonality (review lin. alg.)

A matrix, A^* hermitian transpose. : $(A^*)_{ij} = \overline{A_{ji}}$ c.c.

$A = A^*$ symm?..

Ex. prove $(AB)^* = B^* A^*$, $(A^{-1})^* = (A^*)^{-1}$

x col. vec, x^* row vec.

$x, y \in \mathbb{C}^m$, $x^* y = ?$ inner prod $[\longrightarrow]\begin{bmatrix}\downarrow\end{bmatrix} = \sum_{i=1}^m \bar{x}_i y_i$ 2-norm $\|x\|_2 := \sqrt{x^* x}$ won't always write.

Defn of a norm? i) $\|\alpha x\| = |\alpha| \, \|x\|$ α scalar
ii) $\|x\| = 0 \Rightarrow x = 0$ ← vector.
iii) $\|x+y\| \le \|x\| + \|y\|$ tri

2-norm also has. $|x^* y| \le \|x\| \, \|y\|$ Cauchy-Schwarz

$x^* y = 0$? x, y orthog.

Thm: mutually orthog. set of vectors are lin. indep. (pf: Ex).

⇒ m orthog. vecs in $\mathbb{C}^m$ form basis; if unit length, o.n.b.

Say $\vec{q}_1, \dots \vec{q}_m \in \mathbb{C}^m$ o.n.b., can stack into cols. of Q, then $(Q^* Q)_{ij} = i[\longrightarrow]\begin{bmatrix}\downarrow\end{bmatrix}^{j} = q_i^* q_j = \delta_{ij}$ i, j th entry.

so $Q^* Q = I$ ie $Q^{-1} = Q^*$ $\boxed{\text{cols o.n.b.} \Leftrightarrow \text{unitary}}$

so $Q^* b$ is coeffs of expansion of b in o.n.b $\{q_i\}$

$\|Qx\| = \sqrt{(Qx)^* Qx} \stackrel{?}{=} \sqrt{x^* \cancel{Q^* Q} x} = \|x\|$ preserves lengths. (rotation, poss w/ reflect)

Matrix 2-norm: $\|A\|$ ← 2 implied smallest # c st. $\|Ax\| \le c\|x\|$ $\forall x \in \mathbb{C}^m$

"max growth factor"

or $\|A\| := \sup_{x \ne 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|$

Eg i) 2-norm of diag matrix? Its largest-magnitude entry.

ii) 2-norm of $A = uv^*$? (rank-1). $\|u \underbrace{v^* x}_{\text{scalar}}\| = \underbrace{|v^* x|}_{\text{C-S.}} \|u\| \le \underbrace{\|v\| \, \|u\|}_{= \|A\|.} \|x\|$

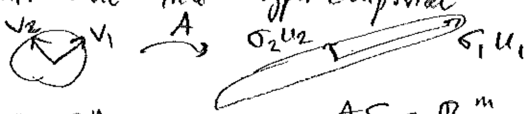
2-norm submultiplicative: $\|ABx\| \underset{\text{why?}}{\le} \|A\| \, \|Bx\| \le \|A\| \, \|B\| \, \|x\|$

Ex. show QA & AQ have same 2-norm as A. So $\|AB\| \le \|A\| \|B\|$. (Thm 3.1)

[lec2] 1/19/12 (2)

Singular Value Decomposition (SVD) — as important as spectral decomp.

Geom fact: every $A \in \mathbb{C}^{m\times n}$ maps unit ball into hyperellipsoid



$S \subset \mathbb{R}^n$ $\qquad$ $AS \subset \mathbb{R}^m$

($\mathbb{C}^n$ harder to plot!)

show ellipse.m. 2×2 case.

