CS 450: Numerical Anlaysis

Lecture 5

Chapter 2 – Linear Systems

Solving Linear Systems

Edgar Solomonik

Department of Computer Science
University of Illinois at Urbana-Champaign

Matrix-vector products

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

$\kappa(A)$ - conditioning

$y = Ax$

x ↑ input, δx - perturbation

$$\frac{\|\delta y\|}{\|y\|} \leq \kappa(A) \frac{\|\delta x\|}{\|x\|}$$

$Ax = b$

b ↑ input, δb - perturbation

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\delta b\|}{\|b\|}$$

$(A + \delta A)x = b$

A ↑ input

$$\Longrightarrow \frac{\|\delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\delta A\|}{\|A\|}$$

Solving Basic Linear Systems

- Solve $\boldsymbol{D}\boldsymbol{x} = \boldsymbol{b}$ if $\boldsymbol{D}$ is diagonal

  $x_i = b_i / d_{ii}$ $\quad O(n)$ work

- Solve $\boldsymbol{Q}\boldsymbol{x} = \boldsymbol{b}$ if $\boldsymbol{Q}$ is orthogonal

  $x = Q^T b$ $\quad O(n^2)$ work

- Given SVD $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$, solve $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

  $U \underbrace{\Sigma V^T x}_{y} = b$, solve $Uy = b \Rightarrow y = U^T b$

  $\Sigma z = y$

  $x = Vz$

Solving Triangular Systems

- $Lx = b$ if L is lower-triangular is solved by forward substitution:

$$\begin{bmatrix} \ell_{11} & & \\ \ell_{21} & \ell_{22} & \\ \vdots & & \ddots \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b \end{bmatrix}$$

$$\begin{aligned} l_{11}x_1 &= b_1 & \qquad & x_1 = b_1 / \ell_{11} \\ l_{21}x_1 + l_{22}x_2 &= b_2 & \Rightarrow \qquad & x_2 = (b_2 - \ell_{21}x_1)/\ell_{22} \\ l_{31}x_1 + l_{32}x_2 + l_{33}x_3 &= b_3 & & x_3 = (b_3 - \ell_{31}x_1 - \ell_{32}x_2)/\ell_{33} \\ &\vdots & & \vdots \end{aligned}$$

- Algorithm can also be formulated recursively by blocks:

same

scalar

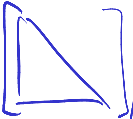
vector

lower-tri. matrix

$$\begin{array}{cc} 1 & n-1 \\ \left[\begin{array}{c|c} \ell_{11} & \\ \hline \ell_{21} & L_{22} \end{array}\right] \end{array} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \Rightarrow \begin{aligned} x_1 &= b_1 / \ell_{11} \\ L_{22} x_2 &= b_2 - \ell_{21} x_1 \end{aligned}$$

Solving Triangular Systems

- **Existence of solution to $Lx = b$:**



if $l_{ii} \neq 0 \ \forall i$, $\mathrm{rank}(L) = n$

$\boxed{\text{if } l_{ii} = 0}$, may not have sol'n

- **Uniqueness of solution:**

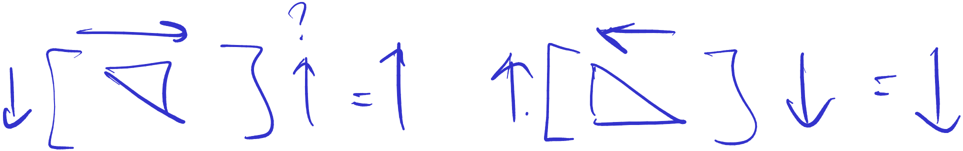
$\rightarrow$ can be non-unique

can minimize $\|x\|$

- **Computational complexity of forward/backward substitution:**

$$T(n) = T(n-1) + 2n = \sum_{i=1}^{n} 2i \approx n^2$$

Backward Substitution



Properties of Triangular Matrices

- $Z = XY$ is lower triangular is X and Y are both lower triangular:

$$\begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} x_{11} & \overset{0}{\overset{\shortparallel}{x_{12}}} \\ x_{21} & X_{22} \end{bmatrix} \begin{bmatrix} y_{11} & \overset{0}{\overset{\shortparallel}{y_{12}}} \\ y_{21} & Y_{22} \end{bmatrix} \Rightarrow \overset{0}{\overset{\shortparallel}{z_{12}}} = x_{11} \overset{0}{\overset{\shortparallel}{y_{12}}} + \underset{0}{\underset{\shortparallel}{x_{12} Y_{22}}}$$

same argument

$$Z_{22} = X_{22} Y_{22}$$

- L^{-1} is lower triangular if it exists:

