SI231b: Matrix Computations

Lecture 6: Solving Linear Equations (Direct Methods)

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Sept. 26, 2022

- Example of LU Factorization
- LU Factorization with Pivoting
- Implementation on Computers
- Computational Complexity of LU Factorization
- General Procedure of Direct Methods

An Example of LU Factorization

Given

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{bmatrix},$$

the LU factorization is given by

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}, \qquad \mathbf{U} = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{bmatrix}$$

Please give **L** and **U** by yourself

LU Factorization with Pivoting

Step k of LU factorization

$$\begin{bmatrix} a_{11}^{(0)} & \times & \times & \cdots & \times & \times \\ 0 & a_{22}^{(1)} & \times & \cdots & \times & \times \\ 0 & 0 & \times & \cdots & \times & \times \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & a_{kk}^{(k-1)} & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} a_{11}^{(0)} & \times & \times & \cdots & \times & \times \\ 0 & a_{22}^{(1)} & \times & \cdots & \times & \times \\ 0 & 0 & \times & \cdots & \times & \times \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & a_{kk}^{(k-1)} & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}$$

- $\mathbf{A}^{(k)} = \mathbf{M}_k \mathbf{A}^{(k-1)}$
- Require: $a_{kk}^{(k-1)} \neq 0$
  - under which condition?
  - if unsatisfied, what to do?

LU Factorization with Pivoting

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

- 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)

Permutation Matrix

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

Example

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \mathbf{Px} = \begin{bmatrix} x_3 \\ x_2 \\ x_1 \end{bmatrix}, \qquad \mathbf{x}^T\mathbf{P} = \begin{bmatrix} x_3 & x_2 & x_1 \end{bmatrix}$$

PA: exchange rows of $\mathbf{A}$

AP: exchange columns of $\mathbf{A}$

Properties:

- $\mathbf{P}$ is an orthogonal matrix, i.e., $\mathbf{P}^T\mathbf{P} = \mathbf{P}\mathbf{P}^T = \mathbf{I}$.
- $\mathbf{P}^{-1} = \mathbf{P}^T$

LU Factorization with Partial Pivoting

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

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

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

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

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

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 $i \geq j$

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?

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
```

$$\ell_{jk} = u_{jk}/u_{kk}$$

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

Operations count:

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

Please give your own explanation

LU Factorization with Partial Pivoting:

U = A, L = I, P = I;
for k = 1 : n-1
    select $i \geq k$ to maximize $|u_{ik}|$
    $u_{k,k:m} \leftrightarrow u_{i,k:m}$ (exchange of rows)
    $\ell_{k,1:k-1} \leftrightarrow \ell_{i,1:k-1}$
    $p_{k,:} \leftrightarrow p_{i,:}$
    for j = k+1 : n
        $\ell_{jk} = u_{jk} / u_{kk}$
        $u_{j,k:n} = u_{j,k:n} - \ell_{jk} u_{k,k:n}$
    end
end

Operations count:

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

General Procedure of Direct Methods

1. compute the LU factorization with partial pivoting, $\mathbf{PA} = \mathbf{LU}$, $\mathcal{O}(\frac{2}{3}n^3)$ flops
2. solve $\mathbf{Lz} = \mathbf{Pb}$ using forward substitution, $\mathcal{O}(n^2)$ flops
3. solve $\mathbf{Ux} = \mathbf{z}$ using backward substitution, $\mathcal{O}(n^2)$ flops

Variant of LU Factorization: LDU Factorization

For LU factorization with partial pivoting $\mathbf{PA} = \mathbf{LU}$

- denote $\mathbf{D} = \mathrm{diag}(u_{11}, u_{22}, \cdots, u_{nn})$
- $\bar{\mathbf{U}} = \mathbf{D}^{-1}\mathbf{U}$: upper-triangular matrix with unit diagonal entries, i.e., $\bar{u}_{ij} = u_{ij}/u_{ii}$ for $i \le j$

Then $\mathbf{PA} = \mathbf{LD\bar{U}}$ gives an LDU factorization of $\mathbf{A}$

Readings

You are supposed to read

- Gene H. Golub and Charles F. Van Loan. *Matrix Computations*, Johns Hopkins University Press, 2013.

  Chapter 3.1 – 3.4

- Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra*, SIAM, 1997.

  Lecture 20 – 22