SOME OPTIMIZATION PROCEDURES USED IN RESPONSE SURFACE METHODOLOGY

by

Bruce R. Glenn

Thesis submitted to the Graduate Faculty of the
Virginia Polytechnic Institute
in candidacy for the degree of

MASTER OF SCIENCE
in
Statistics

APPROVED: _____
Chairman, Dr. Raymond H. Myers

_____ _____
Dr. Klaus Hinkelmann Dr. Boyd Harshbarger

May, 1967
Blacksburg, Virginia

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I INTRODUCTION

A. GENERAL DISCUSSION OF RESPONSE SURFACE METHODOLOGY

Response Surface Methodology is a relatively new field of applied statistics which is essentially used in finding the "best" operating conditions for a physical or chemical process and to explain or describe certain features of the process system. Early applications of response surface methods dealt with chemical processes and their optimization and, as a result, the goal of the procedures is usually given as that of maximizing a yield or minimizing a cost or some other variable having a physical interpretation in a chemical process.

A response surface study will have at least three phases:

(1) Phase One consists of a definition of the problem. This includes the delineation of all process variables and initial limits on these variables.

(2) Phase Two is a search phase in which certain experimental designs on the process variables are used to obtain operating conditions in the vicinity of the true response optimum.

(3) Phase Three is a more detailed examination of the response surface in the near optimal region of operating conditions. This phase utilizes other experimental designs

and mathematical and graphical techniques to present the resulting response surface to the experimenter for interpretation. A general review article and literature survey for response surface methodology is given in reference (22).

This thesis is a literature search into certain aspects of Phase Two. In addition one interpretative technique useful in Phase Three is also presented.

B. NOTATION AND BASIC RELATIONSHIPS

The response variable will be denoted by y and the assumed controllable process variables will be denoted by $x_1, x_2, \cdots, x_k$. Many of the response surface methods are based on the assumption that a response y can be approximated by a first order, second order, or higher order polynomial in $x_1, x_2, \cdots, x_k$. For a first order polynomial the relationship is of the form

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k. \qquad (1.1)$$

The general second order model is of the form

$$y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \cdots$$
$$+ \beta_{k-1,k} x_{k-1} x_k + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \cdots + \beta_{kk} x_k^2 . \qquad (1.2)$$

The coefficients in these models are estimated by least squares[4], usually with data from some designed experiment. The fitted surfaces are of the form, for the first order model:

$$\hat{y} = b_o + \underline{b}' \underline{x}, \tag{1.3}$$

and for the second order model:

$$\hat{y} = b_o + \underline{b}' \underline{x} + \underline{x}' B \underline{x}, \tag{1.4}$$

where $\underline{b}' = (b_1, b_2, \cdots, b_k)$, $\underline{x}' = (x_1, x_2, \cdots, x_k)$ and

$$B = \begin{pmatrix} b_{11} & \frac{1}{2}b_{12} & \frac{1}{2}b_{13} & \cdots & \frac{1}{2}b_{1k} \\ \frac{1}{2}b_{12} & b_{22} & \frac{1}{2}b_{23} & \cdots & \frac{1}{2}b_{2k} \\ \cdots & & & \cdots & \\ \frac{1}{2}b_{1k} & \frac{1}{2}b_{2k} & \frac{1}{2}b_{3k} & \cdots & b_{kk} \end{pmatrix} .$$

The x values, in general, are coded values centered around the origin $(0, 0, \cdots, 0)$.

For the second order model a recommended form of analysis, which is often performed to better interpret the fitted response surface, reduces (1.4) to the "canonical form".[16] This is done by translating the $\underline{x}$ variables to the stationary point (center), $\underline{x}_o$, of the response system, where $\underline{x}_o$ is given by the solution of the set of first order partial derivatives of (1.4). It is easily shown[4] that

$$\underline{x}_o = \tfrac{1}{2} B^{-1} \underline{b}. \tag{1.5}$$

The response system is then rotated to eliminate cross-product terms. The resulting equation is of the form

$$\hat{y}-\hat{y}_o = \mu_1 z_1^2+\mu_1 z_1^2+\cdots+\mu_k z_k^2 \qquad (1.6)$$

where $\hat{y}_o$ is the estimated response at $\underline{x}_o$, $\underline{Z}' = (Z_1, Z_2, \ldots, Z_k)$ are the transformed axes, and $\underline{\mu} = (\mu_1, \mu_2, \ldots, \mu_k)$ are the eigenvalues of the matrix B. The transformation on the $\underline{x}$ variables[4] is given by

$$Z = T'(\underline{x}-\underline{x}_o) \qquad (1.7)$$

where T is a (kxk) orthogonal matrix whose columns are the orthonormal vectors associated with $(\mu_1, \mu_2, \ldots, \mu_k)$.

As an example, consider Figure 1-1. For this surface $\mu_1 < \mu_2 < 0$, hence the response system has a maximum at $\underline{x}_o$. In general, if all of the eigenvalues are negative, the response surface has a maximum. If they are all positive, the response surface has a minimum. When a mixture of positive and negative eigenvalues are obtained, the response system contains ridges or a k-variate saddle point. In the area of experimental interest ridges may be encountered when $\underline{x}_o$ is far outside of this area, even though the eigenvalues are all the same sign. Thus the location of $\underline{x}_o$ and the sign of the eigenvalues all determine the nature of the response system.

When $\underline{x}_o$ is far outside of the area of experimental interest or when the eigenvalues have differing signs, the resulting surface for $k \geq 3$ may become very complicated and

be difficult to understand. These types of surfaces can be subjected to another analysis to better understand the resulting fitted surface. This analysis is called "Ridge Analysis" and is discussed in Chapter II.

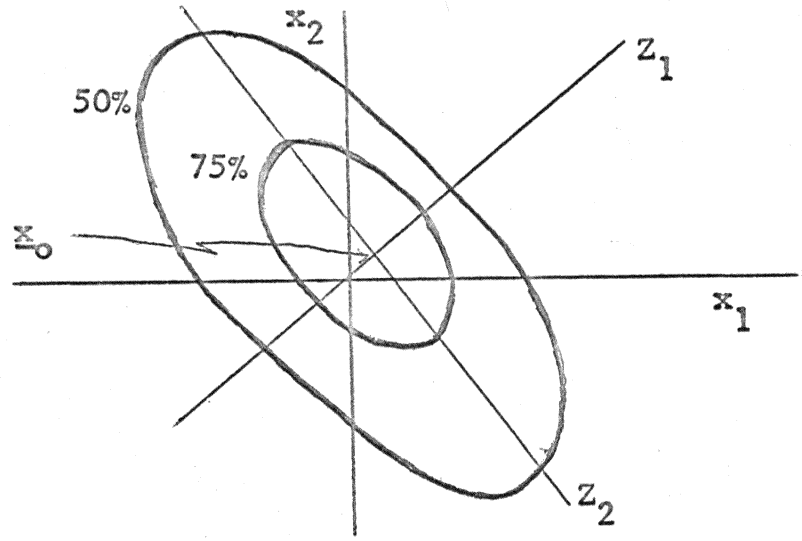


FIGURE 1-1. Contours of a fitted response surface showing the original axes (x_1, x_2) and the canonical form axes (z_1, z_2).

C. SCOPE OF THESIS

The next chapter discusses the "Ridge Analysis" technique for interpreting complicated response surfaces.

Chapters III and IV discuss several of the optimization techniques one can use in Phase Two of the response surface problem. Evolutionary Operation techniques for existing production processes are discussed in Chapter III. Chapter IV discusses optimization techniques useful in the laboratory

or pilot production stages of optimization and covers such topics as the method of steepest ascent, a sequential one factor at a time procedure, "Rotating Square" and "Random" evolutionary operations, some one-factor optimizing techniques, and also includes the results of a study made to compare some of these techniques.

CHAPTER II RIDGE ANALYSIS

A. INTRODUCTION

When a canonical analysis is performed on the estimated second order response surface of (1.4) and the resulting canonical form

$$\hat{y}-\hat{y}_o = \mu_1 z_1^2+\mu_2 z_2^2+\cdots+\mu_k z_k^2 \qquad (1.6)$$

has both positive and negative values for the μ's, the response surface contains ridges or saddle points. Because of the complicated nature of such a system, it is difficult to interpret graphically. "Ridge Analysis" is a method of analysis which provides a graphical interpretation of such a response system. It is a general technique which plots the response on the ridges versus the distance R from the center of the design used to estimate the coefficients of (1.4). In this way one can find the maximum yield for any selected distance R as well as any other intermediate stationary points for that R. This can be done regardless of k; hence, this method provides a means of graphically interpreting the response surface on k variables in just two dimensions.

"Ridge Analysis" was coined by Hoerl in a 1959 paper[22]. Hoerl's paper was expository in contents and did not attempt to mathematically derive the properties of this

method. Draper (1963)[16] provided simple derivations of this technique and gave proofs of the mathematical properties of this method. The following mathematical development and most of the theory discussion is taken from Draper.

For two variables, the complete ridge analysis is not too involved computationally. However for $k \geq 3$, the repeated solution of the set of simultaneous equations (2.2) is best performed on a computer. For this reason this method of analysis is not recommended unless such a computer and associated computer programs are available.

B. DERIVATION OF THE METHOD OF RIDGE ANALYSIS

Ridge Analysis is a method used to investigate the stationary points of $\hat{y}$ as given in (1.6) on a sphere of radius R, where $R^2 = x_1^2+x_2^2+\cdots+x_k^2$ or equivalently

$$\underline{x}'\underline{x}-R^2 = 0. \qquad (2.1)$$

For discussion purposes, assume that one is attempting to maximize y. Then the problem is to maximize y subject to the restriction of (2.1). For a given R, one can find the maximum y and then plot R on the abscissa against the k+1 ordinates $(x_1,x_2,\cdots,x_k)$ and $\hat{y}$. This can be done for varying values of R. In this way one can plot k+1 curves to indicate how the maximum $\hat{y}$ varies with R and also have the coordinates of the maximum response.

This same method can be used to find other stationary ridges such as intermediate maxima or minima or the absolute minimum ridge. This may be helpful if another intermediate maximum ridge provided a response similar to the absolute maximum ridge but this intermediate ridge had other desirable properties such as lower pressures or lower temperatures,etc. To plot these points, we consider the function

$$F = \hat{y}-\lambda(R^2-\underline{x}'\underline{x})$$

$$= b_o+\underline{x}'\underline{b}+\underline{x}'B\underline{x}-\lambda(R^2-\underline{x}'\underline{x})$$

where λ is a Lagrangian multiplier. Taking partials of F with respect to $x_1,x_2,\cdots,x_k$, dividing by 2, equating to zero, and rearranging of the resulting expressions, one obtains the set of equations:

$$\begin{aligned}
&(b_{11}-\lambda)x_1 + \tfrac{1}{2}b_{12}x_2 + \cdots + \tfrac{1}{2}b_{1k}x_k = -\tfrac{1}{2}b_1 \\
&\tfrac{1}{2}b_{12}x_1 + (b_{22}-\lambda)x_2 + \cdots + \tfrac{1}{2}b_{2k}x_k = -\tfrac{1}{2}b_2 \\
&\qquad\cdots \\
&\tfrac{1}{2}b_{1k}x_1 + \tfrac{1}{2}b_{2k}x_2 + \cdots + (b_{kk}-\lambda)x_k = -\tfrac{1}{2}b_k
\end{aligned} \tag{2.2}$$

or in matrix form

$$(B-\lambda I)\underline{x} = -\tfrac{1}{2}\underline{b} \tag{2.3}$$

Equations (2.1) and (2.3) can be solved simultaneously for $\underline{x}' = (x_1,x_2,\cdots,x_k)$ and λ. However, an equivalent technique which is easier computationally is the following:

(1) Regard R as variable, but fix λ instead.

(2) Insert the fixed value of λ in equations (2.2) and solve them for $x_1, x_2, \cdots, x_k$.

(3) Compute $R = (x_1^2 + \cdots + x_k^2)^{\frac{1}{2}} = (\underline{x}'\underline{x})^{\frac{1}{2}}$.

(4) Evaluate $\hat{y} = b_o + \lambda R^2 + \frac{1}{2}\underline{x}'\underline{b}$.

The above expression for $\hat{y}$ will be derived in the next section.

With this technique we have found a stationary value of $\hat{y}$, the coordinates of this stationary value, and the distance R from the center of the design. A plot of R against ordinates $x_1, x_2, \cdots, x_k$, and $\hat{y}$ can now be made. It may turn out that several different values of λ may lead to the same R with different stationary values of $\hat{y}$. The next section will derive some properties of this method of analysis which will enable one to choose values of λ to obtain the absolute maximum ridge, intermediate ridges, or the absolute minimum ridge. Once this is known, different λ's may be used to obtain points on the k+1 curves for each ridge of interest.

C. PROPERTIES OF THE STATIONARY VALUES

It was noted in the introduction chapter that the eigenvalues of B are denoted by μ_i for $i = 1,2,\cdots,k$. Now there exist eigenvectors $\underline{z}_i$ such that $B\underline{z}_i = \mu_i\underline{z}_i$ or $(B-\mu_i I)\underline{z}_i = 0$. It will be shown that the type of stationary point found by applying steps (1) - (4) in the previous

section depends on the values of λ and the μ_i. By comparing λ with the μ_i one can determine if $\hat{y}$ is an absolute maximum, local maximum or minimum, or an absolute minimum.

Before continuing, a result from calculus is needed[36] A method of obtaining the stationary values of a function $f(x_1,x_2,\cdots,x_k)$ subject to restrictions of the form

$$g_j(x_1,x_2,\cdots,x_k) = 0 \qquad j = 1,2,\cdots,m, \qquad (2.4)$$

is to define

$$F = f(x_1,x_2,\cdots,x_k) - \sum_{j=1}^{m} \lambda_j g_j(x_1,x_2,\cdots,x_k). \qquad (2.5)$$

To obtain stationary values of F, one partially differentiates (2.5) with respect to each x_i and equates the resulting expressions to zero. This gives the k equations

$$\frac{\partial F}{\partial x_i} = \frac{\partial f(\underline{x})}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(\underline{x})}{\partial x_i} = 0 \quad i = 1,2,\cdots,k. \qquad (2.6)$$

These k equations and the m equations of (2.4) can be solved simultaneously for $(x_1,x_2,\cdots,x_k)$, and $(\lambda_1,\lambda_2,\cdots,\lambda_m)$. The usual method of solution is to express $(\lambda_1,\cdots,\lambda_m)$ in terms of $(x_1,x_2,\cdots,x_k)$. Alternatively, specific values of λ_j could be used. In either case $F_1(\underline{x})$ given by $F_1(\underline{x}) = F(\underline{x},\lambda)$ is some function of $(x_1,x_2,\cdots,x_k)$ or of $(x_1,x_2,\ldots,x_k)$ and some known constants. Define the matrix

$$M(\underline{x}) = \begin{bmatrix} \frac{\partial^2 F_1}{\partial x_1^2} & \frac{\partial^2 F_1}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F_1}{\partial x_1 \partial x_k} \\ \frac{\partial^2 F_1}{\partial x_1 \partial x_2} & \frac{\partial^2 F_1}{\partial x_2^2} & \cdots & \frac{\partial^2 F_1}{\partial x_2 \partial x_k} \\ \cdots & & \cdots & \\ \frac{\partial^2 F_1}{\partial x_1 \partial x_k} & \frac{\partial^2 F_1}{\partial x_2 \partial x_k} & \cdots & \frac{\partial^2 F_1}{\partial x_k^2} \end{bmatrix} . \qquad (2.7)$$

If $\underline{x}' = (x_1, x_2, \cdots, x_k) = (a_1, a_2, \cdots, a_k) = \underline{a}'$ is a solution of (2.6) and (2.4), let $M(\underline{a})$ be the matrix of second order partial derivatives with the a_i's substituted for the corresponding x_i's in (2.7). Then for $\underline{y}' = (y_1, y_2, \cdots, y_k)$ any (1xk) real vector, if $\underline{y}' M(\underline{a})\underline{y}$ is positive definite, i.e. $\underline{y}'M(\underline{a})\underline{y} > 0$ for all $\underline{y}$, F achieves a local minimum at $\underline{x} = \underline{a}$. If $\underline{y}'M(\underline{a})\underline{y}$ is negative definite, i.e. $\underline{y}'M(a)\underline{y} < 0$ for all $\underline{y}$, F has a local maximum at $\underline{x} = \underline{a}$.

It turns out that the converse of this result is not true. One can have a local minimum without $M(\underline{a})$ being positive definite.

For our case we have

$$f(x_1, x_2, \cdots, x_k) = \hat{y} = b_o + \underline{x}'\underline{b} + \underline{x}'B\underline{x} \qquad (2.8)$$

and

$$g_1(x_1, x_2, \cdots, x_k) = \underline{x}'\underline{x} - R^2 = 0. \qquad (2.1)$$

Hence

$$F = b_o + \underline{x}'\underline{b} + \underline{x}'B\underline{x} - \lambda(\underline{x}'\underline{x} - R^2).$$

As shown before, the set of first derivatives reduced to

$$(B-\lambda I)\underline{x} = -\tfrac{1}{2}\underline{b}. \tag{2.3}$$

For a fixed λ, $M(x_1, x_2, \cdots, x_k)$ is given by

$$M(\underline{x}) = 2(B-\lambda I) \tag{2.9}$$

Now consider two solutions of (2.3),

$\underline{x}_1' = (a_1, a_2, \cdots, a_k)$ for $\lambda = \lambda_1$ and

$\underline{x}_2' = (c_1, c_2, \cdots, c_k)$ for $\lambda = \lambda_2$ with two resulting stationary values of $\hat{y}_1$ and $\hat{y}_2$ on spheres of radius $\underline{x}_1'\underline{x}_1 = R_1^2$ and $\underline{x}_2'\underline{x}_2 = R_2^2$, respectively.

Theorem 2.1 If $R_1 = R_2$ and $\lambda_1 > \lambda_2$, then $\hat{y}_1 > \hat{y}_2$.

