Journal of Multivariate Analysis 102 (2011) 386–392

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Notes

A multivariate ultrastructural errors-in-variables model with equation error

Alexandre G. Patriota [a,*], Heleno Bolfarine [a], Reinaldo B. Arellano-Valle [b]

[a] Departamento de Estatística, Universidade de São Paulo, Rua do Matão, 1010, São Paulo/SP, 05508-090, Brazil
[b] Departamento de Estadística, Pontificia Universidad Católica de Chile, Santiago, Chile

ARTICLE INFO

Article history:
Received 4 August 2009
Available online 15 September 2010

Keywords:
Asymptotic property
Consistent estimator
Measurement error
Multivariate regression

ABSTRACT

This paper deals with asymptotic results on a multivariate ultrastructural errors-in-variables regression model with equation errors. Sufficient conditions for attaining consistent estimators for model parameters are presented. Asymptotic distributions for the line regression estimators are derived. Applications to the elliptical class of distributions with two error assumptions are presented. The model generalizes previous results aimed at univariate scenarios.

1. Introduction

It is well known that maximum-likelihood estimators (MLEs) and ordinary least squares estimators are generally inconsistent if covariates are measured with errors and are intentionally ignored. More specifically, the estimation of the slope parameter of a linear model is attenuated [9] by the presence of measurement errors. It is then expected that the same should occur in multivariate contexts. When variables are subject to measurement errors, we must add to the model appropriate measurement equations in order to capture the measurement error effects. This procedure, when feasible, produces consistent, efficient and asymptotically normally distributed estimators. A careful and rigorous exposition of the inferential process in simple and multiple errors-in-variables models can be seen in [9] and the references therein.

It is however our impression that multivariate ultrastructural errors-in-variables (or measurement errors) regression models have not been as extensively studied in the statistical literature as the multiple and simple regression models. It seems that the majority of the references consider separately the structural and functional versions. For instance, multivariate functional models were studied in [11,10,4,6], while the structural version was studied from a factor analysis point of view [19]. Anemiya and Fuller [1] study both versions of this multivariate model without equation errors.

In this paper we consider a multivariate ultrastructural errors-in-variables regression model with equation errors which, to the best of our knowledge, was not previously exploited in the statistical literature. We present consistent estimators having asymptotic normal distributions. These results are basically attained by using moment estimators and delta methods; see, for instance, [15] and [13, Ch. 3]. The main motivation for considering equation errors comes from epidemiological [14] and astrophysical [12] studies, where equation errors are justified by the influence of factors other than those specified in the model for explaining the variation of the response variable (see also [17]). Multivariate analyses with measurement errors and equation errors are listed as one topic of great methodological challenge for the coming decade in astrophysics, as can be seen on the web page http://nedwww.ipac.caltech.edu/level5/Sept03/Feigelson/Feigelson5.html. Here, we present a multivariate regression model with homoscedastic measurement errors and equation errors. Although astrophysical data

* Corresponding author.
E-mail addresses: patriota.alexandre@gmail.com, patriota@ime.usp.br (A.G. Patriota).

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.08.013

sets typically contain heteroscedastic errors and censoring, we believe that the results of this paper can be used for further generalizations in studying those challenges. According to [8], measurement errors may be among the reasons for high biological false positive rates identified in actual regulatory network models. The authors used a multivariate structural errors-in-variables model with equation errors considering normality and homoscedasticity for modeling gene regulatory networks [20]. Our results generalize this study by relaxing the normality supposition; therefore, the results may also be applied for gene expression data and other regulatory network models. Our paper is intended to be of a theoretical nature and we expect to report on real data applications in future research.

The model proposed in this paper can be seen as a generalization of the multiple regression model considered in [18,9], Section 2.2, and it is also a generalization of the multivariate structural version applied in [8]. The model is defined in its full generality by the following three stochastic equations:

$$\boldsymbol{y}_i = \boldsymbol{a} + \boldsymbol{B}\boldsymbol{x}_i + \boldsymbol{q}_i, \tag{1}$$

$$\boldsymbol{Y}_i = \boldsymbol{y}_i + \boldsymbol{e}_i, \tag{2}$$

$$\boldsymbol{X}_i = \boldsymbol{x}_i + \boldsymbol{u}_i, \tag{3}$$

$i = 1, \ldots, n$. We have in Eq. (1) a multivariate regression model, where $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ and $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are the true response and covariate vectors of dimension $p_1 \times 1$ and $p_2 \times 1$, respectively, $\boldsymbol{a}$ is a $p_1 \times 1$ vector of model intercepts, $\boldsymbol{B}$ is a $p_1 \times p_2$ matrix of slope parameters, and $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ are random vectors of dimension $p_1 \times 1$ representing the equation errors. Moreover, as indicated by Eqs. (2) and (3), $\boldsymbol{Y}_i$ and $\boldsymbol{X}_i$, $i = 1, \ldots, n$, are measurements of the true (unobservable) vectors $\boldsymbol{y}_i$ and $\boldsymbol{x}_i$, $i = 1, \ldots, n$, respectively, where the respective measurement errors are represented by $\boldsymbol{e}_i$ and $\boldsymbol{u}_i$, $i = 1, \ldots, n$. Hence, in order to make inferences on unknown parameters, it is common to suppose that the full vectors $\boldsymbol{r}_i = (\boldsymbol{q}_i^\top, \boldsymbol{e}_i^\top, \boldsymbol{u}_i^\top, (\boldsymbol{x}_i - \boldsymbol{\xi}_i)^\top)^\top$, $i = 1, \ldots, n$, where $\boldsymbol{\xi}_i = E(\boldsymbol{x}_i)$, are independent, and identically and symmetrically distributed (typically assumed to be normally distributed), with

$$E(\boldsymbol{r}_1) = \boldsymbol{0} \quad \text{and} \quad \mathrm{Var}(\boldsymbol{r}_1) = \boldsymbol{\Sigma}_{\mathbf{r}} = \mathrm{diag}(\boldsymbol{\Sigma}_q, \boldsymbol{\Sigma}_e, \boldsymbol{\Sigma}_u, \boldsymbol{\Sigma}_x). \tag{4}$$

Consequently, if we denote the observable vectors by $\boldsymbol{Z}_i = (\boldsymbol{Y}_i^\top, \boldsymbol{X}_i^\top)^\top$, $i = 1, \ldots, n$, we have from the above assumptions that they are independent with mean vectors $\boldsymbol{\mu}_i,\ i = 1, \ldots, n$, and common covariance matrix $\boldsymbol{\Sigma}$ given by

$$\boldsymbol{\mu}_i = \begin{pmatrix} \boldsymbol{a} + \boldsymbol{B}\boldsymbol{\xi}_i \\ \boldsymbol{\xi}_i \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{B}\boldsymbol{\Sigma}_x\boldsymbol{B}^\top + \boldsymbol{\Sigma}_q + \boldsymbol{\Sigma}_e & \boldsymbol{B}\boldsymbol{\Sigma}_x \\ \boldsymbol{\Sigma}_x\boldsymbol{B}^\top & \boldsymbol{\Sigma}_x + \boldsymbol{\Sigma}_u \end{pmatrix}. \tag{5}$$

