THE UNIVERSITY OF MELBOURNE

Student number

Semester 1 Assessment, 2023

School of Mathematics and Statistics

MAST30025 Linear Statistical Models Assignment 1

Submission deadline: **Friday March 24, 5pm**

This assignment consists of 3 pages (including this page) with 5 questions and 33 total marks

Instructions to Students

Writing

- This assignment is worth 6% of your total mark.
- You may choose to either typeset your assignment in LaTeX, or handwrite and scan it to produce an electronic version.
- You may use R for this assignment, but for matrix calculations only (you may not use the `lm` function). If you do, include your R commands and output.
- Write your answers on A4 paper. Page 1 should only have your student number, the subject code and the subject name. Write on one side of each sheet only. Each question should be on a new page. The question number must be written at the top of each page.

Scanning and Submitting

- Put the pages in question order and all the same way up. Use a scanning app to scan all pages to PDF. Scan directly from above. Crop pages to A4.
- Submit your scanned assignment as a single PDF file and carefully review the submission in Gradescope. Scan again and resubmit if necessary.

Question 1 (4 marks)

Prove that if a symmetric matrix A has eigenvalues which are all either 0 or 1, it is idempotent.

Question 2 (6 marks)

We wish to prove (without using Theorem 2.5) that if A, B, and $A+B$ are $n \times n$ idempotent matrices, then $AB = BA = 0$.

(a) Show that $AB + BA = 0$.

(b) By Theorem 2.2, there exists a matrix P which diagonalises A:

$$P^T AP = D = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

where $r = r(A)$.

Write

$$P^T BP = \Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix},$$

where the block matrices are of the same dimensions as above. Show that $\Lambda_{11} = 0$, $\Lambda_{12} = 0$ and $\Lambda_{21} = 0$.

(c) Show that $AB = BA = 0$.

Question 3 (4 marks)

Show directly that for any random vector $\mathbf{y}$ and compatible matrix A, we have $\text{var}\, A\mathbf{y} = A(\text{var}\, \mathbf{y})A^T$.

Question 4 (7 marks)

Let $\mathbf{y} = (y_1, y_2, y_3)^T$ be a 3-dimensional multivariate normal random vector with mean and variance

$$\boldsymbol{\mu} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \quad V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

(a) Find the eigenvalues and eigenvectors of V.

(b) Find the distribution of $z_1 = 3y_1 + 2y_2 + y_3$.

(c) Find the distribution of $z_2 = y_1^2 + \left(\frac{y_2+y_3}{2}\right)^2 + \left(\frac{y_2-y_3}{2}\right)^2$.

Question 5 (12 marks)

A secondary school teacher wants to know if the marks of students in Specialist Mathematics can be predicted from their marks in General Mathematics. A linear model is assumed, and the following data is obtained from nine students:

ID	General Mathematics	Specialist Mathematics
1	86	85
2	85	97
3	89	76
4	82	79
5	84	76
6	86	99
7	84	49
8	78	72
9	92	83

(a) Write down the linear model as a matrix equation, writing out the matrices in full.

(b) Calculate the least squares estimate of the parameters.

(c) Calculate the sample variance s^2.

(d) Calculate the standardised residuals for all students.

(e) Calculate the Cook's distances for all students.

(f) Predict (using a point estimate) the mark of Specialist Mathematics for a student whose mark for General Mathematics is 90.

End of Assignment — Total Available Marks = 33