Discussiones Mathematicae
Probability and Statistics 29 (2009) 223-231

A NOTE ON THE STRONG CONSISTENCY OF LEAST SQUARES ESTIMATES

João Lita da Silva

Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa
Quinta da Torre, 2829–516 Caparica, Portugal

e-mail: jfls@fct.unl.pt

Abstract

The strong consistency of least squares estimates in multiples regression models with i.i.d. errors is obtained under assumptions on the design matrix and moment restrictions on the errors.

Keywords: least squares estimates, linear models, strong consistency.

2000 Mathematics Subject Classification: 60F15.

1. Introduction

Many statisticians have considered the problem of strong consistency of the least squares estimates in multiple regression models. In the seventies, this problem was completely solved under weak moment conditions on the errors, namely, assuming their finite variance (see [5, 6] and [9]). More recently, some other authors had studied this same problem for the case where the variance of the errors is infinite. Nevertheless, these works reveals very restrictive conditions on the design matrix (see [8]) or particular scenarios for the errors (see [11, 12] or [13]).

In this paper, we establish the strong consistency of the least squares estimates for the parameters β_j of the multiple regression model

$$y_i = \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + \varepsilon_i \qquad (i = 1, 2, \ldots) \tag{1.1}$$

under suitable assumptions on the design matrix x_{ij} when the error variance is infinite. Specially, we shall assume that

$$\begin{aligned} &\varepsilon_i \text{ are i.i.d. with } \mathbb{E}\,|\varepsilon_1|^r < \infty \quad \text{for some } \; r \in (0,2) \\ &\text{and } \; \mathbb{E}\,\varepsilon_1 = 0 \; \text{ whenever } \; r \in (1,2) \end{aligned} \tag{1.2}$$

admitting cases where the errors don't have mean value. Let us stress that in [6] or [7] only the errors ε_i with $\sup_i \mathbb{E}\,|\varepsilon_i|^r < \infty$ for some $1 \leqslant r < 2$ are considered leaving the cases where $0 < r < 1$ unsolved. In particular, on the paper [6], the authors establish the strong consistency of the least squares estimates for the case $1 \leqslant r < 2$ using the Hölder inequality (see Corollary 3 and Lemma 4), which is no more useful when $0 < r < 1$.

Throughout this work, we shall let $\mathbf{X}_n$ denote the design matrix

$$\left(x_{ij}\right)_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$$

and let $\mathbf{y}_n = (y_1, \ldots, y_n)'$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$, where prime denotes transpose. For $n \geqslant p$, the least squares estimate $\mathbf{b}_n = (b_{n1}, \ldots, b_{np})'$ of the vector $\boldsymbol{\beta}$ based on the design matrix $\mathbf{X}_n$ and the response vector $\mathbf{y}_n$ is given by

$$\mathbf{b}_n = (\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{X}_n'\mathbf{y}_n \tag{1.3}$$

provided that

$$\mathbf{X}_n'\mathbf{X}_n = \begin{pmatrix} \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} & \ldots & \sum_{i=1}^n x_{i1}x_{ip} \\ \sum_{i=1}^n x_{i1}x_{i2} & \sum_{i=1}^n x_{i2}^2 & \ldots & \sum_{i=1}^n x_{i2}x_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{i1}x_{ip} & \sum_{i=1}^n x_{i2}x_{ip} & \ldots & \sum_{i=1}^n x_{ip}^2 \end{pmatrix}$$

is nonsingular for all $n \geqslant n_0$. From the expression of $\mathbf{b}_n$ it follows that the strong consistency of the least squares estimates is equivalent to

$$(\mathbf{X}_n'\mathbf{X}_n)^{-1}\sum_{i=1}^{n}\mathbf{x}_i\varepsilon_i \xrightarrow{\text{a.s.}} 0$$

where $\mathbf{x}_i = (x_{i1},\ldots,x_{ip})'$.

2. Auxiliary tools

It is well-known that every positive definite matrix $\mathbf{A} = \left(a_{ij}\right)_{1\leqslant i,j\leqslant p}$ satisfies

$$\det(\mathbf{A}) \leqslant a_{11}\ldots a_{pp}$$

and the equality holds if and only if $\mathbf{A}$ is diagonal (see [3], page 477). This classical result due to Hadamard leads us to the following definition.

