Rectangular confidence regions and prediction regions in multivariate calibration

Michael Daniel Lucagbo[a,b], Thomas Mathew[a]

[a]*Department of Mathematics & Statistics, University of Maryland Baltimore County, Baltimore, MD 21250, USA*

[b]*School of Statistics, University of the Philippines Diliman, Quezon City 1101, Philippines*

Corresponding author e-mail: mathew@umbc.edu (Thomas Mathew)

ORCID: 0000-0003-4152-3628 (Thomas Mathew)

ABSTRACT

The multivariate calibration problem deals with inference concerning an unknown value of a covariate vector based on an observation on a response vector. Two distinct scenarios are considered in the multivariate calibration problem: controlled calibration where the covariates are non-stochastic, and random calibration where the covariates are random. Under controlled calibration, a problem of interest is the computation of a confidence region for the unknown covariate vector. Under random calibration, the problem is that of computing a prediction region for the covariate vector. Assuming the standard multivariate normal linear regression model, rectangular confidence and prediction regions are derived using a parametric bootstrap approach. Numerical results show that the regions accurately maintain the coverage probabilities. The results are illustrated using examples. The regions currently available in the literature are all ellipsoidal, and this work is the first attempt to derive rectangular regions.

KEYWORDS: Controlled calibration; multivariate regression; parametric bootstrap; random calibration

1 Introduction

The multivariate calibration problem has received considerable attention in the literature in the setup of a multivariate linear regression model, under the multivariate normality assumption. The problem of interest consists of inference concerning an unknown or unobserved value of the covariate vector after observing the corresponding response. Two scenarios are considered in the literature: the case of controlled calibration, where the covariate vectors are treated as fixed, and the case of random calibration where they are treated as random. Under controlled calibration, it is of interest to compute a confidence region for the unknown covariate vector. Under random calibration, the problem of interest is the computation of a prediction region for the unobserved covariate vector. Some relevant references are Brown (1982), Fujikoshi and Nishii (1984), Davis and Hayakawa (1987), Brown and Sundberg (1987), Oman (1988), Mathew and Kasala (1994), Mathew and Zha (1996), Sundberg (1999) Bellio (2003), Benton, Krishnamoorthy and Mathew (2003), Devijver and Perthame (2018) and Bhaumik and Nordgren (2019). The articles by Osborne (1991) and Sundberg (1994) and the book by Brown (1993) provide a review of some of the relevant results.

Whether it is a multivariate confidence region or a prediction region, the regions available in the literature are all ellipsoidal. However, the pivot statistic used is such that the ellipsoidal shape is not always guaranteed. In fact even in the univariate case, the pivot may not yield a finite interval as the confidence interval; we refer to Brown (1982, 1993) for details. Even if we get an ellipsoidal region, such a region is not appropriate for drawing conclusions on the magnitude of the individual components of the multivariate parameter vector; this could be especially relevant when the components of the multivariate vector represent physical quantities measured in different units. This calls for the computation of a rectangular region. However, the rectangular confidence region or prediction region has to be derived so as to satisfy the coverage probability requirement, taking into account the cross-correlations. This article is an attempt to derive such rectangular regions; we develop a parametric bootstrap algorithm in order to accomplish this. In terms of maintaining the coverage probability, the rectangular regions we have derived are quite satisfactory.

Here is the parametric model and the inference problem. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be n response vectors, where each $\boldsymbol{X}_i$ is $p \times 1$, and let $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ be the corresponding covariate vectors, where each $\boldsymbol{w}_i$ is $q \times 1$, $i = 1, \ldots, n$. Write $\mathbf{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n]$ and $\mathbf{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n]$ so that $\mathbf{X}$ and $\mathbf{W}$ are $p \times n$ and $q \times n$ matrices, respectively. We assume the standard multivariate regression model:

$$\mathbf{X} = \mathbf{B}\mathbf{W} + \mathbf{E}, \tag{1.1}$$

where $\mathbf{E} = [\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n]$ is a $p \times n$ matrix of error terms; the $\boldsymbol{e}_i$s are $p \times 1$ iid multivariate normal random vectors with mean vector $\mathbf{0}$ and unknown covariance matrix $\boldsymbol{\Sigma}$, $i = 1, \ldots, n$. Furthermore, $\mathbf{B}$ is

a $p \times q$ matrix of unknown parameters. We also assume that $\text{rank}\,(\mathbf{W}) = q$ and that $n - q > p$. The least squares estimator of $\mathbf{B}$ and an unbiased estimator of $\boldsymbol{\Sigma}$, denoted by $\hat{\mathbf{B}}$ and $\hat{\boldsymbol{\Sigma}}$, respectively, are given by

$$\hat{\mathbf{B}} = \mathbf{X}\mathbf{W}'\left(\mathbf{W}\mathbf{W}'\right)^{-1},\ \hat{\boldsymbol{\Sigma}} = \frac{1}{n-q}\mathbf{X}\left[\boldsymbol{I}_n - \mathbf{W}'\left(\mathbf{W}\mathbf{W}'\right)^{-1}\mathbf{W}\right]\mathbf{X}'. \tag{1.2}$$

Now suppose we have another observation $\boldsymbol{X}_0$, independent of the $\mathbf{X}$ in (1.1), corresponding to an unknown $q \times 1$ covariate vector $\boldsymbol{w}_0$, such that given $\boldsymbol{w}_0$ we have the distribution

$$\boldsymbol{X}_0 \sim N_p\left(\mathbf{B}\boldsymbol{w}_0, \boldsymbol{\Sigma}\right). \tag{1.3}$$

The multivariate calibration problem consists of inference concerning the unknown vector $\boldsymbol{w}_0$.

In the controlled calibration scenario, the columns of the covariate matrix $\mathbf{W}$ are treated as fixed, whereas in the scenario of random calibration, the columns of $\mathbf{W}$ are treated as iid realizations of a $q \times 1$ random vector. Thus the vector $\boldsymbol{w}_0$ is a fixed unknown parameter in the former case, and a random vector in the latter case, having the same distribution as that of the columns of $\mathbf{W}$. When $\boldsymbol{w}_0$ is fixed, we shall denote it by $\boldsymbol{\theta}$. When it is random, we write it as $\boldsymbol{W}_0$. For the controlled calibration case, we shall address the problem of computing a rectangular confidence region for the parameter vector $\boldsymbol{\theta}$. In the random calibration case, we take up the problem of computing a rectangular prediction region for $\boldsymbol{W}_0$. In each case, a parametric bootstrap approach is developed to get the required regions. The controlled calibration case is discussed in Section 2, and the random calibration case in Section 3. Numerical results indicate that the parametric bootstrap-based confidence regions and prediction regions accurately maintain the coverage probabilities. The results are illustrated with examples.

We conclude this introduction with the following observations. The pivot statistic we have used for the computation of a rectangular confidence region (and also a rectangular prediction region) is a maximum of several univariate pivot statistics. The distribution of such a maximum is then estimated using a parametric bootstrap. Such a pivot statistic has been available in the literature for the computation of simultaneous confidence bands, and the resulting solution is referred to as a "sup-t confidence band". The sup-t approach has been utilized for computing simultaneous confidence bands and simultaneous prediction bands, especially in the context of multivariate time series; see Montiel Olea, Luis and Plagborg-Moller (2019) for the development of simultaneous confidence bands, and Wolf and Wunderli (2015) for the development of simultaneous prediction bands. We refer to these articles for other literature relevant to sup-t bands. A survey of the relevant literature is available in Lütkepohl, Staszewska-Bystrova and Winker (2020).

