3B2v7.51c GML4.3.1 YJMVA : 2156 Prod.Type:COM pp.1–14(col.fig.:NIL) ED:KCThomas PAGN: vs SCAN:

ARTICLE IN PRESS

ACADEMIC PRESS

Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT

Journal of Multivariate Analysis ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Journal of Multivariate Analysis

http://www.elsevier.com/locate/jmva

The efficiency of adjusted least squares in the linear functional relationship

Alexander Kukush[a,b,*] and Erich Otto Maschke[b]

[a] *National Taras Shevchenko, University of Kiev, Vladimirskaya st. 64, 01033 Kiev, Ukraine*
[b] *University of Munich, Germany*

Received 11 April 2001

Abstract

A linear functional errors-in-variables model with unknown slope parameter and Gaussian errors is considered. The measurement error variance is supposed to be known, while the variance of errors in the equation is unknown. In this model a risk bound of asymptotic minimax type for arbitrary estimators is established. The bound lies above that one which was found previously in the case of both variances known. The bound is attained by an adjusted least-squares estimators.
© 2003 Elsevier Science (USA). All rights reserved.

AMS 2000 subject classifications: 62J10; 62J05

Keywords: Adjusted least-squares estimator; Asymptotic efficiency; Gaussian errors; Hájek bound; Linear functional errors-in-variables model

1. Introduction

The linear functional errors-in-variables model has found extensive treatment in the literature—for some reviews see, e.g., [2,3]. Consider such a model, with known measurement error variance and unknown variance of errors in the equation. Then there exists a natural modification of the least squares estimator, see e.g. [2, p. 85], which is consistent and asymptotically normal. We call it an adjusted least squares (ALS) estimator due to [1], where it was developed in a more general setting, namely for a polynomial regression.

*Corresponding author. Lepse Blvd. 83-G, 56, 03126, Kiev, Ukraine. Fax: +0038-044-266-15-58.
E-mail address: alexander_kukush@univ.kiev.ua (A. Kukush).

0047-259X/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0047-259X(03)00048-4

UNCORRECTED PROOF

A natural question arises about asymptotic efficiency of such an estimator. This problem is non-trivial, because the unknown non-stochastic design points are nuisance parameters in the model, and the number of these parameters increases with the sample size. A result of the asymptotic minimax type for estimation in the linear functional error-in-variables model with **both variances known** has been obtained by Nussbaum [6] and by Hasminskii and Ibragimov [4]. It was shown there that the bound of Hájek type is attained by the maximum likelihood estimator.

In the present paper we follow the line of Hasminskii and Ibragimov [4] and establish such a bound for the linear functional model with uncorrelated errors and unknown variance of the errors of the response variable. The asymptotic bound is attained by the ALS estimator. Thus the ALS estimator delivers the smallest possible averaged losses, and it is asymptotically efficient in the sense of [5].

In the next section the linear errors-in-variables model is introduced and the ALS estimator is presented. It is shown that it is asymptotically normal uniformly with respect to designs from a certain class. In Section 3 the asymptotic minimax bound is given. In Section 4 it is shown that the bound is attained by the ALS estimator. The crucial calculations of the inverse Fisher information matrix in the corresponding linear structural model are given in Appendix A. An auxiliary convergence result is proved in Appendix B. The proofs of Lemma 1 and Theorems 1 and 2 are contained in Appendix C.

The symbols E and D stand for the expectation and variance, respectively.

2. The ALS estimator in linear model

Consider a linear functional relationship with errors in the variables and without intercept term:

$$\begin{aligned} y_i &= \beta\xi_i + \varepsilon_i, \\ x_i &= \xi_i + \delta_i, \end{aligned} \tag{1}$$

$i = 1, \ldots, n$, where $(\delta_i, \varepsilon_i)$ are i.i.d. random errors with Gaussian distribution. We suppose that $(\delta_i, \varepsilon_i)$ have the expectation 0 and covariance matrix

$$\Omega = \begin{pmatrix} \sigma_\delta^2 & 0 \\ 0 & v \end{pmatrix}$$

with unknown $v > 0$. Thus we assume that the errors δ_i and ε_i are independent, and the variance of δ_i is known, while the variance v of ε_i is unknown. The design points ξ_i, $i = 1, \ldots, n$, are unobservable non-stochastic variables.

In model (1), the values $v, \xi_1, \ldots, \xi_n$ are nuisance parameters, the number of which grows with the sample size. We are interested only in the slope parameter β. The adjusted least-squares (ALS) estimator of β is given by

$$\hat{\beta} = \frac{\overline{xy}}{\overline{x^2} - \sigma_\delta^2}, \tag{2}$$

UNCORRECTED PROOF

where $\overline{xy} = \frac{1}{n}\sum_1^n x_i y_i$, $\overline{x^2} = \frac{1}{n}\sum_1^n x_i^2$, see [2]. Under normal distributions of errors, the denominator in (2) does not equal 0 a.s., therefore $\hat{\beta}$ is well defined by (2).

Hereafter we use the notations for averaged values like $\overline{\xi\varepsilon} = n^{-1}\sum_{i=1}^n \xi_i \varepsilon_i$, and similar ones. We want to show that the suitably normalized ALS estimators converge in distribution to the normal law uniformly with respect to β, v and ξ_i's (see the definition and properties of uniform convergence in distribution in [5].

