

Generalized Ridge Analysis Under Linear Restrictions, With Particular Applications to Mixture Experiments Problems

Norman R. DRAPER

Statistics Department
University of Wisconsin
Madison, WI 53706
(draper@cs.wisc.edu)

Friedrich PUKELSHEIM

Institut für Mathematik
Universität Augsburg
D-86135 Augsburg, Germany
(pukelsheim@math.uni-augsburg.de)

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 outward 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 ridge trace view enables practitioners to assess and understand the typically complex interplay between the input variables as the response improves. To explore a subspace defined by a linear restriction on the predictors, a situation discussed infrequently in the literature and never in the context of mixture experiments, we show how a modification of ridge regression can be used generally 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, the ridge origin need not be moved 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; Projection; Response surface; Restriction 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 Semple (1997). R. W. Hoerl (1985) provided a wide-ranging discussion.

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, as we show later, ridge analysis can proceed in a mixture space around *any* point without preliminary transformation. We also show that additional linear equalities in the mixture ingredients are easily incorporated into a very general method that 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 whether 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 that entails additional linear equalities and inequalities on the mixture ingredients and 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.

Anik and Sukumar (1981) conducted a study of five ingredients, one of which, x_5 , was held constant at .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

$$x_1 + x_2 + x_3 + x_4 = .9. \quad (1)$$

As a reviewer commented, (1) could be renormalized via $x_i = .9u_i$, 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 that would have to be undone in later calculations. (However, such a renormalization is usually needed when constructing diagrams as we

show in Sec. 4.2.) The essence of our method is that the ridge paths are obtained directly (and more easily) 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 vertices, edge (one-dimensional) centroids, 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 vertices. The method has various subtleties (which we do not describe; see McLean and Anderson 1966 or Cornell 1990) because 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. 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 $x_3 = .08$ subspace. The outer triangle of Figure 1(b) is the same triangle as in Figure 1(a) and makes the point that the $x_3 = .08$ 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 $x_3 = .90$ vertex. We further note that for each pentagon, two pairs of vertices are quite close together. Consequently, 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" (.15, .40, 0, .35) is the average of true vertices (.10, .40, 0, .40) and (.20, .40, 0, .30); "vertex 4" (.11, .40, .08, .31) is the average of true vertices (.10, .40, .08, .32) and (.12, .40, .08, .30); "vertex 5" (.40, .15, 0, .35) is the average of true vertices (.40, .10, 0, .40) and (.40, .20, 0, .30), and "vertex 6" (.40, .11, .08, .31) is the average of true vertices (.40, .10, .08, .32) and (.40, .12, .08, .30).

We revisit this example in Section 4 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,

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \epsilon, \quad (2)$$

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 was given by Prescott, Dean, Draper, and Lewis (2002). 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 containing a constant term, one of several possibilities. The equation resulting from fitting (2) by least squares is

$$\begin{aligned} \hat{y} = & 49.716x_1 + 8.414x_2 + 29.95x_3 + 4.3365x_4 \\ & - 58.671x_1x_2 - 27.83x_1x_3 - 74.902x_1x_4 \\ & + 10.20x_2x_3 + 33.81x_3x_4. \end{aligned} \quad (3)$$

Note that the nonlinear blending term x_2x_4 is missing in (3). When the Scheffé model is used with the design of Table 1, the resulting $X'X$ matrix is singular. Regression of the x_2x_4 column onto the remaining 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

$$(x_1 - x_2)(x_3 + 2x_4 - .7) = 0. \quad (4)$$

For every data point in Table 1, either the first or the second factor of (4) is 0. Because x_2x_4 enters the surface fit in the last position of the terms mentioned earlier, 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 change appropriately. Overall, however, the fitted model is less flexible than it could have been with a better choice of design.

We 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 (1). Later, we add boundary restrictions called for by the exploration.

Table 1. Experimental Design Used by Anik and Sukumar (1981) Together With the Lower and Upper Limits That 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	.10	.10	0	.30	
Upper limit	.40	.40	.08	.70	
Vertices					
1	.10	.10	0	.70	3.0
2	.10	.10	.08	.62	7.3
3	.15	.40	0	.35	4.9
4	.11	.40	.08	.31	8.4
5	.40	.15	0	.35	8.6
6	.40	.11	.08	.31	12.7
Edge centroids (averages of indicated vertices)					
7 (1,2)	.10	.10	.04	.66	5.1
8 (5,6)	.40	.13	.04	.33	10.8
9 (3,4)	.13	.40	.04	.33	6.6
10 (1,3,5)	.216	.216	0	.468	4.4
11 (2,4,6)	.203	.203	.08	.414	7.9
12 (4,6)	.255	.255	.08	.31	9.4
13 (3,5)	.275	.275	0	.35	5.8
Overall centroid					
14	.21	.21	.04	.44	6.3

NOTE: x_1 , polyethylene glycol 400; x_2 , glycerine; x_3 , polysorbate 60; x_4 , water; y , solubility, (mg/mL). Note that $x_1 + x_2 + x_3 + x_4 = .90$ for each point.

