SI231 - Matrix Computations, Fall 2020-21

Homework Set #2

Prof. Yue Qiu and Prof. Ziping Zhao

Name: Tian Yun **Major:** Master in CS

Student No.: 2019232102 **E-mail:** tianyun@shanghaitech.edu.cn

---

Acknowledgements:

1) Deadline: **2020-10-11 23:59:00**
2) Submit your homework at **Gradescope**. Entry Code: **MY3XBJ**. Also, make sure that your gradescope account is your **school e-mail**. Homework #2 contains two parts, the theoretical part the and the programming part.
3) About the the theoretical part:
   (a) Submit your homework in **Homework 2** in gradescope. Make sure that you have assigned the correct pages for the problems in the outline.
   (b) Your homework should be uploaded in the **PDF** format, and the naming format of the file is not specified.
   (c) No handwritten homework is accepted. You need to use LaTeX. (If you have difficulty in using LaTeX, you are allowed to use **Word** for the first and the second homework to accommodate yourself.)
   (d) Use the given template and give your solution in English. Solution in Chinese is not allowed.
4) About the programming part:
   (a) Submit your codes in **Homework 2 Programming part** in gradescope.
   (b) Details of requirements in programming are listed in remarks of Problem 6, please read it carefully before you start to program.
5) **No late submission is allowed.**

I. General Linear System

Problem 1 (6 points + 9 points)

Let $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ -2 & 4 & -6 & 0 \\ 3 & 1 & 14 & -1 \\ -1 & 7 & -5 & 3 \end{bmatrix} \in \mathbb{R}^{4\times 4}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 3 & 1 & 1 \\ 2 & 2 & 2 & -1 \\ 5 & 5 & 2 & 3 \end{bmatrix} \in \mathbb{R}^{4\times 4}$.

1) For $\mathbf{A}$ and $\mathbf{b} = (-1, 2, 5, 3)^T \in \mathbb{R}^4$, find $\mathcal{N}(\mathbf{A})$, $\mathcal{R}(\mathbf{A})$, then solve $\mathbf{Ax} = \mathbf{b}$.
2) For $\mathbf{B}$ and $\mathbf{b} = (1, 1, 1, 2)^T \in \mathbb{R}^4$, solve the linear equation system $\mathbf{Bx} = \mathbf{b}$ with Gauss Elimination, LU decomposition, and LU decomposition with partial pivoting, respectively. (Although not required, you are highly encouraged to write down your solution procedures in detail.)

Solution.

1) Nullspace is the space of all solutions of $\mathbf{A}x = 0$.

$$\mathbf{A}x = \begin{bmatrix} 1 & 0 & 1 & 2 \\ -2 & 4 & -6 & 0 \\ 3 & 1 & 14 & -1 \\ -1 & 7 & -5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0$$

After simplification:

$$[\mathbf{A}|\mathbf{0}] = \left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{8}{3} & 0 \\ 0 & 1 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\mathbf{A}x = \begin{bmatrix} 1 & 0 & 0 & \frac{8}{3} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{3}{8} \\ -\frac{1}{3} \\ \frac{2}{3} \\ 1 \end{bmatrix} = 0$$

$$N(A) = \text{span}\left(\begin{bmatrix} -\frac{3}{8} \\ -\frac{1}{3} \\ \frac{2}{3} \\ 1 \end{bmatrix} \right)$$

Then the column space of A is the space formed by all column vectors in A. The column space of A is represented by C(A)

$$\mathbf{A}x = \begin{bmatrix} 1 & 0 & 1 & 2 \\ -2 & 4 & -6 & 0 \\ 3 & 1 & 14 & -1 \\ -1 & 7 & -5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -2 \\ 3 \\ -1 \end{bmatrix} = x_2 \begin{bmatrix} 0 \\ 4 \\ 1 \\ 7 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -6 \\ 14 \\ -5 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix}$$

$$\mathcal{R}(A) = \text{span}\left(\begin{bmatrix} 1 \\ -2 \\ 3 \\ -1 \end{bmatrix} \begin{bmatrix} 0 \\ 4 \\ 1 \\ 7 \end{bmatrix} \begin{bmatrix} 1 \\ -6 \\ 14 \\ -5 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix}\right) = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ \frac{13}{8} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{12}{16} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ \frac{1}{4} \end{bmatrix}\right)$$

$A\vec{x} = \vec{b}$ When is there a solution?

