Homework Chapter 5

Jinhong Du 15338039

5.5 Consumer finance.

The data below show, for a consumer finance company operating in six cities, the number of competing loan companies operating in the city (X) and the number per thousand of the company's loans made in that city that are currently delinquent (Y):

i	1	2	3	4	5	6
X_i	4	1	2	3	3	4
Y_i	16	5	10	15	13	22

Assume that first-order regression model (2.1) is applicable. Using matrix methods, find

(1) $Y'Y$.

```
data1 <- read.table("CH05PR05.txt",head=FALSE,col.names = c('Y','X'))
Y <- matrix(data1$Y)
n = length(Y)
X <- cbind(rep(1,n),data1$X)
crossprod(Y)
```

```
##      [,1]
## [1,] 1259
```

(2) $X'X$.

```
crossprod(X)
```

```
##      [,1] [,2]
## [1,]    6   17
## [2,]   17   55
```

(3) $X'Y$.

```
crossprod(X,Y)
```

```
##      [,1]
## [1,]   81
## [2,]  261
```

5.18 Consider the following functions of the random variables Y_1, Y_2, Y_3 and Y_4:

$$W_1 = \frac{1}{4}(Y_1 + Y_2 + Y_3 + Y_4)$$

$$W_2 = \frac{1}{2}(Y_1 + Y_2) - \frac{1}{2}(Y_3 - Y_4)$$

(a) State the above in matrix notation.

$$\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix}$$

i.e.

$$W = AY$$

(b) Find the expectation of the random vector W.

$$\begin{aligned} \mathbb{E}W &= \mathbb{E}(AY) \\ &= A\mathbb{E}Y \\ &= A \begin{bmatrix} \mathbb{E}Y_1 \\ \mathbb{E}Y_2 \\ \mathbb{E}Y_3 \\ \mathbb{E}Y_4 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4}(\mathbb{E}Y_1 + \mathbb{E}Y_2 + \mathbb{E}Y_3 + \mathbb{E}Y_4) \\ \frac{1}{2}(\mathbb{E}Y_1 + \mathbb{E}Y_2 - \mathbb{E}Y_3 - \mathbb{E}Y_4) \end{bmatrix} \end{aligned}$$

(c) Find the variance-covariance of W.

$$\begin{aligned} \mathbb{D}(W) &= \mathbb{D}(AY) \\ &= A\mathbb{D}(Y)A \\ &= A \begin{bmatrix} \mathbb{D}Y_1 & Cov(Y_1, Y_2) & Cov(Y_1, Y_3) & Cov(Y_1, Y_4) \\ Cov(Y_2, Y_1) & \mathbb{D}Y_2 & Cov(Y_2, Y_3) & Cov(Y_2, Y_4) \\ Cov(Y_3, Y_1) & Cov(Y_3, Y_2) & \mathbb{D}(Y_3) & Cov(Y_3, Y_4) \\ Cov(Y_4, Y_1) & Cov(Y_4, Y_2) & Cov(Y_4, Y_3) & \mathbb{D}(Y_4) \end{bmatrix} A^T \\ &= A \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix} A^T \\ &= \begin{bmatrix} \sigma^2\{W_1\} & \sigma\{W_1, W_2\} \\ \sigma\{W_2, W_1\} & \sigma^2\{W_2\} \end{bmatrix} \end{aligned}$$

where

$$\begin{aligned} \sigma^2\{W_1\} &= \frac{1}{16}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 + 2\sigma_{12} + 2\sigma_{13} + 2\sigma_{14} + 2\sigma_{23} + 2\sigma_{24} + 2\sigma_{34}) \\ \sigma^2\{W_2\} &= \frac{1}{4}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 + 2\sigma_{12} - 2\sigma_{13} - 2\sigma_{14} - 2\sigma_{23} - 2\sigma_{24} + 2\sigma_{34}) \\ \sigma\{W_1, W_2\} = \sigma\{W_2, W_1\} &= \frac{1}{8}(\sigma_1^2 + \sigma_2^2 - \sigma_3^2 - \sigma_4^2 + 2\sigma_{12} - 2\sigma_{34}) \end{aligned}$$

5.24 Refer to Consumer finance Problems 5.5 and 5.13.

(a) Using matrix methods, obtain the following:

