The Annals of Statistics
1994, Vol. 22, No. 1, 94–105

AN EXACT CONFIDENCE REGION IN MULTIVARIATE CALIBRATION

BY THOMAS MATHEW[1] AND SUBRAMANYAM KASALA

University of Maryland, Baltimore County and University of North Carolina at Wilmington

In the multivariate calibration problem using a multivariate linear model, an exact confidence region is constructed. It is shown that the region is always nonempty and is invariant under nonsingular transformations.

1. Introduction. In this article, we consider the multivariate calibration problem using the multivariate linear model, and we construct an exact confidence region. Our setup is as follows. Let Y be a $p \times N$ random matrix whose columns are independent p-variate normally distributed vectors with

$$E(Y) = BX \quad \text{and} \quad \text{Cov}(\text{vec}(Y)) = I_N \otimes \Sigma, \tag{1.1}$$

where B is a $p \times m$ matrix of unknown parameters, X is an $m \times N$ known matrix of rank m and Σ is a $p \times p$ unknown positive definite matrix. In (1.1), vec (Y) denotes the $pN \times 1$ vector obtained by stacking the columns of Y one below the other. Let $\mathbf{y}$ be another $p \times 1$ normally distributed random vector, distributed independently of Y, satisfying

$$E(\mathbf{y}) = B\theta \quad \text{and} \quad \text{Cov}(\mathbf{y}) = \Sigma, \tag{1.2}$$

where θ is an $m \times 1$ unknown parameter vector. The problem we shall address is the construction of an exact region for θ. We shall assume that $p \geq m$. This condition is clearly necessary for the identifiability of θ [see the end of Section 1.4 in Brown (1982)]. In some applications, the model will contain an intercept. This model, however, can be reduced to one without an intercept; see Remark 2.2 in the next section.

In applications, the linear model (1.1) is known as the calibration curve. Examples where (1.1) and (1.2) are applicable are given in Brown (1982) and in the recent book by Martens and Naes (1989). In the set up (1.1) and (1.2), an exact confidence region for θ has been constructed by Brown (1982). Wood (1982, 1986) and Oman (1988) have pointed out some drawbacks for Brown's confidence region, one of them being that the region can be empty. Oman (1988) has constructed a confidence region that is always nonempty and is applicable to finite samples [for asymptotic results, we refer to Fujikoshi and Nishi (1984), Davis and Hayakawa (1987) and Brown and Sundberg (1987)].

Received February 1992; revised June 1993.

[1]Research supported by Grants AFOSR 89-0237 and AFOSR-F49620-93-1-0001.

AMS 1991 *subject classifications*. Primary 62F25; secondary 62H99.

Key words and phrases. Calibration, confidence region.

Furthermore, Oman's confidence region has the added advantage of being applicable to situations where θ depends nonlinearly on a smaller set of unknown parameters, say ξ, and the problem is to obtain a confidence region for ξ. An application of this type is given in Oman and Wax (1984). However, as Oman (1988) has pointed out, his confidence region is not invariant under the action of the group of $p \times p$ nonsingular matrices acting on Y and $\mathbf{y}$ as $Y \to AY$ and $\mathbf{y} \to A\mathbf{y}$, where A is a $p \times p$ nonsingular matrix. This invariance is clearly a natural requirement since inference problems concerning θ are invariant under the above group action.

In the next section, we construct a confidence region for θ that is exact and has the above invariance property. The problem is first reduced to a suitable canonical form similar to the canonical form for the MANOVA problem. We have also shown that our exact confidence region given in the next section is always nonempty. Unfortunately, our confidence region need not always be an ellipsoid, and it appears quite difficult to obtain a condition under which it will be an ellipsoid, sharing the same drawback with the region of Brown (1982). However, when $m = 1$, so that θ in (1.2) is a scalar, we have derived a sufficient condition under which our confidence region will be an interval. We have also computed the confidence region based on simulated data, for the case $m = 1$, and the region turned out to be an interval in many cases.

2. The confidence region. For convenience, we shall work with the following canonical form of the problem. Let Z be an $(N-m)\times N$ matrix satisfying $XZ' = 0$ and $ZZ' = I_{N-m}$. Then

$$P = \left(X'(XX')^{-1/2} \colon Z'\right) \tag{2.1}$$

is an $N \times N$ orthogonal matrix and

$$E(YP) = \left(B(XX')^{1/2} \colon 0\right) \quad \text{and} \quad \operatorname{Cov}(\operatorname{vec}(YP)) = I_N \otimes \Sigma. \tag{2.2}$$

