A Large-Sample Confidence Region Useful in Characterizing the Stationary Point of a Quadratic Response Surface

Walter H. Carter, Jr.
Vernon M. Chinchilli
Eleanor D. Campbell

Department of Biostatistics
Virginia Commonwealth University
Richmond, VA 23298-0032

A second-order response-surface model is often used to approximate the relationship between a response variable and a set of explanatory variables. The nature of the stationary point of the surface can be determined by considering the eigenvalues of the matrix of the model's second-order terms. Since the elements of this matrix are estimated from the data, however, it follows that the eigenvalues are random variables. Hence the sampling properties of these eigenvalues should be considered in characterizing the nature of the stationary point. In this article, it is shown how a confidence region around these eigenvalues can be used to aid in the characterization of the stationary point and in an improved ridge analysis of the response surface. The delta method is used to construct an approximate confidence region for these eigenvalues. Box and Draper (1987) gave a result that simplifies the calculation of such a confidence region for rotatable or nearly rotatable designs. A simulation study was performed to determine and compare the coverage probabilities for the confidence regions for some frequently used experimental designs. For the cases considered, the approximate procedure developed in this article has the desired small-sample properties and is appropriate for rotatable and nonrotatable designs.

KEY WORDS: Asymptotic distribution; Canonical analysis; Delta method; Eigenvalues; Ridge analysis.

1. INTRODUCTION

It is often the case that a quadratic model provides an adequate approximation, in a restricted region of interest, of the relationship between a response variable, \( y \), and a vector of \( k \) explanatory variables, \( \mathbf{x} = [x_1, x_2, \ldots, x_k]' \). The model can be written as

\[
E(y) = \beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x},
\]

where \( \mathbf{B} = [\beta_{11}, \beta_{12}, \ldots, \beta_{kk}]' \) and

\[
\begin{bmatrix}
\beta_{11} & \beta_{12}/2 & \cdots & \beta_{1k}/2 \\
\beta_{21} & \beta_{22} & \cdots & \beta_{2k}/2 \\
\vdots & \vdots & \ddots & \vdots \\
\beta_{k1} & \beta_{k2} & \cdots & \beta_{kk}
\end{bmatrix}
\]

In practice, one does not know the values of the parameters in (1.1) and as a result must use estimates of them. In the sequel, "\( \cdot \)" will indicate that the parameter(s) in "\( \cdot \)" have been replaced by estimator(s) that are at least consistent and asymptotically normally distributed. In many situations, such estimators will be obtained through use of the method of maximum likelihood.

A complete exploration of the fitted surface involves, typically, the estimation of the stationary point \( \mathbf{x}_0 = -\mathbf{B}^{-1}\mathbf{\beta} \), the estimated response at the stationary point \( \hat{y} = \beta_0 - \mathbf{B}^{-1}\beta_\mathbf{B}/4 \), and a plot of the surface or its contours when \( k \leq 3 \). A canonical analysis (Box and Draper 1987) can be used to further elucidate the properties of the stationary point and the response surface. Such an analysis is especially useful when \( k > 3 \) due to the inability to plot these higher-dimensional surfaces.

The stationary point can be characterized as a point of maximum or minimum response according to whether \( \mathbf{B} \) is negative or positive definite, respectively. When \( \mathbf{B} \) is indefinite, the stationary point is a saddle point. Consequently, information about the eigenvalues of \( \mathbf{B} \) can aid in the characterization of the stationary point and as a result a characterization of the response surface. Indeed, when one or more of the eigenvalues of \( \mathbf{B} \) is near zero, the surface may...
not have a unique stationary point but rather a stationary ridge.

