Response Surface Methodology 16

16.1 Introduction

Response surface methodology was developed by Box and Wilson in 1951 to aid the improvement of manufacturing processes in the chemical industry. The purpose was to optimize chemical reactions to obtain, for example, high yield and purity at low cost. This was accomplished through the use of sequential experimentation involving factors such as temperature, pressure, duration of reaction, and proportion of reactants. The same methodology can be used to model or optimize any response that is affected by the levels of one or more quantitative factors. The models are generalizations of the polynomial regression models studied in Chap. 8.

The general scenario is as follows. The response is a quantitative continuous variable (e.g., yield, purity, cost), and the mean response is a smooth but unknown function of the levels of p factors (e.g., temperature, pressure), and the levels are real-valued and accurately controllable. The mean response, when plotted as a function of the treatment combinations, is a surface in $p + 1$ dimensions, called the *response surface*. For example, Fig. 16.1 shows a response surface for two factors A and B.

We will denote the levels of A by values of x_1 or x_A and the levels of B by values of x_2 or x_B. We will denote a treatment combination by $\mathbf{x} = (x_1, x_2, \ldots, x_p)$ or by $\mathbf{x} = (x_A, x_B, \ldots, x_P)$ and the mean response at $\mathbf{x}$ by $\eta_{\mathbf{x}} = E[Y_{\mathbf{x}}]$. The general response surface model is of the form

$$Y_{\mathbf{x}} = \eta_{\mathbf{x}} + \epsilon_{\mathbf{x}} ,$$

where $\epsilon_{\mathbf{x}}$ is a random error variable.

The objective of obtaining a response surface is twofold:

(i) to locate a feasible treatment combination $\mathbf{x}$ for which the mean response is maximized (or minimized, or equal to a specific target value); and
(ii) to estimate the response surface in the vicinity of this good location or region, in order to better understand the "local" effects of the factors on the mean response.

In general, throughout the chapter we will think about maximizing the response, but we show via an example that exactly the same techniques can be used for minimizing a response. The techniques can easily be adapted when the goal is to have the response close to a target value.

A. Dean et al., *Design and Analysis of Experiments*,
Springer Texts in Statistics, DOI 10.1007/978-3-319-52250-0_16

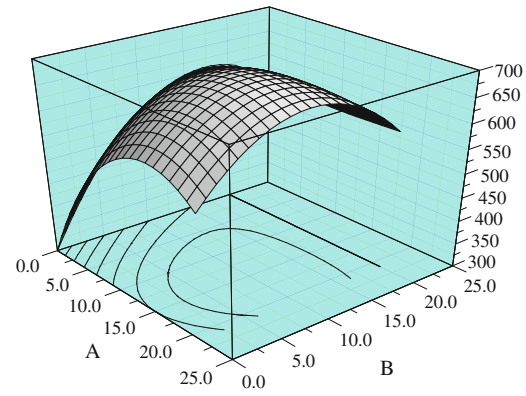


Fig. 16.1 Hypothetical response surface for two factors

One possible approach to achieving the objective involves collecting observations at each location on a grid of treatment combinations spanning the entire experimental region of interest (as suggested by Fig. 16.1). However, the number of observations required by such a comprehensive approach can be very large, and it grows very quickly as the number of factors under study increases. Also, somewhat sophisticated modeling techniques would generally be needed to obtain an adequate fit of a model over the entire region. Instead, it is generally more efficient to conduct a sequence of small "local" experiments with which to search out the location of the peak mean response and then to study its vicinity.

Seeking out the peak is analogous to climbing an unfamiliar mountain under conditions of limited visibility—the mountain is the response surface, and your location on the mountainside is a treatment combination, say $\mathbf{x}_a$. Standing at position $\mathbf{x}_a$, you look around and can see enough to determine in which direction to go to continue a steep ascent. Then you climb in the determined direction as long as it continues to take you up, not looking about lest you lose footing. Then you stop and look around again to determine whether you are at the top of the mountain or in which direction you need to continue your ascent. Of course, when you reach a peak, due to the limited visibility, you may not be sure that you have actually reached the highest peak.

How does one do this experimentally? Looking around with limited visibility is equivalent to analyzing the data of a *local experiment*, consisting of observations on treatment combinations $\mathbf{x}$ close to your current position, $\mathbf{x}_a$. The local terrain is assessed by fitting a local model. Collecting observations in sufficiently close proximity to one another generally allows the local response surface to be well approximated by a rather simple polynomial regression model. When still far from the peak, a first-order model is often adequate. The fitted first-order model is a plane, from which the direction or *path of steepest ascent* is easily determined. Then observations are collected along this path as long as the response continues to increase. When the response stops increasing, another local experiment can be conducted to determine a new path of steepest ascent. This process can be iterated until the first-order model no longer adequately describes the local true surface. For example, close to the peak, the true surface generally exhibits greater curvature, and a first-order regression model becomes inadequate, exhibiting lack of fit. A larger number of observations is needed to fit a higher-order model with which to locate and study the vicinity of the peak. Typically, a second-order model is suitable.

A flow chart describing the steps in this process is shown in Fig. 16.2. While a surface is difficult to envisage in more than three dimensions, the process can work well for any number of factors. How well it works depends on several decisions requiring judgment on the part of the experimenter. The first part of this chapter (Sect. 16.2) looks at the left-hand portion of the flow chart and investigates first-order designs and first-order models, including lack of fit and the path of steepest ascent. Section 16.3

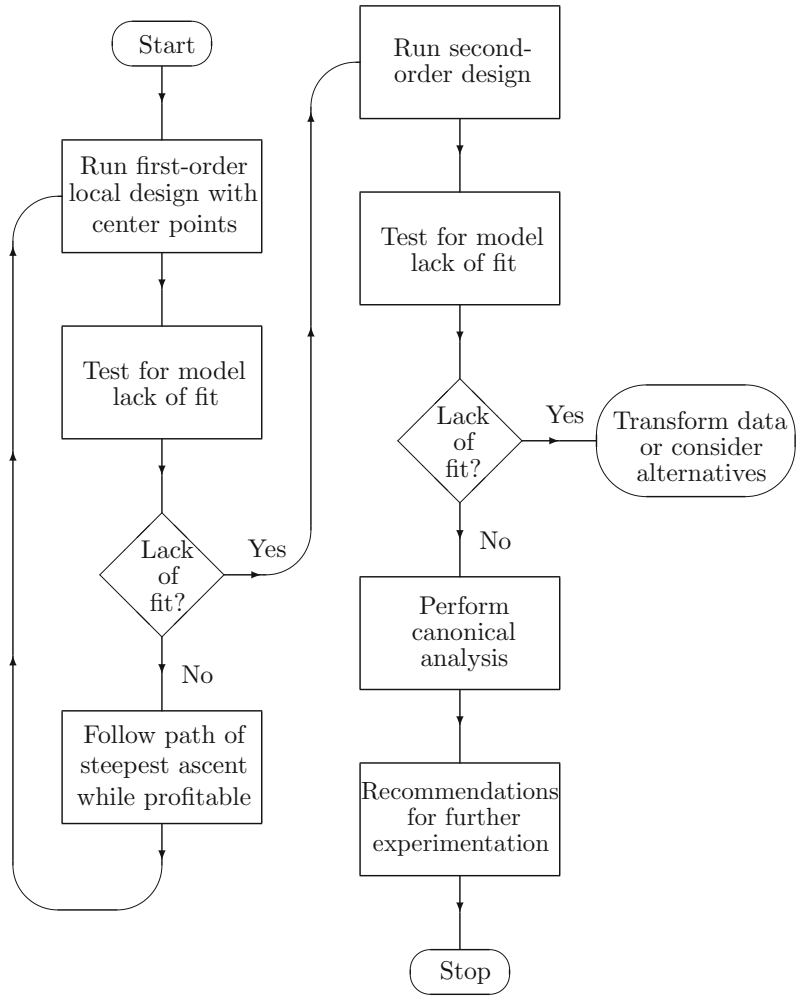


Fig. 16.2 Flow chart for response surface methods

addresses the right-hand portion of the flow chart, which becomes relevant when the vicinity of the peak is reached. Second-order designs and models are described. More details about second-order designs are given in Sect. 16.4, and an experiment conducted in the flour milling industry is described in Sect. 16.5. The collection of observations as a block design is discussed in Sect. 16.6. Sections 16.7 and 16.8 describe the use of the SAS and R software, respectively.

16.2 First-Order Designs and Analysis

16.2.1 Models

Before the peak of the response surface is reached, a small local experiment is conducted to assess the local terrain. If the local experiment is not in the vicinity of the peak, then a first-order regression model often provides an adequate approximation to the local response surface. For p factors, the standard *first-order model* is a first-order polynomial regression model:

$$Y_{\mathbf{x},t} = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \epsilon_{\mathbf{x},t}\,, \tag{16.2.1}$$

where $Y_{\mathbf{x},t}$ denotes the tth observation at treatment combination $\mathbf{x} = (x_1, \ldots, x_p)$, and the error variables $\epsilon_{\mathbf{x},t}$ are assumed to be independent with $N(0, \sigma^2)$ distributions. The parameter β_i is a measure of the local *linear effect* of the ith factor ($i = 1, \ldots, p$).

We often code the levels of each factor in each local experiment so that zero represents the *midrange* of the levels of the factor (the average of the highest and lowest levels included in the experiment) and $+1$ and -1 represent the highest and lowest levels of the factor, respectively. For the ith factor, such coded levels z_i are obtained as

$$z_i = (x_i - m_i)/h_i\,, \tag{16.2.2}$$

where m_i denotes the midrange of the values of x_i of the ith factor, and h_i denotes the *half-range*—half of the range. So, in terms of coded levels, the center of the design corresponds to the point $\mathbf{z}_0 = (0, 0, \ldots, 0)$.

The first-order model (16.2.1) can be rewritten in terms of the coded factor levels as follows:

$$Y_{\mathbf{z},t} = \gamma_0 + \gamma_1 z_1 + \cdots + \gamma_p z_p + \epsilon_{\mathbf{z},t}\,. \tag{16.2.3}$$

The parameters in models (16.2.1) and (16.2.3) are related, since

$$\beta_0 = \gamma_0 - \sum_i m_i \gamma_i / h_i \quad \text{and} \quad \beta_i = \gamma_i / h_i \;\; (i = 1, 2, \ldots, p)\,.$$

A design for estimating the parameters of a first-order model is called a *first-order design*. A first-order design should (i) allow for efficient estimation of each linear effect β_i or γ_i, (ii) allow a test for lack of fit of the first-order model, and (iii) be expandable to a good second-order design.

As long as there is no significant model lack of fit but there are significant linear effects, the fitted first-order model can be used to estimate the path of steepest ascent. If there is significant lack of fit of the first-order model, then additional observations may be collected to augment the first-order design so that a second-order polynomial regression model can be fitted to the data.

If there is no significant model lack of fit and also no significant linear effects, then more data may be needed to increase precision of the parameter estimators. Alternatively, the experimenters may need to change the factors under study or increase the range of levels.

16.2.2 Standard First-Order Designs

Throughout the rest of Sect. 16.2, we consider designs which we refer to as *standard first-order designs*. These designs consist of n_f "factorial points" and n_0 "center points." The *factorial points* consist of the treatment combinations of a 2^p factorial experiment run as a completely randomized design as in Chap. 7, or a 2^{p-s} fractional factorial design of Resolution III or higher. The *center points* are observations collected at the center of the local region under study; that is, at $z_0 = (0, 0, \ldots, 0)$. These are needed to provide error degrees of freedom and to provide adequate power for a test for model lack of fit.

Standard first-order designs are *orthogonal*, which means that

(i) for each factor, the sum of the coded levels used in the design is zero, ($\sum z_i = 0$, summing over observations), so half of the n_f factorial points in the design have each factor at its high level and the other half have each factor at its low level; and
(ii) for each pair of factors, the sum of cross products of the coded levels in the design is zero ($\sum z_i z_j = 0$, summing over observations).

The factorial portion of the design is chosen to be at least Resolution III so that the linear effects can be estimated. Higher resolution allows model lack of fit due to two-factor interaction effects to be tested. The 2^{p-s} orthogonal fractional factorial designs and the Plackett–Burman designs of Chap. 15 are the most efficient designs for estimation of the linear effects.

16.2.3 Least Squares Estimation

The method of least squares (as shown in optional Sect. 8.3) is used to fit a first-order model to the data. Denote the fitted model by

$$\hat{y}_{\mathbf{x}} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p \tag{16.2.4}$$

or, in coded form,

$$\hat{y}_{\mathbf{z}} = \hat{\gamma}_0 + \hat{\gamma}_1 z_1 + \cdots + \hat{\gamma}_p z_p \,. \tag{16.2.5}$$

If a standard first-order design is used, with the extreme levels of each factor coded as $+1$ and -1, then the least squares estimator $\hat{\gamma}_i$ of the linear effect γ_i of the ith factor is

$$\hat{\gamma}_i = (\overline{Y}_{z_i(+1)} - \overline{Y}_{z_i(-1)})/2 \,, \tag{16.2.6}$$

where $\overline{Y}_{z_i(+1)}$ and $\overline{Y}_{z_i(-1)}$ denote the averages of the response variables at the high and the low level of the ith factor, respectively. The parameter $2\gamma_i$ denotes the change in the mean response between the high and low levels of the ith factor. This is the same as the main-effect contrast for the ith factor. The least squares estimator of β_i in the uncoded model is $\hat{\beta}_i = \hat{\gamma}_i / h_i$, where h_i is the half-range of the uncoded levels of the ith factor.

Example 16.2.1 `Paint experiment, continued`

Several experiments were run in Germany by Eibl et al. (1992) on the thickness of a paint coating. The first experiment in the series was examined in Exercise 7 of Chap. 15. To study how to decrease the mean thickness, the experimenters selected the following six factors, each at two levels:

A :	belt speed	B :	tube width	C :	pump pressure
D :	paint viscosity	E :	tube height	F :	heating temperature

They used a 2^{6-3}_{III} fractional factorial design with defining relation

$$I = BCD = ADE = ABCE = ABF = ACDF = BDEF = CEF \,.$$

Table 16.1 Paint thickness for the paint experiment

z_A	z_B	z_C	z_D	z_E	z_F	$y_{\mathbf{z}1}$	$y_{\mathbf{z}2}$	$y_{\mathbf{z}3}$	$y_{\mathbf{z}4}$	$\overline{y}_{\mathbf{z}\cdot}$	$s^2_{\mathbf{z}}$
1	−1	1	−1	−1	−1	1.09	1.12	0.83	0.88	0.9800	0.021400
−1	−1	1	−1	1	1	1.62	1.49	1.48	1.59	1.5450	0.004967
1	1	−1	−1	−1	1	0.88	1.29	1.04	1.31	1.1300	0.042867
−1	1	−1	−1	1	−1	1.83	1.65	1.71	1.76	1.7375	0.005825
−1	−1	−1	1	−1	1	1.46	1.51	1.59	1.40	1.4900	0.006467
1	−1	−1	1	1	−1	0.74	0.98	0.79	0.83	0.8350	0.010700
−1	1	1	1	−1	−1	2.05	2.17	2.36	2.12	2.1750	0.017633
1	1	1	1	1	1	1.51	1.46	1.42	1.40	1.4475	0.002358

Source Eibl et al. (1992). Reprinted with Permission from Journal of Quality Technology © 1992 ASQ, www.asq.org

The experimenters decided to ignore all interactions for this first experiment. Since they wanted to monitor the variation of the paint thickness, they took four observations on each of the 8 treatment combinations in the fraction. The data are shown in Table 16.1, with factor levels coded as −1 and 1. If the data were collected in a completely random order, model (16.2.5) can be fitted using the 32 individual observations.

Using $z_A, \ldots, z_F$ rather than $z_1, \ldots, z_6$ to denote the factor levels, the fitted first-order model for the mean response is

$$\begin{aligned}\hat{y}_{\mathbf{z}} &= \hat{\gamma}_0 + \hat{\gamma}_A z_A + \cdots + \hat{\gamma}_F z_F \\ &= 1.42 - 0.32 z_A + 0.21 z_B + 0.12 z_C + 0.07 z_D - 0.03 z_E - 0.01 z_F\,,\end{aligned}$$

where, for example, the parameter estimate $\hat{\gamma}_D$ is calculated as

$$\hat{\gamma}_D = (\overline{y}_{z_D(+1)} - \overline{y}_{z_D(-1)})/2 = (1.493125 - 1.348125)/2 = 0.0725 \approx 0.07\,,$$

where $\overline{y}_{z_D(+1)}$ is the average of the observations with D at its high level and $\overline{y}_{z_D(-1)}$ is the average of the observations with D at its low level. □

16.2.4 Checking Model Assumptions

Before progressing with the analysis of the fitted model, the model assumptions should be checked. We shall discuss tests for model lack of fit in Sect. 16.2.6.

If there is no model lack of fit, then the error assumptions may be checked using residual plots. If the observations were collected sequentially in a known run order, then the residuals are plotted against run order to check for independence of observations. Residuals are plotted against predicted values to assess equality of error variances. Normality is checked by plotting residuals versus normal scores.