Furthermore, defining

$$\begin{aligned} &Y_1 = YX'(XX')^{-1/2}, \qquad Y_2 = YZ', \qquad S = Y_2Y_2', \\ &B_1 = B(XX')^{1/2} \quad \text{and} \quad \theta_1 = (XX')^{-1/2}\theta, \end{aligned} \tag{2.3}$$

we have

$$\begin{aligned} &E(Y_1) = B_1, \qquad \operatorname{Cov}(\operatorname{vec}(Y_1)) = I_m \otimes \Sigma, \\ &E(\mathbf{y}) = B_1\theta_1, \quad \text{and} \quad S \sim W_p(\Sigma, N-m), \end{aligned} \tag{2.4}$$

where we assume $N - m \geq p$; (2.4) is the canonical form that we shall work with. Note that Y_1, S and $\mathbf{y}$ are independently distributed. Clearly, it is enough to construct a confidence region for θ_1 in (2.4). A confidence region for θ can then be obtained using the transformation in (2.3). In terms of the random variables in (2.4), Brown's (1982) confidence region is based on the statistic

$$\frac{(N-m-p+1)}{p}(1+\theta_1'\theta_1)^{-1}(\mathbf{y}-Y_1\theta_1)'S^{-1}(\mathbf{y}-Y_1\theta_1), \tag{2.5}$$

which follows a central F-distribution with $(p, N-m-p+1)$ degrees of freedom. A $100(1-\alpha)\%$ confidence region for θ_1 (and hence for θ) can be easily obtained using the statistic given in (2.5).

Brown's confidence region can be empty since the minimum value of the expression in (2.5), with respect to θ_1, depends on $\mathbf{y}$ and Y_1 and can be arbitrarily large. Our approach to obtain a nonempty confidence region is to modify (2.5) and come up with a statistic of the form

$$(2.6) \qquad (1+\theta_1'\theta_1)^{-1}(\mathbf{y}-Y_1\theta_1)'S^{-1}U(U'S^{-1}U)^{-1}U'S^{-1}(\mathbf{y}-Y_1\theta_1),$$

where U is a $p \times m$ random quantity (allowed to depend on θ_1 as well) distributed independently of $(\mathbf{y}-Y_1\theta_1)$, so that the minimum value of (2.6) with respect to θ_1 is zero. This latter property will guarantee that our region is always nonempty, as will be proved later in this section. Toward obtaining U in (2.6), we note that the random quantity $(1+\theta_1'\theta_1)^{-1/2}(\mathbf{y}-Y_1\theta_1)$, which appears in (2.5) and (2.6), can be written as $(Y_1\colon \mathbf{y})\mathbf{h}_{\theta_1}$, where $\mathbf{h}_{\theta_1} = (1+\theta_1'\theta_1)^{-1/2}(-\theta_1'\colon 1)'$ is an $(m+1)\times 1$ vector satisfying $\mathbf{h}_{\theta_1}'\mathbf{h}_{\theta_1} = 1$. We shall now construct an $(m+1)\times(m+1)$ orthogonal matrix Q_{θ_1}, depending on θ_1, with $\mathbf{h}_{\theta_1}$ as its last column, and then choose U as the matrix consisting of the first m columns of $(Y_1\colon \mathbf{y})Q_{\theta_1}$. Q_{θ_1} is given by

$$(2.7) \qquad Q_{\theta_1} = \begin{pmatrix} I & -\theta_1 \\ \theta_1' & 1 \end{pmatrix} \begin{pmatrix} (I+\theta_1\theta_1')^{-1/2} & 0 \\ 0 & (1+\theta_1'\theta_1)^{-1/2} \end{pmatrix}$$

and

$$(2.8) \qquad (Y_1\colon \mathbf{y})Q_{\theta_1} = \left((Y_1+\mathbf{y}\theta_1')(I+\theta_1\theta_1')^{-1/2}\colon (\mathbf{y}-Y_1\theta_1)(1+\theta_1'\theta_1)^{-1/2}\right).$$

Let

$$(2.9) \qquad \begin{aligned} Y_1^* &= (Y_1+\mathbf{y}\theta_1')(I+\theta_1\theta_1')^{-1/2}, \\ \mathbf{y}^* &= (\mathbf{y}-Y_1\theta_1)(1+\theta_1'\theta_1)^{-1/2}. \end{aligned}$$

Then

$$(2.10) \qquad E(Y_1^*) = B_1(I+\theta_1\theta_1')^{1/2} \quad \text{and} \quad E(\mathbf{y}^*) = 0,$$

and since Q_{θ_1} is orthogonal, Y_1^* and $\mathbf{y}^*$ are independently distributed (for any given θ_1). Our choice of U in (2.6) is $U = Y_1^*$. We shall now construct an F-ratio whose numerator is (2.6) with $U = Y_1^*$. For this, consider the $p \times p$ random orthogonal matrix Y_0 given by

$$(2.11) \qquad Y_0 = \left(Y_1^*(Y_1^{*\prime}Y_1^*)^{-1/2}\colon Y_2^*\right),$$

where Y_2^* is a $p \times (p-m)$ matrix of rank $(p-m)$ satisfying $Y_2^{*\prime}Y_1^* = 0$ and $Y_2^{*\prime}Y_2^* = I_{p-m}$ (recall our assumption $p \geq m$). Since Y_0 is a function of only

