Ridge Analysis of Mixture Response Surfaces

Norman R. Draper[a,*], Friedrich Pukelsheim[b]

[a]University of Wisconsin, Madison, WI, 53706-1685, USA

[b]Universität Augsburg, D-86135 Augsburg, Germany

Abstract

Previous applications of ridge analysis to second order response surfaces based on mixture ingredients $(x_1, x_2, \ldots, x_q)$ which sum to 1 have required transforming from the origin $(0, 0, \ldots, 0)$, which is *not* in the mixture space, to a point inside the space, most often the centroid $(\frac{1}{q}, \frac{1}{q}, \ldots, \frac{1}{q})$. We show that this transformation is not necessary for tracking the maximum $\hat{y}$ and the minimum $\hat{y}$ paths. In addition, we show that the application of ridge analysis is somewhat simplified if the Scheffé model form is replaced by the K (Kronecker) model form, an alternative, homogeneous model given elsewhere.

AMS 1991 Subject classification: 62J05

Keywords: Kronecker model; Mixture data; Ridge analysis, Response surfaces; Scheffé model

1. Introduction

Ridge analysis was initially suggested by A.E. Hoerl (1959, 1962) in the context of fitted second order response surface models where the factors were not restricted. Some theoretical foundation for the method was later given by Draper (1963). See also A.E. Hoerl (1964) and R.W. Hoerl (1985). The basic method defines a series of paths outward from the origin $(x_1, x_2, \ldots, x_q) = (0, 0, \ldots, 0)$ of the factor space. Suppose the fitted second order surface is written as

$$\hat{y} = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} \tag{1}$$

where

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

and

$$\mathbf{B} = \begin{pmatrix} b_{11} & \frac{1}{2}b_{12} & \ldots & \frac{1}{2}b_{1q} \\ \frac{1}{2}b_{12} & b_{22} & \ldots & \frac{1}{2}b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}b_{1q} & \frac{1}{2}b_{2q} & \ldots & b_{qq} \end{pmatrix}. \tag{2}$$

Then (1) is the matrix format for the second order fitted equation

$$\begin{aligned} \hat{y} = b_0 + b_1x_1 + b_2x_2 + \ldots + b_qx_q + b_{11}x_1^2 + b_{22}x_2^2 + \ldots + b_{qq}x_q^2 \\ + b_{12}x_1x_2 + b_{13}x_1x_3 + \ldots + b_{q-1,q}x_{q-1}x_q. \end{aligned} \tag{3}$$

*Corresponding author.

The outward paths mentioned above are defined by imagining a sphere about the origin $\mathbf{x} = (0, 0, \ldots, 0)'$ of radius R, say. On such a sphere, we can find the point of maximum response and the point of minimum response. In general, there may also be points at which $\hat{y}$ has stationary values that are neither maxima nor minima. As R is increased from zero outwards, the loci of the maximum $\hat{y}$ and the minimum $\hat{y}$ can be followed out, thus giving us (second order) paths of steepest ascent and descent. Typically, the paths of the intermediate stationary values, if any exist, begin at non-zero values of R which depend on the specific response surface being analyzed.

2. Mixtures Ridge Analysis

Mixtures ridge analysis has been featured in only a few papers to date. Related references are Becker (1969), Cornell and Ott (1975), St.John (1984) and R.W. Hoerl (1987). In the last-mentioned paper, ridge analysis is done by first transforming the q mixture variables to $(q-1)$ orthogonal variables, essentially removing the mixture restriction, but also unbalancing the coordinate system. This makes ridge analysis somewhat awkward to apply, and also makes it necessary to transform any conclusions back into the mixture coordinates.

