Existence of LU Decomposition

Theorem

A matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ *has an LU decomposition if every (leading) principal submatrix* $\mathbf{A}(1:k, 1:k)$ *satisfies*

$$\det(\mathbf{A}(1:k,1:k)) \neq 0 \implies \text{all the pivots } a_{kk}^{(k-1)} \neq 0$$

for $k = 1, 2, \ldots, n-1$. *If the LU decomposition of* $\mathbf{A}$ *exists and* $\mathbf{A}$ *is nonsingular, then the LU decomposition is unique and* $\det(\mathbf{A}) = u_{11} \cdots u_{nn}$.

Proof

We first show that if $\det(A(1:k,1:k)) \neq 0 \ \forall k=1,\cdots,n-1$, then $a_{kk}^{(k-1)} \neq 0$, so that LU decomposition exists.

Let $k=1,\cdots,n-1$.

$$A^{(k-1)} = \underbrace{M_{k-1}\cdots M_1}_{W} A \qquad (*)$$

Since each M_i is unit lower triangular, so is W. (Fact 1)

We partition $(*)$ as follows:

$$\left[\begin{array}{c|c} A^{(k-1)}(1:k,1:k) & * \\ \hline * & * \end{array}\right] = \left[\begin{array}{c|c} W(1:k,1:k) & 0 \\ \hline * & * \end{array}\right] \cdot \left[\begin{array}{c|c} A(1:k,1:k) & * \\ \hline * & * \end{array}\right]$$

Proof (cont'd)

This gives $A^{(k-1)}(1:k,1:k) = W(1:k,1:k)\,A(1:k,1:k)$ (#)

Note that $A^{(k-1)}(1:k,1:k)$ is upper triangular.

$$\prod_{i=1}^{k} a_{ii}^{(k-1)} = \det\left(A^{(k-1)}(1:k,1:k)\right) \overset{(\#)}{=} \underbrace{\det(W(1:k,1:k))}_{=1} \cdot \underbrace{\det(A(1:k,1:k))}_{\neq 0}$$

It follows that $a_{kk}^{(k-1)} \neq 0$.

$\Rightarrow$ LU decomposition exists.

Proof (cont'd)

Let A be nonsingular, i.e., $\det(A) \neq 0$.

Assume to the contrary that A has two LU decompositions L_1U_1 and L_2U_2. Note that L_1, U_1, L_2, U_2 nonsingular.

Then,
$$L_2^{-1} L_1 \underbrace{U_1 U_1^{-1}} = \underbrace{L_2^{-1} L_2} U_2 U_1^{-1}$$

$$\underbrace{L_2^{-1} L_1}_{\text{unit lower triangular}} = \underbrace{U_2 U_1^{-1}}_{\text{upper triangular}}$$

Facts 1 & 3

Therefore, the above equation only holds for
$L_2^{-1} L_1 = U_2 U_1^{-1} = I \Rightarrow L_1 = L_2,\ U_1 = U_2.$

Finally, $\det(A) = \underbrace{\det(L)}_{=1} \det(U) = \det(U) = \prod_{i=1}^{n} u_{ii}.$

Matrix Computations
Chapter 2 Linear systems and LU decomposition
Section 2.2 Pivoting for LU Decomposition

Jie Lu
ShanghaiTech University

Pivoting

- Previously, we assume all the pivots are nonzero. What if some $a_{kk}^{(k-1)}$ happens to be zero?

- Gaussian elimination is known to be numerically unstable when a pivot is close to zero
  - Relatively small pivots can cause large entries in $\mathbf{L}$ and $\mathbf{U}$ and thus non-negligible error in solution due to round-off errors

- Pivoting: Find permutations of A with a proper LU decomposition
  - Partial pivoting, complete pivoting, rook pivoting, etc.

Permutation Matrix

A square matrix with exactly one entry of 1 in each row and each column and 0 elsewhere is a permutation matrix

Example: Let $\mathbf{\Pi}$ be a 4×4 permutation matrix and $A \in \mathbb{R}^4$

$$\mathbf{\Pi} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \end{bmatrix} = \begin{bmatrix} — \tilde{\mathbf{a}}_1^T — \\ — \tilde{\mathbf{a}}_2^T — \\ — \tilde{\mathbf{a}}_3^T — \\ — \tilde{\mathbf{a}}_4^T — \end{bmatrix}$$

- $\mathbf{\Pi A}$ is obtained by swapping row 1 and row 4 of $\mathbf{A}$
- $\mathbf{A\Pi}$ is obtained by swapping column 1 and column 4 of $\mathbf{A}$

$$\mathbf{\Pi A} = \begin{bmatrix} — \tilde{\mathbf{a}}_4^T — \\ — \tilde{\mathbf{a}}_2^T — \\ — \tilde{\mathbf{a}}_3^T — \\ — \tilde{\mathbf{a}}_1^T — \end{bmatrix} \qquad \mathbf{A\Pi} = \begin{bmatrix} \mathbf{a}_4 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_1 \end{bmatrix}$$

Permutation Matrix

Example:

