Matrix Algebra Marathon

J2200071 Ryuto Saito

May 12, 2024

1 Spectral Decomposition

1.1 Definition

Definition 1.1.1. *Let $\boldsymbol{A} \in \mathbb{S}^n$. Denote the spectral decomposition of $\boldsymbol{A} \in \mathbb{S}^n$ by $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}}$. Let λ_i be the i-th diagonal entry of $\boldsymbol{\Lambda}$. The matrix power of $\boldsymbol{A}$ by $p \in \mathbb{R}$ is defined as*

$$\boldsymbol{A}^p := \boldsymbol{U}\boldsymbol{\Lambda}^p\boldsymbol{U}^{\mathrm{T}}$$

where

$$\boldsymbol{\Lambda}^p = diag([\lambda_1^p, \ldots, \lambda_n^p]).$$

1.2 Theorem

Theorem 2.18.1.

For any $\boldsymbol{A} \in \mathbb{S}^n$, there exist $\boldsymbol{U} \in \mathbb{O}^{n \times n}$ and $\boldsymbol{\lambda} \in \mathbb{R}^n$ such that

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}} \tag{1}$$

where $\boldsymbol{\Lambda} := diag(\boldsymbol{\lambda})$. This is said to be the spectral decomposition. In this note, we assume $\lambda_1 \geq \cdots \geq \lambda_n$ without loss of generality. Each entry of $\boldsymbol{\lambda}$ is called an eigenvalue, and each column of $\boldsymbol{U}$ is called an eigenvector.

1.3 Exercise

Exercise 2.18.4.

Denote the spectral decomposition of $\boldsymbol{A} \in \mathbb{S}^n$ by $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}}$. Let λ_i be the i-th diagonal enrty of Λ and $\boldsymbol{u}_i$ be the i-th column in $\boldsymbol{U}$. Then, show that $\forall i \in \mathbb{N}_n$,

$$\boldsymbol{u}_i^{\mathrm{T}}\boldsymbol{A}\boldsymbol{u}_i = \lambda_i. \tag{2}$$

Proof. Since $\boldsymbol{U} \in \mathbb{O}^{n \times n}$,

$$\boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{u}_j = \begin{cases} 1 & (j = i) \\ 0 & (j \neq i) \end{cases}.$$

Thus,

$$\begin{aligned}
\boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{A} \boldsymbol{u}_i &= \boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^{\mathrm{T}} \boldsymbol{u}_i \\
&= (\boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{U}) \boldsymbol{\Lambda} (\boldsymbol{U}^{\mathrm{T}} \boldsymbol{u}_i) \\
&= (\boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{U}) \boldsymbol{\Lambda} (\boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{U})^{\mathrm{T}} \\
&= (\boldsymbol{u}_i^{\mathrm{T}} [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n]) \boldsymbol{\Lambda} (\boldsymbol{u}_i^{\mathrm{T}} [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n])^{\mathrm{T}} \\
&= ([\boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{u}_1, \ldots, \boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{u}_n]) \boldsymbol{\Lambda} ([\boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{u}_1, \ldots, \boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{u}_n])^{\mathrm{T}} \\
&= \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{\Lambda} (\boldsymbol{e}_i^{\mathrm{T}})^{\mathrm{T}} \\
&= \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{\Lambda} \boldsymbol{e}_i \\
&= \boldsymbol{e}_i^{\mathrm{T}} (\boldsymbol{\Lambda} \boldsymbol{e}_i) \\
&= \boldsymbol{e}_i^{\mathrm{T}} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
&= \lambda_i.
\end{aligned}$$

□

Exercise 2.18.6.

Denote the spectral decomposition of $\boldsymbol{A} \in \mathbb{S}^n$ by $\boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^{\mathrm{T}}$. Show that

$$tr(\boldsymbol{A}) = tr(\boldsymbol{\Lambda}). \tag{3}$$

Proof. Recall that $tr(\boldsymbol{BC}) = tr(\boldsymbol{CB})$. Then,

$$tr(\boldsymbol{A}) = tr(\boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^{\mathrm{T}}) = tr(\boldsymbol{U}(\boldsymbol{\Lambda} \boldsymbol{U}^{\mathrm{T}})) = tr((\boldsymbol{\Lambda} \boldsymbol{U}^{\mathrm{T}}) \boldsymbol{U}) = tr(\boldsymbol{\Lambda} \boldsymbol{U}^{\mathrm{T}} \boldsymbol{U}). \tag{4}$$

Since $\boldsymbol{U}$ is orthonormal,

$$\boldsymbol{U}^{\mathrm{T}} \boldsymbol{U} = \boldsymbol{I}_n. \tag{5}$$

Substitute (5) into (4), we get

$$tr(\boldsymbol{A}) = tr(\boldsymbol{\Lambda} \boldsymbol{U}^{\mathrm{T}} \boldsymbol{U}) = tr(\boldsymbol{\Lambda} \boldsymbol{I}_n) = tr(\boldsymbol{\Lambda}).$$

Hence, (3) holds. □

Exercise 2.18.9.

