Lecture 5: LU factorization, Cholesky factorization, Gaussian elimination with pivoting

School of Mathematical Sciences, Xiamen University

1. LU factorization

- Definition: Given $\mathbf{A} \in \mathbb{C}^{m\times m}$, an *LU factorization* (if it exists) of $\mathbf{A}$ is a factorization

$$\mathbf{A} = \mathbf{L}\mathbf{U},$$

where $\mathbf{L} \in \mathbb{C}^{m\times m}$ is *unit lower-triangular* and $\mathbf{U} \in \mathbb{C}^{m\times m}$ is *upper-triangular*.

- An approach: find a sequence of unit lower-triangular matrices $\mathbf{L}_k$ such that

$$\mathbf{L}_{m-1}\cdots\mathbf{L}_2\mathbf{L}_1\mathbf{A} = \mathbf{U}$$

and set

$$\mathbf{L} = \mathbf{L}_1^{-1}\mathbf{L}_2^{-1}\cdots\mathbf{L}_{m-1}^{-1}.$$

- A 4×4 example

$$\underset{\mathbf{A}}{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}} \xrightarrow{\mathbf{L}_1} \underset{\mathbf{L}_1\mathbf{A}}{\begin{bmatrix} \times & \times & \times & \times \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix}} \xrightarrow{\mathbf{L}_2} \underset{\mathbf{L}_2\mathbf{L}_1\mathbf{A}}{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix}} \xrightarrow{\mathbf{L}_3} \underset{\mathbf{L}_3\mathbf{L}_2\mathbf{L}_1\mathbf{A}}{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & \mathbf{0} & \boldsymbol{\times} \end{bmatrix}}$$

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

$$\mathbf{L}_1\mathbf{A} = \begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ -4 & & 1 & \\ -3 & & & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ & 1 & 1 & 1 \\ & 3 & 5 & 5 \\ & 4 & 6 & 8 \end{bmatrix}$$

$$\mathbf{L}_2\mathbf{L}_1\mathbf{A} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & -3 & 1 & \\ & -4 & & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ & 1 & 1 & 1 \\ & 3 & 5 & 5 \\ & 4 & 6 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ & 1 & 1 & 1 \\ & & 2 & 2 \\ & & 2 & 4 \end{bmatrix}$$

$$
\mathbf{L}_3\mathbf{L}_2\mathbf{L}_1\mathbf{A} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ & 1 & 1 & 1 \\ & & 2 & 2 \\ & & 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ & 1 & 1 & 1 \\ & & 2 & 2 \\ & & & 2 \end{bmatrix} = \mathbf{U}.
$$

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

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

1.1. General formulas for LU factorization

- Let $\mathbf{u}_k$ denote the kth column of the matrix at the beginning of step k (which matrix? $\mathbf{L}_{k-1} \cdots \mathbf{L}_2 \mathbf{L}_1 \mathbf{A}$).
- The purpose is to eliminate the entries below u_{kk}. To do this we construct the matrix $\mathbf{L}_k$:

$$\mathbf{L}_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -\ell_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -\ell_{mk} & & & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{l}_k & \mathbf{I}_{m-k} \end{bmatrix},$$

where $\boldsymbol{l}_k = \begin{bmatrix} \ell_{k+1,k} & \ell_{k+2,k} & \cdots & \ell_{mk} \end{bmatrix}^\top$ with the *multipliers*

$$\ell_{jk} = \frac{u_{jk}}{u_{kk}}, \quad k+1 \leq j \leq m.$$

Proposition 1

The matrix $\mathbf{L}_k$ can be inverted by negating its subdiagonal entries. We have

$$\mathbf{L}_k^{-1} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \ell_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & \ell_{mk} & & & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{l}_k & \mathbf{I}_{m-k} \end{bmatrix}.$$

Proof. Define the vector

$$\boldsymbol{\ell}_k = \begin{bmatrix} 0 & \cdots & 0 & \ell_{k+1,k} & \cdots & \ell_{mk} \end{bmatrix}^\top.$$

