Stat778 HW03

1. $\Sigma = \begin{pmatrix} 19 & 30 & 2 & 12 \\ 30 & 57 & 5 & 23 \\ 2 & 5 & 38 & 47 \\ 12 & 23 & 47 & 68 \end{pmatrix}$ is given in Example 9.1 on page 484. Answer the followings by quoting results from SAS output. Do not submit the SAS output.

   (1) What is the first principal component for X?

   The first principal component for X is

   $$Y_1 = 0.235376X_1 + 0.440974X_2 + 0.470476X_3 + 0.727006X_4.$$

   (2) Which percentage of total variances in X is explained by the second principal component?

   33.75% of total variances in X is explained by the second principal component of X.

   (3) To have at least 85% of total variances explained, how many principal components should we use?

   Use two principal components such that $97.63\% > 85\%$ of total variances would be explained.

   (4) What is the first principal component for Z, the standardized X?

   The first principal component for Z is

   $$0.495405Z_1 + 0.510276Z_2 + 0.436817Z_3 + 0.550802Z_4.$$

2. In 5.1 (a) on page 261 a sample of size 4 is given by $X = \begin{pmatrix} 2 & 8 & 6 & 8 \\ 12 & 9 & 9 & 10 \end{pmatrix}$ (caution: book uses X' for sample). Answer the followings by quoting results from SAS output. Do not submit the SAS output.

   (1) Sample mean $\overline{X}$ $\qquad \overline{X} = \begin{pmatrix} 6 \\ 10 \end{pmatrix}$

   (2) SSCP matrix $\qquad \text{SSCP} = \begin{pmatrix} 168 & 230 \\ 230 & 406 \end{pmatrix}$

   (3) CSSCP matrix $\qquad \text{CSSCP} = \begin{pmatrix} 24 & -10 \\ -10 & 6 \end{pmatrix}$

   (4) Sample covariance matrix S $\qquad S = \begin{pmatrix} 8 & -3.3333 \\ -3.3333 & 2 \end{pmatrix}$

   (5) Sample correlation matrix R $\qquad R = \begin{pmatrix} 1 & -0.8333 \\ -0.8333 & 1 \end{pmatrix}$

3. $X \in R^{p\times n}$ is a random sample with $X \sim (\mu 1_n', \Sigma, I_n)$. Find $E(\text{SSCP})$.

   $E(\text{SSCP}) = E(XX') = (\mu 1_n')(\mu 1_n')' + \text{tr}(I_n)\Sigma = n\mu\mu' + n\Sigma.$