16.2.5 Analysis of Variance

Suppose that a standard first-order design has been used, with the extreme levels of each factor coded as −1 and +1. Under the first-order model, it follows from Eq. (16.2.6) that

Table 16.2 Analysis of variance for a first order model for the paint experiment

Source of variation	Degrees of freedom	Sum of squares	Mean square	Ratio	p-value	Expected mean square
A	1	3.2640	3.2640	242.07	0.0001	$\sigma^2 + 32\gamma_A^2$
B	1	1.3448	1.3448	99.73	0.0001	$\sigma^2 + 32\gamma_B^2$
C	1	0.4560	0.4560	33.82	0.0001	$\sigma^2 + 32\gamma_C^2$
D	1	0.1540	0.1540	11.42	0.0024	$\sigma^2 + 32\gamma_D^2$
E	1	0.0221	0.0221	1.64	0.2127	$\sigma^2 + 32\gamma_E^2$
F	1	0.0066	0.0066	0.49	0.4902	$\sigma^2 + 32\gamma_F^2$
Error	25	0.3371	0.0135			σ^2
Total	31	5.5846				
Computational formulae						
$ssi = n_f\hat{\gamma}_i^2 = n_f(\overline{y}_{z_i(+1)} - \overline{y}_{z_i(-1)})^2/4$, for $i = A, B, C, \ldots$						
ssE by subtraction		$sstot = \sum_{\mathbf{z}}\sum_t y_{\mathbf{z}t}^2 - n\overline{y}_\cdot^2$				

$$\mathrm{Var}(\hat{\gamma}_i) = \left(\frac{\sigma^2}{n_f/2} + \frac{\sigma^2}{n_f/2}\right)/4 = \sigma^2/n_f\,,$$

for any $i = A, B, C, \ldots$. The sum of squares for testing that the main-effect contrast γ_A is zero (that is, $H_0^A : \gamma_A = 0$ against $H_A^A : \gamma_A \neq 0$) is

$$ssA \;=\; \hat{\gamma}_A^2/(1/n_f) \;=\; n_f\hat{\gamma}_A^2\,,$$

and since there is only one degree of freedom for the A contrast, $msA = ssA$. For the first-order model and a standard first-order design, we have expected mean square

$$E[MSA] \;=\; n_f E[\hat{\gamma}_A^2] \;=\; n_f \mathrm{Var}(\hat{\gamma}_A) + n_f(E[\hat{\gamma}_A])^2 \;=\; \sigma^2 + n_f\gamma_A^2\,.$$

It can also be shown that $msE = ssE/(n - p - 1)$ is an unbiased estimate of σ^2, where ssE is obtained by subtraction in the analysis of variance table. Consequently, the decision rule for testing H_0^A against H_A^A is

$$\text{reject } H_0^A \quad \text{if} \quad msA/msE \;>\; F_{1,n-p-1,\alpha}\,.$$

Similar formulae hold for each main effect. The analysis of variance for the first-order model and a standard first-order design will be illustrated for the paint experiment.

Example 16.2.2 `Paint experiment, continued`

The paint experiment was discussed in Example 16.2.1, and the data were given in Table 16.1. The purpose of the experiment was to study the effects of six factors on paint thickness. The experimental design consisted of four observations on each of the treatment combinations of a 2_{III}^{6-3} design, which is an orthogonal factorial design with $n_f = 32$ factorial points and no center points. The analysis of variance for a first order model is shown in Table 16.2, together with the expected mean squares. The linear effect of each of factors A, B, C, and D is significantly different from zero, but factors E and F appear to have little effect on the response. □

16.2.6 Tests for Lack of Fit

A first-order design allows the experimenter to determine when the first-order model is no longer adequate, provided that there are more design points than first-order model parameters, and the design includes replication at one or more points. There is said to be model *lack of fit* when the model does not adequately represent the mean response as a function of the factor levels. Lack of fit of the first-order model occurs when the local response surface is no longer a plane.

Generic Test

Let n_d denote the number of *distinct* coded treatment combinations $\mathbf{z}$. For each treatment combination for which there is replication, the sample variance $s_{\mathbf{z}}^2$ of the $n_{\mathbf{z}}$ observations at that treatment combination provides an unbiased estimate of the error variance σ^2, whether or not there is model lack of fit. These sample variances can be pooled together to obtain a *sum of squares for pure error*

$$ssPE = \sum_{\mathbf{z}} (n_{\mathbf{z}} - 1) s_{\mathbf{z}}^2 \tag{16.2.7}$$

with $n - n_d$ degrees of freedom, giving a *mean square for pure error*

$$msPE = ssPE/(n - n_d)\,,$$

with $E[msPE] = \sigma^2$. The error sum of squares ssE with $n - p - 1$ degrees of freedom is obtained from fitting the first-order model (Table 16.2), and the difference

$$ssLOF = ssE - ssPE \tag{16.2.8}$$

is called the *sum of squares for lack of fit*. The corresponding mean square is

$$msLOF = ssLOF/(n_d - p - 1)\,.$$

The expected mean square is $E[msLOF] = \sigma^2 + \theta^2$, where θ^2 is a quadratic function of any higher order parameters that are estimable due to the inclusion of more design points than needed to fit the first order model. Then the ratio

$$msLOF/msPE$$

is used to test the null hypothesis of no model lack of fit. The null hypothesis is rejected at level α if this ratio exceeds $F_{n_d-p-1,n-n_d,\alpha}$. This lack-of-fit test is summarized in Table 16.3.

Example 16.2.3 `Paint experiment, continued`

The paint experiment was described in Example 16.2.1. The analysis of variance for the first-order model was shown in Table 16.2, giving $ssE = 0.3371$ with 25 degrees of freedom. There were $n_{\mathbf{z}} = 4$ observations at each of eight factorial points, and the corresponding eight sample variances, each with three degrees of freedom, were given in Table 16.1, p. 570. These eight sample variances can be pooled together to obtain

$$ssPE = \sum_{\mathbf{z}} (4 - 1) s_{\mathbf{z}}^2 = 0.3367$$

Table 16.3 Generic lack-of-fit test for the first-order model

Source of variation	Degrees of freedom	Sum of squares	Mean square	Ratio	Expected mean square
Lack of fit	$n_d - p - 1$	$ssLOF$	$msLOF$	$msLOF/msPE$	$\sigma^2 + \theta^2$
Pure Error	$n - n_d$	$ssPE$	$msPE$		σ^2
Error	$n - p - 1$	ssE			
	Computational formulae				
ssE from Table 16.2, $ssPE = \sum_{\mathbf{z}}(n_{\mathbf{z}} - 1)s^2_{\mathbf{z}}$, $ssLOF$ by subtraction,					
n_d distinct design points, n observations total, p factors,					
θ depends on the nature of estimable model lack of fit					

Table 16.4 Generic lack-of-fit test for the paint experiment

Source of variation	Degrees of freedom	Sum of squares	Mean square	Ratio	p-value
Lack of fit	1	0.0004	0.0004	0.03	0.8594
Pure error	24	0.3367	0.0140		
Error	25	0.3371	0.0135		

based on $n - n_d = 32 - 8 = 24$ degrees of freedom. The sum of squares for lack of fit is

$$ssLOF = ssE - ssPE = 0.3371 - 0.3367 = 0.0004\,,$$

and the test is summarized in Table 16.4. Since the p-value is large, there is no evidence of lack of fit of the first-order model. □

Test for Second-Order Lack of Fit

If the generic test indicates lack of fit of the first-order model, this provides no insight into why the model is not fitting well. To understand the nature of the lack of fit, it can be helpful to consider what the mean square for lack of fit measures in terms of higher-order models. If the first-order model is inadequate, the next possibility is that a second-order model would provide an adequate approximation to the local response surface. If so, then lack of fit of the first-order model is attributable to the presence of either two-factor interactions or to quadratic effects or to both.

If the only lack of fit is due to two-factor interaction effects, this corresponds to a twisting of the response surface. Such lack of fit can be tested if the first-order design allows estimation of two-factor interactions in addition to providing error degrees of freedom. In the paint experiment, for example, it is possible to estimate the AC interaction effect, in addition to the six main effects, provided that all other interaction effects are known to be negligible.

If the center of the experimental design is near the peak of the response surface, then one would expect quadratic effects, or curvature, to be present and a higher mean response near the design center than at the factorial points. Multiple center points $\mathbf{z}_0 = (0, \ldots, 0)$ are usually included in a first-order design, because comparison of the mean response at the center of the design region with the mean response at the factorial points provides an effective test for lack of fit due to quadratic effects.

So, to assess second-order lack of fit we fit a second-order polynomial regression model under the alternative hypothesis. With respect to the coded factor levels, the standard *second-order model* for p factors is

$$Y_{\mathbf{z},t} = \gamma_0 + \sum_i \gamma_i z_i + \sum_i \gamma_{ii} z_i^2 + \sum_{i<j} \gamma_{ij} z_i z_j + \epsilon_{\mathbf{z},t} ,$$

where the parameter γ_i represents the linear effect of the ith factor, γ_{ii} represents the quadratic effect of the ith factor, and γ_{ij} represents the cross product effect between the ith and jth factors.

If the factorial portion of the standard first-order design is either a complete factorial design or a fraction of resolution V or higher, then all two-factor interaction parameters γ_{ij} in the second-order model are estimable (assuming higher-order interactions to be negligible). For testing for second-order lack of fit, we add the sums of squares for these two-factor interactions to obtain a pooled interaction sum of squares, *ssI*. If the factorial portion of the design is a fraction of resolution less than V, then not all two-factor interactions are estimable, and only the sums of squares of those two-factor interactions which are not aliased with main effects may be pooled—one sum of squares from each alias set.

The quadratic-effect parameters are not individually estimable from a standard first-order design. They are aliased with one another, and only their sum can be estimated. It can be shown that

$$E[\overline{Y}_f - \overline{Y}_0] = \sum_{i=1}^{p} \gamma_{ii} ,$$

with

$$\text{Var}(\overline{Y}_f - \overline{Y}_0) = \left(\frac{1}{n_0} + \frac{1}{n_f}\right)\sigma^2 = \frac{n}{n_f n_0}\sigma^2 ,$$

where $\overline{Y}_f$ and $\overline{Y}_0$ denote the average of the n_f factorial points and the average of the n_0 center points, respectively. It follows that the corresponding sum of squares for testing whether or not the sum of the quadratic parameters is zero is

$$ssQ = \frac{n_f n_0}{n}(\overline{Y}_f - \overline{Y}_0)^2 ,$$

with one degree of freedom. The expected mean square is

$$E[MSQ] = \sigma^2 + \frac{n_f n_0}{n}(\sum_{i=1}^{p} \gamma_{ii})^2 .$$

In the generic test for lack of fit of the first-order model, *ssI* and *ssQ* are part of *ssLOF*. Thus, we can write

$$ssLOF = ssI + ssQ + ssH ,$$

where *ssH* is the sum of squares for lack of fit due to a higher-order model. Then lack of fit due specifically to interaction terms and quadratic terms can be investigated separately. The tests are summarized in Table 16.5 for a standard first-order design.

For all tests for lack of fit, an adequate number of pure error degrees of freedom are needed for the test power to be reasonably high. Since $\text{Var}(\overline{Y}_f - \overline{Y}_0) > \sigma^2/n_0$, the test for lack of fit due to quadratic effects will have low power if there are few center points. Typically, 3–6 center points would be used.

Example 16.2.4 Acid copper pattern plating experiment

Poon (1995) conducted a sequence of fractional factorial and response surface experiments each involving as many as seven factors to minimize the coating thickness variation of an acid copper-plating

Table 16.5 Lack-of-fit test for the first-order model, given the data of a standard first-order design, with p factors $A, B, \ldots$ and m alias sets for interaction effects clear of main effects

Source of variation	Degrees of freedom	Sum of squares	Mean square	Ratio	Expected mean square
Interaction	m	$ssI = ssAB + \cdots$	msI	$\frac{msI}{msPE}$	$\sigma^2 + \frac{n_f}{m}\theta_1$
Quadratic	1	ssQ	msQ	$\frac{msQ}{msPE}$	$\sigma^2 + \frac{n_0 n_f}{n}\theta_2^2$
Higher-order	$n_d - p - m - 2$	ssH			
Pure Error	$n - n_d$	$ssPE$	$msPE$		σ^2
Error	$n - p - 1$	ssE			
	Computational formulae				
$ssAB = n_f \hat{\gamma}_{AB}^2 = n_f(\overline{y}_{z_A z_B(+1)} - \overline{y}_{z_A z_B(-1)})^2/4$				ssE from Table 16.2	
$ssQ = (n_0 n_f/n)^2(\overline{y}_f - \overline{y}_0)^2$				$\theta_1 = \gamma_{AB}^2 + \cdots$	
$ssPE = \sum_{\mathbf{z}}(n_{\mathbf{z}} - 1)s_{\mathbf{z}}^2$				$\theta_2 = \gamma_{AA} + \gamma_{BB} + \cdots$	
$ssH = (ssE - ssPE) - ssI - ssQ$					

Table 16.6 Data for the acid copper pattern plating experiment

Anode–cathode separation (in.)		Current density (asf)		Standard deviation (μm)
Coded	Uncoded	Coded	Uncoded	
−1	9.5	−1	31	5.60
−1	9.5	1	41	6.45
1	11.5	−1	31	4.84
1	11.5	1	41	5.19
0	10.5	0	36	4.32
0	10.5	0	36	4.25

Source Poon (1995). Reprinted with permission

process. In the final experiments, conducted in the vicinity of minimum thickness variation, response surface methods were utilized to study the effects of anode-cathode separation (factor A) and cathodic current density (factor B) on the standard deviation of coating thickness. One experiment used the factorial points of a single replicate 2^2 design, augmented by two center points. The response was the standard deviation (in μm) of copper-plating thickness. The coded and uncoded factor levels, together with the resulting data, are given in Table 16.6.

The midrange of levels of factor A is $(11.5+9.5)/2 = 10.5$, and the half-range is $(11.5-9.5)/2.0 = 1.0$. So the coded levels are given by

$$z_A = x_A - 10.5\,.$$

The midrange and half-range of the factor B levels are $(41 + 31)/2 = 36$ and $(41 - 31)/2 = 5$, respectively, so the coded levels of factor B are

$$z_B = (x_B - 36)/5\,.$$

Table 16.7 shows the analysis of variance, including tests for lack of fit, due to a second-order model. The analyses are identical for coded and uncoded factor levels. There are significant quadratic effects—an indication that quadratic terms for either or both of factors A and B are needed to adequately model the response surface. The first-order design is inadequate, then, because not all parameters in

Table 16.7 Analysis of variance and lack-of-fit test for the acid copper pattern plating experiment

Source of variation	Degrees of freedom	Sum of squares	Mean square	Ratio	p-value	Expected mean square
A	1	1.0201	1.0201	1.46	0.3137	$\sigma^2 + n_f\gamma_A^2$
B	1	0.3600	0.3600	0.51	0.5250	$\sigma^2 + n_f\gamma_B^2$
Error	3	2.0986	0.6995			
Total	5	3.4787				
Interaction AB	1	0.0625	0.0625	25.51	0.1244	$\sigma^2 + n_f\gamma_{AB}^2$
Quadratic	1	2.0336	2.0336	830.05	0.0221	$\sigma^2 + \frac{n_0 n_f}{n}\theta^2$
Pure error	1	0.0025	0.0025			
Error	3	2.0986	0.6995			
where $\theta = \gamma_{AA} + \gamma_{BB}$						

the second-order model are estimable. The solution is to collect some additional observations, as will be illustrated in Example 16.3.1. □

16.2.7 Path of Steepest Ascent

If there are significant linear effects and there is no significant lack of fit of the first-order model, then the *path of steepest ascent* may be followed to climb towards the maximum of the response surface.

Given the fitted first-order regression model (16.2.5), the path of steepest ascent from the current position $\mathbf{z}_a$ is determined as follows. If $\hat{\gamma}_i$ is positive, increase z_i to increase predicted mean response $\hat{y}_\mathbf{z}$. If $\hat{\gamma}_i$ is negative, decrease z_i to increase $\hat{y}_\mathbf{z}$. To follow the path of steepest ascent up the fitted response surface, change each z_i in proportion to the magnitude of $\hat{\gamma}_i$. So, if the value z_1 of the first factor is changed by $u\hat{\gamma}_1$ units for some real number u, then the level z_i of the ith factor should be changed by $u\hat{\gamma}_i$ for each other factor i.

The path of steepest ascent is defined above with respect to the coded variables. This presumes that the original variables have been coded in such a way to make the coded scales in some sense comparable. Since the original variable may be measured on scales that are not directly comparable, there is some art to the scaling of the coded variables.

Example 16.2.5 `Paint experiment, continued`

The paint experiment was described in Example 16.2.1, p. 569. The experimenters conducted an experiment to study how to decrease the thickness of a paint coating from about 2 mm to the target 0.8 mm. Four observations were taken at each treatment combination of a 2^{6-3}_{III} design and are shown in Table 16.1, p. 570.

The target thickness is approximately achieved at the experimental design point $\mathbf{z} = (+1, -1, -1, +1, +1, -1)$ so perhaps no further analysis or experimentation is needed. Nevertheless, we will use these data to illustrate how to move efficiently towards the lower target response surface value.

Since a lower mean response is required, we need to identify the path of steepest *descent*. The fitted first-order model is obtained from Example 16.2.1 as

$$\begin{aligned}\hat{y}_\mathbf{z} &= \hat{\gamma}_0 + \hat{\gamma}_A z_A + \cdots + \hat{\gamma}_F z_F \\ &= 1.42 - 0.32 z_A + 0.21 z_B + 0.12 z_C + 0.07 z_D - 0.03 z_E - 0.01 z_F\,.\end{aligned}$$

The analysis of variance conducted in Example 16.2.2 suggests that only factors A, B, C, and D significantly affect the response. So, these four factors should be adjusted in an attempt to reduce paint thickness.

Based on the signs of the parameter estimates in the fitted model, we ought to be able to effect a reduction in mean response if we increase the level of factor A and decrease the level of any of factors B, C, and D. To follow the path of steepest descent, we change the levels of these factors each in proportion to the magnitude of its corresponding parameter estimate, $\hat{\gamma}_i$. So, if we increase z_A by $0.32u$ units for some real number u, then we decrease z_B by $0.21u$ units, decrease z_C by $0.12u$ units, and decrease z_D by $0.07u$ units.

Observations along the path of steepest descent moving away from the center of the current design, $\mathbf{z}_0 = (0, 0, 0, 0, 0, 0)$, consist of treatment combinations $(0.32u, -0.21u, -0.12u, -0.07u, 0, 0)$ corresponding to increasing values of u, such as $u = 3, 3.5, 4, \ldots$. The suggested values of u start at $u = 3$. This value is large enough to move the level of factor A near to the edge of the region of the current local experiment and corresponds to $\hat{y} = 0.9226$. For the value $u = 4$, the extrapolated prediction of the first order model is $\hat{y} = 0.7568$, already less than the target value of 0.8, making the step sizes reasonable or perhaps a bit too large. Certainly, other values of u could also have been chosen. Observations may then be collected along this path setting u equal to each value in turn until the target thickness is achieved, or until the response stops decreasing before reaching the target level. In the latter case, at the point of lowest response along the path another first-order design could be run to determine a new path of steepest descent. □

In the previous example, the effects of factors E and F were not found to be significantly different from zero, so their levels were not changed in following the estimated path of steepest descent. There are a variety of reasons why the effect of a factor may be negligible. The obvious reason is that response is independent of the factor. However, it could also be that the levels used for the factor may be near the optimal value, so the response surface may be relatively flat with respect to small changes in the level of that factor. Alternatively, the levels of the factor may simply be too close together to give rise to a detectable change in the mean response. In subsequent experiments, the levels of such factors can be chosen farther apart to guard against the last scenario.