The matrix $\mathbf{L}_k = \mathbf{I} - \boldsymbol{\ell}_k \mathbf{e}_k^*$, where $\mathbf{e}_k$ is the kth column of the identity matrix $\mathbf{I}$. Obviously, $\mathbf{e}_k^* \boldsymbol{\ell}_k = 0$. Therefore, the statement follows from

$$(\mathbf{I} - \boldsymbol{\ell}_k \mathbf{e}_k^*)(\mathbf{I} + \boldsymbol{\ell}_k \mathbf{e}_k^*) = \mathbf{I} - \boldsymbol{\ell}_k \mathbf{e}_k^* \boldsymbol{\ell}_k \mathbf{e}_k^* = \mathbf{I}. \quad \square$$

Proposition 2

The product $\mathbf{L}_1^{-1}\mathbf{L}_2^{-1}\cdots\mathbf{L}_{m-1}^{-1}$, *i.e., the* L *factor* $\mathbf{L}$, *can be formed by collecting the entries* ℓ_{jk} *in the appropriate places. We have*

$$\mathbf{L} = \begin{bmatrix} 1 & & & & \\ \ell_{21} & 1 & & & \\ \ell_{31} & \ell_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{m1} & \ell_{m2} & \cdots & \ell_{m,m-1} & 1 \end{bmatrix}.$$

Proof. It follows from $\mathbf{L}_k^{-1} = \mathbf{I} + \boldsymbol{\ell}_k\mathbf{e}_k^*$ and $\mathbf{e}_k^*\boldsymbol{\ell}_j = 0$ $(\forall\ j \geq k)$ that

$$\mathbf{L}_k^{-1}\mathbf{L}_{k+1}^{-1} = \mathbf{I} + \boldsymbol{\ell}_k\mathbf{e}_k^* + \boldsymbol{\ell}_{k+1}\mathbf{e}_{k+1}^*.$$

Therefore,

$$\mathbf{L} = \mathbf{L}_1^{-1}\mathbf{L}_2^{-1}\cdots\mathbf{L}_{m-1}^{-1} = \mathbf{I} + \boldsymbol{\ell}_1\mathbf{e}_1^* + \boldsymbol{\ell}_2\mathbf{e}_2^* + \cdots + \boldsymbol{\ell}_{m-1}\mathbf{e}_{m-1}^*. \quad \square$$

Remark 3

- *The matrices $\mathbf{L}_k^{-1}$ are never formed and multiplied explicitly.*
- *The multipliers ℓ_{jk} are computed and stored directly into $\mathbf{L}$.*

1.2. LU factorization algorithm

Algorithm: LU factorization $\mathbf{A} = \mathbf{LU}$

$\mathbf{U} = \mathbf{A}, \quad \mathbf{L} = \mathbf{I}$
for $k = 1$ **to** $m - 1$
  for $j = k + 1$ **to** m
    $\ell_{jk} = u_{jk}/u_{kk}$ (compute multipliers)
    $u_{j,k:m} = u_{j,k:m} - \ell_{jk} u_{k,k:m}$ (elimination)
  end
end

- Exercise. Add a few lines to the above algorithm to obtain $\mathbf{y} = \mathbf{L}_{m-1} \cdots \mathbf{L}_2 \mathbf{L}_1 \mathbf{b}$.

1.3. Gaussian elimination for $\mathbf{A}\mathbf{x} = \mathbf{b}$

- $\mathbf{y} = \mathbf{L}_{m-1} \cdots \mathbf{L}_2 \mathbf{L}_1 \mathbf{b}, \quad \mathbf{U}\mathbf{x} = \mathbf{y}$

Algorithm: Back substitution solving $\mathbf{U}\mathbf{x} = \mathbf{y}$

---

for $k = m$ **downto** 1

$$x_k = \left(y_k - \sum_{j=k+1}^{m} u_{kj} x_j \right) \Big/ u_{kk}$$

end

1.4. Lower-triangular linear system $\mathbf{L}\mathbf{y} = \mathbf{b}$.

Algorithm: Forward elimination solving $\mathbf{L}\mathbf{y} = \mathbf{b}$

---

for $k = 1$ **to** m

$$y_k = \left(b_k - \sum_{j=1}^{k-1} \ell_{kj} y_j \right) / \ell_{kk}$$

end

---

2. Cholesky factorization

- Every Hermitian positive definite matrix $\mathbf{A}$ has a factorization

$$\mathbf{A} = \mathbf{LDL}^*,$$

where $\mathbf{L}$ is the unit lower-triangular matrix in its LU factorization $\mathbf{A} = \mathbf{LU}$ and $\mathbf{D}$ is a diagonal matrix with diagonal entries $d_{ii} > 0$.