Introduce the class F_n of admissible design points $\xi^{(n)} = (\xi_1, \ldots, \xi_n)$. Fix $H > 0$ and a sequence $\{\alpha_n\}$, s.t. $\frac{1}{n} \leqslant \alpha_n \leqslant 1$, $n = 1, 2, \ldots$, and $\alpha_n \to 0$, $n \to \infty$. We set

$$F_n = \left\{ \xi^{(n)} \left| \frac{1}{n}\sum_1^n \xi_k^2 \geqslant H \text{ and } \frac{\max_{1\leqslant k\leqslant n} \xi_k^2}{\sum_1^n \xi_k^2} \leqslant \alpha_n \right. \right\}. \tag{3}$$

Sometimes we need additionally that

$$\liminf_{n\to\infty} \frac{\alpha_n \cdot n}{\ln n} > 2. \tag{4}$$

Under (4), for the r.v. $\tilde{\xi}_k = \tilde{H}^{1/2}\gamma_k$, where γ_k are i.i. $N(0, 1)$, and $\tilde{H} > H$, we have for all $n \geqslant n_0(\omega)$

$$\frac{1}{n}\sum_1^n \tilde{\xi}_k^2 > H$$

and

$$\frac{\max_{1\leqslant k\leqslant n} \tilde{\xi}_k^2}{\sum_1^n \tilde{\xi}_k^2} \cdot \frac{n}{\ln n} \xrightarrow{\text{P}} 2 \tag{5}$$

(see Appendix B). Therefore in this case $\tilde{\xi}^{(n)} = (\tilde{\xi}_1, \ldots, \tilde{\xi}_n) \in F_n$ with probability tending to 1 as $n \to \infty$.

Note also that under (4) for a sequence $\eta_n = n^c$, $c > 0$, or $\eta_n = \ln n, n = 1, 2, \ldots$, we have $\eta^{(n)} = (\eta_1, \ldots, \eta_n) \in F_n$, for all sufficiently large n.

Lemma 1. *Fix $K > 0$ and an interval $[v_1, v_2] \subset (0, +\infty)$. Then*

$$\frac{\overline{\xi^2}}{\sqrt{2\beta^2\sigma_\delta^4 + v\sigma_\delta^2 + \overline{\xi^2}(v + \beta\sigma_\delta^2)}} \sqrt{n}(\hat{\beta} - \beta) \to N(0, 1)$$

in distribution, uniformly with respect to $|\beta| \leqslant K$, $v \in [v_1, v_2]$ and $\xi^{(n)} \in F_n$, $n \geqslant 1$, where F_n is given in (3).

Corollary. *Let l be a bounded continuous function. Then*

$$E_{\beta v \xi^{(n)}} \left\{ l \left(\frac{\overline{\xi^2}\sqrt{n}(\hat{\beta} - \beta)}{\sqrt{2\beta^2\sigma_\delta^4 + v\sigma_\delta^2 + \overline{\xi^2}(v + \beta^2\sigma_\delta^2)}} \right) \right\} \to El(\gamma_1) \tag{6}$$

uniformly with respect to $|\beta| \leqslant K$, $v \in [v_1, v_2]$, $\xi^{(n)} \in F_n$, $n \geqslant 1$, where $\gamma_1 \sim N(0, 1)$.

UNCORRECTED PROOF

Hereafter $E_{\beta v \xi^{(n)}}$ denotes the expectation under the condition that β, v and $\xi^{(n)} = (\xi_1, \ldots, \xi_n)$ are the true values of unknown parameters in the regression model (1). The uniform convergence (6) is an immediate consequence of Lemma 1.

Remark. The convergence in distribution which is stated in Lemma 1, resembles the asymptotic normality result from [3, Theorem 2.3.2]. But there it was assumed that the entire error covariance structure of the linear errors-in-variables model is known up to a scalar multiple, and the maximum likelihood estimator was considered. We mention that estimator (2) is not the maximum likelihood estimator in model (1), with known measurement error variance and unknown variance of errors in the equation.

3. Asymptotic minimax bound

Here we follow the line of [4]. In that paper it was assumed that $v = 1$. But now we consider model (1) with Gaussian errors and **unknown** $v = D\varepsilon_1$. Denote by Λ the class of functions $l(x)$ on $\mathbb{R}$, such that $l(x) = l(-x) \geqslant 0$, $x \in \mathbb{R}$, and $l(x)$ is continuous at $x = 0$ and non-decreasing for $x > 0$.

Theorem 1. *Fix* $l \in \Lambda$, $\beta \in \mathbb{R}$, $v > 0$ *and the set* F_n *of designs given in* (3) *and* (4). *Then for every estimator* β_n *which is based on observations coming from model* (1),

$$\lim_{\delta \to 0} \liminf_{n \to \infty} \sup_{|b-\beta|<\delta, |w-v|<\delta, \xi^{(n)} \in F_n} E_{bw\xi^{(n)}} \{ l(H\sqrt{n}(\beta_n - b), \times (2\beta^2 \sigma_\delta^4 + v\sigma_\delta^2 + H(v + \beta^2 \sigma_\delta^2))^{-\frac{1}{2}}) \} \geqslant El(\gamma_1), \tag{7}$$

where $\gamma_1 \sim N(0,1)$.

4. Asymptotic efficiency of the ALS estimator

Suppose for a moment that in model (1) the variance v is known. Then the corresponding minimax bound can be obtained by a modification of Theorem 1 if the summand $2\beta^2\sigma_\delta^4$ under the root is cancelled, see [4]. The additional summand $2\beta^2\sigma_\delta^4$ in (7) reflects the lack of information in model (1) with unknown variance.

In the case of known v, the maximum likelihood estimator of β attains the corresponding bound, i.e., it is asymptotically efficient in that case. We show now that in the case of unknown v, the ALS estimator attains the bound, but we prove it for bounded loss functions only. Introduce the class $\Lambda_0 = \{l \in \Lambda$: l is bounded and has a finite number of jumps on each bounded interval$\}$. We give two examples of such loss functions: $l(x) = \min(|x|^r, R^r)$, $x \in \mathbb{R}$; $r > 0$, $R > 0$, and $l(x) = I_{[R,+\infty)}(|x|)$, $x \in \mathbb{R}$; $R > 0$.

