Almost sure convergence for weighted sums of pairwise PQD random variables

João Lita da Silva[1]
Department of Mathematics and GeoBioTec
Faculty of Sciences and Technology
NOVA University of Lisbon
Quinta da Torre, 2829-516 Caparica, Portugal

Abstract

We obtain strong laws of large numbers of Marcinkiewicz-Zygmund's type for weighted sums of pairwise positively quadrant dependent random variables stochastically dominated by a random variable $X \in \mathscr{L}_p$, $1 \leqslant p < 2$. We use our results to establish the strong consistency of estimators which emerge from regression models having pairwise positively quadrant dependent errors.

Key words and phrases: Strong law of large numbers, weighted sum, positively quadrant dependent random variables, regression models, strong consistency

2010 Mathematics Subject Classification: 60F15, 62J05, 62J07

1 Introduction

The classical Kolmogorov's strong law of large numbers is, perhaps, the most famous strong limit theorem in Probability Theory. Originally presented in 1933 by Andrei Nikolaevich Kolmogorov (see Kolmogorov 1933), it can be stated as follows: if $X_1, X_2, \ldots$ is a sequence of independent and identically distributed random variables such that $\mathbb{E}\, X_1$ exists then

$$\frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{\text{a.s.}} \mathbb{E}\, X_1.$$

In Walk 2005, an elegant short proof for a Kolmogorov's strong law of large numbers under very general assumptions was given by means of Tauberian theorems. In this paper, we shall follow Walk's statement to establish a Kolmogorov's strong law of large numbers for weighted pairwise positively quadrant dependent random variables. This result improves the corresponding one announced in Louhichi 2000 for sequences of (positively) associated random variables. Further, we shall present strong law of large numbers of Marcinkiewicz-Zygmund's type for weighted pairwise positively quadrant dependent random variables by using a maximal inequality recently announced in Lita da Silva 2018 for these dependent structures.

It is important to stress out that in mathematical statistics, statistical physics or reliability theory, many stochastic models involving dependent random variables have arose over the

[1] *E-mail address:* `jfls@fct.unl.pt; joao.lita@gmail.com`

last decades (see Bulinski and Shashkin 2007, Hutchinson and Lai 1990). For instance, most bivariate distributions in reliability theory are positively quadrant dependent (see Hutchinson and Lai 1990), whence, any efforts to establish formal results for these (and other) dependent structures are always of interest.

In last section, we provide some statistical applications of our assertions, namely, in strong consistency of estimators which can be found in regression models. Our statements not only extend others established lately in this issue (see Lita da Silva and Mexia 2013 and Lita da Silva 2014), but also allow us to consider statistical models having dependent random variables making them more appropriate and realistic.

Throughout, $x \wedge y$ and $x \vee y$ will stand for $\min\{x, y\}$ and $\max\{x, y\}$, respectively. Associated to a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we shall consider the space $\mathscr{L}_p$ $(p > 0)$ of all measurable functions X (necessarily random variables) for which $\mathbb{E}|X|^p < \infty$. For any measurable function X we will define its positive and negative parts by $X^+ = X \vee 0$ and $X^- = (-X) \vee 0$, respectively. Given an event A we shall denote the indicator random variable of the event A by I_A. All over, the function $x \mapsto \log(|x| \vee \mathrm{e})$ will be denoted by $\operatorname{Log} x$. To make the computations be simpler looking, we shall employ the letter C to denote any positive constant that can be explicitly computed, which is not necessarily the same on each appearance; the symbol $C(p)$ has identical meaning with the additional information that the constant depends on p.

2 Almost sure convergence

In the past, many authors have considered limit theorems involving pairwise positively quadrant dependent random variables (see, Birkel 1989, Birkel 1993, Matuła 2005 or Newman 1984 among others). We shall proceed their study by establishing strong limits for weighted sums of these dependent random variables.

The concept of positively quadrant dependence for a pair of random variables due to Lehmann 1966 can be given for sequences of random variables as follows: a sequence $\{X_n,\, n \geqslant 1\}$ of random variables is said to be *pairwise positively quadrant dependent* (pairwise PQD) if

$$\mathbb{P}\{X_k \leqslant x_k, X_j \leqslant x_j\} - \mathbb{P}\{X_k \leqslant x_k\}\, \mathbb{P}\{X_j \leqslant x_j\} \geqslant 0$$

for all reals x_k, x_j and all positive integers k, j such that $k \neq j$.

Let X, Y be random variables and ℓ a positive constant. Throughout, we shall consider the function $g_\ell(t) := (t \wedge \ell) \vee (-\ell)$ and the covariance quantity

$$G_{X,Y}(t) := \operatorname{Cov}\big(g_t(X), g_t(Y)\big) = \int_{-t}^{t} \int_{-t}^{t} \Delta_{X,Y}(x, y)\, \mathrm{d}x\, \mathrm{d}y \tag{2.1}$$

where $\Delta_{X,Y}(x, y) := \mathbb{P}\{X \leqslant x, Y \leqslant y\} - \mathbb{P}\{X \leqslant x\}\, \mathbb{P}\{Y \leqslant y\}$. Further, a random sequence $\{X_n,\, n \geqslant 1\}$ is *stochastically dominated* by a random variable X if there exists a constant $C > 0$ such that $\sup_{n \geqslant 1} \mathbb{P}\{|X_n| > t\} \leqslant C\, \mathbb{P}\{|X| > t\}$ for all $t > 0$ (see, for instance, Lita da Silva 2015).

Now, we state and prove a Kolmogorov's strong law of large numbers for weighted pairwise positively quadrant dependent random variables.

Theorem 1 *Let $\{X_n,\, n \geqslant 1\}$ be a sequence of pairwise PQD random variables stochastically dominated by a random variable $X \in \mathscr{L}_1$. If $\{a_n\}$ is a sequence of constants satisfying* $\sup_{n\geqslant 1} n^{-1} \sum_{k=1}^{n} a_k^2 < \infty$ *and*

$$\sum_{1\leqslant k<j\leqslant\infty} |a_k a_j| \int_j^\infty t^{-3} G_{X_k,X_j}(t)\,\mathrm{d}t < \infty \tag{2.2}$$

then $\sum_{k=1}^{n} a_k(X_k - \mathbb{E}\,X_k)/n \xrightarrow{\text{a.s.}} 0.$

Proof. By writing $a_n = a_n^+ - a_n^-$, we may suppose without loss of generality that a_n is nonnegative for each n. Setting

$$\begin{aligned} X_n' &:= g_n(X_n), \\ X_n'' &:= X_n - g_n(X_n), \\ Y_n' &:= (n \wedge X_n) \vee 0, \\ Y_n'' &:= (-n \vee X_n) \wedge 0 \end{aligned} \tag{2.3}$$

we have $X_n' = Y_n' + Y_n''$ and $X_n = X_n' + X_n''$. From Lemma 1 of Lehmann 1966 we get

$$\begin{aligned} \operatorname{Cov}(Y_k', Y_j'') &\geqslant 0, \\ \operatorname{Cov}(Y_k'', Y_j') &\geqslant 0, \\ \operatorname{Cov}(Y_k'', Y_j'') &\geqslant 0 \end{aligned}$$

because $t \mapsto (n \wedge t) \vee 0$ and $t \mapsto (-n \vee t) \wedge 0$ are nondecreasing functions. Thus,

$$\begin{aligned} &\sum_{k,j=1}^{n} \operatorname{Cov}(a_k Y_k', a_j Y_j') = \\ &\quad = \sum_{k,j=1}^{n} a_k a_j \operatorname{Cov}(Y_k', Y_j') \\ &\quad \leqslant \sum_{k,j=1}^{n} a_k a_j \left[\operatorname{Cov}(Y_k', Y_j') + \operatorname{Cov}(Y_k', Y_j'') + \operatorname{Cov}(Y_k'', Y_j') + \operatorname{Cov}(Y_k'', Y_j'')\right] \\ &\quad = \sum_{k,j=1}^{n} a_k a_j \operatorname{Cov}(X_k', X_j') \\ &\quad = \sum_{k=1}^{n} a_k^2 \mathbb{V}(X_k') + 2 \sum_{1\leqslant k<j\leqslant n} a_k a_j \operatorname{Cov}(X_k', X_j'). \end{aligned}$$

