Chapter 6 Direct Method for Solving Linear Systems

Baodong LIU
baodong@sdu.edu.cn

- **the Linear System of Equations(LSEs):**

$$(I)\begin{cases} E_1 : a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ E_2 : a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad\quad\quad\quad \vdots \quad\quad \vdots \\ E_n : a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

Problem: Given the coefficients
$a_{ij}, b_i,\ i,j = 1, 2, \cdots, n$, to determine $x_1, x_2, \cdots, x_n$

- **Direct techniques** are methods that **gives an answer in a fixed number of steps**, and subject only to the **roundoff error**.

- **the Linear System of Equations(LSEs):**

$$(I)\begin{cases} E_1: a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ E_2: a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \quad \vdots \qquad \vdots \qquad \vdots \quad \vdots \\ E_n: a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

Problem: Given the coefficients $a_{ij}, b_i,\ i,j = 1,2,\cdots,n$, to determine $x_1, x_2, \cdots, x_n$

- **Direct techniques** are methods that **gives an answer in a fixed number of steps**, and subject only to the **roundoff error**.

6.1 Linear System of Equations

Three operations to **simplify** the linear system:

1. **Multiplied operation–数乘**:
   Equation E_i can be multiplied by any nonzero constant λ

$$(\lambda E_i) \to E_i.$$

2. **Multiplied and added operation–倍加**:
   Equation E_j can be multiplied by nonzero contant, and added to Equation E_i in place of E_i, denoted by

$$(\lambda E_j + E_i) \to E_i.$$

3. **Transposition–交换**:
   Equation E_i and E_j can be transposed in order, denoted by

$$E_i \leftrightarrow E_j.$$

Example 1:

To solve the equations:

$$
\begin{array}{rrrrrrrrrrr}
E_1: & x_1 & + & x_2 & & & + & 3x_4 & = & 4, \\
E_2: & 2x_1 & + & x_2 & - & x_3 & + & x_4 & = & 1, \\
E_3: & 3x_1 & - & x_2 & - & x_3 & + & 2x_4 & = & -3, \\
E_4: & -x_1 & + & 2x_2 & + & 3x_3 & - & x_4 & = & 4,
\end{array}
$$

Step 1

Using the equation E_1 to eliminate the coefficients of x_1 from E_2, E_3, E_4 by performing operations:

$$(E_2 - 2E_1) \to (E_2), (E_3 - 3E_1) \to (E_3), (E_4 + E_1) \to (E_4)$$

and the resulting system is

$$\begin{array}{lrcrcrcrcr}
E_1: & x_1 & + & x_2 & & & + & 3x_4 & = & 4, \\
E_2: & & - & x_2 & - & x_3 & - & 5x_4 & = & -7, \\
E_3: & & - & 4x_2 & - & x_3 & - & 7x_4 & = & -15, \\
E_4: & & & 3x_2 & + & 3x_3 & + & 2x_4 & = & 8,
\end{array}$$

Step 2

Using the equation E_2 to eliminate the coefficients of x_2 from E_3, E_4 by performing operations:

$$(E_3 - 4E_2) \to (E_3), (E_4 + 3E_2) \to (E_4),$$

resulting the system

$$\begin{array}{rrrrrrrrrr}
E_1: & x_1 & + & x_2 & & & + & 3x_4 & = & 4, \\
E_2: & & - & x_2 & - & x_3 & - & 5x_4 & = & -7, \\
E_3: & & & & & 3x_3 & + & 13x_4 & = & 13, \\
E_4: & & & & & & - & 13x_4 & = & -13,
\end{array}$$

The system is now transposed in the **triangular form**.

Step 3

Using **backward–substitution(回代法)** method:

1. By the equation E_4 implies $x_4 = 1$
2. E_3 can be solved for x_3 to gives

$$x_3 = \frac{1}{3}(13 - 13x_4) = 0,$$

3. E_2 gives $x_2 = -(-7 + 5x_4 + x_3) = 2$
4. E_1 for x_1 gives $x_1 = 4 - 3x_3 - x_2 = -1$

Notation:

Let $\mathbf{A}$ be an $n \times m(n$ by $m)$ matrix, and

$$\mathbf{A} = (a_{ij})_{n\times m} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

and a_{ij} refers to the entry at the intersection of the ith row and jth column.

and

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

be n-dimensional column vector, and

$$\mathbf{y} = \begin{pmatrix} y_1 & y_2 & \cdots & y_m \end{pmatrix}$$

be m-dimensional row vector.

Now, the linear system of equations (LSEs) can be rewritten as in matrix form:

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

where, $\mathbf{A}$ is the coefficient $n \times n$ matrix, $\mathbf{x}$ and $\mathbf{b}$ are n-dimensional column vectors.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

To solve LSEs, we construct the **Augmented Matrix**

$$\tilde{\mathbf{A}} = [\mathbf{A}, \mathbf{b}] = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & \vdots & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & \vdots & b_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & \vdots & b_n \end{pmatrix}$$

Gaussian Elimination with Backward Substitution Method

Step 1. Assume that $a_{11} \neq 0$, perform the operation

$$(E_i - (a_{i1}/a_{11})E_1) \to E_i$$

to eliminate the coefficients $a_{21}, a_{31}, \cdots, a_{n1}$ of x_1 for each $i = 2, 3, \cdots, n$.

The resulting matrix has the form.

$$\tilde{\mathbf{A}}^{(1)} = [\mathbf{A}^{(1)}, \mathbf{b}^{(1)}] = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & \vdots & b_1 \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & \vdots & b_2^{(1)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & \vdots & b_n^{(1)} \end{pmatrix}$$

where for each $i = 2, 3, \cdots, n$,

$$\begin{aligned} a_{i1}^{(1)} &= 0; \\ a_{ij}^{(1)} &= a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j}, j = 2, 3, \cdots, n; \\ b_i^{(1)} &= b_i - \frac{a_{i1}}{a_{11}} b_1; \end{aligned}$$

Step 2. For $\tilde{A}^{(1)}$, suppose that $a_{22}^{(1)} \neq 0$, do operations

$$(E_i - (a_{i2}^{(1)}/a_{22}^{(1)})E_2) \rightarrow E_i, i = 3, 4, \cdots, n$$

to eliminate the coefficients

$$a_{32}^{(1)}, a_{42}^{(1)}, \cdots, a_{n2}^{(1)}.$$

Thus we obtain

$$\tilde{\mathbf{A}}^{(2)} = [\mathbf{A}^{(2)}, \mathbf{b}^{(2)}] = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & \vdots & b_1 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} & \vdots & b_2^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & \vdots & b_3^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{nn}^{(2)} & \vdots & b_n^{(2)} \end{pmatrix}$$

where for each $i = 3, \cdots, n$,

$$
\begin{aligned}
a_{i2}^{(2)} &= 0; \\
a_{ij}^{(2)} &= a_{ij}^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} a_{2j}^{(1)}, j = 3, \cdots, n; \\
b_i^{(2)} &= b_i^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} b_2^{(1)};
\end{aligned}
$$

