Review of Matrix Algebra

In this course, we usually denote a vector without the transpose symbol as column vectors

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

and a vector with the transpose symbol as row vectors

$$\vec{x}^\top = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$$

Scalar − vector multiplication

if c is scalar, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, then $c\vec{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$.

Vector addition

If $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, then $\vec{x} + \vec{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$.

Inner product

If $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, then their inner product $\vec{x}^\top \vec{y} = \vec{y}^\top \vec{x} = \sum_{i=1}^n x_i y_i$.

Norm

$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$, i.e., $\|\vec{x}\|^2 = \vec{x}^\top \vec{x}$.

Matrix algebra

Let

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{bmatrix} \in \mathbb{R}^{n \times p}.$$

Transpose

$$\boldsymbol{A}^\top = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1p} & a_{2p} & \dots & a_{np} \end{bmatrix} \in \mathbb{R}^{p \times n}$$

Scalar–matrix multiplication

If c is a scalar, then

$$c\boldsymbol{A} = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1p} \\ ca_{21} & ca_{22} & \dots & ca_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \dots & ca_{np} \end{bmatrix}$$

Matrix addition

Let

$$\boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix} \in \mathbb{R}^{n \times p}.$$

Then

$$\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1p} + b_{1p} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2p} + b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \dots & a_{np} + b_{np} \end{bmatrix}$$

Matrix product

If

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nk} \end{bmatrix} \in \mathbb{R}^{n \times k}$$

and

$$\boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \dots & b_{kp} \end{bmatrix} \in \mathbb{R}^{k \times p},$$

then

$$\boldsymbol{A}\boldsymbol{B} = \begin{bmatrix} \sum_{t=1}^{k} a_{1t}b_{t1} & \sum_{t=1}^{k} a_{1t}b_{t2} & \dots & \sum_{t=1}^{k} a_{1t}b_{tp} \\ \sum_{t=1}^{k} a_{2t}b_{t1} & \sum_{t=1}^{k} a_{2t}b_{t2} & \dots & \sum_{t=1}^{k} a_{2t}b_{tp} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=1}^{k} a_{nt}b_{t1} & \sum_{t=1}^{k} a_{nt}b_{t2} & \dots & \sum_{t=1}^{k} a_{nt}b_{tp} \end{bmatrix}$$

Outer product of two vectors

In particular, if $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ and $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} \in \mathbb{R}^p$, then $\vec{x}$ is an $n \times 1$ matrix and $\vec{y}^\top$ is a $1 \times p$ matrix.

Therefore,

$$\vec{x}\vec{y}^\top = \begin{bmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_p \\ x_2y_1 & x_2y_2 & \dots & x_2y_p \\ \vdots & \vdots & \ddots & \vdots \\ x_ny_1 & x_ny_2 & \dots & x_ny_p \end{bmatrix}.$$

This is also called outer product between $\vec{x}$ and $\vec{y}$.

Two ways to represent the matrix products

If we represent

$$\boldsymbol{A} = \begin{bmatrix} \vec{a}_1^\top \\ \vdots \\ \vec{a}_n^\top \end{bmatrix}, \quad \boldsymbol{B} = [\vec{b}_1, \dots, \vec{b}_p],$$

where $\vec{a}_1, \dots, \vec{a}_n, \vec{b}_1, \dots, \vec{b}_p \in \mathbb{R}^k$, then we can represent the product between $\boldsymbol{A}$ and $\boldsymbol{B}$ as

$$\boldsymbol{A}\boldsymbol{B} = \begin{bmatrix} \vec{a}_1^\top\vec{b}_1 & \vec{a}_1^\top\vec{b}_2 & \dots & \vec{a}_1^\top\vec{b}_p \\ \vec{a}_2^\top\vec{b}_1 & \vec{a}_2^\top\vec{b}_2 & \dots & \vec{a}_1^\top\vec{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \vec{a}_n^\top\vec{b}_1 & \vec{a}_n^\top\vec{b}_2 & \dots & \vec{a}_n^\top\vec{b}_p \end{bmatrix}.$$

