Matrix Computations	**Fall 2020**

Lecture 2: LU decomposition

Lecturer: Prof.Yue Qiu & Prof Ziping Zhao — *Scribe: Xinyue Zhang*

1 Preliminary

Lemma 1. *The inverse of a lower(upper) triangular matrix is again lower(upper) triangular.*

Proof. Suppose $\mathbf{L}^{-1} = [\boldsymbol{y}_1, ..., \boldsymbol{y}_n]$, then $\mathbf{L}\mathbf{L}^{-1} = \mathbf{I} = [\mathbf{e}_1, ...\mathbf{e}_n]$. Since $\mathbf{L}\mathbf{L}^{-1} = \mathbf{L}\,[\boldsymbol{y}_1, ..., \boldsymbol{y}_n]$, we have $\mathbf{L}\boldsymbol{y}_i = \mathbf{e}_i$ for $i = 1, ..., n$. Notice that the first $i-1$ entries of $\mathbf{e}_i$ are 0s and $\mathbf{L}$ is lower-triangular, then the first $i-1$ entries of $\boldsymbol{y}_i$ must be 0s which means $\mathbf{L}^{-1}$ is lower-triangular. The proof for upper-triangular matrix is similar. $\square$

Lemma 2. *The product of 2 lower(upper) triangular matrices is also lower(upper) triangular.*

Definition 1. ***Gauss transformation matrix*** *is an $n \times n$ triangular matrix of the form $\mathbf{M}_k = \mathbf{I} - \boldsymbol{\tau}_k \mathbf{e}_k^T$, where $\boldsymbol{\tau}_k = [0, ..., \mu_{k+1}, ..., \mu_n]^T$ is a column with zeros in the first k entries. By observing that $\mathbf{e}_k^T \boldsymbol{\tau}_k = 0$, we have $\mathbf{M}_k^{-1} = \mathbf{I} + \boldsymbol{\tau}_k \mathbf{e}_k^T$. And $\boldsymbol{\tau}$ is called Gauss vector.*

In particular, given $\boldsymbol{x} \in \mathbb{R}^n$ and $x_k \neq 0$, for $k = 1, ..., n$, we can construct $\mathbf{M}_k$ such that

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -\frac{x_{k+1}}{x_k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -\frac{x_n}{x_k} & & & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{1}$$

$$\begin{aligned} \mathbf{M}_1^{-1}\mathbf{M}_2^{-1}...\mathbf{M}_{n-1}^{-1} &= (\mathbf{I} + \boldsymbol{\tau}_1\mathbf{e}_1^T)(\mathbf{I} + \boldsymbol{\tau}_2\mathbf{e}_2^T)...(\mathbf{I} + \boldsymbol{\tau}_{n-1}\mathbf{e}_{n-1}^T) & (2) \\ &= \mathbf{I} + \boldsymbol{\tau}_1\mathbf{e}_1^T + \boldsymbol{\tau}_2\mathbf{e}_2^T + ... + \boldsymbol{\tau}_{n-1}\mathbf{e}_{n-1}^T. & (3) \end{aligned}$$

Definition 2. ***Permutation matrix*** *is a square matrix with exactly one entry of 1 in each row and column and 0 elsewhere.*

Properties:

1. Permutation matrix $\mathbf{P}$ is orthogonal, that is $\mathbf{P}^T\mathbf{P} = \mathbf{P}\mathbf{P}^T = \mathbf{I}$.
2. $\mathbf{P}^{-1} = \mathbf{P}^T$.
3. $\mathbf{P}_1$ and $\mathbf{P}_2$ are 2 permutation matrix, and $\mathbf{P}_1\mathbf{P}_2$ is again a permutation matrix.

2 Direct methods for general linear system

Suppose $\mathbf{A} \in \mathbb{R}^{n \times n}$ has an LU decomposition. Then $\mathbf{A}\boldsymbol{x} = \mathbf{b}$ is equivalent to the following 2 triangular systems $\mathbf{L}\boldsymbol{y} = \mathbf{b}$ and $\mathbf{U}\boldsymbol{x} = \boldsymbol{y}$. First, the lower-triangular system $\mathbf{L}\boldsymbol{y} = \mathbf{b}$ is solved for $\boldsymbol{y}$ by *forward substitution*

$$y_1 = b_1 \quad \text{and} \quad y_i = b_i - \sum_{k=1}^{i-1} \ell_{ik} y_k \quad \text{for} \quad i = 2, ..., n. \tag{4}$$

After $\boldsymbol{y}$ is known, the upper-triangular system $\mathbf{U}\boldsymbol{x} = \boldsymbol{y}$ is solved for $\boldsymbol{x}$ by *back substitution*

$$x_n = \frac{y_n}{u_{nn}} \quad \text{and} \quad x_i = \frac{1}{u_{ii}}\Big(y_i - \sum_{k=i+1}^{n} u_{ik} x_k\Big) \quad \text{for} \quad i = n-1, ..., 1. \tag{5}$$

3 LU decomposition

LU decomposition

If $\mathbf{A}$ is an $n \times n$ matrix such that a zero pivot is never encountered when applying Gaussian elimination with adding to one row a scalar multiple of another, then $\mathbf{A}$ can be factored as the product $\mathbf{A} = \mathbf{LU}$, where the following hold.

