

SMM637 Exercises - Shrinkage Approaches

In this exercise, we will predict the number of applications received using the variables in the `College` data set from the `ISLR` package.

1. Split the data set into a training set and a test set.
2. Fit a linear model using least squares on the training set, and report the test error obtained.
3. Fit a ridge regression model on the training set, with λ chosen by cross-validation. Report the test error obtained.
4. Fit a lasso model on the training set, with λ chosen by cross-validation. Report the test error obtained, along with the number of non-zero coefficient estimates.
5. Comment on the results obtained. How accurately can we predict the number of college applications received? Is there much difference among the test errors resulting from these 3 approaches?