STA312Midterm1Rickey

Rickey Huang

2/15/2022

Question 1

Part a

```
# Loading the dataset
library(faraway)
cheddar <- faraway::cheddar
# Plot the taste against lactic
plot(cheddar$Lactic,cheddar$taste,xlab = "Lactic", ylab = "taste")
```

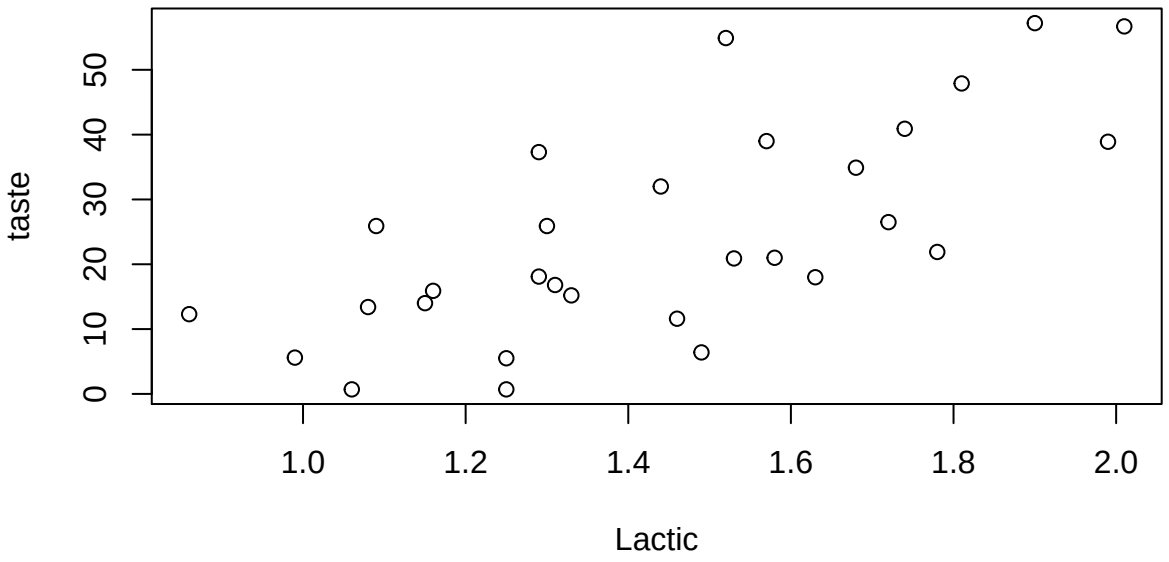

Figure 1: Plot of taste against Lactic

Part b

```
# First find the order of the variable Lactic
ord=order(cheddar$Lactic)
# Using this order to order both the taste and the Lactic
Lord=cheddar$Lactic[ord]
tord=cheddar$taste[ord]
# Creating the line of best fit
lm1 <- lm(tord~Lord)
# Show the model 1
lm1
```

```
##
## Call:
## lm(formula = tord ~ Lord)
##
## Coefficients:
## (Intercept)         Lord
##      -29.86        37.72
```

```
# Creating the parabola of best fit
lm2 <- lm(tord~Lord+I(Lord^2))
# Show the model 2
lm2
```

```
##
## Call:
## lm(formula = tord ~ Lord + I(Lord^2))
##
## Coefficients:
## (Intercept)         Lord    I(Lord^2)
##       22.22       -36.76        25.51
```

```
# Creating the cubic of best fit
lm3 <- lm(tord~Lord+I(Lord^2)+I(Lord^3))
# Show the model 3
lm3
```

```
##
## Call:
## lm(formula = tord ~ Lord + I(Lord^2) + I(Lord^3))
##
## Coefficients:
## (Intercept)         Lord    I(Lord^2)    I(Lord^3)
##       15.88       -22.79        15.58         2.28
```

From the R result, we can see that

$$lm1 : \widehat{taste} = -29.86 + 37.72 \cdot Lactic \tag{1}$$

$$lm2 : \widehat{taste} = 22.22 - 36.76 \cdot Lactic + 25.51 \cdot Lactic^2 \tag{2}$$

$$lm3 : \widehat{taste} = 15.88 - 22.79 \cdot Lactic + 15.58 \cdot Lactic^2 + 2.28 \cdot Lactic^3 \tag{3}$$

