Chapter 32

Response Surface Methodology

We describe regression analyses in this chapter that can deal with quadratic (but not higher, typically) terms or, equivalently, *curvature* in the response surface. Two types of response surface regressions are discussed: central composite response surface designs and optimal response surface designs.

32.1 Response Surface Experiments

We look at the following coding scheme in this section,

$$X_j = \frac{\text{actual level } - \frac{\text{high level} + \text{low level}}{2}}{\frac{\text{high level} - \text{low level}}{2}}$$

Coding schemes are used in the response surface regression analyses, described in later sections.

Exercise 32.1 (Response Surface Experiments)
Consider the effect of air temperature and noise on the rate of oxygen consumption of mice. Two factors are studied, with the following low and high levels.

factor	low level, -1	high level, 1
temperature, X_1	0^o F	30^o F
noise, X_2	10 dB	100 dB

and where the regression equation is

$$Y_i = \beta_0 X_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_{12} X_{i1} X_{i2} + \beta_{11} X_{i1}^2 + \beta_{22} X_{i2}^2 + \varepsilon_{ijk}$$

1. *Using the coding scheme for X_1*

   The coded version of $X_1 = 0$ is

$$X_j = \frac{\text{actual level} - \frac{\text{high level} + \text{low level}}{2}}{\frac{\text{high level} - \text{low level}}{2}}$$
$$= \frac{0 - \frac{30+0}{2}}{\frac{30-0}{2}} =$$

   (choose one) **−1 / 0 / 1**

   and the coded version of $X_1 = 30$ is

$$X_j = \frac{30 - \frac{30+0}{2}}{\frac{30-0}{2}} =$$

   (choose one) **−1 / 0 / 1**

   and the coded version of $X_1 = 20$ is

$$X_j = \frac{20 - \frac{30+0}{2}}{\frac{30-0}{2}} =$$

   (choose one) **−1** / $\mathbf{\frac{1}{3}}$ / **1**

   and the coded version of $X_1 = 40$ (larger than the high value) is

$$X_j = \frac{40 - \frac{30+0}{2}}{\frac{30-0}{2}} =$$

   (choose one) **−1** / $\mathbf{\frac{1}{3}}$ / $\mathbf{\frac{25}{15}}$
   and so, in summary,

temperature, X_1	coded value
0	-1
30	1
20	$\frac{1}{3} = 0.333$
40	$\frac{25}{15} = 1.667$

2. *Using the coding scheme for X_2*

   The coded version of $X_2 = 10$ is

$$X_j = \frac{\text{actual level} - \frac{\text{high level} + \text{low level}}{2}}{\frac{\text{high level} - \text{low level}}{2}} = \frac{10 - \frac{100+10}{2}}{\frac{100-10}{2}} =$$

   (choose one) **−1 / 0 / 1**

   and the coded version of $X_2 = 100$ is

$$X_j = \frac{100 - \frac{100+10}{2}}{\frac{100-10}{2}} =$$

   (choose one) **−1 / 0 / 1**

   and the coded version of $X_2 = 200$ is

$$X_j = \frac{200 - \frac{100+10}{2}}{\frac{100-10}{2}} =$$

   (choose one) **−1** / $\mathbf{\frac{1}{3}}$ / $\mathbf{\frac{145}{45}}$
   and so, in summary,

noise, X_2	coded value
10	-1
100	1
200	$\frac{145}{45} = 3.22$

3. *Coding scheme property*
   True / **False** The coding scheme,

$$X_j = \frac{\text{actual level} - \frac{\text{high level} + \text{low level}}{2}}{\frac{\text{high level} - \text{low level}}{2}}$$

   gives values *in*side the interval $(1, -1)$ for levels between the high and low values, and values *out*side $(-1, 1)$ for levels either below the low value or above the high value.

4. *Review: coding scheme for 2^2 factorial design*
True / **False** The coding scheme for a 2^2 factorial design is,

temperature, X_1	coded value
0	-1
30	1

noise, X_2	coded value
10	-1
100	1

5. *Coding scheme for 3^2 factorial design*
True / **False** The coding scheme for a 3^2 factorial design is,

temperature, X_1	coded value
0	-1
15	0
30	1

noise, X_2	coded value
10	-1
55	0
100	1

where, notice, both factors have three levels each.

