THE EFFECT OF RESPONSE SCALING ON THE DETECTION OF SINGULARITIES IN MULTI (U) FLORIDA UNIV GAINESVILLE DEPT OF STATISTICS A I KHURI SEP 87 TR-291 UNCLASSIFIED N8814-86-K-8859
The Effect of Response Scaling on the Detection of Singularities in Multiresponse Estimation

A. I. Khuri

Department of Statistics
Nuclear Sciences Center, University of Florida
Gainesville, FL 32611

Office of Naval Research
Mathematical Sciences Division (Code 411)
Arlington, VA 22217-5000

Approved for Public Release: Distribution Unlimited

Multiresponse model, Box-Draper estimation criterion, round-off error; linear dependencies, ill conditioning, eigenvalue analysis.

This paper proposes scaling of the responses before applying the eigenvalue analysis, which is used to detect singularities in multiresponse modeling. The paper shows how the eigenvalue analysis can be modified so that it applies to the scaled responses.
THE EFFECT OF RESPONSE SCALING ON THE DETECTION OF SINGULARITIES IN MULTIRESPONSE ESTIMATION

By

André I. Khuri
Department of Statistics
University of Florida
Gainesville, Florida 32611
THE EFFECT OF RESPONSE SCALING ON THE DETECTION OF SINGULARITIES IN
MULTI RESPONSE ESTIMATION

A. I. KHURI
Department of Statistics
University of Florida
Gainesville, Florida 32611

ABSTRACT

Box and Draper (1965) introduced a criterion for the estimation of parameters from a multiresponse model. This criterion can lead to misleading results in the presence of linear relationships among the responses. Box et al. (1973) proposed a procedure for detecting the existence of such relationships when the multiresponse data are subject to round-off errors. The procedure, however, can be adversely affected by large differences in the orders of magnitude of the responses as well as in the units of measurement on which the responses are expressed. It is, therefore, necessary to scale the responses prior to the application of that procedure. In this article, I discuss the effect of scaling on the implementation of the eigenvalue analysis by Box et al. (1973). Two numerical examples are given for illustration.

KEY WORDS: Multiresponse model; Box-Draper estimation criterion; Round-off error; Linear dependencies; Ill conditioning; Eigenvalue analysis.
1. INTRODUCTION

Consider the general multiresponse model

\[ y_{ui} = \mathcal{F}_i(x_u, \theta) + \epsilon_{ui}, \quad u=1,2,\ldots,n; \ i=1,2,\ldots,r, \]  

where

- \( y_{ui} \) is the value of the \( i^{th} \) response at the \( u^{th} \) experimental run,
- \( \mathcal{F}_i(x_u, \theta) \) is the expected value of \( y_{ui} \),
- \( x_u \) is a vector of values of \( k \) input variables, denoted by \( x_1, x_2, \ldots, x_k \), at the \( u^{th} \) experimental run,
- \( \theta \) is a vector of \( p \) unknown parameters, and
- \( \epsilon_{ui} \) is a random error associated with \( y_{ui} \).

Let \( y_i \) and \( \epsilon_i \) denote, respectively, the vector of values of the \( i^{th} \) response and the associated vector of random errors (\( i=1,2,\ldots,r \)). The multiresponse model in (1.1) can be written in the form

\[ Y = \mathbf{F} + \xi, \]  

where \( Y = [y_1 : y_2 : \cdots : y_r] \), \( \mathbf{F} \) is an \( n \times r \) matrix whose \((u,i)^{th}\) element is \( \mathcal{F}_i(x_u, \theta) \) and

\[ \xi = [\xi_1 : \xi_2 : \cdots : \xi_r]. \]

According to Box and Draper (1965), estimates of the elements of \( \theta \) can be obtained by minimizing the determinant, \( |\Gamma(\theta)| \), of the matrix \( \Gamma(\theta) \) with respect to \( \theta \), where

\[ \Gamma(\theta) = (Y - \mathbf{F})'(Y - \mathbf{F}). \]  

We refer to this method of estimation as the Box-Draper estimation criterion. Box et al. (1973) pointed out that when linear relationships exist among the columns of \( Y \), the matrix \( \Gamma(\theta) \) becomes singular, that is, \( |\Gamma(\theta)| = 0 \). In practical situations, the multiresponse data are subject to round-off errors, which cause the determinant of \( \Gamma(\theta) \) to be different from zero and to change as \( \theta \) is changed. Minimization of this determinant under such conditions will produce nonsensical results (Box et al. 1973 and McLean et al. 1979).

In order to detect singularities in \( \Gamma(\theta) \), Box et al. (1973) proposed a procedure in which the
eigenvalues of the matrix $DD'$ are examined, where

$$D = Y'(I_n - \frac{1}{n})$$

(1.4)

In (1.4), $I_n$ and $\frac{1}{n}$ denote the identity matrix and the matrix of ones, respectively, both of order $n \times n$. Box et al. (1973) showed that $m$ linearly independent relationships must exist among the responses if and only if the matrix $DD'$ has a zero eigenvalue of multiplicity $m$. As mentioned earlier, none of the eigenvalues of $DD'$ will be exactly equal to zero because of round-off errors in the multiresponse data. Small eigenvalues of $DD'$ should, therefore, be examined to determine if they correspond to linear dependencies among the responses.

