**LINEAR REGRESSION**: Homework 

*Professor Jingchen Liu*

Fan Yang

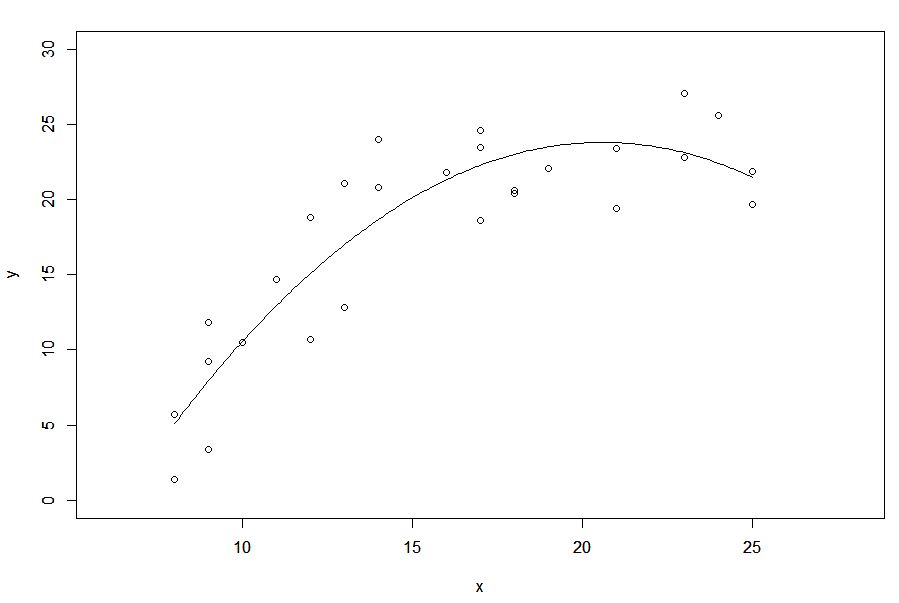
UNI: fy2232

# Problem 1 (8.6)

## (a)



The regression function is 



The quadratic regression function appears to be a good fit here.



The  is 0.8143372

## (b)

Hypothesis:





if , then conclude 

if , then conclude 



Since , we can conclude  that not all of the coefficients are 0, which means there is a regression relation. And the P-value is 1.67764e-09.

## (c)

When  is to be estimated for g levels  with family confidence coefficient , the Bonferroni confidence limits are:



where



and g is the number of confidence intervals in the family.

so 





which is



## (d)

The 99 percent prediction interval for  is:



where



In this situation, ; 



So the prediction interval is 

## (e)

Hypothesis:



if , then conclude 

if , then conclude 





Because , p-value, then we conclude . The quadratic term cannot be dropped from the model.

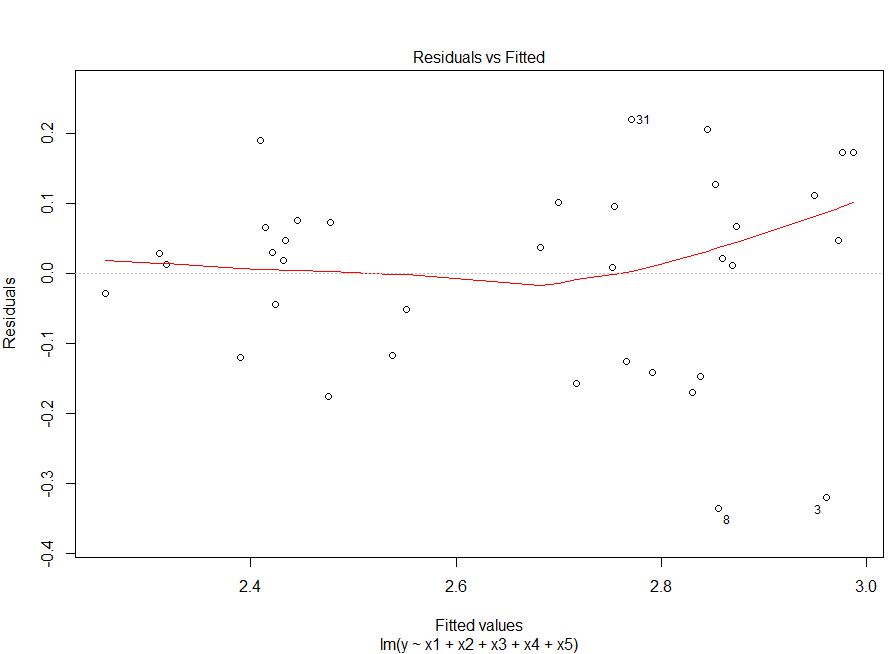
## (f)

The regression function is 

# Problem 2 (8.42)

## (a)







The residuals are some small along the fitted value. The first-order model appear to fit the data well.

