Lecture 20 – Condition number of matrices

Special matrices & their factorization

$$\boxed{Ax = b}$$

$A \to n \times n$

$x \to n \times 1$

$b \to n \times 1$

To solve without inverting

Condition number

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

- Subtractive Cancellative
- Addition of large and small numbers

$\|A\|_e$ Euclidean norm $= \sqrt{\sum_{i,j} a_{i,j}^2}$

$$\left.\begin{aligned} \|A\|_1 &= \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{i,j}| \\ \|A\|_\infty &= \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{i,j}| \end{aligned}\right\} \text{Easier to compute}$$

$\mathrm{Cond}(A) = \|A\| \, \|A^{-1}\|$ numerically expensive

=> Error propagation

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

True model

i.e. no model deficiencies

$b = \tilde{b} + \Delta b$ perturbation/error

How does the system A propagate the error in b?

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

estimate

$\hat{x} = \tilde{x} + \Delta x$

$$\frac{\|\Delta x\|}{\|\tilde{x}\|} \Big/ \frac{\|\Delta b\|}{\|\tilde{b}\|}$$

ratio of relative errors

$$\frac{\|\Delta x\|}{\|\tilde{x}\|} \Big/ \frac{\|\Delta b\|}{\|\tilde{b}\|} = \frac{\|\overset{A^{-1}\Delta b}{\Delta x}\| \, \|\tilde{b}\|}{\|\tilde{x}\| \, \|\Delta b\|} = \frac{\|A^{-1}\Delta b\| \, \|A\tilde{x}\|}{\|\tilde{x}\| \, \|\Delta b\|}$$

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

inequality

$$\|AB\| \leq \|A\| \, \|B\|$$

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

↳ close to 1

too large then

errors in $\tilde{x}$ will be large

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

$\hat{b} - b \quad \hat{b}$

$\hat{\varepsilon} = b - \underbrace{A\hat{x}} \quad - \ 1$

$0 = b - Ax \quad - \ 2$

$1 - 2$

$\hat{\varepsilon} = Ax - A\hat{x}$

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

$\hat{\varepsilon} = A\,\Delta x$

can estimate

$\boxed{\Delta x = \underbrace{A^{-1}} \hat{\varepsilon}}$

large value if cond(A) is poor

cond(A) good

close to 1

Special matrices

$$y = A x$$

y: $m \times 1$, A: $m \times n$, x: $n \times 1$

$m \geq n$

$r = m - n$

redundancy

$$\Phi = (y - Ax)^T (y - Ax)$$

$\Rightarrow$ Normal equation system

$$(A^T A) x = A^T y$$

$$\underset{n \times n}{N} \underset{n \times 1}{x} = \underset{n \times 1}{b}$$

$A^T A \Rightarrow$ inner product of A

$\downarrow$

symmetric matrix

$\boxed{\underline{A A^T} \text{ outer}}$

$LU \rightarrow N$

$$A = L U$$

$\downarrow$

symmetric

$$A = L L^T$$

Sparse matrices

$$\begin{bmatrix} l_{11} & & 0 \\ l_{21} & l_{22} & \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ & & l_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

$l_{11}^2 = a_{11}$

$l_{21} l_{11} = a_{12}$ $\qquad$ $l_{21}^2 + l_{22}^2 = a_{22}$

$l_{31} l_{11} = a_{13}$ $\qquad$ $l_{31} l_{21} + l_{32} l_{22} = a_{23}$ $\qquad$ $l_{31}^2 + l_{32}^2 + l_{33}^2 = a_{33}$

$l_{11} = \sqrt{a_{11}}$

$l_{21} = \dfrac{a_{12}}{l_{11}}$ $\qquad$ $l_{22} = \sqrt{a_{22} - l_{21}^2}$

$l_{31} = \dfrac{a_{13}}{l_{11}}$ $\qquad$ $l_{32} = \dfrac{a_{23} - l_{31} l_{21}}{l_{22}}$ $\qquad$ $l_{33} = \sqrt{a_{33} - l_{31}^2 - l_{32}^2}$

$A = LDL^T$

$A = LL^T$

Cholesky factorization

↓

least squares

LAPACK

BLAS

Decompose

$Ux = d$

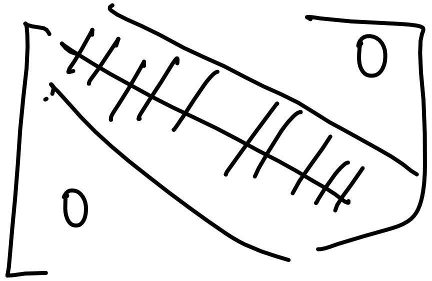
Forward

$Ld = y$

Back

$\tilde{x}$

Tridiagonal system — Banded matrices



$$\begin{bmatrix} f_1 & g_1 & & 0 \\ e_1 & f_2 & g_2 & \\ 0 & e_2 & f_3 & g_3 \\ & & e_3 & f_4 \end{bmatrix}$$

$u_{33} = f_3 - l_{32} u_{23}$

3 → decomposition

loops

loop

LU → Tridiagonal system

$$\begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} = \begin{bmatrix} f_1 & g_1 & 0 \\ e_1 & f_2 & g_2 \\ 0 & e_2 & f_3 \end{bmatrix}$$

$u_{11} = f_1 \qquad u_{12} = g_1 \qquad u_{13} = 0$

$l_{21} u_{11} = e_1 \qquad l_{21} u_{12} + u_{22} = f_2 \qquad \underbrace{l_{21} u_{13}}_{0} + u_{23} = g_2$

$\underbrace{l_{31} u_{11}}_{0} = 0 \qquad \underbrace{l_{31} u_{12}}_{0} + l_{32} u_{22} = e_2 \qquad l_{31} u_{13} + l_{32} u_{23} + u_{33} = f_3$

$u_{11} = f_1 \quad u_{12} = g_1 \quad u_{13} = 0$

$l_{21} = \dfrac{e_1}{u_{11}} \quad u_{22} = f_2 \quad u_{23} = g_2$

$l_{31} = 0 \quad l_{32} = e_2 / u_{22} \quad u_{33} = f_3$