Step $k-1$**:** suppose that we have done $k-1$ steps, and get that

$$\tilde{\mathbf{A}}^{(k-1)} = \left(\begin{array}{ccccccccc}
a_{11} & a_{12} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} & \vdots & b_1 \\
 & a_{22}^{(1)} & \cdots & a_{2k}^{(1)} & a_{2,k+1} & \cdots & a_{2n}^{(1)} & \vdots & b_2^{(1)} \\
 & & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & & & a_{kk}^{(k-1)} & a_{k,k+1}^{(k-1)} & \cdots & a_{kn}^{(k-1)} & \vdots & b_k^{(k-1)} \\
 & & & a_{k+1,k}^{(k-1)} & a_{k+1,k+1}^{(k-1)} & \cdots & a_{k+1,n}^{(k-1)} & \vdots & b_{k+1}^{(k-1)} \\
 & & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & & & a_{nk}^{(k-1)} & a_{n,k+1}^{k-1} & \cdots & a_{nn}^{(k-1)} & \vdots & b_n^{(k-1)}
\end{array}\right)$$

Step k If $a_{kk}^{(k-1)} \neq 0$, then do the kth step

$$(E_i - (a_{ik}^{(k-1)}/a_{kk}^{(k-1)})E_k) \to E_i, i = k+1, k+2, \cdots, n$$

to eliminate the coefficients

$$a_{k+1,k}^{(k-1)}, a_{k+2,k}^{(k-1)}, \cdots, a_{n,k}^{(k-1)},$$

and obtain the new matrix form

$$
\begin{aligned}
\tilde{\mathbf{A}}^{(k)} &= [\mathbf{A}^{(k)}, \mathbf{b}^{(k)}] \\
&= \begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} & \vdots & b_1 \\
 & a_{22}^{(1)} & \cdots & a_{2k}^{(1)} & a_{2,k+1} & \cdots & a_{2n}^{(1)} & \vdots & b_2^{(1)} \\
 & & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & & & a_{kk}^{(k-1)} & a_{k,k+1}^{(k-1)} & \cdots & a_{kn}^{(k-1)} & \vdots & b_k^{(k-1)} \\
 & & & 0 & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} & \vdots & b_{k+1}^{(k)} \\
 & & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & & & 0 & a_{n,k+1}^{k} & \cdots & a_{nn}^{(k)} & \vdots & b_n^{(k)}
\end{pmatrix}
\end{aligned}
$$

where for each $i = k+1, \cdots, n$,

$$\begin{aligned} a_{ik}^{(2)} &= 0; \\ a_{ij}^{(k)} &= a_{ij}^{(k-1)} - \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} a_{kj}^{(k-1)}, j = k+1, \cdots, n; \\ b_i^{(2)} &= b_i^{(k-1)} - \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} b_k^{(k-1)}; \end{aligned}$$

Step n**.** After $(n-1)$th elimination process, we obtain

$$\tilde{\mathbf{A}}^{(n-1)} = [\mathbf{A}^{(n-1)}, \mathbf{b}^{(n-1)}]$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} & \vdots & b_1 \\ & a_{22}^{(1)} & \cdots & a_{2k}^{(1)} & a_{2,k+1} & \cdots & a_{2n}^{(1)} & \vdots & b_2^{(1)} \\ & & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & a_{kk}^{(k-1)} & a_{k,k+1}^{(k-1)} & \cdots & a_{kn}^{(k-1)} & \vdots & b_k^{(k-1)} \\ & & & & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} & \vdots & b_{k+1}^{(k)} \\ & & & & & \ddots & \vdots & \vdots & \vdots \\ & & & & & & {\color{blue}a_{nn}^{(n-1)}} & \vdots & {\color{blue}b_n^{(n-1)}} \end{pmatrix}$$

Result:

- Thus the corresponding linear system of equations is transposed to the new linear system of equations with upper-triangular form.

$$(II)\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &=& b_1 \\ a_{22}x_2 + \cdots + a_{2n}x_n &=& b_2 \\ \ddots \qquad\qquad & & \vdots \\ a_{nn}x_n &=& b_n \end{array}\right.$$

- Thus the **backward substitution** can be performed for solving the new LSEs.

$$\begin{aligned} x_n &= \frac{b_n}{a_{nn}} \\ x_j &= \frac{b_j - \sum\limits_{k=j+1}^{n} a_{jk}x_k}{a_{jj}}, j = n-1, n-2, \cdots, 1 \end{aligned}$$

For augmented matrix $[\mathbf{A}, \mathbf{b}]$ to triangular from:

加/减运算

$$(n-1)(n+1)+(n-2)n+\cdots+1\cdot 3$$
$$=\frac{(2n-3)(n+1)(n+2)}{6}+1$$

乘/除运算

$$(n-1)(n+1+1)+(n-2)(n+1)+\cdots+1\cdot 4$$
$$=\frac{n^3}{3}+n^2-\frac{4n}{3}$$

Notes on Gaussian Elimination Method:

- The Gaussian Elimination Procedure will fail if one of the elements

$$a_{11},\, a_{22}^{(1)},\, \cdots,\, a_{kk}^{(k-1)},\, \cdots,\, a_{nn}^{(n-1)}$$

  is zero, because in this case, the step

$$\left(E_i - \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} E_k\right) \to (E_i)$$

  either cannot be performed, or the backward substitution cannot be accomplished.
- This does not mean that the linear system has no solution, but rather that the technique for finding the solution must be altered.

- To continue the Gaussian Elimination Procedure in the case that some $a_{kk}^{(k-1)} = 0$ for $k = 1, 2, \cdots, n-1$.
  -
    - Searching the kth column of $\tilde{A}_{kk}^{(k-1)}$ from the kth row to the nth row for the first nonzero entry.
    - Suppose that $a_{pk}^{(k-1)} \neq 0$ for some p, $k+1 \leq p \leq n$
    - Perform the operation $(E_k) \leftrightarrow (E_p)$, and then do elimination step with new matrix
  - otherwise if all

$$a_{pk}^{(k-1)} = 0, k+1 \leq p \leq n,$$

then by the solution theory of linear system, the linear system does not have a unique solution.

- To continue the Gaussian Elimination Procedure in the case that some $a_{kk}^{(k-1)} = 0$ for $k = 1, 2, \cdots, n-1$.
    - Searching the kth column of $\tilde{A}_{kk}^{(k-1)}$ from the kth row to the nth row for the first nonzero entry.
    - Suppose that $a_{pk}^{(k-1)} \neq 0$ for some p, $k+1 \leq p \leq n$
    - Perform the operation $(E_k) \leftrightarrow (E_p)$, and then do elimination step with new matrix
  - otherwise if all

$$a_{pk}^{(k-1)} = 0, k+1 \leq p \leq n,$$

then by the solution theory of linear system, the linear system does not have a unique solution.

ALGORITHM 6.1: Gaussian Elimination and Backward Substitution Method

Input: N–dimension, $A(N,N), B(N)$

Output: Solution $x(N)$ or Message that LESs has no unique solution.

Step 1 For $k = 1, 2, \cdots, N-1$, do step 2-4.