Denote the spectral decomposition of $\boldsymbol{A} \in \mathbb{S}^n$ by $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}}$. Let λ_i be the i-th diagonal entry of $\boldsymbol{\Lambda}$. For $\forall p \in \mathbb{R}$, show that

$$tr(\boldsymbol{A}^p) = \sum_{i=1}^{n} \lambda_i^p. \tag{6}$$

Proof. By definition 1.1.1, we have

$$\boldsymbol{A}^p = \boldsymbol{U}\boldsymbol{\Lambda}^p\boldsymbol{U}^{\mathrm{T}} \tag{7}$$

where

$$\boldsymbol{\Lambda}^p = diag([\lambda_1^p, \ldots, \lambda_n^p]).$$

Since $tr(\boldsymbol{BC}) = tr(\boldsymbol{CB})$,

$$\begin{aligned} tr(\boldsymbol{A}^p) &= tr(\boldsymbol{U}\boldsymbol{\Lambda}^p\boldsymbol{U}^{\mathrm{T}}) \\ &= tr(\boldsymbol{U}(\boldsymbol{\Lambda}^p\boldsymbol{U}^{\mathrm{T}})) \\ &= tr((\boldsymbol{\Lambda}^p\boldsymbol{U}^{\mathrm{T}})\boldsymbol{U}) \\ &= tr(\boldsymbol{\Lambda}^p(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U})) \end{aligned}$$

Thus,

$$tr(\boldsymbol{A}^p) = tr(\boldsymbol{\Lambda}^p(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U})). \tag{8}$$

Substitute (5) into (8), we get

$$tr(\boldsymbol{A}^p) = tr(\boldsymbol{\Lambda}^p(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U})) = tr(\boldsymbol{\Lambda}^p) = \sum_{i=1}^{n} \lambda_i^p.$$

□

2 Positive Definite Matrices

2.1 Definition

Definition 2.1.1. *An $n \times n$ symmetric matrix is said to be positive semi-definite if $\forall \boldsymbol{x} \in \mathbb{R}^n$,*

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \geq 0. \tag{9}$$

Definition 2.1.2. *An $n \times n$ symmetric matrix is said to be strictly positive definite if $\forall \boldsymbol{x} \in \mathbb{R}^n$,*

$$\boldsymbol{x} \neq \boldsymbol{0}_n \implies \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0. \tag{10}$$

Exercise 2.19.4.

For any $\boldsymbol{A} \in \mathbb{S}^n_{++}$ and $\boldsymbol{x} \in \mathbb{R}^n$, show that,

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0 \implies \boldsymbol{x} \neq \boldsymbol{0}_n. \tag{11}$$

Proof. Let $\boldsymbol{A} \in \mathbb{S}^n_{++}$ and $\boldsymbol{x} \in \mathbb{R}^n$. Suppose that $\boldsymbol{x} = \boldsymbol{0}_n$. Then,

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} = \boldsymbol{0}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{0} = 0 \leq 0.$$

Therefore, we get

$$\boldsymbol{x} = \boldsymbol{0}_n \implies \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} \leq 0.$$

That is

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} > 0 \implies \boldsymbol{x} \neq \boldsymbol{0}_n.$$

□

Exercise 2.19.6.

For $\forall \boldsymbol{A} \in \mathbb{S}^n_{++}$, every eigen-value is positive.

Proof. From Theorem 2.18.1., for any $\boldsymbol{A} \in \mathbb{S}^n$, there exist $\boldsymbol{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n] \in \mathbb{O}^{n \times n}$ and $\boldsymbol{\lambda} \in \mathbb{R}^n$ such that

$$\boldsymbol{A} = \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^{\mathrm{T}}$$

where $\boldsymbol{u}_i$ is one of the eigenvectors of $\boldsymbol{A}$ and λ_i is the eigenvalue for $\boldsymbol{u}_i$. Recall that

$$\forall i \in \mathbb{N}_n, \boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{A} \boldsymbol{u}_i = \lambda_i. \tag{12}$$

Since $\boldsymbol{A} \in \mathbb{S}^n_{++}$, we have

$$\forall \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x} \neq \boldsymbol{0}_n \implies \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} > 0. \tag{13}$$

By definition, eigenvectors are non-zero. So, from (13), we get

$$\forall i \in \mathbb{N}_n, \boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{A} \boldsymbol{u}_i > 0. \tag{14}$$

Substitute (12) into (14), then

$$\forall i \in \mathbb{N}_n, \lambda_i > 0.$$

Therefore, for $\forall \boldsymbol{A} \in \mathbb{S}^n_{++}$, every eigen-value is positive. □

Exercise 2.19.8.

