Matrix Decomposition

K.Bouanane

12 mars 2023

Determinant and Trace

Determinant

Determinant and Trace

Definition

Consider $\mathbf{A} \in \mathbb{R}^{n \times n}$, a square matrix with n rows and n columns. The determinant of $\mathbf{A}$, denoted $det(\mathbf{A})$, is a function that maps A onto a real number we write

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

Determinant Properties I

The Determinant $\det(\mathbf{A})$ has the following properties :

- For $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$, we have $\det(\mathbf{A}, \mathbf{B}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$.
- $\det(\mathbf{A}) = \det\left(\mathbf{A}^T\right)$: Determinant are invariant to transposition.
- If $\mathbf{A}$ is invertible then $\det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det(\mathbf{A})}$
- Similar matrices possess the same determinant. For a linear mapping $\phi : \mathbf{V} \rightarrow \mathbf{V}$, all transformation matrices $\mathbf{A}\phi$ of ϕ possess the same determinant.
- Adding a multiple of a column|row to another one doesn't affect $\det(\mathbf{A})$.

Determinant Properties II

- Multiplying a column|row by $\lambda \in \mathbb{R}$ scales $\det(\mathbf{A})$ by λ. In particular, we have $\det(\lambda\mathbf{A}) = \lambda^n \det(\mathbf{A})$.
- Swapping two rows|columns changes the sign of $\det(\mathbf{A})$.

Determinants for testing Matrix Invertibility

Determinants for testing Matrix Invertibility I

Consider $\mathbf{A} \in \mathbb{R}^{n\times n}$.

1. If $n = 1$: then $\mathbf{A} = a \in \mathbb{R}$, thus $\mathbf{A}^{-1} = \frac{1}{a}$,which exists if and only if $a \neq 0$.
2. If $n = 2$: then $\mathbf{A} \in \mathbb{R}^{2\times 2}$. From the definition of $\mathbf{A}^{-1}$ we have $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Then the inverse $\mathbf{A}^{-1}$ is

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Hence, $\mathbf{A}$ is invertible if and only if $a_{11}a_{22} - a_{12}a_{21} \neq 0$

Determinants for testing Matrix Invertibility II

The quantity $a_{11}a_{22} - a_{12}a_{21}$ is the determinant of $\mathbf{A}$. i.e.

$$\det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

.

Theorem

For any square Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$.

Determinant Computation

Theorem (Laplace Expansion)

Consider $\mathbf{A} \in \mathbb{R}^{n \times n}$*. Then,for all* $j = 1 \dots n$*.*

- *Expansion along column j :*

$$\det(\mathbf{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det\left(\mathbf{A}_{kj}\right)$$

- *Expansion along row j :*

$$\det(\mathbf{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{jk} \det\left(\mathbf{A}_{jk}\right)$$

Where $\mathbf{A}_{kj} \in \mathbb{R}^{(n-1)\times(n-1)}$ *is the submatrix of A obtained by deleting row k and column j.*

Example

Example : $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$

Example

Example : $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$

$$\begin{aligned}
\det(\mathbf{A}) &= (-1)^{1+1} \times 1 \times \det\left(\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\right) \\
&+ (-1)^{1+2} \times 2 \times \det\left(\begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}\right) \\
&+ (-1)^{1+3} \times 3 \times \det\left(\begin{bmatrix} 3 & 1 \\ 0 & 1 \end{bmatrix}\right) \\
&= -1 - 2 \times 3 + 3 \times 3 = 2.
\end{aligned}$$

Remark

Remark

Guassian Elimination can be used to compute $\det(\mathbf{A})$. We can put $\mathbf{A}$ in a row_echelon Form Therefore $\mathbf{A}$ is transformed to a similar triangular matrix $\tilde{\mathbf{A}}$, and in this case the determinant of $\tilde{\mathbf{A}}$ is the product of its diagonal elements.

Remark

Remark

Guassian Elimination can be used to compute $\det(\mathbf{A})$. We can put $\mathbf{A}$ in a row_echelon Form Therefore $\mathbf{A}$ is transformed to a similar triangular matrix $\tilde{\mathbf{A}}$, and in this case the determinant of $\tilde{\mathbf{A}}$ is the product of its diagonal elements.

