

## Generalized ridge analysis under linear restrictions, with particular applications to mixture problems

Norman R. Draper, Friedrich Pukelsheim

### Angaben zur Veröffentlichung / Publication details:

Draper, Norman R., and Friedrich Pukelsheim. 2002. "Generalized ridge analysis under linear restrictions, with particular applications to mixture problems." *Technometrics* 44 (3): 250–59. <https://doi.org/10.1198/004017002188618437>.

### Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:

**Deutsches Urheberrecht**

Weitere Informationen finden Sie unter: / For more information see:

<https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/>



# GENERALIZED RIDGE ANALYSIS UNDER LINEAR RESTRICTIONS, WITH PARTICULAR APPLICATIONS TO MIXTURE EXPERIMENT PROBLEMS

Norman R. Draper

Statistics Department

University of Wisconsin

Madison, WI 53706-1685

Friedrich Pukelsheim

Institut für Mathematik

Universität Augsburg

D-86135 Augsburg, Germany

Suppose we wish to investigate the behavior of a response  $y$  over a specified region of interest by fitting a second-order response surface. Standard ridge analysis provides a way of following the locus of, for example, a maximum response, moving outwards from the origin of the predictor variable space. Because this approach does not require one to view the fitted regression surface as a whole, this important technique may be applied even when visualization of the surface is difficult in several dimensions. The viewing of a ridge trace enables practitioners to assess and understand the typically complex interplay between the input variables as the response improves. Sometimes, we may wish to explore a subspace defined by a linear restriction on the predictors. This has been discussed only infrequently in the literature and never in the context of mixture experiments. Here we show how a modification of ridge regression can be used in a very general way to investigate second-order mixture surfaces with many ingredients, particularly when the experimental mixture space is itself limited by further linear equalities in addition to the mixture requirement. In some cases, it is not necessary to move the ridge origin into the mixture space to achieve the desired results, and *any* form of the second-order fitted model, whether of Scheffé type, Kronecker type, or something in between, can be accommodated.

Key Words: Kronecker model; Mixture model; Projections; Response surfaces; Restrictions on mixture spaces; Ridge analysis; Scheffé model; Second-order model.

## 1 Introduction

Ridge analysis was first introduced in the context of general response surface methodology by A.E. Hoerl (1959, 1962, 1964). It was further investigated by Draper (1963), who proved

results that Hoerl had suggested without proof, and was then extended by Myers and Carter (1973) for the so called “Dual Response” (DR) problem. Related work has been done by Del Castillo, Fan and Semple (1997, 1999) and by Semple (1997). For a wide-ranging discussion, see R. W. Hoerl (1985).

Only one application of ridge analysis to mixture problems has appeared. Typically,  $q$  nonnegative fractional ingredients  $x_1, x_2, \dots, x_q$  must satisfy the mixture restriction

$$x_1 + x_2 + \dots + x_q = 1,$$

(or some linear restriction that can essentially be reduced to that form). In that one application by R.W. Hoerl (1987), ridge analysis was applied by first invoking a transformation that moved from the  $q$ -dimensional origin  $(0, 0, \dots, 0)$  to the centroid  $(\frac{1}{q}, \frac{1}{q}, \dots, \frac{1}{q})$  of the  $(q - 1)$ -dimensional mixture space. Our ridge paths could also begin from such a centroid but we shall see that ridge analysis can proceed in a mixture space around *any* point without preliminary transformation. We shall see further that additional linear equalities in the mixture ingredients are easily incorporated into a very general method which leads to great flexibility in applying ridge analysis techniques to mixture problems. When linear inequalities are also involved, we can examine the ridge traces and easily determine if the ridges pass into and/or out of the regions defined by the inequalities, by checking the coordinate values of the  $x$ ’s on the paths. When any ingredient value becomes negative, or exceeds the applicable inequalities, a path has gone outside the region and is then of no interest unless it returns.

## 2 A Motivating Experiment

The pharmaceutical mixture example of Anik and Sukumar (1981) is an excellent example of a mixture problem which entails additional linear equalities and inequalities on the mixture ingredients and which thus might profit from this ridge analysis. This work was also motivated by our desire to simplify the application of ridge analysis to mixture problems. Thanks to the help of the reviewers, the method is now very general in its application.

The Anik and Sukumar (1981) paper involved a study of five ingredients, one of which,  $x_5$ , was held constant at 0.10 (10% of the mixture) so that the remaining ingredients,

$x_1, x_2, x_3$ , and  $x_4$  were constrained by the requirement that

$$x_1 + x_2 + x_3 + x_4 = 0.9. \quad (2.1)$$

As a reviewer commented, Eq. (2.1) could be renormalized via  $x_i = 0.9 u_i$ , so that  $u_1 + u_2 + u_3 + u_4 = 1$ . We do not do this, because it introduces a step that is not needed and would have to be undone in later calculations. (However such a renormalization is usually needed when constructing diagrams as we shall see later in Section 4.2.) The essence of our method is that the ridge paths are obtained directly (and more easily so) without any such additional steps.

Anik and Sukumar wanted to examine various combinations of the four ingredients, to fit a quadratic model to a response variable  $y$ , solubility, and to seek the maximum response. Each of the four ingredients was restricted to a range within  $[0, 1]$  as shown in Table 1. Hence the authors decided to use an experimental design based on the “extreme vertices” of the restricted region (see Table 1). This excellent method was first suggested by McLean and Anderson (1966), and Anik and Sukumar (1981) aimed to show how useful the method can be. To implement it, one generates the extreme points (or “corners”) of the region and then selects the design points from (a) vertices, (b) edge (one-dimensional) centroids, (c) face (two-dimensional) centroids, and so on. The last of these groups is the single point represented by the overall centroid, calculated by averaging all the vertices. The method has various subtleties (which we do not describe; see McLean and Anderson, 1966 or Cornell, 1990) due to the fact that the number of extreme vertices (and consequently of the various centroids) depends on the specific ranges of the  $x$ ’s, which determine the consequent region shape. Suffice it to say that Anik and Sukumar (1981) were led to use the specific 14 point experimental design shown in Table 1.

