Chapter 02: Statistical Learning

Stan Piotrowski

September 21 2021

Contents

Exerises

Conceptual

1) This exercises poses a question regarding the performance of flexible statistical learning methods relative to inflexible methods.

   a) For each of these scenarios, we need to consider the variance of the function describing the relationship between the response and the predictor(s) between successive training sets and the bias, or the error introduced when models are used to simplify reality. Both of these metrics factor into the overall test mean squared error (MSE), or the mean squared difference between the predicted value or the value estimated by the function and the observed value (in the case of supervised statistical learning).

   b)

   c) In a scenario where the relationship between the predictors and the response is highly non-linear, we would expect the flexible statistical learning method to be better than an inflexible method (generally). Inflexible learning methods, like ordinary least squares regression, for example, will likely have a low variance between successive training data sets, but have a large bias because we are attempting to explain a non-linear relationship with assumptions in a simplified linear reality.

   d)

Applied