Example : Consider the matrix $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

Trace

Trace

Definition

Trace The trace of a sqeare matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is defined as

$$\operatorname{tr}(\mathbf{A})) = \sum_{i=1}^{n} a_{ii}$$

Trace Properties

The trace satisfies the following properties :

1. $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B}), \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$
2. $\mathrm{tr}(\alpha\mathbf{A}) = \alpha\,\mathrm{tr}(\mathbf{A}), \alpha \in \mathbb{R}, \mathbf{A} \in \mathbb{R}^{n\times n*}$
3. $\mathrm{tr}(\mathbf{I}_n) = n$
4. $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, For $\mathbf{A} \in \mathbb{R}^{n\times k}, \mathbf{B} \in \mathbb{R}^{k\times m}$
5. $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{CAB}) = \mathrm{tr}(\mathbf{BCA})$.

Meaning that the trace is invariant under cyclic permutations.

Characteristic polynomial

Characteristic polynomial

Definition

Characteristic polynomial For $\lambda \in \mathbb{R}$ and a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ we define

$$\mathcal{P}_A(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = c_0 + c_1\lambda + c2\lambda^2 + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n \lambda^n$$

where $c_0, \ldots, c_{n-1} \in \mathbb{R}$

$\mathcal{P}_A(\lambda)$ is called the characteristic polynomial of $\mathbf{A}$.

In particular

$$c_0 = \det(\mathbf{A})$$

$$c_{n-1} = (-1)^{n-1} \operatorname{tr}(\mathbf{A})$$

Example

Example :

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

Definition

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a square matrix . Then a real number $\lambda \in \mathbb{R}$ is an eigenvalue of $\mathbf{A}$ and a non zero vector $\mathbf{x} \in \mathbb{R}^n$ is an *eigenvector* of $\mathbf{A}$ corresponding to λ if

$$\mathbf{Ax} = \lambda \mathbf{x} \tag{1}$$

Eigenvalues and Eigenvectors

Definition

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a square matrix . Then a real number $\lambda \in \mathbb{R}$ is an eigenvalue of $\mathbf{A}$ and a non zero vector $\mathbf{x} \in \mathbb{R}^n$ is an *eigenvector* of $\mathbf{A}$ corresponding to λ if

$$\mathbf{Ax} = \lambda \mathbf{x} \tag{1}$$

Equation (1) is called the eigenvalue Equation.

Eigenvalues and Eigenvectors

Definition

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a square matrix . Then a real number $\lambda \in \mathbb{R}$ is an eigenvalue of $\mathbf{A}$ and a non zero vector $\mathbf{x} \in \mathbb{R}^n$ is an *eigenvector* of $\mathbf{A}$ corresponding to λ if

$$\mathbf{Ax} = \lambda \mathbf{x} \tag{1}$$

Equation (1) is called the eigenvalue Equation.

The zero vector is excluded from because $\mathbf{A0} = \mathbf{0} = \lambda \mathbf{0}$ for every λ.

Eigenvalues and Eigenvectors

The following statements are equivalent

- λ is an eigenvalue of $\mathbf{A}$.
- There exists a non zero vector $\mathbf{x} \in \mathbb{R}^n$ such that $(\mathbf{A} - \lambda \mathrm{I})\mathbf{x} = 0$.
- $\mathrm{rk}(\mathbf{A} - \lambda \mathrm{I}) < n$.

Eigenvalues and Eigenvectors

Theorem

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ *be a square matrix .*
A real number $\lambda \in \mathbb{R}$ *is an eigenvalue of* $\mathbf{A}$ *if and only if* λ *is the root of the characteristic polynomial* $\mathcal{P}_{\mathbf{A}}(\lambda)$.

Some Definitions and Properties

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a square matrix .

- The set of eigenvalues of $\mathbf{A}$ is called the eigenspectrum or spectrum of the matrix $\mathbf{A}$.
- If $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ associated with λ , then vectors $\alpha\mathbf{x}$ are also an eigenvectors of $\mathbf{A}$ associated with λ, for $\alpha \neq 0$.
- For an eigenvalue λ, the subspace spanned by eigenvectors of $\mathbf{A}$ associated with λ is called the eigenspace of $\mathbf{A}$ with respect to λ and is denoted E_λ.

- The Algebraic multiplicity of an eigenvalue λ is the number of times that λ is a root of $\mathcal{P}_{\mathbf{A}}$.
- The Geometric multiplicity of an eigenvalue λ is the number of the linearly independent eigenvectors associated with λ. It is the dimentionality of the eigenspace E_λ.
- The Geometric multiplicity never exceed the Algebraic multiplicity.

