Canonical reduction of second-order fitted models subject to linear restrictions

Norman R. Draper, Friedrich Pukelsheim

Angaben zur Veröffentlichung / Publication details:

Draper, Norman R., and Friedrich Pukelsheim. 2003. "Canonical reduction of second-order fitted models subject to linear restrictions." *Statistics & Probability Letters* 63 (4): 401–10. https://doi.org/10.1016/S0167-7152(03)00119-6.

Canonical reduction of second-order fitted models subject to linear restrictions

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

[a] *Department of Statistics, University of Wisconsin, 1210 West Dayton Street, Madison, WI 53706-1685, USA*
[b] *Universität Augsburg, D-86135 Augsburg, Germany*

Abstract

Canonical reduction of second-order response surfaces is a useful technique for finding the formand shape of surfaces and often for discovering redundancies that enable the surface to be expressible in a simpler formwith fewer canonical predictor variables than there are original predictor variables. Canonical reduction of models subject to linear restrictions has received little attention, possibly due to the apparent difficulty of performing it. An important special application is when the predictor variables are mixture ingredients that must sum to a constant; other linear restrictions may also be encountered in such problems. A possible difficulty in interpretation is that the stationary point may fall outside the permissible restricted space. Here, techniques for performing such a canonical reduction are given, and two mixture examples in the literature are re-examined, and canonically reduced, to illustrate what canonical reduction can and cannot provide.

1. Introduction

Second-order (also called second degree) response surfaces, which can assume various forms (see Box and Draper, 1987, pp. 346–355), are frequently fitted to data arising fromindustrial and research experiments, and it is then necessary to examine exactly what form of surface has been attained. For two or three factors, this can best be done by viewing surface plots, and for more dimensions, cross-sectional plots in two or three dimensions could be drawn. However, a more general tool, which will also reveal informative surface redundancies, is canonical analysis, also called canonical reduction. Canonical reduction is a method of rewriting a fitted second-order equation in a form

* Corresponding author. Tel.: +1-608-262-0926; fax: +1-608-262-0032.
E-mail address: draper@stat.wisc.edu (N.R. Draper).

in which it can be more readily understood. There are two main stages. First, a specific rotation of axes that removes all cross-product terms produces the so-called A canonical form. Then an appropriate change of origin removes all first-order terms as well and gives rise to the so-called B canonical form. This can be easily interpreted to understand the main features of the fitted surface; for specific examples, see Box and Draper (1987, pp. 332–355). In Section 2, we briefly recapitulate the implementation of canonical analysis when there are no restrictions. Sections 3 and 4 set out the modifications needed when linear restrictions on the predictor variables must be observed. These linear restrictions include, but are not limited to, the mixture restriction that applies in standard types of experiments with mixtures of ingredients adding (usually) to 1; see Cornell (2002). Sections 5 and 6 apply the work of Sections 3 and 4 to two abbreviated worked examples from the literature.

2. Canonical reduction with no restrictions

The general fitted second-order response surface can be written as

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

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

$$\mathbf{B} = \begin{bmatrix} b_{11} & \frac{1}{2}b_{12} & \cdots & \frac{1}{2}b_{1q} \\ \cdots & b_{22} & \cdots & \frac{1}{2}b_{2q} \\ \cdots & \cdots & \cdots & \cdots \\ \text{sym} & \cdots & \cdots & b_{qq} \end{bmatrix} \tag{2}$$

is a symmetric matrix. Then (1) is the matrix format for the second-order fitted equation

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

Differentiation of the fitted equation (1) with respect to $\mathbf{x}$ and setting the result equal to a zero vector leads to $\mathbf{b} + 2\mathbf{B}\mathbf{x} = \mathbf{0}$, and the solution,

$$\mathbf{x}_s = -\tfrac{1}{2}\mathbf{B}^{-1}\mathbf{b} \tag{4}$$

defines the stationary point of (1). This is the center of the quadratic system, also.