Theorem 2. *Fix* $l \in \Lambda_0$, $\beta \in \mathbb{R}$, $v > 0$ *and the set* $F_n, n \geqslant 0$, *of designs given in* (3), (4). *Then for the ALS estimator* $\hat{\beta}$ *defined in* (2), *equality in* (7) *holds.*

Thus we showed for model (1) that the ALS estimator $\hat{\beta}$ is asymptotically efficient in the sense of Hájek bound, i.e., $\hat{\beta}$ attains the minimax bound (7). This means that, under suitable normalization, $\hat{\beta}$ has the least possible averaged losses from imprecise estimation of β. The estimator $\hat{\beta}$ is locally asymptotically minimax at the true value point (β, v), i.e., $\hat{\beta}$ approaches the asymptotically minimax estimator in the neighborhood $U_\delta(\beta, v) = \{(b, w)\colon\ |b - \beta| < \delta, |w - v| < \delta\}$, as $\delta \to 0$.

5. Conclusion

We considered a linear functional errors-in-variables model (1) with unknown slope parameter and normal errors. We supposed that the measurement error variance is known while the variance of errors in the equation is unknown. We proved that the adjusted least-squares estimator (2) is asymptotically efficient in the sense of Hájek bound, i.e., the estimator attains the minimax bound established in Theorem 1. This result is a follow-up to Nussbaum [6] and to Hasminskii and Ibragimov [4]. In that papers it was shown that in model (1) with both variances known, the corresponding bound of Hájek type is attained by the maximum likelihood estimator.

Acknowledgments

We are grateful to Silvelyn Zwanzig for a good advice. We thank two anonymous referees for valuable comments.

Appendix A. Auxiliary matrix calculations

Consider a random vector $X \sim N(0, \Sigma)$ with probability density p.

Lemma A.1. *Suppose that the entries of* Σ *are* C^2*-smooth functions of* χ, *and* χ *belongs to an open set* $G \subset \mathbb{R}^d$, *and for each* $\chi \in G$ *the covariance matrix* Σ *is non-singular. Then for each* $i, k = 1, 2, \ldots, d$,

$$-2\mathrm{E}\,\frac{\partial^2 \ln p}{\partial \chi_k \partial \chi_i} = \mathrm{tr}\left(\Sigma^{-1}\frac{\partial \Sigma}{\partial \chi_i}\Sigma^{-1}\frac{\partial \Sigma}{\partial \chi_k}\right).$$

UNCORRECTED PROOF

Proof. Let X be distributed in $\mathbb{R}^m$. Denote

$$l(x,\chi) = \ln p = -\frac{m}{2}\ln(2\pi) - \frac{1}{2}\ln\det\Sigma - \frac{1}{2}x'\Sigma^{-1}x.$$

We have

$$\frac{\partial l}{\partial\chi_i} = -\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\right) + \frac{1}{2}x'\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\Sigma^{-1}x,$$

and

$$\begin{aligned}2\frac{\partial^2 l}{\partial\chi_k\partial\chi_i} = &\operatorname{tr}\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_k}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\right) - \operatorname{tr}\left(\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\chi_k\partial\chi_i}\right)\\ &- x'\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_k}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\Sigma^{-1}x + x'\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\chi_k\partial\chi_i}\Sigma^{-1}x\\ &- x'\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_k}\Sigma^{-1}x.\end{aligned}$$

Now,

$$\begin{aligned}2\mathrm{E}\frac{\partial^2 l}{\partial\chi_k\partial\chi_i} = &\operatorname{tr}\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_k}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\right) - \operatorname{tr}\left(\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\chi_k\partial\chi_i}\right)\\ &- \operatorname{tr}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_k}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\Sigma^{-1}Exx'\\ &+ \operatorname{tr}\Sigma^{-1}\frac{\partial^2\Sigma}{\partial\chi_k\partial\chi_i}\Sigma^{-1}Exx' - \operatorname{tr}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_k}\Sigma^{-1}Exx'\\ = &-\operatorname{tr}\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_k}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\right).\end{aligned}$$

This proves Lemma A.1. □

Now, consider random variables, corresponding to model (1). Let

$$\delta \sim N(0,\sigma_\delta^2), \quad \varepsilon \sim N(0,v), \quad \xi \sim N(0,H)$$

be independent r.v. with positive variances, and

$$x = \xi + \delta, \quad y = \beta\xi + \varepsilon.$$

Here β is a fixed real parameter. Then

$$(x,y)' \sim N(0,\Sigma), \quad \Sigma = \begin{pmatrix} H+\sigma_\delta^2 & H\beta \\ H\beta & H\beta^2+v \end{pmatrix} \tag{A.1}$$

and

$$\det\Sigma = \Delta = H\beta^2\sigma_\delta^2 + (H+\sigma_\delta^2)v, \quad \Delta > 0,$$

$$\Sigma^{-1} = \frac{1}{\Delta}\begin{pmatrix} H\beta^2+v & -H\beta \\ -H\beta & H+\sigma_\delta^2 \end{pmatrix}. \tag{A.2}$$

UNCORRECTED PROOF

The probability density p of random vector (x, y) depends upon the parameter $\chi = (\beta, H, v) \in \mathbb{R} \times (0, +\infty) \times (0, +\infty)$.