$\times\ [a_2 \ a_4 \ a_1 \ a_3]$

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \qquad \mathbf{PA} = \begin{bmatrix} -\tilde{\mathbf{a}}_2^T- \\ -\tilde{\mathbf{a}}_4^T- \\ -\tilde{\mathbf{a}}_1^T- \\ -\tilde{\mathbf{a}}_3^T- \end{bmatrix}, \qquad \mathbf{AP} = \begin{bmatrix} \mathbf{a}_3 & \mathbf{a}_1 & \mathbf{a}_4 & \mathbf{a}_2 \end{bmatrix}$$

Note that $\mathbf{P}$ can be decomposed as

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \underset{P_3}{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}} \cdot \underset{P_2}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}} \cdot \underset{P_1}{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}$$

$$P_1A = \begin{bmatrix} \tilde{a}_3^T \\ \tilde{a}_2^T \\ \tilde{a}_1^T \\ \tilde{a}_4^T \end{bmatrix} \longrightarrow P_2(P_1A) = \begin{bmatrix} \tilde{a}_3^T \\ \tilde{a}_4^T \\ \tilde{a}_1^T \\ \tilde{a}_2^T \end{bmatrix} \longrightarrow P_3(P_2P_1A) = \cdots$$

Interchange Permutations

Let $\mathbf{\Pi}_k \in \mathbb{R}^{n\times n}$, $k = 1, \ldots, m \leq n$ be the $n \times n$ identity matrix $\mathbf{I}$ with row k and row $piv(k)$ swapped, which are called interchange permutations

$\in \{1, \cdots, n\}$

Let $\mathbf{P} = \mathbf{\Pi}_m \cdots \mathbf{\Pi}_1$

Π_k nonsingular

- $\mathbf{\Pi}_k$ is symmetric (but $\mathbf{P}$ may not be symmetric)
- $\mathbf{P}^T = \mathbf{\Pi}_1 \cdots \mathbf{\Pi}_m$
- If $piv = [1, \ldots, m]^T$, then $\mathbf{P} = \mathbf{I}$

$piv(k) = k$

$\Pi_k^{-1} = \Pi_k$

Computation of $\mathbf{Px}$, $\mathbf{x} \in \mathbb{R}^n$

```
for k=1:m  % overwrite x with Px
     x(k)↔ x(piv(k)) % swap entry k and entry piv(k)
end
```

Computation of $\mathbf{P}^T\mathbf{x}$, $\mathbf{x} \in \mathbb{R}^n$

```
for k=m:-1:1
     x(k)↔ x(piv(k))
end
```

No flops needed for permutation (but affect performance nontrivially)

Partial Pivoting

Recall Upper Triangularization in Section 2.1
Given $\mathbf{A}^{(k-1)}$, $k = 1, \ldots, n-1$,

1. Find $piv(k) = arg\ \max_{j \in [k,n]} |\mathbf{A}^{(k-1)}(j,k)|$
2. Let $\mathbf{\Pi}_k \in \mathbb{R}^{n \times n}$ be the interchange permutation that swaps row k and row $piv(k)$ of $\mathbf{I}$
3. Determine the Gauss Transformation $\mathbf{M}_k = \mathbf{I}_n - \boldsymbol{\tau}^{(k)} \mathbf{e}_k^T$, where

$$\boldsymbol{\tau}^{(k)} = \begin{bmatrix} \mathbf{0}_k \\ (\mathbf{\Pi}_k \mathbf{A}^{(k-1)})(k+1:n,k) / (\mathbf{\Pi}_k \mathbf{A}^{(k-1)})(k,k) \end{bmatrix}$$

4. $\mathbf{A}^{(k)} = \mathbf{M}_k (\mathbf{\Pi}_k \mathbf{A}^{(k-1)})$ (which satisfies $\mathbf{A}^{(k)}(k+1:n,k) = \mathbf{0}$)

Upon completing the above process, we have

$$\mathbf{M}_{n-1} \mathbf{\Pi}_{n-1} \cdots \mathbf{M}_1 \mathbf{\Pi}_1 \mathbf{A} = \mathbf{U}$$

Note that all the elements in $\boldsymbol{\tau}^{(k)}(k+1:n)$ are ≤ 1 in absolute value

Partial Pivoting (cont'd)

Example: $\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = A^{(0)}$

$k=1$: $\quad piv(1) = 3 \qquad \Pi_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \Pi_1 A^{(0)} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ 4 & 3 & 3 & 1 \\ 2 & 1 & 1 & 0 \\ 6 & 7 & 9 & 8 \end{bmatrix}$

$$M_1 = \begin{bmatrix} 1 & & & \\ -\frac{1}{2} & 1 & & \\ -\frac{1}{4} & & 1 & \\ -\frac{3}{4} & & & 1 \end{bmatrix}, \quad A^{(1)} = M_1 \Pi_1 A^{(0)} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ 0 & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \\ 0 & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \end{bmatrix}$$

Partial Pivoting (cont'd)

