High Dimensional Data Analysis Individual Assignment: S2, 2024

Department of Econometrics and Business Statistics, Monash University

Due Date: 11th October 2024 at 4:30PM

Questions for ETF3500 and ETF5500 students

1. Consider the $n \times 3$ data matrix $\mathbf{Y}$ The data matrix stores n observations of the 3 dimensional vector $\mathbf{y} = (y_1, y_2, y_3)$. The data matrix has been demeaned. Use this information to answer the following questions. (**5 marks**).

- $\mathbf{S}$ denote the sample covariance matrix of $\mathbf{Y}$. State the sample covariance matrix of $\mathbf{Y}$. Be sure to clearly define any quantities needed to state $\mathbf{S}$, as well as the dimension of $\mathbf{S}$.
- What is the dimension of an arbitrary eigenvector associated with $\mathbf{S}$?
- Let $X = \mathbf{Y}\beta$, where β is a (3×1) vector. What is the dimension of the sample covariance matrix of X? State the sample covariance matrix of X in terms of $\mathbf{S}$.

2. Let $\mathbf{w}$ be the eigenvector of $\mathbf{S}$ corresponding to the largest eigenvalue. In addition, let $\mathbf{v}$ be any arbitrary column vector, whose dimension is the same as that of $\mathbf{w}$. Discuss whether the following matrix operations are conformable. If the operation is conformable, state the dimension of the product. (**5 marks**)

- $\mathbf{w}'\mathbf{Y}$
- $\mathbf{w}\mathbf{w}'$
- $\mathbf{w}\mathbf{v}'$
- $\mathbf{S}'\mathbf{Y}$
- $\mathbf{Y}\mathbf{w} + \mathbf{v}'$

3. Let $\mathbf{X} = \mathbf{Y}\mathbf{C}'$, where

$$\mathbf{C} = \begin{bmatrix} \mathbf{w}' \\ \mathbf{u}' \end{bmatrix},$$

$\mathbf{w}$ is defined in Question 2, and $\mathbf{u}$ is the eigenvector of $\mathbf{S}$ corresponding to the second largest eigenvalue. What does the matrix $\mathbf{X}$ contain? Derive the expression for the sample covariance of $\mathbf{X}$, in terms of $\mathbf{S}$. (**5 marks**)

4. What does it mean for $\mathbf{w}$ and $\mathbf{u}$ to be orthogonal? Prove that the sample covariance matrix of $\mathbf{X}$ is diagonal when $\mathbf{w}$ and $\mathbf{u}$ are orthogonal.

(**5 marks**)

[Hint: Consider expressing the covariance matrix of $\mathbf{X}$ as a function of the eigenvalues, λ_w and λ_u, associated with the eigenvectors $\mathbf{w}$ and $\mathbf{u}$, respectively.]

Questions for ETF5500 students only

5. Let us construct a **new variable A**, whose $(n \times 1)$ vector of observations is constructed by $\mathbf{a} = \frac{1}{\sqrt{\lambda_w}}\mathbf{x}_1$ and a **new variable B**, whose $(n \times 1)$ vector of observations is constructed by $\mathbf{b} = \frac{1}{\sqrt{\lambda_u}}\mathbf{x}_2$. The terms $\mathbf{x}_i$ is the i^{th} column of the matrix $\mathbf{X}$, defined in Question 3; and the terms λ_w and λ_u are defined in Question 4. Let us collect these two new variables into a new $(n \times 2)$ data matrix $\mathbf{Z} = \begin{bmatrix} \mathbf{a} & \mathbf{b} \end{bmatrix}$. Prove that the sample covariance matrix of $\mathbf{Z}$ is an identiy matrix. You MUST use the result from Question 4 in your derivation.

   (**5 marks**)

   [Hint: Write $\mathbf{Z}$ in terms of $\mathbf{X}$, and derive its covariance matrix.]