RESPONSE SURFACE DESIGNS

Rajender Parsad and P.K. Batra
I.A.S.R.I., Library Avenue, New Delhi-110 012
rajender@iasri.res.in; pkbatra@iasri.res.in

1. Introduction
The subject of Design of Experiments deals with the statistical methodology needed for making inferences about the treatment effects on the basis of responses (univariate or multivariate) collected through the planned experiments. To deal with the evolution and analysis of methods for probing into mechanism of a system of variables, the experiments involving several factors simultaneously are being conducted in agricultural, horticultural and allied sciences. Data from experiments with levels or level combinations of one or more factors as treatments are normally investigated to compare level effects of the factors and also their interactions. Though such investigations are useful to have objective assessment of the effects of levels actually tried in the experiment, this seems to have inadequate, especially when the factors are quantitative in nature. The above analysis cannot give any information regarding the possible effects of the intervening levels of the factors or their combinations, i.e., one is not able to interpolate the responses at the treatment combinations not tried in the experiment. In such cases, it is more realistic and informative to carry out investigations with the twin purposes:

a) To determine and to quantify the relationship between the response and the settings of a group of experimental factors.
b) To find the settings of the experimental factors that produces the best value or the best set of values of the response(s).

If all the factors are quantitative in nature, it is natural to think the response as a function of the factor levels and data from quantitative factorial experiments can be used to fit the response surfaces over the region of interest. Response surfaces besides inferring about the twin purposes can provide information about the rate of change of a response variable. They can also indicate the interactions between the quantitative treatment factors. The special class of designed experiments for fitting response surfaces is called response surface designs. A good response surface design should possess the properties viz., detectability of lack of fit, the ability to sequentially build up designs of increasing order and the use of a relatively modest, if not minimum, number of design points. Before formulating the problem mathematically, we shall give examples of some experimental situations, where response surface methodology can be usefully employed.

Example 1: The over-use of nitrogen (N) relative to Phosphorus (P) and Potassium (K) concerns both the agronomic and environmental perspective. Phosphatic and Potassic fertilizers have been in short supply and farmers have been more steadily adopting the use of nitrogenous fertilizers because of the impressive virtual response. There is evidence that soil P and K levels are declining. The technique of obtaining individual optimum doses for the N, P and K through separate response curves may also be responsible for unbalanced fertilizer use. Hence, determining the optimum and balanced dose of N, P and K for different crops has
been an important issue. This optimum and balanced dose should be recommended to farmers in terms of doses from the different sources and not in terms of the values of N, P and K alone, as the optimum combination may vary from source to source. However, in actual practice the values of N, P and K are given in terms of kg/ha rather than the combined doses along with the source of the fertilizers.

**Example 2:** For value addition to the agriculture produce, food-processing experiments are being conducted. In these experiments, the major objective of the experimenter is to obtain the optimum combination of levels of several factors that are required for the product. To be specific, suppose that an experiment related to osmotic dehydration of the banana slices is to be conducted to obtain the optimum combination of levels of concentration of sugar solution, solution to sample ratio and temperature of osmosis. The levels of the various factors are the following

<table>
<thead>
<tr>
<th>Factors</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Concentration of sugar solution</td>
<td>40%, 50%, 60%, 70% and 80%</td>
</tr>
<tr>
<td>2. Solution to sample ratio</td>
<td>1:1, 3:1, 5:1, 7:1 and 9:1</td>
</tr>
<tr>
<td>3. Temperature of osmosis</td>
<td>25°C, 35°C, 45°C, 55°C and 65°C</td>
</tr>
</tbody>
</table>

In this situation, response surface designs for 3 factors each at five equispaced levels can be used.

**Example 3:** Yardsticks (a measure of the average increase in production per unit input of a given improvement measure) of many fertilizers, manures, irrigation, pesticides for various crops are being obtained and used by planners and administrators in the formulation of policies relating to manufacture/import/subsidy of fertilizers, pesticides, development of irrigation projects etc.

The yardsticks have been obtained from the various factorial experiments. However, these will be more reliable and satisfy more statistical properties, if response surface designs for slope estimation are used.

In general response surface methodology is useful for all the factorial experiments in agricultural experimental programme that are undertaken so as to determine the level at which each of these factors must be set in order to optimize the response in some sense and factors are quantitative in nature. To achieve this we postulate that the response is a function of input variables, \( i.e. \)

\[
y_u = \phi(x_{1u}, x_{2u}, \ldots, x_{nu}) + e_u
\]

where \( u = 1, 2, \ldots, N \) represents the \( N \) observations and \( x_{iu} \) is the level of the \( i^{th} \) factor in the \( u^{th} \) observation. The function \( \phi \) describes the form in which the response and the input variables are related and \( e_u \) is the experimental error associated with the \( u^{th} \) observation such that \( E(e_u) = 0 \) and \( \text{Var}(e_u) = \sigma^2 \). Knowledge of function \( \phi \) gives a complete summary of the results of the experiment and also enables us to predict the response for values of the \( x_{iu} \) that are not included in the experiment. If the function \( \phi \) is known then using methods of calculus, one may obtain the values of \( x_1, x_2, \ldots, x_v \) which give the optimum (say, maximum) response.
In practice the mathematical form of $\phi$ is not known; we, therefore, often approximate it, within the experimental region, by a polynomial of suitable degree in variables $x_{iu}$. The adequacy of the fitted polynomial is tested through the usual analysis of variance. Polynomials which adequately represent the true dose-response relationship are called **Response Surfaces** and the designs that allow the fitting of response surfaces and provide a measure for testing their adequacy are called **response surface designs**. If the function $\phi$ in (1.1) is of degree one in $x_{iu}$'s i.e.