- Definition: Given $\mathbf{A} \in \mathbb{C}^{m\times m}$, a *Cholesky factorization* (if it exists) of $\mathbf{A}$ is a factorization

$$\mathbf{A} = \mathbf{R}^*\mathbf{R}$$

where $\mathbf{R} \in \mathbb{C}^{m\times m}$ is *upper-triangular.*

Theorem 4

Every Hermitian positive definite matrix $\mathbf{A} \in \mathbb{C}^{m\times m}$ has a unique Cholesky factorization

$$\mathbf{A} = \mathbf{R}^*\mathbf{R},$$

where $\mathbf{R} \in \mathbb{C}^{m\times m}$ is upper-triangular and $r_{jj} > 0$.

Proof. (By induction on the dimension).

It is easy for the case of dimension 1. Assume it is true for the case of dimension $m-1$. We prove the case of dimension m. Let $\alpha = \sqrt{a_{11}}$. We have

$$\begin{aligned}
\mathbf{A} = \begin{bmatrix} a_{11} & \mathbf{w}^* \\ \mathbf{w} & \mathbf{K} \end{bmatrix} &= \begin{bmatrix} \alpha & \mathbf{0} \\ \mathbf{w}/\alpha & \mathbf{I} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{K} - \mathbf{w}\mathbf{w}^*/a_{11} \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{w}^*/\alpha \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \\
&= \begin{bmatrix} \alpha & \mathbf{0} \\ \mathbf{w}/\alpha & \mathbf{I} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{R}}^*\widehat{\mathbf{R}} \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{w}^*/\alpha \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \\
&\qquad \text{(by } \mathbf{K} - \mathbf{w}\mathbf{w}^*/a_{11} \text{ is HPD and the induction hypothesis)} \\
&= \begin{bmatrix} \alpha & \mathbf{0} \\ \mathbf{w}/\alpha & \widehat{\mathbf{R}}^* \end{bmatrix} \begin{bmatrix} \alpha & \mathbf{w}^*/\alpha \\ \mathbf{0} & \widehat{\mathbf{R}} \end{bmatrix} = \mathbf{R}^*\mathbf{R}.
\end{aligned}$$

The first row of $\mathbf{R}$ is uniquely determined by $r_{11} > 0$ and the factorization itself. The uniqueness of $\mathbf{R}$ follows from the induction hypothesis that $\widehat{\mathbf{R}}$ is unique. □

2.1. A 4×4 example

$$\mathbf{A} = \begin{bmatrix} 4 & 4i & 6 & 2 \\ -4i & 5 & -4i & 5-2i \\ 6 & 4i & 17 & 3-8i \\ 2 & 5+2i & 3+8i & 36 \end{bmatrix}$$

- Compute the upper triangular matrix $\mathbf{R}$ row by row

Step 1: $\begin{bmatrix} 2 & & & \\ -2i & 1 & & \\ 3 & & 1 & \\ 1 & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & 2i & 5 \\ & \times & 8 & -8i \\ & \times & \times & 35 \end{bmatrix} \begin{bmatrix} 2 & 2i & 3 & 1 \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}$

Step 2: $\begin{bmatrix} 1 & 2i & 5 \\ \times & 8 & -8i \\ \times & \times & 35 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -2i & 1 & \\ 5 & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 4 & 2i \\ & \times & 10 \end{bmatrix} \begin{bmatrix} 1 & 2i & 5 \\ & 1 & \\ & & 1 \end{bmatrix}$

Step 3: $\begin{bmatrix} 4 & 2i \\ \times & 10 \end{bmatrix} = \begin{bmatrix} 2 & \\ -1i & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & 9 \end{bmatrix} \begin{bmatrix} 2 & 1i \\ & 1 \end{bmatrix}$

Step 4: $9 = 3 \times 3$

The Cholesky factor $\mathbf{R} = \begin{bmatrix} 2 & 2i & 3 & 1 \\ & 1 & 2i & 5 \\ & & 2 & 1i \\ & & & 3 \end{bmatrix}.$

2.2. Algorithm for Cholesky factorization

Algorithm: Cholesky factorization $\mathbf{A} = \mathbf{R}^*\mathbf{R}$

$\mathbf{R{=}triu(A)}$
for $k = 1$ **to** m
  for $j = k + 1$ **to** m
    $r_{j,j:m} = r_{j,j:m} - \overline{r}_{kj} r_{k,j:m} / r_{kk}$
  end
  $r_{k,k:m} = r_{k,k:m} / \sqrt{r_{kk}}$
end

- Exercise: Design an algorithm to compute $\mathbf{R}^*$ column by column.