16.3 Second-Order Designs and Analysis

16.3.1 Models and Designs

Second-order designs and analysis are used when the test for lack of fit of the first-order model indicates that the vicinity of the maximum (or minimum) of the response surface has been reached and a second-order model should be fitted. For p factors, the standard second-order model is

$$Y_{\mathbf{x},t} = \beta_0 + \sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p} \beta_{ii} x_i^2 + \sum_{i<j} \beta_{ij} x_i x_j + \epsilon_{\mathbf{x},t}\,, \tag{16.3.9}$$

where $Y_{\mathbf{x},t}$ denotes the tth response observed for treatment combination $\mathbf{x} = (x_1, x_2, \ldots, x_p)$. The random-error variables $\epsilon_{\mathbf{x},t}$ are assumed to be independent with $N(0, \sigma^2)$ distributions. The parameter β_i represents the linear effect of the ith factor. The parameter β_{ii} represents the quadratic effect of the ith factor, and β_{ij} represents the cross product effect, or interaction effect, between the ith and jth factors.

With respect to the coded factor levels $z_i = (x_i - m_i)/h_i$, the second-order model is

$$Y_{\mathbf{z},t} = \gamma_0 + \sum_{i=1}^{p} \gamma_i z_i + \sum_{i=1}^{p} \gamma_{ii} z_i^2 + \sum_{i<j} \gamma_{ij} z_i z_j + \epsilon_{\mathbf{z},t} \,. \tag{16.3.10}$$

Experimental designs used to fit a second-order model are referred to as *second-order designs*. A second-order design should (i) allow for efficient estimation of the response surface, in the sense of having $\mathrm{Var}(\widehat{Y}_{\mathbf{z}})$ be small in the design region; (ii) allow a test for lack of fit of the second-order model; and (iii) allow for efficient estimation of all model parameters. Second-order designs must have at least $(p+1)(p+2)/2$ distinct design points; otherwise, not all of the $(p+1)(p+2)/2$ parameters in the second-order model can be estimated. We will consider only such designs in this chapter. Observations at even more points are needed, plus some replication, in order to be able to conduct a generic test for model lack of fit. Other properties of second-order designs that are sometimes desirable include rotatability, orthogonality, and orthogonal blocking—these will be discussed in Sects. 16.4.1–16.4.3.

The method of least squares is used to fit the second-order model to the data. This method is exactly as discussed in optional Sect. 8.3, with each second-order term z_i^2 or $z_i z_j$ being treated as a single regressor. In terms of the uncoded and coded factor levels, the fitted models are, respectively,

$$\hat{y}_{\mathbf{x}} = \hat{\beta}_0 + \sum_i \hat{\beta}_i x_i + \sum_i \hat{\beta}_{ii} x_i^2 + \sum_{i<j} \hat{\beta}_{ij} x_i x_j \tag{16.3.11}$$

and

$$\hat{y}_{\mathbf{z}} = \hat{\gamma}_0 + \sum_i \hat{\gamma}_i z_i + \sum_i \hat{\gamma}_{ii} z_i^2 + \sum_{i<j} \hat{\gamma}_{ij} z_i z_j \,, \tag{16.3.12}$$

where the parameters with hats denote the least squares estimates. Although it is possible to obtain explicit formulae for the least squares estimates for any specific design, the formulae for the quadratic parameter estimates $\hat{\gamma}_{ii}$ are complicated. Consequently, we rely on statistical computer software to obtain the least squares estimates (see Sects. 16.7 and 16.8 for the use of the SAS and R software, respectively).

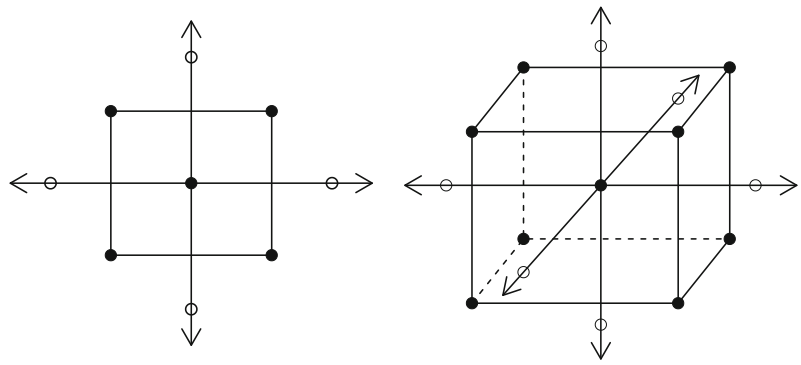
As long as there is no significant lack of fit, the fitted second-order model can be used to study the local response surface. Generally, there will be a unique treatment combination $\mathbf{x}_s = (x_{s1}, x_{s2}, \ldots, x_{sp})$, called the *stationary point*, at which the fitted surface $\hat{y}_{\mathbf{x}}$ is neither increasing or decreasing—the tangent plane is level. At the stationary point, $\hat{y}_{\mathbf{x}}$ is maximized, minimized, or is at a saddle point. The surface near a *saddle point* is reminiscent of a horse saddle—rising up from front to back but sloping down from side to side. A saddle point yields neither a maximum nor a minimum for the fitted model. Instead, these will be found at the boundary of the design region.

If there is significant lack of fit of the second-order model, a higher-order model could be used, or a more local experiment could be run.

16.3.2 Central Composite Designs

Central composite designs were first described by Box and Wilson (1951), and they are nowadays the most popular second-order designs. Each design consists of a standard first-order design with n_f orthogonal factorial points and n_0 center points, augmented by n_a "axial points."

We follow the convention of coding the factor levels so the factorial points have coded levels ± 1 for each factor. However, it should be noted that some software packages will recode the levels in a central



(a) $p = 2$ **(b)** $p = 3$

Fig. 16.3 Central composite designs for $p = 2$ and $p = 3$ factors

composite design before doing the analysis. In SAS, for example, the default is to code the extreme levels of each factor as ± 1, whereas R allows the user to specify the coding. Under our convention, *axial points* are points located at a specified distance α from the design center in each direction on each axis defined by the coded factor levels. On the z_i-axis, for example, two axial points are obtained by setting $z_i = \pm\alpha$, with $z_j = 0$ for all $j \neq i$. Thus, if there are p factors, there are $2p$ distinct axial points. Axial points are also commonly referred to as *star points*. Figure 16.3 shows central composite designs for $p = 2$ and $p = 3$ factors, with axial points represented by unfilled circles or balls.

A central composite design is easily built up from a standard first-order design by the addition of axial points, and possibly some extra factorial and center points. If the factorial portion of the design is a complete factorial or a fractional factorial of resolution V or more, all parameters of the second-order model are estimable. Otherwise, some aliasing will occur, and some terms will need to be omitted from the second-order model. A design should include enough replication, often at the center points, to allow for a test for model lack of fit. The axial points are located at a distance α from the center of the design, where the choice of α depends on the properties required of the design. A popular choice is $\alpha = (n_f)^{1/4}$ (see Sect. 16.4.1).

Example 16.3.1 `Acid copper pattern plating experiment, continued`

In Example 16.2.4, p. 574, a standard first-order design was used to study the effects of anode–cathode separation (factor A) and cathodic current density (factor B) on the standard deviation of a copper-plating thickness. The first-order design involved the $n_f = 4$ factorial points of a single-replicate 2^2 design, augmented by $n_0 = 2$ center points. There was significant lack of fit of the first-order model, so additional observations needed to be taken in order to fit a second-order model. The experimenters augmented the first-order design with four axial points, using $\alpha \approx (n_f)^{1/4} = \sqrt{2}$, say $\alpha = 1.4142$, giving the central composite design and data shown in Table 16.8.

The second-order model is fitted by a computer regression package. In terms of the uncoded factor levels, the fitted model is given by

$$\begin{aligned}\hat{y}_{\mathbf{x}} &= \hat{\beta}_0 + \hat{\beta}_A x_A + \hat{\beta}_B x_B + \hat{\beta}_{AA} x_A^2 + \hat{\beta}_{BB} x_B^2 + \hat{\beta}_{AB} x_A x_B \\ &= 84.1990 - 8.8689 x_A - 1.7526 x_B \\ &\quad + 0.4419 x_A^2 + 0.0286 x_B^2 - 0.0250 x_A x_B \,,\end{aligned}$$

and, in terms of the coded factor levels, $z_A = (x_A - 10.5)$, $z_B = (x_B - 36)/5$, the fitted model is

Table 16.8 Data for the acid copper pattern plating experiment—central composite design

Anode–cathode separation (in.)		Current density (asf)		Standard deviation (μm)
Coded	Uncoded	Coded	Uncoded	
−1.0000	9.5000	−1.0000	31.0000	5.60
−1.0000	9.5000	1.0000	41.0000	6.45
1.0000	11.5000	−1.0000	31.0000	4.84
1.0000	11.5000	1.0000	41.0000	5.19
0.0000	10.5000	0.0000	36.0000	4.32
0.0000	10.5000	0.0000	36.0000	4.25
−1.4142	9.0858	0.0000	36.0000	5.76
1.4142	11.9142	0.0000	36.0000	4.42
0.0000	10.5000	−1.4142	28.9290	5.46
0.0000	10.5000	1.4142	43.0710	5.81

Source Poon (1995). Reprinted with permission

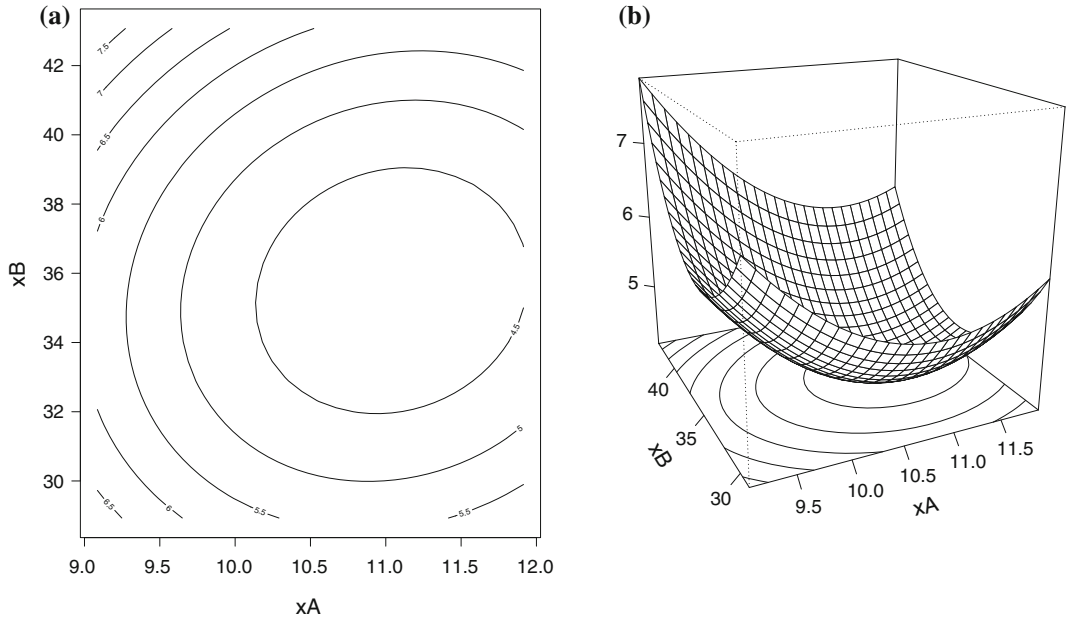


Fig. 16.4 Response surface contour plot and response surface plot of fitted second-order model for the acid copper pattern plating experiment

$$\begin{aligned}\hat{y}_{\mathbf{z}} &= \hat{\gamma}_0 + \hat{\gamma}_A z_A + \hat{\gamma}_B z_B + \hat{\gamma}_{AA} z_A^2 + \hat{\gamma}_{BB} z_B^2 + \hat{\gamma}_{AB} z_A z_B \\ &= 4.2850 - 0.4894 z_A + 0.2119 z_B \\ &\quad + 0.4419 z_A^2 + 0.7144 z_B^2 - 0.1250 z_A z_B \,.\end{aligned}$$

Figure 16.4 shows both a contour plot and a surface plot of the fitted model for uncoded factor levels. The stationary point is in the center of the ellipses. Clearly, the stationary point provides a minimum. The exact location of the stationary point will be determined in Sect. 16.3.5. □

16.3.3 Generic Test for Lack of Fit of the Second-Order Model

If the second-order design includes n_d distinct treatment combinations, with n_d larger than the number of parameters $(p+2)(p+1)/2$, and replication at one or more of these, then a generic test for lack of fit of the second-order model can be conducted, just as for the first-order model (Sect. 16.2.6). The sum of squares for pure error, *ssPE*, and the sum of squares for lack of fit, *ssLOF*, are calculated as in (16.2.7) and (16.2.8). The error sum of squares *ssE* and the error degrees of freedom are obtained from the analysis of variance table of the second-order model. The test proceeds exactly as in Table 16.3 except that the error degrees of freedom are $n - [(p+2)(p+1)/2]$ and the degrees of freedom for lack of fit are then $n_d - [(p+2)(p+1)/2]$. The test will be illustrated for the acid copper-plating experiment in Example 16.3.2 in the next subsection.

16.3.4 Analysis of Variance for a Second-Order Model

Table 16.9 shows an outline analysis of variance table for a central composite design and second-order model, assuming that all parameters are estimable. The degrees of freedom associated with the linear effects have been added (pooled) together, as have those of the quadratic effects and those of the interaction (cross product) effects. Sequential, or Type I, sums of squares are listed for each of these pooled sources of variation. These include the sum of squares for all linear terms, $ss(L)$; the sum of squares for adding all quadratic terms to the model, given that all linear terms are already included, $ss(Q|L)$; and the sum of squares for adding all interaction terms to the model, given that all linear and quadratic terms are already in the model, $ss(I|L, Q)$. Using these sequential sums of squares, the analysis of variance is the same whether factor levels are coded or not. The coefficients a_i, a_{ii}, and a_{ij} listed in the expected mean squares are positive and depend on the design and the model. If coded factor levels are used, the expected mean squares would involve the parameters γ instead of the parameters β, but would have the same form.

If a central composite design is used and factor levels are coded in the usual way, the linear, quadratic and interaction sums of squares are independent of one another, so the corresponding sums of squares are the same, no matter in which order the terms are fitted. Also, the individual linear and interaction (cross product) parameters are estimated independently of one another and of the quadratic effects. The quadratic parameters are estimated independently of each other only if α and the number of center points n_0 are chosen to satisfy certain restrictions (see Sect. 16.4.2).

Table 16.9 Analysis of variance for a central composite design and second-order model

Source of variation	Degrees of freedom	Sum squares of (Type I)	Mean square (Type I)	Ratio	Expected mean square
L	p	ssL	msL	$\frac{msL}{msE}$	$\sigma^2 + \sum_i a_i \beta_i^2$
$Q\|L$	p	$ss(Q\|L)$	$ms(Q\|L)$	$\frac{ms(Q\|L)}{msE}$	$\sigma^2 + \sum_i a_{ii} \beta_{ii}^2$
$I\|L, Q$	$\frac{1}{2}p(p-1)$	$ss(I\|L, Q)$	$ms(I\|L, Q)$	$\frac{ms(I\|L,Q)}{msE}$	$\sigma^2 + \sum_{i<j} a_{ij} \beta_{ij}^2$
Error	df	ssE	msE		σ^2
Total	$n-1$	$sstot$			
Formula: $df = n - \frac{1}{2}(p+2)(p+1)$					

Table 16.10 Analysis of variance for the acid copper pattern plating experiment

Source of variation	Degrees of freedom	Sum of squares	Mean square	Ratio	p-value
Linear	2	2.2751			
A_L	1	1.9159	1.9159	65.99	0.0012
B_L	1	0.3591	0.3591	12.37	0.0245
Quadratic	2	2.4361			
$A_Q \mid B_Q$	1	0.8926	0.8926	30.74	0.0052
$B_Q \mid A_Q$	1	2.3330	2.3330	80.36	0.0009
Interaction	1	0.0625	0.0625	2.15	0.2162
Error	4	0.1161	0.0290		
Total	9	4.8898			

Example 16.3.2 `Acid copper pattern plating experiment, continued`

The data for the central composite design of the acid copper pattern plating experiment were shown in Table 16.8, p. 580. The analysis of variance for the coded data is given in Table 16.10. The table shows the decomposition of the linear sum of squares with respect to the individual linear effects. Each of the two quadratic effects is shown adjusted for the other quadratic effect. If we test each hypothesis at individual level 0.01, the linear effect of factor A is significantly different from zero, as is the adjusted quadratic effect of each factor. Consequently, the model should include these three terms. We would also include the linear effect of B, since the higher-order (quadratic) term is included. The AB-interaction effect, or cross product effect, is not significantly different from zero.

Before settling on a final model, we should check the lack of fit of the second-order model. The only replication consisted of two center-point observations with values 4.32 and 4.25. The sample variance of these two observations is $s_0^2 = 0.00245$, so $ssPE = 0.00245$ with one degree of freedom. From the analysis of variance table, Table 16.10, we see that $ssE = 0.1161$ with 4 degrees of freedom. So,

$$ssLOF = ssE - ssPE = 0.1161 - 0.00245 = 0.11365$$

with $4 - 1 = 3$ degrees of freedom for lack of fit. There is significant lack of fit of the second-order model if

$$msLOF/msPE > F_{3,1,\alpha}\,,$$

for appropriate significance level α. Here,

$$msLOF/msPE = (0.11365/3)/(0.00245/1) = 15.463\,,$$

which is less than $F_{3,1,0.10} = 53.6$, so there is no significant lack of fit of the second-order model, and the model fitted in Example 16.3.1, p. 579, should be a reasonable approximation to the true surface in the local region under study ($9.5 \le x_A \le 11.5;\ 31 \le x_B \le 41$). □

16.3.5 Canonical Analysis of a Second-Order Model

After fitting a second-order model, we need to (i) determine the location of the stationary point and (ii) characterize the stationary point as providing a response surface minimum, maximum, or saddle point. The nature of the response surface at the stationary point may be evident from contour or surface plots, as is the case in Fig. 16.4, or its characterization may be done via canonical analysis. We provide an overview and illustration of canonical analysis in this section, leaving the computations to software (see Sects. 16.7.2 and 16.8.2 for the use of the SAS and R software, respectively).

