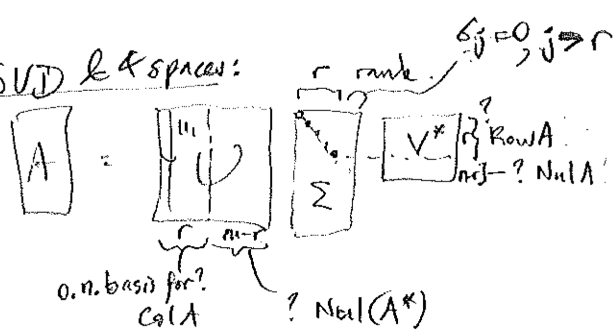
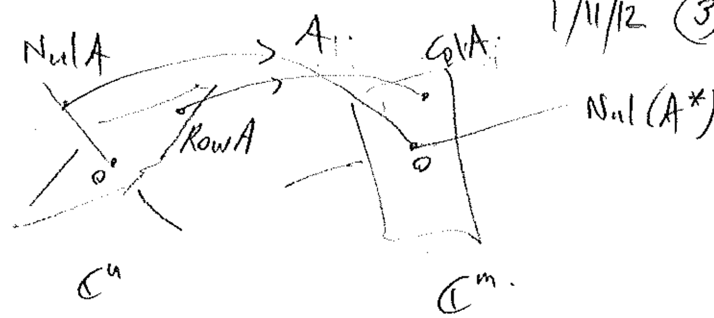
Anatomy: SVD & 4 spaces:



1/11/12 (3)



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

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

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

Lec 3

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

problem is map $f : X \to Y$ (problem input space; space of sol'ns.) $\qquad$ eg. $f(x)$ could return $\begin{cases} \cdot\ 2x & \text{the easy "doubling prob"} \\ \cdot\ \text{vector of roots of poly given } x = \text{vec. of poly coeffs.} \end{cases}$

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

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

if x, f vectors ($x \in \mathbb{C}^n$, $f \in \mathbb{C}^m$), $\frac{\partial f_i}{\partial x_j} = J_{ij}(x)$ 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 # $\kappa := \sup_{\delta x} \frac{\|\delta f\| / \|f\|}{\|\delta x\| / \|x\|} = \frac{\|J(x)\|_2}{\|f\| / \|x\|}$ $\qquad$ important since compute brings in relative errors.

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

Basic ops:
- $f(x) = x/2$ $(m=n=1)$ $J = f' = 1/2$ so $\kappa = \frac{|1/2|}{1/2} = 1$.

do as W/S:
- $f(x) = x^\alpha$ $\quad J = f' = \alpha x^{\alpha-1}$ $\quad \kappa = \frac{|\alpha x^{\alpha-1}|}{x^\alpha / x} = |\alpha|$. $\ll 10^3$ well-cond for reasonable powers.
- $f(x_1, x_2) = x_1 - x_2$ subtraction $(n=2, m=1)$. $J = [1\ \ -1]$ $\quad \|J\| = \sqrt{2}$
  $\kappa = \frac{\sqrt{2}\sqrt{x_1^2 + x_2^2}}{|x_1 - x_2|} \to \infty$ as $x_1 \to x_2 \neq 0$. can be ill-cond.
- $f(x) = \sin x$, for $x \approx 10^{100}$ say: $\|J\| \le 1$ but $\kappa = \frac{\|J\|\,|x|}{|\sin x|} \ge |x|$ = 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 $\begin{bmatrix} 1 & 10^3 \\ 10^{-3} & 1 \end{bmatrix}$ $\qquad \|\delta x\| = 10^{-3}$
  $\lambda = \{1, 1\}$ $\qquad \lambda = \{0, 2\}$ $\qquad \|\delta f\| \sim 1$
  $\hat{\kappa} \sim 10^3$, $\kappa \sim 10^6$

but symm, $\hat{\kappa} \approx 1$.

② 1/11/12

Mat-vec. mult? $f(x) = \underset{m\times m}{A}x \xrightarrow{J=?A} \kappa = \|A\| \dfrac{\|x\|}{\|Ax\|}$ — if A nonsing., $\le \|A^{-1}\|$, why?

pf: $\|x\| = \|A^{-1}Ax\| \le \|A^{-1}\|\,\|Ax\|$ ← defn of 2-norm

so $\kappa \le \|A\|\,\|A^{-1}\|$ with equality if $x = v_m$. so, tight. ← Ex: check.

Solving lin sys? $Ax = b$ so soln is $f(b) = A^{-1}b$ replace A by A^{-1} in above, get again $\kappa \le \|A^{-1}\|\|A\|$. w/ equality if $b = u_1$

call $\|A\|\|A^{-1}\| =: \kappa(A)$ cond # of matrix A $= \dfrac{\sigma_1}{\sigma_m}$ = eccentricity of hyperellipse. ← Ex: check.

What if A perturbed instead, in solving lin sys? input is A, output x (b held const).

consider infinitesimal changes: δA causing δx

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

so $\underbrace{\cancel{Ax}}_{\text{cancels } b} + \delta A x + A\delta x + \cancel{\delta A \delta x} = \cancel{b}$ ignore to 1st order

so $\delta x = -A^{-1}\delta A x$, $\dfrac{\|\delta x\|}{\|x\|} \le \|A^{-1}\|\,\|\delta A\|$

rel. cond #:

⇒ Thm: (12.2) $\dfrac{\|\delta x\|/\|x\|}{\|\delta A\|/\|A\|} \le \|A^{-1}\|\|A\| = \kappa(A)$ again.