According to Abel's identity, we have

$$\sum_{k=j}^{\infty} \frac{a_k^2}{k^2} \leqslant \frac{2}{j} \cdot \sup_{k\geqslant 1} \frac{1}{k} \sum_{m=1}^{k} a_m^2 \leqslant \frac{C}{j}$$

for any $j \geqslant 1$, so that Lemma 1 of Lita da Silva 2015 implies

$$\sum_{n=1}^{\infty} \frac{1}{n^3} \sum_{k=1}^{n} a_k^2 \mathbb{V}(X_k') = \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \frac{a_k^2 \mathbb{V}(X_k')}{n^3}$$

$$
\begin{aligned}
&\leqslant C\sum_{k=1}^{\infty}\frac{a_k^2\mathbb{V}(X_k')}{k^2}\\
&\leqslant C\sum_{k=1}^{\infty}\frac{a_k^2\mathbb{E}|X_k'|^2}{k^2}\\
&\leqslant C\sum_{k=1}^{\infty}\frac{a_k^2\left(\mathbb{E}\,X_k^2 I_{\{|X_k|\leqslant k\}}+k^2\mathbb{P}\left\{|X_k|>k\right\}\right)}{k^2}\\
&\leqslant C\sum_{k=1}^{\infty}\frac{a_k^2\left(\mathbb{E}\,X^2 I_{\{|X|\leqslant k\}}+k^2\mathbb{P}\left\{|X|>k\right\}\right)}{k^2}\\
&= C\sum_{k=1}^{\infty}\frac{a_k^2}{k^2}\int_0^k u\,\mathbb{P}\left\{|X|>u\right\}\mathrm{d}u\\
&= C\int_0^{\infty} u\,\mathbb{P}\left\{|X|>u\right\}\sum_{\{k:\,k>u\}}\frac{a_k^2}{k^2}\,\mathrm{d}u\\
&\leqslant C\int_0^{\infty}\mathbb{P}\left\{|X|>u\right\}\mathrm{d}u\\
&= C\,\mathbb{E}\,|X|<\infty.
\end{aligned}
$$

Since

$$
\begin{aligned}
&\sum_{n=1}^{\infty}\frac{1}{n^3}\sum_{1\leqslant k<j\leqslant n}a_k a_j\mathrm{Cov}(X_k',X_j')=\\
&\quad=\sum_{n=1}^{\infty}\frac{1}{n^3}\sum_{1\leqslant k<j\leqslant n}a_k a_j\int_{-k}^{k}\int_{-j}^{j}\Delta_{X_k,X_j}(x,y)\,\mathrm{d}x\mathrm{d}y\\
&\quad\leqslant\sum_{n=1}^{\infty}\sum_{1\leqslant k<j\leqslant n}\frac{a_k a_j G_{X_k,X_j}(n)}{n^3}\\
&\quad=\sum_{1\leqslant k<j\leqslant\infty}a_k a_j\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{n=1}^{\infty}\frac{I_{\{n\geqslant|x|\vee|y|\vee j\}}}{n^3}\Delta_{X_k,X_j}(x,y)\,\mathrm{d}x\mathrm{d}y\\
&\quad\leqslant C\sum_{1\leqslant k<j\leqslant\infty}a_k a_j\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\Delta_{X_k,X_j}(x,y)}{(|x|\vee|y|\vee j)^2}\,\mathrm{d}x\mathrm{d}y\\
&\quad=C\sum_{1\leqslant k<j\leqslant\infty}a_k a_j\int_j^{\infty}t^{-3}G_{X_k,X_j}(t)\,\mathrm{d}t<\infty
\end{aligned}
\tag{2.4}
$$

by Lemma 4 of Louhichi 2000, we obtain

$$
\sum_{n=1}^{\infty}\frac{1}{n^3}\sum_{k,j=1}^{n}\mathrm{Cov}(a_kY_k',a_jY_j')\leqslant\sum_{n=1}^{\infty}\frac{a_k^2}{n^3}\sum_{k=1}^{n}\mathbb{V}(X_k')+2\sum_{n=1}^{\infty}\frac{1}{n^3}\sum_{1\leqslant k<j\leqslant n}a_k a_j\mathrm{Cov}(X_k',X_j')<\infty.
$$

On the other hand, $a_nY_n'\geqslant 0$ and Lemma 1 of Lita da Silva 2015 yields

$$
\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(a_kY_k')\leqslant\frac{1}{n}\sum_{k=1}^{n}a_k\mathbb{E}|X_k'|
$$

$$\leqslant \frac{C}{n}\sum_{k=1}^{n} a_k \left(\mathbb{E}|X|I_{\{|X|\leqslant k\}} + k\mathbb{P}\left\{|X|>k\right\}\right)$$
$$\leqslant \frac{C\,\mathbb{E}|X|}{n}\sum_{k=1}^{n} a_k$$
$$\leqslant C\,\mathbb{E}|X|\left(\frac{1}{n}\sum_{k=1}^{n} a_k^2\right)^{1/2}$$
$$\leqslant C\,\mathbb{E}|X| < \infty$$

which ensures

$$\frac{1}{n}\sum_{k=1}^{n} a_k(Y_k' - \mathbb{E}\,Y_k') \xrightarrow{\text{a.s.}} 0 \tag{2.5}$$

(see, for instance, Remark 3 of Walk 2005). Noting that $a_n Y_n'' \leqslant 0$, $\sum_{k=1}^{n}\mathbb{E}(-a_k Y_k'')/n \leqslant \sum_{k=1}^{n} a_k \mathbb{E}|X_k'|/n < \infty$ and

$$\sum_{k,j=1}^{n}\mathrm{Cov}(-a_k Y_k'', -a_k Y_j'') = \sum_{k,j=1}^{n}\mathrm{Cov}(a_k Y_k'', a_k Y_j'') \leqslant \sum_{k,j=1}^{n} a_k a_j \mathrm{Cov}(X_k', X_j')$$

one can argue as above to conclude that

$$\frac{1}{n}\sum_{k=1}^{n} a_k\left(\mathbb{E}\,Y_k'' - Y_k''\right) \xrightarrow{\text{a.s.}} 0. \tag{2.6}$$

Hence, (2.5) and (2.6) yield $\sum_{k=1}^{n} a_k(X_k' - \mathbb{E}\,X_k')/n \xrightarrow{\text{a.s.}} 0$. It remains to prove

$$\frac{1}{n}\sum_{k=1}^{n} a_k(X_k'' - \mathbb{E}\,X_k'') \xrightarrow{\text{a.s.}} 0. \tag{2.7}$$

Since

$$\sum_{n=1}^{\infty}\mathbb{P}\left\{X_n \neq X_n'\right\} = \sum_{n=1}^{\infty}\mathbb{P}\left\{|X_n|>n\right\} \leqslant C\sum_{n=1}^{\infty}\mathbb{P}\left\{|X|>n\right\} \leqslant C\,\mathbb{E}|X| < \infty$$

it follows that a.s. $X_n = X_n'$ for all but a finite number of values of n, entailing

$$\frac{1}{n}\sum_{k=1}^{n} a_k X_k'' \xrightarrow{\text{a.s.}} 0.$$

From the dominated convergence theorem we have $\mathbb{E}|X|I_{\{|X|>n\}} = o(1)$ as $n\to\infty$, implying

$$\limsup_{n\to\infty}\left|\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(a_k X_k'')\right| \leqslant \limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} a_k\mathbb{E}|X_k''|$$
$$\leqslant \limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} a_k\mathbb{E}|X_k|I_{\{|X_k|>k\}}$$

$$
\begin{aligned}
&\leqslant \limsup_{n\to\infty} \left(\frac{1}{n}\sum_{k=1}^{n} a_k^2\right)^{1/2} \left(\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}^2|X_k| I_{\{|X_k|>k\}}\right)^{1/2} \\
&\leqslant \limsup_{n\to\infty} C \left(\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}^2|X| I_{\{|X|>k\}}\right)^{1/2} \\
&= 0
\end{aligned}
$$

and (2.7) holds establishing the thesis. □

Remark 1 We observe that condition (2.2) can be replaced by the weaker condition

$$
\sum_{1\leqslant k<j\leqslant\infty} \frac{|a_k a_j|}{j^2} \int_{-k}^{k}\int_{-j}^{j} \Delta_{X_k,X_j}(x,y)\,\mathrm{d}x\mathrm{d}y < \infty \tag{2.8}
$$

by waiving Lemma 4 of Louhichi 2000 in the upper bound (2.4). When $a_n = 1$ for all n, Theorem 1 equipped with (2.8) instead of (2.2) extends Corollary 1 of Lita da Silva 2018 to $p = 1$ and it corresponds to Theorem 3 of Chen and Sung 2019.

Our Theorem 1 extends also Theorem 1 of Louhichi 2000 to sequences of pairwise PQD random variables when $p = 1$. Recall that positively quadrant dependent random variables are not necessarily (positively) associated (see, for instance, Esary et al. 1967 or Tong 1980). Furthermore, the normalising constants in Theorem 1 improve the considered ones in Theorem 2 of Lita da Silva 2018 for $p = 1$. It is worthy to note that for the special case $p = 1$, the previous approach leads to sharped results discarding the direct use of any maximal inequality which is, in fact, the key ingredient in both Louhichi 2000 and Lita da Silva 2018. In particular, Theorem 1 with weights $a_n = 1$ for all n does not require the finiteness of the variance in each random variable unlike Theorem 1 of Birkel 1989.

The statement below gives us the almost sure convergence for weighted sums of pairwise PQD random variables when the moment condition of the random variable X in Theorem 1 is strengthened.

