SI231b: Matrix Computations

Lecture 7: LU Factorization with Pivoting

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Sept. 28, 2022

$$\begin{bmatrix} a_{11}^{(0)} & \times & \times & \cdots & \times & \times \\ 0 & a_{22}^{(1)} & \times & \cdots & \times & \times \\ 0 & 0 & \times & \cdots & \times & \times \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \textcircled{0} & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \textcircled{S} & \times & \times \end{bmatrix} \xrightarrow{\text{pivoting}} \begin{bmatrix} a_{11}^{(0)} & \times & \times & \cdots & \times & \times \\ 0 & a_{22}^{(1)} & \times & \cdots & \times & \times \\ 0 & 0 & \times & \cdots & \times & \times \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \textcircled{S} & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \textcircled{0} & \times & \times \end{bmatrix}$$

- partial pivoting
  - finding $p = \arg\max_{k \le i \le n} \left| a_{ik}^{(k-1)} \right|$
  - let $a_{kk}^{(k-1)} = a_{pk}^{(k-1)}$ (row exchange)
- complete pivoting
  - finding $[p_r, p_c] = \arg\max_{k \le i,\ j \le n} \left| a_{ij}^{(k-1)} \right|$
  - let $a_{kk}^{(k-1)} = a_{p_r p_c}^{(k-1)}$ (row and column exchange)

Step k of LU factorization

1. row exchange: $\tilde{\mathbf{A}}^{(k-1)} = \mathbf{P}_k \mathbf{A}^{(k-1)}$
2. Gaussian elimination: $\mathbf{A}^{(k)} = \mathbf{M}_k \tilde{\mathbf{A}}^{(k-1)}$

In general, the procedure follows

$$\mathbf{M}_{n-1}\mathbf{P}_{n-1}\mathbf{M}_{n-2}\mathbf{P}_{n-2}\cdots\mathbf{M}_1\mathbf{P}_1\mathbf{A} = \mathbf{U}.$$

Denote

$$\begin{aligned}
\tilde{\mathbf{M}}_{n-1} &= \mathbf{M}_{n-1}, \\
\tilde{\mathbf{M}}_{n-2} &= \mathbf{P}_{n-1}\mathbf{M}_{n-2}\mathbf{P}_{n-1}^T, \\
\vdots &= \qquad \vdots \\
\tilde{\mathbf{M}}_k &= \mathbf{P}_{n-1}\mathbf{P}_{n-2}\cdots\mathbf{P}_{k+1}\mathbf{M}_k\mathbf{P}_{k+1}^T\cdots\mathbf{P}_{n-2}^T\mathbf{P}_{n-1}^T
\end{aligned}$$

Note: $\tilde{\mathbf{M}}_k$ has the same structure with $\mathbf{M}_k$ (recall the structure of $\mathbf{M}_k$)

LU Factorization with Partial Pivoting

Following the aforementioned procedure,

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

where

- $\mathbf{P} = \mathbf{P}_{n-1}\mathbf{P}_{n-2}\cdots\mathbf{P}_1$ is again a permutation matrix (why?)
- $\mathbf{L} = \left(\tilde{\mathbf{M}}_{n-1}\tilde{\mathbf{M}}_{n-2}\cdots\tilde{\mathbf{M}}_1\right)^{-1}$ is a lower-triangular matrix with unit diagonals
- sometimes called LUP factorization
- always exists for any square $\mathbf{A}$, no matter $\mathbf{A}$ is nonsingular or not[1]

Another Interpretation

1. permute the rows of $\mathbf{A}$ according to $\mathbf{P}$
2. compute the LU factorization without pivoting to $\mathbf{PA}$

Note: LU factorization with partial pivoting is not carried out in this way, since $\mathbf{P}$ is unknown in advance.

[1] https://arxiv.org/abs/math/0506382

A Simple Example

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

Step 1, 1st row $\longleftrightarrow$ 3rd row of $\mathbf{A}$, then perform Gaussian elimination

$$\tilde{\mathbf{A}}^{(0)} = \mathbf{P}_1\mathbf{A} = \begin{bmatrix} & & 1 & \\ & 1 & & \\ 1 & & & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ 4 & 3 & 3 & 1 \\ 2 & 1 & 1 & 0 \\ 6 & 7 & 9 & 8 \end{bmatrix}$$

$$\mathbf{A}^{(1)} = \mathbf{M}_1\tilde{\mathbf{A}}^{(0)} = \begin{bmatrix} 1 & & & \\ -\frac{1}{2} & 1 & & \\ -\frac{1}{4} & & 1 & \\ -\frac{3}{4} & & & 1 \end{bmatrix} \begin{bmatrix} 8 & 7 & 9 & 5 \\ 4 & 3 & 3 & 1 \\ 2 & 1 & 1 & 0 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \\ & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} \\ & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \end{bmatrix}$$