Suppose $\lambda_1, \lambda_2, \ldots, \lambda_q$ are the eigenvalues of $\mathbf{B}$ and $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_q$ are corresponding normalized and mutually orthogonal eigenvectors. Then the matrix $\mathbf{M} = \{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_q\}$ has $\mathbf{M}' = \mathbf{M}^{-1}$. Because the eigenvalues and vectors are defined by $\mathbf{B}\mathbf{m}_i = \mathbf{m}_i\lambda_i, i = 1, 2, \ldots, q$, we can write $\mathbf{B}\mathbf{M} = \mathbf{M}\Lambda$, where $\Lambda = \textbf{diagonal}(\lambda_1, \lambda_2, \ldots, \lambda_q)$. Premultiplying by $\mathbf{M}' = \mathbf{M}^{-1}$ leads to $\mathbf{M}'\mathbf{B}\mathbf{M} = \Lambda$. Using this result, and by insertion of $\mathbf{M}\mathbf{M}' = \mathbf{I}$ into (1), we obtain

$$\hat{y} = b_0 + (\mathbf{x}'\mathbf{M})(\mathbf{M}'\mathbf{b}) + (\mathbf{x}'\mathbf{M})\mathbf{M}'\mathbf{B}\mathbf{M}(\mathbf{M}'\mathbf{x}). \tag{5}$$

We can now let $\mathbf{X} = \mathbf{M}'\mathbf{x}$ and $\theta = \mathbf{M}'\mathbf{b}$ (or equivalently, $\mathbf{x} = \mathbf{M}\mathbf{X}$ and $\mathbf{b} = \mathbf{M}\theta$) and express (5) as

$$\hat{y} = b_0 + \mathbf{X}'\theta + \mathbf{X}'\Lambda\mathbf{X}. \tag{6}$$

This constitutes the "A canonical form" in which the axes have been rotated to remove cross-product terms, but the new variables $\mathbf{X}$ are measured from the original origin.

The "B canonical form" is obtained by moving the origin to the stationary point (4), by substituting new variables $\mathbf{W} = \mathbf{X} - \mathbf{M}'\mathbf{x}_s = \mathbf{X} - \mathbf{M}'(-\frac{1}{2}\mathbf{B}^{-1}\mathbf{b})$, which implies that $\mathbf{X} = \mathbf{W} - \frac{1}{2}\mathbf{M}'\mathbf{B}^{-1}\mathbf{b}$. After some algebra, this leads to

$$\hat{y} = b_0 - \tfrac{1}{4}\mathbf{b}'\mathbf{B}^{-1}\mathbf{b} + \mathbf{W}'\Lambda\mathbf{W}$$

or

$$\hat{y} = \hat{y}_s + \mathbf{W}'\Lambda\mathbf{W}, \tag{7}$$

where

$$\hat{y}_s = b_0 + \tfrac{1}{2}\mathbf{b}'\mathbf{x}_s = b_0 - \tfrac{1}{4}\mathbf{b}'\mathbf{B}^{-1}\mathbf{b} \tag{8}$$

is the predicted response at the stationary point. Eq. (7) is the B canonical form, most frequently employed in practice. For examples, see Box and Draper (1987, Chapters 10–11).

3. Adding linear restrictions

Suppose we wish to performcanonical reduction subject to a set of linear restrictions of the form

$$\mathbf{A}\mathbf{x} = \mathbf{c}, \tag{9}$$

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 where $\mathbf{c}$ is a given $m \times 1$ vector. For example, suppose we were investigating a mixture problem with ingredients $x_1, x_2, \ldots, x_q$ restricted by

$$\mathbf{1}'\mathbf{x} = \mathbf{x}'\mathbf{1} = x_1 + x_2 + \cdots + x_q = 1, \tag{10}$$

where $\mathbf{1}' = (1, 1, \ldots, 1)$, a $1 \times q$ vector. We could choose

$$\mathbf{A} = \frac{1}{q^{1/2}}(1, 1, \ldots, 1) \quad \text{and} \quad \mathbf{c} = \frac{1}{q^{1/2}}. \tag{11}$$