Eigenvalues and Eigenvectors Computation

To find the eigenvalues and eigenvectors of a matrix $\mathbf{A}$, we proceed as follows :

1. Solve the Equation $\det(\mathbf{A} - \lambda \mathrm{I}) = 0$ to find the eigenvalues.
2. For each eigenvalue λ, solve the homogeneous system $(\mathbf{A} - \lambda \mathrm{I})\mathbf{x} = 0$.

Example1

Find the eigenvalues and eigenvectors of the following matrix

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$$

Example2

Find the eigenvalues and eigenvectors of the following matrix

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

Example3

Find the eigenvalues and eigenvectors of the following matrix

$$\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix}$$

Properties of eigenvalues and eigenvectors I

Let $\mathbf{x}$ be an eigenvector of the square matrix $\mathbf{A}$ with corresponding eigenvalue λ. Then

1. The sum of eigenvalues is the trace of the matrix.
2. The product of the eigenvalues is equal to the determinant of the matrix.
3. The eigenvalues of an invertible matrix are all different from 0.
4. The spectrum of the triangular matrix is the set of its diagonal elements.
5. A square matrix $\mathbf{A}$ and its transpose $\mathbf{A}^\top$ have the same spectrum.

Properties of eigenvalues and eigenvectors II

6. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of the matrix $\mathbf{A}$, then $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ are eigenvalues of the matrix $\mathbf{A}^k$ for any $k \in \mathbb{Z}$. Associated eigenvectors are the same ; $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ and $\mathbf{A}^k$ as well.
7. For any $\gamma \in \mathbb{R}$, $\mathbf{x}$ is an eigenvector of $\mathbf{A} + \gamma\mathbf{I}$ with eigenvalue $\lambda + \gamma$.
8. If $\mathbf{A}$ is invertible, then $\mathbf{x}$ is an eigenvector of $\mathbf{A}^{-1}$ with eigenvalue λ^{-1}.
9. If λ is an eigenvalue of an orthogonal matrix $\mathbf{Q}$, then $\dfrac{1}{\lambda}$ is also an eigenvalue of $\mathbf{Q}$.

Properties of eigenvalues and eigenvectors III

10. If $\mathbf{A}$ is symmetric then all its eigenvalues are real numbers. In this case, the eigenvectors corresponding to **distinct** eigenvalues are orthogonal.

Theorem

The eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ *of an* $n \times n$ *matrix* $\mathbf{A}$ *associated with distinct eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *are linearly independent.*

Theorem

The eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ of an $n \times n$ matrix $\mathbf{A}$ associated with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are linearly independent.

This theorem states that if all eigenvalues of the matrix $\mathbf{A}$ are distinct then the eigenvectors form a basis for $\mathbb{R}^n$.

Definition

A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is called defective if it possesses fewer than n linearly independent eigenvectors.

- A defective matrix has at least one eigenvalue λ with an algebraic multiplicity $m > 1$ and a geometric multiplicity of less than m.

Definition

A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is called defective if it possesses fewer than n linearly independent eigenvectors.

- A defective matrix has at least one eigenvalue λ with an algebraic multiplicity $m > 1$ and a geometric multiplicity of less than m.
- A non-defective matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ does not necessarily require n distinct eigenvalues, but it does require that the eigenvectors form a basis of $\mathbb{R}^n$.

Matrix Symmetrization

Theorem

Consider a matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ *a non square matrix.*

We can always obtain a symmetric, positive semidefinite matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$ *by defining*

$$\mathbf{S} = \mathbf{A}^\top \mathbf{A}$$

Theorem (Spectral Theorem)

If $\mathbf{A} \in \mathbb{R}^{n\times n}$ *is symmetric, then there exists an orthonormal basis of the corresponding vector space* V *consisting of eigenvectors of A, and each eigenvalue is real.*

Matrix Decomposition

LU Decomposition

Illustrative Example

Consider the matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 4 & 6 & 8 \end{pmatrix}$$

That we want to write the echelon form.

Illustrative Example

Consider the matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 4 & 6 & 8 \end{pmatrix}$$

That we want to write the echelon form.
we will proceed by a sequence of elementary and successive operations.

