1. Read Section 3.3 in the ESL textbook by yourself.

2. Repeat the procedure of Problem 2 in Homework 1 and compare the seven methods as follows: (I) full linear model; (II) reduced linear model (4) with \texttt{lcavol}, \texttt{lweight}, \texttt{lbph}, \texttt{svi}; (III) reduced linear model (2) with \texttt{lcavol}, \texttt{lweight}; (IV) subset selection using R function \texttt{step}; (V) Ridge regression; (VI) Lasso; (VII) Lars. That is, randomly partition the prostate cancer data into 67 training data points and 30 testing data points; fit/tune your model on the training data and estimate the mean (absolute) prediction error and mean (squared) prediction error using the testing data; and repeat the procedure for 100 times.

(1) Do pairwise comparison of the seven methods in terms of mean (absolute) prediction error and mean (squared) prediction error, respectively. Announce all significant differences at 5% level.

(2) Does your conclusion change across different partitions of training/testing sets?

(3) Do you conclude Ridge, Lasso, and Lars perform significantly better than the full linear model? If not, why do we still need those shrinkage methods?