If we represent

$$\boldsymbol{A} = \begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_k \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \vec{b}_1^\top \\ \vdots \\ \vec{b}_k^\top \end{bmatrix},$$

where $\vec{a}_1, \dots, \vec{a}_k \in \mathbb{R}^n$ and $\vec{b}_1, \dots, \vec{b}_k \in \mathbb{R}^p$, then we have

$$\boldsymbol{A}\boldsymbol{B} = \vec{a}_1\vec{b}_1^\top + \vec{a}_2\vec{b}_2^\top + \ldots + \vec{a}_k\vec{b}_k^\top.$$

Symmetric matrices

If $\boldsymbol{A}$ is a square matrix, and $\boldsymbol{A} = \boldsymbol{A}^\top$, then $\boldsymbol{A}$ is a symmetric matrix. Therefore, $\boldsymbol{A}$ is symmetric if and only if $a_{ij} = a_{ji}$ for all i and j.

Identity matrix

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{k \times k}$$

Properties: for any $k \times p$ matrix $\boldsymbol{A}$ and $n \times k$ matrix $\boldsymbol{B}$

$$\boldsymbol{IA} = \boldsymbol{A} \quad \text{and} \quad \boldsymbol{BI} = \boldsymbol{B}.$$

Inverse matrix

If $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{k \times k}$ satisfy $\boldsymbol{AB} = \boldsymbol{I}$, then $\boldsymbol{B}$ is said to be the inverse of $\boldsymbol{A}$, denoted as $\boldsymbol{B} = \boldsymbol{A}^{-1}$.

Property 1: If $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{k \times k}$, then

$$\boldsymbol{AB} = \boldsymbol{I} \Leftrightarrow \boldsymbol{BA} = \boldsymbol{I} \Leftrightarrow \boldsymbol{A} = \boldsymbol{B}^{-1} \Leftrightarrow \boldsymbol{B} = \boldsymbol{A}^{-1}$$

Property 2: $\boldsymbol{A}$ is invertible $\Leftrightarrow \det(\boldsymbol{A}) \neq 0$.

Determinant and inverse of a two by two matrix

If $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have its determinant

$$\det(\boldsymbol{A}) = ad - bc.$$

If $\boldsymbol{A}$ is invertible, i.e., $\det(\boldsymbol{A}) = ad - bc \neq 0$, we have its inverse

$$\boldsymbol{A}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

To verify,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} ad + b(-c) & a(-b) + ba \\ cd + d(-c) & c(-b) + da \end{bmatrix} = \boldsymbol{I}_2.$$

Associative property

If $\boldsymbol{A} \in \mathbb{R}^{n \times m}$, $\boldsymbol{B} \in \mathbb{R}^{m \times k}$, and $\boldsymbol{C} \in \mathbb{R}^{k \times p}$, then

$$(\boldsymbol{AB})\,\boldsymbol{C} = \boldsymbol{A}\,(\boldsymbol{BC})\,.$$

For simplicity, we write this product as $\boldsymbol{ABC}$.

Distributive property

If $\boldsymbol{A} \in \mathbb{R}^{n \times k}$ and $\boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{k \times p}$, then we have

$$\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}.$$

Similarly, if $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n \times k}$ and $\boldsymbol{C} \in \mathbb{R}^{k \times p}$, then

$$(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}.$$

Transpose of products

If $\boldsymbol{A} \in \mathbb{R}^{n\times k}$ and $\boldsymbol{B} \in \mathbb{R}^{k\times p}$, then

$$(\boldsymbol{AB})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top.$$

In general, if $\boldsymbol{A}_1 \in \mathbb{R}^{n_1\times n_2}$, $\boldsymbol{A}_2 \in \mathbb{R}^{n_2\times n_3}$, ..., $\boldsymbol{A}_r \in \mathbb{R}^{n_r\times n_{r+1}}$, then

$$(\boldsymbol{A}_1\boldsymbol{A}_2 \dots \boldsymbol{A}_r)^\top = \boldsymbol{A}_r^\top \boldsymbol{A}_{r-1}^\top \dots \boldsymbol{A}_{\boldsymbol{1}}^\top.$$