Therefore, if $\boldsymbol{\Sigma}_x = \boldsymbol{0}$ (where $\boldsymbol{0}$ denotes the zero matrix with appropriate dimensions) then the ultrastructural model (1)–(3) reduces to the functional model. Otherwise, if $\boldsymbol{\xi}_1 = \cdots = \boldsymbol{\xi}_n = \boldsymbol{\xi}_x$, then the ultrastructural model becomes the classical multivariate structural model. Gleser [10] considered a functional multivariate model with $\boldsymbol{\Sigma}_q = \boldsymbol{0}$ and the measurement error vector $(\boldsymbol{e}_i^\top, \boldsymbol{u}_i^\top)^\top$ having null mean vector and covariance matrix of the form $\sigma^2\boldsymbol{\Sigma}_0$, where σ^2 is a parameter to be estimated and $\boldsymbol{\Sigma}_0$ is a known matrix. Anemiya and Fuller [1] consider this model in an ultrastructural version with the covariance matrix of the measurement error replaced by an estimator. As can be seen, the model considered by these authors is not nested with our proposal.

It is clear from (5) that the model (1)–(3) considering (4) is not identifiable when, e.g., the normal assumption is considered. It is only possible to estimate the covariance matrices $\mathrm{Var}(\boldsymbol{q}_i + \boldsymbol{e}_i) = \boldsymbol{\Sigma}_q + \boldsymbol{\Sigma}_e$ and $\mathrm{Var}(\boldsymbol{x}_i + \boldsymbol{u}_i) = \boldsymbol{\Sigma}_x + \boldsymbol{\Sigma}_u$ rather than each of those components separately. However, if the matrices $\boldsymbol{\Sigma}_e$ and $\boldsymbol{\Sigma}_u$ are assumed to be known, the model becomes identifiable. Knowledge of these covariance matrices will then be one of the assumptions considered. In addition, to derive the main results of this paper, we assume that $\boldsymbol{\Sigma}_q$ is a positive definite matrix and that the following assumptions hold:

(A1) there exists a $p_2 \times 1$ vector $\boldsymbol{\xi}$ and a $p_2 \times p_2$ matrix $\boldsymbol{\Sigma}_\xi$ such that $\boldsymbol{\Sigma}_x + \boldsymbol{\Sigma}_\xi$ is positive definite,

$$\sqrt{n}\,(\bar{\boldsymbol{\xi}} - \boldsymbol{\xi}) \to \boldsymbol{0} \quad \text{and} \quad \sqrt{n}\,(\boldsymbol{S}_\xi - \boldsymbol{\Sigma}_\xi) \to \boldsymbol{0},$$

as $n \to \infty$, where $\bar{\boldsymbol{\xi}} = n^{-1}\sum_{k=1}^n \boldsymbol{\xi}_k$ and $\boldsymbol{S}_\xi = n^{-1}\sum_{k=1}^n (\boldsymbol{\xi}_k - \bar{\boldsymbol{\xi}})(\boldsymbol{\xi}_k - \bar{\boldsymbol{\xi}})^\top$;

(A2) $\mathrm{Var}[\mathrm{vech}(\boldsymbol{r}_1\boldsymbol{r}_1^\top)] < \infty$.

Here, for any $p \times q$ matrix $\boldsymbol{C}$ with columns $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_q$, $\mathrm{vec}(\boldsymbol{C})$ is the $(pq) \times 1$ vector defined by $\mathrm{vec}(\boldsymbol{C}) = (\boldsymbol{c}_1^\top, \ldots, \boldsymbol{c}_q^\top)^\top$. Moreover, when $p = q$ and $\boldsymbol{C} = \boldsymbol{C}^\top$, $\mathrm{vech}(\boldsymbol{C})$ denotes the $[p(p+1)/2] \times 1$ vector that contains all the diagonal and (different) subdiagonal elements of $\boldsymbol{C}$. Furthermore, in this last case, there exists a $p^2 \times p(p+1)/2$ duplication matrix $\boldsymbol{D}$ such that $\mathrm{vec}(\boldsymbol{C}) = \boldsymbol{D}\mathrm{vech}(\boldsymbol{C})$, or $\mathrm{vech}(\boldsymbol{C}) = \boldsymbol{D}^+\mathrm{vec}(\boldsymbol{C})$, where $\boldsymbol{D}^+ = (\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top$ is the Moore–Penrose inverse of $\boldsymbol{D}$. We also use the notation $\mathcal{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to represent a d-variate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

Conditions (A1)–(A2) are needed to guarantee convergence in probability and distribution of the estimators and the existence of their limiting covariance matrix. They are conditions usually stated in the ultrastructural measurement error literature, as can be seen, for example, in [5,9,3,2].

This paper is organized as follows. Section 2 establishes the main results of the present article. It presents consistent estimators and, moreover, the joint asymptotic distribution of the line parameter estimators. Section 3 applies the results to the elliptical class of distributions which specializes to results previously derived in the literature.

2. Main results

In this section, the main concern is to develop consistent estimators for the parameters of the novel model defined by (1)–(4). Furthermore, it is also of interest to study their asymptotic distribution. As mentioned in the Introduction, the inconsistency of the maximum-likelihood estimators for the univariate functional case without equation error is well known (see, for instance, [9,5]). However, alternative consistent estimators can be proposed. Schneeweiss [18] obtains the limiting distribution of consistent estimators in a multiple regression model considering the measurement error covariance matrix of the independent variables to be known. The author makes no use of normality assumptions for the error terms. Thus, following this line of thought, Proposition 1 presents consistent estimators for the model parameters in the multivariate context. On the other hand, Proposition 2 establishes the asymptotic distribution for the estimator of $\boldsymbol{\theta} = (\boldsymbol{a}^\top, \mathrm{vec}(\boldsymbol{B})^\top)^\top$. These results are based on the asymptotic behavior of the sample mean vector and covariance matrix of the observable vectors $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$, which are given by

$$\bar{\boldsymbol{Z}} = \begin{pmatrix} \bar{\boldsymbol{Y}} \\ \bar{\boldsymbol{X}} \end{pmatrix} \quad \text{and} \quad \boldsymbol{S_Z} = \begin{pmatrix} \boldsymbol{S_Y} & \boldsymbol{S_{YX}} \\ \boldsymbol{S_{XY}} & \boldsymbol{S_X} \end{pmatrix},$$