$$y_u = \beta_0 + \beta_1 x_{iu} + \beta_2 x_{2u} + \ldots + \beta_v x_{vu} + e_u$$  \hspace{1cm} (1.2)

we call it a first-order response surface in $x_1, x_2, \ldots, x_v$. If (1.1) takes the form

$$y_u = \beta_0 + \sum_{i=1}^{v} \beta_i x_{iu} + \sum_{i=1}^{v} \beta_i x_{i^2u}^2 + \sum_{i=1}^{v-1} \sum_{i'=i+1}^{v} \beta_{ii'} x_{iu} x_{i'u} + e_u$$  \hspace{1cm} (1.3)

We call it a second-order (quadratic) response surface. Henceforth, we shall concentrate on the second order response surface which is more useful in agricultural experiments.

2. **The Quadratic Response Surface**

The general form of a second-degree (quadratic) surface is

$$y_u = \beta_0 + \beta_1 x_{1u} + \beta_2 x_{2u} + \ldots + \beta_v x_{vu} + \beta_{11} x_{1u}^2 + \beta_{22} x_{2u}^2 + \ldots + \beta_{vv} x_{vu}^2 + \beta_{12} x_{1u} x_{2u} + \beta_{13} x_{1u} x_{3u} + \ldots + \beta_{v-1,v} x_{v-1,u} x_{vu} + e_u$$

Let us assume that $x_{iu}$'s satisfy the following conditions:

(A) \[ \sum_{u=1}^{N} \left( \prod_{i=1}^{v} x_{iu}^{\alpha_i} \right) = 0 \] if any $\alpha_i$ is odd, for $\alpha_i = 0,1,2 \text{ or } 3$ and $\sum \alpha_i \leq 4$.

(B) \[ \sum_{u=1}^{N} x_{iu}^2 = \text{constant (for all } i) = N\lambda_2 \text{ (say)} \]

(C) \[ \sum_{u=1}^{N} x_{iu}^4 = \text{constant (for all } i) = CN\lambda_4 \text{ (say)} \]  \hspace{1cm} (2.1)

(D) \[ \sum_{u=1}^{N} x_{iu}^2 x_{i'u}^2 = \text{constant } = N\lambda_4 \text{ (say), for all } i \neq i' \]

We shall estimate the parameters $\beta_i$'s through the method of least squares. Let $b_0, b_1's, b_{ii}'s, b_{ii'}'s$ denote the best linear unbiased estimate of $\beta_0, \beta_i's, \beta_{ii}'s, \beta_{ii'}'s$ respectively. Under the above restrictions on $x_{iu}$'s, the normal equations are found to be:
\[ \sum_{u=1}^{N} y_u = N b_0 + N \lambda_2 \sum_{i=1}^{v} b_{ii} \]
\[ \sum_{u=1}^{N} x_{iu} y_u = N \lambda_2 b_i \]
\[ \sum_{u=1}^{N} x_{iu} x_{i'u} y_u = N \lambda_4 b_{ii'} \]
\[ \sum_{u=1}^{N} x_{i'u}^2 y_u = N \lambda_2 b_0 + CN \lambda_4 b_{ii'} + N \lambda_4 \sum_{i' \neq i}^{v} b_{ii'} \]
\[ = N \lambda_2 b_0 + (C-1)N \lambda_4 b_{ii'} + N \lambda_4 \sum_{i=1}^{v} b_{ii} \]

Solving the above normal equations, we obtain the estimates \( b_i \)'s as
\[
b_0 = \left[ \lambda_4 (C + v - 1) \sum_{u=1}^{N} y_u - \lambda_2 \sum_{i=1}^{v} \sum_{u=1}^{N} x_{iu}^2 y_u \right] / N \Delta \\
b_i = \sum_{u=1}^{N} x_{iu} y_u / N \lambda_2 \\
b_{ii'} = \sum_{u=1}^{N} x_{iu} x_{i'u} y_u / N \lambda_4 \\
b_{ii} = \left[ \sum_{u=1}^{N} x_{iu}^2 y_u - \left( \lambda_2^2 - \lambda_4 \right) \sum_{u=1}^{N} x_{iu}^2 \right] / \left( \Delta / [(C-1)N \lambda_4] \right) \]

where \( \Delta = (C + v - 1) \lambda_4 - v \lambda_2^2 \).

The variances of and covariances between the estimated parameters are as follows:
\[
V(b_0) = \lambda_4 (C + v - 1) \sigma^2 / N \Delta \\
V(b_i) = \sigma^2 / N \lambda_2 \\
V(b_{ii'}) = \sigma^2 / N \lambda_4 \\
V(b_{ii}) = \sigma^2 \left[ 1 + \left( \lambda_2^2 - \lambda_4 \right) \right] / [(C-1)N \lambda_4] \\
Cov(b_0, b_{ii}) = -\lambda_2 \sigma^2 / N \Delta \\
Cov(b_{ii}, b_{ii'}) = (\lambda_2^2 - \lambda_4) \sigma^2 / [(C-1)N \lambda_4 \Delta] \\
(2.4)
\]

