MAST30025 Assignment 2 2021 Michael Le LaTex

Michael Le (998211)

April 20, 2021

Question 1 Solution:
Since

$$\hat{\sigma^2} = \frac{SS_{Res}}{n}$$

is a biased estimator:
<u>Likelihood:</u>
$\mathbf{L}(\beta|\sigma^2) =$

$$\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-\epsilon_i^2}{2\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{\frac{-\sum_{n=1}^{\infty} \epsilon_i^2}{2\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{\frac{-(y-X\beta)^T(y-X\beta)}{2\sigma^2}}$$

<u>Log Likelihood:</u>
$\mathbf{logL}(\beta|\sigma^2) =$

$$\frac{-nlog(2\pi\sigma^2)}{2} - \frac{(y-X\beta)^T(y-X\beta)}{2\sigma^2}$$

Differentiating w.r.t σ^2:

$$\frac{\partial logL(\beta = b|\sigma^2)}{\partial\sigma^2} = \frac{-n}{2}\frac{2\pi}{2\pi\sigma^2} + \frac{1}{2\sigma^4}(y-Xb)^T(y-Xb) = 0$$

$$\frac{(y-Xb)^T(y-Xb)}{2\sigma^4} = \frac{n}{2\sigma^2}$$

$$\hat{\sigma^2} = \frac{(y-Xb)^T(y-Xb)}{n} = \frac{SS_{Res}}{n}$$

which requires formula on the substitution of the ML estimators $\boldsymbol{b} for \beta$.
Question 2 Solution:
Part a:

```r
n = 7
p = 4
X =
matrix(c(rep(1,n),32,19.5,13.3,13.3,5,7.1,34.5,84.9,306.6,562,562,390.6,2175,623.5,10,9,5,5,5,3,7
),n,p)
y = c(37.9,42.2,47.3,43.1,54.8,47.1,40.3)
b = solve(t(X)%*%X,t(X)%*%y)
b
##             [,1]
## [1,] 58.369312708
## [2,] -0.346291960
## [3,] -0.002900359
## [4,] -0.887671692
s2 = sum((y-X%*%b)^2)/(n-p)
s2
## [1] 13.06871
```

Part b:

```r
xst = as.vector(c(1,10,100,6))
xst %*% b + c(-1,1)*qt(0.95,df=n-p)*sqrt(s2 * t(xst) %*% solve(t(X)%*%X) %*% xst)
```

```
## [1] 43.27252 55.30814
```

Part c:

```r
#Attempt 1
tst = c(0,1,0,-1)
#Calculating the Sample Standand Derivation!
s = sqrt(s2)

#Standand error for beta1 - beta3
s*sqrt(t(tst)%*%solve(t(X)%*%X)%*%tst)
##          [,1]
## [1,] 1.388968
```

Part d:

```
#Attempt 1
SSRes = sum((y-X %*% b)^2)
SSReg = sum(y^2)-SSRes
Fstat = (SSReg/p)/(SSRes/(n-p))
pf(Fstat,p,n-p,lower=F)
## [1] 0.000363714

#We reject the null under 5% signifance level!
```

Part e:

```
#Slide 61-63 IFTFRM

SSReg = t(y) %*% X %*% b - sum(y)^2 / n
SSReg
##        [,1]
## [1,] 149.7282
SSRes = s2*(n-p)
SSRes
## [1] 39.20612
Fstat = (SSReg/(p-1))/(SSRes/(n-p))
Fstat
##      [,1]
## [1,] 3.819
pf(Fstat, p-1, n-p, lower.tail = FALSE)
##          [,1]
## [1,] 0.1500833

#We do not reject the null hypothesis of the model relevance!
```

Question 3 Solution:

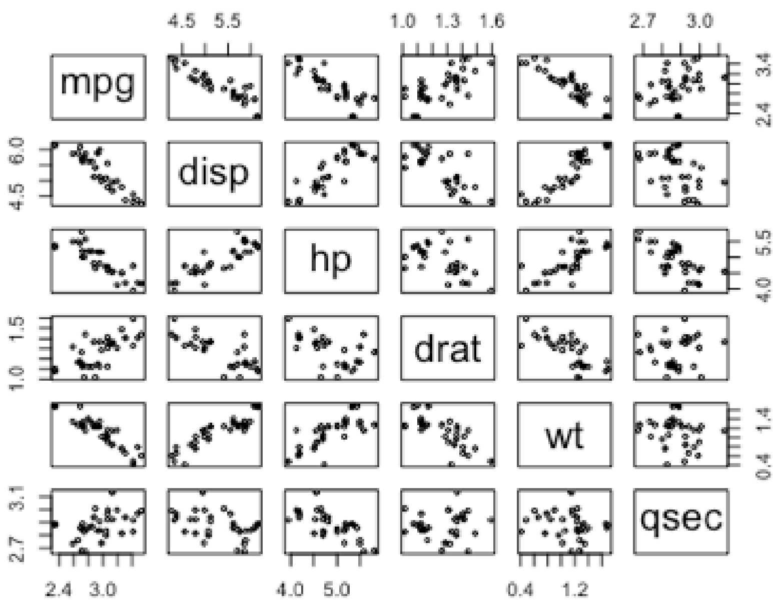
We are given that $\beta = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}$

Now consider the parameters of a reduced model y $=$ Xγ $+$ ϵ, which are γ $= [\beta_0, ...,\beta_r\, 0, ..., 0]^T$ *where r is the number of parameters* in γ_1 and the remaining k-r remaining parameters in β are 0. The reduced model y $=$ X$_1\gamma_1$ $+$ ϵ_1 minimizes SS$_{Res(reduced)}$, *the full model* $y = X\beta + \epsilon_2$ must have SS$_{Res(full)}$.

$SS_Res(full) - SS_Res(reduced) \geq 0$
Which is positive semi-definite. The $SS_Res\ for$
the reduced model is at least the $SS_Res\ for\ the\ full\ model.$

Question 4 Part a Solution:

```
data(mtcars)
mtcars = log(mtcars[, c(1,3:7)])
pairs(mtcars,cex=0.5)
```



Looking at miles per gallon against the other variables, there is evidence of a linear relationship with displacement,gross horsepower,rear axle ratio,weight and a quarter mile time!

Question 4 Part b Solution:

```
model0 = lm(mpg ~ 1, data=mtcars)
add1(model0, scope = ~.+disp+hp+drat+wt+qsec, test = "F")
## Single term additions
##
## Model:
## mpg ~ 1
##        Df Sum of Sq     RSS      AIC  F value    Pr(>F)
## <none>               2.74874  -76.547
## disp    1   2.25596 0.49277 -129.550 137.3427 1.006e-12 ***
## hp      1   1.96733 0.78140 -114.797  75.5310 1.080e-09 ***
## drat    1   1.23131 1.51742  -93.559  24.3435 2.807e-05 ***
## wt      1   2.21452 0.53422 -126.966 124.3596 3.406e-12 ***
## qsec    1   0.47755 2.27119  -80.654   6.3079   0.01763 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

#We take out the displacement variable

```
model1 = lm(mpg ~ 1+disp, data=mtcars)
add1(model1, scope = ~.+hp+drat+wt+qsec, test = "F")
## Single term additions
##
## Model:
## mpg ~ 1 + disp
##        Df Sum of Sq     RSS     AIC F value  Pr(>F)
## <none>              0.49277 -129.55
## hp      1  0.045531 0.44724 -130.65  2.9523 0.09641 .
## drat    1  0.001383 0.49139 -127.64  0.0816 0.77711
## wt      1  0.098796 0.39398 -134.71  7.2722 0.01154 *
## qsec    1  0.000308 0.49247 -127.57  0.0181 0.89382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