Let $\boldsymbol{A}$ be an $n \times n$ symmetric matrix. Denote its spectral decomposition by $\boldsymbol{A} = \boldsymbol{U} diag(\boldsymbol{\lambda}) \boldsymbol{U}^{\mathrm{T}}$, where $\boldsymbol{U} \in \mathbb{O}^{n \times n}$ and $\boldsymbol{\lambda} \in \mathbb{R}^n$. Show that the matrix $\boldsymbol{A}$ is strictly positive definite if $\boldsymbol{\lambda} > 0$.

Proof. Let $\boldsymbol{x} \in \mathbb{R}^n$. Suppose that $\boldsymbol{x} \neq \boldsymbol{0}$. Denote by $\boldsymbol{u}_i$ i-th column in $\boldsymbol{U}$, and let $\boldsymbol{z} \coloneqq \boldsymbol{U}^{\mathrm{T}}\boldsymbol{x}$. Since $\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{I}_n$ for $\boldsymbol{U} \in \mathbb{O}^{n \times n}$ from Exercise 2.17.5.,

$$\begin{aligned}
\|\boldsymbol{z}\| &= \sqrt{\langle \boldsymbol{z}, \boldsymbol{z} \rangle} \\
&= \sqrt{\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}} \\
&= \sqrt{(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{x}} \\
&= \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{x}} \\
&= \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}} \\
&= \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} \\
&= \|\boldsymbol{x}\| \geq 0.
\end{aligned}$$

From Exercise 2.2.6., since $\boldsymbol{x} \neq \boldsymbol{0}$, we get $\|\boldsymbol{x}\| \neq 0$. Hence,

$$\|\boldsymbol{z}\| = \|\boldsymbol{x}\| > 0.$$

Since $\|\boldsymbol{z}\| \neq 0$, $\boldsymbol{z} \neq \boldsymbol{0}$. From Exercise 2.18.3.,

$$\boldsymbol{A} = \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{T}}.$$

Then,

$$\begin{aligned}
\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{x}^{\mathrm{T}} \left(\sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{T}} \right) \boldsymbol{x} \\
&= \sum_{i=1}^{n} \lambda_i \boldsymbol{x}^{\mathrm{T}} \boldsymbol{u}_i \boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{x} \\
&= \sum_{i=1}^{n} \lambda_i \langle \boldsymbol{x}, \boldsymbol{u}_i \rangle \langle \boldsymbol{u}_i, \boldsymbol{x} \rangle \\
&= \sum_{i=1}^{n} \lambda_i \langle \boldsymbol{u}_i, \boldsymbol{x} \rangle \langle \boldsymbol{u}_i, \boldsymbol{x} \rangle \\
&= \sum_{i=1}^{n} \lambda_i \langle \boldsymbol{u}_i, \boldsymbol{x} \rangle^2 \\
&= \sum_{i=1}^{n} \lambda_i z_i^2.
\end{aligned}$$

Since $\boldsymbol{\lambda} > \boldsymbol{0}$ and $\boldsymbol{z} \neq \boldsymbol{0}$,

$$\forall i \in \mathbb{N}_n, \lambda_i > 0$$

and

$$\exists k \in \mathbb{N}_n, z_k \neq 0.$$

Thus,

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{n} \lambda_i z_i^2 \geq \lambda_k z_k^2 > 0.$$

Therefore, the matrix $\boldsymbol{A}$ is strictly positive definite if $\boldsymbol{\lambda} > 0$. □

Exercise 2.19.10.

Show that the determinant of every strictly positive definite matrix is strictry positive.

Proof. Let $\boldsymbol{A} \in \mathbb{S}_{++}^n$. Denote the spectral decomposition of $\boldsymbol{A}$ by $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathbb{T}}$. Let λ_i be the i-th diagonal entry of $\boldsymbol{\Lambda}$. Recall that for $\forall \boldsymbol{A} \in \mathbb{S}_{++}^n$, every eigenvalue is positive. Then,

$$\forall i \in \mathbb{N}_n, \lambda_i > 0. \tag{15}$$

Moreover, from Exercise 2.18.7., we have

$$det(\boldsymbol{A}) = \prod_{i=1}^{n} \lambda_i. \tag{16}$$

Hence, from (15),

$$det(\boldsymbol{A}) = \prod_{i=1}^{n} \lambda_i > 0.$$

Therefore, the determinant of every strictly positive definite matrix is strictry positive. □

