Sensitivity of Linear Systems

Numerical algorithms for

- Square linear systems (Gaussian Elimination*)
- Least-squares problems (QR factorization)
- Eigenvalue problems (QR algorithm)

=> These algorithms are backward stable, or behave like backward stable algorithms in practice*.

How do we understand their accuracy?

=> Condition #'s convert backward error bounds into forward error bounds (accuracy).

If $$\kappa = \lim_{\delta \to 0} \sup_{\|\delta x\| \leq \delta} \left[\frac{\|\delta f\|}{\|f(x)\|} \Big/ \frac{\|\delta x\|}{\|x\|} \right]$$

See Lecture 4 notes.

where $\delta f = f(x+\delta x) - f(x)$,

then $$\underbrace{\frac{\|\delta f\|}{\|f\|}}_{\text{accuracy}} \leq \underbrace{\kappa}_{\text{condition \#}} \underbrace{\frac{\|\delta x\|}{\|x\|}}_{\text{backward error}} + o(\|\delta x\|)$$

Matrix-Vector Multiplication

Compute $\quad \underset{\substack{\uparrow \\ \text{output}}}{b} = f(x) = \overset{\text{"fixed" for now}}{A}\, \underset{\substack{\uparrow \\ \text{input}}}{x}.$

Condition # is

$$\kappa_A(x) = \lim_{\delta \to 0} \sup_{\|\delta x\| \le \delta} \left[\frac{\|A(x+\delta x) - Ax\|}{\|Ax\|} \Big/ \frac{\|\delta x\|}{\|x\|} \right]$$

$$= \lim_{\delta \to 0} \sup_{\|\delta x\| \le \delta} \left[\frac{\|A\delta x\|}{\|\delta x\|} \Big/ \frac{\|Ax\|}{\|x\|} \right]$$

$$= \lim_{\delta \to 0} \|A\| \frac{\|x\|}{\|Ax\|} = \|A\| \frac{\|x\|}{\|b\|}$$

Square Linear System

Solve $\quad Ax = b \quad \overset{\text{Math.}}{\underset{\text{equiv.}}{\Longleftrightarrow}} \quad$ Compute $\quad \underset{\substack{\uparrow \\ \text{output}}}{x} = A^{-1}b = f(\underset{\substack{\uparrow \\ \text{input}}}{b})$

$$\Rightarrow \kappa_{A^{-1}}(b) = \|A^{-1}\| \frac{\|b\|}{\|x\|} \qquad \text{(perturbations in } x\text{)}$$

These bounds depend on x, and $b = Ax$. But

$$\|x\| = \|A^{-1}Ax\| \leq \|A^{-1}\| \|Ax\|$$

$$\Rightarrow \kappa_A(x) = \|A\| \frac{\|x\|}{\|Ax\|} \leq \|A\| \|A^{-1}\|$$

Upper Bound is Independent of x, b

$$\Rightarrow \kappa_{A^{-1}}(b) \leq \|A\| \|A^{-1}\| = \kappa(A)$$

$\kappa(A)$ is called the condition # of A.

What about perturbations to A itself?

$$\underbrace{x}_{\text{output}} = f(A) = (\underbrace{A}_{\text{input}})^{-1} \underbrace{b}_{\text{"fixed"}}$$

i.e.,

$$(\underbrace{A + \delta A}_{\text{perturbed input}})(\underbrace{x + \delta x}_{\text{perturbed output}}) = b$$

$$\Rightarrow \quad Ax = b, \text{ so } (\delta A)x + A(\delta x) = 0$$

$$\Rightarrow \quad \delta x = -A^{-1}(\delta A)x$$

$$\Rightarrow \quad \|\delta x\| \leq \|A^{-1}\| \|\delta A\| \|x\|$$

Condition # is then bounded by

$$\kappa_\delta(A) = \lim_{\delta\to 0} \sup_{\|\delta x\|\le\delta} \left[\frac{\|\delta x\|}{\|x\|} \Big/ \frac{\|\delta A\|}{\|A\|} \right]$$

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

<u>The condition # of a matrix</u>

The upper bound $\kappa(A)$ has an elegant interpretation in terms of the SVD, for 2-norm.

$$A = \underbrace{\begin{bmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} - & v_1^* & - \\ & \vdots & \\ - & v_m^* & - \end{bmatrix}}_{V^*}$$

$$\|A\|_2 = \|\Sigma\|_2 = \sigma_1$$

$$A^{-1} = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_m \\ | & & | \end{bmatrix} \begin{bmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_m \end{bmatrix} \begin{bmatrix} - & u_1^* & - \\ & \vdots & \\ - & u_m^* & - \end{bmatrix}$$

$$\|A^{-1}\|_2 = 1/\sigma_m$$

"how A magnifies"

$$\boxed{\kappa_2(A) = \frac{\sigma_1}{\sigma_m}}$$

how A "shrinks"

"Ill-conditioned" matrices either drastically magnify or shrink some directions relative to others → cancellation errors are significant!

