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

JOÃO TIAO MEXIA

Mathematics Department, Faculty of Science and Technology
New University of Lisbon
Monte da Caparica 2829–516 Caparica, Portugal
{	exttt{e-mail:}} parc@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
{	exttt{e-mail:}} jls@fct.unl.pt

Abstract

Let $Y_i = x_i^T \beta + e_i$, $1 \leq i \leq n$, $n \geq 1$ 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 $\hat{\beta}$ of $\beta$ under additional assumptions on the non-random sequence $x_1, x_2, \ldots$ 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 instance [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 $1 \leq r < 2$. 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,

\begin{equation}
Y_i = x_i^T \beta + e_i, \quad 1 \leq i \leq n, \quad n \geq 1,
\end{equation}

where $x_1, x_2, \ldots$ are known non-random real $\kappa$-vectors,

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

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

Setting

$$
S_n := x_1 x_1^T + \ldots + x_n x_n^T
$$

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

\[
\hat{\beta} = \beta + S_n^{-1} \sum_{i=1}^{n} x_i e_i
\]

so that, \( \hat{\beta} \) is strongly consistent if and only if \( S_n^{-1} \sum_{i=1}^{n} x_i e_i \overset{\text{as}}{\longrightarrow} 0 \).

2. Auxiliary tools

The following notational conventions will be adopted throughout the paper. Given a vector \( x \in \mathbb{R}^n \), the euclidean vector norm will be denoted by

\[
\|x\| := \sqrt{x^T x}
\]

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

\[
\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}.
\]

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

\[
c_1 Z_1 + c_2 Z_2 \overset{d}{=} a Z + b
\]
where $Z_1$ and $Z_2$ are independent copies of $Z$.\footnote{The notation $X \overset{d}{=} Y$ means that the random variables $X$ and $Y$ agree in distribution, that is, $F_X(x) = F_Y(x)$, $x \in \mathbb{R}$.} A random variable $Z$ is called \textit{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 $0 < \alpha \leq 2$, $\sigma \geq 0$, $-1 \leq \lambda \leq 1$ and $\mu$ real such that its characteristic function has the following form,

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

where

\[ \text{sign } t := \begin{cases} 
1 \text{ if } t > 0 \\
0 \text{ if } t = 0 \\
-1 \text{ if } t < 0
\end{cases} \]

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

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

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

\[ Z \sim S_\alpha(\sigma, \lambda, \mu) \]

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

Let us remark that a random variable $Z$ concentrated at one point is always stable, i.e., a constant $\mu$ has degenerate distribution $S_\alpha(0, 0, \mu)$ for
any $0 < \alpha \leq 2$. 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

$$S_2(\sigma, \lambda, \mu) = N(\mu, 2\sigma^2) \quad (\sigma > 0).$$

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

$$Z \sim S_2(\sigma, \lambda, \mu) \quad (\sigma > 0)$$

has probability density function

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

The remaining special cases of probability densities of $\alpha$-stable random variables in closed form are: the Cauchy distribution $S_1(\sigma, 0, \mu)$ whose density is

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

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

$$f(x) = \sqrt{\frac{\sigma}{2\pi(x-\mu)^3}} e^{-\frac{\sigma^2}{2(x-\mu)}}, \quad x > \mu.$$
Some basic properties of stable distributions are listed below: (see [16])

(1) Let $Z_i$ and $Z_2$ be independent random variables with $Z_i \sim S_\alpha(\sigma_i, \lambda_i, \mu_i), i = 1, 2$. Then

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

(2) Let $Z \sim S_\alpha(\sigma, \lambda, \mu)$ and let $a$ be a real constant. Then $Z + a \sim S_\alpha(\sigma, \lambda, \mu + a)$.

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

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

$$aZ \sim S_1 \left( |a| \sigma, \text{sign}(a)\lambda, a\mu - \frac{2}{\pi} \sigma \lambda a \log |a| \right) \quad \text{if} \quad \alpha = 1.$$ 

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

(5) $Z \sim S_1(\sigma, \lambda, \mu)$ is strictly stable if and only if $\lambda = 0$.

(6) $Z \sim S_\alpha(\sigma, \lambda, \mu)$ is symmetric about $\mu$ if and only if $\lambda = 0$.