Definition. A sequence of $p\times p$ positive definite matrices $\mathbf{A}_n = \left(a_{ij}^{(n)}\right)_{1\leqslant i,j\leqslant p}$ is said *asymptotically diagonal dominant* if $\det(\mathbf{A}_n) \asymp a_{11}^{(n)}\ldots a_{pp}^{(n)}$, $n\to\infty$.*

An important tool in proving the strong consistency of $\mathbf{b}_n$ for error structures satisfying (1.2) is the next lemma which is an extension of Marcinkiewicz-Zygmund theorem presented in [1] (page 118).

Lemma 1. *If $\{X_n\}$ are i.i.d. r.v.'s with $\mathbb{E}\,|X_1|^r < \infty$ and $\{a_n\}$ are real numbers such that $a_n = O\left(n^{-1/r}\right)$ for some $0<r<2$ then*

$$\sum_{n=1}^{\infty}(a_nX_n - \mathbb{E}\,Y_n)$$

converges a.s., where $Y_n = a_nX_nI_{\{|X_n|\le n^{1/r}\}}$. Furthermore, if either (i) $0<r<1$ *or* (ii) $1<r<2$ *and $\mathbb{E}\,X_1 = 0$, then $\sum_{n=1}^{\infty}a_nX_n$ converges a.s.*

*$a_n \asymp b_n$, $n\to\infty$ means that $a_n = O(b_n)$ and $b_n = O(a_n)$ as $n\to\infty$.

Proof. Set $A_j = \left\{(j-1)^{1/r} \le |X_1| \le j^{1/r}\right\}$, $j \ge 1$. Then for $\alpha > r > 0$

$$\sum_{n=1}^{\infty} \mathbb{E}\,|Y_n|^{\alpha} \le \sum_{n=1}^{\infty}\sum_{j=1}^{n} |a_n|^{\alpha} \int_{A_j} |X_1|^{\alpha}$$

$$\le C \sum_{n=1}^{\infty}\sum_{j=1}^{n} n^{-\alpha/r} \int_{A_j} |X_1|^{\alpha}$$

$$\le C \sum_{n=1}^{\infty} \left(j^{-\alpha/r} + \frac{r}{\alpha - r} j^{(r-\alpha)/r}\right) \int_{A_j} |X_1|^{\alpha} \tag{2.1}$$

$$\le \sum_{n=1}^{\infty} \frac{\alpha}{\alpha - r} \int_{A_j} |X_1|^{r}$$

$$\le \frac{\alpha}{\alpha - r}\,\mathbb{E}\,|X_1|^{r} < \infty,$$

whence ($\alpha = 2$) $\sum_{n=1}^{\infty}(Y_n - \mathbb{E}\,Y_n)$ converges a.s. by Khintchine's theorem (see [1], page 113). Since

$$\sum_{n=1}^{\infty} \mathbb{P}\left\{a_n X_n \ne Y_n\right\} = \sum_{n=1}^{\infty} \mathbb{P}\left\{|X_1| > n^{1/r}\right\} \le \mathbb{E}\,|X_1|^{r} < \infty$$

the sequences $\{a_n X_n\}$, $\{Y_n\}$ are equivalent and so $\sum_{n=1}^{\infty}(a_n X_n - \mathbb{E}\,Y_n)$ converges a.s.

In case (i), where $0 < r < 1$, $\sum_{n=1}^{\infty} |\mathbb{E}\,Y_n| < \infty$ via (1.1) with $\alpha = 1$. In case (ii), where $1 < r < 2$ and $\mathbb{E}\,X_1 = 0$, we have

$$\begin{aligned}\sum_{n=1}^{\infty}|\mathbb{E}\,Y_n| &\le \sum_{n=1}^{\infty}|a_n|\int_{\{|X_n|>n^{1/r}\}}|X_n|\\ &\le C\sum_{n=1}^{\infty}\sum_{j=n+1}^{\infty}n^{-1/r}\int_{A_j}|X_1|\\ &= C\sum_{j=2}^{\infty}\sum_{n=1}^{j-1}n^{-1/r}\int_{A_j}|X_1|\\ &\le C\,\frac{r}{r-1}\sum_{j=1}^{\infty}(j-1)^{(r-1)/r}\int_{A_j}|X_1|^r\\ &\le C\,\frac{r}{r-1}\sum_{j=1}^{\infty}\int_{A_j}|X_1|^r\\ &= C\,\frac{r}{r-1}\,\mathbb{E}\,|X_1|^r<\infty.\end{aligned}$$

Thus, the second part of Lemma 1 follows from the first. ■

Remark. If $\{X_n, n \ge 1\}$ is a i.i.d. sequence of r.v.'s with $\mathbb{E}\,|X_1| < \infty$ and $\{a_n, n \ge 1\}$ are real numbers such that $a_n = O\left(n^{-1}\right)$ and $\sum_{n=1}^{\infty} a_n$ converges then $\sum_{n=1}^{\infty} a_n X_n$ converges a.s.