The user of response-surface methodology can gain insight into the nature of the stationary point and the nature of the response system from the eigenvalues of $\hat{B}$. Since these eigenvalues are associated with a matrix whose elements are random, however, they in turn are random variables. Consequently, one cannot always be sure of the algebraic sign of these eigenvalues, especially those near 0. In such a situation, as a result of the uncertainty in sign, it follows that there must be uncertainty in the characterization of the stationary point. Carter, Chinchilli, Myers, and Campbell (1986) developed a conservative method for obtaining a confidence region for the eigenvalues of $B$ when the elements of $\hat{B}$ are normally distributed. In that article, the authors presented an example in which the eigenvalues of $\hat{B}$ were all negative, indicating a unique point of maximum response, but the confidence region around the smallest eigenvalue included 0. Such a finding would indicate that instead of a unique point of maximum response the response surface may possess a stationary ridge of maximal responses. This could have important implications to the experimenter in that some of the points of the stationary ridge may be associated with operating conditions that are easier or more economical to achieve. The method used by those authors produces a conservative result and is computer intensive and, therefore, not likely to be used routinely. It is the purpose of the present article to develop a computationally convenient method that uses the eigenvalues of $B$ to construct an approximate 100(1 - $\alpha$)% confidence region for the eigenvalues of $B$.

2. CONSTRUCTION OF AN APPROXIMATE CONFIDENCE REGION FOR THE EIGENVALUES OF $B$

To develop an approximate confidence region for $\lambda$, the vector of eigenvalues of $B$, it is convenient to apply the vec operator to various matrices. By definition, $\text{vec}(A)$ in which $A$ is a $k \times k$ symmetric matrix is the column vector of the $k(k+1)/2$ distinct elements of $A$. Specifically,

$$\text{vec}(B) = [\beta_{11}, \beta_{12}, \ldots, \beta_{kk}] = \beta_{11} : \beta_{12} : \cdots : \beta_{kk} : \cdots : \beta_{kk}^T.$$ 

Let $\beta^*$ denote the column vector of $p = (k + 1)(k + 2)/2$ parameters of the model given in (1.1); that is,

$$\beta^* = [\beta_0 : \beta_1 : \cdots : \beta_k : \beta_{11}, \beta_{12}, \cdots, \beta_{kk}]^T.$$ 

When $\beta^*$ is estimated by $\hat{\beta}$ for which $\beta^* \sim N_p(\beta^*, V^*)$ (at least asymptotically), it follows that $\text{vec}(\hat{B}) \sim N_{k(k+1)/2}(\text{vec}(B), V)$ (at least asymptotically), where $V$ is the lower $k(k+1)/2 \times k(k+1)/2$ submatrix of $V^*$ with multipliers of $1/2$ and $1/2$ in the appropriate cells. Define $H$, a $k(k+1)/2 \times k$ matrix, such that

$$H^T = \begin{bmatrix}
\text{vec}(2d_1, d^T - \text{diag}(d_1, d_1)) \\
\text{vec}(2d_2, d^T - \text{diag}(d_2, d_2)) \\
\vdots \\
\text{vec}(2d_k, d^T - \text{diag}(d_k, d_k))
\end{bmatrix},$$

where $d_i (i = 1, 2, \ldots, k)$ is the $i$th eigenvector of $B$. For convenience, we assume that the eigenvalues of $B$, denoted by $\lambda = [\lambda_1, \ldots, \lambda_k]^T$, are distinct. For each $i = 1, 2, \ldots, k$, $-\lambda_i$ can be substituted for $d_i$, so there are actually $2^k$ realizations of the orthogonal matrix $D = [d_1, d_2, \ldots, d_k]$. To make $D$ unique, we use the convention that the first row of $D$ is positive. It follows directly from Theorem 1 in the Appendix that an approximate 100(1 - $\alpha$)% confidence region for $\lambda$ is given by

$$\hat{C}_\lambda = \{\lambda \in \mathbb{R}^k: (\lambda - \lambda)^T(\hat{H}^T V \hat{H})^{-1}(\lambda - \lambda) \leq \chi^2_{k, \alpha} \}.$$ 