If $\lambda^*$ is a small eigenvalue of $DD'$, then in the presence of round-off error only and for a sufficiently small $\delta$, the expected value of $\lambda^*$ is approximately equal to

$$E(\lambda^*) = (n-1)\sigma_{re}^2$$

(1.5)

where $\sigma_{re}^2 = \delta^2/3$ is the round-off error variance. Formula (1.5) is valid under the assumption that the round-off errors are statistically independent and have the uniform distribution $U(-\delta,\delta)$. See Box et al. (1973). An approximate upper bound on the variance of $\lambda^*$ was given in Khuri and Conlon (1981) as

$$\text{Var}(\lambda^*) \leq \left[ \frac{9nr}{5} + nr(nr-1) - (n-1)^2 \right] \sigma_{re}^4$$

(1.6)

Formulas (1.5) and (1.6) can be used to determine whether a small eigenvalue of $DD'$ should be considered as zero, an indication of a singularity in the matrix $\Gamma(\bar{q})$ in (1.3). Box et al. (1973) discussed possible remedies when such a situation occurs, which include dropping some of the responses that are influential contributors to the singularity.

Quite often, the responses have different units of measurement, which cause them to have widely different orders of magnitude. Furthermore, the round-off errors in all the responses may not be identically distributed as $U(-\delta,\delta)$. All of these factors can seriously affect the eigenvalue analysis described earlier as will be seen in more detail in Section 2. To remedy this numerical inconsistency, the responses should be scaled before proceeding with the eigenvalue analysis. This action will
obviously alter formulas (1.5) and (1.6) for the expected value and variance of a small eigenvalue of $DD'$. A modification of these formulas will be presented in Section 3.

2. SCALING OF THE RESPONSES AND MEASURES OF NEAR SINGULARITY

In a general multiresponse situation, the response variables usually have distinct physical meanings, distinct units of measurement, and widely different values. Such scale imbalance makes it difficult to interpret the results of the eigenvalue analysis as described in Section 1. To avoid this difficulty, the responses must be scaled first.

Let $W$ be an $r \times r$ diagonal matrix whose $i$th diagonal element, $w_i$, is defined as

$$w_i = \left[ \frac{1}{n} \sum_{u=1}^{n} (y_{ui} - \bar{y}_i)^2 \right]^{\frac{1}{2}}, \quad i=1,2,...,r.$$  \hspace{1cm} (2.1)

where $y_{ui}$ is the value of the $i$th response at the $u$th experimental run and $\bar{y}_i = \frac{1}{n} \sum_{u=1}^{n} y_{ui}$. Consider the $n \times r$ matrix $S$ given by

$$S = YW^{-1},$$  \hspace{1cm} (2.2)

where $Y$ is the matrix of multiresponse data in (1.2). The matrix $S$ is a linear transform of $Y$, which results from dividing the $i$th column of $Y$ by $w_i (i=1,2,...,r)$. It is known that estimates of the parameters from the multiresponse model (1.2) are invariant under this scaling convention (see Bates and Watts 1985, p. 330). Using (2.2) in (1.4) we obtain

$$D = WS'(I_n - J_n/n).$$  \hspace{1cm} (2.3)

Let us now define the $r \times n$ matrix, $B$, as
\[ B = S'(I_n - J_n/n). \] (2.4)

From (2.3) and (2.4) we have

\[ B = W^{-1}D. \] (2.5)

or equivalently,

\[ B' = (I_n - J_n/n)YW^{-1}. \] (2.6)

We note that the columns of \( B' \) have unit lengths and that the matrix \( BB' \) is in correlation form.

Furthermore,

\[ BB' = W^{-1}DD'W^{-1} \] (2.7)

Thus, when the responses are scaled as in (2.2), the matrix \( DD' \), used in the eigenvalue analysis of Box et al. (1973), is transformed into the matrix \( BB' \).

In case of a singularity in the matrix \( \Gamma(q) \) (see 1.3) and in the presence of round-off errors in the multiresponse data, both \( DD' \) and \( BB' \) are near singular, that is, their determinants are close to zero. In this case, the columns of \( D' \) as well as those of \( B' \) are nearly linearly dependent, or multicollinear. In other words, if a singularity exists in the matrix \( \Gamma(q) \), then both \( D' \) and \( B' \) will suffer from ill conditioning. Their degrees of ill conditioning, however, can be quite different.

As a measure of ill conditioning of the matrix \( B' \) and that of \( D' \), I use the condition numbers, \( \kappa(B') \) and \( \kappa(D') \), respectively. By definition, the condition number of \( B' \) is

\[ \kappa(B') = \left[ \frac{e_{\text{max}}(BB')}{e_{\text{min}}(BB')} \right]^\frac{1}{2}. \]
where $e_{\max}(BB')$ and $e_{\min}(BB')$ denote the largest and smallest eigenvalues of the matrix $BB'$. The condition number of $D'$ is similarly defined. The larger the condition number, the more ill conditioned the matrix. Another indicator of ill conditioning is provided by the variance inflation factors (VIF). Since the columns of $B'$ are centered and scaled for unit length, the VIF's associated with this matrix are, by definition, the diagonal elements of the matrix $(BB')^{-1}$. Large VIF's indicate ill conditioning and can help diagnose its nature. More specifically, large VIF's correspond to responses that are involved in the multicollinearity in the columns of $B'$.

It was mentioned earlier that the degree of ill conditioning of $D'$ can be quite different from that of $B'$. To see this, let us consider (2.5) and the inequality given in Belsley et al. (1980, p. 182). Then,

$$\kappa(B') \geq \kappa(D')/\kappa(W).$$

Since $W$ is diagonal, its condition number is given by

$$\kappa(W) = \max(w_i)/\min(w_i),$$

where $\min(w_i)$ and $\max(w_i)$ denote, respectively, the smallest and largest of the $w_i$'s defined in (2.1). Inequality (2.8) can be rewritten as

$$\kappa(B') \geq \kappa(D')\left[\min(w_i)/\max(w_i)\right].$$

If $\kappa(D')$ is very large and $\min(w_i)$ is very small as compared to $\max(w_i)$, then $\kappa(B')$ can be small, but need not be. Thus, ill conditioning can be improved by the scaling convention of (2.2), especially
when the \( w_i \)'s in (2.1) are widely different. As a matter of fact, this kind of scaling has "near-optimal" properties in the sense that

\[
\kappa(B') \leq \sqrt{r} \min_{Q \in \Omega} [\kappa(B'Q)], \tag{2.11}
\]

where \( \Omega \) denotes the set of all nonsingular diagonal matrices of order \( r \times r \) (see Belsley et al. 1980, p. 185). In particular, from (2.5) and (2.11) we conclude that

\[
\kappa(B') \leq \sqrt{r} \kappa(D'). \tag{2.12}
\]

Thus, \( \kappa(B') \) cannot be off by more than a factor of \( \sqrt{r} \) from the minimum condition number given on the right-hand side of (2.11).

