AMS526: Numerical Analysis I (Numerical Linear Algebra for Computational and Data Sciences)

Lecture 7: LU Factorization; Gaussian Elimination with Pivoting

Xiangmin Jiao

Stony Brook University

Outline

① Gaussian Elimination(MC§3.2, NLA§20)

2 Gaussian Elimination with Pivoting (NLA§21)

Xiangmin Jiao Numerical Analysis I 2 / 27

Review of Gaussian Elimination

• Given linear system Ax = b, where $A \in \mathbb{R}^{n \times n}$ is nonsingular, Gaussian elimination transforms the linear system to

$$Ux = y$$

where U is upper triangular, and then solves Ux = y by back substitution

- Gaussian elimination performs three types of operations
 - Add a multiple one equation to another
 - 2 Interchange two equations
 - Multiply an equation by a nonzero constant
- Gaussian elimination can be represented as transformation of augmented matrix

$$[A \mid b] \rightarrow [U \mid y]$$

through operations are on rows of matrix

ullet A is nonsingular if and only if U is nonsingular

Xiangmin Jiao Numerical Analysis I 3 / 27

Example of Gaussian Elimination

Consider the linear system

$$3x_1 + 5x_2 = 9$$
$$6x_1 + 7x_2 = 4$$

• Multiply first row by 2 = 6/2 and subtract it from second row, we obtain

$$3x_1 + 5x_2 = 9$$
$$-3x_2 = -4$$

• Then solve upper triangular system using back substitution

LU Factorization

- For Ax = b, suppose we can find A = LU, where L is lower-triangular and U is upper-triangular. Then solve the system in three steps:
 - Find A = LU (LU factorization)
 - 2 Solve Ly = b (forward substitution)
 - **3** Solve Ux = y (back substitution)
- In previous example,

$$\begin{bmatrix} 3 & 5 \\ 6 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 0 & -3 \end{bmatrix}$$

- LU factorization and Gaussian elimination are often used interchangeably
- Note: A = LU may not always exist even if A is non-singular
 - ▶ Why? Division by zero may occur
 - ▶ Remedy: Pivoting may be needed
- We start with simplest case without pivoting

Xiangmin Jiao Numerical Analysis I 5 / 27

Gaussian Elimination and LU Factorization

• Gaussian elimination without row interchanges can be viewed as "triangular triangularization" of nonsingular $A \in \mathbb{R}^{n \times n}$

$$\underbrace{L_{n-1}\cdots L_2L_1}_{I^{-1}}A=U$$

- Then A = LU. It is also called LU factorization (without pivoting)
- For augmented matrix, $L^{-1}[A \mid b] = [A \mid y]$, so $y = L^{-1}b$
- Example of LU factorization of 4×4 matrix A

$$\underbrace{L_1}_{L_1A} \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}}_{L_2A} \underbrace{L_2}_{L_2L_1A} \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & 0 & \mathbf{x} & \mathbf{x} \\ & & & & & & \\ \end{bmatrix}}_{L_2L_1A} \underbrace{L_3}_{L_3L_2L_1A}$$

Example of LU Factorization

Consider

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix}$$

• Then
$$L_1 = \begin{bmatrix} 1 \\ -2 & 1 \\ -4 & 1 \\ -3 & 1 \end{bmatrix}$$
 and $L_1A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 1 \\ 3 & 5 & 5 \\ 4 & 6 & 8 \end{bmatrix}$

Example of LU Factorization Cont'd

•
$$L_1A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 1 \\ 3 & 5 & 5 \\ 4 & 6 & 8 \end{bmatrix}$$

• $L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 \\ -3 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$ and $L_2L_1A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & 2 \\ 2 & 4 \end{bmatrix}$
• $L_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix}$ and $L_3L_2L_1A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & 2 & 0 & 2 \end{bmatrix}$