The following two propositions are equivalent

1. When $\vec{b} \in C(A)$, $A\vec{x} = \vec{b}$ there a solution
2. When $\text{rank}(A) = \text{rank}(A, \vec{b})$, $A\vec{x} = \vec{b}$ there a solution

The general solution of the system of non-homogeneous linear equations $Ax = b$ is the general solution of the system of homogeneous linear equations$Ax = b$, plus any special solution of the system of non-homogeneous linear equations $Ax = b$:

So the solution of $Ax = b$:

$$x = C \begin{bmatrix} -\frac{3}{8} \\ -\frac{1}{3} \\ \frac{2}{3} \\ 1 \end{bmatrix} + \begin{bmatrix} -\frac{5}{3} \\ \frac{2}{3} \\ \frac{2}{3} \\ 1 \end{bmatrix}$$

2) a) gauss decomposition

$$[\mathbf{B}|b] = \left[\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 1 \\ 2 & 3 & 1 & 1 & 1 \\ 2 & 2 & 2 & -1 & 1 \\ 5 & 5 & 2 & 3 & 2 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \frac{1}{6} \\ 0 & 1 & 0 & 0 & \frac{1}{6} \\ 0 & 0 & 1 & 0 & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ 0 \end{bmatrix}$$

b) LU decomposition

$$\mathbf{B}x = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 3 & 1 & 1 \\ 2 & 2 & 2 & -1 \\ 5 & 5 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 \\ 5 & 5 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & -1 & -5 & 3 \\ 0 & 0 & 6 & -5 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

Multiply the two matrices on the left to the right matrix

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ 0 \end{bmatrix}$$

c) LU decomposition with partial pivoting

$$\mathbf{P_1B} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 3 & 1 & 1 \\ 2 & 2 & 2 & -1 \\ 5 & 5 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 5 & 2 & 3 \\ 2 & 3 & 1 & 1 \\ 2 & 2 & 2 & -1 \\ 1 & 2 & 3 & -1 \end{bmatrix}$$

$$\mathbf{B_2} = \mathbf{L_1P_1B_1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.4 & 1 & 0 & 0 \\ -0.4 & 0 & 1 & 0 \\ -0.2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 5 & 2 & 3 \\ 2 & 3 & 1 & 1 \\ 2 & 2 & 2 & -1 \\ 1 & 2 & 3 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 5 & 2 & 3 \\ 0 & 1 & 0.2 & -0.2 \\ 0 & 0 & 1.2 & -2.2 \\ 0 & 1 & 2.6 & -1.6 \end{bmatrix}$$

$$\mathbf{P_2B_2} = \mathbf{B_2}$$

$$\mathbf{B_3} = \mathbf{L_2P_2B_2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 5 & 2 & 3 \\ 0 & 1 & 0.2 & -0.2 \\ 0 & 0 & 1.2 & -2.2 \\ 0 & 1 & 2.6 & -1.6 \end{bmatrix} = \begin{bmatrix} 5 & 5 & 2 & 3 \\ 0 & 1 & 0.2 & -0.2 \\ 0 & 0 & 1.2 & -2.2 \\ 0 & 0 & 2.4 & -1.4 \end{bmatrix}$$

$$\mathbf{P_3B_3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 5 & 2 & 3 \\ 0 & 1 & 0.2 & -0.2 \\ 0 & 0 & 1.2 & -2.2 \\ 0 & 0 & 2.4 & -1.4 \end{bmatrix}$$

$$\mathbf{B_4} = \mathbf{L_3P_3B_3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -0.5 & 1 \end{bmatrix} \begin{bmatrix} 5 & 5 & 2 & 3 \\ 0 & 1 & 0.2 & -0.2 \\ 0 & 0 & 2.4 & -1.4 \\ 0 & 0 & 1.2 & -2.2 \end{bmatrix} = \begin{bmatrix} 5 & 5 & 2 & 3 \\ 0 & 1 & 0.2 & -0.2 \\ 0 & 0 & 2.4 & -1.4 \\ 0 & 0 & 0 & -1.5 \end{bmatrix}$$