Other covariances are zero. From the above expressions it is clear that a necessary condition for the design to exist is that \( \Delta > 0 \). Thus, a necessary condition for a Second Order Design to exist is that
\[
\lambda_4 / \lambda_2^2 > v / (C + v - 1) \]

(2.5)
If $\hat{y}$ is the estimated response at any given experimental point $(x_{10}, x_{20}, \ldots, x_{v0})$, then the variance of $\hat{y}$ is given by

$$
V(\hat{y}) = V(b_0) + V(b_i) \left( \sum_{i=1}^{v} x_{i0}^2 \right) + V(b_{ii}) \left( \sum_{i=1}^{v} x_{i0}^4 \right) + V(b_{ii'}) \left( \sum_{i=1}^{v-1} \sum_{i' = i+1}^{v} x_{i0}^2 x_{i'0}^2 \right) + 2 \text{Cov}(b_0, b_{ii}) \left( \sum_{i=1}^{v} x_{i0}^2 \right) + 2 \text{Cov}(b_{ii}, b_{ii'}) \left( \sum_{i=1}^{v-1} \sum_{i' = i+1}^{v} x_{i0}^2 x_{i'0}^2 \right) (2.6)
$$

If $\sum_{i=1}^{v} x_{i0}^2 = d^2$, where $d$ is the distance of the point $(x_{10}, x_{20}, \ldots, x_{v0})$ from the origin, then we may write

$$
V(\hat{y}) = V(b_0) + d^2[V(b_i) + 2 \text{Cov}(b_0, b_{ii})] + d^4 V(b_{ii}) + \sum_{i=1}^{v-1} \sum_{i' = i+1}^{v} x_{i0}^2 x_{i'0}^2 [V(b_{ii'}) + 2 \text{Cov}(b_{ii}, b_{ii'})] - 2 V(b_{ii}) (2.7)
$$

From the above expression, it is clear that if the coefficient of $\sum_{i=1}^{v-1} \sum_{i' = i+1}^{v} x_{i0}^2 x_{i'0}^2$ is made equal to zero, the variance of the estimated response at $(x_{10}, x_{20}, \ldots, x_{v0})$ will be a function of $d$, the distance of the point $(x_{10}, x_{20}, \ldots, x_{v0})$ from the origin. Now, the coefficient of $\sum_{i=1}^{v-1} \sum_{i' = i+1}^{v} x_{i0}^2 x_{i'0}^2$ is

$$
V(b_{ii'}) + 2 \text{Cov}(b_{ii}, b_{ii'}) - 2 V(b_{ii}) = \sigma^2 \frac{1}{N\lambda_4} \left[ 1 + \frac{2(\lambda_2 - \lambda_4)}{\Delta(C-1)} - \frac{2}{(C-1)} \frac{\lambda_2}{\Delta} \right] (2.8)
$$

Obviously, this is zero, if and only if $C = 3$. Thus, when $C = 3$, the variance of the estimated response at a given point, the response being estimated through a design satisfying (A), (B), (C), (D), (E) becomes a function of the distance of that point from the origin. Such designs are called as Second Order Rotatable Designs (SORD). We may now formally define a SORD:

Let us consider $N$ treatment combinations (points) $\{x_{i'u}, i = 1, 2, \ldots, v, u = 1, 2, \ldots, N\}$ to form a design in $v$ factors, through which a Second-degree surface can be fitted. This design is said to be a SORD if the variance of the estimated response at any given point is a function of the distance of that point from the origin. The necessary and sufficient conditions for a set of points $\{x_{i'u}, i = 1, 2, \ldots, v, u = 1, 2, \ldots, N\}$ to form a SORD are
Response Surface Designs

(A’) \[ \sum_{u=1}^{N} \left( \prod_{u=1}^{V} x_{iu}^{\alpha_i} \right) = 0, \text{ if any } \alpha_i \text{ is odd, for } \alpha_i = 0,1,2 \text{ or } 3 \text{ and } \sum \alpha_i \leq 4. \]

(B’) \[ \sum_{u=1}^{N} x_{iiu}^2 = N\lambda_2 \]

(C’) \[ \sum_{u} x_{iu}^4 = \text{constant} = 3N\lambda_4 \]

(D’) \[ \sum_{u} x_{uu}^2 x_{i'u'u}^2 = N\lambda_4 ; \]

(E’) \[ \frac{\lambda_4}{\lambda_2^2} > v/(v+2) \]

The conditions (A’), (B’) and (D’) are same as conditions (A), (B) and (D) in (2.1).

We now prove the following.

**Lemma:** If a set of points \( \{ x_{iu}, \ i = 1,2,...,v, u = 1,2,...,N \} \), satisfying (A’), (B’), (C’) and (D’) are such that every point is equidistant from the origin, then

\[ \frac{\lambda_4}{\lambda_2^2} = v/(v+2) \quad (2.10) \]

**Proof:** Let \( d \) be the distance of any point from the origin. Then, since all the points are equidistant from the origin, we have

\[ d^2 = \frac{1}{N} \sum_{u=1}^{N} \left( \sum_{i=1}^{v} x_{iu}^2 \right) = v\lambda_2 \]

\[ d^4 = \frac{1}{N} \left( \sum_{u=1}^{N} \left( \sum_{i=1}^{v} x_{iu}^2 \right)^2 \right) \]

and

\[ = \frac{1}{N} \sum_{u=1}^{N} \left[ \sum_{i=1}^{v} x_{iu}^4 + 2 \sum_{i=1}^{v} \sum_{i'=i+1}^{v} x_{iu}^2 x_{i'u'}^2 \right] \]

\[ = 3v\lambda_4 + v(v-1)\lambda_4 \]

Thus, \( v^2\lambda_2^2 = 3v\lambda_4 + v(v-1)\lambda_4 \)

or, \( \lambda_4(v+2) - v\lambda_2^2 = 0 \)

An arrangement of points satisfying (A’), (B’), (C’) and (D’) but not (E’) is called a Second Order Rotatable Arrangement (SORA). A SORA can always be converted to an SORD by adding at least one central point.

