Jared Yu 914640019

HW 2 STA 135

1. Suppose $\boldsymbol{A} \in \mathbb{R}^{k \times k}$ can be written as $\boldsymbol{A} = \boldsymbol{P\Lambda P}^T$ where, $\boldsymbol{P} = [\vec{v}_1 \cdots \vec{v}_k]$ is an orthogonal matrix and $\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix}$ is a diagonal matrix.

   To show that $\boldsymbol{A}$ is a symmetric matrix, we must show that $\boldsymbol{A} = \boldsymbol{A}^T$**.** Since $\boldsymbol{A} = \boldsymbol{P\Lambda P}^T$, then $\boldsymbol{A}^T = (\boldsymbol{P\Lambda P}^T)^T = (\boldsymbol{P}^T)^T\boldsymbol{\Lambda}^T\boldsymbol{P}^T = \boldsymbol{P\Lambda}^T\boldsymbol{P}^T$. Also, $\boldsymbol{\Lambda}$ is a diagonal matrix, so the transpose $\boldsymbol{\Lambda}^T = \boldsymbol{\Lambda}$. Therefore, $\boldsymbol{P\Lambda}^{\mathrm{T}}\boldsymbol{P}^T = \boldsymbol{P\Lambda P}^T$ and so $\boldsymbol{A}^T = \boldsymbol{A}$ or in other words $\boldsymbol{A} = \boldsymbol{A}^T$. We can then conclude that $\boldsymbol{A}$ is a symmetric matrix. ■

   It has been shown that $\boldsymbol{A}$ is a $k \times k$ square symmetric matrix. This implies that $\boldsymbol{A}$ has $k$ pairs of eigenvalues and eigenvectors namely,

$$\lambda_1, e_1 \quad \lambda_2, e_2 \quad \cdots \quad \lambda_k, e_k$$

   where the eigenvectors satisfy $1 = e_1' e_1 = \cdots = e_k' e_k$ $(i)$ and are mutually perpendicular $(ii)$. It must then be shown that $\lambda_1, \cdots, \lambda_k$ are eigenvalues of $\boldsymbol{A}$ and $\vec{v}_1, \cdots, \vec{v}_k$ are corresponding eigenvectors where both $(i)$ and $(ii)$ are also satisfied.

   First, the breakdown of matrix $\boldsymbol{A}$ into eigenvalues $\lambda_1, \cdots, \lambda_k$ and eigenvectors $\vec{v}_1, \cdots, \vec{v}_k$ are shown below.

$$\begin{aligned} \boldsymbol{A} &= \boldsymbol{P\Lambda P}^T \\ &= [\vec{v}_1 \cdots \vec{v}_k] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_k^T \end{bmatrix} \\ &= [\vec{v}_1 \cdots \vec{v}_k] \begin{bmatrix} \lambda_1 \vec{v}_1^T \\ \vdots \\ \lambda_k \vec{v}_k^T \end{bmatrix} \\ &= \lambda_1 \vec{v}_1 \vec{v}_1^T + \lambda_2 \vec{v}_2 \vec{v}_2^T + \cdots + \lambda_k \vec{v}_k \vec{v}_k^T \end{aligned}$$

   It can be demonstrated that these are the eigenvalues and eigenvectors for $\boldsymbol{A}$. Using the example of $i = 1$, if the left-hand side and right-hand side of the equation is multiplied by $\vec{v}_1$, there will be the following result,

$$\begin{aligned} \boldsymbol{A}\vec{v}_1 &= \lambda_1 \vec{v}_1 \underbrace{\vec{v}_1^T \vec{v}_1}_{1} + \sum_{i=2}^{k} \lambda_i \vec{v}_i \underbrace{\vec{v}_i^T \vec{v}_1}_{0} \\ &= \lambda_1 \vec{v}_1. \end{aligned}$$