We now show how ridge analysis can be applied directly to mixtures applications in which $\mathbf{x}'\mathbf{1} = x_1 + x_2 + \ldots + x_q = 1$. For the moment, however, we shall continue to use the general form (1) and (2) for the response surface, particularizing to specific mixture forms when needed. Consider the Lagrangian function

$$F = b_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} - \lambda_1(\mathbf{x}'\mathbf{1} - 1) - \lambda_2(x'x - R^2). \tag{4}$$

When $\lambda_1 = 0$, we fall back to the "usual" second order ridge analysis Lagrangian function, as in A.E. Hoerl (1959, 1962) and Draper (1963). Differentiating (4) with respect to $\mathbf{x}$ (which can be achieved by differentiating (4) with respect to $x_1, x_2, \ldots, x_q$ in turn and rewriting these equations in matrix form) gives

$$\frac{\partial F}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} - \lambda_1\mathbf{1} - 2\lambda_2\mathbf{x}. \tag{5}$$

Setting (5) equal to zero leads to

$$2(\mathbf{B} - \lambda_2\mathbf{I})\mathbf{x} = -(\mathbf{b} - \lambda_1\mathbf{1}). \tag{6}$$

Superficially, it would seem that if $(\mathbf{B} - \lambda_2\mathbf{I})^{-1}$ exists, which will happen as long as λ_2 is not an eigenvalue of $\mathbf{B}$, we obtain solutions for all the stationary points of $\hat{y}$ on the sphere of radius R from the q equations

$$\mathbf{x} = -\frac{1}{2}(\mathbf{B} - \lambda_2\mathbf{I})^{-1}(\mathbf{b} - \lambda_1\mathbf{1}). \tag{7}$$

In fact, however, solutions exist via (6) or (7) for the mixtures problem even *at* those eigenvalues, in general. This is because we must add, to the q equations of (7), the two restrictions

$$\mathbf{1}'\mathbf{x} \equiv \mathbf{x}'\mathbf{1} \equiv x_1 + x_2 + \ldots + x_q = 1 \tag{8}$$

which ensures that the solution lies in the mixture subspace, and

$$\mathbf{x}'\mathbf{x} \equiv x_1^2 + x_2^2 + \ldots + x_q^2 = R^2, \tag{9}$$

which means that the solution is also on a sphere of radius R.

Moreover, for mixtures models, $b_0, \mathbf{B}$ and $\mathbf{b}$ can take only certain forms, as will be explained in Section 4. As in ordinary ridge regression, we could in theory fix R, substitute from (8) and (9) into (7) and then solve (7); it is far simpler to first select a value for λ_2 in (7), however, whereupon we can apply (8) to obtain

$$1 = \mathbf{1}'\mathbf{x} = -\frac{1}{2}\mathbf{1}'(\mathbf{B} - \lambda_2\mathbf{I})^{-1}(\mathbf{b} - \lambda_1\mathbf{1}) \tag{10}$$

which implies that

$$1 + \frac{1}{2}\mathbf{1}'(\mathbf{B} - \lambda_2\mathbf{I})^{-1}\mathbf{b} = \frac{1}{2}\lambda_1\mathbf{1}'(\mathbf{B} - \lambda_2\mathbf{I})^{-1}\mathbf{1} \tag{11}$$

so that

$$\lambda_1 = \frac{1 + \frac{1}{2}\mathbf{1}'(\mathbf{B} - \lambda_2\mathbf{I})^{-1}\mathbf{b}}{\frac{1}{2}\mathbf{1}'(\mathbf{B} - \lambda_2\mathbf{I})^{-1}\mathbf{1}}. \tag{12}$$

We can now in general invoke (7) to give a value for $\mathbf{x}$ for the particular combination of λ_2 (chosen) and λ_1 (from (12)) and evaluate $R^2 = \mathbf{x}'\mathbf{x}$ from (9) and $\hat{y}$ from (1) or (3). We thus have a point on one of the stationary paths defined by $(\lambda_2, \lambda_1, \mathbf{x}, R, \hat{y})$. Furthermore, all such points satisfy (8).