A near stationary region is defined as a region where the surface slopes along the \( v \) variable axes are small compared to the estimate of experimental error. The stationary point of a near
stationary region is the point at which the slope of the response surface is zero when taken in all the directions. The coordinates of the stationary point \( x_0 = (x_{10}, x_{20}, \ldots, x_{v0}) \) are obtained by differentiating the following estimated response equation with respect to each \( x_i \) and equating the derivatives to zero and solving the resulting equations

\[
\hat{Y}(x) = b_0 + \sum_{i=1}^{v} b_i x_i + \sum_{i=1}^{v} b_{ii} x_i^2 + \sum_{i=1}^{v-1} \sum_{i'=i+1}^{v} b_{ii'} x_i x_{i'}
\]  

(2.11)

In matrix notation (2.11) can be written as

\[
\hat{Y}(x) = b_0 + x' b + x' B x
\]

(2.12)

where \( x = (x_1, x_2, \ldots, x_v)' \), \( b = (b_1, b_2, \ldots, b_v)' \) and

\[
B = \begin{bmatrix}
b_{11} & b_{12}/2 & \cdots & b_{1v}/2 \\
b_{12}/2 & b_{22} & \cdots & b_{2v}/2 \\
\vdots & \vdots & \ddots & \vdots \\
b_{1v}/2 & b_{2v}/2 & \cdots & b_{vv}
\end{bmatrix}
\]

From equation (2.12)

\[
\frac{\partial \hat{Y}(x)}{\partial x} = b + 2Bx
\]

(2.13)

The stationary point \( x_0 \) is obtained by equating (2.13) to zero and solving for \( x \), i.e.

\[
x_0 = -\frac{1}{2} B^{-1} b
\]

(2.14)

To find the nature of the surface at the stationary point we examine the second derivative of \( \hat{Y}(x) \). From (2.13)

\[
\frac{\partial^2 \hat{Y}(x)}{\partial x^2} = 2B \quad \text{(since } B \text{ is symmetric)}.
\]

The stationary point is a maximum, minimum or a saddle point according as \( B \) is negative definite, positive definite or indefinite matrix. If \( \lambda_1, \lambda_2, \ldots, \lambda_v \) represent the \( v \) eigenvalues of \( B \). Then it is easy to see that if \( \lambda_1, \lambda_2, \ldots, \lambda_v \) are

(i) All negative, then at \( x_0 \) the surface is a maximum
(ii) All positive, then at \( x_0 \) the surface is a minimum
(iii) of mixed signs, i.e. some are positive and others are negative, then \( x_0 \) is a saddle point of the fitted surface.

Furthermore, if \( \lambda_i \) is zero (or very close to zero), then the response does not change in value in the direction of the axis associated with \( x_i \) variable. The magnitude of \( \lambda_i \) indicates how quickly the response changes in the direction of axis associated with \( x_i \) variable.
The conditions in (2.1) and (2.9) help in fitting of the response surfaces and define some statistical properties of the design like rotatability. However, these conditions need not necessarily be satisfied before fitting a response surface. This can be achieved by using the software packages like the Statistical Analysis System (SAS). PROC RSREG fits a second order response surface design and locates the coordinates of the stationary point, predict the response at the stationary point and give the eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_v \) and the corresponding eigen vectors. It also helps in determining whether the stationary point is a point of maxima, minima or is a saddle point. The lack of fit of a second order response surface can also be tested using LACKFIT option under model statement in PROC RSREG. The lack of fit is tested using the statistic

\[
F = \frac{SS_{LOF}/(N'-p)}{SS_{PE}/(N-N')}
\]

(2.15)

where \( N \) is the total number of observations, \( N' \) is the number of distinct treatments and \( p \) is the number of terms included in the model. \( SS_{PE} \) (sum of squares due to pure error) has been calculated in the following manner: denote the \( l^{th} \) observation at the \( u^{th} \) design point by \( y_{lu} \), where \( l = 1, \ldots, r_u \) ( \( \geq 1 \)), \( u = 1, \ldots, N' \). Define \( \bar{y}_u \) to be average of \( r_u \) observations at the \( u^{th} \) design point. Then, the sum of squares for pure error is

\[
SS_{PE} = \sum_{u=1}^{N'} \sum_{l=1}^{r_u} (y_{lu} - \bar{y}_u)^2
\]

(2.16)

Then sum of squares due to lack of fit \( (SS_{LOF}) = \) sum of squares due to error - \( SS_{PE} \)

The analysis of variance table for a second order response surface design is given below.