## (b)

Full model:





Hypothesis:



if , then conclude 

if , then conclude 





Because , with p-value , then we conclude . All quadratic and interaction terms can be dropped from the regression model.

## (c)

Full model:



Reduced model:



Hypothesis:



if , then conclude 

if , then conclude 





Because , with p-value , then we conclude . Advertising index  and year  can be dropped from the model.

# Problem 3 (8.43)

In order to make prediction, we separate the data to training set and test set. Randomly choose 100 sample as the test set and the rest data as the training set. Then our model will be build on training set and be predicted on test set.



First use the first order model,





As we can see the result summary, the p-value for  (academic year) is large and not significant. So we decide to remove . Let’s test whether  can be removed from the model.

Hypothesis:



if , then conclude 

if , then conclude 



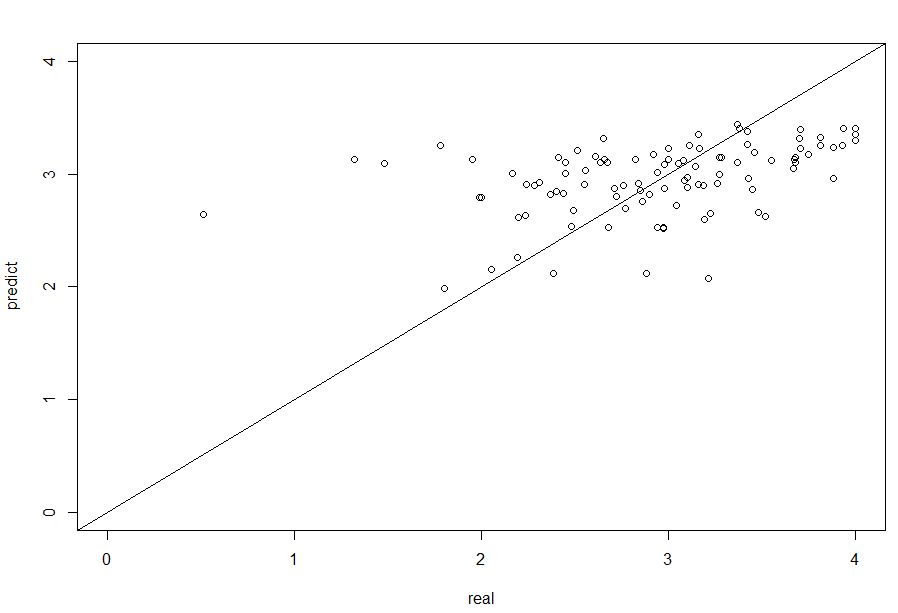


Because , with p-value , then we conclude . .

Therefore, we remove  (academic year).

Now we make the prediction on test set:





This plot is predicting value against real value. We can see the model fit the data well to some degree. Now we try to find a more appropriate model. Let’s introduce quadratic and interaction terms into the regression model.



Hypothesis:





Because , with p-value , then we conclude .

According to the summary of the model:



We notice that only  and  are significant, so we test the following reduced model:



Hypothesis:





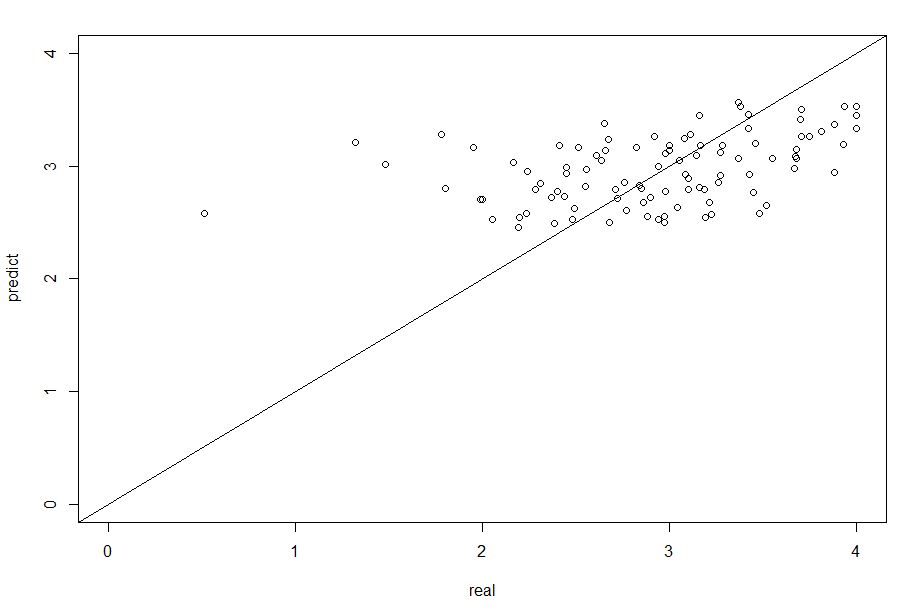
Because , with p-value , then we conclude .