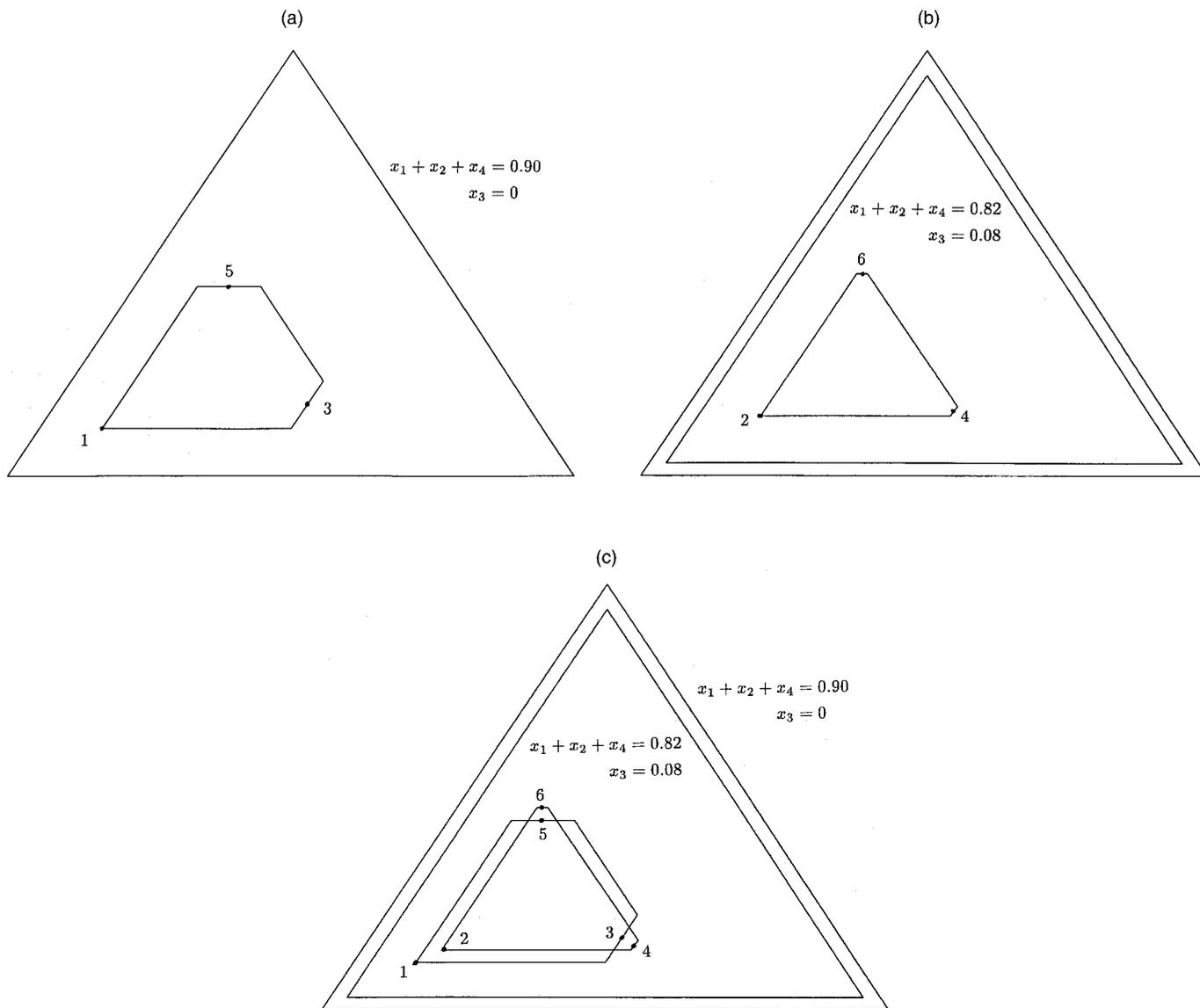


Figure 1. The Triangular Subspaces (a) $x_3 = 0$, Containing the Pentagon Defined by the Restrictions on the Mixture Ingredients, and (b) $x_3 = .08$ (inner triangle), Containing the (different) Pentagon Defined by the Restrictions, (c) The Two Slices are Superimposed as in a View Downward From the $x_3 = .90$ Vertex of the Mixture Space. By joining corresponding pentagon vertices, one defines the entire restricted region.

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 outward 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, in which 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 proceeds as follows. Suppose that the fitted second-order surface is written as

$$\hat{y} = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x}, \quad (5)$$

where

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

and

$$\mathbf{B} = \begin{pmatrix} b_{11} & \frac{1}{2}b_{12} & \cdots & \frac{1}{2}b_{1q} \\ & b_{22} & \cdots & \frac{1}{2}b_{2q} \\ & & \ddots & \vdots \\ sym & & & b_{qq} \end{pmatrix} \quad (6)$$

is symmetric. Then (5) is the matrix format for the second-order fitted equation

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_qx_q + b_{11}x_1^2 + b_{22}x_2^2 + \dots + b_{qq}x_q^2 + b_{12}x_1x_2 + b_{13}x_1x_3 + \dots + b_{q-1,q}x_{q-1}x_q. \quad (7)$$