Note that from (5) and (6), it is apparent that when we set $\lambda_2 = 0$, whereupon from (12),

$$\lambda_1 = \frac{1 + \frac{1}{2}\mathbf{1}'\mathbf{B}^{-1}\mathbf{b}}{\frac{1}{2}\mathbf{1}'\mathbf{B}^{-1}\mathbf{1}}, \tag{13}$$

Eqn. (7) will deliver the stationary point on the mixture space.

How do we know which path we are on — overall maximum $\hat{y}$ on (9), overall minimum $\hat{y}$ on (9), or intermediate stationary values, such as local maxima or minima? The usual matrix of second derivatives

$$\left\{\frac{\partial^2 F}{\partial x_i \partial x_j}\right\} = 2(\mathbf{B} - \lambda_2\mathbf{I}) \tag{14}$$

is not appropriate here, because it does not reflect the fact that the solution has to be on the mixture space. Consider the $(q-1)$ by q matrix

$$\mathbf{T}_q = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1q} \\ t_{21} & t_{22} & \dots & t_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ t_{q-1,1} & t_{q-1,2} & \dots & t_{q-1,q} \end{pmatrix} \tag{15}$$

say, where each row consists of a vector of orthogonal polynomial coefficients, normalized so that the row sum of squares is 1. For $q = 3$ and 4, needed for our examples later, we have

$$\mathbf{T}_3 = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix}, \tag{16}$$

$$\mathbf{T}_4 = \begin{pmatrix} -\frac{3}{\sqrt{20}} & -\frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{3}{\sqrt{20}} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{\sqrt{20}} & \frac{3}{\sqrt{20}} & -\frac{3}{\sqrt{20}} & \frac{1}{\sqrt{20}} \end{pmatrix}. \tag{17}$$

[See, for example, Draper and Smith (1998, p 466).]

The addition of a last row $\mathbf{u}_q' = (\frac{1}{\sqrt{q}}, \frac{1}{\sqrt{q}}, \dots, \frac{1}{\sqrt{q}}) = \frac{1}{\sqrt{q}}\mathbf{1}'$ converts $\mathbf{T}_q$ into a transformation matrix $\mathbf{U}_q = (\mathbf{T}_q', \mathbf{u}_q)'$ which allows the $x_1, x_2, \dots, x_q$ coordinates to be converted into values $z_1, z_2, \dots, z_{q-1}, 1/\sqrt{q}$, via $\mathbf{z} = \mathbf{U}_q\mathbf{x}$ or $\mathbf{x} = \mathbf{U}_q^{-1}\mathbf{z} = \mathbf{U}_q'(z_1, z_2, \dots, z_{q-1}, 1/\sqrt{q})'$, due to the orthogonality of $\mathbf{U}_q$. It follows that, if we define $\mathbf{w} = (z_1, z_2, \dots, z_{q-1})'$, so that $\mathbf{z} = (\mathbf{w}', 1/\sqrt{q})'$,

$$\mathbf{x}'\mathbf{B}\mathbf{x} = \mathbf{z}'\mathbf{U}_q\mathbf{B}\mathbf{U}_q'\mathbf{z} = \mathbf{w}'\mathbf{T}_q\mathbf{B}\mathbf{T}_q'\mathbf{w} + \frac{2}{\sqrt{q}}\mathbf{u}_q'\mathbf{B}\mathbf{T}_q'\mathbf{w} + \frac{1}{q}\mathbf{u}_q'\mathbf{B}\mathbf{u}_q. \tag{18}$$

This means that we can replace (14), after differentiating (18) twice with respect to $\mathbf{w}$, by

$$2(\mathbf{T}_q\mathbf{B}\mathbf{T}_q' - \lambda_2\mathbf{I}). \tag{19}$$

Note that the size of this square matrix is $(q-1)$ not q because $\mathbf{T}_q$ is $(q-1)\times q$. We see that, if (19) is positive definite, we have a minimum, while if (19) is negative definite, we have a maximum. If (19) 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 $\mathbf{T}_q\mathbf{B}\mathbf{T}_q'$ are $\mu_1 \le \mu_2 \le \dots \le \mu_{q-1}$, arranged in order with due regard to sign then on the mixture space:

(a) Choosing $\lambda_2 > \mu_{q-1}$ provides a locus of maximum $\hat{y}$ as R changes, and

(b) Choosing $\lambda_2 < \mu_1$ provides a locus of minimum $\hat{y}$ as R changes.