Now it comes our final model



Let’s make prediction on test set:





This is much better than the first order model fit. So we conclude that



is our final model and

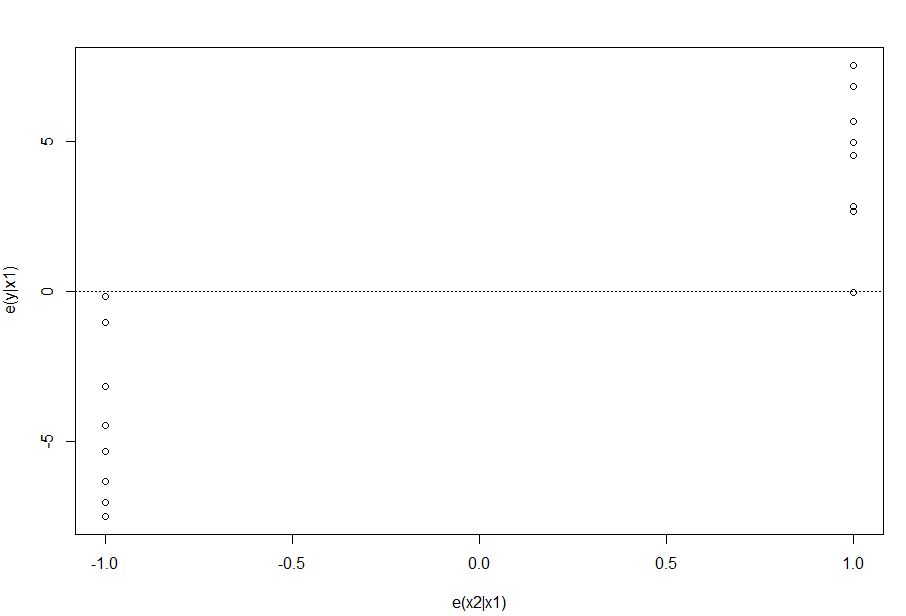


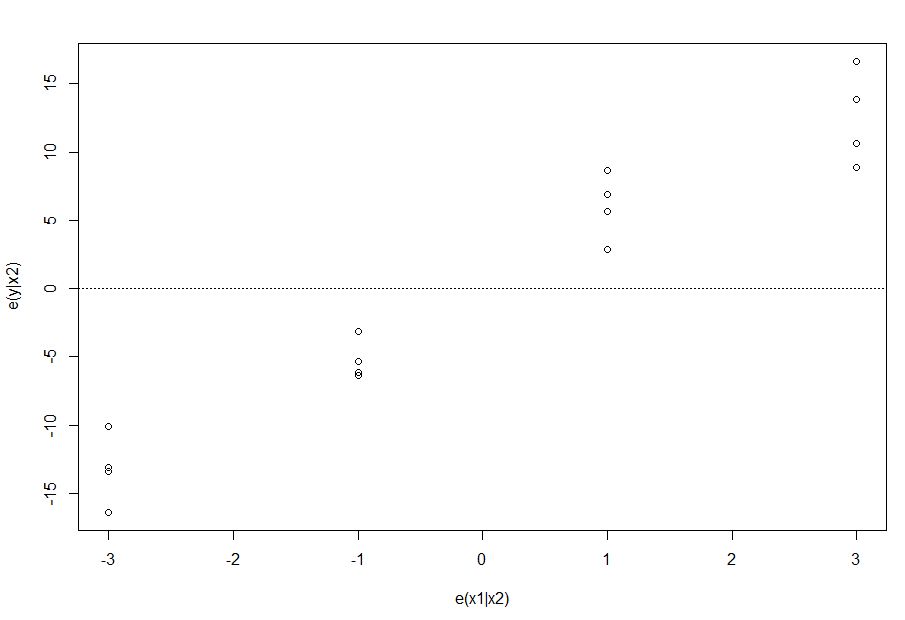
Although the model is much better, its use as a tool for admissions decisions is not much appropriate since admission should take more factors than these two into account.

# Problem 4 (10.5)

## (a)







## (b)

According to the plots above, we can conclude the original regression function is appropriate for the predictor variables. Because the residuals point lay uniformly in the plot and each predictor variables does not show any significant influence.