$B = L^{-1}$

$$\begin{bmatrix} L_{11} & \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} B_{11} & \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I & \\ & I \end{bmatrix}$$

$$L_{21} B_{11} + L_{22} B_{21} = 0$$

$$\Rightarrow L_{11} B_{11} = I \Rightarrow B_{11} = L_{11}^{-1} \quad \Big| \quad B_{21} = -B_{22} L_{21} B_{11}$$

$$Ax = b$$

Direct methods: factorizing A in terms of

$\left\{ \begin{array}{l} \text{diagonal} \\ \text{triangular} \\ \text{orthogonal} \end{array} \right.$ matrices

$$A = XYZW$$

$$X(Y(Z(Wx))) = b$$

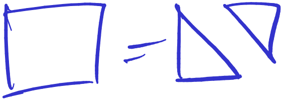
iterative methods - later

good for sparse A

so when $a_{ij} = 0$ for many (i,j)

LU Factorization

- An ***LU factorization*** consists of a unit-diagonal lower-triangular ***factor*** L and upper-triangular factor U such that $A = LU$:



$$\Rightarrow L = \begin{bmatrix} 1 & & & \\ \ell_{21} & 1 & & \\ \ell_{31} & \ell_{32} & 1 & \\ \vdots & \vdots & \ddots & 1 \end{bmatrix}$$

$$Ax = b$$
$$LUx = b$$
$$Ly = b$$
$$Ux = y$$

- Given an LU factorization of A, we can solve the linear system $Ax = b$:

Gaussian Elimination Algorithm

- **Algorithm for factorization is derived from equations given by $A = LU$:**

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} 1 & \\ \overset{}{\underset{?}{l_{21}}} & \underset{?}{L_{22}} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ & U_{22} \end{bmatrix} ?$$

Schur complement

$$u_{11} = a_{11} \qquad u_{12} = a_{12} \qquad l_{21} = a_{21}/u_{11}, \quad S_{22} = A_{22} - l_{21}u_{12}$$

- **The computational complexity of LU is $O(n^3)$:**

$$T(n) = T(n-1) + 2n^2 = \sum_{i=1}^{n} 2i^2 \approx \frac{2}{3}n^3$$

$$\square - 1$$

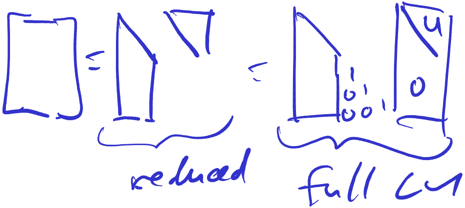
$$L_{22}U_{22} = S_{22}$$

Existence of LU factorization

where A_{11} is singular, then LU does not exist

$\exists\ A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$

- **The LU factorization may not exist:** Consider matrix $\begin{bmatrix} 3 & 2 \\ 6 & 4 \\ 0 & 3 \end{bmatrix}$.



$\begin{bmatrix} 3 & 2 \\ 6 & 4 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & & \\ l_{21} & 1 & \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ & u_{22} \end{bmatrix}$

$= \begin{bmatrix} 1 & \\ 2 & 1 \\ 0 & l_{32} \end{bmatrix} \begin{bmatrix} 3 & 2 \\ & u_{22} \end{bmatrix}$

$\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ u_{22} \end{bmatrix} = 4 = 4 + u_{22}$

$u_{22} = 0$

$\begin{bmatrix} 0 & l_{32} \end{bmatrix} \begin{bmatrix} 2 \\ u_{22} \end{bmatrix} = 3$

$2 \cdot 0 + l_{32} u_{22} = 0$

- **Permutation of rows enables us to transform the matrix so the LU factorization does exist:**

$\exists\ P$ – permutation matrix, so $Pv = \begin{bmatrix} v_3 \\ v_4 \\ v_2 \\ v_1 \end{bmatrix}$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix}$

$P^T = P^{-1}$

$LU = PA$

Gaussian Elimination with Partial Pivoting

- ***Partial pivoting*** **permutes rows to make divisor u_{ii} is maximal at each step:**

$$\begin{bmatrix} a_{11} & & \\ a_{21} & \cdots & \\ a_{31} & & \end{bmatrix} \quad 2 = \underset{i}{\operatorname{argmax}}(|a_{i1}|) \Rightarrow \begin{bmatrix} a_{21} & & \\ a_{11} & \cdots & \\ a_{31} & & \end{bmatrix} = B$$