Step 2 Set p be the smallest integer with $k \leq p \leq N$ and $A(p,k) \neq 0$(If no p can be found, output:"no unique solution exists"; stop).

Step 3 If $p \neq k$, do transposition $E_p \leftrightarrow E_k$.

Step 4 For $i = k+1, \cdots, N$

- 1. Set $m_{i,k} = A(i,k)/A(k,k)$
- 2. Set $B(i) = B(i) - m_{i,k}B(k)$
- 3. For $j = k+1, \cdots, N$, set $A(i,j) = A(i,j) - m_{i,k}A(k,j)$;

ALGORITHM 6.1: Continued

Step 5 If $A(N,N) \neq 0$,set $x(N) = B(N)/A(N,N)$;
Else, output:"no unique solution exists."

Step 6 For $i = N-1, N-2, \cdots, 1$. SET

$$x(i) = \left[B(i) - \sum_{j=i+1}^{N} A(i,j)x(j)\right] / A(i,i)$$

Step 7 Output the solution $x(N)$

Problem:

- When $a_{kk}^{(k-1)} = 0$ in Gaussian Elimination method, we always need to make a row interchange.
- If $a_{kk}^{(k-1)} \neq 0$, but is much small in magnitude compared to a_{ik}, $i = k+1, k+2, \cdots, n$, then the multiplier m_{ik}

$$m_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} >> 1$$

or when performing the backward substitution for

$$x_k = \frac{b_k^{(k-1)} - \sum_{j=k+1}^{n} a_{kj}^{(k-1)} x_j}{a_{kk}^{(k-1)}} >> 1$$

★ To reduce roundoff error, it is necessary to perform row interchange even when $a_{kk}^{(k-1)} \neq 0$.

Gaussian Elimination with Maximal Column (partial) Pivoting Technique—列主元消去法:

Algorithm description:

- For $k = 1, 2, \cdots, n-1$, Choose the smallest integer $p, p \geq k$, such that

$$|a_{pk}^{(k-1)}| = \max_{k \leq i \leq n} |a_{ik}^{(k-1)}|.$$

- If $p \neq k$, make interchange between the rows E_p and E_k: $(E_p) \leftrightarrow (E_k)$, then do Gaussian elimination to make the elements

$$a_{k+1,k}^{(k-1)}, \cdots, a_{n,k}^{(k-1)}$$

to be zeros.

Gaussian Elimination with Maximal Row Pivoting Technique—行主元消去法

Algorithm description:

- For $k = 1, 2, \cdots, n-1$, Choose the smallest integer

$$|a_{kp}^{(k-1)}| = \max_{k \le j \le n} |a_{kj}^{(k-1)}|.$$

- If $p \neq k$, make interchange between pth and kth columns, and $x_k \leftrightarrow x_p$ then do Gaussian elimination to make the elements $a_{k+1,k}^{(k-1)}, \cdots, a_{n,k}^{(k-1)}$ to be zeros.

Gaussian Elimination with Partial Pivoting Technique–部分主元消去法

Algorithm description:

- For $k = 1, 2, \cdots, n-1$, Choose the maximal element

$$|a_{pq}^{(k-1)}| = \max_{k \le i,j \le n} |a_{ij}^{(k-1)}|.$$

- If $p \ge k$ and $q \ge k$, make interchanges between the q and kth columns, and then $E_p \leftrightarrow E_k$, do Gaussian elimination to make the elements $a_{k+1,k}^{(k-1)}, \cdots, a_{n,k}^{(k-1)}$ to be zero.

Gaussian Elimination with scaled Partial Pivoting Technique—比例列主元消去法

- First, define the scaled factors s_i for each row vector $i = 1, 2, \cdots, n$,
$$s_i = \max_{1 \le j \le n} |a_{ij}|.$$
- If for some i we have $s_i = 0$, then the system has no unique solution since all elements in the ith row are zero.
- If no this kind of cases, then choose the integer p, such that
$$\frac{a_{p1}}{s_p} = \max_{1 \le i \le n} \frac{a_{i1}}{s_i}.$$

- Performing the interchange $E_p \leftrightarrow E_1$,then applying Gaussian Elimination Technique to eliminate the elements to be zero in the first column after a_{11}
- In a similar manner, at the step k, find the small integer $r, r \geq k$, such that

$$\frac{a_{rk}}{s_r} = \max_{k \leq i \leq n} \frac{a_{ik}}{s_i}.$$

then do interchange $E_r \leftrightarrow E_k$, and applying Gaussian Elimination Technique to eliminate the elements to be zero in the kth column after a_{kk}

Example:

- Solve the linear system using three-digit rounding arithmetic.

$$\begin{array}{lrcrcrcr} E_1: & 2.11x_1 & - & 4.21x_2 & + & 0.921x_3 & = & 2.01, \\ E_2: & 4.01x_1 & + & 10.2x_2 & - & 1.12x_3 & = & -3.09, \\ E_3: & 1.09x_1 & + & 0.987x_2 & + & 0.832x_3 & = & 4.21. \end{array}$$

- The augmented matrix is in the form

$$[A, b] = \left(\begin{array}{ccccc} 2.11 & -4.21 & 0.921 & \vdots & 2.01 \\ 4.01 & 10.2 & -1.12 & \vdots & -3.09 \\ 1.09 & 0.987 & 0.832 & \vdots & 4.21 \end{array} \right)$$

Partial Pivoting method:

Step 1 $\max_{1\leq i\leq 3}|a_{i1}| = a_{21} = 4.01$, making row interchange $(E_2) \leftrightarrow (E_1)$, get new matrix form

$$A_1 = \begin{pmatrix} 4.01 & 10.2 & -1.12 & \vdots & -3.09 \\ 2.11 & -4.21 & 0.921 & \vdots & 2.01 \\ 1.09 & 0.987 & 0.832 & \vdots & 4.21 \end{pmatrix}$$

Performing Gaussian Elimination, gives

$$A_2 = \begin{pmatrix} 4.01 & 10.2 & -1.12 & \vdots & -3.09 \\ 0 & -9.58 & 1.51 & \vdots & 3.64 \\ 0 & -1.79 & 1.14 & \vdots & 5.05 \end{pmatrix}$$

Step 2 Comparing $a_{22}, a_{3,2}$, get the absolute maximum element $|a_{22}| = 9.58$, no row interchange needs to interchange, do Gaussian Elimination to make a_{32} to be zero

$$A_3 = \begin{pmatrix} 4.01 & 10.2 & -1.12 & \vdots & -3.09 \\ 0 & -9.58 & 1.51 & \vdots & 3.64 \\ 0 & 0 & 0.858 & \vdots & 4.37 \end{pmatrix}$$

Step 3 Making backward substitution, we get the solution

$$x_3 = 5.09, x_2 = 0.422, x_1 = 0.422.$$

Scaling Partial Pivoting Method:

$$A_1 = [A, b] = \begin{pmatrix} 2.11 & \mathbf{-4.21} & 0.921 & \vdots & 2.01 \\ 4.01 & \mathbf{10.2} & -1.12 & \vdots & -3.09 \\ \mathbf{1.09} & 0.987 & 0.832 & \vdots & 4.21 \end{pmatrix}$$