where $\bar{\boldsymbol{X}} = n^{-1}\sum_{i=1}^n \boldsymbol{X}_i$, $\bar{\boldsymbol{Y}} = n^{-1}\sum_{i=1}^n \boldsymbol{Y}_i$, $\boldsymbol{S_X} = n^{-1}\sum_{i=1}^n(\boldsymbol{X}_i - \bar{\boldsymbol{X}})\boldsymbol{X}_i^\top$, $\boldsymbol{S_Y} = n^{-1}\sum_{i=1}^n(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})\boldsymbol{Y}_i^\top$ and $\boldsymbol{S_{XY}} = n^{-1}\sum_{i=1}^n (\boldsymbol{X}_i - \bar{\boldsymbol{X}})\boldsymbol{Y}_i^\top$. The asymptotic properties of $\bar{\boldsymbol{Z}}$ and $\boldsymbol{S_Z}$ are established next in Lemma 1, but they appear in [3]. For this, we note first that the assumptions considered in Section 1 imply that the random vectors $\boldsymbol{\varepsilon}_i = \boldsymbol{Z}_i - \boldsymbol{\mu}_i$, $i = 1, \ldots, n$, are independent and identically distributed (i.i.d.) for all $n \geq 1$, and that they have mean vector $\boldsymbol{0}$ (the null vector) and covariance matrix $\boldsymbol{\Sigma}$, where $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}$ are given in (5). Also, by (A1) we have that $\bar{\boldsymbol{\xi}} \to \boldsymbol{\xi}$ and $\boldsymbol{S}_\xi \to \boldsymbol{\Sigma}_\xi$ as $n \to \infty$. Hence, for $\bar{\boldsymbol{\mu}} = n^{-1}\sum_{i=1}^n \boldsymbol{\mu}_i$ and $\boldsymbol{S}_\mu = n^{-1}\sum_{i=1}^n(\boldsymbol{\mu}_i - \bar{\boldsymbol{\mu}})(\boldsymbol{\mu}_i - \bar{\boldsymbol{\mu}})^\top$ we obtain, as $n \to \infty$, that

$$\bar{\boldsymbol{\mu}} \to \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{a} + \boldsymbol{B\xi} \\ \boldsymbol{\xi} \end{pmatrix} \quad \text{and} \quad \boldsymbol{S}_\mu \to \boldsymbol{\Sigma}_\mu = \begin{pmatrix} \boldsymbol{B\Sigma}_\xi\boldsymbol{B}^\top & \boldsymbol{B\Sigma}_\xi \\ \boldsymbol{\Sigma}_\xi\boldsymbol{B}^\top & \boldsymbol{\Sigma}_\xi \end{pmatrix}. \tag{6}$$

Moreover, $\boldsymbol{\varepsilon}_1 = \boldsymbol{Z}_1 - \boldsymbol{\mu}_1 = \boldsymbol{Ar}_1$, where

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I}_{p_1} & \boldsymbol{I}_{p_1} & \boldsymbol{0} & \boldsymbol{B} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_{p_2} & \boldsymbol{I}_{p_2} \end{pmatrix}, \tag{7}$$

where $\boldsymbol{I}_p$ is the $(p \times p)$ identity matrix. Under the assumption that the distribution of $\boldsymbol{r}_1$ is symmetric we have that $\mathrm{E}[\boldsymbol{\varepsilon}_1 \mathrm{vech}(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^\top)^\top] = \boldsymbol{0}$ and $\boldsymbol{\Lambda} = \mathrm{Var}[\mathrm{vech}(\boldsymbol{\varepsilon}_1\boldsymbol{\varepsilon}_1^\top)] = \mathrm{Var}[\mathrm{vech}(\boldsymbol{Ar}_1\boldsymbol{r}_1^\top\boldsymbol{A}^\top)] = \boldsymbol{D}^+(\boldsymbol{A}\otimes\boldsymbol{A})\boldsymbol{\Lambda_r}(\boldsymbol{A}\otimes\boldsymbol{A})^\top\boldsymbol{D}^{+\top} < \infty$ since by (A2) $\boldsymbol{\Lambda_r} = \mathrm{Var}[\mathrm{vec}(\boldsymbol{r}_1\boldsymbol{r}_1^\top)] < \infty$, where $\boldsymbol{D}^+ = (\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top$, with $\boldsymbol{D}$ being the $[d^2 \times d(d+1)/2]$ corresponding duplication matrix, $d = p_1 + p_2$ and the symbol $\otimes$ indicates the Kronecker product. Thus, the results presented next follow from Theorem 2.1 in [3].

Lemma 1. *Let $\boldsymbol{Z}_i = (\boldsymbol{Y}_i^\top, \boldsymbol{X}_i^\top)^\top$, $i = 1, \ldots, n$, be the observable random vector for the model (1)–(3), so that, by assumption, the random vectors $\boldsymbol{Z}_i - \boldsymbol{\mu}_i$, $i = 1, \ldots, n$, are independent, and identically and symmetrically distributed, with mean vector $\boldsymbol{0}$ and covariance matrix $\boldsymbol{\Sigma}$, where $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}$ are given in (5). Then:*

(i) *under* (A1) *it follows that, as $n \to \infty$,*

$$\bar{\boldsymbol{Z}} = \begin{pmatrix} \bar{\boldsymbol{Y}} \\ \bar{\boldsymbol{X}} \end{pmatrix} \xrightarrow{\text{a.s.}} \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{a} + \boldsymbol{B\xi} \\ \boldsymbol{\xi} \end{pmatrix}$$

and

$$\boldsymbol{S_Z} = \begin{pmatrix} \boldsymbol{S_Y} & \boldsymbol{S_{YX}} \\ \boldsymbol{S_{XY}} & \boldsymbol{S_X} \end{pmatrix} \xrightarrow{\text{a.s.}} \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_\mu = \begin{pmatrix} \boldsymbol{B\Gamma}_{x+\xi}\boldsymbol{B}^\top + \boldsymbol{\Sigma}_q + \boldsymbol{\Sigma}_e & \boldsymbol{B\Gamma}_{x+\xi} \\ \boldsymbol{\Gamma}_{x+\xi}\boldsymbol{B}^\top & \boldsymbol{\Gamma}_{x+\xi} + \boldsymbol{\Sigma}_u \end{pmatrix}$$

where $\boldsymbol{\Gamma}_{x+\xi} = \boldsymbol{\Sigma}_x + \boldsymbol{\Sigma}_\xi$;

(ii) *under* (A1)–(A2) *it follows that, as $n \to \infty$,*

$$\sqrt{n}(\bar{\boldsymbol{Z}} - \boldsymbol{\mu}) \xrightarrow{d} \mathcal{N}_d(\boldsymbol{0}, \boldsymbol{\Sigma})$$

and

$$\sqrt{n}\,\mathrm{vech}(\boldsymbol{S_Z} - \boldsymbol{\Sigma} - \boldsymbol{\Sigma}_\mu) \xrightarrow{d} \mathcal{N}_{d(d+1)/2}\left(\boldsymbol{0}, \boldsymbol{\Lambda} + \boldsymbol{\Lambda_\mu}\right)$$

and they are asymptotically independent, where

$$\boldsymbol{\Lambda} = \boldsymbol{D}^+\{(\boldsymbol{A}\otimes\boldsymbol{A})\boldsymbol{\Lambda_r}(\boldsymbol{A}\otimes\boldsymbol{A})^\top\}\boldsymbol{D}^{+\top} \quad \text{and} \quad \boldsymbol{\Lambda_\mu} = 4\boldsymbol{D}^+(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}_\mu)\boldsymbol{D}^{+\top}, \tag{8}$$

where the matrix $\boldsymbol{A}$ is defined in (7).