(c) Choosing $\mu_1 \leq \lambda_2 \leq \mu_{q-1}$ gives intermediate stationary values.

As in the non-mixture case, when $\lambda_2 = \mu_i$ exactly for $i = 1, 2, \ldots, q-1$, 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 since choosing λ_2 values decreasing from ∞ gives the path of maximum $\hat{y}$, while using values increasing from $-\infty$ gives the path of minimum $\hat{y}$.

3. The Geometry

To see why the solution works *without* moving to an origin on the $\mathbf{1}'\mathbf{x} = 1$ plane, we show geometrically the simplest cases; see Figure 1. Imagine a sphere $x_1^2 + x_2^2 + x_3^2 = R^2$ centered at the origin O in Figure 1(b). When $R < 1/\sqrt{q}$, the sphere will not intersect the mixture space $x_1 + x_2 + x_3 = 1$ so there will be no solutions to (7). When $R = 1/\sqrt{q}$, the sphere just touches the mixtures centroid $(1/\sqrt{q}, 1/\sqrt{q}, \ldots, 1/\sqrt{q}) = (1/\sqrt{q})\mathbf{1}$, which will thus be the only solution point $\mathbf{x}$ of Eqns (7), (8), and (9). It can easily be confirmed that, for this solution, $\lambda_2 = \infty, \lambda_1 = -\infty, x = (1/\sqrt{q})\mathbf{1}$ and $R = 1/\sqrt{q}$. [Or $\lambda_2 = -\infty, \lambda_2 = \infty$, etc.] When $R > 1/\sqrt{q}$, the sphere will intersect the plane $x_1 + x_2 + x_3 = 1$ in a circle centered at the mixture centroid, which is exactly the way we wish to apply ridge analysis in this $q = 3$ mixture space. [Figure 1(a) shows the more elementary $q = 2$ case where the "spheres" are now circles and the "circles" are now pairs of points equally spaced on the line $x_1 + x_2 = 1$ around the centroid $(\frac{1}{2}, \frac{1}{2})$.] For $q \geq 4$, the picture of Figure 1(b) must be mentally extended to higher dimensions. For $q = 4$, for example, the "spheres" cannot be drawn and the "circles" are spheres around the centroid of a pyramid.

Figure 1 about here.

---

4. Choice of Mixture Model.

In actually carrying out the calculations (7) for selected λ_2, and with λ_1 derived from (12) we have to specify the particular form of second order mixture model to fit. Most readers would probably choose the Scheffé model (S-model) which, for second order, consists only of terms in x_i and in $x_i x_j$. In such a case $\mathbf{B}$ in (2) has all diagonal terms zero, while $\mathbf{b}$ is, in general non-zero. Certainly the calculations offer no difficulty if carried out in this form.

A more interesting possibility, we suggest, is to employ the K-model where K stands for Kronecker. This involves using no x_i terms, replacing them by x_i^2 terms to obtain

$$\hat{y} = b_{11}x_1^2 + b_{22}x_2^2 + \ldots + b_{qq}x_q^2 + b_{12}x_1x_2 + b_{13}x_1x_3 + \ldots + b_{q-1,q}x_{q-1}x_q. \qquad (20)$$

This form has been suggested by Draper and Pukelsheim (1998); a comparison between the second order S- and K-models has been provided therein. The advantage in the ridge analysis formulation is that while $\mathbf{B}$ now contains diagonal terms, $\mathbf{b} = \mathbf{0}$. Thus (12) becomes

$$\lambda_1 = 2\left\{\mathbf{1}'(\mathbf{B}-\lambda_2\mathbf{I})^{-1}\mathbf{1}\right\}^{-1} \tag{21}$$

whereupon (7) reduces to

$$\mathbf{x} = \left\{\mathbf{1}'(\mathbf{B}-\lambda_2\mathbf{I})^{-1}\mathbf{1}\right\}^{-1}(\mathbf{B}-\lambda_2\mathbf{I})^{-1}\mathbf{1}. \tag{22}$$

