*Discussiones Mathematicae*
*Probability and Statistics 27(2007) 27–45*

# SUFFICIENT CONDITIONS FOR THE STRONG CONSISTENCY OF LEAST SQUARES ESTIMATOR WITH $\alpha$-STABLE ERRORS

João Tiago Mexia

*Mathematics Department, Faculty of Science and Technology*
*New University of Lisbon*
*Monte da Caparica 2829–516 Caparica, Portugal*

**e-mail:** parcr@fct.unl.pt

and

João Lita da Silva

*New University of Lisbon, Mathematics Department*
*Faculty of Science and Technology*
*Quinta da Torre, 2825–114 Monte da Caparica, Portugal*

**e-mail:** jfls@fct.unl.pt

**Abstract**

Let $Y_i = \mathbf{x}_i^T \boldsymbol{\beta} + e_i$, $1 \leqslant i \leqslant n$, $n \geqslant 1$ be a linear regression model and suppose that the random errors $e_1, e_2, \ldots$ are independent and $\alpha$-stable. In this paper, we obtain sufficient conditions for the strong consistency of the least squares estimator $\widetilde{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ under additional assumptions on the non-random sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ of real vectors.

**Keywords:** linear models, least squares estimator, strong consistency, stability.

**2000 Mathematics Subject Classification:** 60F15.

## 1. Introduction

The strong consistency of the Least Squares Estimator (LSE) has been studied on the last years by many authors (see for instante [6, 10, 11, 12, 13, 14] or [15]). In the papers [10] and [11] are obtained necessary and sufficient conditions for the LSE strong consistency. Nevertheless, the formulation used by the authors of the papers quoted above assumes an i.i.d. random error sequence with null expected value and finite absolute moment of order $1 \leqslant r < 2$. In a first stage, our purpose will be to derive LSE strong consistency from a formulation that is, in some sense, similar to the described previously. As a matter of fact, we will assume that the errors $e_1, e_2, \ldots$ are i.i.d. $\alpha$-stable excluding any hypothesis on their absolute moments. Let us point out that the stability condition for the errors is introduced here to make useful the suitable 'linearity' properties of the stable distributions. On a second stage, we generalized the above assumptions supposing that the random variables $e_1, e_2, \ldots$ are only independent and $\alpha$-stable, obtaining an extension of the results presented in [10] and [11].

Consider the linear model,

$$Y_i = \mathbf{x}_i^T \boldsymbol{\beta} + e_i, \quad 1 \leqslant i \leqslant n, \quad n \geqslant 1, \tag{1.1}$$

where $\mathbf{x}_1, \mathbf{x}_2, \ldots$ are known non-random real $\kappa$-vectors,

$$\boldsymbol{\beta} := \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_\kappa \end{bmatrix}$$

is the vector of (unknown) parameters and $e_i$ is the random error in the $i$-th observation ($i = 1, \ldots, n$).

Setting

$$\mathbf{S}_n := \mathbf{x}_1 \mathbf{x}_1^T + \ldots + \mathbf{x}_n \mathbf{x}_n^T$$

we will assume that $\mathbf{S}_n^{-1}$ exists for $n$ large enough (for small $n$ such that $\mathbf{S}_n^{-1}$ does not exist, $\mathbf{S}_n^{-1}$ can be defined arbitrarily). Denoting the LSE of $\boldsymbol{\beta}$ by $\widetilde{\boldsymbol{\beta}}$

we can write

$$\widetilde{\boldsymbol{\beta}} = \boldsymbol{\beta} + \mathbf{S}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i e_i \tag{1.2}$$

so that, $\widetilde{\boldsymbol{\beta}}$ is strongly consistent if and only if $\mathbf{S}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i e_i \xrightarrow{\text{a.s.}} \mathbf{0}$.

## 2. Auxiliary tools

The following notational conventions will be adopted throughout the paper. Given a vector $\mathbf{x} \in \mathbb{R}^{\kappa}$, the *euclidean vector norm* will be denoted by

$$\|\mathbf{x}\| := \sqrt{\mathbf{x}^T \mathbf{x}}$$

and the matrix norm of an matrix $\mathbf{A} \in \mathcal{M}_{\kappa\times\kappa}(\mathbb{R})$ will be indicated by

$$|\!|\!|\mathbf{A}|\!|\!| := \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}.$$

The *trace* of $\mathbf{A} \in \mathcal{M}_{\kappa\times\kappa}(\mathbb{R})$ will be denoted by tr $(\mathbf{A})$. For any matrix $\mathbf{A} \in \mathcal{M}_{\kappa\times\kappa}(\mathbb{R})$ with real eigenvalues we will indicate by $\nu_{\max}(\mathbf{A})$ and $\nu_{\min}(\mathbf{A})$ the maximum and minimum eigenvalue of $\mathbf{A}$ respectively. Given two symmetric matrices $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{\kappa\times\kappa}(\mathbb{R})$ we write $\mathbf{A} \geqslant \mathbf{B}$ to denote that $\mathbf{A} - \mathbf{B}$ is nonnegative definite; we will use the notation $\mathbf{A} > \mathbf{B}$ to indicate that $\mathbf{A} - \mathbf{B}$ is positive definite.