6. *Coding scheme for 5^2 factorial design*
True / **False** The coding scheme for a 5^2 factorial design is,

temperature, X_1	coded value
-6.21	$-\sqrt{2} = -1.414$
0	-1
15	0
30	1
36.21	$\sqrt{2} = 1.414$

noise, X_2	coded value
-8.64	$-\sqrt{2} = -1.414$
10	-1
55	0
100	1
118.64	$\sqrt{2} = 1.414$

7. *Visualizing coding schemes; composite designs*
   Match the schemes with the graphs given in the figure below.

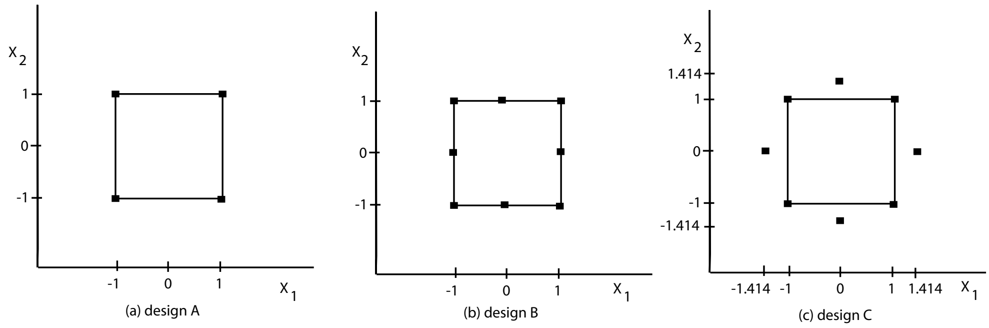

Figure 32.1 (Various designs)

design	graph
(a) 2^2 full factorial	**(a)** design A
(b) 3^2 full factorial, or composite design, $\alpha = 1$	**(b)** design B
(c) 5^2 full factorial, or composite design, $\alpha = \sqrt{2}$	**(c)** design C

design	(a)	(b)	(c)
graph			

Design (b) is not only a 3^2 design but also a *composite* design where $\alpha = 1$. The $\alpha = 1$ because the horizontal (or vertical) distance between the center of the box and sides of the box is one (1). Design (c) is not only a 5^2 design, but also a composite design where $\alpha = \sqrt{2}$.

8. *Central composite designs*
   True / **False** A *central* composite design is one where, in addition to observations to *corner* design points and *star* design points, there are some center design points (used to calculate error), as given in the figure below.

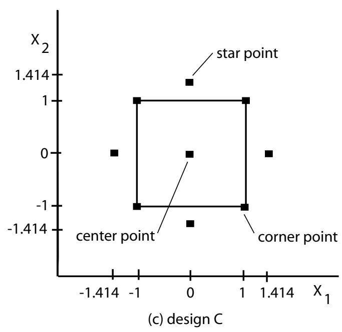

Figure 32.2 (Central composite design)

9. *Number of regression coefficients*
   The response function is given by

$$Y_i = \beta_0 X_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_{12} X_{i1} X_{i2} + \beta_{11} X_{i1}^2 + \beta_{22} X_{i2}^2 + \varepsilon_{ijk}$$

   has (choose one) **4** / **5** / **6** regression coefficients
   where there are (choose one) **1** / **2** / **3** linear main effects
   and (choose one) **2** / **3** / **4** quadratic main effects
   and (choose one) **1** / **2** / **3** two–factor effects

10. *Number of design points and observations*
    On the one hand, at least six (6) design points (not observations!) are required to estimate the six regression coefficients in the model,

$$Y_i = \beta_0 X_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_{12} X_{i1} X_{i2} + \beta_{11} X_{i1}^2 + \beta_{22} X_{i2}^2 + \varepsilon_{ijk}$$

    On the other hand, a 2^2 full factorial design, no replication, provides
    (choose one) **2** / **3** / **4** design points
    (choose one) **2** / **3** / **4** observations (experimental trials)
    and so **is** / **is not** able to estimate the regression coefficients in the model