If this mixture space were further restricted to the plane

$$(\alpha_1, \alpha_2, \ldots, \alpha_q)\mathbf{x} = \alpha, \tag{12}$$

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

$$\mathbf{A} = \begin{bmatrix} \frac{1}{q^{1/2}} & \frac{1}{q^{1/2}} & \cdots & \frac{1}{q^{1/2}} \\ \alpha_1 & \alpha_2 & \cdots & \alpha_q \end{bmatrix} \quad \text{and} \quad \mathbf{c} = \begin{bmatrix} \frac{1}{q^{1/2}} \\ \alpha \end{bmatrix} \tag{13}$$

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 with ingredients adding to 1. The dimension m of $\mathbf{A}$ must be such that $m < q$ in general. When $m = q$, the restrictions define a single point in the x-space.

4. Canonical reduction under linear restrictions

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

$$\begin{aligned}&\mathbf{TA}'=\mathbf{0}\quad(\text{of size }(q-m)\times m),\\&\mathbf{AT}'=\mathbf{0}\quad(\text{of size }m\times(q-m)),\\&\mathbf{TT}'=\mathbf{I}_{q-m}\quad(\text{of size }(q-m)\times(q-m)).\end{aligned}\tag{14}$$

The combined matrix

$$\mathbf{Q}=\begin{bmatrix}\mathbf{A}\\\mathbf{T}\end{bmatrix}\tag{15}$$

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

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

$$\mathbf{z}=\begin{bmatrix}\mathbf{u}\\\mathbf{v}\end{bmatrix}=\mathbf{Qx}=\begin{bmatrix}\mathbf{A}\\\mathbf{T}\end{bmatrix}\mathbf{x}=\begin{bmatrix}\mathbf{Ax}\\\mathbf{Tx}\end{bmatrix}=\begin{bmatrix}\mathbf{c}\\\mathbf{Tx}\end{bmatrix}\tag{16}$$

under restrictions (9). Thus, the transformation fixes the first m coordinates at the desired restricted values, but leaves free the remaining $q-m$ coordinates which specify points in the space restricted by $\mathbf{Ax}=\mathbf{c}$. Consider the inverse of $\mathbf{Q}$. This is of the form

$$\mathbf{Q}^{-1}=\left[\mathbf{A}'(\mathbf{AA}')^{-1},\mathbf{T}'\right].\tag{17}$$

$\mathbf{AA}'$ is non-singular because of our assumption below (9) that the restrictions are linearly independent. It is easy to verify that $\mathbf{QQ}^{-1}=\mathbf{I}_q$ because of conditions (14). It follows that $\mathbf{Q}^{-1}\mathbf{Q}=\mathbf{I}$ also, because the inverse is unique. The quadratic portion of the fitted model function (1) is thus, using $\mathbf{x}=\mathbf{Q}^{-1}\mathbf{z}$, with $\mathbf{z}$ from(16) and $\mathbf{Q}^{-1}$ from(17),

$$\mathbf{x}'\mathbf{Bx}=\mathbf{z}'(\mathbf{Q}^{-1})'\mathbf{BQ}^{-1}\mathbf{z}=[\mathbf{c}',\mathbf{v}']\begin{bmatrix}(\mathbf{AA}')^{-1}\mathbf{A}\\\mathbf{T}\end{bmatrix}\mathbf{B}[\mathbf{A}'(\mathbf{AA}')^{-1},\mathbf{T}']\begin{bmatrix}\mathbf{c}\\\mathbf{v}\end{bmatrix}\tag{18}$$

$$=[\mathbf{c}'(\mathbf{AA}')^{-1}\mathbf{A}+\mathbf{v}'\mathbf{T}]\mathbf{B}[\mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{c}+\mathbf{T}'\mathbf{v}]\tag{19}$$

$$=\mathbf{v}'\mathbf{T}\,\mathbf{B}\,\mathbf{T}'\mathbf{v}+2\mathbf{v}'\mathbf{TBA}'(\mathbf{AA}')^{-1}\mathbf{c}+\mathbf{c}'(\mathbf{AA}')^{-1}\mathbf{AB}\,\mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{c},\tag{20}$$