Part c

```
plot(tord~Lord, xlab = "Lactic", ylab = "taste")
abline(lm1,col = "blue")
points(Lord, lm2$fitted, type = "l", col = "red")
points(Lord, lm3$fitted, type = "l", col = "green")
```

Part d

```
# Find the model matrix X
X <- model.matrix(lm2)
# Define y
y <- tord
# Compute the betahat
betahat <- solve(t(X)%*%X)%*%t(X)%*%y
```

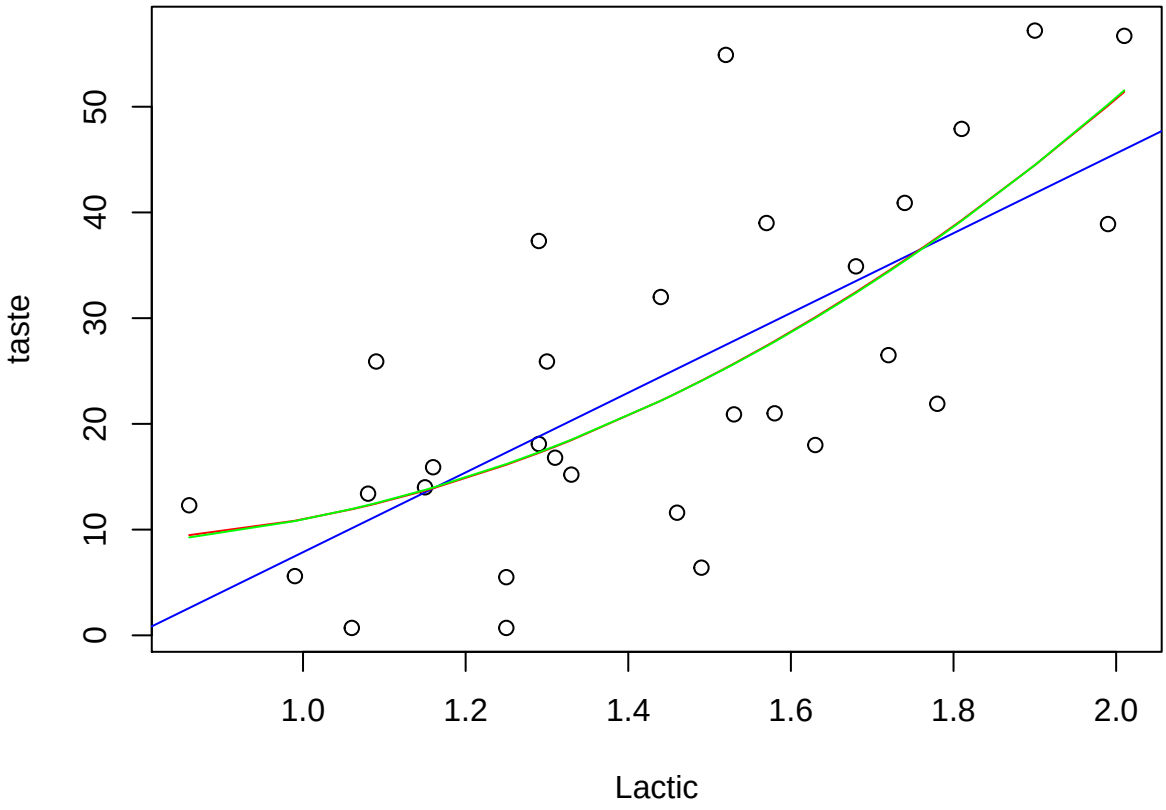

Figure 2: Plots of three models

```
# Return the betahat
betahat
```

```
##                  [,1]
## (Intercept)  22.22495
## Lord        -36.76068
## I(Lord^2)    25.51065
```

From the result, we know that $\hat{\beta} = \begin{bmatrix} 22.22495 \\ -36.76068 \\ 25.51065 \end{bmatrix}$