2.3. Other factorization of HPD matrix

- For any HPD matrix $\mathbf{A}$, there exists a unique HPD matrix $\mathbf{B}$ satisfying

$$\mathbf{A} = \mathbf{B}^2.$$

  $\mathbf{B}$ is called the *square root* of $\mathbf{A}$. (Proof? HPSD case?)

3. Gaussian elimination with partial pivoting (GEPP)

- Partial pivoting: $|u_{ik}| = \max_{k\leq j\leq m} |u_{jk}|$, rows are interchanged.

$$\underset{\text{Pivot selection}}{\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & u_{ik} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \times & \times & \times & \times \end{bmatrix}} \xrightarrow{\mathbf{P}_k} \underset{\text{Row interchange}}{\begin{bmatrix} \times & \times & \times & \times & \times \\ & u_{ik} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \times & \times & \times & \times \\ & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \times & \times & \times & \times \end{bmatrix}} \xrightarrow{\mathbf{L}_k} \underset{\text{Elimination}}{\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix}}$$

- After $m-1$ steps, $\mathbf{A}$ becomes an upper-triangular matrix $\mathbf{U}$:

$$\mathbf{L}_{m-1}\mathbf{P}_{m-1}\cdots\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1\mathbf{A} = \mathbf{U},$$

where $\mathbf{P}_k$ is an elementary permutation matrix ($\mathbf{P}_k = \mathbf{P}_k^\top = \mathbf{P}_k^{-1}$).

Remark 5

Absolute values of all the entries of $\mathbf{L}_k$ in GEPP are ≤ 1 due to the property at step k (after pivot selection and row interchange)

$$|u_{kk}| = \max_{k\leq j\leq m} |u_{jk}|.$$

3.1. A 4×4 Example

- Step 1. Interchange the first and third rows by $\mathbf{P}_1$

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

First elimination by $\mathbf{L}_1$

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

- Step 2. Interchange the second and fourth rows by $\mathbf{P}_2$

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

Second elimination by $\mathbf{L}_2$

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

- Step 3. Interchange the third and fourth rows by $\mathbf{P}_3$

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

Final elimination by $\mathbf{L}_3$

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

- $\mathbf{A}=(\mathbf{L}_3\mathbf{P}_3\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1)^{-1}\mathbf{U}$

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

Is there a unit lower-triangular $\mathbf{L}$? Yes! Let $\mathbf{P}=\mathbf{P}_3\mathbf{P}_2\mathbf{P}_1$. Then
$\mathbf{L}=\mathbf{P}(\mathbf{L}_3\mathbf{P}_3\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1)^{-1}=\mathbf{P}_3\mathbf{P}_2\mathbf{L}_1^{-1}\mathbf{P}_2^{-1}\mathbf{P}_3^{-1}\mathbf{P}_3\mathbf{L}_2^{-1}\mathbf{P}_3^{-1}\mathbf{L}_3^{-1}$.

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

3.2. General formulas for PA = LU

- $\mathbf{U}=\mathbf{L}_{m-1}\mathbf{P}_{m-1}\cdots\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1\mathbf{A}, \quad \mathbf{P}=\mathbf{P}_{m-1}\cdots\mathbf{P}_2\mathbf{P}_1,$
  $\mathbf{L}=\mathbf{P}(\mathbf{L}_{m-1}\mathbf{P}_{m-1}\cdots\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1)^{-1}=\widehat{\mathbf{L}}_1^{-1}\widehat{\mathbf{L}}_2^{-1}\cdots\widehat{\mathbf{L}}_{m-1}^{-1},$
  $\widehat{\mathbf{L}}_k^{-1}=\mathbf{P}_{m-1}\cdots\mathbf{P}_{k+2}\mathbf{P}_{k+1}\mathbf{L}_k^{-1}\mathbf{P}_{k+1}^{-1}\mathbf{P}_{k+2}^{-1}\cdots\mathbf{P}_{m-1}^{-1}.$