2 Rectangular confidence regions under controlled calibration

Recall that the data consist of the $p \times n$ matrix $\mathbf{X}$ and the $p \times 1$ vector $\boldsymbol{X}_0$ following the models:

$$\mathbf{X} = \mathbf{B}\mathbf{W} + \mathbf{E}, \quad \boldsymbol{X}_0 = \mathbf{B}\boldsymbol{\theta} + \boldsymbol{e}_0, \tag{2.1}$$

where $\mathbf{B}$ is an unknown parameter matrix of dimension $p \times q$, $\mathbf{W}$ is the $q \times n$ matrix whose columns are q-dimensional covariate vectors, and $\boldsymbol{\theta}$ is a $q \times 1$ unknown parameter. The n columns of the $p \times n$ matrix $\mathbf{E}$, and the $p \times 1$ column vector $\boldsymbol{e}_0$ are independent, each having the distribution $N_p(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is an unknown $p \times p$ positive definite matrix. We note that since $\mathbf{B}$ is a $p \times q$ matrix, the identifiability of $\boldsymbol{\theta}$ requires the condition $p \geq q$, and we proceed under this assumption. We shall address the problem of computing a rectangular confidence region for $\boldsymbol{\theta}$, and numerically assess the performance of our region based on estimated coverage probabilities. It should be noted that when $\boldsymbol{\theta}$ is a scalar, a parametric bootstrap approach is pursued in Benton, Krishnamoorthy and Mathew (2003) for the interval estimation of $\boldsymbol{\theta}$.

We first note that if $\mathbf{B}$ and $\boldsymbol{\Sigma}$ are known, the generalized least squares estimator of theta, say $\tilde{\boldsymbol{\theta}}$, and its distribution, are given by

$$\widetilde{\boldsymbol{\theta}} = \left(\mathbf{B}'\boldsymbol{\Sigma}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_0 \sim N_q\left(\boldsymbol{\theta}, \left(\mathbf{B}'\boldsymbol{\Sigma}^{-1}\mathbf{B}\right)^{-1}\right). \tag{2.2}$$

Let us now use the following notation: for a vector $\boldsymbol{a}$, the ith component will be denoted by $\boldsymbol{a}_{(i)}$, and for a matrix $\mathbf{A}$, the ith diagonal element will be denoted by $\mathbf{A}_{(ii)}$. From (2.2) we conclude:

$$\frac{\widetilde{\boldsymbol{\theta}}_{(i)} - \boldsymbol{\theta}_{(i)}}{\sqrt{\left(\mathbf{B}'\boldsymbol{\Sigma}^{-1}\mathbf{B}\right)^{-1}_{(ii)}}} \sim N(0,1) \tag{2.3}$$

for each $i = 1, \ldots, q$. Therefore the quantities in (2.3) are identically distributed and are pivots for the corresponding $\boldsymbol{\theta}_{(i)}$; however, they are not independent. In view of (2.3), a $100(1-\alpha)\%$ rectangular confidence region for the $\boldsymbol{\theta}_{(i)}$s can be taken to be of the form $\tilde{\boldsymbol{\theta}}_{(i)} \pm \kappa\sqrt{\left(\mathbf{B}'\boldsymbol{\Sigma}^{-1}\mathbf{B}\right)^{-1}_{(ii)}}$, where the factor κ is the $100(1-\alpha)$th percentile of

$$\max_{1 \leq i \leq q}\left|\frac{\widetilde{\boldsymbol{\theta}}_{(i)} - \boldsymbol{\theta}_{(i)}}{\sqrt{\left(\mathbf{B}'\boldsymbol{\Sigma}^{-1}\mathbf{B}\right)^{-1}_{(ii)}}}\right|. \tag{2.4}$$

Since $\mathbf{B}$ and $\boldsymbol{\Sigma}$ are unknown, we now replace them by their respective estimates $\hat{\mathbf{B}}$ and $\hat{\boldsymbol{\Sigma}}$ given in (1.2). After replacing $\mathbf{B}$ and $\boldsymbol{\Sigma}$ in (2.4) with $\hat{\mathbf{B}}$ and $\hat{\boldsymbol{\Sigma}}$, respectively, we shall compute the $100(1-\alpha)$th percentile

of the resulting quantity using a parametric bootstrap. We shall also appeal to the following results:

$$\text{vec}\left(\hat{\mathbf{B}}\right) \sim N_{pq}\left(\text{vec}\left(\mathbf{B}\right), \left(\mathbf{W}\mathbf{W}'\right)^{-1} \otimes \mathbf{\Sigma}\right), \ \hat{\mathbf{\Sigma}} \sim W\left(n-q, \frac{1}{n-q}\mathbf{\Sigma}\right),$$

where $\hat{\mathbf{B}}$ and $\hat{\mathbf{\Sigma}}$ are also independent. The parametric bootstrap approach for computing a rectangular confidence region for $\boldsymbol{\theta}$ is summarized in Algorithm 1.

Algorithm 1: Rectangular confidence region under controlled calibration

1. Compute the estimates $\hat{\mathbf{B}}$ and $\hat{\mathbf{\Sigma}}$ given in (1.2), and also compute $\hat{\boldsymbol{\theta}} = \left(\hat{\mathbf{B}}'\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{B}}\right)^{-1}\hat{\mathbf{B}}'\hat{\mathbf{\Sigma}}^{-1}\boldsymbol{X}_0$.
2. For $b = 1, \ldots, B$,
   i. Generate $\text{vec}\left(\hat{\mathbf{B}}_b^*\right) \sim N_{pq}\left(\text{vec}\left(\hat{\mathbf{B}}\right), \left(\mathbf{W}\mathbf{W}'\right)^{-1} \otimes \hat{\mathbf{\Sigma}}\right)$,
   $\hat{\mathbf{\Sigma}}_b^* \sim W\left(n-q, \frac{1}{n-q}\hat{\mathbf{\Sigma}}\right)$, and $\boldsymbol{X}_{0b}^* \sim N_p\left(\hat{\mathbf{B}}\hat{\boldsymbol{\theta}}, \hat{\mathbf{\Sigma}}\right)$.
   ii. Compute $\hat{\boldsymbol{\theta}}_b^* = \left(\hat{\mathbf{B}}_b^{*'}\hat{\mathbf{\Sigma}}_b^{*-1}\hat{\mathbf{B}}_b^*\right)^{-1}\hat{\mathbf{B}}_b^{*'}\hat{\mathbf{\Sigma}}_b^{*-1}\boldsymbol{X}_{0b}^*$
   iii. Compute $k_b = \max_{1\le i\le q}\left|\frac{\hat{\boldsymbol{\theta}}_{b(i)}^* - \hat{\boldsymbol{\theta}}_{(i)}}{\sqrt{\left(\hat{\mathbf{B}}_b^{*'}\hat{\mathbf{\Sigma}}_b^{*-1}\hat{\mathbf{B}}_b^*\right)^{-1}_{(ii)}}}\right|$.
3. Compute the $(1-\alpha)$-quantile of $k_1, \ldots, k_B$, call this k.
4. The $100(1-\alpha)\%$ rectangular confidence region for $\boldsymbol{\theta}$ is given by

$$\left[\hat{\boldsymbol{\theta}}_{(1)} \pm k\sqrt{\left(\hat{\mathbf{B}}'\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{B}}\right)^{-1}_{(11)}}\right] \times \cdots \times \left[\hat{\boldsymbol{\theta}}_{(q)} \pm k\sqrt{\left(\hat{\mathbf{B}}'\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{B}}\right)^{-1}_{(qq)}}\right].$$

.