Part e

```
# Find the hat matrix H
H <- X%*%solve(t(X)%*%X)%*%t(X)
# Return the hat matrix
H
```

```
##              1           2           3           4           5
## 1   0.425017578  0.260488247  0.1856092842  0.1659784919  0.156456962
## 2   0.260488247  0.172325086  0.1315144342  0.1207107679  0.115451688
## 3   0.185609284  0.131514434  0.1059610914  0.0991197346  0.095775656
## 4   0.165978492  0.120710768  0.0991197346  0.0933087973  0.090462984
## 5   0.156456962  0.115451688  0.0957756559  0.0904629836  0.087858156
## 6   0.103441908  0.085895755  0.0767835798  0.0742232696  0.072950309
## 7   0.095291753  0.081302857  0.0737969665  0.0716558453  0.070585855
## 8   0.030756347  0.044249375  0.0492154377  0.0503386738  0.050851019
```

```
## 9   0.030756347  0.044249375  0.0492154377  0.0503386738  0.050851019
## 10  0.007167624  0.030255551  0.0396180422  0.0418983935  0.042972795
## 11  0.007167624  0.030255551  0.0396180422  0.0418983935  0.042972795
## 12  0.001760220  0.026995017  0.0373463595  0.0398877483  0.041089086
## 13 -0.003451273  0.023829652  0.0351257433  0.0379168730  0.039239717
## 14 -0.013286527  0.017784429  0.0308377103  0.0340944321  0.035643995
## 15 -0.053372803 -0.008659726  0.0109047819  0.0159145573  0.018322764
## 16 -0.058114376 -0.012230559  0.0079444773  0.0131261346  0.015619857
## 17 -0.063757406 -0.016873042  0.0038870189  0.0092417750  0.011823038
## 18 -0.067637239 -0.020659005  0.0002891589  0.0057153448  0.008335270
## 19 -0.068538695 -0.021730655 -0.0008079949  0.0046194079  0.007241359
## 20 -0.070185412 -0.025065567 -0.0046859451  0.0006333596  0.003209105
## 21 -0.070107314 -0.025661373 -0.0055277664 -0.0002637276  0.002286888
## 22 -0.066778163 -0.027212869 -0.0089708760 -0.0041526149 -0.001809107
## 23 -0.058551243 -0.026385143 -0.0111373237 -0.0070472540 -0.005046627
## 24 -0.048443314 -0.024009923 -0.0119512853 -0.0086471065 -0.007018542
## 25 -0.042213884 -0.022251300 -0.0120518673 -0.0092084132 -0.007798464
## 26 -0.027404096 -0.017592027 -0.0116402335 -0.0098537875 -0.008946242
## 27 -0.014239692 -0.013098300 -0.0107953101 -0.0099202339 -0.009446516
## 28  0.035832703  0.005522001 -0.0055029503 -0.0079719971 -0.009093031
## 29  0.101773870  0.031850980  0.0039257941 -0.0028023960 -0.005958085
## 30  0.118582481  0.038748722  0.0065828019 -0.0012161266 -0.004883702
##                6            7            8            9           10
## 1   0.1034419080  0.095291753  0.030756347  0.030756347  0.007167624
## 2   0.0858957548  0.081302857  0.044249375  0.044249375  0.030255551
## 3   0.0767835798  0.073796966  0.049215438  0.049215438  0.039618042
## 4   0.0742232696  0.071655845  0.050338674  0.050338674  0.041898394
## 5   0.0729503092  0.070585855  0.050851019  0.050851019  0.042972795
## 6   0.0654132728  0.064173888  0.053235264  0.053235264  0.048498362
## 7   0.0641738878  0.063106556  0.053517667  0.053517667  0.049265816
## 8   0.0532352640  0.053517667  0.054581104  0.054581104  0.054199675
## 9   0.0532352640  0.053517667  0.054581104  0.054581104  0.054199675
## 10  0.0484983619  0.049265816  0.054199675  0.054199675  0.055252412
## 11  0.0484983619  0.049265816  0.054199675  0.054199675  0.055252412
## 12  0.0473261275  0.048203803  0.054022195  0.054022195  0.055405972
## 13  0.0461586897  0.047142170  0.053811867  0.053811867  0.055515683
## 14  0.0438382035  0.045020043  0.053292665  