#Take out the weight variable

```
model2 = lm(mpg ~ 1+disp+wt, data=mtcars)
add1(model2, scope = ~.+hp+drat+qsec, test = "F")
## Single term additions
##
## Model:
## mpg ~ 1 + disp + wt
##        Df Sum of Sq     RSS     AIC F value  Pr(>F)
## <none>              0.39398 -134.71
## hp      1  0.078605 0.31537 -139.83  6.9789 0.01334 *
## drat    1  0.007358 0.38662 -133.31  0.5329 0.47146
```

```
## Model:
## mpg ~ 1 + disp + wt
##        Df Sum of Sq     RSS     AIC F value  Pr(>F)
## <none>              0.39398 -134.71
## hp      1  0.078605 0.31537 -139.83  6.9789 0.01334 *
## drat    1  0.007358 0.38662 -133.31  0.5329 0.47146
## qsec    1  0.057788 0.33619 -137.79  4.8130 0.03671 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#We take out horsepower
model3 = lm(mpg ~ 1+disp+wt+hp, data=mtcars)
add1(model3, scope = ~.+drat+qsec, test = "F")
## Single term additions
##
## Model:
## mpg ~ 1 + disp + wt + hp
##        Df Sum of Sq     RSS     AIC F value Pr(>F)
## <none>               0.31537 -139.83
## drat    1 0.0000095 0.31536 -137.83  0.0008 0.9774
## qsec    1 0.0033067 0.31206 -138.17  0.2861 0.5971

#The final variables are disp,wt and hp!
```

Question 4 Part c Solution:

```r
model = step(model0, scope = ~ .+disp+hp+drat+wt+qsec)
## Start:  AIC=-76.55
## mpg ~ 1
##
##        Df Sum of Sq     RSS      AIC
## + disp  1   2.25596 0.49277 -129.550
## + wt    1   2.21452 0.53422 -126.966
## + hp    1   1.96733 0.78140 -114.797
## + drat  1   1.23131 1.51742  -93.559
## + qsec  1   0.47755 2.27119  -80.654
## <none>              2.74874  -76.547
##
## Step:  AIC=-129.55
## mpg ~ disp
##
##        Df Sum of Sq     RSS      AIC
## + wt    1   0.09880 0.39398 -134.710
## + hp    1   0.04553 0.44724 -130.652
## <none>              0.49277 -129.550
## + drat  1   0.00138 0.49139 -127.640
## + qsec  1   0.00031 0.49247 -127.570
## - disp  1   2.25596 2.74874  -76.547
##
## Step:  AIC=-134.71
## mpg ~ disp + wt
##
##        Df Sum of Sq     RSS     AIC
## + hp    1  0.078605 0.31537 -139.83
## + qsec  1  0.057788 0.33619 -137.79
## <none>              0.39398 -134.71
## + drat  1  0.007358 0.38662 -133.31
## - wt    1  0.098796 0.49277 -129.55
## - disp  1  0.140243 0.53422 -126.97
##
## Step:  AIC=-139.83
## mpg ~ disp + wt + hp
##
##        Df Sum of Sq     RSS     AIC
## - disp  1  0.006635 0.32201 -141.16
## <none>              0.31537 -139.83
## + qsec  1  0.003307 0.31207 -138.17
```

```
##
## Step:  AIC=-139.83
## mpg ~ disp + wt + hp
##
##        Df Sum of Sq     RSS     AIC
## - disp  1  0.006635 0.32201 -141.16
## <none>              0.31537 -139.83
## + qsec  1  0.003307 0.31207 -138.17
## + drat  1  0.000010 0.31536 -137.83
## - hp    1  0.078605 0.39398 -134.71
## - wt    1  0.131870 0.44724 -130.65
##
## Step:  AIC=-141.17
## mpg ~ wt + hp
##
##        Df Sum of Sq     RSS     AIC
## <none>              0.32201 -141.16
## + disp  1   0.00664 0.31537 -139.83
## + qsec  1   0.00557 0.31644 -139.72
## + drat  1   0.00112 0.32089 -139.28
## - hp    1   0.21221 0.53422 -126.97
## - wt    1   0.45939 0.78140 -114.80
```

Housepower and Weight are the variables in the final model.