- since A, b stored to 16 digits, expect to get x to $16 - \log_{10}\kappa$ digits acc.

Floating Point

$x \in \mathbb{R}$ digital rep: finite # bits ⇒ finite subset F of $\mathbb{R}$ ⇒ must be { lowest & highest $\pm 10^{308}$ ← in IEEE 'double prec' system. gaps!

eg $[1,2]$ rep by set $\{1, 1+2^{-52}, 1+2\cdot 2^{-52}, \dots 2\}$

$[2,4]$ is twice these (larger gaps!). relative gap 2.2×10^{-16} never exceeded.

(but a poor algorithm can cause this to dominate).

Formally, base $= \beta = 2$, precision $= t = 53$

set is $F = \{0, \pm \dfrac{m}{\beta^t}\beta^e, \underbrace{\pm \text{Inf}, \text{NaN}}\}$ special codes, rather than members of $\mathbb{R}$.

$\beta^{t-1} \le m \le \beta^t$

so $\dfrac{m}{\beta^t} \in [\frac{1}{\beta}, 1]$

$e \in \mathbb{Z}$ exponent (we ignore over/underflow: that there is in fact a largest $|e|$).

Note $\beta F = F$, self-similar.

$\varepsilon_{mach} = \frac{1}{2}\beta^{1-t}$ is largest relative error: ie $\forall x \in \mathbb{R}, \exists x' \in F$ s.t. $|x' - x| \le \varepsilon_{mach}|x|$

β^{1-t}: largest rel. gap.

let such an x' be called $fl(x)$

Then $\forall x \in \mathbb{R}, \exists \varepsilon, |\varepsilon| \le \varepsilon_{mach}$ s.t. $fl(x) = (1+\varepsilon)x$ bounded rel. err.

IEEE double precision $\varepsilon_{mach} = 2^{-53} \approx 1.1\times 10^{-16}$

③ 1/12/12

Arithmetic   let $\oplus\ \ominus\ \otimes\ \oslash$ be analogs of $+ - \times \div$ except done by machine.

let $\circledast$ be any of 4: could require $x \circledast y = fl(x * y)$, ie gives the unique rounding of answer.

But only need weaker: Fund Axiom of Floating pt:

$$\forall x, y \in F \quad \exists \varepsilon \text{ w/ } |\varepsilon| \le \varepsilon_{mach} \quad \text{s.t.} \quad x \circledast y = (1+\varepsilon)(x * y)$$

ie rel. err bnded by ε_{mach}

For $\mathbb{C}$ instead of $\mathbb{R}$, turns out to be $2^{3/2}\varepsilon_{mach}$, similar.

Stability (§14 NLA)   alg getting 'right' ans. even if not exact.

fix: $\begin{cases} \text{problem } f: X \to Y & \text{eg } y = \sin x \text{ or } y \text{ is soln to } Ay = b \quad (\text{here 'x' data is } A, b). \\ \text{fl. pt. sys} \\ \text{algorithm for } f, \text{ also a map } \tilde{f}: X \to Y \end{cases}$

Steps in alg: $\underset{\text{input}}{\overset{\text{data}}{x}} \xrightarrow{\text{round}} fl(x) \xrightarrow[\text{alg}]{\text{machine}} \tilde{f}(x)$

Defn relative error of computation $\dfrac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|}$ ← alg. certainly good if this is $O(\varepsilon_{mach})$

'of the order of',

Formally: $O(\varepsilon_{mach})$ means $\le C\varepsilon_{mach}$ as $\varepsilon_{mach} \to 0$

ie a family of floating pt sys, for some const C, uniformly over all data $x \in X$.

Practically: $< 10^3 \varepsilon_{mach}$ ok, $> 10^8 \varepsilon_{mach}$ not ok.

But if problem f ill-cond, unreasonable to demand this! Why? rounding on input changes $x \to fl(x)$ & if κ v. high, change gets blown up by κ so even if alg. exact, cannot have $O(\varepsilon_{mach})$ rel. err in output.

Instead: defn

Alg Stable if $\forall x \in X$, $\dfrac{\|\tilde{f}(x) - f(\tilde{x})\|}{\|f(\tilde{x})\|} = O(\varepsilon_{mach})$ for some $\tilde{x}$ s.t. $\dfrac{\|\tilde{x} - x\|}{\|x\|} = O(\varepsilon_{mach})$

"nearly right answer to nearly right question"

stronger: Backward stable : $\forall x \in X$, $\tilde{f}(x) = f(\tilde{x})$ for some $\tilde{x}$ s.t. $\dfrac{\|\tilde{x} - x\|}{\|x\|} = O(\varepsilon_{mach})$

"exactly right ans. to nearly right question".

Eg [§15] is $\ominus$ bkw stable?   prob is $f(x_1, x_2) = x_1 - x_2$

alg. is $\tilde{f}(x_1, x_2) = fl(x_1) \ominus fl(x_2)$

$= [x_1(1+\varepsilon_1) - x_2(1+\varepsilon_2)](1+\varepsilon_3) \overset{\text{set}}{=} x_1(1+\varepsilon_4) - x_2(1+\varepsilon_5) = f(\tilde{x}_1, \tilde{x}_2)$ exact for some data rel. close to x_1, x_2.

for $|\varepsilon_i| \le \varepsilon_{mach}$ $i = 1,2,3$. $|\varepsilon_4|, |\varepsilon_5| \le 2\varepsilon_{mach} + O(\varepsilon_{mach}^2)$