It is well-known that a random variable $Z$ (or its distribution) is said to have a *stable* distribution if for any positive numbers $c_1$ and $c_2$, there is a positive number $a$ and a real number $b$ such that

$$c_1 Z_1 + c_2 Z_2 \stackrel{\text{d}}{=} a Z + b \tag{2.1}$$

where $Z_1$ and $Z_2$ are independent copies of $Z$.[1] A random variable $Z$ is called *strictly stable* if (2.1) holds with $b = 0$ (see [16]).

Equivalently (see [16]), a random variable $Z$ is said to have a stable distribution if there are parameters $0 < \alpha \leqslant 2$, $\sigma \geqslant 0$, $-1 \leqslant \lambda \leqslant 1$ and $\mu$ real such that its characteristic function has the following form,

(2.2)
$$\mathbb{E}\left(e^{iZt}\right) = \begin{cases} \exp\left(-\sigma^\alpha\ |t|^\alpha \left(1 - i\lambda(\text{sign } t)\tan\left(\frac{\pi\alpha}{2}\right)\right) + i\mu t\right) & \text{if} \quad \alpha \neq 1 \\ \\ \exp\left(-\sigma\ |t| \left(1 + i\lambda\frac{2}{\pi}(\text{sign } t)\log|t|\right) + i\mu t\right) & \text{if} \quad \alpha = 1 \end{cases},$$

where
$$\text{sign } t := \begin{cases} 1 & \text{if} \quad t > 0 \\ 0 & \text{if} \quad t = 0 \\ -1 & \text{if} \quad t < 0 \end{cases}.$$

Since (2.2) is characterized by four (unique) parameters

$$\alpha \in ]0, 2], \quad \sigma \geqslant 0, \quad \lambda \in [-1, 1], \quad \mu \in \mathbb{R}$$

respectively, the *index of stability*, the *scale parameter*, the *skewness parameter* and the *shift parameter* we will denote stable distributions by $\mathcal{S}_\alpha(\sigma, \lambda, \mu)$ and write

$$Z \sim \mathcal{S}_\alpha(\sigma, \lambda, \mu)$$

to indicate that $Z$ has the stable distribution $\mathcal{S}_\alpha(\sigma, \lambda, \mu)$. A stable random variable $Z$ with index of stability $\alpha$ is called *$\alpha$-stable*.

Let us remark that a random variable $Z$ concentrated at one point is always stable, i.e., a constant $\mu$ has degenerate distribution $\mathcal{S}_\alpha(0, 0, \mu)$ for

[1] The notation $X \overset{\text{d}}{=} Y$ means that the random variables $X$ and $Y$ agree in distribution, that is, $F_X(x) = F_Y(x)$, $x \in \mathbb{R}$.

any $0 < \alpha \leqslant 2$. This degenerate case is of no special interest and, unless stated explicitly, we always assume that $Z$ is non-degenerate. In fact, degenerate distributions have unusual properties, for example, all moments of a degenerate distribution are finite, whereas a non-degenerate stable distribution with $0 < \alpha < 2$ has infinite second order moments.

The distribution functions of $\alpha$-stable random variables are absolutely continuous and their densities have derivatives of any order in every point, but excluding some special cases, their representations can be expressed only through complicated special functions. However, there exist asymptotic expansions of $\alpha$-stable densities in a neighborhood of the origin or of infinity (see [9, 17] or [18]). The main special case is given by

$$\mathcal{S}_2(\sigma, \lambda, \mu) = N(\mu, 2\sigma^2) \qquad (\sigma > 0).$$

In this situation the skewness parameter $\lambda$ is totally irrelevant. Hence, a random variable

$$Z \sim \mathcal{S}_2(\sigma, \lambda, \mu) \qquad (\sigma > 0)$$

has probability density function

$$f_Z(z) = \frac{1}{2\sigma\sqrt{\pi}}\, e^{-\frac{(z-\mu)^2}{4\sigma^2}}, \qquad z \in \mathbb{R}.$$

The remaining special cases of probability densities of $\alpha$-stable random variables in closed form are: the Cauchy distribution $\mathcal{S}_1(\sigma, 0, \mu)$ whose density is

$$f(x) = \frac{\sigma}{\pi\big((x-\mu)^2 - \sigma^2\big)}, \quad x \in \mathbb{R}$$

and the Lévy distribution $\mathcal{S}_{1/2}(\sigma, 1, \mu)$ whose density is

$$f(x) = \sqrt{\frac{\sigma}{2\pi(x-\mu)^3}}\, e^{-\frac{\sigma}{2(x-\mu)}}, \quad x > \mu.$$

Some basic properties of stable distributions are listed below: (see [16])

(1) Let $Z_1$ and $Z_2$ be independent random variables with $Z_i \sim \mathcal{S}_\alpha(\sigma_i, \lambda_i, \mu_i)$, $i = 1, 2$. Then

$$Z_1 + Z_2 \sim \mathcal{S}_\alpha\left((\sigma_1^\alpha + \sigma_2^\alpha)^{1/\alpha}, \frac{\lambda_1\sigma_1^\alpha + \lambda_2\sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}, \mu_1 + \mu_2\right).$$

(2) Let $Z \sim \mathcal{S}_\alpha(\sigma, \lambda, \mu)$ and let $a$ be a real constant. Then $Z + a \sim \mathcal{S}_\alpha(\sigma, \lambda, \mu + a)$.