Therefore, $PB = LU$ where

$$L = L_1'^{-1} L_2'^{-1} L_3'^{-1}$$

with $L_3' = L_3$, $L_2' = P_3 L_2 P_3^{-1}$, and $L_1' = P_3 P_2 L_1 P_2^{-1} P_3^{-1}$,

$$U = \begin{bmatrix} 5 & 5 & 2 & 3 \\ 0 & 1 & 0.2 & -0.2 \\ 0 & 0 & 2.4 & -1.4 \\ 0 & 0 & 0 & -1.5 \end{bmatrix}$$

and

$$P = P_3 P_2 P_1$$

In general, for an $n \times n$ matrix A, the LU factorization provided by Gaussian elimination with partial pivoting can be written in the form:

$$\left(L_{n-1}' \cdots L_2' L_1'\right) \left(P_{n-1} \cdots P_2 P_1\right) A = U$$

where $L_i' = P_{n-1} \cdots P_{i+1} L_i P_{i+1}^{-1} \cdots P_{n-1}^{-1}$ If $L = \left(L_{n-1}' \cdots L_2' L_1'\right)^{-1}$ and $P = P_{n-1} \cdots P_2 P_1$, then $PA = LU$

$PB = LU, B = p^{-1}LU, Bx = b, p^{-1}LUx = b, x = U^{-1}L^{-1}Pb$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{6} \\ \frac{1}{6} \\ 0 \end{bmatrix}$$

II. Understanding Various Matrix Decompositions

Problem 2 (10 points)

Consider the following symmetric matrix $\mathbf{A} \in \mathbb{R}^{4\times 4}$,

$$\mathbf{A} = \begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix}.$$

Give the LU decomposition of $\mathbf{A}$. Then describe under which conditions $\mathbf{A}$ is nonsingular, according to the results of LU decomposition.

Solution.

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a & a & a & a \\ 0 & -a+b & -a+b & -a+b \\ 0 & 0 & -b+c & -b+c \\ 0 & 0 & 0 & d-c \end{bmatrix}$$

First of all: a cannot be equal to 0. Under the premise that a is not equal to 0, b cannot be equal to a, and c cannot be equal to b, and d cannot be equal to c. So a is not equal to 0, then a, b, c, d are different from each other

Problem 3 (5 points + 10 points)

1) Consider a 3×3 matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 4 \\ 1 & 5 & 1 \\ 1 & 1 & 8 \end{bmatrix},$$

find the LDM (also called LDU) decomposition of $\mathbf{A}$, i.e., factor $\mathbf{A}$ as $\mathbf{A} = \mathbf{LDM}^T$ (or $\mathbf{A} = \mathbf{LDU}$), where $\mathbf{L} \in \mathbb{R}^{3\times3}$ is lower triangular with unit diagonal entries, $\mathbf{D} \in \mathbb{R}^{3\times3}$ is a diagonal matrix, and $\mathbf{M} \in \mathbb{R}^{3\times3}$ is lower triangular with unit diagonal entries ($\mathbf{U} \in \mathbb{R}^{3\times3}$ is upper triangular with unit diagonal entries).

2) Consider a 3×3 matrix

$$\mathbf{B} = \begin{bmatrix} 8 & 1 & 1 \\ 1 & 5 & 1 \\ 4 & 2 & 2 \end{bmatrix},$$

find the UL decomposition of $\mathbf{B}$, i.e., factor $\mathbf{B}$ as $\mathbf{B} = \mathbf{UL}$, where $\mathbf{U} \in \mathbb{R}^{3\times3}$ is upper triangular with unit diagonal entries and $\mathbf{L} \in \mathbb{R}^{3\times3}$ is lower triangular.

Hint: $\mathbf{B} = \mathbf{PAP}$, where $\mathbf{P}$ is a unit anti-diagonal matrix [1].

Solution.

