Stability in solving linear systems:

previous example shows that Gaussian elmination needs (partial) pivoting to achieve stability : by this we mean the numerical solution x to $Ax = b$ is not perturbed from the true solution x_t if A is perturbed by round-off errors.

More generally, we also want to know how the solution varies as ① rhs vector b is perturbed $\tilde{b} = b + \delta b$

② $\tilde{A} = A + \delta A$

$$Ax = b$$

$$A\hat{x} = \hat{b}$$

assuming A is nonsingular

$$A(x - \hat{x}) = b - \hat{b} \qquad (x - \hat{x}) = \frac{b - \hat{b}}{A}$$

$$(x-\hat{x}) = A^{-1}(b-\hat{b})$$

$$\|x-\hat{x}\| = \|A^{-1}(b-\hat{b})\| \le \|A^{-1}\| \cdot \|b-\hat{b}\|$$

$$\frac{\|x-\hat{x}\|}{\|x\|} \le \frac{\|A^{-1}\|\,\|b-\hat{b}\|}{\|x\|} = \|A^{-1}\|\,\|A\| \cdot \frac{\|b-\hat{b}\|}{\|A\| \cdot \|x\|}$$

recall that $Ax = b$

$$\|b\| \le \|A\|\,\|x\|$$

$$\therefore \frac{\|x-\hat{x}\|}{\|x\|} \le \|A^{-1}\|\,\|A\| \cdot \frac{\|b-\hat{b}\|}{\|b\|}$$

if $\|A^{-1}\| \cdot \|A\| >> 1$ a small deviation in b causes a lot of difference in $\hat{x}$ from x

$\Rightarrow$ ill-conditioned

if $\|A^{-1}\|\,\|A\|$ ~~and~~ is small $\hat{x}$ is always close to x

$\Rightarrow$ well-conditioned.

$$\text{condition number}(A) \equiv \|A^{-1}\|\,\|A\|$$

Remark: $\|A^{-1}\|\,\|A\| \ge \|A^{-1}A\| = 1$

$\therefore$ condition number $(A) \ge 1$

useful inequalities involved with vector & matrix norms:

$$\|y+z\| \le \|y\| + \|z\| \qquad \big|\,\|y\| - \|z\|\,\big| \le \|y-z\|$$

$$\|A+B\| \le \|A\| + \|B\|$$

$$\|AB\| \le \|A\|\,\|B\|$$

$$\|Az\| \le \|A\|\,\|z\|$$

$$A\,x = b$$

$$\hat{A}\,\hat{x} = b$$

$$(\hat{A}-A+A)\,\hat{x} = b$$

$$A\,\hat{x} = b + (A-\hat{A})\,\hat{x}$$

$$A(x-\hat{x}) = (A-\hat{A})\,\hat{x}$$

$$\|x-\hat{x}\| \cdot (1 - \|A^{-1}\|\,\|A-\hat{A}\|) \le \|A^{-1}\|\,\|A-\hat{A}\|\,\|x\|$$

$$(\hat{A}-A+A)\,\hat{x} = A\,x$$

$$A^{-1}(A \mp (\hat{A}-A))\hat{x} = x$$

$$\|I - A^{-1}(\hat{A}-A)\|\,\|\hat{x}\| \ge \|x\|$$

$$\|x-\hat{x}\| \le \|A^{-1}\|\,\|A-\hat{A}\|\,\|\hat{x}\|$$

$$\|x-\hat{x}\| \le \frac{\|A^{-1}\|\,\|A-\hat{A}\| \cdot \|x\|}{1-\|A\|^{-1}\|A-\hat{A}\|}$$

$$\therefore \quad \frac{\|x-\hat{x}\|}{\|x\|} \le \frac{\mathrm{cond}(A)}{1-\mathrm{cond}(A)\cdot\frac{\|A-\hat{A}\|}{\|A\|}} \cdot \frac{\|A-\hat{A}\|}{\|A\|}$$

again, if cond(A) is large, the system is ill-conditioned

Example:

Calculate the cond(A) for

$$A = \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix} \qquad |c| \neq 1$$

$$A^{-1} = \frac{1}{1-c^2}\begin{bmatrix} 1 & -c \\ -c & 1 \end{bmatrix}$$

$$\|A\| = 1 + |c|$$

$$\|A^{-1}\| = \frac{1+|c|}{|1-c^2|} \qquad \text{cond}(A) = (1+|c|)\cdot\frac{1+|c|}{|1-c^2|}$$

as $|c| \to 1$ $\quad$ $\text{cond}(A) \to \infty$ $\quad \to$ the system is
$\det(A) \to 0$ $\quad$ ill-conditioned.