1. $\mathbf{L}$ is lower triangular and $\mathbf{U}$ is upper triangular.
2. $\ell_{ii} = 1$ and $u_{ii} \neq 0$ for $i = 1, 2, ..., n$.
3. Below the diagonal of $\mathbf{L}$, the entry ℓ_{ii} is the multiple of row j that is subtracted from row i in order to annihilate the (i, j) -position during Gaussian elimination.
4. $\mathbf{U}$ is the final result of Gaussian elimination applied to $\mathbf{A}$.
5. The matrices $\mathbf{L}$ and $\mathbf{U}$ are uniquely determined by properties 1 and 2.

The decomposition of $\mathbf{A}$ into $\mathbf{A} = \mathbf{LU}$ is called the $\mathbf{LU}$ decomposition of $\mathbf{A}$, and the matrices $\mathbf{L}$ and $\mathbf{U}$ are called the LU factors of $\mathbf{A}$.

3.1 Existence

Theorem 3. *If $\mathbf{A}$ is an $n \times n$ matrix and every leading principal submatrix $\mathbf{A}_k$ is non-singular ($\det(\mathbf{A}(1:k, 1:k)) \neq 0$) for $k = 1 : n-1$, then there exists a unit lower triangular $\mathbf{L}$ and an upper triangular $\mathbf{U}$ such that $\mathbf{A} = \mathbf{LU}$.*

Proof. $\Leftarrow$: Assume that A possesses an LU decomposition and partition A as

$$\mathbf{A} = \mathbf{LU} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{0} \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{0} & \mathbf{U}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{11}\mathbf{U}_{11} & \star \\ \star & \star \end{bmatrix} \tag{6}$$

where $\mathbf{L}_{11}$ and $\mathbf{U}_{11}$ are each $k \times k$. Thus $\mathbf{A}_k = \mathbf{L}_{11}\mathbf{U}_{11}$ must be non-singular because $\mathbf{L}_{11}$ and $\mathbf{U}_{11}$ are each non-singular—they are triangular with nonzero diagonal entries.

$\Rightarrow$: Use induction to prove that each $\mathbf{A}_k$ possesses an LU decomposition. For $k = 1$, this statement is clearly true because if $\mathbf{A}_1 = (a_{11})$ is non-singular, then $\mathbf{A}_1 = (1)(a_{11})$ is its LU decomposition. Now assume that $\mathbf{A}_k$ has an LU decomposition and show that this together with the non-singularity condition implies $\mathbf{A}_{k+1}$ must also possess an LU decomposition. If $\mathbf{A}_k = \mathbf{L}_k\mathbf{U}_k$ is the LU decomposition for $\mathbf{A}_k$, then $\mathbf{A}_k^{-1} = \mathbf{U}_k^{-1}\mathbf{L}_k^{-1}$ so that

$$\mathbf{A}_{k+1} = \begin{bmatrix} \mathbf{A}_k & \mathbf{b} \\ \mathbf{c}^T & \alpha_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_k & \mathbf{0} \\ \mathbf{c}^T\mathbf{U}_k^{-1} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{U}_k & \mathbf{L}_k^{-1}\mathbf{b} \\ \mathbf{0} & \alpha_{k+1} - \mathbf{c}^T\mathbf{A}_k^{-1}\mathbf{b} \end{bmatrix}, \tag{7}$$

where $\mathbf{c}^T$ and $\mathbf{b}$ contain the first k components of $\mathbf{A}_{k+1,*}$ and $A_{*,k+1}$, respectively. Observe that this is the LU decomposition for $\mathbf{A}_{k+1}$ because

$$\mathbf{L}_{k+1} = \begin{bmatrix} \mathbf{L}_k & \mathbf{0} \\ \mathbf{c}^T\mathbf{U}_k^{-1} & 1 \end{bmatrix} \text{ and } \mathbf{U}_{k+1} = \begin{bmatrix} \mathbf{U}_k & \mathbf{L}_k^{-1}\mathbf{b} \\ \mathbf{0} & \alpha_{k+1} - \mathbf{c}^T\mathbf{A}_k^{-1}\mathbf{b} \end{bmatrix} \tag{8}$$