Step 1

$$\mathbf{L_1}.\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 4 & 6 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 2 & 4 \end{pmatrix}$$

Step2

$$\mathbf{L}_2(\mathbf{L_1}.\mathbf{A}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{pmatrix} = \mathbf{U}$$

Step2

$$\mathbf{L}_2(\mathbf{L_1}.\mathbf{A}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{pmatrix} = \mathbf{U}$$

Therefore, $\mathbf{L_2 L_1 A} = \mathbf{U}$.

By setting $(\mathbf{L} = \mathbf{L}_2\mathbf{L}_1)^{-1} = \mathbf{L}_1^{-1}\mathbf{L}_2^{-1}$, we have

$$\mathbf{A} = \mathbf{L}\mathbf{U}$$

For the running example, we have

$$\mathbf{L}_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{L}_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

So that

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 2 & 1 \end{pmatrix}$$

LU Decomposition : Main idea

For a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, we construct the matrices $\mathbf{L}_k$ from the multipliers $l_{j,k} = \dfrac{a_{jk}}{a_{kk}}$.

The matrix $\mathbf{L}_k$ has the form

$$L_k = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & \ddots & & \\ & & \vdots & & \ddots & \\ & & -l_{n,k} & & & 1 \end{pmatrix}$$

LU Decomposition : Main idea

Now, let $l_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ l_{k+1,k} \\ \vdots \\ l_{n,k} \end{pmatrix}$.

LU Decomposition : Main idea

Now, let $l_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ l_{k+1,k} \\ \vdots \\ l_{n,k} \end{pmatrix}$. Then $L_k = \mathrm{I} - l_k e_k^\top$

And therefore, $L_k^{-1} = \mathrm{I} - l_k e_k^\top$

Since

And therefore, $L_k^{-1} = \mathrm{I} - l_k e_k^\top$
Since

$$L_k L_k^{-1} = \underbrace{\left(\mathrm{I} - l_k e_k^\top\right)}_{L_k} \underbrace{\left(\mathrm{I} + l_k e_k^\top\right)}_{L_k^{-1}} = \mathrm{I} - l_k e_k^\top l_k e_k^\top = \mathrm{I}$$

With $e_k^\top l_k = \mathbf{0}$.

So, for any k, we have

$$L_k^{-1} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{k+1,k} & \ddots & & \\ & & \vdots & & \ddots & \\ & & l_{n,k} & & & 1 \end{pmatrix}$$

In addition,

$$L_k^{-1}L_{k+1}^{-1} = (\mathrm{I} - l_k e_k^\top)(\mathrm{I} - l_{k+1}e_{k+1}^\top) = \mathrm{I} - l_{k+1}e_{k+1}^\top - l_k e_k^\top + l_{k+1}\underbrace{e_{k+1}^\top l_k}_{=0} e_k^\top$$

Thus

$$L_k^{-1}L_{k+1}^{-1} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{k+1,k} & 1 & & \\ & & \vdots & l_{k+2,k+1} & \ddots & \\ & & \vdots & \vdots & & \\ & & l_{n,k} & l_{n,k+1} & & 1 \end{pmatrix}$$

Therefore, we obtain

$$L = L_1^{-1} L_2^{-1} \ldots L_k^{-1} L_{k+1}^{-1} \ldots L_{n-1}^{-1} L_n^{-1} = \begin{pmatrix} 1 & & & & \\ l_{2,1} & 1 & & & \\ \vdots & l_{3,2} & \ddots & & \\ & \vdots & & 1 & \\ l_{n,1} & l_{n,2} & & l_{n,n-1} & 1 \end{pmatrix}$$

LU Factorization Algorithm

We can summarize the LU Factorization as follows

1. Initialize $U = \mathbf{A}, L = \mathrm{I}$
2. For $k = 1 : n - 1$;

   For $j = k + 1 : n$

   $$L(j,k) = \frac{U(j,k)}{U(k,k)}$$

   $$U(j,k:n) = U(j,k:n) - L(j,k)U(k,k:n)$$

   end

   end

We can use the LU decomposition to solve linear systems $\mathbf{A}\mathbf{x} = \mathbf{b} = \mathbf{L}\mathbf{U}\mathbf{x} = \mathbf{b}$, as follows

We can use the LU decomposition to solve linear systems $\mathbf{Ax} = \mathbf{b} = \mathbf{LUx} = \mathbf{b}$, as follows

- Solve the linear system $\mathbf{Ly} = \mathbf{b}$,
- Solve the linear system $\mathbf{Ux} = \mathbf{y}$.