This result follows because it was stated that $\boldsymbol{P}$ is an orthogonal matrix and so $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P}' = \boldsymbol{I}$. Therefore, the inner product of the vectors within $[\vec{v}_1 \cdots \vec{v}_k]$ and $\begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_k^T \end{bmatrix}$ satisfy $1 = \vec{v}_1^T\vec{v}_1 = \cdots = \vec{v}_k^T\vec{v}_k$ and $0 = \vec{v}_i^T\vec{v}_j$ for $i \neq j$. Then to generalize to all the terms, for $i = 1,2,\cdots,k$,

$$\boldsymbol{A}\vec{v}_i = \sum_{j=1}^{k} \lambda_j \vec{v}_j \vec{v}_j^T \vec{v}_i$$

$$= \lambda_i \vec{v}_i \underbrace{\vec{v}_i^T \vec{v}_i}_{1} + \underbrace{0}_{\text{when } j\neq i}$$

$$= \lambda_i \vec{v}_i$$

The scalar $\lambda_i$ are the eigenvalues and $\vec{v}_i$ are the eigenvectors, because they follow the following formula for finding eigenvalues and eigenvectors from a matrix $\boldsymbol{A}$,

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\Lambda}\boldsymbol{x}.$$

We can then conclude that $\lambda_1, \cdots, \lambda_k$ are eigenvalues of $\boldsymbol{A}$ and $\vec{v}_1, \cdots, \vec{v}_k$ are corresponding eigenvectors.■

2. Suppose $\boldsymbol{A} \in \mathbb{R}^{k\times k}$ is a symmetric matrix, then let

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

be the spectral decomposition. If $\lambda_1, \cdots, \lambda_k$ are nonzero, it will be shown that

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

where

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

Given that

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T,$$

then the inverse,

$$\boldsymbol{A}^{-1} = (\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T)^{-1} = (\boldsymbol{P}^T)^{-1}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^{-1}.$$

We can assume that the inverse of $\boldsymbol{A}$ exists because the eigenvalues are nonzero. Since $\boldsymbol{P}$ is an orthogonal matrix, then $\boldsymbol{P}^T = \boldsymbol{P}^{-1}$ and so,

$$(\boldsymbol{P}^T)^{-1}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^{-1} = (\boldsymbol{P}^{-1})^{-1}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T.$$

Also, the inverse of an inverse of a matrix is the original matrix itself so,

$$(\boldsymbol{P}^{-1})^{-1}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T = \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T.$$

$$\boldsymbol{A}^{-1} = \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T$$

$$= [\vec{v}_1 \cdots \vec{v}_k]\begin{bmatrix}\frac{1}{\lambda_1} & & \\ & \ddots & \\ & & \frac{1}{\lambda_k}\end{bmatrix}\begin{bmatrix}\vec{v}_1^T \\ \vdots \\ \vec{v}_k^T\end{bmatrix}$$

$$= [\vec{v}_1 \cdots \vec{v}_k]\begin{bmatrix}\frac{\vec{v}_1^T}{\lambda_1} \\ \vdots \\ \frac{\vec{v}_k^T}{\lambda_k}\end{bmatrix}$$

$$= \frac{1}{\lambda_1}\vec{v}_1\vec{v}_1^T + \cdots + \frac{1}{\lambda_k}\vec{v}_k\vec{v}_k^T.$$

Then it can be said that,

$$\boldsymbol{A}^{-1} = \frac{1}{\lambda_1}\vec{v}_1\vec{v}_1^T + \cdots + \frac{1}{\lambda_k}\vec{v}_k\vec{v}_k^T = \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T. \blacksquare$$

3. Let $\boldsymbol{A} = \sum_{i=1}^{k} \lambda_i \vec{v}_i \vec{v}_i^T$ be the spectral decomposition with positive eigenvalues $\lambda_1, \cdots, \lambda_k > 0$. Then, setting

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

the following properties will be proven:

a) $\boldsymbol{A}^{\frac{1}{2}}$ is symmetric and $\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T$ is its spectral decomposition;

To show that $\boldsymbol{A}^{\frac{1}{2}}$ is symmetric, we must show that $\boldsymbol{A}^{\frac{1}{2}} = \left(\boldsymbol{A}^{\frac{1}{2}}\right)^T$. The square root of $\boldsymbol{A}$ is denoted by $\boldsymbol{A}^{\frac{1}{2}}$ and can be written out as follows,

$$\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T,$$

where

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

Then the transpose of $\boldsymbol{A}^{\frac{1}{2}}$ is as follows,

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

Since $\boldsymbol{\Lambda}^{\frac{1}{2}}$ is a diagonal matrix, then the transpose of it is itself, therefore,