Remark. The version of this lemma in [3] is proved under the assumption that $\boldsymbol{\Sigma}_\xi$ is positive definite. However, the lemma is still valid under the less restrictive condition that $\boldsymbol{\Sigma}_x + \boldsymbol{\Sigma}_\xi$ is positive definite, as can be seen in [10], where this last condition is considered.

Notice that the above asymptotic distributions depend on the distribution of $\boldsymbol{r}_i$ only through the quantities $\boldsymbol{\Sigma}_{\boldsymbol{r}}$ (see Eq. (4)) and $\boldsymbol{\Lambda}_{\boldsymbol{r}} = \mathrm{Var}[\mathrm{vec}(\boldsymbol{r}_1\boldsymbol{r}_1^\top)]$. In the next section we give some examples related to the class of the elliptical distributions. We present next the main results of the paper.

Proposition 1. *Consider model (1)–(3) with the assumption (4). Suppose also that the covariance matrices $\boldsymbol{\Sigma}_e$ and $\boldsymbol{\Sigma}_u$ are known and that condition* (A1) *holds. Then, parameters $\boldsymbol{a}$, $\boldsymbol{B}$, $\boldsymbol{\xi}$, $\boldsymbol{\Sigma}_q$ and $\boldsymbol{\Gamma}_{x+\xi}$ are consistently estimated by*

$$\begin{aligned}
&\widehat{\boldsymbol{a}} = \bar{\boldsymbol{Y}} - \widehat{\boldsymbol{B}}\bar{\boldsymbol{X}},\\
&\widehat{\boldsymbol{B}} = \boldsymbol{S}_{\boldsymbol{YX}}(\boldsymbol{S}_{\boldsymbol{X}} - \boldsymbol{\Sigma}_u)^{-1},\\
&\widehat{\boldsymbol{\xi}} = \bar{\boldsymbol{X}},\\
&\widehat{\boldsymbol{\Sigma}}_q = \boldsymbol{S}_{\boldsymbol{Y}} - \boldsymbol{S}_{\boldsymbol{YX}}(\boldsymbol{S}_{\boldsymbol{X}} - \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{S}_{\boldsymbol{XY}} - \boldsymbol{\Sigma}_e,\\
&\widehat{\boldsymbol{\Gamma}}_{x+\xi} = \boldsymbol{S}_{\boldsymbol{X}} - \boldsymbol{\Sigma}_u.
\end{aligned}$$

Considering part (i) of Lemma 1, the proof of Proposition 1 is straightforward, since all of these estimators are continuous functions of the sample mean vector $\bar{\boldsymbol{Z}}$ and of the sample covariance matrix $\boldsymbol{S}_{\boldsymbol{Z}}$.

Proposition 2. *In addition, if we consider condition* (A2) *in Proposition 1, then the asymptotic distribution of $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{a}}^\top, \mathrm{vec}(\widehat{\boldsymbol{B}})^\top)^\top$ is given by*

$$\begin{pmatrix} \sqrt{n}(\widehat{\boldsymbol{a}} - \boldsymbol{a}) \\ \sqrt{n}\mathrm{vec}(\widehat{\boldsymbol{B}} - \boldsymbol{B}) \end{pmatrix} \xrightarrow{d} \mathcal{N}_{p_1+p_1p_2}\left(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Phi}_{\boldsymbol{a}} & \boldsymbol{\Phi}_{\boldsymbol{aB}} \\ \boldsymbol{\Phi}_{\boldsymbol{Ba}} & \boldsymbol{\Phi}_{\boldsymbol{B}} \end{pmatrix} \right),$$

where

$$\begin{aligned}
&\boldsymbol{\Phi}_{\boldsymbol{a}} = \boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}^\top + (\boldsymbol{\xi}^\top \otimes \boldsymbol{I}_{p_1})\boldsymbol{\Phi}_{\boldsymbol{B}}(\boldsymbol{\xi} \otimes \boldsymbol{I}_{p_1}),\\
&\boldsymbol{\Phi}_{\boldsymbol{aB}} = \boldsymbol{\Phi}_{\boldsymbol{Ba}}^\top = -(\boldsymbol{\xi}^\top \otimes \boldsymbol{I}_{p_1})\boldsymbol{\Phi}_{\boldsymbol{B}},\\
&\boldsymbol{\Phi}_{\boldsymbol{B}} = \boldsymbol{QD}(\boldsymbol{\Lambda} + \boldsymbol{\Lambda}_\mu)\boldsymbol{D}^\top\boldsymbol{Q}^\top,
\end{aligned}$$

with

$$\boldsymbol{P} = (\boldsymbol{I}_{p_1}, -\boldsymbol{B}), \qquad \boldsymbol{Q} = (\boldsymbol{\Gamma}_{x+\xi}^{-1} \otimes \boldsymbol{I}_{p_1})(\boldsymbol{H}_2 \otimes \boldsymbol{H}_1) - (\boldsymbol{\Gamma}_{x+\xi}^{-1} \otimes \boldsymbol{B})(\boldsymbol{H}_2 \otimes \boldsymbol{H}_2),$$

$\boldsymbol{\Lambda} + \boldsymbol{\Lambda}_\mu$ defined in (8), $\boldsymbol{H}_1 = (\boldsymbol{I}_{p_1}, \boldsymbol{0})$ and $\boldsymbol{H}_2 = (\boldsymbol{0}, \boldsymbol{I}_{p_2})$ matrices such that $\boldsymbol{S}_{\boldsymbol{YX}} = \boldsymbol{H}_1\boldsymbol{S}_{\boldsymbol{Z}}\boldsymbol{H}_2^\top$ and $\boldsymbol{S}_{\boldsymbol{X}} = \boldsymbol{H}_2\boldsymbol{S}_{\boldsymbol{Z}}\boldsymbol{H}_2^\top$.