Improving the conditioning of \( D' \) through the transformation (2.5) is desirable since a severely ill-conditioned matrix is sensitive to round-off errors in its entries (Maron 1982, p. 210). It follows that the eigenvalue analysis of Box et al. (1973) can be safeguarded from the effects of ill conditioning by adopting the scaling convention of (2.2). Of primary importance in that analysis is the magnitude of the smallest eigenvalue of the matrix \( DD' \), where \( D \) is given in (1.4). This can be quite different from the smallest eigenvalue of \( BB' \) as lemma 2.1 shows, especially when the \( w_i \)'s in (2.1) are markedly different.

**Lemma 2.1.** Let \( B, D, \) and \( w_i (i=1,2,...,r) \) be defined as in (2.4), (1.4), and (2.1), respectively. Let \( e_{\min}(\cdot) \) and \( e_{\max}(\cdot) \) denote the smallest and largest eigenvalues of a symmetric matrix. Then,

\[
\min_i (w_i^2) \leq e_{\min}(DD') \leq e_{\min}(BB') \leq \max_i (w_i^2). \tag{2.13}
\]

Proof. See Appendix A.

In (2.13), \( DD' \) and \( BB' \) are computed using rounded-off response values. Consequently, both
matrices are nonsingular even when the responses are linearly dependent. From (2.13) we note that 
\[ e_{\min}(DD') \] can be heavily affected by extreme values of \( w_i (i=1,2,\ldots,r) \). By contrast, \( e_{\min}(BB') \leq 1 \) 
since \( BB' \) is in correlation from (the determinant of this matrix, which is the product of its 
eigenvalues, is less than or equal to one by Theorem 8.7.6 in Graybill 1983).

3. THE EXPECTED VALUE AND VARIANCE OF A SMALL EIGENVALUE OF BB'

IN THE PRESENCE OF ROUND-OFF ERRORS

In this section, the expected value and variance of a small eigenvalue of \( BB' \) in (2.7) are 
derived when round-off error is the only error present. The derived values can be used in an eigenvalue 
analysis to determine whether a small eigenvalue of \( BB' \) is in fact a zero eigenvalue in the absence of 
round-off error. By (2.7), \( BB' \) and \( DD' \) are of the same rank. Hence, if \( BB' \) has a zero eigenvalue, 
then so does \( DD' \), which indicates the presence of a linear relationship among the responses.

Let \( y_{ui} \) be the exact \( i^{th} \) response value at the \( u^{th} \) experimental run, and let \( y_{ui}' \) be the value of 
\( y_{ui} \) rounded off to a certain number of decimal places (\( u=1,2,\ldots,n; i=1,2,\ldots,r \)). In the presence of 
round-off errors only we have the model

\[ Y' = Y + \Delta, \]  
(3.1)

where \( Y' \) and \( Y \) are the \( n \times r \) matrices, \( (y_{ui}') \) and \( (y_{ui}) \), respectively, and \( \Delta \) is the \( n \times r \) matrix \( (\Delta y_{ui}) \) 
of rounding errors in the response values. Let \( \Delta_i \) be the vector of round-off errors for the \( i^{th} \) response 
(\( i=1,2,\ldots,r \)), that is,

\[ \Delta_i = (\Delta y_{1i}', \Delta y_{2i}', \ldots, \Delta y_{ni}')', \quad i=1,2,\ldots,r. \]  
(3.2)
The elements of $\Delta_i$ are assumed to be independently distributed as uniform random variates. Thus,

$$
E(\Delta_i) = 0, \quad i = 1, 2, \ldots, r.
$$

$$
\text{Var}(\Delta_i) = \sigma_i^2 / n,
$$

where $\sigma_i^2 = \delta^2 / 3$ is the rounding error variance for the $i^{th}$ response. We also assume that the $\Delta_i$'s are statistically independent. If the values from the $i^{th}$ response are rounded off to $m_i$ decimal places, then \( \delta_i = 5 \left[ 10^{-(m_i+1)} \right], i = 1, 2, \ldots, r. \)

Consider the scaled multiresponse data matrix $S$ in (2.2). The change, $\Delta S$, in $S$ due to round-off errors is equal to

$$
\Delta S = \Delta^* (W^*)^{-1} -YW^{-1},
$$

where $W^*$ is an $r \times r$ diagonal matrix whose $i^{th}$ diagonal element, $w_i^*$, is the value of $w_i$ that results from using the rounded-off $i^{th}$ response data in (2.1), $i = 1, 2, \ldots, r$. Let $s_{ui}$ denote the $(u,i)^{th}$ element of $S$ ($u = 1, 2, \ldots, n; i = 1, 2, \ldots, r$). From (2.2) this element can be written as

$$
s_{ui} = y_{ui}/\left[ \sum_{v=1}^{n} (y_{vi} - \bar{y}_i)^2 \right]^{1/2}, \quad u = 1, 2, \ldots, n; i = 1, 2, \ldots, r.
$$

where $\bar{y}_i = \frac{1}{n} \sum_{v=1}^{n} y_{vi}$. The right-hand side of (3.5) is a function of $y_{1i}, y_{2i}, \ldots, y_{ni}$ denoted by $s_{ui} y_{1i}, y_{2i}, \ldots, y_{ni}$, $u = 1, 2, \ldots, n$. If $\Delta s_{ui}$ is the $(u,i)^{th}$ element of $\Delta S$, then a first-order approximation of $\Delta s_{ui}$ is given by

$$
\Delta s_{ui} \approx \sum_{v=1}^{n} \frac{\partial s_{ui}}{\partial y_{vi}} \Delta y_{vi}.
$$
where \( \frac{\partial L}{\partial \lambda_i} \) is the partial derivative of \( \lambda_i Y_1', Y_2', \ldots, Y_m' \) with respect to \( \lambda_i \) for \( i = 1, 2, \ldots, r \).