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

$$Z \overset{d}{=} \left( \frac{1 + \lambda}{2} \right)^{1/\alpha} A - \left( \frac{1 - \lambda}{2} \right)^{1/\alpha} B \quad \text{if} \quad \alpha \neq 1.$$
\[ Z \overset{d}{=} \left( \frac{1 + \lambda}{2} \right) A - \left( \frac{1 - \lambda}{2} \right) B + \]
\[ + \sigma \left[ \left( \frac{1 + \lambda}{\pi} \right) \log \left( \frac{1 + \lambda}{2} \right) - \left( \frac{1 - \lambda}{\pi} \right) \log \left( \frac{1 - \lambda}{2} \right) \right] \text{ if } \alpha = 1. \]

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 \( x_i, \ i = 1, 2, \ldots \) be a sequence of real \( \kappa \)-vectors, \( A_i \) a sequence of real symmetric nonsingular \( \kappa \times \kappa \) matrices with \( A_{i+1} \geq A_i > 0 \) and \( P_i \) a sequence of nonsingular \( \kappa \times \kappa \) matrices. Suppose that \( r_i = \sum_{j=1}^{\infty} A_{j-1}^{-1} x_j \) exists (and is finite). Then:

(i) If \( A_\infty := \lim_{n \to \infty} A_n \) exists and is finite, then \( \lim_{n \to \infty} A_n^{-1} \sum_{i=1}^{n} x_i \) exists and is finite.

(ii) If \( \lim_{n \to \infty} \frac{1}{\text{tr} (A_n)} = 0 \) then \( \lim_{n \to \infty} \frac{1}{\text{tr} (A_n)} \sum_{i=1}^{n} x_i = 0. \)

(iii) If \( \lim_{n \to \infty} A_n^{-1} = 0 \) \( \quad \lim_{n \to \infty} P_n r_n = 0 \) and

\[ \limsup_{n \to \infty} \sum_{i=1}^{n} \| A_n^{-1} (A_i - A_{i-1}) P_i^{-1} \| < \infty \text{ then } \lim_{n \to \infty} A_n^{-1} \sum_{i=1}^{n} x_i = 0. \]

**Proof.** The proof can be founded in [1] on page 27. \[ \blacksquare \]
Remark 1. The condition \( \limsup_{n \to \infty} \sum_{i=1}^{n} \| A_n^{-1}(A_i - A_{i-1})P_i^{-1} \| < \infty \)

is implied by the much stronger condition

\[
\limsup_{n \to \infty} \frac{\nu_{\text{max}}(A_n)}{\nu_{\text{min}}(A_n)} < \infty.
\]

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 \( X_i \sim S_\alpha(\sigma_i, \lambda_i, \mu_i) \).

**Proposition 2.1.** Let \( \{X_i, i \geq 1\} \) be independent random variables \( X_i \sim S_\alpha(\sigma_i, \lambda_i, \mu_i) \), \( 0 < \alpha \leq 2 \). Then,

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

**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,

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

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

we get, for all \( n \in \mathbb{N} \),
\[ \sum_{i=1}^{n} X_i \sim S_{\alpha}(a_n, b_n, c_n) \]

provided that \( X_i \sim S_{\alpha}(\sigma_i, \lambda_i, \mu_i) \).

Therefore, we can write

\[ \sum_{i=1}^{n} X_i \overset{d}{=} a_n Y_n + c_n \quad \text{if} \quad \alpha \neq 1 \]

and

\[ \sum_{i=1}^{n} X_i \overset{d}{=} a_n Y_n + \frac{2}{\pi} b_n a_n \log a_n + c_n \quad \text{if} \quad \alpha = 1 \]

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

\[ Y_n \overset{d}{=} \left( \frac{1 + b_n}{2} \right)^{1/\alpha} A_n - \left( \frac{1 - b_n}{2} \right)^{1/\alpha} B_n \quad \text{if} \quad \alpha \neq 1 \]

and

\[ Y_n \overset{d}{=} \left( \frac{1 + b_n}{2} \right) A_n - \left( \frac{1 - b_n}{2} \right) B_n + \frac{1 + b_n}{\pi} \log \left( \frac{1 + b_n}{2} \right) - \frac{1 - b_n}{\pi} \log \left( \frac{1 - b_n}{2} \right) \quad \text{if} \quad \alpha = 1. \]
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 \( \delta_n(1,1,0) \) 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 \( \sum_{i=1}^{n} X_i \xrightarrow{a.s.} S \) (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

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

there exist a subsequence \((b_{\eta_n})\) 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

\[
\lim_{n \to \infty} \sum_{i=1}^{n} \sigma_i^\alpha < \infty
\]
Sufficient conditions for the strong consistency of ... 37

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

$$\sum_{i=1}^{\infty} \sigma_i^0 < \infty$$

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

$$\frac{\lambda_1 \sigma_1^0 + \ldots + \lambda_n \sigma_n^0 + \ldots}{\sigma_1^0 + \ldots + \sigma_n^0 + \ldots}$$

converge, \(\sum_{i=1}^{\infty} \mu_i\) converge.

\[\square\]