Our examples will be analyzed using the K-approach; either approach gives the same numerical solutions, of course. The proof of this follows from the fact that $\mathbf{B}_S + \frac{1}{2}\mathbf{b}\mathbf{1}' + \frac{1}{2}\mathbf{1}\mathbf{b}' = \mathbf{B}_K$, where subscript S denotes the Scheffé form of $\mathbf{B}$ and subscript K denotes the Kronecker form of $\mathbf{b}$. We can now premultiply both sides by $\mathbf{x}'$, postmultiply both sides of the result by $\mathbf{x}$, set $\mathbf{1}'\mathbf{x} = \mathbf{x}'\mathbf{1} = 1$ and we obtain $\mathbf{x}'\mathbf{B}_S\mathbf{x} + \frac{1}{2}(\mathbf{x}'\mathbf{b} + \mathbf{b}'\mathbf{x}) = \mathbf{x}'\mathbf{B}_K\mathbf{x}$, the bracketed terms being identical and reducing to $b_1x_1 + \ldots + b_qx_q$.

All solutions $\mathbf{x}$ will be on the subspace $\mathbf{1}'\mathbf{x} = 1$, but depending on the λ_2 value chosen, some points will have coordinates that exceed 1 or that are negative. Since the mixture space is such that $0 \le x_i \le 1$, solutions that violate these reductions are not relevant. We shall discuss this in our examples.

5. Examples

Example 1. [Kurotori, 1966; see also Draper and Smith, 1998, pp. 418–419.] A set of 10 experimental runs was performed on a propellant problem with three ingredients ($q = 3$). In the original experiment, a restricted region with $x_1 \ge 0.2$, $x_2 \ge 0.4$, and $x_3 \ge 0.2$ was explored, and the best predictions were found to be on the $x_1 = 0.20$ boundary. In the present paper we do not restrict the surface to the smaller region, but follow the maximum $\hat{y}$ path from the centroid (which is outside the region explored) into the restricted region. Note, in this regard, that the ridge paths may exist mathematically even when they might not be practically relevant.

Fitting the surface in K-model form, we obtain the following equation

$$\hat{y} = -2.732\,x_1^2 - 3.340\,x_2^2 - 17.259\,x_3^2 + 3.249\,x_1x_2 + 14.694\,x_1x_3 + 28.813\,x_2x_3$$

whereupon

$$\mathbf{B} = \begin{pmatrix} -2.732 & 1.624 & 7.347 \\ 1.624 & -3.340 & 14.406 \\ 7.347 & 14.406 & -17.259 \end{pmatrix}$$

with eigenvalues $(-27.50, -4.24, 8.40)$. However, these eigenvalues are not the ones that affect movement on the mixture surface. With $\mathbf{T}_3$ defined as in (16) we find the eigenvalues

Table 1. Ridge path (x_1, x_2, x_3) of maximum $\hat{y}$ as λ_2 decreases and R increases, Kurotori data

λ_2	x_1	x_2	x_3	R	$\hat{y}$
∞	0.333	0.333*	0.333	0.577	2.603
50	0.321	0.359*	0.320	0.578	2.714
20	0.302	0.384*	0.314	0.581	2.800
10	0.276	0.412	0.312	0.586	2.883
9	0.271	0.417	0.312	0.587	2.900
7	0.258	0.429	0.313	0.590	2.928
5	0.239	0.447	0.314	0.596	2.969
3	0.208	0.474	0.318	0.608	3.021
1	0.152*	0.522	0.326	0.634	3.082
0	0.101*	0.564	0.335	0.664	3.099
−1	0.015*	0.635	0.350	0.725	3.050
−2	−0.165†	0.781	0.383	0.886	2.635

Restrictions in the original data set: $x_1 \geq 0.2$, $x_2 \geq 0.4$, $x_3 \geq 0.2$

* Point lies outside the original restricted data space.