The experimental design of Anik and Sukumar requires additional explanation to avoid potential confusion. Figure 1(a) shows the triangular subspace  $x_3 = 0$ ; within it, the other restrictions create the five-sided figure. The inner triangle and the pentagon in Figure 1(b) play the same respective roles for the  $x_3 = 0.08$  subspace. The outer triangle of Figure 1(b) is the same triangle as in Figure 1(a), and makes the point that the  $x_3 = 0.08$  slice of the four dimensional simplex is smaller than the  $x_3 = 0$  slice. Figure 1(c) shows the two slices superimposed as they would be seen in a birds-eye view from the  $x_3 = 0.90$  vertex. We further

Table 1. Experimental Design Used by Anik and Sukumar (1981) Together With the Lower and Upper Limits Which Define the Mixture Space of Interest, and the Response Data Obtained From the Experiment

Point no.	$x_1$	$x_2$	$x_3$	$x_4$	$y$
Lower limit	0.10	0.10	0	0.30	—
Upper limit	0.40	0.40	0.08	0.70	—
Vertices					
1	0.10	0.10	0	0.70	3.0
2	0.10	0.10	0.08	0.62	7.3
3	0.15	0.40	0	0.35	4.9
4	0.11	0.40	0.08	0.31	8.4
5	0.40	0.15	0	0.35	8.6
6	0.40	0.11	0.08	0.31	12.7
Edge centroids (averages of indicated vertices)					
7 (1,2)	0.10	0.10	0.04	0.66	5.1
8 (5,6)	0.40	0.13	0.04	0.33	10.8
9 (3,4)	0.13	0.40	0.04	0.33	6.6
10 (1,3,5)	0.216	0.216	0	0.468	4.4
11 (2,4,6)	0.203	0.203	0.08	0.414	7.9
12 (4,6)	0.255	0.255	0.08	0.31	9.4
13 (3,5)	0.275	0.275	0	0.35	5.8
Overall centroid					
14	0.21	0.21	0.04	0.44	6.3

$x_1$  = Polyethylene glycol 400

$x_2$  = Glycerine

$x_3$  = Polysorbate 60

$x_4$  = Water

$y$  = Solubility, mg/ml

Note that  $x_1 + x_2 + x_3 + x_4 = 0.90$  for each point.

note that, for each pentagon, two pairs of vertices are quite close together. Because of this, Anik and Sukumar (1981, p. 898) averaged these close pairs of points, and called the resulting averages “vertices” of their region. Thus, in Table 1, “vertex 3,” (0.15, 0.40, 0, 0.35) is the average of true vertices (0.10, 0.40, 0, 0.40) and (0.20, 0.40, 0, 0.30); “vertex 4,” (0.11, 0.40, 0.08, 0.31) is the average of true vertices (0.10, 0.40, 0.08, 0.32) and (0.12, 0.40, 0.08, 0.30); “vertex 5,” (0.40, 0.15, 0, 0.35) is the average of true vertices (0.40, 0.10, 0, 0.40) and (0.40, 0.20, 0, 0.30); and “vertex 6,” (0.40, 0.11, 0.08, 0.31) is the average of true vertices (0.40, 0.10, 0.08, 0.32) and (0.40, 0.12, 0.08, 0.30).

We shall revisit this example to illustrate how ridge analysis can be applied to mixture experiments with regions restricted by linear equalities and inequalities in the ingredients. We follow the approach of the original authors in fitting a second-order (or quadratic) Scheffé model

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \epsilon \quad (2.2)$$

via least squares using the data in Table 1; see Scheffé (1958, 1963). A discussion of the various equivalent second-order model forms that can be fitted in a mixture problem is given by Prescott, Dean, Draper and Lewis (2000). For purposes of interpretation, it does not matter *which* of the several alternative possible models is fitted, because the resulting response contours will be identical in every case. The ridge paths are exactly the same if other choices are made; in fact, Anik and Sukumar (1981) fitted a model with a constant term in it, one of several possibilities. The equation resulting from fitting (2.2) by least squares is

$$\begin{aligned} \hat{y} = & 49.716x_1 + 8.414x_2 + 29.95x_3 + 4.3365x_4 \\ & -58.671x_1x_2 - 27.83x_1x_3 - 74.902x_1x_4 \\ & +10.20x_2x_3 + 33.81x_3x_4. \end{aligned} \quad (2.3)$$

Note that the nonlinear blending term  $x_2x_4$  is missing in (2.3). When the Scheffé model is used with the design of Table 1, the resulting  $\mathbf{X}'\mathbf{X}$  matrix is singular. Regression of the  $x_2x_4$  column onto the remaining  $\mathbf{X}$ -columns produces an exact fit on the columns  $x_1, x_2, x_1x_3, x_1x_4$  and  $x_2x_3$ . After rearrangement of terms and factorization, the exact fit equation can be written as

$$(x_1 - x_2)(x_3 + 2x_4 - 0.7) = 0. \quad (2.4)$$

For every data point in Table 1, either the first or the second factor of (2.4) is zero. Because  $x_2x_4$  enters the surface fit in the last position of the terms mentioned above, we chose to eliminate it. The contours of the fitted response surface and the associated ridge paths are not affected by which term is eliminated, but substitution of specific numbers into the formulas of Section 3 will, of course, change appropriately. Overall, however, the fitted model is less flexible than it could have been with a better choice of design.

We shall explore the ridges of this surface in two ways, both covered by the theory in Section 3. First, we seek the ridges that emanate from a selected “focal point” of the space restricted by (2.1). Later we add boundary restrictions called for by the exploration.

### 3 Ridge Analysis with Multiple Linear Restrictions, Including Applications to Mixture Experiments

#### 3.1 Ridge Analysis; Basic Method

In its original, unrestricted form (Hoerl 1959, 1962, 1964), ridge analysis was used on a second-order fitted response to obtain a set of paths, going outwards from the origin  $(x_1, x_2, \dots, x_q) = (0, 0, \dots, 0)$  of the factor space. Two of these paths provided the maximum response (path of steepest ascent) and the minimum response (path of steepest descent) on spheres of increasing radius  $R$ , beginning at the origin. Other paths, where the response was neither a maximum nor a minimum, but was locally (on the sphere) stationary, could also be found. These other paths, which might be of interest in practical problems, for example, if they provide good, but not optimum, response values at lower cost, typically do not start at the origin, but appear suddenly when certain radii values (which depend on the specific response surface under study) are attained.

The basic ridge analysis method goes as follows. Suppose the fitted second-order surface is written as

$$\hat{y} = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} \quad (3.1.1)$$

where

$$\mathbf{x}' = (x_1, x_2, \dots, x_q), \quad \mathbf{b}' = (b_1, b_2, \dots, b_q),$$

and where

$$\mathbf{B} = \begin{pmatrix} b_{11} & \frac{1}{2}b_{12} & \cdots & \frac{1}{2}b_{1q} \\ & b_{22} & \cdots & \frac{1}{2}b_{2q} \\ & & \ddots & \vdots \\ sym & & & b_{qq} \end{pmatrix}, \quad (3.1.2)$$

is symmetric. Then (3.1.1) is the matrix format for the second-order fitted equation

$$\begin{aligned} \hat{y} = & b_0 + b_1x_1 + b_2x_2 + \cdots + b_qx_q + b_{11}x_1^2 + b_{22}x_2^2 + \cdots + b_{qq}x_q^2 \\ & + b_{12}x_1x_2 + b_{13}x_1x_3 + \cdots + b_{q-1,q}x_{q-1}x_q. \end{aligned} \quad (3.1.3)$$

The stationary values of (3.1.5) subject to being on a sphere, centered at the origin,

$$\mathbf{x}'\mathbf{x} \equiv x_1^2 + x_2^2 + \cdots + x_q^2 = R^2 \quad (3.1.4)$$

are obtained by considering the Lagrangian function

$$F = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} - \lambda(\mathbf{x}'\mathbf{x} - R^2). \quad (3.1.5)$$

Differentiating (3.1.5) with respect to  $\mathbf{x}$  (which can be achieved by differentiating with respect to  $x_1, x_2, \dots, x_q$  in turn and rewriting these equations in matrix form) gives

$$\frac{\partial F}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} - 2\lambda\mathbf{x}. \quad (3.1.6)$$

Setting (3.1.6) equal to a zero vector leads to

$$2(\mathbf{B} - \lambda\mathbf{I})\mathbf{x} = -\mathbf{b}. \quad (3.1.7)$$

We can now select a value for  $\lambda$ . If  $(\mathbf{B} - \lambda\mathbf{I})^{-1}$  exists, which will happen as long as  $\lambda$  is not an eigenvalue of  $\mathbf{B}$ , we obtain a solution  $\mathbf{x}$  for a stationary point of  $\hat{y}$ ,

$$\mathbf{x} = -\frac{1}{2}(\mathbf{B} - \lambda\mathbf{I})^{-1}\mathbf{b}, \quad (3.1.8)$$

and can then find the radius  $R$ , from (3.1.4), associated with the solution  $\mathbf{x}$  from (3.1.8).