(3) Let $Z \sim \mathcal{S}_\alpha(\sigma, \lambda, \mu)$ and let $a$ be a non-zero real constant. Then

$$aZ \sim \mathcal{S}_\alpha(|a|\,\sigma, \text{sign }(a)\lambda, a\mu) \qquad \text{if} \quad \alpha \neq 1$$

$$aZ \sim \mathcal{S}_1\left(|a|\,\sigma, \text{sign }(a)\lambda, a\mu - \frac{2}{\pi}\sigma\lambda a \log|a|\right) \quad \text{if} \quad \alpha = 1.$$

(4) Let $Z \sim \mathcal{S}_\alpha(\sigma, \lambda, \mu)$ with $\alpha \neq 1$. Then $Z$ is strictly stable if and only if $\mu = 0$.

(5) $Z \sim \mathcal{S}_1(\sigma, \lambda, \mu)$ is strictly stable if and only if $\lambda = 0$.

(6) $Z \sim \mathcal{S}_\alpha(\sigma, \lambda, \mu)$ is symmetric about $\mu$ if and only if $\lambda = 0$.

(7) Let $Z$ have distribution $\mathcal{S}_\alpha(\sigma, \lambda, 0)$ with $\alpha < 2$. Then there exist two i.i.d. random variables $A$ and $B$ with common distribution $\mathcal{S}_\alpha(\sigma, 1, 0)$ such that

$$Z \stackrel{\text{d}}{=} \left(\frac{1+\lambda}{2}\right)^{1/\alpha} A - \left(\frac{1-\lambda}{2}\right)^{1/\alpha} B \ \text{ if } \ \alpha \neq 1$$

$$Z \stackrel{\mathrm{d}}{=} \left(\frac{1+\lambda}{2}\right) A - \left(\frac{1-\lambda}{2}\right) B +$$

$$+\sigma \left[ \left(\frac{1+\lambda}{\pi}\right) \log\left(\frac{1+\lambda}{2}\right) - \left(\frac{1-\lambda}{\pi}\right) \log\left(\frac{1-\lambda}{2}\right) \right] \text{ if } \alpha = 1.$$

The next statement is a powerful tool in matrix analysis and generalizes a result which is a *sine qua non* for probability theory: the scalar version of the Kronecker lemma.

**Lemma 2.1** Kronecker's lemma (matrix case). *Let* $\mathbf{x}_i$, $i = 1, 2, \ldots$ *be a sequence of real* $\kappa$*-vectors,* $\mathbf{A}_i$ *a sequence of real symmetric nonsingular* $\kappa \times \kappa$ *matrices with* $\mathbf{A}_{i+1} \geqslant \mathbf{A}_i > \mathbf{O}$ *and* $\mathbf{P}_i$ *a sequence of nonsingular* $\kappa \times \kappa$ *matrices. Suppose that* $\mathbf{r}_i = \sum_{j=i}^{\infty} \mathbf{A}_j^{-1} \mathbf{x}_j$ *exists (and is finite). Then:*

(i) *If* $\mathbf{A}_\infty := \lim_{n\to\infty} \mathbf{A}_n$ *exists and is finite, then* $\lim_{n\to\infty} \mathbf{A}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i$ *exists and is finite.*

(ii) *If* $\lim_{n\to\infty} \frac{1}{\mathrm{tr}\ (\mathbf{A}_n)} = 0$ *then* $\lim_{n\to\infty} \frac{1}{\mathrm{tr}\ (\mathbf{A}_n)} \sum_{i=1}^{n} \mathbf{x}_i = \mathbf{0}$.

(iii) *If* $\lim_{n\to\infty} \mathbf{A}_n^{-1} = \mathbf{O}$, $\lim_{n\to\infty} \mathbf{P}_n \mathbf{r}_n = \mathbf{0}$ *and*

$$\limsup_{n\to\infty} \sum_{i=1}^{n} \left\| \mathbf{A}_n^{-1} (\mathbf{A}_i - \mathbf{A}_{i-1}) \mathbf{P}_i^{-1} \right\| < \infty \text{ then } \lim_{n\to\infty} \mathbf{A}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i = \mathbf{0}.$$

***Proof.*** The proof can be founded in [1] on page 2.7. ■

**Remark 1.** The condition $\limsup_{n\to\infty}\sum_{i=1}^{n}\left\|\mathbf{A}_n^{-1}\big(\mathbf{A}_i-\mathbf{A}_{i-1}\big)\mathbf{P}_i^{-1}\right\|<\infty$ is implied by the much stronger condition

$$\limsup_{n\to\infty}\frac{\nu_{\max}(\mathbf{A}_n)}{\nu_{\min}(\mathbf{A}_n)}<\infty.$$

For some remarks on a matrix version of Kronecker's lemma, see [2].

