MAST30025 Linear Statistical Models
Assignment 3

Student Code: 1079860

May 2023

Question 1

(a)

$r(A^c A) \leq min[r(A^c),\ r(A)] \leq r(A)$

$r(A) = r(AA^c A) \leq min[r(A),\ r(A^c A)] \leq r(A^c A)$

$\implies r(A^C A) = r(A)$

(b)

$(I - A(A^T A)^c A^T)(I - A(A^T A)^c A^T)$

$= I^2 - IA(A^T A)^c A^T - A(A^T A)^c A^T I + A(A^T A)^c A^T A(A^T A)^c A^T$

$= I - 2(A(A^T A)^c A^T) + A(A^T A)^c A^T A(A^T A)^c A^T$

$= I - 2(A(A^T A)^c A^T) + [A(A^T A)^c A^T A]\ (A^T A)^c A^T$

$= I - 2(A(A^T A)^c A^T) + A(A^T A)^c A^T$

$I - A(A^T A)^c A^T\ \ is\ \ idempotent,\ \ I\ \ and\ \ A(A^T A)^c A^T\ \ are\ \ idempotent\ \ and\ \ symmetric$

(c)

$r(A(A^TA)^cA^T) \leq min[r(A), r((A^TA)^c), r(A^T)] \leq r(A)$

$r(A) = r(A(A^TA)^cA^TA) \leq min[r(A), r(A(A^TA)^cA^T)] \leq r(A(A^TA)^cA^T)$

$\implies r(A(A^TA)^cA^T) = r(A)$

$A(A^TA)^cA^T \; is \;\; n \times n \;\; so \;\; I \; must \;\; be \;\; n \times n$

$\implies r(I) = n$

$As \;\; I - A(A^TA)^cA^T \;\; is \;\; symmetric \;\; and \;\; n \times n, \;\; it \;\; is \;\; diagonalised \;\; by \;\; P, \;\; so$
$r(I - A(A^TA)^cA^T) = r(P^T(I - A(A^TA)^cA^T)P) = r(I - D)$

$Since \;\; A(A^TA)^cA^T \;\; is \;\; idempotent, \;\; it \;\; has \;\; r(A(A^TA)^cA^T) = r(A) \;\; 1s \;\; on \;\; the \;\; diagonal$
$of \;\; its \;\; diagonal \;\; matrix \;\; D$

$\implies r(I-D) = r(I)-r(D) = n-r(A) \;\; as \;\; the \;\; rank \;\; of \;\; a \;\; diagonal \;\; matrix =$
$the \;\; number \;\; of \;\; non-zero \;\; entries \;\; on \;\; its \;\; diagonal$

Question 2

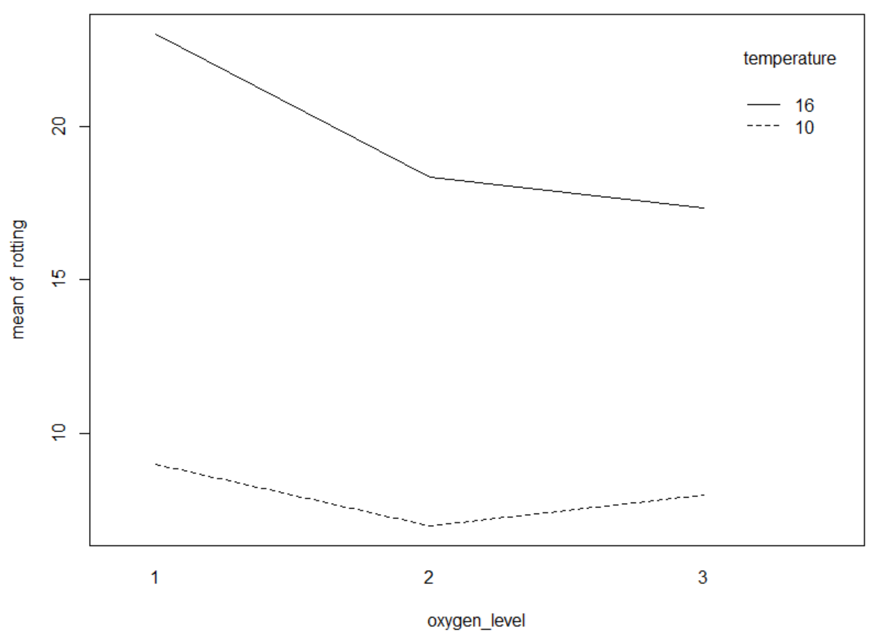
(a)