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 \(E(e_1) = 0\) and \(0 < E(e_1^2) < \infty\) then a sufficient condition for the strong consistency of \(\hat{\beta}\) is,

$$S_n^{-1} = \left(\sum_{i=1}^{n} x_i x_i^T\right)^{-1} \rightarrow O$$

as \(n \rightarrow \infty\). 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. \(S_2(\sigma, \lambda, 0)\),

(b) \(\lim_{n \rightarrow \infty} S_n^{-1} = O\),

then \(\hat{\beta}\) is strongly consistent.
Remark 2. Let us note that the necessity of the condition \( \lim_{n \to \infty} S_n^{-1} = 0 \) 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. \( \mathcal{N}_2(\sigma, \lambda, \mu) \) with \( \mu \neq 0 \),

(b) \( \lim_{n \to \infty} S_n^{-1} = 0 \), \( \limsup_{n \to \infty} \sum_{i=1}^{n} \| S_n^{-1} x_i x_i^T \| < \infty \) and \( \sum_{i=1}^{\infty} S_n^{-1} x_i \| < \infty \),

then \( \tilde{\beta} \) is strongly consistent.

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

\[
\tilde{\beta} = \beta + S_n^{-1} \sum_{i=1}^{n} x_i (e_i - \mu) + \mu S_n^{-1} \sum_{i=1}^{n} x_i.
\]

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 \( 1 \leq r < 2 \). 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. $\mathcal{G}_\alpha(\sigma, \lambda, \mu)$,

(b) $\lim_{n \to \infty} S_n^{-1} = \mathcal{O}$, $\limsup_{n \to \infty} \sum_{i=1}^{n} \| S_n^{-1} x_i x_i^T \| < \infty$ and $\sum_{i=1}^{\infty} \| S_i^{-1} x_i \|^{\min(\alpha, 1)} < \infty$.

then $\tilde{\beta}$ is strongly consistent.

Proof. The identity (1.2) guarantees,

$$\tilde{\beta} = \beta + S_n^{-1} \sum_{i=1}^{n} x_i e_i = \beta + S_n^{-1} \sum_{i=1}^{n} x_i e_i^* + \mu S_n^{-1} \sum_{i=1}^{n} x_i$$

with $e_i^* \sim \mathcal{G}_\alpha(\sigma, \lambda, 0)$. On the one hand, by Kronecker's lemma (matrix case) we have $\mu S_n^{-1} \sum_{i=1}^{n} x_i \xrightarrow{n \to \infty} 0$ provided that $\| \sum_{i=1}^{\infty} S_i^{-1} x_i \| < \infty$ and $\limsup_{n \to \infty} \sum_{i=1}^{n} \| S_n^{-1} x_i x_i^T \| < \infty$. On the other hand, each component of the vector $\sum_{i=1}^{n} S_i^{-1} x_i e_i^*$ has distribution

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

where $a_{ij}$ is the $j$-th component of the vector $a_i := S_i^{-1} x_i$. Since

$$\sum_{i=1}^{\infty} |a_{ij}|^\alpha \leq \sum_{i=1}^{\infty} \| S_i^{-1} x_i \|^{\alpha}$$

we obtain the almost sure convergence of $\sum_{i=1}^{\infty} S_i^{-1} x_i e_i^*$ through the Proposition 2.1 and Kronecker's lemma (matrix case) permit us to conclude

$S_n^{-1} \sum_{i=1}^{n} x_i e_i^*$ a.s. $\to 0$. ■
On the next result we will use the notation $|x|_{\min} := \min(|x_1|, \ldots, |x_\kappa|)$ where $x$ 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. $S_1(\sigma, \lambda, \mu),$

(b) $\lim_{n \to \infty} S_n^{-1} = 0$, $\limsup_{n \to \infty} \sum_{i=1}^{n} \|S_n^{-1} x_i x_i^T\| < \infty$ and

$$\sum_{i=1}^{\infty} \|S_i^{-1} x_i\| \log |S_i^{-1} x_i|_{\min} \text{ converges},$$

then $\tilde{\beta}$ is strongly consistent.

