MATH 668 Homework 4 Solutions

1. (a) By Step 3 of the orthogonalization steps in the Section 7.10 notes,
$\hat{\boldsymbol{\beta}}_1 = \begin{pmatrix} 0.07 \\ -0.05 \\ 0.02 \end{pmatrix}$ and $\hat{\beta}_3 = \hat{\beta}_3^* - \boldsymbol{u}^\top \hat{\boldsymbol{\beta}}_1 = 0.7 - (0.65, 3.10, 1.78) \begin{pmatrix} 0.07 \\ -0.05 \\ 0.02 \end{pmatrix} = 0.7739$ so that
$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\boldsymbol{\beta}}_1 \\ \hat{\beta}_3 \end{pmatrix} = \begin{pmatrix} 0.07 \\ -0.05 \\ 0.02 \\ 0.7739 \end{pmatrix}$.

(b) Here, we test $H_0 : \boldsymbol{\beta}_1 = \mathbf{0}$ versus $H_1 : \boldsymbol{\beta}_1 \neq \mathbf{0}$.
The test statistic is $F = \dfrac{SS(\boldsymbol{\beta}_1|\beta_3)/3}{SSE/(10-4)} = \dfrac{.10863/3}{.20137/6} = 1.0789$

since $SS(\boldsymbol{\beta}_1|\beta_3) = \hat{\boldsymbol{\beta}}^\top \mathbf{X}^\top \boldsymbol{y} - \hat{\beta}_3^* \boldsymbol{x}^\top \boldsymbol{y} = (0.07, -0.05, 0.02, 0.7739) \begin{pmatrix} 1.16 \\ 1.35 \\ 2.16 \\ 0.70 \end{pmatrix} - 0.70(0.70) = 0.10863$ and

$SSE = \boldsymbol{y}^\top \boldsymbol{y} - \hat{\boldsymbol{\beta}}^\top \mathbf{X}^\top \boldsymbol{y} = 0.80 - (0.07, -0.05, 0.02, 0.7739) \begin{pmatrix} 1.16 \\ 1.35 \\ 2.16 \\ 0.70 \end{pmatrix} = 0.20137.$

The critical value is $F_{.05,3,6} = 4.76$.
So, we fail to reject H_0 at level .05 since $F < F_{.05,3,6}$.

2. (a) The maximum likelihood estimate $\hat{\boldsymbol{\beta}} = \begin{pmatrix} -2659.867 \\ 13.49277 \\ 11.43641 \\ 18.61282 \\ -0.01608840 \\ 0.0009565257 \\ 0.003438604 \\ -0.03395494 \\ -0.05570193 \\ -0.01386106 \end{pmatrix}$ can be computed using the built-in R function `lm` as follows.

```
setwd("C:\\Users\\ryan\\Desktop\\S18\\668\\data")
D=read.table("YieldData.txt",header=TRUE)
y=D[,1]
x1=D[,2]
x2=D[,3]
x3=D[,4]
x4=x1*x1
x5=x2*x2
x6=x3*x3
x7=x1*x2
x8=x1*x3
x9=x2*x3
beta.hat=lm(y~x1+x2+x3+x4+x5+x6+x7+x8+x9)$coef
beta.hat
```

```
##   (Intercept)            x1            x2            x3            x4
## -2.659867e+03  1.349277e+01  1.143641e+01  1.861282e+01 -1.608840e-02
```

```
##            x5            x6            x7            x8            x9
##  9.565257e-04  3.438604e-02 -3.395494e-02 -5.570193e-02 -1.386106e-02
```

Then this can be used to compute $\hat{\sigma}^2 = 4.802127$ directly with the following R command.

```
X=cbind(1,x1,x2,x3,x4,x5,x6,x7,x8,x9)
SSE=sum((y-X%*%beta.hat)^2)
n=length(y);n
```

```
## [1] 19
```

```
sigma2.hat=SSE/n
sigma2.hat
```

```
## [1] 4.802127
```

The MLE of $\boldsymbol{\beta}$ can be computed directly with some difficulty because of some numerical issues. If we try directly, the calculation fails.

```
solve(t(X)%*%X)%*%t(X)%*%y
```

However, we can rescale the columns of `X` to avoid these computational issues.

```
Xs=cbind(1,x1/10,x2,x3,x1*x1/100,x2*x2,x3*x3,x1*x2/10,x1*x3/10,x2*x3)
beta.hat.s=solve(t(Xs)%*%Xs)%*%t(Xs)%*%y
beta.hat=beta.hat.s/c(1,10,1,1,100,1,1,10,10,1)
beta.hat
```

```
##             [,1]
##    -2.659867e+03
##     1.349277e+01
## x2  1.143641e+01
## x3  1.861282e+01
##    -1.608840e-02
##     9.565257e-04
##     3.438604e-02
##    -3.395494e-02
##    -5.570193e-02
##    -1.386106e-02
```

