A2.R

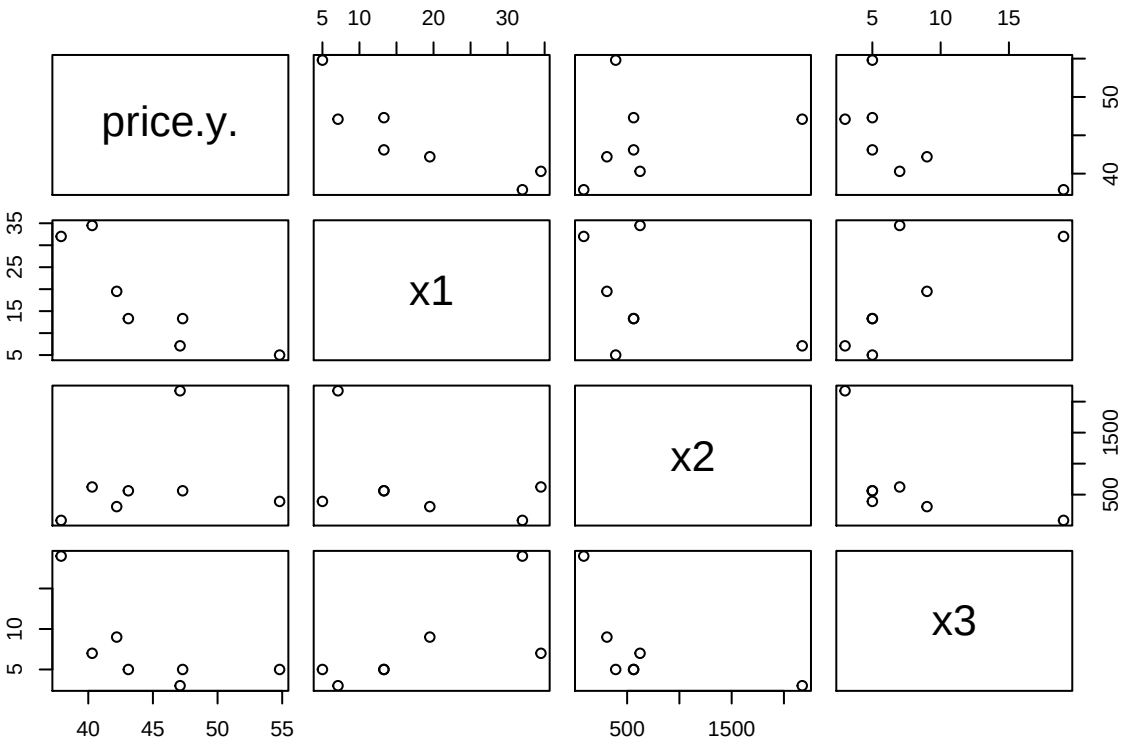
kschy

2021-04-30

```
#Question 2

#Pre Question
#Loading Data
setwd("~/Desktop/University Works & Notes/2021/Semester 1/Linear Statistical Models/Assignment/A2/R Cod
q2data <- data.matrix(read.csv(file = "Q2_Data.csv"))
q2frame <- read.csv(file="Q2_Data.csv")

#Viewing data frame
pairs(q2frame)
```




```
#Defininting X and Y
y <- matrix(q2data[,1],7,1)
y
```

```
##      [,1]
## [1,] 37.9
## [2,] 42.2
## [3,] 47.3
```

```
## [4,] 43.1
## [5,] 54.8
## [6,] 47.1
## [7,] 40.3
```

```
x <- matrix(c(rep(1,7),q2data[,-1]),7,4)
x
```

```
##      [,1] [,2]   [,3] [,4]
## [1,]    1 32.0   84.9   19
## [2,]    1 19.5  306.6    9
## [3,]    1 13.3  562.0    5
## [4,]    1 13.3  562.0    5
## [5,]    1  5.0  390.6    5
## [6,]    1  7.1 2175.0    3
## [7,]    1 34.5  623.5    7
```

```
df <- 7-4

#Finding Beta
b <- solve(t(x)%*%x,t(x)%*%y)
b
```

```
##              [,1]
## [1,] 54.776606226
## [2,] -0.389598784
## [3,] -0.001973937
## [4,] -0.242767764
```

```
#Finding variance
#sum-Square
e <- (y-x%*%b)
SSres <- sum(e^2)
s2 <- SSres/(df)
s <- sqrt(s2)

#Beta Variance
C2x <- solve(t(x)%*%x)*s2
diag(C2x)
```

```
## [1] 1.964791e+01 3.378471e-02 7.330554e-06 1.870117e-01
```

```
#Part B Computing CI

alpha <- 0.1
x.star <- c(1,10,100,6)
y.star <- x.star%*%b

ta <- qt(1-alpha/2, df)
ta
```

```
## [1] 2.353363
```

```
#90CI for x1=10, x2= 100 ,x3 =6

CI = c(y.star - s*sqrt(t(x.star)%*%solve(t(x)%*%x)%*%x.star),y.star + s*sqrt(t(x.star)%*%solve(t(x)%*%x
CI
```

```
## [1] 46.59336 51.85988
```

```r
#Part C
C <- c(0,1,0,-1)
cdelta.star <- matrix(0)

Cb.var <- t(C)%*%solve(t(x)%*%x)%*%C*s2