Step1. Since $s_1 = 4.21, s_2 = 10.2, s_3 = 1.09$, so

$$\frac{|a_{11}|}{s_1} = \frac{2.11}{4.21} = 0.501, \frac{|a_{21}|}{s_2} = \frac{4.01}{10.2} = 0.393,$$

$$\frac{|a_{31}|}{s_3} = \frac{1.09}{1.09} = 1.$$

Making row interchange $(E_3) \leftrightarrow (E_1)$, obtain

$$A_2 = \begin{pmatrix} 1.09 & 0.987 & 0.832 & \vdots & 4.21 \\ 4.01 & 10.2 & -1.12 & \vdots & -3.09 \\ 2.11 & -4.21 & 0.921 & \vdots & 2.01 \end{pmatrix}$$

and then performing Gaussian Elimination to eliminate a_{21}, a_{31}, obtain

$$A_3 = \begin{pmatrix} 1.09 & 0.987 & 0.832 & \vdots & 4.21 \\ 0 & \mathbf{6.57} & -4.18 & \vdots & -18.6 \\ 0 & \mathbf{-6.12} & -0.689 & \vdots & -6.16 \end{pmatrix}$$

Step2. It is easy to see that

$$\frac{|a_{22}|}{s_2} = \frac{6.57}{10.2} = 0.644 < \frac{|a_{32}|}{s_3} = \frac{6.12}{4.21} = 1.45.$$

(备注：因换行关系，此处$s_3 = s_1$.)

Making row interchange $(E_3) \leftrightarrow (E_2)$, obtain

$$A_4 = \begin{pmatrix} 1.09 & 0.987 & 0.832 & \vdots & 4.21 \\ 0 & \mathbf{-6.12} & -0.689 & \vdots & -6.16 \\ 0 & \mathbf{6.57} & -4.18 & \vdots & -18.6 \end{pmatrix}$$

and then performing Gaussian Elimination to eliminate a_{32}, obtain

$$A_5 = \begin{pmatrix} 1.09 & 0.987 & 0.832 & \vdots & 4.21 \\ 0 & -6.12 & -0.689 & \vdots & -6.16 \\ 0 & 0 & -4.92 & \vdots & -25.2 \end{pmatrix}$$

Step 3 Making backward substitution, we get the solution

$$x_3 = 5.12, x_2 = 0.430, x_1 = 0.431.$$

To solve the LSEs: $\mathbf{Ax} = \mathbf{b}$, if $\mathbf{A}$ has been factored into the triangular form $\mathbf{A} = \mathbf{LU}$, with the form

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{pmatrix},$$

$$\mathbf{U} = \begin{pmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{pmatrix}$$

- Then the system can be written as $\mathbf{LUx} = \mathbf{b}$, and we can solve for $\mathbf{x}$ more easily by using a two-step process:
  1. First we let
  $$\mathbf{y} = \mathbf{Ux}$$
  and solve the system
  $$L\mathbf{y} = \mathbf{b}$$
  for $\mathbf{y}$.
  2. Once $\mathbf{y}$ is known, the same as the backward substitution method can be used to solve linear system
  $$\mathbf{Ux} = \mathbf{y}$$
  to determine $\mathbf{x}$.

- To examine which matrices have an $\mathbf{LU}$ factorization and find how it is determined, first suppose that Gaussian elimination can be performed on the system $\mathbf{Ax} = \mathbf{b}$ without row interchanges.
- The first step in the Gaussian elimination process consists of performing, for each $j = 2, 3, \cdots, n$, the operations

$$(E_j - m_{j,1}E_1) \to (E)_j, \texttt{where} \quad m_{j,1} = \frac{a_{j1}}{a_{11}}$$

The first Gaussian transformation matrix

- It is simultaneously accomplished by multiplying the original matrix $\mathbf{A}$ on the left by the matrix

$$\mathbf{M}^{(1)} = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ -m_{21} & 1 & 0 & \ddots & & 0 \\ -m_{31} & 0 & 1 & & & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ -m_{n1} & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

where

$$m_{j,1} = \frac{a_{j1}}{a_{11}}, j = 2, 3, \cdots, n$$

- We denote the product of this matrix with $\mathbf{A}^{(1)} \equiv \mathbf{A}$ by $\mathbf{A}^{(2)}$ and with $\mathbf{b}$ by $\mathbf{b}^{(2)}$, so

$$\mathbf{A}^{(2)}\mathbf{x} = \mathbf{M}^{(1)}\mathbf{A}\mathbf{x} = \mathbf{M}^{(1)}\mathbf{b} = \mathbf{b}^{(2)}.$$

The Second Gaussian transformation matrix

- In a similar manner we construct $\mathbf{M}^{(2)}$

$$\mathbf{M}^{(2)} = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \ddots & & 0 \\ 0 & -m_{32} & 1 & & & 0 \\ \vdots & \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & -m_{n2} & 0 & \cdots & 0 & 1 \end{pmatrix}$$

where

$$m_{j,2} = \frac{a_{j2}^{(2)}}{a_{22}^{(2)}}, j = 3, 4, \cdots, n$$

- Let

$$\begin{aligned} \mathbf{A}^{(3)}\mathbf{x} &= \mathbf{M}^{(2)}\mathbf{A}^{(2)}\mathbf{x} = \mathbf{M}^{(2)}\mathbf{M}^{(1)}\mathbf{A}^{(1)}\mathbf{x} = \mathbf{M}^{(2)}\mathbf{b}^{(2)} \\ &= \mathbf{b}^{(3)} = \mathbf{M}^{(2)}\mathbf{M}^{(1)}\mathbf{b}^{(1)}. \end{aligned}$$

In general, with $\mathbf{A}^{(k)}\mathbf{x} = \mathbf{b}^{(k)}$ already formed, multiply by the kth Gaussian transformation matrix

$$\mathbf{M}^{(k)} = \begin{pmatrix} 1 & \cdots & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & 0 & \ddots & & 0 \\ 0 & \ddots & 1 & & & 0 \\ \vdots & & -m_{k+1,k} & \ddots & \ddots & \vdots \\ \vdots & & & \vdots & \ddots & \vdots \\ 0 & \cdots & -m_{n,k} & \cdots & 0 & 1 \end{pmatrix}$$

where

$$m_{j,k} = \frac{a_{jk}^{(k)}}{a_{kk}^{(k)}}, j = k+1, k+2, \cdots, n$$

to obtain

$$\begin{aligned} \mathbf{A}^{(k+1)}\mathbf{x} &= \mathbf{M}^{(k)} A^{(k)}\mathbf{x} = \mathbf{M}^{(k)} \cdots \mathbf{M}^{(1)}\mathbf{A}\mathbf{x} \\ &= \mathbf{M}^{(k)}\mathbf{b}^{(k)} = \mathbf{b}^{(k+1)} = \mathbf{M}^{(k)} \cdots \mathbf{M}^{(1)}\mathbf{b}. \end{aligned}$$