(2.1)

where $\lambda$ is the vector of eigenvalues of $\hat{B}$, $\hat{H}$ is the estimator of $H$ formed from eigenvectors $d_i$ of $\hat{B}$, and $\chi^2_{k, \alpha}$ is the 1 - $\alpha$th quantile of a chi-squared distribution with $k$ df. The same theorem can be used to obtain an approximate 100(1 - $\alpha$)% confidence interval for the $i$th ordered eigenvalue of $B$, $\hat{\lambda}_i$, where $\hat{\lambda}_1 < \hat{\lambda}_2 < \cdots < \hat{\lambda}_k$. The interval is given by

$$\hat{\lambda}_i - Z_{1 - \alpha/2}[e_i^T \hat{H}^T V \hat{H}e_i]^{1/2} \leq \hat{\lambda}_i \leq \hat{\lambda}_i + Z_{1 - \alpha/2}[e_i^T \hat{H}^T V \hat{H}e_i]^{1/2}.$$ 

(2.2)

where $\hat{\lambda}_i$ is the $i$th ordered eigenvalue of $\hat{B}$, $e_i$ is a $k \times 1$ vector of zeroes with a 1 in the $i$th row, and $Z_{1 - \alpha/2}$ is the $1 - \alpha/2$th quantile of the standard normal distribution. That $\hat{\lambda}_i$ estimates $\lambda_i$ is a consequence of Lemma 1 in the appendix.

In the special case in which a linear model with additive iid $N(0, \sigma^2)$ error terms is used to approximate the response relationship, $V^* = \sigma^2(X'X)^{-1}$, so each element of $V$ is a multiple of $\sigma^2$. In this situation,
the following procedures are recommended instead of (2.1) and (2.2), because they will have better coverage in small samples. Instead of (2.1), one should use

$$\tilde{C}_k = \{ \lambda \in \mathbb{R}^k : (\hat{\lambda} - \lambda)'(\hat{H}'V\hat{H})^{-1}(\hat{\lambda} - \lambda) \leq ks^2F_{k,n-p,1-a} \}.$$  

(2.3)

where

$$s^2 = y'(1 - X(X'X)^{-1}X'y)/(n - p)$$

and

$$F_{k,n-p,1-a}$$

is the 100(1 - a) percentile of the central F distribution with k and n - p df. Instead of (2.2), one should use

$$\tilde{\lambda}_i \leq \lambda_i \leq \tilde{\lambda}_i + t_{1-a/2,n-p}s[\hat{e}'\hat{H}'V\hat{H}\hat{e}]^{1/2}.$$  

(2.4)

3. AN ILLUSTRATION

Myers (1976, pp. 78–84) considered an example in which it was of interest to relate the sealing strength of a breadwrapper stock to a set of three explanatory variables. It was assumed that the linear model given by (1.1) with additive iid N(0, \sigma^2) error would adequately describe the underlying relationship. The method of least squares was used to estimate the model parameters from data collected from a rotatable, uniform precision central composite design.

For this example, the eigenvalues of B are -1.271, -1.117, and -0.562. Without considering the variability of the elements of B, it would be concluded that the surface had a unique stationary point and that the stationary point was associated with a maximum response. Carter et al. (1986) reported a conservative 90% confidence interval on the largest eigenvalue, \lambda_1, of B to be \([-9.43, 2.12]\). Since the interval includes 0, the authors pointed out that it is possible that the surface is a stationary ridge as opposed to one with a unique maximum point. To arrive at this interval, 10,240 different points were chosen from the 90% confidence region, C, around B*. From each of these, the B matrix was formed and the largest and smallest values of \lambda_1 were determined and used as endpoints of the conservative interval. Theoretically the endpoints of the interval should be \([\min \lambda_1, \max \lambda_1]\). By employing the Nelder–Mead (1965) simplex research procedure, a more accurate estimate of the conservative 90% confidence interval for \lambda_1 was found to be \([-1.667, 0.964]\). This clearly indicates the importance of employing a search over C as opposed to merely evaluating a large number of points from C.