'tall' case $m \geq n$ & full rank. $(=n)$: left sing. vecs $u_1 \cdots u_n$ unit orthog.

sing. vals. $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_n > 0$ semiaxes.

so what is $\|A\|$? $= \sigma_1$

right sing. vecs v_j. $j=1\cdots n$, unit preimages of $\sigma_j u_j$ also orthog amazingly! $\Rightarrow$ o.n.b. for $\mathbb{C}^n$.

rank $r<n$ then $\sigma_1, \cdots \sigma_r > 0$ while $\sigma_{r+1} = \cdots \sigma_n = 0$.

So $Av_j = \sigma_j u_j$ $\quad j = 1\cdots n$.

$$A \underbrace{\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}}_{V} = \underbrace{\begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}}_{\hat{U}} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}}_{\hat{\Sigma}}$$

(postmult by V^*) so $A = \underset{m\times n}{\hat{U}} \underset{n\times n}{\hat{\Sigma}} \underset{n\times n}{V^*}$ reduced SVD.

May complete $\underset{m\times n}{\hat{U}}$ to $\underset{m\times m}{U}$ o.n.b. for $\mathbb{C}^m$.

& pad $\underset{n\times n}{\hat{\Sigma}}$ to $\underset{m\times n}{\Sigma}$

then $A = \underset{U}{\square_{m\times m}} \; \underset{\Sigma}{\square_{m\times n}} \; \underset{V}{\square_{n\times n}}$. (analogous if $m<n$)

Thm (e.g. 4.1 in NLA): every $A \in \mathbb{C}^{m\times n}$ has SVD, $\{\sigma_j\}$ unique,

& if σ_j simple, u_j & v_j unique up to phase.

Every matrix is rotation (w/ refl.) $\rightarrow$ coord stretch $\rightarrow$ rot. (w. refl.) even nonsymm or nonsquare ones. (cf. spectral decomp may not exist.)

or: every matrix is diag in correct bases for $\mathbb{C}^n$ & $\mathbb{C}^m$ (V^* projects to coeffs of onb)

If A square invertible, $A^{-1} = (U\Sigma V^*)^{-1} \overset{?}{=} V \underset{\text{diag}\{\sigma_j^{-1}\}}{\Sigma^{-1}} U^*$ is the SVD of A^{-1} (if reorder Σ^{-1}!)

What is $\|A^{-1}\|$? largest sing val of A^{-1} $= \sigma_n^{-1}$ ← smallest singval of A.

→ WS (PTO).

needed for WS.

break

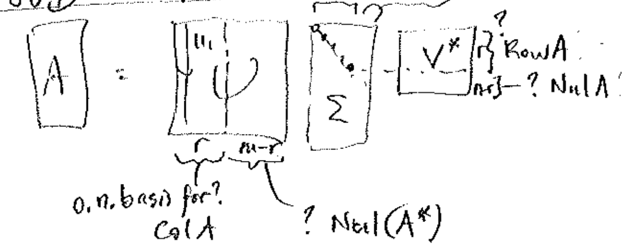
Notes on SVD: existence pf either by induction (NLA Ch. 4 — beautiful proof: grad studs read)

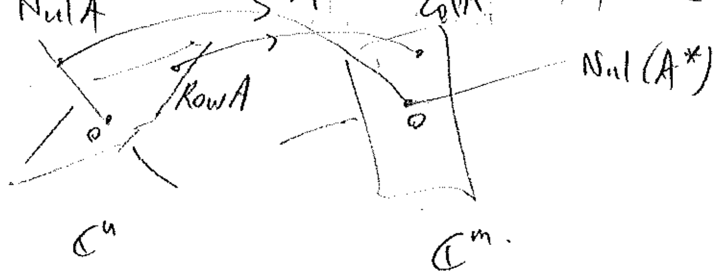
or $\cdot$ A^*A has eigvals $\{\sigma_j^2\}$ & complete eigvecs v_j o.n.b. why? A^*A symm.

AA^* " " " u_j "

eigvals $\lambda_j \geq 0$ & then define $\sigma_j = \sqrt{\lambda_j}$.

