CS 422 Homework0

Xingli Li, Illinois Institute of Technology

1. Exercises

Question 1.1

Section 3.7 (Exercises), page 123

Exercise 6.

Using (3.4), argue that in the case of simple linear regression, the least squares line always passes through the point (x̄, ȳ).

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

Question 1.2

Section 3.7 (Exercises), page 120

Exercises 1, 3, 4-a.

1 .Describe the null hypotheses to which the p-values given in Table 3.4 correspond. Explain what conclusions you can draw based on these p-values. Your explanation should be phrased in terms of sales, TV, radio, and newspaper, rather than in terms of the coefficients of the linear model.

3 .Suppose we have a data set with five predictors, X1 = GPA, X2 = IQ,X3 = Gender (1 for Female and 0 for Male), X4 = Interaction between GPA and IQ, and X5 = Interaction between GPA and Gender. The response is starting salary after graduation (in thousands of dollars). Suppose we use least squares to fit the model, and get β̂0 = 50, β̂1 = 20, β̂2 = 0.07, β̂3 = 35, β̂4 = 0.01, β̂5 = −10.

   (a) Which answer is correct, and why?

```
  i. For a fixed value of IQ and GPA, males earn more on average  than females.
  ii. For a fixed value of IQ and GPA, females earn more on  average than males.
  iii. For a fixed value of IQ and GPA, males earn more on average  than females
provided that the GPA is high enough.
  iv. For a fixed value of IQ and GPA, females earn more on average than males
provided that the GPA is high enough.
```

   (b) Predict the salary of a female with IQ of 110 and a GPA of 4.0.

   (c) True or false: Since ti. For a fixed value of IQ and GPA, males earn more on average than females.

        ii. For a fixed value of IQ and GPA, females earn more on average than males.

        iii. For a fixed value of IQ and GPA, males earn more on average than females provided that the GPA is high enough.

        iv. For a fixed value of IQ and GPA, females earn more on average than males provided that the GPA is high enough.

he coefficient for the GPA/IQ interaction term is very small, there is very little evidence of an interaction effect. Justify your answer.

4 .I collect a set of data (n = 100 observations) containing a single predictor and a quantitative response. I then fit a linear regression model to the data, as well as a separate cubic regression, i.e. $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$

(a) Suppose that the true relationship between X and Y is linear, i.e. Y = β0 + β1X + •. Consider the training residual sum of squares (RSS) for the linear regression, and also the training RSS for the cubic regression. Would we expect one to be lower than the other, would we expect them to be the same, or is there not enough information to tell? Justify your answer.