Theorem 2 *Let $1 < p < 2$ and $\{X_n,\, n \geqslant 1\}$ be a sequence of pairwise PQD random variables stochastically dominated by a random variable $X \in \mathscr{L}_p$. If $\{a_n\}$ is a sequence of constants satisfying* $\sup_{n\geqslant 1} n^{-1}\sum_{k=1}^{n} a_k^2 < \infty$ *and*

$$
\sum_{1\leqslant k<j\leqslant\infty} |a_k a_j| \int_{\frac{j^{1/p}}{\mathrm{Log}^{2/p} j}}^{\infty} \frac{G_{X_k,X_j}(t)}{t^3 \mathrm{Log}^2 t}\,\mathrm{d}t < \infty, \tag{2.9}
$$

then $\sum_{k=1}^{n} a_k(X_k - \mathbb{E}\,X_k)/(n^{1/p}\mathrm{Log}^{2(p-1)/p}\, n) \xrightarrow{\text{a.s.}} 0$.

Proof. As in the proof of Theorem 1, we shall assume $a_n \geqslant 0$ for all n. Considering $X_n' := g_{n^{1/p}/\mathrm{Log}^{2/p} n}(X_n)$ and $X_n'' := X_n - g_{n^{1/p}/\mathrm{Log}^{2/p} n}(X_n)$ it follows that $\{a_n X_n',\, n \geqslant 1\}$ is a sequence of pairwise PQD random variables. From Lemma 1 of Lita da Silva 2018 we obtain, for each $\varepsilon > 0$ and a fixed n_0,

$$
\sum_{n=n_0}^{\infty} \frac{1}{n}\mathbb{P}\left\{\max_{1\leqslant k\leqslant n}\left|\sum_{j=1}^{k} a_j(X_j' - \mathbb{E}\,X_j')\right| > \varepsilon n^{1/p}\mathrm{Log}^{2(p-1)/p}\, n\right\} \leqslant
$$

$$
\begin{aligned}
&\leqslant \frac{C}{\varepsilon^2}\sum_{n=n_0}^{\infty}\sum_{j=1}^{n}\frac{\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\mathrm{Cov}\left[\sum_{i=1}^{j}a_i(X_i'-\mathbb{E}\,X_i'),a_j(X_j'-\mathbb{E}\,X_j')\right]\\
&\leqslant \frac{C}{\varepsilon^2}\sum_{n=n_0}^{\infty}\sum_{j=1}^{n}\frac{\mathrm{Log}^{2(2-p)/p}n\,\mathbb{E}(a_jX_j')^2}{n^{1+2/p}}+\frac{C}{\varepsilon^2}\sum_{n=n_0}^{\infty}\sum_{1\leqslant i<j\leqslant n}\frac{\mathrm{Log}^{2(2-p)/p}n\,\mathrm{Cov}(a_iX_i',a_jX_j')}{n^{1+2/p}}\\
&\leqslant \frac{C}{\varepsilon^2}\sum_{n=n_0}^{\infty}\sum_{j=1}^{n}\frac{a_j^2\,\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\left(\mathbb{E}\,X_j^2I_{\left\{|X_j|\leqslant\frac{j^{1/p}}{\mathrm{Log}^{2/p}j}\right\}}+\frac{j^{2/p}}{\mathrm{Log}^{4/p}j}\mathbb{P}\left\{|X_j|>\frac{j^{1/p}}{\mathrm{Log}^{2/p}j}\right\}\right)+\\
&\quad \frac{C}{\varepsilon^2}\sum_{n=n_0}^{\infty}\sum_{1\leqslant i<j\leqslant n}\frac{a_ia_j\,\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\int_{-i^{1/p}/\mathrm{Log}^{2/p}i}^{i^{1/p}/\mathrm{Log}^{2/p}i}\int_{-j^{1/p}/\mathrm{Log}^{2/p}j}^{j^{1/p}/\mathrm{Log}^{2/p}j}\Delta_{X_i,X_j}(x,y)\,\mathrm{d}x\mathrm{d}y\\
&\leqslant \frac{C}{\varepsilon^2}\sum_{n=n_0}^{\infty}\sum_{j=1}^{n}\frac{a_j^2\,\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\left(\mathbb{E}\,X_j^2I_{\left\{|X_j|\leqslant\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right\}}+\frac{n^{2/p}}{\mathrm{Log}^{4/p}n}\mathbb{P}\left\{|X_j|>\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right\}\right)+\\
&\quad \frac{C}{\varepsilon^2}\sum_{n=n_0}^{\infty}\sum_{1\leqslant i<j\leqslant n}\frac{a_ia_j\,\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\int_{-n^{1/p}/\mathrm{Log}^{2/p}n}^{n^{1/p}/\mathrm{Log}^{2/p}n}\int_{-n^{1/p}/\mathrm{Log}^{2/p}n}^{n^{1/p}/\mathrm{Log}^{2/p}n}\Delta_{X_i,X_j}(x,y)\,\mathrm{d}x\mathrm{d}y\\
&\leqslant \frac{C}{\varepsilon^2}\sum_{n=1}^{\infty}\sum_{j=1}^{n}\frac{a_j^2\,\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\left(\mathbb{E}\,X^2I_{\left\{|X|\leqslant\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right\}}+\frac{n^{2/p}}{\mathrm{Log}^{4/p}n}\mathbb{P}\left\{|X|>\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right\}\right)+\\
&\quad \frac{C}{\varepsilon^2}\sum_{n=1}^{\infty}\sum_{1\leqslant i<j\leqslant n}\frac{a_ia_j\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}G_{X_i,X_j}\left(\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right)\\
&\leqslant \frac{C}{\varepsilon^2}\sum_{n=1}^{\infty}\frac{\mathrm{Log}^{2(2-p)/p}n}{n^{2/p}}\mathbb{E}\,X^2I_{\left\{|X|\leqslant\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right\}}+\frac{C}{\varepsilon^2}\sum_{n=1}^{\infty}\frac{1}{\mathrm{Log}^2 n}\mathbb{P}\left\{|X|>\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right\}+\\
&\quad \frac{C}{\varepsilon^2}\sum_{n=1}^{\infty}\sum_{1\leqslant i<j\leqslant n}\frac{a_ia_j\,\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}G_{X_i,X_j}\left(\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right)
\end{aligned}
$$

since $\sup_{n\geqslant 1}\sum_{j=1}^{n}a_j^2/n<\infty$. Supposing

$$
A_j=\left\{(j-1)^{1/p}/\mathrm{Log}^{2/p}(j-1)<|X|\leqslant j^{1/p}/\mathrm{Log}^{2/p}j\right\},\quad j\geqslant 1 \tag{2.10}
$$

we have

$$
\begin{aligned}
\sum_{n=1}^{\infty}\frac{\mathrm{Log}^{2(2-p)/p}n}{n^{2/p}}\,\mathbb{E}\,X^2I_{\left\{|X|\leqslant\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right\}}&=\sum_{n=1}^{\infty}\sum_{j=1}^{n}\frac{\mathrm{Log}^{2(2-p)/p}n}{n^{2/p}}\,\mathbb{E}\,X^2I_{A_j}\\
&=\sum_{j=1}^{\infty}\sum_{n=j}^{\infty}\frac{\mathrm{Log}^{2(2-p)/p}n}{n^{2/p}}\,\mathbb{E}\,X^2I_{A_j}\\
&\leqslant C\sum_{j=1}^{\infty}\frac{\mathrm{Log}^{2(2-p)/p}j}{j^{2/p-1}}\,\mathbb{E}\,X^2I_{A_j}\\
&\leqslant C\sum_{j=1}^{\infty}\mathbb{E}\,|X|^p\,I_{A_j}\\
&=C\,\mathbb{E}\,|X|^p<\infty
\end{aligned}
$$

and

$$\sum_{n=1}^{\infty}\frac{1}{\mathrm{Log}^2 n}\mathbb{P}\left\{|X|>\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right\}=\sum_{n=1}^{\infty}\frac{1}{\mathrm{Log}^2 n}\,\mathbb{P}\left\{|X|^p>\frac{n}{\mathrm{Log}^2 n}\right\}\leqslant C\,\mathbb{E}\,|X|^p<\infty.$$

In order to prove

$$\sum_{n=1}^{\infty}\sum_{1\leqslant i<j\leqslant n}\frac{a_i a_j\,\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}G_{X_i,X_j}\left(\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right)<\infty \tag{2.11}$$

we have

$$\sum_{n=1}^{\infty}\sum_{1\leqslant i<j\leqslant n}\frac{a_i a_j\,\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}G_{X_i,X_j}\left(\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\right)=$$
$$\sum_{1\leqslant i<j\leqslant\infty}a_i a_j\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_{n=1}^{\infty}\frac{\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\,I_{\left\{\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\geqslant|x|\right\}}\,I_{\left\{\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\geqslant|y|\right\}}I_{\{n\geqslant j\}}\Delta_{X_i,X_j}(x,y)\,\mathrm{d}x\mathrm{d}y. \tag{2.12}$$