Both  $R$  and  $\mathbf{x}$  are functions of  $\lambda$ .

The theory in Draper (1963) tells us that, if we select values of  $\lambda$  from  $+\infty$  downwards, we shall be on the “maximum  $\hat{y}$ ” path. Values of  $\lambda$  from  $-\infty$  upwards give us the “minimum  $\hat{y}$ ” path. Intermediate paths lie in the ranges of  $\lambda$  between the eigenvalues of  $\mathbf{B}$ .

We next discuss how these methods can be widened in general to facilitate, among other applications, their use in mixture experiments.



### 3.2 Ridge Analysis Around a Selected Focus

Ridge analysis can be started from any selected “focal point”, or “focus”, which we denote here by  $\mathbf{f}$ . (In mixture experiments, for example,  $\mathbf{f}$  could be chosen as a central point, perhaps even the exact centroid, of some pre-defined restricted region in which the experimental runs were confined.) When  $\mathbf{f} \neq \mathbf{0}$ , Eq. (3.1.4),  $\mathbf{x}'\mathbf{x} = R^2$ , would be replaced by

$$(\mathbf{x} - \mathbf{f})'(\mathbf{x} - \mathbf{f}) = R^2. \quad (3.2.1)$$

Note that, if  $\mathbf{f}$  were an unconstrained mixture region centroid with all coordinates identical, that is, if  $\mathbf{f} = (1/q, 1/q, \dots, 1/q)' = (1/q)\mathbf{1}'$ , then

$$\begin{aligned} R^2 = (\mathbf{x} - \mathbf{f})'(\mathbf{x} - \mathbf{f}) &= \mathbf{x}'\mathbf{x} - 2\mathbf{f}'\mathbf{x} + \mathbf{f}'\mathbf{f} \\ &= \mathbf{x}'\mathbf{x} - 2/q + 1/q \\ &= \mathbf{x}'\mathbf{x} - 1/q. \end{aligned} \quad (3.2.2)$$

In this special case, there would be no need to move to the focus at all, as the restriction is now  $\mathbf{x}'\mathbf{x} = R^2 + 1/q$ , essentially a redefinition of the radius value. The physical meaning of this is that any sphere centered at the origin  $(0, 0, \dots, 0)$  eventually expands so that its intersection with the mixture space is a subsphere centered at the mixture space centroid. (For a diagram, see Draper and Pukelsheim 1998, p. 135.)

### 3.3 Adding Linear Restrictions

Suppose we wish to perform ridge analysis subject to a set of linear restrictions of the form

$$\mathbf{A}\mathbf{x} = \mathbf{c} \quad (3.3.1)$$

where  $\mathbf{A}$  is a given  $m \times q$  matrix of linearly independent rows, normalized so that the sum of squares of each row is 1, and  $\mathbf{c}$  is a given  $m \times 1$  vector. For example, if we were investigating a mixture problem with ingredients  $(x_1, x_2, \dots, x_q)$  restricted by

$$\mathbf{1}'\mathbf{x} \equiv \mathbf{x}'\mathbf{1} \equiv x_1 + x_2 + \dots + x_q = 1, \quad (3.3.2)$$

we could choose  $\mathbf{A} = (1/q^{1/2})(1, 1, \dots, 1)$  and  $\mathbf{c} = 1/q^{1/2}$ . If this mixture space were further restricted to the plane

$$(\alpha_1, \alpha_2, \dots, \alpha_q)\mathbf{x} = \alpha, \quad (3.3.3)$$

where all  $\alpha$ 's were pre-specified and  $\alpha_1^2 + \alpha_2^2 + \dots + \alpha_q^2 = 1$ , then  $m = 2$ ,

$$\mathbf{A} = \begin{bmatrix} 1/q^{1/2} & 1/q^{1/2} & \dots & 1/q^{1/2} \\ \alpha_1 & \alpha_2 & \dots & \alpha_q \end{bmatrix}, \text{ and } \mathbf{c} = \begin{bmatrix} 1/q^{1/2} \\ \alpha \end{bmatrix} \quad (3.3.4)$$

and so on. (Of course, any set of non-contradictory, linearly independent linear restrictions can be adopted. We are not confined only to mixture problems adding to 1, although mixtures are our emphasis here.) The dimension  $m$  of  $\mathbf{A}$  must be such that  $m < q$  in general. When  $m = q$  we are reduced to a single point in the  $x$ -space and no paths are possible, or rather, all paths coalesce into a single point. Note that, because  $\mathbf{f}$  must lie in the restricted space,  $\mathbf{A}\mathbf{f} = \mathbf{c}$ .

Under conditions (3.2.1) and (3.3.1), we now consider the Lagrangian function

$$G = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} - \lambda[(\mathbf{x} - \mathbf{f})'(\mathbf{x} - \mathbf{f}) - R^2] - \boldsymbol{\theta}'(\mathbf{A}\mathbf{x} - \mathbf{c}) \quad (3.3.5)$$

where  $\lambda$ , and the elements  $(\theta_1, \theta_2, \dots, \theta_m)$  forming  $\boldsymbol{\theta}'$  are Lagrangian multipliers. Differentiation with respect to  $\mathbf{x}$  leads to

$$\frac{\partial G}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} - 2\lambda(\mathbf{x} - \mathbf{f}) - \mathbf{A}'\boldsymbol{\theta}, \quad (3.3.6)$$

and setting (3.3.6) equal to a zero vector implies

$$2(\mathbf{B} - \lambda\mathbf{I})\mathbf{x} = \mathbf{A}'\boldsymbol{\theta} - \mathbf{b} - 2\lambda\mathbf{f}. \quad (3.3.7)$$

For many given values of  $\lambda$  (the specific choices will be discussed below) we can write a solution for  $\mathbf{x}$  as

$$\mathbf{x} = \frac{1}{2}(\mathbf{B} - \lambda\mathbf{I})^{-1}(\mathbf{A}'\boldsymbol{\theta} - \mathbf{b} - 2\lambda\mathbf{f}). \quad (3.3.8)$$

This  $\mathbf{x}$  must satisfy (3.3.1), which implies that

$$\mathbf{c} = \frac{1}{2}\mathbf{A}(\mathbf{B} - \lambda\mathbf{I})^{-1}\mathbf{A}'\boldsymbol{\theta} - \frac{1}{2}\mathbf{A}(\mathbf{B} - \lambda\mathbf{I})^{-1}(\mathbf{b} + 2\lambda\mathbf{f}), \quad (3.3.9)$$

whereupon

$$\boldsymbol{\theta} = \left\{ \mathbf{A}(\mathbf{B} - \lambda\mathbf{I})^{-1}\mathbf{A}' \right\}^{-1} \left\{ 2\mathbf{c} + \mathbf{A}(\mathbf{B} - \lambda\mathbf{I})^{-1}(\mathbf{b} + 2\lambda\mathbf{f}) \right\}. \quad (3.3.10)$$

This leads us into the following solution sequence:

1. Choose values of  $\lambda$  appropriate for the desired path (to be explained below).
2. Solve (3.3.10) for  $\boldsymbol{\theta}$ .
3. Obtain  $\mathbf{x}$  from (3.3.8).

