Statistical Sciences

Honours Multivariate Analysis
Continuous Assessment 1

Instructions:

- You will be divided into groups for this assessment. Only 1 submission per group is required.
- Your **.pdf** report may be compiled using any software you like (Rmarkdown, LaTeX, MSWord, etc.), as long as the presentation is neat.
- Do NOT paste R output verbatim, this will be penalised. If you want to include R output, typeset it properly or present it in a table.
- To help the reader easily assimilate the information, round values to a small number of decimal places (unless there is a reason for expressing a more exact value).

Researchers have suggested that a change in skull size over time is evidence of the interbreeding of a resident population with immigrant populations. Four measurements were made of male Egyptian skulls for five different time periods representing approximate time of skull formation. These periods are 4000 B.C., 3300 B.C., 1850 B.C., 200 B.C., and 150 A.D., respectively. The variables are as follows:

$$\begin{aligned} X_1 &= \text{maximal breadth of skull (mm)} \\ X_2 &= \text{basibregmatic height of skull (mm)} \\ X_3 &= \text{basialveolar length of skull (mm)} \\ X_4 &= \text{nasal height of skull (mm)} \end{aligned}$$

The data are given in `CA1.csv`.

1. Compute and report the sample mean vectors for each of the five time periods.
2. Provide a heat map of the correlation matrix for each time period and briefly interpret. Are there any noticeable changes over the time periods?
3. Calculate the angle between the deviation vectors for X_1 and X_3 in period 1. Explain why this value is to be expected by referring to the appropriate value from question 2. *For a bonus mark, plot all the deviation vectors for period 1 across the first two observations.*
4. Suppose researchers are interested in the quantity $Y_i = 3X_4 - X_1$ for time periods $i = 1, \ldots, 5$. Use your answers from question 1 and an appropriate vector $\boldsymbol{b}$ to determine the sample means $\bar{y}_1, \ldots, \bar{y}_5$. Clearly show your vector $\boldsymbol{b}$. Also give the covariance matrix of $\boldsymbol{Y} = \begin{bmatrix} Y_1 & Y_2 & Y_3 & Y_4 & Y_5 \end{bmatrix}'$.