

Linear Regression Subset Selection Reading Assignment

Read the beginning of Chapter 6 in ISLR and Section 6.1 on Subset Selection. Note the Algorithms 6.1, 6.2 and 6.3 for best, forward stepwise and backward stepwise regression. Pay attention to the problems that might be encountered with these approaches. Review or learn about the common criteria for choosing an optimal model: $MSE C_p$, AIC, BIC and adjusted R^2 . Read the section on Validation and Cross-Validation.

The assignment lab is in Section 6.5 Subset Selection Methods. The lab may clarify some of the Section 6.1 reading. This lab demonstrates that for the same number of variables the best, forward and backward variable selection methods do not always pick all of the same variables. Read through the lab in parallel with running the class provided version of the R scripts. The class version add comments, suggests some graph redesigns, provide a partial redesign, and make some minor adjustments such as rounding so R results match the text results. Bring questions about the script and methods to class. The script only has one plot due. The main assignment happens later when selecting variables for a crime data set.

.