4. Evaluate  $R^2$  as in (3.2.1).

Then the point  $\mathbf{x}$  will be on the desired path of stationary values and will lie on a sphere of radius  $R$ . The question is now whether the chosen value of  $\lambda$  places us on the maximum path, the minimum path, or some intermediate path.

### 3.4 Determining the Ridge Paths Under Linear Restrictions

In the unrestricted ridge analysis described in Section 3.1, the matrix of second derivatives

$$\left\{ \frac{\partial F}{\partial x_i \partial x_j} \right\} = 2(\mathbf{B} - \lambda \mathbf{I}) \quad (3.4.1)$$

is key in determining which path is selected. The eigenvalues of  $\mathbf{B}$ , that is, the values that result from solving

$$|\mathbf{B} - \lambda \mathbf{I}| = 0 \quad (3.4.2)$$

form the dividing points for the various paths of stationary values. In general, there are  $q$  eigenvalues and  $2q$  paths, see Draper (1963). Those eigenvalues are not appropriate for the restricted problem, however. What are needed instead are the eigenvalues of a lower dimension matrix which makes allowance for the linear restrictions.

We recall that, if there are  $m$  restrictions, as in (3.3.1),  $\mathbf{A}$  is a given  $m \times q$  matrix with  $m$  linearly independent rows of length  $q$ , normalized to make the sum of squares of each row equal to 1.

Let  $\mathbf{T}$  be a  $(q - m) \times q$  matrix each of whose  $(q - m)$  rows is orthogonal to every row of  $\mathbf{A}$ , and such that  $\mathbf{T}\mathbf{T}' = \mathbf{I}_{q-m}$ . Another way of saying this is that the columns of  $\mathbf{A}'$  form a basis for the restriction space, and those of  $\mathbf{T}'$  form an *orthonormal* basis for the space orthogonal to  $\mathbf{A}'$ . It follows that

$$\begin{aligned} \mathbf{T}\mathbf{A}' &= \mathbf{0}, & (\text{of size } (q - m) \times m) \\ \mathbf{A}\mathbf{T}' &= \mathbf{0}, & (\text{of size } m \times (q - m)) \\ \mathbf{T}\mathbf{T}' &= \mathbf{I}_{q-m}. \end{aligned} \quad (3.4.3)$$

The combined matrix,

$$\mathbf{Q} = \begin{bmatrix} \mathbf{A} \\ \mathbf{T} \end{bmatrix},$$

is then a  $q \times q$  matrix which provides a transformation of the coordinate system  $(x_1, x_2, \dots, x_q)$  into coordinates  $z_1, z_2, \dots, z_q$ , via  $\mathbf{z} = \mathbf{Q}\mathbf{x}$ , whereupon  $\mathbf{x} = \mathbf{Q}^{-1}\mathbf{z}$ .

If we partition  $\mathbf{z}' = (z_1, z_2, \dots, z_m, z_{m+1}, \dots, z_q)$  into  $\mathbf{z}' = (\mathbf{u}', \mathbf{v}')$ , where  $\mathbf{u}' = (z_1, z_2, \dots, z_m)$  and  $\mathbf{v}' = (z_{m+1}, \dots, z_q)$ ,

$$\mathbf{z} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{T} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{Ax} \\ \mathbf{T}\mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \mathbf{T}\mathbf{x} \end{bmatrix} \quad (3.4.4)$$

under the restrictions (3.3.1). Consider the inverse of  $\mathbf{Q}$ . This is of the form

$$\mathbf{Q}^{-1} = \begin{bmatrix} \mathbf{A}'(\mathbf{AA}')^{-1}, \mathbf{T}' \end{bmatrix}. \quad (3.4.5)$$

$\mathbf{AA}'$  is nonsingular because of our assumption below (3.3.1) that the restrictions are linearly independent. We verify (3.4.5) by writing

$$\mathbf{Q}\mathbf{Q}^{-1} = \begin{bmatrix} \mathbf{A} \\ \mathbf{T} \end{bmatrix} \begin{bmatrix} \mathbf{A}'(\mathbf{AA}')^{-1}, \mathbf{T}' \end{bmatrix} = \mathbf{I}_q \quad (3.4.6)$$

as a result of conditions (3.4.3). It follows that  $\mathbf{Q}^{-1}\mathbf{Q} = \mathbf{I}$  also, since the inverse is unique.

The first quadratic portion of the Lagrangian function (3.3.5) is thus, using  $\mathbf{x} = \mathbf{Q}^{-1}\mathbf{z}$ , with  $\mathbf{z}$  from (3.4.4) and  $\mathbf{Q}^{-1}$  from (3.4.5),

$$\begin{aligned} \mathbf{x}'\mathbf{B}\mathbf{x} &= \mathbf{z}'(\mathbf{Q}^{-1})'\mathbf{B}\mathbf{Q}^{-1}\mathbf{z} \\ &= [\mathbf{c}', \mathbf{v}'] \begin{bmatrix} (\mathbf{AA}')^{-1}\mathbf{A} \\ \mathbf{T} \end{bmatrix} \mathbf{B} \begin{bmatrix} \mathbf{A}'(\mathbf{AA}')^{-1}, \mathbf{T}' \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \mathbf{v} \end{bmatrix} \end{aligned} \quad (3.4.7)$$

$$= [\mathbf{c}'(\mathbf{AA}')^{-1}\mathbf{A} + \mathbf{v}'\mathbf{T}] \mathbf{B} [\mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{c} + \mathbf{T}'\mathbf{v}] \quad (3.4.8)$$

$$= \mathbf{v}'\mathbf{T}\mathbf{B}\mathbf{T}'\mathbf{v} + 2\mathbf{v}'\mathbf{T}\mathbf{B}\mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{c} + \mathbf{c}'(\mathbf{AA}')^{-1}\mathbf{A}\mathbf{B}\mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{c}, \quad (3.4.9)$$

after reduction. From the result (3.4.9), if we set  $\mathbf{B} = \mathbf{I}$  as a special case and apply (3.4.3), we obtain for the second quadratic portion of (3.3.5)

$$\lambda\mathbf{x}'\mathbf{x} = \lambda\mathbf{v}'\mathbf{v} + \mathbf{0} + \lambda\mathbf{c}'(\mathbf{AA}')^{-1}\mathbf{c}. \quad (3.4.10)$$

Differentiating the transformed version of (3.3.5) twice with respect to  $\mathbf{v}$ , noting that constants and terms linear in  $\mathbf{v}$  drop out, we obtain

$$\left\{ \frac{\partial G}{\partial v_i \partial v_j} \right\} = 2(\mathbf{T}\mathbf{B}\mathbf{T}' - \lambda\mathbf{I}), \quad (3.4.11)$$

in place of (3.4.1).