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

It is now concluded that $\boldsymbol{A}^{\frac{1}{2}}$ is symmetric because $\boldsymbol{A}^{\frac{1}{2}} = \left(\boldsymbol{A}^{\frac{1}{2}}\right)^T$. ■

Next, given that $\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T$ we can write,

$$\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T = [\vec{v}_1 \quad \cdots \quad \vec{v}_k]\begin{bmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_k}\end{bmatrix}\begin{bmatrix}\vec{v}_1^T \\ \vdots \\ \vec{v}_k^T\end{bmatrix} = [\vec{v}_1 \quad \cdots \quad \vec{v}_k]\begin{bmatrix}\sqrt{\lambda_1}\vec{v}_1^T \\ \vdots \\ \sqrt{\lambda_k}\vec{v}_k^T\end{bmatrix}$$
$$= \sqrt{\lambda_1}\vec{v}_1\vec{v}_1^T + \cdots + \sqrt{\lambda_k}\vec{v}_k\vec{v}_k^T = \sum_{i=1}^{k}\sqrt{\lambda_i}\,\vec{v}_i\vec{v}_i^T.$$

Then to show that $\sqrt{\lambda_i}' s$ are the eigenvalues and $\vec{v}_i$ are the corresponding eigenvectors, it will be demonstrated that for $j = 1, \cdots, k$,

$$\boldsymbol{A}^{\frac{1}{2}}v_j = \sum_{i=1}^{k}\sqrt{\lambda_i}\,\vec{v}_i\vec{v}_i^T\,\vec{v}_j$$
$$\boldsymbol{A}^{\frac{1}{2}}v_j = \sqrt{\lambda_j\vec{v}_j\vec{v}_j^T\,\vec{v}_j} + 0$$

The above follows since it was stated that $\boldsymbol{P}$ is an orthogonal matrix and so $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P}' = \boldsymbol{I}$. Therefore, the inner product of the vectors within $[\vec{v}_1 \cdots \vec{v}_k]$ and $\begin{bmatrix}\vec{v}_1^T \\ \vdots \\ \vec{v}_k^T\end{bmatrix}$ satisfy $1 = \vec{v}_1^T\vec{v}_1 = \cdots = \vec{v}_k^T\vec{v}_k$ and $0 = \vec{v}_i^T\,\vec{v}_j$ for $i \neq j$. It then continues that,

$$\boldsymbol{A}^{\frac{1}{2}}v_j = \sqrt{\lambda_j\vec{v}_j}.$$

The scalar $\sqrt{\lambda_j}$ are the eigenvalues and $\vec{v}_j$ are the eigenvectors, because they follow the following formula for finding eigenvalues and eigenvectors from a matrix $\boldsymbol{A}$,

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\Lambda}\boldsymbol{x}.$$

Therefore, $\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T = \sum_{i=1}^{k}\sqrt{\lambda_i}\,\vec{v}_i\vec{v}_i^T$ is the spectral decomposition of $\boldsymbol{A}^{\frac{1}{2}}$. ■

b) $\boldsymbol{A}^{\frac{1}{2}}\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{A}$;
To prove the above, it will be shown that,

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

To show that $\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}} = \boldsymbol{\Lambda}$,

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

It has now been shown that $\boldsymbol{A}^{\frac{1}{2}}\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{A}$. ■

c) Denote $\boldsymbol{A}^{-\frac{1}{2}} = \left(\boldsymbol{A}^{\frac{1}{2}}\right)^{-1}$. Then

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

where

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

So, given that,

$$\boldsymbol{A}^{-\frac{1}{2}} = \left(\boldsymbol{A}^{\frac{1}{2}}\right)^{-1},$$

the following can also be stated,

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

The step $(\boldsymbol{P}^T)^{-1} = (\boldsymbol{P}^{-1})^{-1}$ follows from the fact that $P$ is an orthogonal matrix and $\boldsymbol{P}^T = \boldsymbol{P}^{-1}$. Also, $\left(\boldsymbol{\Lambda}^{\frac{1}{2}}\right)^{-1} = \boldsymbol{\Lambda}^{-\frac{1}{2}}$, since $\boldsymbol{\Lambda}^{\frac{1}{2}}$ is a diagonal matrix with elements $\sqrt{\lambda_i}$, so the inverse of this diagonal matrix is $\boldsymbol{\Lambda}^{-\frac{1}{2}}$ where the elements are $\frac{1}{\sqrt{\lambda_i}}$.