Proof: We know that the following are true:

$$(B-\lambda_1 I)\underline{x}_1 = -\tfrac{1}{2}\underline{b}, \tag{2.1.1}$$

$$(B-\lambda_2 I)\underline{x}_2 = -\tfrac{1}{2}\underline{b}, \tag{2.1.2}$$

$$\underline{x}_1'\underline{x}_1 = \underline{x}_2'\underline{x}_2 = R^2, \tag{2.1.3}$$

$$\hat{y}_1 = \underline{x}_1'B\underline{x}_1 + \underline{x}_1'\underline{b} + b_o \tag{2.1.4}$$

and

$$\hat{y}_2 = \underline{x}_2'B\underline{x}_2 + \underline{x}_2'\underline{b} + b_o. \tag{2.1.5}$$

Premultiplying (2.1.1) by $\underline{x}_1'$ and (2.1.2) by $\underline{x}_2'$ gives

$$\underline{x}_1'(B-\lambda_1 I)\underline{x}_1 = -\tfrac{1}{2}\underline{x}_1'\underline{b}$$

and

$$\underline{x}_2'(B-\lambda_2 I)\underline{x}_2 = -\tfrac{1}{2}\underline{x}_2'\underline{b} .$$

Subtracting these two equations gives

$$\underline{x}_1'B\underline{x}_1-\lambda_1\underline{x}_1'\underline{x}_1-\underline{x}_2'B\underline{x}_2+\lambda_2\underline{x}_2'\underline{x}_2 = -\tfrac{1}{2}(\underline{x}_1'-\underline{x}_2')\underline{b}$$

or

$$\underline{x}_1'B\underline{x}_1-\underline{x}_2'B\underline{x}_2+\tfrac{1}{2}(\underline{x}_1-\underline{x}_2)'\underline{b} = (\lambda_1-\lambda_2)R^2 \qquad (2.1.6)$$

Now subtracting (2.1.5) from (2.1.4) gives

$$\hat{y}_1-\hat{y}_2 = \underline{x}_1'B\underline{x}_1-\underline{x}_2'B\underline{x}_2+(\underline{x}_1-\underline{x}_2)'\underline{b} ,$$

so

$$\hat{y}_1-\hat{y}_2 = \tfrac{1}{2}(\underline{x}_1-\underline{x}_2)'\underline{b}+(\lambda_1-\lambda_2)R^2 . \qquad (2.1.7)$$

Premultiplying (2.1.1) by $\underline{x}_2'$, (2.1.2) by $\underline{x}_1'$ and subtracting gives

$$\underline{x}_2'(B-\lambda_1 I)\underline{x}_1-\underline{x}_1'(B-\lambda_2 I)\underline{x}_2 = -\tfrac{1}{2}(\underline{x}_2-\underline{x}_1)'\underline{b} ,$$

or

$$\underline{x}_2'B\underline{x}_1-\underline{x}_1'B\underline{x}_2-\lambda_1\underline{x}_2'\underline{x}_1+\lambda_2\underline{x}_1'\underline{x}_2 = \tfrac{1}{2}(\underline{x}_1-\underline{x}_2)'\underline{b}$$

which, since $\underline{x}_2'B\underline{x}_1 = \underline{x}_1'B\underline{x}_2$ and $\underline{x}_1'\underline{x}_2 = \underline{x}_2'\underline{x}_1$, gives

$$(\lambda_2-\lambda_1)\underline{x}_1'\underline{x}_2 = \tfrac{1}{2}(\underline{x}_1-\underline{x}_2)'\underline{b} . \qquad (2.1.8)$$

Hence, from (2.1.7) and (2.1.8) we get

$$\hat{y}_1-\hat{y}_2 = (\lambda_1-\lambda_2)(R^2-\underline{x}_1'\underline{x}_2) . \qquad (2.1.9)$$

Now consider $(R^2-\underline{x}_1'\underline{x}_2)$.

$$R^2-\underline{x}_1'\underline{x}_2 = (a_1^2+a_2^2+\cdots+a_k^2)^{\frac{1}{2}}(c_1^2+c_2^2+\cdots+c_k^2)^{\frac{1}{2}}-(a_1c_1+a_2c_2 +\cdots+a_kc_k)\geq 0$$

due to one form of the Cauchy-Schwarz inequality.

Hence

$$\hat{y}_1-\hat{y}_2 \geq \lambda_1-\lambda_2 ,$$

and since $\lambda_1 > \lambda_2$, $\hat{y}_1-\hat{y}_2 > 0$ or $\hat{y}_1 > \hat{y}_2$.

Theorem 2.2 If $R_1 = R_2$, $M(\underline{x}_1)$ is positive definite, and $M(\underline{x}_2)$ is indefinite, then $\hat{y}_1 < \hat{y}_2$.

Proof: From (2.9), $M(\underline{x}_i) = 2(B-\lambda_i I)$, $i = 1,2$ for the stationary point $\underline{x}_i'$. Let $\underline{y}$ be any kx1 real vector. Since $M(\underline{x}_2)$ is indefinite,

$$\underline{y}'(B-\lambda_2 I)\underline{y} \leq 0 \quad \text{for at least one } \underline{y}=\underline{q}\neq\underline{0}$$

so

$$\underline{q}'B\underline{q}-\lambda_2\underline{q}'\underline{q} \leq 0 .$$

Since $M(\underline{x}_1)$ is positive definite,

$$\underline{y}'(B-\lambda_1 I)\underline{y} > 0 \quad \text{for all } \underline{y} \text{ including } \underline{y}=\underline{q},$$

so

$$\underline{q}'B\underline{q}-\lambda_1\underline{q}'\underline{q} > 0 .$$

Hence

$$\lambda_2\underline{q}'\underline{q} \geq \underline{q}'B\underline{q} > \lambda_1\underline{q}'\underline{q}$$

so $\lambda_2 > \lambda_1$ since $\underline{q}'\underline{q} > 0$. Thus by theorem 2.1, $\hat{y}_1 < \hat{y}_2$. Similarly, if $R_1 = R_2$, $M(\underline{x}_1)$ is negative definite, and $M(\underline{x}_2)$ is indefinite, then $\hat{y}_1 > \hat{y}_2$.

Theorem 2.3 If $\lambda_1 > \mu_i$ for all i, then $\underline{x}_1'$ is a point at which $\hat{y}$ attains a local maximum on the sphere of radius R; if $\lambda_1 < \mu_i$ for all i, then $\underline{x}_1'$ is a point at which $\hat{y}$ attains a local minimum on the sphere of radius R.

Proof: Let $\underline{y}$ be any (kx1) real vector.

Define

$$\underline{z} = T\underline{y} \quad \text{or} \quad \underline{y} = T'\underline{z}$$

where T is a (kxk) orthogonal matrix whose columns are the orthonormal eigenvectors of B. Consider

$$\begin{aligned} \tfrac{1}{2}\underline{y}'M(\underline{x}_1)\underline{y} &= \underline{y}'(B-\lambda_1 I)\underline{y} \\ &= \underline{z}'T(B-\lambda_1 I)T'\underline{z} \\ &= \underline{z}'TBT'\underline{z}-\underline{z}'T\lambda_1 T'\underline{z} \\ &= \underline{z}'TBT'\underline{z}-\lambda_1\underline{z}'I\underline{z} \ . \end{aligned}$$

Now $\underline{z}'TBT'\underline{z} = \underline{z}'\Lambda\underline{z}$

where Λ is a diagonal matrix given by

$$\Lambda = \begin{pmatrix} \mu_1 & 0 & 0 & \cdots & 0 \\ 0 & \mu_2 & 0 & \cdots & 0 \\ \cdots & & & \cdots & \\ 0 & 0 & 0 & \cdots & \mu_k \end{pmatrix} .$$

Hence

$$\tfrac{1}{2}\underline{y}'M(\underline{x}_1)\underline{y} = \underline{z}'\begin{pmatrix} \mu_1-\lambda_1 & 0 & \cdots & 0 \\ 0 & \mu_2-\lambda_1 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & \mu_k-\lambda_1 \end{pmatrix}\underline{z}$$

For λ_1 greater than all the μ_i, this expression is negative definite and hence a local maximum is achieved at $\underline{x}_1$. If λ_1 is less than all μ_i, this expression is positive definite and a local minimum is achieved at $\underline{x}_1$. Hence the theorem is proved.

Suppose, as R increases, we trace a locus of stationary points (the absolute maximum or minimum or a local maximum or minimum) and examine the changing values of $\hat{y}$. We then have the following result.

Theorem 2.4 As R increases, $\hat{y}$ changes in one of the following ways (when the response surface is quadratic):

(a) decreases monotonically,

(b) increases monotonically,

(c) passes through a maximum and then decreases monotonically, or

(d) passes through a minimum and then increases monotonically.

If (c) or (d) happens, it is because the locus has passed through the center of the quadratic system.

Proof: We know for every $\underline{x}$ that satisfies

$$(B-\lambda I)\underline{x} = -\tfrac{1}{2}\underline{b} \ , \qquad (2.3)$$

that

$$\begin{aligned}\hat{y} &= \underline{x}'B\underline{x}+\underline{x}'\underline{b}+b_o \\ &= \underline{x}'(-\tfrac{1}{2}\underline{b}+\lambda\underline{x})+\underline{x}'\underline{b}+b_o \\ &= \lambda\underline{x}'\underline{x}+\tfrac{1}{2}\underline{x}'b+b_o \ ,\end{aligned}$$

so

$$\hat{y} = \lambda R^2+\tfrac{1}{2}\underline{x}'\underline{b}+b_o \ . \qquad (2.4.1)$$

Hence $\partial\hat{y}/\partial R = 2\lambda R$, which is zero when $R = 0$ or when $\lambda = 0$. When $R = 0$, we are at the origin (of the design, $\underline{x}' = (0,0,\ldots,0)$) and $\hat{y} = b_o$ is the starting value for the locus of absolute maximum and absolute minimum of $\hat{y}$. When $R = 0$, $\hat{y}$ is stationary with respect to R only when $\lambda = 0$. When $\lambda = 0$, equation (2.3) reduces to

$$B\underline{x} = -\tfrac{1}{2}\underline{b} \ ,$$

so

$$\underline{x}_o = -\tfrac{1}{2}B^{-1}\underline{b}$$

where $\underline{x}_o$ is the center of the quadratic response system for which $\partial\hat{y}/\partial x_i = 0$, $i=1,2,\ldots,k$. Since $\partial\hat{y}/\partial R$ has the same sign as λ, y is monotonic when λ lies between eigenvalues of the same sign. When eigenvalues μ_j and μ_{j+1} have differing signs, $\lambda = 0$ somewhere in the interval and (c) and (d) conditions arise.

Theorem 2.5 $\partial^2 R/\partial\lambda^2 > 0$ for all $R \neq 0$. $\partial^2 R/\partial\lambda^2$ is zero when $R = 0$.

Proof: We know that solutions $\underline{x}$ satisfy

$$(B-\lambda I)\underline{x} = -\tfrac{1}{2}\underline{b} \tag{2.5.1}$$

and

$$\underline{x}'\underline{x} = R^2 \tag{2.5.2}$$

Differentiating these equations with respect to λ gives

$$(B-\lambda I)\frac{\partial \underline{x}}{\partial \lambda} - \underline{x} = 0$$

or

$$(B-\lambda I)\frac{\partial \underline{x}}{\partial \lambda} = \underline{x} \tag{2.5.3}$$

and

$$2\underline{x}'\frac{\partial \underline{x}}{\partial \lambda} = 2R\frac{\partial R}{\partial \lambda}$$

or

$$\underline{x}'\frac{\partial \underline{x}}{\partial \lambda} = R\frac{\partial R}{\partial \lambda} \tag{2.5.4}$$

Differentiating again with respect to λ gives, from (2.5.3)

$$(B-\lambda I)\frac{\partial^2 \underline{x}}{\partial \lambda^2} - \frac{\partial \underline{x}}{\partial \lambda} = \frac{\partial \underline{x}}{\partial \lambda}$$

or

$$(B-\lambda I)\frac{\partial^2 \underline{x}}{\partial \lambda^2} = 2\frac{\partial \underline{x}}{\partial \lambda} \tag{2.5.5}$$

and from (2.5.4)

$$\underline{x}'\left(\frac{\partial^2 \underline{x}}{\partial \lambda^2}\right)+\frac{\partial \underline{x}'}{\partial \lambda}\frac{\partial \underline{x}}{\partial \lambda} = R\left(\frac{\partial^2 R}{\partial \lambda^2}\right)+\left(\frac{\partial R}{\partial \lambda}\right)^2 . \qquad (2.5.6)$$

If one premultiplies (2.5.3) by $\partial^2 \underline{x}'/\partial \lambda^2$, one obtains

$$\frac{\partial^2 \underline{x}'}{\partial \lambda^2}(B-\lambda I)\frac{\partial \underline{x}}{\partial \lambda} = \frac{\partial^2 \underline{x}'}{\partial \lambda^2}\underline{x} . \qquad (2.5.7)$$

Premultiplying (2.5.5) by $\partial \underline{x}'/\partial \lambda$ gives

$$\frac{\partial \underline{x}'}{\partial \lambda}(B-\lambda I)\frac{\partial^2 \underline{x}}{\partial \lambda^2} = 2\frac{\partial \underline{x}'}{\partial \lambda}\frac{\partial \underline{x}}{\partial \lambda} \qquad (2.5.8)$$

Transposing (2.5.7) and subtracting it from (2.5.8) gives

$$\underline{x}'\frac{\partial^2 \underline{x}}{\partial \lambda^2} - 2\frac{\partial \underline{x}'}{\partial \lambda}\frac{\partial \underline{x}}{\partial \lambda} = 0$$

or

$$\underline{x}'\frac{\partial^2 \underline{x}}{\partial \lambda^2} = 2\frac{\partial \underline{x}'}{\partial \lambda}\frac{\partial \underline{x}}{\partial \lambda} .$$

Substituting the right hand side for the left hand side of the above equation in (2.5.6) we obtain

$$2\frac{\partial \underline{x}'}{\partial \lambda}\frac{\partial \underline{x}}{\partial \lambda} + \frac{\partial \underline{x}'}{\partial \lambda}\frac{\partial \underline{x}}{\partial \lambda} = R\left(\frac{\partial^2 R}{\partial \lambda^2}\right)+\left(\frac{\partial R}{\partial \lambda}\right)^2$$

or

$$R\left(\frac{\partial^2 R}{\partial \lambda^2}\right) = 3\frac{\partial \underline{x}'}{\partial \lambda}\frac{\partial \underline{x}}{\partial \lambda} - \left(\frac{\partial R}{\partial \lambda}\right)^2 . \qquad (2.5.9)$$

Now from (2.5.4)

$$\frac{\partial R}{\partial \lambda} = \frac{\underline{x}'}{R} \frac{\partial \underline{x}}{\partial \lambda} .$$

Hence (2.5.9) becomes

$$R\frac{\partial^2 R}{\partial \lambda^2} = 3\frac{\partial \underline{x}'}{\partial \lambda} \frac{\partial \underline{x}}{\partial \lambda} - (\underline{x}'\frac{\partial \underline{x}}{\partial \lambda})^2 \frac{1}{R^2} .$$

Multiplying both sides by R^2 we obtain

$$R^3 \frac{\partial^2 R}{\partial \lambda^2} = 3R^2 \frac{\partial \underline{x}'}{\partial \lambda} \frac{\partial \underline{x}}{\partial \lambda} - (\underline{x}'\frac{\partial \underline{x}}{\partial \lambda})^2$$

so

$$R^3 \frac{\partial^2 R}{\partial \lambda^2} = 2R^2 \frac{\partial \underline{x}'}{\partial \lambda} \frac{\partial \underline{x}}{\partial \lambda} + \left\{ R^2 \frac{\partial \underline{x}'}{\partial \lambda} \frac{\partial \underline{x}}{\partial \lambda} - (\underline{x}'\frac{\partial \underline{x}}{\partial \lambda})^2 \right\} \qquad (2.5.10)$$

The first term on the right hand side of (2.5.10) is always positive and is zero only when $R = 0$ or $\partial \underline{x}/\partial \lambda = 0$. The second term is always positive, due to a result of Hardy, Littlewood and Polya, 1952,[19] and is zero only when $\underline{x} = \underline{0}$, (hence $R = 0$) or when $\partial \underline{x}/\partial \lambda = 0$.

When $\partial \underline{x}/\partial \lambda = 0$, $\underline{x} = \underline{0}$ from equation (2.5.3), and again $R = 0$. Hence $\partial^2 R/\partial \lambda^2 > 0$ except when $R = 0$. When $R = 0$, $\partial^2 R/\partial \lambda^2 = 0$.

Now that this result has been established, we can further consider the relationship between λ and R by plotting the graph of λ versus R. By such a plotting we shall be able

to determine the requirements on λ to obtain points on the absolute maximum and/or minimum ridges.

D. GRAPH OF R VERSUS λ

As before, order the μ_i such that $\mu_1 < \mu_2 < \ldots < \mu_k$. Since $(B-\lambda I)\underline{x} = -\frac{1}{2}\underline{b}$, as $\lambda \rightarrow \mu_i$, $\underline{x} \rightarrow \pm \infty$ so $R \rightarrow +\infty$. As $\lambda \rightarrow \pm \infty$, $\underline{x} \rightarrow \underline{0}$ so $R \rightarrow 0$. Since $\partial^2 R/\partial \lambda^2 > 0$, the graph of R on the ordinate plotted against λ has the following shape:

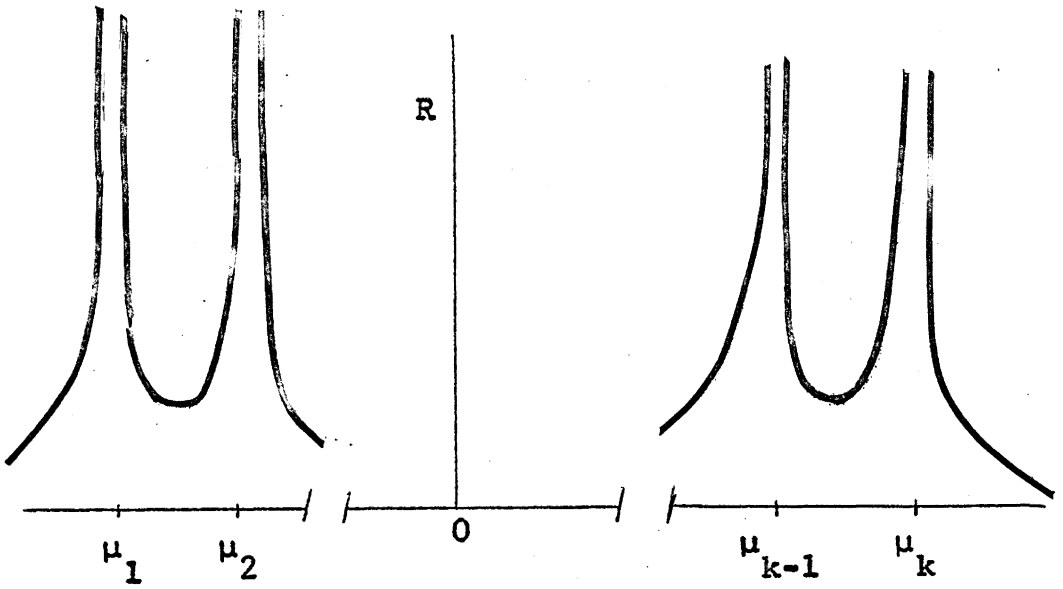


FIGURE 2-1. Graph of R against λ where μ_i's are ranked eigenvalues of B.

Considering this graph, there will be at most 2k values of λ for which stationary values of $\hat{y}$ will exist. (Each eigenvalue provides two branches). For values of $\lambda > \mu_k$ there will only be one stationary value of $\hat{y}$ as R varies, since as $\lambda \rightarrow \infty$, $\underline{x} \rightarrow \underline{0}$ so $R \rightarrow 0$. There are no eigenvalues

larger than μ_k to turn the graph up to R = + ∞. Hence the absolute maximum ridge can be plotted by using values of $\lambda > \mu_k$. Similarly the absolute minimum ridge can be found by using values of $\lambda < \mu_1$.

When values of λ are chosen between eigenvalues a stationary value will occur which is on a local minimum or maximum ridge depending on whether it is on the right or left of the u-shaped curve.

E. SUMMARY OF RIDGE ANALYSIS

To perform a ridge analysis, one follows these steps:

(1) Estimate the coefficients of $\hat{y} = b_o + \underline{x}'\underline{b} + \underline{x}'B\underline{x}$

(2) Compute the eigenvalues of B.