Cb.var
```

```
##           [,1]
## [1,] 0.316463
```

```r
Cb.ste <- sqrt(Cb.var)
Cb.ste
```

```
##           [,1]
## [1,] 0.5625504
```

```r
#Double Checking
model1 <- lm(price.y. ~ x1+x2+x3, data = q2frame)
summary(model1)
```

```
##
## Call:
## lm(formula = price.y. ~ x1 + x2 + x3, data = q2frame)
##
## Residuals:
##        1        2        3        4        5        6        7
##  0.37073 -2.18931  0.02825 -4.17175  3.95625  0.11116  1.89468
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 54.776606   4.432596  12.358  0.00114 **
## x1          -0.389599   0.183806  -2.120  0.12423
## x2          -0.001974   0.002707  -0.729  0.51873
## x3          -0.242768   0.432448  -0.561  0.61374
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.723 on 3 degrees of freedom
## Multiple R-squared:  0.7799, Adjusted R-squared:  0.5598
## F-statistic: 3.543 on 3 and 3 DF,  p-value: 0.1632
```

```r
#Question 4
data(mtcars)
mtcars
```

```
##                    mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
## Duster 360        14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
## Merc 240D         24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
## Merc 230          22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
## Merc 280          19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
## Merc 280C         17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
```

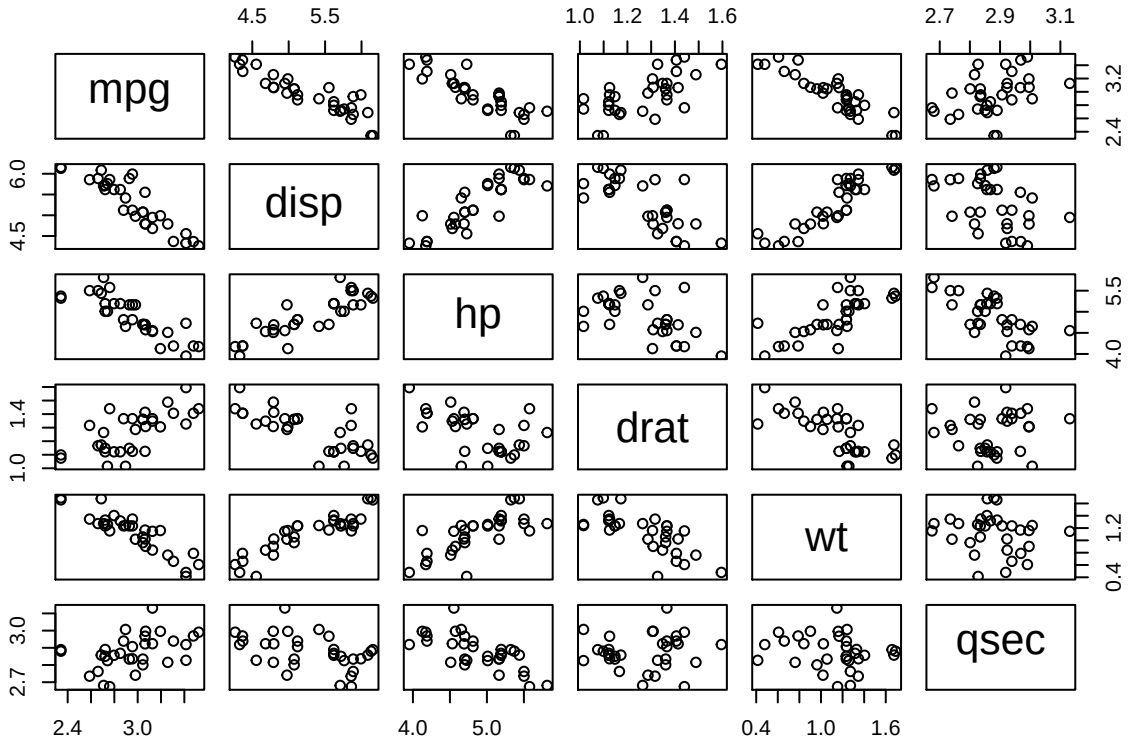
```
