ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ

Χειμερινό εξάμηνο 2024-25
(ΜΥΥ104-ΠΛΥ104)

Κωνσταντίνος Σκιάνης
Επίκουρος Καθηγητής

Περιεχόμενα

Motivation for LU Decomposition

- Direct methods for solving systems of linear equations often rely on matrix factorizations.
- The LU decomposition provides a systematic way to break down a matrix A into simpler building blocks.
- It enables efficient solutions to $Ax = b$ for multiple right-hand sides.

Definition of LU Decomposition

LU Decomposition

For a given $n \times n$ matrix A, an **LU decomposition** (or factorization) is a representation:

$$A = LU$$

where:

- L is a lower-triangular matrix with ones on the diagonal.
- U is an upper-triangular matrix.

Definition of LU Decomposition

LU Decomposition

For a given $n \times n$ matrix A, an **LU decomposition** (or factorization) is a representation:

$$A = LU$$

where:

- L is a lower-triangular matrix with ones on the diagonal.
- U is an upper-triangular matrix.

We already know the matrix U that results from the elimination process; it is an upper triangular matrix where the pivots are located on its main diagonal (we assume that row exchanges are not needed).

Definition of LU Decomposition

LU Decomposition

For a given $n \times n$ matrix A, an **LU decomposition** (or factorization) is a representation:

$$A = LU$$

where:

- L is a lower-triangular matrix with ones on the diagonal.
- U is an upper-triangular matrix.

We already know the matrix U that results from the elimination process; it is an upper triangular matrix where the pivots are located on its main diagonal (we assume that row exchanges are not needed).

Key Question: Under what conditions does A admit an LU decomposition?

Conditions for A to admit an LU decomposition

1. **Square Matrix:**
   - A must be a square matrix ($n \times n$). For non-square matrices, variations of LU decomposition (e.g., $PA = LU$) might be considered.
2. **No Zero Leading Principal Minors:**
   - All leading principal minors (determinants of the upper-left $k \times k$ submatrices of A, where $k = 1, 2, \ldots, n$) must be nonzero. This ensures that the Gaussian elimination process used to compute the LU decomposition does not encounter a zero pivot.
3. **Pivoting or Permutation (if necessary):**
   - If A does not satisfy the condition on leading principal minors, LU decomposition may still exist if row interchanges are allowed. This leads to a *partial pivoting version* of the decomposition, where $PA = LU$, with P being a permutation matrix.
4. **Stability of Numerical Algorithms:**
   - In numerical computation, stability considerations often require pivoting (row swaps), even if A theoretically satisfies the condition for LU decomposition without pivoting. This ensures numerical accuracy.

Constructing the LU Factorization

To find L and U:

1. Perform Gaussian elimination on A without row interchanges.
2. The multipliers used to eliminate entries below the pivots fill in L.
3. The resulting echelon form is U.

Constructing the LU Factorization

To find L and U:

1. Perform Gaussian elimination on A without row interchanges.
2. The multipliers used to eliminate entries below the pivots fill in L.
3. The resulting echelon form is U.

If at any step a pivot is zero, we may need to reorder the equations (pivoting), leading to a $PA = LU$ factorization.

Step-by-Step Procedure for LU Decomposition

Given Matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

Step 1: Eliminate entries below a_{11} (first column):
For each $i = 2, 3, \ldots, n$, compute the multiplier:

$$l_{i1} = \frac{a_{i1}}{a_{11}}.$$

Update the entries of row i as:

$$a_{ij} \leftarrow a_{ij} - l_{i1} \cdot a_{1j}, \quad \text{for } j = 1, 2, \ldots, n.$$

Result after Step 1:

$$A^{(1)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & \cdots & a_{3n}^{(1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & a_{n3}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix}.$$

Step 2: Eliminate entries below $a_{22}^{(1)}$ (second column):
For each $i = 3, 4, \ldots, n$, compute the multiplier:

$$l_{i2} = \frac{a_{i2}^{(1)}}{a_{22}^{(1)}}.$$

Update the entries of row i as:

$$a_{ij}^{(1)} \leftarrow a_{ij}^{(1)} - l_{i2} \cdot a_{2j}^{(1)}, \quad \text{for } j = 2, 3, \ldots, n.$$