To finish this section, we present an important result that gives us the almost sure convergence for sums of independent random variables $X_i\sim\mathcal{S}_\alpha(\sigma_i,\lambda_i,\mu_i)$.

**Proposition 2.1.** *Let $\{X_i,\, i\geqslant 1\}$ be independent random variables $X_i\sim\mathcal{S}_\alpha(\sigma_i,\lambda_i,\mu_i)$, $0<\alpha\leqslant 2$. Then,*

$$\sum_{i=1}^{\infty}X_i \quad \text{converge} \quad \text{a.s.} \iff \sum_{i=1}^{\infty}\sigma_i^\alpha<\infty \quad \text{and} \quad \sum_{i=1}^{\infty}\mu_i \ \text{converge.}$$

***Proof.*** We establish the proof only for $0<\alpha<2$ since the case $\alpha=2$ is obvious in view of the form of characteristic functions of normal law. We show first that if $\sum_{i=1}^{\infty}\sigma_i^\alpha<\infty$ and $\sum_{i=1}^{\infty}\mu_i$ converge then $\sum_{i=1}^{\infty}X_i$ converges almost surely. Setting,

$$a_n:=(\sigma_1^\alpha+\ldots+\sigma_n^\alpha)^{1/\alpha}\,,\ \ b_n:=\frac{\lambda_1\,\sigma_1^\alpha+\ldots+\lambda_n\,\sigma_n^\alpha}{\sigma_1^\alpha+\ldots+\sigma_n^\alpha}$$

$$\text{and}\ \ c_n:=\mu_1+\ldots+\mu_n$$

we get, for all $n\in\mathbb{N}$,

$$\sum_{i=1}^{n} X_i \sim \mathcal{S}_\alpha\left(a_n, b_n, c_n\right)$$

provided that $X_i \sim \mathcal{S}_\alpha(\sigma_i, \lambda_i, \mu_i)$.

Therefore, we can write

$$\sum_{i=1}^{n} X_i \stackrel{\mathrm{d}}{=} a_n\, Y_n + c_n \quad \text{if} \quad \alpha \neq 1 \tag{2.3}$$

and

$$\sum_{i=1}^{n} X_i \stackrel{\mathrm{d}}{=} a_n\, Y_n + \frac{2}{\pi}\, b_n\, a_n\, \log a_n + c_n \quad \text{if} \quad \alpha = 1 \tag{2.4}$$

where $\{Y_n, n \geqslant 1\}$ is a sequence of random variables satisfying $Y_n \sim \mathcal{S}_\alpha(1, b_n, 0)$. Hence, there exist two i.i.d. random variables $A_n$ and $B_n$ with common distribution $\mathcal{S}_\alpha(1, 1, 0)$ such that,

$$Y_n \stackrel{\mathrm{d}}{=} \left(\frac{1+b_n}{2}\right)^{1/\alpha} A_n - \left(\frac{1-b_n}{2}\right)^{1/\alpha} B_n \quad \text{if} \quad \alpha \neq 1 \tag{2.5}$$

and

$$\begin{aligned} Y_n \stackrel{\mathrm{d}}{=} \left(\frac{1+b_n}{2}\right) A_n - \left(\frac{1-b_n}{2}\right) B_n + \frac{1+b_n}{\pi} \log\left(\frac{1+b_n}{2}\right) - \\ - \frac{1-b_n}{\pi} \log\left(\frac{1-b_n}{2}\right) \quad \text{if} \quad \alpha = 1. \end{aligned} \tag{2.6}$$

Since $\sum_{i=1}^{\infty} \sigma_i^{\alpha} < \infty$ and $\sum_{i=1}^{\infty} \mu_i$ is convergent we obtain the convergence in law of $Y_n$ since $A_n$ and $B_n$ have both distribution $\mathcal{S}_{\alpha}(1,1,0)$ for all $n \in \mathbb{N}$ (let us note that $b_n$ is a convergent sequence). According to (2.3) and (2.4), the series $\sum_{i=1}^{n} X_i$ converges in law which implies the existence of a (proper) random variable $S$ such that $\sum_{i=1}^{n} X_i \xrightarrow{\text{a.s.}} S$ (see [4]).

We now show that, if $\sum_{i=1}^{\infty} X_i$ converges almost surely then $\sum_{i=1}^{\infty} \sigma_i^{\alpha} < \infty$ and $\sum_{i=1}^{\infty} \mu_i$ converge. Since

$$|b_n| = \left| \frac{\lambda_1 \, \sigma_1^{\alpha} + \ldots + \lambda_n \, \sigma_n^{\alpha}}{\sigma_1^{\alpha} + \ldots + \sigma_n^{\alpha}} \right| \leqslant 1, \quad \forall n \in \mathbb{N}$$

there exist a subsequence $\left(b_{\eta_n}\right)$ converging to a point of $[-1,1]$ which implies the converge in law of the following random sequences:

$$\left(\frac{1+b_{\eta_n}}{2}\right)^{1/\alpha} A_{\eta_n} - \left(\frac{1-b_{\eta_n}}{2}\right)^{1/\alpha} B_{\eta_n} \quad \text{if} \quad \alpha \neq 1$$

and

$$\left(\frac{1+b_{\eta_n}}{2}\right) A_{\eta_n} - \left(\frac{1-b_{\eta_n}}{2}\right) B_{\eta_n} + \frac{1+b_{\eta_n}}{\pi} \log\left(\frac{1+b_{\eta_n}}{2}\right) -$$

$$- \frac{1-b_{\eta_n}}{\pi} \log\left(\frac{1-b_{\eta_n}}{2}\right) \quad \text{if} \quad \alpha = 1 \, .$$