† Point lies outside the main simplex.

of $\mathbf{T}_3\mathbf{B}\mathbf{T}_3'$ to be $(-27.28, -3.86)$. The path for the maximum $\hat{y}$ will thus be mapped out for values of $\lambda \geq -3.86$. [Because this problem is not well conditioned, slightly different numbers may be obtained by different programs.]

Table 1 shows some selected calculations for this path, moving out from the centroid $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. We see that the path enters the restricted subspace across the $x_2 = 0.40$ boundary, and exits it across the $x_1 = 0.20$ boundary later. The maximum predicted $\hat{y}$ in or on the restricted subspace is at about (0.20, 0.48, 0.32), close to the $\lambda_2 = 3$ entry of Table 1.

In examining the path, we can ignore points which violate the conditions of the practical problem. Note that a path could in theory pass outside the mixture space (or a defined restricted sub-region of it) and then come back in. The practical interpretation of such behavior would be to follow the path to the border and then move along the border until the path returned to the valid part of the mixture region.

Figures 2 and 3 are helpful in understanding the calculations made for Table 1. Figure 2 shows how the radius R varies with λ_2, leaping to infinity at the eigenvalues -27.28 and -3.86 of $\mathbf{T}_3\mathbf{B}\mathbf{T}_3'$. Table 1 corresponds to the upper part of Figure 2 only, of course; this upper part is smooth and so not shown. Since the sphere of radius R passes outside the mixture space when $R > 1$, only the lower portion of Figure 2 is meaningful in practice.

The radius $R \geq 1/\sqrt{q}$ always [here 0.577]. When $\lambda_2 = -\infty$ or ∞, we are at the centroid exactly. The two separate curves forming the U-shaped curve between the eigenvalues may (or may not) go as low as $1/\sqrt{q}$ [here they do not]. Compare with a similar diagram in Draper (1963).

Figures 2 and 3 about here.

Figure 3 shows how λ_1 varies as a function of λ_2; as λ_2 increases, λ_1 basically decreases but there are three sections in the plot ($q-1$ in general) with divisions at the eigenvalues of $\mathbf{T}_3\mathbf{B}\mathbf{T}_3'$. At each eigenvalue, the curve passes instantaneously from $-\infty$ to ∞. [We note again that the eigenvalues of $\mathbf{B}$ itself are not relevant to these calculations.]

Example 2. [Draper et al., 1993; Draper and Smith, 1998, pp. 419–422.] The data consist of the 36 observations in four blocks. A second order model with three added blocking variables was used.

Fitting the surface in K-model form, we get

$$\begin{aligned}\hat{y} = &-14.89B_1 - 21.78B_2 - 20.11B_3\\ &+ 400.40x_1^2 + 449.32x_2^2 + 398.90x_3^2 + 403.49x_4^2\\ &+ 946.17x_1x_2 + 988.86x_1x_3 + 954.56x_1x_4\\ &+ 811.33x_2x_3 + 823.69x_2x_4 + 746.39x_3x_4.\end{aligned}$$

The first three terms do not depend on the x's and do not contribute to the ridge analysis except for the fitted value calculations. For the purposes of this example, we simply choose to omit them, i.e., set $B_1 = B_2 = B_3 = 0$, because the ridge calculations are not thus affected [beyond $\hat{y}$ levels]. Now

$$\mathbf{B} = \begin{pmatrix} 400.403 & 473.083 & 494.431 & 477.278\\ 473.083 & 449.319 & 405.667 & 411.847\\ 494.431 & 405.667 & 398.903 & 373.194\\ 477.278 & 411.847 & 373.194 & 403.486 \end{pmatrix}$$

with eigenvalues $(-128.9, 16.62, 30.81, 1733.6)$. These eigenvalues are not needed. The eigenvalues of $\mathbf{T}_4\mathbf{B}\mathbf{T}_4'$ where $\mathbf{T}_4$ is given by (17), are $(-126.3, 16.78, 30.81)$. Figure 4, parallel to Figure 2 but for Example 2, shows that R goes to ∞ at these eigenvalues. For the path of maximum predicted response, we need λ_2 values for which $30.81 \leq \lambda_2 \leq \infty$. Table 2 shows some selected calculations for this path, moving out from the centroid $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, until the path goes outside the mixture space [on the last line].