One-sided confidence regions

To compute one-sided upper confidence limits for $\boldsymbol{\theta}$, we simply change the definition of k_b in Step 2 (iii) of Algorithm 1 to

$$k_b = \max_{1\le i\le q}\frac{\hat{\boldsymbol{\theta}}_{b(i)}^* - \hat{\boldsymbol{\theta}}_{(i)}}{\sqrt{\left(\hat{\mathbf{B}}_b^{*'}\left(\hat{\mathbf{\Sigma}}_b^*\right)^{-1}\hat{\mathbf{B}}_b^*\right)^{-1}_{(ii)}}},$$

and take k to be the α-quantile of $k_1, \ldots, k_B$. The resulting $100(1-\alpha)\%$ confidence region having only upper limits is given by

$$\left(-\infty, \hat{\boldsymbol{\theta}}_{(1)} + k\sqrt{\left(\hat{\mathbf{B}}'\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{B}}\right)^{-1}_{(11)}}\right] \times \cdots \times \left(-\infty, \hat{\boldsymbol{\theta}}_{(q)} + k\sqrt{\left(\hat{\mathbf{B}}'\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{B}}\right)^{-1}_{(qq)}}\right].$$

Similarly, to compute one-sided lower confidence limits for $\boldsymbol{\theta}$, we use

$$k_b = \min_{1 \le i \le q} \frac{\hat{\boldsymbol{\theta}}^*_{b(i)} - \hat{\boldsymbol{\theta}}_{(i)}}{\sqrt{\left(\hat{\mathbf{B}}^{*'}_b \left(\hat{\boldsymbol{\Sigma}}^*_b \right)^{-1} \hat{\mathbf{B}}^*_b \right)^{-1}_{(ii)}}}$$

and take k to be the α-quantile of $k_1, \ldots, k_B$. The resulting $100(1-\alpha)\%$ confidence region having only lower limits is given by

$$\left[\hat{\boldsymbol{\theta}}_{(1)} + k \sqrt{\left(\hat{\mathbf{B}}' \hat{\boldsymbol{\Sigma}}^{-1} \hat{\mathbf{B}} \right)^{-1}_{(11)}}, \infty \right) \times \cdots \times \left[\hat{\boldsymbol{\theta}}_{(q)} + k \sqrt{\left(\hat{\mathbf{B}}' \hat{\boldsymbol{\Sigma}}^{-1} \hat{\mathbf{B}} \right)^{-1}_{(qq)}}, \infty \right).$$

Inclusion of an intercept term

The methodology outlined in Algorithm 1 assumes that the vector $\boldsymbol{\theta}$ is a completely unknown parameter, which is the case when the multivariate linear regression model does not include an intercept term. However, in many applications the model will include an intercept term; see Brown (1982) and Brown and Sundberg (1987). Instead of (2.1), we now have the models

$$\mathbf{X} = \boldsymbol{a} \mathbf{1}_n^{'} + \mathbf{B}\mathbf{W} + \mathbf{E}, \quad \boldsymbol{X}_0 = \boldsymbol{a} + \mathbf{B}\boldsymbol{\theta} + \boldsymbol{e}_0, \tag{2.5}$$

where $\boldsymbol{a}$ is a $(p \times 1)$ intercept vector, and we make the same assumptions as for the model (2.1). For inference concerning $\boldsymbol{\theta}$, it is possible to have a reduced form of the models in (2.5), where the intercept term is absent, as noted in Mathew and Kasala (1994). For this, let $\mathbf{Z}$ be an $n \times (n-1)$ matrix such that $\mathbf{Q} = [(1/\sqrt{n})\, \mathbf{1}_n, \mathbf{Z}]$ is an orthogonal matrix. Now define the "adjusted quantities"

$$\mathbf{X}_a = \mathbf{X}\mathbf{Z}, \quad \mathbf{W}_a = \mathbf{W}\mathbf{Z}, \; \boldsymbol{X}_{0,a} = \sqrt{1 + \frac{1}{n}} \left(\boldsymbol{X}_0 - \frac{1}{n} \mathbf{X} \mathbf{1}_n \right), \text{ and } \boldsymbol{\theta}_a = \sqrt{1 + \frac{1}{n}} \left(\boldsymbol{\theta} - \frac{1}{n} \mathbf{W} \mathbf{1}_n \right). \tag{2.6}$$

It is easy to verify that the following models hold for the quantities $\mathbf{X}_a$ and $\boldsymbol{X}_{0,a}$:

$$\mathbf{X}_a = \mathbf{B}\mathbf{W}_a + \mathbf{E}_a, \quad \boldsymbol{X}_{0,a} = \mathbf{B}\boldsymbol{\theta}_a + \boldsymbol{e}_{0,a}, \tag{2.7}$$

where $\boldsymbol{e}_{0,a}$ and the $(n-1)$ columns of $\mathbf{E}_a$ are independent and identically distributed as $N_p(\mathbf{0}, \boldsymbol{\Sigma})$. Thus we are back to the models in (2.1) with n replaced by $n-1$. Algorithm 1 can now be implemented for computing a rectangular confidence region for $\boldsymbol{\theta}_a$ (the sample size n appearing in Algorithm 1 has to be replaced with $n-1$). From the definition of $\boldsymbol{\theta}_a$ given in (2.6) it should be clear that from the rectangular confidence region for $\boldsymbol{\theta}_a$, we can obtain the corresponding region for $\boldsymbol{\theta}$. Here we also want to note that even though the choice of the matrix $\mathbf{Z}$ is not unique, the estimated confidence factor k does not depend on the choice of the matrix $\mathbf{Z}$. In order to see this, we note that the orthogonality of

$\mathbf{Q}$ implies $\mathbf{Z}\mathbf{Z}' = \boldsymbol{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'$. Making use of this property, it can be verified that the estimates $\hat{\mathbf{B}}$ and $\hat{\boldsymbol{\Sigma}}$ computed using the model (2.7) are free of the particular choice of $\mathbf{Z}$. Consequently, the estimated confidence factor k is also free of the choice of the matrix $\mathbf{Z}$.

Numerical results

In order to estimate the coverage probabilities of the proposed rectangular confidence regions, we shall work with the following canonical form corresponding to (2.1); see Mathew and Zha (1996):

$$\mathbf{X}_1 = \mathbf{X}\mathbf{W}'\left(\mathbf{W}\mathbf{W}'\right)^{-1/2}, \mathbf{B}_1 = \mathbf{B}\left(\mathbf{W}\mathbf{W}'\right)^{1/2}, \boldsymbol{\theta}_1 = \left(\mathbf{W}\mathbf{W}'\right)^{-1/2}\boldsymbol{\theta}, \tag{2.8}$$

along with $\hat{\boldsymbol{\Sigma}}$ defined in (1.2). Then we have

$$\mathrm{vec}\left(\mathbf{X}_1\right) \sim N_{pq}\left(\mathrm{vec}\left(\mathbf{B}_1\right), \boldsymbol{I}_q \otimes \boldsymbol{\Sigma}\right), \ \hat{\boldsymbol{\Sigma}} \sim W\left(n-q, \frac{1}{n-q}\boldsymbol{\Sigma}\right), \boldsymbol{X}_0 \sim N_p\left(\mathbf{B}_1\boldsymbol{\theta}_1, \boldsymbol{\Sigma}\right). \tag{2.9}$$