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>S.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Due to regression coefficients</td>
<td>( 2v+\binom{v}{2} )</td>
<td>( \hat{b}<em>0 \sum</em>{u=1}^{N} y_u + \sum_{i=1}^{v} \hat{b}<em>i \left( \sum</em>{u=1}^{N} x_{iu} y_u \right) + \sum_{i=1}^{v} \hat{b}<em>{ii} \left( \sum</em>{u=1}^{N} x_{iiu}^2 y_u \right) + \sum_{i \neq i'} \hat{b}<em>{ii'} \left( \sum</em>{u=1}^{N} x_{iu} x_{u'i} y_u \right) ) - ( CF )</td>
</tr>
<tr>
<td>Error</td>
<td>( N-2v-\binom{v}{2}-1 )</td>
<td>By subtraction = ( SSE )</td>
</tr>
<tr>
<td>Total</td>
<td>( N-1 )</td>
<td>( \sum_{u=1}^{N} y_u^2 - CF )</td>
</tr>
</tbody>
</table>

In the above table \( CF = \) correction factor = \( \left( \frac{\text{Grand Total}}{N} \right)^2 \). For testing the lack of fit the sum of squares is obtained using (2.16) and then sum of squares is obtained by subtracting the sum of squares due to pure error from sum of squares due to error. The sum of squares due to
lack of fit and sum of squares due to pure error are based on \( N' - 2v - \left(\frac{v}{2}\right) - 1 \) and \( N - N' \) degrees of freedom respectively.

It is suggested that in the experiments conducted to find a optimum combination of levels of several quantitative input factors, at least one level of each of the factors should be higher than the expected optimum. It is also suggested that the optimum combination should be determined from response surface fitting rather than response curve fitting, if the experiment involves two or more than two factors.

3. Construction of Second Order Rotatable designs

A second order response surface design is at least resolution V fractional factorial design. Here

3.1 Central Composite Rotatable Designs

Let there be \( v \) factors in the design. A class of SORD for \( v \) factors can be constructed in the following manner. Construct a factorial \( v \)-factors with levels \( \pm \alpha \) containing \( 2^p \) combinations, where \( 2^p \) is the smallest fraction of \( 2^v \) without confounding any interaction of third order or less. Next, another \( 2v \) points of the following type are considered: \((\pm \beta 0 0 \ldots 0), (0 \pm \beta 0 \ldots 0), (0 0 \ldots \pm \beta)\). These \( N = 2^p + 2v \) points, give rise to a SORD in \( v \) factors with levels \( \pm \alpha, \pm \beta, 0 \). We have for this design,

\[
\sum_{u=1}^{N} x_{iu}^2 = 2^p \alpha^2 + 2 \beta^2
\]

\[
\sum_{u=1}^{N} x_{iu}^4 = 2^p \alpha^4 + 2 \beta^4
\]

\[
\sum_{u=1}^{N} x_{i0}^2 x_{i0}^2 = 2^p \alpha^4.
\]

On applying the condition of rotatability, we have

\[
3.2^p \alpha^4 = 2^p \alpha^4 + 2 \beta^4
\]

\[
\Rightarrow \beta^4 = \alpha^4 2^p
\]

\[
or \ \beta^2 / \alpha^2 = 2^{p/2}.
\]

This equation gives a relationship between \( \beta \) and \( \alpha \). For determining \( \alpha \) and \( \beta \) uniquely, we either fix \( \alpha = 1 \) or \( \lambda_2 = 1 \). For \( \alpha = 1 \), \( \Rightarrow \beta^2 = 2^{p/2} \).
**Example.** Let $v = 4$. Then the points of the SORD are

\[
\begin{array}{cccc}
-\alpha & -\alpha & -\alpha & -\alpha \\
-\alpha & -\alpha & -\alpha & \alpha \\
-\alpha & -\alpha & \alpha & -\alpha \\
-\alpha & -\alpha & \alpha & \alpha \\
-\alpha & \alpha & -\alpha & -\alpha \\
-\alpha & \alpha & -\alpha & \alpha \\
-\alpha & \alpha & \alpha & -\alpha \\
-\alpha & \alpha & \alpha & \alpha \\
\alpha & -\alpha & -\alpha & -\alpha \\
\alpha & -\alpha & -\alpha & \alpha \\
\alpha & -\alpha & \alpha & -\alpha \\
\alpha & -\alpha & \alpha & \alpha \\
\alpha & \alpha & -\alpha & -\alpha \\
\alpha & \alpha & -\alpha & \alpha \\
\alpha & \alpha & \alpha & -\alpha \\
\alpha & \alpha & \alpha & \alpha \\
\beta & 0 & 0 & 0 \\
-\beta & 0 & 0 & 0 \\
0 & \beta & 0 & 0 \\
0 & -\beta & 0 & 0 \\
0 & 0 & \beta & 0 \\
0 & 0 & -\beta & 0 \\
0 & 0 & 0 & \beta \\
0 & 0 & 0 & -\beta \\
0 & 0 & 0 & 0 \\
\end{array}
\]

There are 25 points – a central point has been added because, all the non-central points are equidistant from the origin, as $\beta = 2\alpha$, here.

### 3.2 Construction of SORD using BIB Designs

If there exists a BIB design $D$ with parameters $v^*, b^*, r^*, k^*, \lambda^*$ such that $r^* = 3\lambda^*$, then a SORD with each factor at 3 levels can be constructed.