after reduction. This gives us the quadratic portion of $\hat{y}$ in terms of $\mathbf{v}$ and $\mathbf{c}$. For the linear portion of $\hat{y}$, we obtain

$$\mathbf{x}'\mathbf{b} = \mathbf{z}'(\mathbf{Q}^{-1})'\mathbf{b} = [\mathbf{c}', \mathbf{v}'] \begin{bmatrix} (\mathbf{AA}')^{-1}\mathbf{A} \\ \mathbf{T} \end{bmatrix} \mathbf{b} = \mathbf{c}'(\mathbf{AA})^{-1}\mathbf{Ab} + \mathbf{v}'\mathbf{Tb}. \tag{21}$$

The transformed form of (1) is now b_0 + Eq. (21) + Eq. (20), namely

$$\hat{y} = d_0 + \mathbf{v}'\{\mathbf{Tb} + 2\mathbf{TBA}'(\mathbf{AA}')^{-1}\mathbf{c}\} + \mathbf{v}'\mathbf{T}\,\mathbf{B}\,\mathbf{T}'\mathbf{v}, \tag{22}$$

where

$$d_0 = b_0 + \mathbf{c}'(\mathbf{AA}')^{-1}\mathbf{Ab} + \mathbf{c}'(\mathbf{AA}')^{-1}\mathbf{AB}\,\mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{c}. \tag{23}$$

Differentiating this transformed version (22) of $\hat{y}$ once with respect to $\mathbf{v}$, and setting the result to zero, we obtain the stationary point in the restricted space from

$$2\mathbf{T}\,\mathbf{B}\,\mathbf{T}'\mathbf{v} + 2\mathbf{TBA}'(\mathbf{AA}')^{-1}\mathbf{c} + \mathbf{Tb} = \mathbf{0}, \tag{24}$$

with solution

$$\mathbf{v}_s = -(\mathbf{T}\,\mathbf{B}\,\mathbf{T}')^{-1}\{\tfrac{1}{2}\mathbf{Tb} + \mathbf{TBA}'(\mathbf{AA}')^{-1}\mathbf{c}\}. \tag{25}$$

Substituting (25) into (22) and canceling several terms leads to the predicted response at the restricted stationary point as

$$\hat{y}_s = d_0 - \mathbf{v}_s'(\mathbf{TBT}')\mathbf{v}_s, \tag{26}$$

where d_0 is given in (23). We note that, in the no-restrictions case with $\mathbf{T}=\mathbf{I}, \mathbf{A}=\mathbf{0}, \mathbf{c}=\mathbf{0}$, so that $\mathbf{v}=\mathbf{x}$, Eq. (25) reduces to Eq. (4), as it should. Under the same specialization, Eq. (26) reduces to (8). Because

$$\mathbf{x} = \mathbf{Q}^{-1}\mathbf{z} = \mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{c} + \mathbf{T}'\mathbf{v}, \tag{27}$$

we can obtain the stationary point in the space subject to the restrictions as

$$\mathbf{x}_s = \mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{c} + \mathbf{T}'\mathbf{v}_s \tag{28}$$

in terms of the original coordinates. This point is not usually, of course, the stationary point in the full x-space, which is often unattainable because of the restrictions and so would no longer be relevant anyway. Note that $\mathbf{Ax}_s = \mathbf{AA}'(\mathbf{AA}')^{-1}\mathbf{c} + \mathbf{AT}'\mathbf{v}_s = \mathbf{c}$, by applying (14), as it should, because $\mathbf{x}_s$ is in the restricted space. It is also clear that $\mathbf{x}_s$ is the same whatever choice is made for $\mathbf{T}$. For if another choice, $\mathbf{T}_2 = \mathbf{PT}$ is made, where $\mathbf{P}$ is a square $(m-q)\times(m-q)$ non-singular matrix such that $\mathbf{T}_2$ satisfies (14), it will be found that the $\mathbf{P}$'s cancel out when $\mathbf{T}_2$ is inserted in (28) in place of $\mathbf{T}$. The result of (28) does, of course, depend on $\mathbf{A}, \mathbf{c}$, and the coefficients of the fitted model.