Note that the size of this square matrix (3.4.11) is  $(q-m)$ , not  $q$ , because  $\mathbf{T}$  is  $(q-m) \times q$ . We see that, when  $\lambda$  is such that (3.4.11) is positive definite, we have a minimum, while if (3.4.11) is negative definite, we have a maximum. If (3.4.11) is indefinite, intermediate stationary values are indicated. In fact, the theory at this point is a complete parallel of that in Draper (1963). If the eigenvalues of  $\mathbf{TBT}'$  are  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_{q-m}$ , arranged in order with due regard to sign then, subject to the restrictions  $\mathbf{Ax} = \mathbf{c}$ :

- (a) Choosing  $\lambda > \mu_{q-m}$  provides a locus of maximum  $\hat{y}$  as  $R$  changes;
- (b) Choosing  $\lambda < \mu_1$  provides a locus of minimum  $\hat{y}$  as  $R$  changes;
- (c) Choosing  $\mu_1 \leq \lambda \leq \mu_{q-m}$  gives intermediate stationary values.

As in the unrestricted case, when  $\lambda = \mu_i$  exactly for  $i = 1, 2, \dots, q-m$ ,  $R$  is infinite. See Draper, 1963.

Note that we do not need these eigenvalues to obtain the paths, but only to distinguish between paths. For the loci of maximum  $\hat{y}$  and the minimum  $\hat{y}$ , the eigenvalues are not necessary at all since choosing  $\lambda$  values decreasing from  $\infty$  gives the path of maximum  $\hat{y}$ , while using values increasing from  $-\infty$  gives the path of minimum  $\hat{y}$ . However, knowing the eigenvalues helps us in selecting appropriate  $\lambda$  values for intermediate paths.

We now apply these results to the mixture problem described by Anik and Sukumar (1981).

## 4 Generalized Ridge Analysis of the Experiment

The foregoing section describes, in a very general context, the calculation details necessary to find the ridge paths as they stream from a selected focus. The important sequence of operations for this lies below equation (3.3.10). We now apply this theory to the Anik and Sukumar (1981) data set. For this, we have  $q = 4$ , and from (2.3) and (2.1)

$$b_0 = 0, \tag{4.1}$$

$$\mathbf{b} = \begin{bmatrix} 49.716 \\ 8.414 \\ 29.95 \\ 4.3365 \end{bmatrix} \tag{4.2}$$

$$\mathbf{B} = \begin{bmatrix} 0 & -29.3355 & -13.915 & -37.451 \\ -29.3355 & 0 & 5.1 & 0 \\ -13.915 & 5.1 & 0 & 16.905 \\ -37.451 & 0 & 16.905 & 0 \end{bmatrix} \quad (4.3)$$

$$\mathbf{A} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}), \quad \mathbf{c} = 0.9/2 = 0.45. \quad (4.4)$$

#### 4.1 The First Set of Ridge Paths

We choose, as the focus  $\mathbf{f}$  of the ridge system, the centroid of the points 1–6 in Table 1, namely  $\mathbf{f} = (0.21, 0.21, 0.04, 0.44)'$ . The distances from  $\mathbf{f}$  to the six points 1, 2, ..., 6 of Table 1 are respectively 0.966, 0.763, 0.703, 0.773, 0.703, and 0.804; these values will give some comparative perspective to the  $R$  values in Table 2. The eigenvalues of  $\mathbf{B}$  are not relevant here because of the restriction (4.4). Instead, we need the eigenvalues of the matrix  $\mathbf{TBT}'$  in (3.4.11). An appropriate  $\mathbf{T}$  takes the form

$$\mathbf{T} = \begin{bmatrix} -0.6708204 & -0.2236068 & 0.2236068 & 0.6708204 \\ 0.5 & -0.5 & -0.5 & 0.5 \\ -0.2236068 & 0.6708204 & -0.6708204 & 0.2236068 \end{bmatrix}. \quad (4.5)$$

The reasoning behind this calculation is explained in Section 3. The rows of  $\mathbf{T}$  consist of the first-, second- and third-order orthogonal polynomials, normalized so that the sum of the squared elements in each row equals 1. (See, for example, Draper and Smith, 1998, p. 466.) The three rows of  $\mathbf{T}$  are orthogonal to one another, the sum of their squares equals 1, and are all orthogonal to  $\frac{1}{2}\mathbf{1}' = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$  which is the normalized vector of coefficients of the  $x$ 's in the mixture restriction  $x_1 + x_2 + x_3 + x_4 = 0.90$ . The three eigenvalues of  $\mathbf{TBT}'$  are  $(-20.04, 2.52, 46.87)$  and the radius  $R$  becomes infinite when  $\lambda$  takes these eigenvalues. The ridge path of maximum  $\hat{y}$  (path A) will be given by choosing  $\lambda$  values from  $\infty$  (where the solution will be  $\mathbf{x} = \mathbf{f}$ , and where  $R = 0$ ) to 46.87 (where the solution will be  $\mathbf{x} = \infty$ ). The ridge of minimum  $\hat{y}$  (path F) will be given by choosing  $\lambda$  values from  $-\infty$  (where the solution will be  $\mathbf{x} = \mathbf{f}$ ) to  $-20.04$  (where the solution will be  $\mathbf{x} = -\infty$ ). Other  $\lambda$  values between the eigenvalues will deliver four more paths  $B, C, D$ , and  $E$  of stationary values of  $\hat{y}$  (see Draper, 1963).

Table 2 shows a selected representative set of values of  $\lambda$  (which we choose initially), of  $(x_1, x_2, x_3, x_4)$  on the paths designated, and of the resultant  $R$  and  $\hat{y}$  values, derived from

the calculations given in Section 3. Path A, the maximum  $\hat{y}$  path begins at the selected focus  $\mathbf{f}$ , where  $R = 0$  and  $\hat{y} = 6.27$ , and moves quickly (see the  $x_3$  values) to the  $x_3 = 0.08$  boundary and beyond, while the values of  $x_1$ ,  $x_2$  and  $x_4$  change only slowly. This clearly shows the importance of variable  $x_3$  and, unless the range of  $x_3$  can be extended past the  $x_3 = 0.08$  value, indicates that further exploration of the fitted surface needs to be carried out on the  $x_3 = 0.08$  face of the restricted region.