Figures 4 and 5 about here.

Table 2. Ridge path of maximum $\hat{y}$ as R increases. Bread data.

λ_2	x_1	x_2	x_3	x_4	R	$\hat{y}$
∞	0.250	0.250	0.250	0.250	0.500	433
600	0.289	0.255	0.230	0.226	0.503	436
200	0.335	0.272	0.205	0.187	0.514	440
100	0.369	0.308	0.187	0.136	0.533	442
50	0.389	0.414	0.182	0.016	0.596	447
48	0.389	0.425	0.184	0.003	0.605	448
46	0.389	0.437	0.186	0.013*	0.614	448

* Point lies outside the mixture region.

For both Examples 1 and 2, diagrams of the response contours appear in Draper and Smith (1998, pp. 419 and 421). In the case of the second of these diagrams, the effects of three non-significant terms have been omitted in the equation used, but this does not have a material effect on the contours drawn. [The ridge analysis could be redone in terms of the reduced model with very similar results.] In both Examples 1 and 2, the ridge paths obtained are consistent with the diagrams.

Acknowledgments

We are grateful to the Alexander von Humboldt-Stiftung for support through a Max-Planck-Award for cooperative research.

References

Becker, N.G. (1969), Regression problems when the predictor variables are proportions, *J. Roy. Statist. Soc. Ser. B* **31,** 107–112.

Cornell, J.A. and Ott, L. (1975), The use of gradients to aid in the interpretation of mixture response surfaces, *Technometrics* **17,** 409–424.

Draper, N.R. (1963), Ridge analysis of response surfaces, *Technometrics* **5,** 469–479.

Draper, N.R. and Pukelsheim, F. (1998), Mixture models based on homogeneous polynomials, *J. Statist. Plann. Inference* **71,** 303–311.

Draper, N.R., Prescott, P., Lewis, S.M., Dean, A.M., John, P.W.M. and Tuck, M.G. (1993), Mixture designs for four components in orthogonal blocks, *Technometrics* **35,** 268–276. (1995) **37** 131–132. See also *Technometrics* **37** (1995), 131–132.

Draper, N.R. and Smith, H. (1998), *Applied Regression Analysis, Third Edition* (Wiley, New York).

Hoerl, A.E. (1959), Optimum solution of many variables equations, *Chem. Eng. Progress* **55**(11), 69–78.

Hoerl, A.E. (1962), Applications of ridge analysis to regression problems, *Chem. Eng. Progress* **58**(3), 54–59.

Hoerl, A.E. (1964), Ridge analysis, *Chem. Eng. Progress Symp. Ser.* **60,** 67–77.

Hoerl, R.W. (1985), Ridge analysis 25 years later, *Amer. Statist.* **39,** 186–192.

Hoerl, R.W. (1987), The application of ridge techniques to mixture data: ridge analysis, *Technometrics* **29,** 161–172.

Kurotori, I.S. (1966), Experiments with mixtures of components having lower bounds, *Indust. Qual. Control* **22,** 592–596.

St.John, R.C. (1984), Experiments with mixtures, ill-conditioning, and ridge regression, *J. Qual. Technology* **16,** 81–87.

Figure captions

Figure 1. Ridge analysis geometry in the mixtures case for (a) $q = 2$, (b) $q = 3$.

Figure 2. How R varies with λ_2. Kurotori data.

Figure 3. How λ_1 varies with λ_2. Kurotori data.

Figure 4. How R varies with λ_2. Bread data.