Math440 homework 1 due Friday January 22

If we follow the variation of some quantity over time, we are dealing with a time series. Time series are incredibly common: examples range from stock market movements to the icon on your computer that displays the CPU's utilization. What makes time series so different is that they allow us to see not only a single quantity by itself but at the same time give us some context for this quantity. On the face of it, time series analysis is a bivariate problem. However, time series raise a different set of issues and a rather specialized set of methods has been developed to deal with them. One issue will be the smoothing of data.

 # 1 For time series, the simplest smoothing algorithm that we can devise is what is called the running or moving average. The idea is straightforward: for any odd number of consecutive points, replace the centermost value with the average of the other points, ie.

$$s_i = \frac{1}{2k+1} \sum_{j=-k}^{k} x_{i+j}$$

This approach has a potential problem that will reveal itself to you by doing this homework problem. Dr. Martin has emailed you the data set, hmwk1.txt, which gives the time evolution of a certain quantity of interest to him. Using R, plot the original data. In addition, using R's ma() function, plot a smoothed version of the data. What attribute of the original data is hidden by referring only to the smoothed data?

#2 Dr. Martin has emailed to you the data set, co2.txt. The data represents 16 years of collecting monthly CO_2 data on the island of Hawaii, with the first year of CO_2 starting in January of 1958. A) Using R, plot the data. B) Using R's stl() command, plot the 1. trend, 2. seasonal, and 3. irregular omponents of the data.

For the following problem you may use R's lm function plus any amount of Mathematica to attempt to answer the question.

#3 The following table gives the depth Z of water in feet for surface points with rectangular coordinates X, Y in meters. The depth of measurements were made at low tide. Your ship has a draft of 5 feet. What region should you avoid within the rectangle $[75, 200] \times [-50, 150]$?

X	Y	Z
129.0	7.5	4
140.0	141.5	8
108.5	28.0	6
88.0	147.0	8
185.5	22.5	6
195.0	137.5	8
105.5	85.5	8
157.5	−6.5	9
107.5	−81.0	9
77.0	3.0	8
162.0	−66.5	9
162.0	84.0	4
117.5	−38.5	9

#4 My full name is James Elder Martin – numerically this corresponds to the vector 10, 1, 13, 5, 19, 5, 12, 4, 5, 18, 13, 1, 18, 20, 9, 14. Type the numerical representation of your own full name into R as a single vector. Then issue the command, source('batch.R'), and then (instructions for next calling batch.R are detailed in the comments of the file, batch.R) produce a vector containing batch averages, using a batch size, that is, a k, of 3.