MA 573 — PROJECT 2

Due: Friday, September 15

Problem 1.

In this problem, we will model heat generated during the hardening of cement. Data from [1] is compiled in Table 1. Here y denotes heat with units of calories/gram cement and $x_1 - x_4$ respectively denote the percentage of tricalcium aluminate, tricalcium silicate, tetracalcium aluminoferite and dicalcium phosphate.

(a) Consider first the linear model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon.$$

Estimate the parameters, plot the residual, and determine confidence intervals of two standard deviations as well as 95% confidence intervals.

(b) Perform the same analysis using linear models that incorporate only x_1 as well as x_1 and x_2. How do your results compare with those obtained in (a)?

Obs. No.	x_1	x_2	x_3	x_4	y
1	7	26	6	60	78.5
2	1	29	15	52	74.3
3	11	56	8	20	104.3
4	11	31	8	47	87.6
5	7	52	6	33	95.9
6	11	55	9	22	109.2
7	3	71	17	6	102.7
8	1	31	22	44	72.5
9	2	54	18	22	93.1
10	21	47	4	26	115.9
11	1	40	23	34	83.8
12	11	66	9	12	113.3
13	10	68	8	12	109.4

Table 1: Cement data from [1].

Problem 2.

This is a variation of the Project on pages 58 and 59 of your book but with data having unknown parameters and statistical properties. This data is contained in the file `Data` where the first and second columns respectively contain simulated temporal and displacement measurements. After downloading the file, you can load it into MATLAB using the command `>> load Data`.

Consider the model

$$\frac{d^2y}{dt^2} + C\frac{dy}{dt} + Ky = 0$$

$$y(0) = 2 \; , \; \frac{dy}{dt}(0) = 10.$$

(i) Estimate the parameters C and K using the ordinary least squares method. You should include a plot of your model fit to the data. Use a continuous line type for the model and a discrete line type for the data.

(ii) Plot the residuals using a discrete line type (e.g., `x`). Do your errors appear to be iid?

(iii) Determine an estimate $\hat{s}^2$ for the variance σ_0^2 and compute the covariance matrix $\widehat{\text{cov}(Q)}$ using both the analytic expressions and a finite difference approximation. What are the standard errors $SE_k(\hat{q}) = \sqrt{\widehat{\text{cov}(Q)}_{kk}}$ for $k = 1, 2$. Do the parameters appear to be correlated?

(iv) Determine the 95% confidence intervals for each parameter.

References

[1] G. Hald, *Statistical Theory with Engineering Applications*, John Wiley and Sons, New York, 1952.