Therefore, the almost sure convergence of $\sum_{i=1}^{\infty} X_i$ leads to

$$\lim_{n \to \infty} \sum_{i=1}^{\eta_n} \sigma_i^{\alpha} < \infty$$

(see Theorem 14.2 of [3] in page 193) so that,

$$\sum_{i=1}^{\infty} \sigma_i^{\alpha} < \infty$$

since this is a series of non-negative terms. Thus,

$$\frac{\lambda_1 \, \sigma_1^{\alpha} + \ldots + \lambda_n \, \sigma_n^{\alpha} + \ldots}{\sigma_1^{\alpha} + \ldots + \sigma_n^{\alpha} + \ldots} \quad \text{converge,} \qquad \sum_{i=1}^{\infty} \mu_i \quad \text{converge.}$$

■

## 3. THE STRONG CONSISTENCY OF LSE

From the sixties, the strong consistency of LSE has attracted much attention from many statisticians. In the middle and late of seventies the problem was satisfactorily solved for the case where the random errors possesses a finite variance. In fact, a paper by Lai, Robbins and Wei (1979) showed that if the $e_1, e_2, \ldots$ are i.i.d. with $\mathbb{E}\big(e_1\big) = 0$ and $0 < \mathbb{E}\big(e_1^2\big) < \infty$ then a sufficient condition for the strong consistency of $\widetilde{\boldsymbol{\beta}}$ is,

$$\mathbf{S}_n^{-1} = \left( \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T \right)^{-1} \longrightarrow \mathbf{O}$$

as $n \to \infty$. Hence, the particular case of the linear regression model (1.1) given by i.i.d. strictly stable errors with index of stability $\alpha = 2$ is completely under cover by the Lai, Robbins & Wei's result.

**Theorem 3.1.** *Consider the linear regression model* (1.1). *If*

(a) *the random variables* $e_1, e_2, \ldots$ *are i.i.d.* $\mathcal{S}_2(\sigma, \lambda, 0)$,

(b) $\lim\limits_{n \to \infty} \mathbf{S}_n^{-1} = \mathbf{O}$,

*then* $\widetilde{\boldsymbol{\beta}}$ *is strongly consistent.*

**Remark 2.** Let us note that the necessity of the condition $\lim_{n\to\infty} \mathbf{S}_n^{-1} = \mathbf{O}$ on the LSE strong consistency it had been proved earlier by H. Drygas (1976) on the same assumptions for the error sequence: $e_1, e_2, \ldots$ i.i.d. with $\mathbb{E}(e_1) = 0$ and $0 < \mathbb{E}(e_1^2) < \infty$.

More generally,

**Theorem 3.2.** *Consider the linear regression model* (1.1). *If*

(a) *the random variables* $e_1, e_2, \ldots$ *are i.i.d.* $\mathcal{S}_2(\sigma, \lambda, \mu)$ *with* $\mu \neq 0$,

(b) $\lim\limits_{n\to\infty} \mathbf{S}_n^{-1} = \mathbf{O}$, $\limsup\limits_{n\to\infty} \sum\limits_{i=1}^{n} \left\| \mathbf{S}_n^{-1} \mathbf{x}_i \mathbf{x}_i^T \right\| < \infty$ *and* $\left\| \sum\limits_{i=1}^{\infty} \mathbf{S}_i^{-1} \mathbf{x}_i \right\| < \infty$,

*then* $\widetilde{\boldsymbol{\beta}}$ *is strongly consistent.*

***Proof.*** From (1.2) we get

$$\widetilde{\boldsymbol{\beta}} = \boldsymbol{\beta} + \mathbf{S}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i (e_i - \mu) + \mu \mathbf{S}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i.$$

Since $e_i - \mu \sim \mathcal{S}_2(\sigma, \lambda, 0)$ the thesis is a consequence of Theorem 3.1 and Kronecker's lemma (matrix case). ■

The next theorem can be interpreted as an alternative path to the results presented in [10] and [11]. Indeed, the authors of the papers quoted above suppose an i.i.d. error random sequence where each term has null expected value and finite absolute moment of order $1 \leqslant r < 2$. Our results will establishes the strong consistency of the LSE assuming only that the error random sequence is i.i.d. $\alpha$-stable ($\alpha < 2$).