Since $p^2t^p\mathrm{Log}^2\,t$ is an asymptotic inverse of $t^{1/p}/\mathrm{Log}^{2/p}t$ (see Bingham et al. 1987, page 28), we get

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\,I_{\left\{\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\geqslant|x|\right\}}\,I_{\left\{\frac{n^{1/p}}{\mathrm{Log}^{2/p}n}\geqslant|y|\right\}}I_{\{n\geqslant j\}}\leqslant\\ \leqslant\sum_{n=1}^{\infty}\frac{\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\,I_{\left\{C(p)n\geqslant p^2|x|^p\mathrm{Log}^2|x|\right\}}\,I_{\left\{C(p)n\geqslant p^2|y|^p\mathrm{Log}^2|y|\right\}}I_{\{n\geqslant j\}}\\ =\sum_{n=1}^{\infty}\frac{\mathrm{Log}^{2(2-p)/p}n}{n^{1+2/p}}\,I_{\left\{n\geqslant\frac{p^2|x|^p\mathrm{Log}^2|x|}{C(p)}\vee\frac{p^2|y|^p\mathrm{Log}^2|y|}{C(p)}\vee j\right\}}\\ \leqslant C(p)\cdot\frac{\mathrm{Log}^{2(2-p)/p}\left[\frac{p^2|x|^p\mathrm{Log}^2|x|}{C(p)}\vee\frac{p^2|y|^p\mathrm{Log}^2|y|}{C(p)}\vee j\right]}{\left[\frac{p^2|x|^p\mathrm{Log}^2|x|}{C(p)}\vee\frac{p^2|y|^p\mathrm{Log}^2|y|}{C(p)}\vee j\right]^{2/p}}.\end{aligned} \tag{2.13}$$

Putting

$$m=\sup_{j\geqslant1}\frac{\mathrm{Log}^{2(2-p)/p}j}{j^{2/p}},$$
$$u(t)=\frac{\mathrm{Log}^{2(2-p)/p}\left[\frac{p^2t^p\mathrm{Log}^2\,t}{C(p)}\right]}{t^2\mathrm{Log}^{4/p}t}$$

it follows

$$u(t)\sim\frac{p^{2(2-p)/p}}{t^2\mathrm{Log}^2\,t},\quad t\to\infty,$$
$$u(t)\sim\frac{1}{t^2\mathrm{Log}^2\,t},\quad t\to0^+$$

and there is a constant $M(p) > 1$ such that

$$\sup_{t>0} \frac{u(t)}{\frac{1}{t^2 \mathrm{Log}^2 t}} \leqslant M(p) < \infty.$$

Hence, for all $x \neq 0$ and $y \neq 0$,

$$\begin{aligned}
&\frac{\mathrm{Log}^{2(2-p)/p}\left[\frac{p^2|x|^p \mathrm{Log}^2|x|}{C(p)} \vee \frac{p^2|y|^p \mathrm{Log}^2|y|}{C(p)} \vee j\right]}{\left[\frac{p^2|x|^p \mathrm{Log}^2|x|}{C(p)} \vee \frac{p^2|y|^p \mathrm{Log}^2|y|}{C(p)} \vee j\right]^{2/p}} = \\
&= \int_0^m I_{\{t \leqslant u(|x|)\}} I_{\{t \leqslant u(|y|)\}} I_{\left\{t \leqslant \frac{\mathrm{Log}^{2(2-p)/p} j}{j^{2/p}}\right\}} \,\mathrm{d}t \\
&\leqslant \int_0^m I_{\left\{t \leqslant \frac{M(p)}{x^2 \mathrm{Log}^2|x|}\right\}} I_{\left\{t \leqslant \frac{M(p)}{y^2 \mathrm{Log}^2|y|}\right\}} I_{\left\{t \leqslant \frac{\mathrm{Log}^{2(2-p)/p} j}{j^{2/p}}\right\}} \,\mathrm{d}t \\
&= \int_0^m I_{\{|x| \leqslant v^{-1}(t)\}} I_{\{|y| \leqslant v^{-1}(t)\}} I_{\left\{t \leqslant \frac{\mathrm{Log}^{2(2-p)/p} j}{j^{2/p}}\right\}} \,\mathrm{d}t
\end{aligned} \tag{2.14}$$

where $v^{-1}(t)$ denotes the inverse of $v(t) = M(p)/(t^2 \mathrm{Log}^2 t)$, $t > 0$ and according to Fubini's theorem, we obtain

$$\begin{aligned}
&\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \frac{\mathrm{Log}^{2(2-p)/p} n}{n^{1+2/p}} I_{\left\{\frac{n^{1/p}}{\mathrm{Log}^{2/p} n} \geqslant |x|\right\}} I_{\left\{\frac{n^{1/p}}{\mathrm{Log}^{2/p} n} \geqslant |y|\right\}} I_{\{n \geqslant j\}} \Delta_{X_i, X_j}(x, y) \,\mathrm{d}x \mathrm{d}y \\
&\leqslant C(p) \int_0^m I_{\left\{t \leqslant \frac{\mathrm{Log}^{2(2-p)/p} j}{j^{2/p}}\right\}} G_{X_i, X_j}\left[v^{-1}(t)\right] \mathrm{d}t \\
&= C(p) \int_0^{\frac{\mathrm{Log}^{2(2-p)/p} j}{j^{2/p}}} G_{X_i, X_j}\left[v^{-1}(t)\right] \mathrm{d}t \\
&= -C(p) \int_{v^{-1}\left(\frac{\mathrm{Log}^{2(2-p)/p} j}{j^{2/p}}\right)}^{\infty} v'(s) G_{X_i, X_j}(s) \,\mathrm{d}s \\
&\leqslant C(p) \int_{v^{-1}\left(\frac{\mathrm{Log}^{2(2-p)/p} j}{j^{2/p}}\right)}^{\infty} \left[\frac{1}{s^3 \mathrm{Log}^2 s} + \frac{1}{(s \vee \mathrm{e}) s^2 \mathrm{Log}^3 s}\right] G_{X_i, X_j}(s) \,\mathrm{d}s \qquad (2.15) \\
&= C(p) \int_{\frac{p\sqrt{M(p)} j^{1/p}}{\mathrm{Log}^{2/p} j} + o\left(\frac{j^{1/p}}{\mathrm{Log}^{2/p} j}\right)}^{\infty} \left[\frac{1}{s^3 \mathrm{Log}^2 s} + \frac{1}{(s \vee \mathrm{e}) s^2 \mathrm{Log}^3 s}\right] G_{X_i, X_j}(s) \,\mathrm{d}s \\
&\leqslant C(p) \int_{\frac{p\sqrt{M(p)} j^{1/p}}{\mathrm{Log}^{2/p} j} + o\left(\frac{j^{1/p}}{\mathrm{Log}^{2/p} j}\right)}^{\infty} \frac{G_{X_i, X_j}(s)}{s^3 \mathrm{Log}^2 s} \,\mathrm{d}s \\
&\leqslant C(p) \int_{\frac{j^{1/p}}{\mathrm{Log}^{2/p} j}}^{\infty} \frac{G_{X_i, X_j}(s)}{s^3 \mathrm{Log}^2 s} \,\mathrm{d}s
\end{aligned}$$

for j large enough, because

$$v^{-1}(t) \sim \frac{\sqrt{M(p)}}{\sqrt{t}\left|\log(\sqrt{t} \wedge \mathrm{e})\right|}, \quad t \to 0^+$$

and $s \mapsto G_{X_i,X_j}(s)$ is a nonnegative and nondecreasing function. Thus, gathering (2.12), (2.13), (2.14) and (2.15) we obtain (2.11) by using (2.9). Hence,

$$\sum_{n=1}^{\infty} \frac{1}{n} \mathbb{P}\left\{ \max_{1\leqslant k\leqslant n} \left| \sum_{j=1}^{k} a_j (X'_j - \mathbb{E}\, X'_j) \right| > \varepsilon n^{1/p} \mathrm{Log}^{2(p-1)/p}\, n \right\} < \infty$$

and Theorem 2.1 of Yang et al. 2008 yields $\sum_{k=1}^{n} a_k (X'_k - \mathbb{E}\, X'_k)/(n^{1/p}\mathrm{Log}^{2(p-1)/p}\, n) \xrightarrow{\text{a.s.}} 0$. It remains to show

$$\frac{1}{n^{1/p}\mathrm{Log}^{2(p-1)/p}\, n} \sum_{k=1}^{n} a_k (X''_k - \mathbb{E}\, X''_k) \xrightarrow{\text{a.s.}} 0. \tag{2.16}$$