## Merc 450SE          16.4   8 275.8 180 3.07 4.070 17.40  0  0    3    3
## Merc 450SL          17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
## Merc 450SLC         15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
## Cadillac Fleetwood  10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
## Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
## Chrysler Imperial   14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
## Fiat 128            32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
## Honda Civic         30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
## Toyota Corolla      33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
## Toyota Corona       21.5   4 120.1  97 3.70 2.465 20.01  1  0    3    1
## Dodge Challenger    15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
## AMC Javelin         15.2   8 304.0 150 3.15 3.435 17.30  0  0    3    2
## Camaro Z28          13.3   8 350.0 245 3.73 3.840 15.41  0  0    3    4
## Pontiac Firebird    19.2   8 400.0 175 3.08 3.845 17.05  0  0    3    2
## Fiat X1-9           27.3   4  79.0  66 4.08 1.935 18.90  1  1    4    1
## Porsche 914-2       26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
## Lotus Europa        30.4   4  95.1 113 3.77 1.513 16.90  1  1    5    2
## Ford Pantera L      15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
## Ferrari Dino        19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
## Maserati Bora       15.0   8 301.0 335 3.54 3.570 14.60  0  1    5    8
## Volvo 142E          21.4   4 121.0 109 4.11 2.780 18.60  1  1    4    2
```

```
mtcars.new = log(mtcars[, c(1,3:7)])
mtcars.new
```

```
