THE UNIVERSITY OF MELBOURNE

Student number 1067799

Semester 1 Assessment, 2021

School of Mathematics and Statistics

MAST30025 Linear Statistical Models Assignment 1

Submission deadline: **Friday March 26, 5pm**

This assignment consists of 10 pages (including this page)

Instructions to Students

- If you have a printer, print the assignment template.

Writing

- There are 5 questions with marks as shown. The total number of marks available is 40.
- 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 the page.

Scanning

- 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. Check PDF is readable.

Submitting

- Go to the Gradescope window. Choose the Canvas assignment for this assignment. Submit your file as a single PDF document only. Get Gradescope confirmation on email.
- It is your responsibility to ensure that your assignments are submitted correctly and on time, and problems with online submissions are not a valid excuse for submitting a late or incorrect version of an assignment.

Question 1 (7 marks)

Which of the following statements are true, and which are false? For the ones which are true, give a proof. For the ones which are false, give a counterexample.

(a) For a symmetric matrix A, if $A^2 = A^3$, A is idempotent.

(b) For a symmetric matrix A, if $A = A^3$, A is idempotent.

(a) True

if $A^2 = A$, matrix A is idempotent.

let $\Lambda = P^T A P$, then $A = P\Lambda P^T$, P is an orthogonal matrix that diagonalise A with $PP^T = I$

$A^2 = P\lambda P^T \cdot P\lambda P^T = P\Lambda^2 P^T$ $\qquad \Lambda = \begin{bmatrix} \lambda_{11} & & \\ & \lambda_{ii} & \\ & & \lambda_{nn} \end{bmatrix}$

$A^3 = P\lambda^2 P^T \cdot P\lambda P^T = P\Lambda^3 P^T$

if $A^2 = A^3$, $P\Lambda^2 P^T = P\Lambda^3 P^T$

$\lambda_{ii}^2 = \lambda_{ii}^3$, for $i \in [1, n]$

$\lambda_{ii} = 0$ or $\lambda_{ii} = 1$

$\Rightarrow \lambda_{ii}^2 = \lambda_{ii}$ $\Rightarrow \Lambda^2 = \Lambda$ $\Rightarrow P\Lambda^2 P^T = P\Lambda P^T$

$\Rightarrow A^2 = A$, Therefore, A is idempotent.

(b) False

Counter example:

if $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\qquad A^3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = A$.

$A^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \neq A$.

A is not always idempotent.

Question 2 (4 marks)

Let $A_1, A_2, \ldots, A_m$ be a set of symmetric $k \times k$ matrices. Suppose that there exists an orthogonal matrix P such that $P^T A_i P$ is diagonal for all i. Show that $A_i A_j = A_j A_i$ for every pair $i, j = 1, 2, \ldots, m$.

As $P^T A_i P$ is diagonal for all i,

$$\underset{(k\times k)}{P^T A_i P} = \begin{bmatrix} \lambda_{i1} & & \\ & \ddots & \\ & & \lambda_{im} \end{bmatrix}, \quad \underset{(k\times k)}{P^T A_j P} = \begin{bmatrix} \lambda_{j1} & & \\ & \ddots & \\ & & \lambda_{jm} \end{bmatrix}$$

Since P is orthogonal, $PP^T = I$

$$P^T A_i P \cdot P^T A_j P = \begin{bmatrix} \lambda_{i1}\lambda_{j1} & & \\ & \ddots & \\ & & \lambda_{im}\lambda_{jm} \end{bmatrix}$$

$$= P^T A_i \underline{PP^T} A_j P.$$

$$= P^T A_i A_j P$$

$$P^T A_j P \cdot P^T A_i P = \begin{bmatrix} \lambda_{j1}\cdot\lambda_{i1} & & \\ & \ddots & \\ & & \lambda_{jm}\lambda_{im} \end{bmatrix}$$

$$= P^T A_j P P^T A_i P$$

$$= P^T A_i A_j P.$$

This two terms are equal as λs are scalar. && they both are kxk matrix

$$\Rightarrow \quad P^T A_i A_j P = P^T A_j A_i P$$

$$P \cdot P^T A_i A_j P \cdot P^T = P \cdot P^T A_j A_i P \cdot P^T$$

$$I \, A_i A_j \, I = I \, A_j A_i \, I$$

$\Rightarrow$ $A_i A_j = A_j A_i$ for every pair of $i, j = 1..m$

Question 3 (4 marks)

Let A be an $n \times p$ matrix with rank p. Show that $A^T A$ is a positive definite matrix.