The stationary values of (7), subject to being on a sphere centered at the origin,

$$\mathbf{x}'\mathbf{x} \equiv x_1^2 + x_2^2 + \dots + x_q^2 = R^2, \quad (8)$$

are obtained by considering the Lagrangian function

$$F = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} - \lambda(\mathbf{x}'\mathbf{x} - R^2). \quad (9)$$

Differentiating (9) 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

$$\frac{\partial F}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} - 2\lambda\mathbf{x}. \quad (10)$$

Setting (10) equal to a zero vector leads to

$$2(\mathbf{B} - \lambda\mathbf{I})\mathbf{x} = -\mathbf{b}. \quad (11)$$

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

$$\mathbf{x} = -\frac{1}{2}(\mathbf{B} - \lambda\mathbf{I})^{-1}\mathbf{b}, \quad (12)$$

and can then find the radius R , from (8), associated with the solution \mathbf{x} from (12). Both R and \mathbf{x} are functions of λ .

The theory of Draper (1963) tells us that if we select values of λ from $+\infty$ downward, then we shall be on the "maximum \hat{y} " path. Values of λ from $-\infty$ upward yield the "minimum \hat{y} " path. Intermediate paths lie in the ranges of λ 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 predefined restricted region in which the experimental runs were confined.) When $\mathbf{f} \neq \mathbf{0}$, (8), $\mathbf{x}'\mathbf{x} = R^2$, would be replaced by

$$(\mathbf{x} - \mathbf{f})'(\mathbf{x} - \mathbf{f}) = R^2. \quad (13)$$

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

$$\begin{aligned} R^2 &= (\mathbf{x} - \mathbf{f})'(\mathbf{x} - \mathbf{f}) = \mathbf{x}'\mathbf{x} - 2\mathbf{f}'\mathbf{x} + \mathbf{f}'\mathbf{f} \\ &= \mathbf{x}'\mathbf{x} - 2/q + 1/q \\ &= \mathbf{x}'\mathbf{x} - 1/q. \end{aligned} \quad (14)$$

In this special case, the focus need not be moved at all, because 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 2000, p. 135.)

3.3 Adding Linear Restrictions

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

$$\mathbf{A}\mathbf{x} = \mathbf{c}, \quad (15)$$

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

$$\mathbf{1}'\mathbf{x} \equiv \mathbf{x}'\mathbf{1} \equiv x_1 + x_2 + \dots + x_q = 1, \quad (16)$$

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

$$(\alpha_1, \alpha_2, \dots, \alpha_q)\mathbf{x} = \alpha, \quad (17)$$

where all α 's were prespecified and $\alpha_1^2 + \alpha_2^2 + \dots + \alpha_q^2 = 1$, then

$$\mathbf{A} = \begin{bmatrix} 1/q^{1/2} & 1/q^{1/2} & \dots & 1/q^{1/2} \\ \alpha_1 & \alpha_2 & \dots & \alpha_q \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} 1/q^{1/2} \\ \alpha \end{bmatrix} \quad (18)$$

($m = 2$), and so on. (Of course, any set of noncontradictory, linearly independent linear restrictions can be adopted. We are not confined only to mixtures where the components add 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 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 (13) and (15), we now consider the Lagrangian function

$$G = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} - \lambda[(\mathbf{x} - \mathbf{f})'(\mathbf{x} - \mathbf{f}) - R^2] - \boldsymbol{\theta}'(\mathbf{A}\mathbf{x} - \mathbf{c}), \quad (19)$$

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

$$\frac{\partial G}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} - 2\lambda(\mathbf{x} - \mathbf{f}) - \mathbf{A}'\boldsymbol{\theta}, \quad (20)$$

and setting (20) equal to a zero vector implies that

$$2(\mathbf{B} - \lambda\mathbf{I})\mathbf{x} = \mathbf{A}'\boldsymbol{\theta} - \mathbf{b} - 2\lambda\mathbf{f}. \quad (21)$$

For many given values of λ (the specific choices are discussed later), we can write a solution for \mathbf{x} as

$$\mathbf{x} = \frac{1}{2}(\mathbf{B} - \lambda\mathbf{I})^{-1}(\mathbf{A}'\boldsymbol{\theta} - \mathbf{b} - 2\lambda\mathbf{f}). \quad (22)$$

This \mathbf{x} must satisfy (15), which implies that

$$\mathbf{c} = \frac{1}{2}\mathbf{A}(\mathbf{B} - \lambda\mathbf{I})^{-1}\mathbf{A}'\boldsymbol{\theta} - \frac{1}{2}\mathbf{A}(\mathbf{B} - \lambda\mathbf{I})^{-1}(\mathbf{b} + 2\lambda\mathbf{f}), \quad (23)$$

whereupon

$$\boldsymbol{\theta} = \{\mathbf{A}(\mathbf{B} - \lambda\mathbf{I})^{-1}\mathbf{A}'\}^{-1}\{2\mathbf{c} + \mathbf{A}(\mathbf{B} - \lambda\mathbf{I})^{-1}(\mathbf{b} + 2\lambda\mathbf{f})\}. \quad (24)$$

This leads to the following solution sequence:

1. Choose values of λ appropriate for the desired path (explained later).

2. Solve (24) for θ .
3. Obtain x from (22).
4. Evaluate R^2 as in (13).

Then the point 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 λ 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,

$$\left\{ \frac{\partial F}{\partial x_i \partial x_j} \right\} = 2(\mathbf{B} - \lambda \mathbf{I}), \tag{25}$$

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

$$|\mathbf{B} - \lambda \mathbf{I}| = 0, \tag{26}$$

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; instead, we need the eigenvalues of a lower-dimension matrix that makes allowance for the linear restrictions.