Box and Draper (1987, p. 354) indicated without discussion that whenever an (approximately) rotatable design is used to fit a model such as that given in (1.1), the variances of the eigenvalues of B, \lambda_i (i = 1, 2, . . . , k) are (approximately) equal to the variances of the estimates of the pure quadratic terms in the model. Peterson (1988, p. 12) suggested a proof of this by noting that \bar{d}_i = \lambda_i = 1, 2, . . . , k and using the results of Myers (1976, pp. 146, 147) that give the variances of the predicted response at any point x for a rotatable design. By assuming that x = \bar{d} and recalling that \bar{d}_i = 1, it follows that var \bar{d}_i = var \hat{\beta}_i. Since the design used in this example is rotatable, this result is applicable. The estimated 90% confidence interval for \lambda_1 calculated with this variance is given by \([-1.082, -0.042]\). It is interesting to note that this approximate interval excludes 0.

Equation (2.3) was used to arrive at the approximate 90% confidence region for the vector of eigenvalues of B. Figure 1 presents a series of six two-dimensional plots of the confidence region. The plots were obtained by searching over a grid for all values of \lambda_1 and \lambda_3 satisfying (2.3) for fixed values of \lambda_2.

Equation (2.4) can be used to provide confidence intervals around each of the eigenvalues by proper choice of the vector of 0s and 1s in this expression. For example, the 100(1 - a)% confidence interval

![Figure 1. The 90% Confidence Region for \lambda for the Bread-Wrapper Example.](image-url)
around $\lambda_3$ is estimated by

$$\hat{\lambda}_3 - t_{1-a/2, n-p} s^2 [\hat{H}' V \hat{H} e]^{1/2} \leq \hat{\lambda}_3 \leq \hat{\lambda}_3 + t_{1-a/2, n-p} s^2 [\hat{H}' V \hat{H} e]^{1/2}, \quad (3.1)$$

where $t_{1-a/2, n-p}$ is the 100(1 - $a/2$) percentile of the Student-$t$ distribution with $n - p$ df. (If this approach is taken to calculate a confidence interval around each eigenvalue separately, some adjustment for multiple comparisons—e.g., Bonferroni—should be employed.) The estimated 90% confidence interval around the largest eigenvalue of $B$ is given by $[-1.321, .197].$

4. SIMULATION STUDY

For many of the designs developed for fitting second-order models—for example, the central composite design and the Box-Behnken designs—the number of experimental points required is only slightly higher than the number of model parameters. Rotatable or nearly rotatable designs are frequently used but not always; thus the use of the relationship mentioned by Box and Draper (1987, p. 354) is not always appropriate to calculate confidence intervals for the eigenvalues of $B$. Consequently, it is important to study the small-sample properties of the confidence regions given by (2.3) and (2.4) as well as those resulting from use of the Box-Draper result for rotatable and nonrotatable designs. If the asymptotic results are applicable to the sample sizes required by the designs mentioned previously and if nonrotatability affects the coverage probabilities, it would appear that the use of the methodology developed in this article would be justified.

Although the results of simulation studies must be considered as special cases, such studies provide a mechanism for obtaining a glimpse of how various procedures might behave in practice. In the current situation, the small-sample properties of the various confidence regions were evaluated by way of a simulation study. The parameters of the underlying model were those estimated from the bread-wrapper-stock data reported by Myers (1976). Independent and identically normally distributed error terms with 0 mean when $\sigma^2 = 1, 3, 5$ were generated and added to the value of the mean response at the treatment combinations required by:

1. A rotatable central composite design ($\alpha = 1.682$) with a single centerpoint
2. A uniform precision, rotatable central composite design ($\alpha = 1.682$) with six centerpoints
3. A central composite design in which $\alpha = 1.000$ and with a single centerpoint
4. A Box-Behnken design with no centerpoint
5. A central composite design in which $\alpha = 5.000$ and with a single centerpoint

The rationale behind the choice of these designs is as follows:

1. Each of the designs requires a relatively small number of observations. It is of interest to determine how close the coverage probabilities of the various approximated confidence regions are to their nominal levels. It is known that the performances of these designs—for example, the Box-Behnken design—as measured by the ability to estimate parameters or predict unobserved responses, can be improved by the addition of centerpoints. The purpose of the simulation study is to gain an indication of the adequacy of sample size. Thus, if the design without centerpoints permits approximate confidence regions with coverage probabilities nearly equal to the nominal level, the use of the approximate methods developed here would seem justified.