In response surface methods, it is customary to perform the canonical analysis using the model fit to the coded data. We think of each coded treatment combination $\mathbf{z}$ as a point in p-dimensional space, $\mathbf{z} = (z_1, z_2, \ldots, z_p)$. Then the stationary point that we are trying to find is the point $\mathbf{z}_s = (z_{s1}, z_{s2}, \ldots, z_{sp})$ at which the fitted response surface $\hat{y}_\mathbf{z}$ is neither increasing nor decreasing—the tangent plane is level. The stationary point can be obtained via calculus as the critical point of the fitted surface $\hat{y}_\mathbf{z}$. The stationary point $\mathbf{x}_s$ for the model fit to the uncoded data can be obtained from $\mathbf{z}_s$ by simply uncoding each factor level z_{si}. In view of Eq. (16.2.2), this uncoding is accomplished by taking $x_{si} = h_i \times z_{is} + m_i$, for $i = 1, 2, \ldots, p$.

The second step—characterizing the response surface at the stationary point as a minimum, maximum, or saddle point—may be accomplished by putting the fitted second-order response surface into canonical form. To accomplish this, we change to a new coordinate system of points in two steps. First we set $\mathbf{v} = \mathbf{z} - \mathbf{z}_s$, so that $v_i = z_i - z_{si}$ for $i = 1, 2, \ldots, p$. This moves the coordinate system so that the stationary point is at the origin with respect to the v_i-axes, so the stationary point is now $\mathbf{v}_s = (0, 0, \ldots, 0)$. The other points $\mathbf{v} = \mathbf{z} - \mathbf{z}_s$ measure position relative to the stationary point $\mathbf{z}_s$. This eliminates all linear terms from the model. As the second step, the v_i-axes are rotated to obtain w_i-axes, with the rotation chosen to eliminate the cross product terms from the model.

In terms of each of these coordinate systems, the fitted model has the following equivalent representations:

$$\hat{y}_\mathbf{z} = \hat{\gamma}_0 + \sum_{i=1}^{p} \hat{\gamma}_i z_i + \sum_{i=1}^{p} \hat{\gamma}_{ii} z_i^2 + \sum_{i<j} \hat{\gamma}_{ij} z_i z_j \,,$$

$$\hat{y}_\mathbf{v} = \hat{y}_{\mathbf{v}_s} + \sum_{i=1}^{p} \hat{\gamma}_{ii} v_i^2 + \sum_{i<j} \hat{\gamma}_{ij} v_i v_j \,,$$

$$\hat{y}_\mathbf{w} = \hat{y}_{\mathbf{w}_s} + \sum_{i=1}^{p} \hat{\lambda}_{ii} w_i^2 \,,$$

where $\hat{y}_{\mathbf{v}_s}$ and $\hat{y}_{\mathbf{w}_s}$ are equal and each denotes the predicted response at the stationary point.

The last equation is said to be in *canonical form*, and in this form, we can immediately tell whether the stationary point is a maximum, a minimum, or a saddle point. If all of the $\hat{\lambda}_{ii}$'s are negative, then the fitted model is concave down and has a maximum at the stationary point. If all of the $\hat{\lambda}_{ii}$'s are positive, then the fitted model is concave up and has a minimum at the stationary point. If some of the $\hat{\lambda}_{ii}$'s are positive and some are negative, the stationary point is a saddle point. The $\hat{\lambda}_{ii}$ are called the *canonical coefficients*.

If a specific $\hat{\lambda}_{ii}$ is relatively large in magnitude, then $\hat{y}_\mathbf{w}$ will change rapidly for changes away from the stationary point $\mathbf{w}_s = (0, 0, \ldots, 0)$ in the w_i direction. Thus, if the stationary point is a saddle point and if $\hat{\lambda}_{\ell\ell}$ is the largest positive $\hat{\lambda}_{ii}$, movement in either direction away from the stationary point along the w_ℓ-axis provides a path of steepest ascent. On the other hand, if a specific $\hat{\lambda}_{ii}$ is relatively

small in magnitude, then $\hat{y}_{\mathbf{w}}$ is relatively unaffected by changes away from the stationary point along the w_i-axis.

Example 16.3.3 `Acid copper pattern plating experiment, continued`

In Example 16.3.1, p. 579, a second-order model was fitted to data collected from a central composite design. The experiment was run in order to study the effects of anode-cathode separation (factor A) and cathodic current density (factor B) on the standard deviation of copper-plating thickness. In terms of the coded factor levels $z_A = (x_A - 10.5)$, $z_B = (x_B - 36)/5$, the fitted model was

$$\begin{aligned}\hat{y}_{\mathbf{z}} &= \hat{\gamma}_0 + \hat{\gamma}_A z_A + \hat{\gamma}_B z_B + \hat{\gamma}_{AA} z_A^2 + \hat{\gamma}_{BB} z_B^2 + \hat{\gamma}_{AB} z_A z_B \\ &= 4.2850 - 0.4894 z_A + 0.2119 z_B \\ &\quad + 0.4419 z_A^2 + 0.7144 z_B^2 - 0.1250 z_A z_B \,.\end{aligned}$$

The following additional results are provided without computational details, since we are leaving those to software (see Sects. 16.7.2 and 16.8.2).

The stationary point is $\mathbf{z}_s = (z_{sA}, z_{sB}) = (0.5395, -0.1011)$. Using these values of z_A and z_B in the fitted model, we obtain the predicted response at the stationary point to be $\hat{y}_{\mathbf{z}_s} = 4.1423$.

The canonical coefficients are $\hat{\lambda}_{11} = 0.7280$ and $\hat{\lambda}_{22} = 0.4282$. Since both canonical coefficients are positive, the stationary point minimizes the estimated standard deviation of coating thickness. Now, $\hat{\lambda}_{11}$ is larger than $\hat{\lambda}_{22}$—nearly twice as large—so the surface will rise more rapidly as we move away from $\mathbf{z}_s$ in the w_1 direction than in the w_2 direction.

The w_1 canonical axis consists of all points (z_A, z_B) of the form

$$(z_{sA},\ z_{sB}) = (0.5395, -0.1011) + u(-0.2134, 0.9770)\,.$$

The point $(-0.2134, 0.9770)$ has been scaled to be one unit from the origin (i.e. $(-0.2134, 0.9770)$ is a vector of length one), so a unit change in u corresponds to a step of size one along the w_1-axis. Since the second component of this point is nearly one, the w_1-axis is nearly parallel to the z_2-axis, or equivalently, to the B axis. This means that the coded level of B must be controlled more precisely than the coded level of A in order to maintain a minimum response. This conclusion is suggested by examining the fitted equation, since the coefficient of z_B^2 is somewhat larger than those of z_A^2 and $z_A z_B$.

Likewise, the w_2 canonical axis consists of all points (z_A, z_B) of the form

$$(z_{sA},\ z_{sB}) = (0.5395, -0.1011) + q(-0.9770, -0.2134)\,,$$

where a unit change in q corresponds to a step of size one along the w_2-axis. Since the first component of the point $(-0.9770, -0.2134)$ has magnitude nearly one, the w_2-axis is nearly parallel to the z_A-axis (or z_1-axis). □

The canonical analysis has been described and illustrated here in terms of the coded factor levels. The SAS and R software likewise do the canonical analysis in terms of coded factor levels, though SAS software codes the levels somewhat differently, which impacts the canonical coefficients.

Canonical Analysis Formulas (Optional)

This subsection requires the knowledge of matrices and vectors. Consider the fitted second-order model

$$\hat{y}_{\mathbf{z}} = \hat{\gamma}_0 + \sum_i \hat{\gamma}_i z_i + \sum_i \hat{\gamma}_{ii} z_i^2 + \sum_{i<j} \hat{\gamma}_{ij} z_i z_j$$

for p factors. Let $\mathbf{b}$ denote the $p \times 1$ vector of linear parameter estimates, with ith entry $\hat{\gamma}_i$. Let $\mathbf{B}$ denote the $p \times p$ matrix with ith diagonal element $\hat{\gamma}_{ii}$ and with off-diagonal (ij)th entry $\hat{\gamma}_{ij}/2$. Then the least squares fitted model can be written in matrix terms as

$$\hat{y}_{\mathbf{z}} = \hat{\gamma}_0 + \mathbf{z}'\mathbf{b} + \mathbf{z}'\mathbf{B}\mathbf{z}\,.$$

Furthermore, the stationary point is

$$\mathbf{z}_s = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b}\,,$$

with corresponding predicted mean response

$$\hat{y}_{\mathbf{z}_s} \;=\; \hat{\gamma}_0 - \mathbf{z}_s'\mathbf{B}\mathbf{z}_s \;=\; \hat{\gamma}_0 + \frac{1}{2}\mathbf{z}_s'\mathbf{b}\,.$$

The *canonical coefficients* $\hat{\lambda}_{ii}$ are the *eigenvalues* of the matrix $\mathbf{B}$. The *eigenvectors* of $\mathbf{B}$ determine the *canonical axes*, the canonical axis w_i being the normalized eigenvector of $\mathbf{B}$ corresponding to the eigenvalue $\hat{\lambda}_{ii}$. Obtaining the canonical coefficients and canonical axes using SAS and R software will be illustrated in Sects. 16.7.2 and 16.8.2, respectively.

16.4 Properties of Second-Order Designs: CCDs

In this section we discuss some desirable properties—rotatability, orthogonality, and orthogonal blocking—of second-order designs. The discussion here focuses on central composite designs (CCDs) because their properties can be controlled by judicious choice of the number of center points n_0 and the distance α of the axial points from the design center. In addition to rotatability, orthogonal blocking, and orthogonality, a design should include enough center points (say 3–6) to provide a reasonably sensitive test for lack of fit.

16.4.1 Rotatability

A design is *rotatable* if the variance $\mathrm{Var}(\widehat{Y}_{\mathbf{z}})$ of the predicted response is the same for all coded points $\mathbf{z} = (z_1, z_2, \ldots, z_p)$ at any given distance $d = (\sum_i z_i^2)^{1/2}$ from the design center, $\mathbf{z}_0 = (0, 0, \ldots, 0)$. Thus, there is the same amount of information about the response surface at the same distance d in any direction from the design center. This is a reasonable requirement of a design, since data are generally collected without knowing in which direction from the design center the stationary point of the fitted surface will be located.

Rotatable Central Composite Designs

Suppose we take a central composite design for p factors, with one observation at each axial point located a distance α from the design center, and with one observation at each of the n_f factorial points.

It can be shown that such a central composite design is rotatable if

$$\alpha = (n_f)^{1/4}, \tag{16.4.13}$$

and if each axial point is observed r_a times, then the requirement for rotatability becomes

$$\alpha = (n_f / r_a)^{1/4}.$$

The details can be found in the articles by Box and Hunter (1957) and Draper (1982).

Example 16.4.1 `Acid copper pattern plating experiment, continued`

In Example 16.3.1, p. 579, a central composite design was used for $p = 2$ factors. The design involved one observation at each $n_f = 4$ factorial points and $n_a = 4$ axial points, plus two center points. If the model, in terms of coded factor levels, is fitted using $\alpha = (n_f)^{1/4} = \sqrt{2}$, the design is rotatable with respect to the coded factor levels. For example, it can be verified that the estimate of the variance is

$$\widehat{\text{Var}}(\widehat{Y}_{\mathbf{z}}) = 0.0182$$

at each point $\mathbf{z} = (z_1, z_2)$ at distance $\sqrt{2}$ from the design center. This includes each factorial point and each axial point. For comparison, $\widehat{\text{Var}}(\widehat{Y}) = 0.0145$ at the center point and $\widehat{\text{Var}}(\widehat{Y}) = 0.0100$ at the points $(-1, 0)$, $(1, 0)$, $(0, -1)$, and $(0, 1)$, which are each a distance 1.0 from the design center. □

16.4.2 Orthogonality

The second-order model (16.3.10) includes $(p + 1)(p + 2)/2$ parameters, including the intercept γ_0. A second-order design is *orthogonal* if the sums of squares, $ss(\gamma_i|\gamma_0)$ $(i = 1, 2, \ldots, p)$, $ss(\gamma_{ii}|\gamma_0)$ $(i = 1, 2, \ldots, p)$, and $ss(\gamma_{ij}|\gamma_0)$ $(1 \le i < j \le p)$, each adjusted for the intercept γ_0, are independent. In the analysis of variance of an orthogonal design, the sums of squares associated with these $(p + 2)(p + 1)/2 - 1$ parameters are independent, and do not depend on the order in which the parameters are entered into the model. Orthogonality is advantageous if the experimenter is interested in evaluating which of the linear, quadratic, and cross product effects are significantly different from zero.

Orthogonal Central Composite Designs

Suppose we take a central composite design with one observation at each of the n_f factorial points and $2p$ axial points, and with n_0 observations at the center. As shown by Khuri and Cornell (1987), p. 119, the design is orthogonal if

$$(n_f + 2\alpha^2)^2 = n_f n,$$

where n is the total number of observations; that is, $n = n_f + 2p + n_0$. So, a central composite design with n_f factorial points and $2p$ axial points can be made orthogonal by appropriate choice of α or n_0. For example, if the number of center points is fixed at n_0, then n is fixed, and a central composite design is orthogonal if

$$\alpha = \left(\frac{\sqrt{n_f n} - n_f}{2} \right)^{1/2}. \tag{16.4.14}$$

If a central composite design is to be rotatable and n_0 is not fixed, then we would choose $\alpha = (n_f)^{1/4}$, and the design would also be orthogonal if the number of center points was chosen to be

$$n_0 = 4\sqrt{n_f} + 4 - 2p\,. \tag{16.4.15}$$

This may not be achievable, since n_0 must be an integer. Rounding (16.4.15) to the nearest integer gives a rotatable design that is nearly orthogonal.

Example 16.4.2 `Flour production experiment`

In Sect. 16.5, we will consider the last of four experiments described by Tuck et al. (1993) to develop robust bread flours. This experiment was run using a central composite design for three factors, with one observation at each of $n_f = 8$ factorial points and $2p = 6$ axial points. From Eq. (16.4.15), since $\sqrt{n_f} = \sqrt{8}$ is not an integer, the design with $n_f = 8$ cannot be both orthogonal and rotatable. The experimenters used only $n_0 = 2$ center points, giving $n = 16$ observations in total. From Eq. (16.4.14), the design is orthogonal if

$$\alpha = \left(\frac{\sqrt{(8)(16)} - 8}{2}\right)^{1/2} = 1.2872\,.$$

This value of α was used by the experimenters. □

16.4.3 Orthogonal Blocking

If a second-order design is conducted as a block design, then the second-order model (16.3.10) is modified to include additive block effects. For p factors, the model is

$$Y_{h,\mathbf{z},t} = \gamma_0 + \theta_h + \sum_{i=1}^{p} \gamma_i z_i + \sum_{i=1}^{p} \gamma_{ii} z_i^2 + \sum_{i<j} \gamma_{ij} z_i z_j + \epsilon_{h,\mathbf{z},t}\,, \tag{16.4.16}$$

where $Y_{h,\mathbf{z},t}$ denotes the tth observation at coded treatment combination $\mathbf{z} = (z_1, z_2, \ldots, z_p)$ in block h, and the error variables $\epsilon_{h,\mathbf{z},t}$ are independent with $N(0, \sigma^2)$ distributions. The parameter θ_h denotes the effect of the hth block, and the other parameters are defined as in the second-order model (16.3.10).

A design is said to have *orthogonal blocking* if the least squares estimates of the linear, quadratic, and cross product effect parameters are the same under model (16.4.16), which includes block effects, as under the model (16.3.10) without block effects; that is, the linear, quadratic, and cross product effects are estimated independently of the block effects. The primary advantage of orthogonal blocking as compared with nonorthogonal blocking is that an orthogonally blocked design gives the smallest values of $\mathrm{Var}(\widehat{Y})$, $\mathrm{Var}(\hat{\gamma}_i)$, $\mathrm{Var}(\hat{\gamma}_{ii})$, and $\mathrm{Var}(\hat{\gamma}_{ij})$. A second advantage is that a rotatable design conducted with orthogonal blocking is still rotatable.

Given a design in b blocks with orthogonal blocking, the analysis under the block design model (16.4.16) is almost the same as it would be under model (16.3.10) for the design with no blocking. However, the sum of squares for blocks is extracted from the sum of squares for error, and there are $b - 1$ degrees of freedom for blocks giving $b - 1$ fewer degrees of freedom for error. The sum of squares for blocks is

$$ss\theta = \sum_{h=1}^{b} k_h(\overline{y}_{h..} - \overline{y}_{...})^2 = \sum_{h=1}^{b} y_{h..}^2/k_h - y_{...}^2/n\,,$$

where $y_{h..}$ is the sum of the observations in the hth block, k_h is the size of the hth block, and $y_{...}$ is the sum of all n observations in the design.

In their 1957 article, Box and Hunter developed the following general conditions under which a second-order design can be blocked orthogonally.

(1) Each block must be a first-order orthogonal design: that is, (i) for each block and each factor i, the sum of coded levels of the factor, $\sum z_i$, is zero; and (ii) for each block and each pair of factors i and j, the sum of cross products, $\sum z_i z_j$, is zero. (Each sum is over all the observations in the block.)
(2) For each block and each factor i, the sum of squares $\sum z_i^2$ of the coded levels of the ith factor in the block must be proportional to the number of observations in the block.