0.053292665  0.055603556
## 15  0.0314184774  0.033375508  0.048088364  0.048088364  0.052951613
## 16  0.0292226997  0.031263258  0.046715093  0.046715093  0.051899397
## 17  0.0259650066  0.028097733  0.044408822  0.044408822  0.049992202
## 18  0.0227504819  0.024935627  0.041806911  0.041806911  0.047690360
## 19  0.0216885666  0.023882351  0.040873910  0.040873910  0.046835380
## 20  0.0174888703  0.019673047  0.036813418  0.036813418  0.042976966
## 21  0.0164509375  0.018621671  0.035716173  0.035716173  0.041902738
## 22  0.0113332203  0.013370488  0.029737217  0.029737217  0.035873857
## 23  0.0063354151  0.008128803  0.022937041  0.022937041  0.028748736
## 24  0.0024235076  0.003942292  0.016905623  0.016905623  0.022259348
## 25  0.0004963327  0.001851316  0.013692821  0.013692821  0.018751556
## 26 -0.0033004593 -0.002326077  0.006873033  0.006873033  0.011209778
## 27 -0.0060976902 -0.005455133  0.001413279  0.001413279  0.005093024
## 28 -0.0142303732 -0.014821787 -0.016739816 -0.016739816 -0.015625115
## 29 -0.0219745412 -0.024157670 -0.037553662 -0.037553662 -0.039895070
## 30 -0.0236427061 -0.026228132 -0.042540297 -0.042540297 -0.045770738
##               11           12           13           14           15
```

```
##  1   0.007167624  0.00176022 -0.003451273 -0.013286527 -0.0533728027
##  2   0.030255551  0.02699502  0.023829652  0.017784429 -0.0086597260
##  3   0.039618042  0.03734636  0.035125743  0.030837710  0.0109047819
##  4   0.041898394  0.03988775  0.037916873  0.034094432  0.0159145573
##  5   0.042972795  0.04108909  0.039239717  0.035643995  0.0183227644
##  6   0.048498362  0.04732613  0.046158690  0.043838203  0.0314184774
##  7   0.049265816  0.04820380  0.047142170  0.045020043  0.0333755081
##  8   0.054199675  0.05402220  0.053811867  0.053292665  0.0480883640
##  9   0.054199675  0.05402220  0.053811867  0.053292665  0.0480883640
##  10  0.055252412  0.05540597  0.055515683  0.055603556  0.0529516130
##  11  0.055252412  0.05540597  0.055515683  0.055603556  0.0529516130
##  12  0.055405972  0.05563632  0.055820436  0.056049949  0.0540062908
##  13  0.055515683  0.05582044  0.056076708  0.056443810  0.0549965148
##  14  0.055603556  0.05604995  0.056443810  0.057073937  0.0567836016
##  15  0.052951613  0.05400629  0.054996515  0.056783602  0.0620041338
##  16  0.051899397  0.05303363  0.054103123  0.056047899  0.0621154224
##  17  0.049992202  0.05122785  0.052399431  0.054550355  0.0617989521
##  18  0.047690360  0.04900594  0.050259415  0.052580024  0.0609023977
##  19  0.046835380  0.04817283  0.049449115  0.051818183  0.0604746387
##  20  0.042976966  0.04437801  0.045723110  0.048245501  0.0581190650
##  21  0.041902738  0.04331371  0.044670407  0.047221001  0.0573690372
##  22  0.035873857  0.03729865  0.038679686  0.041310520  0.0526520913
##  23  0.028748736  0.03012765  0.031476951  0.034086743  0.0463238010
##  24  0.022259348  0.02355857  0.024842113  0.027362146  0.0401010009
##  25  0.018751556  0.01999661  0.021233811  0.023684656  0.0366028782
##  26  0.011209778  0.01231783  0.013435440  0.015699294  0.0288331875
##  27  0.005093024  0.00607326  0.007077616  0.009158687  0.0223291548
##  28 -0.015625115 -0.01515729 -0.014613805 -0.013299856 -0.0006634469
##  29 -0.039895070 -0.04013308 -0.040232151 -0.040013483 -0.0288768043
##  30 -0.045770738 -0.04619187 -0.046458402 -0.046527696 -0.0358554308
##                