By replacing the curly bracket in (22) using (25), we can re-write (22) in the form

$$\hat{y} = d_0 - 2\mathbf{v}'\mathbf{TBT}'\mathbf{v}_s + \mathbf{v}'\mathbf{T}\,\mathbf{B}\,\mathbf{T}'\mathbf{v}. \tag{29}$$

Subtracting (26) from(29) and then factorizing the result gives

$$\hat{y} - \hat{y}_s = \mathbf{v}_s'(\mathbf{TBT}')\mathbf{v}_s - 2\mathbf{v}'\mathbf{TBT}'\mathbf{v}_s + \mathbf{v}'\mathbf{T}\,\mathbf{B}\,\mathbf{T}'\mathbf{v}$$
$$= (\mathbf{v} - \mathbf{v}_s)'\,\mathbf{TBT}'(\mathbf{v} - \mathbf{v}_s). \tag{30}$$

Suppose, we now move the origin to the stationary point of Eq. (25) by choosing new coordinate values $\mathbf{Z} = \mathbf{v} - \mathbf{v}_s$ and substituting $\mathbf{v} = \mathbf{Z} + \mathbf{v}_s$ into Eq. (30). This leads immediately to the form

$$\hat{y} = \hat{y}_s + \mathbf{Z}'\mathbf{T}\,\mathbf{B}\,\mathbf{T}'\mathbf{Z}, \tag{31}$$

where $\hat{y}_s$ is defined in Eq. (26). There are no first-order terms in $\mathbf{Z}$ in this equation, which means that we can convert to canonical formby repeating the same steps as used previously between Eqs. (4) and (5), but with the role of $\mathbf{B}$ now being given to the matrix $\mathbf{TBT}'$. Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_{q-m}$ are the $q - m$ eigenvalues of $\mathbf{TBT}'$ and let $\mathbf{M} = \{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_{q-m}\}$ denote a corresponding matrix of orthonormal eigenvectors. Write $\Lambda = \textbf{diagonal}\{\lambda_1, \lambda_2, \ldots, \lambda_{q-m}\}$. Note that we employ the same sort of notation given previously between Eqs. (4) and (5); however, the dimensions are reduced by the number of linear restrictions, and so are now $q - m$ rather than q. The rotated orthogonal axes are now of the form

$$\mathbf{W} = \mathbf{M}'\mathbf{Z} = \mathbf{M}'(\mathbf{v} - \mathbf{v}_s) = \mathbf{M}'\mathbf{T}(\mathbf{x} - \mathbf{x}_s) \tag{32}$$

fromEqs. (27) and (28), and the surface can now be expressed as

$$\hat{y} = \hat{y}_s + \mathbf{W}'\Lambda\mathbf{W} = \hat{y}_s + \sum_{i=1}^{q-m} \lambda_i W_i^2. \tag{33}$$

The stationary point lies where all $W_i = 0$. Individual axes are determined by the conditions that $q - m - 1$ (that is, all but one) of the W_i are zero.

The numerical coefficients attached to the elements of $\mathbf{x}$ in (32) are the rows of $\mathbf{M}'\mathbf{T}$. However $\mathbf{M}'\mathbf{TA}' = \mathbf{0}$; see (14). Thus, the evaluation of the product $\mathbf{M}'\mathbf{TA}'$, which should be $\mathbf{0}$ within rounding error, provides a useful numerical check that the calculations (32) have been correctly performed. Specifically, in examples where $\mathbf{A}$ contains a row $(1/q^{1/2}, 1/q^{1/2}, \ldots, 1/q^{1/2})$, the coefficients of $x_1, x_2, \ldots, x_q$, in each W_i will add to zero.