Proof. First, after some algebra, we can write

$$\widehat{\boldsymbol{a}} - \boldsymbol{a} = \widehat{\boldsymbol{P}}(\bar{\boldsymbol{Z}} - \boldsymbol{\mu}) - (\boldsymbol{\xi}^\top \otimes \boldsymbol{I}_{p_1})\mathrm{vec}(\widehat{\boldsymbol{B}} - \boldsymbol{B}). \tag{9}$$

Moreover, since $\boldsymbol{S}_{\boldsymbol{YX}} = \boldsymbol{H}_1\boldsymbol{S}_{\boldsymbol{Z}}\boldsymbol{H}_2^\top$ and $\boldsymbol{S}_{\boldsymbol{X}} = \boldsymbol{H}_2\boldsymbol{S}_{\boldsymbol{Z}}\boldsymbol{H}_2^\top$, where $\boldsymbol{H}_1 = (\boldsymbol{I}_{p_1}, \boldsymbol{0})$ and $\boldsymbol{H}_2 = (\boldsymbol{0}, \boldsymbol{I}_{p_2})$, we have that $\widehat{\boldsymbol{B}} = \widehat{\boldsymbol{B}}(\boldsymbol{S}_{\boldsymbol{Z}}) = \boldsymbol{H}_1\boldsymbol{S}_{\boldsymbol{Z}}\boldsymbol{H}_2^\top(\boldsymbol{H}_2\boldsymbol{S}_{\boldsymbol{Z}}\boldsymbol{H}_2^\top - \boldsymbol{\Sigma}_u)^{-1}$ is a continuous function of the sample covariance matrix $\boldsymbol{S}_{\boldsymbol{Z}}$. We define a_k as the kth element of $\mathrm{vech}(\boldsymbol{S}_{\boldsymbol{Z}})$. Hence, by applying the matrix derivatives methodology (see, e.g., [13,16]) we have that

$$\frac{\partial \widehat{\boldsymbol{B}}}{\partial a_k} = \boldsymbol{H}_1\left(\frac{\partial \boldsymbol{S}_{\boldsymbol{Z}}}{\partial a_k}\right)\boldsymbol{H}_2^\top(\boldsymbol{S}_{\boldsymbol{X}} - \boldsymbol{\Sigma}_u)^{-1} - \boldsymbol{S}_{\boldsymbol{YX}}(\boldsymbol{S}_{\boldsymbol{X}} - \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{H}_2\left(\frac{\partial \boldsymbol{S}_{\boldsymbol{Z}}}{\partial a_k}\right)\boldsymbol{H}_2^\top(\boldsymbol{S}_{\boldsymbol{X}} - \boldsymbol{\Sigma}_u)^{-1},$$

leading to

$$\frac{\partial \mathrm{vec}(\widehat{\boldsymbol{B}})}{\partial a_k} = [((\boldsymbol{S}_{\boldsymbol{X}} - \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{H}_2) \otimes \boldsymbol{H}_1]\boldsymbol{D}\frac{\partial \mathrm{vech}(\boldsymbol{S}_{\boldsymbol{Z}})}{\partial a_k} - [((\boldsymbol{S}_{\boldsymbol{X}} - \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{H}_2) \otimes (\boldsymbol{S}_{\boldsymbol{YX}}(\boldsymbol{S}_{\boldsymbol{X}} - \boldsymbol{\Sigma}_u)^{-1}\boldsymbol{H}_2)]\boldsymbol{D}\frac{\partial \mathrm{vech}(\boldsymbol{S}_{\boldsymbol{Z}})}{\partial a_k}.$$

That is, the Jacobian at $\boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\boldsymbol{\mu}}$ reduces to

$$\left.\frac{\partial\{\mathrm{vec}(\widehat{\boldsymbol{B}})\}}{\partial\{\mathrm{vech}(\boldsymbol{S}_{\boldsymbol{Z}})^\top\}}\right|_{\boldsymbol{S}_{\boldsymbol{Z}} = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_{\boldsymbol{\mu}}} = (\boldsymbol{\Gamma}_{x+\xi}^{-1} \otimes \boldsymbol{I}_{p_1})(\boldsymbol{H}_2 \otimes \boldsymbol{H}_1)\boldsymbol{D} - (\boldsymbol{\Gamma}_{x+\xi}^{-1} \otimes \boldsymbol{B})(\boldsymbol{H}_2 \otimes \boldsymbol{H}_2)\boldsymbol{D} = \boldsymbol{QD},$$

where $\boldsymbol{Q} = (\boldsymbol{\Gamma}_{x+\xi}^{-1} \otimes \boldsymbol{I}_{p_1})(\boldsymbol{H}_2 \otimes \boldsymbol{H}_1) - (\boldsymbol{\Gamma}_{x+\xi}^{-1} \otimes \boldsymbol{B})(\boldsymbol{H}_2 \otimes \boldsymbol{H}_2)$. Therefore, by the delta method we have that $\sqrt{n}\,\mathrm{vec}(\widehat{\boldsymbol{B}} - \boldsymbol{B}) \xrightarrow{d} \mathcal{N}_{p_1p_2}(\boldsymbol{0}, \boldsymbol{\Phi}_{\boldsymbol{B}})$ where

$$\boldsymbol{\Phi}_{\boldsymbol{B}} = \boldsymbol{QD}(\boldsymbol{\Lambda} + \boldsymbol{\Lambda}_\mu)\boldsymbol{D}^\top\boldsymbol{Q}^\top.$$

Moreover, from Lemma 1(ii) it follows that $\sqrt{n}\,(\bar{\boldsymbol{Z}}-\boldsymbol{\mu})$ and $\sqrt{n}\,\mathrm{vec}(\widehat{\boldsymbol{B}}-\boldsymbol{B})$ are asymptotically independent, with $\sqrt{n}\,(\bar{\boldsymbol{Z}}-\boldsymbol{\mu}) \xrightarrow{d} \mathcal{N}_d\,(\boldsymbol{0},\boldsymbol{\Sigma})$; thus, the proof follows by noting from (9) that

$$\begin{pmatrix}\sqrt{n}(\widehat{\boldsymbol{a}}-\boldsymbol{a})\\ \sqrt{n}\mathrm{vec}(\widehat{\boldsymbol{B}}-\boldsymbol{B})\end{pmatrix}=\begin{pmatrix}\widehat{\boldsymbol{P}} & -(\boldsymbol{\xi}^\top\otimes\boldsymbol{I}_{p_1})\\ \boldsymbol{0} & \boldsymbol{I}_{p_1p_2}\end{pmatrix}\begin{pmatrix}\sqrt{n}(\bar{\boldsymbol{Z}}-\boldsymbol{\mu})\\ \sqrt{n}\mathrm{vec}(\widehat{\boldsymbol{B}}-\boldsymbol{B})\end{pmatrix}.\quad\square$$

Notice that for the structural case, we have $\boldsymbol{\xi}_i=\boldsymbol{\xi}$ for all $i=1,\ldots,n$ and $\boldsymbol{\Sigma}_\xi=\boldsymbol{0}$; then $\boldsymbol{\Gamma}_{x+\xi}=\boldsymbol{\Sigma}_x$ and $\boldsymbol{\Lambda}_{\boldsymbol{\mu}}=0$. For the functional case, we have $\boldsymbol{\Sigma}_x=\boldsymbol{0}$; then $\boldsymbol{\Gamma}_{x+\xi}=\boldsymbol{\Sigma}_\xi$. Defining

$$\boldsymbol{\Phi}=\begin{pmatrix}\boldsymbol{\Phi}_{\boldsymbol{a}} & \boldsymbol{\Phi}_{\boldsymbol{aB}}\\ \boldsymbol{\Phi}_{\boldsymbol{Ba}} & \boldsymbol{\Phi}_{\boldsymbol{B}}\end{pmatrix},$$

then it may be consistently estimated by replacing the unknown parameters with their consistent estimators given in Proposition 1. The ith element of $\widehat{\boldsymbol{\theta}}$ is asymptotically normally distributed with standard errors given by the square root of the ith diagonal element of $\boldsymbol{\Phi}$ divided by $\sqrt{n}$. Thus, we can formulate statistical hypotheses for the individual coefficients or, more generally, for contrasts of the form

$$\mathrm{H}_0:\boldsymbol{C\theta}=\boldsymbol{d}\quad\text{Versus}\quad\mathrm{H}_1:\boldsymbol{C\theta}\neq\boldsymbol{d},$$

which can be tested by using Wald-type statistics conveniently expressed as

$$n(\boldsymbol{C}\widehat{\boldsymbol{\theta}}-\boldsymbol{d})^\top\left[\boldsymbol{C}\widehat{\boldsymbol{\Phi}}\boldsymbol{C}^\top\right]^{-1}(\boldsymbol{C}\widehat{\boldsymbol{\theta}}-\boldsymbol{d}).\tag{10}$$