Example of LU Factorization Cont'd

•
$$L_1 = \begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ -4 & 1 & \\ -3 & & 1 \end{bmatrix}$$
, $L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & -3 & 1 & 0 \\ & -4 & 0 & 1 \end{bmatrix}$, $L_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 \\ & & 1 & 0 & 0 \\ & & & 1 & 0 \\ & & & & -1 & 1 \end{bmatrix}$ and $L_3L_2L_1A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ & 1 & 1 & 1 \\ & & 2 & 2 \\ & & 0 & 2 \end{bmatrix}$

• Then,

$$\begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 & 1 \\ 4 & 3 & 1 \\ 3 & 4 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ & 2 & 2 \\ & 0 & 2 \end{bmatrix}.$$

$$L = L_1^{-1} L_2^{-1} L_3^{-1} \qquad U$$

What is Matrix L_1 ?

• In Gaussian elimination, at first step A is transformed to $A^{(1)}$ by

$$\ell_{i1} = a_{i1}/a_{11}, \qquad i = 2, \dots, n$$

$$a_{ij}^{(1)} = a_{ij} - \ell_{i1}a_{1j}, \qquad i = 2, \dots, n, \quad j = 2, \dots, n$$

• In matrix form, we have $A^{(1)} = L_1 A$, where

$$L_1 = \left[egin{array}{cccc} 1 & & & & & \ -\ell_{2,1} & 1 & & & & \ -\ell_{3,1} & & 1 & & & \ dots & & \ddots & & \ -\ell_{n,1} & & & 1 \end{array}
ight]$$

What is Matrices L_k ?

- At step k, eliminate entries below $a_{kk}^{(k-1)}$:
- Let $a_{:,k}^{(k-1)}$ be kth column of $L_{k-1}\cdots L_1 A$, then

$$a_{:,k}^{(k-1)} = [a_{1k}^{(k-1)}, a_{2k}^{(k-1)}, \cdots, a_{kk}^{(k-1)}, a_{k,k+1}^{(k-1)}, \cdots a_{n,k}^{(k-1)}]^T$$

$$L_k a_{:,k}^{(k-1)} = [a_{1k}^{(k-1)}, a_{2k}^{(k-1)}, \cdots, a_{kk}^{(k-1)}, 0, \cdots 0]^T$$

• The multipliers $\ell_{jk} = a_{jk}^{(k-1)}/a_{kk}^{(k-1)}$ appear in L_k

Forming The *L* Matrix

- Let $\ell_k = [\underbrace{0,\cdots,0}_k,\ell_{k+1,k},\cdots,\ell_{n,k}]^T$ and $e_k = [\underbrace{0,\cdots,0}_{k-1},1,\cdots,0]^T$, then $L_k = I \ell_k e_k^T$
- Luckily, L matrix contains the multipliers $\ell_{jk} = x_{jk}/x_{kk}$

$$L = L_1^{-1} L_2^{-1} \cdots L_{n-1}^{-1} = \begin{bmatrix} 1 & & & & \\ \ell_{21} & 1 & & & \\ \ell_{31} & \ell_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{n1} & \ell_{n2} & \dots & \ell_{n,n-1} & 1 \end{bmatrix}$$

and is said to be a unit lower triangular matrix

Proof of Structure of L Matrix

• First, $L_k^{-1} = I + \ell_k e_k^T$, because $e_k^T \ell_k = 0$ and

$$(I - \ell_k e_k^T)(I + \ell_k e_k^T) = I - \ell_k e_k^T \ell_k e_k^T = I$$

• Second, $L_1^{-1}L_2^{-1}\cdots L_{k+1}^{-1}=I+\sum_{i=1}^{k+1}\ell_ie_i^T$, since (prove by induction)

$$(I + \sum_{j=1}^{k} \ell_{j} e_{j}^{T})(I + \ell_{k+1} e_{k+1}^{T}) = I + \sum_{j=1}^{k+1} \ell_{j} e_{j}^{T} + \sum_{j=1}^{k} \ell_{j} (e_{j}^{T} \ell_{k+1}) e_{k+1}^{T}$$

where $e_i^T \ell_{k+1} = 0$ for j < k+1

• In other words, L is "union" of nonzero entries in $L_1^{-1}, L_2^{-1}, \cdots, L_{n-1}^{-1}$