Remark 6

The elementary permutation matrix $\mathbf{P}_k$ *in* GEPP *has the form*

$$\mathbf{P}_k = \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{P}}_k \end{bmatrix},$$

where $\widehat{\mathbf{P}}_k \in \mathbb{R}^{(m-k+1)\times(m-k+1)}$ *is an elementary permutation matrix.*

Remark 7

The unit lower triangular matrix $\widehat{\mathbf{L}}_k^{-1}$ *in* GEPP *has the same sparsity pattern as that of* $\mathbf{L}_k^{-1}$. *The sparsity pattern is*

$$\begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \bigstar & \mathbf{I}_{m-k} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 0 & \mathbf{0} \\ \mathbf{0} & \bigstar & \mathbf{0} \end{bmatrix} + \mathbf{I}.$$

The matrix $\widehat{\mathbf{L}}_k^{-1}$ *is equal to* $\mathbf{L}_k^{-1}$ *but with the* $\bigstar$*'s entries permuted.*

Remark 8

By Proposition 2, the product $\widehat{\mathbf{L}}_1^{-1}\widehat{\mathbf{L}}_2^{-1}\cdots\widehat{\mathbf{L}}_{m-1}^{-1}$ *is unit lower triangular, and the matrices* $\widehat{\mathbf{L}}_k^{-1}$ *are never formed and multiplied explicitly.*

Remark 9

The multipliers ℓ_{jk} *are computed and stored in the appropriate places. It is not a one-step process, but a dynamic adjustment one.*

Remark 10

We have the LU *factorization of* $\mathbf{PA} = \mathbf{LU}$ (*note that the permutation matrix* $\mathbf{P}$ *is not known ahead of time*) *where* $\mathbf{P} = \mathbf{P}_{m-1}\cdots\mathbf{P}_2\mathbf{P}_1$,

$$\mathbf{U} = \mathbf{L}_{m-1}\mathbf{P}_{m-1}\cdots\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1\mathbf{A},$$

$$\mathbf{L} = \mathbf{P}(\mathbf{L}_{m-1}\mathbf{P}_{m-1}\cdots\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1)^{-1} = \widehat{\mathbf{L}}_1^{-1}\widehat{\mathbf{L}}_2^{-1}\cdots\widehat{\mathbf{L}}_{m-1}^{-1},$$

$$\widehat{\mathbf{L}}_k^{-1} = \mathbf{P}_{m-1}\cdots\mathbf{P}_{k+2}\mathbf{P}_{k+1}\mathbf{L}_k^{-1}\mathbf{P}_{k+1}^{-1}\mathbf{P}_{k+2}^{-1}\cdots\mathbf{P}_{m-1}^{-1}.$$

3.3. Algorithm for $\mathbf{PA} = \mathbf{LU}$

Algorithm: LU factorization $\mathbf{PA} = \mathbf{LU}$ in GEPP for $\mathbf{Ax} = \mathbf{b}$

$\mathbf{U} = \mathbf{A},\ \mathbf{L} = \mathbf{I},\ \mathbf{P} = \mathbf{I}$
for $k = 1$ **to** $m - 1$
  Select $i \geq k$ to maximize $|u_{ik}|$ (select pivot)
  $p_{k,:} \leftrightarrow p_{i,:}$ (update $\mathbf{P}$)
  $\ell_{k,1:k-1} \leftrightarrow \ell_{i,1:k-1}$ (update $\widehat{\mathbf{L}}_1^{-1}, \widehat{\mathbf{L}}_2^{-1}, \ldots, \widehat{\mathbf{L}}_{k-1}^{-1}$)
  $u_{k,k:m} \leftrightarrow u_{i,k:m}$ (interchange two rows)
  for $j = k + 1$ **to** m
    $\ell_{jk} = u_{jk}/u_{kk}$ (compute multipliers)
    $u_{j,k:m} = u_{j,k:m} - \ell_{jk} u_{k,k:m}$ (elimination)
  end
end

3.4. GEPP for $\mathbf{Ax} = \mathbf{b}$.

- $\mathbf{y} = \mathbf{L}_{m-1}\mathbf{P}_{m-1} \cdots \mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1\mathbf{b}, \quad \mathbf{Ux} = \mathbf{y}$.
- Exercise. Add a few lines to the above algorithm to obtain $\mathbf{y}$.

