3. Matrix decompositions and its applications

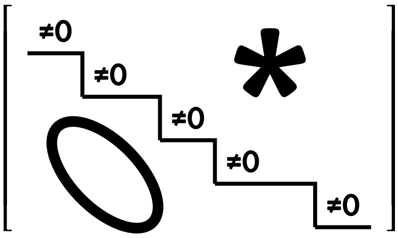
We have already discussed the skeleton decomposition, slightly talked about the Singular Value Decomposition and Spectral decomposition. Now we should investigate some others useful decompositions and give the concrete examples of SVD since it is the most important method to data science and linear algebra that can solve a set of problems, e.g. dimension reduction and low-rank approximations. Let us begin with the LU decomposition.

LU decomposition

Lower-Upper (LU) decomposition or factorization factors a matrix as the product of a lower triangular matrix and an upper triangular matrix.

Suppose $A = M_{n\times n}(F)$ is a square matrix. For simplicity, let's talk about the homogeneity system of linear equations $A\vec{x} = \vec{0}$. We can use the classical approach such as Gaussian elimination to obtain the Row Echelon Form of the matrix A:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \xrightarrow[\text{elimination}]{\text{Gaussian}}$$



Matrix obtained with Gaussian elimination is the upper-triangular one. The steps of Gaussian elimination process are quite simple. Suppose the element of the matrix A $a_{11} \neq 0$. Than we need to subtract from other rows the chosen one (in our case it is the row A_1 with element a_{11}) multiplied with some number:

$$A_n \hookrightarrow A_n - \lambda A_1,$$

where $\lambda = \dfrac{a_{n1}}{a_{11}}$. After that we need to perform approximately n^2 operations to obtain the desired upper triangular matrix. In fact, when we apply such operations we perform matrix multiplication:

$$A \hookrightarrow A' = T \cdot A = \underbrace{T \cdot I}_{T} \cdot A',$$

where the matrix $T = \begin{bmatrix} 1 & & \bigcirc & \\ & 1 & & \\ \bigcirc & & \ddots & \\ -\lambda & & & 1 \end{bmatrix} = L_1$ which is lower triangular matrix with 1's on the diagonal. So the complete pipeline:

$$A \longrightarrow U = L_1 L_2 \cdot L_k \cdot A,$$
$$A = (L_1 \dots L_k)^{-1} U = LU,$$

where, again, L is a lower triangular matrix and U is upper triangular one.

But actually it is not a general case. On the very first step we have assumed that the element a_{11} of the row A_1 is non-zero element and repeated it for each further step. But it can be equal to zero and the same steps wont be enough. In such cases we need to permute the rows to make the decomposition. Permuting is equivalent to multiplying on the matrix of permutation. Hence, for general matrices the decomposition can be written in the following way:

$$A = PLU,$$

where the matrix P is the permutation matrix (permutation of rows of identity matrix I), e.g.:

$$\begin{bmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & \dots & \dots & 0 \\ 0 & 1 & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & 1 & \dots & 0 \end{bmatrix}$$

Note

Useful properties of LU decomposition

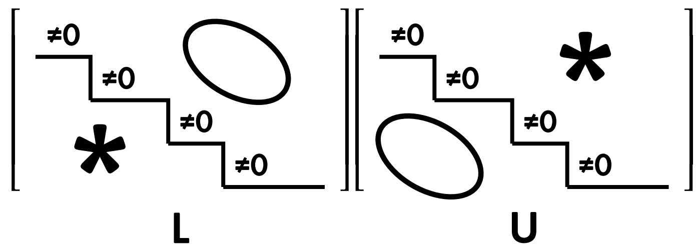
- The determinant of the decomposed matrix A can be computed from finding the product of the diagonal elements of both the L and U matrices from the LU decomposition;

$$\det \boldsymbol{A} = \det \boldsymbol{L} \det \boldsymbol{U} = \left(\prod_{i=1}^{n} l_{ii}\right)\left(\prod_{i=1}^{n} u_{ii}\right)$$

- The knowledge of the LU decomposition of matrix A allows us to solve the system of linear equations $A\vec{x} = \vec{b}$.