Remarks

- The lower matrix $\mathbf{L}$ has all its diagonal elements equal to 1.
- LU factorization is very effective and cheap method (computationally), especially when solving different linear systems that are associated with the same matrix $\mathbf{A}$.

$$\mathbf{AX} = \mathbf{B} \iff \mathbf{LUX} = \mathbf{B}$$

$\mathbf{A} \in \mathbb{R}^{n\times n}, \mathbf{X}, \mathbf{B} \in \mathbb{R}^{n\times k}$.

Example

Solve the linear system

$$\begin{aligned} x_1 + x_2 - x_3 &= 4 \\ x_1 - 2x_2 + 3x_3 &= -6 \\ 2x_1 + 3x_2 + x_3 &= 7 \end{aligned}$$

The problem of pivot equal to zero

At iteration k in the elimination process, we can encounter a zero pivot.

The problem of pivot equal to zero

At iteration k in the elimination process, we can encounter a zero pivot.
In this case, we have to swap between rows before multiplying by the matrix L_k.

The problem of pivot equal to zero

At iteration k in the elimination process, we can encounter a zero pivot.
In this case, we have to swap between rows before multiplying by the matrix L_k.
This is exactly equivalent to multiply by the permutation matrix P_k that permits to swap two specific rows. Therefore

$$L_n P_n L_{n-1} P_{n-1} \dots L_2 P_2 L_1 P_1 \mathbf{A} = U$$

Example

Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 5 \\ 4 & 6 & 8 \end{pmatrix}$$

Cholesky Decomposition

Theorem

A symmetric, positive definite matrix $\mathbf{A}$ can be factorized into a product $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$, where $\mathbf{L}$ is a lower triangular matrix with positive diagonal elements :

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} l_{11} & \dots & 0 \\ \vdots & \ddots & \dots \\ l_{n1} & \dots & l_{nn} \end{pmatrix} \begin{pmatrix} l_{11} & \dots & l_{n1} \\ \vdots & \ddots & \dots \\ 0 & \dots & l_{nn} \end{pmatrix} \tag{2}$$

The matrix $\mathbf{L}$ is uniquely defined and is called the Cholesky factor of $\mathbf{A}$.

How to find the matrix $\mathbf{L}$

From Formula 2, we can write

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} l_{11}^2 & l_{11}l_{12} & \dots & l_{n1}l_{11} \\ l_{11}l_{12} & l_{12}^2 + l_{22}^2 & \dots & l_{21}l_{1n} + l_{22}l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1}l_{11} & l_{21}l_{1n} + l_{22}l_{2n} & \dots & \sum_{i=1}^{n} l_{in}^2 \end{pmatrix}$$

We can then compute elements of matrix $\mathbf{L}$:

$$l_{11} = \sqrt{a_{11}}; l_{21} = \frac{1}{l_{11}} a_{21}; \ldots, l_{n1} = \frac{1}{l_{11}} a_{n1}$$

$$l_{22} = \sqrt{a_{22} - l_{21}^2}; \ldots; l_{nn} = \sqrt{a_{nn} - \sum_{i=1}^{n-1} l_{in}^2}$$

Cholesky Algorithm

The algorithm can be represented as follows

$$
\begin{aligned}
l_{11} &= \sqrt{a_{11}}, \\
l_{j1} &= \frac{a_{j1}}{l_{11}}, \quad j \in [2, n], \\
l_{ii} &= \sqrt{a_{ii} - \sum_{p=1}^{i-1} l_{ip}^2}, \quad i \in [2, n], \\
l_{ji} &= \left(a_{ji} - \sum_{p=1}^{i-1} l_{ip} l_{jp} \right) / l_{ii}, \quad i \in [2, n-1], j \in [i+1, n].
\end{aligned}
$$

Example

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Eigendecomposition and diagonalization

Definitions

Definition

Consider a square matrix $\mathbf{A}$.

We say that $\mathbf{A}$ is diagonalizable , if it similar to a diagonal matrix $\mathbf{D}$. This means that there exists an invertible matrix $\mathbf{P}$, such that

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

Eigendecomposition

Consider a square $\mathbf{A} \in \mathbb{R}^{n \times n}$ that is diagonalizable. Let $\mathbf{P} = [P_1, P_2, \dots, P_n]$ the invertible matrix for which we have

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

and $\mathbf{D} = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}$.

