Standard Errors for the Eigenvalues in Second-Order Response Surface Models
Author(s): Søren Bisgaard and Bruce Ankenman
Source: *Technometrics*, Vol. 38, No. 3 (Aug., 1996), pp. 238-246
Published by: American Statistical Association and American Society for Quality
Stable URL: http://www.jstor.org/stable/1270607
Accessed: 07-08-2014 16:15 UTC

Standard Errors for the Eigenvalues in Second-Order Response Surface Models

Søren BISGAARD

Center for Quality and Productivity Improvement
and Department of Industrial Engineering
University of Wisconsin-Madison
Madison, WI 53705

Bruce ANKENMAN

Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60208

In the interpretation of the geometry of second-order response surface models, standard errors and confidence intervals for the eigenvalues of the second-order coefficient matrix play an important role. In this article, we propose a new method for estimating the standard errors, and hence approximate confidence intervals, of these eigenvalues. The method is simple in both concept and execution. It involves the refitting of a full quadratic model after rotating the coordinate system to coincide with the canonical axes. The estimated standard errors of the pure quadratic terms from this refitting are then used as approximate standard errors of the eigenvalues. Because this approach is based on the canonical form, it is geometrically intuitive and easily taught. Our method is intended as a way for practitioners to get quick estimates of the standard errors of the eigenvalues. In our justification of the approach, we show that it is equivalent to using the delta method proposed by Carter, Chinchilli, and Campbell.

KEY WORDS: Canonical analysis; Confidence intervals; Delta method; Response surface methodology.

Response surface methodology, first introduced by Box and Wilson (1951), is frequently used to find conditions that either maximize or minimize a given quality characteristic or, more broadly, explore near-stationary conditions including ridges. When near-stationary conditions are reached, second-order polynomial models are typically fit to the data using ordinary least squares (OLS). Using matrix notation, the standard quadratic model is

$$E(y) = \beta_0 + \mathbf{x}'\beta + \mathbf{x}'B\mathbf{x}, \tag{1}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_k)'$ is a vector of k factors, $\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$, and

$$B = \begin{bmatrix} \beta_{11} & \frac{1}{2}\beta_{12} & \cdots & \frac{1}{2}\beta_{1k} \\ \frac{1}{2}\beta_{12} & \beta_{22} & \cdots & \frac{1}{2}\beta_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\beta_{1k} & \frac{1}{2}\beta_{2k} & \cdots & \beta_{kk} \end{bmatrix}. \tag{2}$$

To distinguish between (1) and its estimated equivalent, we use the conventional hat notation for estimated parameters.

If (1) contains three or fewer factors, contour plots of the fitted surface can aid in the geometric interpretation of the response surface. Many textbooks, including that of Box and Draper (1987, pp. 347–351), provide examples of such plots and discuss their practical significance. For example, by studying the contour plot, the experimenter can determine if the fitted surface has a maximum, minimum, or saddlepoint close to the experimental region or whether it describes some type of ridge system. When there is a unique optimum, it is easy to determine its location from the plot. If there is a rising ridge in the surface, the plot can be used to find the ridge and to decide the direction for future experiments.

A stationary ridge system, in which there is a line or plane of optimal or nearly optimal points, is a case of particular interest. Because all of the points on the stationary ridge optimize the quality characteristic under study, we can choose a point on the ridge that is most cost-effective or optimize some secondary criteria. It is this potential ability to produce high quality at reduced cost that makes the identification and exploration of ridge systems so important. In two or three dimensions, a stationary ridge can be easily identified and exploited using a contour plot like those just discussed.

When more than three factors are involved, plots are difficult to make and other methods are needed. Box and Wilson (1951) pointed out that the shape of the response surface locally can be interpreted by performing a *canonical* analysis in which the signs and magnitudes of the eigenvalues of B, the matrix of second-order parameters in (1), are considered. In particular, when all of the eigenvalues are positive, the stationary point is a minimum, whereas eigenvalues that are all negative indicate a maximum. If the eigenvalues have mixed signs, it indicates a saddlepoint. An eigenvalue of 0 indicates the presence of a ridge in the surface. Canonical analysis was discussed in detail by Box and Wilson (1951), Box and Hunter (1954), and Box and Youle (1955), and in many textbooks including those by Box and Draper (1987), Khuri and Cornell (1987), and Myers (1976).

Because response surface models are typically based on experimental data, the eigenvalues of the estimated coefficient matrix $\hat{B}$ are subject to error. This further complicates

the determination of the shape of the response surface. To provide standard errors and large-sample confidence intervals for the eigenvalues of $\hat{B}$, Carter, Chinchilli, and Campbell (1990), referred to hereafter as CCC, suggested using the delta method. They furthermore produced simulation results showing that their approximate confidence intervals using the delta method are accurate under a fairly wide range of conditions.

The method presented by CCC for finding confidence intervals for the eigenvalues involves a significant amount of vector and matrix manipulation, which may not be familiar to many practitioners. The purpose of this article is therefore to present and justify an alternative and very simple method. Our approach, which we call the *double linear regression* (DLR) *method,* involves two successive linear regressions, each fitting a full quadratic model. The first regression is used to fit a standard quadratic model followed by a determination of the canonical axes. In the second regression a full quadratic model is fit based on a new set of coordinates rotated to match the canonical axes. After this rotation and refitting, the standard errors of the pure quadratic terms that are produced as standard output from regression routines are used to approximate the standard errors of the eigenvalues. The method's simple geometry makes it easy to explain and use. It is also of theoretical interest because, as we will show in Section 5, our approach is equivalent to the delta method proposed by CCC.

In the remainder of the article, we will first briefly discuss some previously proposed methods for constructing confidence intervals for the eigenvalues of B. As a way of defining our notation, we then, in Section 2, provide a short summary of the canonical form of a second-order response surface model. Following that, we present the DLR method in Section 3 and use an example from Box (1954) in Section 4 to demonstrate its usefulness and simplicity. Finally, in Section 5 we show that the confidence intervals obtained by the DLR method are identical to those obtained by the delta method.

1. PREVIOUSLY SUGGESTED APPROACHES

The most accessible method for supplying confidence statements for the eigenvalues of B was suggested by Box and Draper (1987, p. 354). They simply argued that, if the experimental design is approximately rotatable, then it follows that the variance of the estimated eigenvalues is approximately equal to the variance of the diagonal elements of $\hat{B}$. The advantage of the Box–Draper approach is that it requires no further calculations beyond those already associated with fitting of the original second-order model. It is applicable only to approximately rotatable designs, however.

A more general approach, suggested by Carter, Chinchilli, Myers, and Campbell (1986), for finding confidence intervals for the eigenvalues of B is based on the observation that the eigenvalues are functions of the parameters in B. They then proceeded to use a computer-intensive search method to produce conservative Scheffé-type (ST) intervals. Because the ST intervals are conservative as well as computer intensive, however, CCC later suggested an alternative approach using the delta method. The resulting intervals are therefore based on first-order approximations and are supported by large-sample theory. Although this latter approach is less computer intensive, it involves a fair amount of vector and matrix manipulation that may be discouraging to practitioners. We will provide more details in Section 5.