(b) Here is the F-test in Theorem 8.2.1.
We test $H_0 : \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$.
The test statistic can be computed $F = \dfrac{\boldsymbol{y}^\top(\mathbf{H} - \mathbf{H}_1)\boldsymbol{y}/6}{\boldsymbol{y}^\top(\mathbf{H} - \mathbf{H}_1)\boldsymbol{y}/(19-9-1)} = 3.098427$ using the following R code.

```
X1=cbind(1,x1,x2,x3)
beta.hat.1=solve(t(X1)%*%X1)%*%t(X1)%*%y
Q1=sum((y-X1%*%beta.hat.1)^2)
Q=SSE
F=((Q1-Q)/6)/(Q/(19-10))
F
```

```
## [1] 3.098427
```

The critical value is $F_{.05,6,9} = 3.373754$ as computed below in R.

```
qf(.95,6,9)
```

```
## [1] 3.373754
```

So, we fail to reject H_0 at level .05 since $F < F_{.05,6,9}$.

3. (a) We know $\mathbf{X}^\top\mathbf{X} = \begin{pmatrix} \boldsymbol{j}^\top \\ \mathbf{X}_1^\top \end{pmatrix} (\boldsymbol{j}, \mathbf{X}_1) = \begin{pmatrix} \boldsymbol{j}^\top\boldsymbol{j} & \boldsymbol{j}^\top\mathbf{X}_1 \\ \mathbf{X}_1^\top\boldsymbol{j} & \mathbf{X}_1^\top\mathbf{X}_1 \end{pmatrix}$. Then we use the inverse formula for block matrices (Theorem 2.5.3) to find that the inverse of the lower right block of the inverse of $\mathbf{X}^\top\mathbf{X}$ is $\mathbf{X}_1^\top\mathbf{X}_1 - \mathbf{X}_1^\top\boldsymbol{j}(\boldsymbol{j}^\top\boldsymbol{j})^{-1}\boldsymbol{j}^\top\mathbf{X}_1 = \mathbf{X}_1^\top\mathbf{X}_1 - \mathbf{X}_1^\top\dfrac{\boldsymbol{j}\boldsymbol{j}^\top}{\boldsymbol{j}^\top\boldsymbol{j}}\mathbf{X}_1 = \mathbf{X}_1^\top\mathbf{X}_1 - \mathbf{X}_1^\top\dfrac{1}{n}\mathbf{J}\mathbf{X}_1 = \mathbf{X}_1^\top\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)\mathbf{X}_1$.
It follows that
$\mathbf{C}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{C}^\top = (\mathbf{0}, \mathbf{I}) \begin{pmatrix} ? & ? \\ ? & \left(\mathbf{X}_1^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{X}_1\right)^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0}^\top \\ \mathbf{I} \end{pmatrix} = \left(\mathbf{X}_1^\top\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)\mathbf{X}_1\right)^{-1}$
so that $\left(\mathbf{C}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{C}^\top\right)^{-1} = \mathbf{X}_1^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{X}_1$.

(b) $\mathbf{X}^\top\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)\mathbf{X} = \begin{pmatrix} \boldsymbol{j}^\top \\ \mathbf{X}_1^\top \end{pmatrix}\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)(\boldsymbol{j}, \mathbf{X}_1) = \begin{pmatrix} \boldsymbol{j}^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right) \\ \mathbf{X}_1^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right) \end{pmatrix}(\boldsymbol{j}, \mathbf{X}_1) = \begin{pmatrix} \boldsymbol{j}^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\boldsymbol{j} & \boldsymbol{j}^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{X}_1 \\ \mathbf{X}_1^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\boldsymbol{j} & \mathbf{X}_1^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{X}_1 \end{pmatrix} =$
$\begin{pmatrix} \boldsymbol{j}^\top\boldsymbol{j} - \frac{1}{n}\boldsymbol{j}^\top\mathbf{J}\boldsymbol{j} & \boldsymbol{j}^\top\mathbf{X}_1 - \frac{1}{n}\boldsymbol{j}^\top\mathbf{J}\mathbf{X}_1 \\ \mathbf{X}_1^\top\boldsymbol{j} - \frac{1}{n}\mathbf{X}_1^\top\mathbf{J}\boldsymbol{j} & \mathbf{X}_1^\top\mathbf{X} - \mathbf{X}_1^\top\frac{1}{n}\mathbf{J}\mathbf{X}_1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{j}^\top\boldsymbol{j} - \frac{1}{n}\boldsymbol{j}^\top\boldsymbol{j}\boldsymbol{j}^\top\boldsymbol{j} & \boldsymbol{j}^\top\mathbf{X}_1 - \frac{1}{n}\boldsymbol{j}^\top\boldsymbol{j}\boldsymbol{j}^\top\mathbf{X}_1 \\ \mathbf{X}_1^\top\boldsymbol{j} - \frac{1}{n}\mathbf{X}_1^\top\boldsymbol{j}\boldsymbol{j}^\top\boldsymbol{j} & \mathbf{X}_1^\top\mathbf{X}_1 - \mathbf{X}_1^\top\frac{1}{n}\mathbf{J}\mathbf{X}_1 \end{pmatrix} =$
$\begin{pmatrix} n - \frac{1}{n}n^2 & \boldsymbol{j}^\top\mathbf{X}_1 - \frac{1}{n}n\boldsymbol{j}^\top\mathbf{X}_1 \\ \mathbf{X}_1^\top\boldsymbol{j} - \frac{1}{n}\mathbf{X}_1^\top\boldsymbol{j}n & \mathbf{X}_1^\top\mathbf{X}_1 - \mathbf{X}_1^\top\frac{1}{n}\mathbf{J}\mathbf{X}_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{X}_1^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{X}_1 \end{pmatrix}$