A Simple Example

Step 2: 2nd row $\longleftrightarrow$ 4th row of $\mathbf{A}^{(1)}$, then repeat Gaussian elimination

$$\tilde{\mathbf{A}}^{(1)} = \mathbf{P}_2\mathbf{A}^{(1)} = \begin{bmatrix} 1 & & & \\ & & & 1 \\ & & 1 & \\ & 1 & & \end{bmatrix} \begin{bmatrix} 8 & 7 & 9 & 5 \\ & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \\ & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} \\ & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \end{bmatrix} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} \\ & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \end{bmatrix}$$

$$\mathbf{A}^{(2)} = \mathbf{M}_2\tilde{\mathbf{A}}^{(1)} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & \frac{3}{7} & 1 & \\ & \frac{2}{7} & & 1 \end{bmatrix} \begin{bmatrix} 8 & 7 & 9 & 5 \\ & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} \\ & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \end{bmatrix} = \begin{bmatrix} 8 & 7 & 9 & 5 \\ & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ & & -\frac{2}{7} & \frac{4}{7} \\ & & -\frac{6}{7} & -\frac{2}{7} \end{bmatrix}$$

Now, it's your turn to give $\mathbf{P}_3$, $\mathbf{M}_3$ and the final $\mathbf{P}$, $\mathbf{L}$, and $\mathbf{U}$

A Simple Example

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

In practice, the permutation matrix **P**

- is not represented explicitly as a matrix or the product of permutation matrices
- an equivalent effect can be achieved via a permutation vector

Note: $|\ell_{ij}| \leq 1$ for $i \geq j$

Following the aforementioned procedure,

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

where

- $\mathbf{P} = \mathbf{P}_{n-1}\mathbf{P}_{n-2}\cdots\mathbf{P}_1$ is again a permutation matrix (why?)
- $\mathbf{L} = \left(\tilde{\mathbf{M}}_{n-1}\tilde{\mathbf{M}}_{n-2}\cdots\tilde{\mathbf{M}}_1\right)^{-1}$ is a lower-triangular matrix with unit diagonals

$$\tilde{\mathbf{M}}_k = \tilde{\mathbf{P}}_{k+1}\mathbf{M}_k\tilde{\mathbf{P}}_{k+1}^T = \mathbf{I} - \tilde{\mathbf{P}}_{k+1}\tau^{(k)}\mathbf{e}_k^T \qquad \textit{why?}$$

Here $\tilde{\mathbf{P}}_{k+1} = \mathbf{P}_{n-1}\mathbf{P}_{n-2}\cdots\mathbf{P}_{k+1}$

Then, we obtain

$$\tilde{\mathbf{M}}_k^{-1} = \mathbf{I} + \tilde{\mathbf{P}}_{k+1}\tau^{(k)}\mathbf{e}_k^T$$

$$\mathbf{L} = \mathbf{I} + \sum_{i=1}^{n-1}\tilde{\mathbf{P}}_{k+1}\tau^{(k)}\mathbf{e}_k^T$$

For $\mathbf{A} \in \mathbb{R}^{n \times n}$, and a permutation matrix $\mathbf{P}_1$

$$\mathbf{P}_1\mathbf{A} = \left[\begin{array}{c|c} a_{11}^{(0)} & \mathbf{v}^T \\ \hline \mathbf{u} & \mathbf{A}_1' \end{array}\right] = \underbrace{\left[\begin{array}{c|c} 1 & 0 \\ \hline 1/a_{11}^{(0)}\mathbf{u} & \mathbf{I}_{n-1} \end{array}\right]}_{\mathbf{L}_1} \underbrace{\left[\begin{array}{c|c} a_{11}^{(0)} & \mathbf{v}^T \\ \hline 0 & \mathbf{A}_1' - 1/a_{11}^{(0)}\mathbf{u}\mathbf{v}^T \end{array}\right]}_{\mathbf{U}_1}$$

