Tells you fund. spaces, o.n.b. for them:
Ran A = Span$\{u_1, \dots u_r\}$, Nul A = Span$\{v_{r+1}, \dots v_n\}$, etc.

SVD
- exists for all (real, non symm) A, unlike EVD.
- tells the most about action of A.
- also exists for operators (func. anal.).
- can be computed in $O(nm \max(n,m))$, Golub–Kahan 1965.
  $\approx O(n^3)$ if square.

1:10 ?

Floating pt arithmetic. Convergence is one type of error ; this is the other.

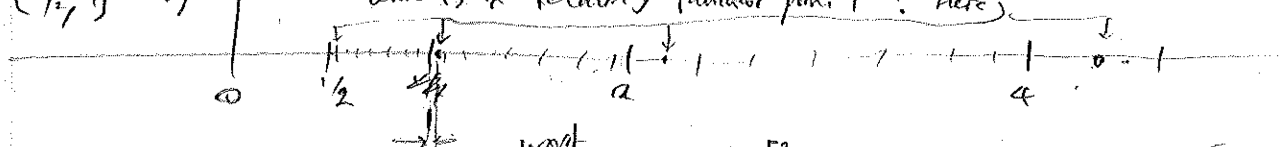
Modern CPUs work only to relative error $\approx 10^{-16}$ ("double precision" – e.g. Matlab).
10^{-8} ("single prec" – not used in science, only gaming since faster!).

$x \in \mathbb{R}$ approximated $fl(x) \in F$, rep by finite # bits (64 in double prec, so $|F| \le 2^{64} \approx 1.8 \times 10^{19}$).

e.g. $x \in [1,2]$ by set $\{1, 1+2^{-52}, 1+2\cdot2^{-52}, 1+3\cdot2^{-52}, \dots 2\}$
↖ 4.5×10^{15} elements.

$[2,4]$ by ? double these.
$[1/2, 1]$ by ?

where is x relatively furthest from F? here



worst relative dist. $2^{-53} = \varepsilon_{mach} \approx 1.1 \times 10^{-16}$ ↪ prints as 1.1e16 in C, Matlab etc.

52 bits for where in $[1,2]$ "mantissa" M → how many bits
1 bit sign $\pm$
11 bits for exponent $-1023 < E \le 1023$. "floating pt". $[1/2,1]: E=-1$, $[1,2]: E=0$, $[2,4]: E=1$, etc.

$fl(x) = \pm M 2^E$ ← mean 10^{308} is largest, 10^{-308} smallest

are also codes for $0, \pm$Inf, ~~Nd~~ NaN (not a number) → e.g. if do $0/0$. → $1/0$

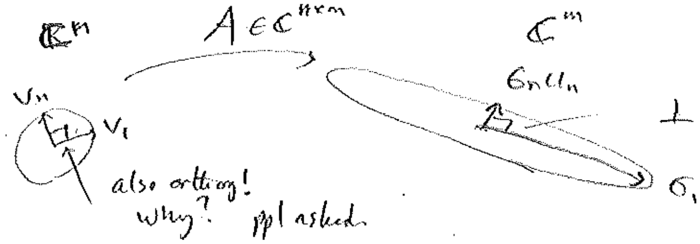
Let $\oplus, \ominus, \otimes, \oslash$ be machine ops.
All arithmetic obeys:
Fund. Axiom: $\forall x,y \in F$, $\exists\, \varepsilon$ w/ $|\varepsilon| \le \varepsilon_{mach}$ st. $x \oplus y = (1+\varepsilon)(x+y)$
($\ominus, \otimes, \oslash$ likewise with $-, \times, \div$)

~~1:30~~. Did example of Catastrophic Cancellation:
$x = 1+10^{-10}$, $y = 1$ $\quad x, y \xrightarrow{\text{rounding}} fl(x), fl(y) \xrightarrow{\text{arith}} fl(x) \ominus fl(y)$
$fl(x) = x(1+\varepsilon_1)$, $fl(y) = 1$
$fl(x) \ominus fl(y) = (x - 1 + \varepsilon_1 x)(1+\varepsilon_2)$, where $x-1 = 10^{-10}$
rel. err. $\frac{\varepsilon_1 x}{10^{-10}} \approx 10^{-6}$.