Let y be $p \times 1$ vector. $A = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{bmatrix}$ $y = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}$

$y^T A^T A y = (Ay)^T A y$, Ay is a $n \times 1$ vector.

$$Ay = \begin{bmatrix} \sum_p a_{1i} \cdot y_i \\ \vdots \\ \sum_p a_{ni} \cdot y_i \end{bmatrix}$$

$(Ay)^T (Ay) = \sum_N \left(\sum_p a_{ji} y_i \right)^2$, $i \in [1, p]$, $j \in [1, n]$

This term is always greater than zero if $Ay \neq 0$.

Since $\text{rank}(A) = p$, A has independent p columns,

$\sum_p a_{ji} \cdot y_i$, $j \in [1, n]$, cannot be equal to zero as long as y is not zero.

Then it is proved that $y^T A^T A y = 0$ only if $y = 0$,
$y^T A^T A y > 0$ if $y \neq 0$.

Therefore, $A^T A$ is a positive definite matrix

Question 4 (10 marks)

Let $x_1, x_2, x_3 \sim N(\mu, \sigma^2)$ be a sequence of independent normal random variables, $\bar{x} = \frac{x_1+x_2+x_3}{3}$, $\mathbf{x} = (x_1, x_2, x_3)$, and $\mathbf{y} = (x_1 - \bar{x}, x_2 - \bar{x}, x_3 - \bar{x})^T$.

3x1 3x1x1x3

(a) Let $\mathbf{y} = A\mathbf{x}$. Find A.

(b) Find the rank of A.

(c) Find $E(\mathbf{y}^T\mathbf{y})$.

(d) Using Theorem 3.5, find the distribution of $\mathbf{y}^T\mathbf{y}/\sigma^2$.

(a) $$\underset{(3\times1)}{y} = \begin{bmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \end{bmatrix} = \begin{bmatrix} \frac{2x_1 - x_2 - x_3}{3} \\ \frac{-x_1 + 2x_2 - x_3}{3} \\ \frac{-x_1 - x_2 + 2x_3}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3}x_1 - \frac{1}{3}x_2 - \frac{1}{3}x_3 \\ -\frac{1}{3}x_1 + \frac{2}{3}x_2 - \frac{1}{3}x_3 \\ -\frac{1}{3}x_1 - \frac{1}{3}x_2 + \frac{2}{3}x_3 \end{bmatrix}$$

$$\underset{(3\times1)}{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

As $y = Ax$, A is a 3×3 matrix.

Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & I \end{bmatrix}$, $Ax = \begin{bmatrix} ax_1 + bx_2 + cx_3 \\ dx_1 + ex_2 + fx_3 \\ gx_1 + hx_2 + Ix_3 \end{bmatrix}$

$$\Rightarrow A = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

(b) $$A \xrightarrow{x_2 \to x_2 - x_3} \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ 0 & 3 & -3 \\ -1 & -1 & 2 \end{bmatrix} \xrightarrow{x_3 \to x_3 + \frac{1}{2}x_1} \begin{bmatrix} 2 & -1 & -1 \\ 0 & 3 & -3 \\ 0 & -\frac{3}{2} & \frac{3}{2} \end{bmatrix}$$

$$\xrightarrow{x_3 \to x_3 + \frac{1}{2}x_2} \begin{bmatrix} 2 & -1 & -1 \\ 0 & 3 & -3 \\ 0 & 0 & 0 \end{bmatrix}$$

$\text{rank}(A) = 2$

(c) Let $z_1, z_2, z_3 \sim N(0,1)$, $z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$

Since $x_1, x_2, x_3 \sim N(\mu, \sigma^2)$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$,