    A 2^2 full factorial design, two replications, provides
    (choose one) **2** / **3** / **4** design points
    (choose one) **4** / **6** / **8** observations
    and so **is** / **is not** able to estimate the regression coefficients in the model

    and a composite design with $\alpha = 1$, no replications, provides
    (choose one) **7** / **8** / **9** design points
    (choose one) **7** / **8** / **9** observations
    and so **is** / **is not** able to estimate the regression coefficients in the model

    and a *central* composite design with $\alpha = 1$, two replications, provides
    (choose one) **7** / **8** / **9** design points
    (choose one) **14** / **16** / **18** observations
    and so **is** / **is not** able to estimate the regression coefficients in the model

    and a central composite design with $\alpha = \sqrt{2}$, two replications, provides
    (choose one) **7** / **8** / **9** design points
    (choose one) **14** / **16** / **18** observations
    and so **is** / **is not** able to estimate the regression coefficients in the model

11. *More on the number of observations required in composite designs*
The number of observations (experimental trials) required for a 2^{k-f} fractional factorial design (where k is the number of factors and f is the level of fractionation) with n_c *replications* for *each* corner point, n_s *replications* for *each* star point and n_0 *replications* for *each* center point, is

$$n_T = 2^{k-f} n_c + 2kn_s + n_0$$

For example, for a 2^2 factorial design, no replications,
$k =$ (choose one) **1** / **2** / **3**
$f =$ (choose one) **0** / **1** / **2**
$\alpha =$ (choose one) **1** / **2** / **3**
$n_c =$ (choose one) **0** / **1** / **2**
$n_s =$ (choose one) **0** / **1** / **2**
$n_0 =$ (choose one) **0** / **1** / **2**
and so number of observations is

$$n_T = 2^{k-f} n_c + 2kn_s + n_0 = 2^{2-0}(1) + 2(2)(0) + 0 = 4$$

12. *And still more on the number of observations required in composite designs*
A 2_V^{4-1} fractional factorial with $\alpha = \sqrt{2}$ composite design, and with two replications for each of the corner and star points and four replications for the center point ($n_c = n_s = 2$, $n_0 = 4$), requires

$$n_T = 2^{k-f} n_c + 2kn_s + n_0 = 2^{4-1}(1) + 2(4)(2) + 4 =$$

(choose one) **27** / **28** / **29** observations

13. *Rotatable central composite designs*
A central composite design is *rotatable* ($\sigma^2\{\hat{Y}_h\}$ same for any $\mathbf{X}_h$ that is a given distance from the center point) if

$$\alpha = \left(\frac{2^{k-f}(n_c)}{n_s}\right)^{1/4}$$

For example, for a 2^2 factorial ($k = 2$, $f = 0$) where $n_c = n_s = 1$,

$$\alpha = \left(\frac{2^{2-0}(1)}{1}\right)^{1/4} =$$

(choose one) **1** / $\mathbf{\sqrt{2}}$ / **2**

32.2 Central Composite Response Surface Designs

SAS program: att14-32-2-mice-compositeresponse

Central composite response surface designs are two–level full or fractional factorial designs that have a few extra observations added to allow estimation of second–order response surface models.

Exercise 32.2 (Central Composite Response Surface Designs)
Consider the effect of air temperature (X_1) and noise (X_2) on the rate of oxygen consumption (Y) of mice. A two–factor rotatable central composite design with four replications is used,

ROC, Y	X_1	X_2
7.6	-1	-1
8.2	1	-1
8.0	-1	1
9.2	1	1
5.4	-1.414	0
5.9	-1.414	0
6.3	0	-1.414
8.5	0	1.414
7.3	0	0
7.3	0	0
8.4	0	0
9.5	0	0

and where the regression equation is

$$\begin{aligned} Y_i = \beta_0 X_0 &+ \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_{12} X_{i1} X_{i2} \\ &+ \beta_{11} X_{i1}^2 + \beta_{22} X_{i2}^2 + \varepsilon_{ijk} \end{aligned}$$

1. *Two–factor rotatable central composite design*
   True / False
   From SAS, the second–order response function is where the estimated coefficients are

coefficient	b_q	p–value
b_0	8.12474	< 0.0001
b_1	0.31342	0.5138
b_2	0.56394	0.2584
b_{12}	0.15	0.8222
b_{11}	-0.80624	0.1616
b_{22}	0.06902	0.8958

2. *Residuals*
   From SAS, the residuals look randomly scattered and so this indicates (choose one) **constant** / **non–constant** variance.