1) $$\mathbf{A} = \mathbf{LU} = \begin{bmatrix} 2 & 2 & 4 \\ 1 & 5 & 1 \\ 1 & 1 & 8 \end{bmatrix} = \begin{bmatrix} 8 & 1 & 1 \\ 1 & 5 & 1 \\ 4 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & -1 \\ 0 & 0 & 6 \end{bmatrix} = \mathbf{LDU} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -\frac{1}{4} \\ 0 & 0 & 1 \end{bmatrix}$$

2) $$\mathbf{B} = \mathbf{PAP} = \mathbf{PLUP} = (\mathbf{PLP^T})(\mathbf{PUP^T}) = \begin{bmatrix} 8 & 1 & 1 \\ 1 & 5 & 1 \\ 4 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & -1 \\ 0 & 0 & 6 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & -1 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Let $B = PAP$ where P is the permutation matrix with 1 's on the anti-diagonal and 0's elsewhere. Thus $P = P^T = P^{-1}$, and B is orthogonally similar to A

If $A = LU$ is a factorization with lower triangular L having $1'$ s along the diagonal, and U an upper triangular matrix, then:

$$B = \left(PLP^T\right)\left(PUP^T\right)$$

[1]**Anti-diagonal matrix:** An anti-diagonal matrix is a square matrix where all the entries are zero except those on the diagonal going from the lower left corner to the upper right corner, known as the anti-diagonal. For example,

$$\operatorname{adiag}(a_1, \ldots, a_n) = \begin{bmatrix} 0 & 0 & \cdots & 0 & a_1 \\ 0 & 0 & \cdots & a_2 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_n & 0 & \cdots & \cdots & 0 \end{bmatrix},$$

and consequently, unit anti-diagonal matrix means $\operatorname{adiag}(1, \ldots, 1)$, also known as the **exchange matrix** or the **permutation matrix**.

Note that PLP^T is an upper triangular matrix with 1 's along the diagonal, and PUP^T is a lower triangular matrix, so the above is a factorization of the desired form.

$$\mathbf{B} = \left(PLP^T\right)\left(PUP^T\right) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 & 4 \\ 0 & 4 & -1 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\mathbf{B} = \left(PLP^T\right)\left(PUP^T\right) = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ -1 & 4 & 0 \\ 4 & 2 & 2 \end{bmatrix}$$

Problem 4 (7 points + 6 points + 7 points + 5 points)

Given a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, suppose that the LDM (LDU) decomposition of $\mathbf{A}$ exists, prove that

1) the LDM (LDU) decomposition of $\mathbf{A}$ is *uniquely* determined;
2) if $\mathbf{A}$ is a symmetric matrix, then its LDM (LDU) decomposition must be $\mathbf{A} = \mathbf{LDL}^T$, which is called LDL (LDL^T) decomposition in this case;
3) $\mathbf{A}$ is a symmetric and positive definite matrix if and only if its Cholesky decomposition exists (i.e., there exists a matrix $\mathbf{G} \in \mathbb{R}^{n\times n}$ such that $\mathbf{A} = \mathbf{GG}^T$, where $\mathbf{G}$ is lower triangular with *positive* diagonal entries);
4) if $\mathbf{A}$ is a symmetric and positive definite matrix, then its Cholesky decomposition is *uniquely* determined.

Hints:

1) The existence of the LDM (LDU) decomposition implies the non-singularity of the matrix.
2) You can directly utilize the following lemmas,
   - the inverse (if it exists) of a lower (resp. upper) triangular matrix is also lower (resp. upper) triangular;
   - the product of two lower (resp. upper) triangular matrices is lower (resp. upper) triangular;
   - also, if such two lower (resp. upper) triangular matrices have unit diagonal entries, then their product also has unit diagonal entries.

Solution.

1) First of all, we want to prove that the LU decomposition is unique.

   To prove LU is unique. Suppose that

$$L_1U_1 = L_2U_2$$

   where L_1, L_2 are unit lower triangular and U_1, U_2 are upper triangular and invertible. (They must be invertible if L_iU_i is so.) We now rearrange the equation to get

$$L_2^{-1}L_1 = U_2U_1^{-1}$$

   Now the left side is unit lower triangular and the right side is upper triangular, but only the identity matrix satisies both conditions at once. Therefore we have

$$L_2^{-1}L_1 = I = U_2U_1^{-1}$$

   hence

$$L_1 = L_2, U_1 = U_2$$

   proving the LU is uniqueness.

   If the LU decomposition is unique, then we change the upper and lower triangular matrices into triangular matrices with diagonal elements are 1. because the integers are the same, then LDU is unique.

