Metrika (2006)
DOI 10.1007/s00184-006-0029-z

ORIGINAL ARTICLE

Chi-Lun Cheng · Alexander Kukush

Non-existence of the first moment of the adjusted least squares estimator in multivariate errors-in-variables model

Received: 5 January 2004 / Accepted: 21 March 2005

Abstract The linear normal multivariate errors-in-variables model is considered. The model involves an equation error. It is proved in both structural and functional cases that the first moment of the adjusted least squares estimator does not exist.

Keywords Adjusted least squares · Equation error model · Functional model · Infinite first moment · Linear multivariate error-in-variables model · Structural model

AMS Subject Classifications 62J05 · 62H12 · 62H10

1 Introduction

All statistical analysis assumes that the observations are obtained without any error. Such an assumption is quite often violated in many statistical applications because of presence of measurement errors in the observation, see Fuller (1987) and Cheng and Van Ness (1999) for details on different issues related with measurement error modeling. It is shown in Fuller (1987, p 28, Problem 13) that the maximum likelihood estimator of a slope parameter does not posses finite expectation in the univariate structural normal model when the measurement error variance of the independent variable is known but the variance of error in response variable is unknown. The maximum likelihood estimator under this case coincides with the adjusted least squares (ALS) estimator. However, it is not clear that non-existence of the finite first moment is true for the ALS estimator in the functional model

C.-L. Cheng
Institute of Statistical Science, Academia Sinica, 115 Taipei, Republic of China

A. Kukush (✉)
Department of Mechanics and Mathematics, National Taras Shevchenko University, Vladimirskaya st. 64, 01033 Kiev, Ukraine
E-mail: alexander_kukush@univ.kiev.ua

too. We have attempted to generalize this result for multivariate normal functional model in this paper.

The plan of the paper is as follows. In section 2, we describe the multivariate model and consider the ALS estimator. We prove that its first moment does not exist in section 3. Some concluding remarks are made in section 3.

2 Multivariate model

Let us consider the model for the observations as

$$\beta_0^{\mathrm{T}} + x^{\mathrm{T}}\beta \approx y^{\mathrm{T}}. \tag{1}$$

where $\beta_0 \in \mathbf{R}^{d\times 1}$ and $\beta \in \mathbf{R}^{n\times d}$ are non-stochastic regression parameters to be estimated; $x \in \mathbf{R}^{n\times 1}$ and $y \in \mathbf{R}^{d\times 1}$ are measured with errors. Further, we assume that there exist $\beta_0 \in \mathbf{R}^{d\times 1}$ and $\beta \in \mathbf{R}^{n\times d}$ such that

$$\beta_0^{\mathrm{T}} + \xi_i^{\mathrm{T}}\beta = \eta_i^{\mathrm{T}}, \qquad i = 1, \ldots, m, \tag{2}$$

and we observe

$$x_i = \xi_i + \delta_i, \quad y_i = \eta_i + e_i, \quad i = 1, \ldots, m. \tag{3}$$

Here ξ_i and η_i are the true values which are unobservable. The measurement error in the independent variable is δ_i and e_i is the combined measurement error of the dependent variable and the equation error. Thus the models (2) and (3) describe the multivariate equation error model. This type of multivariate model is considered in Gleser (1992). Next, we assume that $m \geq n + 1$ so that the models (2) and (3) are identifiable. If ξ_i are non-stochastic (fixed) values, it is called a *functional* model. On the other hand, if ξ_i are independently and identically distributed random variable, then it is called a *structural* model.

It is convenient to rewrite (2) and (3) in a matrix form as

$$\mathbf{1}_m\beta_0^{\mathrm{T}} + K\beta = H, \tag{4}$$

$$X = K + \Delta, \quad Y = H + E, \tag{5}$$

and the model (1) corresponds to an over-determined set of equations $\mathbf{1}_m\beta_0^T + X\beta \approx Y$.

Here $\mathbf{1}_m := (1, \ldots, 1)^{\mathrm{T}} \in \mathbf{R}^{m\times 1}$, $\{X, K, \Delta\} \subset \mathbf{R}^{m\times n}$ and $\{Y, H, E\} \subset \mathbf{R}^{m\times d}$. We denote in (5)

$$X^{\mathrm{T}} := [x_1, \ldots, x_m], \quad Y^{\mathrm{T}} := [y_1, \ldots, y_m],$$

and similarly ξ_i^{T}, δ_i^{T}, η_i^{T} and e_i^{T} are rows of K, Δ, H and E, respectively.