Let us suppose that in the absence of round-off error, the matrix \( BB' \) in (2.7) has a zero eigenvalue of multiplicity \( m \geq 1 \). This means that \( m \) linearly independent relationships must exist among the responses. Let \( \lambda_1, \lambda_2, \ldots, \lambda_m \) be an orthonormal set of eigenvectors of \( BB' \) for the eigenvalue zero. If the rounded-off multiresponse data matrix, \( Y' \), is used in the computation of \( B \) in (2.5), then the matrix \( BB' \) will have \( m \) small eigenvalues. Let \( \lambda_j^* \) (\( j = 1, 2, \ldots, m \)) denote the \( j \)th smallest of these eigenvalues. Then, for sufficiently small \( \delta_1, \delta_2, \ldots, \delta_r \) we have

\[
\lambda_j^* \sim \lambda_j^* (\Delta S)' (I_n - J_n/n)(\Delta S) \lambda_j^*, \quad j = 1, 2, \ldots, m.
\]

(3.7)

where \( \delta_1, \delta_2, \ldots, \delta_r \) are the parameters of the uniform distributions associated with the round-off errors from the \( r \) responses (see Wilkinson 1963, p. 138; Khuri and Colon 1981, p. 372). The symbol \( \sim \) in (3.7) means that the two sides are of the same order of magnitude. The eigenvectors, \( \tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_m \), that appear in (3.7) can be approximated with \( \tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_m \), an orthonormal set of eigenvectors of \( BB' \) which correspond to \( \lambda_1^*, \lambda_2^*, \ldots, \lambda_m^* \), respectively. We can then write

\[
\lambda_j^* \sim \tilde{\lambda}_j^*(\Delta S)' (I_n - J_n/n)(\Delta S) \tilde{\lambda}_j^*, \quad j = 1, 2, \ldots, m.
\]

(3.8)

Using the expression in (3.8), it can be shown (see Appendix B) that we approximately have

\[
E(\lambda_j^*) = (n-2) \sum_{i=1}^{r} \sigma_i^2 \tilde{\lambda}_i^2 / \tilde{w}_i^2, \quad j = 1, 2, \ldots, m.
\]

(3.9)

where \( \sigma_i^2 = \sigma_i^2 \) is the rounding error variance for response \( i (i = 1, 2, \ldots, r) \).
\[
\text{Var} (\lambda_j) \leq \left[ \frac{3n}{5} + n(n-1) \right] \sum_{i=1}^{r} \tau_i^2 \sigma_i^4 + 2n^2 \sum_{i < \ell} \tau_i \tau_\ell \sigma_i^2 \sigma_\ell^2 - (n-2)^2 \left[ \sum_{i=1}^{r} \sigma_i^2 \bar{a}_{ij}^2 / w_i^2 \right]^2, \quad j=1,2,\ldots,m. \tag{3.10}
\]

where \( \bar{a}_{ij} \) is the \( i \)th element of \( \bar{a}_j \) (\( i=1,2,\ldots,r; j=1,2,\ldots,m \)) and \( \tau_i \) is given by

\[
\tau_i = \left[ (n-1) w_i^2 + n\bar{y}_i^2 \right] / w_i^4, \quad i=1,2,\ldots,r. \tag{3.11}
\]

If \( \lambda_j \) is of the same order of magnitude as \( E(\lambda_j) \), then we may treat \( \lambda_j \) as a zero eigenvalue and conclude that a linear dependency exists among the responses. The elements of the corresponding eigenvector, \( \bar{a}_j \), may be used to identify the linear dependency. The upper bound in (3.10), denoted by \( u_j \), can be used to obtain the standardized values

\[
\eta_j = \left[ \lambda_j - E(\lambda_j) \right] / \sqrt{u_j}, \quad j=1,2,\ldots,m. \tag{3.12}
\]

Large values of \( \eta_j \) clearly indicate that \( \lambda_j \) does not correspond to a zero eigenvalue of \( \mathbf{B} \mathbf{B}' \). Small values of \( \eta_j \), however, do not necessarily imply a zero eigenvalue since \( u_j \) is just an upper bound to the true variance of \( \lambda_j \).

4. NUMERICAL EXAMPLES

Two examples are presented in this section to illustrate the implementation of the eigenvalue analysis described in Section 3.

4.1 Example 1

Let us consider the data on the thermal isomerization of \( \alpha \)-pinene at 189.5°, which was reported in Fuguitt and Hawkins (1947) and analyzed by Box et al. (1973). The data are reproduced
in Table 1. In this example, values from the five responses have been recorded to the nearest .10, hence \( \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = .05 \), where the \( \delta_i \)'s are the parameters associated with the uniform distributions of round-off errors. The matrix \( BB' \) is given in Table 2 along with its corresponding variance inflation factors and the condition number, \( \kappa(B') \). The eigenvalues of \( BB' \), their corresponding expected values, values of \( \eta_j \) (see 3.12), and the matrix of orthonormal eigenvectors of \( BB' \) are given in Table 3.