2. For the rotatable central composite design there are two ways to increase the total sample size and maintain rotatability: (a) increase the number of centerpoints and (b) replicate the entire experiment. The former method is clearly more economical than the latter. It is of interest to compare the coverage probabilities associated with the approximate confidence regions to gain an indication of the necessity to replicate the entire experiment to obtain adequate sample sizes.

3. Designs 3 and 4 each have been described as approximately rotatable. Presumably, the sense in which they are nearly rotatable is that the ratio of pure fourth to mixed fourth design moments is approximately equal to 3. The importance of rotatability has to do with the constancy of variables calculated at points $x$ equidistant from the design center. [Several authors have considered the development of criteria for measuring the nearness to spherical prediction variance contours, e.g., Khuri (1988) and Draper and Gutman (1988). Giovannitti-Jensen and Myers (1989) considered another highly informative approach to this problem.] Since Box and Draper (1987) indicated that the variances of the estimated pure quadratic regression coefficients for approximately rotatable designs, it is of interest to compare the performance of the approximate confidence regions developed in this article to those calculated from the Box-Draper approach.

4. Design 5 was chosen because it is not rotatable; that is, the ratio of the pure fourth design moments to the mixed fourth design moments is 157.25. (The value of $\alpha = 5$ was arbitrarily chosen.) It is of interest to compare the coverage probabilities of the approximate confidence regions developed here to that of the confidence region suggested by Box and Draper (1987). This will provide an indication of the ro-
Figure 2. Results From the Simulation Study for the $\alpha = 1.682$ (rotatable) Central Composite Design With A Single Centerpoint.
bustness of their method to departures from rotatability in addition to a comparison between the various methods of estimating the confidence regions. With the exception of the uniform precision rotatable central composite design, all designs were replicated 1, 2, 3, 4, and 6 times. The simulation study was designed so that for all ($\sigma^2$, experimental design, replications) triples, 500 experiments were generated. Specifically, the approximate 90% confidence region, defined in (3.1), around $\lambda(3)$: the 90% confidence interval around $\lambda(3)$ using the result mentioned by Box and Draper (1987, p. 354); and the approximate 90% confidence region, defined in (2.3), around $\lambda$ were estimated. The number of times the 90% confidence region contained the true vector or value, as appropriate, was recorded. To obtain an estimate of the variability associated with the simulation results, the 500 experiments were blocked into five sets of 100 experiments. The results of the study are summarized in Figures 2-6.

From studying these figures, the following conclusions can be drawn:

1. With few exceptions, the sample size associated with a single replication of the experiments considered seems to be adequate for the large-sample properties to hold. For the central composite design in which $\sigma = 1.000$ and $\sigma^2 = 5$, the experiment had to be replicated three times before the coverage probability for $\lambda(3)$ reached 90%, whereas when $\lambda$ was under consideration, one replication was sufficient (Fig. 4). For the Box–Behnken design when $\sigma^2 = 3$, 5, three and four replications, respectively, were required to give 90% coverage probability for $\lambda(3)$, whereas for $\lambda$ one replication was adequate when $\sigma^2 = 3$, four replications were required when $\sigma^2 = 5$ (Fig. 5).

2. For rotatable designs the coverage probabilities for the large-sample and Box–Draper intervals for $\lambda(3)$ are identical (Figs. 2 and 3). One replication of the experiment results in an adequate sample size for the asymptotic results to hold except possibly when $\sigma^2 = 3$ or 5 and interest is on the vector of eigenvalues $\lambda$ as opposed to the largest eigenvalue (Figs. 2 and 3).

