ETF 5500 Assignment 2

Arindom Baruah (32779267)

Question 1

Part A

Here, we have a data matrix Y with a dimension of $n \times 3$ which has a mean of 0.

For a vector Y which has been demeaned, the sample covariance matrix S can be obtained through the following computation.

$$S = \frac{1}{n-1}(Y^T Y) \tag{1}$$

To maintain matrix conformability, the dimension of S will be a 3×3 matrix.

In order to define the covariance matric S as defined by Equation 1, we require the following quanities:

1. A data matrix Y with a mean of 0.
2. The total number of observations n which is given by the number of rows in Y.

Part B

An Eigen Value problem is linear in nature and defined by Equation 2.

$$SW = \lambda W \tag{2}$$

Where,

S = Covariance Matrix
W = Eigen Vector
λ = Eigen Value

As the Eigen Vector is a column vector, hence, to satisfy matrix conformability in Equation 2, the dimension of W must be a 3×1 vector.

Part C

We are given the linear combination as stated through Equation 3.

$$X = \beta Y \tag{3}$$

We know that Y is a $n \times 1$ vector and β is stated as a 3×1 vector. To obtain the Variance-Covariance matrix, we perform the following computations.

$$\mathrm{Cov}(X) = \frac{1}{n-1} X^T X$$

$$\implies \mathrm{Cov}(X) = \sum_{i=1}^{n} \frac{(Y_i \beta)^2}{n-1}$$

$$\implies \mathrm{Cov}(X) = \beta^T \Big[\sum_{i=1}^{n} \frac{(Y_i)^2}{n-1} \Big] \beta$$

For a data matrix Y_i with mean $= 0$, the term $\sum_{i=1}^{n} \dfrac{(Y_i)^2}{n-1}$ is considered as the covariance and is denoted by S.

$$\boxed{\implies \mathrm{Cov}(X) = \beta^T S \beta} \tag{4}$$

Equation 4 provides us with the final form of the variance-covariance matrix.

We know, the dimension of β^T is 1×3 and that of β is 3×1. In order to maintain matrix conformability, the value dimension of the variance covariance matrix $\mathrm{Var}(X)$ will be 1×1. This suggests that the **resultant matrix is a scalar value.**

Exercise 2

The dimensions of the matrices of interest are as follows:

W is a 3×1 matrix
Y is a $n \times 1$ matrix
S is a 3×3 matrix

Part A

W^TY matrix will be conformable **if the matrix Y has a $n = 3$ observations.**

The dimension of this matrix will be 1×3 .

Part B

WW^T will be a conformable matrix with dimension 3×3.

Part C

WV^T will be a conformable matrix with dimension 3×3.

Part D

S^TY matrix will be conformable **if the matrix Y has a $n = 3$ observations.**

The dimension of this new matrix will be 3×3.

Part E

$YW + V^T$ is **not a conformable matrix** as the number of columns for matrix Y and the number of rows for matrix W are not identical.

Exercise 3

We are given with the equation $X = YC^T$

Where

$$C = \begin{bmatrix} W^T \\ U^T \end{bmatrix}$$

and is a 2×3 dimensional matrix.

W is the eigen vector of S corresponding to the largest eigen value while U is the eigen vector of S corresponding to the second-largest eigen value.

Performing a matrix multiplication to obtain X gives us a $n \times 2$ matrix where n is the number of observations.

Content of matrix X

The matrix multiplication performs a linear combination such that the data contained in matrix Y is now projected along the eigen vectors relating to the largest and the second-largest eigen value. **In the context of principal component analysis (PCA), X contains the data which is projected onto the top 2 principal components.**

Derivation for sample covariance of X

We know, the variance-covariance matrix is calculated as follows:

$$\text{Cov}(X) = \frac{(X^T X)}{n-1}$$
$$\implies \text{Cov}(X) = \sum_{i=1}^{n} \frac{(YC^T)^2}{n-1}$$

$$\implies \text{Cov}(X) = C^T \Big[\sum_{i=1}^{n} \frac{(Y_i)^2}{n-1} \Big] C$$

As previously defined in Equation 1, we can use the matrix S in the above equation as follows.

$$\boxed{\implies \text{Cov}(X) = C^T S C} \tag{5}$$

Equation 5 is the expression of the sample variance-covariance matrix $\text{Cov}(X)$ in terms of C and S.

Exercise 4

Vectors W and U are said to be orthogonal when they are at complete right angles to each other. This means that performing a matrix dot product, $W \cdot U$ will **yield a result of 0.**

In the context of PCA, vectors W and U (also called principal components) are **uncorrelated** and capture unique signals in the data.