Under the null hypotheses, (10) has a limit $\chi^2(m)$ distribution where $m=\mathrm{rank}(\boldsymbol{C})$ corresponds to the number of linear restrictions.

It is important to remark that the estimation of $\boldsymbol{\Phi}_{\boldsymbol{B}}$ may involve plugging in estimators for both $\boldsymbol{\Sigma}_x$ and $\boldsymbol{\Sigma}_\xi$. As it is not possible to estimate these quantities separately, one must choose working with the structural ($\boldsymbol{\Sigma}_\xi=\boldsymbol{0}$) or the functional ($\boldsymbol{\Sigma}_x=\boldsymbol{0}$) version. We present one example with this characteristic in the next section.

3. The elliptical class of distributions

We shall use a definition for an elliptical distribution that is the same as the one presented in [7]. That is, we say that an $s\times 1$ random vector $\boldsymbol{Y}$ has a multivariate elliptical distribution with location parameter $\boldsymbol{\mu}$ and a positive definite scale matrix $\boldsymbol{\Psi}$ if its density function exists and is given by

$$f_{\boldsymbol{Y}}(\boldsymbol{y})=|\boldsymbol{\Psi}|^{-1/2}g\left[(\boldsymbol{y}-\boldsymbol{\mu})^\top\boldsymbol{\Psi}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right],\tag{11}$$

where $g:\mathbb{R}\to[0,\infty)$ is the density generator and it is such that $\int_0^\infty u^{\frac{s}{2}-1}g(u)<\infty$. We use the notation $\boldsymbol{Y}\sim El_s(\boldsymbol{\mu},\boldsymbol{\Psi},g)$. It is possible to show that the characteristic function of $\boldsymbol{Y}$ is given by $\psi(t)=\mathrm{E}(\exp(\mathrm{i}\boldsymbol{t}^\top\boldsymbol{Y}))=\exp(\mathrm{i}\boldsymbol{t}^\top\boldsymbol{\mu})\varphi(\boldsymbol{t}^\top\boldsymbol{\Psi t})$, where $\boldsymbol{t}\in\mathbb{R}^s$ and $\varphi:[0,\infty)\to\mathbb{R}$. Then, if φ is twice differentiable at zero, we have that $\mathrm{E}(\boldsymbol{Y})=\boldsymbol{\mu}$ and $\mathrm{Var}(\boldsymbol{Y})=\delta\boldsymbol{\Psi}$, where $\delta=-2\varphi'(0)$. A detailed description of the elliptical multivariate class given in (11) can be found in [7].

For the general elliptical situation, we consider two cases.

Case 1: Let $\boldsymbol{r}_i\overset{\text{i.i.d.}}{\sim}El_{2d}(\boldsymbol{0},\boldsymbol{\Psi},g)$; then $\mathrm{Var}(\boldsymbol{r}_1)=\delta\boldsymbol{\Psi}=\boldsymbol{\Sigma}_{\boldsymbol{r}}$. Since $\boldsymbol{Z}_1-\boldsymbol{\mu}_1=\boldsymbol{A}\boldsymbol{r}_1$, we have that $\boldsymbol{Z}_1-\boldsymbol{\mu}_1\sim El_d(\boldsymbol{0},\boldsymbol{A\Psi A}^\top,g)$, where $\mathrm{Var}(\boldsymbol{Z}_1)=\delta\boldsymbol{A\Psi A}^\top=\boldsymbol{\Sigma}$. As mentioned before, the asymptotic covariance matrix of $\sqrt{n}(\widehat{\boldsymbol{a}}-\boldsymbol{a},\mathrm{vec}(\widehat{\boldsymbol{B}}-\boldsymbol{B}))$ just depends on $\boldsymbol{r}_1$, only through $\boldsymbol{\Sigma}_{\boldsymbol{r}}$ and $\boldsymbol{\Lambda}_{\boldsymbol{r}}$. Therefore, following [3] (see page 209, Eq. 2.3) and by using matrix properties, we have that

$$\boldsymbol{\Lambda}_{\boldsymbol{r}}=\mathrm{Var}[\mathrm{vec}(\boldsymbol{r}_1\boldsymbol{r}_1^\top)]=(1+\kappa)(\boldsymbol{I}_{4d^2}+\boldsymbol{K}_{2d,2d})(\boldsymbol{\Sigma}_{\boldsymbol{r}}\otimes\boldsymbol{\Sigma}_{\boldsymbol{r}})+\kappa\mathrm{vec}(\boldsymbol{\Sigma}_{\boldsymbol{r}})\mathrm{vec}(\boldsymbol{\Sigma}_{\boldsymbol{r}})^\top,$$

where $\boldsymbol{K}_{s_1,s_2}$ is a commutation matrix such that for the s_1-vector $\boldsymbol{a}$ and the s_2-vector $\boldsymbol{b}$, $\boldsymbol{K}_{s_1,s_2}(\boldsymbol{b}\otimes\boldsymbol{a})=(\boldsymbol{a}\otimes\boldsymbol{b})$ and $\kappa=\varphi''(0)/[\varphi'(0)]^2-1$ is the kurtosis parameter assumed to be known. According to Lemma 4.1 in [2], the following relationships hold: $\delta=\mathrm{E}[U/(2d)]$ and $\delta^2(\kappa+1)=\mathrm{E}\{U^2/[4d(d+1)]\}$, where U has the same distribution as $\boldsymbol{r}_1^\top\boldsymbol{\Psi}^{-1}\boldsymbol{r}_1$.

The following example illustrates *Case 1* in the context of the simple linear regression model with measurement errors.

