Lab 8

Maria Barouti

Exercise 1 - Resampling Methods

a) Load the data `Auto.rda` and install the package `ISLR`. Then split the data (half training - half testing) and use `lm` to predict `mpg` using the variables `weight`, `horsepower`, `acceleration`. Estimate the mean-squared error (MSE).

b) (Leave-One-Out Cross-Validation, LOOCV). Install `boot` and access the outputs of the cross-validation tool. Note: `glm()` - to perform logistic regression by passing in the `family="binomial"` argument. But if we use `glm()` to fit a model without passing in the family argument, then it performs linear regression, just like the `lm()` function.

c) Initiate a for loop which iteratively fits polynomial regressions for polynomials of order `i = 1` to `i = 10`. Compute the associated cross-validation error, and store it in the `i`th element of the vector `cv.error`.

d) Use the `cv.glm()` function to implement k-fold cross validation. Use `k = 10`, a common choice for `k`, on the Auto data set.