COMP1021 Mathematics for Computer Science
Linear Algebra (Part 2)
Practical - Week 12

Karl Southern (With thanks to William K. Moses Jr.)

January 2025

Instructions: Work on these problems in the practical sessions for the week specified. First try them on your own. If you're stuck, try discussing things with others. If you get the answer, still discuss with others to see if maybe you missed something. If you run into major roadblocks, ask the demonstrators for hints.

Solutions will be posted on Learn Ultra at the end of the week. Make sure you're all set with the solutions and understand them before the next practical.

Purpose of this practical: This practical is about LU decomposition and all the ideas surrounding it. The goal of this practical is to get you familiar with decomposing a matrix into its LU factorization and subsequently using that form to solve a system of linear equations. This practical also gets you familiar with which matrices can be decomposed into LU form and which cannot. And what to do for matrices that can't be decomposed into LU form.

1. Consider the following matrices A, B, C, D, and F:

$$A = \begin{pmatrix} 2 & 6 & -8 \\ 0 & 5 & 3 \\ 4 & 7 & 9 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 6 & -8 \\ 0 & 5 & 3 \\ 0 & -5 & 25 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 3 & -4 \\ 0 & 5 & 3 \\ 4 & 7 & 9 \end{pmatrix}$$

$$D = \begin{pmatrix} 2 & 6 & -8 \\ 0 & 0 & 28 \\ 0 & -5 & 25 \end{pmatrix}, \quad F = \begin{pmatrix} 4 & 7 & 9 \\ 0 & 5 & 3 \\ 1 & 3 & -4 \end{pmatrix}$$

For each of the following equations, find an elementary matrix E that satisfies the equation:

(a) $EA = B$ (b) $EB = A$ (c) $EA = C$ (d) $EC = A$

(e) $EB = D$ (f) $ED = B$ (g) $EC = F$ (h) $EF = C$

Answer:

(a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(e) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ (g) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (h) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

2. Consider a matrix A that can be decomposed into the LU form. If U is the result of elementary matrices E_1, E_2, …, E_k being applied to A in that order, then what is L equal to in terms of these elementary matrices and why?

   Answer: $L = E_1^{-1}E_2^{-1}E_3^{-1}$. Why? Because $U = E_3E_2E_1A$, so left multiplying E_3^{-1}, then E_2^{-1}, then E_1^{-1}, we see that $E_1^{-1}E_2^{-1}E_3^{-1}U = A$.

3. Solve the linear system $A\mathbf{x} = \mathbf{b}$ by using the LU method, where

$$A = \begin{pmatrix} 3 & -6 & -3 \\ 2 & 0 & 6 \\ -4 & 7 & 4 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} -3 \\ -22 \\ 3 \end{pmatrix}.$$

   Answer: First find an LU-decomposition of A.

Transforming A to U	Building L from I
$\begin{pmatrix} 3 & -6 & -3 \\ 2 & 0 & 6 \\ -4 & 7 & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
$1/3 \times R_1$	
$\begin{pmatrix} 1 & -2 & -1 \\ 2 & 0 & 6 \\ -4 & 7 & 4 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
add $(-2) \times R_1 \to R_2$ add $4 \times R_1 \to R_3$	
$\begin{pmatrix} 1 & -2 & -1 \\ 0 & 4 & 8 \\ 0 & -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 & 0 \\ 2 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}$
$1/4 \times R_2$	
$\begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \\ 0 & -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 & 0 \\ 2 & 4 & 0 \\ -4 & 0 & 1 \end{pmatrix}$
add $1 \times R_2 \to R_3$	
$\begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 & 0 \\ 2 & 4 & 0 \\ -4 & -1 & 1 \end{pmatrix}$
$1/2 \times R_3$	
$\begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 3 & 0 & 0 \\ 2 & 4 & 0 \\ -4 & -1 & 2 \end{pmatrix}$

So, the decomposition is (don't forget to quickly check it by multiplying L and U)

$$A = \begin{pmatrix} 3 & -6 & -3 \\ 2 & 0 & 6 \\ -4 & 7 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 2 & 4 & 0 \\ -4 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = LU$$

Now, re-write $A\mathbf{x} = \mathbf{b}$ as $LU\mathbf{x} = \mathbf{b}$. Denote $U\mathbf{x} = \mathbf{y}$ and solve $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$:

$$\begin{array}{rrrl} 3y_1 & & & = -3 \\ 2y_1 & +4y_2 & & = -22 \\ -4y_1 & -y_2 & +2y_3 & = 3 \end{array}$$

Solving by forward substitution, we get $y_1 = -1, y_2 = -5, y_3 = -3$.

Now solve $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$.

$$\begin{array}{rrrl} x_1 & -2x_2 & -x_3 & = -1 \\ & x_2 & +2x_3 & = -5 \\ & & x_3 & = -3 \end{array}$$

By backward substitution, get $x_3 = -3, x_2 = 1, x_1 = -2$. (Don't forget to check this solution by substituting into the original system.)

4. We talked in class about why using LU decomposition is preferable to Gaussian elimination when you must solve multiple systems of linear equations using the same linear mapping. To really drive that home, actually calculate the number of operations it would take to use Gaussian elimination to solve a system of equations using an $n \times n$ linear mapping. Let's assume that that each of the following constitute a single operation: (i) multiplication of two numbers (ii) division of two numbers (iii) addition of two numbers (iv) subtraction of two numbers. **A loose (but not too loose) upper bound on the number of operations is fine.** (Why is this fine? Because what we truly care about is the asymptotic time to utilize this method.) Subsequently, calculate how many operations it would take to compute the LU decomposition of the same $n \times n$ linear mapping, and then calculate how many operations it would take to solve a linear system using the computed LU decomposition.