Next, it can be shown that,

$$\boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^{-1} = [\vec{v}_1 \quad \cdots \quad \vec{v}_k]\begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{\lambda_k}} \end{bmatrix}\begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_k^T \end{bmatrix} = [\vec{v}_1 \quad \cdots \quad \vec{v}_k]\begin{bmatrix} \frac{\vec{v}_1^T}{\sqrt{\lambda_1}} \\ \vdots \\ \frac{\vec{v}_k^T}{\sqrt{\lambda_k}} \end{bmatrix} = \frac{1}{\sqrt{\lambda_1}}\vec{v}_1\vec{v}_1^T + \cdots + \frac{1}{\sqrt{\lambda_k}}\vec{v}_k\vec{v}_k^T$$

$$= \sum_{i=1}^{k} \frac{1}{\sqrt{\lambda_i}} \vec{v}_i \vec{v}_i^T. \blacksquare$$

d) $\boldsymbol{A}^{-\frac{1}{2}}\boldsymbol{A}^{-\frac{1}{2}} = \boldsymbol{A}^{-1}$

So, given $\boldsymbol{A}^{-\frac{1}{2}}\boldsymbol{A}^{-\frac{1}{2}}$, it can be written that,

$$\boldsymbol{A}^{-\frac{1}{2}}\boldsymbol{A}^{-\frac{1}{2}} = \left(\boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T\right)\left(\boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T\right) = \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{I}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T = \boldsymbol{P}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P}^T = \boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^T = \boldsymbol{A}^{-1}.$$

To show that $\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Lambda}^{-\frac{1}{2}} = \boldsymbol{\Lambda}^{-1}$,

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

It has now been shown that $\boldsymbol{A}^{-\frac{1}{2}}\boldsymbol{A}^{-\frac{1}{2}} = \boldsymbol{A}^{-1}$. ■

4. Considering a $p$-variate sample that is of size $n$:

$$\vec{x}_1, \cdots, \vec{x}_n.$$

Then for some $\boldsymbol{C} \in \mathbb{R}^{q\times p}$ and $\vec{a} \in \mathbb{R}^q$, the following linear transformation is given,

$$\vec{y}_i = \boldsymbol{C}\vec{x}_i + \vec{a}, \quad i = 1, \cdots, n.$$

Also, for some $\boldsymbol{D} \in \mathbb{R}^{r\times p}$ and $\vec{b} \in \mathbb{R}^r$, the following linear transformation is given,

$$z_i = \boldsymbol{D}\vec{x}_i + \vec{b}, \quad i = 1, \cdots, n.$$

For any $j = 1, \cdots, q$ and $k = 1, \cdots, r$, denote by $s_{Y_jZ_k}$, the sample covariance between $\{y_{ij}\}_{i=1}^n$ and $\{z_{ik}\}_{i=1}^n$. Define the matrix

$$\boldsymbol{S}_{\vec{Y},\vec{Z}} = \begin{bmatrix} s_{Y_1,Z_1} & s_{Y_1,Z_2} & \cdots & s_{Y_1,Z_r} \\ s_{Y_2,Z_1} & s_{Y_2,Z_2} & \cdots & s_{Y_2,Z_r} \\ \vdots & \vdots & \ddots & \vdots \\ s_{Y_q,Z_1} & s_{Y_q,Z_2} & \cdots & s_{Y_q,Z_r} \end{bmatrix}.$$

It must be shown then that

$$\boldsymbol{S}_{\vec{Y},\vec{Z}} = \boldsymbol{C}\boldsymbol{S}_{\vec{X}}\boldsymbol{D}^T,$$

where $\boldsymbol{S}_{\vec{X}}$ is the sample covariance matrix of $\vec{x}_1, \cdots, \vec{x}_n$.