Question 4 Part d Solution:

```
model
##
## Call:
## lm(formula = mpg ~ wt + hp, data = mtcars)
##
## Coefficients:
## (Intercept)          wt          hp
##      4.8347     -0.5623     -0.2553
```

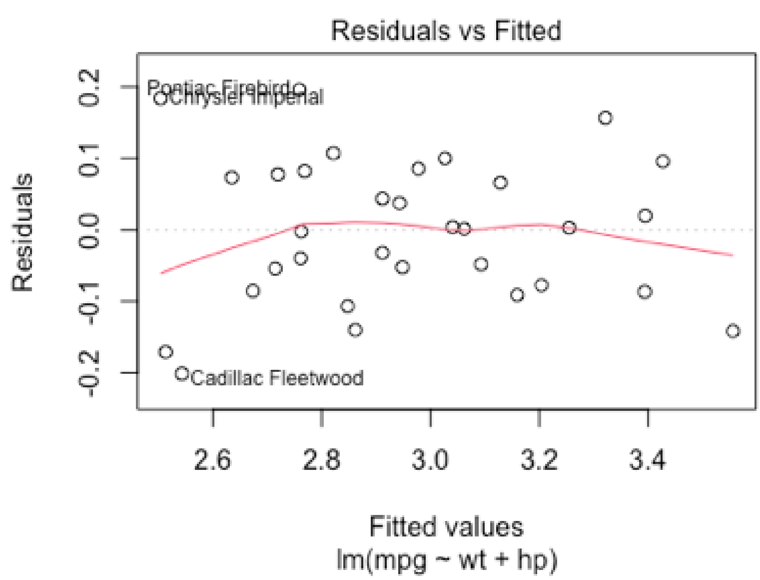
Were are dealing with a log transformation. In the final model is

$$log(mpg) = 4.8347 - 0.2553log(hp) - 0.5623log(wt) + \epsilon$$

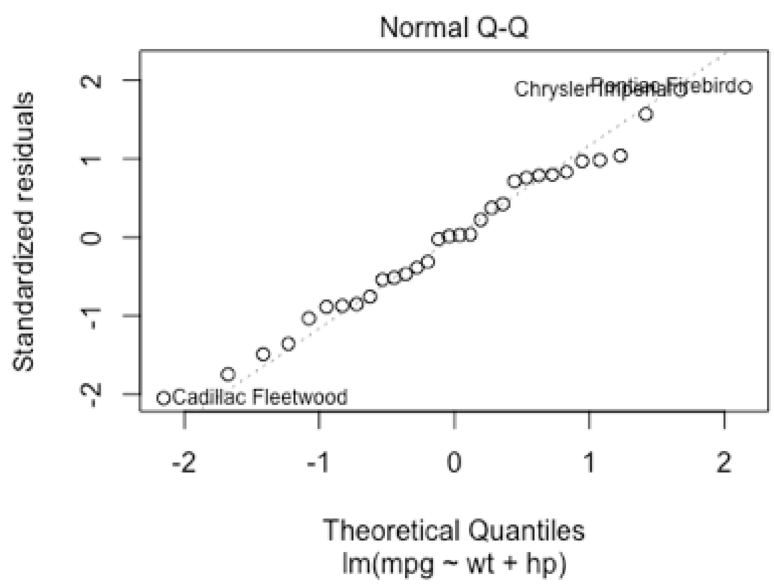
Take the exponential on both sides of the linear model mpg $= \exp^{4.8347} hp^{-0.2553} wt^{-0.5623} \epsilon^{'}$. Where $\epsilon^{'}$ = exp(e).