In a recent article, Peterson (1993) discussed the related problem of inference for the eigenvalues when the response surface is constrained, as in mixture experiments and ridge analysis. In summary, he used a polar coordinate transformation of the linear model to produce what he calls conservative and approximate confidence band (CB) intervals for both the mean response at the stationary point and the maximum or minimum eigenvalue of B. He also showed that the conservative CB intervals are easier to find than the ST intervals. The CB intervals, however, are at least as large as the already conservative ST intervals. To compensate for this conservatism, Peterson (1993) proposed an approximate CB interval that is based on work by Clarke (1987). He showed that, in the case of rotatable designs, this approximate CB interval for the maximum (minimum) eigenvalue is identical to that of Box and Draper (1987) and CCC.

Both CCC and Peterson (1993) provided simulation results that suggest that the approximate confidence intervals that they propose for the eigenvalues may be adequate for many situations, including the case of relatively small sample sizes. Unfortunately, except for the method suggested by Box and Draper, which is restricted to rotatable designs, none of the methods currently available are readily accessible to the average industrial practitioner. The need for a simple yet general method is therefore the motivation for us to suggest the DLR method. Before presenting the details of this method, however, we will, in Section 2, provide a short review of the canonical form of a quadratic response surface model.

2. THE CANONICAL FORM OF A RESPONSE SURFACE

Geometrically, the canonical form of a second-order response surface model (1) amounts to rotating the basis vectors of a new coordinate system so that they coincide with the normalized eigenvectors of $\hat{B}$, the matrix of second-order coefficients from the fitted surface. In other words, the new axes are parallel to the main axes of the second-order response surface. If we denote the ith normalized eigenvector of $\hat{B}$ by $\hat{\mathbf{d}}_i$, then the matrix

$$\hat{D} = [\hat{\mathbf{d}}_1, \hat{\mathbf{d}}_2, \ldots, \hat{\mathbf{d}}_k] \tag{3}$$

defines this rotation and the new canonical variables $\mathbf{z}$ can be expressed as functions of the original variables $\mathbf{x}$ as follows:

$$\mathbf{z} = \hat{D}'\mathbf{x}. \tag{4}$$

If we now let $\hat{\lambda}_i$ be the eigenvalue corresponding to $\hat{\mathbf{d}}_i$, then in the rotated coordinate system the fitted surface is given

by

$$\hat{y} = \hat{\beta}_0 + \mathbf{z}'\hat{\phi} + \mathbf{z}'\hat{\Lambda}\mathbf{z}, \tag{5}$$

where $\hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_k)' = \hat{D}'\hat{\beta}$ and $\hat{\Lambda} = \hat{D}'\hat{B}\hat{D} = \text{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_k)$. Notice that, because $\hat{\Lambda}$ is a diagonal matrix, the canonical model does not contain any cross-product terms, hence simplifying its interpretation. As explained in the Introduction, the signs and magnitudes of the eigenvalues of $\hat{B}$, which are the pure quadratic terms in the canonical model, enable the experimenter to interpret the amount of curvature along each canonical axis and hence to understand the shape of the surface.

3. THE DOUBLE LINEAR REGRESSION METHOD

As an algorithm, the first step of the DLR method for finding the standard errors of the eigenvalues of $\hat{B}$ is to fit a standard full quadratic model (1). Using the parameter estimates from this fitting, we then calculate $\hat{D}$, the rotation matrix of eigenvectors of $\hat{B}$. Next, as in standard canonical analysis, the original variables $\mathbf{x}$ are replaced by the canonical variables $\mathbf{z}$ using (4). The final step, which is the only step beyond standard canonical analysis, is to again use linear regression to refit the following *full* second-order model using the new rotated coordinates $\mathbf{z}$ as a basis:

$$E(y) = \beta_0^* + \mathbf{z}'\beta^* + \mathbf{z}'B^*\mathbf{z}, \tag{6}$$

where $\beta^* = (\beta_1^*, \beta_2^*, \ldots, \beta_k^*)'$ and

$$B^* = \begin{bmatrix} \beta_{11}^* & \frac{1}{2}\beta_{12}^* & \cdots & \frac{1}{2}\beta_{1k}^* \\ \frac{1}{2}\beta_{12}^* & \beta_{22}^* & \cdots & \frac{1}{2}\beta_{2k}^* \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}\beta_{1k}^* & \frac{1}{2}\beta_{2k}^* & \cdots & \beta_{kk}^* \end{bmatrix}.$$

It can be shown (Bisgaard and Ankenman 1994) that $\hat{B}^*$ is diagonal up to numerical rounding and that the diagonal elements of $\hat{B}^*$ are equal to the eigenvalues of $\hat{B}$ from the first regression, again up to numerical rounding. The standard error of the ith diagonal term, $\hat{\beta}_{ii}^*$, then provides us with an approximate standard error for $\hat{\lambda}_i$, the ith eigenvalue of $\hat{B}$. Thus, estimates of the standard errors of the eigenvalues will be calculated by any linear regression routine used to perform this second regression. We now describe the DLR method in step-by-step detail.

Suppose that, for a given experiment, there is a single response, y, and k explanatory variables denoted $\mathbf{x} = (x_1, x_2, \ldots x_k)'$. If there are n observations, the design matrix, X_D, is an $n \times k$ matrix such that the (i, j)th element of X_D is the level of x_j in the ith observation. The full second-order model for this problem has $p = (k+1)(k+2)/2$ parameters. The steps of the DLR method are as follows:

1. Fit a full second-order model, $E(y) = \beta_0 + \mathbf{x}'\beta + \mathbf{x}'B\mathbf{x}$, using OLS to get the estimates $\hat{\beta}_0, \hat{\beta}$, and $\hat{B}$.
2. Calculate the eigenvalues, $\hat{\lambda}_i$ for $i = 1, 2, \ldots, k$, and the corresponding eigenvectors of $\hat{B}$. The matrix of eigenvectors is the rotation matrix, $\hat{D}$, from the standard canonical relationships (a) $\hat{\lambda}_i$ is the ith eigenvalue of $\hat{B}$ and (b) $\hat{D} = [\hat{\mathbf{d}}_1, \hat{\mathbf{d}}_2, \ldots, \hat{\mathbf{d}}_k]$, where $\hat{\mathbf{d}}_i$ is the ith normalized eigenvector of $\hat{B}$.
3. Rotate the coordinate basis vectors and redefine the design points in the new coordinate system, $\mathbf{z} = \hat{D}'\mathbf{x} = (z_1, z_2, \ldots, z_k)'$. This step is conveniently accomplished for the entire dataset by postmultiplying the original design matrix by $\hat{D}$ to produce a new design matrix: $Z_D = X_D\hat{D}$.
4. Using the new design matrix, Z_D, again fit a full second-order model, $E(y) = \beta_0^* + \mathbf{z}'\beta^* + \mathbf{z}'B^*\mathbf{z}$, with OLS.
5. For $i \neq j$, the $\hat{\beta}_{ij}^*$'s from step 4 will be very close to 0. For $i = 1, 2, \ldots, k$, $\hat{\beta}_{ii}^*$ will be equal to $\hat{\lambda}_i$ up to numerical rounding, and $\text{se}(\hat{\beta}_{ii}^*)$, the standard error of $\hat{\beta}_{ii}^*$ provided by a standard linear regression routine, is used as an approximate standard error for the $\hat{\lambda}_i$.
6. As in standard linear regression, the following equation can then be used to produce an approximate $100(1-\alpha)\%$ confidence interval for λ_i:

$$\hat{\lambda}_i \pm t_{1-\alpha/2, n-p}\text{se}(\hat{\beta}_{ii}^*), \tag{7}$$

where n is the number of observations, p is the number of parameters in the model, and $t_{1-\alpha/2,n-p}$ is the $1-\alpha/2$ quantile of Student-t distribution with $n-p$ df. Note that, if all of the eigenvalues are to be tested individually, simultaneous confidence interval adjustment of (7) might be considered. For example, the Bonferroni adjustment replaces $t_{1-\alpha/2,n-p}$ with $t_{1-\alpha/2k,n-p}$, when there are k eigenvalues being compared to 0.

At first glance, the estimates of the parameters obtained in the regression of step 4 may appear to be conditional on the rotation matrix estimated in step 2. If this were so, the standard-error estimates would be overly optimistic. Because a *full* quadratic model including all the cross-product terms is fit in the second regression, however, the fit is unconditional. Hence we are using only a fixed rotation, which "incidentally" is chosen to correspond to the canonical axes of the fitted response surface. Our approach is therefore no more controversial than the standard practice of centering and scaling regression models or the use of stepwise regression modeling (e.g., see Draper and Smith 1981). A more theoretical justification showing the exact equivalence between our approach and the delta method used by CCC is provided later in the article, however.

In Section 4 we will show a detailed example of the use of the DLR method. Let us first retrace the final steps of the second regression, however, and review how the standard errors of the $\hat{\beta}_{ii}^*$'s in step 4 are estimated. The matrix of regressor columns for fitting (6) with linear regression is an $n \times p$ matrix as follows:

$$Z = [\mathbf{1}, \mathbf{z}_1, \ldots, \mathbf{z}_k, \mathbf{z}_{12}, \ldots, \mathbf{z}_{(k-1)k}, \mathbf{z}_{11}, \ldots, \mathbf{z}_{kk}], \tag{8}$$

where $\mathbf{1}$ is a column of n ones, $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k$ are the k columns of the rotated design matrix Z_D, and the other $k(k+1)/2$ columns are generated through element-by-element multiplication of all possible pairs of columns of Z_D. Thus, for any $r = 1, 2, \ldots, k$ and $s = r(r+1, \ldots, k)$, $\mathbf{z}_{rs}$ is generated by the element-by-element multiplication of $\mathbf{z}_r$ and $\mathbf{z}_s$. Note that, for convenience, we have

permuted the columns of Z so that the columns associated with the pure quadratic terms appear in the far right columns of the matrix. Assuming additive independent and identically distributed $N(0, \sigma^2)$ errors, the standard error of $\hat{B}^*_{ii}$, and thus the approximate standard error of $\hat{\lambda}_i$, is the square root of the ith diagonal element of the lower $k \times k$ submatrix of $s^2(Z'Z)^{-1}$, where s^2 is the estimate of the residual variance.

We notice in passing that the DLR method has been presented using the A-canonical form involving only a rotation of the coordinate axes. The alternative B-canonical form involves a rotation of the axes and a shift of the origin of the coordinate system to the stationary point. It can be shown (Ankenman 1994) that the DLR method will provide the same estimates of standard error for the eigenvalues if the B-canonical form is used. The A-canonical form, however, requires fewer calculations and thus is recommended for most situations.

Because the DLR method produces the same standard errors as those produced by the delta method of CCC, the advantage that we claim for the DLR method is not increased accuracy but ease of calculation and geometric intuition. With reference to ease of calculation, one of the referees pointed out that certain software packages, including SAS, allow for the estimation of the variance of linear combinations of the linear model coefficients. If we treat the elements of the eigenvectors as constants as we do in the DLR method, then each eigenvalue estimate can be written as a linear combination of the vector of second-order coefficients $\hat{\beta}_q = (\hat{\beta}_{11}, \hat{\beta}_{12}, \ldots, \hat{\beta}_{1k} \vdots \hat{\beta}_{22}, \hat{\beta}_{23}, \ldots, \hat{\beta}_{2k} \vdots \ldots \vdots \hat{\beta}_{kk})'$. Specifically, using the matrix-to-vector operator, vech, which is defined later in Section 5, we find that $\hat{\lambda}_i = \text{vech}'(\hat{\mathbf{d}}_i \hat{\mathbf{d}}_i')\hat{\beta}_q$. Estimating the variance of this expression will lead to the same confidence intervals as the DLR method and the delta method. The ease of this calculation will depend on the user's familiarity with the software package. The DLR method relies only on standard matrix operations and linear regression, however, and thus can be performed in almost any regression package. In addition, because it is based on the geometrically intuitive idea of rotation, it is easily taught and retained. In fact, once a practitioner has used and understands the DLR method, there should be no need to refer to the steps provided previously even when switching from one software package to another.

4. AN EXAMPLE

In this section we will provide a detailed example of the steps of the double linear regression approach using an experiment originally published by Box (1954). The objective of the experiment was to maximize the yield of a chemical reaction. It involved five factors—stage one temperature (x_1), stage one reaction time (x_2), stage one concentration (x_3), stage two temperature (x_4), and stage two reaction time (x_5). The experimental design is quite unusual because it was actually run as three sequential experiments in which the experimental design for each step was determined by the analysis from the previous step. The initial step was a standard 2^{5-1} fractional factorial. The analysis of this experiment led the experimenters to believe that the optimum was likely to be near a certain corner of the design. Based on this, they decided to add enough design points to estimate a second-order surface but to add the points near that corner. Thus, after two steps they had a noncentral composite design with 21 design points. They were now estimating 21 parameters with 21 observations, however, so they felt it was necessary to add some confirmatory design points in a third step. Because they had estimates of the stationary point and the canonical axes of the fitted surface, they decided to add the final 11 experimental points as center and starpoints around this estimated stationary point and on the canonical axes. This, of course, makes the levels in the original variables for step 3 quite nonstandard.