The first and second smallest eigenvalues of \( BB' \) in Table 3 are of the same order of magnitude as their expected values. The associated values of \( \eta_j \) (j=4,5) are very close to zero. These two eigenvalues can, therefore, be regarded as equal to zero. The elements of the corresponding eigenvectors (the last two columns of the matrix in Table 3) can be used to define two linearly independent relationships among the five responses. The third smallest eigenvalue of \( BB' \), namely .0027773, is not of the same order of magnitude as \( E(\lambda_3) \). The associated \( \eta_3 \) value, however, is relatively small. This eigenvalue corresponds to a linear relationship among the expected values of the responses (see Box et al. 1973, Section 6). The remaining two eigenvalues, namely .881599 and 4.11521, are much larger and have large \( \eta_j \) values (j=1,2). Hence, they are not associated with any linear relationships among the responses or among their expected values.

The large variance inflation factors and condition number in Table 2 indicate that the matrix \( B' \) is severely ill conditioned, hence it is very sensitive to round-off errors. We note that the responses \( y_1, y_2, y_4 \) and \( y_5 \) are quite involved in the linear dependencies since their corresponding variance inflation factors are extremely large. This is also supported by an examination of the elements of the last two columns of the matrix of eigenvectors in Table 3. In both columns, the third element, which corresponds to \( y_3 \), is the smallest in absolute value.

4.2 Example 2

Research was conducted by Ahmed et al. (1983) at the University of Florida in order to develop acceptable fish patties from sheepshead harvested along the Florida coast. Deboned and
washed sheepshead flesh was mixed with varying proportions of $x_1 = \text{sodium chloride}$, $x_2 = \text{sodium tripolyphosphate}$, and $x_3 = \text{sodium alginate}$. Of interest was the determination of the effects of these three input variables on the texture quality of cooked patties. This was measured by the values of four response variables, namely, $y_1 = \text{breaking force in grams}$, $y_2 = \text{texture firmness}$, $y_3 = \text{texture preference}$, and $y_4 = \text{flavor}$. The last three responses were measured using a nine-point rating scale with 1 being least desirable and 9 most desirable. The response values are given in Table 4.

The variance inflation factors in this example are 2.374, 28.653, 24.244, 1.768, and the condition number of the matrix $B'$ is $\kappa(B') = 12.552$, an indication of moderate ill conditioning. The values of $\eta_j$ in (3.12) are $\eta_1 = 1909.61$, $\eta_2 = 392.93$, $\eta_3 = 214.702$, and $\eta_4 = 12.12$. Hence, no eigenvalue of $BB'$ can be regarded as equal to zero.

It is interesting here to note that the eigenvalues of the matrix $BB'$ are .0192, .3397, .6202, and 3.0210. By contrast, the eigenvalues of the matrix $DD'$ for the unscaled responses are .2871, 6.8863, 15.8812, and 2,009, 864 with a condition number, $\kappa(D') = (2,009,864/.2871)^{1/2} = 2645.86$. The latter number greatly exceeds $\kappa(B')$. This shows that scaling of the responses can improve the conditioning of the matrix $BB'$, which, in turn, reduces its sensitivity to round-off errors.

5. CONCLUDING REMARKS

The scaling convention proposed in (2.2) is recommended to be used prior to the application of the eigenvalue analysis for detecting linear dependencies among the responses. Scaling removes inconsistencies in the units of measurement by putting all response variables on a common scale. An eigenvalue analysis applied directly to the original data, e.g., the data of Example 2 in Section 4.2, is not always meaningful because the responses may not have the same scale.

As was observed in Example 2, round-off error can be reduced by scaling. To see this in general, we note that if the $w_i$'s in (3.9) exceed unity and if the $\sigma_i^2$'s are equal to $\sigma_{Fe}^2$, then the
expected value of \( \lambda_j \) is less than \((n-1)\sigma^2 \). The latter quantity is given in Box et al. (1973) as the expected value of a small eigenvalue of \( DD' \).

Acknowledgements

This research was partially supported by Grant N00014-86-K-0059 from the office of Naval Research.
Appendix A: Proof of the Double Inequality (2.13)

From (2.5) we have that

$$DD' = WBB'W.$$  \hspace{1cm} (A.1)

where $W$ is the $r \times r$ diagonal matrix whose $i^{th}$ diagonal element, $w_i$, is given in (2.1). Then, from (A.1) and Lemma 2 in Khuri (1986, p. 355) we can write

$$e_{\min} (DD') \geq e_{\min} (BB') e_{\min} (W^2)$$
$$= e_{\min} (BB') \min_i (w_i^2).$$  \hspace{1cm} (A.2)

Let $e_{\min}^+ (\cdot)$ denote the smallest positive eigenvalue of a symmetric matrix. Since the nonzero eigenvalues of $WBB'W$ are equal to those of $B'BW^2B$, then

$$e_{\min} (WBB'W) = e_{\min}^+ (B'BW^2B)$$
$$\leq e_{\min}^+ (B'B) e_{\max} (W^2)$$  \hspace{1cm} (see Khuri 1986, p. 356) \hspace{1cm} (A.3)
$$= e_{\min} (BB') \max_i (w_i^2).$$

The double inequality in (2.13) now follows from (A.1), (A.2), and (A.3).
Appendix B: The Derivation of the Expected Value in (3.9) and the Upper Bound in (3.10)

i) The Expected Value in (3.9)

From (3.5) we have that

$$g_u(y_{1i}, y_{2i}, ..., y_{ni}) = y_{ui} / \left[ \sum_{v=1}^{n} (y_{vi} - \bar{y}_i)^2 \right]^{\frac{1}{2}}, \quad u=1,2,...,n; \; i=1,2,...,r. \quad (B.1)$$