The advantage of using the canonical quantities for estimating the coverage probabilities is that we do not have to specify values for the covariates. We shall use the following true values of the respective parameters:

$$\mathbf{B}_1 = \begin{pmatrix} 1 & 2 & 6 & 2 \\ 3 & 4 & 7 & 9 \end{pmatrix}', \quad \boldsymbol{\theta}_1 = \begin{pmatrix} 0.7 & 0.9 \end{pmatrix}', \quad \boldsymbol{\Sigma} = (1-\rho)\,\boldsymbol{I}_p + \rho\mathbf{1}_p\mathbf{1}_p' \tag{2.10}$$

with $\rho = 0.9, 0.5, 0.3, -0.3$. In addition, we shall consider the sample sizes $n = 30$, 50, and 100, and a 95% nominal level. The estimated coverage probabilities for the two-sided rectangular confidence region for $\boldsymbol{\theta}_1$, estimated using 5000 simulated samples, are shown in Table 1. Also included are the expected values of the factor k and the expected volumes of the confidence region. We see that at least for the scenarios considered in the simulations, the coverage probabilities are quite satisfactory even for the small sample size situation. The correlations do not appear to affect the coverage probabilities. However, the correlations do have an effect on the expected values of the factor k and the expected volumes of the confidence region; the expected volume decreases with increasing correlation. Furthermore, both of these quantities are decreasing with the sample size, as expected.

Here we would like to make the following comment on the expected volume. According to a general result due to Gleser and Hwang (1987), there does not exist a confidence region having finite expected volume and positive confidence in the controlled calibration setup. However, it should be noted that the covariates (the entries of the matrix $\mathbf{W}$) represent physical quantities, and will be bounded with

Table 1: Estimated coverage probabilities along with the expected values of the factor k and the expected volumes of rectangular two-sided confidence regions under controlled calibration for the parameter choices in (2.10); nominal level = 95%

n	ρ	Coverage	$E(k)$	E(Volume)	ρ	Coverage	$E(k)$	E(Volume)
100		0.9526	2.8820	0.1876		0.9502	2.9167	0.5986
50	0.90	0.9492	3.0059	0.1971	0.30	0.9492	3.0310	0.6131
30		0.9536	3.1853	0.2111		0.9568	3.2209	0.6363
100		0.9520	2.9088	0.5296		0.9414	2.8279	0.6294
50	0.50	0.9554	3.0182	0.5453	-0.30	0.9462	2.9640	0.6410
30		0.9510	3.2084	0.5714		0.9492	3.1358	0.6601

known bounds. Consequently, the parameter of interest $\boldsymbol{\theta}$ in the controlled calibration problem will be a bounded quantity, with known bounds for each component. If w_{ij} denotes the (ij)th element of $\mathbf{W}$, and $\boldsymbol{\theta}_{(i)}$ denotes the ith component of $\boldsymbol{\theta}$, then it is reasonable to assume that

$$\min_{1\leq j\leq n} w_{ij} \leq \boldsymbol{\theta}_{(i)} \leq \max_{1\leq j\leq n} w_{ij},$$

$i = 1, \ldots, q$. This amounts to assuming that $\boldsymbol{\theta}$ is "like the columns of $\mathbf{W}$." In terms of the canonical form (2.8), we can assume that each component of $\boldsymbol{\theta}_1$ is bounded by 0 and 1. Thus, after computing the rectangular confidence region for $\boldsymbol{\theta}_1$, our recommendation is to intersect each marginal interval with the interval $(0, 1)$. This will of course result in a finite expected volume for the rectangular confidence region. This is the expected volume reported in Table 1. We also note that the canonical form (2.8) is used only for the purpose of simulation and cannot be used to derive rectangular confidence regions for $\boldsymbol{\theta}$, since a rectangular region for $\boldsymbol{\theta}_1$ cannot be translated into a rectangular region for $\boldsymbol{\theta}$.

We shall now illustrate our results using two examples. The first example is the paint finish data example taken from Brown (1982). This data set has been analyzed in several articles on calibration, and has become a classic example to illustrate calibration. Our second example is a more recent example, taken from Atkins et al. (2015) and Bhaumik and Nordgren (2019).

Example 1: Analysis of paint finish data

We now apply the proposed procedure to compute a rectangular controlled calibration confidence region on a data set taken from Brown (1982). This data set has been analyzed in several articles on the calibration problem. The data came from a paint finish experimental study where the factors are pigment level and viscosity. Data were obtained corresponding to three values of the pigment level (0 percent, 0.15 percent, and 0.30 percent) and three values of the viscosity (30, 33, 36 seconds in an efflux cup). In our analysis, we shall assume the three values of both the pigment level and viscosity to be 0, 1 and 2, without loss of generality. The effects of these factors were assessed using six response variables: three spectrometer measurements of incident light, integrated reflectance with normal light, and peak-

height and band-width on a recording goniophotometer. Based on an initial screening, it turned out that four of the six response variables could be discarded without incurring any loss of information. The variables retained were one of the three spectrometer measurements of incident light and the integrated reflectance with normal light. Each of the nine factor combinations was replicated four times, resulting in 36 bivariate observations. The 36 observations are displayed in Brown (1982, Table 2), and the two responses under consideration are the variables Y_1 and Y_4 given in Brown (1982, Table 2). Of the 36 bivariate observations, 9 observations are extracted for calibration; that is, for estimating the pigmentation and viscosity, assuming them to be unknown. These 9 observations are given in Table 2. The remaining 27 observations in Brown (1982, Table 2) were used to fit the regression model.

We use the model (2.5) with an intercept, consistent with Brown (1982). Consequently, the region is computed following the steps explained in the subsection on models with an intercept term. The least squares estimates of $\mathbf{B}$ and $\mathbf{\Sigma}$ are:

$$\hat{\mathbf{B}} = \begin{bmatrix} 0.0300 & -0.1278 \\ 1.3694 & -1.6922 \end{bmatrix} \quad \hat{\mathbf{\Sigma}} = \begin{bmatrix} 0.00543 & 0.02270 \\ 0.02270 & 0.41690 \end{bmatrix}.$$

We shall compute a rectangular confidence region for the two components of the vector $\boldsymbol{\theta}$ corresponding to the 9 observations given in Table 2. The table also gives the rectangular confidence regions obtained by applying Algorithm 1. We recall that the three values of both the pigment level and viscosity, at which data were obtained, are assumed to be 0, 1 and 2. Since each component of the parameter vector $\boldsymbol{\theta}$ is bounded between 0 and 2, the intervals in Table 2 have been truncated to be between 0 and 2. By referring to Brown (1982, Table 2), we can verify that all of the resulting confidence regions contain the corresponding true values of the pigment level and viscosity. Being of rectangular shape, the confidence regions in Table 2 can be used to draw inference concerning the pigment level and viscosity, individually.