Lemma A.2. *The Fisher information matrix I of the density $p(x, y; \chi)$ has the form*

$$2I = \frac{1}{\Delta^2} A$$

with

$$A = \begin{pmatrix} 2H^2(\Delta + 2\beta^2\sigma_\delta^4) & 2H\beta\sigma_\delta^2(v + \beta^2\sigma_\delta^2) & 2H\sigma_\delta^2(H + \sigma_\delta^2)\beta \\ 2H\beta\sigma_\delta^2(v + \beta^2\sigma_\delta^2) & (v + \beta^2\sigma_\delta^2)^2 & \beta^2\sigma_\delta^4 \\ 2H\sigma_\delta^2(H + \sigma_\delta^2)\beta & \beta^2\sigma_\delta^4 & (H + \sigma_\delta^2)^2 \end{pmatrix}.$$

Proof. By Lemma A.1

$$(2I)_{ik} = -2\mathrm{E}\frac{\partial^2 \ln p}{\partial\chi_k\partial\chi_i} = \mathrm{tr}\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_i}\Sigma^{-1}\frac{\partial\Sigma}{\partial\chi_k}\right). \tag{A.3}$$

Using (A.1) and (A.2) we have consequently

$$\Sigma^{-1}\frac{\partial\Sigma}{\partial\beta} = \frac{H}{\Delta}\begin{pmatrix} -H\beta & v - H\beta^2 \\ H + \sigma_\delta^2 & \beta(H + 2\sigma_\delta^2) \end{pmatrix},$$

$$\Sigma^{-1}\frac{\partial\Sigma}{\partial H} = \frac{1}{\Delta}\begin{pmatrix} v & v\beta \\ \beta\sigma_\delta^2 & \beta^2\sigma_\delta^2 \end{pmatrix},$$

$$\Sigma^{-1}\frac{\partial\Sigma}{\partial v} = \frac{1}{\Delta}\begin{pmatrix} 0 & -H\beta \\ 0 & H + \sigma_\delta^2 \end{pmatrix}.$$

Now, a direct calculation of the entries of $2I$ using (A.3) accomplishes the proof. □

Lemma A.3. *The Fisher information matrix I of the density $p(x, y; \chi)$ is non-singular, and*

$$(I^{-1})_{11} = \frac{\Delta + 2\beta^2\sigma_\delta^4}{H^2} = \frac{2\beta^2\sigma_\delta^4 + v\sigma_\delta^2 + H(v + \beta^2\sigma_\delta^2)}{H^2}.$$

Proof. Due to Lemma A.2 we have to show non-singularity of A and to compute $(A^{-1})_{11}$. Find the algebraic complements in A for the entries of the top row.

$$A_{11} = \Delta(\Delta + 2\beta^2\sigma_\delta^4), \quad A_{12} = -2H\sigma_\delta^2(H + \sigma_\delta^2)\beta\Delta,$$

$$A_{13} = -2H\beta\sigma_\delta^2(v + \beta^2\sigma_\delta^2)\Delta.$$

Then

UNCORRECTED PROOF

$$\begin{aligned}\det A &= a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}\\ &= 2H^2(\varDelta + 2\beta^2\sigma_\delta^4)^2\varDelta - 4H^2\beta^2\sigma_\delta^4(v + \beta^2\sigma_\delta^2)(H + \sigma_\delta^2)\varDelta\\ &\quad - 4H^2\sigma_\delta^4(H + \sigma_\delta^2)\beta^2(v + \beta^2\sigma_\delta^2)\varDelta\\ &= 2H^2\varDelta[(\varDelta + 2\beta^2\sigma_\delta^4)^2 - 4\beta^2\sigma_\delta^4(H + \sigma_\delta^2)(v + \beta^2\sigma_\delta^2)]\\ &= 2H^2\varDelta[(\varDelta + 2\beta^2\sigma_\delta^4)^2 - 4\beta^2\sigma_\delta^4(\varDelta + \beta^2\sigma_\delta^4)]\\ &= 2H^2\varDelta^3,\end{aligned}$$

and $\det A > 0$. Therefore I is also non-singular. At last

$$(A^{-1})_{11} = \frac{A_{11}}{\det A} = \frac{\varDelta + 2\beta^2\sigma_\delta^4}{2H^2\varDelta^2}$$

and by Lemma A.2

$$(I^{-1})_{11} = 2\varDelta^2(A^{-1})_{11} = \frac{\varDelta + 2\beta^2\sigma_\delta^4}{H^2}. \qquad \square$$

Appendix B. Proof of convergence (5)

Lemma B.1. *Let* γ_k, $k = 1, 2, \dots$ *be i.i.* $N(0,1)$ *distributed random values. Then*

$$\frac{\max_{1\leqslant k\leqslant n}\gamma_k^2}{\sum_1^n \gamma_k^2}\frac{n}{\ln n} \xrightarrow{\mathrm{P}} 2. \tag{B.1}$$

Proof. As $\sum_1^n \gamma_k^2/n \to 1$, a.s., (B.1) is equivalent to

$$\eta_n = \frac{\max_{1\leqslant k\leqslant n}\gamma_k^2}{\ln n} \xrightarrow{\mathrm{P}} 2. \tag{B.2}$$

Find the d.f. of η_n. A r.v. γ_k^2 has a d.f. F with density

$$f(t) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{t}{2}}, \quad t > 0.$$

$\max_{1\leqslant k\leqslant n}\gamma_k^2$ has a d.f. $F^n(t)$, and the d.f. F_n of η_n is given by

$$F_n(t) = F^n(t\ln n), \quad n \geqslant 2.$$

To prove (B.2) it is sufficient to show that for each $\varepsilon > 0$,

$$F_n(2-\varepsilon) \to 0, \quad \text{and} \quad F_n(2+\varepsilon) \to 1, \quad n \to \infty. \tag{B.3}$$