We recall that, with m restrictions as in (15), \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 $\mathbf{T}\mathbf{T}' = \mathbf{I}_{q-m}$. That is, 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

$$\begin{aligned} \mathbf{T}\mathbf{A}' &= \mathbf{0}, & \text{of size } (q - m) \times m, \\ \mathbf{A}\mathbf{T}' &= \mathbf{0}, & \text{of size } m \times (q - m), \\ \mathbf{T}\mathbf{T}' &= \mathbf{I}_{q-m}. \end{aligned} \tag{27}$$

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 $\mathbf{v}' = (z_{m+1}, \dots, z_q)$,

$$\mathbf{z} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{T} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{A}\mathbf{x} \\ \mathbf{T}\mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \mathbf{T}\mathbf{x} \end{bmatrix} \tag{28}$$

under the restrictions (15). Consider the inverse of \mathbf{Q} , which is of the form

$$\mathbf{Q}^{-1} = [\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}, \mathbf{T}']. \tag{29}$$

$\mathbf{A}\mathbf{A}'$ is nonsingular because of our assumption after (15) that the restrictions are linearly independent. We verify (29) by writing

$$\mathbf{Q}\mathbf{Q}^{-1} = \begin{bmatrix} \mathbf{A} \\ \mathbf{T} \end{bmatrix} [\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}, \mathbf{T}'] = \mathbf{I}_q \tag{30}$$

as a result of conditions (27). It follows that $\mathbf{Q}^{-1}\mathbf{Q} = \mathbf{I}$ also, because the inverse is unique.

Thus, using $\mathbf{x} = \mathbf{Q}^{-1}\mathbf{z}$, with \mathbf{z} from (28) and \mathbf{Q}^{-1} from (29), the first quadratic portion of the Lagrangian function (19) is

$$\begin{aligned} \mathbf{x}'\mathbf{B}\mathbf{x} &= \mathbf{z}'(\mathbf{Q}^{-1})'\mathbf{B}\mathbf{Q}^{-1}\mathbf{z} \\ &= [\mathbf{c}', \mathbf{v}'] \begin{bmatrix} (\mathbf{A}\mathbf{A}')^{-1}\mathbf{A} \\ \mathbf{T} \end{bmatrix} \mathbf{B} [\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}, \mathbf{T}'] \begin{bmatrix} \mathbf{c} \\ \mathbf{v} \end{bmatrix} \end{aligned} \tag{31}$$

$$= [\mathbf{c}'(\mathbf{A}\mathbf{A}')^{-1}\mathbf{A} + \mathbf{v}'\mathbf{T}'] \mathbf{B} [\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}\mathbf{c} + \mathbf{T}'\mathbf{v}] \tag{32}$$

$$\begin{aligned} &= \mathbf{v}'\mathbf{T}\mathbf{B}\mathbf{T}'\mathbf{v} + 2\mathbf{v}'\mathbf{T}\mathbf{B}\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}\mathbf{c} \\ &\quad + \mathbf{c}'(\mathbf{A}\mathbf{A}')^{-1}\mathbf{A}\mathbf{B}\mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}\mathbf{c}, \end{aligned} \tag{33}$$

after reduction. From the result (33), if we set $\mathbf{B} = \mathbf{I}$ as a special case and apply (27), then we obtain, for the second quadratic portion of (19),

$$\lambda \mathbf{x}'\mathbf{x} = \lambda \mathbf{v}'\mathbf{v} + 0 + \lambda \mathbf{c}'(\mathbf{A}\mathbf{A}')^{-1}\mathbf{c}. \tag{34}$$

Differentiating the transformed version of (19) twice with respect to \mathbf{v} , and noting that constants and terms linear in \mathbf{v} drop out, we obtain

$$\left\{ \frac{\partial G}{\partial v_i \partial v_j} \right\} = 2(\mathbf{T}\mathbf{B}\mathbf{T}' - \lambda \mathbf{I}) \tag{35}$$

in place of (25).