(c) Since $C\hat{\boldsymbol{\beta}} = (\mathbf{0}, \mathbf{I})\begin{pmatrix} \hat{\beta}_0 \\ \hat{\boldsymbol{\beta}}_1 \end{pmatrix} = \hat{\boldsymbol{\beta}}_1$,

we see from (a) that $(\mathbf{C}\hat{\boldsymbol{\beta}})^\top\left(\mathbf{C}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{C}^\top\right)^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}}) = \hat{\boldsymbol{\beta}}_1^\top\mathbf{X}_1^\top\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)\mathbf{X}_1\boldsymbol{\beta}_1$

and we see from (b) that $\hat{\boldsymbol{\beta}}^\top\mathbf{X}^\top\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)\mathbf{X}\hat{\boldsymbol{\beta}} = \left(\hat{\beta}_0, \hat{\boldsymbol{\beta}}_1^\top\right)\begin{pmatrix} 0 & 0 \\ 0 & \mathbf{X}_1^\top\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{X}_1 \end{pmatrix}\begin{pmatrix} \hat{\beta}_0 \\ \hat{\boldsymbol{\beta}}_1 \end{pmatrix} = \hat{\boldsymbol{\beta}}_1^\top\mathbf{X}_1^\top\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)\mathbf{X}_1\boldsymbol{\beta}_1$

so it follows that $(\mathbf{C}\hat{\boldsymbol{\beta}})^\top\left(\mathbf{C}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{C}^\top\right)^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}}) = \hat{\boldsymbol{\beta}}^\top\mathbf{X}^\top\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)\mathbf{X}\hat{\boldsymbol{\beta}}$.

Note that this shows that the test of $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ in Theorem 8.4.2 with $\mathbf{C} = (\mathbf{0}, \mathbf{I})$ is equivalent to the test of $H_0 : \boldsymbol{\beta}_1 = \mathbf{0}$ in Theorem 8.1.1 since
$\hat{\boldsymbol{\beta}}^\top\mathbf{X}^\top\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)\mathbf{X}\hat{\boldsymbol{\beta}} = \boldsymbol{y}^\top\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\left(\mathbf{I} - \dfrac{1}{n}\mathbf{J}\right)\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\boldsymbol{y} =$
$\boldsymbol{y}^\top\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\boldsymbol{y} - \boldsymbol{y}^\top\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\dfrac{1}{n}\mathbf{J}\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\boldsymbol{y} =$
$\boldsymbol{y}^\top\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\boldsymbol{y} - \boldsymbol{y}^\top\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\dfrac{1}{n}\mathbf{J}\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\boldsymbol{y} = \boldsymbol{y}^\top\mathbf{H}\boldsymbol{y} - \boldsymbol{y}^\top\mathbf{H}\left(\dfrac{1}{n}\mathbf{J}\right)\mathbf{H}\boldsymbol{y} =$
$\boldsymbol{y}^\top\mathbf{H}\boldsymbol{y} - \boldsymbol{y}^\top\left(\dfrac{1}{n}\mathbf{J}\right)\boldsymbol{y} = \boldsymbol{y}^\top\left(\mathbf{H} - \dfrac{1}{n}\mathbf{J}\right)\boldsymbol{y}$.
We will see that $\mathbf{HJH} = \mathbf{J}$ in Chapter 9.