ASSIGNMENT #5

EPsy 8252

This assignment is intended to have you exercise some of your R programming skills. Please submit your responses to each of the questions below in a printed document. Only provide your responses to the question asked. You do not need to include any R syntax and output unless it is specifically required in the question.

Any graphics you include should be resized so that they do not take up more room than necessary and all should have an appropriate caption. Any equations should be appropriately typeset within the document.There are 10 points possible for the assignment (each question is worth one point unless otherwise noted).

Regression Estimation

Least squares estimation optimizes the criterion of the sum of squared error in computing the estimates for a given set of regression parameters. In this set of exercises, you are going to examine the estimates we obtain for the set of parameters for the model $\text{Wage}_i = \beta_0 + \beta_1(\text{Age}_i) + \epsilon_i$ when we change the criterion for model fit (mis-fit). In this set of exercises, you will be computing the sum of the absolute errors, $\left|Y_i - \hat{Y}_i\right|$, as a measure of the model fit (mis-fit). Use the following data set to help you answer the following questions,

1. Write a function that computes the sum of the absolute errors given the estimated the regression coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$. Include the syntax for your function in your word-processed document. Be sure that your function includes comments. **(2pts)**
2. Provide the output of your function (i.e., the sum of the absolute errors) for the parameter estimates of $\hat{\beta}_0 = 3$ and $\hat{\beta}_1 = 2$.
3. Carry out a grid search to estimate the coefficients (to the nearest hundredth) when we minimize the sum of the absolute errors. Report the parameter estimates. **(2pts)**
4. Compare the estimated effect of age from this estimation to that from the Least Squares estimate. How different are the interpretations for the effect of age on wage when we use a different criterion for measuring the model fit (or misfit)?

Gradient Descent

In the next set of exercises, you are going to program a function to estimate the conventional regression coefficients (i.e., optimizing the sum of squared errors) using the gradient descent algorithm presented in the class notes. In class we used this algorithm to compute the estimate for a single parameter (β_0). Here you will need to extend that to compute the estimate for both β_0 and β_1.

5. Write a function that incorporates the gradient descent algorithm to estimate the β_0 and β_1 regression coefficients for the wage–age data above (`myData`). This function should take inputs of the starting values for the two coefficients and a value for the tolerance. The output of the function should give the estimates for the parameters at each stage of the algorithm. For full credit present the commented syntax for your function, and the output when you use the starting values of $\hat{\beta}_0 = 3$ and $\hat{\beta}_1 = 2$ and a tolerance of 0.0001. **(4pts.)**

Here is a lecture about the gradient descent algorithm for linear regression models from Andrew Ng's *Machine Learning* course on Coursera https://class.coursera.org/ml-003/lecture/9. This may or may not help you do the actual R programming, but it will likely help you think about the algorithm more deeply. Note that he uses the terminology associated with Machine Learning (i.e., cost function, etc.) in the lecture.