The sample covariance matrix $\boldsymbol{S}_{\vec{X}}$ can be expressed as follows,

$$\boldsymbol{S}_{\vec{X}} = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,p} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p,1} & s_{p,2} & \cdots & s_{p,p} \end{bmatrix},$$

where

$$s_j^2 = s_{jj} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2$$

and

$$s_{jk} = s_{kj} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)\left(x_{ik} - \bar{x}_k\right).$$

It will be further expressed using different notation. First, the sample mean vector $\bar{\vec{x}}$ will be shown below, the $\boldsymbol{X}$ matrix is as follows:

$$\boldsymbol{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix}$$

Then the sample mean vector can be seen below:

$$\bar{\vec{x}} = \begin{bmatrix} \bar{\vec{x}}_1 \\ \bar{\vec{x}}_2 \\ \vdots \\ \bar{\vec{x}}_p \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} x_{i1} \\ \frac{1}{n}\sum_{i=1}^{n} x_{i2} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n} x_{ip} \end{bmatrix} = \frac{1}{n}\sum_{i=1}^{n}\begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{bmatrix} = \frac{1}{n}\sum_{i=1}^{n}\vec{x}_i$$

Then sample covariance matrix $\boldsymbol{S}_{\vec{X}}$ can be rewritten as,

$$
\begin{aligned}
&\boldsymbol{S}_{\vec{X}} \\
&= \begin{bmatrix}
\frac{1}{n-1}\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2 & \frac{1}{n-1}\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2) & \cdots & \frac{1}{n-1}\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2) \\
\frac{1}{n-1}\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)(x_{i1}-\bar{x}_1) & \frac{,1}{n-1}\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)^2 & \cdots & \frac{1}{n-1}\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)(x_{ip}-\bar{x}_p) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{n-1}\sum_{i=1}^{n}(x_{ip}-\bar{x}_p)(x_{i1}-\bar{x}_1) & \frac{1}{n-1}\sum_{i=1}^{n}(x_{ip}-\bar{x}_p)(x_{i2}-\bar{x}_2) & \cdots & \frac{1}{n-1}\sum_{i=1}^{n}(x_{ip}-\bar{x}_p)^2
\end{bmatrix} \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\begin{bmatrix}
(x_{i1}-\bar{x}_1)^2 & (x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2) & \cdots & (x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2) \\
(x_{i2}-\bar{x}_2)(x_{i1}-\bar{x}_1) & (x_{i2}-\bar{x}_2)^2 & \cdots & (x_{i2}-\bar{x}_2)(x_{ip}-\bar{x}_p) \\
\vdots & \vdots & \ddots & \vdots \\
(x_{ip}-\bar{x}_p)(x_{i1}-\bar{x}_1) & (x_{ip}-\bar{x}_p)(x_{i2}-\bar{x}_2) & \cdots & (x_{ip}-\bar{x}_p)^2
\end{bmatrix} \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\begin{bmatrix} x_{i1}-\bar{x}_1 \\ x_{i2}-\bar{x}_2 \\ \vdots \\ x_{ip}-\bar{x}_p \end{bmatrix}\begin{bmatrix} x_{i1}-\bar{x}_1 & x_{i2}-\bar{x}_2 & \cdots & x_{ip}-\bar{x}_p \end{bmatrix} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\vec{x}_i-\bar{\vec{x}}\right)\left(\vec{x}_i-\bar{\vec{x}}\right)^T \\
&= \frac{1}{n-1}\left[\left(\vec{x}_1-\bar{\vec{x}}\right)\left(\vec{x}_1-\bar{\vec{x}}\right)^T + \cdots + \left(\vec{x}_n-\bar{\vec{x}}\right)\left(\vec{x}_n-\bar{\vec{x}}\right)^T\right] \\
&= \frac{1}{n-1}\begin{bmatrix} \vec{x}_1-\bar{\vec{x}} & \vec{x}_2-\bar{\vec{x}} & \cdots & \vec{x}_n-\bar{\vec{x}} \end{bmatrix}\begin{bmatrix} \left(\vec{x}_1-\bar{\vec{x}}\right)^T \\ \left(\vec{x}_2-\bar{\vec{x}}\right)^T \\ \vdots \\ \left(\vec{x}_n-\bar{\vec{x}}\right)^T \end{bmatrix}
\end{aligned}
$$