##                           mpg     disp       hp     drat        wt     qsec
## Mazda RX4           3.044522 5.075174 4.700480 1.360977 0.9631743 2.800933
## Mazda RX4 Wag       3.044522 5.075174 4.700480 1.360977 1.0560527 2.834389
## Datsun 710          3.126761 4.682131 4.532599 1.348073 0.8415672 2.923699
## Hornet 4 Drive      3.063391 5.552960 4.700480 1.124930 1.1678274 2.967333
## Hornet Sportabout   2.928524 5.886104 5.164786 1.147402 1.2354715 2.834389
## Valiant             2.895912 5.416100 4.653960 1.015231 1.2412686 3.006672
## Duster 360          2.660260 5.886104 5.501258 1.166271 1.2725656 2.762538
## Merc 240D           3.194583 4.988390 4.127134 1.305626 1.1600209 2.995732
## Merc 230            3.126761 4.947340 4.553877 1.366092 1.1474025 3.131137
## Merc 280            2.954910 5.121580 4.812184 1.366092 1.2354715 2.906901
## Merc 280C           2.879198 5.121580 4.812184 1.366092 1.2354715 2.939162
## Merc 450SE          2.797281 5.619676 5.192957 1.121678 1.4036430 2.856470
## Merc 450SL          2.850707 5.619676 5.192957 1.121678 1.3164082 2.867899
## Merc 450SLC         2.721295 5.619676 5.192957 1.121678 1.3297240 2.890372
## Cadillac Fleetwood  2.341806 6.156979 5.323010 1.075002 1.6582281 2.889260
## Lincoln Continental 2.341806 6.131226 5.370638 1.098612 1.6908336 2.880321
## Chrysler Imperial   2.687847 6.086775 5.438079 1.172482 1.6761615 2.857619
## Fiat 128            3.478158 4.365643 4.189655 1.406097 0.7884574 2.968875
## Honda Civic         3.414443 4.326778 3.951244 1.595339 0.4793350 2.918851
## Toyota Corolla      3.523415 4.264087 4.174387 1.439835 0.6070445 2.990720
## Toyota Corona       3.068053 4.788325 4.574711 1.308333 0.9021918 2.996232
## Dodge Challenger    2.740840 5.762051 5.010635 1.015231 1.2584610 2.825537
## AMC Javelin         