Y_1^*, Y_0 is distributed independently of S and $\mathbf{y}^*$. Consequently, conditionally given Y_0, we have

$$
\begin{aligned}
&\mathbf{u} = Y_0'\mathbf{y}^* \sim N(0, Y_0'\Sigma Y_0),\\
(2.12)\qquad & \\
&W = Y_0'SY_0 \sim W_p(Y_0'\Sigma Y_0, N-m),
\end{aligned}
$$

and furthermore, $\mathbf{u}$ and W are independently distributed (conditionally given Y_0). To derive a confidence region for θ_1, we shall first derive the distribution of T^* given by

$$
(2.13)\qquad T^* = \frac{(N-m-p+1)}{m}\frac{\mathbf{y}^{*\prime}S^{-1}Y_1^*(Y_1^{*\prime}S^{-1}Y_1^*)^{-1}Y_1^{*\prime}S^{-1}\mathbf{y}^*}{1+\mathbf{y}^{*\prime}Y_2^*(Y_2^{*\prime}SY_2^*)^{-1}Y_2^{*\prime}\mathbf{y}^*}.
$$

We note that $\mathbf{y}^{*\prime}S^{-1}Y_1^*(Y_1^{*\prime}S^{-1}Y_1^*)^{-1}Y_1^{*\prime}S^{-1}\mathbf{y}^*$, which appears in the numerator of (2.13), is the expression in (2.6) with $U = Y_1^*$. We shall show that conditionally given Y_1^* [i.e., conditionally given Y_0 in (2.11)], T^* has a central F-distribution with $(m, N-m-p+1)$ degrees of freedom, which is also its unconditional distribution. Toward this, using the orthogonality of Y_0, we note that $Y_1^{*\prime}Y_0 = ((Y_1^{*\prime}Y_1^*)^{1/2}\colon 0)$. Hence,

$$
\begin{aligned}
&\mathbf{y}^{*\prime}S^{-1}Y_1^*(Y_1^{*\prime}S^{-1}Y_1^*)^{-1}Y_1^{*\prime}S^{-1}\mathbf{y}^*\\
&\quad = \mathbf{y}^{*\prime}Y_0Y_0'S^{-1}Y_0Y_0'Y_1^*(Y_1^{*\prime}Y_1^*)^{-1/2}\\
(2.14)\qquad &\qquad \times [(Y_1^{*\prime}Y_1^*)^{-1/2}Y_1^{*\prime}Y_0Y_0'S^{-1}Y_0Y_0'Y_1^*(Y_1^{*\prime}Y_1^*)^{-1/2}]^{-1}\\
&\qquad \times (Y_1^{*\prime}Y_1^*)^{-1/2}Y_1^{*\prime}Y_0Y_0'S^{-1}Y_0Y_0'\mathbf{y}^*\\
&\quad = \mathbf{u}'W^{-1}\begin{pmatrix} I \\ 0 \end{pmatrix}\left[(I\colon 0)W^{-1}\begin{pmatrix} I \\ 0 \end{pmatrix}\right]^{-1}(I\colon 0)W^{-1}\mathbf{u},
\end{aligned}
$$

where $\mathbf{u}$ and W are given by (2.12). In order to simplify (2.14) further, let us partition $\mathbf{u}$ and W as

$$
(2.15)\qquad \mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix} \quad \text{and} \quad W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix},
$$

where $\mathbf{u}_1$ and $\mathbf{u}_2$ are, respectively, $m \times 1$ and $(p-m) \times 1$ vectors and W_{11}, W_{12} and W_{22} are, respectively, $m \times m, m \times (p-m)$ and $(p-m) \times (p-m)$ matrices. After expressing W^{-1} using the formula for the inverse of a partitioned matrix [see Rao (1974), page 33] and doing straightforward algebraic computations, (2.14) simplifies to

$$
\begin{aligned}
(2.16)\qquad &\mathbf{y}^{*\prime}S^{-1}Y_1^*(Y_1^{*\prime}S^{-1}Y_1^*)^{-1}Y_1^{*\prime}S^{-1}\mathbf{y}^*\\
&\quad = (\mathbf{u}_1 - W_{12}W_{22}^{-1}\mathbf{u}_2)'W_{11.2}^{-1}(\mathbf{u}_1 - W_{12}W_{22}^{-1}\mathbf{u}_2),
\end{aligned}
$$

where $W_{11.2} = W_{11} - W_{12}W_{22}^{-1}W_{21}$. From (2.15), (2.12) and (2.11), we have $\mathbf{u}_2 = Y_2^{*\prime}\mathbf{y}^*$ and $W_{22} = Y_2^{*\prime}SY_2^*$. Hence, using (2.16), T^* given in (2.13) can be expressed as

$$(2.17)\quad T^* = \frac{(N-m-p+1)}{m}\,\frac{(\mathbf{u}_1 - W_{12}W_{22}^{-1}\mathbf{u}_2)'W_{11.2}^{-1}(\mathbf{u}_1 - W_{12}W_{22}^{-1}\mathbf{u}_2)}{1+\mathbf{u}_2'W_{22}^{-1}\mathbf{u}_2}.$$

A statistic similar to (2.17) appears in Section 4.4.1 in Srivastava and Khatri [(1979), page 117] and Section 6.8 in Giri [(1977), page 131], where it is shown that such a statistic has a central F-distribution with $(m,\ N-m-p+1)$ degrees of freedom. Thus, from (2.12) and (2.17), we see that conditionally given Y_0, T^* follows the central F-distribution with $(m, N-m-p+1)$ degrees of freedom, which is also its unconditional distribution. We now prove the following theorem.