are lower- and upper-triangular matrices, respectively, and $\mathbf{L}$ has 1's on its diagonal while the diagonal entries of $\mathbf{U}$ are nonzero. The fact that

$$\alpha_{k+1} - \mathbf{c}^T\mathbf{A}_k^{-1}\mathbf{b} \neq 0 \tag{9}$$

follows because $\mathbf{A}_{k+1}$ and $\mathbf{L}_{k+1}$ are each non-singular, so $\mathbf{U}_{k+1} = \mathbf{L}_{k+1}^{-1}\mathbf{A}_{k+1}$ must also be non-singular. Therefore, the non-singularity of the leading principal submatrices implies that each $\mathbf{A}_k$ possesses an LU decomposition, and hence $\mathbf{A}_n = \mathbf{A}$ must have an LU decomposition. $\square$

Proposition 1. *If every leading principal submatrix is non-singular, then no zero pivots encountered.*

Proof. Suppose $k-1$ steps in LU decomposition have been executed. And we have $\mathbf{M}_{k-1}...\mathbf{M}_1\mathbf{A} = \mathbf{A}^{(k-1)}$. Then $\det(\mathbf{M}_{k-1}...\mathbf{M}_1\mathbf{A}) = \det(A)$, so $\det(\mathbf{A}(1:k, 1:k)) = a_{11}^{(k-1)}...a_{kk}^{(k-1)}$. If the leading principal submatrix is non-singular, then the kth pivot $a_{kk}^{(k-1)}$ is nonzero. $\square$

3.2 Uniqueness

Theorem 4. *If A is non-singular and every leading principal submatrix of it is non-singular, then the LU decomposition is unique.*

Proof. Since LU factors have nonzero diagonals, then they must be non-singular. Suppose there exists 2 LU decompositions for a non-singular matrix $\mathbf{A}$, we have

$$L_1U_1 = A = L_2U_2,$$

$$L_2^{-1}L_1 = U_2U_1^{-1}. \tag{10}$$

From lemma 1 and lemma 2, we know $L_2^{-1}L_1$ is lower triangular and $U_2U_1^{-1}$ is upper triangular, Equation (10) implies $L_2^{-1}L_1 = D = U_2U_1^{-1}$ where D is a diagonal matrix. However, the diagonal entries of L_1 and L_2 are ones, so it must be the case that $L_2^{-1}L_1 = I = U_2U_1^{-1}$, that is $L_1 = L_2$ and $U_1 = U_2$. $\square$

4 Pivoting

When row interchanges are allowed, zero pivots can always be avoided when the original matrix $\mathbf{A}$ is non-singular. Consequently, we may conclude that for every non-singular matrix $\mathbf{A}$, there exists a permutation matrix $\mathbf{P}$ (a product of elementary interchange matrices) such that $\mathbf{PA}$ has an LU decomposition.

1. Partial pivoting: do row or column interchange so that no multiplier is greater than 1 in absolute value.
2. Complete pivoting: do row and column interchange so that no multiplier is greater than 1 in absolute value.

5 The LDM decomposition

Suppose the square matrix $\mathbf{A}$ has LU decomposition $\mathbf{A} = \mathbf{LU}$, then the LDM decomposition can be written as $\mathbf{A} = \mathbf{LDM}$, where $\mathbf{L}$ and $\mathbf{U}$ are lower- and upper-triangular matrices with 1's on both of their diagonals. $\mathbf{U} = \mathbf{DM}$ is easily remedied by factoring the diagonal entries out of the upper factor as shown below:

$$\begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix} = \begin{bmatrix} u_{11} & 0 & \dots & 0 \\ 0 & u_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} 1 & u_{12}/u_{11} & \dots & u_{1n}/u_{11} \\ 0 & 1 & \dots & u_{2n}/u_{22} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}. \tag{11}$$

It is uniquely determined, and when $\mathbf{A}$ is symmetric, the LDU decomposition is $\mathbf{A} = \mathbf{LDL}^T$.