$k=2$: $\quad piv(2) = 4. \qquad \Pi_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

$$\Pi_2 A^{(1)} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} \\ 0 & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \end{bmatrix} \qquad M_2 = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & \frac{3}{7} & 1 & \\ & \frac{2}{7} & & 1 \end{bmatrix}$$

$$A^{(2)} = M_2 \Pi_2 A^{(1)} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & 0 & -\frac{2}{7} & \frac{4}{7} \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} \end{bmatrix}$$

Partial Pivoting (cont'd)

$k=3$: $\quad piv(3) = 4 \qquad \Pi_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

$$\Pi_3 A^{(2)} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} \\ 0 & 0 & -\frac{2}{7} & \frac{4}{7} \end{bmatrix} \qquad M_3 = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & -\frac{1}{3} & 1 \end{bmatrix}$$

$$A^{(3)} = M_3 \Pi_3 A^{(2)} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & \frac{2}{3} \end{bmatrix}$$

Computation of $\mathbf{L}$ with Partial Pivoting

Define $\mathbf{P} = \mathbf{\Pi}_{n-1} \cdots \mathbf{\Pi}_1$ and for each $k = 1, \ldots, n-1$,

When $k = n-1$

$$\tilde{\mathbf{M}}_k = (\mathbf{\Pi}_{n-1} \cdots \mathbf{\Pi}_{k+1}) \mathbf{M}_k (\mathbf{\Pi}_{k+1} \cdots \mathbf{\Pi}_{n-1})$$

$\tilde{M}_k = M_k$

Note: $\tilde{\mathbf{M}}_k$ **is a Gauss transformation**

$$\tilde{\mathbf{M}}_k = (\mathbf{\Pi}_{n-1} \cdots \mathbf{\Pi}_{k+1}) \cdot (\mathbf{I} - \boldsymbol{\tau}^{(k)} \mathbf{e}_k^T) \cdot (\mathbf{\Pi}_{k+1} \cdots \mathbf{\Pi}_{n-1}) = \mathbf{I} - \tilde{\boldsymbol{\tau}}^{(k)} \mathbf{e}_k^T$$

with $\tilde{\boldsymbol{\tau}}^{(k)} = \mathbf{\Pi}_{n-1} \cdots \mathbf{\Pi}_{k+1} \boldsymbol{\tau}^{(k)}$ (Why?)

$$\tau^{(k)} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \times \\ \# \\ \square \end{bmatrix} \Big\} k \qquad \tilde{\tau}^{(k)} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \square \\ \# \\ \times \end{bmatrix}$$

$$\tilde{M}_k = \overbrace{\Pi_{n-1} \cdots \overbrace{\Pi_{k+1} \underbrace{\cancel{I}\, \Pi_{k+1}}_{I}}^{I} \cdots \Pi_{n-1}}^{I}$$

$$- \underbrace{\Pi_{n-1} \cdots \Pi_{k+1} \tau^{(k)}}_{\tilde{\tau}^{(k)}} \underbrace{e_k^T \boxed{\Pi_{k+1} \cdots \Pi_{n-1}}}_{=\, e_k^T}$$

interchange permutation only for $k+1:\cdots:n$ ~~columns~~ rows

$$e_k^T = [0 \cdots 0 \; 1 \; 0 \cdots 0]$$

kth

Computation of $\mathbf{L}$ with Partial Pivoting (cont'd)

Example: Let $n = 4$

$$\tilde{\mathbf{M}}_3\tilde{\mathbf{M}}_2\tilde{\mathbf{M}}_1\mathbf{PA} = \mathbf{M}_3 \cdot (\mathbf{\Pi}_3\mathbf{M}_2\mathbf{\Pi}_3) \cdot (\mathbf{\Pi}_3\mathbf{\Pi}_2\mathbf{M}_1\mathbf{\Pi}_2\mathbf{\Pi}_3) \cdot (\mathbf{\Pi}_3\mathbf{\Pi}_2\mathbf{\Pi}_1)\mathbf{A}$$
$$= \mathbf{M}_3\mathbf{\Pi}_3\mathbf{M}_2\Pi_2\mathbf{M}_1\Pi_1\mathbf{A} = \mathbf{U}$$

We can easily extend this to general n and obtain

$$\tilde{\mathbf{M}}_{n-1}\cdots\tilde{\mathbf{M}}_1\mathbf{PA} = \mathbf{U}$$

In addition, let

$$\mathbf{L} = \tilde{\mathbf{M}}_1^{-1}\cdots\tilde{\mathbf{M}}_{n-1}^{-1} = (\mathbf{I} + \tilde{\boldsymbol{\tau}}^{(1)}\mathbf{e}_1^T)\cdots(\mathbf{I} + \tilde{\boldsymbol{\tau}}^{(n-1)}\mathbf{e}_{n-1}^T) = \mathbf{I} + \sum_{k=1}^{n-1}\tilde{\boldsymbol{\tau}}^{(k)}\mathbf{e}_k^T$$

absolute value

where each entry of $\mathbf{L}$ is ≤ 1 (Why?)

Therefore, LU decomposition with pivoting is equivalent to

$$\mathbf{PA} = \mathbf{LU}$$