THEOREM. *Consider the models* (1.1) *and* (1.2) *and assume that $p \geq m$ and $N-m \geq p$. Define the following*:

$$\mathbf{y}_\theta = [1+\theta'(XX')^{-1}\theta]^{-1/2}[\mathbf{y} - YX'(XX')^{-1}\theta];$$

$$Y_{1\theta} = [YX' + \mathbf{y}\theta'];$$

$$S = Y[I - X'(XX')^{-1}X]Y';$$

$$T(\theta) = \frac{\mathbf{y}_\theta'S^{-1}Y_{1\theta}(Y_{1\theta}'S^{-1}Y_{1\theta})^{-1}Y_{1\theta}'S^{-1}\mathbf{y}_\theta}{1+\mathbf{y}_\theta'S^{-1}\mathbf{y}_\theta}.$$

Then the region

$$\left\{\theta \colon T(\theta) \leq \left[1 + \frac{N-m-p+1}{m}\,\frac{1}{F_\alpha(m, N-m-p+1)}\right]^{-1}\right\}$$

is a $100(1-\alpha)\%$ *confidence region for* θ, *where* $F_\alpha(m, N-m-p+1)$ *is the* $100(1-\alpha)$*th percentile of the central F distribution with* $(m, N-m-p+1)$ *degrees of freedom. Furthermore, the region is nonempty and is invariant under the transformation* $Y \to AY$ *and* $\mathbf{y} \to A\mathbf{y}$, *where A is any* $p \times p$ *nonsingular matrix.*

PROOF. We note that Y_1^* and $\mathbf{y}^*$ given in (2.9) satisfy

$$\mathbf{y}^* = \mathbf{y}_\theta \quad \text{and} \quad Y_1^* = Y_{1\theta}(XX')^{-1/2}[I + (XX')^{-1/2}\theta\theta'(XX')^{-1/2}]^{-1/2},$$

where $Y_{1\theta}$ and $\mathbf{y}_\theta$ are given in the statement of the theorem. Furthermore, $T(\theta)$ remains unchanged if we use Y_1^* instead of $Y_{1\theta}$. Since T^* given in (2.13) follows the central F-distribution with $(m, N-m-p+1)$ degrees of freedom, the expression for the confidence region follows if we can show that

$$\frac{1}{T(\theta)} = 1 + \frac{(N-m-p+1)}{m}\,\frac{1}{T^*},$$

or, equivalently,

$$\frac{(N-m-p+1)}{m}\frac{T(\theta)}{1-T(\theta)} = T^*.$$

This follows by using the expression for T^* in (2.13) and noting that

$$S^{-1} - S^{-1}Y_1^*(Y_1^{*\prime}S^{-1}Y_1^*)^{-1}Y_1^{*\prime}S^{-1} = Y_2^*(Y_2^{*\prime}SY_2^*)^{-1}Y_2^{*\prime}$$

[see Rao (1974), page 77]. Since $T(\theta)$ is invariant under the transformation mentioned in the theorem, so is the confidence region. Thus, it only remains to show that the region is nonempty.

For this, it is enough to show that there exists θ for which $T(\theta) = 0$, or, equivalently, there exists θ_1 for which $T^* = 0$ [θ_1 is defined in (2.3)]. From the expression for T^* in (2.13), it is clear that $T^* = 0$ if and only if

$$Y_1^{*\prime}S^{-1}\mathbf{y}^* = 0,$$

where Y_1^* and $\mathbf{y}^*$ are given in (2.9). We thus have to show that there exists θ_1 satisfying

$$(Y_1 + \mathbf{y}\theta_1')'S^{-1}(\mathbf{y} - Y_1\theta_1) = 0. \tag{2.18}$$

It is enough to prove this with $S = I$ (since we can consider $S^{-1/2}\mathbf{y}$ and $S^{-1/2}Y_1$ in the place of $\mathbf{y}$ and Y_1). After simplifying the left-hand side of (2.18) with $S = I$, we thus have to establish the existence of θ_1 satisfying

$$Y_1'\mathbf{y} - \{Y_1'Y_1 - (\mathbf{y}'\mathbf{y})I\}\theta_1 + (\mathbf{y}'Y_1\theta_1)\theta_1 = 0. \tag{2.19}$$

Consider the singular value decomposition

$$Y_1 = E\begin{pmatrix}\Delta\\ 0\end{pmatrix}F', \tag{2.20}$$

where E and F are, respectively, $p \times p$ and $m \times m$ orthogonal matrices, Δ is an $m \times m$ diagonal matrix with positive diagonal elements and 0 represents a $(p-m) \times m$ matrix of zeros. Write

$$\begin{aligned}\Delta &= \operatorname{diag}(\delta_1, \dots, \delta_m),\\ F'\theta_1 &= \eta = (\eta_1, \dots, \eta_m)',\\ E'\mathbf{y} &= \mathbf{g} = (\mathbf{g}_1', \mathbf{g}_2')' = (g_{11}, \dots, g_{1m}, g_{21}, \dots, g_{2(p-m)})'.\end{aligned} \tag{2.21}$$