(3) Plot the absolute maximum ridge by substituting values of λ greater than the largest eigenvalue in $(B-\lambda I)\underline{x} = -\frac{1}{2}\underline{b}$ and solve for $\underline{x}$. Values of λ less than the smallest eigenvalue will provide the absolute minimum ridge. Intermediate ridges will be found by using values between eigenvalues.

(4) For each solution $\underline{x}$, compute $R = (\underline{x}'\underline{x})^{\frac{1}{2}}$

(5) For each solution, calculate $\hat{y} = b_o + \lambda R^2 + \frac{1}{2}\underline{x}'\underline{b}$

(6) Perform steps (3) - (5) until a sufficient number of points on the graphs of R against $x_1, x_2, \ldots, x_k$, and $\hat{y}$ have been found to adequately graph the relationship.

F. NUMERICAL EXAMPLE

Frankel[17] was investigating the yield of mercaptobenzothiazole(MBT) as a function of time and temperature. A rotatable octagonal design with three center points and an additional replication of one of the design points was run. A second order response surface was estimated to be

$$\hat{y} = 82.17-1.01x_1-8.61x_2+1.40x_1^2-8.76x_2^2-7.20x_1x_2 \quad (2.10)$$

where

$$x_1 = \frac{\text{Time(hr)}-12}{8} \quad \text{and} \quad x_2 = \frac{\text{Temp}(^{\circ}\text{C})-250}{30} .$$

The contours of this response surface is given in Figure 2-2. As can be seen, the resulting surface is a saddle. For this example, equation (2.3) becomes

$$\begin{pmatrix}(1.40-\lambda) & -3.60\\ -3.60 & (-8.76-\lambda)\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} = \begin{pmatrix}.505\\ 4.305\end{pmatrix}. \quad (2.11)$$

Solving for x_1 and x_2 one obtains

$$x_1 = \frac{11.0742-0.505\lambda}{\lambda^2+7.36\lambda-25.224}$$

and (2.12)

$$x_2 = \frac{7.845-4.305\lambda}{\lambda^2+7.36\lambda-25.224} .$$

The eigenvalues of

$$B = \begin{pmatrix} 1.40 & -3.60 \\ -3.60 & -8.76 \end{pmatrix}$$

are

$$\mu_1 = -9.90625$$
$$\mu_2 = 2.54625 \qquad (2.13)$$

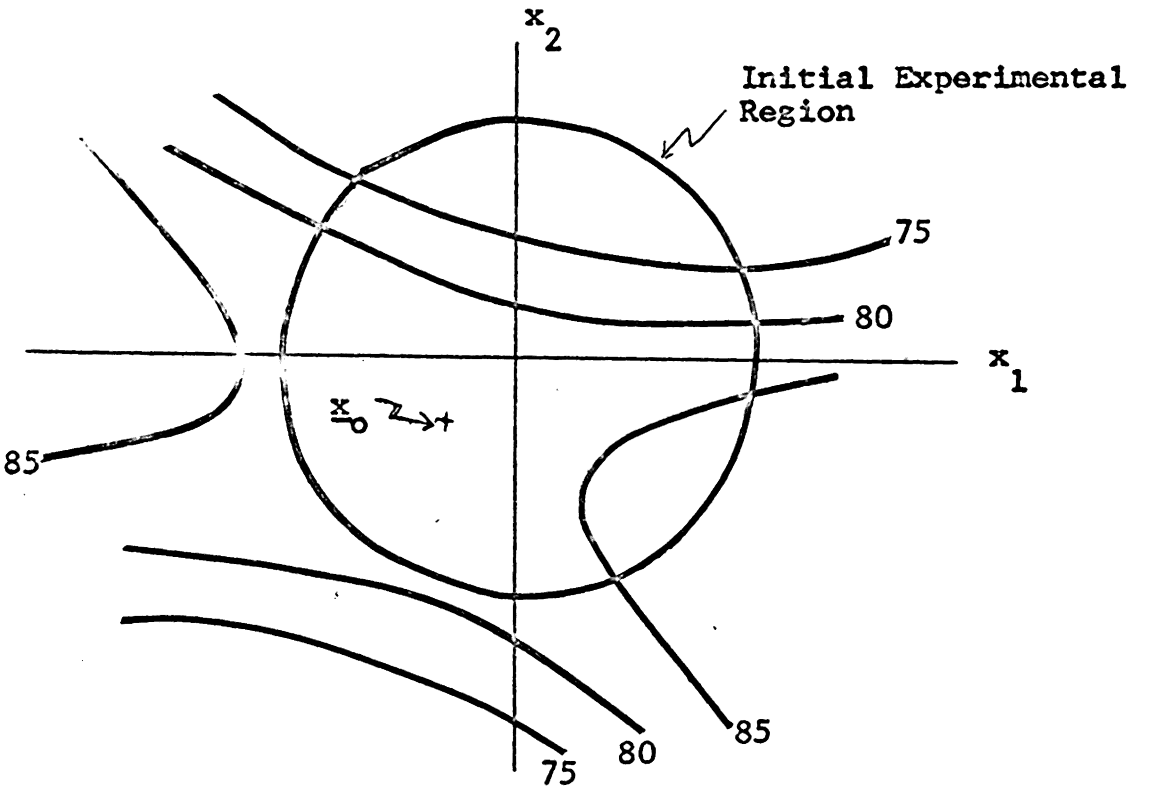


FIGURE 2-2. Yield of MBT as a function of time (x_1) and temperature (x_2).

Hence to find the locus of the absolute maximum ridge, values of $\lambda > 2.54625$ are substituted in (2.12) and solved for x_1 and x_2. R^2 is then computed by

$$R^2 = x_1^2 + x_2^2,$$

and $\hat{y}$ is found by use of equation (2.4.1)

$$\hat{y} = \lambda R^2 + \tfrac{1}{2}\underline{x}'\underline{b} + b_o$$

$$= \lambda R^2 - 0.505x_1 - 4.305x_2 + 82.17.$$

For example, let $\lambda = 4.0$, then

$$x_1 = \frac{11.0742 - 0.505(4.0)}{(4.0)^2 + (7.36)(4.0) - 25.224} = \frac{9.0542}{20.216} = .44787,$$

$$x_2 = \frac{7.845 - (4.305)(4.0)}{20.216} = \frac{-9.375}{20.216} = -.46374,$$

$$R^2 = .44787^2 + (-.46374)^2 = .41564,$$

and

$$\hat{y} = (4)(.41564) - 0.505(.44787) + (4.305)(.46374) + 82.17$$

$$= 85.60 .$$

Various values of $\lambda > 2.54625$ were used to obtain the plot of R versus $\hat{y}$. Figure 2-3 contains the graph of $\hat{y}$ versus R for all four ridges. Ridges R_{AMAX} and R_{AMIN} are the absolute maximum and minimum ridges and the intermediate or secondary ridges are labeled R_{SMAX} and R_{SMIN}.

To obtain these graphs values of λ were used as follows:

For R_{AMAX} $\lambda > 2.55$,
R_{SMAX} $0 \leq \lambda < 2.5$,
R_{SMIN} $-9.9 < \lambda \leq 0$, and
R_{AMIN} $\lambda < -9.9$.

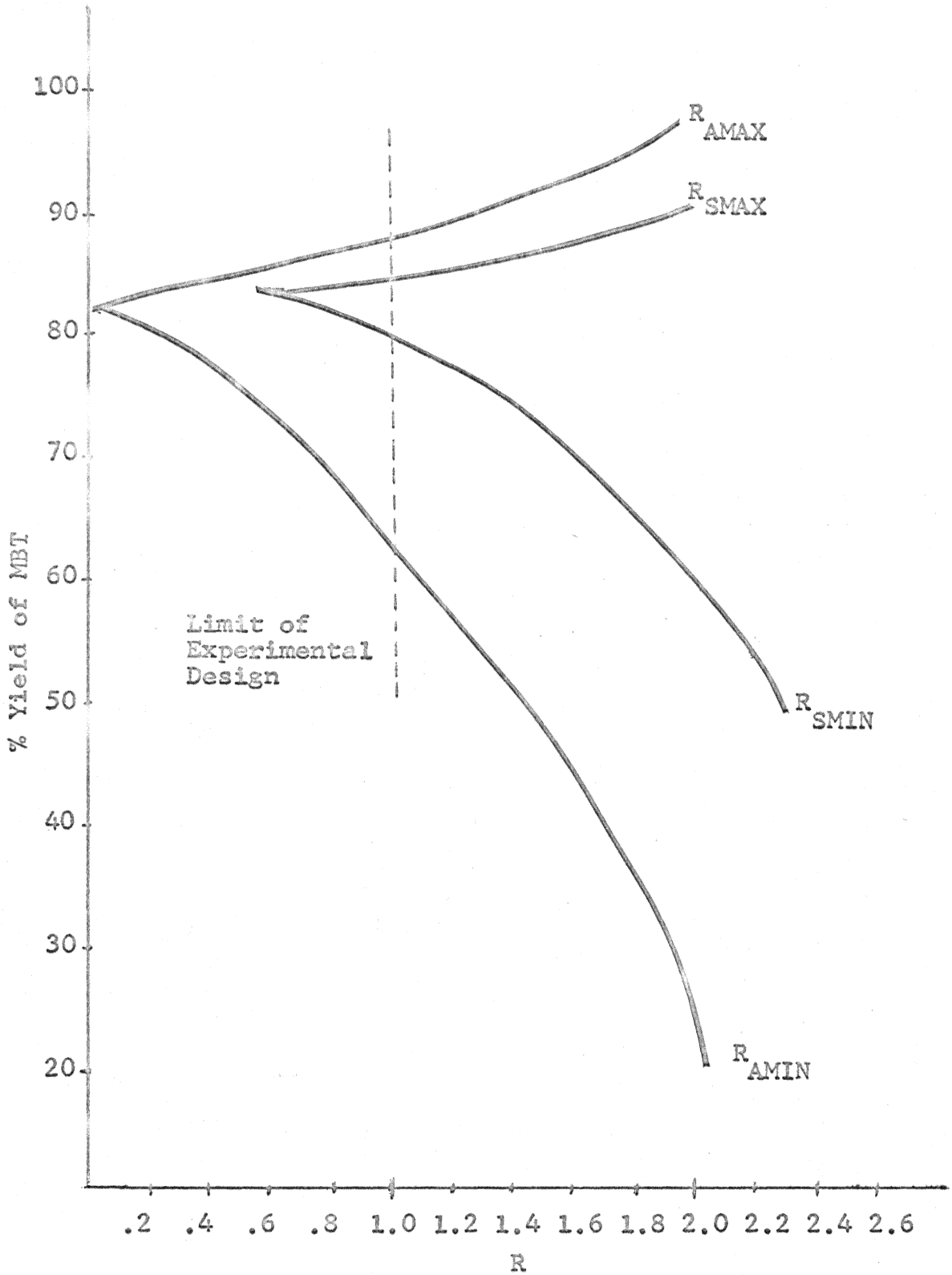


FIGURE 2-3. Yield of MBT versus R for Ridge Analysis Example.

As would be expected, the secondary ridges do not appear until R equals the distance from the center of the design to $\underline{x}_o$. In this example

$$\underline{x}_o = (-.439, -.311)$$

with yield $\hat{y}_o = 83.73$, so

$$R = ((-.439)^2 + (-.311)^2)^{\frac{1}{2}} = .538$$

is the minimum distance needed to obtain secondary ridges.

Figure 2-4 contains the plots of x_1 and x_2 against R for the absolute maximum ridge. Figure 2-5 plots x_1 and x_2 against R for the secondary maximum ridge. Figures 2-6 and 2-7 plot x_1 and x_2 against R for the secondary and absolute minimum ridges respectively.

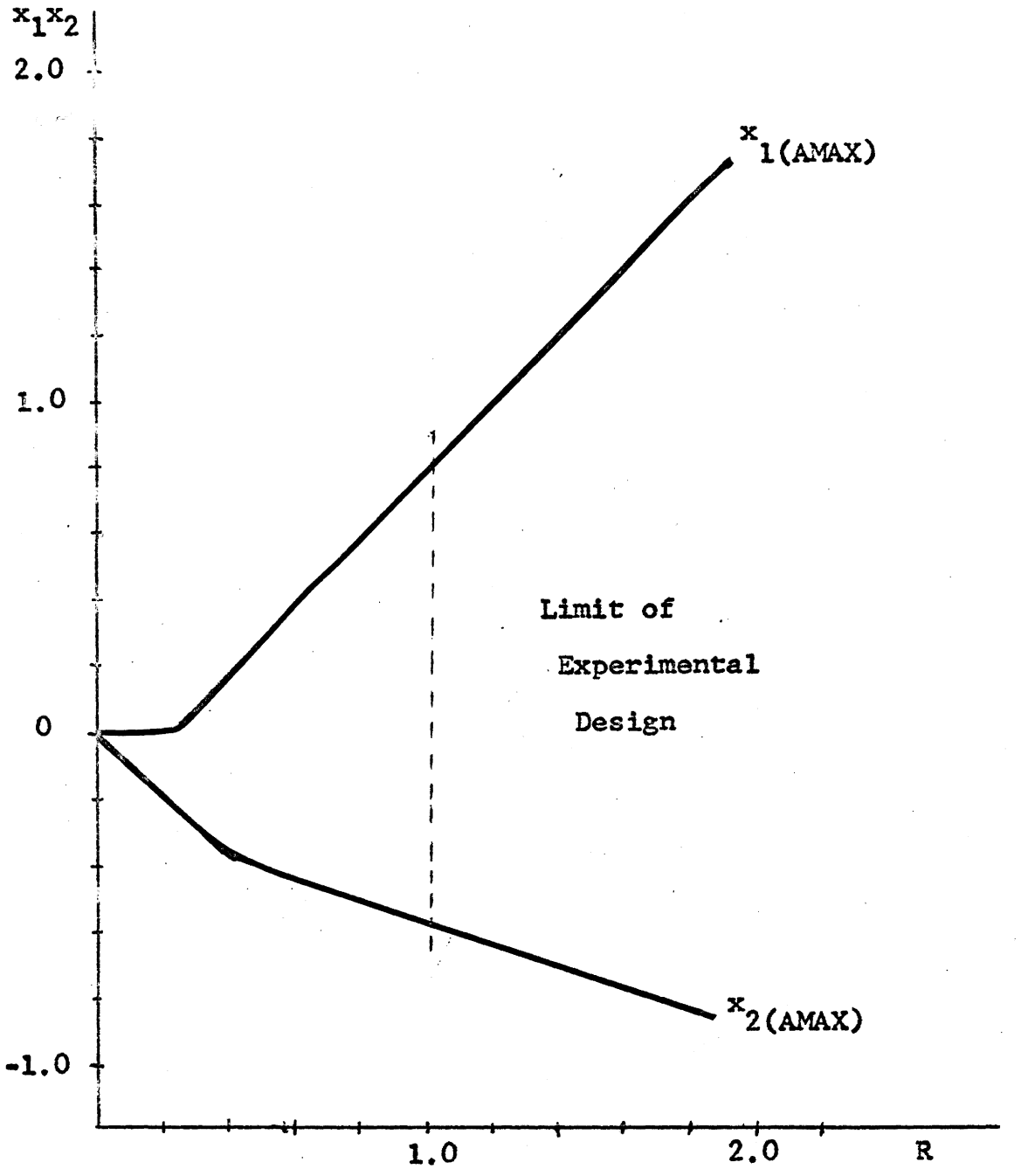


FIGURE 2-4. R versus x_1 and x_2 for the absolute maximum ridge.

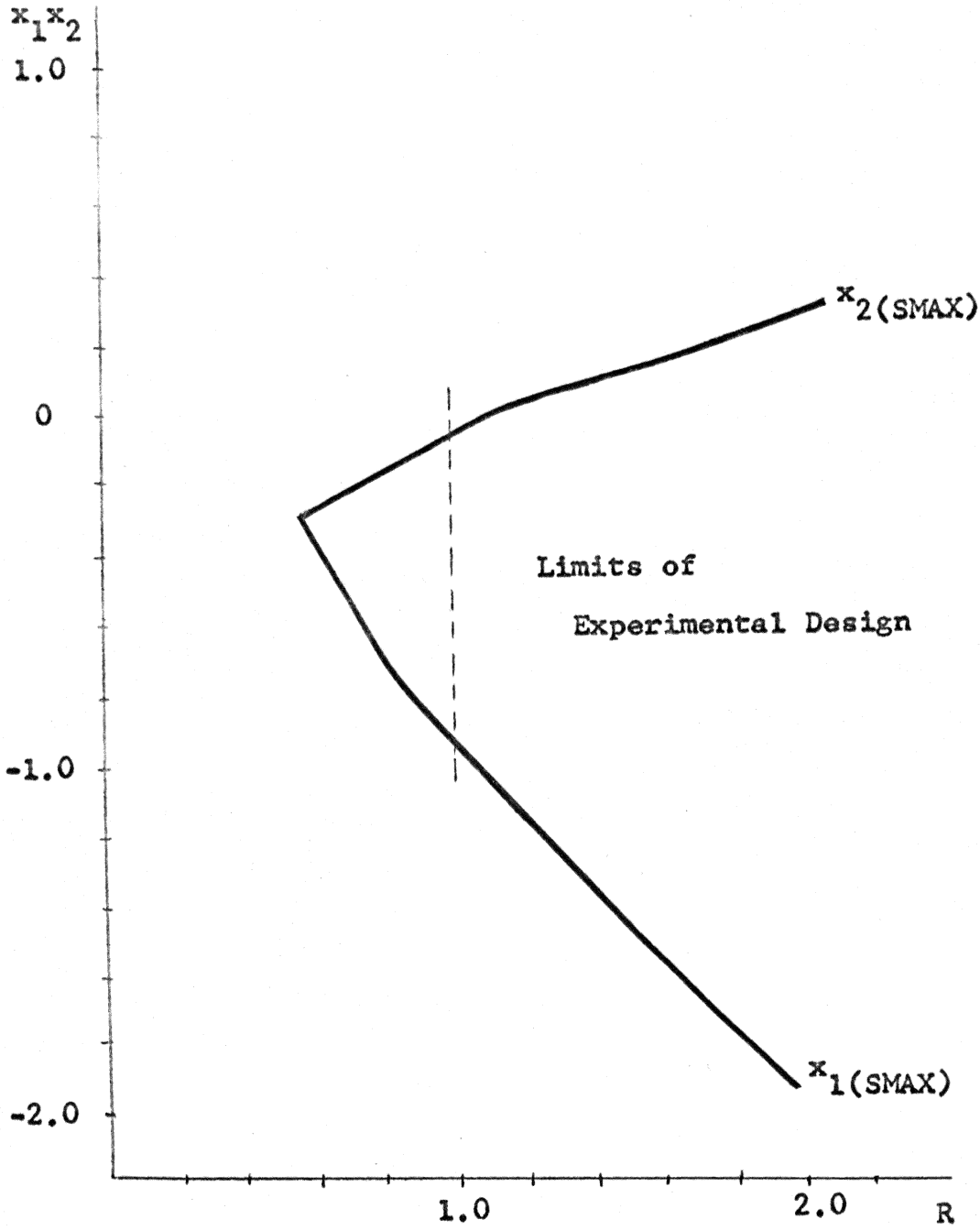


FIGURE 2-5. R versus x_1 and x_2 for the secondary maximum ridge.

