Statistics for Engineering, 4C3/6C3

Assignment 7

Kevin Dunn, kevin.dunn@mcmaster.ca **Due date: 7 March 2014**

Note: Source code will **not be graded** in this assignment. The purpose of this assignment is to show your understanding of the software output.

Question 1 [11]

Your company is developing a microgel-hydrogel composite, used for controlled drug delivery with a magnetic field. A previous employee did the experimental work but she has since left the company. You have been asked to analyze the existing experimental data.

The response variable, y = sodium fluorescein (SF) released [mg], per gram of gel and the data collected, in the original units are:

Experiment	Order	M = microgel weight [%]	H = hydrogel weight [%]	y
1	4	4	10	119
2	1	8	10	93
3	6	4	16	154
4	3	8	16	89
5	2	6	13	85
6	5	6	13	88
7	9	3.2	13	125
8	7	8.8	13	111
9	10	6	17.2	136
10	8	6	8.8	98

1. What was likely the reason the experimenter added experiments 5 and 6?
2. Why might the experimenter have added experiments 7, 8, 9 and 10 after the first six? Provide a rough sketch of the design, and all necessary calculations to justify your answer.
3. What is the name of the type of experimental design chosen by the employee for *all 10 experiments in the table*?
4. Using these data, you wish to estimate a nonlinear approximation of the response surface using a model with quadratic terms. Write out the equation of such a model that can be calculated from these 10 experiments (*also read the next question*).
5. Write out
   - the $\mathbf{X}$ matrix,
   - the corresponding symbolic entries in $\mathbf{b}$
   - and the $\mathbf{y}$ vector

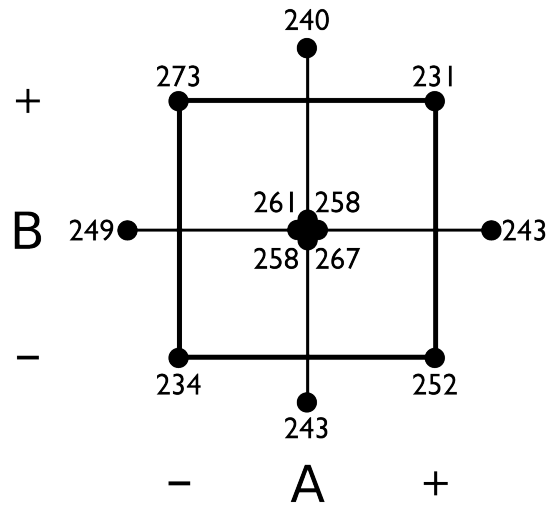
   that you would use to solve the set of linear equations $\mathbf{b} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$ to obtain the parameter estimates of the model you proposed in the previous part. You must use data from all 10 experiments.
6. How many degrees of freedom will be available to estimate the standard error and confidence intervals?
7. Now calculate the coefficients in the linear model using computer software. Which terms in the model are statistically significant?

   Note: "*linear*" implies the model is linear in the coefficients, not in the terms; that is because the coefficients in front of the nonlinear terms can still be found from solving a set of linear equations).

Question 2 [7]

The following diagram shows data from a central composite design. The factors were run at their standard levels, and there were 4 runs at the center point.

1. Calculate the parameters for a suitable quadratic model in these factors. Show your matrices for $\mathbf{X}$ and $\mathbf{y}$.
2. Draw a response surface plot of **A** *vs* **B** over a suitably wide range beyond the experimental region.
3. Where would you move **A** and **B** if your objective is to increase the response value?
   (a) Report your answer in coded units.
   (b) Report your answer in real-world units, if the full factorial portion of the experiments were ran at:
      - **A** = *stirrer speed*, 200rpm and 340 rpm
      - **B** = *stirring time*, 30 minutes and 40 minutes



You might feel more comfortable setting up the problem in MATLAB. You can use the contour plot functions in MATLAB to visualize the results.

If you are using R, you can use the `rbind(...)` or `cbind(...)` functions to build up your $\mathbf{X}$ matrix row-by-row or column-by-column. The equivalent of meshgrid in R is the `expand.grid(...)` function. Pee the R code on the course website that shows how to generate surface plots in R.

Question 3 [30]

The next question will be posted soon. It is due as a separate hand-in.