By taking the partial derivatives of $g_u$ with respect to its arguments we get

$$\frac{\partial g_u}{\partial y_{vi}} = \begin{cases} \frac{1}{w_i} \cdot (y_{vi} - y_i) y_{ui} / w_i^3, & v = u \\ (y_{vi} - y_i) y_{ui} / w_i^3, & v \neq u \end{cases} \quad (B.2)$$

where $w_i$ is defined in (2.1). Let $\varphi_{ui}$ be the $n \times 1$ vector

$$\varphi_{ui} = \left( \frac{\partial g_u}{\partial y_{1i}}, \frac{\partial g_u}{\partial y_{2i}}, ..., \frac{\partial g_u}{\partial y_{ni}} \right)' \quad (B.3)$$

We can then write (3.6) in the form

$$\Delta s_{ui} \approx \varphi_{ui}' \Delta_i, \quad (B.4)$$

where $\Delta_i$ is defined in (3.2). A first-order approximation of $\Delta S = (\Delta s_{ui})$ in (3.4) can then be expressed as

$$\Delta S \approx \left[ \varphi_1 \Delta_1 : \varphi_2 \Delta_2 : \cdots : \varphi_r \Delta_r \right] \quad (B.5)$$
where

\[ \phi_i = [\phi_{i1}, \phi_{i2}, \ldots, \phi_{in}]'. \]  

From (3.3) and (B.6) we get

\[
\begin{align*}
E(\Delta_i) &= 0, \\
Var(\Delta_i) &= \phi_i' \sigma_i^2 \phi_i, \\
\end{align*}
\]

where \( \sigma_i^2 = \sigma^2_i / 3 \) is the round-off error variance for the \( i^{th} \) response (\( i = 1, 2, \ldots, r \)).

Let us now write the expression in (3.8) as

\[
\lambda_j^* \sim b_j^* (I_n - \frac{1}{n} I_n)b_j, \quad j = 1, 2, \ldots, m.
\]

where

\[
b_j = (\Delta S) \hat{\delta}_j \\
\approx \sum_{i=1}^{r} \hat{a}_{ij} \phi_i \Delta_i, \quad j = 1, 2, \ldots, m,
\]

where \( \hat{a}_{ij} \) is the \( i^{th} \) element of \( \hat{\delta}_j \). From (B.7), (B.9), and the fact that the \( \Delta_i \)'s are statistically independent we approximately have

\[
\begin{align*}
E(b_j) &= 0, \\
Var(b_j) &= \sum_{i=1}^{r} \hat{a}_{ij}^2 \sigma_i^2 \phi_i' \phi_i', \\
\end{align*}
\]

The expected value of \( \lambda_j^* \) in (B.8) can then be approximately written as

\[
E(\lambda_j^*) = \text{tr} \left[ (I_n - \frac{1}{n} I_n) \sum_{i=1}^{r} \sigma_i^2 \hat{a}_{ij}^2 \phi_i' \phi_i' \right]
\]

-17-
Since the \( u \)th row of \( \phi_i \) (\( u=1,2,\ldots,n \), \( i=1,2,\ldots,r \)) is of the form given by \((B.2)\) and \((B.3)\), then it is easy to show that

\[
\left( I_n - \frac{1}{n} \right) \phi_i = \frac{1}{w_i} \left( I_n - \frac{1}{n} \right) \phi_i = \left( I_n - \frac{1}{n} \right) H_i, \quad i=1,2,\ldots,r
\]  

\( (B.12) \)

where \( H_i \) is a symmetric \( n \times n \) matrix whose \((\mu,\nu)\)th element is

\[
h_{\mu\nu}^{(i)} = (y_{\mu i} - \bar{y}_i)(y_{\nu i} - \bar{y}_i), \quad \mu,\nu = 1,2,\ldots,n; \quad i=1,2,\ldots,r.
\]  

\( (B.13) \)

We note that \( J_n H_i = 0 \) for \( i=1,2,\ldots,r \). It follows that

\[
\text{tr} \left[ \phi_i' \left( I_n - \frac{1}{n} \right) \phi_i \right] = (n-2)/w_i^2, \quad i=1,2,\ldots,r.
\]  

\( (B.14) \)

Formula (3.9) follows from \((B.11)\) and \((B.14)\).

\( \textbf{i}) \) The Upper Bound in (3.10)

From (3.8) we approximately have

\[
\text{Var}(\lambda_j^*) = E \left[ \hat{\lambda}_j^*(\Delta \hat{S})' \left( I_n - \frac{1}{n} \right) (\Delta \hat{S}) \hat{\lambda}_j^* \right] - \left[ E(\hat{\lambda}_j^*) \right]^2, \quad j=1,2,\ldots,m.
\]  

\( (B.15) \)

We note that

\[
E \left[ \hat{\lambda}_j^*(\Delta \hat{S})' \left( I_n - \frac{1}{n} \right) (\Delta \hat{S}) \hat{\lambda}_j^* \right] \leq E \left[ \hat{\lambda}_j^*(\Delta \hat{S})' (\Delta \hat{S}) \hat{\lambda}_j^* \right]^2, \quad j=1,2,\ldots,m.
\]  

\( (B.16) \)
since \((\Delta S)'(\Delta S) \cdot (\Delta S)'(1_n - 1_n/n)(\Delta S)\) is a positive semidefinite matrix. We also note that

\[
E\left[ \hat{a}_j'(\Delta S)'(\Delta S)\hat{a}_j \right]^2 \leq E\left[ \text{tr}\left[ (\Delta S)'(\Delta S) \right] \right] \cdot j, 1, 2, \ldots, m. \tag{B.17}
\]

since \(\hat{a}_j'(\Delta S)'(\Delta S)\hat{a}_j\) is less than or equal to the largest eigenvalue of \((\Delta S)'(\Delta S)\). Lancaster (1969, p. 109).