Let $N^*$ be the $v^* \times b^*$ incidence matrix of $D$. Then $N^*$ is a matrix of order $b^* \times v^*$, every row of which contains exactly $k^*$ unities and every column contains exactly $r^*$ unities, rest positions being filled up by zeros. In $N^*$, replace the unity by $\alpha$. Then, we get $b^*$ combinations involving $\alpha$ and zero. Next, each of these combinations are ‘multiplied’ with those of a $2^k$ factorial with levels $\pm 1$ where, the term ‘multiplication’ means the multiplication of the corresponding entries in the two combinations, zero entries remaining unaltered. Thus, if $(\alpha \alpha 0)$ is multiplied by $(-1 -1)$ we get $(-\alpha -\alpha 0)$. The procedure of multiplication gives rise to $b^* 2^k$ points each of $v^*$-dimension. These points evidently satisfy all the conditions (A’), (B’), (C’) and (D’); however, since each point in the
arrangement is at the same distance from the origin, we have to take at least one central point
to get a SORD in $v = v^*$ factors. The levels of the factors are $\pm \alpha, 0$. The value of $\alpha$ can be
determined by fixing $\lambda_2 = 1$.

SORD’s can be constructed using BIB designs, even when $r^* \neq 3\lambda^*$. In the case, where
$r^* < 3\lambda^*$ the set of $b \cdot 2^{k^*}$ points obtained using $N^*$ is to be augmented with further
$2v^*$ points of the type
$$(\pm \beta 0 0 \ldots 0), (0 \pm \beta 0 \ldots 0), (0 0 \ldots \pm \beta)$$

For the $N$ points ($N = b \cdot 2^{k^*} + 2v^*$), we have
$$
\sum_{u} x_{i u}^4 = r \cdot 2^{k^*} \alpha^4 + 2\beta^4
$$
$$
\sum_{u} x_{i u}^2 x_{i u}^2 = \lambda \cdot 2^{k^*} \alpha^4.
$$

Thus $2\beta^4 + r \cdot 2^{k^*} \alpha^4 = 3\lambda \cdot 2^{k^*} \alpha^4
\frac{k^* - 1}{k^* - 1}$

or, $\beta^2 / \alpha^2 = (3\lambda - r^*)^{1/2} 2^{2}$

When $r^* > 3\lambda^*$, the points augmented are of type $(\pm \beta \pm \beta \ldots \pm \beta)$ and $2^p$ in number,
where $2^p$ is the smallest fraction of $2^{v^*}$ factorial with levels $\pm \beta$, such that no interaction of
order three or less is confounded. In this case,
$$
\sum_{u} x_{i u}^4 = r \cdot 2^{k^*} \alpha^4 + 2^p \beta^4
$$
$$
\sum_{u} x_{i u}^2 x_{i u}^2 = \lambda \cdot 2^{k^*} \alpha^4 + 2^p \beta^4.
$$

Thus, $3\lambda \cdot 2^{k^*} \alpha^4 + 3.2^{p} \beta^4 = r \cdot 2^{k^*} \alpha^4 + 2^p \beta^4$

or, $2^{p+1} \beta^4 = (r^* - 3\lambda^*)2^{k^*} \alpha^4$

which gives $\beta^2 / \alpha^2 = (r^* - 3\lambda^*)^{1/2} 2^{(k^* - p - 1)/2}$

In both the cases, we get $v^*$-factor SORD with each factor at five levels

4. **Practical Exercise**

**Exercise 1:** Consider an experiment that was conducted to investigate the effects of three
fertilizer ingredients on the yield of a crop under fields conditions using a second order
rotatable design. The fertilizer ingredients and actual amount applied were nitrogen (N), from
0.89 to 2.83 kg/plot; phosphoric acid (P$_2$O$_5$) from 0.265 to 1.336 kg/plot; and potash (K$_2$O),
from 0.27 to 1.89 kg/plot. The response of interest is the average yield in kg per plot. The
levels of nitrogen, phosphoric acid and potash are coded, and the coded variables are defined as

$$ X_1 = (N - 1.629)/0.716, \quad X_2 = (P_2O_5 - 0.796)/0.311, \quad X_3 = (K_2O - 1.089)/0.482 $$

II-281
The values 1.629, 0.796 and 1.089 kg/plot represent the centres of the values for nitrogen, phosphoric acid and potash, respectively. Five levels of each variable are used in the experimental design. The coded and measured levels for the variables are listed as:

<table>
<thead>
<tr>
<th>Levels of $x_i$</th>
<th>-1.682</th>
<th>-1.000</th>
<th>0.000</th>
<th>+1.000</th>
<th>+1.682</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>0.425</td>
<td>0.913</td>
<td>1.629</td>
<td>2.345</td>
<td>2.833</td>
</tr>
<tr>
<td>P$_2$O$_5$</td>
<td>0.266</td>
<td>0.481</td>
<td>0.796</td>
<td>1.111</td>
<td>1.326</td>
</tr>
<tr>
<td>K$_2$O</td>
<td>0.278</td>
<td>0.607</td>
<td>1.089</td>
<td>1.571</td>
<td>1.899</td>
</tr>
</tbody>
</table>

Six center point replications were run in order to obtain an estimate of the experimental error variance. The complete second order model to be fitted to yield values is:

$$Y = \beta_0 + \sum_{i=1}^{3} \beta_i x_i + \sum_{i=1}^{3} \beta_{ii} x_i^2 + \sum_{i=1}^{3} \sum_{i'=2}^{3} \beta_{ii'} x_i x_{i'} + e$$

The following table lists the design settings of $x_1$, $x_2$ and $x_3$ and the observed values at 15 design points N, P$_2$O$_5$, K$_2$O and yield are in kg.