The design matrix for all three steps taken as one large experiment, X_D, and the vector of observed yields, $\mathbf{y}$, are shown as follows:

$$X_D = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5]$$

$$= \begin{bmatrix}
-1 & -1 & -1 & -1 & 1 \\
1 & -1 & -1 & -1 & -1 \\
-1 & 1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & 1 \\
-1 & -1 & 1 & -1 & -1 \\
1 & -1 & 1 & -1 & 1 \\
-1 & 1 & 1 & -1 & 1 \\
1 & 1 & 1 & -1 & -1 \\
-1 & -1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1 & 1 \\
-1 & 1 & -1 & 1 & 1 \\
1 & 1 & -1 & 1 & -1 \\
-1 & -1 & 1 & 1 & 1 \\
1 & -1 & 1 & 1 & -1 \\
-1 & 1 & 1 & 1 & -1 \\
1 & 1 & 1 & 1 & 1 \\
3 & -1 & -1 & 1 & 1 \\
1 & -3 & -1 & 1 & 1 \\
1 & -1 & -3 & 1 & 1 \\
1 & -1 & -1 & 3 & 1 \\
1 & -1 & -1 & 1 & 3 \\
1.23 & -.56 & -.03 & .69 & .70 \\
.77 & -.82 & 1.48 & 1.88 & .77 \\
1.69 & -.30 & -1.55 & -.50 & .62 \\
2.53 & .64 & -.10 & 1.51 & 1.12 \\
-.08 & -1.75 & .04 & -.13 & .27 \\
.78 & -.06 & .47 & -.12 & 2.32 \\
1.68 & -1.06 & -.54 & 1.50 & -.93 \\
2.08 & -2.05 & -.32 & 1.00 & 1.63 \\
.38 & .93 & .25 & .38 & -.24 \\
.15 & -.38 & -1.20 & 1.76 & 1.24 \\
2.30 & -.74 & 1.13 & -.38 & .15
\end{bmatrix}$$

and

$$\mathbf{y} = \begin{bmatrix} 49.8 \\ 51.2 \\ 50.4 \\ 52.4 \\ 49.2 \\ 67.1 \\ 59.6 \\ 67.9 \\ 59.3 \\ 70.4 \\ 69.6 \\ 64.0 \\ 53.1 \\ 63.2 \\ 58.4 \\ 64.3 \\ 63.0 \\ 63.8 \\ 53.5 \\ 66.8 \\ 67.4 \\ 72.3 \\ 57.1 \\ 53.4 \\ 62.3 \\ 61.3 \\ 64.8 \\ 63.4 \\ 72.5 \\ 72.0 \\ 70.4 \\ 71.8 \end{bmatrix}.$$

The matrix of regressor columns for fitting the full quadratic model is

$$X = [\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_{11}, \mathbf{x}_{12}, \mathbf{x}_{13}, \mathbf{x}_{14}, \mathbf{x}_{15}, \mathbf{x}_{22}, \mathbf{x}_{23}, \mathbf{x}_{24}, \mathbf{x}_{25}, \mathbf{x}_{33}, \mathbf{x}_{34}, \mathbf{x}_{35}, \mathbf{x}_{44}, \mathbf{x}_{45}, \mathbf{x}_{55}],$$

where, for example, $\mathbf{x}_1$ is the column associated with stage one temperature and $\mathbf{x}_{12}$ is the column associated with the interaction between stage one temperature and stage one reaction time. Using OLS to fit the model in (1) to the data, we find that

$$\hat{B} = \begin{bmatrix} -1.61 & -.95 & 1.05 & -.18 & -.02 \\ -.95 & -1.35 & .30 & -.08 & -.55 \\ 1.05 & .30 & -2.58 & -1.77 & -.38 \\ -.18 & -.08 & -1.77 & -2.34 & .20 \\ -.02 & -.55 & -.38 & .20 & -1.42 \end{bmatrix},$$

$$\hat{\beta} = \begin{bmatrix} 3.26 \\ 1.58 \\ 1.16 \\ 3.47 \\ 1.49 \end{bmatrix}, \qquad \hat{\beta}_0 = 68.72.$$

With these estimates, we find the matrix of eigenvalues of $\hat{B}$, $\hat{\Lambda} = \text{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4, \hat{\lambda}_5)$ and their corresponding eigenvectors, $\hat{D} = [\hat{\mathbf{d}}_1, \hat{\mathbf{d}}_2, \hat{\mathbf{d}}_3, \hat{\mathbf{d}}_4, \hat{\mathbf{d}}_5]$ to be

$$\hat{\Lambda} = \begin{bmatrix} -4.46 & 0 & 0 & 0 & 0 \\ 0 & -2.62 & 0 & 0 & 0 \\ 0 & 0 & -1.78 & 0 & 0 \\ 0 & 0 & 0 & -.40 & 0 \\ 0 & 0 & 0 & 0 & -.04 \end{bmatrix}$$

and

$$\hat{D} = \begin{bmatrix} -.28 & .64 & -.25 & .37 & -.56 \\ -.14 & .60 & .25 & -.73 & .16 \\ .74 & .00 & .23 & -.19 & -.60 \\ .59 & .43 & -.38 & .22 & .52 \\ .03 & .21 & .82 & .50 & .19 \end{bmatrix}.$$

In this experiment the stationary point of the fitted surface is $\mathbf{x}_s = -\frac{1}{2}\hat{B}^{-1}\hat{\beta} = (2.52, -1.10, 1.27, -.32, .53)'$ and the estimated response at this point is $\hat{y}_s = \hat{\beta}_0 + \frac{1}{2}\mathbf{x}_s'\hat{\beta} = 72.51$. When the stationary point is within the experimental region, as it is in this case, the B-canonical form is often fit for model simplification. We can write out the B-canonical model as follows: $\hat{y} = 72.51 - 4.46w_1^2 - 2.62w_2^2 - 1.78w_3^2 - .40w_4^2 - .04w_5^2$, where $w = (w_1, w_2, w_3, w_4, w_5)' = \hat{D}'(\mathbf{x} - \mathbf{x}_s)$ and the coefficients of the equation are the eigenvalues of $\hat{B}$. In Box's original analysis of the experiment, he concluded that $\hat{\lambda}_4$ and $\hat{\lambda}_5$ are small and thus that there is stationary ridge with a maximum on the plane that includes the fourth and fifth canonical axes. With this conclusion, the model is reduced to $\hat{y} = 72.51 - 4.46w_1^2 - 2.62w_2^2 - 1.78w_3^2$. He then produced a table of nearly alternative conditions for maximal yield, which could be used to maintain high yield while increasing throughput of the reaction or decreasing cost in some other way.

We now apply the DLR method to this example and obtain approximate confidence intervals for the eigenvalues. This will allow us to determine if the model reduction proposed previously is supported by statistical evidence. The first regression of the DLR method has already been done, and we are now ready to put the model into canonical form and perform the second regression. In both the A-canonical form and the B-canonical form, the eigenvalues of $\hat{B}$ are the second-order coefficients. Because the DLR method provides the same standard errors for both canonical forms and the A-canonical form requires less manipulation, we will use the A-canonical form for calculating confidence intervals for the eigenvalues.