3. *Lack of fit test*
   From SAS, the ANOVA table is

Source	Sum Of Squares	Degrees of Freedom	Mean Squares
Regression	7.93093	5	1.58619
Error	9.79574	6	1.63262
Lack of Fit	6.468238	3	2.156079
Pure Error	3.3275	3	1.109167
Total	17.72667	11	

$H_0: \ \mu = \beta_0 X_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_{12} X_{i1} X_{i2} + \beta_{11} X_{i1}^2 + \beta_{22} X_{i2}^2$ versus
$H_a: \ \mu \neq \beta_0 X_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_{12} X_{i1} X_{i2} + \beta_{11} X_{i1}^2 + \beta_{22} X_{i2}^2$
(In other words, H_o : no lack of fit versus H_a : lack of fit)
The test statistic is

$$\begin{aligned} F^* &= \frac{SSE - SSPE}{df_E - df_{PE}} \div \frac{SSPE}{df_{PE}} \\ &= \frac{SSLF}{df_{LF}} \div \frac{SSPE}{df_{PE}} \\ &= \frac{6.468238}{3} \div \frac{3.3275}{3} \\ &= 1.94 \end{aligned}$$

The critical value is $F(1-\alpha; df_{LF}, df_{PE}) = F(0.99; 3, 3) = 29.5$
since $F^* = 1.94 < F = 29.5$
accept null; that is, there (choose one) **is** / **is no** lack of fit
(In other words, the model appears to be a good one.)

32.3 Optimal Response Surface Designs

Not covered.

32.4 Analysis of Response Surface Experiments

SAS program: att14-32-4-mice-surface-analysis

The analysis of second–order response surfaces includes

- estimation of response function (done above)
- model visualization
- identification of optimum operating conditions

Exercise 32.3 (Analysis of Response Surface Experiments)
Consider the effect of air temperature (X_1) and noise (X_2) on the rate of oxygen consumption (Y) of mice. A two–factor rotatable central composite design with four replications is used,

ROC, Y	X_1	X_2
7.6	−1	−1
8.2	1	−1
8.0	−1	1
9.2	1	1
5.4	−1.414	0
5.9	−1.414	0
6.3	0	−1.414
8.5	0	1.414
7.3	0	0
7.3	0	0
8.4	0	0
9.5	0	0

and where the regression equation is

$$\begin{aligned} Y_i = \beta_0 X_0 &+ \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_{12} X_{i1} X_{i2} \\ &+ \beta_{11} X_{i1}^2 + \beta_{22} X_{i2}^2 + \varepsilon_{ijk} \end{aligned}$$

1. *Three–dimensional plot and contour plot for response surface*
   Both the three–dimensional plot and contour plot for the response surface indicate a surface which is at a maximum for large values of X_1 and X_2, at, say, roughly, $(X_1, X_2) =$ (choose one) $\mathbf{(-1,-1)}$ / $\mathbf{(-1.4, 1.4)}$ / $\mathbf{(1.4, 1.4)}$

2. *Calculation of maximum for the response surface*
   True / False

$$\mathbf{B} = \begin{bmatrix} -0.80624 & 0.15/2 \\ 0.15/2 & 0.06902 \end{bmatrix}, \quad \mathbf{b}^* = \begin{bmatrix} 0.31342 \\ 0.56394 \end{bmatrix}$$

   and so

$$\mathbf{X_s} = -\frac{1}{2}\mathbf{B^{-1}b}^* = \begin{bmatrix} -0.17 \\ -3.90 \end{bmatrix}$$

   Since $(X_1, X_2) = (-0.17, -3.90)$ is well outside the range of the analysis (and does not match the point found by looking at the plots above), the model should be reformulated to better fit the data.

32.5 Sequential Search for Optimum Conditions–Method of Steepest Ascent

Not covered.