Orthogonal Blocking of Central Composite Designs

For a central composite design, we first divide the observations into two blocks: an *axial-points block* consisting of the n_a axial points plus n_{0a} center points, and a *factorial-points block* consisting of the n_f factorial points plus n_{0f} center points. This division into blocks is natural if, for example, a first-order design results in lack of fit, so that axial and additional center points are added at a later date to build up to a second-order design. Each of the blocks is a first-order orthogonal design, meeting condition (1) for orthogonal blocking. Concerning condition (2), the sum of squares $\sum z_i^2$ of the coded levels of each factor is $2\alpha^2$ in the axial block and n_f in the factorial block. So, condition (2) requires that

$$\frac{2\alpha^2}{n_f} = \frac{n_a + n_{0a}}{n_f + n_{0f}} .$$

Solving for α, a central composite design has orthogonal blocking if

$$\alpha = \left(\frac{n_f (n_a + n_{0a})}{2(n_f + n_{0f})} \right)^{1/2} . \tag{16.4.17}$$

The design is also rotatable if $\alpha = (n_f)^{1/4}$, in which case we require

$$n_{0f} = (\sqrt{n_f}/2)(n_a + n_{0a}) - n_f . \tag{16.4.18}$$

If the numbers of center points, n_{0a} and n_{0f}, in the blocks can be chosen to satisfy this equation, then the design will be rotatable and can be orthogonally blocked. When this is not possible, it is preferable to maintain orthogonal blocking but to relax rotatability. To accomplish this, the numbers n_{0a} and n_{0f} can be chosen such that Eq. (16.4.18) is approximately satisfied, and then α can be computed from Eq. (16.4.17).

It is sometimes possible to block a central composite design orthogonally in more than two blocks. The axial block cannot be further subdivided, but the factorial block can sometimes be divided into 2^m factorial blocks while maintaining orthogonal blocking if the number of factorial center points n_{0f} is divisible by 2^m so the factorial center points can be equally divided among the 2^m factorial blocks. This is done by confounding interaction effects between three or more factors. Box and Hunter (1957, p. 233) provide a table of blocking arrangements for rotatable and near-rotatable central composite designs. Notice that if center points are spread across b blocks, then they provide $b - 1$ fewer pure error degrees of freedom than they would in a design that is not blocked.

Example 16.4.3 `PAH recovery experiment`

Barnabas et al. (1995) used a central composite design to study the effects of four factors—pressure, temperature, extraction time, and methanol content—on the total recovery of polycyclic aromatic hydrocarbons (PAHs) when extracted from soil. The design was composed of $n_f = 2^4 = 16$ factorial points and $n_a = 2p = 8$ axial points. Taking $\alpha = 16^{1/4} = 2$ would give a rotatable design. From Eq. (16.4.18),

$$n_{0f} = (\sqrt{16}/2)(8 + n_{0a}) - 16 = 2n_{0a},$$

so use of twice as many factorial center points as axial center points would give a rotatable design that could be orthogonally blocked.

The experimenters chose to use $n_{0a} = 2$ axial center points and $n_{0f} = 4$ factorial center points. This gave an axial block of size 10 and a factorial block of size 20. They then subdivided the factorial block into two blocks each of size 10 by confounding the four-factor interaction and including two of the four factorial center points in each factorial block. The resulting design was rotatable with orthogonal blocking. Analysis of the design is discussed in Sects. 16.7.2 and 16.8.2 using the SAS and R software packages, respectively. The design itself is shown in Tables 16.16 (p. 596) and 16.19 (p. 604), where the first ten observations comprise the first factorial block, the second ten the second factorial block, and the final ten the axial block. □

16.5 A Real Experiment: Flour Production Experiment, Continued

Tuck et al. (1993) described a series of four related experiments, involving quality improvement in the milling industry. The collective purpose of the experiments was to develop a bread flour that would give high loaf volume despite fluctuations in the bread-making process. We consider here their fourth experiment.

Bread flour consists of wheat plus a small number of minor ingredients. Their fourth experiment was concerned with the effects of three such ingredients (labeled design factors B, C, and D) on loaf volume. An orthogonal central composite design, involving eight factorial points, six axial points, and two center points, was used. For the axial points, the value $\alpha = 1.2872$ was used to make the design orthogonal (see Example 16.4.2).

When a product consists of a mixture of ingredients, and the total volume of the mixture is held constant, the fractions associated with the ingredients in the mixture necessarily sum to one. This has implications for the model and data analysis. However, in this experiment, the minor ingredients constituted such a small portion of the mixture that the total volume did not need to remain fixed, and standard response surface methods could be used to study the design factors.

There were a number of sources of variation in the production process that constituted noise factors. The production factors were paired in order to keep the experiment small. So, noise factor G represented oven bake and proof time, noise factor J represented yeast and water level, and noise factor K represented degree of mixing and moulding pressure. Each of these composite factors had two levels, "high" and "low." A 2^{3-1}_{III} fraction in the composite noise factors was used, with defining relation

$$I = GJK.$$

The experimental design used was a product array. It included $16 \times 4 = 64$ observations—each of the 16 design factor combinations of the central composite design was observed with each of the four noise factor combinations of the noise array. Also, the noise factors were difficult to change, so each

Table 16.11 Flour production experiment: average specific volume $y_{h\mathbf{z}}$ of loaves on half-day h; $\alpha = 1.2872$

z_B	z_C	z_D	$y_{1\mathbf{z}}$	$y_{2\mathbf{z}}$	$y_{3\mathbf{z}}$	$y_{4\mathbf{z}}$	$\overline{y}_{\cdot\mathbf{z}}$	$100\log_{10}(s_{\mathbf{z}})$
−1	−1	−1	586	399	418	404	451.75	195.36
−1	−1	1	615	411	435	421	470.50	198.60
−1	1	−1	611	422	431	439	475.75	195.63
−1	1	1	639	436	444	454	493.25	198.88
1	−1	−1	603	422	400	430	463.75	197.17
1	−1	1	622	411	425	436	473.50	199.79
1	1	−1	634	471	436	425	491.50	198.68
1	1	1	673	433	423	462	497.75	207.19
α	0	0	618	414	419	477	482.00	197.80
$-\alpha$	0	0	586	421	420	455	470.50	189.60
0	α	0	621	426	427	458	483.00	196.94
0	$-\alpha$	0	629	412	412	426	469.75	202.68
0	0	α	631	411	433	453	482.00	200.35
0	0	$-\alpha$	587	413	419	430	462.25	192.15
0	0	0	604	432	416	438	472.50	194.53
0	0	0	602	425	407	439	468.25	195.48

Source Tuck et al. (1993). Copyright © 1993 Blackwell Publishers. Reprinted with permission

noise factor combination constituted a different block, and in each block the design factor treatment combinations $(z_B z_C z_D)$ were randomly ordered. Observations were collected over two days using half-days as blocks, with the four blocks collected in the order $(z_G z_J z_K) = 111, 100, 010, 001$. As a result, noise factor effects are also confounded with changes in conditions from half-day to half-day. For each observation, three loaves were baked from a single dough, then the average specific volume of the three loaves recorded. The resulting data $y_{h\mathbf{z}}$ are shown in Table 16.11.

For each of the 16 treatment combinations $\mathbf{z}$ of the central composite design in turn, the sample mean $\overline{y}_{\cdot\mathbf{z}}$ and the log sample variance ($\times 100$) were computed from the observations $y_{h\mathbf{z}}$ in the four blocks ($h = 1, 2, 3, 4$). The effects of the design factors on these two response variables were studied separately by fitting second-order response surface regression models to each set of 16 responses.

The analysis of variance for fitting the second-order model to the response $\overline{y}_{\cdot\mathbf{z}}$ is shown in Table 16.12. Because the design is orthogonal, the effects can be assessed for significance independently of their order of entry into the model. The only effects that are significantly different from zero at an individual significance level of 0.01 are the main effects of factors C and D. The overall significance level for the nine tests is at most 0.09. The experimenters decided also to retain the main effect of factor B, for which $p = 0.0204$. If the corresponding first-order model is fitted to $\overline{y}_{\cdot\mathbf{z}}$, we obtain

$$\hat{\overline{y}}_{\cdot\mathbf{z}} = 475.50 + 4.42 z_B + 10.24 z_C + 6.87 z_D \,.$$

The coefficients of z_B, z_C, and z_D are all positive. Thus, increasing the level of design factors B, C, and D has a positive effect on the mean loaf specific volume.

The analysis of variance for the response $100\log_{10}(s_{\mathbf{z}})$ is shown in Table 16.13. No effects can be regarded as significantly different from zero at an individual 0.01 significance level. However, in this setting it would not be particularly bad to make a Type I error, and if we raise the individual significance level we would select the linear effects of factors B and D and the quadratic effect of C as being the important effects. If the corresponding model is fitted and the linear effect of C is also included, we obtain

Table 16.12 Flour production experiment: analysis of variance for $\overline{y}_{\cdot\mathbf{z}}$

Source of variation	Degrees of freedom	Sum of squares	Mean square	Ratio	p-value
z_B	1	221.4366	221.4366	9.77	0.0204
z_C	1	1185.3603	1185.3603	52.30	0.0004
z_D	1	533.2415	533.2415	23.53	0.0029
z_B^2	1	48.0081	48.0081	2.12	0.1958
z_C^2	1	50.4906	50.4906	2.23	0.1862
z_D^2	1	1.1997	1.1997	0.05	0.8257
$z_B z_C$	1	3.4453	3.4453	0.15	0.7101
$z_B z_D$	1	51.2578	51.2578	2.26	0.1833
$z_C z_D$	1	2.8203	2.8203	0.12	0.7363
Error	6	135.9897	22.6650		
Total	15	2233.2500			

Table 16.13 Flour production experiment: analysis of variance for $100 \log_{10}(s_{\mathbf{z}})$

Source of variation	Degrees of freedom	Sum of squares	Mean squares	Ratio	p-value
z_B	1	54.9174	54.9174	8.24	0.0284
z_C	1	0.3730	0.3730	0.06	0.8208
z_D	1	70.1514	70.1514	10.53	0.0176
z_B^2	1	0.6409	0.6409	0.10	0.7669
z_C^2	1	61.4625	61.4625	9.23	0.0229
z_D^2	1	7.8587	7.8587	1.18	0.3191
$z_B z_C$	1	8.7175	8.7175	1.31	0.2963
$z_B z_D$	1	2.6931	2.6931	0.40	0.5484
$z_C z_D$	1	4.3269	4.3269	0.65	0.4511
Error	6	39.9752	6.6625		
Total	15	251.1164			

$$\begin{aligned} \widehat{100 \log_{10}}(s_{\mathbf{z}}) &= 195.19 + 0.18 z_C + 3.35 z_C^2 + 2.20 z_B + 2.49 z_D \\ &\approx 195.19 + 3.35 (z_C + 0.027)^2 + 2.20 z_B + 2.49 z_D \,. \end{aligned}$$

Taking the two fitted models, we see that not only does the mean response increase as the levels of factors B, C, and D are increased, but so does the variability. The minimum variability with respect to factor C is achieved at $z_C = -0.027$. However, the amount of factor C in the loaf cannot be negative, and so the minimum variability is achieved when the amount of factor C, as well as factors B and D, is zero.

The end result was that the experimenters set $z_C = 0$ to achieve low variability and adjusted the level of factor B (which has the slightly smaller effect on the variance, and may have been less costly than factor D) to raise mean response to the desired level.

16.6 Box–Behnken Designs

A central composite design has five levels for each factor, ± 1, $\pm\alpha$, 0. For a given experiment, circumstances may dictate the use of fewer levels, but at least three levels per factor are needed for quadratic terms to be estimable in the second-order model. Use of 3^p factorial designs or regular 3^{p-s} fractional factorial designs might be considered. These tend to be large, however, and the smaller ones tend to be of resolution III or IV so that two-factor interactions are confounded with main effects or other two-factor interactions. For fitting a second-order response model a different type of design, called a *Box–Behnken design*, is often preferred, since interaction parameter estimates are not completely confounded, and in many cases, these designs are considerably smaller than 3^{p-s} fractional factorial designs.

A Box–Behnken design for p factors is constructed by a composition of an incomplete block design for p treatments in b blocks of size k and a 2^k factorial design having factor levels coded $+1$ and -1. The method of composition is illustrated in Example 16.6.1. In addition to the points generated by the composition, center points must be added to the design for all model parameters to be estimable.

A list of Box–Behnken designs can be found in the article of Box and Behnken (1960). The designs have p factors with each factor observed at 3 levels, for $p = 3$–7, 9–12, and 16. The designs for $p = 4$ and 7 are rotatable, and the others are nearly rotatable. The designs for $p = 4$–7, 9, 10, 12, and 16 allow orthogonal blocking. All of the designs possess a high degree of orthogonality, the only correlation being between the estimators of the intercept and the quadratic terms.

Example 16.6.1 `Construction of a Box-Behnken Design:` $p = 4$

Suppose we require a second-order design for $p = 4$ factors, each observed at three levels, and with a total of 27 observations. As illustrated by Box and Behnken (1960), a Box–Behnken design can be constructed from a composition of a balanced incomplete block design in $b = 6$ blocks of size $k = 2$ and a 2^2 factorial design as follows. The balanced incomplete block design, shown below left, consists of all possible combinations of four treatment labels taken two at a time. Shown to its right are the $v = 4$ treatment combinations of a 2^2 design, with factor levels coded $+1$ and -1. These two designs are composed as follows. In each of the six blocks of the incomplete block design, the treatment labels are replaced by the symbol ± 1 and the blank "$-$" by 0 to give the Box–Behnken design represented in condensed form (and without center points) below right.

$$\begin{bmatrix} 1 & 2 & - & - \\ - & - & 3 & 4 \\ 1 & - & 3 & - \\ - & 2 & - & 4 \\ 1 & - & - & 4 \\ - & 2 & 3 & - \end{bmatrix} \text{ with } \begin{bmatrix} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \text{ gives } \begin{bmatrix} \pm 1 & \pm 1 & 0 & 0 \\ 0 & 0 & \pm 1 & \pm 1 \\ \pm 1 & 0 & \pm 1 & 0 \\ 0 & \pm 1 & 0 & \pm 1 \\ \pm 1 & 0 & 0 & \pm 1 \\ 0 & \pm 1 & \pm 1 & 0 \end{bmatrix}$$

The same design, but expanded out and augmented with three center points, is shown in Table 16.14. The first ± 1 in each row of the condensed design is replaced by the first column of levels of the 2^2 design, the second ± 1 is replaced by the second column of levels, and each 0 is replaced by a column of $v = 4$ zeros. With the addition of three center points, this gives the design with 27 treatment

Table 16.14 Box–Behnken design: $p = 4$ factors, $n = 27$ treatment combinations

$$\begin{bmatrix} -1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

combinations shown as the 27 rows of Table 16.14. Although this design has the same number of treatment combinations as a 3_{IV}^{4-1} design, it does not have complete confounding of the two-factor interactions in pairs. □

In general, the composition of an incomplete block design for p treatments in b blocks of size k with a factorial design with $v = 2^k$ treatment combinations yields a Box–Behnken design for p factors with bv treatment combinations. The ith of the k treatment labels in each block is replaced by the ith of the k columns of the factorial design, and each "$-$" is replaced by a column of v zeros.

In general, if the incomplete block design is a balanced incomplete block design with $r = 3\lambda$, as in Example 16.6.1, then the resulting Box–Behnken design is rotatable—otherwise not. If there does not exist a balanced incomplete block design with $r = 3\lambda$, then one can either use a balanced incomplete block design with $r \neq 3\lambda$ or use a partially balanced incomplete block design. If a partially balanced incomplete block design is used, each pair of treatment labels must occur together in at least one block for all second-order model parameters to be estimable.

Orthogonal Blocking

Many Box–Behnken designs can be blocked orthogonally. The requirements for orthogonal blocking of a second-order design were given in Sect. 16.4.3, and these imply that a Box–Behnken design can be blocked orthogonally under either of two circumstances.

First, if the blocks of the incomplete block design in the composition can be partitioned into equireplicate sets, then the same partition of observations in the Box–Behnken design provides orthogonal blocking as long as the same number of center points is included in each block. Such is the case for the design of Example 16.6.1, since each pair of blocks in the balanced incomplete block design includes every treatment label exactly once. For the resulting Box–Behnken design in Table 16.14, each bracketed set of nine treatment combinations is a corresponding block with one center point included.

The second situation that allows orthogonal blocking occurs when interactions involving three or more factors can be confounded in the generating factorial design. An example follows:

Example 16.6.2 Example of orthogonal blocking

For $p = 4$ factors, the balanced incomplete block design with blocks consisting of the four combinations of three treatment labels can be combined with the 2^3 factorial design as follows.

$$\begin{bmatrix} 1 & 2 & 3 & - \\ 1 & 2 & - & 4 \\ 1 & - & 3 & 4 \\ - & 2 & 3 & 4 \end{bmatrix} \text{ with } \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \text{ gives } \begin{bmatrix} \pm 1 & \pm 1 & \pm 1 & 0 \\ \pm 1 & \pm 1 & 0 & \pm 1 \\ \pm 1 & 0 & \pm 1 & \pm 1 \\ 0 & \pm 1 & \pm 1 & \pm 1 \end{bmatrix},$$

where the ith occurrence of ± 1 in any row of the combined design is replaced by the ith column of the factorial design, and each 0 in the combined design is replaced by a column of eight 0's. The resulting 32-run Box–Behnken design can be partitioned into two blocks of size 16 by confounding the three-factor interaction in the generating factorial design. Thus, treatment combinations in the combined design are divided into two blocks, the division depending on whether they include an even or odd number of factors at level "−1." An equal number of center points must be added to each block. □

16.7 Using SAS Software

In this section we illustrate the analysis of a standard first-order design and a central composite design using the SAS procedures GLM and RSREG, respectively.