Table 2: Controlled calibration 95% rectangular confidence regions for pigmentation and viscosity at nine different values of the response vector

Response	Covariate	Confidence region	Response	Covariate	Confidence region
$(1.87, 40.68)'$	Pigmentation	0 to 1.60	$(1.52, 35.65)'$	Pigmentation	0 to 1.95
	Viscosity	0 to 1.13		Viscosity	1.33 to 2
$(1.79, 39.83)'$	Pigmentation	0 to 1.45	$(1.92, 38.17)'$	Pigmentation	0.91 to 2
	Viscosity	0 to 1.69		Viscosity	0 to 1.17
$(1.61, 37.36)'$	Pigmentation	0 to 1.59	$(1.85, 37.17)'$	Pigmentation	0.97 to 2
	Viscosity	0.62 to 2		Viscosity	0 to 1.71
$(1.78, 38.73)'$	Pigmentation	0 to 2	$(1.62, 34.16)'$	Pigmentation	0.87 to 2
	Viscosity	0 to 1.96		Viscosity	0.98 to 2
$(1.68, 38.64)'$	Pigmentation	0 to 1.37			
	Viscosity	0.05 to 2			

Example 2: The link to learning study

Our second example is based on data from the "Link to Learning Study" reported in Atkins et al. (2015). The study was focused on children living in high-poverty urban communities, and examined whether their behavior at school and at home can be improved by community mental health services that target the empirical predictors of learning. This application and the relevant data analysis are also reported in Bhaumik and Nordgren (2019); these authors have investigated a prediction problem and a calibration problem, and have addressed the problem of computing ellipsoidal confidence regions. We shall consider the following five-dimensional response vector, as given in Bhaumik and Nordgren (2019, Section 2.4): Curriculum-based measures (CBM), Academic Competence Evaluation Scales (ACES), Homework Problem Checklist (HPC), Social Skills Rating System: Social Skills rated by parent (SOC) and Social Skills Rating System: Problem Behaviors rated by parent (PB). The covariates considered are a parent self-efficacy score (PSES) and the child's grade. The data analyzed in Bhaumik (2019) consist of 175 observations. We shall fit the model (2.5) using 172 observations and use the remaining three for calibration; these three observations are given in Table 3. Assuming an unknown value for the corresponding covariate vector (PSES, Grade), a rectangular confidence region will be constructed.

We shall once again use the model (2.5) with an intercept. The least squares estimates of $\mathbf{B}$ and $\mathbf{\Sigma}$ are:

$$\hat{\mathbf{B}} = \begin{bmatrix} 0.5199 & 0.5247 \\ 0.3226 & -0.0307 \\ -0.4356 & 0.0692 \\ 0.4412 & 0.0140 \\ -0.5496 & 0.0575 \end{bmatrix} \quad \hat{\mathbf{\Sigma}} = \begin{bmatrix} 2.92267 & 0.51584 & -0.20249 & -0.06435 & 0.24461 \\ 0.51584 & 0.70314 & -0.10948 & -0.21458 & 0.23672 \\ -0.20249 & -0.10948 & 0.42504 & -0.11444 & 0.39074 \\ -0.06435 & -0.21458 & -0.11444 & 1.05164 & -0.54653 \\ 0.24461 & 0.23672 & 0.39074 & -0.54653 & 1.72339 \end{bmatrix}.$$

We shall compute a rectangular confidence region for the two components of the vector $\boldsymbol{\theta}$ = (PSES, Grade)$'$ corresponding to the three observations given in Table 3. The table also gives the rectangular confidence regions obtained by applying Algorithm 1.

We would like to make one comment concerning the results reported in Table 3. In the 172 observations used to fit the model (2.5), the PSES values were between 2.33 and 6, and the Grade values were between 0 and 7. Thus we have assumed that these intervals also provide bounds for the respective parameters. Consequently, each confidence interval of the rectangular confidence regions reported in Table 3 has been truncated to be within the respective bounds (similar to the previous example). The fact that confidence

Table 3: Controlled calibration 95% rectangular confidence regions for PSES and Grade for three different values of the response vector

Response	Covariate	True value	Confidence region
$(5.017, 2.89, 0.15, 5.8, 1.4)'$	PSES	4.71	4.27 to 6
	Grade	4	0 to 7
$(0.284, 1.72, 0.75, 4.1, 2.0)'$	PSES	4.71	2.33 to 6
	Grade	2	0 to 6.17
$(0.567, 1.79, 2.35, 4.6, 4.6)'$	PSES	4.57	2.33 to 5.44
	Grade	2	0 to 7

regions can be outside the parameter space is a phenomenon encountered in the calibration problem, even in the univariate case. In this example also, we note that the rectangular confidence regions include the true parameter values, which are also given in Table 3.

3 Rectangular prediction regions under random calibration

In the case of random calibration, we shall consider models with an intercept term. Thus let $\boldsymbol{X}_i$, $i = 1, 2, \ldots, n$, and $\boldsymbol{X}_0$ be independent $p \times 1$ random vectors following the models

$$\begin{aligned} \boldsymbol{X}_i &= \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{W}_i + \boldsymbol{e}_i, \quad i = 1, \ldots, n \\ \boldsymbol{X}_0 &= \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{W}_0 + \boldsymbol{e}_0, \end{aligned} \tag{3.1}$$

where the $\boldsymbol{W}_i$, $i = 1, \ldots, n$, and $\boldsymbol{W}_0$ are $q \times 1$ covariate vectors, $\boldsymbol{\beta}_0$ is the $p \times 1$ intercept vector (unknown), $\boldsymbol{\beta}_1$ is a $p \times q$ matrix of unknown parameters, and $\boldsymbol{e}_i$ and $\boldsymbol{e}_0$ are $p \times 1$ vectors of error terms. If we express the above model in matrix form as in (2.1), then $\mathbf{B} = [\boldsymbol{\beta}_0, \boldsymbol{\beta}_1]$ is a $p \times (q+1)$ matrix, $\mathbf{X}$ is a $p \times n$ matrix of response vectors, and the covariate matrix $\mathbf{W}$ becomes a $(q+1) \times n$ matrix whose first row is a vector of ones. We shall ignore this inconsistency in dimension between the model (2.1) considered for controlled calibration, and the models (3.1) for random calibration. In (3.1), we assume that $\left[\boldsymbol{W}_i', \boldsymbol{e}_i'\right]'$, $i = 0, 1, \ldots, n$, are independently distributed as

$$\begin{bmatrix} \boldsymbol{W}_i \\ \boldsymbol{e}_i \end{bmatrix} \sim N_{p+q} \left(\begin{bmatrix} \boldsymbol{\mu}_w \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} \end{bmatrix} \right), \tag{3.2}$$

$i = 0, 1, \ldots, n$, where $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ are unknown positive definite matrices. It follows from (3.1) and (3.2) that $E(\boldsymbol{X}_i) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{\mu}_w$, $V(\boldsymbol{X}_i) = \boldsymbol{\beta}_1 \boldsymbol{\Psi} \boldsymbol{\beta}_1' + \boldsymbol{\Sigma}$, and $\mathrm{Cov}(\boldsymbol{W}_i, \boldsymbol{X}_i) = \boldsymbol{\Psi} \boldsymbol{\beta}_1'$. Hence,

$$\begin{bmatrix} \boldsymbol{W}_i \\ \boldsymbol{X}_i \end{bmatrix} \sim N_{p+q} \left(\begin{bmatrix} \boldsymbol{\mu}_w \\ \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{\mu}_w \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Psi} & \boldsymbol{\Psi} \boldsymbol{\beta}_1' \\ \boldsymbol{\beta}_1 \boldsymbol{\Psi}' & \boldsymbol{\beta}_1 \boldsymbol{\Psi} \boldsymbol{\beta}_1' + \boldsymbol{\Sigma} \end{bmatrix} \right), \tag{3.3}$$