For $t > 0$ we have

UNCORRECTED PROOF

YJMVA : 2156

ARTICLE IN PRESS

$$F_n(t) = [1 - (1 - F(t \ln n))^{\frac{1}{1-F(t \ln n)}}]^{n(1-F(t \ln n))} = A_n^{B_n}.$$

Here $A_n \to e^{-1}$, $n \to \infty$, and

$$\begin{aligned}\lim_{n\to\infty} B_n &= \lim_{z\to+\infty} \frac{1 - F(t \ln z)}{1/z} = \lim_{z\to+\infty} \frac{-f(t \ln z)\frac{t}{z}}{-1/z^2} \\ &= \frac{\sqrt{t}}{\sqrt{2\pi}} \lim_{z\to+\infty} \frac{z}{\sqrt{\ln z}} e^{-\frac{t}{2}\ln z} \\ &= \sqrt{\frac{t}{2\pi}} \lim_{z\to+\infty} \frac{z^{1-\frac{t}{2}}}{\sqrt{\ln z}} = \begin{cases} +\infty & \text{if } t<2, \\ 0 & \text{if } t>2. \end{cases}\end{aligned}$$

Therefore $A_n^{B_n} \to 0$ if $t<2$, and $A_n^{B_n} \to 1$ if $t>2$. This proves (B.3), and (B.2) holds true. Lemma is proved. □

Appendix C. Proof of the main results

C.1. Proof of Lemma 1

Substituting (1) in (2), we have that

$$\overline{\xi^2}\sqrt{n}(\hat\beta - \beta) = \frac{\sqrt{n}[-\beta(\overline{\xi\delta} + \overline{\delta^2} - \sigma_\delta^2) + \overline{\xi\varepsilon} + \overline{\delta\varepsilon}]}{1 + \frac{2\overline{\xi\delta}}{\overline{\xi^2}} + \frac{\overline{\delta^2} - \sigma_\delta^2}{\overline{\xi^2}}}. \tag{C.1}$$

Consider firstly the denominator

$$E\left(\frac{\overline{\xi\delta}}{\overline{\xi^2}}\right)^2 = \frac{\sigma_\delta^2}{n\overline{\xi^2}},$$

but $\frac{1}{\overline{\xi^2}} \leqslant \frac{1}{H}$ for $\xi^{(n)} \in F_n$, therefore $\overline{\xi\delta}/\overline{\xi^2}$ converges to 0 in probability uniformly for $\xi^{(n)} \in F_n$, $n \geqslant 1$. And

$$E\left(\frac{\overline{\delta^2} - \sigma_\delta^2}{\overline{\xi^2}}\right)^2 \to 0, \quad n \to \infty$$

uniformly for $\xi^{(n)} \in F_n$, therefore $(\overline{\delta^2} - \sigma_\delta^2)/\overline{\xi^2}$ also converges to 0 in probability uniformly for $\xi^{(n)} \in F_n$. Thus the denominator of (C.1) converges to 1 in probability uniformly for $\xi^{(n)} \in F_n$. To prove Lemma 1, it is enough to show that

$$\frac{\frac{1}{\sqrt{n}} \sum_1^n [-\beta(\xi_i\delta_i + \delta_i^2 - \sigma_\delta^2) + \xi_i\varepsilon_i + \delta_i\varepsilon_i]}{\sqrt{2\beta^2\sigma_\delta^4 + v\sigma_\delta^2 + \overline{\xi^2}(v + \beta^2\sigma_\delta^2)}} \to N(0,1) \tag{C.2}$$

UNCORRECTED PROOF

uniformly for $|\beta|\leqslant K,\ v\in[v_1,v_2],\ \xi^{(n)}\in F_n$. Denote

$$\varphi_i = -\beta(\xi_i\delta_i+\delta_i^2-\sigma_\delta^2)+\xi_i\varepsilon_i+\delta_i\varepsilon_i,\quad i\geqslant 1.$$

Then $E\varphi_i=0,\ B_n^2=\sum_1^n D\varphi_i=[2\beta^2\sigma_\delta^4+v\sigma_\delta^2+\overline{\xi^2}(v+\beta^2\sigma_\delta^2)]n.$

We bound Liapunov's ratio

$$\frac{1}{(B_n^2)^{3/2}}\sum_1^n E|\varphi_i|^3. \tag{C.3}$$

For $|\beta|\leqslant K,\ v\in[v_1,v_2],\ \xi^{(n)}\in F_n$ consider for instance the moments of the first summand of φ_i.

$$(B_n^{-2})^{3/2}\sum_1^n E|\beta\xi_i\delta_i|^3\leqslant\frac{\text{const}}{(\sum_1^n\xi_i^2)^{3/2}}\sum_1^n|\xi_i|^3$$
$$\leqslant\text{const}\left(\frac{\max_{1\leqslant i\leqslant n}\xi_i^2}{\sum_1^n\xi_i^2}\right)^{1/2}\leqslant\text{const}\ \alpha_n^{1/2},$$

with α_n given in (3). Similar calculations for other summands of φ_i show that Liapunov's ratio (C.3) converges to 0 uniformly for $|\beta|\leqslant K,\ v\in[v_1,v_2],\ \xi^{(n)}\in F_n$. Now, by Ibragimov and Has'minskii [5, Theorem 15, p. 369] uniform convergence (C.2) holds.