By virtue of the Kronecker's lemma, convergence (2.16) holds if

$$\sum_{k=1}^{\infty} \frac{1}{k^{1/p}\mathrm{Log}^{2(p-1)/p}\, k} |a_k (X''_k - \mathbb{E}\, X''_k)| < \infty \quad \text{a.s.} \tag{2.17}$$

Since, for all $k \geqslant 3$,

$$\begin{aligned}
&\frac{1}{k^{1/p}\mathrm{Log}^{2(p-1)/p}\, k} - \frac{1}{(k+1)^{1/p}\mathrm{Log}^{2(p-1)/p}(k+1)} \\
&\quad = \int_k^{k+1} \frac{\mathrm{Log}^{(2-2p)/p}x + 2(p-1)\mathrm{Log}^{(2-3p)/p}x}{px^{1+1/p}}\, \mathrm{d}x \\
&\quad \leqslant \frac{2p-1}{p} \int_k^{k+1} \frac{\mathrm{Log}^{(2-2p)/p}x}{x^{1+1/p}}\, \mathrm{d}x \\
&\quad \leqslant \frac{2p-1}{p} \cdot \frac{1}{k^{1+1/p}\mathrm{Log}^{2(p-1)/p}k}
\end{aligned}$$

it follows

$$\begin{aligned}
&\sum_{k=3}^{\infty} \frac{1}{k^{1/p}\mathrm{Log}^{2(p-1)/p}\, k} \mathbb{E}|a_k (X''_k - \mathbb{E}\, X''_k)| \\
&\quad \leqslant 2 \sum_{k=3}^{\infty} \frac{1}{k^{1/p}\mathrm{Log}^{2(p-1)/p}\, k} \mathbb{E}|a_k X''_k| \\
&\quad \leqslant 2 \sum_{k=3}^{\infty} \frac{1}{k^{1/p}\mathrm{Log}^{2(p-1)/p}\, k} \mathbb{E}|a_k X_k| I_{\{|X_k| > k^{1/p}/\mathrm{Log}^{2/p}k\}} \\
&\quad \leqslant C \sum_{k=3}^{\infty} \frac{|a_k|}{k^{1/p}\mathrm{Log}^{2(p-1)/p}\, k} \mathbb{E}|X| I_{\{|X| > k^{1/p}/\mathrm{Log}^{2/p}k\}} \\
&\quad = C \sum_{k=3}^{\infty} \sum_{j=k}^{\infty} \frac{|a_k|}{k^{1/p}\mathrm{Log}^{2(p-1)/p}\, k} \mathbb{E}|X| I_{A_{j+1}} \\
&\quad = C \sum_{j=3}^{\infty} \sum_{k=3}^{j} \frac{|a_k|}{k^{1/p}\mathrm{Log}^{2(p-1)/p}\, k} \mathbb{E}|X| I_{A_{j+1}} \\
&\quad \leqslant C \sum_{j=3}^{\infty} \left[\sum_{k=3}^{j} \left(\frac{1}{k^{1/p}\mathrm{Log}^{2(p-1)/p}\, k} - \frac{1}{(k+1)^{1/p}\mathrm{Log}^{2(p-1)/p}(k+1)} \right) \sum_{m=1}^{k} |a_m| \right.
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{1}{(j+1)^{1/p}\mathrm{Log}^{2(p-1)/p}(j+1)} \sum_{m=1}^{j} |a_m| \Bigg] \mathbb{E}|X| I_{A_{j+1}} \\
&\leqslant C(p) \sum_{j=3}^{\infty} \Bigg[\sum_{k=3}^{j} \frac{1}{k^{1+1/p}\mathrm{Log}^{2(p-1)/p} k} \sum_{m=1}^{k} |a_m| \\
&+ \frac{1}{(j+1)^{1/p}\mathrm{Log}^{2(p-1)/p}(j+1)} \sum_{m=1}^{j} |a_m| \Bigg] \mathbb{E}|X| I_{A_{j+1}} \\
&\leqslant C(p) \sum_{j=1}^{\infty} \Bigg[\sum_{k=1}^{j} \frac{1}{k^{1/p}\mathrm{Log}^{2(p-1)/p} k} + \frac{1}{(j+1)^{1/p-1}\mathrm{Log}^{2(p-1)/p}(j+1)} \Bigg] \mathbb{E}|X| I_{A_{j+1}} \\
&\leqslant C(p) \sum_{j=1}^{\infty} \Bigg[\frac{1}{j^{1/p-1}\mathrm{Log}^{2(p-1)/p} j} + \frac{1}{(j+1)^{1/p-1}\mathrm{Log}^{2(p-1)/p}(j+1)} \Bigg] \mathbb{E}|X| I_{A_{j+1}} \\
&\leqslant C(p) \sum_{j=1}^{\infty} \frac{1}{j^{1/p-1}\mathrm{Log}^{2(p-1)/p} j} \mathbb{E}|X| I_{A_{j+1}} \\
&\leqslant C(p) \sum_{j=1}^{\infty} \frac{1}{j^{1/p-1}\mathrm{Log}^{2(p-1)/p} j} \cdot \frac{j^{(1-p)/p}}{\mathrm{Log}^{2(1-p)/p} j} \mathbb{E}|X|^p I_{A_{j+1}} \\
&\leqslant C(p) \mathbb{E}|X|^p < \infty
\end{aligned}
$$

where A_{j+1} is defined in (2.10). Therefore, a.s. convergence (2.17) is assured and

$$
\frac{1}{n^{1/p}\mathrm{Log}^{2(p-1)/p} n} \sum_{k=1}^{n} a_k (X_k'' - \mathbb{E}\, X_k'') \xrightarrow{\text{a.s.}} 0.
$$

The proof is complete. □

Remark 2 For any $r \geqslant s > 0$, $\sum_{k=1}^{n} |a_k|^s / n \leqslant (\sum_{k=1}^{n} |a_k|^r / n)^{s/r}$ by Hölder's inequality which implies that assumption $\sup_{n\geqslant 1} \sum_{k=1}^{n} a_k^2 / n < \infty$ in both Theorems 1 and 2 can be replaced by the (stronger) condition $\sup_{n\geqslant 1} \sum_{k=1}^{n} |a_k|^q / n < \infty$, $q \geqslant 2$.

The example below provides us random variables satisfying (2.2) or (2.9) but not possessing finite second moments.

Example 1 Let $1 \leqslant p < 2$ and $\{X_n,\, n \geqslant 1\}$ be a sequence of random variables such that for every $k \neq j$, (X_k, X_j) has Farlie-Gumbel-Morgenstern bivariate distribution, i.e.

$$
F_{X_k,X_j}(x,y) = F_{X_k}(x)F_{X_j}(y) + \rho F_{X_k}(x)F_{X_j}(y)\big[1 - F_{X_k}(x)\big]\big[1 - F_{X_j}(y)\big], \quad 0 \leqslant \rho \leqslant 1.
$$

Thus, $\{X_n,\, n \geqslant 1\}$ is a pairwise PQD sequence (see, for instance, Lai and Xie 2000). Supposing that X_n has probability density function $f_{X_n}(t) = n(q-1)(nt)^{-q} I_{(1/n,\infty)}(t)$, $p+1 < q \leqslant 3$, it follows that $\{X_n,\, n \geqslant 1\}$ is stochastically dominated by X_1 and $\mathbb{E}\,|X_1|^p = (q-1)/(q-p-1)$. Further, $\mathbb{E}\, X_n^2 = \infty$ for all n and after standard computations we obtain $G_{X_k,X_j}(t) \leqslant \rho\, C(q)/(kj)$. Hence,

$$
\sum_{1\leqslant k<j\leqslant\infty} |a_k a_j| \int_j^{\infty} t^{-3} G_{X_k,X_j}(t)\, \mathrm{d}t \leqslant \rho\, C(q) \sum_{1\leqslant k<j\leqslant\infty} \left(\frac{1}{j^3} + \frac{1}{kj^2} \right) < \infty
$$

when $p = 1$, and

$$\sum_{1\leqslant k<j\leqslant\infty} |a_k a_j| \int_{\frac{j^{1/p}}{\mathrm{Log}^{2/p} j}}^{\infty} \frac{G_{X_k,X_j}(t)}{t^3 \mathrm{Log}^2 t}\,\mathrm{d}t \leqslant \rho\, C(p,q) \sum_{1\leqslant k<j\leqslant\infty} \mathrm{Log}^{(4-2p)/p} j \left(\frac{1}{j^{1+2/p}} + \frac{1}{kj^{2/p}}\right) < \infty$$

with $C(p,q)$ a positive constant depending only on p and q whenever $1 < p < 2$, by noting that $\sup_{n\geqslant 1} n^{-1}\sum_{k=1}^{n} a_k^2 < \infty$ entails

$$|a_k a_j| \leqslant \frac{a_k^2 + a_j^2}{2} \leqslant \frac{\sum_{\ell=1}^{k} a_\ell^2 + \sum_{\ell=1}^{j} a_\ell^2}{2} \leqslant C(k+j).$$