Figure 2, derived from the A path details in Table 2, shows how the coordinates  $x_1, x_2, x_3$  and  $x_4$  and the predicted maximum response value  $\hat{y}$  change versus  $R$ . Such a diagram could be drawn also for any of the ridge paths we provide, and is considered by many scientists to be the best way to view the ridge results. It enables practitioners to assess and understand the typically complex interplay between the mixture ingredients as the response improves. It also permits the addition of a “cost” curve for the ingredients, or of any other curves measuring selected qualities of the changing mixture. For reasons of space, however, we provide only this one example, since it duplicates the information in the corresponding table. We recall that closed form expressions for the dependency of  $x_i$  and  $\hat{y}$  upon  $R$  are not available. However numerical computer calculations are feasible and these provided the details for constructing the smooth lines of Figure 2. Alternatively, a satisfactory working diagram can be obtained by plotting the values given in Table 2.

Intermediate paths  $B$  and  $C$  have no points of practical interest. The  $x_1$  values are negative from the eigenvalue  $\lambda = 46.87$  until about  $\lambda = 41.5$ , where the  $x_3$  value reaches a minimum of about  $x_3 = 0.358$ , well above the  $x_3$  upper limit for the experimental region. The minimum  $R$  value of about 0.379 is attained at about  $\lambda = 40$ .

Intermediate paths  $D$  and  $E$  are also of no practical interest, having negative  $x_1$  and  $x_4$  values throughout. Their minimum radius lies beyond the range of  $R$  shown in Figure 2.

The minimum  $\hat{y}$  path  $F$  begins, like path  $A$ , at the selected focus  $\mathbf{f}$ , where  $\lambda = -\infty$ ,  $R = 0$  and  $\hat{y} = 6.27$ . As might be anticipated from the behavior of path  $A$ , path  $E$  goes quickly to the (opposite)  $x_3 = 0$  boundary, after which it is of no practical interest because  $x_3$  must be non-negative. We show selected  $\lambda$  values, to a point where the predicted  $\hat{y}$  has turned negative, in Table 2.

Table 2. Ridge Paths for the Anik and Sukumar (1981) Data, Applying Only the  
Mixture Restriction  $x_1 + x_2 + x_3 + x_4 = 0.9$

Path	$\lambda$	$x_1$	$x_2$	$x_3$	$x_4$	$R$	$\hat{y}$
$A$ (max)	$\infty$	0.210	0.210	0.040	0.440	0	6.27
	2000	0.209	0.207	0.048	0.436	0.010	6.64
	1000	0.208	0.204	0.056	0.432	0.020	7.02
	750	0.207	0.202	0.062	0.429	0.026	7.27
	500	0.206	0.199	0.072	0.423	0.038	7.75
	400	0.205	0.196	0.080	0.419	0.048	8.10
	300	0.204	0.191	0.092	0.413	0.062	8.66
	250	0.203	0.187	0.102	0.408	0.074	9.10
	100	0.201	0.152	0.181	0.366	0.170	12.48
	62	0.230	0.107	0.243	0.320	0.259	15.40
	50	0.441	0.020	0.244	0.195	0.437	21.94
	48	0.920	-0.131	0.168	-0.057	0.940	55.58
$B, C, D, E$ : Do not occur within the experimental region							
$F$ (min)	-90	0.248	0.273	-0.194	0.573	0.279	-6.26
	-100	0.243	0.266	-0.165	0.556	0.244	-4.55
	-200	0.224	0.238	-0.052	0.490	0.109	1.69
	-436	0.216	0.223	0.000	0.461	0.048	4.32
	-500	0.215	0.221	0.005	0.459	0.041	4.58
	-700	0.213	0.218	0.016	0.453	0.029	5.08
	-900	0.213	0.216	0.021	0.450	0.023	5.35
	$-\infty$	0.210	0.210	0.040	0.440	0.000	6.27



## 4.2 The Second Set of Ridge Paths

Because we are interested in maximizing  $\hat{y}$ , we now need to explore the surface on the  $x_3 = 0.08$  plane. (Had we been interested in minimizing  $\hat{y}$  we would have gone to the  $x_3 = 0$  plane instead.)

The theory of Section 3 can again be applied, but now with the addition of the linear equality  $x_3 = 0.08$ . This means that Eq. (4.4) is replaced by

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0.45 \\ 0.08 \end{bmatrix}. \quad (4.6)$$

Repeating the calculations with these new restrictions requires us to use a new  $\mathbf{T}$ ,

$$\mathbf{T} = \begin{bmatrix} 0.267261 & 0.534523 & 0 & -0.801784 \\ 0.771517 & -0.617213 & 0 & -0.154303 \end{bmatrix} \quad (4.7)$$

which leads to the eigenvalues of  $\mathbf{T}\mathbf{B}\mathbf{T}'$  being  $(-0.49, 45.01)$ . (The two rows of  $\mathbf{T}$  are  $(1, 2, 0, -3)$  and  $(5, -4, 0, 1)$  renormalized to have sum of squares 1.) There are now four ridge paths which we designate  $A$  (maximum  $\hat{y}$ ),  $B, C$  and  $D$  (minimum  $\hat{y}$ ). A new focus needs to be chosen.

The current restricted region is shown in Figure 1(b). The design points 2, 4 and 6 from Table 1 lie on this pentagon and we choose  $\mathbf{f} = (0.203, 0.203, 0.08, 0.413)'$ , their centroid. This point lies at distances  $R = 0.253, 0.241, 0.241$  from points 2, 4, 6 respectively, and these numbers can be compared with the values of  $R$  that we see on the ridge paths shown in Table 3. We recall that  $x_3 = 0.08$  throughout, and we show the paths in Figure 3. The maximum  $\hat{y}$  path crosses the  $x_4 = 0.30$  boundary when  $\lambda$  is about 65.95. As in all steepest ascent studies when a boundary is met, one must now move along this boundary. We postpone this for the moment to discuss the other three ridge traces. Neither path  $B$  nor path  $C$  lies within the restricted region and their details are not given. The minimum  $\hat{y}$  path  $D$  moves downward until  $x_1$  is about 0.16 and then turns, crossing the  $x_2 = 0.10$  boundary at roughly this  $x_1 = 0.16$  level; see Figure 3.

## 4.3 The Third Set of Ridge Paths

To move along the boundary  $x_4 = 0.30$ , we designate a new focus  $\mathbf{f}$  and a new matrix  $\mathbf{T}$ . The end points of the restricted region along the boundary are the corner points

Table 3. Ridge Paths for the Anik and Sukumar (1981) Data,  
Under the Restrictions  $x_3 = 0.08$ ,  $x_1 + x_2 + x_4 = 0.82$ .