Note that the size of this square matrix (35) is $(q - m)$, not q , because \mathbf{T} is $(q - m) \times q$. We see that when λ is such that (35) is positive definite, we have a minimum, whereas if (35) is negative definite, we have a maximum. If (35) is indefinite, intermediate stationary values are indicated. In fact, the theory at this point is a complete parallel of that of Draper (1963). If the eigenvalues of $\mathbf{T}\mathbf{B}\mathbf{T}'$ 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 $\mathbf{A}\mathbf{x} = \mathbf{c}$, the following conditions hold:

- 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 among paths. For the loci of maximum \hat{y} and minimum \hat{y} , the eigenvalues are not necessary, because choosing λ values decreasing from ∞ gives the path of maximum \hat{y} , whereas using values increasing from $-\infty$ gives the path of minimum \hat{y} . However, knowing the eigenvalues helps us select appropriate λ 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 (repetitive) operations for this lies below (24).] We now apply this theory to the Anik and Sukumar (1981) dataset. Here, $q = 4$, and from (3) and (1),

$$b_0 = 0, \tag{36}$$

$$b = \begin{bmatrix} 49.716 \\ 8.414 \\ 29.95 \\ 4.3365 \end{bmatrix}, \tag{37}$$

$$B = \begin{bmatrix} 0 & -29.3355 & -13.915 & -37.451 \\ -29.3355 & 0 & 5.1 & 0 \\ -13.915 & 5.1 & 0 & 16.905 \\ -37.451 & 0 & 16.905 & 0 \end{bmatrix}, \tag{38}$$

$$A = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right) \quad \text{and} \quad c = .9/2 = .45. \tag{39}$$

4.1 The First Set of Ridge Paths

We choose the centroid of the points 1–6 in Table 1 as the focus f of the ridge system, namely $f = (.21, .21, .04, .44)'$. The distances from f to the six points 1, 2, . . . , 6 of Table 1 are .966, .763, .703, .773, .703, and .804; these values will give some comparative perspective to the R values in Table 2. The eigenvalues of B are not relevant here because of the restriction (39). Instead, we need the eigenvalues of the matrix TBT' in (35). An appropriate T takes the form

$$T = \begin{bmatrix} -.6708204 & -.2236068 & .2236068 & .6708204 \\ .5 & -.5 & -.5 & .5 \\ -.2236068 & .6708204 & -.6708204 & .2236068 \end{bmatrix}. \tag{40}$$

The reasoning behind this calculation is explained in Section 3. The rows of 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, e.g., Draper and Smith 1998, p. 466.) The three rows of 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 $x_1 + x_2 + x_3 + x_4 = .90$. The three eigenvalues of TBT' are $(-20.04, 2.52, 46.87)$, and the radius R becomes infinite when λ takes these eigenvalues. The ridge path of maximum \hat{y} (path A) is obtained given by choosing λ values from ∞ (where the solution will be $x = f$, and where $R = 0$) to 46.87 (where the solution will be $x = \infty$). The ridge of minimum \hat{y} (path F) will be given by choosing λ values from $-\infty$ (where the solution will be $x = f$) to -20.04 (where the solution will be $x = -\infty$). Other λ 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 λ (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 f , where $R = 0$ and $\hat{y} = 6.27$, and moves quickly (see the x_3 values) to the $x_3 = .08$ boundary and beyond, whereas 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 $x_3 = .08$ value, indicates that further exploration of the fitted surface needs to be carried out on the $x_3 = .08$ face of the restricted region.

Figure 2, derived from the path A 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 also be drawn for any of the ridge paths that 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, because it duplicates the information in the corresponding table. We recall that closed-form expressions for the dependency of x_i and \hat{y} on 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 $x_3 = .358$, well above the x_3 upper limit for the experimental region. The minimum R value of about .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 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 nonnegative. Table 2 gives selected λ values, to a point where the predicted \hat{y} has turned negative.

4.2 The Second Set of Ridge Paths

Because we are interested in maximizing \hat{y} , we now explore the surface on the $x_3 = .08$ 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 $x_3 = .08$. This means that (39) is replaced by

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad c = \begin{bmatrix} .45 \\ .08 \end{bmatrix}. \tag{41}$$

Table 2. Ridge Paths for the Anik and Sukumar (1981) Data, Applying Only the Mixture Restriction $x_1 + x_2 + x_3 + x_4 = .9$

Path	λ	x_1	x_2	x_3	x_4	R	\hat{y}
A (max)	∞	.210	.210	.040	.440	0	6.27
	2000	.209	.207	.048	.436	.010	6.64
	1000	.208	.204	.056	.432	.020	7.02
	750	.207	.202	.062	.429	.026	7.27
	500	.206	.199	.072	.423	.038	7.75
	400	.205	.196	.080	.419	.048	8.10
	300	.204	.191	.092	.413	.062	8.66
	250	.203	.187	.102	.408	.074	9.10
	100	.201	.152	.181	.366	.170	12.48
	62	.230	.107	.243	.320	.259	15.40
	50	.441	.020	.244	.195	.437	21.94
48	.920	-.131	.168	-.057	.940	55.58	
B, C, D, and E do not occur within the experimental region							
F (min)	-90	.248	.273	-.194	.573	.279	-6.26
	-100	.243	.266	-.165	.556	.244	-4.55
	-200	.224	.238	-.052	.490	.109	1.69
	-436	.216	.223	.000	.461	.048	4.32
	-500	.215	.221	.005	.459	.041	4.58
	-700	.213	.218	.016	.453	.029	5.08
	-900	.213	.216	.021	.450	.023	5.35
	$-\infty$.210	.210	.040	.440	.000	6.27