Proof that Cov(X) is a diagonal matrix

From Equation 5, we know that $\mathrm{Cov}(X) = C^T SC$.

Additionally, W and U are orthogonal.

Based on the spectral theorem, we can write the variance-covariance matrix S as follows.

$$S = \Sigma \lambda_i v_i v_i^T$$

Where λ_i is an eigen value and v_i is an eigen vector corresponding to the eigen value.

So, based on the above information, we can write Equation 5 as follows.

$$\mathrm{Cov}(X) = \frac{1}{n-1} C^T (\lambda_w W W^T + \lambda_u U U^T) C$$

$$\mathrm{Cov}(X) = \frac{1}{n-1} \begin{bmatrix} W^T \\ U^T \end{bmatrix} \quad (\lambda_w W W^T + \lambda_u U U^T) \begin{bmatrix} W^T U^T \end{bmatrix} \tag{6}$$

All dot products between W and U reduce to 0 due to the orthogonality condition between the two vectors.

Additionally, for normalised vectors, $vv^T = 1$. As a result, Equation 6 reduces to the following.

$$\mathrm{Cov}(X) = \frac{1}{n-1} (\lambda_w W + \lambda_u U) \tag{7}$$

Converting Equation 7 to the matrix form, we can write it as follows.

$$\mathrm{Cov}(X) = \frac{1}{n-1} \begin{bmatrix} \lambda_w & 0 \\ 0 & \lambda_u \end{bmatrix} \tag{8}$$

The final form of the variance-covariance matrix is diagonal, as shown in Equation 8. This diagonal structure arises from the orthogonality condition, which ensures that all off-diagonal elements representing covariances are zero.

Exercise 5

Here, we have observations give as follows.

$a = \dfrac{x_1}{\sqrt{\lambda_u}}$ and $b = \dfrac{x_2}{\sqrt{\lambda_u}}$

Z is a datamatrix consisting of the vectors a and b.

Now, the sample variance-covariance matrix Cov (Z) will be calculated as follows.

$$\mathrm{Cov}(Z) = \frac{1}{n-1} Z^T Z$$

$$\mathrm{Cov}(Z) = \frac{1}{n-1} \begin{bmatrix} a^T a & a^T b \\ b^T a & b^T b \end{bmatrix} \tag{9}$$

From @covmatfinal, we know that the variances of x_1 and x_2 are given by their eigen values. Using this result, we obtain the variances of the vectors a and b as follows.

$$\mathrm{Var(a)} = \mathrm{Var}\Big(\frac{x_1}{\sqrt{\lambda_w}}\Big)$$

Now, implementing the variance scaling rule Var(cX) = c^2Var(X) and utilising results from Equation 8 to the above equation, we get the result as follows.

$$\mathrm{Var(a)} = \frac{1}{\lambda_w}\mathrm{Var}(x_1)$$

$$\implies \mathrm{Var(a)} = \frac{\lambda_w}{\lambda_w} = 1$$

Similary, we obtain the value of $\mathrm{Var}(b) = 1$.

For a demeaned vector upon normalisation, we know that,

$$\mathrm{Var}(a) = \frac{1}{n-1} a^T a$$

$$\implies \frac{1}{n-1} a^T a = 1$$

$$\implies a^T a = n - 1 \tag{10}$$

Similarly, $b^T b = n - 1$

Next, we obtain the covariances between a and b.

$$\text{Cov}(a,b) = \text{Cov}(\frac{x_1}{\lambda_w}, \frac{x_2}{\lambda_u})$$

$$\implies \text{Cov}(a,b) = \frac{1}{\sqrt{\lambda_w \lambda_u}} \text{Cov}(x_1, x_2)$$

However, we know that our vectors x_1 and x_2 are the columns in the matrix X, representing the two principal components PC1 and PC2. Hence, x_1 and x_2 are uncorrelated and will have a covariance $\text{Cov}(x_1, x_2) = 0$.

As a result, we obtain the following.

$$\text{Cov}(a,b) = a^T b = b^T a = 0 \tag{11}$$

Replacing the relevant terms in Equation 9 by Equation 10 and Equation 11, we obtain the following result.

$$\text{Cov}(Z) = \frac{1}{n-1} \begin{bmatrix} n-1 & 0 \\ 0 & n-1 \end{bmatrix}$$

$$\implies \text{Cov}(Z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{I} \tag{12}$$

As observed in Equation 12, the sample variance-covariance matrix of Z is an identity matrix.