$x = Az + \mu$, $AA^T = \sigma^2$, $x \sim MVN(\mu, \sigma^2)$

Since $y = Ax$, $y \sim MVN(\mu_y, A\sigma^2 A^T)$

$$\mu_y = \begin{bmatrix} \mu(x_1) - \mu(\bar{x}) \\ \mu(x_2) - \mu(\bar{x}) \\ \mu(x_3) - \mu(\bar{x}) \end{bmatrix} = \begin{bmatrix} \mu - \mu \\ \mu - \mu \\ \mu - \mu \end{bmatrix} = 0$$

NOTE: $A^T = A$ according to observation,

$var(y) = A \cdot \sigma^2 \cdot A^T = A^2 \sigma^2$

According to theorem of MVN distribution,

$$E(y^T y) = tr(A^2\sigma^2) + 0$$

$$= tr\left[\frac{1}{9}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \sigma^2\right]$$

$$= tr\left[\frac{1}{9}\begin{bmatrix} 6 & \cdot & \cdot \\ \cdot & 6 & \cdot \\ \cdot & \cdot & 6 \end{bmatrix} \sigma^2\right]$$

$$= \frac{1}{9} \times 18 \cdot \sigma^2$$

$$= 2\sigma^2$$

(d) $\frac{y^T y}{\sigma^2} = \frac{(AX)^T (AX)}{\sigma^2} = \frac{X^T A^T A X}{\sigma^2} = \frac{X^T}{\sigma} \cdot A^2 \cdot \frac{X}{\sigma^2}$

As $X \sim MVN(\mu, \sigma^2)$, $\frac{X}{\sigma} \sim MVN(\frac{\mu}{\sigma}, I)$

According to theorem 3.5, $X/\sigma \sim MVN(\frac{\mu}{\sigma}, I)$ and

A is a $n \times n$ symmetric matrix.

Then $\frac{X^T}{\sigma} A \frac{X}{\sigma} \sim \chi^2_{\text{rank}(A), \lambda}$, $\lambda = \frac{1}{2} \cdot \frac{\mu^T}{\sigma} \cdot A \cdot \frac{\mu}{\sigma}$

$= \frac{\mu^T A \mu}{2\sigma^2}$

As $A\mu = 0$, proved in part (c), $\lambda = 0$

Then, $\frac{y^T y}{\sigma^2} \sim \chi^2_{2,0}$.

Question 5 (15 marks)

The table below shows prices in US cents per pound received by fishermen and vessel owners for various species of fish and shellfish in 1970 and 1980. (Taken from Moore & McCabe, Introduction to the Practice of Statistics, 1989.)

Type of fish	Price (1970)	Price (1980)
Cod	13.1	27.3
Flounder	15.3	42.4
Haddock	25.8	38.7
Menhaden	1.8	4.5
Ocean perch	4.9	23.0
Salmon, chinook	55.4	166.3
Salmon, coho	39.3	109.7
Tuna, albacore	26.7	80.1
Clams, soft-shelled	47.5	150.7
Clams, blue hard-shelled	6.6	20.3
Lobsters, american	94.7	189.7
Oysters, eastern	61.1	131.3
Sea scallops	135.6	404.2
Shrimp	47.6	149.0

We will model the 1980 price of fish, based on the 1970 price.

(a) The linear model is of the form $\mathbf{y} = X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$. Write down the matrices and vectors involved in this equation.

(b) Find the least squares estimators of the parameters.

(c) Calculate the sample variance s^2.

(d) A fisherman sold ocean trout for 28c/pound in 1970. Predict the price for ocean trout in 1980.

(e) Calculate the standardised residual for sea scallops.

(f) Calculate the Cook's distance for sea scallops.

(g) Does sea scallops fit the linear model? Justify your argument.