[R code]

```
rot_df = data.frame(rotting = c(13, 11, 3, 26, 19, 24, 10, 4, 7, 15,
                                22, 18, 15, 2, 7, 20, 24, 8),
                    oxygen_level = c(rep(1, 6), rep(2, 6), rep(3, 6)),
                    temperature = c(rep(c(10, 16), 3, each=3)) )

rot_df$oxygen_level = factor(rot_df$oxygen_level)
rot_df$temperature = factor(rot_df$temperature)

with(rot_df, (interaction.plot(oxygen_level, temperature, rotting)))
```



It may be reasonable to assume that there is no interaction due to the lines being relatively parallel despite the low sample size, although its hard to be certain.

(b)

[R code]

```
> y = rot_df$rotting
> n = length(y)
> X = matrix(c(rep(1, n), rep(0, n*5)), n, 6)
> X[cbind(1:n, as.numeric(rot_df$oxygen_level)+1)] = 1
> X[cbind(1:n, as.numeric(rot_df$temperature)+4)] = 1
>
> library(Matrix)
>
> r=rankMatrix(X)[1]
>
> # find conditional inverse of XtX
>
> XtX = t(X) %*% X
>
> M = XtX[2:5, 2:5]
>
> det(M)
[1] 972
>
> XtXc = matrix(0, 6, 6)
>
> XtXc[2:5,2:5] = t(solve(M))
>
> XtXc = t(XtXc)
>
> all.equal(XtX %*% XtXc %*% XtX,XtX)
[1] TRUE
>
> b = XtXc %*% t(X) %*% y
>
> s2 = sum((y - X %*% b)^2) / (n-r)
```

```
> X
      [,1] [,2] [,3] [,4] [,5] [,6]
 [1,]    1    1    0    0    1    0
 [2,]    1    1    0    0    1    0
 [3,]    1    1    0    0    1    0
 [4,]    1    1    0    0    0    1
 [5,]    1    1    0    0    0    1
 [6,]    1    1    0    0    0    1
 [7,]    1    0    1    0    1    0
 [8,]    1    0    1    0    1    0
 [9,]    1    0    1    0    1    0
[10,]    1    0    1    0    0    1
[11,]    1    0    1    0    0    1
[12,]    1    0    1    0    0    1
[13,]    1    0    0    1    1    0
[14,]    1    0    0    1    1    0
[15,]    1    0    0    1    1    0
[16,]    1    0    0    1    0    1
[17,]    1    0    0    1    0    1
[18,]    1    0    0    1    0    1
>
> s2
[1] 26.12698
```

(c)

[R code]

```
> tt = c(0, 0, 0, 0, 1, −1) # temp10mean − temp16mean
>
> ta = qt(0.975, n−r)
>
> halfwidth = ta * sqrt(s2 * t(tt) %*% XtXc %*% tt)
>
> tt %*% b + c(−1, 1) * halfwidth
[1] −16.723555  −6.387556
```

$$95\% \ confidence \ interval \ for \ temp10effect - temp16effect:$$

$[-16.724, -6.388]$

(d)

$H_0 : \tau_1 = \tau_2 = \tau_3 = 0$

[R code]

```
> C = matrix(c(0,0,1,1,−1,0,0,−1,0,0,0,0), 2, 6)
>
> all.equal(round(C %*% XtXc %*% t(X) %*% X, 3), C)
[1] TRUE
>
> numer = t(C %*% b) %*% solve(C %*% XtXc %*% t(C)) %*% C%*%b
>
> Fstat = (numer / 2 )/ s2
>
> pf(Fstat, 2, n−r, lower=F)
          [,1]
[1,] 0.4481124
```

$p - value = 0.448 \implies Cannot \ reject \ H_0 \ at \ 5\% \ significance \ level$

(e)

This would be a complete block design study with oxygen level as the factor of interest and temperature as the blocking factor.