Since $\mathbf{y}'\mathbf{y} = \mathbf{g}'\mathbf{g}$, (2.19) is to

$$\Delta\mathbf{g}_1 - \{\Delta^2 - (\mathbf{g}'\mathbf{g})I\}\eta - (\mathbf{g}_1'\Delta\eta)\eta = 0. \tag{2.22}$$

Thus, in order to show the existence of θ_1 satisfying (2.19), we have to show the existence of η satisfying (2.22). Let

$$h(\eta) = \mathbf{g}_1'\Delta\eta. \tag{2.23}$$

Then (2.22) is equivalent to

$$\delta_i g_{1i} - (\delta_i^2 - \mathbf{g}'\mathbf{g})\eta_i - h(\eta)\eta_i = 0, \qquad i = 1, 2, \ldots, m, \tag{2.24}$$

or, equivalently,

$$\eta_i = \frac{\delta_i g_{1i}}{\delta_i^2 - \mathbf{g}'\mathbf{g} + h(\eta)}, \qquad i = 1, 2, \ldots, m. \tag{2.25}$$

Using (2.25), $h(\eta)$ in (2.23) can be written as

$$h(\eta) = \mathbf{g}_1'\Delta\eta = \sum_{i=1}^{m} \delta_i g_{1i}\eta_i = \sum_{i=1}^{m} \frac{\delta_i^2 g_{1i}^2}{\delta_i^2 - \mathbf{g}'\mathbf{g} + h(\eta)}.$$

Thus, if we can show the existence of $h(\eta)$ satisfying

$$h(\eta) = \sum_{i=1}^{m} \frac{\delta_i^2 g_{1i}^2}{\delta_i^2 - \mathbf{g}'\mathbf{g} + h(\eta)}, \tag{2.26}$$

then, for such an $h(\eta)$, η_i given by (2.25) satisfies (2.24) and (2.26), and this will prove the existence of η satisfying (2.22). To show the existence of $h(\eta)$ satisfying (2.26), we have to show that the equation

$$x = \sum_{i=1}^{m} \frac{\delta_i^2 g_{1i}^2}{\delta_i^2 - \mathbf{g}'\mathbf{g} + x} \tag{2.27}$$

has a solution x. Let

$$f(x) = \sum_{i=1}^{m} \frac{\delta_i^2 g_{1i}^2}{\delta_i^2 - \mathbf{g}'\mathbf{g} + x} - x. \tag{2.28}$$

To establish the existence of x satisfying $f(x) = 0$, we consider three different cases: (i) If $f(0) = 0$, the required value of x is $x = 0$. (ii) If $f(0) > 0$, choose x_1 positive and large enough so that $f(x_1) < 0$ [the existence of such an x_1 should be clear from expression (2.28)]. Hence $f(x)$ must vanish inside $(0, x_1)$. (iii) If $f(0) < 0$, choose x_2 negative and large enough so that $f(x_2) > 0$. Hence $f(x)$ must vanish inside $(x_2, 0)$. This completes the proof of the theorem. □

REMARK 2.1. In terms of the canonical representation (2.3), the group action mentioned in the theorem is $Y_1 \to AY_1$, $\mathbf{y} \to A\mathbf{y}$ and $S \to ASA'$. Consequently, the action on Y_1^* and $\mathbf{y}^*$ in (2.9) is $Y_1^* \to AY_1^*$ and $\mathbf{y}^* \to A\mathbf{y}^*$. It should be noted that the matrix Y_0 and, in particular, the matrix Y_2^* in (2.11) are not invariant under the above group action. However, T^* in (2.13) is invariant since the quantity $T(\theta)$ in the theorem is invariant, and

$$\frac{(N - m - p + 1)}{m} \frac{T(\theta)}{1 - T(\theta)} = T^*,$$

as noted in the proof of the theorem.

Unfortunately, the shape of the confidence region given in the theorem is far from clear, since $T(\theta)$ is a rather complicated function of θ. However, in the special case when $p = m, Y_{1\theta}$ is a $p \times p$ matrix and $T(\theta)$ simplifies to $T(\theta) = \mathbf{y}'_\theta S^{-1}\mathbf{y}_\theta/(1 + \mathbf{y}'_\theta S^{-1}\mathbf{y}_\theta)$. In this case, the confidence region is based on $\mathbf{y}'_\theta S^{-1}\mathbf{y}_\theta$ [which has a central F-distribution with $(p, N - 2p + 1)$ degrees of freedom] and is the confidence region given in Brown (1982). When $p = m$, Brown (1982) has also given a condition under which his confidence region will be an ellipsoid. This condition depends on the data and need not always be satisfied. Consequently, the confidence region that we have constructed need not be an ellipsoid, in general. Furthermore, it appears quite difficult to obtain a condition under which it is an ellipsoid. However, in the univariate case, that is, when $m = 1$, a sufficient condition can be obtained so that the confidence region for θ (a scalar when $m = 1$) is an interval [see Brown and Sundberg (1987) and Lieftinck-Koeijers (1988) for applications where $m = 1$ and $p \geq 1$]. We shall now derive this condition. Let

$$c_\alpha(m) = \left[1 + \frac{(N - m - p + 1)}{m}\frac{1}{F_\alpha(m, N - m - p + 1)}\right]^{-1}. \tag{2.29}$$