16           17            18            19           20
##  1  -0.058114376 -0.063757406 -0.0676372390 -0.0685386951 -0.0701854116
##  2  -0.012230559 -0.016873042 -0.0206590053 -0.0217306554 -0.0250655669
##  3   0.007944477  0.003887019  0.0002891589 -0.0008079949 -0.0046859451
##  4   0.013126135  0.009241775  0.0057153448  0.0046194079  0.0006333596
##  5   0.015619857  0.011823038  0.0083352704  0.0072413592  0.0032091045
##  6   0.029222700  0.025965007  0.0227504819  0.0216885666  0.0174888703
##  7   0.031263258  0.028097733  0.0249356268  0.0238823511  0.0196730473
##  8   0.046715093  0.044408822  0.0418069113  0.0408739102  0.0368134182
##  9   0.046715093  0.044408822  0.0418069113  0.0408739102  0.0368134182
##  10  0.051899397  0.049992202  0.0476903599  0.0468353802  0.0429769657
##  11  0.051899397  0.049992202  0.0476903599  0.0468353802  0.0429769657
##  12  0.053033629  0.051227855  0.0490059432  0.0481728309  0.0443780069
##  13  0.054103123  0.052399431  0.0502594149  0.0494491150  0.0457231098
##  14  0.056047899  0.054550355  0.0525800238  0.0518181831  0.0482455009
##  15  0.062115422  0.061798952  0.0609023977  0.0604746387  0.0581190650
##  16  0.062377019  0.062283882  0.0616081066  0.0612533728  0.0591870610
##  17  0.062283882  0.062530703  0.0622008334  0.0619627236  0.0603695178
##  18  0.061608107  0.062200833  0.0622345562  0.0621215741  0.0610485303
##  19  0.061253373  0.061962724  0.0621215741  0.0620521909  0.0611629912
##  20  0.059187061  0.060369518  0.0610485303  0.0611629912  0.0610614522
##  21  0.058508639  0.059811025  0.0606249905  0.0607877745  0.0608962217
##  22  0.054145464  0.056057410  0.0575756183  0.0579941908  0.0592309955
##  23  0.048163848  0.050701878  0.0529734576  0.0536714397  0.0561673124
```

```
## 24  0.042213278  0.045264073  0.0481737213  0.0491122383  0.0527094770
## 25  0.038849568  0.042160711  0.0454011840  0.0464656374  0.0506449297
## 26  0.031345294  0.035185066  0.0391107708  0.0404384354  0.0458445758
## 27  0.025037344  0.029280527  0.0337407897  0.0352757836  0.0416569585
## 28  0.002617661  0.008106772  0.0142768233  0.0164848267  0.0260734399
## 29 -0.025045771 -0.018257191 -0.0102181779 -0.0072606324  0.0059589212
## 30 -0.031905314 -0.024820693 -0.0163447384 -0.0132102347  0.0008737072
##               21           22           23           24           25
## 1  -0.0701073138 -0.066778163 -0.058551243 -0.048443314 -0.0422138842
## 2  -0.0256613727 -0.027212869 -0.026385143 -0.024009923 -0.0222512997
## 3  -0.0055277664 -0.008970876 -0.011137324 -0.011951285 -0.0120518673
## 4  -0.0002637276 -0.004152615 -0.007047254 -0.008647107 -0.0092084132
## 5   0.0022868884 -0.001809107 -0.005046627 -0.007018542 -0.0077984645
## 6   0.0164509375  0.011333220  0.006335415  0.002423508  0.0004963327
## 7   0.0186216711  0.013370488  0.008128803  0.003942292  0.0018513163
## 8   0.0357161733  0.029737217  0.022937041  0.016905623  0.0136928214
## 9   0.0357161733  0.029737217  0.022937041  0.016905623  0.0136928214
## 10  0.0419027382  0.035873857  0.028748736  0.022259348  0.0187515562
## 11  0.0419027382  0.035873857  0.028748736  0.022259348  0.0187515562
## 12  0.0433137071  0.037298646  0.030127646  0.023558572  0.0199966095
## 13  0.0446704072  0.038679686  0.031476951  0.024842113  0.0212338106
## 14  0.0472210005  0.041310520  0.034086743  0.027362146  0.0236846564
## 15  0.0573690372  