Notice that if the weights $\{a_n\}$ satisfy $a_n = 1$ for all n and $\{X_n,\, n \geqslant 1\}$ is a sequence of pairwise PQD random variables such that

$$\Delta_{X_k,X_j}(x,y) = \Delta_{X_1,X_j}(x,y) \tag{2.18}$$

for any $1 \leqslant k < j$ and every $x, y \in \mathbb{R}$, then condition (2.2) can be simplified to

$$\sum_{j=2}^{\infty} \int_{j}^{\infty} \frac{G_{X_1,X_j}(v)}{v^2}\,\mathrm{d}v < \infty \tag{2.19}$$

(see Remark (6) of Louhichi 2000) or even to the less restrictive assumption

$$\sum_{j=2}^{\infty} \frac{G_{X_1,X_j}(j)}{j} < \infty$$

because

$$\begin{aligned}
&\sum_{n=1}^{\infty} \frac{1}{n^3} \sum_{1\leqslant k<j\leqslant n} \int_{-k}^{k}\int_{-j}^{j} \Delta_{X_k,X_j}(x,y)\,\mathrm{d}x\mathrm{d}y = \\
&\quad = \sum_{n=1}^{\infty} \frac{1}{n^3} \sum_{1\leqslant k<j\leqslant n} \int_{-k}^{k}\int_{-j}^{j} \Delta_{X_1,X_j}(x,y)\,\mathrm{d}x\mathrm{d}y \\
&\quad \leqslant \sum_{n=2}^{\infty}\sum_{j=2}^{n} \frac{G_{X_1,X_j}(n)}{n^2} \\
&\quad = \sum_{j=2}^{\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \frac{I_{\{n\geqslant |x|\vee|y|\vee j\}}}{n^2} \Delta_{X_1,X_j}(x,y)\,\mathrm{d}x\mathrm{d}y \\
&\quad \leqslant C \sum_{j=2}^{\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\Delta_{X_1,X_j}(x,y)}{|x|\vee|y|\vee j}\,\mathrm{d}x\mathrm{d}y \\
&\quad = C \sum_{j=2}^{\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{1} I_{\{|x|\leqslant 1/u\}} I_{\{|y|\leqslant 1/u\}} I_{\{u\leqslant 1/j\}} \Delta_{X_1,X_j}(x,y)\,\mathrm{d}u\,\mathrm{d}x\mathrm{d}y \\
&\quad = C \sum_{j=2}^{\infty} \int_{0}^{1/j} G_{X_1,X_j}\left(\frac{1}{u}\right)\mathrm{d}u
\end{aligned}$$

$$= C \sum_{j=2}^{\infty} \int_{j}^{\infty} \frac{G_{X_1,X_j}(v)}{v^2} \, \mathrm{d}v$$

and

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{n^3} \sum_{1 \leqslant k < j \leqslant n} \int_{-k}^{k} \int_{-j}^{j} \Delta_{X_k,X_j}(x,y) \, \mathrm{d}x\mathrm{d}y &= \sum_{n=1}^{\infty} \frac{1}{n^3} \sum_{1 \leqslant k < j \leqslant n} \int_{-k}^{k} \int_{-j}^{j} \Delta_{X_1,X_j}(x,y) \, \mathrm{d}x\mathrm{d}y \\ &\leqslant \sum_{n=2}^{\infty} \sum_{j=2}^{n} \frac{(j-1) G_{X_1,X_j}(j)}{n^3} \\ &\leqslant C \sum_{j=2}^{\infty} \frac{G_{X_1,X_j}(j)}{j}\end{aligned}$$

(recall that $t \mapsto G_{X_k,X_j}(t)$ is nondecreasing). We emphasise that (2.18) holds if for all $n, m \geqslant 1$,

$$(X_n, X_{n+m}) \overset{\mathrm{d}}{=} (X_1, X_{1+m}). \tag{2.20}$$

In addition, one can demonstrate that for any sequence $\{X_n,\ n \geqslant 1\}$ of pairwise PQD random variables satisfying condition (2.20),

$$\sum_{j=1}^{\infty} j \int_{j+1}^{\infty} v^{-3} G_{1,j+1}(v) \, \mathrm{d}v < \infty \tag{2.21}$$

and

$$\sum_{1 \leqslant k < j \leqslant \infty} j^{-2} \int_{-k}^{k} \int_{-j}^{j} \Delta_{X_1,X_j}(x,y) \, \mathrm{d}x\mathrm{d}y < \infty$$

are both equivalent by employing the same proof of Chen and Sung 2019 for (positively) associated random sequences (see Appendix of Chen and Sung 2019 for details). Obviously, (2.19) implies (2.21).

Similarly, under the assumptions of Theorem 2, and the extra conditions $a_n = 1$ for all n, (2.18), condition (2.9) can be simplified to

$$\sum_{j=2}^{\infty} \int_{\frac{j^{1/p}}{\mathrm{Log}^{2/p} j}}^{\infty} \frac{G_{X_1,X_j}(t)}{t^3 \mathrm{Log}^2 t} \, \mathrm{d}t < \infty.$$

Moreover, in this scenario one can still relax (2.9) to

$$\sum_{j=2}^{\infty} \frac{\mathrm{Log}^{2(2-p)/p} j}{j^{2/p-1}} G_{X_1,X_j}\left(\frac{j^{1/p}}{\mathrm{Log}^{2/p} j} \right) < \infty.$$

Let us point out that the identical distribution of $\{X_n,\ n \geqslant 1\}$ is not a sufficient condition to obtain (2.18) as the next example shows.

Example 2 Considering the following joint probability function of (X_k, X_j), $k < j$,

$X_k \backslash X_j$	0	1	
0	$\frac{1}{4}+\frac{1}{2^{k+j}}$	$\frac{1}{4}-\frac{1}{2^{k+j}}$	$\frac{1}{2}$
1	$\frac{1}{4}-\frac{1}{2^{k+j}}$	$\frac{1}{4}+\frac{1}{2^{k+j}}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	

we have $\mathbb{P}\{X_n = 0\} = 1/2 = \mathbb{P}\{X_n = 1\}$ for each $n \geqslant 1$ and

$$\Delta_{X_k,X_j}(x,y) = \frac{1}{2^{k+j}} \neq \frac{1}{2^{1+j}} = \Delta_{X_1,X_j}(x,y), \qquad k > 1$$

for all $0 < x, y < 1$.

3 Applications

3.1 Linear errors-in-variables regression model

Consider the simple linear errors-in-variables regression model,

$$\begin{cases} \eta_n = \alpha + \beta x_n + \varepsilon_n \\ \xi_n = x_n + \delta_n \end{cases} \qquad (n \geqslant 1) \tag{3.1}$$

where α, β are unknown parameters, $x_1, x_2, \ldots$ are (non-random) constants and $\{\varepsilon_n,\, n \geqslant 1\}$, $\{\delta_n,\, n \geqslant 1\}$ are two sequences of random variables. Recall that the model (3.1) not only furnishes an approximation to real world situations but also it helps us understand the theoretical underpinnings of methods for other models (see Fuller 1987). Rewriting (3.1) as an ordinary regression model having stochastic regressors and errors $\varepsilon_k - \beta\delta_k$, i.e.

$$\eta_n = \alpha + \beta\xi_n + (\varepsilon_n - \beta\delta_n) \qquad (n \geqslant 1),$$

formally, we can obtain the least-squares estimators of β and α as

$$\widehat{\beta}_n := \frac{\sum_{k=1}^n \left(\xi_k - n^{-1}\sum_{j=1}^n \xi_j\right)\left(\eta_k - n^{-1}\sum_{j=1}^n \eta_j\right)}{\sum_{k=1}^n \left(\xi_k - n^{-1}\sum_{k=1}^n \xi_k\right)^2} \tag{3.2}$$

and

$$\widehat{\alpha}_n := \frac{1}{n}\sum_{k=1}^n \eta_k - \frac{\widehat{\beta}_n}{n}\sum_{k=1}^n \xi_k, \tag{3.3}$$

respectively (see Liu and Chen 2005).

In Liu and Chen 2005, necessary and sufficient conditions were given to ensure the strong consistency of $\widehat{\beta}_n$ and $\widehat{\alpha}_n$ assuming that $\{(\varepsilon_n, \delta_n),\, n \geqslant 1\}$ is a sequence of independent random vectors, $\{\varepsilon_n,\, n \geqslant 1\}$ is a sequence of i.i.d. random variables and $\{\delta_n,\, n \geqslant 1\}$ is a sequence of i.i.d. random variables satisfying $\mathbb{E}\,\varepsilon_1 = \mathbb{E}\,\delta_1 = 0$, $0 < \mathbb{E}\,\delta_1^2 < \infty$, $0 < \mathbb{E}\,\varepsilon_1^2 < \infty$. Later, admitting that $\{(\varepsilon_n, \delta_n),\, n \geqslant 1\}$ is a sequence of stationary α-mixing random vectors, sufficient conditions were given in Fan et al. 2010 to get the strong consistency of $\widehat{\alpha}_n$ and

$\widehat{\beta}_n$. More recently, necessary and sufficient conditions for the strong consistency of these estimators were obtained in Hu et al. 2017 when $\{(\varepsilon_n, \delta_n),\, n \geqslant 1\}$ is a sequence of identically distributed ψ-mixing random vectors.