With $m = 1$, the confidence region for θ given in the theorem is $\{\theta : T(\theta) \leq c_\alpha(1)\}$. When $m = 1$, note that X' is a column vector, say $X' = \mathbf{x}$, and $Y'_{1\theta}S^{-1}Y_{1\theta}$ is a scalar. $T(\theta)$ can then be simplified by direct algebraic computations and it can be verified that $T(\theta) \leq c_\alpha(1)$ is equivalent to

$$f(\theta) \leq 0, \tag{2.30}$$

where

$$f(\theta) = a_4\theta^4 + 2a_3\theta^3 + a_2\theta^2 + 2a_1\theta + a_0, \tag{2.31}$$

with

$$\begin{aligned} a_4 &= (\mathbf{x}'\mathbf{x})^{-2}(\mathbf{x}'Y'S^{-1}\mathbf{y})^2 \\ &\quad - c_\alpha(1)(\mathbf{x}'\mathbf{x})^{-1}\mathbf{y}'S^{-1}\mathbf{y}[1 + (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'Y'S^{-1}Y\mathbf{x}], \\ a_3 &= (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'Y'S^{-1}\mathbf{y}[\{(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'Y'S^{-1}Y\mathbf{x} - \mathbf{y}'S^{-1}\mathbf{y}\} \\ &\qquad \times \{1 - c_\alpha(1)\} - c_\alpha(1)], \\ a_2 &= [\{(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'Y'S^{-1}Y\mathbf{x} - \mathbf{y}'S^{-1}\mathbf{y}\}^2 - 2(\mathbf{x}'\mathbf{x})^{-1}(\mathbf{x}'Y'S^{-1}\mathbf{y})^2] \\ &\quad - c_\alpha(1)[\mathbf{y}'S^{-1}\mathbf{y}(1 + \mathbf{y}'S^{-1}\mathbf{y}) \\ &\qquad + (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'Y'S^{-1}Y\mathbf{x}\{1 + (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'Y'S^{-1}Y\mathbf{x}\} \\ &\qquad - 4(\mathbf{x}'\mathbf{x})^{-1}(\mathbf{x}'Y'S^{-1}\mathbf{y})^2], \\ a_1 &= \mathbf{x}'Y'S^{-1}\mathbf{y}[\{\mathbf{y}'S^{-1}\mathbf{y} - (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'Y'S^{-1}Y\mathbf{x}\}\{1 - c_\alpha(1)\} - c_\alpha(1)] \\ &= -(\mathbf{x}'\mathbf{x})a_3, \\ a_0 &= (\mathbf{x}'Y'S^{-1}\mathbf{y})^2 - c_\alpha(1)\mathbf{x}'Y'S^{-1}Y\mathbf{x}(1 + \mathbf{y}'S^{-1}\mathbf{y}). \end{aligned} \tag{2.32}$$

Note that

$$\{\theta: T(\theta) \leq c_\alpha(1)\} = \{\theta: f(\theta) \leq 0\}, \tag{2.33}$$

which is an interval if $f(\theta)$ is a convex function of θ, or, equivalently, $d^2f(\theta)/d\theta^2 \geq 0$, for all θ. From (2.31), it is clear that $d^2f(\theta)/d\theta^2$ is quadratic in θ and is nonnegative for all θ if and only if

$$a_4 \geq 0, \quad a_2 \geq 0 \quad \text{and} \quad 2a_2a_4 - 3a_3^2 \geq 0. \tag{2.34}$$

Once the data is available to us, a_4, a_3 and a_2 can be easily computed and the conditions in (2.34) can be easily verified.