2) Let A be symmetric and $A = LDU'$ be a LU-decomposition of A where D is a diagonal matrix and $u'_{i,i} = 1$ $A = LDU' = (LDU')^T = U^T DL^T.U^T$ is a lower left triangular matrix and L^T is a upper right triangular matrix. This means $(LDU')^T$ is another LU-decomposition of A. But the decomposition is unique which implies $L = U'^T$ and $L^T = U'$. Hence $A = LDL^T$
3) Suppose a matrix A factors as $A = L^* L$. Then

$$\begin{aligned} x^* A x &= x^* L^* L x \\ &= (Lx)^*(Lx) \\ &= \|Lx\|^2 \\ &\geq 0 \end{aligned}$$

This shows that A is positive semidefinite. If we further assume that L is square and triangular with positive real diagonal entries, then L is invertible, so $Lx = 0 \iff x = 0$. In this case, we see that A is positive definite.

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

4) The Cholesky decomposition of a positive-definite matrix A, is a decomposition of the form

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

Let A be our positive definite matrix. Suppose A has Cholesky factorizations $A = R^* R = S^* S$, for R, S upper-triangular matrices with positive diagonal entries. Then we can write

$$\langle Ax, x \rangle = \langle Rx, Rx \rangle = \langle Sx, Sx \rangle$$

Pick $x = e_1$, the first coordinate vector. Then $\langle Ax, x \rangle = A_{11} = \|Re_1\|^2 = \|Se_1\|^2$, and since R, S are upper triangular, this uniquely defines the upper left entry of R, S, so they must be the same. Now, the k th entry in the first row of R is given by

$$\langle Re_k, e_1 \rangle = \frac{1}{\sqrt{A_{11}}} \langle Re_k, Re_1 \rangle = \frac{1}{\sqrt{A_{11}}} \langle Ae_k, e_1 \rangle$$

so we in fact know the whole first row of R and S, and they must be the same. Now, we reduce our A to a new $(n-1) \times (n-1)$ submatrix, which will also be positive definite, and then repeat the procedure to uniquely determine the subsequent rows.

Problem 5 (10 points + 5 points)

Consider matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ in the following form,

$$\mathbf{A} = \begin{bmatrix} b_1 & c_1 & 0 & 0 & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & 0 & 0 \\ 0 & a_3 & b_3 & c_3 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & 0 & \cdots & 0 & a_n & b_n \end{bmatrix},$$

where a_j, b_j, and c_j are non-zero entries. The matrix in such form is known as a **Tridiagonal Matrix** in the sense that it contains three diagonals.

1) LU decomposition is particularly efficient in the case of tridiagonal matrices. Find the LU decomposition of $\mathbf{A}$ (derivation is expected) and try to complete the Algorithm 1.

Algorithm 1: LU decomposition for tridiagonal matrices

Input : Tridiagonal matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$.

Output: LU decomposition of $\mathbf{A}$.

```
v_1 = b_1,
For k = 2 : n
  l_k = a_k / v_{k-1}
   Save l_k in matrix L.
  v_k = b_k - l_k e_{k-1}
   Save v_k, c_k in matrix U.
End
Return L, U
```

$$\begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & 0 \\ & \ddots & \ddots & \ddots & \\ & 0 & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{pmatrix} = \begin{pmatrix} 1 & & & & 0 \\ l_2 & 1 & & & \\ & l_3 & 1 & & \\ & & \ddots & \ddots & \\ 0 & & & l_n & 1 \end{pmatrix} \begin{pmatrix} v_1 & c_1 & & & 0 \\ & v_2 & c_2 & & \\ & & \ddots & \ddots & \\ & & & v_{n-1} & c_{n-1} \\ 0 & & & & v_n \end{pmatrix}$$

2) Consider symmetric tridiagonal matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{and } \mathbf{B} = \begin{bmatrix} a & a & 0 \\ a & a+b & b \\ 0 & b & b+c \end{bmatrix},$$

and give the LU decompositions and the LDL^T (also known as the LDL) decompositions of $\mathbf{A}$ and $\mathbf{B}$ respectively.

Solution.

1) As shown above

2) $$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} = LDL^T = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} a & a & 0 \\ a & a+b & b \\ 0 & b & b+c \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a & a & 0 \\ 0 & b & b \\ 0 & 0 & c \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} a & a & 0 \\ a & a+b & b \\ 0 & b & b+c \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

III. Programming

Problem 6 (5 points + 15 points)

In this problem, we explore the efficiency of the LU method together with the classical linear system solvers we have learnt in linear algebra.