This implies the statement of Lemma 1. □

C.2. *Proof of Theorem 1*

Let H be a lower bound from (3). Consider the sequence $\{\tilde{\xi}_i\}$ of i.i.d. $(0,\tilde{H})$-normal random variables, independent of $\{\delta_i,\varepsilon_i;\ i\geqslant 1\}$. The variance $\tilde{H}$ is unknown, we know only that $\tilde{H}>H$. In Section 2 it was mentioned that $\tilde{\xi}^{(n)}=(\tilde{\xi}_1,\ldots,\tilde{\xi}_n)\in F_n,\ n\geqslant n_0(w)$. Now, consider the problem of estimation of the parameters $\tilde{H},\beta,v$ on the basis of independent observations

$$x_i=\tilde{\xi}_i+\delta_i,\quad y_i=\beta\tilde{\xi}_i+\varepsilon_i,\quad i=1,\ldots,n. \tag{C.4}$$

Denote

$$\tilde{\Delta}=\tilde{H}v+\tilde{H}\beta^2\sigma_\delta^2+v\sigma_\delta^2. \tag{C.5}$$

Observations (C.4) are Gaussian with density function

$$p(x,y;\beta,\tilde{H},v)$$
$$=\frac{1}{2\pi\sqrt{\tilde{\Delta}}}\exp\left[-\frac{1}{2\tilde{\Delta}}\{x^2(\beta^2\tilde{H}+v)-2xy\beta\tilde{H}+y^2(\tilde{H}+\sigma_\delta^2)\}\right].$$

The Fisher information matrix I of the density has the form (see Lemma A.2)

$$2I=\frac{1}{\tilde{\Delta}^2}\tilde{A},$$

UNCORRECTED PROOF

where $\tilde{A}$ is obtained from the expression in Lemma A.2 by substituting there $\tilde{H}$ instead of H and $\tilde{\Delta}$ instead of Δ.

As $\det \tilde{A} > 0$ (see Lemma A.3 in Appendix A), observations (C.4) satisfy Le Cam's LAN conditions with the norming factors $n^{-1/2}I^{-1/2}$.

Denote $z = (1,\, 0,\, 0)'$. We are interested in

$$
\begin{aligned}
&(I^{-1/2})^2_{11} + (I^{-1/2})^2_{12} + (I^{-1/2})^2_{13} \\
&\quad = (I^{-1/2}z,\, I^{-1/2}z) = (I^{-1}z,\, z) = (I^{-1})_{11} = \frac{\tilde{\Delta} + 2\beta^2\sigma_\delta^4}{\tilde{H}^2},
\end{aligned} \tag{C.6}
$$

see Lemma A.3. Consider the class of bounded loss functions

$$
\Lambda_b = \{ l \in \Lambda\colon\ l \text{ is bounded} \}
$$

and let

$$
B = \frac{(\tilde{\Delta} + 2\beta^2\sigma_\delta^4)I}{H^2},
$$

with $\tilde{\Delta}$ given in (C.5). We apply [5, Theorem 12.1, p. 162] to model (C.4) with $\varepsilon = 1/n,\ \varphi(\varepsilon) = (nI)^{-1/2}$, and the loss function

$$
w(x) = l((B^{-1/2}x)_1), \quad l \in \Lambda_b,
$$

where $(x)_1$ denotes the first component of the vector $x \in \mathbb{R}^3$. We mention that $w(\varphi^{-1}(\varepsilon)x) = l(H\sqrt{n}x_1(\tilde{\Delta} + 2\beta^2\sigma_\delta^4)^{-1/2})$, and by (C.6) we have

$$
\sum_1^3 (B^{-1/2})^2_{1i} = (B^{-1})_{11} = \left(\frac{H}{\tilde{H}}\right)^2. \tag{C.7}
$$

Taking into account (C.7) we find that for every estimators β_n,

$$
\begin{aligned}
&\lim_{\delta\to 0}\ \liminf_{n\to\infty}\ \sup_{(|b-\beta|<\delta,|w-v|<\delta,|h-\tilde{H}|<\delta)} E_{bwh}\{ l(H\sqrt{n}(\beta_n - b), \\
&(\tilde{H}(v + \beta^2\sigma_\delta^2) + v\sigma_\delta^2 + 2\beta^2\sigma_\delta^4)^{-\frac{1}{2}})\} \geqslant E\{ l((B^{-1/2}\gamma)_1)\} = E\, l\left(\frac{H}{\tilde{H}}\gamma_1\right).
\end{aligned} \tag{C.8}
$$

Here γ is a standard normal random vector in $\mathbb{R}^3$, and E_{bwh} denotes the expectation under the condition that in model (C.4) $\beta = b,\ D\varepsilon_1 = w,\ D\tilde{\xi}_1 = h$.

For $l \in \Lambda_b$, denote $||l|| = \sup_{t\in\mathbb{R}} l(t)$, and

$$
g_n = l(H \cdot \sqrt{n}(\beta_n - b) \cdot (2\beta^2\sigma_\delta^4 + v\sigma_\delta^2 + H(v + \beta^2\sigma_\delta^2))^{-\frac{1}{2}}).
$$

For $\delta_0 = \frac{\tilde{H}-H}{2}$, $P_h\{\tilde{\xi}^{(n)} \in F_n\} \to 1,\ n \to \infty$ uniformly for $h \in (\tilde{H} - \delta_0, \tilde{H} + \delta_0)$. We have now for $\delta \leqslant \delta_0$,

UNCORRECTED PROOF

$$\sup_{(|h-\tilde{H}|<\delta)} E_{bwh}(g_n)$$
$$\leqslant \sup_{(|h-\tilde{H}|<\delta)} E_{bwh}\left\{g_n I(\tilde{\xi}^{(n)} \in F_n)\right\} + \sup_{(|h-\tilde{H}|<\delta)} E_{bwh}\{g_n I(\tilde{\xi}^{(n)} \bar{\in} F_n)\}$$
$$\leqslant \sup_{(\xi^{(n)} \in F_n)} E_{bw\xi^{(n)}}(g_n) + ||l|| \cdot \sup_{(|h-\tilde{H}|<\delta_0)} P_h\{\tilde{\xi}^{(n)} \bar{\in} F_n\}.$$