3. It appears, except when $\sigma^2 = 5$, that a single replication of the central composite design in which $\alpha = 1.000$ results in a large enough sample size for either the method developed in this article or the method of Box and Draper to apply to $\lambda(3)$ (Fig. 4). When $\lambda$ is of interest, a single replication of the experiment seems to be adequate for each of the values of $\sigma^2$ considered (Fig. 4). A single replication of the Box–Behnken design when $\sigma^2 = 1$ appears to provide a large enough sample size for the coverage probabilities associated with the use of either method of approximating a confidence interval around $\lambda(3)$ and the approximate region around $\lambda$ to achieve the nominal level (Fig. 5). When $\sigma^2 = 3$, a single replication of the experiment is associated with a confidence interval calculated by the Box–Draper method whose coverage probability is not different from the nominal level. The approximate interval for $\lambda(3)$ developed in this article does not achieve the nominal level until the experiment has been replicated three times (Fig. 5). When $\sigma^2 = 5$, a single replication is not adequate for either method of computing approximate confidence intervals for $\lambda(3)$. Here, the Box–Draper method produces intervals that achieve the nominal coverage probability after two replications of the experiment. The intervals associated with the method developed in this article do not achieve the nominal level until the experiment has been replicated four times (Fig. 5). It is interesting to note that

---

Figure 3. Results From the Simulation Study For the $\alpha = 1.002$ Central Composite Design With Six Centerpoints—a Uniform Precision Design.
Figure 4. Results From the Simulation Study for the \( \alpha = 1.000 \) Central Composite Design With a Single Centerpoint.
Figure 5. Results From the Simulation Study for the Box-Behnken Design Without a Centerpoint.
Figure 6. Results From the Simulation Study for the $\sigma = 5.000$ Central Composite Design With a Single Centerpoint.
in all cases considered involving these two designs the Box–Draper method produced the interval associated with the higher point estimate for the coverage probability (Figs. 4 and 5).

4. For the nonrotatable design, the central composite design in which \( \alpha = 5.000 \), the intervals calculated using the Box–Draper result fail. The method developed in this article provides an interval around \( \lambda_{(3)} \) and a region around \( \lambda \) that achieve the nominal coverage probability after a single replication of the experiment (Fig. 6).

5. CONCLUSION

The methodology developed here results in an easily programmed procedure for computing a large-sample confidence region for the eigenvalues of the B matrix associated with a quadratic response surface. It was shown by Carter et al. (1986) that such a result can be useful in characterizing the stationary point and in performing an improved ridge analysis. Specifically, these authors showed that an improved ridge analysis of a quadratic response surface resulted when choices for the Lagrange multipliers were restricted to be larger (smaller) than the upper (lower) bound of the confidence interval about the largest (smallest) eigenvalue of B when locating a constrained maximum (minimum) response. The confidence intervals developed in this article can be used for this purpose, and they are more readily obtained.

The advantages offered by the new procedure are computational ease and a method for calculating confidence regions that achieves nearly nominal coverages. Since the new result is based on large-sample theory, it is important to determine if the sample sizes required by the designs developed for and frequently used to fit a quadratic response surface are adequate. In addition, Box and Draper (1987, p. 354) gave a result for rotatable and nearly rotatable designs that readily permits the construction of a confidence interval for the individual eigenvalues of B. The sensitivity of the coverage probabilities of these intervals to departures from rotatability is not known, however. To gain an indication of the appropriateness of the approximate method in some commonly used designs and to gain an indication of the effect of nonrotatability on the intervals constructed using the result of Box and Draper (1987), a simulation study was performed. Although any simulation study must be considered as a special case, some interesting observations were noted. For the rotatable designs considered, the two methods resulted in confidence intervals of about \( \lambda = 1.000 \) and a single centerpoint and the Box–Behnken design without a centerpoint for the example considered (Figs. 4 and 5). When a nonrotatable central composite design was considered, the method of Box and Draper produced confidence intervals whose coverage probabilities were far from the nominal value. In this situation, the coverage probabilities for the approximated confidence regions developed in this article were not different from the nominal level (Fig. 6).

For the rotatable central composite design with a single centerpoint when \( \sigma^2 = 3 \), the sample size associated with a single replication of the experiment was not sufficiently large for the approximate method developed here to produce a confidence region around \( \lambda \) with the nominal value for the coverage probability (Fig. 2). It is interesting to note that when the number of centerpoints is increased to six, the coverage probability associated with the confidence region around \( \lambda \) does not differ from the nominal level (Fig. 3). This serves as another illustration of the importance of the centerpoint in these designs.