(a) y: 14x1 vector, $\begin{bmatrix} y_1 \\ \vdots \\ y_{14} \end{bmatrix}$, y_i, $i \in [1,14]$ are the price of fishes in 1970

X: 14x2 vector, $\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_{14} \end{bmatrix}$, x_i, $i \in [1,14]$ are the price of fishes in 1980

β: 2x1 vector, $\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$

ε: 14x1 vector, $\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{14} \end{bmatrix}$

13.1
15.3
25.8
1.8
4.9
55.4
39.3
26.7
47.5
6.6
94.7
61.1
135.6
47.6

(b) $b = (X^T X)^{-1} X^T y$

The calculation using R:

```
> x <- matrix(c(rep(1, 14),13.1, 15.3, 25.8, 1.8, 4.9, 55.4, 39.3, 26.7, 47.5, 6.6, 94.7, 61.1, 13
5.6, 47.6),14,2)
> y <- c(27.3, 42.4, 38.7, 4.5, 23, 166.3, 109.7, 80.1, 150.7, 20.3, 189.7, 131.3, 404.2, 149)
> y
 [1]  27.3  42.4  38.7   4.5  23.0 166.3 109.7  80.1 150.7  20.3 189.7 131.3 404.2 149.0
> (b<- solve(t(x)%*%x, t(x)%*%y))
          [,1]
[1,] -1.233836
[2,]  2.701553
```

Therefore, $b = \begin{bmatrix} -1.23 \\ 2.70 \end{bmatrix}$

(c)

```
> (b<- solve(t(x)%*%x, t(x)%*%y))
          [,1]
[1,] -1.233836
[2,]  2.701553
> (e <- y - x%*%b)
             [,1]
 [1,]  -6.8565106
 [2,]   2.3000724
 [3,] -29.7662361
 [4,]   0.8710405
 [5,]  10.9962256
 [6,]  17.8677893
 [7,]   4.7627957
 [8,]   9.2023660
 [9,]  23.6100596
[10,]   3.7035852
[11,] -64.9032511
[12,] -32.5310639
[13,]  39.1032233
[14,]  21.6399042
> (SSRes <- sum(e^2))
[1] 9325.833
> (s2 <- SSRes/(14-2)
+ )
[1] 777.1528
```

The calculation using R is given here.

The sample variance $s^2 = 777.1528$

d) The linear model is $y = b_1 x_1 + b_2 x_2 + \varepsilon$

$= -1.23 + 2.7X + \varepsilon$

if $x = 28$, $y = -1.23 + 2.7 \times 28$

$= 74.37$

e) $\hat{E}(y_{\text{seascallops}}) = -1.23 + 2.7 \times 135.6 = 364.89$

$e_s = y_s - \hat{E}(y) = 404.2 - 364.89 = 39.31$

$z_s = \dfrac{e_s}{\sqrt{s^2(1 - H_{13,13})}} = \dfrac{39.31}{\sqrt{777.1528 \times (1 - 0.5559672)}} = 2.11613$

