LINEAR REGRESSION: Homework #5

Professor Jingchen Liu

Fan Yang
UNI: fy2232

Problem 1

(a)

Suppose $A = \begin{bmatrix} a_{11} & a_{1n} \\ ... & ...... \\ a_{k1} & a_{kn} \end{bmatrix}$ and $X = \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix}$.

Then the i^{th} element in $Y = AX$ is $a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n$

So

$$\begin{aligned} Cov(Y_i, Y_j) &= Cov(a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n, a_{j1}x_1 + a_{j2}x_2 + ... + a_{jn}x_n) \\ &= a_{i1}a_{j1}Cov(x_1, x_1) + a_{i1}a_{j2}Cov(x_1, x_2) + ... + a_{i1}a_{jn}Cov(x_1, x_n) \\ &\quad + a_{i2}a_{j1}Cov(x_2, x_1) + a_{i2}a_{j2}Cov(x_2, x_2) + ... + a_{i2}a_{jn}Cov(x_2, x_n) \\ &\quad + ... \\ &\quad + a_{in}a_{j1}Cov(x_n, x_1) + a_{in}a_{j2}Cov(x_n, x_2) + ... + a_{in}a_{jn}Cov(x_n, x_n) \end{aligned}$$

While

$$(A\Sigma)_{ij} = a_{i1}Cov(x_1, x_j) + a_{i2}Cov(x_2, x_j) + ... + a_{in}Cov(x_n, x_j)$$

$$\begin{aligned} (A\Sigma A^T)_{ij} &= a_{j1}[a_{i1}Cov(x_1, x_1) + a_{i2}Cov(x_2, x_1) + ... + a_{in}Cov(x_n, x_1)] \\ &\quad + a_{j2}[a_{i1}Cov(x_1, x_2) + a_{i2}Cov(x_2, x_2) + ... + a_{in}Cov(x_n, x_2)] \\ &\quad + ... \\ &\quad + a_{jn}[a_{i1}Cov(x_1, x_n) + a_{i2}Cov(x_2, x_n) + ... + a_{in}Cov(x_n, x_n)] \\ &= a_{i1}a_{j1}Cov(x_1, x_1) + a_{i1}a_{j2}Cov(x_1, x_2) + ... + a_{i1}a_{jn}Cov(x_1, x_n) \\ &\quad + a_{i2}a_{j1}Cov(x_2, x_1) + a_{i2}a_{j2}Cov(x_2, x_2) + ... + a_{i2}a_{jn}Cov(x_2, x_n) \\ &\quad + ... \\ &\quad + a_{in}a_{j1}Cov(x_n, x_1) + a_{in}a_{j2}Cov(x_n, x_2) + ... + a_{in}a_{jn}Cov(x_n, x_n) \end{aligned}$$

which means the $(i, j)^{\text{th}}$ element in covariance matrix of Y equals to that of $A\Sigma A^T$

So, covariance matrix of Y = $A\Sigma A^T$.

Problem 2

$$t^* = \frac{\hat{\beta}_1}{s(\hat{\beta_1})} = \frac{\hat{\beta}_1/\sigma(\hat{\beta}_1)}{s(\hat{\beta}_1)/\sigma(\hat{\beta}_1)}$$

As we know $\hat{\beta}_1/\sigma(\hat{\beta}_1)$ follows *N* (0, 1) distribution;

and $s(\hat{\beta}_1)/\sigma(\hat{\beta}_1)$ follows $\sqrt{\dfrac{\chi^2(n-2)}{n-2}}$.