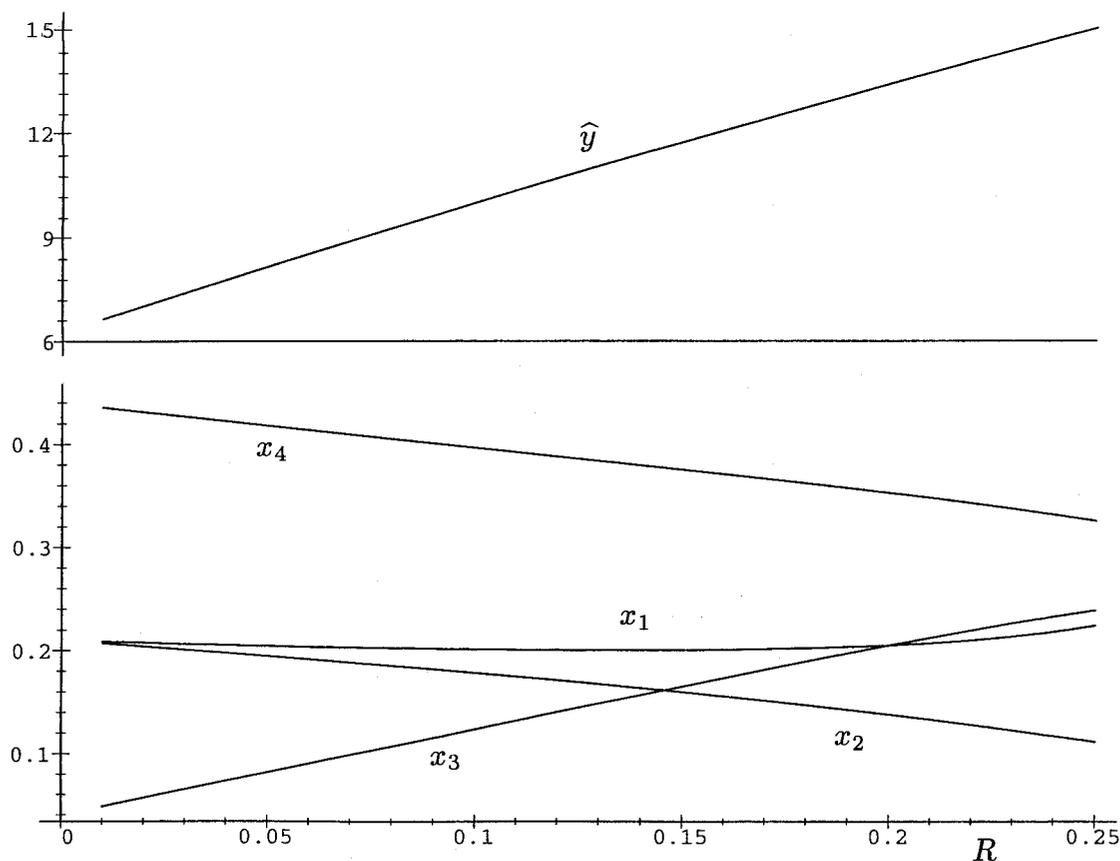


Figure 2. The Maximum Predicted Response \hat{y} and Its Corresponding Positional Coordinates (x_1, x_2, x_3, x_4) Plotted Against R , the Distance the Point Lies From the Focus $f = (.210, .210, .04, .440)$ in the Space $x_1 + x_2 + x_3 + x_4 = .90$. The numerical details are given in Table 2.

Repeating the calculations with these new restrictions requires a new T ,

$$T = \begin{bmatrix} .267261 & .534523 & 0 & -.801784 \\ .771517 & -.617213 & 0 & -.154303 \end{bmatrix} \quad (42)$$

which leads to the eigenvalues of TBT' of $(-.49, 45.01)$. [The two rows of 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 as 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 $\mathbf{f} = (.203, .203, .08, .413)'$, their centroid. This point lies at distances $R = .253, .241, \text{ and } .241$ from points 2, 4, and 6, 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 $x_3 = .08$ throughout, and we show the paths in Figure 3. The maximum \hat{y} path crosses the $x_4 = .30$ boundary when λ 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 .16 and then turns, crossing the $x_2 = .10$ boundary at roughly this $x_1 = .16$ level; see Figure 3.

4.3 The Third Set of Ridge Paths

To move along the boundary $x_4 = .30$, we designate a new focus \mathbf{f} and a new matrix T . The endpoints of the restricted region along the boundary are the corner points $(.40, .12, .08, .30)$, near design point 6 in Table 1, and $(.12, .40, .08, .30)$, near design point 4; see Figure 1. We choose their centroid, namely $\mathbf{f} = (.26, .26, .08, .30)'$. T is now a normalized row vector orthogonal to the rows of A in $A\mathbf{x} = \mathbf{c}$, namely

$$\begin{bmatrix} .5 & .5 & .5 & .5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} .45 \\ .08 \\ .30 \end{bmatrix}. \quad (43)$$

Necessarily, $T = (.707107, -.707107, 0, 0)$, or the vector with signs reversed. The sole eigenvalue of 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, $x_3 = .08$ and $x_4 = .30$, and so $x_1 + x_2 = .52$. 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 $x_1 = .40$ 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 $x_1 = .12$ when the second corner point is reached. (It is not the minimum region response, which we would find by exploring the $x_2 = .10$ boundary, choosing $\mathbf{f} = (.25, .10, .08, .47)'$ and $T = (.707107, 0, 0, -.707107)$ 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, in which case geometrical visualization is often difficult. In this article the technique is extended in a general way to mixture response surfaces. The focus from which the curved paths emanate can be freely chosen, and any linear equality restrictions, including the usual mixture restriction $x_1 + x_2 + \dots + x_q = 1$, 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 pseudocomponents, 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 because the path of maximum \hat{y} entered the $x_3 = .08$ face of the restricted region at $(.205, .196, .08, .419)$, 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 pathologic examples.

