

Lab 10

Exercise - Ridge Regression and Lasso

- a) Load the data `Boston` generate the correlation matrix for all variables in this dataset. Use the `cor` function to perform this task.
- b) Apply ridge regression.
- c) Observe how the slopes change with respect to the parameter λ . Remember that for $\lambda = 0$, we get LSE. Large λ forces them toward 0.
- d) Use the `select` function to choose a good λ value. (Hint: Fit ridge regression with various λ and compare prediction performance) What is the best λ to choose?
- e) An alternative way to study ridge regression and Lasso is to use the library `glmnet`. This package requires though the X variables in a matrix form. Remember to use `install.packages("glmnet")`.
- f) Which is the best λ to choose? Run cross-validation and for the selected values of λ plot the the mean cross-validation error (`cvm`), its estimated standard error (`cvstd`), the lower curve (`cvlo=cvm-cvstd`), and the upper curve (`cvup=cvm+cvstd`).
- g) Which λ minimized the MSE? Also, predict the regression coefficients for this value of `lambda`.
- h) Perform Lasso by selecting `alpha=1`.
- i) For Lasso, perform cross-validation (`nfolds` is 10) to find the optimum λ . Also, predict the regression coefficients for this value of `lambda`.
- j) For Lasso and the optimum value of `lambda`, perform validation set approach to compute the test MSE.