From (B.4) and (B.17) we approximately have

\[
E\left[ \hat{a}_j'(\Delta S)'(\Delta S)\hat{a}_j \right]^2 \leq E\left[ \sum_{u=1}^{p} \sum_{i=1}^{r} (\phi_{ui}\Delta_i)^2 \right] \cdot j, 1, 2, \ldots, m. \tag{B.18}
\]

Note that

\[
E(\Delta_i')^2 = E\left[ \sum_{u=1}^{p} (\Delta y_{ui})^4 + 2 \sum_{u<v} (\Delta y_{ui})^2(\Delta y_{vi})^2 \right], \quad i, 1, 2, \ldots, r. \tag{B.19}
\]

Since \(\Delta y_{ui}\) has the uniform distribution \(U(-\delta_i, \delta_i)\), \(i, 1, 2, \ldots, r\), then for \(u=1, 2, \ldots, n\),

\[
E(\Delta y_{ui})^2 = \sigma_i^2 = \delta_i^2/3, \quad i, 1, 2, \ldots, r.
\]

\[
E(\Delta y_{ui})^4 = (9/5)\sigma_i^4, \quad i, 1, 2, \ldots, r.
\]

Furthermore, \(\Delta y_{ui}\) and \(\Delta y_{vi}\) are statistically independent for \(u \neq v\). Formula (B.19) can therefore be
\[ E(\Delta'_i \Delta'_i)^2 = \left[ \frac{9n}{5} + n(n-1) \right] \sigma^4_i, \quad i=1,2,...,r. \] (B.20)

From (B.2) and (B.3) we also have
\[ \varphi'_u \varphi'_u = \frac{1}{w_i^2} - \frac{2}{w_i^4} (\bar{y}_{ui} - \bar{y}) y_{ui} + \frac{y_{ui}^2}{w_i^4}, \quad u=1,2,...,n; i=1,2,...,r. \] (B.21)

Thus,
\[ \sum_{u=1}^{n} \varphi'_u \varphi'_u = \tau_i, \quad i=1,2,...,r. \] (B.22)

where
\[ \tau_i = \left[ (n-1)w_i^2 + n\bar{y}_i^2 \right]/w_i^4, \quad i=1,2,...,r. \] (B.23)

Using (B.20), (B.22), and (B.23), inequality (B.18) can be expressed as
\[
E \left[ \tilde{S}_j '(\Delta \tilde{S})'(\Delta \tilde{S}) \tilde{S}_j \right]^2 \leq E \left[ \sum_{i=1}^{r} \tau_i \Delta'_i \Delta'_i \right]^2 \\
= E \left[ \sum_{i=1}^{r} \tau_i^2 (\Delta'_i \Delta_i)^2 + 2 \sum_{i<\ell} \tau_i \tau_\ell (\Delta'_i \Delta_i)(\Delta'_\ell \Delta_\ell) \right] \\
= \left[ \frac{9n}{5} + n(n-1) \right] \sum_{i=1}^{r} \tau_i^2 \sigma^4_i + 2 \sum_{i<\ell} \tau_i \tau_\ell E(\Delta'_i \Delta_i)E(\Delta'_\ell \Delta_\ell) \\
= \left[ \frac{9n}{5} + n(n-1) \right] \sum_{i=1}^{r} \tau_i^2 \sigma^4_i + 2n^2 \sum_{i<\ell} \tau_i \tau_\ell \sigma^2_i \sigma^2_\ell. \] (B.24)
since by (3.3),

\[ E(\Delta'_i \Delta_i) = \text{tr}(\sigma_i^2 I_n) = n\sigma_i^2, \quad i=1,2,\ldots,r. \]

From (B.15), (B.16), and (B.24) we finally get

\[ \text{Var}(\lambda_j^*) \leq \left[ \frac{3n}{2} + n(n-1) \right] \sum_{i=1}^{r} \tau_i^2 \sigma_i^4 + 2n^2 \sum_{i<j} \tau_i \tau_j \sigma_i^2 \sigma_j^2 \cdot \left[ E(\lambda_j^*) \right]^2, \quad j=1,2,\ldots,m. \]  

(B.25)

By substituting the mean of \( \lambda_j^* \) from (3.9) in (B.25) we obtain the upper bound given in (3.10).

REFERENCES


Polynomial Regression Functions". Technometrics, 23, 363-375.


Table 1. Data* for the Isomerization of α-Pinene at 189.5° (Example 1)

<table>
<thead>
<tr>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
<th>$y_4$</th>
<th>$y_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-pinene</td>
<td>dipentene</td>
<td>ailo-oicinene</td>
<td>pyronene</td>
<td>dimer</td>
</tr>
<tr>
<td>88.35</td>
<td>7.3</td>
<td>2.3</td>
<td>.4</td>
<td>1.75</td>
</tr>
<tr>
<td>75.4</td>
<td>15.6</td>
<td>4.5</td>
<td>.7</td>
<td>2.8</td>
</tr>
<tr>
<td>65.1</td>
<td>23.1</td>
<td>5.3</td>
<td>1.1</td>
<td>5.8</td>
</tr>
<tr>
<td>50.4</td>
<td>32.9</td>
<td>6.0</td>
<td>1.5</td>
<td>9.3</td>
</tr>
<tr>
<td>37.5</td>
<td>42.7</td>
<td>6.0</td>
<td>1.9</td>
<td>12.0</td>
</tr>
<tr>
<td>25.9</td>
<td>49.1</td>
<td>5.9</td>
<td>2.2</td>
<td>17.0</td>
</tr>
<tr>
<td>14.0</td>
<td>57.4</td>
<td>5.1</td>
<td>2.6</td>
<td>21.0</td>
</tr>
<tr>
<td>4.5</td>
<td>53.1</td>
<td>3.8</td>
<td>2.9</td>
<td>25.7</td>
</tr>
</tbody>
</table>