**Table 2: Central Composite Rotatable Design Settings in the Coded Variables $x_1$, $x_2$ and $x_3$, the original variables N, P$_2$O$_5$, K$_2$O and the Average Yield of a Crop at Each Setting**

<table>
<thead>
<tr>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>N</th>
<th>P$_2$O$_5$</th>
<th>K$_2$O</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0.913</td>
<td>0.481</td>
<td>0.607</td>
<td>5.076</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>2.345</td>
<td>0.481</td>
<td>0.607</td>
<td>3.798</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>0.913</td>
<td>1.111</td>
<td>0.607</td>
<td>3.798</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>2.345</td>
<td>1.111</td>
<td>0.607</td>
<td>3.469</td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0.913</td>
<td>0.481</td>
<td>1.571</td>
<td>4.023</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>2.345</td>
<td>0.481</td>
<td>1.571</td>
<td>4.905</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>0.913</td>
<td>1.111</td>
<td>1.571</td>
<td>5.287</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2.345</td>
<td>1.111</td>
<td>1.571</td>
<td>4.963</td>
</tr>
<tr>
<td>-1.682</td>
<td>0</td>
<td>0</td>
<td>0.425</td>
<td>0.796</td>
<td>1.089</td>
<td>3.541</td>
</tr>
<tr>
<td>1.682</td>
<td>0</td>
<td>0</td>
<td>2.833</td>
<td>0.796</td>
<td>1.089</td>
<td>3.541</td>
</tr>
<tr>
<td>0</td>
<td>-1.682</td>
<td>0</td>
<td>1.629</td>
<td>0.266</td>
<td>1.089</td>
<td>5.436</td>
</tr>
<tr>
<td>0</td>
<td>1.682</td>
<td>0</td>
<td>1.629</td>
<td>1.326</td>
<td>1.089</td>
<td>4.977</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>-1.682</td>
<td>1.629</td>
<td>0.796</td>
<td>0.278</td>
<td>3.591</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1.682</td>
<td>1.629</td>
<td>0.796</td>
<td>1.899</td>
<td>4.693</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.629</td>
<td>0.796</td>
<td>1.089</td>
<td>4.563</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.629</td>
<td>0.796</td>
<td>1.089</td>
<td>4.599</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.629</td>
<td>0.796</td>
<td>1.089</td>
<td>4.599</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.629</td>
<td>0.796</td>
<td>1.089</td>
<td>5.188</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.629</td>
<td>0.796</td>
<td>1.089</td>
<td>4.959</td>
</tr>
</tbody>
</table>
OPTIONS LINESIZE = 72;
DATA RP;
INPUT N P K YIELD;
CARDS;
....
....
....
;
PROC RSREG;
MODEL YIELD = N P K /LACKFIT NOCODE;
RUN;

Response Surface for Variable YIELD
Response Mean 4.464050
Root MSE 0.356424
R-Square 0.8440
Coef. of Variation 7.9843

<table>
<thead>
<tr>
<th>Regression</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>R-Square</th>
<th>F-Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>3</td>
<td>1.914067</td>
<td>0.2350</td>
<td>5.022</td>
<td>0.0223</td>
</tr>
<tr>
<td>Quadratic</td>
<td>3</td>
<td>3.293541</td>
<td>0.4044</td>
<td>8.642</td>
<td>0.0040</td>
</tr>
<tr>
<td>Crossproduct</td>
<td>3</td>
<td>1.666539</td>
<td>0.2046</td>
<td>4.373</td>
<td>0.0327</td>
</tr>
<tr>
<td>Total Regression</td>
<td>9</td>
<td>6.874147</td>
<td>0.8440</td>
<td>6.012</td>
<td>0.0049</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regression</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>R-Square</th>
<th>F-Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of Fit</td>
<td>5</td>
<td>0.745407</td>
<td>0.149081</td>
<td>1.420</td>
<td>0.3549</td>
</tr>
<tr>
<td>Pure Error</td>
<td>5</td>
<td>0.524973</td>
<td>0.104995</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Error</td>
<td>10</td>
<td>1.270380</td>
<td>0.127038</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Parameter      | d.f. | Estimate  | Std Error | T-ratio | Prob > |T| |
|----------------|------|-----------|-----------|---------|---------|
| INTERCEPT      | 1    | 6.084180  | 1.543975  | 3.941   | 0.0028  |
| N              | 1    | 1.558870  | 0.854546  | 1.824   | 0.0981  |
| P              | 1    | -6.009301 | 2.001253  | -3.003  | 0.0133  |
| K              | 1    | -0.897830 | 1.266909  | -0.709  | 0.4947  |
| N*N            | 1    | -0.738716 | 0.183184  | -4.033  | 0.0024  |
| P*N            | 1    | -0.142436 | 0.558725  | -0.255  | 0.8039  |
| P*P            | 1    | 2.116594  | 0.945550  | 2.238   | 0.0491  |
| K*N            | 1    | 0.784166  | 0.365142  | 2.148   | 0.0573  |
| K*P            | 1    | 2.411414  | 0.829973  | 2.905   | 0.0157  |
| K*K            | 1    | -0.714584 | 0.404233  | -1.768  | 0.1075  |

<table>
<thead>
<tr>
<th>Factor</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>4</td>
<td>2.740664</td>
<td>0.685166</td>
<td>5.393</td>
<td>0.0141</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>1.799019</td>
<td>0.449755</td>
<td>3.540</td>
<td>0.0477</td>
</tr>
<tr>
<td>K</td>
<td>4</td>
<td>3.807069</td>
<td>0.951767</td>
<td>7.492</td>
<td>0.0047</td>
</tr>
</tbody>
</table>
### Canonical Analysis of Response Surface