   Answer: Consider using Gaussian elimination to solve $A\mathbf{x} = \mathbf{b}$. We will first count the number of operations we need to perform on A and then add in the extra operations performed on b at the end. To get A into the row echelon form, we first make the leading coefficient of row one 1 through a division and process the remaining $n-1$ numbers. So n operations. Then for each of the $n-1$ rows below, for each element in that row, we take the product of the first row's element with some constant and subtract it from that element, resulting in $2n$ operations per row. Totally we have $2n(n-1)+n < 2n^2$ operations. For the second row, there might be one row swap and the subsequent cost of repeating the above operations is $2n(n-1)$. For row i, aside from the possible row swap, the cost is at most $2n(n-i+1)$. So total cost to process all n rows would be upper bounded by $\sum_{i=1}^{n}(2n(n-i+1)+1) \leq 2n^3$. Applying all the above described operations to the n elements of $\mathbf{b}$, we see that we require an additional at most n^2 operations.

   For LU decomposition of a matrix A, we again apply Gaussian elimination to A to get U, resulting in at most $2n^3$ operations. We get the corresponding series of elementary matrices E_1 to E_k, $k \leq 2n$. Calculating an E^-1 for a given E requires one operation, and since there are at most $2n$ elementary matrices, that's a total of at most $2n$ operations. Due to the nature of the elementary operations used to get U, once we calculate their inverses, we immediately have the values of the matrix L (so no extra operations needed). (Why is this true? The best way to build your intuition is to play around with multiplying several elementary matrices together. What sequences of elementary operations would make this not true? For what sequences would this be true?) Totally we take some $2n^3 + 2n \leq 3n^3$ operations.

   Once we have an LU decomposition, let's look at the number of operations to solve $A\mathbf{x} = \mathbf{b}$. recall that we substitute $U\mathbf{x} = \mathbf{y}$ and then solve $L\mathbf{y} = \mathbf{b}$. Solving this takes at most n^2 operations (actually much less). Why? It's a triangular system. Subsequently solving $U\mathbf{x} = \mathbf{y}$ also takes at most n^2 operations for similar reasons for a total of at most $2n^2$ operations.

   Again, what was the point of all of this? To observe the asymptotic relationship between Gaussian elimination and LU decomposition. Let's say that the cost of using Gaussian elimination to solve a system of equations is T_G, the cost of performing LU decomposition on a matrix is $T_{LU-decompose}$, and the cost of using a pre-existing LU decomposition of A to solve a system of equations is L_{LU-use}. We see that $T_G = \Theta(T_{LU-decompose})$ and $T_{LU-use} = o(T_G)$. If we want to solve a single system of equations, either method takes the same amount of time asymptotically. But if we have many systems to solve using the same A, then it's much better to use LU decomposition since we only have to decompose A once and the cost of using it is much less than that of using Gaussian elimination.

5. Prove that the matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has no LU decomposition.

   Answer:
   Hint: If the matrix did have an LU decomposition, what would be the implications on the actual values in the LU matrices.

   Proof: For contradiction, assume that

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \ell_{11} & 0 \\ \ell_{21} & \ell_{22} \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{pmatrix}.$$

   Then we have $\ell_{11}u_{11} = 0$, $\ell_{11}u_{12} = 1$ and $\ell_{21}u_{11} = 1$. The first equality means that at least one of ℓ_{11} and u_{11} is 0, which would contradict at least one of the other two equations.

6. Recall the theorem we saw in class: Let A be a square matrix and let U be its (non-reduced) row echelon form, obtained by Gaussian elimination. If A and U are as above and **no row exchanges** were performed while obtaining U from A, then A can be factored $A = LU$, where L is lower triangular.

   Let's build some insight into why this is true.

   (a) Consider the elementary matrix E corresponding to the elementary row operation $R_3 = R_3 - 2R_1$. For this elementary matrix, what is its inverse? Find the inverses for different types of elementary matrices to see if you notice a pattern.
   Answer:

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}, E^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

   Pattern: If row addition by a, then inverse has $-a$ in same position. E.g.,

$$E = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

   Pattern: If row multiplication by a, then inverse has $1/a$ in same position. E.g.,

$$E = \begin{pmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E^{-1} = \begin{pmatrix} 1/a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

   Pattern: Elementary matrix E corresponds to row switching, then $E = E^{-1}$.

   (b) Consider the product of two lower triangular matrices. Can you say something about whether the product is also lower triangular or not?
   Answer: It is also a lower triangular matrix.

   (c) Recall that applying Gaussian elimination to a matrix A to get an upper triangular matrix U corresponds to left multiplying A by a sequence of elementary matrix operations, say E_1 to E_k. Using this and what you've seen so far, can you argue about why the theorem is true.
   Answer: Notice that all elementary row operations to get A to U are only of the type corresponding to row addition and multiplication of a row. Hence all the corresponding

elementary matrices are lower triangular. By the patterns we observed, we see that their inverses are also lower triangular. We finally noticed that the product of two lower triangular matrices is also a lower triangular matrix. Hence, if no row switches are used, we're guaranteed to get a lower triangular matrix if we multiply the inverses of the elementary matrices together.

(d) With regards to the theorem, what exactly is the problem when one of elementary matrices corresponds to a row exchange?

Answer: The elementary matrix corresponding to a row exchange is not lower triangular. The inverse of such a matrix is itself, and as such not lower triangular. So we don't have the guarantee that that the product of inverses will be a lower triangular matrix.

7. Implement the LU decomposition algorithm in python.