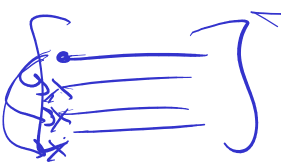
$$PA = B \quad \text{where } P_{12} = \begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix}$$

- **A row permutation corresponds to an application of a *row permutation matrix*** $\boldsymbol{P}_{jk} = \boldsymbol{I} - (\boldsymbol{e}_j - \boldsymbol{e}_k)(\boldsymbol{e}_j - \boldsymbol{e}_k)^T$**:**

Complete Pivoting

- ***Complete pivoting*** **permutes rows and columns to make divisor u_{ii} is maximal at each step:**

ℓ_{21} – multipliers used to eliminate column



Complete pivoting, selects $\underset{(i,j)}{\arg\max}\, |a_{ij}|$ as pivot element

partial

Schur complement $A_{22} - \ell_{21} u_{12}$

$\|\ell_{21}\|_\infty \le 1$

- **Complete pivoting is noticeably more expensive than partial pivoting:**

$|\ell_{21}| = |a_{21}|/u_{11}$

$\|\ell_{21}\|_\infty = \frac{\|a_{21}\|}{|u_{11}|}$

$\|u_{12}\|_\infty \le |u_{11}|$

$\|\ell_{21} u_{12}\|_\infty \le \|a_{21}\|$

$$A = \begin{bmatrix} U \end{bmatrix} \begin{bmatrix} \Sigma \end{bmatrix} \begin{bmatrix} V^T \end{bmatrix}$$

σ_{max} σ_{min}

$x + \delta x$

σ_{min} σ_{max}

σ_{n-1}

ϵ

σ_{n-2}

$$\begin{bmatrix} 0 & a_1 \\ 0 & A \end{bmatrix} = \begin{bmatrix} 1 & \\ 0 & L \end{bmatrix} \cdot \begin{bmatrix} 0 & a_1 \\ & U \end{bmatrix}$$

$$A = LU$$

$$A = \begin{bmatrix} \; \end{bmatrix} \begin{bmatrix} \; \end{bmatrix} \} k$$

$$\begin{bmatrix} 0 \end{bmatrix} = \begin{bmatrix} 1 & \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \end{bmatrix}$$

Round-off Error in LU

- **Lets consider factorization of** $\begin{bmatrix} \epsilon & 1 \\ 1 & 1 \end{bmatrix}$ **where** $\epsilon < \epsilon_{\textbf{mach}}$**:**

- **Permuting the rows of** A **in partial pivoting gives** $PA = \begin{bmatrix} 1 & 1 \\ \epsilon & 1 \end{bmatrix}$

Error Analysis of LU

- **The main source of round-off error in LU is in the computation of the Schur completement:**

- **When computed in floating point, absolute backward error δA in LU (so $\hat{\boldsymbol{L}}\hat{\boldsymbol{U}} = \boldsymbol{A} + \boldsymbol{\delta A}$) is** $|\delta a_{ij}| \leq \epsilon_{\mathsf{mach}}(|\hat{\boldsymbol{L}}| \cdot |\hat{\boldsymbol{U}}|)_{ij}$

Helpful Matrix Properties

- **Matrix is *diagonally dominant*, so $\sum_{i \neq j} |a_{ij}| \leq |a_{ii}|$:**

- **Matrix is *symmetric positive definite (SPD)*, so $\forall_{\boldsymbol{x} \neq 0}, \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} > 0$:**

- **Matrix is symmetric but indefinite:**

- **Matrix is *banded*, $a_{ij} = 0$ if $|i - j| > b$:**

Solving Many Linear Systems

- **Suppose we have computed $A = LU$ and want to solve $AX = B$ where B is $n \times k$ with $k < n$:**

- **Supposed we have computed $A = LU$ and now want to solve a perturbed system $(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^T)\boldsymbol{x} = \boldsymbol{b}$:**
  Can use the *Sherman-Morrison-Woodbury* formula

$$(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^T)^{-1} = \boldsymbol{A}^{-1} + \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^T\boldsymbol{A}^{-1}}{1 - \boldsymbol{v}^T\boldsymbol{A}^{-1}\boldsymbol{u}}$$