The process ends with the formation of $\mathbf{A}^{(n)}\mathbf{x} = \mathbf{b}^{(n)}$, where $\mathbf{A}^{(n)}$ is the upper triangular matrix

$$\mathbf{A}^{(n)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1k}^{(1)} & a_{1,k+1}^{(1)} & \cdots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} & a_{2,k+1}^{(2)} & \cdots & a_{2n}^{(2)} \\ & & \ddots & \vdots & \vdots & \vdots & \vdots \\ & & & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \cdots & a_{kn}^{(k)} \\ & & & & a_{k+1,k+1}^{(k+1)} & \cdots & a_{k+1,n}^{(k+1)} \\ & & & & & \ddots & \vdots \\ & & & & & & a_{nn}^{(n)} \end{pmatrix}$$

given by

$$\mathbf{A}^{(n)} = \mathbf{M}^{(n-1)}\mathbf{M}^{(n-2)}\cdots\mathbf{M}^{(1)}\mathbf{A}.$$

- Note that $\mathbf{M}^{(n-1)}, \mathbf{M}^{(n-2)}, \cdots, \mathbf{M}^{(1)}$ are non-singular. matrices
- Let $\mathbf{U} = \mathbf{A}^{(n)}$, then $\mathbf{U}$ is an upper-triangular(上三角) matrix.
- This means that

$$\mathbf{U} = \mathbf{M}^{(n-1)}\mathbf{M}^{(n-2)}\cdots\mathbf{M}^{(1)}\mathbf{A}.$$

or

$$[\mathbf{M}^{(1)}]^{-1}\cdots[\mathbf{M}^{(n-2)}]^{-1}[\mathbf{M}^{(n-1)}]^{-1}\mathbf{U} = \mathbf{A}.$$

- Let

$$\mathbf{L}^{(k)} = [\mathbf{M}^{(k)}]^{-1} = \begin{pmatrix} 1 & \cdots & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & 0 & \ddots & & 0 \\ 0 & \ddots & 1 & & & 0 \\ \vdots & & m_{k+1,k} & \ddots & \ddots & \vdots \\ \vdots & & & \vdots & \ddots & \vdots \\ 0 & \cdots & m_{n,k} & \cdots & 0 & 1 \end{pmatrix}$$

- So The lower-triangular matrix L in the factorization of A is the product of the matrices $L^{(k)}$:

$$\mathbf{L} = \mathbf{L}^{(1)}\mathbf{L}^{(2)}\cdots\mathbf{L}^{(n-1)} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{n,n-1} & 1 \end{pmatrix}$$

Theorem 6.18

If Gaussian elimination can be performed on the linear system $\mathbf{Ax} = \mathbf{b}$ without row interchanges, then the matrix $\mathbf{A}$ can be factored into the product of a lower-triangular $\mathbf{L}$ and an upper-triangular matrix $\mathbf{U}$,

$$\mathbf{A} = \mathbf{LU}$$

where

$$\mathbf{U} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \cdots & \cdots & a_{2n}^{(2)} \\ & & \ddots & \vdots & \vdots \\ & & & \ddots & \vdots \\ & & & & a_{nn}^{(n)} \end{pmatrix},$$

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ m_{21} & 1 & 0 & \cdots & 0 \\ m_{31} & m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{n,n-1} & 1 \end{pmatrix}$$

Some Remarks on LU Factorization:— 直接求解如何设计？

$$\mathbf{LU} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} = \mathbf{A}.$$

Algorithm: Direct LU Factorization

To factor the $n \times n$ matrix $\mathbf{A} = (a_{ij})$ into the product of the lower-triangular matrix $\mathbf{L} = (l_{ij})$ and the upper-triangular matrix $\mathbf{U} = (u_{ij})$, where the main diagonal of either $\mathbf{L}$ or $\mathbf{U}$ consists of all ones. Here we suppose that the diagonal $l_{11} = l_{22} = \cdots = l_{nn} = 1$ of $\mathbf{L}$.

INPUT:

- dimension n; the entries $a_{ij}, 1 \leq i, j \leq n$ of $\mathbf{A}$;
- the diagonal $l_{11} = l_{22} = \cdots = l_{nn} = 1$ of $\mathbf{L}$, or

OUTPUT: the entries of $\mathbf{L}$

$$l_{ij}, 1 \leq j \leq i, 1 \leq i \leq n$$

and the entries of $\mathbf{U}$

$$u_{ij}, i \leq j \leq n, 1 \leq i \leq n$$

Step 1 Select l_{11} and u_{11} satisfying $l_{11}u_{11} = a_{11}$.
If $l_{11}u_{11} = 0$ then OUTPUT ('Factorization impossible'); STOP

Step 2 For $j = 2, \cdots, n$, set

$$u_{1j} = a_{1j}/l_{11}(\texttt{First row of } \mathbf{U}.);$$

$$l_{j1} = a_{j1}/u_{11}.(\texttt{First column of } \mathbf{L})$$

Step 3 For $i = 2, \cdots, n-1$ do Steps 4 and 5.

Step 4 Select l_{ii} and u_{ii} satisfying

$$l_{ii}u_{ii} = a_{ii} - \sum_{k=1}^{i-1} l_{ik}u_{ki}.$$

If $l_{ii}u_{ii} = 0$ then OUTPUT ('Factorization impossible'); STOP.

Step 5 For $j = i+1, \cdots, n$, set

$$u_{ij} = \frac{1}{l_{ii}}[a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}], (i\texttt{th row of } \mathbf{U})$$

$$l_{ji} = \frac{1}{u_{ii}}[a_{ji} - \sum_{k=1}^{i-1} l_{jk} u_{ki}], (i\texttt{th column of } \mathbf{L})$$

Step 6 Select l_{nn} and u_{nn} satisfying

$$l_{nn} u_{nn} = a_{nn} - \sum_{k=1}^{n-1} l_{nk} u_{kn}.$$

(**Note:** If $l_{nn} u_{nn} = 0$, then $\mathbf{A} = \mathbf{LU}$ but $\mathbf{A}$ is singular.)

Step 7 OUTPUT
(l_{ij} for $j = 1, \cdots, i$ and $i = 1, \cdots, n$);
(u_{ij} for $j = i, \cdots, n$ and $i = 1, \cdots, n$); STOP.

