Indian Institute of Technology Guwahati
Statistical Inference and Multivariate Analysis (MA324)
Problem Set 12

1. Let $\boldsymbol{X} = (X_1, \ldots, X_p)'$ be a p-dimensional random vector and $\boldsymbol{Y} = (Y_1, \ldots, Y_p)'$ be the corresponding principal components. Show that correlation coefficient between Y_i and X_k is $\dfrac{e_{ik}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}}$.

2. Determine the population principal components Y_1 and Y_2 for the variance-covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}.$$

   Also, calculate the proportion of total population variance explained by the first principal component.

3. Convert the variance-covariance matrix of the previous problem to a correlation matrix $\boldsymbol{\rho}$. Determine the principal components Y_1 and Y_2 base on $\boldsymbol{\rho}$ and compute the proportion of total population variance explained by Y_1.

4. Find the principal components and the proportion of the total variance explained by each when the variance-covariance matrix is

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2 & \sigma^2\rho & 0 \\ \sigma^2\rho & \sigma^2 & \sigma^2\rho \\ 0 & \sigma^2\rho & \sigma^2 \end{bmatrix}.$$

5. Show that the canonical correlations are invariant under non-singular linear transformation of the $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ variables of the form $\boldsymbol{C}\boldsymbol{X}^{(1)}$ and $\boldsymbol{D}\boldsymbol{X}^{(2)}$.

6. Let

$$\boldsymbol{\rho}_{12} = \begin{bmatrix} \rho & \rho \\ \rho & \rho \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\rho}_{11} = \boldsymbol{\rho}_{22} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

   corresponding to the equal correlation structure, where $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ each have two components. Determine the canonical variates corresponding to the nonzero canonical correlation.