Question 4 Part e Solution:

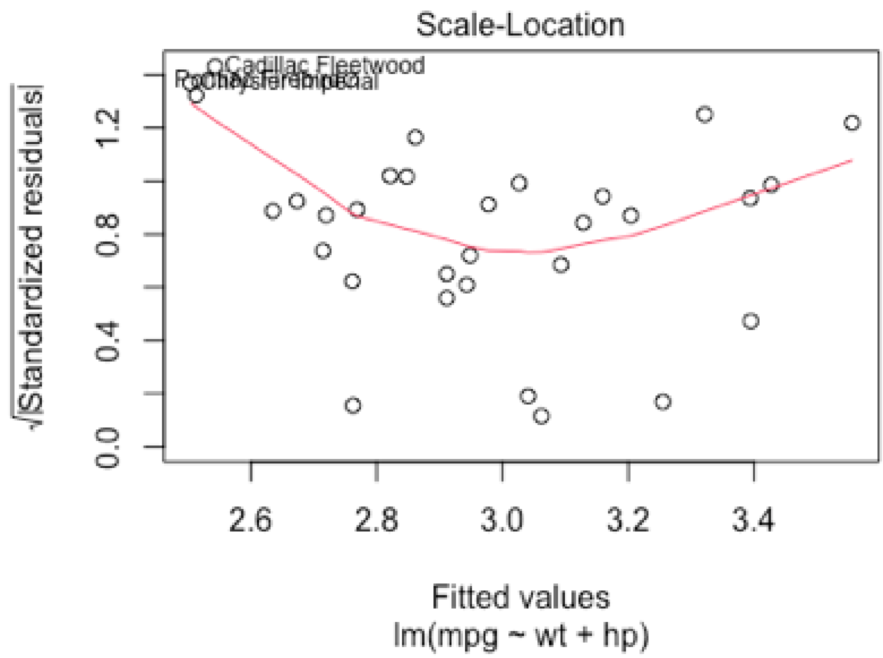
```
plot(model, which=1)
```



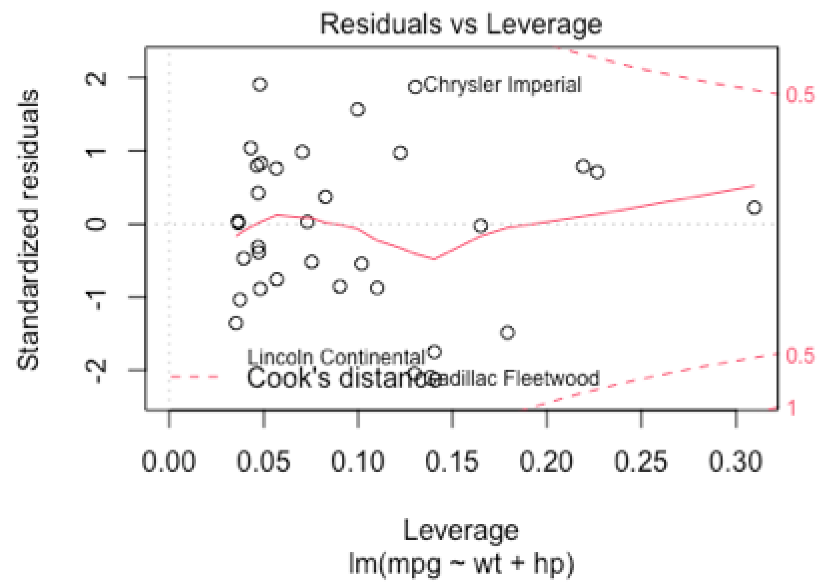
```r
plot(model, which=2)
```



```
plot(model, which=3)
```



```
plot(model, which=5)
```



#Diagnostic plots show a reasonable fit to linear model assumptions. About the only area of concern is a slight positive trend for higher fitted values and moderate leverages, but this does not appear to be too alarming.

Diagnostic plots show a reasonable fit to linear model assumptions. About the only area of concern is a slight positive trend for higher fitted values and moderate leverages, but this does not appear to be too alarming.