**Theorem 3.3.** *Let* $0 < \alpha < 2$, $\alpha \neq 1$ *and consider the linear regression model* (1.1). *If*

(a) *the random variables* $e_1, e_2, \ldots$ *are i.i.d.* $\mathcal{S}_\alpha(\sigma, \lambda, \mu)$,

(b) $\lim_{n\to\infty} \mathbf{S}_n^{-1} = \mathbf{O}$, $\limsup_{n\to\infty} \sum_{i=1}^{n} \left\| \mathbf{S}_n^{-1} \mathbf{x}_i \mathbf{x}_i^T \right\| < \infty$ *and* $\sum_{i=1}^{\infty} \left\| \mathbf{S}_i^{-1} \mathbf{x}_i \right\|^{\min(\alpha,1)} < \infty$,

*then* $\widetilde{\boldsymbol{\beta}}$ *is strongly consistent.*

***Proof.*** The identity (1.2) guarantees,

$$\widetilde{\boldsymbol{\beta}} = \boldsymbol{\beta} + \mathbf{S}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i e_i = \boldsymbol{\beta} + \mathbf{S}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i e_i^* + \mu\, \mathbf{S}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i$$

with $e_i^* \sim \mathcal{S}_\alpha(\sigma, \lambda, 0)$. On the one hand, by Kronecker's lemma (matrix case) we have $\mu\, \mathbf{S}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i \longrightarrow \mathbf{0}$ as $n \to \infty$ provided that $\left\| \sum_{i=1}^{\infty} \mathbf{S}_i^{-1} \mathbf{x}_i \right\| < \infty$ and $\limsup_{n\to\infty} \sum_{i=1}^{n} \left\| \mathbf{S}_n^{-1} \mathbf{x}_i \mathbf{x}_i^T \right\| < \infty$. On the other hand, each component of the vector $\sum_{i=1}^{n} \mathbf{S}_i^{-1} \mathbf{x}_i e_i^*$ has distribution

$$\mathcal{S}_\alpha \left( \sigma \left( \sum_{i=1}^{n} |a_{ij}|^\alpha \right)^{1/\alpha}, \frac{\lambda \sum_{i=1}^{n} \operatorname{sign}(a_{ij})\, |a_{ij}|^\alpha}{\sum_{i=1}^{n} |a_{ij}|^\alpha}, 0 \right), \quad j = 1, \ldots, \kappa$$

where $a_{ij}$ is the $j$-th component of the vector $\mathbf{a}_i := \mathbf{S}_i^{-1} \mathbf{x}_i$. Since

$$\sum_{i=1}^{\infty} |a_{ij}|^\alpha \leqslant \sum_{i=1}^{\infty} \left\| \mathbf{S}_i^{-1} \mathbf{x}_i \right\|^\alpha$$

we obtain the almost sure convergence of $\sum_{i=1}^{\infty} \mathbf{S}_i^{-1} \mathbf{x}_i e_i^*$ through the Proposition 2.1 and Kronecker's lemma (matrix case) permit us to conclude $\mathbf{S}_n^{-1} \sum_{i=1}^{n} \mathbf{x}_i e_i^* \xrightarrow{\text{a.s.}} \mathbf{0}$. ■

On the next result we will use the notation $|\mathbf{x}|_{\min} := \min\big(|x_1|, \ldots, |x_\kappa|\big)$ where $\mathbf{x}$ is an non-random $\kappa$-vector.

**Theorem 3.4.** *Consider the linear regression model* (1.1). *If*

(a) *the random variables* $e_1, e_2, \ldots$ *are i.i.d.* $\mathcal{S}_1(\sigma, \lambda, \mu)$,

(b) $\lim\limits_{n\to\infty} \mathbf{S}_n^{-1} = \mathbf{O}$, $\limsup\limits_{n\to\infty} \sum\limits_{i=1}^{n} \left\|\!\left|\mathbf{S}_n^{-1}\mathbf{x}_i\mathbf{x}_i^T\right|\!\right\| < \infty$ *and*

$\sum\limits_{i=1}^{\infty} \left\|\mathbf{S}_i^{-1}\mathbf{x}_i\right\| \log\left|\mathbf{S}_i^{-1}\mathbf{x}_i\right|_{\min}$ *converges,*

*then* $\widetilde{\boldsymbol{\beta}}$ *is strongly consistent.*

***Proof.*** The relation (1.2) gives us

$$\widetilde{\boldsymbol{\beta}} = \boldsymbol{\beta} + \mathbf{S}_n^{-1}\sum_{i=1}^{n}\mathbf{x}_i e_i = \boldsymbol{\beta} + \mathbf{S}_n^{-1}\sum_{i=1}^{n}\mathbf{x}_i e_i^* + \mu\,\mathbf{S}_n^{-1}\sum_{i=1}^{n}\mathbf{x}_i$$

with $e_i^* \sim \mathcal{S}_1(\sigma, \lambda, 0)$. Each component of the vector $\sum_{i=1}^{n}\mathbf{S}_i^{-1}\mathbf{x}_i e_i^*$ has distribution

$$\mathcal{S}_1\left(\sigma\sum_{i=1}^{n}|a_{ij}|\,,\ \frac{\lambda\sum\limits_{i=1}^{n}\operatorname{sign}\,(a_{ij})\,|a_{ij}|}{\sum\limits_{i=1}^{n}|a_{ij}|},\ -\frac{2\sigma\lambda}{\pi}\sum_{i=1}^{n}a_{ij}\log|a_{ij}|\right), \quad j = 1, \ldots, \kappa$$