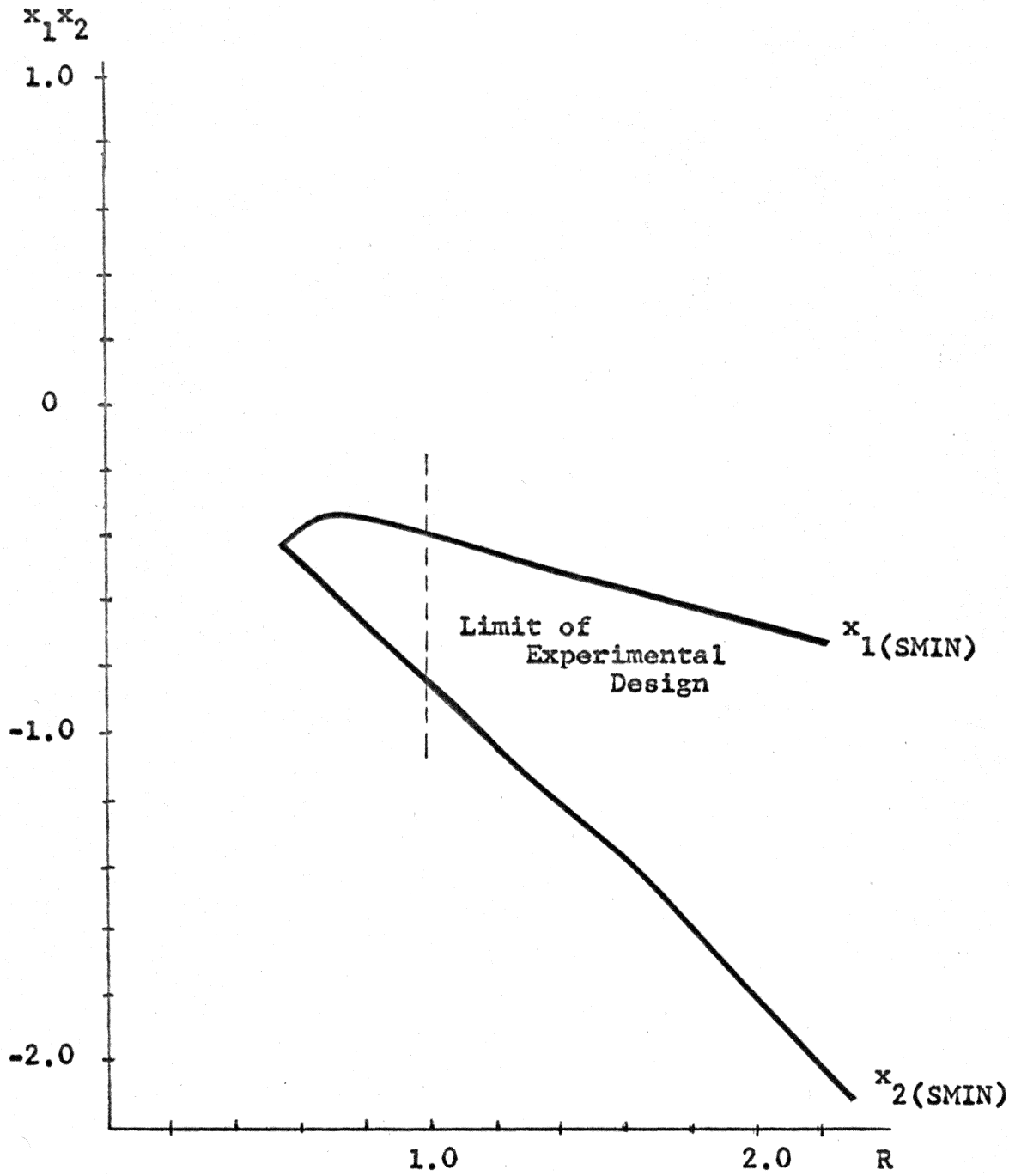


FIGURE 2-6. R versus x_1 and x_2 for the secondary minimum ridge.

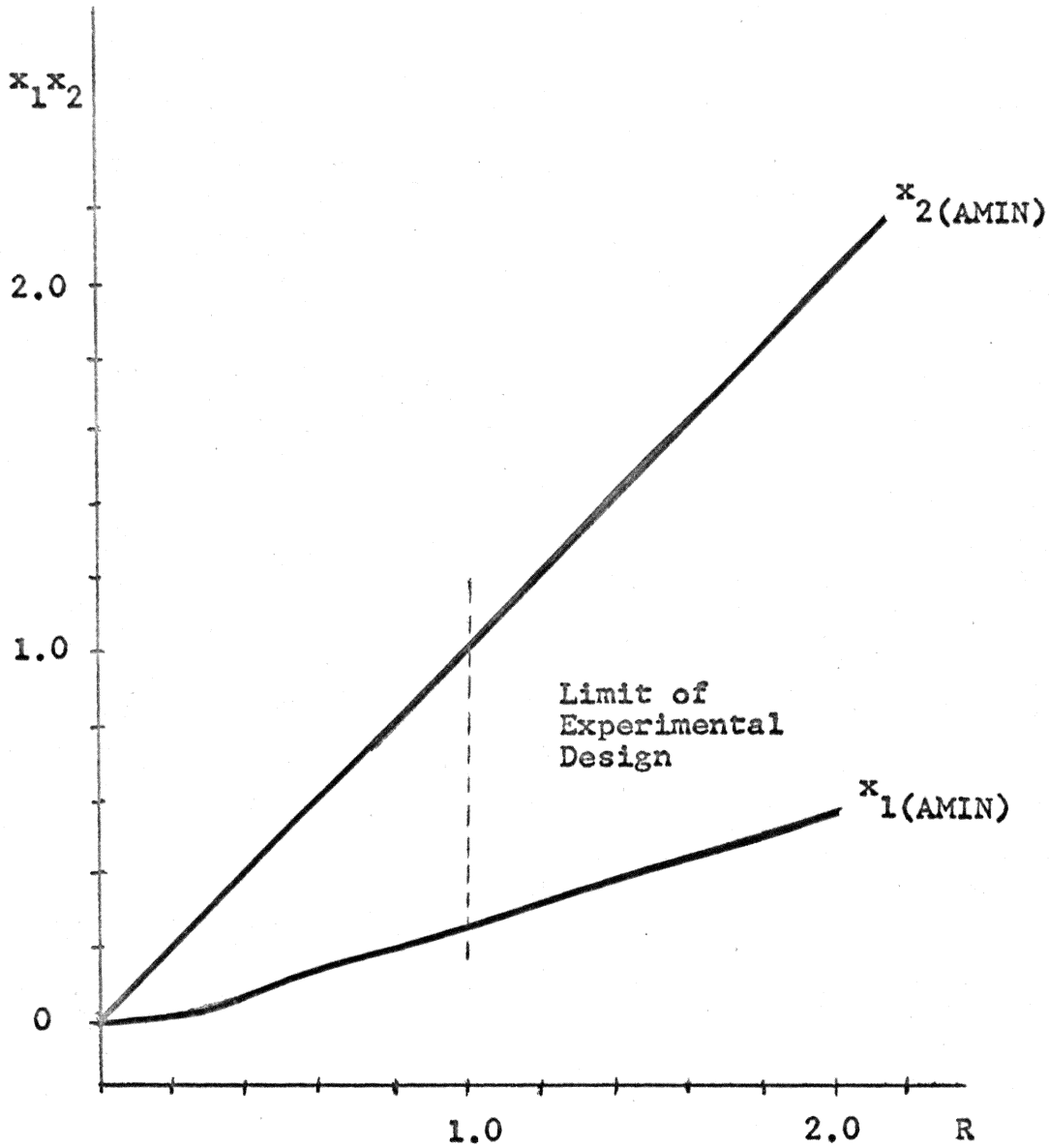


FIGURE 2-7. R versus x_1 and x_2 for the absolute minimum ridge.

CHAPTER III OPTIMIZATION TECHNIQUES FOR PRODUCTION PROCESSES

The optimization of existing production processes must be performed in such a manner that the output is not dramatically changed. To do this specific experimental procedures have been developed in which only small changes to the existing process are allowed. Two of these procedures are discussed in this chapter. The next chapter considers other techniques, some of which can be applied to production processes, but which are usually applied to initial experimentation or pilot-production experimentation.

A. EVOLUTIONARY OPERATION

1. Basic Philosophy

In 1957 George E. P. Box[3] proposed a method for improving the operation of a process. This method is called Evolutionary Operation or EVOP. It is a method of process operation which has a built-in procedure for increasing productivity of the process. Although it was developed for a chemical type process it has more general application and several other industries have used it profitably.

EVOP is a method of production which uses some simple statistical concepts during the regular normal routine of production. The method consists of running a simple experiment, usually a factorial, by the production personnel themselves. The basic philosophy of EVOP is that it is inefficient

to run an industrial process for the product alone. The process should be run so as to generate product plus information on how to improve the product.

Box makes the important point that EVOP is not a substitute for more fundamental designed experiments aimed at a better understanding of the process. These experiments (and process theory) are always necessary to obtain basic information about the actual operation of the process. On the contrary, EVOP may point out areas where more fundamental research is needed. EVOP was specifically designed to improve upon the "best" operating conditions found through designed experiments and to make continued improvement on the process.

One example of the use of EVOP concerns the resulting "scale-up" problems usually found when the process graduates from the laboratory, to pilot production, and finally to full-scale production. It is usually found that considerable modification of the operating conditions is required to obtain yields near those obtained in the laboratory.

EVOP has been compared to the process of evolution in biology. Living organisms advance by two methods:

1. Mutation
2. Natural Selection.

Chemical processes also advance by two similar methods. A discovery of a new route to the final product is equivalent to genetic mutation. Adjustment of the operating conditions corresponds to natural selection. Adjustments that "work"

are retained and adjustments that harm the product are avoided in the future. Thus, the process is slowly optimized. Evolutionary operations employ a method of purposeful adjustment to speed up this optimizing process.

2. Operation of EVOP

Routine production is normally run at rigidly defined operating conditions for the process. This is called the "works process" and is the best set of operating conditions found for the process. Any method of introducing variation in the works process must provide safeguards which will ensure that the risk of producing any appreciable unsatisfactory product is small. For this reason EVOP is based on the premise that only small changes are allowable in applying the method. Since production must continue during the EVOP procedure, the effects of small changes in operating conditions can be detected by continued replication of the basic experiment.

The application of EVOP first involves the selection of the responses which are to be optimized. This may be yield, cost/pound, tensile strength, etc. Generally several responses are considered because the product is usually too complicated to be represented by only one response. Optimization is then simultaneously attempted. Some of the responses, expecially properties of the product, are not optimized but are observed so that the effects of the variation in operating conditions may be seen.

As an example consider the manufacture of a liquid product. The main response to be optimized could be the cost per gallon. In addition, suppose that the level of an impurity must not exceed .5% and the fluidity of the product must lie between the limits of 55 and 80. These last two responses are not to be optimized but must be satisfied by the product. Measurements of the level of impurity and the fluidity would be made and analyzed as well as the cost/gallon to make sure the product was acceptable.

Next, the particular operating conditions which will be systematically varied must be chosen. To keep the experimental designs simple usually only two or three process variables are chosen for study. The limits on the variables are then specified, through considerations previously mentioned.

The most common experimental design used for EVOP is that of a 2^k factorial plus a center point. For two variables, the design points are given in Figure 3-1. To start, the center point is the works process for the two variables of interest.

The routine of plant operation consists of repeatedly running the production process at these five design points in the order 1, 2, 3, 4 and 5. Each group of 5 runs is called a cycle. Randomization of the order of running these design points is rarely performed and would only tend to confuse the production personnel. At the end of each cycle the results

of the responses are posted to an Information Board which is kept up-to-date on a cycle basis. The Plant Manager can inspect the Information Board at any given stage and make one of the following decisions:

1. Wait for further information.
2. Adopt one of the design points 2, 3, 4 or 5 as the new "works process" and start the cycle around it as a center point.
3. Change the levels of the process variables by increasing (or decreasing) their range.
4. Substitute new variables for one or more of the variables under study or add variables.

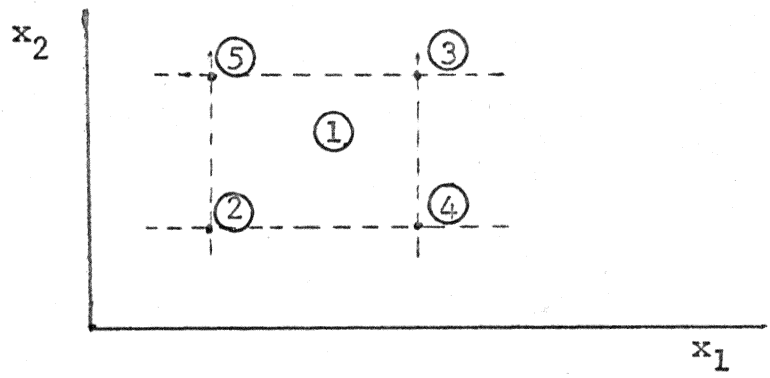
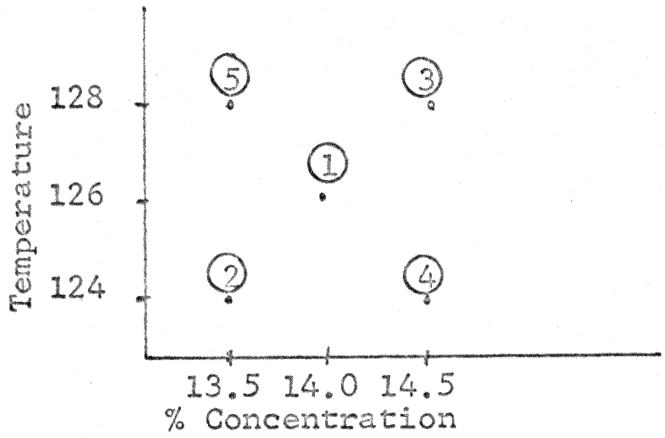


FIGURE 3-1. Design points for an EVOP study in two variables.

As an example of an Information Board, consider the previous example of responses where the variables under study are per cent concentration of one of the ingredients and temperature of the reaction. Figure 3-2 shows the Information Board at the end of 16 cycles. In this example the effect of concentration has clearly exceeded the approximate 95% con-



Phase 1.

Cycle 16.

	Cost (cents)/gal	Impurity %	Fluidity
Requirement	Minimize	Less than .5%	Between 55 & 80
Running Averages	32.6 33.9 / 32.8 / 32.4 33.4	0.29 0.35 / 0.27 / 0.17 0.19	73.2 76.2 / 71.3 / 60.2 67.6
95% error limits on averages	±0.7	±0.03	±1.1
Effects			
Concentration	1.2 ±0.7	0.04 ±0.03	5.2 ±1.1
Temperature	0.4 ±0.7	0.14 ±0.03	10.8 ±1.1
C X T	0.1 ±0.7	0.02 ±0.03	-2.2 ±1.1
Change in Mean	0.2 ±0.6	-.02 ±0.03	-1.6 ±1.0
Standard Deviation	1.44	0.059	2.12

FIGURE 3-2. Example of an Information Board after 16 cycles.

fidence limits, called error limits in EVOP. The calculation of these error limits are discussed later. An examination of the impurity and fluidity responses shows that a shifting of the works process to point 2 would not harm the product. In fact it would also tend to minimize the impurity level. In such a case, the decision would usually be made to shift to point 2 as a new works process. The EVOP study would then enter a new phase (Phase Two in this example).

The phase number on the Information Board indicates the experiment number. Whenever any change in levels, substitution or addition of variables is made, the phase number is increased.

3. Calculation of Effects

The "effects" are calculated from the running averages. If $\bar{y}_1$, $\bar{y}_2$, $\bar{y}_3$, $\bar{y}_4$ and $\bar{y}_5$ are the running averages for the r^{th} cycle, the effects of concentration, temperature, and interaction are calculated as in a 2^2 factorial experiment ($\bar{y}_1$, the center point is ignored). The main effects are interpreted as the difference in average response in going from the low level of the variable to the high level when averaged over all the other variables. The interaction effect measures how the levels of the variables interact with one another. If the interaction effect is zero, the variables act independently of each other.

The change in mean (CIM) is the difference between the

center point and the average yield for all of the points $\bar{y}_1$, $\bar{y}_2$, $\bar{y}_3$, $\bar{y}_4$ and $\bar{y}_5$. The change in mean is used for assessing non-linearity and the cost of running the EVOP study. The formulas for calculation of the effects used are (for two variables):

Effect	Formula	Limits of error
Concentration	$\frac{1}{2}(\bar{y}_3+\bar{y}_4-\bar{y}_2-\bar{y}_5)$	$\pm \frac{2s}{\sqrt{r}}$
Temperature	$\frac{1}{2}(\bar{y}_3+\bar{y}_5-\bar{y}_2-\bar{y}_4)$	$\pm \frac{2s}{\sqrt{r}}$
C X T	$\frac{1}{2}(\bar{y}_2+\bar{y}_3-\bar{y}_4-\bar{y}_5)$	$\pm \frac{2s}{\sqrt{r}}$
CIM	$\frac{1}{5}(\bar{y}_2+\bar{y}_3+\bar{y}_4+\bar{y}_5-4\bar{y}_1)$	$\pm \frac{1.78s}{\sqrt{r}}$

The limits of error are approximately 95% confidence limits on the estimated effects. They are ± 2 times the estimated standard error of the effect.

The change in mean, CIM, as defined before is

$$\text{CIM} = \frac{\bar{y}_1+\bar{y}_2+\bar{y}_3+\bar{y}_4+\bar{y}_5}{5} - y_1 = \frac{1}{5}(\bar{y}_2+\bar{y}_3+\bar{y}_4+\bar{y}_5-4\bar{y}_1).$$

The variance of CIM is given by

$$\text{var(CIM)} = \frac{1}{25}\sigma^2_{\bar{y}}(1+1+1+1+16) = \frac{4}{5}\sigma^2_{\bar{y}}$$

$$= \frac{4}{5r}\sigma^2_y$$

Hence $\pm$ 2 times the standard error for the CIM limits is

$$\pm 2 \sqrt{\frac{4}{5r} s^2} = \frac{\pm 1.78\ s}{\sqrt{r}} .$$

4. Discussion of Change in Mean

If the true response surface were quadratic, the CIM is an estimate of a function of the pure quadratic terms. To show this, consider a two-variable EVOP where x_1 and x_2 are coded so that the values of x_1 and x_2 are given as in the following table:

Point	x_1	x_2	Response
1	0	0	y_1
2	-1	-1	y_2
3	1	1	y_3
4	1	-1	y_4
5	-1	1	y_5

For the model

$$y = \mu+\beta_1 x_1+\beta_2 x_2+\beta_{12} x_1 x_2+\beta_{11} x_1^2+\beta_{22} x_2^2$$

the expected values of y_1, y_2, y_3, y_4, and y_5 are

$$E(y_1) = \mu$$

$$E(y_2) = \mu-\beta_1-\beta_2+\beta_{12}+\beta_{11}+\beta_{22}$$

$$E(y_3) = \mu+\beta_1+\beta_2+\beta_{12}+\beta_{11}+\beta_{22}$$

$$E(y_4) = \mu+\beta_1-\beta_2-\beta_{12}+\beta_{11}+\beta_{22}$$

$$E(y_5) = \mu-\beta_1+\beta_2-\beta_{12}+\beta_{11}+\beta_{22}$$

Hence

$$E(CIM) = E\left\{\frac{1}{5}(y_2+y_3+y_4+y_5-4y_1)\right\}$$

$$= \frac{1}{5}(4\beta_{11}+4\beta_{22}) = \frac{4}{5}(\beta_{11}+\beta_{22})$$

Substitution of $\bar{y}_i$ for y_i would not change this result.