Eigendecomposition

We can write

$$\mathbf{AP} = \mathbf{PD}$$

And more specifically, we have

$$\mathbf{A}[P_1, P_2, \ldots, P_n] = [P_1, P_2, \ldots, P_n] \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix}$$

$$[\mathbf{A}P_1, \mathbf{A}P_2, \ldots, \mathbf{A}P_n] = [\lambda_1 P_1, \lambda_2 P_2, \ldots, \lambda_n P_n]$$

This means that

$$\mathbf{A}P_i = \lambda_i P_i, \quad i = 1..n$$

This means that

$$\mathbf{A}P_i = \lambda_i P_i, \quad i = 1..n$$

Thus, P_i is an eigenvector of $\mathbf{A}$ associated with $\lambda_i, \quad i = 1..n$.

Note that in order to get the diagonalization of $\mathbf{A}$, the matrix $\mathbf{P}$ must be invertible.

Note that in order to get the diagonalization of $\mathbf{A}$, the matrix $\mathbf{P}$ must be invertible. This means that $P_i, i = 1..n$ must be linearly independent and thus form a basis of $\mathbb{R}^n$.

Eigendecomposition Theorem

Theorem

A square matrix can be factorized into

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$$

where $\mathbf{P} \in \mathbb{R}^{n \times n}$ *and* $\mathbf{D}$ *is a diagonal matrix whose diagonal elements are the eigenvalues of* $\mathbf{A}$ *if and only if the eigenvectors of* $\mathbf{A}$ *form a basis of* $\mathbb{R}^n$*.*

This theorem implies that only non-defective matrices can be diagonalized and that the columns of$\mathbf{P}$ are the n eigenvectors of $\mathbf{A}$.

For symmetric matrices we have the following result

Theorem

Every symmetric matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$ *can be diagonalized.*

Examples

Consider the three matrices we computed eigenvalues and eigenvectors previously :

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix}$$

Issues with the eigendecomposition

The eigendecomposition concerns only square non defective matrices, since it relies on the existence of an eigenvectors basis for $\mathbb{R}^n$. In the case of rectangular or defective matrices, we cannot perform an eigendecomposition.

Singular Value Decomposition

The Singular Value Decomposition is a matrix decomposition that can be applied to any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$. It is a decomposition of the form

$$\mathbf{A} = \mathbf{U}.\Sigma.\mathbf{V}^\top$$

Where
$\mathbf{U} \in \mathbb{R}^{m \times m}$ is an orthogonal matrix with columns $u_i, i = 1..m$, and an orthogonal matrix $\mathbf{V} \in \mathbb{R}^{n \times n}$ with columns $v_j, j = 1..n$. While $\Sigma \in \mathbb{R}^{m \times n}$ such that $\Sigma_{ii} = \sigma_i \geq 0$ and $\Sigma_{ij} = 0$ for $i \neq j$.

- Elements $\sigma_k > 0, k = 1..r$ are called singular values of $\mathbf{A}$. Where r is the rank of the matrix $\mathbf{A}$.
- Vectors $u_i, i = 1..m$ and $v_j, j = 1..n$ are called left and right singular vectors, respectively.
- The singular matrix Σ is unique and it has the same size as $\mathbf{A}$. This means that Σ has a diagonal submatrix that contains the singular values and needs additional zero padding.
- However, the decomposition SVD itself is not unique.

If $m > n$, then Σ has the form

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

If $m > n$, then Σ has the form

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

If $m < n$, then Σ has the form

$$\Sigma = \begin{pmatrix} \sigma_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & \sigma_m & 0 & \dots & 0 \end{pmatrix}$$

Computing $\mathbf{V}$, the right singular matrix of $\mathbf{A}$.

Consider a rectangular matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$.

Computing $\mathbf{V}$, the right singular matrix of $\mathbf{A}$.

Consider a rectangular matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$.

According to the Spectral theorem, the symmetric matrix $\mathbf{A}^\top \mathbf{A} \in \mathbb{R}^{n \times n}$ is diagonalizable. Thus it admits an eigendecomposition of the form

Computing $\mathbf{V}$, the right singular matrix of $\mathbf{A}$.