$i = 0, 1, \ldots, n$. The observed data consist of $(\boldsymbol{X}_i, \boldsymbol{W}_i)$, $i = 1, \ldots, n$, along with $\boldsymbol{X}_0$. We shall address the computation of a rectangular prediction region for the $q \times 1$ vector $\boldsymbol{W}_0$. The unknown parameters $\boldsymbol{\mu}_w$, $\boldsymbol{\beta}_0$, $\boldsymbol{\beta}_1$, $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ are estimated as follows:

$$\hat{\boldsymbol{\mu}}_w = \bar{\boldsymbol{W}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{W}_i,\ \hat{\mathbf{B}} = \left[\hat{\boldsymbol{\beta}}_0, \hat{\boldsymbol{\beta}}_1\right] = \mathbf{X}\mathbf{W}'\left(\mathbf{W}\mathbf{W}'\right)^{-1}$$
$$\hat{\boldsymbol{\Psi}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\boldsymbol{W}_i - \bar{\boldsymbol{W}}\right)\left(\boldsymbol{W}_i - \bar{\boldsymbol{W}}\right)',\ \hat{\boldsymbol{\Sigma}} = \frac{1}{n-q-1}\mathbf{X}\left[\boldsymbol{I}_n - \mathbf{W}'\left(\mathbf{W}\mathbf{W}'\right)^{-1}\mathbf{W}\right]\mathbf{X}'. \quad (3.4)$$

Recall that $\mathbf{W}$ is a $(q+1) \times n$ matrix whose first row is a vector of ones. From (3.3), we see that the conditional distribution of $\boldsymbol{W}_i$ given $\boldsymbol{X}_i$ is

$$\boldsymbol{W}_i | \boldsymbol{X}_i \sim N_q\left(\boldsymbol{\mu}_{\boldsymbol{W}|\boldsymbol{X}_i}, \boldsymbol{\Sigma}_{\boldsymbol{W}|\boldsymbol{X}}\right), \quad (3.5)$$

$$\text{where } \boldsymbol{\mu}_{\boldsymbol{W}|\boldsymbol{X}_i} = \boldsymbol{\mu}_w + \boldsymbol{\Psi}\boldsymbol{\beta}_1'\left(\boldsymbol{\beta}_1\boldsymbol{\Psi}\boldsymbol{\beta}_1' + \boldsymbol{\Sigma}\right)^{-1}\left(\boldsymbol{X}_i - \boldsymbol{\beta}_0 - \boldsymbol{\beta}_1\boldsymbol{\mu}_w\right)$$
$$\text{and } \boldsymbol{\Sigma}_{\boldsymbol{W}|\boldsymbol{X}} = \boldsymbol{\Psi} - \boldsymbol{\Psi}\boldsymbol{\beta}_1'\left(\boldsymbol{\beta}_1\boldsymbol{\Psi}\boldsymbol{\beta}_1' + \boldsymbol{\Sigma}\right)^{-1}\boldsymbol{\beta}_1\boldsymbol{\Psi}'. \quad (3.6)$$

for each $i = 0, 1, \ldots, n$. We have used the notation $\boldsymbol{\mu}_{\boldsymbol{W}|\boldsymbol{X}_i}$ to emphasize that the conditional mean depends on $\boldsymbol{X}_i$; however, the conditional covariance matrix does not depend on the specific $\boldsymbol{X}_i$ and therefore we have used the notation $\boldsymbol{\Sigma}_{\boldsymbol{W}|\boldsymbol{X}}$. Estimates of $\boldsymbol{\mu}_{\boldsymbol{W}|\boldsymbol{X}_i}$ and $\boldsymbol{\Sigma}_{\boldsymbol{W}|\boldsymbol{X}}$, say $\hat{\boldsymbol{\mu}}_{\boldsymbol{W}|\boldsymbol{X}_i}$ and $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{W}|\boldsymbol{X}}$, respectively, can be obtained by replacing the unknown parameters by the corresponding estimates given in (3.4). From (3.5) it follows that,

$$\frac{\boldsymbol{W}_{0(j)} - \boldsymbol{\mu}_{\boldsymbol{W}|\boldsymbol{X}_0(j)}}{\sqrt{\left(\boldsymbol{\Sigma}_{\boldsymbol{W}|\boldsymbol{X}}\right)_{(jj)}}} \sim N(0,1) \quad (3.7)$$

$j = 1, \ldots, q$, where, as before, for a vector $\boldsymbol{a}$, $\boldsymbol{a}_{(j)}$ denotes its jth component, and for a matrix $\mathbf{A}$, $\mathbf{A}_{(jj)}$ denotes its jth diagonal element. If the true values of the parameters were known, then the prediction factor for a two-sided rectangular prediction region for $\boldsymbol{W}_0$ can be taken to be the $100(1-\alpha)$th percentile of the distribution of

$$\max_{1 \le j \le q}\left|\frac{\boldsymbol{W}_{0(j)} - \boldsymbol{\mu}_{\boldsymbol{W}|\boldsymbol{X}_0(j)}}{\sqrt{\left(\boldsymbol{\Sigma}_{\boldsymbol{W}|\boldsymbol{X}}\right)_{(jj)}}}\right|. \quad (3.8)$$

However, since the parameters are unknown, we replace them by their corresponding estimates (see (3.4)), resulting in the quantity

$$\max_{1 \le j \le q}\left|\frac{\boldsymbol{W}_{0(j)} - \hat{\boldsymbol{\mu}}_{\boldsymbol{W}|\boldsymbol{X}_0(j)}}{\sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{W}|\boldsymbol{X}}\right)_{(jj)}}}\right|. \quad (3.9)$$

The quantity in (3.9) is our pivot statistic for computing a rectangular prediction region for $\boldsymbol{W}_0$. A parametric bootstrap approach can now be employed for estimating the required $100(1-\alpha)$th percentile, similar to what was done in the case of controlled calibration; Algorithm 2 summarizes the steps.

Algorithm 2. Rectangular two-sided prediction region under random calibration

1. Compute the estimates $\hat{\boldsymbol{\mu}}_w$, $\hat{\boldsymbol{\beta}}_0$, $\hat{\boldsymbol{\beta}}_1$, $\hat{\boldsymbol{\Psi}}$, $\hat{\boldsymbol{\Sigma}}$ given in (3.4).

2. For $b = 1, \ldots, B$,

   i. Generate

$$\begin{bmatrix} \boldsymbol{W}^*_{ib} \\ \boldsymbol{X}^*_{ib} \end{bmatrix} \sim N_{p+q}\left(\begin{bmatrix} \hat{\boldsymbol{\mu}}_w \\ \hat{\boldsymbol{\beta}}_0 + \hat{\boldsymbol{\beta}}_1 \hat{\boldsymbol{\mu}}_x \end{bmatrix}, \begin{bmatrix} \hat{\boldsymbol{\Psi}} & \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\beta}}_1^{'} \\ \hat{\boldsymbol{\beta}}_1\hat{\boldsymbol{\Psi}}^{'} & \hat{\boldsymbol{\beta}}_1\hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\beta}}_1 + \hat{\boldsymbol{\Sigma}} \end{bmatrix} \right), \quad i = 0, 1, 2, \ldots, n.$$

   ii. Compute $\hat{\boldsymbol{\mu}}^*_{wb}$, $\hat{\boldsymbol{\beta}}^*_{0b}$, $\hat{\boldsymbol{\beta}}^*_{1b}$, $\hat{\boldsymbol{\Psi}}^*_b$, $\hat{\boldsymbol{\Sigma}}^*_b$ using (3.4), but using the quantities $\begin{bmatrix} \boldsymbol{W}^*_{ib} \\ \boldsymbol{X}^*_{ib} \end{bmatrix}$, $i = 1, 2, \ldots, n$. Note that $\begin{bmatrix} \boldsymbol{W}^*_{0b} \\ \boldsymbol{X}^*_{0b} \end{bmatrix}$ is not included in the above computation.