In order to broaden further the dependence structure of the random components in the model (3.1), we shall establish sufficient conditions for the strong consistency of both estimators, $\widehat{\alpha}_n$ and $\widehat{\beta}_n$, under sequences $\{\varepsilon_n,\, n \geqslant 1\}$ and $\{\delta_n,\, n \geqslant 1\}$ of pairwise PQD random variables.

Here, $\overline{x}_n := \sum_{k=1}^n x_k/n$ and other similar notations, such as $\overline{\delta}_n$ or $\overline{\xi}_n$ are defined in the same way.

Theorem 3 *Suppose that in model* (3.1), $\{\varepsilon_n,\, n \geqslant 1\}$ *is a sequence of pairwise PQD random variables stochastically dominated by a random variable* $\varepsilon \in \mathscr{L}_2$,

$$\sum_{1\leqslant k<j\leqslant\infty} \int_j^\infty t^{-2} \left[G_{\varepsilon_k^+,\varepsilon_j^+}(\sqrt{t}) + G_{\varepsilon_k^-,\varepsilon_j^-}(\sqrt{t})\right] \mathrm{d}t < \infty \tag{3.4}$$

and $\{\delta_n,\, n \geqslant 1\}$ *is a sequence of pairwise PQD random variables stochastically dominated by a random variable* $\delta \in \mathscr{L}_2$,

$$\sum_{1\leqslant k<j\leqslant\infty} \int_j^\infty t^{-2} \left[G_{\delta_k^+,\delta_j^+}(\sqrt{t}) + G_{\delta_k^-,\delta_j^-}(\sqrt{t})\right] \mathrm{d}t < \infty.$$

If $n/\sum_{k=1}^n (x_k - \overline{x}_n)^2 = o(1)$ *as* $n \to \infty$, *then* $\widehat{\beta}_n \xrightarrow{\text{a.s.}} \beta$. *Additionally, if* $n\,|\overline{x}_n|(|\overline{x}_n| \vee 1)/\sum_{k=1}^n (x_k - \overline{x}_n)^2 = o(1)$ *as* $n \to \infty$ *then* $\widehat{\alpha}_n \xrightarrow{\text{a.s.}} \alpha$.

Proof. Supposing $\varepsilon_n^+ := \varepsilon_n \vee 0$ and $\varepsilon_n^- := (-\varepsilon_n) \vee 0$, it is straightforward to see that $\left\{(\varepsilon_n^+)^2,\, n \geqslant 1\right\}$ is a sequence of pairwise PQD random variables stochastically dominated by ε^2. Since

$$\begin{aligned}
G_{(\varepsilon_k^+)^2,(\varepsilon_j^+)^2}(t) &= \int_{-t}^t \int_{-t}^t \left[\mathbb{P}\left\{(\varepsilon_k^+)^2 \leqslant x, (\varepsilon_j^+)^2 \leqslant y\right\} - \mathbb{P}\left\{(\varepsilon_k^+)^2 \leqslant x\right\}\mathbb{P}\left\{(\varepsilon_j^+)^2 \leqslant y\right\}\right] \mathrm{d}x\mathrm{d}y \\
&= \int_0^t \int_0^t \left[\mathbb{P}\left\{\varepsilon_k^+ \leqslant \sqrt{x}, \varepsilon_j^+ \leqslant \sqrt{y}\right\} - \mathbb{P}\left\{\varepsilon_k^+ \leqslant \sqrt{x}\right\}\mathbb{P}\left\{\varepsilon_j^+ \leqslant \sqrt{y}\right\}\right] \mathrm{d}x\mathrm{d}y \\
&= 4\int_0^{\sqrt{t}} \int_0^{\sqrt{t}} uv \left[\mathbb{P}\left\{\varepsilon_k^+ \leqslant u, \varepsilon_j^+ \leqslant v\right\} - \mathbb{P}\left\{\varepsilon_k^+ \leqslant u\right\}\mathbb{P}\left\{\varepsilon_j^+ \leqslant v\right\}\right] \mathrm{d}u\mathrm{d}v \\
&\leqslant 4t\int_0^{\sqrt{t}} \int_0^{\sqrt{t}} \left[\mathbb{P}\left\{\varepsilon_k^+ \leqslant u, \varepsilon_j^+ \leqslant v\right\} - \mathbb{P}\left\{\varepsilon_k^+ \leqslant u\right\}\mathbb{P}\left\{\varepsilon_j^+ \leqslant v\right\}\right] \mathrm{d}u\mathrm{d}v \\
&= 4tG_{\varepsilon_k^+,\varepsilon_j^+}(\sqrt{t})
\end{aligned}$$

we obtain $\sum_{k=1}^n \left[(\varepsilon_k^+)^2 - \mathbb{E}(\varepsilon_k^+)^2\right]/n \xrightarrow{\text{a.s.}} 0$ via Theorem 1. Thus,

$$\limsup_{n\to\infty} \left|\frac{1}{n}\sum_{k=1}^n (\varepsilon_k^+)^2\right| \leqslant \limsup_{n\to\infty} \left\{\left|\frac{1}{n}\sum_{k=1}^n \left[(\varepsilon_k^+)^2 - \mathbb{E}(\varepsilon_k^+)^2\right]\right| + \frac{1}{n}\sum_{k=1}^n \mathbb{E}(\varepsilon_k^+)^2\right\} \leqslant \mathbb{E}\,\varepsilon^2 \quad \text{a.s.}$$

By analogous reasoning we can conclude $\limsup_{n\to\infty} \left|\sum_{k=1}^n (\varepsilon_k^-)^2/n\right| \leqslant \mathbb{E}\,\varepsilon^2$ a.s. and so

$$\limsup_{n\to\infty} \left|\frac{1}{n}\sum_{k=1}^n \varepsilon_k^2\right| = \limsup_{n\to\infty} \left|\frac{1}{n}\sum_{k=1}^n \left[(\varepsilon_k^+)^2 + (\varepsilon_k^-)^2\right]\right| \leqslant 2\,\mathbb{E}\,\varepsilon^2 \quad \text{a.s.} \tag{3.5}$$

Similarly,

$$\limsup_{n\to\infty}\left|\frac{1}{n}\sum_{k=1}^{n}\delta_k^2\right| \leqslant 2\,\mathbb{E}\,\delta^2 \quad \text{a.s.} \tag{3.6}$$

Setting $s_n := \sum_{k=1}^{n}(x_k-\overline{x}_n)^2$, it follows

$$\left|\frac{\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)\varepsilon_k}{s_n}\right| \leqslant \frac{n}{s_n}\cdot\frac{1}{n}\sum_{k=1}^{n}|\varepsilon_k\delta_k| + \frac{n}{s_n}|\overline{\delta}_n\,\overline{\varepsilon}_n| \leqslant \frac{2n}{s_n}\left(\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2\right)^{1/2}\left(\frac{1}{n}\sum_{k=1}^{n}\delta_k^2\right)^{1/2} \xrightarrow{\text{a.s.}} 0 \tag{3.7}$$

and

$$\frac{\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)^2}{s_n} = \frac{n}{s_n}\cdot\frac{1}{n}\sum_{k=1}^{n}\delta_k^2 - \frac{n}{s_n}\cdot\overline{\delta}_n^2 \leqslant \frac{n}{s_n}\cdot\frac{1}{n}\sum_{k=1}^{n}\delta_k^2 \xrightarrow{\text{a.s.}} 0 \tag{3.8}$$

from (3.5) and (3.6). Moreover,

$$\left|\frac{\sum_{k=1}^{n}(x_k-\overline{x}_n)\varepsilon_k}{s_n}\right| \leqslant \left(\frac{n}{s_n}\right)^{1/2}\left(\frac{\sum_{k=1}^{n}\varepsilon_k^2}{n}\right)^{1/2} \xrightarrow{\text{a.s.}} 0$$

and

$$\left|\frac{\sum_{k=1}^{n}(x_k-\overline{x}_n)\delta_k}{s_n}\right| \leqslant \left(\frac{n}{s_n}\right)^{1/2}\left(\frac{\sum_{k=1}^{n}\delta_k^2}{n}\right)^{1/2} \xrightarrow{\text{a.s.}} 0,$$

yielding

$$\frac{\sum_{k=1}^{n}(x_k-\overline{x}_n)(\varepsilon_k-\beta\delta_k)}{s_n} \xrightarrow{\text{a.s.}} 0. \tag{3.9}$$