16.7.1 Analysis of a Standard First-Order Design

The acid copper pattern plating experiment of Poon (1995) was introduced in Example 16.2.4 (p. 574). This small experiment involved four factorial points and two center points. A SAS program using the GLM procedure for the analysis of this standard first-order design is shown in Table 16.15. After reading

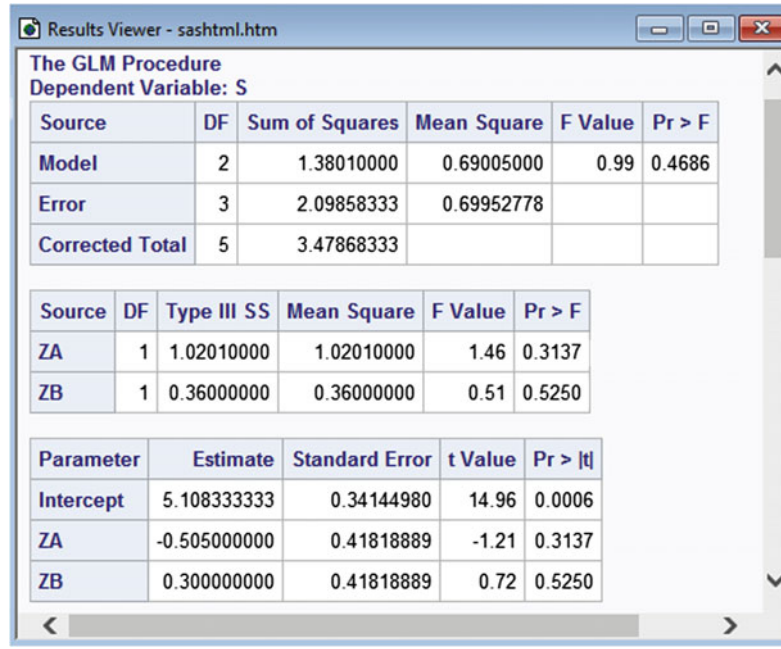
Table 16.15 SAS program for first-order response surface regression

```
* Enter data of the first-order design and code levels;
DATA COPPER;
  INPUT XA XB S;
  ZA = (XA - 10.5);
  ZB = (XB - 36)/5;
  LINES;
   9.5  31  5.60
   9.5  41  6.45
  11.5  31  4.84
  11.5  41  5.19
  10.5  36  4.32
  10.5  36  4.25
;
* Analysis of the first-order design;
PROC GLM;
  MODEL S = ZA ZB;
* Add model terms to test for lack of fit;
PROC GLM;
  MODEL S = ZA ZB ZA*ZB ZA*ZA;
```

the data and coding the factor levels, there are two calls of PROC GLM. Neither of these calls includes a CLASS statement, since the goal is to fit a regression model to the levels of the quantitative factors and not to compare the effects of their levels.

In the first call, the first-order model (16.2.3) is fitted, generating the output shown in Fig. 16.5. Neither main effect is significantly different from zero, indicating either that the experimental region is in the vicinity of the peak, or that neither factor affects the response.

In the second call of PROC GLM, the interaction term and one quadratic term are added to the model to test for lack of fit of the first-order model—the model would contain too many parameters if both quadratic terms were added. Some of the resulting output is shown in Fig. 16.6. At an overall level of



Results Viewer - sashtml.htm

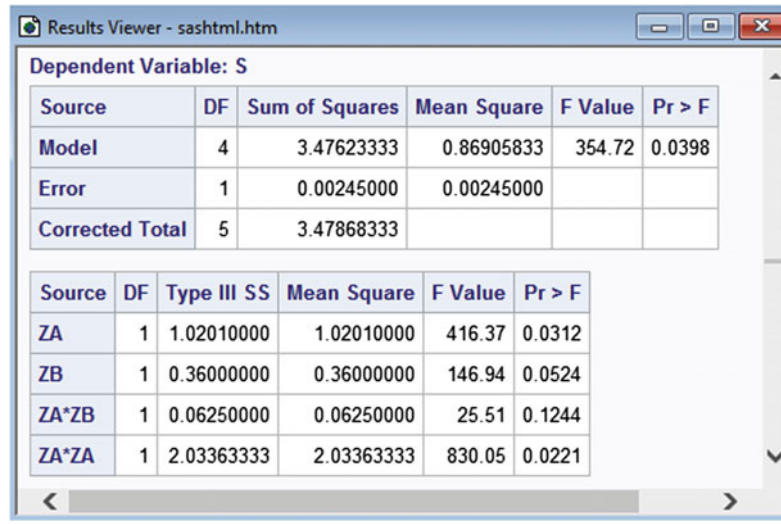
The GLM Procedure
Dependent Variable: S

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1.38010000	0.69005000	0.99	0.4686
Error	3	2.09858333	0.69952778		
Corrected Total	5	3.47868333			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ZA	1	1.02010000	1.02010000	1.46	0.3137
ZB	1	0.36000000	0.36000000	0.51	0.5250

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	5.108333333	0.34144980	14.96	0.0006
ZA	-0.505000000	0.41818889	-1.21	0.3137
ZB	0.300000000	0.41818889	0.72	0.5250

Fig. 16.5 SAS output from the first call of PROC GLM: analysis of variance and parameter estimates for a first-order design



Results Viewer - sashtml.htm

Dependent Variable: S

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	3.47623333	0.86905833	354.72	0.0398
Error	1	0.00245000	0.00245000		
Corrected Total	5	3.47868333			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ZA	1	1.02010000	1.02010000	416.37	0.0312
ZB	1	0.36000000	0.36000000	146.94	0.0524
ZA*ZB	1	0.06250000	0.06250000	25.51	0.1244
ZA*ZA	1	2.03363333	2.03363333	830.05	0.0221

Fig. 16.6 SAS output from the second call of PROC GLM: test for lack of fit of the first-order model

Table 16.16 SAS program for response surface regression (PAH recovery experiment)

```
DATA PAH;
  INPUT RUN B1  B2 PRES TEMP   ET  MC     Y;
  LINES;
         1   1   0  250   55 47.5  15 391.8
         2   1   0  150   85 47.5  15 413.6
         3   1   0  250   55 22.5   5  68.7
         4   1   0  250   85 47.5   5 143.0
         5   1   0  150   85 22.5   5 104.0
         6   1   0  150   55 22.5  15 309.1
         7   1   0  200   70 35.0  10 400.6
         8   1   0  250   85 22.5  15 402.5
         9   1   0  150   55 47.5   5  77.7
        10   1   0  200   70 35.0  10 426.5
        11   0   1  250   85 47.5  15 457.5
        12   0   1  150   55 22.5   5  56.9
        13   0   1  250   85 22.5   5  94.1
        14   0   1  250   55 22.5  15 409.7
        15   0   1  150   55 47.5  15 410.9
        16   0   1  150   85 22.5  15 375.8
        17   0   1  150   85 47.5   5 110.5
        18   0   1  200   70 35.0  10 387.8
        19   0   1  250   55 47.5   5 103.0
        20   0   1  200   70 35.0  10 399.1
        21  -1  -1  200   70 35.0  10 416.9
        22  -1  -1  200   40 35.0  10 359.8
        23  -1  -1  200   70 10.0  10 276.1
        24  -1  -1  200   70 60.0  10 462.3
        25  -1  -1  100   70 35.0  10 311.5
        26  -1  -1  200   70 35.0  10 346.5
        27  -1  -1  200   70 35.0   0  46.8
        28  -1  -1  200   70 35.0  20 418.7
        29  -1  -1  200  100 35.0  10 413.9
        30  -1  -1  300   70 35.0  10 429.4
;
* Sort by independent variables for lack of fit test;
PROC SORT; BY B1 B2 PRES TEMP ET MC;
* Response surface regression, including contour plots;
ODS GRAPHICS ON; * Needed for contour plots;
PROC RSREG PLOTS = SURFACE;
  MODEL Y = B1 B2 PRES TEMP ET MC / COVAR = 2 LACKFIT;
  RUN;
ODS GRAPHICS OFF;
```

Source The data in the program are reprinted from Barnabas et al. (1995) with permission. Copyright © 1995 American Chemical Society

0.10 for the four tests (each being done at individual level $\alpha^* = 0.025$), the quadratic term `ZA*ZA` is significantly different from zero, indicating the presence of significant curvature. This fact caused the experimenters to add axial points to the first-order design to obtain a central composite design (see Example 16.3.1).

16.7.2 Analysis of a Second-Order Design

The SAS procedure RSREG is used to fit a second-order response surface regression model. This is illustrated in Table 16.16 in the context of the PAH recovery experiment that was introduced in Example 16.4.3 (p. 589). A rotatable central composite design with orthogonal blocking was used to study the effects of four factors—pressure (PRES), temperature (TEMP), extraction time (ET), and methanol content (MC)—on the total recovery of polycyclic aromatic hydrocarbons (Y) when extracted from soil.

The SAS program shown in Table 16.16 reads the run number, the levels of the block indicator variables, the uncoded levels of the four factors, and the data into data set ONE. Until now, we have always declared a block variable to be a classification variable via the CLASS statement and listed its levels as $1, 2, \ldots, b$. However, PROC RSREG does not recognize classification variables, and if a single block factor were included in the model, it would be interpreted as a single covariate—a quantitative variable possessing one degree of freedom. We have included in the model the pair of covariates (B1, B2), for which we have selected the three coded pairs of levels (1, 0), (0, 1) and $(-1, -1)$. The three pairs of levels distinguish the three blocks and provide two block degrees of freedom.

Only the factor *names* need be listed in the MODEL statement in RSREG, as all quadratic and cross product terms in the factors are automatically included in the model. To avoid treatment–block interactions from being included, B1 and B2 are declared to be covariates. This is done via the option COVAR = 2, which indicates that the first two listed independent variables are covariates and should not be included in any interactions.

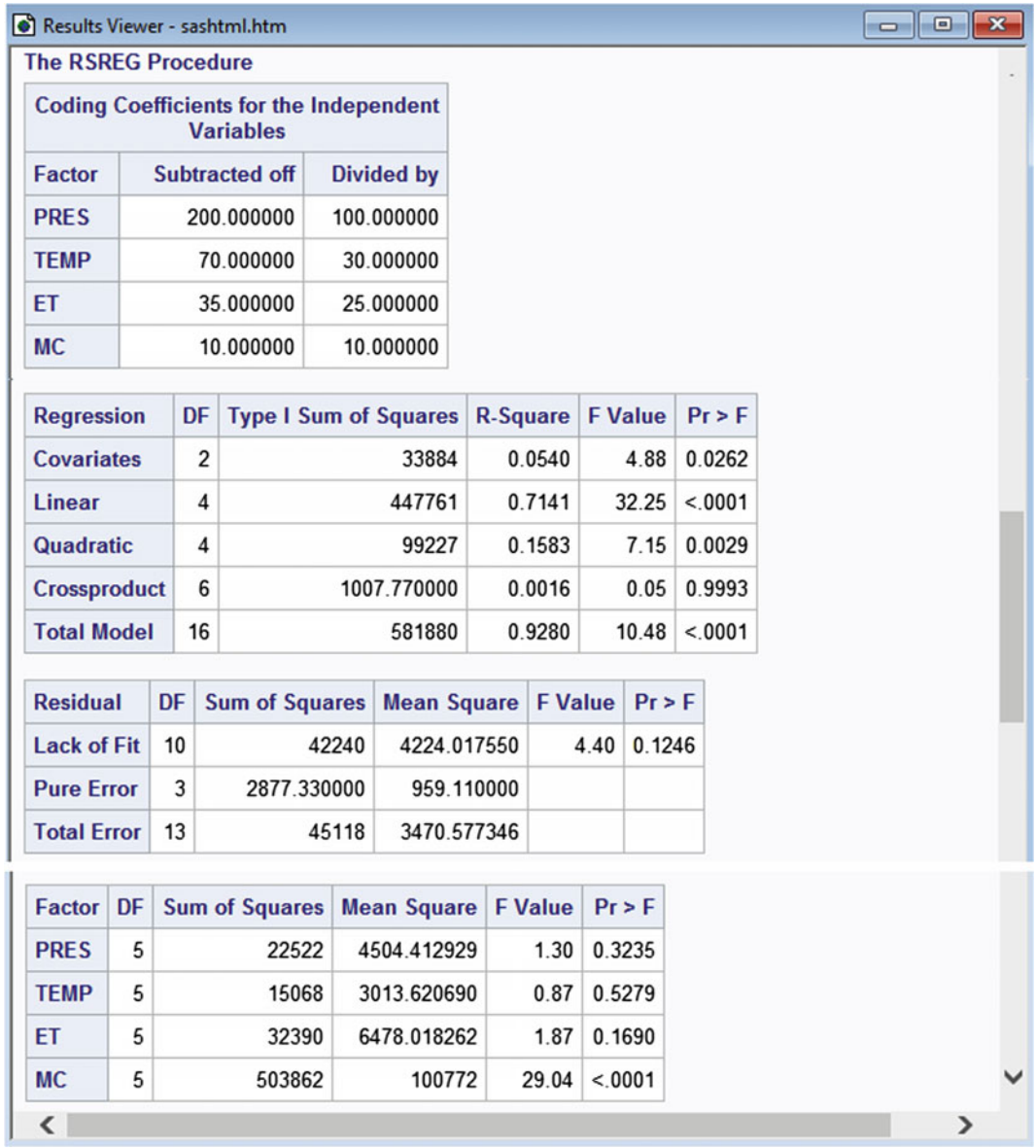
A generic test for model lack of fit can optionally be requested if the SAS data set has been sorted by the independent variables in the model to cluster replicated observations. PROC SORT is used to sort the data, and a test for lack of fit is requested via the option LACKFIT in the model statement of PROC RSREG.

PROC RSREG codes the levels of each factor so that $+1$ and -1 represent the extreme levels of each factor. For example, the axial points of a central composite design would typically be coded ± 1 by SAS software rather than the conventional $\pm\alpha$. Figure 16.7 shows how SAS codes the factor levels, as well as the resulting analysis of variance table. The analysis of variance table includes Type I sums of squares for covariates, linear terms, quadratic terms, and cross product terms, adding the terms to the model in that order. These Type I sums of squares are the same, whether coded or uncoded factor levels are specified in the model statement. Observe that the cross product terms are not significantly different from zero, and there is no significant lack of fit of the model.

Type III sums of squares are also provided for each factor, pooling together the sums of squares for all terms—linear, quadratic and interaction—involving the factor. This information can be used for assessing whether any single factor can be removed from the model. These Type III sums of squares are also the same using either the coded or uncoded factor levels. The Type III sums of squares indicate that the factor methanol content (MC) is needed in the model, but perhaps not the other factors. Further analysis could explore what additional terms are needed, if any.

Figure 16.8 contains the parameter estimates and corresponding t-tests. Clearly the linear and quadratic methanol content terms are needed in the model, providing some clarification to the analysis of variance results.

In Fig. 16.9, the canonical analysis is shown, including the stationary point (Critical Value) in terms of both coded and uncoded factor levels; the predicted value at the stationary point; the canonical coefficients (Eigenvalues); classification of the stationary point as a maximum, minimum, or saddle point; and the direction of each canonical axis (Eigenvectors). The canonical coefficients and axes are with respect to the coded factor levels.


Results Viewer - sashtml.htm

The RSREG Procedure

Coding Coefficients for the Independent Variables

Factor	Subtracted off	Divided by
PRES	200.000000	100.000000
TEMP	70.000000	30.000000
ET	35.000000	25.000000
MC	10.000000	10.000000

Regression	DF	Type I Sum of Squares	R-Square	F Value	Pr > F
Covariates	2	33884	0.0540	4.88	0.0262
Linear	4	447761	0.7141	32.25	<.0001
Quadratic	4	99227	0.1583	7.15	0.0029
Crossproduct	6	1007.770000	0.0016	0.05	0.9993
Total Model	16	581880	0.9280	10.48	<.0001

Residual	DF	Sum of Squares	Mean Square	F Value	Pr > F
Lack of Fit	10	42240	4224.017550	4.40	0.1246
Pure Error	3	2877.330000	959.110000		
Total Error	13	45118	3470.577346		

Factor	DF	Sum of Squares	Mean Square	F Value	Pr > F
PRES	5	22522	4504.412929	1.30	0.3235
TEMP	5	15068	3013.620690	0.87	0.5279
ET	5	32390	6478.018262	1.87	0.1690
MC	5	503862	100772	29.04	<.0001

Fig. 16.7 SAS output from `PROC RSREG`: coding of factor levels, analysis of variance, and lack-of-fit test

For this experiment, all eigenvalues (canonical coefficients) are negative, so the stationary point is a maximum. The eigenvectors are each scaled to be of length one. The last eigenvalue, with value −227.865046, is the largest in magnitude. For the corresponding eigenvector, the primary component is that of `MC` with value 0.994301. So, the fitted model has greatest curvature at the stationary point when moving in either direction determined by this fourth eigenvector, which is nearly parallel to the `MC`-axis. This is evident from the contour plots in Fig. 16.10, where `MC` is the x-axis variable of plots (b), (e) and (f). Such a panel of contour plots is generated by inclusion of `PLOTS = SURFACE` as an option of `PROC RSREG` in Table 16.16, whereas changing the option to `PLOTS = 3D` would

Results Viewer - sashtml.htm

Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|	Parameter Estimate from Coded Data
Intercept	1	-1238.272778	552.190337	-2.24	0.0430	396.233333
PRES	1	3.998267	2.492765	1.60	0.1327	37.300000
TEMP	1	12.755444	8.744702	1.46	0.1684	31.783333
ET	1	12.320000	9.182487	1.34	0.2027	54.966667
MC	1	65.649667	21.967351	2.99	0.0105	263.066667
PRES*PRES	1	-0.008863	0.004499	-1.97	0.0706	-88.625000
TEMP*PRES	1	-0.002117	0.019637	-0.11	0.9158	-6.350000
TEMP*TEMP	1	-0.080250	0.049994	-1.61	0.1325	-72.225000
ET*PRES	1	-0.004660	0.023565	-0.20	0.8463	-11.650000
ET*TEMP	1	0.003067	0.078549	0.04	0.9695	2.300000
ET*ET	1	-0.143800	0.071991	-2.00	0.0671	-89.875000
MC*PRES	1	0.023100	0.058912	0.39	0.7013	23.100000
MC*TEMP	1	-0.014500	0.196372	-0.07	0.9423	-4.350000
MC*ET	1	0.066200	0.235646	0.28	0.7832	16.550000
MC*MC	1	-2.263250	0.449945	-5.03	0.0002	-226.325000
B1	1	-27.073333	15.210911	-1.78	0.0985	-27.073333
B2	1	-20.293333	15.210911	-1.33	0.2051	-20.293333

Fig. 16.8 SAS output from `PROC RSREG`: parameter estimates for uncoded and coded factor levels

generate response surface plots. Note that such graphics require `ODS GRAPHICS ON`, and `PROC RSREG` must run before turning `ODS GRAPHICS OFF`.