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

$$\tilde{\beta} = \beta + S_n^{-1} \sum_{i=1}^{n} x_i e_i = \beta + S_n^{-1} \sum_{i=1}^{n} x_i e_i^* + \mu S_n^{-1} \sum_{i=1}^{n} x_i$$

with $e_i^* \sim S_1(\sigma, \lambda, 0)$. Each component of the vector $\sum_{i=1}^{n} S_i^{-1} x_i e_i^*$ has distribution

$$S_1 \left( \sigma \sum_{i=1}^{n} |a_{ij}|, \frac{\lambda}{n} \sum_{i=1}^{n} \text{sign}(a_{ij}) |a_{ij}| \right), \quad -2\sigma \lambda \sum_{i=1}^{n} a_{ij} \log |a_{ij}|, \quad j = 1, \ldots, \kappa.$$
where $a_{ij}$ is the $j$-th component of $a_i := S_i^{-1} x_i$. Therefore, we have

$$S_n^{-1} \sum_{i=1}^{n} x_i e_i^* \xrightarrow{a.s.} 0$$

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

Since

$$\left\| \sum_{i=1}^{\infty} S_i^{-1} x_i \right\| < \infty \quad \text{and} \quad \limsup_{n \to \infty} \sum_{i=1}^{n} \left\| S_i^{-1} x_i x_i^T \right\| < \infty$$

we get $\mu S_n^{-1} \sum_{i=1}^{n} x_i \to 0$ as $n \to \infty$ and the thesis is established.

\[ \blacksquare \]

**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 $\lim_{n \to \infty} S_n^{-1} = 0$, $\limsup_{n \to \infty} \sum_{i=1}^{n} \| S_n^{-1} x_i x_i^T \| < \infty$ and $\sum_{i=1}^{\infty} \| S_i^{-1} x_i \|^{2/\alpha} < \infty$.

The previous results can be extended to the case where the random variables $e_1, e_2, \ldots$ are (only) independent and $e_i \sim S_0(\sigma_i, \lambda_i, \mu_i)$. 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 $0 < \alpha \leq 2$, $\alpha \neq 1$.

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

(a) the random variables $e_1, e_2, \ldots$ are independent such that $e_i \sim S_0(\sigma_i, \lambda_i, \mu_i)$,
(b) \( \lim_{n \to \infty} S_n^{-1} = O, \limsup_{n \to \infty} \sum_{i=1}^{n} \|S_i^{-1}x_i x_i^T\| < \infty, \sum_{i=1}^{\infty} \left( \sigma_i \|S_i^{-1}x_i\| \right) < \infty \)

and \( \left\| \sum_{i=1}^{\infty} \mu_i S_i^{-1} x_i \right\| < \infty, \)

then \( \tilde{\beta} \) is strongly consistent.

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

\[
S_\alpha \left( \sum_{i=1}^{n} (\sigma_i |a_{ij}|)^\alpha \right)^{1/\alpha} \frac{\sum_{i=1}^{n} (\sigma_i |a_{ij}|)^\alpha \lambda_i \text{sign}(a_{ij})}{\sum_{i=1}^{n} (\sigma_i |a_{ij}|)^\alpha}, \quad j = 1, \ldots, \kappa
\]

where \( a_{ij} \) is the \( j \)-th component of \( a_i := S_i^{-1} x_i \) so that, from Proposition 2.1 we obtain the almost sure convergence of \( \sum_{i=1}^{\infty} S_i^{-1} x_i e_i \).

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 \( e_i \sim S_1(\sigma_i, \lambda_i, \mu_i) \),

(b) \( \lim_{n \to \infty} S_n^{-1} = O, \limsup_{n \to \infty} \sum_{i=1}^{n} \|S_i^{-1}x_i x_i^T\| < \infty, \sum_{i=1}^{\infty} \sigma_i \|S_i^{-1}x_i\| \log |S_i^{-1}x_i|_{\min} \)

converges and \( \left\| \sum_{i=1}^{\infty} \mu_i S_i^{-1} x_i \right\| < \infty, \)

then \( \tilde{\beta} \) 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

\[ \sum_{i=1}^{\infty} S_i^{-1} x_i e_i \]

since the Kronecker’s lemma (matrix case) establishes the thesis. Each component of the vector \( \sum_{i=1}^{n} S_i^{-1} x_i e_i \) has distribution,

\[
S_1 \left( \sum_{i=1}^{n} \sigma_i |a_{ij}|, \frac{\sum_{i=1}^{n} \sigma_i |a_{ij}| \lambda_i \text{sign} (a_{ij})}{\sum_{i=1}^{n} \sigma_i |a_{ij}|}, \sum_{i=1}^{n} \left( \mu_i - \frac{2\sigma_i \lambda_i}{\pi} \log |a_{ij}| \right) a_{ij} \right)
\]

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

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

\[
\lim_{n \to \infty} S_n^{-1} = O, \quad \limsup_{n \to \infty} \sum_{i=1}^{n} \| S_n^{-1} x_i x_i^T \| < \infty
\]

and

\[
\sum_{i=1}^{\infty} \left( \sigma_i \| S_i^{-1} x_i \| \right)^\alpha < \infty.
\]

References


Received 21 March 2007
Revised 18 November 2007