NOTE:

$$
\begin{bmatrix} \left(\vec{x}_1-\bar{\vec{x}}\right)^T \\ \left(\vec{x}_2-\bar{\vec{x}}\right)^T \\ \vdots \\ \left(\vec{x}_n-\bar{\vec{x}}\right)^T \end{bmatrix} = \begin{bmatrix} \vec{x}_1^T \\ \vdots \\ \vec{x}_n^T \end{bmatrix} - \begin{bmatrix} \bar{\vec{x}}^T \\ \vdots \\ \bar{\vec{x}}^T \end{bmatrix}
$$

Then,

$$
\boldsymbol{X} - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}\bar{\vec{x}}^T = \boldsymbol{X} - 1_n\bar{\vec{x}}^T.
$$

$$
\boldsymbol{S}_{\vec{X}} = \frac{1}{n-1}\left(\boldsymbol{X} - 1_n\bar{\vec{x}}^T\right)^T\left(\boldsymbol{X} - 1_n\bar{\vec{x}}^T\right).
$$

Next, $\boldsymbol{C}\boldsymbol{S}_{\vec{X}}\boldsymbol{D}^T$ will be expanded:

$$
\begin{aligned}
\boldsymbol{C}\boldsymbol{S}_{\vec{X}}\boldsymbol{D}^T &= \frac{1}{n-1}\boldsymbol{C}\left(\boldsymbol{X} - 1_n\bar{\vec{x}}^T\right)^T\left(\boldsymbol{X} - 1_n\bar{\vec{x}}^T\right)\boldsymbol{D}^T \\
&= \frac{1}{n-1}\boldsymbol{C}\left(\boldsymbol{X}^T - \bar{\vec{x}}1_n^T\right)\left[\boldsymbol{D}\left(\boldsymbol{X} - 1_n\bar{\vec{x}}^T\right)^T\right]^T
\end{aligned}
$$

$$= \frac{1}{n-1}\left[\boldsymbol{C}\left(\boldsymbol{X}^T - \bar{\vec{x}}1_n^T\right)\right]\left[\boldsymbol{D}\left(\boldsymbol{X}^T - \bar{\vec{x}}1_n^T\right)\right]^T$$

$$= \frac{1}{n-1}\left(\boldsymbol{C}\boldsymbol{X}^T - \boldsymbol{C}\bar{\vec{x}}1_n^T\right)\left(\boldsymbol{D}\boldsymbol{X}^T - \boldsymbol{D}\bar{\vec{x}}1_n^T\right)^T$$

$$= \frac{1}{n-1}\{\boldsymbol{C}[\vec{x}_1 \quad \vec{x}_2 \quad \cdots \quad \vec{x}_n] - \boldsymbol{C}[\bar{\vec{x}} \quad \bar{\vec{x}} \quad \cdots \quad \bar{\vec{x}}]\}\{\boldsymbol{D}[\vec{x}_1 \quad \vec{x}_2 \quad \cdots \quad \vec{x}_n] - \boldsymbol{D}[\bar{\vec{x}} \quad \bar{\vec{x}} \quad \cdots \quad \bar{\vec{x}}]\}^T$$

$$= \frac{1}{n-1}\{[\boldsymbol{C}\vec{x}_1 + \vec{a} \quad \boldsymbol{C}\vec{x}_2 + \vec{a} \quad \cdots \quad \boldsymbol{C}\vec{x}_n + \vec{a}] - [\boldsymbol{C}\bar{\vec{x}} + \vec{a} \quad \boldsymbol{C}\bar{\vec{x}} + \vec{a} \quad \cdots \quad \boldsymbol{C}\bar{\vec{x}} + \vec{a}]\}\left\{[\boldsymbol{D}\vec{x}_1 + \vec{b} \quad \boldsymbol{D}\vec{x}_2 + \vec{b} \quad \cdots \quad \boldsymbol{D}\vec{x}_n + \vec{b}] - [\boldsymbol{D}\bar{\vec{x}} + \vec{b} \quad \boldsymbol{D}\bar{\vec{x}} + \vec{b} \quad \cdots \quad \boldsymbol{D}\bar{\vec{x}} + \vec{b}]\right\}^T$$