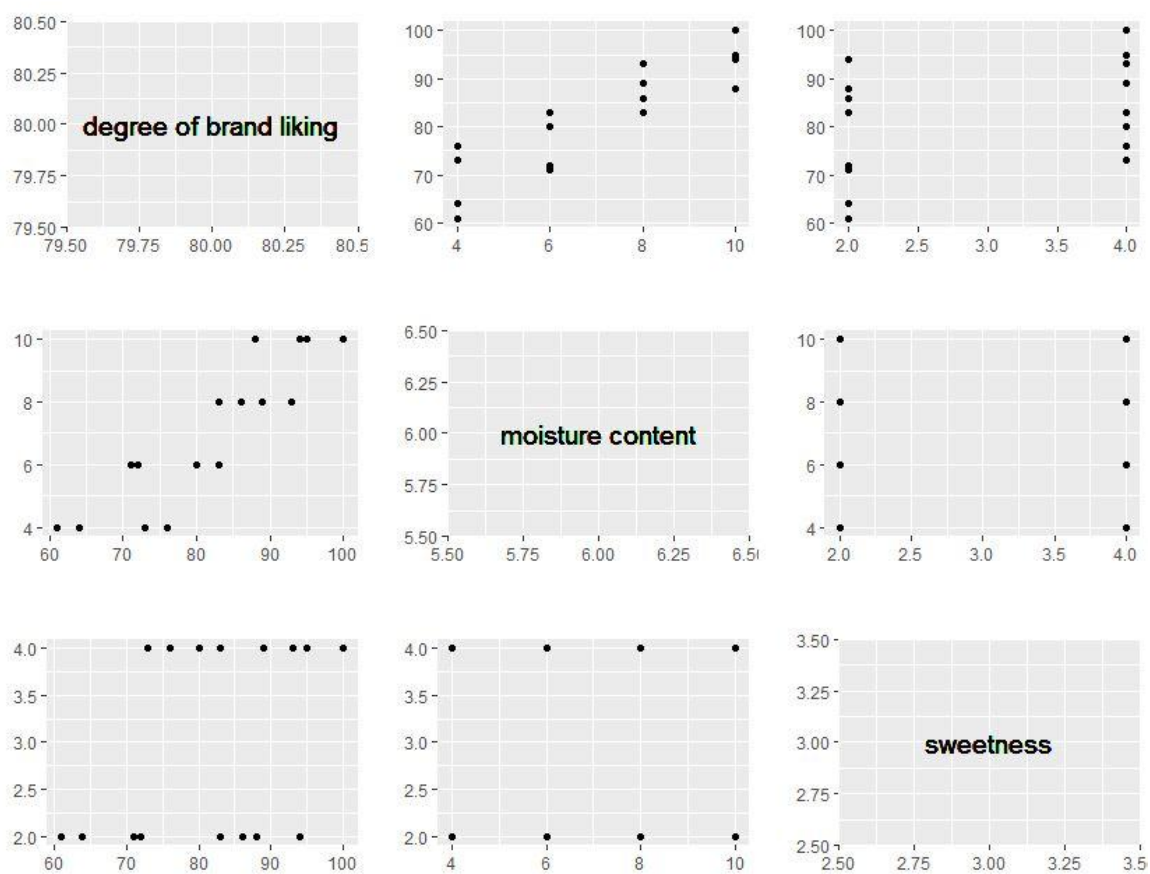
When we use $(t^*)^2$, $\left(\dfrac{\hat{\beta}_1}{\sigma(\hat{\beta}_1)}\right)^2$ follows $\dfrac{\chi^2(1)}{1}$ distribution;

and $\left(\dfrac{s(\hat{\beta}_1}{\sigma(\hat{\beta}_1)}\right)^2$ follows $\dfrac{\chi^2(n-2)}{n-2}$.

Therefore, T^2 follows a *F* distribution *F* (1, n-2). So *t*-test and *F*-test are equivalent in the sense that the $T^2 = F$.

Problem 3 (6.5)

(a)



```
> cor(d5)
                      degree_of_brand_liking moisture_content sweetness
degree_of_brand_liking             1.0000000        0.8923929 0.3945807
moisture_content                   0.8923929        1.0000000 0.0000000
sweetness                          0.3945807        0.0000000 1.0000000
```

From the graphs above, we conclude that x_1 and x_2 are uncorrelated. x_1 and y represent a likely linear relation while x_2 seems to have weak relation with y.

(b)

```
> model5 <- lm(V1~V2+V3, data=d5)
> model5
Call:
lm(formula = V1 ~ V2 + V3, data = d5)
Coefficients:
(Intercept)           V2           V3
     37.650        4.425        4.375
```

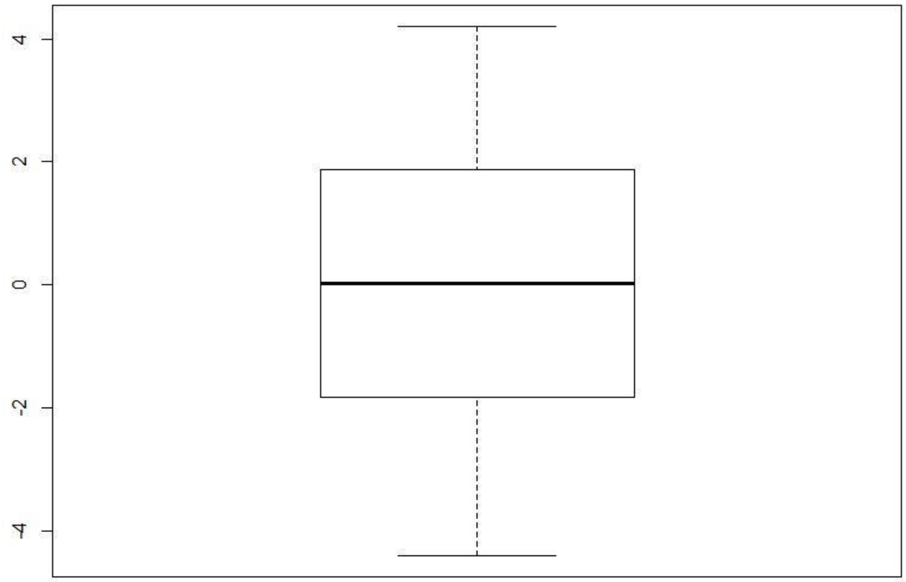
$$Y = 37.650 + 4.425 \times X_1 + 4.375 \times X_2$$