where $a_{ij}$ is the $j$-th component of $\mathbf{a}_i := \mathbf{S}_i^{-1}\mathbf{x}_i$. Therefore, we have

$$\mathbf{S}_n^{-1}\sum_{i=1}^{n}\mathbf{x}_i e_i^* \xrightarrow{\text{a.s.}} \mathbf{0}$$

which is a consequence of the Proposition 2.1 and the Kronecker's lemma (matrix case) provided that $\lim_{i\to\infty}\mathbf{S}_i^{-1}\mathbf{x}_i = \mathbf{0}$ (note that our assumptions guarantees $\left\|\mathbf{S}_i^{-1}\mathbf{x}_i\mathbf{x}_i^T\right\| < \infty$ as $i \to \infty$ ).
Since

$$\left\|\sum_{i=1}^{\infty}\mathbf{S}_i^{-1}\mathbf{x}_i\right\| < \infty \quad \text{and} \quad \limsup_{n\to\infty}\sum_{i=1}^{n}\left|\!\left|\!\left|\mathbf{S}_n^{-1}\mathbf{x}_i\mathbf{x}_i^T\right|\!\right|\!\right| < \infty$$

we get $\mu\,\mathbf{S}_n^{-1}\sum_{i=1}^{n}\mathbf{x}_i \longrightarrow \mathbf{0}$ as $n \to \infty$ and the thesis is established. ■

**Remark 3.** If the random errors $e_1, e_2, \ldots$ are i.i.d. strictly stable with index of stability $0 < \alpha < 2$ then the strong consistency of the LSE is still valid assuming $\lim_{n\to\infty}\mathbf{S}_n^{-1} = \mathbf{O}$, $\limsup_{n\to\infty}\sum_{i=1}^{n}\left|\!\left|\!\left|\mathbf{S}_n^{-1}\mathbf{x}_i\mathbf{x}_i^T\right|\!\right|\!\right| < \infty$ and $\sum_{i=1}^{\infty}\left\|\mathbf{S}_i^{-1}\mathbf{x}_i\right\|^{\alpha} < \infty$.

The previous results can be extended to the case where the random variables $e_1, e_2, \ldots$ are (only) independent and $e_i \sim \mathcal{S}_\alpha(\sigma_i, \lambda_i, \mu_i)$. In this sense, we finish this work with a possible generalization where the assumption of identical distribution for random errors $e_1, e_2, \ldots$ is completely excluded. For a better presentation of our ideas and their associated results let us start with the case $0 < \alpha \leqslant 2$, $\alpha \neq 1$.

**Theorem 3.5.** *Let* $0 < \alpha \leqslant 2$, $\alpha \neq 1$ *and consider the linear regression model* (1.1). *If*

(a) *the random variables* $e_1, e_2, \ldots$ *are independent such that* $e_i \sim \mathcal{S}_\alpha(\sigma_i, \lambda_i, \mu_i)$,

(b) $\lim_{n\to\infty} \mathbf{S}_n^{-1} = \mathbf{O}$, $\limsup_{n\to\infty} \sum_{i=1}^{n} \left\| \mathbf{S}_n^{-1}\mathbf{x}_i\mathbf{x}_i^T \right\| < \infty$, $\sum_{i=1}^{\infty} \left(\sigma_i \left\|\mathbf{S}_i^{-1}\mathbf{x}_i\right\|\right)^{\alpha} < \infty$

*and* $\left\| \sum_{i=1}^{\infty} \mu_i \mathbf{S}_i^{-1}\mathbf{x}_i \right\| < \infty$,

*then* $\widetilde{\boldsymbol{\beta}}$ *is strongly consistent.*

***Proof.*** From (1.2) we only have to prove the almost sure convergence of the series $\sum_{i=1}^{\infty} \mathbf{S}_i^{-1}\mathbf{x}_i e_i$ since the Kronecker's lemma (matrix case) establishes the thesis. Each component of the vector $\sum_{i=1}^{n} \mathbf{S}_i^{-1}\mathbf{x}_i e_i$ has distribution,

$$\mathcal{S}_\alpha\left(\left(\sum_{i=1}^{n}\left(\sigma_i\left|a_{ij}\right|\right)^{\alpha}\right)^{1/\alpha}, \frac{\sum_{i=1}^{n}\left(\sigma_i\left|a_{ij}\right|\right)^{\alpha}\lambda_i \text{sign}\,(a_{ij})}{\sum_{i=1}^{n}\left(\sigma_i\left|a_{ij}\right|\right)^{\alpha}}, \sum_{i=1}^{n}\mu_i a_{ij}\right), \quad j=1,\ldots,\kappa$$

where $a_{ij}$ is the $j$-th component of $\mathbf{a}_i := \mathbf{S}_i^{-1}\mathbf{x}_i$ so that, from Proposition 2.1 we obtain the almost sure convergence of $\sum_{i=1}^{\infty} \mathbf{S}_i^{-1}\mathbf{x}_i e_i$. ■

The case $\alpha = 1$ is described next.