1) Derive the complexity of the LU decomposition. Particularly, how many flops does the LU decomposition require? The corresponding pseudo code (in Matlab) is provided as follows:

```
function [L,U]= Naive_lu(A)
    n = size(A,1)
    L = eye(n)
    U = A
    for k=1:n-1
        for j=k+1:n
            L(j,k)=U(j,k)/U(k,k)
            U(j,k:n)=U(j,k:n)-L(j,k)*U(k,k:n)
        end
    end
    for k=2:n
        U(k,1:k-1)=0
    end
end
```

2) **Programming part:** Randomly generate a non-singular matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ and a vector $\mathbf{b} \in \mathbb{R}^{n\times 1}$, then program the following methods to solve $\mathbf{A}\mathbf{x} = \mathbf{b}$:

- **The inverse method:** Use the inverse of $\mathbf{A}$ to solve the problem, which can be written as,

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}\,.$$

- **Cramer rule:** Suppose $\mathbf{x} = [x_1, \ldots, x_n]^T$, and we denote $\mathbf{A}_{-i}(\mathbf{b})$ the matrix that we replace the i-th column of $\mathbf{A}$ with $\mathbf{b}$. Then we have

$$x_i = \frac{\det(\mathbf{A}_{-i}(\mathbf{b}))}{\det(\mathbf{A})}, i = 1, \ldots, n\,.$$

- **Gauss Elimination:** We perform row operations on the augmented matrix $[\mathbf{A}|\mathbf{b}]$, and use back substitution to obtain the solution $\mathbf{x}$.
- **LU decomposition.** We first find the LU decomposition of $\mathbf{A}$, then we solve $\mathbf{L}\mathbf{y} = \mathbf{b}$ and $\mathbf{U}\mathbf{x} = \mathbf{y}$.

In your homework, you are required to submit the time-consuming plot (**one figure**) of given methods against the size of matrix $\mathbf{A}$ (i.e., n), where $n = 100, 150, \ldots, 1000$ (You can try larger n and see what will happen, but be careful with the memory use of your PC!).

Remarks: (Important!)

- Coding languages are restricted, but do not use any bulit-in function. For example, do not use Matlab functions such as A/b, inv(A) or lu(A). Otherwise, your results will contradict the complexity analysis, and your scores will be discounted. You can implement the simplest version of these methods by yourself.
- When handing in your homework in gradescope, package all your codes into your_student_id+hw2_code.zip and upload. In the package, you also need to include a file named README.txt/md to clearly identify the function of each file.
- Make sure that your codes can run and are consistent with your solutions.

Solution.

1) 1. For the first step $L(j,k) = U(j,k)/U(k,k)$:

$$\frac{[(n-1)+1]*(n-1)}{2} = \frac{n(n-1)}{2}$$

2. For the step $U(j,k:n) = U(j,k:n) - L(j,k)*U(k,k:n)$:

$$2[(n-1)n + (n-2)(n-1) + ... + 2*1] = \sum_{1}^{n} n^2 - n = 2\frac{n(n+1)(2n+1)}{6} - 2\frac{(n+1)n}{2}$$

