Conditioning

$Ax = b$.

(1) Defining $\underset{(N\times N)}{A}$ → random
→ Hilbert matrix $\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \cdots & & & \frac{1}{n+1} \\ \vdots & & & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{bmatrix}$

(2) Start with exact solution $X_{true} = [1, 1, \dots, 1]$

(3) Compute $b = A @ X_{true}$

(4) Solve $\left.\begin{matrix} b \\ A \end{matrix}\right\} \to \boxed{} \longrightarrow \boxed{X_{solve}}$ output.

input

(5) Computer error $\| X_{solve} - X_{true} \|$.

Sensitivity of solutions of linear systems.

change input $b \to b + \Delta b$.

$\Longrightarrow$ How large is Δx ?

$$A(x+\Delta x) = b + \Delta b \quad \Longrightarrow \quad A\Delta x = \Delta b.$$

$$\frac{\text{Output relative error}}{\text{Input relative error}} = \frac{\|\Delta x\| / \|x\|}{\|\Delta b\| / \|b\|} = \frac{\|\Delta x\|\,\|b\|}{\|\Delta b\|\,\|x\|}$$

$$= \frac{\|A^{-1}\Delta b\|\,\|Ax\|}{\|\Delta b\|\,\|x\|} \leq \frac{\|A^{-1}\|\,\cancel{\|\Delta b\|}\,\|A\|\,\cancel{\|x\|}}{\cancel{\|\Delta b\|}\,\cancel{\|x\|}} = \|A^{-1}\|\,\|A\|$$

$$\Longrightarrow \quad \frac{\text{output relative error}}{\text{Input relative error}} \leq \|A^{-1}\|\,\|A\|$$

$$\frac{\|\Delta x\|}{\|x\|} \leq \underbrace{\|A^{-1}\|\|A\|}_{cond(A)} \frac{\|\Delta b\|}{\|b\|}$$

change input $A \to A+E$

$$(A+E)(x+\Delta x) = b$$

$$\cancel{Ax} + A\Delta x + Ex + E\Delta x = \cancel{Ax}$$

$$(A+E)\Delta x = -Ex.$$

$$\Delta x = -(A+E)^{-1}Ex$$

$$\frac{\|\Delta x\|\,\|A\|}{\|x\|\,\|E\|} \leq \frac{\|(A+E)^{-1}\|\,\cancel{\|E\|}\,\cancel{\|x\|}\,\|A\|}{\cancel{\|x\|}\,\cancel{\|E\|}}$$

$$\frac{\|\Delta x\|}{\|x\|} \leq \|(A+E)^{-1}\|\,\|A\|\,\frac{\|E\|}{\|A\|}$$

$$\approx \underbrace{\|A^{-1}\|\,\|A\|}_{cond(A)}$$

large : ill-conditioned

small : well-conditioned.

Small condition numbers mean not a lot of error amplification. $\Longrightarrow$ we want small cond(A).

$$cond(A) = \|A\|\|A^{-1}\| \geq \|I\| = 1.$$

can be defined as $cond_\infty(A)$

$cond_2(A)$

1. $cond(A) \geq 1$.

2. $cond(\gamma A) = cond(A)$

3. $\forall$ Diagonal matrix D, $cond(D) = \frac{\max(d_i)}{\min(d_i)}$.

4. $cond(A)$ is large. A is nearly singular.

near singularity is not good.

5. orthogonal matrix A. $cond(A) = \|A^{-1}\|\|A\| = \|A^T\|\|A\| = 1$.

always optimal.

Residual versus error.

$Ax = b$.

$\hat{x} = (x + \Delta x)$

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

error vector: $e = \Delta x = \hat{x} - x$.

residual vector: $r = b - A\hat{x}$.

Relative residual: $\frac{\|r\|}{\|A\|\|x\|}$ when A is ill it will be smaller.

Relative error: $\frac{\|\Delta x\|}{\|x\|}$

$\frac{\|r\|}{\|A\|\|\hat{x}\|} \leq C\,\epsilon_m$

constant C $\begin{cases} \text{"large" when LU/Gaussian performed without pivoting} \\ \text{"small" when use partial pivoting.} \end{cases}$

ϵ_m: machine Epsilon

$$r = Ax - A\hat{x} = -A\Delta x$$

$$\frac{\|\Delta x\|}{\|x\|} = \frac{\|A^{-1} r\|}{\|x\|} \leqslant \frac{\|A^{-1}\| \|r\|}{\|x\|} = \|A^{-1}\| \|A\| \frac{\|r\|}{\|x\| \|A\|}$$

$$\Longrightarrow \frac{\|\Delta x\|}{\|x\|} \leqslant \mathrm{cond}(A) \frac{\|r\|}{\|x\| \|A\|}$$

if entries in A and b are accurate to S decimal digits, and $\mathrm{cond}(A) = 10^{w}$

$$\frac{\|\Delta x\|}{\|x\|} \leqslant 10^{w} \cdot \frac{\|\Delta b\|}{\|b\|} \leqslant 10^{w} \cdot 10^{-s} = 10^{w-s} = 10^{-(s-w)}$$

$\Longrightarrow$ $\hat{x}$ will be accurate to $S-W$ decimal digits.

Example: $\mathrm{cond}(A) = 10^{7}$ $\quad W = 7$

working with IEEE double precision

$2^{-52} \approx 2.2 \times 10^{-16}$ $\quad S = 16.$

$\Longrightarrow S - W = 9.$