Example 1. Let's consider the univariate measurement error model, that is, when $p_1=p_2=1$,

$$y_i=a+bx_i+q_i,\qquad Y_i=y_i+e_i\quad\text{and}\quad X_i=x_i+u_i$$

with $\boldsymbol{r}_i=(q_i,e_i,u_i,(x_i-\xi_i))^\top\sim El_4(\boldsymbol{0},\boldsymbol{\Psi},g)$, where $\mathrm{Var}(\boldsymbol{r}_i)=\delta\boldsymbol{\Psi}=\mathrm{diag}(\sigma_q^2,\sigma_e^2,\sigma_u^2,\sigma_x^2)$. Define the sample moments $S_{XY}=\sum_{i=1}^n(X_i-\bar{X})Y_i/n$, $S_X=\sum_{i=1}^n(X_i-\bar{X})X_i/n$, $\bar{Y}=\sum_{i=1}^nY_i/n$ and $\bar{X}=\sum_{i=1}^nX_i/n$. Then, under the suppositions (A1) and (A2), $\widehat{b}=S_{YX}/(S_X-\sigma_u^2)$ is a consistent estimator for b and

$$\sqrt{n}(\widehat{b}-b)\xrightarrow{d}\mathcal{N}_1\,(0,\phi_b)$$

where

$$\phi_b = \frac{(3\kappa+2)b^2\sigma_u^4 + (\kappa+1)[\sigma_u^2(b^2\sigma_x^2+\sigma_q^2+\sigma_e^2) + \sigma_x^2(\sigma_q^2+\sigma_e^2)] + \sigma_\xi^2(b^2\sigma_u^2+\sigma_q^2+\sigma_e^2)}{(\sigma_x^2+\sigma_\xi^2)^2}.$$

The asymptotic variance ϕ_b is computed by specializing the matrices $\boldsymbol{Q}$, $\boldsymbol{D}$, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Lambda_\mu}$ of Proposition 2 under $p_1 = p_2 = 1$ and the additional supposition $\boldsymbol{r}_i \sim El_4(\boldsymbol{0}, \boldsymbol{\Psi}, g)$. Taking $\sigma_e^2 = 0$ we get the same result as in Theorem 3.3 of [3]; in addition, if $\sigma_e^2 = \kappa = 0$ we obtain the same results as in [5]. In this example, when $\kappa \neq 0$, the estimation of the asymptotic variance ϕ_b depends on both σ_x^2 and σ_ξ^2, so the user must decide which approach to use: functional or structural.

Case 2: Suppose that $\boldsymbol{q}_i \overset{\text{i.i.d.}}{\sim} El_{p_1}(\boldsymbol{0}, \boldsymbol{\Psi_q}, g_q)$, $\boldsymbol{e}_i \overset{\text{i.i.d.}}{\sim} El_{p_1}(\boldsymbol{0}, \boldsymbol{\Psi_e}, g_e)$, $\boldsymbol{u}_i \overset{\text{i.i.d.}}{\sim} El_{p_2}(\boldsymbol{0}, \boldsymbol{\Psi_u}, g_u)$ and $\boldsymbol{v}_i = \boldsymbol{x}_i - \boldsymbol{\xi}_i \overset{\text{i.i.d.}}{\sim} El_{p_2}(\boldsymbol{0}, \boldsymbol{\Psi_x}, g_x)$ are independent random vectors; then $\mathrm{Var}(\boldsymbol{r}_i) = \boldsymbol{\Sigma_r} = \mathrm{diag}(\delta_q\boldsymbol{\Psi}_q, \delta_e\boldsymbol{\Psi}_e, \delta_u\boldsymbol{\Psi}_u, \delta_x\boldsymbol{\Psi}_x)$, where δ_q, δ_e, δ_u and δ_x depend on their characteristic functions, respectively (as defined previously). We must compute $\mathrm{Var}(\mathrm{vec}(\boldsymbol{r}_1\boldsymbol{r}_1^\top)) = \mathrm{Var}(\boldsymbol{r}_1 \otimes \boldsymbol{r}_1)$. Notice that there exists a matrix $\boldsymbol{J}$ such that $(\boldsymbol{r}_1 \otimes \boldsymbol{r}_1) = \boldsymbol{J}\boldsymbol{w}$, where

$$\boldsymbol{w} = (\boldsymbol{w}_1^\top, \boldsymbol{w}_2^\top, \boldsymbol{w}_3^\top, \boldsymbol{w}_4^\top, \boldsymbol{w}_5^\top, \boldsymbol{w}_6^\top, \boldsymbol{w}_7^\top)^\top$$

with $\boldsymbol{w}_1^\top = (\boldsymbol{q}_1^\top \otimes \boldsymbol{q}_1^\top, \boldsymbol{e}_1^\top \otimes \boldsymbol{e}_1^\top, \boldsymbol{u}_1^\top \otimes \boldsymbol{u}_1^\top, \boldsymbol{v}_1^\top \otimes \boldsymbol{v}_1^\top)^\top$, $\boldsymbol{w}_2 = (\boldsymbol{e}_1 \otimes \boldsymbol{q}_1)$, $\boldsymbol{w}_3 = (\boldsymbol{u}_1 \otimes \boldsymbol{q}_1)$, $\boldsymbol{w}_4 = (\boldsymbol{v}_1 \otimes \boldsymbol{q}_1)$, $\boldsymbol{w}_5 = (\boldsymbol{u}_1 \otimes \boldsymbol{e}_1)$, $\boldsymbol{w}_6 = (\boldsymbol{v}_1 \otimes \boldsymbol{e}_1)$, $\boldsymbol{w}_7 = (\boldsymbol{v}_1 \otimes \boldsymbol{u}_1)$. Then, $\mathrm{Var}(\boldsymbol{r}_1 \otimes \boldsymbol{r}_1) = \boldsymbol{J}\mathrm{Var}(\boldsymbol{w})\boldsymbol{J}^\top$, where

$$\mathrm{Var}(\boldsymbol{w}) = \mathrm{diag}(\boldsymbol{\Sigma}_{w_1}, \boldsymbol{\Sigma}_{e,q}, \boldsymbol{\Sigma}_{u,q}, \boldsymbol{\Sigma}_{v,q}, \boldsymbol{\Sigma}_{u,e}, \boldsymbol{\Sigma}_{v,e}, \boldsymbol{\Sigma}_{v,u})$$

with $\boldsymbol{\Sigma}_{w_1} = \mathrm{diag}(\boldsymbol{\Lambda_q}, \boldsymbol{\Lambda_e}, \boldsymbol{\Lambda_u}, \boldsymbol{\Lambda_x})$, $\boldsymbol{\Sigma}_{a,b} = \boldsymbol{\Sigma}_a \otimes \boldsymbol{\Sigma}_b$,

$$\boldsymbol{\Lambda_a} = (1+\kappa_a)(\boldsymbol{I}_{p_a^2} + \boldsymbol{K}_{p_a,p_a})(\boldsymbol{\Sigma_a} \otimes \boldsymbol{\Sigma_a}) + \kappa_a \mathrm{vec}(\boldsymbol{\Sigma_a})\mathrm{vec}(\boldsymbol{\Sigma_a})^\top,$$

p_a being the dimension of the generic vector $\boldsymbol{a}$, κ_a the kurtosis parameter of $\boldsymbol{a}$ and $\boldsymbol{\Sigma_a}$ the variance matrix of $\boldsymbol{a}$ for $\boldsymbol{a} = \boldsymbol{q}, \boldsymbol{e}, \boldsymbol{u}, \boldsymbol{x}$.