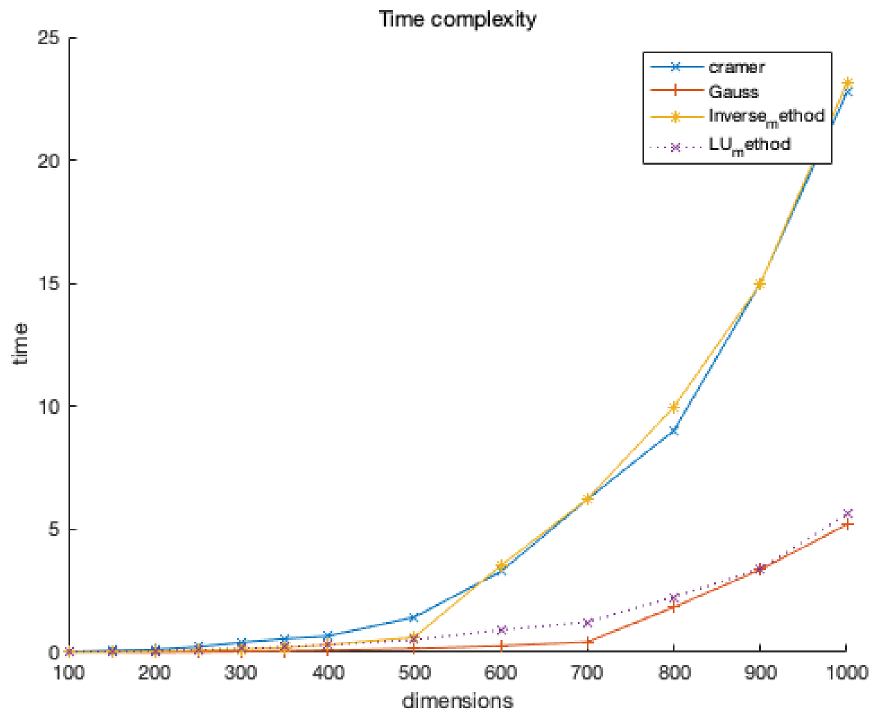
3. For the step $U(k,1:k-1) = 0$:

$$\frac{[(n-1)+1]*(n-1)}{2} = \frac{n(n-1)}{2}$$

$$SUM = \frac{n(n-1)}{2} + 2\frac{n(n+1)(2n+1)}{6} - (n+1)n + \frac{n(n-1)}{2} = \frac{2}{3}n^3 + n^2 - \frac{5}{3}n$$

So the time complexity is $O(\frac{2}{3}n^3)$

2)



IV. Roundoff Error

Problem 7 (Bonus Problem: 10 points + 8 points + 2 points)

Given a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, consider the roundoff error in the process of solving $\mathbf{Ax} = \mathbf{b}$ by Gaussian elimination in three stages:

1. Decompose $\mathbf{A}$ into $\mathbf{LU}$, in a machine with roundoff error $\mathbf{E}$, $\bar{\mathbf{L}}$ and $\bar{\mathbf{U}}$ are computed instead, i.e.,

$$\mathbf{A} + \mathbf{E} = \bar{\mathbf{L}}\bar{\mathbf{U}}\,.$$

2. Solving $\mathbf{Ly} = \mathbf{b}$, numerically with roundoff error $\delta\bar{\mathbf{L}}$, $\hat{\mathbf{y}} = \mathbf{y} + \delta\mathbf{y}$ are computed instead, i.e.,

$$(\bar{\mathbf{L}} + \delta\bar{\mathbf{L}})(\mathbf{y} + \delta\mathbf{y}) = \mathbf{b}\,.$$

3. Solving $\mathbf{Ux} = \mathbf{y}$, numerically with roundoff error $\delta\bar{\mathbf{U}}$, $\hat{\mathbf{x}} = \mathbf{x} + \delta\mathbf{x}$ are computed instead, i.e.,

$$(\bar{\mathbf{U}} + \delta\bar{\mathbf{U}})(\mathbf{x} + \delta\mathbf{x}) = \hat{\mathbf{y}}\,.$$

Finally, we can get the computed solution $\hat{\mathbf{x}}$ and

$$\begin{aligned}\mathbf{b} =& (\bar{\mathbf{L}} + \delta\bar{\mathbf{L}})(\bar{\mathbf{U}} + \delta\bar{\mathbf{U}})(\mathbf{x} + \delta\mathbf{x}) \\ =& (\mathbf{A} + \delta\mathbf{A})(\mathbf{x} + \delta\mathbf{x})\,.\end{aligned}$$

1) Prove that the relative error of $\mathbf{x}$ has an upper bound as follows,

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{1}{1 - \kappa(\mathbf{A})\frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}}\kappa(\mathbf{A})\frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|},$$

where $\kappa(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$ denotes the condition number of matrix $\mathbf{A}$ (Suppose $\mathbf{A}$ and $\mathbf{A}+\delta\mathbf{A}$ are nonsingular and $\|\mathbf{A}^{-1}\|\|\delta\mathbf{A}\| < 1$), and $\|\cdot\|$ can be any norm.