For the A-canonical form, the transformation $\mathbf{z} = \hat{D}'\mathbf{x}$ is applied to each trial in the design matrix X_D by postmultiplying X_D by $\hat{D}$. Thus the design matrix for the second regression in the DLR method is

Table 1. The Estimated Eigenvalues, Their Approximate Standard Errors Calculated Using the DLR Method and the Associated Approximate 95% Confidence Intervals for the Chemical Experiment of Box (1954)

Canonical parameter	DLR regression parameter	Estimate	Standard error	Approx. 95% confidence interval
λ_1	β^*_{11}	−4.46	.25	(−5.02, −3.90)
λ_2	β^*_{22}	−2.62	.24	(−3.15, −2.10)
λ_3	β^*_{33}	−1.78	.26	(−2.36, −1.20)
λ_4	β^*_{44}	−0.40	.15	(−.73, −.07)
λ_5	β^*_{55}	−0.04	.24	(−.57, .48)

$$Z_D = X_D\hat{D} = [\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4, \mathbf{z}_5]$$

$$= \begin{bmatrix}
-.89 & -1.46 & .98 & .82 & .67 \\
-1.50 & -.61 & -1.17 & .57 & -.82 \\
-1.21 & -.69 & -.15 & -1.63 & .61 \\
-1.73 & 1.01 & .98 & .10 & -.13 \\
.55 & -1.88 & -.21 & -.55 & -.90 \\
.03 & -.18 & .92 & 1.19 & -1.65 \\
.33 & -.26 & 1.94 & -1.01 & -.22 \\
-.29 & .59 & -.21 & -1.27 & -1.71 \\
.24 & -1.02 & -1.43 & .28 & 1.33 \\
-.27 & .68 & -.30 & 2.01 & .59 \\
.02 & .60 & .72 & -.19 & 2.02 \\
-.60 & 1.45 & -1.43 & -.45 & .53 \\
1.78 & -.59 & .66 & .89 & .50 \\
1.16 & .26 & -1.49 & .64 & -.98 \\
1.45 & .18 & -.47 & -1.56 & .45 \\
.94 & 1.89 & .66 & .17 & -.30 \\
-.84 & 1.96 & -.81 & 2.74 & -.53 \\
.00 & -.51 & -.81 & 3.47 & .28 \\
-1.76 & .68 & -.75 & 2.38 & 1.79 \\
.91 & 1.55 & -1.06 & 2.45 & 1.63 \\
-.22 & 1.11 & 1.34 & 3.00 & .96 \\
.13 & .90 & -.15 & 1.37 & -.27 \\
2.12 & .99 & -.16 & 1.40 & -.33 \\
-1.87 & .81 & -.16 & 1.33 & -.20 \\
.04 & 2.89 & -.16 & 1.38 & -.26 \\
.22 & -1.10 & -.15 & 1.35 & -.27 \\
.13 & .91 & 1.84 & 1.37 & -.36 \\
.13 & .89 & -2.16 & 1.37 & -.17 \\
.09 & .88 & -.17 & 3.35 & -.47 \\
.17 & .91 & -.15 & -.62 & -.06 \\
.19 & .89 & -.06 & 1.56 & 1.72 \\
.07 & .90 & -.25 & 1.17 & -2.25
\end{bmatrix}.$$

After this coordinate rotation, we now fit the following full quadratic model in the new coordinate system:

$$\begin{aligned}
E(y) = {} & \beta^*_0 + \beta^*_1 z_1 + \beta^*_2 z_2 + \beta^*_3 z_3 + \beta^*_4 z_4 + \beta^*_5 z_5 \\
& + \beta^*_{12} z_1 z_2 + \beta^*_{13} z_1 z_3 + \beta_{14} z_1 z_4 + \beta^*_{15} z_1 z_5 \\
& + \beta^*_{23} z_2 z_3 + \beta^*_{24} z_2 z_4 + \beta^*_{25} z_2 z_5 + \beta^*_{34} z_3 z_4 \\
& + \beta^*_{35} z_3 z_5 + \beta^*_{45} z_4 z_5 + \beta^*_{11} z_1^2 + \beta^*_{22} z_2^2 \\
& + \beta^*_{33} z_3^2 + \beta^*_{44} z_4^2 + \beta^*_{55} z_5^2.
\end{aligned}$$

The regressor matrix, defined in (8), for fitting this model with linear regression is formed by element-by-element multiplication of the appropriate columns of Z_D to get

$$Z = [\mathbf{1}, \mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4, \mathbf{z}_5, \mathbf{z}_{12}, \mathbf{z}_{13}, \mathbf{z}_{14}, \mathbf{z}_{15}, \mathbf{z}_{23}, \mathbf{z}_{24}, \mathbf{z}_{25}, \mathbf{z}_{34}, \mathbf{z}_{35}, \mathbf{z}_{45}, \mathbf{z}_{11}, \mathbf{z}_{22}, \mathbf{z}_{33}, \mathbf{z}_{44}, \mathbf{z}_{55}].$$

The pure quadratic coefficients from this regression, $\hat{\beta}^*_{11}, \hat{\beta}^*_{22}, \hat{\beta}^*_{33}, \hat{\beta}^*_{44}$, and $\hat{\beta}^*_{55}$, are equal to the eigenvalues of $\hat{B}$ from the first regression. Now let S_λ denote the usual OLS estimate of the covariance matrix for these pure quadratic coefficients. Then S_λ is the lower 5×5 submatrix of $s^2(Z'Z)^{-1}$, where $s^2 = 2.04$ is the estimate of the residual variance. The standard errors of the eigenvalues are then approximated by taking the square root of the appropriate diagonal element of S_λ. For this example,

$$S_\lambda = \begin{bmatrix}
.064 & .028 & .030 & .016 & .028 \\
.028 & .058 & .028 & .017 & .028 \\
.030 & .028 & .069 & .016 & .027 \\
.016 & .017 & .016 & .023 & .015 \\
.028 & .028 & .027 & .015 & .057
\end{bmatrix}.$$

Using S_λ, we can produce an approximate $100(1-\alpha)\%$ confidence interval for λ_i by $\hat{\lambda}_i \pm t_{1-\alpha/2,n-p}(\mathbf{e}_i' S_\lambda \mathbf{e}_i)^{1/2}$, where $\mathbf{e}_i$ is a vector consisting entirely of zeros except for a one in the ith position and $t_{1-\alpha/2,n-p}$ is the $1-\alpha/2$ quantile of Student-t distribution with $n-p$ df. Notice that this computation is already provided by standard regression programs. For this example, we use $t_{.975,11} = 2.201$ to construct the confidence intervals. The estimated eigenvalues, their standard errors, and the approximate 95% confidence interval for each are summarized in Table 1. The individual confidence intervals for the eigenvalues are then plotted in Figure 1.

The confidence intervals for the first three eigenvalues show that they are clearly negative. The confidence interval for $\hat{\lambda}_5$ includes 0 and thus indicates that there may be a stationary ridge in the surface with a maximum along the fifth canonical axis. The confidence interval for $\hat{\lambda}_4$ is the most interesting because it comes very close to, but does not include, 0. In Box's original analysis of the experiment, he concluded that $\hat{\lambda}_4$ is small and thus that the stationary ridge is most likely a plane. The DLR method shows, however, that, although $\hat{\lambda}_4$ is indeed small with respect to the first three eigenvalues, the standard error of $\hat{\lambda}_4$ is also small and hence λ_4 may be significant. At a 95% confidence level, the DLR method shows significant curvature along the fourth

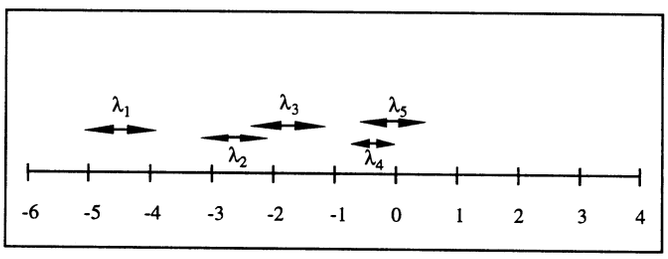

Figure 1. The Individual Approximate Confidence Intervals for the Eigenvalues in the Final Example Calculated by the DLR Method.