3. Strong consistency

In this section we shall prove the main result of this paper.

Theorem 1. *Suppose that in model* (1.1), $\varepsilon_1, \varepsilon_2, \dots$ *are random variables satisfying* (1.2) *and* $\{x_{ij}\}$ $(i = 1, 2, \dots ; j = 1, \dots, p)$ *is an arbitrary double* *array of constants. If* $\mathbf{X}_n'\mathbf{X}_n$ *is nonsingular for all* $n \geqslant n_0$ *and asymptotically diagonal dominant with constants* x_{ij} *satisfying* $\sum_{k=1}^{n} x_{kj}^2 \to \infty$ *for all* j *and*

(i) $$\frac{x_{nj}}{\left(\sum_{k=1}^{n} x_{ki}^2 \sum_{k=1}^{n} x_{kj}^2\right)^{1/2}} = O\left(n^{-1/r}\right) \text{ for all } i,j = 1,\ldots,p \text{ when } r \neq 1$$

or

(ii) $$\frac{x_{nj}}{\left(\sum_{k=1}^{n} x_{ki}^2 \sum_{k=1}^{n} x_{kj}^2\right)^{1/2}} = O\left(n^{-1}\right) \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{x_{nj}}{\left(\sum_{k=1}^{n} x_{ki}^2 \sum_{k=1}^{n} x_{kj}^2\right)^{1/2}}$$

converges for all $i,j = 1,\ldots,p$ *whenever* $r = 1$,

then $\mathbf{b}_n \xrightarrow{\text{a.s.}} \boldsymbol{\beta}$.

Proof. Setting $\mathbf{C}_n = \left(c_{ij}^{(n)}\right)_{1 \leqslant i,j \leqslant p} = (\mathbf{X}_n'\mathbf{X}_n)^{-1}$ we have

$$\det(\mathbf{X}_n'\mathbf{X}_n)\, c_{ij}^{(n)} = (-1)^{i+j} \sum_{\sigma \in S_{\hat{i}\hat{j}}} \operatorname{sgn}(\sigma) \prod_{\substack{m=1 \\ m \neq i \\ \sigma(m) \neq j}}^{p} \sum_{k=1}^{n} x_{km} x_{k\sigma(m)}$$

where the sum is computed over all bijections σ of $\{1,\ldots,i-1,i+1,\ldots,p\}$ into $\{1,\ldots,j-1,j+1,\ldots,p\}$. Thus, using Cauchy-Schwarz inequality we get

$$\det(\mathbf{X}_n'\mathbf{X}_n)\left|c_{ii}^{(n)}\right| \leqslant (p-1)! \prod_{\substack{m=1 \\ m \neq i}}^{p} \sum_{k=1}^{n} x_{km}^2$$

and

$$\det(\mathbf{X}_n'\mathbf{X}_n)\left|c_{ij}^{(n)}\right| \leqslant$$

$$\leqslant (p-1)! \left(\sum_{k=1}^{n} x_{ki}^2\right)^{1/2} \left(\sum_{k=1}^{n} x_{kj}^2\right)^{1/2} \prod_{\substack{m=1 \\ m \neq i,j}}^{p} \sum_{k=1}^{n} x_{km}^2, \quad i \neq j.$$

Since $\det(\mathbf{X}'_n\mathbf{X}_n) \asymp \sum_{k=1}^n x_{k1}^2 \ldots \sum_{k=1}^n x_{kp}^2$ as $n \to \infty$ it is sufficient to prove that

$$\frac{1}{\left(\sum_{k=1}^n x_{ki}^2 \sum_{k=1}^n x_{kj}^2\right)^{1/2}} \sum_{m=1}^n x_{mj}\,\varepsilon_m \xrightarrow{\text{a.s.}} 0 \qquad (i,j=1,\ldots,p). \tag{3.1}$$