We shall now report some confidence regions based on simulated data in the univariate case, that is, $m = 1$. The purpose of the simulation is to show that the conditions in (2.34) can hold and, hence, the confidence region obtained using the theorem can be an interval. In the simulation, the quantities p, N and α were chosen to have values $p = 2, N = 12$ and $\alpha = 0.05$. Then, $F_\alpha(m, N - m - p + 1) = F_{0.05}(1, 10) = 4.96$ and $c_\alpha(m) = c_{0.05}(1) = 0.33$ [from (2.29)]. Furthermore, in the notation used in (2.32), we chose a model with $\mathbf{x}'\mathbf{x} = 1$. (Note that since $m = 1$ and $N = 12, \mathbf{x}$ is a 12×1 vector.) Hence, for the simulation, one can consider the independent bivariate normal random vectors $\mathbf{y}_1 = Y\mathbf{x}$ and $\mathbf{y}$, having the distributions $N(\mathbf{b}, \Sigma)$ and $N(\mathbf{b}\theta, \Sigma)$, respectively, and the Wishart matrix $S \sim W_2(\Sigma, 11)$, where $\mathbf{b}$ is a 2×1 unknown parameter vector and Σ is an unknown 2×2 positive definite matrix. Since $\mathbf{x}'\mathbf{x} = 1$, the components of $\mathbf{x}$ lie in the interval $[-1, 1]$ and we shall consider only values of θ in this interval for the simulation. $\Sigma = I$ was used in the simulation along with the following arbitrarily chosen values of $\mathbf{b}$ and θ: $\mathbf{b} = (10, 1)'$ and $(5, 1)'$ and $\theta = -0.5, 0.1, 0.2, 0.4$ and 0.8. A simulated value of S was obtained from the preceding Wishart distribution with $\Sigma = I$ and simulated values of $\mathbf{y}_1$ and $\mathbf{y}$ were obtained from the appropriate normal distributions for each value of $\mathbf{b}$ and θ with $\Sigma = I$. In Table 1, we give the confidence region for θ, along with the value of the coefficients $(a_4, a_3, a_2, a_1, a_0)$ in (2.32), computed using the simulated data. The values of these coefficients are reported so that one can verify if the conditions in (2.34) are satisfied by these coefficients. The confidence regions we have computed turned out to be intervals in all the cases except for the parameter values $\mathbf{b} = (10, 1)'$ and $\theta = -0.5$ and 0.8. One can also verify that except for the cases corresponding to these parameter values, (2.34) holds in all the other cases. In the two cases where the confidence regions were not intervals, the regions turned out to be unions of disjoint line segments, with one of the line segments containing the true value of the parameter.

REMARK 2.2. In some applications, one has to consider models where an intercept is present. Brown (1982) and Brown and Sundberg (1987) consider models having an intercept. In this case, instead of (1.1) and (1.2), the model to be used is

$$E(Y) = \alpha\mathbf{1}_N' + BX, \qquad \mathrm{Cov}(\mathrm{vec}(Y)) = I_N \otimes \Sigma \tag{2.35}$$

$$E(\mathbf{y}) = \alpha + B\theta, \qquad \mathrm{Cov}(\mathbf{y}) = \Sigma, \tag{2.36}$$

TABLE 1

95% confidence regions for θ in the model $\mathbf{y}_1 \sim N(\mathbf{b}, \Sigma), \mathbf{y} \sim N(\mathbf{b}\theta, \Sigma)$ and $S \sim W_2(\Sigma, 11)$, with $m = 1, p = 2, N = 12$ and $\Sigma = I$

$\mathbf{b}$	θ	$(\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_2, \mathbf{a}_1, \mathbf{a}_0)$	**Confidence region**
$(10, 1)'$	−0.5	(−1.6985, −34.3095, 8.6649, 34.3095, 17.7511)	$(-\infty, -40.5016) \cup (-0.7573, -0.2952) \cup (1.1541, \infty)$
	0.1	(0.5943, 10.5277, 280.9991, −10.5277, −6.4714)	(−0.1193, 0.1921)
	0.2	(0.3188, 1.9607, 18.8100, −1.9607, −1.6542)	(−0.2132, 0.3965)
	0.4	(0.3641, 1.6397, 11.3348, −1.6397, −1.2592)	(−0.2230, 0.4636)
	0.8	(0.5764, 8.2280, -12.2437, −8.2280, 7.0015)	$(-29.2414, -0.9062) \cup (0.3744, 1.2244)$
$(5, 1)'$	−0.5	(0.0544, −0.1875, 1.1826, 0.1875, −0.5391)	(−0.7404, 0.5735)
	0.1	(0.0375, 0.6671, 19.3229, −0.6671, −1.8349)	(−0.2779, 0.3402)
	0.2	(0.1454, 1.2592, 17.3799, −1.2592, −1.6988)	(−0.2520, 0.3804)
	0.4	(0.1575, 1.1280, 12.3982, −1.1280, −1.4876)	(−0.2721, 0.4280)
	0.8	(0.0853, 0.1341, 0.3266, −0.1341, −0.3544)	(−0.8736, 0.9608)

where α is the $p \times 1$ intercept vector, $\mathbf{1}_N$ is an $N \times 1$ vector of 1's, and the other quantities in (2.35) and (2.36) are as in (1.1) and (1.2). Multivariate normality is once again assumed. We shall now show that the above models can be reduced to models the type (1.1) and (1.2), that is, models without an intercept. Toward this, let Z be an $N \times (N-1)$ matrix such that $R = ((1/\sqrt{N})\mathbf{1}_N, Z)$ is an $N \times N$ orthogonal matrix. Since R is orthogonal, the columns of YR are independent, having the common covariance matrix Σ, where Y is as given in (2.35). Considering the model for YR and using the definition of R, it follows from (2.35) that

$$E\left(\frac{1}{\sqrt{N}}Y\mathbf{1}_N\right) = \sqrt{N}\alpha + \frac{1}{\sqrt{N}}BX\mathbf{1}_N \quad \text{and} \quad E(YZ) = BXZ, \tag{2.37}$$

where we have used the fact that $\mathbf{1}_N'Z = 0$. From (2.36) and (2.37), we get

$$\begin{aligned} &E\left(\mathbf{y} - \tfrac{1}{N}Y\mathbf{1}_N\right) = B\left(\theta - \tfrac{1}{N}X\mathbf{1}_N\right) \quad \text{and} \\ &\mathrm{Cov}\left(\mathbf{y} - \tfrac{1}{N}Y\mathbf{1}_N\right) = \left(1 + \tfrac{1}{N}\right)\Sigma. \end{aligned} \tag{2.38}$$