(c)

```
> model5$residuals
    1     2     3     4     5     6     7     8     9
-0.10  0.15 -3.10  3.15 -0.95 -1.70 -1.95  1.30  1.20
   10    11    12    13    14    15    16
-1.55  4.20  2.45 -2.65 -4.40  3.35  0.60
```

We can see the boxplot of residuals below. The median of residuals lays nearly to the mean of residuals which is 0, and the plot shows a strong symmetric property. Most of the values lay between -2 and 2 which is a small variation.

Therefore, our models seems to be a good fit from the point of residuals.



(f)

Hypothesis:

$$H_0: \ E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 x_2$$
$$H_1: \ E(Y) \neq \beta_0 + \beta_1 X_1 + \beta_2 x_2$$

if $F^* \leq F(0.99, c-p, n-c)$, then conclude H_0

if $F^* > F(0.99, c-p, n-c)$, then conclude H_1

```
> anova(lm(d5$V1~d5$V2+d5$V3), lm(d5$V1~factor(d5$V2)*factor(d5$V3)))
Analysis of Variance Table
Model 1: d5$V1 ~ d5$V2 + d5$V3
Model 2: d5$V1 ~ factor(d5$V2) * factor(d5$V3)
  Res.Df  RSS Df Sum of Sq     F Pr(>F)
1     13 94.3
2      8 57.0  5      37.3 1.047  0.453
```

Now $F^* = 1.047 <\leq F(0.99, 5, 8) = 6.632$, so we conclude H_0 that the regression function is linear.

Problem 4 (6.7)

(a)

```
> SSE=sum((d5$V1-model5$fitted.values)^2)
> SST=sum((d5$V1-mean(d5$V1))^2)
> 1-SSE/SST
[1] 0.952059
```

We get the $R^2 = 0.952059$, which means there are about 95.21% of total variation can be explained by our model.

(b)

```
> cor(d5$V1,model5$fitted.values)
[1] 0.9757351
```

We get the coefficient of simple determination $R^2 = 0.9757351$, and this is different from the R^2 in part (a).

Problem 5 (6.8)

(a)

```
> X = cbind(rep(1,nrow(d5)), d5$V2, d5$V3)
> Xh = c(1,5,4)
> MSE = sum((model5$residuals)^2) / model5$df.residual
> s = sqrt(MSE*(t(Xh)%*%solve((t(X)%*%X))%*%Xh))
> Ynew = predict(model5, data.frame(V2=5,V3=4),level=0.99)
> Ynew+qt(0.995,model5$df.residual)*s
         [,1]
[1,] 80.66889
> Ynew-qt(0.995,model5$df.residual)*s
         [,1]
[1,] 73.88111
```

The interval estimate is $[73.88111,\ 80.66889]$

(b)

```
> s_pred = sqrt(MSE*(1+t(Xh)%*%solve((t(X)%*%X))%*%Xh))
> Ynew+qt(0.995,model5$df.residual)*s_pred
         [,1]
[1,] 86.06923
> Ynew-qt(0.995,model5$df.residual)*s_pred
         [,1]
[1,] 68.48077
```

The prediction interval is $[68.48077,\ 86.06923]$.

Problem 6 (6.25)

Suppose the original data is

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}_{n\times 1} \qquad X = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ & & \dots & \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{bmatrix}_{n\times 4} \qquad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}_{4\times 1}$$

Since we know that $\beta_2 = 4$, we can make the following transform:

$$\widetilde{Y} = \begin{bmatrix} y_1 - 4x_{12} \\ y_2 - 4x_{22} \\ \dots \\ y_n - 4x_{n2} \end{bmatrix}_{n\times 1} \qquad \widetilde{X} = \begin{bmatrix} 1 & x_{11} & x_{13} \\ 1 & x_{21} & x_{23} \\ & \dots & \\ 1 & x_{n1} & x_{n3} \end{bmatrix}_{n\times 3} \qquad \widetilde{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_3 \end{bmatrix}_{3\times 1}$$

Then we only need to fit the model $\widetilde{Y} = \widetilde{X}\widetilde{\beta} + \epsilon$.