<table>
<thead>
<tr>
<th>Factor</th>
<th>Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1.758160</td>
</tr>
<tr>
<td>P</td>
<td>0.656278</td>
</tr>
<tr>
<td>K</td>
<td>1.443790</td>
</tr>
</tbody>
</table>

**Predicted value at stationary point**: 4.834526 kg

<table>
<thead>
<tr>
<th>Eigenvectors</th>
<th>N</th>
<th>P</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.561918</td>
<td>0.021051</td>
<td>0.937448</td>
<td>0.347487</td>
</tr>
<tr>
<td>-0.504592</td>
<td>0.857206</td>
<td>-0.195800</td>
<td>0.476298</td>
</tr>
<tr>
<td>-1.394032</td>
<td>-0.514543</td>
<td>-0.287842</td>
<td>0.807708</td>
</tr>
</tbody>
</table>

Stationary point is a saddle point.

The eigenvalues obtained are $\lambda_1, \lambda_2$ and $\lambda_3$ as 2.561918, -0.504592, -1.394032. As $\lambda_2$ and $\lambda_3$ are negative, therefore, take $W_2 = W_3 = 0$. Let

$$M = \begin{bmatrix} 0.021051 & 0.857206 & -0.514543, \\ 0.937448 & -0.195800 & -0.287842, \\ 0.34787 & 0.476298 & 0.807708; \end{bmatrix}$$

denotes the matrix of eigenvectors. The estimated response at the stationary points be 4.834526 kg/plot. Let the desired response be $Y_{\text{des}}=5.0$ kg/plot. Therefore, let $W_1$, obtained from the equation is $\sqrt{\text{difference}/2.561918}=AX_1$, say. To obtain various different sets of many values of $W_1$, generate a random variable, $u$, which follows uniform distribution and multiply this value with $2u-1$ such that $W_1$ lies within the interval, (-AX1, AX1). Now to get a combination of $x_i$’s that produces the desired response obtain $X = M \times W + x_0$.

```
PROC IML;
W=J(3,1,0);
Ydes=5.0;
W2=0;
W3=0;
Dif=Ydes - 4.834526;
Ax1=Sqrt(dif/2.561918);
u= uniform(0);
W1= ax1*(2*u-1); print w1;
w[1,] = w1;
w[2,] = 0;
w[3,] = 0;
m = {0.021051 0.857206 -0.514543,
     0.937448 -0.195800 -0.287842,
     0.34787 0.476298 0.807708};
xest = {1.758160, 0.656278, 1.443790};
x = m*W + xest;
print x;
run;
```
Combination of N, P, K estimated to produce 5.0 kg/plot of Beans.

<table>
<thead>
<tr>
<th>Y</th>
<th>N</th>
<th>P</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>1.760</td>
<td>0.730</td>
<td>1.471</td>
</tr>
<tr>
<td>1.762</td>
<td>0.815</td>
<td>1.503</td>
<td></td>
</tr>
<tr>
<td>1.754</td>
<td>0.460</td>
<td>1.371</td>
<td></td>
</tr>
</tbody>
</table>

One can select a practically feasible combination of N, P and K.

5. Response Surface Designs for Slope Estimation

The above discussion relates to the response surface designs for response optimization. In many practical situations, however, the experimenter is interested in estimation of the rate of change of response for given value of independent variable(s) rather than optimization of response. This problem is frequently encountered e.g., in estimating rates of reaction in chemical experiments; rates of growth of biological populations; rates of changes in response of a human being or an animal to a drug dosage, rate of change of yield per unit of fertilizer dose. Efforts have been made in the literature for obtaining efficient designs for the estimation of differences in responses i.e., for estimating the slope of a response surface.

Many researchers with different approaches have taken up the problem of designs for estimating the slope of a response surface. We confine ourselves to two main approaches, namely

- Slope Rotatability
- Minimax Designs

The designs possessing the property that the estimate of derivative is equal for all points equidistant from the origin are known as slope rotatable designs. For a second order response surface, the rate of change of response due to \( i^{th} \) independent variable is given by

\[
\frac{\partial \hat{y}(x)}{\partial x_i} = b_i + 2b_{ii}x_i + \sum_{i' \neq i} b_{ii'}x_i'
\]

For second order design satisfying (2.1) we have

\[
\text{Cov}(b_i, b_{ii}) = \text{Cov}(b_{ii}, b_{ii'}) = \text{Cov}(b_{ii}, b_{ii'})
\]

Thus variance of \( \frac{\partial \hat{y}(x)}{\partial x_i} \) is given by

\[
\text{Var}\left( \frac{\partial \hat{y}(x)}{\partial x_i} \right) = \text{Var}(b_i) + \rho^2\text{Var}(b_{ii}) + x_i^2[4\text{Var}(b_{ii})-\text{Var}(b_{ii'})]
\]

Thus in order to obtain slope rotatable design, the design must satisfy

- Conditions of symmetry (2.1)
- \( \frac{\lambda_4}{\lambda_2^2} > \frac{v}{c+v-1} \)
- \( 4\text{Var}(b_{ii}) = \text{Var}(b_{ii'}) \)

It is important to note here that no rotatable design can be slope rotatable.

A minimax design is one that minimizes the variance of the estimated slope maximized over all points in the design.
Some Useful References