在存在换行情况下，如何进行矩阵分解？

观察：permutation matrix (置换矩阵)

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \ \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

$$\mathbf{PA} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix};$$

$$\mathbf{AP} = \begin{pmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{22} \\ a_{31} & a_{33} & a_{32} \end{pmatrix};$$

结论：

1. 置换矩阵是单位矩阵换行而成；

2. 选主元而需要的矩阵行变换等价于利用相同变换的置换矩阵左乘需要矩阵；

3. 置换矩阵$\mathbf{P}^{-1}=\mathbf{P}^{T}$

4. 在需要因选主元而进行换行的情况下，相应操作可描述为:

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

$$\mathbf{Ax}=\mathbf{b},\to\mathbf{PAx}=\mathbf{Pb}\Rightarrow\mathbf{LUx}=\mathbf{Pb}$$

5. 高斯选主元过程可以描述为：

$$\mathbf{M}^{(n-1)}\mathbf{P}^{(n-1)}\mathbf{M}^{(n-2)}\mathbf{P}^{(n-2)}\cdots\mathbf{M}^{(1)}\mathbf{P}^{(1)}\mathbf{A}=\mathbf{A}^{(n)}=\mathbf{U}.$$

$$\begin{aligned}
\mathbf{A} &= \mathbf{A}^{(n)}=(\mathbf{M}^{(n-1)}\mathbf{P}^{(n-1)}\mathbf{M}^{(n-2)}\mathbf{P}^{(n-2)}\cdots\mathbf{M}^{(1)}\mathbf{P}^{(1)})^{-1}\mathbf{U} \\
&= {\mathbf{P}^{(1)}}^{-1}{\mathbf{M}^{(1)}}^{-1}\cdots{\mathbf{P}^{(n-1)}}^{-1}{\mathbf{M}^{(n-1)}}^{-1}\mathbf{U} \\
&= {\mathbf{P}^{(1)}}^{T}{\mathbf{M}^{(1)}}^{-1}\cdots{\mathbf{P}^{(n-1)}}^{T}{\mathbf{M}^{(n-1)}}^{-1}\mathbf{U}=\mathbf{LU}.
\end{aligned}$$

Example on LU Factorization

Example: Determine $\mathbf{LU}$ factorization of $\mathbf{A}$:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{pmatrix}$$

- First Gaussian elimination:

$$\mathbf{M}^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix},$$

and

$$\mathbf{A}^{(2)} = \mathbf{M}^{(1)}\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & -4 & -1 & -7 \\ 0 & 3 & 3 & 2 \end{pmatrix}$$

Example on LU Factorization

Second Gaussian elimination:

$$\mathbf{M}^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -4 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{pmatrix},$$

and

$$\mathbf{A}^{(3)} = \mathbf{M}^{(2)}\mathbf{A}^{(2)} = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix} = \mathbf{U}$$

Thus

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{pmatrix},$$

Example on LU Factorization

可以验证：$\mathbf{A} = \mathbf{LU}$

Example: Find the LU factoriazation for matrix A

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 1 & 1 & -1 & 2 \\ -1 & -1 & 2 & 0 \\ 1 & 2 & 0 & 2 \end{pmatrix}$$

按照列主元消去法:
第一步: $E_2 \leftrightarrow E_1$

$$\mathbf{P}^{(1)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{P}^{(1)}\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ -1 & -1 & 2 & 0 \\ 1 & 2 & 0 & 2 \end{pmatrix}$$

$$\mathbf{M}^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \mathbf{A}^{(2)} = \mathbf{M}^{(1)}\mathbf{P}^{(1)}\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

第二步：$E_4 \leftrightarrow E_2$

$$\mathbf{P}^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \mathbf{A}^{(3)} = \mathbf{P}^{(2)}\mathbf{A}^{(2)} = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & 1 \end{pmatrix},$$

第二列无需消元

第三步：无需换行，

$$\mathbf{M}^{(3)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

$$\mathbf{A}^{(3)} = \mathbf{M}^{(3)}\mathbf{A}^{(3)} = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix} = \mathbf{U}$$

$$\begin{aligned} \mathbf{L} &= \mathbf{P}^{(1)^{-1}}\mathbf{M}^{(1)^{-1}}\mathbf{P}^{(2)^{-1}}\mathbf{M}^{(3)^{-1}} \\ &= \mathbf{P}^{(1)^{T}}\mathbf{M}^{(1)^{-1}}\mathbf{P}^{(2)^{T}}\mathbf{M}^{(3)^{-1}} \\ &= \begin{pmatrix} 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \end{aligned}$$

备注：利用MATLAB 调用lu(A),可得相同的计算结果.

In this section, we will discuss two special matrices, which Gaussian Elimination can be performed without row interchanges.

Definition 6.18

The $n \times n$ matrix $\mathbf{A}$ is said to be **strictly diagonally dominant**(严格对角占优) when

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|$$

holds for each $i = 1, 2, 3, \cdots, n$.

Example 1: Consider two matrices

$$\mathbf{A} = \begin{bmatrix} 7 & 2 & 0 \\ 3 & 5 & -1 \\ 0 & 5 & -6 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 6 & 4 & -3 \\ 4 & -2 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

- It can be seen that $\mathbf{A}$ is a **nonsymmetric matrix** and **strictly diagonally dominant**. But $\mathbf{A}^T$ is not strictly diagonally dominant
- The **symmetric matrix B** is not strictly diagonally dominant, nor, of course, $\mathbf{B}^T = \mathbf{B}$ is also not.

Theorem 6.19

- A strictly diagonally dominant matrix $\mathbf{A}$ is **nonsingular**.
- Moreover, in this case, Gaussian elimination can be performed on any linear system of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$ to obtain its unique solution without row or column interchanges, and the computations are stable with respect to the growth of roundoff errors.

Symmetric and Positive Definite Matrix–对称正定矩阵

Definition 6.20

A matrix $\mathbf{A}$ is positive definite if it is **symmetric** and if $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for every n-dimensional column vector $\mathbf{x} \neq \mathbf{0}$.

- **Note that:** not all authors require symmetry of a positive definite matrix.

Theorem 6.21

If $\mathbf{A}$ is an $n \times n$ positive definite matrix, then

1. $\mathbf{A}$ is nonsingular
2. $a_{ii} > 0$ for each $i = 1, 2, \cdots, n$.
3. $\max_{1 \le k,j \le n} |a_{k,j}| \le \max_{1 \le i \le n} |a_{ii}|$
4. $(a_{ij})^2 < a_{ii} a_{jj}$ for each $i \neq j$

Definition 6.22

A leading principal submatrix(—主子阵) of a matrix $\mathbf{A}$ is a matrix of the form

$$\mathbf{A}_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$$

Theorem 6.23

A symmetric matrix $\mathbf{A}$ is positive definite if and only if each of its leading principal submatrices has a positive determinant.

Theorem 6.24

The symmetric matrix $\mathbf{A}$ is positive definite if and only if Gaussian elimination without row interchanges can be performed on the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ with all pivot elements positive. Moreover, in this case, the computations are stable with respect to the growth of roundoff errors.

- Some interesting facts that are uncovered in constructing the proof of Theorem 6.24 are presented in the following corollaries.

Corollary 6.25

The matrix $\mathbf{A}$ is **positive definite** if and only if $\mathbf{A}$ can be factored in the form $\mathbf{L}\mathbf{D}\mathbf{L}^T$, where $\mathbf{L}$ is lower triangular with 1s on its diagonal and $\mathbf{D}$ is a diagonal matrix with positive diagonal entries.

Corollary 6.26

The matrix $\mathbf{A}$ is **positive definite** if and only if $\mathbf{A}$ can be factored in the form $\mathbf{L}\mathbf{L}^T$, where $\mathbf{L}$ is lower triangular with nonzero diagonal entries.

$\mathbf{LL}^T$ Factorization—Choleski Algorithm

For a $n \times n$ symmetric and positive definite matrix $\mathbf{A}$ with the form

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{12} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn} \end{pmatrix}$$

where $\mathbf{A}^T = \mathbf{A}$.