In applications with linear restrictions, canonical reduction can give rise to several possible outcomes. Our first concern is typically whether or not the stationary point lies within the appropriate mixture space, which would be the full mixture region if only the mixture restriction applied. If it does, we can then investigate the shape taken by the fitted surface within the restricted space. It will then also be informative to evaluate the coordinates where the new W-axes intersect, and exit, the boundaries of the restricted region. When the stationary point is outside the region of interest, the behavior of the surface around the stationary point is usually not relevant, particularly when the stationary point is far away. However, it may still be useful to see if any of the canonical axes passes through the restricted region and, if any do, to determine predicted values of the fitted response surface along such axes. In cases where canonical reduction is not fruitful, performing a full ridge analysis is then the best course of action; see Hoerl (1987), Peterson (1993) and Draper and Pukelsheim(2002) . We next provide two illustrative examples of canonical reduction in mixture problems.

5. Three ingredients mixture example (Bures et al., 1992)

We illustrate some of the details above via a three-ingredient ($q=3$) example in the usual mixture space $x_1+x_2+x_3=1$ ($m=1$), leading to a canonical reduction in $q-m=2$ dimensions. Consider the fitted mixture model in three ingredients $\hat{y} = -0.00658x_1 - 0.00243x_2 + 0.00367x_3 + 0.34265x_1x_2 + 0.47074x_1x_3 + 0.14115x_2x_3$, derived fromselected data for the proportional shrinkage of a mixture of three ingredients forming an artificial medium for growing plants. We have $q = 3, b_0 = 0, \mathbf{b}' = (-0.00658, -0.00243, 0.00367)$,

$$\mathbf{B} = \begin{bmatrix} 0 & 0.171325 & 0.23537 \\ 0.171325 & 0 & 0.070575 \\ 0.23537 & 0.070575 & 0 \end{bmatrix},$$

$$\mathbf{A} = \left(\frac{1}{3^{1/2}}, \frac{1}{3^{1/2}}, \frac{1}{3^{1/2}}\right), \quad \mathbf{c} = \frac{1}{3^{1/2}}$$

and

$$\mathbf{T} = \begin{bmatrix} -\frac{1}{2^{1/2}} & 0 & \frac{1}{2^{1/2}} \\ \frac{1}{6^{1/2}} & -\frac{2}{6^{1/2}} & \frac{1}{6^{1/2}} \end{bmatrix}.$$

The rows of $\mathbf{T}$ are simply the normalized orthogonal polynomials of first and second order. Then fromEq. (25), $\mathbf{v}_s' = (0.0104, 0.3882)$ so that, fromEq. (28), $\mathbf{x}_s' = (0.484, 0.016, 0.499)$, which is a point just inside the mixture space near the $x_2 = 0$ boundary; see Fig. 1. The two eigenvalues of $\mathbf{TBT}'$ are $(-0.2550, -0.0632)$, with a corresponding matrix of eigenvectors of

$$\mathbf{M} = \begin{bmatrix} 0.9474 & 0.3200 \\ -0.3200 & 0.9474 \end{bmatrix}.$$

The canonical formof the fitted surface is now

$$\hat{y} = \hat{y}_s - 0.2550\, W_1^2 - 0.0632\, W_2^2,$$

where $\hat{y}_s = 11.62$. This formindicates (see Box and Draper, 1987, p. 338) a set of elliptical contours centered at the point $\mathbf{x}_s$. The negative signs indicate a *maximum* response of 11.62 at the stationary point $W_1 = W_2 = 0$. The lengths of the major axes of the ellipses are in the ratio of $(0.255)^{-1/2} : (0.0632)^{-1/2}$, which is approximately 1:2; see Fig. 1. The $q - m = 3 - 1 = 2$ canonical axes in the space $x_1 + x_2 + x_3 = 1$ are given via (32) as

$$W_1 = -0.801x_1 + 0.261x_2 + 0.539x_3 + 0.114,$$

$$W_2 = 0.160x_1 - 0.774x_2 + 0.613x_3 - 0.371.$$

(We note that the x-coefficients in each W add to zero, within rounding error, as explained below Eq. (33).) By setting $W_1 = 0$, then setting each $x_i = 0$ in turn, $i = 1, 2, 3$, while maintaining the