**Theorem 3.6.** *Consider the linear regression model* (1.1). *If*

(a) *the random variables* $e_1, e_2, \ldots$ *are independent and such that* $e_i \sim \mathcal{S}_1(\sigma_i, \lambda_i, \mu_i)$,

(b) $\lim_{n\to\infty} \mathbf{S}_n^{-1} = \mathbf{O}$, $\limsup_{n\to\infty} \sum_{i=1}^{n} \left\| \mathbf{S}_n^{-1}\mathbf{x}_i\mathbf{x}_i^T \right\| < \infty$, $\sum_{i=1}^{\infty} \sigma_i \left\|\mathbf{S}_i^{-1}\mathbf{x}_i\right\| \log \left|\mathbf{S}_i^{-1}\mathbf{x}_i\right|_{\min}$

*converges and* $\left\| \sum_{i=1}^{\infty} \mu_i \mathbf{S}_i^{-1}\mathbf{x}_i \right\| < \infty$,

*then* $\widetilde{\boldsymbol{\beta}}$ *is strongly consistent.*

***Proof.*** One more time, making useful the relation (1.2) it will be sufficient to prove the almost sure convergence of the series

$$\sum_{i=1}^{\infty} \mathbf{S}_i^{-1} \mathbf{x}_i e_i \tag{3.1}$$

since the Kronecker's lemma (matrix case) establishes the thesis. Each component of the vector $\sum_{i=1}^{n} \mathbf{S}_i^{-1} \mathbf{x}_i e_i$ has distribution,

$$\mathcal{S}_1 \left( \sum_{i=1}^{n} \sigma_i \left| a_{ij} \right|, \frac{\sum_{i=1}^{n} \sigma_i \left| a_{ij} \right| \lambda_i \text{sign}\ (a_{ij})}{\sum_{i=1}^{n} \sigma_i \left| a_{ij} \right|}, \sum_{i=1}^{n} \left( \mu_i - \frac{2\sigma_i \lambda_i}{\pi} \log \left| a_{ij} \right| \right) a_{ij} \right)$$

with $j = 1, \ldots, \kappa$ and where $a_{ij}$ is the $j$-th component of $\mathbf{a}_i := \mathbf{S}_i^{-1} \mathbf{x}_i$. The almost sure convergence of the series (3.1) is a consequence of the Proposition 2.1 provided that $\lim_{i \to \infty} \mathbf{S}_i^{-1} \mathbf{x}_i = \mathbf{0}$. ■

**Remark 4.** If the random errors $e_1, e_2, \ldots$ are independent, $\alpha$-stable and symmetric about the origin then the strong consistency of the LSE is valid assuming

$$\lim_{n \to \infty} \mathbf{S}_n^{-1} = \mathbf{O}, \quad \limsup_{n \to \infty} \sum_{i=1}^{n} \left\| \mathbf{S}_n^{-1} \mathbf{x}_i \mathbf{x}_i^T \right\| < \infty$$

$$\text{and} \quad \sum_{i=1}^{\infty} \left( \sigma_i \left\| \mathbf{S}_i^{-1} \mathbf{x}_i \right\| \right)^{\alpha} < \infty.$$

## REFERENCES

[1] B.D.O. Anderson and J.B. Moore, *On martingales and least squares linear system identification*, Technical report EE7522 (1975).

[2] B.D.O. Anderson and J.B. Moore, *A matrix Kronecker lemma*, Linear Algebra and its Applications **15** (1976), 227–234.

[3] P. Billingsley, *Probability and Measure*, (third edition) John Wiley & Sons 1995.

[4] Y.S. Chow and H. Teicher, *Probability Theory: Independence*, Interchangeability, Martingales, Springer 1997.

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

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

[7] W. Feller, *An Introduction to Probability Theory and Its Applications - Volume I*, (third edition) John Wiley & Sons 1968.

[8] W. Feller, *An Introduction to Probability Theory and Its Applications - Volume II*, (second edition) John Wiley & Sons 1971.

[9] I.A. Ibragimov and Yu.V. Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen (1971).

[10] M. Jin, *Some new results of the strong consistency of multiple regression coefficients*, p. 514–519 in: *"Proceedings of the Second Asian Mathematical Conference 1995"* (Tangmanee, S. & Schulz, E. eds.), World Scientific.

[11] M. Jin and X. Chen, *Strong consistency of least squares estimate in multiple regression when the error variance is infinite*, Stat. Sin. **9** (1) (1999), 289–296.

[12] T.L. Lai, H. Robbins and C.Z. Wei, *Strong consistency of least squares estimates in multiple regression*, Proc. Natl. Acad. Sci. USA **75** (7) (1978), 3034–3036.

[13] 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.

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

[15] J.T. Mexia and J. Lita da Silva, *Least squares estimator consistency: a geometric approach*, Discussiones Mathematicae - Probability and Statistics **26** (1) (2006), 19–45.

[16] G. Samorodnitsky and M.S. Taqqu, *Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman & Hall 1994.

[17] V.V. Uchaikin and V.M. Zolotarev, *Chance and Stability, Stable Distributions and Their Applications*, Ultrech 1999.

[18] V.M. Zolotarev, *One-Dimensional Stable Distributions*, American Mathematical Society, Providence, R.I. 1986.

Received 21 March 2007
Revised 18 November 2007