Hence, the CIM provides an estimate of $c \sum_{i=1}^{k} \beta_{ii}$ where c is $2^k/(2^k+1)$.

As such the CIM gives one some idea of the non-linearity of the local response surface. If the local response surface is planar, the expected value of the CIM is zero. If the response surface were locally quadratic with a maximum, say, at the works process, the CIM would be negative. If the response surface had a minimum at the works process (such as for a response like cost), the CIM would be positive. In this latter case the CIM would be the cost incurred by running the EVOP production rather than all of the production at the works process. For the above example, the cost of EVOP was 0.2 $\pm$ 0.6 cents per gallon. In the case of a maximum, the CIM would estimate the loss in yield incurred by running the EVOP procedure. Except when the process has been brought very close to the optimum, the cost incurred by EVOP is very small compared to the cost saved

in the operation of the process if any improvement has been made at all by EVOP.

The Information Board also shows the standard deviation of the response and may also show the previous estimate obtained at the end of the previous cycle. EVOP uses a sequential estimate for σ and this additional information gives an indication of the stability of this estimate of σ. In some cases, however, it may be desirable to omit this information from the board, as was done in the example.

5. Three-variable EVOP

For three variables in the EVOP study, the design points are shown in Figure 3-3. Note that the center point of the design is replicated so that 10 observations are needed to complete a cycle. If the runs need to be separated into two blocks, one of the center points is run in each block.

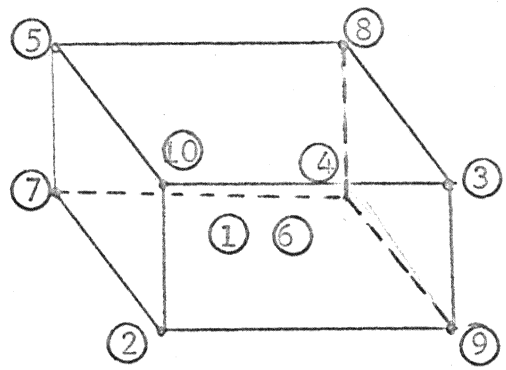


FIGURE 3-3. Design points for a 3-variable EVOP study.

6. Estimation of s

In starting an EVOP study there may be no reliable estimate of the response variance. However, after several cycles there is no lack of data and the response error variance can be estimated, for example, by the method of least-squares.

However, this would involve the squaring of considerable number of observations and would be fairly difficult computationaly, especially for routine analysis by plant personnel. Instead a method due to Box and J. S. Hunter[5] utilizes a range method to calculate an estimate of the error standard deviation. Since a multitude of responses are available after a few cycles, there seems little point in not using this slightly less efficient estimate.

Suppose N sets of conditions are run in an EVOP study. (N=5 for the two variable case and N=10 for the three variable EVOP. However, this technique is quite general.)

At the end of (r-1) cycles, there are N(r-1) observations and they can be classified in the following two-way table along with the data from the r^{th} cycle:

Average for N conditions at end of (r-1) cycles	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	$\cdots$	$\bar{y}_N$
New data from r^{th} cycle	y_{1r}	y_{2r}	y_{3r}	$\cdots$	y_{Nr}
Differences $\delta_i=(\bar{y}_i-y_{ir})$	δ_1	δ_2	δ_3	$\cdots$	δ_N

Now

$$\delta_i = \sum_{j=1}^{r-1} y_{ij}/r-1 - y_{ir},$$

where y_{ij} is the response of the i^{th} condition for the j^{th} cycle.

Assume for a given cycle j, y_{ij} is normally distributed with mean $\mu+\alpha_i$ and variance σ^2, i=1,2,...,N; j=1,2,..., r. Then

$$\text{var}(\delta_i) = \sigma_\delta^2 = \frac{1}{(r-1)^2} \sum_{j=1}^{r-1} \text{var } y_{ij} + \text{var } y_{ir}$$

$$= \frac{\sigma^2}{(r-1)} + \sigma^2$$

$$= \frac{r}{r-1} \sigma^2$$

Hence,

$$\sigma^2 = \frac{r-1}{r} \sigma_\delta^2$$

and

$$\sigma = \sqrt{\frac{r-1}{r}}\, \sigma_\delta$$

Now σ_δ can be estimated from the range R of the δ_i. An unbiased estimate of σ_δ is[4]

$$\hat{\sigma}_\delta = \frac{R}{d_2}$$

where d_2 depends on N. Values of d_2 can be found in any quality control textbook. Hence, σ can be estimated by

$$s_r = \sqrt{\frac{r-1}{r}}\, \frac{R}{d_2} = f(N,r)R$$

The quantity $\sqrt{\frac{r-1}{r}} \cdot \frac{1}{d_2}$ is called f(k,n) in Box and Hunters' notation and values of f(N,r) are given below for selected values of N and r.

Table 3-4. Values of f(N,r) for use in estimation of s for EVOP.

N/r	2	3	4	5	6	7	8	9	10
5	.30	.35	.37	.38	.39	.40	.40	.40	.41
10	.23	.26	.28	.29	.30	.30	.30	.31	.31

For each cycle after the second, an estimate of σ is made, and the running average $s_r = \sum_{j=2}^{r} s_j/r-1$ is used for the estimate of σ for the r^{th} cycle in the calculation of error limits on the variable effects and CIM.

Since the estimate of σ is not too reliable for the first several cycles, the prior phase estimate can be used for these first few cycles. By the time enough data have been accumulated to make a decision on the process, usually a sufficient number of cycles has been run to provide a reliable estimate from the data of this phase. For phase 1, an estimate of σ can often be obtained from past process history or quality control data.

7. Advantages and Disadvantages of Box's EVOP

One advantage of using the factorial plus center point design is that, if the CIM indicates considerable non-linearity, only a few additional points need be added to obtain a central composite design[2] for the estimation of a quadratic response surface. This estimated quadratic response surface can be used to locate the optimum response conditions and provide an estimate of the optimum attainable. Chapter 1 gives the required analysis for this procedure.

The basic EVOP design allows the addition of more variables very easily and, for three or more variables, blocking can be accomplished by confounding higher-order interactions. In addition, for this design the levels of the factors need not be quantitive.

There are several disadvantages to this design for an EVOP procedure. The effect of the variables is clearly dependent on the allowable range of the variable. If the range is too short, a large number of cycles may be required before a prudent plant manager will increase the range of the variable.

If a large number of variables, k, are of interest, a complete cycle would require at least 2^k+1 production runs and this may be too many. However, fractional designs can be used to help alleviate this problem.

8. EVOP Committee

Box recommended that an EVOP committee be formed consisting of production and technical personnel (including a statistician, if possible). This committee, along with the plant manager or process foreman, can choose the variables to be studied and set their levels, discuss and interpret the results of the Information Board, and provide other technical inputs to the plant manager to aid him in his decisions in the EVOP study. The EVOP procedure is rather conservative since a decision to move the works process would not usually be made unless some effect exceeded its error limits or unless a large number of cycles showed no significant effect. In the latter case, the plant manager must decide whether to increase the variable ranges or else add or substitute new variables in the EVOP study. The EVOP committee, with its varied type of personnel, can be most helpful to the plant manager in such situations by providing alternate recommendations to him.

It is very important to realize that EVOP is a philosophy of production and in almost all situations, EVOP should be considered as a permanent method of production. This is especially true for the type of process whose inputs (raw material, operating personnel) change with time. If EVOP is always in use, the production process will automati-

cally have a chance to follow the optimum operating conditions as the inputs vary slowly with time. A more automative EVOP procedure to do this is presented in the next section.

B. SIMPLEX EVOLUTIONARY OPERATION

In the EVOP procedure of the preceding section there were no precise rules regulating when or where to move the works process in the EVOP experimentation. This section will present an automatic EVOP procedure in which some simple rules will be given to control adjustments of the operating conditions.(33)

1. Simplex Design

The basic design in this EVOP is a simplex. A simplex is an orthogonal first order design which requires only one more experimental point than the number of variables under consideration. Thus for k variables, N=k+1. The columns of the design matrix are orthogonal and the design is formed by using a regular sided figure with N=k+1 vertices so situated that the cosine of the angle formed by any two vertices with the center of the design is constant and equal to -1/k. For k=3, the design is an equilateral triangle; for k=4, the design is a regular tetrahedron.

For this application a slight variation of the standard

simplex design for k variables is used. Let the works process be coded $(0,0,\cdots,0)$; a regular simplex of unit edge is then specified by the (k+1) x k design matrix:

$$D_o = \begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ p & q & q & \cdot & \cdot & \cdot & q \\ q & p & q & \cdot & \cdot & \cdot & q \\ \cdot & \cdot & \cdot & & & & \\ q & q & q & \cdot & \cdot & \cdot & p \end{pmatrix}$$

where $p = \frac{1}{k\sqrt{2}}(k-1+\sqrt{k+1})$

and $q = \frac{1}{k\sqrt{2}}(\sqrt{k+1} - 1)$.

The rows of this matrix give the coordinates of the k+1 design points. Table 3-5 gives values of p and q for some values of k.

Table 3-5. Values of p and q for different values of k in a Simplex Design.

k	p	q
2	.966	.259
3	.943	.236
4	.926	.219
5	.912	.205
6	.901	.194

2. "Simplex EVOP"

In "Simplex EVOP" the operating conditions are changed on every production run and the new operating con-

ditions are chosen so as to make a new simplex design using k of the previous (k+1) points and the new run. This new point is chosen to maximize (or minimize) the response.

The direction of maximum response out of a simplex would proceed in some direction from the center of the design out of the side opposite to the lowest response. Hence the point corresponding to the lowest response in the simplex should be discarded and the new run should be made so as to make a simplex out of the remaining k experimental points and the new run. For a response which is to be minimized, the highest observation would be discarded. To illustrate, consider Figure 3-6 for two variables (a maximization problem). In Figure 3-6 the numbers of the production runs are circled and the response is given beside each experimental point. In this figure trial 5 moved away from the general direction of the maximum due to components of random error in trials 2 and 3 . Note, however, that only two trials were required to correct for this error.

Note that the direction of advance is determined solely on a ranking scale. Hence absolute numbers for the response are not required. Since this procedure is limited to only one response, multiple observations on the product must be handled by comparing the characteristics of the product as a whole and the least favorable operating conditions then chosen for the discarded point.

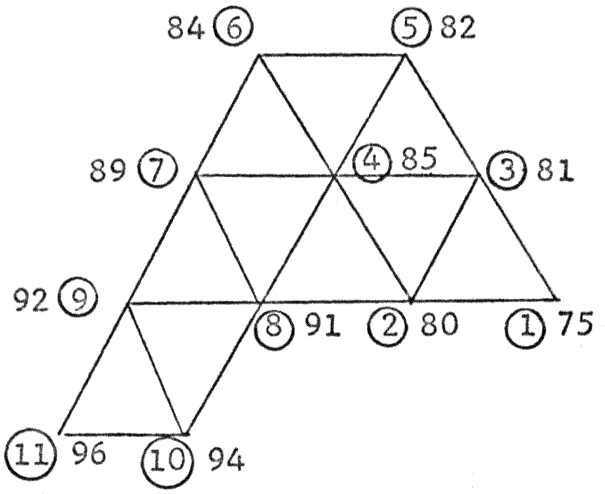


FIGURE 3-6. Illustration of a two variable simplex EVOP.

Since this sequential procedure can move away from the works process very quickly when the response at the works process is quite low compared to the optimum, a "unit" on each of the variables should be chosen quite small so as to reduce the risk of a bad production run. This procedure is not scale invariant. A unit on each of the variables should be chosen in such a way that a unit change is of equal interest to the experimenter.

To obtain the coordinates of the new experimental point the following procedure is followed:

Let the rows of D_o be denoted by the vector $\underline{d}_i'$, $i=1,\cdots,k+1$. The coordinates, a vector, of the new run if the lowest observation occurred at $\underline{d}_i'$ is given by

$$\underline{d}_i'^* = \frac{2}{k}(\underline{d}_1'+\underline{d}_2'+ \cdots +\underline{d}_{i-1}'+\underline{d}_{i+1}'+ \cdots +\underline{d}_{k+1}')-\underline{d}_i' .$$

In words, the coordinate of the new point is twice the

average of the coordinates of the common points minus the coordinate of the rejected point.

Once the design matrix has been obtained, the usual procedure would involve uncoding the design points and proceeding to new points in uncoded units. This can be easily accomplished in tabular form.

3. Simplex EVOP Rules

The Simplex EVOP procedure for a maximation is given by application of the following rules. For a minimization, replace "lowest" with "highest" throughout these rules.

Rule 1. Ascertain the lowest reading y_i of y_1, y_2, y_3, ..., y_{k+1}. Complete a new simplex by excluding $\underline{d}_i'$ and running at $\underline{d}_i'^*$.

Since the responses are subject to error, there is a chance that the system of simplices may become anchored to a spuriously high response. To reduce this possibility, Rule 2 is applied.

Rule 2. If a result has occurred in k+1 successive simplices and not then eliminated by Rule 1, do not form a new simplex but discard this result and replace it by a new observation at that same point. Then apply Rule 1.

The philosophy behind this rule is that if the point is at a true optimum the replication will also be high and the system of simplices will again be clustered about it.

If it were high due to error, the replication will probably not be so high and would be eliminated in due course.

A spuriously low response will generally be eliminated quickly but may cause some oscillation from one simplex to a previous one. Hence some benefit can be obtained by applying:

Rule 3. If y_i is the lowest reading in the simplex S_o, and y_{i*} is the lowest reading in the new simplex, S_1, do not move back to S_o. Instead reject the second lowest reading of S_1.

The application of these rules causes the system of simplices to circle continuously about the optimum rather than oscillate over a limited range. This is especially helpful if the optimum changes in time or changes as a function of raw materials since the system is always free to follow the optimum. Rule 3 makes progress possible if the system straddles a ridge in the variable space.

The three rules given above may be summarized by the following:

Move by rejecting the lowest response unless (a) another response is too "old", in which case we replicate the "old" response, or (b) such a move would cause us to return to the previous simplex, in which case we try the next favorable direction of movement.

A new variable may be added at any time by simply

running one additional point and then proceeding as before. Deletion of a variable however, requires one to initiate experimentation with a new design matrix.

If there are contraints on the levels of the variables of the form $a \leq x \leq b$, whenever this Simplex EVOP leads to a point that would violate the contraint, the second most favorable direction would be used.

4. Discussion of Simplex EVOP

Spendley, Hext, and Himsworth[33] showed by computer simulation that the rate of advance of Simplex EVOP is inversely proportional to the error deviation. Now a replication of n observations at a point reduces the error standard deviation by a factor $\sqrt{n}$ and the rate of advance gains by a factor of $\sqrt{n}$. However, on a per observation basis, the expected gain is reduced by this factor $\sqrt{n}$. Hence replication not only has no value but is positively detrimental.

Spendley, et al, also evaluated the efficiency of the simplex procedure compared to a very simply defined evolutionary plane-climbing procedure. This procedure was one in which an observation at the current position is compared with one at a unit distance in some randomly chosen direction. If this new observation is higher, the system moves to this point. If the new observation is lower, the system is moved in the diametrically opposite direction. Since they were

able to theoretically evaluate this method of advance (see (33) for details) the efficiency of the simplex procedure (in terms of expected advance per observation) was compared to it. They found that the efficiency of Simplex EVOP increased in a direct proportion as k, the number of variables, increased.

Hence the best use of Simplex EVOP uses as many variables as possible with no replication. As such, Simplex EVOP seems to be a very good alternative to EVOP. The only disadvantage of Simplex EVOP compared to EVOP is that Simplex EVOP requires quantitive process variables and EVOP does not.

Simplex EVOP has been successfully utilized in industry in several different processes.(15)

CHAPTER IV EXPERIMENTAL OPTIMIZATION TECHNIQUES

In an experimental situation the experimenter may have a good idea of the levels of his experimental factors which produce a near optimal response. In such cases Phase Two of the general response problem is not required and the experimenter can proceed to Phase Three for further exploration. Often, however, the experimenter is not so fortunate in having such detailed information about the response system. He must then utilize some search technique to locate the levels of the factors to provide a near optimal response. In this chapter several search techniques are discussed, starting with the simplest case of only one experimental factor.

A. ONE FACTOR TECHNIQUES

1. Hotelling's Procedure

In 1941, Hotelling[25] considered the problem of finding the optimum response when a single factor is involved. His method consists of the following steps:

(1) A study is made to indicate the general range on the factor in which the optimum is most likely to lie.

(2) An intermediate experiment is run to provide gross estimates of the parameters of the response equation

$$y = \beta_o + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \cdots.$$

For example, 6 equally spaced levels of x could be used and a 5^{th} degree polynomial could be estimated. The usual derivative procedure can then be used to find the optimum.

(3) A final experiment is run in the neighborhood of the estimated optimum to further refine one's estimate. This experiment assumes y can be approximated by a quadratic equation

$$y = \beta_o + \beta_1(x-m) + \beta_2(x-m)^2$$

where m is the value of x at the maximum. Hotelling shows how to allocate the N samples to make any cubic bias zero and any quartic bias a minimum.

2. Stochastic Techniques

More recently Kiefer and Wolfowitz[29] have provided a method of finding the optimum (maximum for this discussion). Let y(x) be the regression function of the response as a function of x. The K-W technique involves the determination of the average slope of the function of y(x) by a very specialized method. The average derivative is calculated by observing the response at two points a distance c_n on either side of x_n as

$$\Delta_n = \frac{y(x_n+c_n) - y(x_n-c_n)}{2c_n}$$

See figure 4-1 for an illustration of this method.

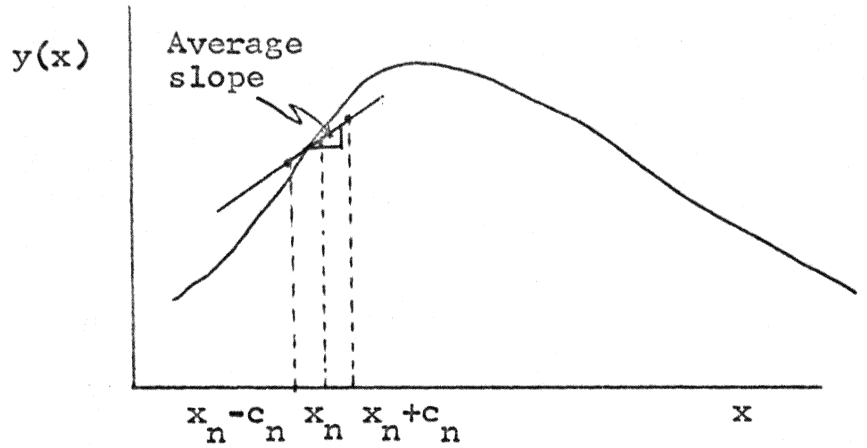


FIGURE 4-1. Kiefer-Wolfowitz estimation of average slope.