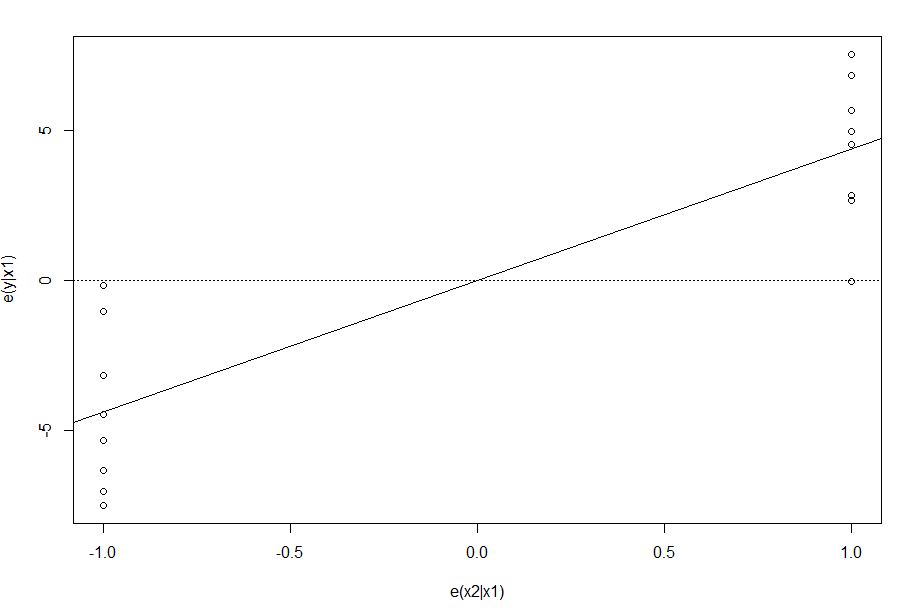
## (c)



Obtain:







The plot shows the line through the origin with slope equal to the regression coefficient for the predictor variable if it were added to the fitted model. This plot provides some useful additional information. The plot follows the prototype in Figure 10.1 a, suggesting that  is of little additional help in the model when  is already present.

# Problem 5 (10.9)

## (a)



We shall use the Bonferroni simultaneous test procedure with a family significance level of . We therefore require:



if , then conclude case i is not outlier;

if , then conclude case i not outlier;

Conclusion is that there is no outlier among the 16 cases.

## (b)

The diagonal elements of the hat matrix are provided in part (a) as hii.

There are only two levels of the diagonal elements of the hat matrix which are 0.1375 and 0.2375.

## (c)

Leverage values greater than 2p/n are considered by this rule to indicate outlying

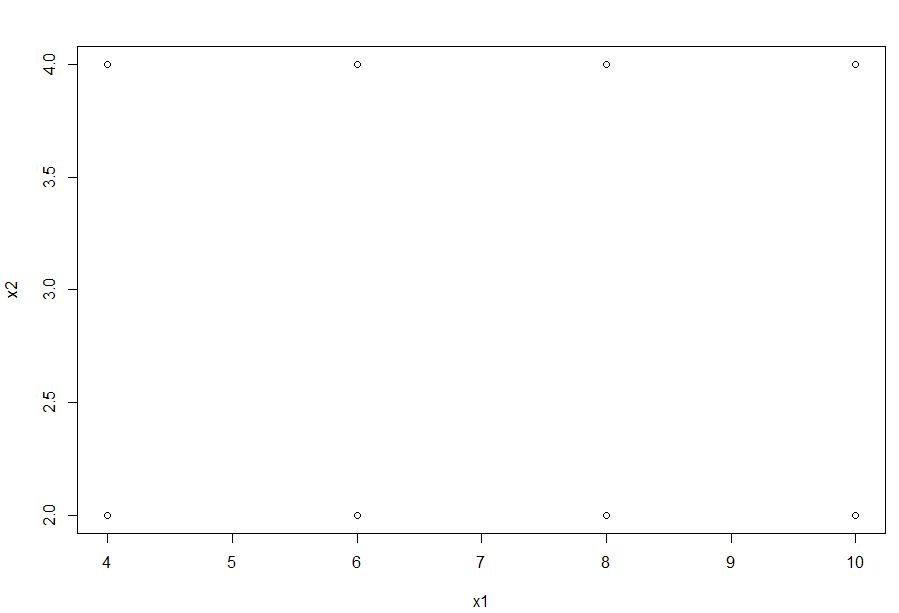
cases with regard to their X values. Therefore,





So there is not any of the observations outlying with regard to their X values according to the rule of thumb.

## (d)



Visually this prediction does not involve an extrapolation beyond the range of the data.



If  is much larger than the leverage values for the cases in the data set, an extrapolation is indicated.

Therefore, the small  leads to the conclusion that no extrapolation is needed.

## (g)



The larger either  or  is, the larger  is.

The  case can be influential:

(1) by having a large residual  and only a moderate leverage value . or

(2) by having a large leverage value  with only a moderately sized residual , or (3) by having both a large residual  and a large leverage value  .

Therefore, none of the case is influential.