3.5. Growth factor of GEPP

- Define the *growth factor* for $\mathbf{A}$ as the ratio $\rho = \dfrac{\max_{ij} |u_{ij}|}{\max_{ij} |a_{ij}|}$.

Proposition 11

The growth factor ρ *of* GEPP *applied to any matrix* $\mathbf{A} \in \mathbb{C}^{m\times m}$ *satisfies* $\rho \le 2^{m-1}$.

Proof. See Exercise 22.1. □

- The upper bound in the above proposition is sharp.

  Consider the 5×5 matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} 1 & & & & 1 \\ -1 & 1 & & & 1 \\ -1 & -1 & 1 & & 1 \\ -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 \end{bmatrix}.$$

The L and U factors are given by

$$\mathbf{L} = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ -1 & -1 & 1 & & \\ -1 & -1 & -1 & 1 & \\ -1 & -1 & -1 & -1 & 1 \end{bmatrix},$$

and

$$\mathbf{U} = \begin{bmatrix} 1 & & & & 1 \\ & 1 & & & 2 \\ & & 1 & & 4 \\ & & & 1 & 8 \\ & & & & 16 \end{bmatrix}.$$

The growth factor $\rho = 2^{5-1} = 16$.

It is easy to construct an $m \times m$ matrix such that $\rho = 2^{m-1}$.

4. Gaussian elimination with complete pivoting (GECP)

- Both rows and columns are interchanged
- After $m-1$ steps, $\mathbf{A}$ becomes an upper-triangular matrix $\mathbf{U}$:

$$\mathbf{L}_{m-1}\mathbf{P}_{m-1}\cdots\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1\mathbf{A}\mathbf{Q}_1\mathbf{Q}_2\cdots\mathbf{Q}_{m-1}=\mathbf{U}.$$

Remark 12

We have the LU *factorization (note that the permutation matrices* $\mathbf{P}$ *and* $\mathbf{Q}$ *are not known ahead of time)*

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

where $\mathbf{P}=\mathbf{P}_{m-1}\cdots\mathbf{P}_2\mathbf{P}_1,\quad \mathbf{Q}=\mathbf{Q}_1\mathbf{Q}_2\cdots\mathbf{Q}_{m-1},$

$$\mathbf{U}=\mathbf{L}_{m-1}\mathbf{P}_{m-1}\cdots\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1\mathbf{A}\mathbf{Q}_1\mathbf{Q}_2\cdots\mathbf{Q}_{m-1},$$

$$\mathbf{L}=\mathbf{P}(\mathbf{L}_{m-1}\mathbf{P}_{m-1}\cdots\mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1)^{-1}=\widehat{\mathbf{L}}_1^{-1}\widehat{\mathbf{L}}_2^{-1}\cdots\widehat{\mathbf{L}}_{m-1}^{-1},$$

$$\widehat{\mathbf{L}}_k^{-1}=\mathbf{P}_{m-1}\cdots\mathbf{P}_{k+2}\mathbf{P}_{k+1}\mathbf{L}_k^{-1}\mathbf{P}_{k+1}^{-1}\mathbf{P}_{k+2}^{-1}\cdots\mathbf{P}_{m-1}^{-1}.$$

4.1. Algorithm for $\mathbf{PAQ} = \mathbf{LU}$

Algorithm: LU factorization $\mathbf{PAQ} = \mathbf{LU}$ in GECP

The details are left as an exercise.

4.2. GECP for $\mathbf{Ax} = \mathbf{b}$

- $\mathbf{y} = \mathbf{L}_{m-1}\mathbf{P}_{m-1} \cdots \mathbf{L}_2\mathbf{P}_2\mathbf{L}_1\mathbf{P}_1\mathbf{b}, \quad \mathbf{Uz} = \mathbf{y}, \quad \mathbf{x} = \mathbf{Qz}$

5. Further reading

- Shufang Xu, Li Gao, and Pingwen Zhang
  Numerical Linear Algebra.
  Second Edition, Peking University Press, 2013
- Exercise: Try your best to modify the pseudocode of the algorithms in this lecture to save storage and to improve efficiency.