The center of the next pair of trials is centered around x_{n+1} where

$$x_{n+1} = x_n + a_n\Delta_n = x_n + a_n \frac{y(x_n+c_n) - y(x_n-c_n)}{2c_n}$$

where a_n is one of a sequence of positive numbers which determine the step size, and $2c_n$ is the distance between the last pair of trials. An example of the sequences $\{a_n\}$ and $\{c_n\}$ are:

$$a_n = \frac{1}{n}$$

$$c_n = n^{-1/3}$$

In general, the sequences $\{a_n\}$ and $\{c_n\}$ must have the following properties:

i. $\lim_{n\to\infty} a_n = 0$

ii. $\lim_{n\to\infty} c_n = 0$

iii. $\sum_{n=1}^{\infty} a_n = \infty$

$$\text{iv.} \quad \sum_{n=1}^{\infty} \left(\frac{a_n}{c_n}\right)^2 < \infty$$

Properties i and ii are necessary to assure one that the process will converge. Property iii gives sufficient steps so that one will always straddle the peak. Property iv is used to cancel out the cumulative error effect. With some very mild restrictions on y(x), this sequence of trials will converge to m with probability one. The proof of this is given in Kiefer and Wolfowitz[29] and will not be shown here.

Because of the method of estimating the average slope, the K-W procedure can sometimes be slow in convergence. Consider Figure 4-2, for example, in which the steps in approaching m from the left would be large. However when m was "overshot" on the right, the steps would be very slow in converging back to m. Thus the K-W technique would spend most of its time trying to climb the low slope on the right.

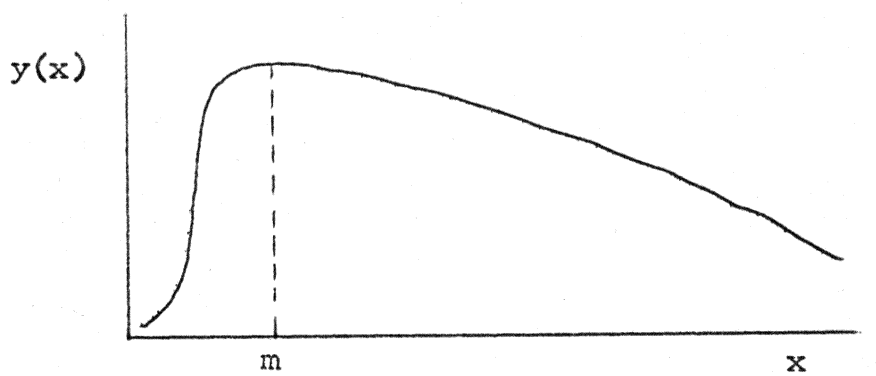


FIGURE 4-2. Example of a curve for which the K-W procedure is slow to converge.

Kesten[27] modified the K-W procedure to converge more swiftly in cases such as illustrated in Figure 4-2. Kesten's procedure shortens the step size only when a change in sign of the average slope is found. This procedure converges with probability 1 with the additional restrictions of $a_{n+1}<a_n$ for n=1,2,···,and for c_n a constant.

To illustrate these procedures, let a_n be the harmonic series 1, 1/2, 1/3, ···, 1/n,··· . The following table provides the step sizes for both procedures for a hypothetical search.

Note that after 8 pairs of trials, the K-W procedure has a step size reduced to 1/8 and may still be fairly far from m. The Kesten procedure has a step size of only 1/5 and will move back to m much quicker.

Table 4-3. Comparison of Kiefer-Wolfowitz and Kesten's convergence procedures.

Trial	1	2	3	4	5	6	7	8	Total movement
Sign of Δ_n	+	+	+	-	-	+	-	+	
K-W	1	$\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{4}$	$-\frac{1}{5}$	$\frac{1}{6}$	$-\frac{1}{7}$	$\frac{1}{8}$	$1\ \frac{149}{280}$
Kesten	1	1	1	$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{4}$	$\frac{1}{5}$	$2\ \frac{17}{60}$

One large drawback to both of these procedures is that there is no stopping rule for convergence on m. Kiefer

and Wolfowitz mention this as an unsolved problem.

More details on these procedures and others may be found in Wilde[34].

B. SEQUENTIAL ONE-FACTOR AT A TIME PROCEDURE

Friedman and Savage[19] described a method of locating an optimal response by the following steps:

(1) Order the factors in some manner. Presumably this would be done according to the investigator's intuition or prior knowledge regarding the effect of each factor on the response. The most important factors would be ordered and investigated first since less important factors may not show a significant effect until the important ones are near their optimal values.

(2) Use the best estimate of the optimal factor combination as an initial starting point.

(3) Vary the levels of the first factor until an approximate optimal response was reached. (Hotelling's methods could be used for this.)

(4) Using the factor combination for optimal response in (3), vary the second most important factor as in (3). Continue in this manner until all factors have been so investigated.

(5) Repeat this procedure starting at the factor combination of step (4).

(6) If the changes in the second cycle of this procedure showed a significant improvement, it would be advisable to proceed along the path defined by the two sets of local optima.

As one gets closer to the optimal response, Friedman and Savage suggest that the levels of the factors be made closer in order to better map the region of optimal response.

An illustration of this procedure is given in Figure 4-4 for two variables. Point P_1 is the original factor combination and x_2 is investigated first. P_3 is the factor

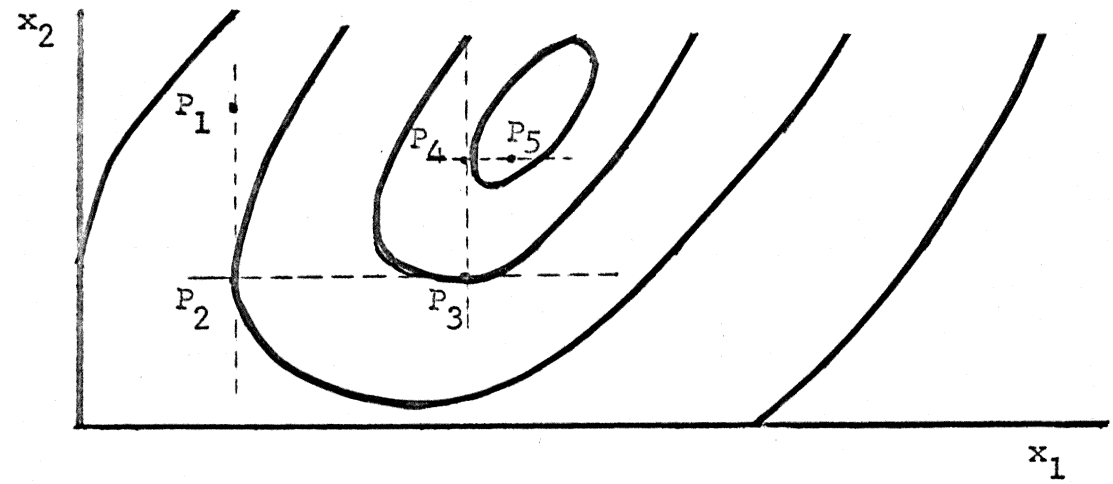


FIGURE 4-4 Illustration of Friedman and Savage Sequential One-Factor at a Time Procedure.

combination found by the end of the first round. The point P_5 is the optimal factor combination after the second round. Application of step (6) would lead the investigation along the path from P_3 to P_5 and obtain the optimal response in this manner if the points P_3 and P_5 were, in fact, located with sufficient accuracy.

If this method were applied without the application of step (6), the optimum would be approached very slowly when some mild interaction among the factors was present. Because of this feature, this procedure has fallen into disfavor.

C. METHOD OF STEEPEST ASCENT

1. Introduction

Historically the method of steepest ascent as an optimization technique seems to have been first suggested by Cauchy[14] in 1847. Box and Wilson[6] (1951) recommend this procedure within the experimental framework of response surface methodology. Another term for this procedure is the "optimal gradient" method.

The method of steepest ascent assumes that the response surface in some small subregion of the whole experimental space can be adequately represented by a hyperplane in the k factors. From a point P_o on this hyperplane we will proceed a distance R to a point P_1. P_1 will be located in such a manner so as to maximize the response on this hyperplane.

Let the response be estimated by the equation

$$y(x_1,x_2,\cdots,x_k) = b_o + \sum_{i=1}^{k} b_i x_i .$$

Let the point P_o be considered the origin of this k-dimensional space. At point P_1, R units from P_o, R^2 is given by

$$\sum_{i=1}^{k} x_i^2 = R^2 \qquad (4.1)$$

We wish to maximize y subject to the restriction of (4.1). Using a Lagrangian multipler $\frac{1}{2}\lambda$, maximize

$$\varphi(\underline{x},\lambda) = b_o + \sum_{i=1}^{k} b_i x_i - \tfrac{1}{2}\lambda\left(\sum_{i=1}^{k} x_i^2 - R^2\right)$$

Taking partials of φ with respect to x_i and equating to zero gives:

$$\frac{\partial \varphi}{\partial x_i} = b_i - \lambda x_i = 0 \qquad i = 1, 2, \cdots, k.$$

Hence

$$x_i = \frac{b_i}{\lambda} \qquad i = 1, 2, \cdots, k. \qquad (4.2)$$

Thus, P_1 is located along the path on which the coordinates are directly proportional to the first derivatives of the estimated plane.

The experimenter substitutes various values of λ into equations (4.2) and tries experimentally either to obtain a maximum response on that line or he will experiment along that line as far as he believes in the assumption of a local planar relationship for y. At that point he again will perform an experiment to estimate the response plane and will determine the new steepest ascent path. This procedure will be continued until the investigator feels he is "near" the

optimum response. At that point, he should perform an experiment to estimate a higher order (at least 2) response curve, perform a canonical analysis, draw contours, etc. for a more informative insight into the response relationship.

Figure 4-5 illustrates the steepest ascent method. The path of steepest ascent is perpendicular to the assumed local parallel contour lines.

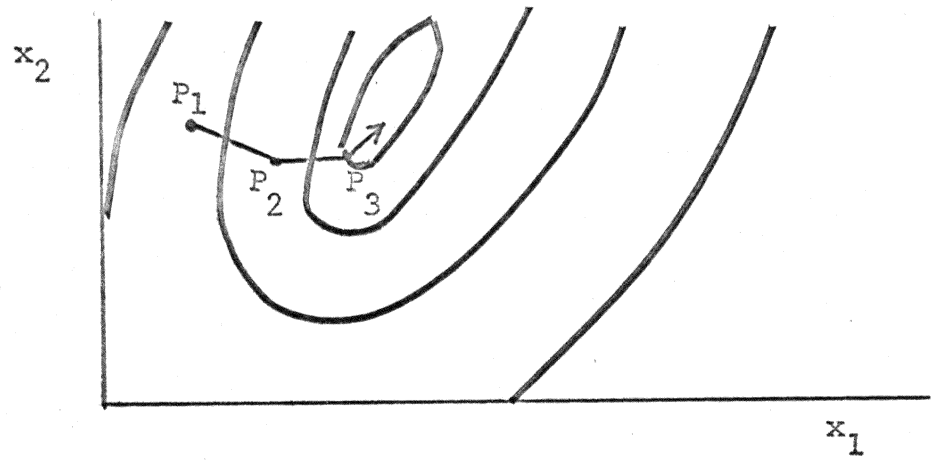


FIGURE 4-5. Illustration of Steepest Ascent Method.

2. Properties of the Method of Steepest Ascent

This procedure is not scale-invariant. Only after the relative scales for $x_1, x_2, \cdots, x_k$ have been predetermined does the concept of "distance" have any meaning. To illustrate, consider the following example due to Buehler, Shah and Kempthorne.(9)

Suppose $y(x_1,x_2)$ is a function that depends on variables x_1 and x_2 according to

$$y(x_1,x_2) = 190-(x_1-c_1)^2-(x_2-c_2)^2. \qquad (4.3)$$

Now y has a maximum of 190 at (C_1,C_2). If one makes measurements without error in the immediate neighborhood of $P_o = (x_1,x_2) = (C_1-9,C_2-1)$, y will be found to be:

$$y = 190 - (C_1-9-C_1)^2 - 9(C_2-1-C_2)^2 = 100$$

at P_o.

The linear approximation equation would be estimated by

$$\hat{y} = \left(\frac{\partial y}{\partial x_1}\right)_{P_o}\left[x_1-(C_1-9)\right] + \left(\frac{\partial y}{\partial x_2}\right)_{P_o}\left[x_2-(C_2-1)\right] + 100$$

Now

$$\left(\frac{\partial y}{\partial x_1}\right)_{P_o} = -2(x_1-C_1)\Big|_{x_1 = C_1-9} = 18$$

and

$$\left(\frac{\partial y}{\partial x_2}\right)_{P_o} = -18(x_2-C_2)\Big|_{x_2 = C_2-1} = 18$$

hence

$$\hat{y} = 18(x_1-C_1+9) + 18(x_2-C_2+1) + 100. \qquad (4.4)$$

Equation (4.4) is the approximating plane the experimenter would estimate in the neighborhood of $P_o = (C_1-9,C_2-1)$.

The method of steepest ascent would regard (C_1-9, C_2-1) as the new origin and would call for experimentation along the path whose coordinates were directly proportional to the first order coefficients. In this case the path would be the line :

$$\frac{x_1 - (C_1-9)}{18} = \frac{x_2 - (C_2-1)}{18}$$

or

$$x_1 - C_1 + 9 = x_2 - C_2 + 1. \qquad (4.5)$$

Now let us change the scale of measurement and see how this effects the path of steepest ascent. If x_1 were degrees centigrade, let 5^oC be a change of one unit of x_1'. For x_2, a pressure in pounds per square inch, let a unit of x_2' be equal to 10 pounds per square inch. The basic relationship (4.3) now becomes

$$y(x_1',x_2') = 190 - 25(x_1'-C_1')^2 - 900(x_2'-C_2')^2,$$

where $5C_1' = C_1$ and $10C_2' = C_2$. The point $P_o = (C_1-9, C_2-1)$ becomes $P_o' = (C_1'-9/5, C_2'-1/10)$.

The response at P_o' is

$$y = 190 - 25(-9/5)^2 - 900(-1/10)^2 = 100$$

as before. The partials of y evaluated at P_o' are

$$\left(\frac{\partial y}{\partial x_1'}\right)_{P_o'} = -50(x_1'-C_1')\Big|_{x_1' = C_1' - 9/5} = 90$$

and

$$\left(\frac{\partial y}{\partial x_2'}\right)_{P_o'} = -1800(x_2'-C_2')\Big|_{x_2' = C_2 - 1/10} = 180.$$

The approximating plane at $(C_1'-9/5, C_2'-1/10)$ is

$$\hat{y} = 90\left[x_1' - (C_1'-9/5)\right] + 180\left[x_2'-(C_2'-1/10)\right] + 100 .$$

The path of steepest ascent is now given by

$$\frac{x_1' - (C_1'-9/5)}{90} = \frac{x_2' - (C_2'-1/10)}{180} . \qquad (4.6)$$

Substituting $x_1' = x_1/5$, $x_2' = x_2/10$, $C_1' = C_1/5$, $C_2' = C_2/10$ into (4.6), this expression is reduced to

$$(x_1-C_1+9) = \tfrac{1}{4}(x_2-C_2+1) \qquad (4.7)$$

which differs from the first path (4.5) by the coefficient $\frac{1}{4}$. Thus, by just changing the scale of the factors a considerably different path was determined. For this reason the choice of scale is very important.

Several writers[6,16] suggest the following rule for choice of scales: the best units are those for which a unit change in one factor <u>at the optimum</u> gives the same change in response as a unit change in any other factor at the optimum. The effect of such an advantageous choice of scale is to make the response contours circles. From any point in the factor space, the gradient direction will then pass through the optimum.

Unfortunately one does not usually know what choice of units to use to obtain circular contours. Often some

estimated b_i may be small compared to experimental error. This might arise because of one of the following situations:

(1) The average level chosen for that factor is near a conditional maximum for the response as a function of that factor.

(2) The choice of units is too small, or

(3) The response system is independent of the factor.

In this case, the next determination of the path should use a slight variant in the design. For the factor in question the average level should be moved somewhat off the previous path of steepest ascent and the unit level increased. Then if the system is independent of this factor, the coefficient will again be small. If (1) or (2) above were true, the coefficient should be larger since a real effect should be discovered. This will be illustrated in the numerical example which is given in the next section.

Brooks and Mickey[8] investigated the problem of the number of trials needed to estimate the direction of steepest ascent. Since there is experimental error, the gradient direction is determined with some error. If one had more trials than the minimum of k+1 (a Simplex Design) required to estimate the first-order coefficients, the coefficients would be estimated with greater precision and the error in the gradient direction would thus be reduced.

If θ is the error in the estimated gradient direction

and a step of S units is to be taken, then $S \cdot \cos \theta$ is the component of this step in the direction of the true gradient. If $\cos \theta$ were near 1.0, the gradient was well estimated, if $\cos \theta$ were near zero, the gradient direction was not well estimated. Brooks and Mickey regarded each trial as a unit of effort. For N trials in a set, the improvement per unit of effort is $(S \cos \theta)/N$. They tried to choose N to maximize $E((S \cos \theta)/N)$ which is equivalent to maximizing $(1/N)E(\cos \theta)$ regarded as a function of N. They considered designs which yield normally distributed, uncorrelated estimates of the first-order coefficients with the same variance. They derived $E(\cos \theta)$ under these conditions (see (8) for details) and found that the maximum is achieved when $N=k+1$, the minimum number of trials required to estimate the coefficients. This result, it turns out, does not depend on σ^2, the error variance, and, hence, is a convenient result for applications. Hence, on a per observation basis, the maximum gain would be achieved by using a Simplex design for estimation of the gradient direction.

The method of steepest ascent is inherently self-defeating as one gets closer and closer to the optimum. The partial derivatives of y evaluated at points close to the optimum are near zero. For this reason, this method should be discontinued when the experimenter feels he has reached a near optimum region in the factor space. At that point,

he should utilize other designs to explore this region.[4]

3. Example of the Method of Steepest Ascent

To illustrate the method of steepest ascent consider the following example from Box and Wilson[6].

An investigation of a chemical process was made to maximize the yield of D for a given amount of material A in the chemical reaction of the type

$$A+B+C \rightarrow D+\text{other products}$$

This reaction utilized a solvent E. The yield presently obtainable was about 45% and the experimental error was of the order of 1%. Five factors were studied in a ½ fractional replication of a 2^5 design. The factors and their coded levels (-1 and +1) are as follows:

			-1	+1	(Units)
x_1	-	amount of solvent E	200	250	c.c.
x_2	-	proportion of C to A	4.0	4.5	mol./mol.
x_3	-	concentration of C	90	93	%
x_4	-	time of reaction	1	2	hours
x_5	-	proportion of B to A	3.0	3.5	mol./mol.