0.052652091  0.046323801  0.040101001  0.0366028782
## 16  0.0585086389  0.054145464  0.048163848  0.042213278  0.0388495676
## 17  0.0598110248  0.056057410  0.050701878  0.045264073  0.0421607105
## 18  0.0606249905  0.057575618  0.052973458  0.048173721  0.0454011840
## 19  0.0607877745  0.057994191  0.053671440  0.049112238  0.0464656374
## 20  0.0608962217  0.059230995  0.056167312  0.052709477  0.0506449297
## 21  0.0607876613  0.059430825  0.056717267  0.053569579  0.0516701224
## 22  0.0594308253  0.059773747  0.059022955  0.057634848  0.0566783031
## 23  0.0567172667  0.059022955  0.060588504  0.061308043  0.0614901797
## 24  0.0535695794  0.057634848  0.061308043  0.063964307  0.0651983420
## 25  0.0516701224  0.056678303  0.061490180  0.065198342  0.0670053102
## 26  0.0472199813  0.054240231  0.061499186  0.067478216  0.0705250205
## 27  0.0433125520  0.051952318  0.061195082  0.069023451  0.0730823555
## 28  0.0286597430  0.042726158  0.058684070  0.072812280  0.0803303436
## 29  0.0096111525  0.029956366  0.053775009  0.075330792  0.0869423063
## 30  0.0047811744  0.026637401  0.052358445  0.075717949  0.0883252575
##               26           27           28           29           30
## 1  -0.027404096 -0.014239692  0.0358327029  0.101773870  0.1185824811
## 2  -0.017592027 -0.013098300  0.0055220015  0.031850980  0.0387487217
## 3  -0.011640233 -0.010795310 -0.0055029503  0.003925794  0.0065828019
## 4  -0.009853788 -0.009920234 -0.0079719971 -0.002802396 -0.0012161266
## 5  -0.008946242 -0.009446516 -0.0090930306 -0.005958085 -0.0048837018
## 6  -0.003300459 -0.006097690 -0.0142303732 -0.021974541 -0.0236427061
## 7  -0.002326077 -0.005455133 -0.0148217872 -0.024157670 -0.0262281324
## 8   0.006873033  0.001413279 -0.0167398162 -0.037553662 -0.0425402972
## 9   0.006873033  0.001413279 -0.0167398162 -0.037553662 -0.0425402972
## 10  0.011209778  0.005093024 -0.0156251154 -0.039895070 -0.0457707380
## 11  0.011209778  0.005093024 -0.0156251154 -0.039895070 -0.0457707380
## 12  0.012317835  0.006073260 -0.0151572903 -0.040133079 -0.0461918664
## 13  0.013435440  0.007077616 -0.0146138053 -0.040232151 -0.0464584022
## 14  0.015699294  0.009158687 -0.0132998556 -0.040013483 -0.0465276955
## 15  0.028833187  0.022329155 -0.0006634469 -0.028876804 -0.0358554308
```

```
## 16  0.031345294  0.025037344  0.0026176610 -0.025045771 -0.0319053139
## 17  0.035185066  0.029280527  0.0081067725 -0.018257191 -0.0248206933
## 18  0.039110771  0.033740790  0.0142768233 -0.010218178 -0.0163447384
## 19  0.040438435  0.035275784  0.0164848267 -0.007260632 -0.0132102347
## 20  0.045844576  0.041656958  0.0260734399  0.005958921  0.0008737072
## 21  0.047219981  0.043312552  0.0286597430  0.009611152  0.0047811744
## 22  0.054240231  0.051952318  0.0427261576  0.029956366  0.0266374010
## 23  0.061499186  0.061195082  0.0586840704  0.053775009  0.0523584452
## 24  0.067478216  0.069023451  0.0728122795  0.075330792  0.0757179492
## 25  0.070525021  0.073082356  0.0803303436  0.086942306  0.0883252575
## 26  0.076733207  0.081489603  0.0962743910  0.111832581  0.1153949863
## 27  0.081489603  0.088048297  0.1090268559  0.131959127  0.1373205063
## 28  0.096274391  0.109026856  0.1513698866  0.199841372  0.2114450723
## 29  0.111832581  0.131959127  0.1998413719  0.278977526  0.2980916471
## 30  0.115394986  0.137320506  0.2114450723  0.298091647  0.3190469613
```