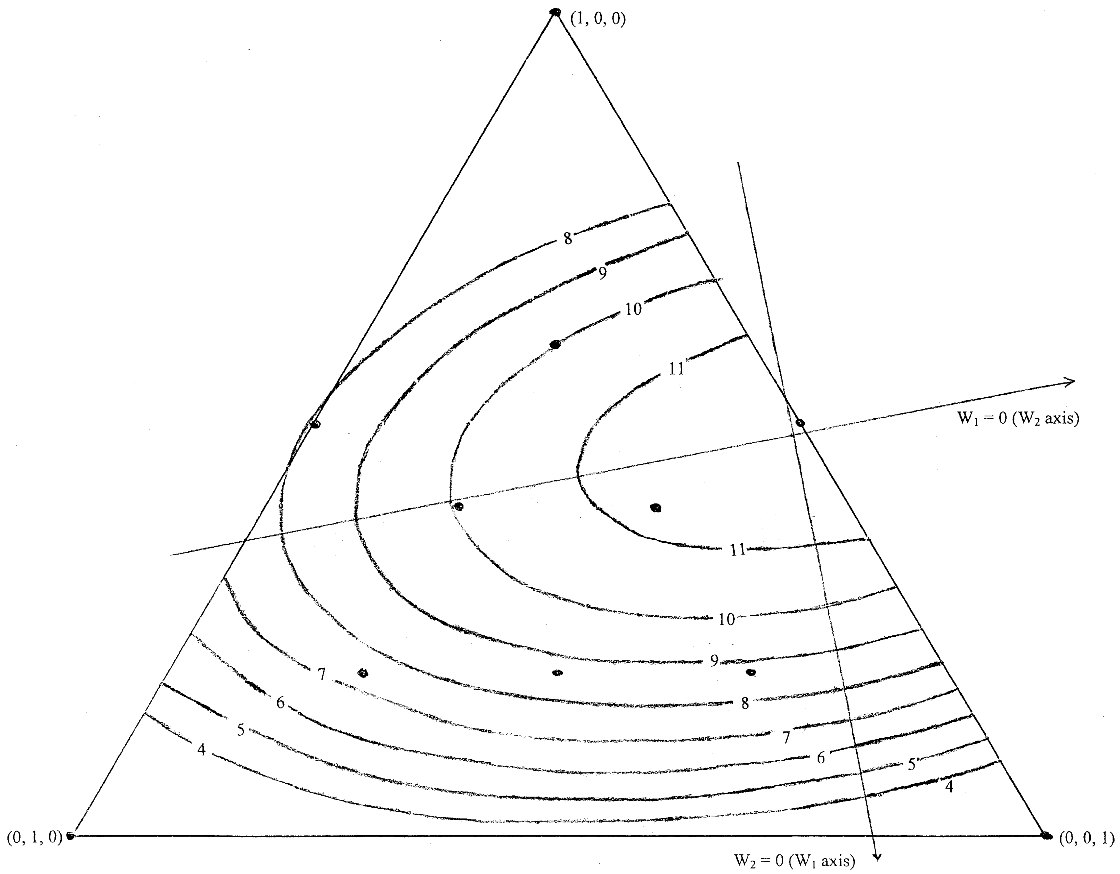

Fig. 1. Canonical reduction for Example 1. (The dots are the locations of the data points.)

mixture restriction, we obtain points of intersection with the mixture boundaries with coordinates

$$(0, 2.35, -1.35), (0.49, 0, 0.51) \quad \text{and} \quad (0.35, 0.65, 0),$$

indicating that the W_2-axis, defined by $W_1 = 0$, meets the $x_1 = 0$ boundary well outside the mixture region but intersects the $x_2 = 0$ and $x_3 = 0$ boundaries at valid mixture points.