Question 5 Part a Solution:

$\sum_{i=1}^{n} \varepsilon_i^2 + \lambda \sum_{j=0}^{k} b_j^2 = (y - Xb)^T (y - Xb) + \lambda b^T b$

$=$ y$^T y - 2(X^T y)^T b + b^T (X^T X) b + \lambda b^T b$

Now were differentiating w.r.t to b:

$$\frac{\partial(\varepsilon^T \varepsilon + \lambda b^T b)}{\partial b}$$

$=$ 0 - 2X$^T y + X^T X b + (X^T X)^T b + 2\lambda I b$

$=$ 0 - 2X$^T y + 2(X^T X) b + 2\lambda I b$

since X$^T X and (X^T X)^T are symmetric$!

Now equate them to 0 and solve b:

0 - 2X$^T y + 2(X^T X) b + 2\lambda I b = 0$

X$^T y = b(X^T X + \lambda I)$

$$b = (X^T X + \lambda I)^{-1} X^T y$$

Question 5 Part b Solution:

Using Theorem 4.4 (Gauss-Markov Theorem)

$$E[b] = (X^T X + \lambda I)^{-1} X^T E[y])$$
$$= (X^T X + \lambda I)^{-1} X^T X \beta \geq \beta$$

is biased!

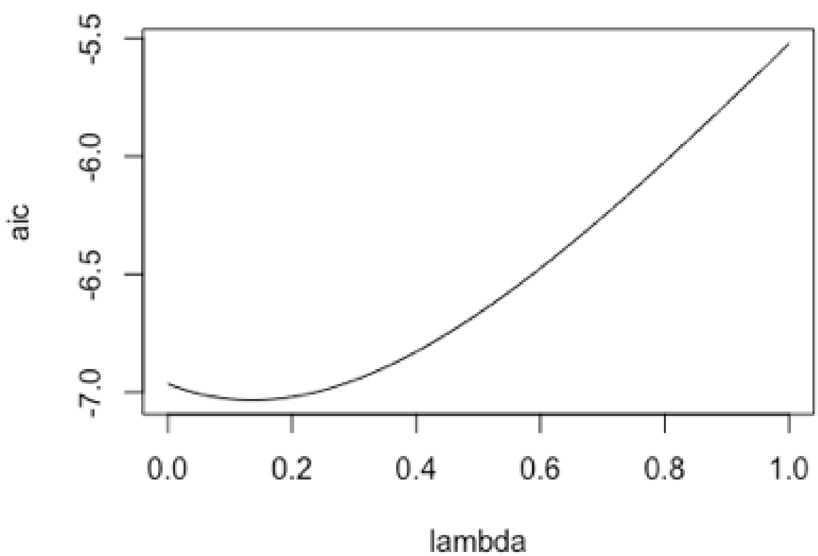
So b is an unbiased estimator for beta, we know that E[b]=β. Therefore (X$^T X$+λI)$^{-1} X^T X \beta$ which means $\lambda = 0$. Then (X$^T X)^{-1} X^T X = I\beta = \beta$

Question 5 Part c Solution:

Solution:

```
Xs <- scale(X[,-1],center=T,scale=T)
ys <- scale(y,center=T,scale=F)
p = 4
p <- p-1
solve(t(Xs)%*%Xs + diag(rep(0.5,p)),t(Xs)%*%ys)
##          [,1]
## [1,] 0.3494789
## [2,] 1.7899861
## [3,] 0.3432961
```

```
n = 8
lambda <- seq(0,1,0.001)
aic <- c()
for (l in lambda) {
    b <- solve(t(Xs)%*%Xs + diag(rep(l,p)),t(Xs)%*%ys)
    ssres <- sum((ys-Xs%*%b)^2)
    H <- Xs %*% solve(t(Xs)%*%Xs + diag(rep(l,p))) %*% t(Xs)
    aic <- c(aic, n*log(ssres/n) + 2*sum(diag(H)))
}
lambda[which.min(aic)]
## [1] 0.136
plot(lambda,aic,type='l')
```



END OF ASSIGNMENT