ISLR | Chapter 6 Exercises

Marshall McQuillen

7/28/2018

Conceptual

1

- **A**. For a model with k predictors, Best Subset Selection will always have the best *training* RSS. The reason for this is, given a fixed k, there are $\binom{p}{k}$ possible models, and Best Subset Selection considers all of those $\binom{p}{k}$ possibilities.

  In Forward Stepwise Selection, of the total $\binom{p}{k}$ possible models, only the models that contain the $(k-1)$ model produced by Forward Stepwise Selection will be considered for the "best" k-variable model.

  In Backward Stepwise Selection, of the total $\binom{p}{k}$ possible models, the predictors in the "best" k-variable model *must* be a subset of the model with $(k+1)$ predictors.

  In short, Best Subset Selection will have the best (lowest) *training* RSS for a model with k predictors because it considers **all** the possible $\binom{p}{k}$ models, whereas Forward and Backward Stepwise Selection only consider a **subset** of all the possible $\binom{p}{k}$ models.

- **B**. There is no definitive answer for which subset selection method will have the lowest *testing* RSS (overfitting). If there is a large number of predictors, Best Subset Selection has the possibility of finding a model that has a low training RSS but a high testing RSS. Cross validation could be used to estimate the testing error of three models (one for Best Subset Selection, one for Forward Stepwise Selection and one for Backward Stepwise Selection) and a decision on which model has the lowest testing RSS could be made in consideration of the CV results.

- **C**.

  i. True.

  ii. True.

  iii. False, the predictors in the k-variable model identified by Backward Subset Selection are **not** a subset of the predictors in the $(k+1)$-variable model identified by Forward Subset Selection.

  iv. False, the predictors in the k-variable model identified by Forward Stepwise Selection are **not** a subset of the predictors in the $(k+1)$-variable model identified by Backward Stepwise Selection.

  v. False, the predictors in the k-variable model identified by Best Subset Selection are **not necessarily** a subset of the predictors in the $(k+1)$-variable model identified by Best Subset Selection.