Result after Step 2:

$$A^{(2)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}.$$

Step k: Eliminate entries below $a_{kk}^{(k-1)}$ (k-th column):
For each $i = k+1, k+2, \ldots, n$, compute the multiplier:

$$l_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}.$$

Update the entries of row i as:

$$a_{ij}^{(k-1)} \leftarrow a_{ij}^{(k-1)} - l_{ik} \cdot a_{kj}^{(k-1)}, \quad \text{for } j = k, k+1, \ldots, n.$$

Final Result: After $n-1$ steps, A is transformed into the upper triangular matrix U, and the multipliers l_{ik} form the entries of the lower triangular matrix L.

$$A = LU, \quad \text{where}$$

$$L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix}, \quad U = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{bmatrix}.$$

Example of an LU Decomposition

Consider the matrix

$$A = \begin{pmatrix} 2 & 4 & 2 \\ 4 & 9 & 6 \\ 2 & 6 & 9 \end{pmatrix}.$$

One possible LU decomposition is:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}.$$

Example of an LU Decomposition

Consider the matrix

$$A = \begin{pmatrix} 2 & 4 & 2 \\ 4 & 9 & 6 \\ 2 & 6 & 9 \end{pmatrix}.$$

One possible LU decomposition is:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}.$$

Note how L captures the elimination steps and U stores the multipliers and final reduced form.

Step 1: Initialize L and U

Start with L as the identity matrix and U as A:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 2 & 4 & 2 \\ 4 & 9 & 6 \\ 2 & 6 & 9 \end{pmatrix}.$$

Step 2: Eliminate a_{21} and a_{31} (First Column)

Compute the multipliers:

$$l_{21} = \frac{U_{21}}{U_{11}} = \frac{4}{2} = 2, \quad l_{31} = \frac{U_{31}}{U_{11}} = \frac{2}{2} = 1.$$

Update L with these multipliers:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Perform row operations on U:

$$R_2 \leftarrow R_2 - l_{21} \cdot R_1, \quad R_3 \leftarrow R_3 - l_{31} \cdot R_1.$$

The updated U becomes:

$$U = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 2 & 7 \end{pmatrix}.$$

Step 3: Eliminate a_{32} (Second Column)

Compute the multiplier:

$$l_{32} = \frac{U_{32}}{U_{22}} = \frac{2}{1} = 2.$$

Update L with this multiplier:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}.$$

Perform the row operation on U:

$$R_3 \leftarrow R_3 - l_{32} \cdot R_2.$$

The updated U becomes:

$$U = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}.$$

Final L and U

After completing the elimination steps, we have:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}.$$

Verification

Verify that $A = LU$:

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}.$$

Compute the product:

$$LU = \begin{pmatrix} 2 & 4 & 2 \\ 4 & 9 & 6 \\ 2 & 6 & 9 \end{pmatrix}.$$

This matches the original matrix A, confirming that the decomposition is correct.

Pivoting and $PA = LU$ Factorization

- Sometimes A does not admit an LU decomposition directly (e.g., if a pivot is zero).
- Introducing a permutation matrix P allows row exchanges, resulting in:

$$PA = LU.$$

- P is obtained from the identity matrix by permuting rows.

Pivoting and $PA = LU$ Factorization

- Sometimes A does not admit an LU decomposition directly (e.g., if a pivot is zero).
- Introducing a permutation matrix P allows row exchanges, resulting in:
$$PA = LU.$$
- P is obtained from the identity matrix by permuting rows.

This ensures numerical stability and existence of the factorization in a broader set of cases.

Applications of LU Decomposition

- **Solving Systems:** Once we have $A = LU$, solving $Ax = b$ involves:

$$LUx = b.$$

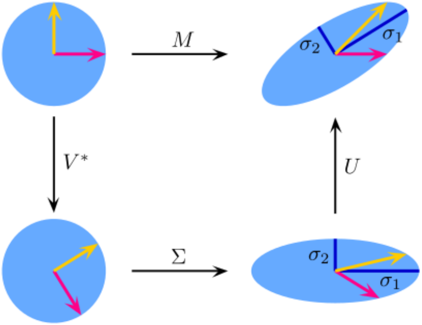
  Set $y = Ux$, then solve $Ly = b$ and then $Ux = y$.
- **Inverse and Determinants:** - If A is invertible, A^{-1} can be computed from L and U. - The determinant of A is the product of the diagonal entries of U (up to the sign adjustments from P if pivoting is used).

Summary

- The LU decomposition is a fundamental tool for factorizing a matrix into a product of simpler lower and upper-triangular factors.
- It streamlines the solution of linear systems and the computation of determinants and inverses.
- When direct LU factorization is not possible, consider the $PA = LU$ variant.

