STAT 440, Homework 6

due 4.23.2015

Load data. This can be done by using the command `load("regression.Rdata")`. You should now have two variables `x` and `y` with the obvious interpretations.

1. Write a function that takes as inputs the `x` and `y` coordinates of the data along with the bandwidth, `h`, and a gridpoint (`xgrid`), and returns the value of the regression curve at xgrid using kernel regression. Use this function to plot the data along with the regression curves for a variety of bandwidth values. Just using these plots, identify what you think is a reasonable bandwidth.

2. Write a loop which will output the empirical loss over a grid of reasonable bandwidth values. Identify the minimum bandwidth value.

3. Using the minimum bandwidth value, find the residuals of the kernel regression performed. Use exploratory data analysis and discuss your findings.