2.721295 5.717028 5.010635 1.147402 1.2340169 2.850707
## Camaro Z28          2.587764 5.857933 5.501258 1.316408 1.3454724 2.735017
## Pontiac Firebird    2.954910 5.991465 5.164786 1.124930 1.3467736 2.836150
## Fiat X1-9           3.306887 4.369448 4.189655 1.406097 0.6601073 2.939162
## Porsche 914-2       3.258097 4.789989 4.510860 1.488400 0.7608058 2.815409
## Lotus Europa        3.414443 4.554929 4.727388 1.327075 0.4140944 2.827314
## Ford Pantera L      2.760010 5.860786 5.575949 1.439835 1.1537316 2.674149
```

```
## Ferrari Dino         2.980619 4.976734 5.164786 1.286474 1.0188473 2.740840
## Maserati Bora        2.708050 5.707110 5.814131 1.264127 1.2725656 2.681022
## Volvo 142E           3.063391 4.795791 4.691348 1.413423 1.0224509 2.923162
```

```
#Part A: Plotting Data
pairs(mtcars.new)
```




```
#Part B: Forward Selection

basemodel <- lm(mpg~1, data=mtcars.new)
add1(basemodel, 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
```

```
q4model2 <- lm(mpg ~ disp, data=mtcars.new)
add1(q4model2, scope = ~.+hp+drat+wt+qsec, test="F")
```

```
## Single term additions
##
## Model:
```

```
## mpg ~ 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
```

```
q4model3 <- lm(mpg~disp+wt, data=mtcars.new)
add1(q4model3, scope = ~.+hp+drat+qsec, test="F")
```

```
## Single term additions
##
## Model:
## mpg ~ 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
```

```
q4model4 <- lm(mpg~disp+hp+wt, data=mtcars.new)
add1(q4model4, scope = ~.+drat+qsec, test="F")
```

```
## Single term additions
##
## Model:
## mpg ~ disp + hp + wt
##        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
```