Similarly, when we let $W_2 = 0$, and then set each $x_i = 0$ in turn, $i = 1, 2, 3$, again while maintaining the mixture restriction, we obtain the intersections

$$(0, 0.17, 0.83), (0.53, 0, 0.47), (1.23, -0.23, 0),$$

so that the W_1-axis cuts across the $x_1 = 0$ and $x_2 = 0$ boundaries within the mixture space restrictions; however, the W_1-axis meets the $x_3 = 0$ boundary outside the mixture space. By substituting for the centroid $(x_1, x_2, x_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ in the W-equations above, we find that it lies at $(W_1, W_2) = (0.114, -0.371)$. This enables us to determine which direction of each axis is the positive one in the canonical equation. (These directions, the choices of which are made by the convention written into the programused to obtain eigenvalues and eigenvectors, are needed to interpret the locations of points on the response surface in the restricted space.) All these remarks are illustrated in Fig. 1.

6. Five ingredients mixture example (Kissell and Marshall, 1962)

This example gave rise to a fitted second-order response surface of form (1) with

$$\mathbf{b} = \begin{bmatrix} -1605003 \\ 4487 \\ 559 \\ -7418 \\ -13347 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 0 & 1731252 & 1674333 & 1427295 & 1904909 \\ 1731252 & 0 & -6202 & 912 & 7783 \\ 1674333 & -6202 & 0 & 15718 & 4486 \\ 1427295 & 912 & 15718 & 0 & 41439 \\ 1904909 & 7783 & 4486 & 41439 & 0 \end{bmatrix},$$

$$\mathbf{A} = \left[\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right], \quad \mathbf{c} = \frac{1}{\sqrt{5}}$$

and

$$\mathbf{T} = \begin{bmatrix} -\frac{2}{10^{1/2}} & -\frac{1}{10^{1/2}} & 0 & \frac{1}{10^{1/2}} & \frac{2}{10^{1/2}} \\ \frac{2}{14^{1/2}} & -\frac{1}{14^{1/2}} & -\frac{2}{14^{1/2}} & -\frac{1}{14^{1/2}} & \frac{2}{14^{1/2}} \\ -\frac{1}{10^{1/2}} & \frac{2}{10^{1/2}} & 0 & -\frac{2}{10^{1/2}} & \frac{1}{10^{1/2}} \\ \frac{1}{70^{1/2}} & -\frac{4}{70^{1/2}} & \frac{6}{70^{1/2}} & -\frac{4}{70^{1/2}} & \frac{1}{70^{1/2}} \end{bmatrix}.$$

Again the rows of **T** are normalized orthogonal polynomials, here of orders 1–4. The stationary point is at $\mathbf{x}_s = (0.335, -1.872, 9.084, -3.783, -2.763)$ well outside the mixture region. None of the major axes passes through the region of interest, which is a subregion of the mixture space defined by $x_1 + x_2 + x_3 + x_4 + x_5 = 1$. Canonical reduction tells us relatively little here, and it would be necessary to use ridge analysis to negotiate the restricted region instead. See Hoerl (1987), Peterson (1993) and Draper and Pukelsheim(2002) .

Acknowledgements

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

References

Box, G.E.P., Draper, N.R., 1987. Empirical Model Building and Response Surfaces. Wiley, New York.
Bures, S., Pokorny, F.A., Ware, G.O., 1992. Shrinkage in ternary mixtures of container ingredients. Proceedings of the 1991 Kansas State University Conference on Applied Statistics in Agriculture, Manhattan, KS, USA, pp. 43–52.
Cornell, J.A., 2002. Experiments with Mixtures, 3rd Edition. Wiley, New York.
Draper, N.R., Pukelsheim, F., 2002. Generalized ridge analysis under linear restrictions, with particular applications to mixture experiments problems. Technometrics 44 (3), 250–259, (see also the Letter to the Editor in Technometrics 45 (2), 185).
Hoerl, R.W., 1987. The application of ridge techniques to mixture data: ridge analysis. Technometrics 29 (2), 161–172.
Kissell, L.T., Marshall, B.D., 1962. Multi-factor responses of cake quality to basic ingredient ratios. Cereal Chem39, 16–30.
Peterson, J.J., 1993. A general approach to ridge analysis with confidence intervals. Technometrics 35 (2), 204–214.