  The steps to solve the system of linear equations using LU decomposition:

  1. Find the LU decomposition:

  $A = LU =$ 

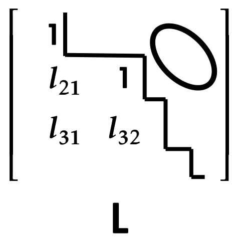
  2. We can write the initial system using the decomposition $LUx = b$ or $L(Ux) = b$. Then we can write the system:

  $$\begin{cases} Ly = b, \\ Ux = y \end{cases}$$

  3. We can immediately find $y_1, \dots, y_n$:

  $$\begin{cases} l_{11}y_1 = b_1 & \rightsquigarrow \text{find } y_1 \\ l_{21}y_1 + l_{22}y_2 = b_2 & \rightsquigarrow \text{find } y_2 \\ \vdots & \vdots \\ l_{n1}y_1 + l_{n2}y_2 + \dots + l_{nn}y_n = b_n & \rightsquigarrow \text{find } y_n \end{cases}$$

  4. We can find the elements of matrices L and U in a fast way:

  $$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & a_{12} & \dots & a_{nn} \end{bmatrix} =$$  U

Hence, we have:

$$
\begin{aligned}
a_{11} &= u_{11} \\
a_{12} &= u_{12} \\
a_{13} &= u_{13} \\
a_{21} &= l_{21} \cdot u_{11} && \rightsquigarrow l_{21} \\
a_{22} &= l_{21} \cdot u_{12} + u_{22} && \rightsquigarrow u_{22}
\end{aligned}
$$

Example 1: Find the LU decomposition of the matrix:

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

Let's find the elements of matrices L and U:

$$\begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$$

Hence, the elements $u_{11} = 2,\ u_{12} = 3$. Then

$$
\begin{aligned}
-1 &= l_{21} \cdot u_{11} = 2l_{21} && \Longrightarrow l_{21} = -\frac{1}{2} \\
-2 &= l_{21} \cdot u_{12} + u_{22} = -\frac{3}{2} + u_{22} && \Longrightarrow u_{22} = -\frac{1}{2}
\end{aligned}
$$

Cholesky Decomposition

Probably, the most useful case for LU decomposition is the decomposition for symmetric positive-definite matrices ($a_{ij} = a_{ji}$ and all eigenvalues are non-negative $\lambda_i \geq 0$). Suppose, A is a symmetric (or Hermitian if A is complex) positive-definite matrix, we can arrange matters so that U is the conjugate transpose of L.

Note

- The symmetric matrix: $a_{ij} = a_{ji} \forall i \neq j$;
- The positive definite:

  All eigenvalues of matrix are greater than zero:

$$\forall i : \ \lambda_i(A) \geq 0$$

or

$$\forall \vec{x} : \ \vec{x}^\top A\vec{x} \geq 0$$

The Cholesky decomposition of a Hermitian positive-definite matrix A is a decomposition of the form:

$$A = LL^*$$

where L is a lower triangular matrix with real and positive diagonal entries $l_{ii} \geq 0$, and $L*$ its conjugate transpose. A closely related variant of this decomposition is the LDL decomposition:

$$A = LDL^*$$

where L is a lower unit triangular (unitriangular) matrix, and D is a diagonal matrix:

$$D = \begin{bmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_{nn} \end{bmatrix}, \ \forall i \in \{1, \dots, n\} \ d_{ii} \geq 0.$$

The way to find such decomposition is similar to LU one:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \dots & \dots & 0 \\ l_{21} & l_{22} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & \dots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & \dots & l_{n1} \\ 0 & l_{22} & \dots & l_{n2} \\ \vdots & \ddots & \ddots & \ddots \\ \vdots & \dots & \dots & \dots \end{bmatrix}$$

So the steps are:

$$\begin{array}{ll} l_{11}^2 = a_1 1 & \rightsquigarrow l_{11} = \sqrt{a_{11}} \\ l_{11} \cdot l_{21} = a_{12} & \rightsquigarrow l_{21} \\ l_{21} \cdot l_{11} = a_{21} & \\ \dots & \dots \end{array}$$

If the system of linear equations have the solution then the system is equivalent to solving the normal system:

$$Ax = b \Longleftrightarrow A^*Ax = A^*b,$$

where $N = A^*A$ is symmetric non-negative definite.