By Lemma 1

$$\sum_{m=1}^n \frac{x_{mj}}{\left(\sum_{k=1}^m x_{ki}^2 \sum_{k=1}^m x_{kj}^2\right)^{1/2}}\,\varepsilon_m \quad \text{converges a.s.} \qquad (i,j=1,\ldots,p)$$

and Kronecker's lemma permit us to conclude (3.1) which establish the thesis. ■

$$\frac{}{2\alpha_i+1} \quad \text{and} \quad \sum_{k=1}^n x_{ki}x_{kj} \sim \frac{n^{\alpha_i+\alpha_j+1}}{\alpha_i+\alpha_j+1}, \quad i \neq j$$

which implies that $\mathbf{X}'_n\mathbf{X}_n$ is asymptotically diagonal dominant. Indeed, using induction on p we have

$$\det(\mathbf{X}'_n\mathbf{X}_n) = \sum_{j=1}^p \sum_{k=1}^n x_{kp}x_{kj} \cdot (-1)^{p+j} \det\big(\mathbf{X}'_n\mathbf{X}_n(p|j)\big)$$

where the modulus of each term of $(-1)^{p+j}\det\big(\mathbf{X}'_n\mathbf{X}_n(p|j)\big)$, $j=1,\ldots,p-1$ is bounded by

$$\left(\sum_{k=1}^n x_{kp}^2\right)^{1/2} \left(\sum_{k=1}^n x_{kj}^2\right)^{1/2} \prod_{\substack{m=1\\ m\neq j,p}}^p \sum_{k=1}^n x_{km}^2.$$

The conclusion now follows since

$$\frac{\sum_{k=1}^{n} x_{kp}x_{kj}}{\left(\sum_{k=1}^{n} x_{kp}^2\right)^{1/2}\left(\sum_{k=1}^{n} x_{kj}^2\right)^{1/2}} \sim \frac{\sqrt{2\alpha_p+1}\sqrt{2\alpha_j+1}}{\alpha_p+\alpha_j+1} < 1, \quad j=1,\dots,p-1$$

and $\sum_{k=1}^n x_{kp}^2 \cdot \det\left(\mathbf{X}_n'\mathbf{X}_n(p|p)\right) \sim \sum_{k=1}^n x_{k1}^2 \cdots \sum_{k=1}^n x_{kp}^2$ as $n \to \infty$. The assumptions (i) or (ii) of Theorem 1 are also satisfied.

References

[1] Y.S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales* (third edition) Springer 1997.

[2] Chen Xiru, *Consistency of LS estimates of multiple regression under a lower order moment condition*, Sci. Chin. **38** (12) (1995), 1420–1431.

[3] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press 1985.

[4] H. Drygas, *Consistency of the least squares and Gauss-Markov estimators in regression models,* Z. Wahrscheinlichkeitstheorie Verw. Gebiete **17** (1971), 309–326.

[5] H. Drygas, *Weak and strong consistency of the least squares estimators in regression model,* Z. Wahrscheinlichkeitstheorie Verw. Gebiete **34** (1976), 119–127.

[6] C. Gui-Jing, T.L. Lai and C.Z. Wei, *Convergence systems and strong, consistency of least squares estimates in regression models*, J. Multivariate Anal. **11** (1981), 319–333.

[7] C. GuiJing, *Extension of Lai-Robbins-Wei's theorem*, Acta Mathematicae Applicatae Sinica **1** (1) (1984), 2–7.

[8] J. Mingzhong, Some new results of the strong consistency of multiple regression coefficients, in: S. Tangmanee & E. Schulz, eds. World Scientific, *Proceedings of the Second Asian Mathematical Conference 1995* (R. Nakhon 1995), 514–519.

[9] T.L. Lai, H. Robbins and C.Z. Wei, *Strong consistency of least squares estimates in multiple regression II*, J. Multivariate Anal. **9** (1979), 343–362.

[10] B.M. Makarov, M.G. Goluzina, A.A. Lodkin and A.N. Podkorytov, *Selected Problems in Real Analysis* (American Mathematical Society, Providence R.I. 1992).

[11] J.T. Mexia, P. Corte Real, M.L. Esquível e J. Lita da Silva, *Convergência do estimador dos mínimos quadrados em modelos lineares*, Estatística Jubilar. Actas do XII Congresso da Sociedade Portuguesa de Estatística, Edições SPE (2005), 455–466.

[12] J.T. Mexia e J. Lita da Silva, *A consistência do estimador dos mínimos quadrados em domínios de atracção maximais*, Ciência Estatística. Actas do XIII Congresso Anual da Sociedade Portuguesa de Estatística, Edições SPE (2006), 481–492.

[13] J.T. Mexia and J. Lita da Silva, *Sufficient conditions for the strong consistency of least squares estimator with α-stable errors*, Discussiones Mathematicae - Probability and Statistics **27** (2007), 27–45.

Received 15 November 2009