Define

$$Y_0 = YZ, \qquad X_0 = XZ, \qquad \mathbf{y}_0 = \sqrt{\left(1 + \frac{1}{N}\right)}\left(\mathbf{y} - \frac{1}{N}Y\mathbf{1}_N\right)$$

and

$$\theta_0 = \sqrt{\left(1 + \frac{1}{N}\right)}\left(\theta - \frac{1}{N}X\mathbf{1}_N\right). \tag{2.39}$$

From (2.37), (2.38) and (2.39), we get

$$\begin{aligned} &E(Y_0) = BX_0, \quad \mathrm{Cov}(Y_0) = I_{N-1} \otimes \Sigma, \\ &E(\mathbf{y}_0) = B\theta_0, \quad \mathrm{Cov}(\mathbf{y}_0) = \Sigma, \end{aligned} \tag{2.40}$$

and Y_0 and $\mathbf{y}_0$ are independent. Hence (2.35) and (2.36) reduce to models of the type of (1.1) and (1.2) without an intercept. Once we obtain a confidence region for θ_0 in (2.40), a confidence region for θ can be obtained using the transformation in (2.39).

3. Concluding remarks. In the multivariate calibration problem, confidence regions that are currently available for the finite sample case have serious drawbacks. Apart from the fact that the shapes of these regions, in general, are unknown, the regions can be empty, or noninvariant. The confidence region derived in this article, being nonempty and invariant, overcomes some of these drawbacks. It appears that a confidence region satisfying these natural requirements was not available prior to our work. In the univariate case, we have derived a condition under which our region will be an interval.

However, the shape of our region, in general, is unknown. Being nonempty and invariant, our region should certainly be of use in hypothesis testing, that is, to verify if θ in (1.2) has a prespecified value. An ellipsoidal confidence region that is also nonempty and invariant is currently available only in the asymptotic case. Derivation of such a region in the finite sample case is still an open problem.

Acknowledgments. Thanks are due to an Associate Editor and the referees for several helpful suggestions. We are also grateful to Ms. Wenxing Zha for assistance with the numerical computations.

REFERENCES

BROWN, P. J. (1982). Multivariate calibration (with discussion). *J. Roy. Statist. Soc. Ser. B* **44** 287–321.

BROWN, P. J. and SUNDBERG, R. (1987). Confidence and conflict in multivariate calibration. *J. Roy. Statist. Soc. Ser. B* **49** 46–57.

DAVIS, A. W. and HAYAKAWA, T. (1987). Some distribution theory relating to confidence regions in multivariate calibration. *Ann. Inst. Statist. Math.* **39** 141–152.

FUJIKOSHI, Y. and NISHI, R. (1984). On the distribution of a statistic in multivariate inverse regression analysis. *Hiroshima Math. J.* **14** 215–225.

GIRI, N. C. (1977). *Multivariate Statistical Inference*. Academic, New York.

LIEFTINCK-KOEIJERS, C. A. J. (1988). Multivariate calibration: A generalization of the classical estimator. *J. Multivariate Anal.* **25** 31–44.

MARTENS, H. and NAES, T. (1989). *Multivariate Calibration*. Wiley, New York.

OMAN, S. D. (1988). Confidence regions in multivariate calibration. *Ann. Statist.* **16** 174–187.

OMAN, S. D. and WAX, Y. (1984). Estimating fetal age by ultrasound measurements: An example of multivariate calibration. *Biometrics* **40** 947–960.

RAO, C. R. (1974). *Linear Statistical Inference and Its Applications*. Wiley, New York.

SRIVASTAVA, M. S. and KHATRI, C. G. (1979). *An Introduction to Multivariate Statistics*. North-Holland, Amsterdam.

WOOD, J. T. (1982). Estimating the age of an animal; an application of multivariate calibration. In *Proceedings of the International Biometrics Conference* **11** 117–122.

WOOD, J. T. (1986). Confidence regions from multivariate calibration. In *Proceedings of the Pacific Statistical Congress* (I. S. Francis, B. F. J. Manly and F. C. Lam, eds.) 258–260. North-Holland, Amsterdam.

DEPARTMENT OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF MARYLAND
BALTIMORE, MARYLAND 21228

DEPARTMENT OF MATHEMATICAL
SCIENCES
UNIVERSITY OF NORTH CAROLINA
WILMINGTON, NORTH CAROLINA 28403