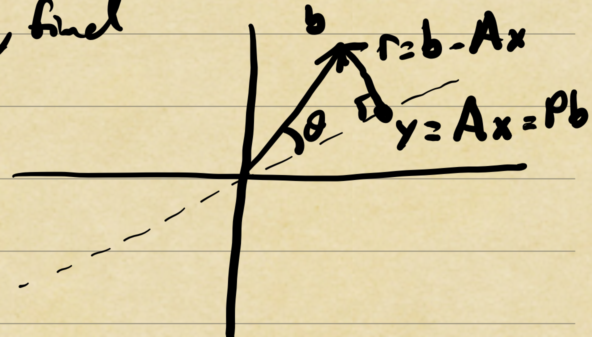
Least-Squares Problems

$\|\cdot\| = \|\cdot\|_2$ in this section. A of full col. rank.

$$\begin{array}{c c} & n \\ m & \underbrace{\begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix}}_{A} \end{array} = \begin{array}{c} m \\ \underbrace{\begin{bmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{bmatrix}}_{U} \end{array} \begin{array}{c} n \\ \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \end{array} \begin{array}{c c} n & \\ \begin{bmatrix} — & v_1^* & — \\ & \vdots & \\ — & v_n^* & — \end{bmatrix} & n \end{array}$$

minimize $\|Ax - b\|_2$, i.e., find

$$x_* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \|Ax - b\|_2$$



Diagonalize: $y = V^* x$, $d = U^* b$

$$\| \underbrace{U \Sigma V^*}_{A} x - b \| = \| \Sigma y - d \|$$

[orthogonal transform doesn't change length]

$$\Rightarrow \left[\begin{array}{ccc} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{array}\right] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \left[\begin{array}{c} d_1 \\ \vdots \\ d_n \\ \hline d_{n+1} \\ \vdots \\ d_m \end{array}\right] \quad \text{Solve } \Sigma y_* = d_{1:n}$$

$$\} = \Sigma y_* - d$$

so $\|Ax_* - b\| = \|d_{n+1:m}\|$ and

$$x_* = V y_* = V \Sigma^{-1}_{1:n,1:n} d_{1:n}$$

$$= \underbrace{\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_n^{-1} \end{bmatrix} \begin{bmatrix} - & u_1^* & - \\ & \vdots & \\ - & u_n^* & - \end{bmatrix}}_{\text{pseudo inverse } A^+} \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

linear map $\Rightarrow x_* = A^+ b$

Jacobian $x_* + \delta x_* = A^+(b + \delta b)$

$$J = A^+$$

Condition # $\kappa_A(b) = \dfrac{\|A^+\|}{\|x_*\|/\|b\|} = \|A^+\| \dfrac{\|b\|}{\|Ax\|} \dfrac{\|Ax\|}{\|x\|}$

$$= \|A^+\| \frac{1}{\cos\theta} \frac{\|A\|}{\eta}$$

where

$$\eta = \frac{\|A\|\|x\|}{\|Ax\|}$$ measures how

$\|Ax\|$ compares to its "maximum allowed" value

$$\boxed{\kappa_A(b) = \frac{\|A\|\|A^+\|}{\eta\cos\theta} = \frac{\sigma_1/\sigma_n}{\eta\cos\theta}}$$

$1 \leq \eta \leq \kappa(A)$, $\quad 0 \leq \theta \leq \frac{\pi}{2}$, $\quad 1 \leq \sigma_1/\sigma_n < \infty$.

$\kappa_A(b)$ may be large when A is ill-conditioned ($\sigma_1/\sigma_n \gg 1$) or when $\theta \approx \frac{\pi}{2}$.

Corresponds to 2-step soln:

1) Orthogonal projection onto range(A).

2) Solve reduced $n \times n$ system for y_*.

$\theta \approx \frac{\pi}{2}$, step 1 is ill-conditioned.
$\sigma_1/\sigma_n \gg 1$, step 2 is ill-conditioned.