2.1 The ALS estimator

Now, we state the various assumptions about the model (2) and (3). Hereafter, bold **E** denotes the expectation, and cov stands for the variance–covariance matrix of a random vector.

i. $\mathbf{E}(\Delta) = 0$ and $\mathbf{E}(E) = 0$.
ii. E is independent of Δ and K.
iii. $\{\delta_i, i = 1, \dots, m\}$ are i.i.d. and Gaussian with $\mathrm{cov}(\delta_i) = V_\delta$, where V_δ is a known positive definite matrix.

To construct the ALS estimator, we obtain from (2)

$$\eta_i^{\mathrm{T}} - \bar{\eta}^{\mathrm{T}} = (\xi_i^{\mathrm{T}} - \bar{\xi}^{\mathrm{T}})\beta,$$

or in matrix notation,

$$H - \bar{H} = (K - \bar{K})\beta.$$

Here bar means average, e.g., $\bar{\eta} := \dfrac{1}{m} \sum_{i=1}^{m} \eta_i$ and $\bar{H}^{\mathrm{T}} := [\bar{\eta}, \dots, \bar{\eta}] \in \mathbf{R}^{n \times m}$. Now, the ALS estimator of β is

$$\hat{\beta} := [(X^{\mathrm{T}} - \bar{X}^{\mathrm{T}})(X - \bar{X}) - mV_\delta]^{-1}(X^{\mathrm{T}} - \bar{X}^{\mathrm{T}})(Y - \bar{Y}). \tag{6}$$

The ALS estimator of β_0 is found from the equation as $\hat{\beta}_0 = \bar{y} - \hat{\beta}^{\mathrm{T}}\bar{x}$.

It is worth noting that the ALS estimator is the maximum likelihood estimator for the normal structural model. But the maximum likelihood estimation breaks down in the normal functional model. However, the ALS estimator is regarded as method of moments estimator in the normal functional model. Generally speaking, ALS estimator is widely used in both functional and structural models with or without normality assumption. This is the reason that we investigate the moment property of the ALS estimator. See, for example, Fuller (1987) and Cheng and Van Ness (1999) for more detailed discussion.

3 Non-existence of the first moment

In this section, we consider the multivariate normal function model, i.e., we assume in continuation to (i)–(iii) in section 2.1 that,

iv. $\{\xi_i, i = 1, \dots, m\}$ are non-random vectors and the matrix $S_{KK} := (K^{\mathrm{T}} - \bar{K}^{\mathrm{T}})(K - \bar{K})$ is positive definite.

Theorem 1 *Under the conditions* (i), (ii), (iii) *and* (iv), *let* $\beta_1, \dots, \beta_p$ *be* $(n \times d)$ *matrices such that the following condition holds:*

$$\mathrm{rank}([\beta_1, \dots, \beta_p]) = n. \tag{7}$$

Then there exists a number i, $i = 1, \dots, p$, *such that* $\mathbf{E}_{\beta_0, \beta_i} \|\hat{\beta}\| = \infty$, *where* $\hat{\beta}$ *is given by* (6).

Proof Without loss of generality, we assume that $V_\delta = I_n$. We confine our attention to the case for $n \geq 2$ only. We prove this by contradiction. Suppose that for each $i = 1, \ldots, p$, $\mathbf{E}_{\beta_0,\beta_i} \|\hat{\beta}\| < \infty$. Fix $\beta \in \{\beta_1, \ldots, \beta_p\}$. From (6) we have

$$E[\hat{\beta} \mid \Delta] = (S_{XX} - mI_n)^{-1}(X^{\mathrm{T}} - \bar{X}^{\mathrm{T}})(H - \bar{H}) = (S_{XX} - mI_n)^{-1} S_{XK}\beta.$$