(1) vector of estimated regression coefficients,
(2) vector of residuals,
(3) SSR,
(4) SSE,
(5) estimated variance-covariance matrix of b,
(6) point estimate of $\{Y_h\}$ when $X_h = 4$,
(7) $s^2\{pred\}$ when $X_h = 4$.

```
fit <- lm('Y~X',data1)
b <- as.matrix(fit$coefficients)
res <- as.matrix(fit$residuals)
J = matrix(rep(1,n*n),nrow=n,ncol=n)
SSR <- t(b)%*%crossprod(X,Y) - t(Y)%*%J%*%Y/n
SSE <- crossprod(Y) - t(b)%*%crossprod(X,Y)
MSE <- SSE/fit$df.residual
s2b <- solve(crossprod(X)) * MSE[1,1]
Xh <- matrix(c(1,4),nrow = 2,ncol = 1)
EYh <- t(b) %*% Xh
s2_pred <- (1+t(Xh)%*%solve(crossprod(X))%*%Xh)*MSE[1,1]
print("Regression coefficients matrix is ")
```

```
## [1] "Regression coefficients matrix is "
```

```
print(b)
```

```
##                  [,1]
## (Intercept) 0.4390244
## X           4.6097561
```

```
print("Residual is ")
```

```
## [1] "Residual is "
```

```
print(res)
```

```
##          [,1]
## 1 -2.87804878
## 2 -0.04878049
## 3  0.34146341
## 4  0.73170732
## 5 -1.26829268
## 6  3.12195122
```

```
print(sprintf("SSR is %f",SSR))
```

```
## [1] "SSR is 145.207317"
```

```
print(sprintf("SSE is %f",SSE))
```

```
## [1] "SSE is 20.292683"
```

```
print("estimated variance-covariance matrix of b is")
```

```
## [1] "estimated variance-covariance matrix of b is"
```

```
print(s2b)
```

```
##           [,1]       [,2]
## [1,]  6.805473 -2.1035098
## [2,] -2.103510  0.7424152
```

```
print(sprintf("point estimate of E{Yh} when Xh=4 is %f",EYh))
```

```
## [1] "point estimate of E{Yh} when Xh=4 is 18.878049"
```

```
print(sprintf("s2{pred} when Xh=4 is %f",s2_pred))
```

```
## [1] "s2{pred} when Xh=4 is 6.929209"
```

(b) From your estimated variance—covariance matrix in part (a5), obtain the following:

(1) $s\{b_0, b_1\}$;
(2) $s^2\{b_0\}$;
(3) $s\{b_1\}$.

```
print(sprintf("s{b0,b1} is %f",s2b[1,2]))
```

```
## [1] "s{b0,b1} is -2.103510"
```

```
print(sprintf("s2{b0} is %f",s2b[1,1]))
```

```
## [1] "s2{b0} is 6.805473"
```

```
print(sprintf("s{b1} is %f",sqrt(s2b[2,2])))
```

```
## [1] "s{b1} is 0.861635"
```

(c) Find the hat matrix H.

```
H <- X%*%solve(crossprod(X))%*%t(X)
print("The hat matrix H is ")
```

```
## [1] "The hat matrix H is "
```

```
print(H)
```

```
##             [,1]       [,2]       [,3]      [,4]      [,5]        [,6]
## [1,]  0.36585366 -0.1463415 0.02439024 0.1951220 0.1951220  0.36585366
## [2,] -0.14634146  0.6585366 0.39024390 0.1219512 0.1219512 -0.14634146
## [3,]  0.02439024  0.3902439 0.26829268 0.1463415 0.1463415  0.02439024
## [4,]  0.19512195  0.1219512 0.14634146 0.1707317 0.1707317  0.19512195
## [5,]  0.19512195  0.1219512 0.14634146 0.1707317 0.1707317  0.19512195
## [6,]  0.36585366 -0.1463415 0.02439024 0.1951220 0.1951220  0.36585366
```

(d) Find $s^2\{e\}$.

```
s2e <- (diag(n)-H)*MSE[1,1]
print("s2{e} is ")
```

```
## [1] "s2{e} is "
```

```
print(s2e)
```

```
##             [,1]       [,2]       [,3]       [,4]       [,5]       [,6]
## [1,]  3.2171327  0.7424152 -0.1237359 -0.9898870 -0.9898870 -1.8560381
## [2,]  0.7424152  1.7323022 -1.9797739 -0.6186794 -0.6186794  0.7424152
## [3,] -0.1237359 -1.9797739  3.7120761 -0.7424152 -0.7424152 -0.1237359
## [4,] -0.9898870 -0.6186794 -0.7424152  4.2070196 -0.8661511 -0.9898870
## [5,] -0.9898870 -0.6186794 -0.7424152 -0.8661511  4.2070196 -0.9898870
## [6,] -1.8560381  0.7424152 -0.1237359 -0.9898870 -0.9898870  3.2171327
```

5.28 Consider model (4.10) for regression through the origin and the estimator b_1 given in (4.14). Obtain (4.14) by utilizing (5.60) with X suitably defined.

Let

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

Then

$$\begin{aligned} b_1 &= (X^T X)^{-1} X^T Y \\ &= \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2} \end{aligned}$$