- To factor the positive definite matrix $\mathbf{A}$ into the form $\mathbf{L}\mathbf{L}^T$, where $\mathbf{L}$ is a lower triangular matrix with form as follows

$$\mathbf{L} = \begin{pmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & \cdots & 0 \\ l_{31} & l_{32} & l_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{nn} \end{pmatrix},$$

- The problem is to determine the elements $l_{ij}, j = 1, 2, \cdots, i$, for each $i = 1, 2, \cdots, n$.

- Let us first view the relationship by the definition of equal matrices

$$
\begin{aligned}
\mathbf{A} &= \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{12} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn} \end{pmatrix} = \mathbf{L}\mathbf{L}^T \\
&= \begin{pmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & \cdots & 0 \\ l_{31} & l_{32} & l_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & {}_{n2} & l_{n3} & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} & \cdots & l_{n1} \\ 0 & l_{22} & l_{32} & \cdots & l_{n2} \\ 0 & 0 & l_{33} & \cdots & l_{n3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & l_{nn} \end{pmatrix}
\end{aligned}
$$

- First, determining l_{11} by

$$l_{11} = \sqrt{a_{11}}$$

and $l_{i1}, i = 2, 3, \cdots, n$, by

$$l_{i1} l_{11} = a_{i1}, \texttt{thus} \quad l_{i1} = \frac{a_{1i}}{l_{11}}.$$

- Second to determine l_{22}, by

$$l_{21}^2 + l_{22}^2 = a_{22}$$

thus

$$l_{22} = [a_{22} - l_{21}^2]^{1/2}.$$

and then the second column $l_{i2}, i = 3, 4, \cdots, n$ of matrix L satisfies

$$l_{21} l_{i1} + l_{22} l_{i2} = a_{2i}, i = 3, 4, \cdots, n.$$

so we obtain

$$l_{i2} = \frac{a_{2i} - l_{21} l_{i1}}{l_{22}}, i = 3, 4, \cdots, n$$

- With similar idea, we can drive for any $k, k = 2, 3, \cdots, n-1$

$$l_{kk} = [a_{kk} - \sum_{j=1}^{k-1} l_{kj}^2]^{1/2},$$

and for $i = k+1, k+2, \cdots, n$,

$$l_{ik} = \frac{a_{ki} - \sum_{j=1}^{k-1} l_{kj} l_{ij}}{l_{kk}}$$

- Specially, when $k = n$, we have

$$l_{nn} = [a_{nn} - \sum_{j=1}^{n-1} l_{nj}^2]^{1/2}.$$

Choleski's Algorithm

To factor the positive definite $n \times n$ matrix $\mathbf{A}$ into $\mathbf{LL}^T$, where L is lower triangular:

INPUT the dimension n; entries a_{ij}, for $i = 1, 2, \cdots, n, j = 1, 2, \cdots, i$ of $\mathbf{A}$.

OUTPUT the entries l_{ij}, for $j = 1, 2, \cdots, i$ and $i = 1, 2, \cdots, n$ of $\mathbf{L}$

Step 1 Set $l_{11} = \sqrt{a_{11}}$.

Step2 For $j = 2, \cdots, n$,
set $l_{j1} = a_{1j}/l_{11}$.

Step 3 For $i = 2, \cdots, n-1$ do Steps 4 and 5.

Step 4 Set $l_{ii} = [a_{ii} - \sum_{j=1}^{i-1} l_{ij}^2]^{1/2}$.

Step 5 For $j = i+1, i+2, \cdots, n$, set $l_{ji} = \dfrac{a_{ij} - \sum\limits_{k=1}^{i-1} l_{ik} l_{jk}}{l_{ii}}$

Step 6 Set $l_{nn} = [a_{nn} - \sum\limits_{k=1}^{n-1} l_{nk}^2]^{1/2}$.

Step 7 OUTPUT l_{ij} for $j = 1, 2, \cdots, i$ and $i = 1, 2, \cdots, n$.
STOP!

$\mathbf{LDL}^T$ Factorization

- To factor the positive definite matrix $\mathbf{A}$ into the form $\mathbf{LDL}^T$, where $\mathbf{L}$ is a lower triangular matrix with form:

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & {}_{n2} & l_{n3} & \cdots & 1 \end{pmatrix},$$

and $\mathbf{D}$ is a diagonal matrix with positive entries on the diagonal, which can be formed as follows

$$\mathbf{D} = \begin{pmatrix} d_{11} & 0 & 0 & \cdots & 0 \\ 0 & d_{22} & 0 & \cdots & 0 \\ 0 & 0 & d_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_{nn} \end{pmatrix}$$

By the relationship, we can see that

$$
\begin{aligned}
\mathbf{A} &= \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{12} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nn} \end{pmatrix} = \mathbf{LDL}^T \\
&= \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} d_{11} & 0 & 0 & \cdots & 0 \\ 0 & d_{22} & 0 & \cdots & 0 \\ 0 & 0 & d_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_{nn} \end{pmatrix} \begin{pmatrix} 1 & l_{21} & l_{31} & \cdots \\ 0 & 1 & l_{32} & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & \cdots \end{pmatrix} \\
&= \begin{pmatrix} d_{11} & 0 & 0 & \cdots & 0 \\ l_{21} d_{11} & d_{22} & 0 & \cdots & 0 \\ l_{31} d_{11} & l_{32} d_{22} & d_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} d_{11} & l_{n2} d_{22} & l_{n3} & \cdots & d_{nn} \end{pmatrix} \begin{pmatrix} 1 & l_{21} & l_{31} & \cdots & l_{n1} \\ 0 & 1 & l_{32} & \cdots & l_{n2} \\ 0 & 0 & 1 & \cdots & l_{n3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}
\end{aligned}
$$

- So we can easily get that

$$d_{11} = a_{11}$$

for $j = 2, \cdots, n$,

$$l_{j1} = a_{1j}/d_{11}.$$

- For each i, $i = 2, 3, \cdots, n$

$$d_{ii} = a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 d_{kk}$$

and

$$l_{ji} = \frac{a_{ij} - \sum_{k=1}^{i-1} l_{ik} l_{jk} d_{kk}}{d_{ii}}$$

for each $j = i+1, i+2, \cdots, n$.

LDL^T Algorithm

To factor the positive definite $n \times n$ matrix $\mathbf{A}$ into $\mathbf{LDL}^T$, where $\mathbf{L}$ is lower triangular with 1s along the diagonal:

INPUT the dimension n; entries a_{ij}, for $i = 1, 2, \cdots, n, j = 1, 2, \cdots, i$ of $\mathbf{A}$.

OUTPUT the entries l_{ij}, for $j = 1, 2, \cdots, i-1$ and $i = 1, 2, \cdots, n$ of $\mathbf{L}$, and d_i for $1 \le i \le n$ of $\mathbf{D}$.

Step 1 Set $d_{11} = a_{11}$.

Step 2 For $j = 2, \cdots, n$, set $l_{j1} = a_{1j}/d_{11}$.

Step 3 For $i = 2, \cdots, n$, do Steps 4 and 5.

Step 4 Set $d_{ii} = a_{ii} - \sum_{j=1}^{i-1} l_{ij}^2 d_{jj}$.

