Chapter2-Exercises

Anthony Chau

12/8/2019

Contents

Conceptual Exercises **2**

Question 1 2

Question 2 3

Conceptual Exercises

Question 1

For each of parts (a) through (d), indicate whether we would generally expect the performance of a flexible statistical learning method to be better or worse than an inflexible method. Justify your answer.

(a) The sample size n is extremely large, and the number of predictors p is small.

(b) The number of predictors p is extremely large, and the number of observations n is small.

(c) The relationship between the predictors and response is highly non-linear.

(d) The variance of the error terms, i.e. $\sigma^2 =$ `Var()`, is extremely high.

Solution:

(a) We would expect the performance of a flexible statistical learning method to be better than an inflexible method because the flexible methods can model the data more accurately given the large amount of observations. Also, since the number of predictors is small, the flexible methods would not estimate a large number of parameters.

(b) The performance of flexible methods is worse than inflexible methods. Given the large number of predictors and low number of observations, flexible methods would most likely result in overfitting as well as have low interprebility.

(c) The performance of flexible methods is better than inflexible methods because the flexible methods can better capture non-linear relationships over inflexible methods (such as linear regression).

(d) Flexible methods would perform better than inflexible methods because the flexible methods could include more potential variables that could be useful in predicting the response. The inclusion of those variables can offset the presence of a high variance for the error terms.

Question 2