*Source: Box et al. (1973).
Table 2. The Matrix $BB'$ and its Measures of Ill Conditioning (Example 1)

<table>
<thead>
<tr>
<th></th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
<th>$y_4$</th>
<th>$y_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_1$</td>
<td>1.0</td>
<td>-0.9997</td>
<td>-0.3679</td>
<td>-0.9996</td>
<td>-0.9852</td>
</tr>
<tr>
<td>$y_2$</td>
<td>-0.9997</td>
<td>1.0</td>
<td>0.3805</td>
<td>0.9993</td>
<td>0.9817</td>
</tr>
<tr>
<td>$y_3$</td>
<td>-0.3679</td>
<td>0.3805</td>
<td>1.0</td>
<td>0.3557</td>
<td>0.2105</td>
</tr>
<tr>
<td>$y_4$</td>
<td>-0.9996</td>
<td>0.9993</td>
<td>0.3557</td>
<td>1.0</td>
<td>0.9868</td>
</tr>
<tr>
<td>$y_5$</td>
<td>-0.9852</td>
<td>0.9817</td>
<td>0.2105</td>
<td>0.9868</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Variance inflation: 224.115 138.165 370.77 3781.55 20,624.1

Condition number: $\kappa(B') = 1258.08$
Table 3. Values of $\lambda_j$, $E(\lambda_j)$, $\eta_j$, and the Eigenvectors of $BB'$ (Example 1)

<table>
<thead>
<tr>
<th>$\lambda_j$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.11521</td>
<td>.881599</td>
<td>.0027773</td>
<td>.0004080</td>
<td>.0000026</td>
<td></td>
</tr>
<tr>
<td>0.0002361</td>
<td>0.0004034</td>
<td>0.0000732</td>
<td>0.0006083</td>
<td>0.0000068</td>
<td></td>
</tr>
<tr>
<td>$\eta_j$</td>
<td>153.018</td>
<td>32.7703</td>
<td>100.552</td>
<td>-.007451</td>
<td>-.000157</td>
</tr>
<tr>
<td>Orthnormal eigenvectors</td>
<td>.492148</td>
<td>-.059831</td>
<td>.109986</td>
<td>-.399853</td>
<td>-.763041</td>
</tr>
<tr>
<td>of $BB'$</td>
<td>-.492328</td>
<td>.045061</td>
<td>-.491572</td>
<td>.39444</td>
<td>-.598629</td>
</tr>
<tr>
<td></td>
<td>-.206855</td>
<td>-.966695</td>
<td>.14714</td>
<td>-.010939</td>
<td>-.030677</td>
</tr>
<tr>
<td></td>
<td>-.491659</td>
<td>.073606</td>
<td>-.253694</td>
<td>-.82648</td>
<td>.073646</td>
</tr>
<tr>
<td></td>
<td>-.480505</td>
<td>.233392</td>
<td>.812557</td>
<td>.036687</td>
<td>-.230319</td>
</tr>
</tbody>
</table>
Table 4. Values* of the Breaking Force and Sensory Response Variables (Example 2)

<table>
<thead>
<tr>
<th>$y_1$ (g)</th>
<th>$y_2^{**}$ Texture firmness</th>
<th>$y_3^{**}$ Texture preference</th>
<th>$y_4^{**}$ Flavor</th>
</tr>
</thead>
<tbody>
<tr>
<td>637.5</td>
<td>4.25</td>
<td>4.25</td>
<td>5.13</td>
</tr>
<tr>
<td>1020.8</td>
<td>4.75</td>
<td>4.88</td>
<td>5.38</td>
</tr>
<tr>
<td>1529.2</td>
<td>5.75</td>
<td>5.50</td>
<td>5.63</td>
</tr>
<tr>
<td>1445.8</td>
<td>6.63</td>
<td>6.25</td>
<td>6.50</td>
</tr>
<tr>
<td>345.0</td>
<td>2.75</td>
<td>3.38</td>
<td>3.88</td>
</tr>
<tr>
<td>441.7</td>
<td>3.38</td>
<td>3.50</td>
<td>5.00</td>
</tr>
<tr>
<td>576.7</td>
<td>4.88</td>
<td>5.13</td>
<td>5.00</td>
</tr>
<tr>
<td>531.7</td>
<td>5.38</td>
<td>5.15</td>
<td>5.75</td>
</tr>
<tr>
<td>380.0</td>
<td>4.63</td>
<td>4.63</td>
<td>3.25</td>
</tr>
<tr>
<td>575.0</td>
<td>5.38</td>
<td>5.50</td>
<td>5.63</td>
</tr>
<tr>
<td>676.7</td>
<td>5.25</td>
<td>5.38</td>
<td>4.88</td>
</tr>
<tr>
<td>845.0</td>
<td>5.75</td>
<td>5.38</td>
<td>6.13</td>
</tr>
<tr>
<td>1161.7</td>
<td>6.38</td>
<td>6.00</td>
<td>5.50</td>
</tr>
<tr>
<td>585.0</td>
<td>3.13</td>
<td>3.25</td>
<td>5.25</td>
</tr>
<tr>
<td>886.7</td>
<td>5.25</td>
<td>5.50</td>
<td>5.75</td>
</tr>
<tr>
<td>1115.0</td>
<td>5.50</td>
<td>5.13</td>
<td>5.75</td>
</tr>
<tr>
<td>825.0</td>
<td>5.13</td>
<td>5.25</td>
<td>3.88</td>
</tr>
</tbody>
</table>

*Source: Ahmed et al. (1983)

**The response values are based on a scale from 1 to 9.