Inverse of products

If $\boldsymbol{A}_1, \dots, \boldsymbol{A}_r \in \mathbb{R}^{k\times k}$, and for any $i = 1, \dots, r$, $\boldsymbol{A}_i$ is invertible. Then we have

$$(\boldsymbol{A}_1\boldsymbol{A}_2 \dots \boldsymbol{A}_r)^{-1} = \boldsymbol{A}_r^{-1} \boldsymbol{A}_{r-1}^{-1} \dots \boldsymbol{A}_{\boldsymbol{1}}^{-1}.$$

Orthogonal matrices

If $\boldsymbol{Q} \in \mathbb{R}^{k\times k}$ satisfies $\boldsymbol{Q}^\top\boldsymbol{Q} = \boldsymbol{I}$, we say $\boldsymbol{Q}$ is an orthogonal matrix.

Property 1: We have the following equivalence:

$$\boldsymbol{Q} \text{ is orthogonal } \Leftrightarrow \boldsymbol{Q}^\top\boldsymbol{Q} = \boldsymbol{I} \Leftrightarrow \boldsymbol{Q}\boldsymbol{Q}^\top = \boldsymbol{I} \Leftrightarrow \boldsymbol{Q}^\top = \boldsymbol{Q}^{-1}.$$

Property 2: Let $\boldsymbol{Q} = [\vec{q}_1, \dots, \vec{q}_k]$, where $\vec{q}_1, \dots, \vec{q}_k \in \mathbb{R}^k$. We have the following equivalence:

$$\boldsymbol{Q} \text{ is orthogonal if and only if } \vec{q}_1, \dots, \vec{q}_k \text{ are unit and pairwise orthogonal.}$$

This can be seen by

$$\boldsymbol{Q} \text{ is orthogonal } \Leftrightarrow \boldsymbol{Q}^\top\boldsymbol{Q} = \boldsymbol{I}_k \Leftrightarrow \begin{bmatrix} \vec{q}_1^\top\vec{q}_1 & \dots & \vec{q}_1^\top\vec{q}_n \\ \vdots & \ddots & \vdots \\ \vec{q}_n^\top\vec{q}_1 & \dots & \vec{q}_n^\top\vec{q}_n \end{bmatrix} = \boldsymbol{I}_k$$

The last equality means $\vec{q}_i^\top\vec{q}_j = 1$ if $i = j$ and $\vec{q}_i^\top\vec{q}_j = 0$ if $i \neq j$.

Eigenvalues and eigenvectors

For any square matrix $\boldsymbol{A} \in \mathbb{R}^{k\times k}$, if

$$\boldsymbol{A}\vec{x} = \lambda\vec{x}$$

for some scalar λ and some vector $\vec{x} \neq \vec{0}$, then λ is called an eigenvalue of $\boldsymbol{A}$, and $\vec{x}$ is called an eigenvector of $\boldsymbol{A}$ corresponding to λ.

Spectral decomposition of symmetric matrices

If $\boldsymbol{A}$ is a $k \times k$ symmetric matrix, then $\boldsymbol{A}$ has k pairs of eigenvalues and eigenvectors, namely

$$(\lambda_1, \vec{v}_1), (\lambda_2, \vec{v}_2), \dots, (\lambda_k, \vec{v}_k),$$

where $\vec{v}_1, \dots, \vec{v}_k$ are unit and pairwise perpendicular. In other words, the square matrix

$$\boldsymbol{V} = [\vec{v}_1, \dots, \vec{v}_k] \in \mathbb{R}^{k\times k}$$

is an orthogonal matrix. Furthermore, we have the following spectral decomposition

$$\boldsymbol{A} = \lambda_1\vec{v}_1\vec{v}_1^\intercal + \dots + \lambda_k\vec{v}_k\vec{v}_k^\intercal.$$

By letting

$$\boldsymbol{P} = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_k \end{bmatrix}, \quad \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix}$$