Based on these findings, it appears that the methodology developed here provides a useful tool in the analysis of a quadratic response surface.

ACKNOWLEDGMENT

We acknowledge the contributions of the referees, the associate editor, and the editor in the preparation of this manuscript. The result of their efforts is a more complete and readable article.

APPENDIX: LARGE SAMPLE DISTRIBUTION OF \( \lambda \)

**Lemma 1.** Let \( \lambda_{(1)} < \cdots < \lambda_{(k)} \) denote the ordered and distinct eigenvalues of B, and let \( \mathbf{d}_{(1)}, \ldots, \mathbf{d}_{(k)} \) denote the corresponding eigenvectors such that \( D = [\mathbf{d}_{(1)}, \mathbf{d}_{(2)}, \ldots, \mathbf{d}_{(k)}] \) has its first row positive. Suppose that \( \hat{B} \), the estimator of B, is the maximum likelihood estimator and that it has an exact multivariate normal distribution. Let \( \hat{\lambda}_{(1)} \leq \cdots \leq \hat{\lambda}_{(k)} \) denote the ordered set of eigenvalues of \( \hat{B} \) and \( \mathbf{d}_{(1)}, \ldots, \mathbf{d}_{(k)} \) the corresponding eigenvectors. Then the maximum likelihood estimator of \( (\lambda_{(1)}, \mathbf{d}_{(1)}) \) is \( (\hat{\lambda}_{(1)}, \mathbf{d}_{(1)}) \).

**Proof.** The transformation from B to \( (\lambda_{(1)}, \mathbf{d}_{(1)}, \ldots, \mathbf{d}_{(k)}) \) is one-to-one. Then, according to corollary 3.2.1 of Anderson (1984), the maximum likelihood estimator of \( (\lambda_{(1)}, \ldots, \lambda_{(k)}, \mathbf{d}_{(1)}, \ldots, \mathbf{d}_{(k)}) \) is \( (\hat{\lambda}_{(1)}, \ldots, \hat{\lambda}_{(k)}, \mathbf{d}_{(1)}, \ldots, \mathbf{d}_{(k)}) \).

**Theorem 1.** Under the conditions of the lemma, \( n^{1/2}(\lambda - \lambda) \) has an asymptotic \( k \)-variate normal distribution with null mean vector and variance–covariance matrix \( nH'VH \), where H and V are as defined in Section 2.
Proof. Let \( t \in \mathbb{R}^k \) be any nonnull vector, and assume that \( f(B) = n^{1/2} t' \lambda \). Using a Taylor series expansion, it can be shown that

\[
f(\hat{B}) = f(B) + \text{vec}'(\hat{B} - B) \times \left\{ \frac{\partial f(B)}{\partial \text{vec}(B)} \right\} + o_p(n^{-1/2}),
\]

where \( o_p(n^{-1/2}) \) is the collection of terms which converge to 0 in probability when divided by \( n^{-1/2} \). The first partial derivatives of \( f(B) \) with respect to \( \text{vec}(B) \) are given by \( n^{1/2} H_t \), so that

\[
f(\hat{B}) = f(B) + n^{1/2} \text{vec}'(\hat{B} - B)H_t + o_p(n^{-1/2}).
\]

Therefore, for large \( n \), \( f(\hat{B}) \) is equivalent to a linear combination of the elements of \( \text{vec}(\hat{B} - B) \), which has the multivariate normal distribution described in Section 2. From this, it follows that \( n^{1/2} t' (\hat{\lambda} - \lambda) \) has an asymptotic normal distribution with null mean and variance \( nt'H_tVH_t \). Since this is true for every non-null \( t \in \mathbb{R}^k \), the statement of the theorem follows (see sec. 1.5.2 of Serfling 1980).

[Received July 1988. Revised May 1990.]

REFERENCES


Myers, R. H. (1976), Response Surface Methodology, Blacksburg, VA: Author (distributed by Edwards Brothers, Ann Arbor, MI).