16.8 Using R Software

In this section we illustrate the analysis of a standard first-order design and a central composite design using the response surface methods function `rsm` of the `rsm` library. We then illustrate generation of central composite and Box–Behnken designs using functions of the `rsm` package.

16.8.1 Analysis of a Standard First-Order Design

The acid copper pattern plating experiment of Poon (1995) was introduced in Example 16.2.4 (p. 574). This small experiment involved four factorial points and two center points. An R program and output for the analysis of this standard first-order design is shown in Tables 16.17 and 16.18. In Table 16.17,

Fig. 16.9 SAS output from PROC RSREG: canonical analysis

Results Viewer - sashtml.htm

The RSREG Procedure
Canonical Analysis of Response Surface Based on Coded Data

Factor	Critical Value	
	Coded	Uncoded
PRES	0.259473	225.947276
TEMP	0.195926	75.877786
ET	0.347209	43.680235
MC	0.605224	16.052237

Predicted value at stationary point: 493.335662

Eigenvalues	Eigenvectors			
	PRES	TEMP	ET	MC
-71.256548	-0.233360	0.964432	0.121729	-0.024413
-84.167652	0.712899	0.255316	-0.652941	0.016008
-93.760753	0.655835	0.067269	0.744877	0.102535
-227.865046	-0.084838	0.012632	-0.063313	0.994301

Stationary point is a maximum.

the data are read from file, coded using the `coded.data` function of the `rsm` package, saved as `copper1`, and displayed.

In the R program continuation in Table 16.18, the first-order analysis is generated using the `rsm` function. In the statement

```
model1 = rsm(s ~ FO(zA, zB), data = copper1)
```

the syntax `FO(zA, zB)` fits the first order model in both coded response variables, using the coded data `copper1`, saving the results as `model1`. Then the `summary(model1)` command generates pertinent information, including: parameter estimates, standard errors, and corresponding t-tests; the analysis of variance table, including a lack-of-fit test; and the direction of steepest ascent with respect to coded and uncoded variables. Finally, the command

```
steepest(model1, dist=seq(0, 5, by = 1), descent = F)
```

provides predicted response at steps along the path of steepest ascent, stepping from the origin (which is the design center point for coded data) at distances from zero to five in unit increments, showing the location of each step in terms of the coded and uncoded predictor variables.

Based on the t tests, neither main effect is significantly different from zero, indicating either that the experimental region is in the vicinity of the peak, or that neither factor affects the response. The lack-of-fit test yields a p-value of 0.034, indicating significant lack-of-fit of the first order model, though the test does not distinguish whether this is due to interaction or quadratic effects. The reader may verify that the significant lack-of-fit is due to a quadratic effect. This fact caused the experimenters to add axial points to the first-order design to obtain a central composite design (see Example 16.3.1).

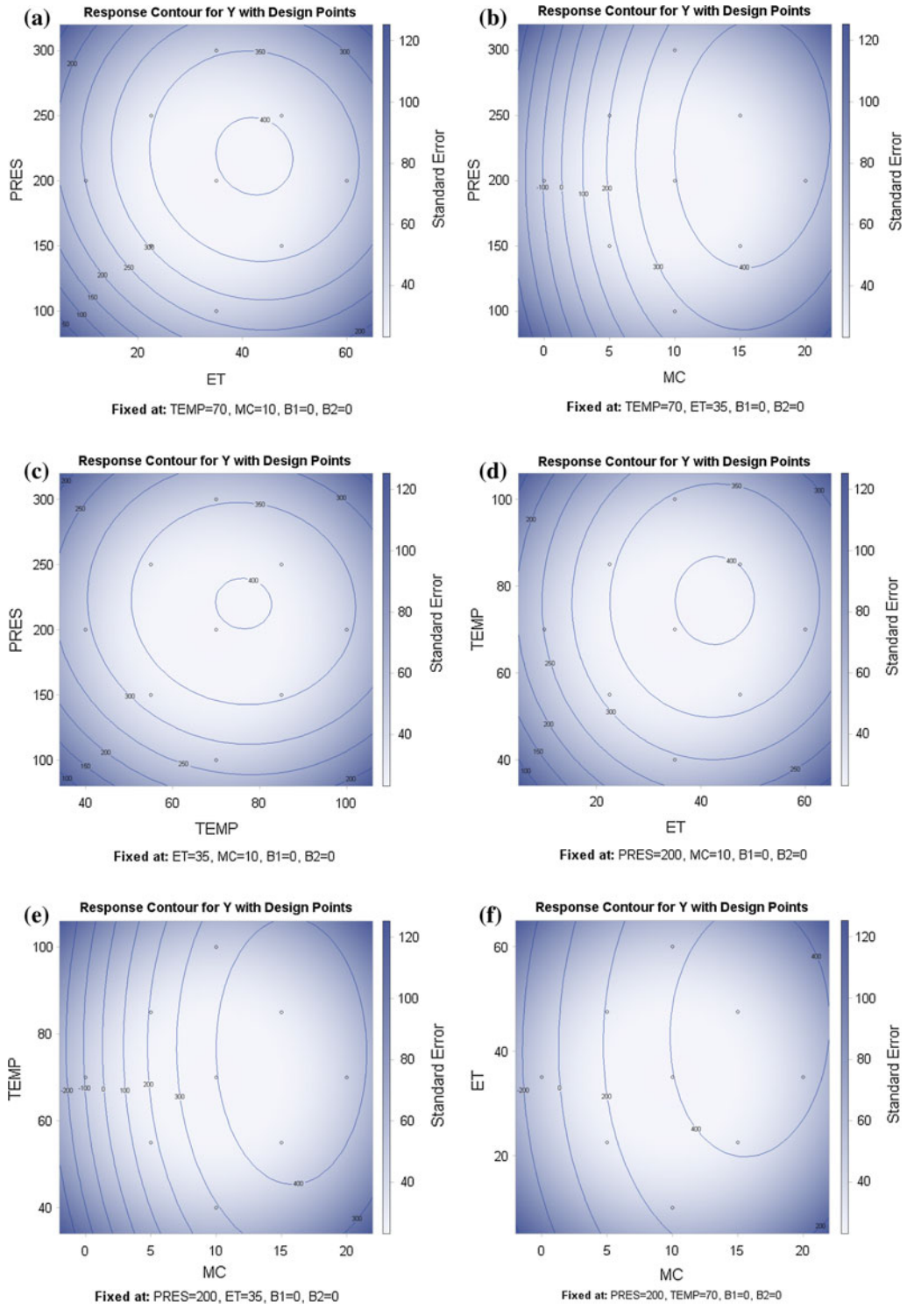


Fig. 16.10 Response surface contour plots for the PAH recovery experiment

Table 16.17 R program and output for first-order response surface regression: data entry and coding

```
> # Read first 6 observations from file
> copper.data = head(read.table("data/copper.txt", header = T), 6)
> # Code data
> # install.packages("rsm")
> library(rsm)
> copper1 = coded.data(copper.data, zA ~ xA - 10.5, zB ~ (xB - 36)/5)
> copper1

    xA xB    s
1  9.5 31 5.60
2  9.5 41 6.45
3 11.5 31 4.84
4 11.5 41 5.19
5 10.5 36 4.32
6 10.5 36 4.25

Data are stored in coded form using these coding formulas ...
zA ~ xA - 10.5
zB ~ (xB - 36)/5
```

16.8.2 Analysis of a Second-Order Design

In the previous section, the syntax `FO` was used with the `rsm` function to fit a first order response surface regression model. Analogously, the syntax `SO` is used to fit a second order model. This is illustrated in the R program beginning in Table 16.19, in the context of the PAH recovery experiment that was introduced in Example 16.4.3 (p. 589). A rotatable central composite design with orthogonal blocking was used to study the effects of four factors—pressure (`Pres`), temperature (`Temp`), extraction time (`ET`), and methanol content (`MC`)—on the total recovery of polycyclic aromatic hydrocarbons (`y`) when extracted from soil.

The R program beginning in Table 16.19 reads the data from a file `pah.txt` which contains the run number, the block level, the uncoded levels of the four factors, and the response variable. A factor variable for blocks is created and all the information is saved in the data set `pah.data`. The function `coded.data` of the response surface methods package `rsm` then codes the levels of each of the factors, saving the coded data as `pah.ccd.data`. The data are then displayed (see Table 16.19), including the coding formulas. The coding formulas used here code the extreme levels of each factor as ± 1, so the results presented here are consistent with those in the prior SAS software section, though one could certainly choose instead to code the levels of the factorial points as ± 1.

The R program is continued in Table 16.20, where the output shown is all generated by the following two program lines.

```
model2 = rsm(y ~ fBlock + SO(zP, zT, zET, zMC), data = pah.ccd.data)
summary(model2)
```

The first line fits the response surface regression model, saving the results as `model2`. Block effects are included in the model additively, whereas the syntax `SO(zP, zT, zET, zMC)` causes inclusion of all terms up to second order in the four factors, including linear, interaction, and quadratic effects. The `summary` command then generates the output shown in Table 16.20, as well as the canonical analysis shown in Table 16.21.

Table 16.18 R program and output for first-order response surface regression: parameter estimates and analysis of variance

```
> # First-order model analysis
> model1 = rsm(s ~ FO(zA, zB), data = copper1)
> summary(model1)

Call:
rsm(formula = s ~ FO(zA, zB), data = copper1)

            Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.108      0.341   14.96  0.00065
zA            -0.505      0.418   -1.21  0.31373
zB             0.300      0.418    0.72  0.52495

Multiple R-squared:  0.397, Adjusted R-squared:  -0.00545
F-statistic: 0.986 on 2 and 3 DF,  p-value: 0.469

Analysis of Variance Table

Response: s
            Df Sum Sq Mean Sq F value Pr(>F)
FO(zA, zB)   2  1.380   0.690    0.99  0.469
Residuals    3  2.099   0.700
Lack of fit  2  2.096   1.048  427.78  0.034
Pure error   1  0.002   0.002

Direction of steepest ascent (at radius 1):
      zA       zB
-0.85974  0.51074

Corresponding increment in original units:
      xA       xB
-0.85974  2.55368

> steepest(model1, dist = seq(0, 5, by = 1), descent = F)

Path of steepest ascent from ridge analysis:
  dist     zA    zB |     xA     xB |  yhat
1    0  0.000 0.000 | 10.500 36.000 | 5.108
2    1 -0.860 0.511 |  9.640 38.555 | 5.696
3    2 -1.720 1.021 |  8.780 41.105 | 6.283
4    3 -2.579 1.532 |  7.921 43.660 | 6.870
5    4 -3.439 2.043 |  7.061 46.215 | 7.458
6    5 -4.299 2.554 |  6.201 48.770 | 8.046
```

In Table 16.20, the analysis of variance table includes Type I sums of squares for blocks, linear or first order terms, two way interaction terms, and pure quadratic terms, adding the terms to the model in that order. These Type I sums of squares would be the same modeling either coded or uncoded factor levels. A lack-of-fit test is also provided. Observe that the cross product terms are not significantly different from zero, and there is no significant lack of fit of the model. The linear and quadratic components are clearly significant. Looking at the parameter estimates and corresponding t-tests in Table 16.20, clearly the linear and quadratic methanol content terms are needed in the model, providing some clarification to the analysis of variance results.

Table 16.19 R program and output for second-order response surface regression: data entry and coding

```
> pah.data = read.table("data/pah.txt", header = T)
> pah.data$fBlock = factor(pah.data$Block)
> library(rsm)
> pah.ccd.data = coded.data(pah.data, zP ~ (Pres - 200)/100,
+   zT ~ (Temp - 70)/30, zET ~ (ET - 35)/25, zMC ~ (MC - 10)/10)
> pah.ccd.data
   Run Block Pres Temp   ET MC     y fBlock
1    1     1  250   55 47.5 15 391.8      1
2    2     1  150   85 47.5 15 413.6      1
3    3     1  250   55 22.5  5  68.7      1
4    4     1  250   85 47.5  5 143.0      1
5    5     1  150   85 22.5  5 104.0      1
6    6     1  150   55 22.5 15 309.1      1
7    7     1  200   70 35.0 10 400.6      1
8    8     1  250   85 22.5 15 402.5      1
9    9     1  150   55 47.5  5  77.7      1
10  10     1  200   70 35.0 10 426.5      1
11  11     2  250   85 47.5 15 457.5      2
12  12     2  150   55 22.5  5  56.9      2
13  13     2  250   85 22.5  5  94.1      2
14  14     2  250   55 22.5 15 409.7      2
15  15     2  150   55 47.5 15 410.9      2
16  16     2  150   85 22.5 15 375.8      2
17  17     2  150   85 47.5  5 110.5      2
18  18     2  200   70 35.0 10 387.8      2
19  19     2  250   55 47.5  5 103.0      2
20  20     2  200   70 35.0 10 399.1      2
21  21     3  200   70 35.0 10 416.9      3
22  22     3  200   40 35.0 10 359.8      3
23  23     3  200   70 10.0 10 276.1      3
24  24     3  200   70 60.0 10 462.3      3
25  25     3  100   70 35.0 10 311.5      3
26  26     3  200   70 35.0 10 346.5      3
27  27     3  200   70 35.0  0  46.8      3
28  28     3  200   70 35.0 20 418.7      3
29  29     3  200  100 35.0 10 413.9      3
30  30     3  300   70 35.0 10 429.4      3

Data are stored in coded form using these coding formulas ...
zP ~ (Pres - 200)/100
zT ~ (Temp - 70)/30
zET ~ (ET - 35)/25
zMC ~ (MC - 10)/10
```

Source The data in the program are reprinted from Barnabas et al. (1995) with permission. Copyright © 1995 American Chemical Society

In Table 16.21, the canonical analysis is shown, including: the stationary point expressed in terms of both coded and uncoded units; the eigenvalues, or canonical coefficients; and the eigenvectors, giving the direction of each canonical axis. The canonical coefficients and axes are with respect to the coded factor levels.

For this experiment, all eigenvalues (canonical coefficients) are negative, so the stationary point is a maximum. The eigenvectors are each scaled to be of length one. The last eigenvalue, with value −227.865, is the largest in magnitude. For the corresponding eigenvector, the primary component is

Table 16.20 R program and output for second-order response surface regression: parameter estimates and analysis of variance

```
> # Second-order model and analysis
> model2 = rsm(y ~ fBlock + SO(zP, zT, zET, zMC), data = pah.ccd.data)
> summary(model2)

Call:
rsm(formula = y ~ fBlock + SO(zP, zT, zET, zMC), data = pah.ccd.data)

            Estimate Std. Error t value Pr(>|t|)
(Intercept)   369.16      28.46   12.97  8.2e-09
fBlock2         6.78      26.35    0.26  0.80094
fBlock3        74.44      26.35    2.83  0.01431
zP             37.30      24.05    1.55  0.14492
zT             31.78      24.05    1.32  0.20911
zET            54.97      24.05    2.29  0.03972
zMC           263.07      24.05   10.94  6.3e-08
zP:zT          -6.35      58.91   -0.11  0.91581
zP:zET        -11.65      58.91   -0.20  0.84630
zP:zMC         23.10      58.91    0.39  0.70133
zT:zET          2.30      58.91    0.04  0.96945
zT:zMC         -4.35      58.91   -0.07  0.94226
zET:zMC        16.55      58.91    0.28  0.78319
zP^2          -88.63      44.99   -1.97  0.07056
zT^2          -72.23      44.99   -1.61  0.13246
zET^2         -89.88      44.99   -2.00  0.06714
zMC^2        -226.33      44.99   -5.03  0.00023

Multiple R-squared:  0.928, Adjusted R-squared:  0.839
F-statistic: 10.5 on 16 and 13 DF,  p-value: 0.0000582

Analysis of Variance Table

Response: y
                      Df Sum Sq Mean Sq F value    Pr(>F)
fBlock                 2  33884   16942    4.88    0.0262
FO(zP, zT, zET, zMC)   4 447761  111940   32.25 0.0000012
TWI(zP, zT, zET, zMC)  6   1008     168    0.05    0.9993
PQ(zP, zT, zET, zMC)   4  99227   24807    7.15    0.0029
Residuals             13  45118    3471
Lack of fit           10  42240    4224    4.40    0.1246
Pure error             3   2877     959
```

that of MC with value 0.994301. So, the fitted model has greatest curvature at the stationary point when moving in either direction determined by this fourth eigenvector, which is nearly parallel to the MC-axis, as is evident from the SAS contour plots in Fig. 16.10 (p. 601), where MC is the x-axis variable of plots (b), (e) and (f). To generate similar plots using R, add the following code to the end of the R program in Tables 16.19 and 16.20.

```
par(mfrow = c(3, 2))
contour(model2, ~ zP + zT + zET + zMC,
              at = round(xs(model2), 3), las = 1)
```

Table 16.21 R program output for second-order response surface regression: canonical analysis

```
Stationary point of response surface:
     zP      zT     zET     zMC
0.25947 0.19593 0.34721 0.60522

Stationary point in original units:
   Pres    Temp      ET      MC
225.947  75.878  43.680  16.052

Eigenanalysis:
$values
[1] -71.257  -84.168  -93.761 -227.865

$vectors
          [,1]       [,2]        [,3]        [,4]
zP    0.233360   0.712899   -0.655835   -0.084838
zT   -0.964432   0.255316   -0.067269    0.012632
zET  -0.121729  -0.652941   -0.744877   -0.063313
zMC   0.024413   0.016008   -0.102535    0.994301
```

The `contour` statement above generates six black-and-white contour plots—one for each pair of the four factors listed. The `par` statement causes the six contour plots to be arranged in a 3×2 layout; otherwise, one would obtain six separate plots. The option `at = xs(model2)` would use the stationary point rather than the center point to fix the levels of the unplotted factors, whereas the option `at = round(xs(model2),3)` does the same but rounds each component to 3 digits, providing cleaner subheadings. For color plots, include the option `image = T`.

16.8.3 Generating Designs

In this section, we illustrate some capabilities of the R software to generate response surface designs, using functions of the `rsm` package. Table 16.22 contains sample R code generating most of the designs used in examples in this chapter.