For $\delta \leqslant \delta_0$, this implies that

$$\sup_{(|b-\beta|<\delta, |w-v|<\delta, \xi^{(n)} \in F_n)} E_{bw\xi^{(n)}}(g_n)$$
$$\geqslant \sup_{(|b-\beta|<\delta, |w-v|<\delta, |h-\tilde{H}|<\delta)} E_{bwh}(g_n) - o(1), \quad n \to \infty. \tag{C.9}$$

Then the following chain of inequalities holds, see (C.9)

$$Q = \lim_{\delta\to 0} \liminf_{n\to\infty} \sup_{(|b-\beta|<\delta, |w-v|<\delta, \xi^{(n)} \in F_n)} E_{bw\xi^{(n)}}(g_n)$$
$$\geqslant \lim_{\delta\to 0} \liminf_{n\to\infty} \sup_{(|b-\beta|<\delta, |w-v|<\delta, |h-\tilde{H}|<\delta)} E_{bwh}(g_n)$$
$$\geqslant \lim_{\delta\to 0} \liminf_{n\to\infty} \sup_{(|b-\beta|<\delta, |w-v|<\delta, |h-\tilde{H}|<\delta)} E_{bwh}\{l(H \cdot \sqrt{n}(\beta_n - b)\cdot$$
$$(2\beta^2\sigma_\delta^4 + v\sigma_\delta^2 + \tilde{H}(v + \beta^2\sigma_\delta^2))^{-\frac{1}{2}})\}.$$

Here we used the properties of l and the inequality $H < \tilde{H}$. By (C.8) we have

$$Q \geqslant E\, l\left(\frac{H}{\tilde{H}}\gamma_1\right) \to E\, l(\gamma_1), \quad \tilde{H} \to H+.$$

We proved (7) for all bounded $l \in \Lambda$. At last consider an unbounded loss function $f \in \Lambda$. Denote by $\phi(l)$, $l \in \Lambda$, the left-hand side of inequality (7) and by f_c the truncated function $f_c(t) = \min(c, f(t))$, $t \in \mathbb{R}$; $c > 0$. The function f_c belongs to the class Λ_b, therefore

$$\phi(f) \geqslant \phi(f_c) \geqslant E f_c(\gamma_1) \to E f(\gamma_1), \quad c \to +\infty.$$

This proves Theorem 1. □

C.3. *Proof of Theorem 2*

First we show that for $l \in \Lambda_0$, $\beta \in \mathbb{R}$, and $v > 0$,

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \sup_{(|b-\beta|<\delta, |w-v|<\delta, \xi^{(n)} \in F_n)} E_{bw\xi^{(n)}}\{l(\overline{\xi^2}\sqrt{n}(\hat{\beta}_n - b),$$
$$(2\beta^2\sigma_\delta^4 + v\sigma_\delta^2 + \overline{\xi^2}(v + \beta^2\sigma_\delta^2))^{-\frac{1}{2}})\} \leqslant E\, l(\gamma_1). \tag{C.10}$$

We follow the line of [5, p. 366]. Denote $\hat{\Delta}_n = \sqrt{n\overline{\xi^2}}(\hat{\beta}_n - b)$,

UNCORRECTED PROOF

$k(\beta, v, z) = (z(2\beta^2\sigma_\delta^4 + v\sigma_\delta^4) + v + \beta\sigma_\delta^2)^{-1/2}$. Then the argument of l on the left-hand side of (C.10) equals $\hat{\Delta}_n \cdot k(b, v, 1/\overline{\xi^2})$. Let $\delta_0 = v/2$. By Lemma 1

$$\hat{\Delta}_n \cdot k(b, w, 1/\overline{\xi^2}) \to N(0, 1)$$

in distribution $P_{bw\xi^{(n)}}$, uniformly with respect to $b \in [\beta - \delta_0, \beta + \delta_0]$, $w \in [v - \delta_0, v + \delta_0]$ and $\xi^{(n)} \in F_n$, $n \geqslant 1$. Therefore for certain n_0, the random variables $\{\hat{\Delta}_n \cdot k(b, w, 1/\overline{\xi^2})$: $b \in [\beta - \delta_0, \beta + \delta_0]$, $w \in [v - \delta_0, v + \delta_0]$, and $\xi^{(n)} \in F_n$, $n \geqslant n_0\}$ are stochastically bounded. Then for given $\varepsilon > 0$ it is possible to construct an open set G_0 containing all the jump points of l, such that

$$\limsup_{n\to\infty} \sup_{(|b-\beta|<\delta_0,\ |w-v|<\delta_0, \xi^{(n)}\in F_n)} P_{bw\xi^{(n)}}(\hat{\Delta}_n \cdot k(b, w, 1/\overline{\xi^2}) \in G_0) \leqslant \varepsilon. \quad \text{(C.11)}$$