The hat matrix is shown in the R result, which is a huge 30×30 matrix.

Part f

```
# Find the SSE
SSE <- t(y)%*%(diag(30)-H)%*%y
# Define n and p
n <- 30
p <- 3
# Find an estimate for the sigmasq
sigmasq <- SSE[1]/(n - p)
#Return the estimate for sigmasq
sigmasq
```

```
## [1] 136.4015
```

Under Gauss-Markov assumptions, we know that $E(SSE) = \sigma^2(n-p)$, then $\sigma^2 \approx \frac{SSE}{n-p} = 136.4015$.

Part g

```
# Find variance of betahat using matrices
varbetahat <- sigmasq*solve(t(X)%*%X)
# Show the result
varbetahat
```

```
##             (Intercept)      Lord  I(Lord^2)
## (Intercept)    2170.416 -3019.018  1008.8339
## Lord          -3019.018  4263.008 -1442.6485
## I(Lord^2)      1008.834 -1442.649   494.1271
```

From the result, $Var(\hat{\beta}) = \begin{bmatrix} 2170.416 & -3019.018 & 1008.8339 \\ -3019.018 & 4263.008 & -1442.6485 \\ 1008.834 & -1442.649 & 494.1271 \end{bmatrix}$

Part h

```
# Find the standard error for beta2
SEbeta2 <- sqrt(4246.008)
# Show the result
SEbeta2
```

```
## [1] 65.1614
```

From the code above, we can see that $SE_{\hat{\beta}_2} = 65.1614$.

Question 2

```
# Creating the dataset
x1 <- runif(100,0,10)
x2 <- runif(100,5,15)
# Creating two lists each of length 1000 with NA as entries
sebeta2hats <- vector(mode = "double", length = 1000)
beta2hats <- vector(mode = "double", length = 1000)
# Construct sebeta2hats and beta2hats
#Set seed in order to keep the randomized process unchanged
set.seed(2022)
for (i in seq(1,1000)){
  # Creating 1000 random y for each time of iteration
  y <- 2*x1-x2+rnorm(100,0,5)
  # Fitting model with each y
  modeli <- lm(y~x1+x2)
  # Storing the betahats into the vector
  beta2hats[i] <- modeli$coefficient[2]
  # Finding the covariance matrix for each iteration
  covi <- vcov(modeli)
  # Storing the standard error of the beta2hat into the vector
  sebeta2hats[i] <- sqrt(diag(covi)[2])
}
# Finding the standard deviation of the beta2hat list
sd(beta2hats)
```

```
## [1] 0.1834729
```

```
# Finding the mean of the sebeta2hats list
mean(sebeta2hats)
```

```
## [1] 0.1821
```

From the result, we can see that the standard deviation of $\hat{\beta}_2$ and mean of the $SE_{\hat{\beta}_2}$ are almost the same. the standard deviation of $\hat{\beta}_2$ measures how these 1000 $\hat{\beta}_2$ deviate from the real value of β_2, while $SE_{\hat{\beta}_2}$ measures the average standard deviation for $each\hat{\beta}_2$ of the 1000 models. Both of the measures are exploring how our estimated $\hat{\beta}_2$ is away from the real value, but they calculate the standard deviation for the whole population and for each individual separately. This makes the two values be almost the same, since, in a large population like 1000 case, the trend is the consistent.