Therefore the expectation, $E\{(S_{XX} - mI_n)^{-1} S_{XK}[\beta_1, \ldots, \beta_p]\}$ is finite and from (7), we obtain the existence of

$$E\{(S_{XX} - mI_n)^{-1} S_{XK}\}. \tag{8}$$

Let $K^{\mathrm{T}} - \bar{K}^{\mathrm{T}} := [d_1, \ldots, d_m]$ and $c_i := (d_1(i), \ldots, d_m(i))^{\mathrm{T}}$, $i = 1, \ldots, n$, where $d_j(i)$ is the ith component of the column vector d_j. Denoting $\delta(i) := (\delta_1(i), \ldots, \delta_m(i))^{\mathrm{T}}$ and $h_i := \delta(i) - \big(\delta(i), \frac{1}{\sqrt{m}}\mathbf{1}_m\big)\frac{1}{\sqrt{m}}\mathbf{1}_m$, we have

$$S_{XX} = G(h_1 + c_1, \ldots, h_n + c_n),$$

where G is the corresponding Gram matrix. Let $h_i + c_i := u_i$ and $U^{\mathrm{T}} := [u_1, \ldots, u_n]$. Then $S_{XX} = UU^{\mathrm{T}}$, $S_{XK} = U[d_1, \ldots, d_m]$ and (8) takes the form

$$\mathbf{E}\{(UU^{\mathrm{T}} - mI_n)^{-1} U[d_1, \ldots, d_m]\}.$$

By the condition (iv) $d_1 \neq 0$, and we see that

$$\mathbf{E}\{(UU^{\mathrm{T}} - mI_n)^{-1} Ud_1\}. \tag{9}$$

should be finite.

The random matrix U has independent rows $u_1, \ldots, u_n$ with positive Gaussian densities $\rho_1(u_1), \ldots, \rho_n(u_n)$. Now,

$$\det(UU^{\mathrm{T}} - mI_n) = (\|u_1\|^2 - m)A_{11} + \sum_{i,j=2}^{n} (u_i, u_1)(u_j, u_1)B_{ij}.$$

where A_{11} is algebraic complement to $(1, 1)$ element of $UU^{\mathrm{T}} - mI_n$, and up to a sign B_{ij} is a minor, which is obtained from $UU^{\mathrm{T}} - mI_n$ by crossing two columns with numbers 1 and j and two rows with numbers 1 and i (if $n = 2$ then $B_{22} := -1$). The absolute value of the first component of the vector $(UU^{\mathrm{T}} - mI_n)^{-1}Ud_1$ is

$$F := \frac{1}{|\det(UU^{\mathrm{T}} - mI_n)|}\left|(u_1, d_1)A_{11} + \sum_{i,j=2}^{n} (u_i, u_1)(u_j, d_1)B_{ij}\right|.$$

Let $T := \{(u_1, \ldots, u_n) : u_1 \in \mathbf{R}^{\mathrm{m}}, |u_j|^2 \geq m + 1, 2 \leq j \leq n$, and $|(u_i, u_j)| < \frac{1}{n-1}, i, j \geq 2, i \neq j, d_1 \notin \mathrm{span}(u_2, \ldots, u_n)\}$. We mention that for Gaussian vectors $u_2, \ldots, u_n$, the relation $d_1 \notin \mathrm{span}(u_2, \ldots, u_n)$ holds a.s., because $m \geq n + 1$ and $d_1 \neq 0$. Consider the integral

$$\int_T F(u_1, \ldots, u_n)\rho_1(u_1) \cdots \rho_n(u_n)du_1 \cdots du_n$$

$$= \int_{T_0} \rho_2(u_2) \cdots \rho_n(u_n)du_2 \cdots du_n \int_{\mathbf{R}^m} \rho_1(u_1)F(u_1, \ldots, u_n)du_1. \tag{10}$$

where T_0 is a base of T, i.e., $T = \mathbf{R}^m \times T_0$. We use the spherical coordinates $r, \varphi_1, \ldots, \varphi_{m-1}$ in the inner integral with respect to u_1. The inner integral of (10) equals

$$\int\limits_{[0,\pi]^{m-2}\times[0,2\pi]} J(\varphi_1, \ldots, \varphi_{m-1}) d\varphi_1 \cdots \varphi_{m-1}$$

$$\times \int\limits_0^\infty \frac{r^{m-1}\rho_1(r, \varphi_1, \ldots, \varphi_{m-1}) r \left| \psi_1 A_{11} + \sum_{i,j=2}^n f_i \psi_j B_{ij} \right|}{\left| (r^2 - m) A_{11} + r^2 \sum_{i,j=2}^n f_i f_j B_{ij} \right|} dr. \tag{11}$$