```
summary(q4model4)
```

```
##
## Call:
## lm(formula = mpg ~ disp + hp + wt, data = mtcars.new)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.196932 -0.086109  0.005329  0.073336  0.220450
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.94620    0.26867  18.410  < 2e-16 ***
## disp        -0.07792    0.10152  -0.768  0.44919
## hp          -0.21299    0.08063  -2.642  0.01334 *
## wt          -0.47880    0.13993  -3.422  0.00193 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.1061 on 28 degrees of freedom
## Multiple R-squared:  0.8853, Adjusted R-squared:  0.873
## F-statistic: 72.01 on 3 and 28 DF,  p-value: 2.805e-13
```

```
#Part C:
AICbasemodel <- lm(mpg ~ disp+hp+drat+wt+qsec ,data=mtcars.new)
q4modelAIC <- step(AICbasemodel, scope = ~., steps=4)
```

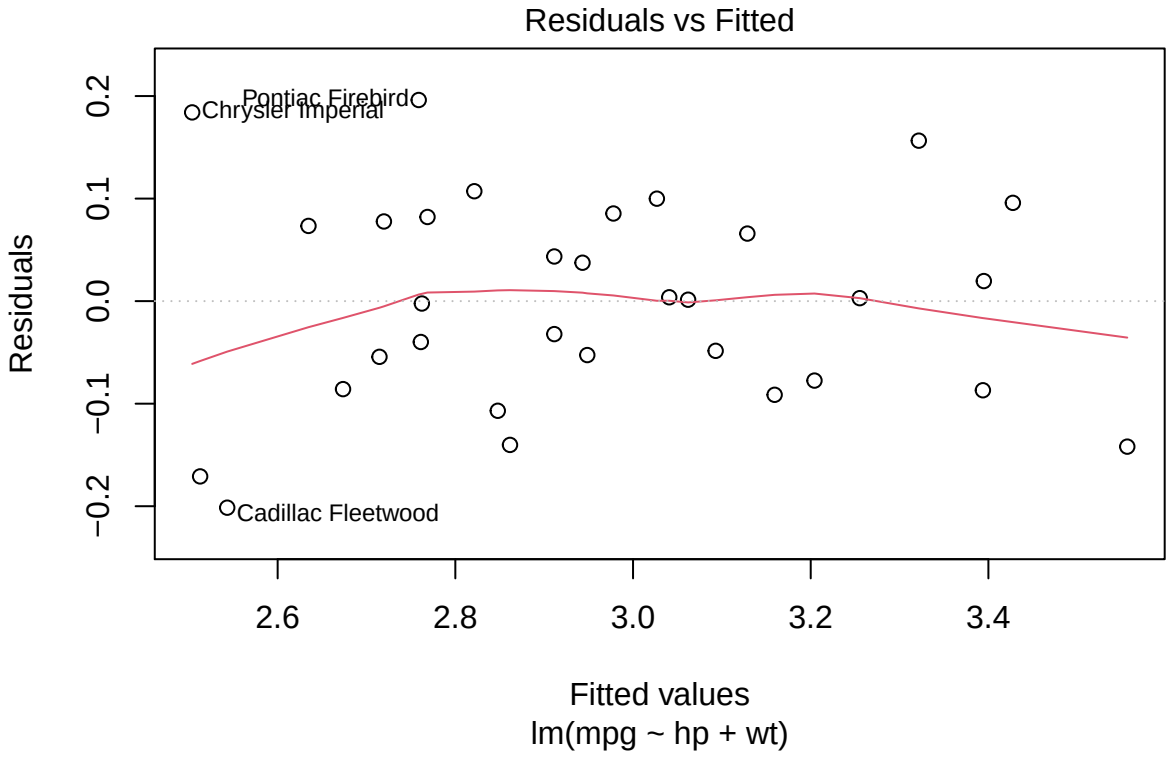
```
## Start:  AIC=-136.21
## mpg ~ disp + hp + drat + wt + qsec
##
##        Df Sum of Sq     RSS     AIC
## - drat  1  0.000402 0.31207 -138.17
## - disp  1  0.002104 0.31377 -138.00
## - qsec  1  0.003699 0.31536 -137.83
## <none>              0.31166 -136.21
## - hp    1  0.023697 0.33536 -135.87
## - wt    1  0.103076 0.41474 -129.07
##
## Step:  AIC=-138.17
## mpg ~ disp + hp + wt + qsec
##
##        Df Sum of Sq     RSS     AIC
## - qsec  1  0.003307 0.31537 -139.83
## - disp  1  0.004372 0.31644 -139.72
## <none>              0.31207 -138.17
## - hp    1  0.024123 0.33619 -137.79
## + drat  1  0.000402 0.31166 -136.21
## - wt    1  0.103779 0.41584 -130.98
##
## Step:  AIC=-139.83
## mpg ~ disp + hp + wt
##
##        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 ~ hp + wt
##
##        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
```