Then the above decomposition can be written as

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\top.$$

This decomposition is called a spectral decomposition. In fact

$$\begin{aligned} \boldsymbol{A} &= \lambda_1 \vec{v}_1 \vec{v}_1^\intercal + \dots + \lambda_k \vec{v}_k \vec{v}_k^\intercal \\ &= \vec{v}_1(\lambda_1 \vec{v}_1^\intercal) + \dots + \vec{v}_k(\lambda_k \vec{v}_k^\intercal) \\ &= [\vec{v}_1, \dots, \vec{v}_k] \begin{bmatrix} \lambda_1 \vec{v}_1^\intercal \\ \vdots \\ \lambda_k \vec{v}_k^\intercal \end{bmatrix} \\ &= [\vec{v}_1, \dots, \vec{v}_k] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \begin{bmatrix} \vec{v}_1^\intercal \\ \vdots \\ \vec{v}_k^\intercal \end{bmatrix} \\ &= \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\intercal. \end{aligned}$$

On the other hand, suppose that $\boldsymbol{A} \in \mathbb{R}^{k \times k}$ can be written as

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\top$$

where

$$\boldsymbol{P} = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_k \end{bmatrix}$$

is an orthogonal matrix, and

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix}$$

is a diagonal matrix. Then

- $\boldsymbol{A}$ is a symmetric matrix;
- $\lambda_1, \dots, \lambda_k$ are eigenvalues of $\boldsymbol{A}$, and $\vec{v}_1, \dots, \vec{v}_k$ are corresponding eigenvectors, respectively.
- $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\top$ is a spectral decomposition.

Proof Since

$$\begin{aligned} \boldsymbol{A} &= \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\intercal \\ &= [\vec{v}_1, \dots, \vec{v}_k] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \begin{bmatrix} \vec{v}_1^\intercal \\ \vdots \\ \vec{v}_k^\intercal \end{bmatrix} \\ &= [\vec{v}_1, \dots, \vec{v}_k] \begin{bmatrix} \lambda_1 \vec{v}_1^\intercal \\ \vdots \\ \lambda_k \vec{v}_k^\intercal \end{bmatrix} = \vec{v}_1(\lambda_1 \vec{v}_1^\intercal) + \dots + \vec{v}_k(\lambda_k \vec{v}_k^\intercal) = \sum_{i=1}^{k} \lambda_i \vec{v}_i \vec{v}_i^\intercal, \end{aligned}$$

we have for any $j = 1, \dots, k$,

$$\begin{aligned}
\boldsymbol{A}\vec{v}_j &= (\sum_{i=1}^{k} \lambda_i \vec{v}_i \vec{v}_i^{\intercal})\vec{v}_j \\
&= \sum_{i=1}^{k} \lambda_i \vec{v}_i (\vec{v}_i^{\intercal} \vec{v}_j) \\
&= \sum_{i \neq j} \lambda_i \vec{v}_i (\vec{v}_i^{\intercal} \vec{v}_j) + \lambda_j \vec{v}_j (\vec{v}_j^{\intercal} \vec{v}_j) \\
&= \lambda_j \vec{v}_j (\vec{v}_j^{\intercal} \vec{v}_j) = \lambda_j \vec{v}_j.
\end{aligned}$$

■

Example

Let $\boldsymbol{A} = \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix}$. Give the eigenvalues and corresponding unit vectors.
Solution: Let

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} 1-\lambda & -5 \\ -5 & 1-\lambda \end{vmatrix} = 0$$

which is equivalent to

$$(1-\lambda)^2 - 25 = 0 \Leftrightarrow (\lambda+4)(\lambda-6) = 0 \Leftrightarrow \lambda_1 = -4,\ \lambda_2 = 6.$$