Then repeat the above procedure to $\mathbf{A}_1' - 1/a_{11}^{(0)}\mathbf{u}\mathbf{v}^T$, i.e.,

$$\begin{aligned} \mathbf{P}_2'\left(\mathbf{A}_1' - 1/a_{11}^{(0)}\mathbf{u}\mathbf{v}^T\right) &= \left[\begin{array}{c|c} a_{22}^{(1)} & \mathbf{w}^T \\ \hline \mathbf{s} & \mathbf{A}_2' \end{array}\right] \\ &= \left[\begin{array}{c|c} 1 & 0 \\ \hline 1/a_{22}^{(1)}\mathbf{s} & \mathbf{I}_{n-2} \end{array}\right] \left[\begin{array}{c|c} a_{22}^{(1)} & \mathbf{w}^T \\ \hline 0 & \mathbf{A}_2' - 1/a_{22}^{(1)}\mathbf{s}\mathbf{w}^T \end{array}\right] \end{aligned}$$

Denote $\mathbf{P}_2 = \begin{bmatrix} 1 & \\ & \mathbf{P}_2' \end{bmatrix}$, we obtain (next page)

An Alternative Approach for LU Factorization with Partial Pivoting

$$\mathbf{P}_2\mathbf{P}_1\mathbf{A} = \underbrace{\begin{bmatrix} 1 & & \\ & 1 & \\ \frac{1}{a_{11}^{(0)}}\mathbf{P}_2'\mathbf{u} & \frac{1}{a_{22}^{(1)}}\mathbf{s} & \mathbf{I}_{n-2} \end{bmatrix}}_{\mathbf{L}_2} \underbrace{\begin{bmatrix} a_{11}^{(0)} & & \mathbf{v}^T \\ & a_{22}^{(1)} & \mathbf{w}^T \\ & & \mathbf{A}_2' - \frac{1}{a_{22}^{(1)}}\mathbf{s}\mathbf{w}^T \end{bmatrix}}_{\mathbf{U}_2}$$

- following the above notations, $\mathbf{L} = \mathbf{L}_{n-1}$, $\mathbf{U} = \mathbf{U}_{n-1}$
- $\mathbf{P}_k$ only acts on the first $(k-1)$ columns of $\mathbf{L}_k$
- algorithm style, suitable for computer implementation

Remark:

- Gaussian elimination tells why you can perform an LU factorization, and when does it exist
- the recursive approach tells how you can compute the LU factorization on a modern computer

Example

Please compute an LU factorization with partial pivoting using the method introduced in the last page for

$$\begin{bmatrix} 4 & 2 & 3 \\ -3 & 1 & 4 \\ 2 & 4 & 5 \end{bmatrix}$$

LU with complete pivoting:

In matrix form, at each stage before Gaussian elimination

- permutation of rows with $\mathbf{P}_k$ on the left
- permutation of columns with $\mathbf{Q}_k$ on the right

$$\mathbf{M}_{n-1}\mathbf{P}_{n-1}\mathbf{M}_{n-2}\mathbf{P}_{n-2}\cdots\mathbf{M}_1\mathbf{P}_1\mathbf{A}\mathbf{Q}_1\mathbf{Q}_2\cdots\mathbf{Q}_{n-1} = \mathbf{U}.$$

By

- using the same definition of $\mathbf{L}$, $\mathbf{P}$ with LU factorization with partial pivoting,
- denoting $\mathbf{Q} = \mathbf{Q}_1\mathbf{Q}_2\cdots\mathbf{Q}_{n-1}$,

the LU factorization with complete pivoting can be represented by

$$\mathbf{PAQ} = \mathbf{LU}$$

Too computationally expensive, why?

Computational Complexity of LU Factorization

LU Factorization without Pivoting:

```
U = A, L = I;
for k = 1 : n-1
    for j = k+1 : n
        ℓ_jk = u_jk / u_kk
        u_{j,k:n} = u_{j,k:n} − ℓ_jk u_{k,k:n}
    end
end
U = triu(U)
```

Operations count:

- $\mathcal{O}\left(\frac{2}{3}n^3\right)$ flops

Please give your own explanation

Computational Complexity of LU Factorization with Partial Pivoting

LU Factorization with Partial Pivoting:

```
U = A, L = I, P = I;
for k = 1 : n-1
     select i ≥ k to maximize |u_ik|
     u_{k,k:m} ↔ u_{i,k:m}  (exchange of rows)
     ℓ_{k,1:k-1} ↔ ℓ_{i,1:k-1}
     p_{k,:} ↔ p_{i,:}
     for j = k+1 : n
          ℓ_jk = u_jk / u_kk
          u_{j,k:n} = u_{j,k:n} - ℓ_jk u_{k,k:n}
     end
end
U = triu(U)
```

Operations count:

- $\mathcal{O}\left(\frac{2}{3}n^3\right)$ flops, flops count of partial pivoting?