LU Factorization without Pivoting

• Factorize $A \in \mathbb{R}^{n \times n}$ into A = LU

Gaussian elimination without pivoting
$$U \leftarrow A, L \leftarrow I$$
 for $k = 1: n-1$ for $j = k+1: n$
$$\ell_{jk} \leftarrow u_{jk}/u_{kk}$$

$$u_{j,k:n} \leftarrow u_{j,k:n} - \ell_{jk}u_{k,k:n}$$

- Flop count $\sim \sum_{k=1}^{n} 2(n-k)(n-k) \sim 2 \sum_{k=1}^{n} k^2 \sim 2n^3/3$
- ullet In practice, L overwrites lower-triangular part of A and U overwrites upper-triangular part of A
- Question: What happens if u_{kk} is 0?
- Answer: The algorithm will break due to division by zero!

Theorem of LU Factorization

Theorem

Let A be an $n \times n$ matrix whose leading principal submatrices are all nonsingular. Then A can be factorized in exactly one way into a product A = LU, such that L is unit lower triangular and U is upper triangular.

- However, the requirement of nonsingular leading principal submatrices is too strong!
- More importantly, what happens if u_{kk} is **nearly** 0 (or a submatrix is **nearly** singular)?

Outline

① Gaussian Elimination(MC§3.2, NLA§20)

② Gaussian Elimination with Pivoting (NLA§21)

Xiangmin Jiao Numerical Analysis I 16 / 27

Gaussian Elimination with Partial Pivoting

- At step k, we divide by u_{kk} (i.e., $a_{kk}^{(k-1)}$), which would break if u_{kk} is 0 (or close to 0), which can happen even if A is nonsingular
- Other nonzero entry in kth column below diagonal can be pivot

and we permute (interchange) row i with row k

Partial Pivoting

- We choose nonzero entry with largest magnitude as pivot
- kth step of Gaussian elimination of partial pivoting

$$\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & \times_{ik} & \mathbf{x} & \mathbf{x} \\ & \times & \times & \times \end{bmatrix} \xrightarrow{P_k} \begin{bmatrix} \times & \times & \times & \times \\ & x_{kk} & \mathbf{x} & \mathbf{x} \\ & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ & \times & \times & \times \end{bmatrix} \xrightarrow{L_k} \begin{bmatrix} \times & \times & \times & \times \\ & x_{kk} & \times & \times \\ & 0 & \mathbf{x} & \mathbf{x} \\ & 0 & \mathbf{x} & \mathbf{x} \end{bmatrix}$$
Pivot selection Row interchange Elimination

and we interchange row i with row k

- ullet P_k is a permutation matrix, obtained by interchanging two rows of I
- Product of permutation matrices is also a permutation matrix
- For any permutation matrix P, all its entries are zeros and ones, and $PP^T = P^TP = I$

Xiangmin Jiao Numerical Analysis I 18 / 27

Example of LU with Partial Pivoting

• Consider
$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix}$$

$$\bullet \begin{bmatrix} & & 1 & \\ & 1 & & \\ 1 & & & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ 4 & 3 & 3 & 1 \\ 2 & 1 & 1 & 0 \\ 6 & 7 & 9 & 8 \end{bmatrix}$$

$$\bullet \ \, L_1 = \left[\begin{array}{ccc} 1 & & & \\ -\frac{1}{2} & 1 & & \\ -\frac{1}{4} & & 1 \\ -\frac{3}{4} & & 1 \end{array} \right] \text{ and } L_1 P_1 A = \left[\begin{array}{cccc} 8 & 7 & 9 & 5 \\ & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \\ & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} \\ & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \end{array} \right]$$