~~Stability~~ ?

M116 lec 4. ① 4/3/14.

Note re SVD: ($m \geq n$)



$\perp$ by geom. of hyperellipse.

Sketch:

Set $\sigma_1 = \|A\|_2$ by compactness, $\exists$ unit vectors u_1, v_1 st. $A u_1 = \sigma_1 v_1$

Extend v_1 to onb. $\{v_j\}$ for $\mathbb{C}^n$ (any old way.) $\to$ stack into V_1 unitary
u_1 " $\{u_j\}$ " $\mathbb{C}^m$ U_1 "

then $U_1^* A V_1 = \begin{bmatrix} \sigma_1 & w^* \\ 0 & B \end{bmatrix} =: S$ — unknown row vec. : if can show $w = \vec{0}$ then $A\{v_2, \dots, v_n\} \perp u_1$, so $\{v_2, \dots, v_n\}$ all $\perp v_1$

col vec 0 why? since $u_2, \dots, u_m \perp u_1$

Continue by induction... see [Ch. 4 NLA]

$\left\| \begin{bmatrix} \sigma_1 & w^* \\ 0 & B \end{bmatrix} \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2 \geq \sigma_1^2 + w^* w = \sqrt{\sigma_1^2 + w^* w} \left\| \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|$ ← pf.

↑ a vector. so $\|S\|_2 \geq \sqrt{\sigma_1^2 + w^* w}$

19 min.

→ Fl pt. (lec 3 ②)

but $\|S\|_2 = \|A\|_2$ by unitarity of V_1, U_1 $\Rightarrow w \equiv 0$. ∃ Geom. pf?

30 min ~~Old example~~

Stability:

did Lenya: V = eigvec matrix of $A^* A$, $\Rightarrow$ unitary.

$f: X \to Y$ math prob

$\tilde{f}: X \to Y$ numerical alg, on fl arith.

rel. err. $\varepsilon = \dfrac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|}$ would be nice to be $O(\varepsilon_{mach})$, eg. $\leq 10^3 \varepsilon_{mach}$.

Unreasonable if κ large since even rounding error in input $x \to fl(x)$ amplified to $\kappa \varepsilon_{mach}$ in output rel. err.

Defn:

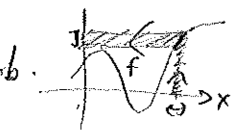
Alg. Stable: if $\forall x \in X$, $\dfrac{\|\tilde{f}(x) - f(\tilde{x})\|}{\|f(\tilde{x})\|} = O(\varepsilon_{mach})$ for some $\tilde{x}$ st. $\dfrac{\|\tilde{x} - x\|}{\|x\|} = O(\varepsilon_{mach})$.

"nearly right ans. to nearly right question."

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

"exactly right ans. to nearly right question".

We'd like these in algs: allows error but only appropriate to ill-cond. of the math prob.



Thm (15.1): rel. err. $\varepsilon = O(\kappa \varepsilon_{mach})$. if alg bkw stab.

Do one example:

- Is $f(x) = x \oplus 1$ bkw st? $\tilde{f}(x) = (x(1+\varepsilon_1) + 1)\cdot(1+\varepsilon_2) \overset{set}{=} \underbrace{x(1+\varepsilon_3)}_{\tilde{x}} + 1$
  ↑ from $fl(x)$ ↑ from $\oplus$ ↖ exact $\tilde{x}+1$

Can ε_3 be found at each x, uniformly $O(\varepsilon_{mach})$?