Above, the $a's$ and $b's$ are added then subtracted back, so it is equivalent to adding $0$. Then, let the following be convenient notations:

$$\bar{\vec{y}} = \frac{1}{n}\sum_{i=1}^{n} \vec{y}_i \text{ and } \bar{\vec{z}} = \frac{1}{n}\sum_{i=1}^{n} \vec{z}_i$$

$$= \frac{1}{n-1}\{[\vec{y}_1 \quad \vec{y}_2 \quad \cdots \quad \vec{y}_n] \quad - \quad [\bar{\vec{y}} \quad \bar{\vec{y}} \quad \cdots \quad \bar{\vec{y}}]\}\{[\vec{z}_1 \quad \vec{z}_2 \quad \cdots \quad \vec{z}_n] \quad \cdots \quad [\bar{\vec{z}} \quad \bar{\vec{z}} \quad \cdots \quad \bar{\vec{z}}]\}^T$$

$$= \frac{1}{n-1}\left\{\left[\vec{y}^T - \bar{\vec{y}}1_n^T\right] \quad \left[\vec{z}^T - \bar{\vec{z}}1_n^T\right]^T\right\}$$

$$= \frac{1}{n-1}\left(\vec{y} - 1_n\bar{\vec{y}}^T\right)^T\left(\vec{z} - 1_n\bar{\vec{z}}^T\right)$$

$$= \boldsymbol{S}_{\vec{Y},\vec{Z}}$$

It has thus been shown that $\boldsymbol{S}_{\vec{Y},\vec{Z}} = \boldsymbol{C}\boldsymbol{S}_{\vec{X}}\boldsymbol{D}^T$, using the formula earlier shown to derive $\boldsymbol{S}_{\vec{X}} = \frac{1}{n-1}\left(\boldsymbol{X} - 1_n\bar{\vec{x}}^T\right)^T\left(\boldsymbol{X} - 1_n\bar{\vec{x}}^T\right)$. ■

5. Given that $\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ and the observed sample covariance matrix

$$\boldsymbol{S} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}.$$

Then the variance covariance matrix is as follows,

$$\boldsymbol{\Sigma} = \mathrm{Cov}(\boldsymbol{X}) = \begin{bmatrix} Cov(X_1, X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Cov(X_2, X_2) \end{bmatrix}$$

a) To find a scalar $\alpha$ such that $X_2 + \alpha X_1$ has zero sample variance, the following will be shown: First, the population variance will be given,

$$Var(X_2 + \alpha X_1) = Var(X_2) + \alpha^2 Var(X_1) + 2\alpha Cov(X_1, X_2).$$

Then, let $Y = X_2 + \alpha X_1$, so the sample variance is:

$$\boldsymbol{S}_Y = \boldsymbol{S}_{X_2} + \alpha^2 \boldsymbol{S}_{X_1} + 2\alpha \boldsymbol{S}_{X_1,X_2}$$
$$= 3 + \alpha^2(2) + 2\alpha(1)$$

where the above values come from the variance covariance matrix $\boldsymbol{\Sigma}$. Then using the quadratic formula,

$$\alpha = \frac{-2 \pm \sqrt{(-2)^2 - 4(2)(3)}}{2(2)} = \frac{-2 \pm \sqrt{-20}}{4}.$$

The result is a complex number with no real number valued solution for $\alpha$. The result then is that no such scalar exists for $\alpha$.

b) To find $\alpha$ such that $X_1$ and $X_2 + \alpha X_1$ has zero sample correlation, first the population correlation will be shown:

$$Corr(X_1, X_2 + \alpha X_1) = \frac{Cov(X_1, X_2 + \alpha X_1)}{\sqrt{Var(X_1)Var(X_2 + \alpha X_1)}}$$

Setting this value to zero leads to the following:

$$\frac{Cov(X_1, X_2 + \alpha X_1)}{\sqrt{Var(X_1)Var(X_2 + \alpha X_1)}} = 0$$

$$Cov(X_1, X_2 + \alpha X_1) = 0$$

$$Cov(X_1, X_2) + Cov(X_1, \alpha X_1) = 0$$

$$Cov(X_1, X_2) + \alpha Var(X_1) = 0$$

$$\alpha Var(X_1) = -Cov(X_1, X_2)$$

$$\alpha = -\frac{Cov(X_1, X_2)}{Var(X_1)}$$

Then using the sample variance covariance matrix, the following is the result for $\alpha$:

$$\boxed{\alpha = -\frac{1}{2}}$$