Example of LU with Partial Pivoting Cont'd

Example of LU with Partial Pivoting Cont'd

$$\bullet \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 \\ & & 1 \end{bmatrix} \begin{bmatrix} 8 & 7 & 9 & 5 \\ & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ & & -\frac{2}{7} & \frac{4}{7} \\ & & -\frac{6}{7} & -\frac{2}{7} \end{bmatrix} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ & -\frac{6}{7} & -\frac{2}{7} \\ & & -\frac{2}{7} & \frac{4}{7} \end{bmatrix}$$

$$\bullet L_{3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & -\frac{1}{3} & 1 \end{bmatrix} \text{ and } L_{3}P_{3}L_{2}P_{2}L_{1}P_{1}A = \begin{bmatrix} 8 & 7 & 9 & 5 \\ \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ & & -\frac{6}{7} & -\frac{2}{7} \\ & & \frac{2}{3} \end{bmatrix}$$

• Then,

$$\begin{bmatrix} & 1 & \\ & & 1 \\ 1 & \\ 1 & \\ P & A \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{3}{4} & 1 & 0 & 0 \\ \frac{1}{2} & -\frac{2}{7} & 1 & 0 \\ \frac{1}{4} & -\frac{3}{7} & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 8 & 7 & 9 & 5 \\ & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ & & -\frac{6}{7} & -\frac{2}{7} \\ & & & \frac{2}{3} \end{bmatrix}$$

Matrix Notation of Partial Pivoting

- In terms of matrices, it becomes $\underbrace{L_{n-1}P_{n-1}\cdots L_2P_2L_1P_1}_{L^{-1}P}A=U$
- PA = LU or $A = P^T LU$. This is LU factorization with partial pivoting
- $P = P_{n-1} \cdots P_2 P_1$ and $L = (L'_{n-1} \cdots L'_2 L'_1)^{-1}$, where $L'_k = P_{n-1} \cdots P_{k+1} L_k P_{k+1}^{-1} \cdots P_{n-1}^{-1}$
- It is easy to verify that $L_{n-1}P_{n-1}\cdots L_2P_2L_1P_1=\left(L'_{n-1}\cdots L'_2L'_1\right)\left(P_{n-1}\cdots P_2P_1\right)$
- $L'_k = I P_{n-1} \cdots P_{k+1} \ell_k e_k^T$, and $(L'_k)^{-1} \equiv I + P_{n-1} \cdots P_{k+1} \ell_k e_k^T$.
- In other words, L_k' is obtained by continued permutation of $\ell_k e_k^T$ along with $A^{(k+)}$, and lower-triangular part of L is "union" of permuted $\ell_k e_k^T$

Algorithm of Gaussian Elimination with Partial Pivoting

• Factorize $A \in \mathbb{R}^{n \times n}$ into $A = P^T L U$

Gaussian elimination with partial pivoting

$$\begin{array}{l} \textit{U} \leftarrow \textit{A}, \textit{L} \leftarrow \textit{I}, \; \textit{P} \leftarrow \textit{I} \\ \textit{for} \; \textit{k} = 1 : \textit{n} - 1 \\ \quad \textit{i} \leftarrow \arg\max_{i \geq k} |\textit{u}_{ik}| \\ \quad \textit{u}_{k,k:n} \leftrightarrow \textit{u}_{i,k:n} \\ \quad \ell_{k,1:k-1} \leftrightarrow \ell_{i,1:k-1} \\ \quad \textit{p}_{k,:} \leftrightarrow \textit{p}_{i,:} \\ \textit{for} \; \textit{j} = \textit{k} + 1 : \textit{n} \\ \quad \ell_{jk} \leftarrow \textit{u}_{jk}/\textit{u}_{kk} \\ \quad \textit{u}_{j,k:n} \leftarrow \textit{u}_{j,k:n} - \ell_{jk}\textit{u}_{k,k:n} \end{array}$$

- Flop count $\sim \sum_{k=1}^{n} 2(n-k)(n-k) \sim 2 \sum_{k=1}^{n} k^2 \sim 2n^3/3$, same as without pivoting
- Question: Can u_{kk} be 0?