Path	$\lambda$	$x_1$	$x_2$	$x_3$	$x_4$	$R$	$\hat{y}$
$A$ (max)	$\infty$	0.203	0.203	0.08	0.414	0.000	8.12
	1000	0.207	0.203	0.08	0.410	0.005	8.16
	500	0.211	0.202	0.08	0.407	0.010	8.21
	200	0.225	0.200	0.08	0.395	0.029	8.41
	150	0.236	0.197	0.08	0.387	0.042	8.57
	125	0.246	0.194	0.08	0.380	0.055	8.74
	100	0.265	0.189	0.08	0.366	0.079	9.10
	90	0.279	0.184	0.08	0.357	0.097	9.39
	80	0.301	0.177	0.08	0.342	0.124	9.90
	75	0.317	0.171	0.08	0.332	0.144	10.32
	70	0.341	0.162	0.08	0.317	0.173	10.97
	66	0.367	0.152	0.08	0.301	0.205	11.80
	65.95	0.368	0.152	0.08	0.300	0.206	11.82
	60	0.433	0.127	0.08	0.260	0.287	14.31
	55	0.549	0.081	0.08	0.190	0.429	20.13
	52	0.698	0.021	0.08	0.101	0.613	30.32
$B, C$ : Do not occur within the region $x_3 = 0.08$ , $x_1 + x_2 + x_4 = 0.82$							
$D$ (min)	-6	0.158	0.028	0.08	0.634	0.286	6.86
	-7	0.156	0.058	0.08	0.606	0.245	7.00
	-9	0.154	0.098	0.08	0.568	0.194	7.18
	-9.15	0.154	0.100	0.08	0.566	0.191	7.19
	-10	0.154	0.111	0.08	0.555	0.176	7.23
	-20	0.156	0.168	0.08	0.496	0.101	7.51
	-30	0.161	0.184	0.08	0.475	0.077	7.62
	-40	0.166	0.192	0.08	0.462	0.063	7.69
	-50	0.169	0.196	0.08	0.455	0.054	7.73
	-70	0.175	0.200	0.08	0.445	0.043	7.80
	-100	0.181	0.202	0.08	0.437	0.033	7.86
	-500	0.197	0.204	0.08	0.419	0.008	8.05
	-1000	0.200	0.204	0.08	0.416	0.004	8.08
	$-\infty$	0.203	0.203	0.08	0.414	0.000	8.12

$(0.40, 0.12, 0.08, 0.30)$ , near design point 6 in Table 1, and  $(0.12, 0.40, 0.08, 0.30)$ , near design point 4; see Figure 1. We choose their centroid, namely  $\mathbf{f} = (0.26, 0.26, 0.08, 0.30)$ .  $\mathbf{T}$  is now a normalized row vector orthogonal to the rows of  $\mathbf{A}$  in  $\mathbf{Ax} = \mathbf{c}$ , namely

$$\begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.08 \\ 0.30 \end{bmatrix}. \quad (4.8)$$

Necessarily,  $\mathbf{T} = (0.707107, -0.707107, 0, 0)$ , or the vector with signs reversed. The sole eigenvalue of  $\mathbf{TBT}'$  is 29.3355, which is  $-\frac{1}{2}b_{12}$  where  $b_{12}$  is the regression coefficient associated with  $x_1x_2$ . Only the path A of maximum  $\hat{y}$  ( $\lambda > 29.3355$ ) and the path B of minimum  $\hat{y}$  ( $\lambda < 29.3355$ ) exist. On these paths,  $x_3 = 0.08, x_4 = 0.30$  and so  $x_1 + x_2 = 0.52$ . Thus we can show the paths most simply by quoting only the  $x_1$  value, as we have done in Table 4. Path A is shown only to the point  $x_1 = 0.40$  when the first corner point is reached and we attain the maximum predicted response,  $\hat{y} = 12.81$ , subject to the restrictions. Path B is shown only to  $x_1 = 0.12$  when the second corner point is reached. (It is *not*, of course, the minimum region response, which we would find by exploring the  $x_2 = 0.10$  boundary, choosing  $\mathbf{f} = (0.25, 0.10, 0.08, 0.47)$  and  $\mathbf{T} = (0.707107, 0, 0, -0.707107)$  or the vector with signs reversed.) We see that, by a triple application of the ridge analysis technique, we have come to the predicted maximum response in the restricted region, improving from  $\hat{y} = 8.09$  in Table 2 to  $\hat{y} = 11.82$  in Table 3 to  $\hat{y} = 12.81$  in Table 4.

## 5 Summary and Discussion

Ridge analysis, due to A.E. Hoerl (1959, 1962, 1964), can be applied to response surfaces, most usefully those of second-order, to provide a curved direction of steepest ascent for  $\hat{y}$  in the space of the predictor variables  $x_1, x_2, \dots, x_q$ . It is also possible to determine a path of steepest descent or paths of intermediate stationary values, by finding the stationary values of the fitted response  $\hat{y}$  on a sphere of radius  $R$ , and following the solutions as  $R$  expands. This technique is especially useful on surfaces where  $q$  is large, when geometrical visualization is often difficult. In this paper, the technique is extended to mixture response surfaces in a very general way. The focus from which the curved paths emanate can be freely chosen, and any linear equality restrictions, including the usual mixture restriction

Table 4. Ridge Paths for the Anik and Sukumar (1981) Data,  
under the Restrictions  $x_3 = 0.08$ ,  $x_4 = 0.30$ ,  $x_1 + x_2 = 0.52$ .

Path	$\lambda$	$x_1$	$R$	$\hat{y}$
$A$ (max)	$\infty$	0.260	0.000	9.45
	1000	0.264	0.006	9.51
	500	0.268	0.012	9.58
	250	0.278	0.025	9.75
	100	0.316	0.079	10.51
	90	0.325	0.092	10.72
	80	0.338	0.110	11.03
	75	0.346	0.122	11.25
	70	0.357	0.137	11.53
	65	0.371	0.157	11.91
	60	0.389	0.182	12.45
	57.5	0.400	0.198	12.81
$B$ (min)	1.15	0.120	0.198	8.39
	0	0.125	0.190	8.38
	-10	0.160	0.142	8.45
	-20	0.180	0.113	8.56
	-40	0.203	0.081	8.74
	-100	0.229	0.043	9.02
	-200	0.243	0.024	9.19
	-750	0.255	0.007	9.37
	$-\infty$	0.260	0.000	9.45

$x_1 + x_2 + \dots + x_q = 1$ , can be incorporated into the analysis. In an illustration using data from Anik and Sukumar (1981), this technique is applied in three stages to take account of tightening restrictions on the best path caused by factor space limitations, and to find the point in that space of maximum predicted response. An advantage of this method is that there is no need to change the initial  $x$ -coordinate system, nor to use pseudo-components, in any of the resulting calculations.

We now briefly discuss points that arise in connection with this work:

1. The exact choice of focus  $\mathbf{f}$  is not a crucial feature of the restricted steepest ascent/descent procedure we have described. After the first stage in our example, one might have argued that, since the path of maximum  $\hat{y}$  entered the  $x_3 = 0.08$  face of the restricted region at  $(0.205, 0.196, 0.08, 0.419)$ , we should start again there. However, steepest ascent is a very flexible procedure, and a rigid method for choosing  $\mathbf{f}$  would be inappropriate. Choosing some central point of the region is always safe, barring pathological examples.

2. The formulas we have given also can be applied to steepest ascent subject to linear restrictions when the model is a first-order mixture model  $\hat{y} = b_1x_1 + \dots + b_qx_q$ . In this case,  $b_0 = 0$ ,  $\mathbf{B} = \mathbf{0}$  in (3.3.5) through (3.3.10). The “eigenvalues of  $\mathbf{B}$ ” are all zero and, by the choice of  $\mathbf{f}$ ,  $\mathbf{A}\mathbf{f} = \mathbf{c}$ . The solution reduces to

$$\mathbf{x} = \mathbf{f} + (2\lambda)^{-1}(\mathbf{I} - \mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}\mathbf{A})\mathbf{b}. \quad (5.1)$$

The choice of  $\lambda \in [0, \infty]$  gives the straight line steepest-ascent direction, and  $\lambda \in [-\infty, 0]$  gives the steepest-descent direction. Note that, when there are no linear conditions on  $\mathbf{x}$ ,  $\mathbf{A} = \mathbf{0}$  and  $\mathbf{x} - \mathbf{f}$  is proportional to  $\mathbf{b}$  as required.