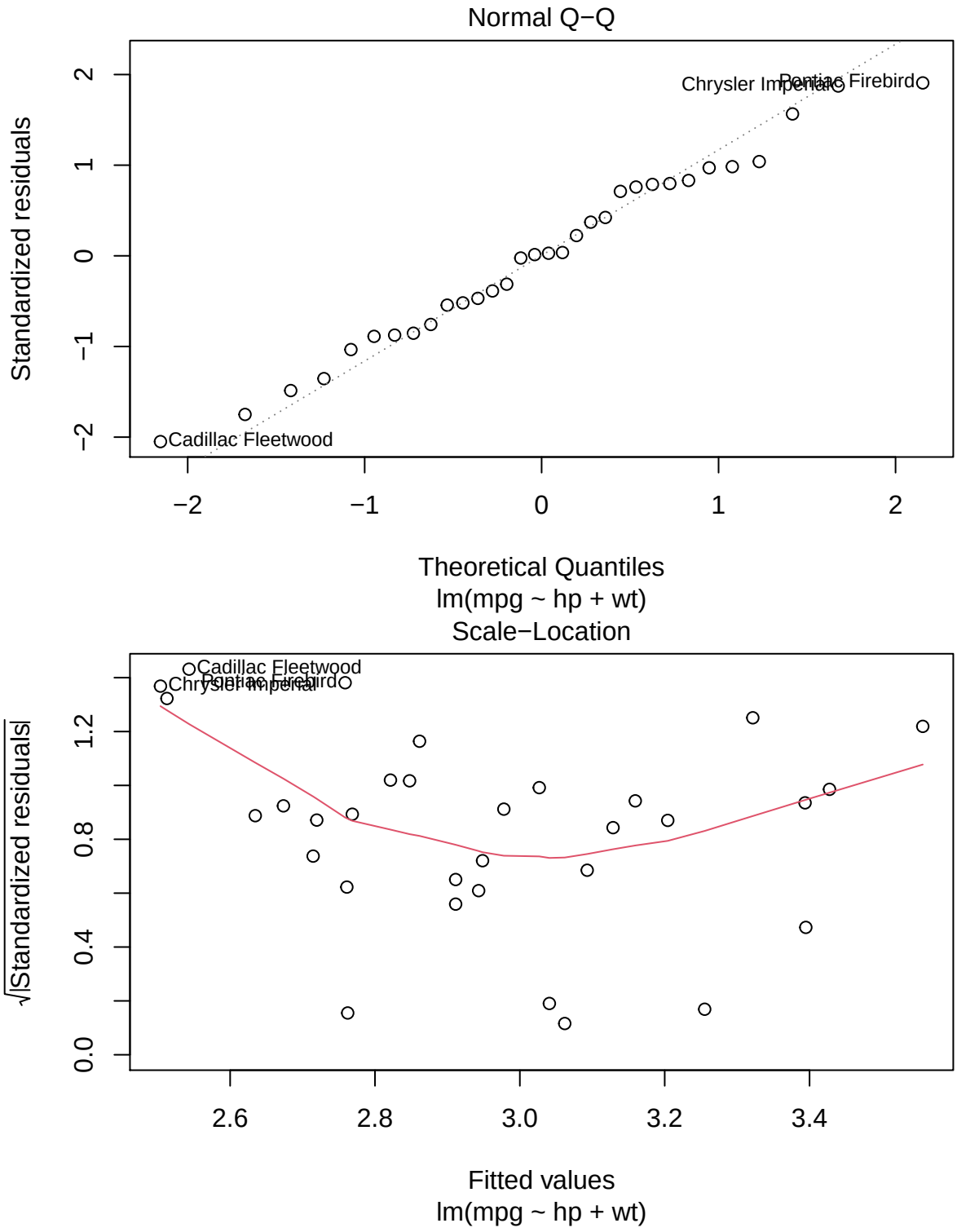
```
#Part D:
summary(q4modelAIC)
```

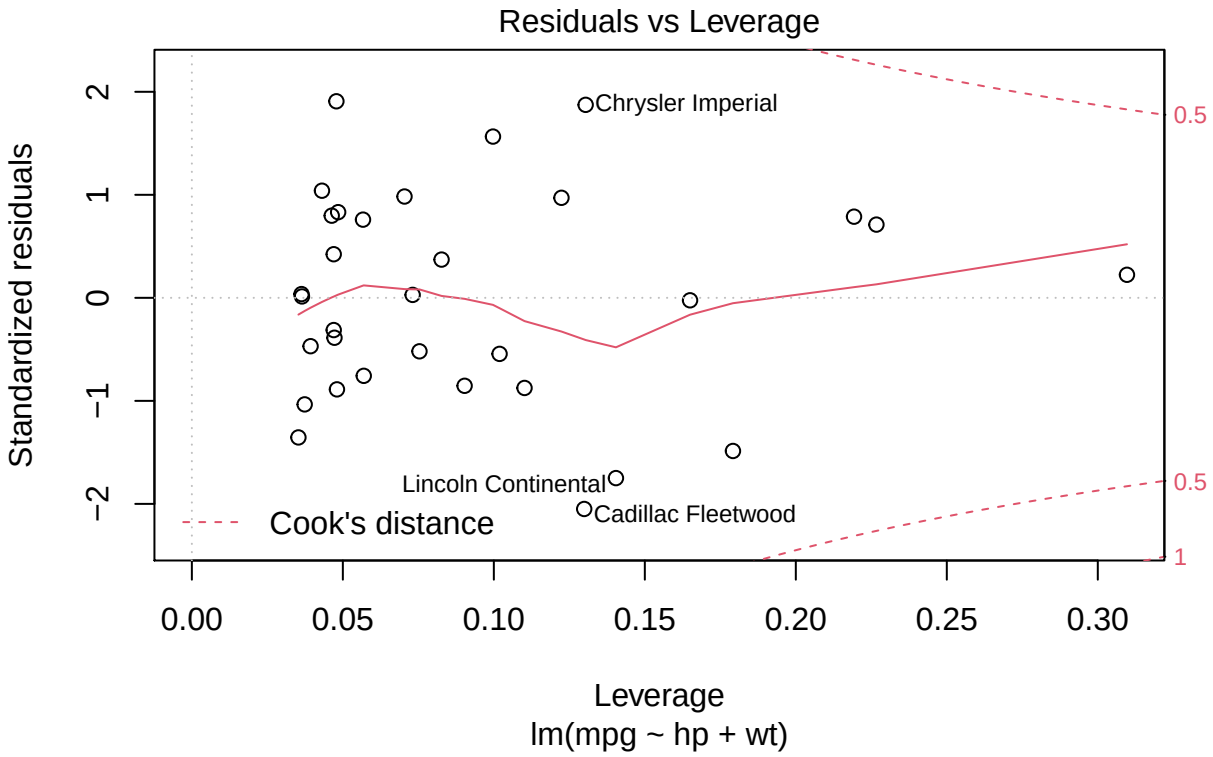
```
##
```

```
## Call:
## lm(formula = mpg ~ hp + wt, data = mtcars.new)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.201432 -0.079563  0.002145  0.078784  0.196150
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.83469    0.22440  21.545  < 2e-16 ***
## hp          -0.25532    0.05840  -4.372 0.000145 ***
## wt          -0.56228    0.08742  -6.432  4.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1054 on 29 degrees of freedom
## Multiple R-squared:  0.8829, Adjusted R-squared:  0.8748
## F-statistic: 109.3 on 2 and 29 DF,  p-value: 3.138e-14
```

```
#Part E:
plot(q4modelAIC)
```




Normal Q−Q
Standardized residuals
Chrysler Imperial
Pontiac Firebird
Cadillac Fleetwood
−2
−1
0
1
2
Theoretical Quantiles
lm(mpg ~ hp + wt)
Scale−Location
√|Standardized residuals|
Cadillac Fleetwood
Chrysler Imperial
Pontiac Firebird
0.0
0.4
0.8
1.2
2.6
2.8
3.0
3.2
3.4
Fitted values
lm(mpg ~ hp + wt)




```r
#Question 5:

#Part C:

f1 <-function(lambda){
  q2data <- data.matrix(read.csv(file = "Q2_Data.csv"))
  q2frame <- read.csv(file="Q2_Data.csv")
  y <- matrix(q2data[,1],7,1)
  x <- matrix(c(rep(1,7),q2data[,-1]),7,4)
  x <- scale(x[,-1],center=T,scale=T)
  y <- scale(y,center=T,scale=T)
  lambda = matrix(c(lambda,0,0,0,lambda,0,0,0,lambda),3,3)
  H <- x%*%solve(t(x)%*%x+lambda)%*%t(x)
  df <- sum(diag(H))
  b <- solve((t(x)%*%x)+lambda)%*%t(x)%*%y
  e <- (y-X%*%b)
  SSres <- sum(e^2)
  n <- dim(y)[1]
  gof <- n*log(SSres/n)+2*df
  return(gof)
}
```