

ISL Conceptual Exercises Ch 3

Everton Lima

March 4, 2016

1

Each p value in Table 3.4 corresponds to the probability of the respective coefficient being zero. The conclusion drawn is that given the effect of both TV and radio has on sales there is no evidence that newspaper is related to sales.

2

Both KNN classifier and KNN regression use information about the neighborhood of data points to draw conclusions from, however each method has unique goals. In the case of KNN classifier we are interested in finding to which class a new data point belongs to, thus we assign it to the class of most frequent point class in the neighborhood of this point. On the other hand, in KNN regression our target is not a class label but rather a real number, so we assign a new point value to the mean of the points in the neighborhood.

3

3.a

Female : $85 + 20 \times GPA + 0.07 \times IQ + 0.01 \times GPA : IQ - 10 \times GPA$

Male : $50 + 20 \times GPA + 0.07 \times IQ + 0.01 \times GPA : IQ$

When GPA is large Females will have a lower salary. The correct alternative is **iii**.

3.b

$$50 + 20 \times 4 + 0.07 \times 110 + 35 + 0.01 \times 4 \times 110 - 10 \times 4 = 137.1$$

3.c

This may not be the case. While the coefficient is small it may have a high t-statistic and thus small p-value implying that an interaction effect exists.

4

4.a

I would expect the RSS to be lower when fitting cubic linear regression on the training set, since fitting polynomial coefficients would provide a tighter fit to the data.

4.b

In the test set the RSS would tend to be higher for the cubic linear model. Since the true relationship is linear the reduction of bias is not offset by a reduction in variance thus RSS is larger for the cubic linear model.

4.c

Since the cubic model is more flexible I would expect it to perform well on the training set.

4.d

There is not enough information in this case; Depending on how non-linear is the true relationship either model may perform better. In the case the true relationship is non-linear but closer to linear than cubic then the linear model will perform better. The opposite is true when the true relationship is closer to cubic than linear.

5

$$a_i = x_i / (\sum x_i^2)$$

6

A linear model with one predictor is defined by $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ and if (\bar{y}, \bar{x}) is a point on the line then $\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$ then $\bar{y} - \hat{\beta}_1 \bar{x} = \hat{\beta}_0$ which is true by the definition of $\hat{\beta}_0$ on equation 3.4.