Table 2. The Estimated Eigenvalues, Their Approximate Standard Errors Calculated Using the DLR Method and the Associated Approximate 95% Bonferroni Intervals for the Chemical Experiment of Box (1954)

Canonical parameter	DLR regression parameter	Estimate	Standard error	Approx. 95% confidence interval
λ_1	β^*_{11}	−4.46	.25	(−5.24, −3.68)
λ_2	β^*_{22}	−2.62	.24	(−3.37, −1.88)
λ_3	β^*_{33}	−1.78	.26	(−2.59, −.97)
λ_4	β^*_{44}	−.40	.15	(−.87, .07)
λ_5	β^*_{55}	−.04	.24	(−.78, .70)

canonical axis and thus suggests that the optimal response can be found only along the fifth canonical axis, not on an entire plane.

An alternative is to use the Bonferroni adjustment for multiple comparisons. In that case we use

$$\hat{\lambda}_i \pm t_{1-\alpha/2k,n-p}(\mathbf{e}_i' S_\lambda \mathbf{e}_i)^{1/2} \quad \forall i,$$

where $t_{1-\alpha/2k,n-p}$ is the $1-\alpha/2k$ quantile of a Student-t distribution with $n-p$ df. For this example, we use $\alpha = .05$, so $t_{.995,11} = 3.106$. The estimated eigenvalues, their standard errors and the approximate 95% confidence interval for each are summarized in Table 2. The individual confidence intervals for the eigenvalues are then plotted in Figure 2.

As before, the confidence intervals for the first three eigenvalues show that they are clearly negative, and the confidence interval for λ_5 includes 0. The confidence interval for λ_4, however, now covers 0 by a small margin. In view of this small margin and the conservative nature of the Bonferroni approximation, doubt may still remain about the significance of λ_4.

In cases in which there is doubt about the statistical conclusions, the practitioner must decide the best course of action based on knowledge of the context of the experiment. If the amount of curvature along the fourth canonical axis is not of practical importance despite being statistically significant, then the experimenter can conclude that there is an approximate plane of optimal points and proceed accordingly. If the experimenter is attempting to make a conclusion about the physical mechanism involved in the experiment or if small changes in the response along the canonical axis are potentially important, however, it would be unwise to conclude that λ_4 can be removed from the model.

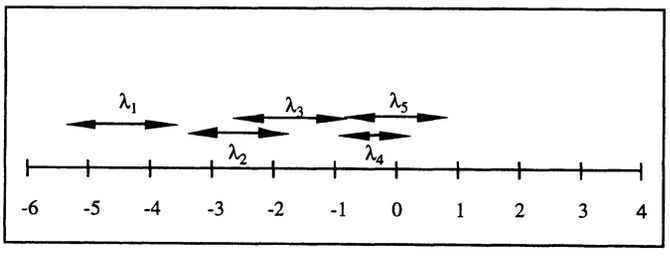

Figure 2. The Individual Approximate 95% Bonferroni Intervals for the Eigenvalues in the Chemical Example Calculated by the DLR Method.

In Section 5 we show that the DLR method, despite its simplicity, provides the same confidence intervals as the delta method. A more detailed proof was given by Bisgaard and Ankenman (1994).

5. EQUIVALENCE TO THE DELTA METHOD

The delta method is a method for approximating the expected value and the variance of functions of random variables. Suppose that a vector of random variables, $\mathbf{W}_n = (W_{n1}, \ldots, W_{nr})'$, has mean vector μ with covariance matrix $\boldsymbol{\Sigma}$. Furthermore, for any vector $\mathbf{w} = (w_1, \ldots, w_r)$, let $\mathbf{g}(\mathbf{w}) = (g_1(\mathbf{w}), \ldots, g_m(\mathbf{w}))'$ be a vector-valued function with real-valued component functions, $g_j(\mathbf{w})$, and nonzero differential at $\mathbf{w} = \mu$. Now let $U = \partial \mathbf{g}/\partial \mathbf{w}$, which is to say $U = \{u_{ij}\}_{r\times m}$, such that, for $i = 1, \ldots, r$ and $j = 1, \ldots, m$,

$$u_{ij} = \left.\frac{\partial g_j}{\partial w_i}\right|_{w=\mu}. \tag{9}$$

Then the expectation of $\mathbf{g}(\mathbf{W}_n)$ is approximately given by $\mathbf{g}(\mu)$ and the covariance matrix of $\mathbf{g}(\mathbf{W}_n)$ is approximately given by $U'\Sigma U$. Furthermore, Rao (1973, p. 388) showed that, if $\mathbf{W}_n$ is asymptotically normally distributed—that is, $\mathbf{W}_n \sim AN(\mu, \Sigma)$—then $\mathbf{g}(\mathbf{W}_n) \sim AN(\mathbf{g}(\mu), U'\Sigma U)$.

In the response surface context presented earlier, the eigenvalue estimators represented on the diagonal of $\hat{\Lambda}$ are the roots of the characteristic polynomial of $\hat{B}$. Thus they are functions of the estimators of the second-order coefficients of (1):

$$\hat{\beta}_q = (\hat{\beta}_{11}, \hat{\beta}_{12}, \ldots, \hat{\beta}_{1k} \vdots \hat{\beta}_{22}, \hat{\beta}_{23}, \ldots, \hat{\beta}_{2k} \vdots \ldots \vdots \hat{\beta}_{kk})'. \tag{10}$$

Hence using the mean and covariance matrix of $\hat{\beta}_q$ provided by a standard OLS fit of (1), we can in principle use the delta method to find approximate standard errors of the estimators in $\hat{\Lambda}$.

This is in essence what was done by CCC. Specifically, they first used corollary 3.2.1 of Anderson (1984) to show that, because the transformation from $\hat{B}$ to $(\hat{\lambda}_1, \ldots, \hat{\lambda}_k, \hat{\mathbf{d}}_1, \ldots, \hat{\mathbf{d}}_k)$ is one-to-one when each $\hat{\mathbf{d}}_i$ vector is required to have a positive first element and $\lambda_i < \lambda_{i+1}$ for all i, then $(\hat{\lambda}_1, \ldots, \hat{\lambda}_k, \hat{\mathbf{d}}_1, \ldots, \hat{\mathbf{d}}_k)$ is the maximum likelihood estimator of $(\lambda_1, \ldots, \lambda_k, \mathbf{d}_1, \ldots, \mathbf{d}_k)$. They used this result to show that the $\hat{\lambda}_i$'s are asymptotically distributed as a k-variate normal distribution with mean vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$. They also provided a procedure, involving the matrix-to-vector operator of a symmetric matrix for calculating large-sample confidence intervals for specific eigenvalues.