Example:

$$19x_1 + 20x_2 = b_1$$
$$20x_1 + 21x_2 = b_2$$

$$A\,x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad A = \begin{bmatrix} 19 & 20 \\ 20 & 21 \end{bmatrix}$$

$$\|A\| = 41$$

$$A^{-1} = \begin{bmatrix} -21 & 20 \\ 20 & -19 \end{bmatrix} \qquad \|A^{-1}\| = 41 \qquad \text{cond}(A) = 41^2 = 1681 \gg 1$$

Theorem $\quad r_\sigma(A) \leq \|A\|$

$$\text{Cond}(A) = \|A\| \, \|A^{-1}\| \geq r_\sigma(A) \, r_\sigma(A^{-1})$$

$$\therefore \text{Cond}(A) \geq \frac{\max_{\lambda \in \sigma(A)} |\lambda|}{\min_{\lambda \in \sigma(A)} |\lambda|} \equiv \text{cond}(A)_*$$

Example: Consider the linear system

$$7x_1 + 10x_2 = b_1$$

$$5x_1 + 7x_2 = b_2$$

$$A = \begin{bmatrix} 7 & 10 \\ 5 & 7 \end{bmatrix} \qquad A^{-1} = \begin{bmatrix} -7 & 10 \\ 5 & -7 \end{bmatrix}$$

$$\text{cond}(A) = 17 \cdot 17 = 289$$

$$\text{Cond}(A)_* = \frac{7+5\sqrt{2}}{|7-5\sqrt{2}|} \sim 198$$

Let $|\lambda_l| = \min_{\lambda \in \sigma(A)} |\lambda| \qquad |\lambda_u| = \max_{\lambda \in \sigma(A)} |\lambda|$

$$\text{cond}(A)_* = \frac{|\lambda_u|}{|\lambda_l|}$$

Example: $A = \begin{bmatrix} 1 & 0 \\ 0 & 10^{-10} \end{bmatrix}$ ill-conditioned?

Stiffness in ODE (PDE) & ill-conditioned systems?

$$\partial_t u = \Delta u$$

$$\partial_t u = -k^2 u$$

$\lambda u = -k^2 u$ $\quad \lambda = -k^2 \rightarrow$ stiff ODE

condition # $\quad \dfrac{\max(k^2)}{\min(k^2)} \gg 1$

Application of the above stability analysis to Gaussian elimination:

1. $LU = A + E$

$$\|E\|_\infty \le n^2 \cdot \|A\|_\infty \cdot \delta \cdot \overbrace{\boxed{\frac{1}{\|A\|_\infty} \max_{i,j,k} |a_{ij}^{(k)}|}}^{\rho}$$

δ is the unit round on the computer

Note: δ is a positive floating-point number
it is the smallest s.t. $fl(1+\delta) > 1$

2. $$\frac{\|x - \hat{x}\|_\infty}{\|x\|_\infty} \le \frac{\mathrm{cond}(A)_\infty}{1 - \mathrm{cond}(A)_\infty \frac{\|\delta A\|_\infty}{\|A\|_\infty}} \cdot \frac{\|\delta A\|_\infty}{\|A\|_\infty}$$ from previous results

with $$\frac{\|\delta A\|_\infty}{\|A\|_\infty} \le 1.01 \cdot \{(n^3 + 3n^2)\rho u\}$$

$\le nu$ smaller upper bound

Iteration Method : $Ax = b$

$A = N + P$, $\begin{cases} N \text{ is a nonsingular matrix} \\ Nz = f \text{ is easy to solve for general } f \end{cases}$

$P = A - N$

The iteration method is defined as

$$N \cdot x^{(k+1)} = b + P \cdot x^{(k)} \quad k = 0, 1, 2, \dots$$

Jacobi method: $N = \begin{bmatrix} a_{11} & & 0 \\ & a_{22} & \\ 0 & \ddots & a_{nn} \end{bmatrix}$ diagonal

Gauss-Seidel method: $N = \begin{bmatrix} a_{11} & & & 0 \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$ lower triangular

error analysis: $N e^{(k+1)} = P e^{(k)}$

$$e^{(k+1)} = N^{-1} P e^{(k)}$$

$$\| e^{(k+1)} \| \leq \| N^{-1} P \| \, \| e^{(k)} \| = \| M \| \, \| e^{(k)} \|$$

$$\| e^{(k+1)} \| \leq \| M \|^{k} \| e^{(0)} \|$$

error converges to zero if $\|M\| < 1$

For the Jacobi method, $N \Rightarrow$ the diagonal matrix from A

$$\|M\| = \max_{1 \le i \le n} \sum_{\substack{j=1 \\ j \ne i}}^{n} \left| \frac{a_{ij}}{a_{ii}} \right|$$

for stability $\boxed{\sum_{\substack{j=1 \\ j \ne i}}^{n} |a_{ij}| < |a_{ii}| \qquad 1 \le i \le n.}$

More generally, $N x^{(k+1)} = b + P x^{(k)}$ $k = 0, 1, 2, \cdots$

will converge for all b and all initial guess $x^{(0)}$ iff

all eigenvalues λ of $M = N^{-1} P$ satisfy $|\lambda| < 1$

partial proof: Assume the eigen value λ of M satisfies $|\lambda| \ge 1$

choose $e^{(0)}$ to be the eigenvector associated with λ.

$$\|e^{(k)}\| = \|M^k e^{(0)}\| = \|\lambda^k e^{(0)}\| = |\lambda|^k \cdot \|e^{(0)}\|$$

$\|e^{(k)}\| \ne 0$ as $|\lambda| \ge 1$

therefore, for convergence $|\lambda| < 1$ is needed.