The following example illustrates *Case 2* in the context of the simple linear regression model with measurement error.

Example 2. Let's consider the same model as in Example 1, with $\boldsymbol{r}_i = (q_i, e_i, u_i, (x_i - \xi_i))^\top$ where (I) $q_i \overset{\text{i.i.d.}}{\sim} El_1(0, \psi_q, g_q)$, (II) $e_i \overset{\text{i.i.d.}}{\sim} El_1(0, \psi_e, g_e)$, (III) $u_i \overset{\text{i.i.d.}}{\sim} El_1(0, \psi_u, g_u)$ and (IV) $v_i = (x_i - \xi_i) \overset{\text{i.i.d.}}{\sim} El_1(0, \psi_x, g_x)$ are independent random variables. Consider the same sample moments as were defined in the previous example. Then, $\widehat{b} = S_{YX}/(S_X - \sigma_u^2)$ is also a consistent estimator for b and, under the suppositions (A1) and (A2) together with (I)–(IV), we have

$$\sqrt{n}(\widehat{b} - b) \xrightarrow{d} \mathcal{N}_1(0, \phi_b)$$

where

$$\phi_b = \frac{(3\kappa_u+2)b^2\sigma_u^4 + (\sigma_e^2+\sigma_q^2)\sigma_u^2 + (b^2\sigma_u^2+\sigma_q^2+\sigma_e^2)(\sigma_\xi^2+\sigma_x^2)}{(\sigma_x^2+\sigma_\xi^2)^2}.$$

The asymptotic variance ϕ_b is computed by specializing the matrices $\boldsymbol{Q}$, $\boldsymbol{D}$, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Lambda_\mu}$ of Proposition 2, under $p_1 = p_2 = 1$ and the additional suppositions (I)–(IV). Notice that, when q_i, e_i, u_i and x_i are independent random variables, the asymptotic variance of $\widehat{b}$ just depends on the distribution of u_i. Taking $\sigma_e^2 = 0$ we obtain $\phi_b = \{(3\kappa_u+2)b^2\sigma_u^4 + \sigma_q^2\sigma_u^2 + (b^2\sigma_u^2+\sigma_q^2)(\sigma_x^2+\sigma_\xi^2)\}/(\sigma_x^2+\sigma_\xi^2)^2$; in addition, if $\sigma_e^2 = \kappa_u = 0$ we attain $\phi_b = \{2b^2\sigma_u^4 + \sigma_q^2\sigma_u^2 + (b^2\sigma_u^2+\sigma_q^2)(\sigma_x^2+\sigma_\xi^2)\}/(\sigma_x^2+\sigma_\xi^2)^2$ which is the very same asymptotic variance as was derived in [5].

In general, Zografos [21] showed that, if $\boldsymbol{z} \sim El_{2d}(\boldsymbol{0}, \boldsymbol{\Psi}, g)$, then the kurtosis parameter is given by

$$\kappa = \frac{\pi^d \int_0^\infty w^{d+1} g(w)\mathrm{d}w}{4d(d+1)\Gamma(d)\delta^2} - 1$$

where δ was defined previously. Then, assumption (A2) is always true under the elliptical class of distributions if φ is twice differentiable at zero, such that $\varphi'(0) \neq 0$ and $\int_0^\infty w^{d+1}g(w)\mathrm{d}w < \infty$.

Acknowledgments

We gratefully acknowledge helpful suggestions from two anonymous referees and important comments from the associate editor. Grants from FAPESP and CNPq (Brazil) and FONDECYT 1085241 (Chile) are also acknowledged.

References

[1] Y. Anemiya, W.A. Fuller, Estimation for the multivariate errors-in-variables model with estimated error covariance matrix, The Annals of Statistics 12 (1984) 497–509.
[2] R.B. Arellano-Valle, H. Bolfarine, L. Gasco, Measurement error models with nonconstant covariance matrix, Journal of Multivariate Analysis 82 (2002) 395–415.
[3] R.B. Arellano-Valle, H. Bolfarine, R. Vilca-Labra, Ultrastructural elliptical model, The Canadian Journal of Statistics 24 (1996) 207–216.

[4] N.N. Chan, T.K. Max, Estimation of multivariate linear functional relationships, Biometrika 70 (1983) 263–267.
[5] C.L. Cheng, J.W. Van Ness, On the unreplicated ultrastructural model, Biometrika 78 (1991) 442–445.
[6] P.F. Dahm, W.A. Fuller, Generalized least squares estimation of the functional multivariate linear errors-in-variables model, Journal of Multivariate Analysis 19 (1986) 132–141.
[7] K.T. Fang, S. Kotz, K.W. Ng, Symmetric Multivariate and Related Distributions, Chapman and Hall, London, 1990.
[8] A. Fujita, A.G. Patriota, J.R. Sato, S. Miyano, The impact of measurement errors in the identification of regulatory networks, BMC Bioinformatics 10 (2009) 412.
[9] W. Fuller, Measurement Error Models, Wiley, Chichester, 1987.
[10] L.J. Gleser, Estimation in a multivariate "errors in variables" regression model: large sample results, The Annals of Statistics 09 (1981) 24–44.
[11] J.D. Healy, Maximum likelihood estimation of a multivariate linear functional relationship, Journal of Multivariate Analysis 10 (1980) 243–251.
[12] B.C. Kelly, Some aspects of measurement error in linear regression of astronomical data, The Astrophysical Journal 665 (2007) 1489–1506.
[13] T. Kollo, D. von Rosen, Advanced Multivariate Statistics with Matrices, in: Mathematics and its Applications, Springer, Dordrecht, New York, 2005.
[14] S.B. Kulathinal, K. Kuulasmaa, D. Gasbarra, Estimation of an errors-in-variables regression model when the variances of the measurement error vary between the observations, Statistics in Medicine 21 (2002) 1089–1101.
[15] E.L. Lehmann, G. Casella, Theory of Point Estimation, 2nd ed., Springer-Verlag, New York, 1998.
[16] J.R. Magnus, H. Neudecker, Matrix Differential Calculus with Applications in Statistics and Econometrics, third ed., Wiley, Chichester, 2007.
[17] A.G. Patriota, H. Bolfarine, M. de Castro, A heteroscedastic structural errors-in-variables model with equation error, Statistical Methodology 6 (2009) 408–423.
[18] H. Schneeweiss, Consistent estimation of a regression with errors in the variables, Metrika 23 (1976) 101–115.
[19] C.M. Theobald, J.R. Mallison, Comparative calibration, linear structural relationships and congeneric measurements, Biometrics 34 (1978) 39–45.
[20] M. Xiong, J. Li, X. Fang, Identification of genetic networks, Genetics 166 (2004) 1037–1052.
[21] K. Zografos, On Mardia's and Song's measures of kurtosis in elliptical distributions, Journal of Multivariate Analysis 99 (2008) 858–879.