Consider a rectangular matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$.
According to the Spectral theorem, the symmetric matrix $\mathbf{A}^\top \mathbf{A} \in \mathbb{R}^{n \times n}$ is diagonalizable. Thus it admits an eigendecomposition of the form

$$\mathbf{P}\mathbf{D}\mathbf{P}^{-1}$$

where $\mathbf{P}$ is an orthogonal matrix, composed of the orthonormal eigenbasis. The $\lambda_i \geq 0$ are eigenvalues of $\mathbf{A}^\top \mathbf{A}$.

Computing $\mathbf{V}$, the right singular matrix of $\mathbf{A}$.

Therefore, we have

Computing $\mathbf{V}$, the right singular matrix of $\mathbf{A}$.

Therefore, we have

$$
\begin{aligned}
\mathbf{A}^\top \mathbf{A} &= \mathbf{P}\mathbf{D}\mathbf{P}^{-1} \\
&= \mathbf{P}\mathbf{D}\mathbf{P}^\top \\
&= \mathbf{P} \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \mathbf{P}^\top
\end{aligned} \tag{3}
$$

Computing $\mathbf{V}$, the right singular matrix of $\mathbf{A}$.

On the other hand, we have

$$\begin{aligned}
\mathbf{A}^\top\mathbf{A} =& \left(\mathbf{U}.\Sigma.\mathbf{V}^\top\right)^\top \left(\mathbf{U}.\Sigma.\mathbf{V}^\top\right) \\
=& \mathbf{V}.\Sigma^\top.\mathbf{U}^\top\mathbf{U}.\Sigma.\mathbf{V}^\top \\
=& \mathbf{V}.\Sigma^2.\mathbf{V}^\top
\end{aligned} \tag{4}$$

By comparing 3 and 4, we obtain

By comparing 3 and 4, we obtain

$$\mathbf{A}^\top \mathbf{A} = \mathbf{P} \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \mathbf{P}^\top = \mathbf{V}.\Sigma^2.\mathbf{V}^\top$$

Therefore,

By comparing 3 and 4, we obtain

$$\mathbf{A}^\top\mathbf{A} = \mathbf{P}\begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}\mathbf{P}^\top = \mathbf{V}.\Sigma^2.\mathbf{V}^\top$$

Therefore,

$$\mathbf{V}^\top = \mathbf{P}^\top \tag{5}$$

$$\sigma_i^2 = \lambda_i \tag{6}$$

By comparing 3 and 4, we obtain

$$\mathbf{A}^\top \mathbf{A} = \mathbf{P} \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \mathbf{P}^\top = \mathbf{V}.\Sigma^2.\mathbf{V}^\top$$

Therefore,

$$\mathbf{V}^\top = \mathbf{P}^\top \tag{5}$$

$$\sigma_i^2 = \lambda_i \tag{6}$$

The eigenvectors of the matrix $\mathbf{A}^\top \mathbf{A}$ are the right singular vectors of $\mathbf{A}$.

By comparing 3 and 4, we obtain

$$\mathbf{A}^\top \mathbf{A} = \mathbf{P} \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \mathbf{P}^\top = \mathbf{V}.\Sigma^2.\mathbf{V}^\top$$

Therefore,

$$\mathbf{V}^\top = \mathbf{P}^\top \tag{5}$$

$$\sigma_i^2 = \lambda_i \tag{6}$$

The eigenvectors of the matrix $\mathbf{A}^\top\mathbf{A}$ are the right singular vectors of $\mathbf{A}$. The eigenvalues of $\mathbf{A}^\top\mathbf{A}$ are the squared singular values of $\mathbf{A}$.

Finding $\mathbf{U}$, the left singular matrix of $\mathbf{A}$.

Similarly, for the matrix $\mathbf{A}\mathbf{A}^\top \in \mathbb{R}^{m\times m}$, we have

$$\begin{aligned}\mathbf{A}\mathbf{A}^\top =& \left(\mathbf{U}.\Sigma.\mathbf{V}^\top\right)\left(\mathbf{U}.\Sigma.\mathbf{V}^\top\right)^\top \\ =&\mathbf{U}.\Sigma.\mathbf{V}^\top.\mathbf{V}.\Sigma^\top.\mathbf{U}^\top \\ =&\mathbf{U}.\Sigma^2.\mathbf{U}^\top\end{aligned}$$

Finding $\mathbf{U}$, the left singular matrix of $\mathbf{A}$.