3. A reviewer pointed out that a move to a selected focus  $\mathbf{f}$  could be accompanied by changing to pseudo-components, if desired. This would involve a preliminary transformation of the form  $\mathbf{z} = u\mathbf{x} - \mathbf{v}$  which might improve conditioning for the design used.

4. In our example, the paths of intermediate stationary values were of no practical interest; in other examples, they may well be. As a reviewer pointed out, “a secondary maximum, ... that would give us near-optimal properties ... may be in a very distant location in design space ... [and] could have other advantages in terms of cost, ease of operation, safety, etc. [and might improve] additional responses”. We fully agree, but add that, because of the mixture restrictions, such locations often fall outside permissible operating conditions. Certainly, these other paths need to be examined in all cases.

5. A reviewer questioned whether the stage-by-stage following of the optimum  $\hat{y}$  path to, and along, boundaries of the restricted region necessarily leads to the overall optimum. As a specific check of the example of Table 4, which gives the maximum  $\hat{y} = 12.81$  value at the true vertex  $(0.40, 0.12, 0.08, 0.30)$ , we calculated the predicted response values at all ten true vertices of the restricted region. Among these ten  $\hat{y}$  values, the second largest is 12.63 and occurs at the vertex  $(0.40, 0.10, 0.08, 0.32)$ , the vertex closest to the maximum. More generally it would be possible to use the methods of this paper on any selected subregion, including the faces of the bounding polyhedron. In cases where boundaries cut off the path of the maximum ridge quickly, and where secondary paths begin within the restricted region, it would be possible for the true restricted maximum to lie on another path. In our example, there are no secondary paths within the restricted region, so this cannot occur.

6. The contours of Figure 3 are drawn here *only* to show the paths, and thereby display what the method achieves. One does not actually need the contours, as examination of the coordinates in Tables 2, 3 and 4 makes clear. This would be especially important in a high-dimensional mixture space, when contours could be drawn only in sections.

## Acknowledgments

We are grateful to the Alexander von Humboldt-Stiftung for support through a Max-Planck-Award for cooperative research. We thank the Editor, an Associate Editor, and two referees, all of whom made excellent comments that greatly improved the original version of this article.

## References

- ANIK, S.T. and SUKUMAR, L. (1981). "Extreme Vertexes Design in Formulation Development: Solubility of Butoconazole Nitrate in a Multicomponent System". *Journal of Pharmaceutical Sciences*, 70, No. 8, pp. 897–900.
- CORNELL, J.A. (1990). *Experiments with Mixtures*, 2nd ed, Wiley, New York.
- DEL CASTILLO, E., FAN, S.K. and SEMPLE, J. (1997). "The Computation of Global Optima in Dual Response Systems". *Journal of Quality Technology*, 29, pp. 347–353.

- DEL CASTILLO, E., FAN, S.K. and SEMPLE, J. (1999). "Optimization of Dual Response Systems: A Comprehensive Procedure for Degenerate and Nondegenerate Problems". *European Journal of Operations Research*, 112, pp. 174–186.
- DRAPER, N.R. (1963). "Ridge Analysis of Response Surfaces". *Technometrics*, 5, pp. 469–479.
- DRAPER, N.R. and PUKELSHEIM, F. (2000). "Ridge Analysis of Mixture Response Surfaces". *Statistics & Probability Letters*, 48, pp. 131–140.
- DRAPER, N.R. and SMITH, H. (1998). *Applied Regression Analysis*, Wiley, New York.
- HOERL, A.E. (1959). "Optimum Solution of Many Variables Equations". *Chemical Engineering Progress*, 55(11), pp. 69–78.
- HOERL, A.E. (1962). "Applications of Ridge Analysis to Regression Problems". *Chemical Engineering Progress*, 58(3), pp. 54–59.
- HOERL, A.E. (1964). "Ridge Analysis". *Chemical Engineering Progress Symposium Series*, 60, pp. 67–77.
- HOERL, R.W. (1985). "Ridge Analysis 25 Years Later". *The American Statistician*, 39, pp. 186–192.
- HOERL, R.W. (1987). "The Application of Ridge Techniques to Mixture Data: Ridge Analysis". *Technometrics*, 29, pp. 161–172.
- McLEAN, R.A. and ANDERSON, V.L. (1966). "Extreme Vertices Design of Mixture Experiments". *Technometrics*, 8, pp. 447–454. (Discussion by J.W. Gorman, pp. 455–456.)
- MYERS, R.H. and CARTER, W.H., Jr. (1973). "Response Surface Techniques for Dual Response Systems". *Technometrics*, 15, pp. 301–317.
- PRESCOTT, P., DEAN, A.M., DRAPER, N.R. and LEWIS, S.M. (2000). "Mixture Experiments: Ill-conditioning and Quadratic Model Formulation". Manuscript submitted for publication.

- SCHEFFÉ, H. (1958). “Experiments with Mixtures”. *Journal of the Royal Statistical Society, Series B*, 20, pp. 344–360.
- SCHEFFÉ, H. (1963). “The Simplex-Centroid Design for Experiments with Mixtures”. *Journal of the Royal Statistical Society, Series B*, 25, pp. 235–263.
- SEMPLE, J. (1997). “Optimality Conditions and Solutions Procedures for Nondegenerate Dual Response Systems”. *IIE Transactions*, 29, pp. 743–752.



## Captions for the figures

Figure 1. (a) The triangular subspace  $x_3 = 0$  contains the pentagon defined by the restrictions on the mixture ingredients. (b) The triangular subspace  $x_3 = 0.08$  (inner triangle) contains the (different) pentagon defined by the restrictions. (c) The two slices are superimposed as in a view downwards from the  $x_3 = 0.90$  vertex of the mixture space. By joining corresponding pentagon vertices, one defines the entire restricted region.

Figure 2. The maximum predicted response  $\hat{y}$  and its corresponding positional coordinates  $(x_1, x_2, x_3, x_4)$  are plotted against  $R$ , the distance the point lies from the focus  $\mathbf{f} = (0.210, 0.210, 0.04, 0.440)'$  in the space  $x_1 + x_2 + x_3 + x_4 = 0.90$ . The numerical details are given in Table 2.

Figure 3. The fitted contours defined by Eq. (2.3) when  $x_3 = 0.08$  are shown in the subspace  $x_1 + x_2 + x_4 = 0.82$ . The ridge paths of maximum  $\hat{y}$  and minimum  $\hat{y}$  on spheres of radius  $R$  emanate from the focus  $\mathbf{f} = (0.203, 0.203, 0.08, 0.413)'$ ; numerical details are in Table 3. The pentagon is the inner one of Figure 1(b).