For $\lambda_1 = -4$, to find the eigenvector, we let

$$(\boldsymbol{A} - \lambda_1 \boldsymbol{I})\, \vec{v}_1 = \vec{0}.$$

To solve this linear equation, consider the augmented matrix:

$$\left[\begin{array}{cc|c} 5 & -5 & 0 \\ -5 & 5 & 0 \end{array}\right]$$

We will implement Gaussian elimination. By elementary row operations, this augmented matrix is modified to the reduced row echelon form:

$$\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

The solutions are

$$t \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ t \in \mathbb{R}.$$

To choose a unit solution, we let $\vec{v}_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$.
For $\lambda_2 = 6$, to find the eigenvector, we let

$$(\boldsymbol{A} - \lambda_2 \boldsymbol{I})\, \vec{v}_2 = \vec{0}.$$

To solve this linear equation, consider the augmented matrix:

$$\left[\begin{array}{cc|c} -5 & -5 & 0 \\ -5 & -5 & 0 \end{array}\right]$$

We will implement Gaussian elimination. By elementary row operations, this augmented matrix is modified to the reduced row echelon form:

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

The solutions are

$$t \begin{bmatrix} 1 \\ -1 \end{bmatrix},\ t \in \mathbb{R}.$$

To choose a unit solution, we let $\vec{v}_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$.

To summarize, we get two eigenpairs $(\lambda_1, \vec{v}_1)$ and $(\lambda_2, \vec{v}_2)$. It is also easy to verify that $\vec{v}_1^\top \vec{v}_2 = 0$.

From this example, we have

$$\begin{aligned} \lambda_1 \vec{v}_1 \vec{v}_1^\top + \lambda_2 \vec{v}_2 \vec{v}_2^\top &= -4 \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} + 6 \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} \\ &= -4 \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} + 6 \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix} = \boldsymbol{A}. \end{aligned}$$

Spectral decomposition and matrix inverse

Suppose $\boldsymbol{A} \in \mathbb{R}^{k \times k}$ is symmetric matrix. Let

$$\boldsymbol{A} = \lambda_1 \vec{v}_1 \vec{v}_1^\intercal + \ldots + \lambda_k \vec{v}_k \vec{v}_k^\intercal = \boldsymbol{P} \boldsymbol{\Lambda} \boldsymbol{P}^\intercal$$

be the spectral decomposition. Then $\boldsymbol{A}$ is invertible if and only if $\lambda_1, \ldots \lambda_k$ are all nonzero, and

$$\boldsymbol{A}^{-1} = \frac{1}{\lambda_1} \vec{v}_1 \vec{v}_1^\intercal + \ldots + \frac{1}{\lambda_k} \vec{v}_k \vec{v}_k^\intercal = \boldsymbol{P} \boldsymbol{\Lambda}^{-1} \boldsymbol{P}^\intercal,$$

where it's obvious that

$$\boldsymbol{\Lambda}^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & & \\ & \ddots & \\ & & \frac{1}{\lambda_k} \end{bmatrix}.$$

Proof

$$\begin{aligned} (\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\intercal)\left(\boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^\intercal\right) &= \boldsymbol{P}\boldsymbol{\Lambda}(\boldsymbol{P}^\intercal\boldsymbol{P})\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^\intercal \\ &= \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{I}_k\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^\intercal \\ &= \boldsymbol{P}(\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{-1})\boldsymbol{P}^\intercal \\ &= \boldsymbol{P}\boldsymbol{P}^\intercal = \boldsymbol{I}_k. \end{aligned}$$

■

Recall that we have derived the spectral decomposition

$$\begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix} = -4 \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} + 6 \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}$$

This implies that

$$\begin{aligned} \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix}^{-1} &= -\frac{1}{4} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} + \frac{1}{6} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} \\ &= -\frac{1}{8} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{1}{12} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{24} & -\frac{5}{24} \\ -\frac{5}{24} & -\frac{1}{24} \end{bmatrix} \end{aligned}$$

This is consistent with

$$\begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix}^{-1} = \frac{1}{1(1)-(-5)(-5)} \begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix} = -\frac{1}{24} \begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}.$$

Spectral decomposition and matrix square-root

Positive definite matrix Suppose the symmetric matrix $\boldsymbol{A}$ has spectral decomposition

$$\boldsymbol{A} = \lambda_1 \vec{v}_1 \vec{v}_1^\top + \ldots + \lambda_k \vec{v}_k \vec{v}_k^\top.$$

Then $\boldsymbol{A}$ is said to be positive semidefinite if and only if $\lambda_1, \ldots, \lambda_k \geq 0$, and $\boldsymbol{A}$ is said to be positive definite if and only if $\lambda_1, \ldots, \lambda_k > 0$.