Anatomy: SVD & 4 spaces: 1/11/12 ③





rank $r := \#\{j : \sigma_j > 0\}$

numerical rank $r_\varepsilon := \#\{j : \sigma_j > \varepsilon\}$ $\quad \varepsilon \sim \sigma_1 \varepsilon_{mach}$ (machine precision!) $\sim 10^{?} \leftarrow -16$

Qu: what do think σ_j's of Vandermonde did? Shot down to ε when $m \sim 40$.

Conditioning (§12 NLA) : property of a math problem (vs. Stability: property of alg. used to solve it).

problem is map $f : X \to Y$ (X: problem input space; Y: space of solns.) $\quad$ eg. $f(x)$ could return $\{ \cdot\ 2x$ "doubling prob", the easy. $\cdot$ vector of roots of poly given x = vec. of poly coeffs.

f well-cond if infinitesimal pert δx (one symbol!) to x causes 'small' pert $\delta f := f(x+\delta x) - f(x)$ in soln.

Abs. cond. # $\hat{\kappa} = \hat{\kappa}(x) := \lim_{\delta \to 0} \sup_{\|\delta x\| \leq \delta} \frac{\|\delta f\|}{\|\delta x\|} \overset{\text{abbrev}}{=} \sup_{\delta x} \frac{\|\delta f\|}{\|\delta x\|}$ (2-norms)

if x, f vectors ($\in \mathbb{C}^n$, $\in \mathbb{C}^m$), $\frac{\partial f_i}{\partial x_j} = J_{ij}(x)$ is Jacobean matrix $J \in \mathbb{C}^{m \times n}$

As $\|\delta x\| \to 0$ have $\delta f \approx J(x)\,\delta x$ so $\hat{\kappa}(x) = \|J(x)\|$ matrix 2-norm.

more useful is:
Rel. cond # $\kappa := \sup_{\delta x} \frac{\|\delta f\| / \|f\|}{\|\delta x\| / \|x\|} = \frac{\|J(x)\|_2}{\|f\| / \|x\|}$ $\quad$ important since computer brings in relative errors.

$\kappa < 10^3$ 'well cond'
$\gg 10^3$ ill-cond

Basic ops:
- $f(x) = x/2$ $(m=n=1)$ $J = f' = 1/2$ so $\kappa = \frac{|1/2|}{1/2} = 1$.
- $f(x) = x^\alpha$ $\quad J = f' = \alpha x^{\alpha-1}$ $\quad \kappa = \frac{|\alpha x^{\alpha-1}|}{x^\alpha / x} = |\alpha|$. $\ll 10^3$ well-cond for reasonable powers.
- $f(x_1,x_2) = x_1 - x_2$ subtraction $(n=2, m=1)$. $J = [1\ -1]$ $\|J\| = \sqrt{2}$
  $\kappa = \frac{\sqrt{2}\sqrt{x_1^2+x_2^2}}{|x_1 - x_2|} \to \infty$ as $x_1 \to x_2 \neq 0$. can be ill-cond.
- $f(x) = \sin x$, for $x \approx 10^{100}$ say: $\|J\| \leq 1$ but $\kappa = \frac{\|J\|\,|x|}{|\sin x|} \geq |x|$ = huge. (but abs cond # ≤ 1).
- finding poly roots ill-cond
- eigvals of nonsymm matrices: eg $A = \begin{bmatrix} 1 & 10^3 \\ & 1 \end{bmatrix}$ vs $\begin{bmatrix} 1 & 10^3 \\ 10^{-3} & 1 \end{bmatrix}$ $\quad \|\delta x\| = 10^{-3}$
  $\lambda = \{1, 1\}$ $\quad \lambda = \{0, 2\}$ $\quad \|\delta f\| \sim 1$
  but symm, $\hat{\kappa} \approx 1$. $\quad \hat{\kappa} \sim 10^3,\ \kappa \sim 10^6$