The function `cube` generates first order designs consisting of factorial (or cube) points and center points. For example, the statement

```
cube(basis = ~ A+B+C, generators = c(D ~ B*C, E ~ A*D, F ~ A*B),
     reps = 4, n0 = 0)
```

generates the design for the paint experiment introduced in Example 16.2.1. The syntax `basis = ~ A+B+C` causes inclusion of all eight combinations of A, B and C at two levels each, then the `generators` option defines the levels of D, E and F in terms of prior variables using the generators BCD, ADE and ABF. Each of the resulting eight treatment combinations is replicated four times, since `reps = 4`, and the design includes `n0 = 0` center points.

Central composite designs can be generated either in one or two steps, as illustrated in the code in Table 16.22 for the acid copper plating experiment of Example 16.2.4. The `ccd` function generates a complete central composite design. Alternatively, the `cube` function generates the first order design, the `star` function generates the axial or star points and any additional center points, and the `djoin` function joins these factorial and axial parts together. The default is for the design to have two blocks—one for each of the factorial and axial parts—unless `oneblock = T` is specified. The design generated

Table 16.22 Sample R code for generating response surface designs used in the chapter

```
library(rsm)

# Paint experiment
cube(basis = ~ A+B+C, generators = c(D ~ B*C, E ~ A*D, F ~ A*B),
     n0 = 0, reps = 4)

# Paint experiment: same design, different notation for factors
cube(3, generators = c(x4 ~ x2*x3, x5 ~ x1*x4, x6 ~ x1*x2),
     n0 = 0, reps = 4)

# Acid copper pattern plating experiment: rotatable CCD
# Create design in 2 parts, then join the parts
dsgn1 = cube(2, n0 = 2, reps = 1,
             coding = list(x1 ~ (A - 10.5)/1, x2 ~ (B - 36)/5))
dsgn2 = star(dsgn1, n0 = 0, alpha = "rotatable", reps = 1)
dsgn12 = djoin(dsgn1, dsgn2)
dsgn12 # Show design in randomized order
stdorder(dsgn12) # Show design in standard order
# Create the design all at once: defaults is factorial and axial blocks
dsgn = ccd(2, n0 = c(2, 0), alpha = "rotatable", oneblock = T,
  randomize = F, coding = list(x1 ~ (A - 10.5)/1, x2 ~ (B - 36)/5))
varfcn(dsgn, ~ SO(x1, x2), contour = T) # Contour of scaled variances
# Add data to the coded data set containing the design
dsgn$s = c(5.60, 4.84, 6.45, 5.19, 4.32, 4.25, 5.76, 4.42, 5.46, 5.81)
# Data analysis
model2 = rsm(s ~ SO(x1, x2), data = dsgn)
summary(model2)
contour(model2, ~ x1 + x2, at = round(xs(model2), 3), image = T,
              las = 1)

# Flour experiment: orthogonal CCD
ccd(basis = ~ B+C+D, n0 = 2, alpha = 1.2872, randomize = F)

# PAH experiment: rotatable CCD, with orthogonal blocking
ccd(4, alpha = "rotatable", n0 = 2, blocks = ~ x1*x2*x3*x4,
       randomize = F)

# Flour experiment: noise array only
cube(basis = ~ G+J, generators = c(K ~ G*J), n0 = 0, randomize = F)

# Box-Behnken example design: 4 factors, 1 ctr pt per block
bbd(4, n0 = 3, block = F, randomize = F)
# Same design, except 3 blocks
bbd(4, n0 = 1, block = T, randomize = F)
```

is in terms of coded variables, though the `coding` option allows the user to specify the coding formula relating each variable to its coded levels. By default, the design is randomized separately within each block, in which case the function `stdorder` can be used to display the design in standard, or unrandomized, order. One can specify a specific value of α (e.g. `alpha = 1.2872` for the flour experiment), or request that the value be set corresponding to a desired design property, such as: `alpha = "rotatable"` for a rotatable design; `alpha = "orthogonal"` for orthogonal blocking (not orthogonality); `alpha = "spherical"` for the axial and factorial points to be the same distance from the center points, and `alpha = "faces"` for the axial points to be on the faces of the cube (same as `alpha = 1`). Given a design, the variance function `varfcn` can generate a contour plot of the scaled variance for a given design and model, so one can see rotatability or non-rotatability, for example.

Using the `bbd` function, one can generate Box–Behnken designs for 3–7 factors, including designs with orthogonal blocking for either four factors and three blocks or five factors and two blocks.

For sake of completeness, the code for the acid copper pattern plating experiment illustrates one way to add data to the coded design data set, as needed for data analysis.

Exercises

1. **Paint experiment, continued**

   The paint experiment of Eibl et al. (1992) was discussed in Example 16.2.1 (p. 569), where the first-order model was fitted to the data. For the fitted first-order model, do the following.

   (a) Plot the residuals versus run order, and use the plot to check the independence assumption. (The order of the observations was not randomized in this experiment. Rather, the observations were collected in the order they are shown row by row in Table 16.1, p. 570.)
   (b) Plot the residuals versus the predicted values, and use the plot to check the assumption of equal variance.
   (c) Plot the residuals versus their normal scores, and use the plot to check the normality assumption.
   (d) Verify that the design is orthogonal.

2. **Paint followup experiment**

   The data of the second paint experiment described by Eibl et al. (1992) are given in Table 16.23. This experiment involves factors A–D, as these had significant effects in the first experiment (Example 16.2.1). The factors are

   A: belt speed B: tube width
   C: pump pressure D: paint viscosity

   All four factors are at lower levels than in the first experiment. Lowering the levels of factors B–D was indicated by the analysis of the first experiment. Lowering the level of factor A was based on a conjecture of the experimenters.

   (a) The experiment consists of two replicates of a half-fraction. Find the defining relation for the half-fraction.
   (b) Fit the first-order model, recoding the factor levels as ± 1.
   (c) Test for lack of fit of the first-order model.

Table 16.23 Paint thickness $y_{\mathbf{z}l}$ for the paint followup experiment

z_A	z_B	z_C	z_D	$y_{\mathbf{z}1}$	$y_{\mathbf{z}2}$
−1.5	0	−2	0	1.71	1.61
0.5	0	−2	0	0.91	1.30
−1.5	−2	0	0	1.71	1.60
0.5	−2	0	0	1.15	1.29
−1.5	0	0	−2	1.33	1.06
0.5	0	0	−2	1.74	1.98
−1.5	−2	−2	−2	0.64	0.78
0.5	−2	−2	−2	1.51	1.18

Source Eibl et al. (1992). Reprinted with Permission from Journal of Quality Technology © 1992 ASQ, www.asq.org

(d) What would you recommend the experimenters do next?

3. **Fractionation experiment**

Sosada (1993) studied the effects of extraction time, solvent volume, ethanol concentration, and temperature on the yield and phosphatidylcholine enrichment (PCE) of deoiled rapeseed lecithin when fractionated with ethanol. Initially, a single-replicate 2^4 experiment was conducted, augmented by three center points.

(a) The results for the 16 factorial points are shown as the first 16 runs in Table 16.24. Fit the first-order model for the response variable "PCE" and conduct the corresponding analysis of variance.
(b) The design also included $n_0 = 3$ center-point observations of PCE. The sample variance of these three observations was $s_0^2 = 1.120$. Test the first-order model for lack of fit, using a 5% level of significance. (Hint: Since the factorial points include no replication, $msPE = s_0^2$, and ssE based on all 19 runs is equal to ssE from the factorial portion of the design plus $(n_0 - 1)s_0^2$.)
(c) Based on the results of parts (a) and (b), what subsequent experimentation would you recommend?

4. **Fractionation experiment, continued**

The fractionation experiment was described in Exercise 3, where the response PCE was used. Consider here, instead, the analysis of "Yield".

(a) Fit the first-order model for the response variable "Yield" based on the initial first-order 2^4 factorial design, shown as the first 16 runs in Table 16.24. Conduct the corresponding analysis of variance.
(b) At the design center point, three additional observations were collected, for which the sample variance was $s_0^2 = 0.090$. Test the first-order model for lack of fit, using a 5% level of significance. (Hint: Since the factorial points include no replication, $msPE = s_0^2$, and ssE based on all 19 runs is equal to ssE from the factorial portion of the design plus $(n_0 - 1)s_0^2$.)
(c) Based on the results of parts (a) and (b), what subsequent experimentation would you recommend?

Table 16.24 Purified lecithin yield and phosphatidylcholine enrichment (PCE), given extraction time (z_1), solvent volume (z_2), ethanol concentration (z_3), and temperature (z_4); fractionation experiment

Run	z_1	z_2	z_3	z_4	Yield	PCE
1	1	1	1	1	27.6	43.8
2	−1	−1	1	1	16.6	27.2
3	1	−1	−1	1	15.4	23.6
4	−1	1	−1	1	17.4	26.2
5	1	−1	1	−1	17.0	27.8
6	−1	1	1	−1	19.0	30.2
7	1	1	−1	−1	17.4	25.2
8	−1	−1	−1	−1	12.6	18.8
9	1	−1	1	1	18.6	28.8
10	−1	1	1	1	22.4	36.8
11	1	1	−1	1	21.4	33.4
12	−1	−1	−1	1	14.0	21.0
13	1	1	1	−1	24.0	38.0
14	−1	−1	1	−1	15.6	23.6
15	1	−1	−1	−1	13.0	20.2
16	−1	1	−1	−1	14.4	22.6
17	0	0	0	0	22.6	
18	$\sqrt{2}$	0	0	0	23.4	
19	$-\sqrt{2}$	0	0	0	20.6	
20	0	$\sqrt{2}$	0	0	22.6	
21	0	$-\sqrt{2}$	0	0	13.4	
22	0	0	$\sqrt{2}$	0	20.6	
23	0	0	$-\sqrt{2}$	0	15.6	
24	0	0	0	$\sqrt{2}$	21.0	
25	0	0	0	$-\sqrt{2}$	17.6	

Source Sosada (1993). Copyright © 1993 American Oil Chemists Society. Reprinted with permission

5. **Fractionation experiment, continued**

The fractionation experiment was described in Exercise 3, and analysis of the first-order model for "Yield" was considered in Exercise 4. Based on the analysis of the first-order design, the experimenter chose to augment the 16 factorial points of the first-order design into a 25-run central composite design, the yields from which are shown in Table 16.24.

(a) Determine whether the central composite design used is rotatable or orthogonal.
(b) Fit the second-order response surface model and determine which effects are significantly different from zero.
(c) Conduct a canonical analysis and discuss the results with respect to the following items. What is the nature of the critical point? Noting that the objective is to increase yield, in what direction should one move in subsequent experimentation?

6. **Film viscosity experiment**

Cuq et al. (1995, *Journal of Food Science*) used a central composite design to study the effects of protein concentration (g/100 g solution), pH, and temperature (°C), denoted by P, H, and T,

respectively, on the apparent viscosity Y (mPa) of film-forming solution, in the development of edible packaging films based on fish myofibrillar proteins. The data are shown in Table 16.25.

(a) Is this central composite design rotatable or orthogonal?
(b) Fit the second-order model to the data using the coded factor levels, and check the model assumptions. Would you recommend that a transformation of the data be taken?
(c) Fit the second-order model to the natural log of the data, $\ln(y)$, using the coded factor levels.
(d) Conduct the test for lack of fit of the second-order model for $\ln(y)$.
(e) Check the model assumptions for $\ln(y)$.
(f) Conduct the canonical analysis for $\ln(y)$.
(g) Conduct the analysis of variance for $\ln(y)$.
(h) Compute the coefficient of multiple determination R^2 for the second-order model for $\ln(y)$.
(i) Assess the results of the experiment, based on the model for $\ln(y)$.

7. **Flour production experiment, continued**

Consider again the flour production experiment of Sect. 16.5. The data were given in Table 16.11 (p. 590), along with the statistics $\overline{y}_{.\mathbf{z}}$ and $100\log_{10}(s_{\mathbf{z}})$ computed for the observations at each design-factor combination $\mathbf{z}$.

(a) Plot $\log_{10}(s_{\mathbf{z}})$ versus $\log_{10}(\overline{y}_{.\mathbf{z}})$, and use the methods of Sect. 5.6.2 to determine an appropriate variance-stabilizing transformation for these data. (Use of $\log_{10}$ is equivalent to use of ln for choosing a transformation.)

Table 16.25 Apparent viscosity $y_{\mathbf{z}l}$ of film-forming solution, for combinations of levels of protein concentration (g/100 g solution), pH, and temperature (°C)

Design point	P		H		T		
	z_P	x_P	z_H	x_H	z_T	x_T	y
1	−1	1.25	−1	2.75	−1	20	50
2	1	2.75	−1	2.75	−1	20	48
3	−1	1.25	1	3.25	−1	20	16700
4	1	2.75	1	3.25	−1	20	560
5	−1	1.25	−1	2.75	1	40	320
6	1	2.75	−1	2.75	1	40	18
7	−1	1.25	1	3.25	1	40	19000
8	1	2.75	1	3.25	1	40	5000
9	−2	0.50	0	3.00	0	30	12700
10	2	3.50	0	3.00	0	30	182
11	0	2.00	−2	2.50	0	30	14
12	0	2.00	2	3.50	0	30	27800
13	0	2.00	0	3.00	−2	10	133
14	0	2.00	0	3.00	2	50	4300
15	0	2.00	0	3.00	0	30	57
16	0	2.00	0	3.00	0	30	70
17	0	2.00	0	3.00	0	30	58
18	0	2.00	0	3.00	0	30	56

Source Cuq et al. (1995). Copyright © 1995 Inst. of Food Technologists. Reprinted with permission

Table 16.26 Resin impurity content $y_{\mathbf{z}t}$ (ppm)

Design point	Time	Temp.	$y_{\mathbf{x},t}$
1	7.0	232.4	18.5
2	3.0	220.0	22.5
3	11.0	220.0	17.2
4	1.3	190.0	42.2
5	7.0	190.0	28.6
6	7.0	190.0	19.8
7	7.0	190.0	23.6
8	7.0	190.0	24.1
9	7.0	190.0	24.2
10	12.7	190.0	19.1
11	3.0	160.0	54.1
12	11.0	160.0	33.8
13	7.0	147.6	55.4

(b) Repeat the first analysis of variance of Sect. 16.5, for which the response variable was $\overline{y}_{\cdot\mathbf{z}}$, after applying the transformation determined in part (a) to the observations $y_{h\mathbf{z}}$. Compare your conclusions with those reached in Sect. 16.5.
(c) Repeat the second analysis of variance of Sect. 16.5, for which the response variable was $100 \log_{10}(s_{\mathbf{z}})$, after applying the transformation determined in part (a) to the observations $y_{h\mathbf{z}}$. Compare your conclusions to those reached in Sect. 16.5.

8. **Central composite design**

Consider using a central composite design for three factors, to include eight factorial points and six axial points.

(a) Determine the value of α to make the design rotatable.
(b) Investigate how α and the number of center points should be chosen to make the design both rotatable and orthogonal, if possible. If this is not possible, how can the design be made rotatable and nearly orthogonal?
(c) Investigate whether the design can be rotatable with orthogonal blocking. If not, then investigate whether orthogonal blocking is possible. If so, how many blocks could be used? Investigate whether orthogonal blocking and near rotatability is possible.

9. **Central composite design**

Repeat Exercise 8 for a central composite design for four factors, to include 16 factorial points and eight axial points.

10. **Resin impurity experiment**

An experiment was conducted using a design close to a central composite design to study the effects of drying time (hours) and temperature (°C) on the content y (ppm) of undesirable compounds in a resin. The data are shown in Table 16.26.

Table 16.27 Resin degradation (ppm) for the resin moisture experiment

Design point	T		H		P		
	z_T	x_T	z_H	x_H	z_P	x_P	y
1	−1	150	−1	0	0	4	83
2	−1	150	0	50	−1	1	103
3	−1	150	0	50	1	9	94
4	−1	150	1	100	0	4	98
5	0	185	−1	0	−1	1	51
6	0	185	−1	0	1	9	48
7	0	185	1	100	−1	1	106
8	0	185	1	100	1	9	108
9	1	220	−1	0	0	4	36
10	1	220	0	50	−1	1	153
11	1	220	0	50	1	9	107
12	1	220	1	100	0	4	87
13	0	185	0	50	0	4	80
14	0	185	0	50	0	4	81
15	0	185	0	50	0	4	77
16	0	185	0	50	0	4	80
17	0	185	0	50	0	4	82

(a) Determine the coded levels of time and temperature, as well as the values of n_f, n_a, n_0. What values of α for each factor were selected by the experimenters for the axial points? Why is the design not quite a central composite design?
(b) Fit the second-order model, using coded factor levels.
(c) Test for model lack of fit.
(d) Check the equal variance and normality assumptions of the model using residual plots.
(e) Conduct the canonical analysis.
(f) Conduct the analysis of variance.
(g) Summarize the results.

11. **Resin moisture experiment**

A Box–Behnken design was used to determine whether specific drying conditions for a process could yield a resin that is sufficiently devoid of moisture and low-molecular-weight components. The three factors T, H, and P under study were temperature (150, 185, 220°C), relative humidity (0, 50, 100%), and air pressure (1, 5, 9 torr). The response variable y was a measure of product degradation (ppm). The design and data are shown in Table 16.27.

(a) Fit the second-order model, using coded factor levels.
(b) Test for model lack of fit.
(c) Check the equal-variance and normality model assumptions using residual plots.
(d) Conduct the canonical analysis.
(e) Conduct the analysis of variance.
(f) Summarize the results.

12. **Box–Behnken design**

(a) Construct a Box–Behnken design for three factors based on the balanced incomplete block design for three treatments in three blocks of size two and the 2^2 factorial design.
(b) Determine whether the design constructed in part (a) is rotatable.
(c) For the design constructed in part (a), determine whether orthogonal blocking is possible.

13. **Box–Behnken design**

(a) Construct a Box–Behnken design for five factors based on the balanced incomplete block design for five treatments in 10 blocks of size two.
(b) Determine whether the design constructed in part (a) is rotatable.
(c) For the design constructed in part (a), determine whether orthogonal blocking is possible.