The analysis of the data provided the following estimating plane:

$$\hat{y} = 48.5+(7.9)x_1-(2.2)x_2+(6.0)x_3+(0.4)x_4+(0.4)x_5 .$$

The coordinates of the path are proportional to the

coefficients in terms of units of the design. A convenient method used to obtain the path coordinates is to define a "step" in one of the factors. The investigator decided that a step of 10 cc in x_1 would be a sufficient step. The unit for x_1 is 25 cc so, in terms of the design units, this is equivalent to a step of 10/25 or 0.4 units of x_1. Hence we set up the following equation to estimate λ:

$$\frac{7.9}{\lambda} = 0.4 ,$$

so $\lambda = 19.75$. This proportionality factor is now applied to all of the coefficients of the other factors which is then multiplied by the unit of the factor to obtain steps in terms of the original units. This gives

$$\Delta x_1 = 10 \text{c.c.}$$

$$\Delta x_2 = \frac{-2.2}{19.75}(.25) = -0.028 \text{ mol./mol.}$$

$$\Delta x_3 = \frac{6.0}{19.75}(1.5) = 0.456\%$$

$$\Delta x_4 = \frac{0.4}{19.75}(.5) = 0.011 \text{ hours}$$

$$\Delta x_5 = \frac{0.4}{19.75}(.25) = 0.005 \text{ mol./mol.}$$

The path is now obtained by successively adding the Δx_i's to the center point of the design in original units. The path is given by the following table.

Table 4-6. Path of Steepest Ascent.

	x_1	x_2	x_3	x_4	x_5
origin	225	4.25	91.5	1.5	3.25
Δx_i	10	-0.028	0.456	0.011	0.005
Path	235	4.22	92.0	1.51	3.25
	245	4.19	92.4	1.52	3.26
	255	4.17	92.9	1.53	3.26
	265	4.14	93.3	1.54	3.27
$\hat{y}$=80.0%	275	4.11	93.8	1.56	3.27
	285	4.08	94.2	1.57	3.28
$\hat{y}$=79.4%	295	4.06	94.7	1.57	3.28
	305	4.03	95.1	1.60	3.29

Additional experiments were performed at x_1=275cc and x_1=295cc with yields of 80.0% and 79.4%. Since this was near the expected optimum the investigator decided to run another experiment in the neighborhood of the point whose x_1 coordinate was 295cc.

Since the coefficients of x_4 and x_5 were small compared to their errors (error for all coefficients was $\pm$0.4) the path coordinates for x_4 and x_5 were not used and the unit for these factors was also increased. The unit change was decreased for x_1, x_2, and x_3 since the investigator felt he might be in a region where the response function has some large curvature. The next experiment was run at the following levels:

	-1	1	(Units)
x_1	280	310	cc
x_2	3.85	4.15	mol./mol.
x_3	94	96	%
x_4	2	4	hours
x_5	3.5	5.5	mol./mol.

A similiar 8 point, ½ fractional factorial of a 2^5 design was used. The resulting plane was given by

$$\hat{y} = 70.7-(2.8)x_1+(0.1)x_2-(2.3)x_3-(1.7)x_4-(0.4)x_5 .$$

Note that all of the coefficients changed sign except that of x_5 compared to the previous estimating plane. The new path of steepest ascent calculated as before is as follows:

	x_1	x_2	x_3	x_4	x_5	Yield
origin	295	4.0	95.0	3.0	4.5	
Path	285	4.0	94.5	2.6	4.4	80.8%
	275	4.0	93.9	2.2	4.3	
	265	4.0	93.4	1.8	4.2	84.0%
	255	4.0	92.8	1.4	4.1	81.5%

The point whose x_1=265cc gave the highest yield and can now be the center point for additional designs for further exploration of the response surface.

Further examples of this method can be found in Homme and Othmer[24] and Remmers and Dunn[31].

D. RANDOM EVOLUTIONARY OPERATION (REVOP)

Random Evolutionary Operation is a procedure that uses the Random Balance experimental methods developed by Satterthwaite[32]. The discussion that follows is from Lowe[30] in which he references a personal communication from Satterthwaite for this development.

REVOP claims to be an effective method of optimization when the number of variables is large and when the functional relationships among the variables is not known and anticipated to be complex. REVOP is essentially a non-parametric method as it uses the results on a ranking scale. The number of experiments required to find a set of optimum values appears to be independent of the number of variables and of the complexity of the functional relationship.

REVOP involves random choices of the levels of the variables under consideration. As such the data are unbalanced and are somewhat inefficient for evaluation of an assumed functional relationship. REVOP requires a starting point - the works process, a measurable response, measurable variables for study and a feed-back procedure for process optimization.

The REVOP procedure is as follows:

(1) A list is made of all variables acting on the response and the allowable range for each variable is set.

(2) The lower limit of each variable is coded as zero (0.0) and the upper limit as 10.0. Then if these limits are symmetrical about the works process, the starting point for each variable is coded as 5.0. The code 2.0, for example, represents a level of the variable at 20% of the range from its lower limit. The first experiment, E_o, is run at values 5.0 for each variable. A worksheet is set up for this procedure similar to Figure 4-7.

(3) To choose the levels for the next experiment, digits from 0 to 9 are selected at random for each variable. These random digits are placed on the line d_1 on the worksheet. The step increment $D_1=d_1^2/C$ is then completed for each variable, where C is chosen so that the maximum D_1 is about 2.0. This limits the change in level for the variables to a maximum change of 20% of its allowable range. The values D_1 for each variable are given a + or - sign at random.

(4) Step increments (D_1) are added to the previous levels and run so long as the response continues to improve. When no improvement is made, another set of random digits, d_2, is selected, $D_2=d_2^2/C$ is calculated (C may change), + and -'s assigned to D_2 and D_2 is added (algebraically) to the previous best set of conditions. This procedure is continued until an optimum is reached.

If the newly calculated variable levels show negative values, they are changed in sign to show the positive value.

Table 4-7. Example of a REVOP worksheet for a maximization.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	y
E_0	5·0	5·0	5·0	5·0	5·0	5·0	5·0	5·0	5·0	5·0	14·2
d_1	4	9	1	5	7	5	4	8	5	9	
$D_1 = d_1^2/40$	+0·4	−2·0	+0·0	−0·6	−1·2	+0·6	−0·4	−1·6	−0·6	−2·0	
$E_1 = E_0 + D_1$	5·4	3·0	5·0	4·4	3·8	5·6	4·6	3·4	4·4	3·0	22·3
$E_2 = E_1 + D_1$	5·8	1·0	5·0	3·8	2·6	6·2	4·2	1·8	3·8	1·0	16·0
d_2	0	5	5	4	5	5	5	0	4	3	
$D_2 = d_2^2/12$	−0·0	+2·1	−2·1	−1·3	−2·1	+2·1	−2·1	−0·0	+1·3	+0·7	
$E_3 = E_1 + D_2$	5·4	5·1	2·9	3·1	1·7	7·7	2·5	3·4	5·7	3·7	31·8
$A_4 = E_3 + D_3$	5·4	7·2	0·8	1·8	−0·4	9·8	0·4	3·4	7·0	4·4	
$E_4 =$	5·4	7·2	0·8	1·8	0·4	9·8	0·4	3·4	7·0	4·4	39·6
$A_5 = A_4 + D_3$	5·4	9·3	−1·3	0·5	−2·5	11·9	−1·7	3·4	8·3	5·1	
$E_5 =$	5·4	9·3	1·3	0·5	2·5	11·9	1·7	3·4	8·3	5·1	31·3
d_3	1	4	8	7	1	6	0	3	5	0	
$D_3 = d_3^2/30$	−0·0	+0·5	+2·1	+1·8	+0·0	−1·2	+0·0	+0·3	+0·8	−0·0	
$A_6 = A_4 + D_3$	5·4	7·7	2·9	3·6	−0·4	8·6	0·4	3·7	7·8	4·4	
$E_6 =$	5·4	7·7	2·9	3·6	0·4	8·6	0·4	3·7	7·8	4·4	30·8
$A_7 = A_6 - D_3$	5·4	6·7	−1·3	0·0	0·4	11·0	0·4	3·1	6·2	4·4	
$E_7 =$	5·4	6·7	1·3	0·0	0·4	11·0	0·4	3·1	6·2	4·4	38·0
d_4	3	8	9	7	6	7	4	9	5	1	
$D_4 = d_4^2/40$	−0·2	+1·6	+2·0	−1·2	−0·9	−1·2	−0·4	+2·0	−0·6	−0·0	
$A_8 = A_6 + D_4$	5·2	8·8	2·8	0·6	−1·3	8·6	0·0	5·4	6·4	4·4	
$E_8 =$	5·2	8·8	2·8	0·6	1·3	8·6	0·0	5·4	6·4	4·4	23·5
$A_9 = A_8 - D_4$	5·6	5·6	−1·2	3·0	0·5	11·0	0·8	1·4	7·6	4·4	
$E_9 =$	5·6	5·6	1·2	3·0	0·5	11·0	0·8	1·4	7·6	4·4	32·8
d_5	9	7	3	1	2	6	1	7	1	8	
$D_5 = d_5^2/40$	−2·0	−1·2	−0·2	−0·0	−0·1	−0·9	+0·0	+1·2	+0·0	+1·6	
$A_{10} = A_8 + D_5$	3·4	6·0	0·6	1·8	−0·5	9·1	0·4	4·6	7·0	6·0	
$E_{10} =$	3·4	6·0	0·6	1·8	0·5	9·1	0·4	4·6	7·0	6·0	55·1
$A_{11} = A_{10} + D_5$	1·4	4·8	0·4	1·8	0·4	8·2	0·4	5·8	7·0	7·6	
$E_{11} =$	1·4	4·8	0·4	1·8	0·4	8·2	0·4	5·8	7·0	7·6	71·6
$A_{12} = A_{11} + D_5$	−0·6	3·6	0·2	1·8	0·3	7·3	0·4	7·0	7·0	9·2	
$E_{12} =$	0·6	3·6	0·2	1·8	0·3	7·3	0·4	7·0	7·0	9·2	88·15
$A_{13} = A_{12} + D_5$	−2·6	2·4	0·0	1·8	0·2	6·4	0·4	8·2	7·0	10·8	
$E_{13} =$	2·6	2·4	0·0	1·8	0·2	6·4	0·4	8·2	7·0	10·8	82·3
d_6	1	1	7	4	2	6	9	3	8	1	
$D_6 = d_6^2/40$	−0·0	+0·0	−1·2	−0·4	−0·1	−0·9	+2·0	+0·2	+1·6	+0·0	
$A_{14} = A_{12} + D_6$	−0·6	3·6	−1·0	1·4	0·2	6·4	2·4	7·2	8·6	9·2	
$E_{14} =$	0·6	3·6	1·0	1·4	0·2	6·4	2·4	7·2	8·6	9·2	90·42
$A_{15} = A_{14} + D_6$	−0·6	3·6	−2·2	1·0	0·1	5·5	4·4	7·4	10·2	9·2	
$E_{15} =$	0·6	3·6	2·2	1·0	0·1	5·5	4·4	7·4	10·2	9·2	84·5
d_7	4	3	3	6	1	2	8	8	5	9	
$D_7 = d_7^2/40$	−0·4	−0·2	+0·2	−0·9	+0·0	+0·1	−1·6	+1·6	−0·6	−2·0	
$A_{16} = A_{14} + D_7$	−1·0	3·4	−0·8	0·5	0·2	6·5	0·8	8·8	9·2	7·2	
$E_{16} =$	1·0	3·4	0·8	0·5	0·2	6·5	0·8	8·8	9·2	7·2	82·6
$A_{17} = A_{16} - D_7$	−0·2	3·8	−1·2	2·3	0·2	6·3	4·0	5·6	8·0	11·2	
$E_{17} =$	0·2	3·8	1·2	2·3	0·2	6·3	4·0	5·6	8·0	11·2	92·58

If the level is greater than 10.0, and thereby outside the original range of experimentation, a decision must be made whether to permit the calculated value or to restrict it to a value of 10.0. This extension of range could also be allowed for negative values.

The advantages of REVOP are that any number of variables may be handled and the calculations are extremely simple.

REVOP's disadvantages lie in the unbalanced nature of experimentation in which no separation of the variable effects is possible. In addition, the calculated step interval D_i may not be measurable on the process control instrumentation.

E. ROTATING-SQUARE EVOLUTIONARY OPERATION (ROVOP)

Lowe[30] discusses the method of evolutionary operation entitled ROVOP (abbreviation for Rotating Square Evolutionary Operation). It is designed to eliminate uncertainty because of the range of the variables involved in the experimentation. Like EVOP the ROVOP design is simple for two or three variables but gets very complicated for situations with more variables.

ROVOP starts, for two variables, with the basic 2^2 plus center point design, where the design matrix is coded by the usual 0 and ± 1 notation for the levels of the factors.

After this cycle of 5 runs is completed, the coded design matrix is multiplied by the factor of $\sqrt{2}$ and then the design matrix is rotated through an angle of 45°. This defines the new treatment levels. Experimentation at the four <u>new</u> treatment levels and the center point are then run for the second cycle. For the next cycle and each succeeding cycle of 5 runs, the previous cycle's coded (0 and ±1's) design matrix is multiplied by $\sqrt{2}$ and then rotated through an angle of 45°. Figure 4-8 below illustrates this procedure for two variables.

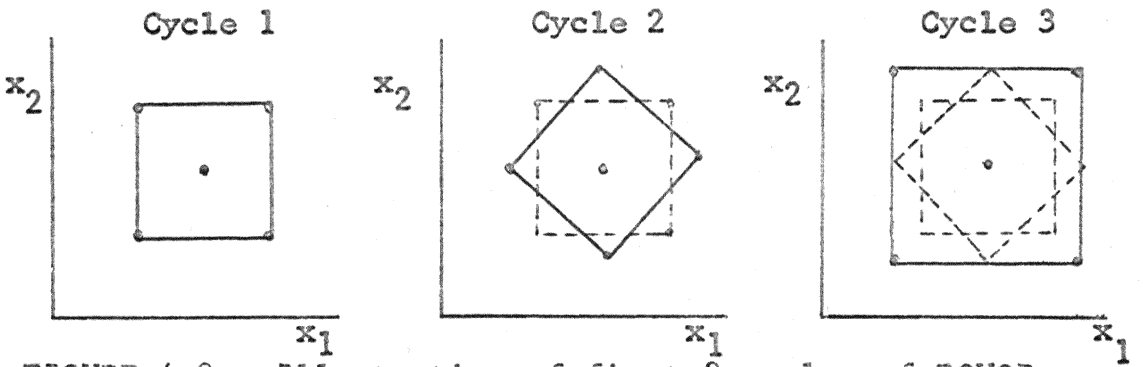


FIGURE 4-8. Illustration of first 3 cycles of ROVOP.

After each cycle (beyond the first) the response is analysed by multiple regression techniques for the second order model. The mean squares due to the regression is calculated and compared to the residual mean squares. When the regression is significant at some predetermined significance level, the response surface can no longer be considered to be constant over the experimental region. In such a case

the decision is made to move experimentation in the direction of optimization.

The stationary point of the quadratic system, $\underline{x}_o$, would be found by solving the equations of first-order partial derivatives set equal to zero; see equation (1.5). A new phase of experimentation would be initiated by moving in the direction of optimum response. Movement is not made by going to the center of the quadratic system. Instead ROVOP movement is made conservatively by requiring that the new phase original design matrix have at least 2 of its design points on or within the experimental region of the cycles of the just completed phase. The range of the variables should be reduced by at least a factor of $\sqrt{2}$ over the last cycle of the previous phase. (This would not be necessary for variables which were not significant.) The design matrix for this next phase would be recoded to the 0, ± 1 notation and the analysis would also start over again. Examples of movement are shown in Figure 4-9.

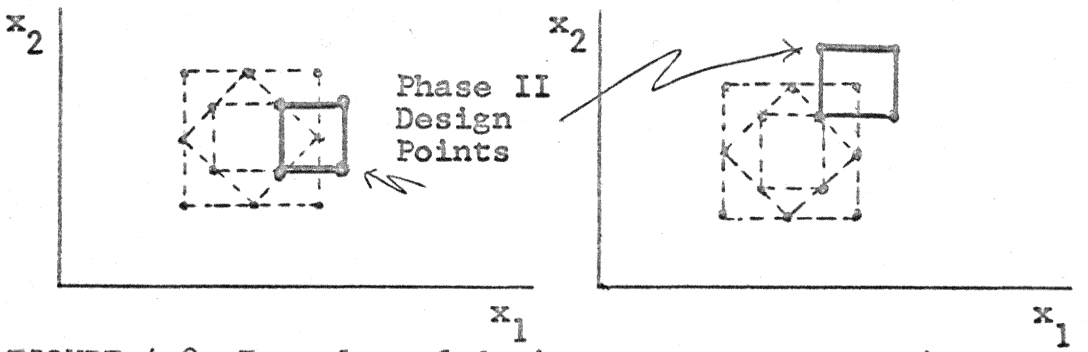


FIGURE 4-9. Examples of design center movement in two variable ROVOP.

This method of optimization would gradually approach and straddle the optimum. At this point, the stationary point of the fitted response system would be located well within the experimental region and other designs should be used to further explore the region. Continuation of ROVOP beyond this point will provide adjustment for long-term change only, such as raw material changes, and may not be worthwhile.

For a k variable ROVOP, the basic design is a 2^k plus center point design and the same procedure is followed as in the two variable (k=2) case. For k=3, successive cycles are expanding cubes.

Since the analysis of ROVOP data is not straightforward, a computer would usually be required for fast feedback to production personnel. Response contours could also be plotted by the computer for better understanding of the estimated response surface.

If several phases of ROVOP were required before optimum conditions were reached, a wide area of the variables would have been explored and an overall response surface could be fitted to all of the data. However, care must be used in so doing as the second-order model may not be too good an approximation to the true underlying functional relationship if the area explored is too wide. In addition, if the process in not stable in time, older data may no

longer represent the current response surface and so older data may have to be deleted.

In summary, the advantages of ROVOP are:

(1) Too small an initial range of the variables is not critical.

(2) The rotating pattern of experimentation provides estimation of quadratic terms very quickly.

(3) Procedure provides a good exploration of the experimental region.

There are several disadvantages of ROVOP. They are:

(1) Analysis of data requires a computer.

(2) Only quantitive variables can be used.

(3) For more than three variables, the designs require many points (2^k+1) for each cycle before a new analysis is made.

(4) ROVOP is not a technique for use on a permanent basis.

This last disadvantage is of such importance that it perhaps should not be included as an evolutionary operation technique. However ROVOP certainly has its use in experimental situations or in some EVOP situations where too much conservative action is evident on the EVOP committee with regard to range of the variables under study. However, as an EVOP technique it perhaps changes the levels of the factors too drastically for a production process.

