LONDON TAUGHT COURSE CENTRE

LTCC Basic Course — Statistical Modelling and Estimation

Exam Question **April 2010**

You should attempt all parts of the question. Software such as Maple or Mathematica may be used for inverting matrices.

An experiment was carried out to maximise the yield of a chemical reaction. There were three continuous explanatory variables

X_1 = reaction time in hours,

X_2 = reaction temperature in degrees Celsius and

X_3 = pressure.

The response variable Y was the percentage yield of the reaction.

The experiment used a central-composite design with $n = 11$ runs. The design in which the values of the explanatory variables are coded to lie between -1 and 1 and the responses are shown below:

	Variable			
Run	X_1	X_2	X_3	Y
1	-1	-1	-1	27.9
2	-1	1	1	21.2
3	1	-1	1	94.2
4	1	1	-1	38.1
5	-1	0	0	31.6
6	1	0	0	60.7
7	0	-1	0	49.6
8	0	1	0	32.1
9	0	0	-1	31.1
10	0	0	1	48.6
11	0	0	0	39.5

A suitable model for these data is the second-order polynomial regression or second-order response surface model with expected response

$$E(Y_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_{11} x_{1i}^2 + \beta_{22} x_{2i}^2 + \beta_{33} x_{3i}^2 + \beta_{12} x_{1i} x_{2i} + \beta_{13} x_{1i} x_{3i} + \beta_{23} x_{2i} x_{3i}$$

where the variables $Y_1, \ldots, Y_n$ are assumed to be uncorrelated and to have the same variance $V(Y_i) = \sigma^2$, $i = 1, \ldots, n$.

(a) Write down the model in matrix form, that is write down the design matrix $\mathbf{X}$, the vector of parameters $\boldsymbol{\beta}$ and the variance-covariance matrix $V(\mathbf{Y})$.

(b) For the given data is the least squares estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ uniquely determined? Justify your answer, then find $\hat{\boldsymbol{\beta}}$.

(c) Test at the 5% level of significance if the full second-order model can be simplified to the model with expected response

$$E(Y_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_{11} x_{1i}^2 + \beta_{12} x_{1i} x_{2i} + \beta_{13} x_{1i} x_{3i} + \beta_{23} x_{2i} x_{3i}.$$

In other words, we wish to test simultaneously if $\beta_{22} = 0$ and if $\beta_{33} = 0$.

(d) Suppose that due to some mishap the observations for runs 9 and 10 in the table given at the start of the question were missing, so that only $n = 9$ observations would be available and the design matrix would only have nine rows. In this case which of the parameters β_1 and β_3 in the full second-order model would be estimable and which one would not be estimable.