2. The formulas that 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, $b_0 = 0$ and $\mathbf{B} = \mathbf{0}$ in (19)–(24). The “eigenvalues of \mathbf{B} ” are all 0 and, by the choice of \mathbf{f} , $A\mathbf{f} = \mathbf{c}$. The solution reduces to

$$\mathbf{x} = \mathbf{f} + (2\lambda)^{-1}(\mathbf{I} - A'(AA')^{-1}A)\mathbf{b}. \quad (44)$$

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} , $A = 0$, 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 pseudocomponents, if desired. This would involve a preliminary transformation of the form $\mathbf{z} = \mathbf{u}\mathbf{x} - \mathbf{v}$, 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,

Table 3. Ridge Paths for the Anik and Sukumar (1981) Data, Under the Restrictions $x_3 = .08$ and $x_1 + x_2 + x_4 = .82$.

Path	λ	x_1	x_2	x_3	x_4	R	\hat{y}
A (max)	∞	.203	.203	.08	.414	0	8.12
	1000	.207	.203	.08	.410	.005	8.16
	500	.211	.202	.08	.407	.010	8.21
	200	.225	.200	.08	.395	.029	8.41
	150	.236	.197	.08	.387	.042	8.57
	125	.246	.194	.08	.380	.055	8.74
	100	.265	.189	.08	.366	.079	9.10
	90	.279	.184	.08	.357	.097	9.39
	80	.301	.177	.08	.342	.124	9.90
	75	.317	.171	.08	.332	.144	10.32
	70	.341	.162	.08	.317	.173	10.97
	66	.367	.152	.08	.301	.205	11.80
	65.95	.368	.152	.08	.300	.206	11.82
	60	.433	.127	.08	.260	.287	14.31
	55	.549	.081	.08	.190	.429	20.13
	52	.698	.021	.08	.101	.613	30.32
B and C do not occur within the region $x_3 = .08, x_1 + x_2 + x_4 = .82$							
D (min)	-6	.158	.028	.08	.634	.286	6.86
	-7	.156	.058	.08	.606	.245	7.00
	-9	.154	.098	.08	.568	.194	7.18
	-9.15	.154	.100	.08	.566	.191	7.19
	-10	.154	.111	.08	.555	.176	7.23
	-20	.156	.168	.08	.496	.101	7.51
	-30	.161	.184	.08	.475	.077	7.62
	-40	.166	.192	.08	.462	.063	7.69
	-50	.169	.196	.08	.455	.054	7.73
	-70	.175	.200	.08	.445	.043	7.80
	-100	.181	.202	.08	.437	.033	7.86
	-500	.197	.204	.08	.419	.008	8.05
	-1000	.200	.204	.08	.416	.004	8.08
	$-\infty$.203	.203	.08	.414	0	8.12