F. COMPARISON OF SOME OPTIMIZATION TECHNIQUES

1. Introduction

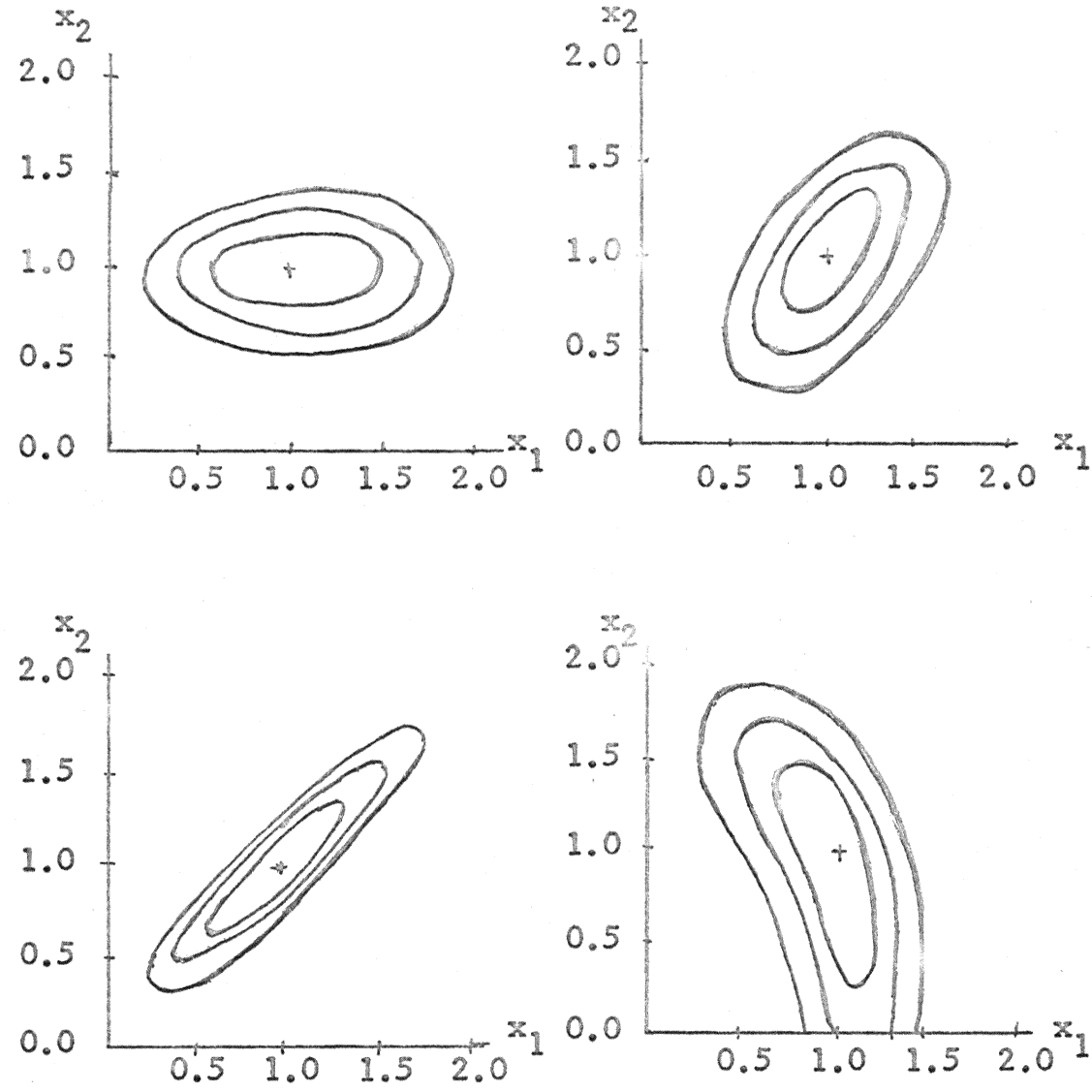
Brooks[7], in his doctoral dissertation, compared several of these optimization techniques for the case of two variables (k=2). He used four types of response surfaces each of which had an expected maximum yield of 1.0 at the point (1.0, 1.0). Three contours, 0.25, 0.50, and 0.75, of these four response surfaces, are shown in Figure 4-10.

On Surface 1, as one can see, the two factors are independent. Surface 2 is Surface 1 rotated through an angle of about 37 degrees. Surface 3 contains a sharp ridge and has large areas of low flat response. Surface 4 contains a curvilinear ridge.

For each response surface, the entire experimental region of the 2 unit square was divided up into nine overlapping subregions. Each of these subregions was a unit square (1x1) and was located independently and randomly within the 2x2 square in a stratified pattern.

Experimentation was conducted in each subregion with N=16 and N=30. The experimental error was assumed to be normally distributed with a mean of zero and a standard deviation of 0.03.

The measure of effectiveness used to compare the maximum seeking methods was the average maximum response



Surface 1 $y=(\frac{1}{2}+\frac{1}{2}x_1)^4 x_2^4 \exp\left[2-(\frac{1}{2}+\frac{1}{2}x_1)^4-x_2^4\right]+\epsilon$

Surface 2 $y=(0.3+0.4x_1+0.3x_2)^4(0.8-0.6x_1+0.8x_2)^4\exp\left[2-(0.3+0.4x_1+0.3x_2)^4-(0.8-0.6x_1+0.8x_2)^4\right]+\epsilon$

Surface 3 $y=x_1^2\exp\left[1-x_1^2-20.25(x_1-x_2)^2\right]+\epsilon$

Surface 4 $y=(0.3x_1^2+0.7x_2^2)^3\exp\left[1-0.6(x_1-x_2)^2-(0.3x_1+0.7x_2)^3\right]+\epsilon$

FIGURE 4-10. Response surfaces used in comparison study. Contours are 0.25, 0.50, and 0.75 responses.

observed in the nine subregions for each response surface.

Brookes considered the following methods of maximation: factorial experiments, sequential-one-factor at a time procedure, method of steepest ascent, and a random method. Spendly, Hext, and Himsworth[33] extended this procedure to the simplex design method of optimization.

2. Factorial Experiments

For the case of N=16, a 4x4 factorial design was used. When N=30, both a 5x6 and a 6x5 design were used to compare the effect of orientation of the design. The results are given in the following table.

Table 4-11. Achievement averages for factorial experiments.

Response Surface	Factorial Design		
	4x4	5x6	6x5
1	0.9355	0.9706	0.9585
2	0.9705	0.9810	0.9705
3	0.9196	0.9542	0.9191
4	0.9691	0.9727	0.9705
Average	0.9487	0.9696	0.9547

A local quadratic surface was also fitted by using 9 responses in the neighborhood of the highest observed response. For N=16, this average estimated maximum was 0.9541. For N=30, this maximum was 0.9602.

3. Sequential One-factor at a Time Procedure

In each subregion for the N=16 case, a set of four equally spaced observations was made at the middle of one of the factors, x_1 say, and a cubic equation was estimated to fit the four responses. The value of x_2, x_2^*, which maximized $\hat{y}$ was found in the usual fashion; i. e., by differentiating this cubic equation with respect to x_2, equating to zero, and solving. Four equally spaced observations were then made holding $x_2 = x_2^*$ and another cubic equation was fitted and x_1^* was found which maximized $\hat{y}$. This completed the first round. The second round then started with the step interval reduced. This procedure is illustrated in Figure 4-12.

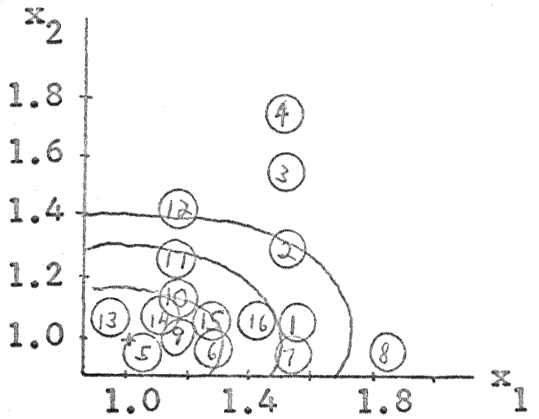


FIGURE 4-12. Illustration of Sequential One-factor at a Time Procedure for N=16.

For N=30, there were three rounds with two sets of 5 trials in each round.

In both cases, N=16 and N=30, the procedure was tried both with x_1 and x_2 as the first variable. The results of this procedure are given in Table 4-13. Note that the N=30 results are only slightly better than the N=16 results.

Table 4-13. Results of the Sequential One-factor at a Time Procedure.

First Variate	Surface	N=16	N=30
x_1	1	0.9831	0.9898
	2	0.9822	0.9899
	3	0.8978	0.9290
	4	0.9779	0.9799
x_2	1	0.9629	0.9875
	2	0.9641	0.9826
	3	0.9297	0.9512
	4	0.9732	0.9704
	Average	0.9587	0.9726

4. Method of Steepest Ascent Procedure

Initially a 2x2 design about the midpoint of the subregion is used to estimate the gradient direction. Trials are then made in the gradient direction until the response obtained is lower than the previous one or the border of the subregion is obtained. At that point, another 2x2 design is used and a new gradient direction is found. Trials are made in the new gradient direction (with a smaller step size) and this procedure is continued until the number of tirals (16 or 30) are all completed. Figure 4-14 illustrates this procedure for N=16.

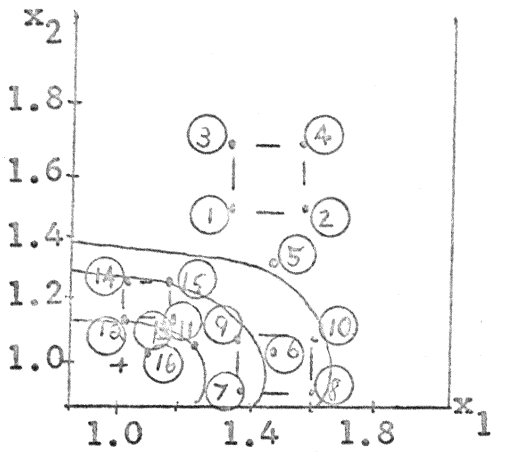


FIGURE 4-14. Illustration of the method of steepest ascent for N=16.

The highest observed response was taken as the maximum with the results as given in the following table.

Table 4-15. Results of Method of Steepest Ascent.

Surface	N=16	N=30
1	0.9564	0.9781
2	0.9640	0.9758
3	0.9381	0.9622
4	0.9369	0.9763
Average	0.9489	0.9731

5. Random Method

Each subregion was explored by a simple random sampling method and by a stratified procedure. The stratified procedure subdivided the region into 16 or 30 equal sections and a trial was made at random within this small plot. The results of this procedure are given in Table 4-16. As can

be seen there was little difference between the simple random and stratified random procedures for N=30.

Table 4-16. Results of the Random method of Sampling.

Surface	Simple Random		Stratified Random	
	N=16	N=30	N=16	N=30
1	0.8952	0.9476	0.9162	0.9708
2	0.8662	0.9338	0.9151	0.9685
3	0.9002	0.9369	0.8910	0.9062
4	0.8788	0.9477	0.9749	0.9738
Average	0.8851	0.9415	0.9243	0.9548

6. Simplex Design Procedure

For this procedure the size of the simplex design (an equilateral triangle) was chosen so that any point in the experimental subregion could just be reached within the allotted number of trials. The same response surfaces and experimental subregions as previously discussed were used. Spendly, Hext, and Himsworth[33], however, do not give their results broken down by surface type. The overall average maximum for 16 trials was 0.9442 and for N=30, 0.9191 was achieved. Some of their runs (about 5%) were subject to the influence of random errors and were not able to free themselves from the effects of these errors within the allotted number of trials. The average achievement of those 95% of the runs which did not get "bogged down" was 0.9711 and 0.9716 for N=16 and 30, respectively.

7. Comparison Among Procedures

Table 4-17 summarizes these results.

Table 4-17. Summary of Results of Optimum Seeking Methods.

Method	N=16	N=30
Factorial	0.9487	0.9622
Sequential One-factor	0.9587	0.9726
Steepest Ascent	0.9489	0.9731
Simple Random	0.8851	0.9415
Stratified Random	0.9243	0.9548
Simplex	0.9442	0.9191
Simplex (95% of runs)	(0.9711)	(0.9716)

Each of these tabulated averages is the average of the highest observation of 36 experiments. Since each experiment was only run once, these results cannot be considered true Monte Carlo results but are perhaps indicative of the performance of these methods.

Of the non-sequential methods, which were the Factorial and Sampling procedures, the Factorial provides the better estimate of the maximum. However it should be pointed out that the factorial design was applied to an area in which the experimenter knew the true maximum was located and he also had well defined limits on his variables.

For the sequential methods with N=30 there seem to be no practical differences among the methods unless the simplex procedure had some dominating random errors. For N=16, the simplex procedure without dominating errors is

clearly best with the one-factor at a time procedure second best.

Buehler, Shah, and Kempthorne[9] compared the method of steepest ascent to the one-factor at a time procedure and concluded that the average performance of the latter method to the steepest ascent method is inferior over any choice of scale in the steepest ascent as long as there are some mild interactions among the factors. This is an interesting result since the one factor at a time procedure is scale invariant and the method of steepest ascent is not.

These same authors[10,11,12] have proposed the "method of parallel tangents" and they compare this procedure with steepest ascent and some other procedures. The interested reader can pursue this topic in the references given above.

V ACKNOWLEDGEMENTS

The author wishes to acknowledge Dr. Raymond H. Myers, thesis advisor, for his help and guidance in writing this thesis. Thanks are also extended to Dr. Klaus Hinkelmann for his suggestions and remarks.

Much appreciation is extended to the author's employer, the U. S. Naval Avionics Facility, Indianapolis, Indiana, who made this year of graduate work possible by providing financial support to him.

Appreciation is also given to Mrs. Claude Boyd Loadholt, who typed the final manuscript.

VI REFERENCES

1. Anderson, R. L., 1953. Recent advances in finding best operating conditions. Journal American Statistical Association, 48, 789.

2. Box, G. E. P., 1952. Multifactor designs of first order. Biometrika, 39, 49.

3. Box, G. E. P. 1957. Evolutionary Operation: A method for increasing industrial productivity. Applied Statistics, VI, No. 2, 81.

4. Box, G. E. P. and Hunter, J. S., 1957. Multifactor experimental designs for exploring response surfaces. Annals of Mathematical Statistics, 28, 195.

5. Box, G. E. P. and Hunter, J. S., 1959. Condensed calculations for evolutionary operation. Technometrics, 1, No. 1, 77.

6. Box, G. E. P. and Wilson, K. B., 1951. On the experimental attainment of optimum conditions. J. Royal Statistical Society, Series B, 13, pp. 1-45.

7. Brooks, S. H., 1959. A comparison of maximum-seeking methods. Operations Research, 7, 430.

8. Brooks, S. H. and Mickey, M. R., 1961. Optimum estimation of gradient direction in steepest ascent experiments. Biometrics, 17, 48.

9. Buehler, R. J., Shah, B. V., and Kempthorne, O., 1961. Some properties of steepest ascent and related procedures for finding optimum conditions. Iowa State Statistical Lab. ONR contract Technical Report No. 1.

10. Buehler, R. J., Shah, B. V., and Kempthorne, O., 1961. The method of parallel tangents (PARTAN) for finding an optimum. Iowa State Statistical Lab. ONP contract Technical Report No. 2.

11. Buehler, R. J., Shah, B. V., and Kempthorne, O., 1961. Some further properties of the method of parallel tangents and conjugate gradients. Iowa State Statistical Lab. ONR contract Technical Report No. 3.

12. Buehler, R. J., Shah, B. V., and Kempthorne, O., 1964. Methods of parallel tangents. Chemical Engineering Progress Symposium Series, Vol 60, No. 50.

13. Bunde, T. A., 1959. The application of random balance designs. Technometrics, 1, No. 2., 139.

14. Cauchy, A., 1847. Méthode générale pour la résolution des systémes d'equations simutlanées. C. R. Acad. Sci. Paris 25, 536-538.

15. Carpenter, B. H., and Sweeny, H. C., 1965. Process improvement with "Simplex" self-directing evolutionary operation. Chemical Engineering, 72, July 5, 1965, 117.

16. Davies, O. L., Editor. Design and Analysis of Industrial Experiments, Hafner Publishing Co., New York, N. Y., (1956).

17. Draper, N. R., 1963. "Ridge Analysis" of response surfaces. Technometrics, 5, No. 4., 469.

18. Frankel, S. A., 1961. Statistical design of experiments for process development of MBT. Rubber Age, 89, 453.

19. Friedman, M., and Savage, I. J., 1947. Planning experiments seeking maxima. Chapter 13 of Techniques of Statistical Analysis, edited by Eisenhart, Hastay, and Wallis. McGraw-Hill Book Co. New York, N. Y. (1947).

20. Hardy, G. H., Littlewood, J. E., and Polya, G., Inequalitites, Cambridge University Press, (1952).

21. Hicks, C. R., Fundamental Concepts in the Design of Experiments. Holt, Rinehart and Winston, New York, (1964).

22. Hill, W. J. and Hunter, W. G., 1966. A review of response surface methodology: a literature survey. Technometrics, 8, No. 4., 571.

23. Hoerl, A. E., 1959. Optimum solution of many variables equations. Chemical Engineering Progress, 55, No. 11., 69.

24. Homme, A. C. and Othmer, D. F., 1961. Sulphuric acid-optimized condition in catalyst manufacturing. Industrial Engineering Chemistry, 53, 979.

25. Hotelling, H., 1941. Experimental determination of the maximum of a function. Annals of Mathematical Statistics, 12, 20.

26. Hunter, W. G. and Kittrell, J. R., 1965. Evolutionary operation: a review. Department of Statistics, University of Wisconsin, Technical Report No. 55, December 1965.

27. Kaplan, W. Advanced Calculus. Addison-Wesley Publishing Co., Reading, Mass., (1952).

28. Keston, H., 1958. Accelerated stochastic approximation. Annals of Mathematical Statistics, 29, 41.

29. Kiefer, J. and Wolfowitz, J., 1952. Stochastic estimation of the maximum of a regression function. Annals of Mathematical Statistics, 22, 400.

30. Lowe, C. W., 1964. Some techniques of evolutionary operation. Trans. Instn. Chem. Engrs., 42,T334.

31. Remmers, E. G., and Dunn, G. G., 1961. Process improvement of a fermentation product. Indus. Engineering Chemistry, 53, 743.

32. Satterthwaite, F. E., 1959. Random balance experimentation. Technometrics, 1, No. 2., 11.

33. Spendley, W., Hext, G. R. and Himsworth, F. R., 1962. Sequential application of simplex designs in optimizing and evolutionary operation. Technometrics, 4, No. 4., 441.

34. Wilde, D. J. Optimum Seeking Methods. Prentice-Hall, Inc., Englewood Cliffs, N. J. (1964).

The vita has been removed from the scanned document

SOME OPTIMIZATION PROCEDURES USED IN RESPONSE SURFACE METHODOLOGY

by

Bruce R. Glenn

ABSTRACT

This thesis is a literature survey into selected optimization topics of response surface methodology. In a typical response surface problem one of the main problems to be solved is to find those levels of the controllable variables to provide an optimum response such as highest yield or lowest cost. Several methods for attaining this optimum response are discussed, such as: the method of steepest ascent, sequential one-factor at a time procedure, a random balance method also known as random evolutionary operation (REVOP), rotating square evolutionary operation (ROVOP), and some one-factor optimization techniques. In addition two methods for optimizing an existing production process are discussed: evolutionary operation (EVOP) and simplex EVOP.

A technique useful in the interpretation of a complex response surface is also discussed. This method is called "Ridge Analysis" and the development of this technique is presented and a worked example is given.