To ease the discussion of matrix-to-vector operators, we will now review pertinent material from Henderson and Searle (1979). The vech operator is defined for an arbitrary *symmetric* matrix,

$$C = \begin{bmatrix} c_{11} & & & \text{sym.} \\ c_{21} & c_{22} & & \\ \vdots & \vdots & \ddots & \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{bmatrix},$$

as vech $C = (c_{11}, c_{21}, \ldots, c_{p1} \vdots c_{22}, \ldots, c_{p2} \vdots \ldots \vdots c_{pp})'$. Note that the standard vec operator, as defined by Henderson and Searle (1979) and others, stacks the columns of any $m \times n$ matrix below each other to form an $mn \times 1$ vector. The vech operator, on the other hand, takes into account the redundancy of a symmetric matrix. In their procedure, CCC used the notation "vec" for what Henderson and Searle (1979) referred to as the vech operator. To avoid confusion, we will consistently use the more common notation of Henderson and Searle (1979).

We now define two matrices that relate the two matrix-to-vector operators. Recalling that k is the number of factors under study, the matrices G and L are such that, for any $k \times k$ symmetric matrix C,

$$\text{vec } C = G \text{ vech } C \tag{11}$$

and

$$L = (G'G)^{-1}G'. \tag{12}$$

Henderson and Searle (1979) showed that G is unique and that vech $C = L$ vec C. For further useful properties of the vec and vech operators, see Magnus and Neudecker (1988), Graybill (1983), Searle (1982), and Graham (1981).

We now apply the delta method to obtain standard errors for the estimated eigenvalues. First note that, for G as defined in (11), $\hat{\beta}_q$ from (10) can be written as $\hat{\beta}_q = G'$ vec $\hat{B}$. Because $\hat{\Lambda} = \hat{D}'\hat{B}\hat{D}$, it follows from Henderson and Searle (1979, eq. 5) that vec $\hat{\Lambda} = (\hat{D}' \otimes \hat{D}')$vec $\hat{B}$, where $\otimes$ indicates the standard Kronecker product.

If we now define the vector $\hat{\lambda}_q = G'$ vec $\hat{\Lambda} =$ vech $\hat{\Lambda}$, the eigenvalue estimators are specific elements of $\hat{\lambda}_q$ as follows: $\hat{\lambda}_q = (\hat{\lambda}_1, 0, \ldots, 0 \vdots \hat{\lambda}_2, 0, \ldots, 0 \vdots \ldots \vdots \hat{\lambda}_k)'$. Because Henderson and Searle (1979, eq. 24) showed that $G'(\hat{D}' \otimes \hat{D}') = G'(\hat{D}' \otimes \hat{D}')L'G'$, then

$$\begin{aligned} G' \text{ vec } \hat{\Lambda} &= G'(\hat{D}' \otimes \hat{D}')\text{vec } \hat{B} \\ &= G'(\hat{D}' \otimes \hat{D}')L'G' \text{ vec } \hat{B}. \end{aligned}$$

We differentiate to find that

$$\frac{\partial \hat{\lambda}_q}{\partial \hat{\beta}_q} = \frac{\partial G' \text{ vec } \hat{\Lambda}}{\partial G' \text{ vec } \hat{B}} = L(\hat{D} \otimes \hat{D})G.$$

We then apply the delta method to get the estimated covariance matrix for $\hat{\lambda}_q$ as follows:

$$\left(\frac{\partial \hat{\lambda}_q}{\partial \hat{\beta}_q}\right)' V_X \frac{\partial \hat{\lambda}_q}{\partial \hat{\beta}_q} = G'(\hat{D}' \otimes \hat{D}')L'V_X L(\hat{D} \otimes \hat{D})G, \tag{13}$$

where V_X is the estimated covariance matrix of $\hat{\beta}_q$ and thus is the lower $[k(k+1)/2] \times [k(k+1)/2]$ submatrix of $s^2(X'X)^{-1}$, from the initial fitting of (1). By partitioning X, the regressor matrix from the first regression of the DLR method, as follows:

$$\begin{aligned} X = [X_1 \vdots X_2] = [[\mathbf{1}, \mathbf{x}_1, \ldots, \mathbf{x}_k] \\ \vdots [\mathbf{x}_{11}, \mathbf{x}_{12}, \ldots, \mathbf{x}_{1k} \vdots \mathbf{x}_{22}, \mathbf{x}_{23}, \ldots, \mathbf{x}_{2k} \vdots \ldots \vdots \mathbf{x}_{kk}]], \end{aligned} \tag{14}$$

we can get an expression for V_X as follows: $V_X = s^2[X_2'X_2 - X_2'X_1[X_1'X_1]^{-1}X_1'X_2]^{-1}$.

We now use the DLR method to find an estimated covariance matrix for $\hat{\lambda}_q$. It will be convenient here to leave Z, the regressor matrix for the second DLR regression, in its natural order; thus

$$Z = [\mathbf{1}, \mathbf{z}_1, \ldots, \mathbf{z}_k \vdots \mathbf{z}_{11}, \ldots, \mathbf{z}_{1k} \vdots \mathbf{z}_{22}, \ldots, \mathbf{z}_{2k}, \vdots \ldots \vdots \mathbf{z}_{kk}]. \tag{15}$$

For X_1 and X_2 as defined in (14), Appendix 1 shows that

$$Z = [X_1 M \vdots X_2 F], \tag{16}$$

where

$$M = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{D} \end{bmatrix}$$

and $F = G'(\hat{D} \otimes \hat{D})L'$.

Let us denote the estimated covariance matrix for the second-order terms of the second regression as S_q, which is the lower $[k(k+1)/2] \times [k(k+1)/2]$ submatrix of $s^2(Z'Z)^{-1}$. Using (16) and some manipulation, S_q can be reduced as follows:

$$\begin{aligned} S_q &= s^2[F'[X_2'X_2 - X_2'X_1[X_1'X_1]^{-1}X_1'X_2]F]^{-1} \\ &= F^{-1}V_X(F^{-1})'. \end{aligned}$$

It follows from Magnus and Neudecker (1988, p. 50, eq. 12) that $F^{-1} = G'(\hat{D}' \otimes \hat{D}')L'$. Thus

$$S_q = F^{-1}V_X(F^{-1})' = G'(\hat{D}' \otimes \hat{D}')L'V_X L(\hat{D} \otimes \hat{D})G.$$

Therefore, S_q, the estimated covariance matrix of all the second-order terms of the second regression of the DLR method, is identical to the estimated covariance matrix for $\hat{\lambda}_q$ using the delta method (13) and hence the two methods are equivalent. Note that, when CCC applied the delta method, they used the vector vech $\hat{B}$ instead of $\hat{\beta}_q = G'$ vec $\hat{B}$. These two vectors differ only by some constants, and thus the results are the same.

