PRINCIPAL COMPONENT ANALYSIS

1. Following is the dispersion matrix of the scores in the 3 subjects for 55 students.

X_1 : Statistics, X_2 : Economics, X_3 : Mathematics

$$\begin{pmatrix} 4.735 & 0.562 & 1.469 \\ 0.562 & 0.142 & 0.217 \\ 1.469 & 0.217 & 0.570 \end{pmatrix}$$

Find the principal components and their variances.

2. Consider the following three variate dataset with ten observations. Each observation consists of three measurements on a wafer: thickness, horizontal displacement and vertical displacement.

$$X = \begin{bmatrix} 7 & 4 & 3 \\ 4 & 1 & 8 \\ 6 & 3 & 5 \\ 8 & 6 & 1 \\ 8 & 5 & 7 \\ 7 & 2 & 9 \\ 5 & 3 & 3 \\ 9 & 5 & 8 \\ 7 & 4 & 5 \\ 8 & 2 & 2 \end{bmatrix}$$

(a) Begin with an adjusted data matrix. Construct new variables as weighted averages of the original variables so that their variances are maximized, subject to the constraint that the covariance between the ith and jth variables is zero for all i and j $(i \neq j)$.

(b) Compute the correlation between an ith variable, constructed in (a) and the jth original variable. Also calculate the total sample variance.