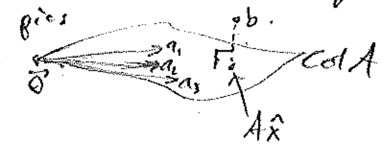
$1 + x + \varepsilon_1 x + \varepsilon_2 x + \varepsilon_2 + O(\varepsilon^2) = 1 + x + \varepsilon_3 x$ so $\varepsilon_3 = \underbrace{\varepsilon_1 + \varepsilon_2}_{\checkmark} + \left(\dfrac{\varepsilon_2}{x}\right) \to \infty$ as $x \to 0$, $\Rightarrow$ not BS.

But, is stable.

Th. (2) 4/3/14

- Lots of common algs are BS, eg computing {EVD, SVD (Golub-Kahan) or,
  least-sq: $\hat{x} = \arg\min_{x \in \mathbb{C}^n} \|Ax - b\|_2$, done by SVD.

Note: • LSQ prob has $\kappa \leq \kappa(A)$ as usual.
• tons of apps.



↗ in Matlab: $x = A \backslash b$.

How? i) Do SVD.
ii) pick numerical rank $r = \#\{\sigma_j \geq \varepsilon\}$
iii) Proj b onto Col A: coeffs $U_\varepsilon^* b$. let $U_\varepsilon = [u_1 \cdots u_r]$
iv) Map back to coeffs of onb. $\{v_j\}$: $\Sigma^{-1} U_\varepsilon^* b$
v) so $\hat{x} = V \Sigma^{-1} U_\varepsilon^* b$ — called pseudo inverse A^+; Σ^{-1} = diag $\{\sigma_1^{-1}, \cdots \sigma_r^{-1}\}$.

* 'Normal eqns' solve $A^*A x = A^* b$ (A^*A square) is not BS if $\kappa(A)$ large.

- Gaussian elim. (equiv to factorization $A = L U$) (A: $m \times m$, L lower-tri, U upper-tri) BS if $\|L\|, \|U\|$ not large.
  Ie, given $\tilde{L}\tilde{U} = A + \delta A$ w/ $\|\delta A\| = \|L\| \|U\| O(\varepsilon_{mach})$

check: (code) [L,U] = lu(A) ; norm(A - L*U)/norm(A).

Most A have small L, U entries (after permutation = 'pivoting') eg $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$

ex: try $A = \begin{bmatrix} 1 & & & 1 \\ -1 & 1 & & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \rightarrow U = \begin{bmatrix} 1 & & & 1 \\ & 1 & & 2 \\ & & 1 & 4 \\ & & & 8 \end{bmatrix}$ size 2^m. but it's observed 'never' to happen for other matrices (Driscoll).

m=10; A = tril(-ones(m,m)) + eye(m) → show; play by adding random tiny matrix → destroys growth!

- If BS algo composed: $x \xrightarrow{\tilde{f}} y \xrightarrow{\hat{g}} z$ then is BS — did this earlier after BS

1:10 hr. got here)

QR factorization.
- decomposition of $A \in \mathbb{C}^{m \times n}$ into $A = Q R$ (A: $m \times n$, Q: $m \times n$, R: $n \times n$ upper-tri.; Q o.n. columns, $q_1 \cdots q_n$)
- useful for LSQ. Idea of matrix factorizations transformed 20th C. computing
- ∃ BS algs.

$A = \begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix}$ $A = QR$ says: $a_1 = r_{11} q_1$
$a_2 = r_{12} q_1 + r_{22} q_2$ etc : mix of only 1st two. reminds of? $\{q_i\}_{i=1}^j$ o.n.b for 1st j cols of A.

Alg 1: Gram-Schmidt: also proves existence of QR:

```
for j = 1..n
    v_j = a_j
    for i = 1..j-1
        v_j = v_j - (q_i* a_j) q_i
    end
    q_j = v_j / ||v_j||
end
```

$q_i^* a_j$: call scalar r_{ij}. $\|v_j\|$: call r_{jj}. v_j is a working vector.