Spectral Decomposition

Suppose A is symmetric matrix $A = A^\top$ (or at least normal), then there is a diagonal matrix:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_n \end{bmatrix}, \ \lambda_i - \text{ eigenvalues of A,}$$

and unitary matrix:

$$U = \begin{bmatrix} | & | & & | \\ u^1 & u^2 & \dots & u^n \\ | & | & & | \end{bmatrix}$$
$$U^\top = U^{-1}, \quad |U^i| = 1, \quad U^i \perp U^I$$

where u^i is an eigenvector for with corresponding eigenvalue $\lambda = \lambda_i$. Then we can decompose the matrix A:

$$A = U\Lambda U^{-1}$$
$$A = U\Lambda U^\top$$

It is very easy to find the solutions of system of linear equations using this decomposition:

$$A\vec{x} = \vec{b} \to U\Lambda U^\top \vec{x} = \vec{b} \Longrightarrow$$
$$\Rightarrow U^\top \vec{x} = \Lambda^{-1} U^\top \vec{b},$$

$$\Lambda^{-1} = \begin{bmatrix} \lambda_1^{-1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_n^{-1} \end{bmatrix}$$

Hence, the solution:

$$\vec{x} = U\Lambda^{-1} U^\top \vec{b}$$

Singular Value Decomposition

For the arbitrary matrices and even non-square ones we can find the Singular Value Decomposition. Firstly, let's talk about the singular values. To find the singular values we need to find the eigenvalues of normal matrix $A^\top A$.

$$A \rightsquigarrow A^\top A \rightsquigarrow \lambda_i(A^\top A) \geq 0,$$

where $i = 1, \dots, n$. The singular values of matrix A then:

$$\begin{matrix} \sigma_1 = \sqrt{\lambda_1(A^*A)}, \\ \vdots \\ \sigma_n = \sqrt{\lambda_n(A^*A)} \end{matrix}, \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$$

If matrix is symmetric $A = A^\top$:

$$\sigma_i = \sqrt{\lambda_i(A)^2} = |\lambda_i(A)|$$

If the matrix is also positive definite $\forall i \ \lambda_i(A) \geq 0$, then $\sigma_i = \lambda_i$. We can form from the obtained singular values the diagonal matrix of Σ:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

The singular value decomposition:

$$A = \underset{m\times n}{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}} = U\Sigma V^\mathsf{T} = \underset{m\times m}{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \dots & u_m \\ | & | & & | \end{bmatrix}} \times \underset{m\times n}{\begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \\ 0 & 0 & \dots & 0 \end{bmatrix}} \times \underset{n\times n}{\begin{bmatrix} - & V_1^T & - \\ - & V_2^T & - \\ & \vdots & \\ - & V_n^T & - \end{bmatrix}}$$

QR decomposition

Definition: QR decomposition

Let A be a square invertable matrix. We call following decomposition a QR decomposition

$$A = QR,$$

where Q is an unitary (orthogonal) matrix and R is an upper triangular matrix.

Theorem: How to find QR decomposition

Given square invertable matrix $A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_m \\ | & | & & | \end{bmatrix}$ with columns $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$, the QR decomposition of A can be found using Gram–Schmidt process

$$\begin{aligned}
\vec{u}_1 &= \vec{a}_1, & \vec{e}_1 &= \frac{\vec{u}_1}{||\vec{u}_1||} \\
\vec{u}_2 &= \vec{a}_2 - \mathrm{proj}_{\vec{u}_1} \vec{a}_2, & \vec{e}_2 &= \frac{\vec{u}_2}{||\vec{u}_2||} \\
\vec{u}_3 &= \vec{a}_3 - \mathrm{proj}_{\vec{u}_1} \vec{a}_3 - \mathrm{proj}_{\vec{u}_2} \vec{a}_3, & \vec{e}_3 &= \frac{\vec{u}_3}{||\vec{u}_3||} \\
&\vdots & &\vdots \\
\vec{u}_m &= \vec{a}_m - \sum_{j=1}^{m-1} \mathrm{proj}_{\vec{u}_j} \vec{a}_m, & \vec{e}_m &= \frac{\vec{u}_m}{||\vec{u}_m||},
\end{aligned}$$

where

$$\mathrm{proj}_{\vec{u}} \vec{a} = \frac{\vec{u}^* \vec{a}}{\vec{u}^* \vec{u}} \vec{u}.$$

Then

$$Q = \begin{bmatrix} | & | & & | \\ \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_m \\ | & | & & | \end{bmatrix}, \qquad R = \begin{bmatrix} \vec{e}_1^* \vec{a}_1 & \vec{e}_1^* \vec{a}_2 & \vec{e}_1^* \vec{a}_3 & \cdots & \vec{e}_1^* \vec{a}_m \\ 0 & \vec{e}_2^* \vec{a}_2 & \vec{e}_2^* \vec{a}_3 & \cdots & \vec{e}_2^* \vec{a}_m \\ 0 & 0 & \vec{e}_3^* \vec{a}_3 & \cdots & \vec{e}_3^* \vec{a}_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \vec{e}_m^* \vec{a}_m \end{bmatrix}.$$