Matrix square-root Suppose $\boldsymbol{A}$ is positive definite matrix with spectral decomposition

$$\boldsymbol{A} = \lambda_1 \vec{v}_1 \vec{v}_1^\top + \ldots + \lambda_k \vec{v}_k \vec{v}_k^\top = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\top.$$

For

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix}$$

we define

$$\boldsymbol{\Lambda}^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_k} \end{bmatrix}.$$

Then define the square root of $\boldsymbol{A}$ as

$$\boldsymbol{A}^{\frac{1}{2}} = \sqrt{\lambda_1} \vec{v}_1 \vec{v}_1^\top + \ldots + \sqrt{\lambda_k} \vec{v}_k \vec{v}_k^\top = \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^\top.$$

Properties

1. $\boldsymbol{A}^{\frac{1}{2}}$ is symmetric and $\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^\top$ is its spectral decomposition;
2. $\boldsymbol{A}^{\frac{1}{2}}\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{A}$;
3. Denote $\boldsymbol{A}^{-\frac{1}{2}} = \left(\boldsymbol{A}^{\frac{1}{2}}\right)^{-1}$. Then

$$\boldsymbol{A}^{-\frac{1}{2}} = \sum_{i=1}^{k} \frac{1}{\sqrt{\lambda_i}} \vec{v}_i \vec{v}_i^\top = \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^\top;$$

4. $\boldsymbol{A}^{-\frac{1}{2}}\boldsymbol{A}^{-\frac{1}{2}} = \boldsymbol{A}^{-1}$.

Proof 1.

$$\left(\boldsymbol{A}^{\frac{1}{2}}\right)^\intercal = (\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^\intercal)^\intercal = (\boldsymbol{P}^\intercal)^\intercal(\boldsymbol{\Lambda}^{\frac{1}{2}})^\intercal \boldsymbol{P}^\intercal)^\intercal = \boldsymbol{A}^{\frac{1}{2}}$$

2.

$$\begin{aligned} \boldsymbol{A}^{\frac{1}{2}}\boldsymbol{A}^{\frac{1}{2}} &= \left(\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^\intercal\right)\left(\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^\intercal\right) \\ &= \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}(\boldsymbol{P}^\intercal\boldsymbol{P})\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^\intercal \\ &= \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{I}_k\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^\intercal \\ &= \boldsymbol{P}(\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}})\boldsymbol{P}^\intercal \\ &= \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\intercal = \boldsymbol{A}. \end{aligned}$$

3. By definition,

$$\boldsymbol{\Lambda}^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{\lambda_k}} \end{bmatrix}.$$

Then given $\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^\top$ is spectral decomposition, there holds

$$\boldsymbol{A}^{-\frac{1}{2}} = (\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^\top)^{-1} = \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^\intercal.$$

4. The same as 2. ■

Examples Example 1: Let

$$\boldsymbol{A} = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}.$$

We can calculate its eigenvalues and eigenvectors

$$\lambda_1 = 4, \quad \vec{v}_1 = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}, \quad \lambda_2 = 6, \quad \vec{v}_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

We obtain the spectral decomposition of $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

Since $\boldsymbol{A}$ is positive definite, we have

$$\boldsymbol{A}^{-\frac{1}{2}} = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{6}+3}{12} & \frac{\sqrt{6}-3}{12} \\ \frac{\sqrt{6}-3}{12} & \frac{\sqrt{6}+3}{12} \end{bmatrix}$$

Example 2: Let

$$\boldsymbol{A} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

We have obtained its spectral decomposition

$$\boldsymbol{A} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}.$$

Since $\boldsymbol{A}$ is not positive definite, its inverse square root does not exist.