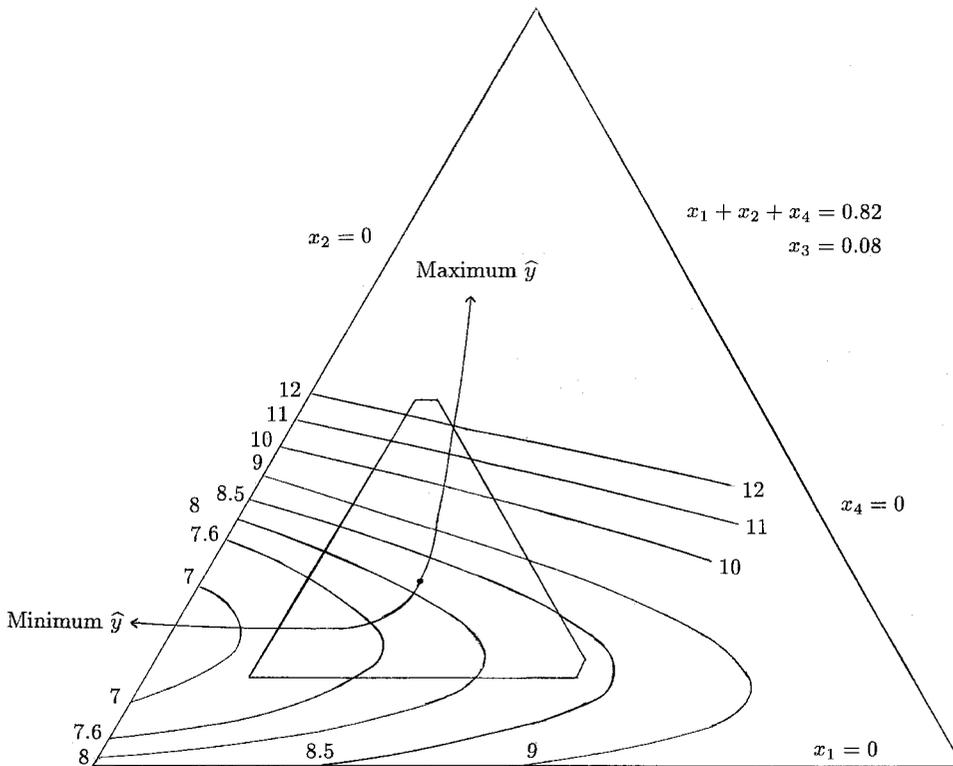


Figure 3. The Fitted Contours Defined by (3) When $x_3 = .08$ Shown in the Subspace $x_1 + x_2 + x_4 = .82$. The ridge paths of maximum \hat{y} and minimum \hat{y} on spheres of radius R emanate from the focus $\mathbf{f} = (.203, .203, .08, .413)'$; numerical details are given in Table 3. The pentagon is the inner one of Figure 1(b).

Table 4. Ridge Paths for the Anik and Sukumar (1981) Data, Under the Restrictions $x_3 = .08$, $x_4 = .30$, and $x_1 + x_2 = .52$

Path	λ	x_1	R	\hat{y}
A (max)	∞	.260	0	9.45
	1000	.264	.006	9.51
	500	.268	.012	9.58
	250	.278	.025	9.75
	100	.316	.079	10.51
	90	.325	.092	10.72
	80	.338	.110	11.03
	75	.346	.122	11.25
	70	.357	.137	11.53
	65	.371	.157	11.91
	60	.389	.182	12.45
	57.5	.400	.198	12.81
	B (min)	1.15	.120	.198
0		.125	.190	8.38
-10		.160	.142	8.45
-20		.180	.113	8.56
-40		.203	.081	8.74
-100		.229	.043	9.02
-200		.243	.024	9.19
-750		.255	.007	9.37
$-\infty$.260	0	9.45

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 (.40, .12, .08, .30), we calculated the predicted response values at all 10 true vertices of the restricted region. Among these 10 \hat{y} values, the second largest is 12.63 and occurs at the vertex (.40, .10, .08, .32), the vertex closest to the maximum. More generally, it would be possible to use the methods of this article 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, where contours could be drawn only in sections.

ACKNOWLEDGMENTS

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

[Received October 2001. Revised December 2001.]

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, 897-900.
- Cornell, J. A. (1990), *Experiments with Mixtures* (2nd ed.), New York: Wiley.
- Del Castillo, E., Fan, S. K., and Semple, J. (1997), "The Computation of Global Optima in Dual Response Systems," *Journal of Quality Technology*, 29, 347-353.
- (1999), "Optimization of Dual Response Systems: A Comprehensive Procedure for Degenerate and Nondegenerate Problems," *European Journal of Operations Research*, 112, 174-186.
- Draper, N. R. (1963), "Ridge Analysis of Response Surfaces," *Technometrics*, 5, 469-479.
- Draper, N. R., and Pukelsheim, F. (2000), "Ridge Analysis of Mixture Response Surfaces," *Statistics & Probability Letters*, 48, 131-140.
- Draper, N. R., and Smith, H. (1998), *Applied Regression Analysis* (3rd ed.), New York: Wiley.
- Hoerl, A. E. (1959), "Optimum Solution of Many Variables Equations," *Chemical Engineering Progress*, 55(11), 69-78.
- (1962), "Applications of Ridge Analysis to Regression Problems," *Chemical Engineering Progress*, 58(3), 54-59.
- (1964), "Ridge Analysis," *Chemical Engineering Progress Symposium Series*, 60, 67-77.
- Hoerl, R. W. (1985), "Ridge Analysis 25 Years Later," *The American Statistician*, 39, 186-192.
- (1987), "The Application of Ridge Techniques to Mixture Data: Ridge Analysis," *Technometrics*, 29, 161-172.
- McLean, R. A., and Anderson, V. L. (1966), "Extreme Vertices Design of Mixture Experiments," (with discussion), *Technometrics*, 8, 447-456.
- Myers, R. H., and Carter, W. H. Jr. (1973), "Response Surface Techniques for Dual Response Systems," *Technometrics*, 15, 301-317.
- Prescott, P., Dean, A. M., Draper, N. R., and Lewis, S. M. (2002), "Mixture Experiments: Ill-Conditioning and Quadratic Model Formulation," *Technometrics*, 44, 260-268.
- Scheffé, H. (1958), "Experiments with Mixtures," *Journal of the Royal Statistical Society, Ser. B*, 20, 344-360.
- (1963), "The Simplex-Centroid Design for Experiments With Mixtures," *Journal of the Royal Statistical Society, Ser. B*, 25, 235-263.
- Semple, J. (1997), "Optimality Conditions and Solutions Procedures for Nondegenerate Dual Response Systems," *Institute of Industrial Engineers Transactions*, 29, 743-752.