6. CONCLUSIONS

The DLR method introduced in this article provides a simple new approach for estimating the standard error of the eigenvalues in the canonical form of a second-order response surface model. Its key advantage, beyond being simple to execute, is that it is based on the simple idea of rotating the original coordinate system such that the new axes coincide with the canonical axes of the response surface, thus providing a simple geometric understanding of the method. It is intended to provide quick estimates of the standard error of the eigenvalues using standard linear regression software packages. The approximate confidence intervals of the eigenvalues constructed using these standard errors can be used to interpret the shape of the response

surface. In this article we have shown that the method is a linear approximation and is equivalent to the delta method used by Carter et al. (1990). Because of its simple geometry, however, our method is substantially more accessible to the average practitioner.

ACKNOWLEDGMENTS

This work was supported by a grant from the Alfred P. Sloan Foundation and the National Science Foundation (EEC 8721545). We thank Douglas Bates, George Box, Howard Fuller, Ilya Gertsbakh, Spencer Graves, and John Peterson for helpful discussions while writing this article. We also thank the Center for Quality and Productivity Improvement's reports committee, the *Technometrics* editors, and referees for helpful comments on an earlier draft.

APPENDIX: A USEFUL MATRIX RESULT

Lemma 1. Let G and L be as defined in (11) and (12), respectively. Moreover, let $\hat{D}$ be as in (3) and let X be defined and partitioned as in (14). Then Z, as defined in (15), can be partitioned as follows: $Z = [X_1 M \vdots X_2 F]$, where

$$M = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{D} \end{bmatrix},$$

and $F = G'(\hat{D} \otimes \hat{D})L'$.

Proof. From the definition $\mathbf{z} = \hat{D}'\mathbf{x}$, it follows that the first $k+1$ columns of Z are

$$X_1 \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{D} \end{bmatrix} = X_1 M,$$

and thus we only need to show that the last $k(k+1)/2$ columns are $X_2 F$.

Let $X_D = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k]$ be the $n \times k$ design matrix from the original model in (1), and let $\mathbf{e}_i$ be the n-length basis vector of all zeros except for a one in the ith position. It is readily verified that, if $K = \sum_{i=1}^{n} \mathbf{e}_i \otimes \mathbf{e}_i \mathbf{e}_i'$, then the matrix X in (14) can be formed by $X = [\mathbf{1} \vdots X_D \vdots K'(X_D \otimes X_D)L']$. For $X = [X_1 \vdots X_2]$ as partitioned in Equation (14), we note that $X_1 = [\mathbf{1} \vdots X_D]$ and $X_2 = K'(X_D \otimes X_D)L'$. Similarly, in the $\mathbf{z}$-coordinate space the design matrix is $X_D \hat{D}$ and thus $Z = [\mathbf{1} \vdots X_D \hat{D} \vdots K'(X_D \hat{D} \otimes X_D \hat{D})L']$. As we did with X, we partition $Z = [Z_1 \vdots Z_2]$, where $Z_1 = [\mathbf{1} \vdots X_D \hat{D}]$ and $Z_2 = K'(X_D \hat{D} \otimes X_D \hat{D})L'$.

Because $X_2 = K'(X_D \otimes X_D)L'$, using the properties of the Kronecker product and Henderson and Searle (1979, eq. 24), we find that

$$\begin{aligned} Z_2 &= K'(X_D \hat{D} \otimes X_D \hat{D})L' \\ &= K'(X_D \otimes X_D)(\hat{D} \otimes \hat{D})L' \\ &= K'(X_D \otimes X_D)L'G'(\hat{D} \otimes \hat{D})L' \\ &= X_2 F, \end{aligned}$$

where $F = G'(\hat{D} \otimes \hat{D})L'$. Thus, the statement of the lemma follows.

[Received March 1994. Revised February 1996.]

REFERENCES

Anderson, T. W. (1984), *An Introduction to Multivariate Statistical Analysis* (2nd ed.), New York: John Wiley.

Ankenman, B. E. (1994), "Inference for the Eigenvalues in Second Order Response Surface Models," unpublished Ph.D. thesis, University of Wisconsin-Madison, Dept. of Industrial Engineering.

Bisgaard, S., and Ankenman, B. E. (1994), "Standard Errors of the Eigenvalues of Second Order Response Surface Models," Report 113: University of Wisconsin-Madison, Center for Quality and Productivity Improvement.

Box, G. E. P. (1954), "The Exploration and Exploitation of Response Surfaces: Some General Considerations and Examples," *Biometrics*, 10, 16–60.

Box, G. E. P., and Draper, N. R. (1987), *Empirical Model-Building and Response Surfaces*, New York: John Wiley.

Box, G. E. P., and Hunter, J. S. (1954), "A Confidence Region for the Solution of a Set of Simultaneous Equations With an Application to Experimental Design," *Biometrika*, 41, 109–199.

Box, G. E. P., and Wilson, K. B. (1951), "On the Experimental Attainment of Optimum Conditions," *Journal of the Royal Statistical Society*, Ser. B, 13, 1–45.

Box, G. E. P., and Youle, P. V. (1955), "The Exploration and Exploitation of Response Surfaces: An Example of the Link Between the Fitted Surface and the Basic Mechanism of the System," *Biometrics*, 11, 287–323.

Carter, W. H., Jr., Chinchilli, V. M., and Campbell, E. D. (1990), "A Large-Sample Confidence Region Useful in Characterizing the Stationary Point of a Quadratic Response Surface," *Technometrics*, 32, 425–435.

Carter, W. H., Jr., Chinchilli, V. M., Myers, R. H., and Campbell, E. D. (1986), "Confidence Intervals and Improved Ridge Analysis of Response Surfaces," *Technometrics*, 28, 339–346.

Clarke, G. P. Y. (1987), "Approximate Confidence Limits for a Parameter Function in Nonlinear Regression," *Journal of the American Statistical Association*, 82, 221–230.

Draper, N. R., and Smith, H. (1981), *Applied Regression Analysis* (2nd ed.), New York: John Wiley.

Graham, A. (1981), *Kronecker Products and Matrix Calculus: With Applications*, Chichester, U.K.: Halstead Press, Ellis Harwood.

Graybill, F. A. (1983), *Matrices With Applications in Statistics* (2nd ed.), Belmont, CA: Wadsworth.

Henderson, H. V., and Searle, S. R. (1979), "Vec and Vech Operators for Matrices With Some Uses in Jacobians and Multivariate Statistics," *Canadian Journal of Statistics*, 7, 65–81.

Khuri, A. I., and Cornell, J. A. (1987), *Response Surfaces: Designs and Analyses*, New York: Marcel Dekker.

Magnus, J. R., and Neudecker, H. (1988), *Matrix Differential Calculus With Applications in Statistics and Econometrics*, Chichester, U.K.: John Wiley.

Myers, R. H. (1976), *Response Surface Methodology*, Boston: Allyn and Bacon.

Peterson, J. J. (1993), "A General Approach to Ridge Analysis With Confidence Intervals," *Technometrics*, 35, 204–214.

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications* (2nd ed.), New York: John Wiley.

Searle, S. R. (1982), *Matrix Algebra Useful for Statistics*, New York: John Wiley.