Here $r^{m-1} J(\varphi_1, \ldots, \varphi_{m-1})$ is the Jacobian of the spherical coordinates, A_{11} and B_{ij} are functions of $u_2, \ldots, u_n$, and f_i, ψ_j are the functions of $\varphi_1, \ldots, \varphi_{m-1}$. Now, the expression $r \mid \psi_1 A_{11} + \sum_{i,j=2}^n f_i \psi_j B_{ij} \mid$ corresponds to the numerator of F, i.e., to $\mid (u_1, f) \mid$, with $f := d_1 A_{11} + \sum_{i=2}^n u_i \sum_{j=2}^n (u_j, d_1) B_{ij}$. For $(u_2, \ldots, u_n) \in T_0$, $A_{11} > 0$ and $f \neq 0$ because $d_1 \notin \operatorname{span}(u_2, \ldots, u_n)$. Therefore $\psi_1 A_{11} + \sum_{i,j=2}^n f_i \psi_j B_{ij} \neq 0$.

In the inner integral of (11), we have singularity of the form $\dfrac{1}{\mid r - r_0 \mid}$, $r_0 > 0$ provided $A_{11} + \sum_{i,j=2}^n f_i f_j B_{ij} > 0$. But the latter inequality holds for $(\varphi_1, \ldots, \varphi_{m-1}) \in D_0$, where D_0 is domain with $|f_i| \leq \varepsilon_0$, $2 \leq i \leq n$ and ε_0 is sufficiently small. Therefore

$$\int\limits_{D_0} J(\varphi_1 \ldots \varphi_{m-1}) d\varphi_1 \cdots d\varphi_{m-1}$$

$$\times \int\limits_0^\infty \frac{r^{m-1}\rho_1 r \left| \psi_1 A_{11} + \sum_{i,j=2}^n f_i \psi_j B_{ij} \right|}{\left| (r^2 - m) A_{11} + r^2 \sum_{i,j=2}^n f_i f_j B_{ij} \right|} dr = \infty,$$

which implies that the integral (10) is divergent. But this contradicts to the finiteness of (9). □

Remark In case of structural model, i.e., ξ_i are IID random variables, we first condition on ξ and then the theorem above shows that the conditional expectation of the ALS estimator fails to exist. It implies that the first moment of the ALS estimator of the structural model does not exist regardless of the distribution of ξ.

4 Concluding remarks

We have considered a linear multivariate errors-in-variables model specified by (2) and (3), in that both the structural and functional case, the expectation of the ALS estimator does not exist. This is an extension of well-known properties of the estimator in a univariate normal structural model.

The ALS estimator is widely used as stated at the end of section 2. Despite its good asymptotic properties, see Fuller (1987), the ALS estimator fails to have finite first moment. Consequently, many important criteria such as mean squared error, cannot be used for a comparison among estimators. To overcome this difficulty, Fuller (1987, section 2.5) suggested modified estimators that possess finite moments. However, the modified estimators have complicated distribution. Cheng et al. (2000) introduced a small sample modification of the ALS estimator for a polynomial functional model. This modification is more stable than the ALS estimator, but the expectation of this modified estimator does not exist as well, due to the singularity that is still preserved. To find more satisfactory modified estimators seem to be an interesting problem for further research. But it is beyond the scope of the present note.

Acknowledgements The authors are grateful to the editor and the referee for their comments on an earlier draft of this paper. C.-L. Cheng and A. Kukush are partially supported by a grant from the National Science Council of Taiwan, R.O.C. A. Kukush is also supported by a postdoctoral research fellowship of the Belgian office for Scientific, Technical Affairs.

References

Cheng CL, Van Ness JW (1999) Statistical regression with measurement error. Arnold, London

Cheng CL, Schneeweiss H, Thamerus M (2000) A small sample estimator for a polynomial regression with errors in the variables. J Roy Stat Soc B 62:699–709

Fuller WA (1987) Measurement error models. Wiley, New York

Gleser LJ (1992) The importance of assessing measurement reliability in multivariate regression. J Am Stat Assoc 87:696–707