Similarly, for the matrix $\mathbf{A}\mathbf{A}^\top \in \mathbb{R}^{m \times m}$, we have

$$\begin{aligned}\mathbf{A}\mathbf{A}^\top &= \left(\mathbf{U}.\Sigma.\mathbf{V}^\top\right)\left(\mathbf{U}.\Sigma.\mathbf{V}^\top\right)^\top \\ &= \mathbf{U}.\Sigma.\mathbf{V}^\top.\mathbf{V}.\Sigma^\top.\mathbf{U}^\top \\ &= \mathbf{U}.\Sigma^2.\mathbf{U}^\top\end{aligned}$$

On the other hand, the eigendecomposition of $\mathbf{A}\mathbf{A}^\top$ allows us to write

$$\mathbf{A}\mathbf{A}^\top = \mathbf{Q}\Lambda\mathbf{Q}^\top$$

Finding $\mathbf{U}$, the left singular matrix of $\mathbf{A}$.

Similarly, for the matrix $\mathbf{A}\mathbf{A}^\top \in \mathbb{R}^{m\times m}$, we have

$$\begin{aligned}\mathbf{A}\mathbf{A}^\top =& \left(\mathbf{U}.\Sigma.\mathbf{V}^\top\right)\left(\mathbf{U}.\Sigma.\mathbf{V}^\top\right)^\top \\ =&\mathbf{U}.\Sigma.\mathbf{V}^\top.\mathbf{V}.\Sigma^\top.\mathbf{U}^\top \\ =&\mathbf{U}.\Sigma^2.\mathbf{U}^\top\end{aligned}$$

On the other hand, the eigendecomposition of $\mathbf{A}\mathbf{A}^\top$ allows us to write

$$\mathbf{A}\mathbf{A}^\top = \mathbf{Q}\Lambda\mathbf{Q}^\top$$

Therefore, vectors $u_i, i = 1..m$ are the eigenvectors of the matrix $\mathbf{A}\mathbf{A}^\top$ and Σ^2 are eigenvalues of $\mathbf{A}\mathbf{A}^\top$.

Finding $\mathbf{U}$, the left singular matrix of $\mathbf{A}$.

How to find vectors $u_i, i = 1..m$ that correspond to vectors $v_j, j = 1..n$?

Finding $\mathbf{U}$, the left singular matrix of $\mathbf{A}$.

How to find vectors $u_i, i = 1..m$ that correspond to vectors $v_j, j = 1..n$?
Back to the SVD formula :

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

Finding $\mathbf{U}$, the left singular matrix of $\mathbf{A}$.

How to find vectors $u_i, i = 1..m$ that correspond to vectors $v_j, j = 1..n$?
Back to the SVD formula :

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

Since $\mathbf{V}$ is an orthogonal matrix, we can write :

$$\mathbf{AV} = \mathbf{U}\Sigma$$

Thus, for $\sigma_k > 0$, we have

$$[\mathbf{A}v_1, \mathbf{A}v_2, \ldots, \mathbf{A}v_r] = [\sigma_1 u_1, \sigma_1 u_2, \ldots, \sigma_r u_r]$$

Finding $\mathbf{U}$, the left singular matrix of $\mathbf{A}$.

Finding $\mathbf{U}$, the left singular matrix of $\mathbf{A}$.

And therefore

$$\mathbf{A}v_k = \sigma_k u_k, i = 1..r \Leftrightarrow u_k = \frac{1}{\sigma_k}\mathbf{A}v_k, k = 1..r$$

Finding $\mathbf{U}$, the left singular matrix of $\mathbf{A}$.

And therefore

$$\mathbf{A}v_k = \sigma_k u_k, i = 1..r \Leftrightarrow u_k = \frac{1}{\sigma_k}\mathbf{A}v_k, k = 1..r$$

If $r < m$, then we complete $\mathbf{U}$ by finding orthogonal vectors $u_i, i = r+1, .., m$ that verify $\mathbf{A}\mathbf{A}^\top u_i = 0$.

SVD Algorithm

To summarize the differents steps for SVD :

1. Compute $\mathbf{A}^\top \mathbf{A}$
2. Perform an eigendecomposition of $\mathbf{A}^\top \mathbf{A}$ and deduce the right singular vectors as eigenvectors of $\mathbf{A}^\top \mathbf{A}$
3. Compute the singular values $\sigma_k, k = 1..r$
4. Compute the left singular vectors u_i using the formula $u_i = \dfrac{1}{\sigma_i}\mathbf{A}v_i, i = 1..r$

Example1

$$\mathbf{A} = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$$

Example2

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$