The function $l \in \Lambda_0$ has finite number of jumps on each bounded interval, let $\{t_k,\ k \geqslant 1\}$ be the jump points of l. We can choose the set G_0 satisfying (C.11) of the form

$$G_0 = \bigcup_{k\geqslant 1} (t_k - \tau_0, t_k + \tau_0),$$

with certain $\tau_0 > 0$. Now, the function $k(b, w, z)$ is uniformly continuous on a compact set $K = [\beta - \delta_0, \beta + \delta_0] \times [v - \delta_0, v + \delta_0] \times [0, H^{-1}]$, and $1/\overline{\xi^2} \leqslant H^{-1}$, for $\xi^{(n)} \in F_n$; we mention that under true values b, w and $\xi^{(n)}$, the variables

$$\{\hat{\Delta}_n\colon b \in [\beta - \delta_0,\ \beta + \delta_0],\ w \in [v - \delta_0, v + \delta_0],\ \xi^{(n)} \in F_n,\ n \geqslant n_0\}$$

are stochastically bounded. Then for any $\tau_1 \in (0, \tau_0)$ and $G = \bigcup_{k\geqslant 1} (t_k - \tau_1, t_k + \tau_1)$ we induce from (C.11) that

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{(|b-\beta|<\delta, |w-v|<\delta, \xi^{(n)}\in F_n)} P_{bw\xi^{(n)}}(\hat{\Delta}_n \cdot k(b, v, 1/\overline{\xi^2}) \in G) \leqslant \varepsilon.$$

There is only finite number of $\{t_k\}$ on each bounded interval, therefore we may and do assume that $P(\gamma_1 \in G) \leqslant \varepsilon$, where $\gamma_1 \sim N(0, 1)$.

The function l is continuous on the closed set $\mathbb{R}\backslash G$, and it is possible to construct a function $\tilde{l} \in C(\mathbb{R})$, such that $\tilde{l}(x) = l(x)$, $x \in \mathbb{R}\backslash G$, and $0 \leqslant \tilde{l}(x) \leqslant \sup_{t\in\mathbb{R}} l(t)$, $x \in \mathbb{R}$. By corollary of Lemma 1,

$$E_{bw\xi^{(n)}}\{\tilde{l}(\hat{\Delta}_n \cdot k(b, w, 1/\overline{\xi^2}))\} \to E\tilde{l}(\gamma_1)$$

uniformly with respect to $b \in [\beta - \delta_0, \beta + \delta_0]$, $w \in [v - \delta_0, v + \delta_0]$, $\xi^{(n)} \in F_n$, $n \geqslant 1$. Then we have

UNCORRECTED PROOF

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \sup_{(|b-\beta|<\delta,|w-v|<\delta,\xi^{(n)}\in F_n)} E_{bw\xi^{(n)}}\{l(\hat{\Delta}_n \cdot k(\beta, v, 1/\overline{\xi^2}))\}$$

$$\leqslant \lim_{\delta\to 0} \liminf_{n\to\infty} \sup_{(|b-\beta|<\delta,|w-v|<\delta,\xi^{(n)}\in F_n)} E_{bw\xi^{(n)}}\{\tilde{l}(\hat{\Delta}_n \cdot k(b, w, 1/\overline{\xi^2}))$$

$$+ |l-\tilde{l}|(\hat{\Delta}_n \cdot k(\beta, v, 1/\overline{\xi^2})) + |\tilde{l}(\hat{\Delta}_n \cdot k(\beta, v, 1/\overline{\xi^2})) - \hat{l}(\hat{\Delta}_n \cdot k(b, w, 1/\overline{\xi^2}))|\}$$

$$\leqslant E\tilde{l}(\gamma_1) + \varepsilon \cdot \sup_{t\in\mathbb{R}} l(t) \leqslant El(\gamma_1) + 2\varepsilon \sup_{t\in\mathbb{R}} l(t).$$

Here the term containing $|\tilde{l}(\hat{\Delta}_n \cdot k(\beta, v, 1/\overline{\xi^2})) - \hat{l}(\hat{\Delta}_n \cdot k(b, w, 1/\overline{\xi^2}))|$ tends to 0, because $\hat{\Lambda}_n$'s are stochastically bounded, $\tilde{l}$ is bounded and continuous, and the function $k(b, w, z)$ is uniformly continuous on K. Thus we proved (C.10).

The function $\varphi(u) = \frac{u}{\sqrt{A^2+u^2}}$, $u \geqslant 0$, is increasing, and $\overline{\xi^2} \geqslant H$, for $\xi^{(n)} \in F_n$. Therefore from (C.10) we get

$$\lim_{\delta\to 0} \liminf_{n\to\infty} \sup_{(|b-\beta|<\delta,|w-v|<\delta,\xi^{(n)}\in F_n)} E_{bw\xi^{(n)}}\{l(H\sqrt{n}(\hat{\beta}_n - b),$$

$$(2\beta^2\sigma_\delta^4 + v\sigma_\delta^2 + H(v + \beta^2\sigma_\delta^2))^{-\frac{1}{2}})\} \leqslant E\, l(\gamma_1). \qquad \text{(C.12)}$$

But it follows from (7) that actually in (C.12) equality holds, and the theorem is proved. □

References

[1] C.-L. Cheng, H. Schneeweiss, Polynomial regression with errors in the variables, J. R. Statist. Soc. B 60 (1998) 189–199.

[2] C.-L. Cheng, J.W. Van Ness, Statistical Regression with Measurement Error, Arnold, London, 1999.

[3] W.A. Fuller, Measurement Error Models, Wiley, New York, 1987.

[4] R.Z. Hasminskii, I.A. Ibragimov, On Asymptotic Efficiency in the Presence of an Infinite-dimensional Nuisance Parameter, Springer Lecture Notes in Mathematics, Vol. 1021, Springer, Berlin, 1983, pp. 195–229.

[5] I.A. Ibragimov, R.Z. Has'minskii, Statistical Estimation: Asymptotic Theory, Springer, Berlin, 1981.

[6] M. Nussbaum, An asymptotic minimax risk bound for estimation of a linear functional relationship, J. Multivariate Anal. 14 (3) (1984) 300–314.

UNCORRECTED PROOF