3 Hadamard Product

3.1 Definition

Definition 3.1.1. *Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^{m\times n}$. The Hadamard product between $\boldsymbol{X}$ and $\boldsymbol{Y}$, denote by $\boldsymbol{Z}$, is defined as*

$$\boldsymbol{Z} = \boldsymbol{X} \bigodot \boldsymbol{Y} \iff \forall i \in \mathbb{N}_m, \forall j \in \mathbb{N}_n, Z_{i,j} = X_{i,j}Y_{i,j}. \tag{17}$$

3.2 Exercise

Exercise 2.20.3.

Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Show the equalities:

$$\boldsymbol{x} \bigodot \boldsymbol{y} = diag(\boldsymbol{x})\boldsymbol{y} = diag(\boldsymbol{y})\boldsymbol{x} = diag(\boldsymbol{x} \bigodot \boldsymbol{y})\mathbf{1}_n. \tag{18}$$

Proof. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. By the definition of Hadamard product,

$$\boldsymbol{x} \bigodot \boldsymbol{y} = \begin{bmatrix} x_1 y_1 \\ \vdots \\ x_n y_n \end{bmatrix}.$$

Moreover,

$$diag(\boldsymbol{x})\boldsymbol{y} = \begin{bmatrix} x_1 & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 \\ \vdots \\ x_n y_n \end{bmatrix},$$

$$diag(\boldsymbol{x})\boldsymbol{y} = \begin{bmatrix} y_1 & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & y_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 x_1 \\ \vdots \\ y_n x_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 \\ \vdots \\ x_n y_n \end{bmatrix},$$

and

$$diag(\boldsymbol{x} \bigodot \boldsymbol{y})\boldsymbol{1}_n = diag(\begin{bmatrix} x_1 y_1 \\ \vdots \\ x_n y_n \end{bmatrix})\boldsymbol{1}_n = \begin{bmatrix} x_1 y_1 & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & x_n y_n \end{bmatrix} \boldsymbol{1}_n = \begin{bmatrix} x_1 y_1 \\ \vdots \\ x_n y_n \end{bmatrix}.$$

From these equalities, we get (18). □

4 Vec Operator

4.1 Definition

Definition 4.1.1. *Let* $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n] \in \mathbb{R}^{m \times n}$*. The vec operator is defined as*

$$vec(\boldsymbol{A}) \coloneqq \begin{bmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{bmatrix}. \tag{19}$$

4.2 Exercise

Exercise 2.21.4.

Let $\boldsymbol{X} \in \mathbb{R}^{m \times n}$. Show that

$$\|\boldsymbol{X}\|_{\mathrm{F}} = \|vec(\boldsymbol{X})\|. \tag{20}$$

Proof. Let $\boldsymbol{X} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n] \in \mathbb{R}^{m \times n}$. By the definition of Frobenius norm,

$$\|\boldsymbol{X}\|_{\mathrm{F}} = \sqrt{\langle \boldsymbol{X}, \boldsymbol{X} \rangle} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} X_{i,j}^2}.$$

Furthermore,

$$\begin{aligned}
\|vec(\boldsymbol{X})\| &= \sqrt{\langle vec(\boldsymbol{X}), vec(\boldsymbol{X})\rangle} \\
&= \sqrt{vec(\boldsymbol{X})^{\mathrm{T}} vec(\boldsymbol{X})} \\
&= \sqrt{[\boldsymbol{x}_1^{\mathrm{T}}, \cdots, \boldsymbol{x}_n^{\mathrm{T}}] \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix}} \\
&= \sqrt{\sum_{j=1}^{n} \boldsymbol{x}_j^{\mathrm{T}} \boldsymbol{x}_j}.
\end{aligned}$$

Thus, we get

$$\|vec(\boldsymbol{X})\| = \sqrt{\sum_{j=1}^{n} \boldsymbol{x}_j^{\mathrm{T}} \boldsymbol{x}_j} \tag{21}$$

Since $\boldsymbol{x}_j$ is j-th column in $\boldsymbol{X}$,

$$\boldsymbol{x}_j^{\mathrm{T}} \boldsymbol{x}_j = \sum_{i=1}^{m} X_{i,j}^2. \tag{22}$$

Substitute (22) into (21), we get

$$\begin{aligned}
\|vec(\boldsymbol{X})\| &= \sqrt{\sum_{j=1}^{n} \left(\sum_{i=1}^{m} X_{i,j}^2 \right)} \\
&= \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} X_{i,j}^2}.
\end{aligned}$$

Hence, (20) holds. □