Hint: The following equation might be useful,

$$\|(\mathbf{I}-\mathbf{B})^{-1}\| = \|\sum_{k=0}^{\infty}\mathbf{B}^k\| \leq \sum_{k=0}^{\infty}\|\mathbf{B}\|^k \leq \frac{1}{1-\|\mathbf{B}\|}.$$

where $\mathbf{I}-\mathbf{B}$ is nonsingular and $\lim_{n\to\infty}\mathbf{B}^n = \mathbf{0}$.

2) Consider a linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 10^{-10} & 10^{-10} \\ 1 & 10^{-10} & 10^{-10} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2(1+10^{-10}) \\ -10^{-10} \\ 10^{-10} \end{bmatrix}$$

find the solution $\mathbf{x}$, and calculate the condition number of $\mathbf{A}$ with the matrix infinite norm[2], i.e. $\kappa_\infty(\mathbf{A}) = \|\mathbf{A}\|_\infty\|\mathbf{A}^{-1}\|_\infty$. Suppose $|\delta\mathbf{A}| < 10^{-18}|\mathbf{A}|$[3], use $\kappa_\infty(\mathbf{A})$ to verify that

$$\|\delta\mathbf{x}\| < 10^{-7}\|\mathbf{x}\|.$$

3) Discuss what you have observed from the previous 2 questions. What are the main factors that influence the relative error of the computed solution? Does the ill-conditioned matrix (i.e. the condition number is large) always lead to a large error of the solution?

Solution.

1) Suppose the original equation has undergone some changes after being disturbed

$$(A+\delta A)(x+\delta x) = b+\delta b$$

Because $b = Ax$,So we get:

$$(A+\delta A)\delta x = \delta b - \delta A x$$

Because $\left\|A^{-1}\right\|\|\delta A\| < 1$

$$\left\|\left(I+A^{-1}\delta A\right)^{-1}\right\| \leqslant \frac{1}{1-\|A^{-1}\|\|\delta A\|}$$

$$A+\delta A = A\left(I+A^{-1}\delta A\right)$$

$$\begin{aligned} \delta x &= (A+\delta A)^{-1}(\delta b - \delta A x) \\ &= \left(I+A^{-1}\delta A\right)^{-1}A^{-1}(\delta b - \delta A x) \end{aligned}$$

$$\begin{aligned} \|\delta x\| &\leqslant \left\|\left(I+A^{-1}\delta A\right)^{-1}\right\|\left\|A^{-1}\right\|(\|\delta b\| + \|\delta A\|\|x\|) \\ &\leqslant \frac{\left\|A^{-1}\right\|}{1-\|A^{-1}\|\|\delta A\|}(\|\delta b\| + \|\delta A\|\|x\|) \end{aligned}$$

[2]If $\mathbf{A} \in \mathbb{R}^{n\times n}$, then the matrix infinite norm is $\|\mathbf{A}\|_\infty = \max_{1<i<n}\sum_{j=1}^{n}|a_{i,j}|$.

[3]$|\mathbf{A}| \leq |\mathbf{B}|$ means each element in $\mathbf{A}$ is relative smaller to the corresponding element of $\mathbf{A}$

$$\frac{\|\delta x\|}{\|x\|} \leqslant \frac{\left\|A^{-1}\right\| \|A\|}{1 - \|A^{-1}\| \|\delta A\|} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right)$$

Because $\|b\| \leqslant \|A\| \|x\|$

$$\frac{\|\hat{x} - x\|}{\|x\|} = \frac{\|\delta x\|}{\|x\|} \leq \frac{1}{1 - \kappa(A) \frac{\|\delta A\|}{\|A\|}} \kappa(A) \frac{\|\delta A\|}{\|A\|},$$

2) $\kappa_\infty(\mathbf{A}) = \|\mathbf{A}\|_\infty \|\mathbf{A}^{-1}\|_\infty = 2 * 10^{10}$

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{1}{1 - \kappa(\mathbf{A}) \frac{\|\delta \mathbf{A}\|}{\|\mathbf{A}\|}} \kappa(\mathbf{A}) \frac{\|\delta \mathbf{A}\|}{\|\mathbf{A}\|},$$

$|\delta \mathbf{A}| < 10^{-18} |\mathbf{A}|$

$\frac{|\delta \mathbf{A}|}{|\mathbf{A}|} < 10^{-18}$

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{1}{1 - 10^{-18} * 2 * 10^{10}} 2 * 10^{10} * 10^{-18} < 10^{-7}$$

3)