6 Exercise

6.1 Find LU decomposition by Gauss Elimination

Problem: Do LU decomposition on the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{bmatrix} \tag{12}$$

Solution:

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{bmatrix} \xrightarrow[row3-3*row1]{row2-2*row1} \begin{bmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 12 & 16 \end{bmatrix} \xrightarrow{row3-4*row2} \begin{bmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{bmatrix} = \mathbf{U}. \tag{13}$$

Then we can write the row operations in matrix representation $\mathbf{N}_3\mathbf{N}_2\mathbf{N}_1\mathbf{A} = \mathbf{U}$, and

$$\mathbf{N}_3\mathbf{N}_2\mathbf{N}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -4 & 1 \end{bmatrix}. \tag{14}$$

So that $\mathbf{A} = \mathbf{N}_1^{-1}\mathbf{N}_2^{-1}\mathbf{N}_3^{-1}\mathbf{U} = \mathbf{LU}$, where

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}. \tag{15}$$

6.2 Use partial pivoting

Problem: Use partial pivoting on the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -3 & 4 \\ 4 & 8 & 12 & -8 \\ 2 & 3 & 2 & 1 \\ -3 & -1 & 1 & -4 \end{bmatrix} \tag{16}$$

Solution: Note that the components from **L** and **U** are overwritten in **A**, and the multipliers ℓ_{ij} are shown in boldface. Adjoin a "permutation counter column" **p** that is initially set to the natural order 1,2,3,4. Permuting components of **p** as the various row interchanges are executed will accumulate the desired permutation. The matrix **P** is obtained by executing the final permutation residing in **p** to the rows of an appropriate size identity matrix:

$$[\mathbf{A}|\mathbf{p}] = \left[\begin{array}{cccc|c} 1 & 2 & -3 & 4 & 1 \\ 4 & 8 & 12 & -8 & 2 \\ 2 & 3 & 2 & 1 & 3 \\ -3 & -1 & 1 & -4 & 4 \end{array}\right] \xrightarrow{\mathbf{P}_1} \left[\begin{array}{cccc|c} 4 & 8 & 12 & -8 & 2 \\ 1 & 2 & -3 & 4 & 1 \\ 2 & 3 & 2 & 1 & 3 \\ -3 & -1 & 1 & -4 & 4 \end{array}\right] \tag{17}$$

$$\xrightarrow{\mathbf{M}_1} \left[\begin{array}{cccc|c} 4 & 8 & 12 & -8 & 2 \\ \mathbf{1/4} & 0 & -6 & 6 & 1 \\ \mathbf{1/2} & -1 & -4 & 5 & 3 \\ \mathbf{-3/4} & 5 & 10 & -10 & 4 \end{array}\right] \xrightarrow{\mathbf{P}_2} \left[\begin{array}{cccc|c} 4 & 8 & 12 & -8 & 2 \\ \mathbf{-3/4} & 5 & 10 & -10 & 4 \\ \mathbf{1/2} & -1 & -4 & 5 & 3 \\ \mathbf{1/4} & 0 & -6 & 6 & 1 \end{array}\right] \tag{18}$$

$$\xrightarrow{\mathbf{M}_2} \left[\begin{array}{cccc|c} 4 & 8 & 12 & -8 & 2 \\ \mathbf{-3/4} & 5 & 10 & -10 & 4 \\ \mathbf{1/2} & \mathbf{-1/5} & -2 & 3 & 3 \\ \mathbf{1/4} & \mathbf{0} & -6 & 6 & 1 \end{array}\right] \xrightarrow{\mathbf{P}_3} \left[\begin{array}{cccc|c} 4 & 8 & 12 & -8 & 2 \\ \mathbf{-3/4} & 5 & 10 & -10 & 4 \\ \mathbf{1/4} & \mathbf{0} & -6 & 6 & 1 \\ \mathbf{1/2} & \mathbf{-1/5} & -2 & 3 & 3 \end{array}\right] \tag{19}$$

$$\xrightarrow{\mathbf{M}_3} \left[\begin{array}{cccc|c} 4 & 8 & 12 & -8 & 2 \\ \mathbf{-3/4} & 5 & 10 & -10 & 4 \\ \mathbf{1/4} & \mathbf{0} & -6 & 6 & 1 \\ \mathbf{1/2} & \mathbf{-1/5} & \mathbf{1/3} & 1 & 3 \end{array}\right]. \tag{20}$$

Therefore, $\mathbf{M}_3\mathbf{P}_3\mathbf{M}_2\mathbf{P}_2\mathbf{M}_1\mathbf{P}_1\mathbf{A} = \mathbf{U}, \mathbf{LU} = \mathbf{PA}$ and

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3/4 & 1 & 0 & 0 \\ 1/4 & 0 & 1 & 0 \\ 1/2 & -1/5 & 1/3 & 1 \end{bmatrix}, \mathbf{U} = \begin{bmatrix} 4 & 8 & 12 & -8 \\ 0 & 5 & 10 & -10 \\ 0 & 0 & -6 & 6 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{21}$$

References

[1] Horn, Roger A., and Charles R. Johnson. Matrix analysis. Cambridge university press, 2012.

[2] Meyer, Carl D. Matrix analysis and applied linear algebra. Vol. 71. Siam, 2000.