   iii. Compute $\hat{\boldsymbol{\mu}}^*_{\boldsymbol{W}|\boldsymbol{X}_0,b}$ and $\hat{\boldsymbol{\Sigma}}^*_{\boldsymbol{W}|\boldsymbol{X},b}$ using (3.6) with $\hat{\boldsymbol{\mu}}^*_{wb}$, $\hat{\boldsymbol{\beta}}^*_{0b}$, $\hat{\boldsymbol{\beta}}^*_{1b}$, $\hat{\boldsymbol{\Psi}}^*_b$, and $\hat{\boldsymbol{\Sigma}}^*_b$ replacing the corresponding parameters and with $\boldsymbol{X}^*_{0b}$ replacing $\boldsymbol{X}_0$.

   iv. Compute $k_b = \max\limits_{1 \le j \le q} \left| \frac{\boldsymbol{W}^*_{0b(j)} - \hat{\boldsymbol{\mu}}^*_{\boldsymbol{W}|\boldsymbol{X}_0,b(j)}}{\sqrt{\left(\hat{\boldsymbol{\Sigma}}^*_{\boldsymbol{W}|\boldsymbol{X},b}\right)_{(jj)}}} \right|$.

3. Compute the $(1-\alpha)$-quantile of $k_1, \ldots, k_B$, call this k.

4. The $100\,(1-\alpha)$-prediction region for $\boldsymbol{W}_0$ is given by

$$\left[\hat{\boldsymbol{\mu}}_{\boldsymbol{W}|\boldsymbol{X}_0(1)} \pm k\sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{W}|\boldsymbol{X}}\right)_{(11)}}\right] \times \cdots \times \left[\hat{\boldsymbol{\mu}}_{\boldsymbol{W}|\boldsymbol{X}_0(q)} \pm k\sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{W}|\boldsymbol{X}}\right)_{(qq)}}\right]$$

   where $\hat{\boldsymbol{\mu}}_{\boldsymbol{W}|\boldsymbol{X}_0}$ and $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{W}|\boldsymbol{X}}$ are computed using the formulas in (3.6) with each parameter replaced by the corresponding estimate in (3.4).

As in the case of controlled calibration, Algorithm 2 can be modified to obtain one-sided prediction limits having only upper limits, or only lower limits. In order to get upper limits, we modify the definition of k_b in Step 2(iv) to be

$$k_b = \max_{1 \le j \le q} \frac{\boldsymbol{W}^*_{0b(j)} - \hat{\boldsymbol{\mu}}^*_{\boldsymbol{W}|\boldsymbol{X}_0,b(j)}}{\sqrt{\left(\hat{\boldsymbol{\Sigma}}^*_{\boldsymbol{W}|\boldsymbol{X},b}\right)_{(jj)}}},$$

and we take k to be the $(1-\alpha)$-quantile of $k_1, \ldots, k_B$. The resulting $100(1-\alpha)\%$ one-sided prediction region with upper limits only is given by

$$\left(-\infty, \hat{\boldsymbol{\mu}}_{\boldsymbol{W}|\boldsymbol{X}_0(1)} + k\sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{W}|\boldsymbol{X}}\right)_{(11)}}\right] \times \cdots \times \left(-\infty, \hat{\boldsymbol{\mu}}_{\boldsymbol{W}|\boldsymbol{X}_0(q)} + k\sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{W}|\boldsymbol{X}}\right)_{(qq)}}\right].$$

Similarly, in order to compute lower limits, we use the following expression in Step 2(iv) of Algorithm 2:

$$k_b = \min_{1 \le j \le q} \frac{\boldsymbol{W}^*_{0b(j)} - \hat{\boldsymbol{\mu}}^*_{\boldsymbol{W}|\boldsymbol{X}_0,b(j)}}{\sqrt{\left(\hat{\boldsymbol{\Sigma}}^*_{\boldsymbol{W}|\boldsymbol{X},b}\right)_{(jj)}}}.$$

and then we take k to be the α-quantile of $k_1, \ldots, k_B$. The resulting $100(1-\alpha)\%$ one-sided prediction region with lower limits is given by

$$\left[\hat{\boldsymbol{\mu}}_{\boldsymbol{W}|\boldsymbol{X}_0(1)} + k\sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{W}|\boldsymbol{X}}\right)_{(11)}}, \infty\right) \times \cdots \times \left[\hat{\boldsymbol{\mu}}_{\boldsymbol{W}|\boldsymbol{X}_0(q)} + k\sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{W}|\boldsymbol{X}}\right)_{(qq)}}, \infty\right).$$

Numerical results

We now report some numerical results on the coverage probabilities of the proposed prediction regions. We shall consider the sample sizes $n = 30$, 50, and 100, and a 95% nominal level. With regard to dimensions, we have made two choices: $p = 1$, $q = 1$ (the univariate case) and $p = 4$, $q = 2$. For the univariate case, we have chosen the parameter values $\boldsymbol{\mu}_w = 5$, $\boldsymbol{\beta}_0 = 1$, $\boldsymbol{\beta}_1 = 2$, $\boldsymbol{\Psi} = 0.5$, $\boldsymbol{\Sigma} = 0.5$, and for the multivariate case we have chosen

$$\boldsymbol{\mu}_w = \begin{bmatrix} 5 \\ 6 \end{bmatrix}, \quad \boldsymbol{\beta}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\beta}_1 = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}, \quad \boldsymbol{\Psi} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \boldsymbol{\Sigma} = 0.5\boldsymbol{I}_4 + 0.5\mathbf{1}_4\mathbf{1}_4'.$$

The estimated coverage probabilities for the two-sided rectangular prediction region are shown in Table 4. Also included are the expected values of the factor k and the expected volumes of the prediction region. We see that at least for the scenarios considered in the simulations, the coverage probabilities are quite satisfactory even for the small sample size situation. We also note that the expected values of the factor k, and the expected volumes, are decreasing with the sample size, as expected.

Application: analysis of wheat quality data

The example that we now use to compute a random calibration region is the wheat quality data from Brown (1982, Table 1). The response vector consists of four (derived) infrared reflectance response

Table 4: Estimated coverage probabilities along with the expected values of the factor k and the expected volumes of rectangular two-sided prediction regions under random calibration for the parameter choices in Section 3; nominal level = 95%

		Coverage	$E(k)$	E(Volume)
$p = 1,\ q = 1$	$n = 100$	0.9472	1.9962	1.2576
	$n = 50$	0.9484	2.0497	1.2840
	$n = 30$	0.9482	2.1185	1.3196
$p = 4,\ q = 2$	$n = 100$	0.9458	2.3033	10.4481
	$n = 50$	0.9460	2.4501	11.3783
	$n = 30$	0.9534	2.6866	12.8865

variables, and the covariates consist of percent water and percent protein. In his paper, Brown (1982) notes that the covariate vector, namely $\boldsymbol{W} = (\%\text{ water}, \%\text{ protein})'$ is a random variable. Furthermore, 21 observations are given in Brown (1982, Table 1). The first 16 observations are used to build the statistical model while the last 5 are used for calibration, as done in Brown (1982). Our goal is to compute a 95% rectangular prediction regions for $\boldsymbol{W} = (\%\text{ water}, \%\text{ protein})'$.