Singular Value Decomposition

Illustration of the singular value decomposition $U\Sigma V^*$



$$M = U \cdot \Sigma \cdot V^*$$

Top: The action of M, indicated by its effect on the unit disc D and the two canonical unit vectors e_1 and e_2.
Left: The action of V^*, a rotation, on D, e_1, and e_2.
Bottom: The action of Σ, a scaling by the singular values σ_1 horizontally and σ_2 vertically.
Right: The action of U, another rotation.

Geometric Meaning

- Consider the transformation defined by $A : \mathbb{R}^n \to \mathbb{R}^m$.
- The matrix V provides an orthonormal basis for the input space. In this basis, A stretches or shrinks along the coordinate axes defined by V.
- The singular values σ_i describe how much A stretches these basis vectors.
- The matrix U gives an orthonormal basis of the output space aligned with these stretched directions.

$$Av_i = \sigma_i u_i.$$

Singular Value Decomposition Example

We are given the matrix:

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}.$$

The goal is to compute the Singular Value Decomposition (SVD) of A in the form:

$$A = U\Sigma V^\top,$$

where:

- U is an orthogonal matrix,
- Σ is a diagonal matrix with non-negative singular values,
- V is an orthogonal matrix.

Step 1: Compute $A^\top A$

First, compute $A^\top A$:

$$A^\top A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}^\top \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix}.$$

Step 2: Eigenvalues and Eigenvectors of $A^\top A$

Solve $\det(A^\top A - \lambda I) = 0$ to find the eigenvalues:

$$\det\left(\begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det\begin{pmatrix} 10-\lambda & 6 \\ 6 & 10-\lambda \end{pmatrix}.$$

$$\det = (10-\lambda)^2 - 36 = \lambda^2 - 20\lambda + 64.$$

Solve $\lambda^2 - 20\lambda + 64 = 0$:

$$\lambda = 16, \quad \lambda = 4.$$

The eigenvalues of $A^\top A$ are $\lambda_1 = 16$ and $\lambda_2 = 4$.
The singular values of A are:

$$\sigma_1 = \sqrt{\lambda_1} = 4, \quad \sigma_2 = \sqrt{\lambda_2} = 2.$$

Step 3: Compute V

Find the eigenvectors of $A^\top A$ corresponding to $\lambda_1 = 16$ and $\lambda_2 = 4$.
For $\lambda_1 = 16$:

$$\begin{pmatrix} 10 - 16 & 6 \\ 6 & 10 - 16 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This simplifies to:

$$\begin{pmatrix} -6 & 6 \\ 6 & -6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The eigenvector is:

$$v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

For $\lambda_2 = 4$:

$$\begin{pmatrix} 10-4 & 6 \\ 6 & 10-4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This simplifies to:

$$\begin{pmatrix} 6 & 6 \\ 6 & 6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The eigenvector is:

$$v_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

Thus, V is:

$$V = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Step 4: Compute U

Compute U using $U = AV\Sigma^{-1}$, where:

$$\Sigma = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}.$$

Final Result

The Singular Value Decomposition of A is:

$$A = U\Sigma V^{\top},$$

where:

$$U = \begin{pmatrix} \cdots & \cdots \\ \cdots & \cdots \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, \quad V^{\top} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Low-Rank Approximation

- The SVD provides the best low-rank approximation to a given matrix.
- Truncate the SVD after k terms:

$$A \approx U_k \Sigma_k V_k^T$$

  where Σ_k contains only the top k singular values.
- This approximation minimizes the Frobenius norm error:

$$\|A - U_k \Sigma_k V_k^T\|_F = \min_{\text{rank}(B) \le k} \|A - B\|_F.$$

Data Compression

- SVD is used extensively in data compression, image processing, and noise reduction.
- For example, compressing an image by storing only the top singular values and corresponding singular vectors drastically reduces storage while preserving image quality.

Principal Component Analysis (PCA)

- PCA can be derived using the SVD of a data matrix.
- Given data with mean zero:

$$X \in \mathbb{R}^{m\times n}, \quad X = U\Sigma V^T.$$

- The columns of V (eigenvectors of X^TX) are the principal directions.
- The singular values determine the importance of each principal component.

Summary

- SVD is a powerful decomposition that applies to any $m \times n$ matrix.
- It reveals the intrinsic structure of the linear transformation defined by the matrix.
- SVD is critical for dimensionality reduction, noise filtering, and data analysis techniques like PCA.