Question 3

Question 4

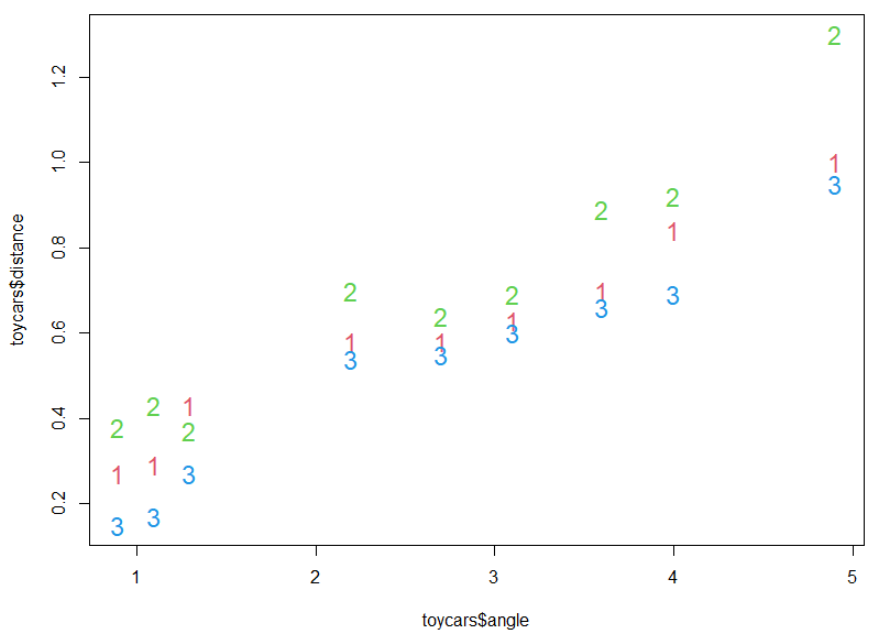
(a)

[R code]

```
toycars = read.csv("toycars.csv")

toycars$car = factor(toycars$car)

plot(toycars$angle, toycars$distance, pch=array(toycars$car),
col=as.numeric(toycars$car)+1, cex=1.5)
```



It seems that the distance travelled by the car increases as the angle increases. It would also seem that the type of car doesn't seem to have as much of an effect as the angle in this experiment, although we can still say that car 3 generally travels the least distance, while car 1 travels slightly further and car 2 travels the furthest.

(b)

[R code]

```
> imodel = lm(toycars$distance ~ toycars$car * toycars$angle, data=toycars)
>
> amodel = lm(toycars$distance ~ toycars$car + toycars$angle, data=toycars)
>
> anova(amodel, imodel)
Analysis of Variance Table

Model 1: toycars$distance ~ toycars$car + toycars$angle
Model 2: toycars$distance ~ toycars$car * toycars$angle
  Res.Df      RSS Df Sum of Sq      F Pr(>F)
1     23 0.105657
2     21 0.093271  2  0.012386 1.3944   0.27
```

No significant interaction present between the type of toy car and the angle

(c)

[R code]

```
> fullmodel = imodel
>
> drop1(fullmodel, scope = ~ ., test="F")
Single term deletions

Model:
toycars$distance ~ toycars$car * toycars$angle
                            Df Sum of Sq     RSS      AIC  F value
Pr(>F)
<none>                                   0.09327 -141.038
toycars$car                  2   0.01979 0.11307 -139.842
2.2284   0.1325
toycars$angle                1   0.44593 0.53920  -95.664 100.4023 1.87e-09
toycars$car:toycars$angle    2   0.01239 0.10566 -141.672
1.3944   0.2700

<none>
toycars$car
toycars$angle               ***
toycars$car:toycars$angle
---
Signif. codes:  0   ***   0.001   **   0.01   *   0.05   .   0.1
>
> backmodel2 = lm(toycars$distance ~ toycars$car + toycars$angle, data=toycars)
>
> drop1(backmodel2, scope = ~ ., test="F")
Single term deletions

Model:
toycars$distance ~ toycars$car + toycars$angle
              Df Sum of Sq     RSS      AIC F value    Pr(>F)
<none>                     0.10566 -141.672
toycars$car    2   0.16945 0.27511 -119.833  18.444 1.662e-05 ***
toycars$angle  1   1.65108 1.75673  -67.774 359.416 1.547e-15 ***
---
Signif. codes:  0   ***   0.001   **   0.01   *   0.05   .   0.1
>
> # all tests are significant, we stop at backmodel2
```

(d)

$H_0 : \tau_1 - \tau_3 = 0.05$

[R code]

```
> linearHypothesis(backmodel2, c(0,0,-1, 0), 0.05)
Linear hypothesis test

Hypothesis:
- toycars$car3 = 0.05

Model 1: restricted model
Model 2: toycars$distance ~ toycars$car + toycars$angle

  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1     24 0.11033
2     23 0.10566  1 0.0046722 1.0171 0.3237
---
Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1
```

p-value = 0.3237 > 0.05, so we cannot reject H_0 at the 5% significance level

(e)

