CS416 Spring 2007
Prof. Wright

Assignment #8

Due April 20, 2007

1. A classic small data set in statistics is the "Iowa Wheat" data set, which shows the yield of the wheat harvest in Iowa for the years 1930-1962 together with the average temperatures and rainfalls for the months of the growing season. (The table can be found on the class web site.) The 1st column indicates the year and the final column indicates yield for that year. The 2nd, 4th, 6th, and 8th columns indicates the rainfalls for the four months of the growing season, while the 3rd, 5th, 7th, and 9th columns show the average temperatures for each of the four months. Note in particular the low yields and high temperatures during the "dust bowl" years in the 1930s.

   (a) By minimizing a least-squares function, by setting up and solving normal equations in a Matlab code called `wheat.m`, determine how the yield is predicted from the eight temperature/rainfall variables. That is, find coefficients $x_1, x_2, \ldots, x_8, x_9$ such that $y \approx \sum_{j=1}^{8} x_j z_j + x_9$ in the least-squares sense over the 33 years of data, where the z_j represent the eight temperature/rainfall values for each year and y represents the yield for each year. (The variable x_9 represents the constant offset, or "intercept.")

   You can load the data into your Matlab code using the Matlab "load" command, after editing the data file appropriately. (Do a "help load" to get the details.)

   (b) Using your solution $x_1, x_2, \ldots, x_8, x_9$, predict the yield that would be obtained in a year in which the data is as shown in the following table.

data	value
rain0	18.2
temp1	64.4
rain1	8.1
temp2	68.0
rain2	3.5
temp3	74.2
rain3	4.45
temp4	75.2

   (Add this calculation to your file `wheat.m`.)

2. Question 4 from page 528 in the text. The "line" referred to in this question is the straight line $y = az + b$, where a and b are to be determined by solving a least-squares problem.

3. Question 15 from page 529 in the text.

Hand in hard copies of your codes `wheat.m` and your output in the file `wheatOut.txt`, along with your written answers. Put your codes and output file in a directory called `homework8`. From the parent directory of `homework8`, run the following command:

```
handin -c cs416-1 -a hwk8 -d homework8
```