The parameter estimates obtained from the first 16 observations in Brown (1982, Table 1) are as follows:

$$\hat{\boldsymbol{\mu}}_w = \begin{bmatrix} 9.5424 \\ 11.2605 \end{bmatrix} \quad \hat{\boldsymbol{\beta}}_0 = \begin{bmatrix} 389.2855 \\ 159.4307 \\ 334.4639 \\ 457.0545 \end{bmatrix}, \quad \hat{\boldsymbol{\beta}}_1 = \begin{bmatrix} 0.4488 & -3.0699 \\ -0.3728 & -4.5855 \\ -24.2729 & -1.8848 \\ -23.8768 & 0.1460 \end{bmatrix},$$

$$\hat{\boldsymbol{\Psi}} = \begin{bmatrix} 0.2255 & 0.1820 \\ 0.1820 & 2.1028 \end{bmatrix}, \quad \hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 10.1550 & 4.3372 & 0.9803 & 4.1743 \\ 4.3372 & 2.2465 & 0.6628 & 1.4670 \\ 0.9803 & 0.6628 & 1.9769 & 1.4422 \\ 4.1743 & 1.4670 & 1.4422 & 3.1187 \end{bmatrix}.$$

Table 5 gives the prediction regions obtained using Algorithm 2, corresponding to the five response vectors given in the table. The rectangular regions that we have obtained all contain their respective true values of the covariate vector and are reasonably narrow, especially for % water.

4 Discussion

Our work is the first attempt at deriving rectangular confidence regions and prediction regions in the multivariate calibration problem. As already noted, such regions are necessary if we want to draw inference concerning the individual components of a parameter vector while taking into account the multivariate nature of the data; i.e., take into account the correlations. The Bonferroni inequality is often used to

Table 5: Random calibration prediction regions for % water and % protein at five different settings of infrared reflectance

Response point	Covariate	Prediction region
$(363, 113, 88, 231)'$	Water	9.37 to 9.60
	Protein	8.85 to 9.64
$(363, 110, 101, 248)'$	Water	8.72 to 8.94
	Protein	9.79 to 10.57
$(366, 114, 79, 224)'$	Water	9.72 to 9.95
	Protein	8.75 to 9.54
$(350, 96, 85, 235)'$	Water	9.17 to 9.40
	Protein	12.32 to 13.11
$(355, 97, 63, 216)'$	Water	10.08 to 10.31
	Protein	12.38 to 13.16

derive such simultaneous confidence intervals and prediction intervals, resulting in a solution that is known to be conservative. However, our parametric bootstrap approach has resulted in solutions that appear to be quite satisfactory in terms of maintaining the coverage probability. In addition, our calibration region is guaranteed to be rectangular, by construction. We are emphasizing this since the calibration regions available in the literature, though meant to be ellipsoidal, are not necessarily so.

There are two other scenarios where rectangular confidence and prediction regions can be of interest in multivariate calibration, but not addressed in the present article. One of them is the setup where the covariates enter the model non-linearly (for example, polynomial regression). Ellipsoidal regions in this case are derived in Oman (1988) and in Mathew and Zha (1996). Yet another scenario calls for the computation of *multiple use* confidence regions in controlled calibration for unknown values of the covariate vector. That is, the same data set is to be used to compute confidence regions for an unspecified sequence of values of the covariate vector, corresponding to a sequence of response vectors; see Mathew and Zha (1996), Mathew, Sharma and Nordström (1998), and Section 10.6 in Krishnamoorthy and Mathew (2009) for details and additional references. The confidence regions derived by these authors are ellipsoidal, and rectangular multiple use confidence regions are certainly of interest. We hope to address this in a future communication.

Acknowledgements The authors are grateful to Dr. Dulal Bhaumik for providing the data for the second example in Section 2.

Declarations

The authors have no relevant financial or non-financial interests to disclose.

References

Atkins, M. S., Shernoff, E. S., Frazier, S. L., Schoenwald, S. K., Cappella, E., Marinez-Lora, A., Mehta, T. G., Lakind, D., Cua, G., Bhaumik, R., & Bhaumik, D. (2015). Redesigning community mental health services for urban children: Supporting schooling to promote mental health, *Journal of Consulting and Clinical Psychology*, 83, 839–852.

Bellio, R. (2003). Likelihood methods for controlled calibration, *Scandinavian Journal of Statistics*, 30, 339-353.

Benton, D., Krishnamoorthy K., & Mathew, T. (2003). Inferences in multivariate-univariate calibration problems, *The Statistician*, 52, 15-39. Brown, P. J. (1982). Multivariate calibration, *Journal of the Royal Statistical Society, Series B*, 44, 287-321.

Bhaumik, D. K., & Nordgren, R. (2019). Prediction and calibration for multiple correlated variables, *Journal of Multivariate Analysis*, 173, 313–327.

Brown, P. J. (1982). Multivariate calibration, *Journal of the Royal Statistical Society, Series B*, 44, 287-321.

Brown, P. J. (1993). *Measurement, Regression, and Calibration*. Oxford University Press.

Brown, P. J., & Sundberg, R. (1987). Confidence and conflict in multivariate calibration, *Journal of the Royal Statistical Society, Series B*, 49, 46-57.

Davis, A., & Hayakawa, T. (1987). Some distribution theory relating to confidence regions in multivariate calibration, *Annals of the Institute of Statistical Mathematics*, 39, 141-152.

Devijver, E., & Perthame, E. (2018). Prediction Regions through Inverse Regression. Retrieved from: https://arxiv.org/abs/1807.03184v1. Retrieved March 31, 2020.

Fujikoshi, Y., & Nishii, R. (1984). On the distributiono of a statistic in multivariate inverse regression analysis, *Hiroshima Mathematical Journal*, 14, 215-225.

Gleser, L. J., & Hwang, J. T. (1987). The nonexistence of $100(1-\alpha)\%$ confidence sets of finite expected diameter in errors-in-variables and related models, *The Annals of Statistics*, 15, 1351-1362.

Krishnamoorthy, K., & Mathew, T. (2009). *Statistical Tolerance Regions: Theory, Applications and Computation*. John Wiley & Sons.

Lütkepohl, H., Staszewska-Bystrova, A., & Winker, P. (2020). Constructing joint confidence bands for impulse response functions of VAR models - A review, *Econometrics and Statistics*, 13, 69-83.

Mathew, T., & Kasala, S. (1994). An exact confidence region in multivariate calibration, *The Annals of Statistics*, 22, 94-105.

Mathew, T., Sharma, M. K., & Nordström, K. (1998). Tolerance regions and multiple use confidence regions in multivariatec alibration, *The Annals of Statistics*, 1989-2013.

Mathew, T., & Zha, W. (1996a). Conservative confidence regions in multivariate calibration, *The Annals of Statistics*, 24, 707-725.

Mathew, T., & Zha, W. (1996b). Multiple use confidence regions in multivariate calibration, *Journal of the American Statistical Association*, 92, 1141-1150.

Montiel Olea, J., Luis, J., & Plagborg-Moller, M. (2019). Simultaneous confidence bands: theory, implementation, and an application to SVARs, *Journal of Applied Econometrics*, 34, 1-17.

Oman, S. D. (1988). Confidence regions in multivariate calibration, *The Annals of Statistics*, 16, 174-187.

Osborne, C. (1991). Statistical calibration: a review, *International Statistical Review*, 59, 309-336.

Sundberg, R. (1994). Most modern calibration is multivariate. In 17th International Biometric Conference, Hamilton, Canada. pp. 395-405.

Sundberg, R. (1999). Multivariate calibration - direct and indirect regression methodology (with discussion), *Scandinavian Journal of Statistics*, 26, 161-207.

Wolf, M., & Wunderli, D. (2015). Bootstrap joint prediction regions. *Journal of Time Series Analysis*, 36, 352-376.