$H_0 : \tau_2 = \tau_3$

[R code]

```
> linearHypothesis(fullmodel, c(0,1,-1, 0, 0, 0), 0)
Linear hypothesis test

Hypothesis:
toycars$car2 - toycars$car3 = 0

Model 1: restricted model
Model 2: toycars$distance ~ toycars$car * toycars$angle

  Res.Df      RSS Df Sum of Sq      F  Pr(>F)
1     22 0.108767
2     21 0.093271  1  0.015497 3.4891 0.07579 .
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1
```

p-value = 0.0758 > 0.05, so we cannot reject H_0 at the 5% significance level

Question 5

(a)

Mu can be regarded as a nuisance parameter since we are not interested in it specifically but it still accounts for some variation in our model

(b)

$$X_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} 1 & 0 & x_{11} \\ 1 & 0 & x_{12} \\ 1 & 0 & x_{13} \\ 0 & 1 & x_{21} \\ 0 & 1 & x_{22} \\ 0 & 1 & x_{23} \end{bmatrix}$$

(c)

[R code]

```
> X1 = matrix(c(rep(1, 6)), 6, 1)
>
> H1 = X1 %*% ginv(t(X1) %*% X1) %*% t(X1)
>
> fractions(diag(6) - H1)
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]  5/6 -1/6 -1/6 -1/6 -1/6 -1/6
[2,] -1/6  5/6 -1/6 -1/6 -1/6 -1/6
[3,] -1/6 -1/6  5/6 -1/6 -1/6 -1/6
[4,] -1/6 -1/6 -1/6  5/6 -1/6 -1/6
[5,] -1/6 -1/6 -1/6 -1/6  5/6 -1/6
[6,] -1/6 -1/6 -1/6 -1/6 -1/6  5/6
```

$$
\begin{aligned}
X_{2|1} = [I - H_1]X_2 &= \begin{bmatrix} \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & x_{11} \\ 1 & 0 & x_{12} \\ 1 & 0 & x_{13} \\ 0 & 1 & x_{21} \\ 0 & 1 & x_{22} \\ 0 & 1 & x_{23} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & x_{11} - \bar{x} \\ \frac{1}{2} & -\frac{1}{2} & x_{12} - \bar{x} \\ \frac{1}{2} & -\frac{1}{2} & x_{13} - \bar{x} \\ -\frac{1}{2} & \frac{1}{2} & x_{21} - \bar{x} \\ -\frac{1}{2} & \frac{1}{2} & x_{22} - \bar{x} \\ -\frac{1}{2} & \frac{1}{2} & x_{23} - \bar{x} \end{bmatrix}
\end{aligned}
$$

(d)

$\vec{b_2} = (X_{2|1}^T X_{2|1})^c X_{2|1}^T \vec{y}$

[R code]

```
> xs = c(2,4,8,7,6,4)
>
> y = c(4,2,10,8,8,12)
>
> xbar = mean(xs)
>
> X21 = matrix(c(0.5,0.5,0.5,-0.5,-0.5,-0.5,-0.5,-0.5,-0.5,0.5,0.5,0.5,
        xs[1]-xbar,xs[2]-xbar,xs[3]-xbar,xs[4]-xbar,xs[5]-xbar,
        xs[6]-xbar), 6,3)
>
> b2 = ginv(t(X21) %*% X21) %*% t(X21) %*% y
>
> b2
           [,1]
[1,] -1.6857143
[2,]  1.6857143
[3,]  0.6285714
```

$$\vec{b_2} = \begin{bmatrix} -1.6857 \\ 1.6857 \\ 0.6286 \end{bmatrix}$$

(e)

$\vec{b_1} = (X_1^T X_1)^c (X_1^T \vec{y} - X_1^T X_2 \vec{b_2})$

[R code]

```
> b1 = ginv(t(X1) %*% X1) %*% (t(X1) %*% y - t(X1) %*% X2 %*% b2)
>
> b1
             [,1]
[1,]  4.085714
>
> X = cbind(X1, X2)
>
> b = rbind(b1,b2)
>
> t(X) %*% X %*% b
      [,1]
[1,]    44
[2,]    16
[3,]    28
[4,]   248
>
> t(X) %*% y
      [,1]
[1,]    44
[2,]    16
[3,]    28
[4,]   248
```

$\vec{b_1} = 4.0857$

$X^T X \vec{b} = X^T \vec{y}$