An Alternative Implementation

ullet In practice, L and U overwrite A, and P is represented by a vector

Gaussian elimination with partial pivoting (alternative) $p \leftarrow [1,2,\cdots,n]$ for k=1:n-1 $i \leftarrow \arg\max_{i\geq k}|a_{ik}|$ $a_{k,1:n} \leftrightarrow a_{i,1:n}$ $p_k \leftrightarrow p_i$ $a_{k+1:n,k} \leftarrow a_{k+1:n,k}/a_{k,k}$ $A_{k+1:n,k+1:n} \leftarrow A_{k+1:n,k+1:n} - a_{k+1:n,k} * a_{k,k+1:n}$

- Using LU factorization with partial pivoting to solve Ax = b:
 - \bullet $A = P^T LU$; (LU factorization with partial pivoting)
 - 2 Ly = Pb; (Forward substitution, where $(Pb)_i = b(p_i)$)
 - Ux = y; (Back substitution)
- If the augmented matrix is used, then first two steps are merged.

Xiangmin Jiao Numerical Analysis I 24 / 27

Complete Pivoting

• More generally, pivot can be chosen from entries (i,j), $i \geq k$, $j \geq k$

$$\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & \mathbf{x} & \mathbf{x}_{ij} & \mathbf{x} \\ & \times & \times & \times \end{bmatrix} \rightarrow \begin{bmatrix} \times & \times & \times & \times \\ & \mathbf{x} & \mathbf{0} & \mathbf{x} \\ & \times & \mathbf{x}_{ij} & \times \\ & \mathbf{x} & \mathbf{0} & \mathbf{x} \end{bmatrix}$$

and we then permute row i with row k, column j with column k

• In matrix operations, complete can be expressed as

$$\underbrace{L_{n-1}P_{n-1}\cdots L_2P_2L_1P_1}_{L^{-1}P}A\underbrace{Q_1Q_2\cdots Q_{n-1}}_{Q}=U$$

• Therefore, PAQ = LU where $P = P_{n-1} \cdots P_2 P_1$ and $L = (L'_{n-1} \cdots L'_2 L'_1)^{-1}$

Complete Pivoting

- By choosing largest absolute value among these entries, complete pivoting gives better stability theoretically (to be discussed later)
- Complete pivoting is typically not used in practice, because it increases cost in searching pivot and complexity of implementation
 - ► complete pivoting incurs $O(n^3)$ comparisons, whereas partial pivoting incurs $O(n^2)$
- Complete pivoting can be useful in solving underdetermined systems
 - System is underdetermined if A is singular, and it has an infinite number of solutions
 - One can compute

$$PAQ^T = L[U_1|U_2],$$

where U_1 is nonsingular and upper-triangular, and U_2 has zeros along diagonal

▶ It allows finding a solution, but not necessarily an optimal solution

Alternative Linear Solvers

- Gauss-Jordan elimination: At the kth step, use pivot to eliminate nonzeros in column k both above and below diagonal. It is more expensive than Gaussian elimination. How much more?
- Cramer's rule $x_i = \det(A^{(i)})/\det(A)$, where $A^{(i)}$ is obtained from A by replacing ith column by b. This is overly expensive and unstable, so it is never used except for n = 2 or 3
- Strassen's method (and similar recursive algorithms) for matrix-matrix multiplication can be modified to solve linear systems in $O(n^s)$ flops for s < 3, but are not advantageous in practice

These are direct methods. Other important classes of methods include

- iterative methods (later) and
- multigrid methods