Step 5 For $j = i+1, i+2, \cdots, n$, set

$$l_{ji} = \frac{a_{ij} - \sum_{k=1}^{i-1} l_{ik} l_{jk} d_{kk}}{d_{ii}}$$

Step6 OUTPUT l_{ij} for $1 \le j \le i-1, 1 \le i \le n$ and $d_{ii}, 1 \le i \le n$.STOP!

- To solve the linear system $\mathbf{Ax} = \mathbf{b}$ with the $\mathbf{LDL}^T$ factorization method of $\mathbf{A}$.
- Let

$$\mathbf{Ly} = \mathbf{b}, \mathbf{Dz} = \mathbf{y}, \mathbf{L}^T\mathbf{x} = \mathbf{z}.$$

- Solve the linear system $\mathbf{LY} = \mathbf{b}$ with listed algorithm

  Step 1 Set $y_1 = b_1$.

  Step 2 Set $y_i = b_i - \sum_{j=1}^{i-1} l_{ij} y_j, i = 2, \cdots, n.$

- Solve $\mathbf{Dz} = \mathbf{y}$, Set

$$z_i = \frac{y_i}{d_{ii}}, \; i = 1, 2, \cdots, n.$$

- Finally, solve $\mathbf{L}^T\mathbf{x} = \mathbf{z}$.

  Step 4 Set $x_n = z_n$.

  Step 5 Set $x_i = z_i - \sum_{j=i+1}^{n} l_{ji} x_j, \; i = n-1, \cdots, 1$

  Step 6 OUTPUT $(x_i, i = 1, \cdots, n)$; STOP.

Using Choleski $\mathbf{L}\mathbf{L}^T$ factorization to solve $\mathbf{A}\mathbf{x} = \mathbf{b}$

- This system $\mathbf{A}\mathbf{x} = \mathbf{L}\mathbf{L}^T\mathbf{x} = \mathbf{b}$ can be factorized into two subsystem: $\mathbf{L}\mathbf{y} = \mathbf{b}, \mathbf{L}^T\mathbf{x} = \mathbf{y}$.

  Step 1 Set $y_1 = b_1/l_{11}$.

  Step 2 For $i = 2, 3, \cdots, n$, set

$$y_i = (b_i - \sum_{j=1}^{i-1} l_{ij} y_j)/l_{ii}.$$

  Step 3 Set $x_n = y_n/l_{nn}$.

  Step 4 For $i = n-1, \cdots, 1$, set

$$x_i = (y_i - \sum_{j=i+1}^{n} l_{ji} x_j)/l_{ii}.$$

  Step 5 OUTPUT $(x_i, i = 1, \cdots, n)$; STOP.

Definition 6.28

An $n \times n$ matrix $\mathbf{A}$ is called a **band matrix**(带状矩阵), if integers p and q, with $1 < p, q < n$, exist having the property that $a_{ij} = 0$ whenever $i + p \leq j$ or $j + q \leq i$. The **bandwidth**（带宽）of a band matrix is defined as $w = p + q - 1$.

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1p} & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ a_{q1} & & \ddots & & \ddots & 0 \\ 0 & \ddots & & \ddots & & a_{n-p+1,n} \\ \vdots & \ddots & \ddots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n-q+1} & \cdots & a_{nn} \end{bmatrix}$$

- The matrix of bandwidth 3, occurring when $p = q = 2$, and is called **tridiagonal–三对角矩阵** with the form

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & \cdots & \cdots & 0 \\ a_{21} & a_{22} & a_{23} & \ddots & & & \vdots \\ 0 & a_{32} & a_{33} & a_{34} & \ddots & \vdots & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & a_{n-1.n} \\ 0 & \cdots & \cdots & \cdots & 0 & a_{n.n-1} & a_{nn} \end{bmatrix}$$

- The matrix of bandwidth 5, occurring when $p = q = 3$.

LU Factorization for Tridiagonal Matrix

Suppose that the matrices $\mathbf{L}, \mathbf{U}$ can be found in the form

$$\mathbf{L} = \begin{bmatrix} l_{11} & 0 & \cdots & \cdots & \cdots & 0 \\ l_{21} & l_{22} & \ddots & & & \vdots \\ 0 & l_{32} & l_{33} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & l_{n,n-1} & l_{nn} \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} 1 & u_{12} & 0 & \cdots & \cdots & 0 \\ 0 & 1 & u_{23} & \ddots & & \vdots \\ \vdots & 0 & 1 & u_{34} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 & u_{n-1,n} \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix}$$

In similar manner, we can gives that the entries in $\mathbf{L}, \mathbf{U}$,

$$\begin{aligned} a_{11} &= l_{11}, \\ a_{i,i-1} &= l_{i,i-1}, i = 2, 3, 4, \cdots, n \\ a_{ii} &= l_{i,i-1} u_{i-1,i} + l_{i,i}, i = 2, 3, \cdots, n \\ a_{i,i+1} &= l_{ii} u_{i,i+1}, i = 2, 3, \cdots, n \end{aligned}$$

Algorithm for Tridiagonal Linear System with Crout Factorization(LU)

INPUT the dimension n, the entries of $\mathbf{A}$

OUTPUT the solution $x_1, x_2, \cdots, x_n$.

Solving $\mathbf{Lz} = \mathbf{b}$

STEP 1 set $l_{11} = a_{11}; u_{12} = a_{12}/l_{11}, z_1 = a_{1,n+1}/l_{11}$.

STEP 2 For $i = 2, \cdots, , n-1$, set

$$\begin{aligned} l_{i,i-1} &= a_{i,i-1}; \\ l_{ii} &= a_{ii} - l_{i,i-1}u_{i-1,i}; \\ u_{i,i+1} &= a_{i,i+1}/l_{ii} \\ z_i &= (a_{i,n+1} - l_{i,i-1}z_{i-1}/l_{ii}.(\mathbf{Lz} = \mathbf{b}) \end{aligned}$$

STEP 3 set

$$\begin{aligned} l_{n,n-1} &= a_{n,n-1}; \\ l_{nn} &= a_{nn} - l_{n,n-1}u_{n-1,n}; \\ z_n &= (a_{n,n+1} - l_{n,n-1}z_{n-1}/l_{nn}. \end{aligned}$$

Solving $\mathbf{U}\mathbf{x} = \mathbf{z}$

STEP 4 set $x_n = z_n$

STEP 5 For $i = n-1, \cdots, 1$, set

$$x_i = z_i - u_{i,i+1}x_{i+1}.$$

STEP 6 OUTPUT $(x_1, x_2, \cdots, x_n)$,STOP.

思考问题：

- 算法的复杂性估计：每个算法的计算次数（乘除法或加减法）？
- 大型带状矩阵为稀疏矩阵，如何存储才能有效地减少存储空间？以及如何查询？
- 原始矩阵与计算矩阵如何存储才能减少存储空间？
- 当n 足够大时，除了用矩阵的直接分解方法求解方程组，还有没有好的方法？
- 带状矩阵多出现在偏微分方程数值计算方法中，是常见的矩阵形式.