(The calculation of $H_{13,13}$ See next page)

```
> (H <- x%*%solve(t(x)%*%x)%*%t(x))
            [,1]         [,2]         [,3]        [,4]        [,5]       [,6]       [,7]
 [1,]  0.11396694  0.110624640  0.094672751  0.13113421  0.12642461 0.04970362 0.07416318
 [2,]  0.11062464  0.107544949  0.092846423  0.12644305  0.12210349 0.05141058 0.07394832
 [3,]  0.09467275  0.092846423  0.084129856  0.10405344  0.10147998 0.05955744 0.07292284
 [4,]  0.13113421  0.126443053  0.104053438  0.15522970  0.14861943 0.04093605 0.07526679
 [5,]  0.12642461  0.122103488  0.101479976  0.14861943  0.14253059 0.04334131 0.07496403
 [6,]  0.04970362  0.051410579  0.059557437  0.04093605  0.04334131 0.08252382 0.07003197
 [7,]  0.07416318  0.073948319  0.072922840  0.07526679  0.07496403 0.07003197 0.07160437
 [8,]  0.09330545  0.091586549  0.083382721  0.10213433  0.09971225 0.06025574 0.07283494
 [9,]  0.06170552  0.062469470  0.066115616  0.05778157  0.05885805 0.07639427 0.07080352
[10,]  0.12384192  0.119723727  0.100068722  0.14499445  0.13919154 0.04466033 0.07479800
[11,] -0.01000202 -0.003603902  0.026932570 -0.04286508 -0.03384955 0.11301634 0.06619375
[12,]  0.04104402  0.043431380  0.054825586  0.02878169  0.03214569 0.08694639 0.06947528
[13,] -0.07213842 -0.060858159 -0.007020536 -0.13007796 -0.11418304 0.14475029 0.06219926
[14,]  0.06155359  0.062329484  0.066032601  0.05756833  0.05866164 0.07647186 0.07079375
             [,8]       [,9]       [,10]        [,11]      [,12]        [,13]      [,14]
 [1,]  0.093305447 0.06170552  0.12384192 -0.010002020 0.04104402 -0.072138423 0.06155359
 [2,]  0.091586549 0.06246947  0.11972373 -0.003603902 0.04343138 -0.060858159 0.06232948
 [3,]  0.083382721 0.06611562  0.10006872  0.026932570 0.05482559 -0.007020536 0.06603260
 [4,]  0.102134329 0.05778157  0.14499445 -0.042865080 0.02878169 -0.130077960 0.05756833
 [5,]  0.099712246 0.05885805  0.13919154 -0.033849550 0.03214569 -0.114183042 0.05866164
 [6,]  0.060255739 0.07639427  0.04466033  0.113016338 0.08694639  0.144750286 0.07647186
 [7,]  0.072834942 0.07080352  0.07479800  0.066193748 0.06947528  0.062199265 0.07079375
 [8,]  0.082679536 0.06642814  0.09838401  0.029549982 0.05580223 -0.002405883 0.06635001
 [9,]  0.066428143 0.07365098  0.05944838  0.090041278 0.07837361  0.104243884 0.07368571
[10,]  0.098384007 0.05944838  0.13600930 -0.028905550 0.03399047 -0.105466475 0.05926119
[11,]  0.029549982 0.09004128 -0.02890555  0.227309989 0.12959328  0.346256817 0.09033210
[12,]  0.055802232 0.07837361  0.03399047  0.129593280 0.09313182  0.173976424 0.07848213
[13,] -0.002405883 0.10424388 -0.10546647  0.346256817 0.17397642  0.555967177 0.10475662
[14,]  0.066350011 0.07368571  0.05926119  0.090332102 0.07848213  0.104756624 0.07372098
> H13 <- H[13, 13]
> H13
[1] 0.5559672
> z13 <- 39.31/sqrt(s2*(1-H13))
> z13
[1] 2.11613
```

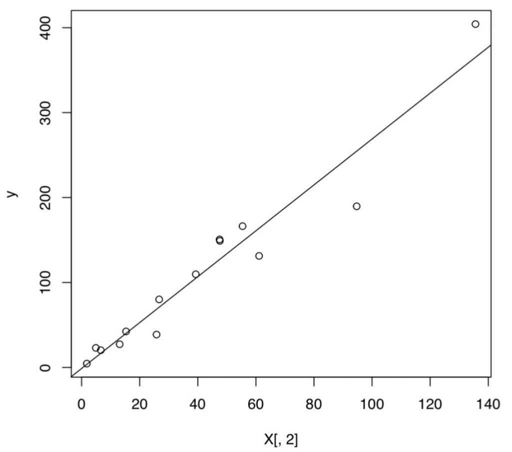
f) $D_{13} = \frac{1}{k+1} z_{13}^2 \left(\frac{H_{13,13}}{1 - H_{13,13}} \right)$

$= \frac{1}{1+1} \times 2.116^2 \times \left(\frac{0.5559672}{1 - 0.5559672} \right)$

≈ 2.80

g) As $D_{13} = 2.8 > 1$, it is considered as large, it does not fit the model.

(g)
```
> plot(X[,2],y)
> abline(b[1],b[2])
```



The Cook's distance certainly indicates it should be of some concern; however looking at the plot, it seems that the fit is actually okay. There is considerable evidence for heteroskedasticity — the variance increases with x (the design variable). Sea scallops has (by far) the largest x and so may be prone to a larger variance than the remaining points. The high Cook's distance therefore comes primarily from a very high leverage, rather than a bad fit to the model.

End of Assignment — Total Available Marks = 40