Assignment 7 Solutions

1. Please do the following exercises from 'Applied Multivariate Statistical Analysis' by Johnson and Wichern. Use matrix calculations in R or by hand.

   - 7.1

solution: This question is a univariate regression example. We need to find the following:

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\boldsymbol{z}^T\boldsymbol{z})^{-1}\boldsymbol{z}^T\boldsymbol{y},\\ \hat{\boldsymbol{y}} &= \boldsymbol{z}\hat{\boldsymbol{\beta}},\\ \hat{\boldsymbol{\epsilon}} &= \boldsymbol{y}-\hat{\boldsymbol{y}}.\end{aligned}$$

```
> y <- matrix(c(15,9,3,25,7,13), 6,1)
> z <- cbind(rep(1,6), c(10, 5, 7, 19, 11, 8))
> n <- nrow(y)
> p <- ncol(y)

> ## calculate beta.hat
> beta.hat <- solve(t(z)%*%z)%*%t(z)%*%y
> beta.hat
           [,1]
[1,] -0.6666667
[2,]  1.2666667

> ## calculate y.hat (the fitted values)
> y.hat <- z%*%beta.hat
> y.hat
          [,1]
[1,] 12.000000
[2,]  5.666667
[3,]  8.200000
[4,] 23.400000
[5,] 13.266667
[6,]  9.466667
```

```
> ## calculate epsilon.hat (the residuals)
> eps.hat <- y-y.hat
> eps.hat
          [,1]
[1,]  3.000000
[2,]  3.333333
[3,] -5.200000
[4,]  1.600000
[5,] -6.266667
[6,]  3.533333

> ## calculate the cross products of the residuals
> eps.hat.eps.hat <- t(eps.hat)%*%eps.hat
> eps.hat.eps.hat
         [,1]
[1,] 101.4667
```

- Using the data in 7.9 answer the same questions as in 7.1. Thus find the least squares solutions for $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{\epsilon}}$, and $\hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}}$.

solution: This question is a multivariate regression example. The only difference between this and the univariate is the dimension of the matrices change (except for $\boldsymbol{z}$). For this data $m = 2$, or in the usual notation $p = 2$.

$$
\begin{aligned}
\boldsymbol{y} &= [\boldsymbol{y}_{(1)}|\boldsymbol{y}_{(2)}], \\
\hat{\boldsymbol{\beta}} &= (\boldsymbol{z}^T\boldsymbol{z})^{-1}\boldsymbol{z}^T\boldsymbol{y} \\
&= (\boldsymbol{z}^T\boldsymbol{z})^{-1}\boldsymbol{z}^T[\boldsymbol{y}_{(1)}|\boldsymbol{y}_{(2)}] \\
&= [\hat{\boldsymbol{\beta}}_{(1)}|\hat{\boldsymbol{\beta}}_{(2)}], \\
\hat{\boldsymbol{y}} &= \boldsymbol{z}\hat{\boldsymbol{\beta}} \\
&= [\hat{\boldsymbol{y}}_{(1)}|\hat{\boldsymbol{y}}_{(2)}], \\
\hat{\boldsymbol{\epsilon}} &= \boldsymbol{y} - \hat{\boldsymbol{y}} \\
&= [\hat{\boldsymbol{\epsilon}}_{(1)}|\hat{\boldsymbol{\epsilon}}_{(2)}].
\end{aligned}
$$

```
> y <- matrix(c(5,-3,3,-1,4,-1,2,2,1,3), 5,2, byrow=T)
> z <- cbind(rep(1,5), c(-2,-1,0,1,2))
> n <- nrow(y)
> p <- ncol(y)

> ## calculate beta.hat
> beta.hat <- solve(t(z)%*%z)%*%t(z)%*%y
> beta.hat
     [,1]         [,2]
[1,]  3.0 1.110223e-16
[2,] -0.9 1.500000e+00
```

```
> ## calculate y.hat (the fitted values)
> y.hat <- z%*%beta.hat
> y.hat
     [,1]          [,2]
[1,]  4.8 -3.000000e+00
[2,]  3.9 -1.500000e+00
[3,]  3.0  1.110223e-16
[4,]  2.1  1.500000e+00
[5,]  1.2  3.000000e+00

> ## calculate epsilon.hat (the residuals)
> eps.hat <- y-y.hat
> eps.hat
     [,1]          [,2]
[1,]  0.2  4.440892e-16
[2,] -0.9  5.000000e-01
[3,]  1.0 -1.000000e+00
[4,] -0.1  5.000000e-01
[5,] -0.2 -4.440892e-16

> ## calculate the cross products of the residuals
> eps.hat.eps.hat <- t(eps.hat)%*%eps.hat
> eps.hat.eps.hat
     [,1] [,2]
[1,]  1.9 -1.5
[2,] -1.5  1.5
```