Thus, (3.8) entails

$$\left|\frac{\sum_{k=1}^{n}(x_k-\overline{x}_n)(\delta_k-\overline{\delta}_n)}{s_n}\right| \leqslant \left[\frac{\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)^2}{s_n}\right]^{1/2} \xrightarrow{\text{a.s.}} 0$$

and also

$$\frac{\sum_{k=1}^{n}(\xi_k-\overline{\xi}_n)^2}{s_n} = 1 + \frac{2\sum_{k=1}^{n}(x_k-\overline{x}_n)(\delta_k-\overline{\delta}_n)}{s_n} + \frac{\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)^2}{s_n} \xrightarrow{\text{a.s.}} 1. \tag{3.10}$$

Since

$$\begin{aligned}
&\widehat{\beta}_n - \beta = \\
&= \frac{\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)\varepsilon_k + \sum_{k=1}^{n}(x_k-\overline{x}_n)(\varepsilon_k-\beta\overline{\delta}_k) - \beta\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)^2}{\sum_{k=1}^{n}(\xi_k-\overline{\xi}_n)^2} \\
&= \frac{s_n}{\sum_{k=1}^{n}(\xi_k-\overline{\xi}_n)^2}\left[\frac{\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)\varepsilon_k}{s_n} + \frac{\sum_{k=1}^{n}(x_k-\overline{x}_n)(\varepsilon_k-\beta\overline{\delta}_k)}{s_n} - \beta\frac{\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)^2}{s_n}\right]
\end{aligned}$$

we obtain $\widehat{\beta}_n \xrightarrow{\text{a.s.}} \beta$ from (3.7), (3.8), (3.9) and (3.10). On the other hand,

$$\widehat{\alpha}_n - \alpha = (\beta-\widehat{\beta}_n)\overline{x}_n + (\beta-\widehat{\beta}_n)\overline{\delta}_n - \beta\overline{\delta}_n + \overline{\varepsilon}_n.$$

According to Theorem 1, $\overline{\varepsilon}_n \xrightarrow{\text{a.s.}} 0$ and $\overline{\delta}_n \xrightarrow{\text{a.s.}} 0$. Hence, it suffices to prove

$$(\beta-\widehat{\beta}_n)\overline{x}_n \xrightarrow{\text{a.s.}} 0. \tag{3.11}$$

We have

$$\left|\frac{\overline{x}_n}{s_n}\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)\varepsilon_k\right| \leqslant \frac{n|\overline{x}_n|}{s_n}\cdot\frac{1}{n}\sum_{k=1}^{n}|\varepsilon_k\delta_k| + \frac{n|\overline{x}_n|}{s_n}\cdot|\overline{\delta}_n\,\overline{\varepsilon}_n| \leqslant \\ \leqslant \frac{2n|\overline{x}_n|}{s_n}\cdot\left(\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2\right)^{1/2}\left(\frac{1}{n}\sum_{k=1}^{n}\delta_k^2\right)^{1/2} \xrightarrow{\text{a.s.}} 0 \tag{3.12}$$

and

$$\left|\frac{\overline{x}_n}{s_n}\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)^2\right| = \frac{n|\overline{x}_n|}{s_n}\cdot\frac{1}{n}\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)^2 = \\ = \frac{n|\overline{x}_n|}{s_n}\left(\frac{1}{n}\sum_{k=1}^{n}\delta_k^2-\overline{\delta}_n^2\right) \leqslant \frac{n|\overline{x}_n|}{s_n}\cdot\frac{1}{n}\sum_{k=1}^{n}\delta_k^2 \xrightarrow{\text{a.s.}} 0. \tag{3.13}$$

Moreover,

$$\left|\frac{\overline{x}_n\sum_{k=1}^{n}(x_k-\overline{x}_n)\varepsilon_k}{s_n}\right| \leqslant \frac{|\overline{x}_n|}{s_n}\cdot\left[\sum_{k=1}^{n}(x_k-\overline{x}_n)^2\right]^{1/2}\left(\sum_{k=1}^{n}\varepsilon_k^2\right)^{1/2} = \\ = \left(\frac{n\,\overline{x}_n^2}{s_n}\right)^{1/2}\cdot\left(\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2\right)^{1/2} \xrightarrow{\text{a.s.}} 0$$

and

$$\left|\frac{\overline{x}_n\sum_{k=1}^{n}(x_k-\overline{x}_n)\delta_k}{s_n}\right| \leqslant \frac{|\overline{x}_n|}{s_n}\cdot\left[\sum_{k=1}^{n}(x_k-\overline{x}_n)^2\right]^{1/2}\left(\sum_{k=1}^{n}\delta_k^2\right)^{1/2} = \\ = \left(\frac{n\,\overline{x}_n^2}{s_n}\right)^{1/2}\cdot\left(\frac{1}{n}\sum_{k=1}^{n}\delta_k^2\right)^{1/2} \xrightarrow{\text{a.s.}} 0$$

imply

$$\frac{\overline{x}_n\sum_{k=1}^{n}(x_k-\overline{x}_n)(\varepsilon_k-\beta\delta_k)}{s_n} \xrightarrow{\text{a.s.}} 0. \tag{3.14}$$

Thus,

$$\overline{x}_n(\beta-\widehat{\beta}_n) = -\frac{s_n}{\sum_{k=1}^{n}(\xi_k-\overline{\xi}_n)^2}\cdot \\ \left[\frac{\overline{x}_n\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)\varepsilon_k}{s_n} + \frac{\overline{x}_n\sum_{k=1}^{n}(x_k-\overline{x}_n)(\varepsilon_k-\beta\overline{\delta}_k)}{s_n} - \beta\frac{\overline{x}_n\sum_{k=1}^{n}(\delta_k-\overline{\delta}_n)^2}{s_n}\right]$$

and (3.11) holds from (3.10), (3.12), (3.13) and (3.14). The proof is complete. □

Remark 3 Let us note that if $\overline{x}_n$ is bounded then condition $n|\overline{x}_n|(|\overline{x}_n|\vee 1)/\sum_{k=1}^{n}(x_k-\overline{x}_n)^2 = o(1)$, $n\to\infty$ can be dropped (that is, $n/\sum_{k=1}^{n}(x_k-\overline{x}_n)^2 = o(1)$, $n\to\infty$ is sufficient to obtain strong consistency of both estimators $\widehat{\alpha}_n$ and $\widehat{\beta}_n$).

3.2 Multiple regression model

Consider the multiple regression model

$$\mathbf{y}_n = \mathbf{X}_n \boldsymbol{\beta} + \boldsymbol{\varepsilon}_n \tag{3.15}$$

where $\mathbf{X}_n = \left(x_{ij}\right)_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$ is a known $n \times p$ matrix of rank p, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$ is the p-dimensional parameter vector, $\boldsymbol{\varepsilon}_n = (\varepsilon_1, \ldots, \varepsilon_n)'$ the n-dimensional error vector and $\mathbf{y}_n = (y_1, \ldots, y_n)'$ the n-dimensional observation vector with prime denoting transpose. For $n \geqslant p$,

$$\widehat{\boldsymbol{\beta}}_n = \boldsymbol{\beta} + (\mathbf{X}_n' \mathbf{X}_n)^{-1} \mathbf{X}_n' \boldsymbol{\varepsilon}_n$$

is the least-squares estimate of $\boldsymbol{\beta}$.

3.2.1 Non-stochastic regressors

The strong consistency of the least squares estimates in multiple regression models having non-stochastic regressors was studied in the past by many authors (see, Drygas 1976, Gui-Jing et al. 1981 or Lai et al. 1979, among others). In the following, the strong consistency for least-squares estimators of unknown parameter vector is given. It extends Theorem 1 of Lita da Silva and Mexia 2013 to sequences $\{\varepsilon_n,\, n \geqslant 1\}$ of pairwise PQD random variables.

Theorem 4 *Suppose that in model* (3.15), $\{\varepsilon_n,\, n \geqslant 1\}$ *is a sequence of identically distributed pairwise PQD random variables such that* $\varepsilon_1 \in \mathscr{L}_r$ *for some* $1 \leqslant r < 2$ *and* $\mathbb{E}\,\varepsilon_1 = 0$. *If* $\mathbf{X}_n' \mathbf{X}_n$ *is non-singular for some* $n \geqslant n_0$*, the design levels* $\{x_{ij},\, 1 \leqslant j \leqslant p, i \geqslant 1\}$ *satisfy* $\sup_{n \geqslant 1} \sum_{k=1}^{n} x_{kj}^2 / n < \infty$ *for every* j*,*

(i) $\left[(\mathbf{X}_n' \mathbf{X}_n)^{-1}\right]_{jj} = O\left(n^{-1}\right)$ *as* $n \to \infty$ *for all* j *and*

$$\sum_{1 \leqslant k < \ell \leqslant \infty} |x_{ki} x_{\ell j}| \int_{\ell}^{\infty} t^{-3} G_{\varepsilon_k, \varepsilon_\ell}(t)\, \mathrm{d}t < \infty \qquad (i, j = 1, \ldots, p)$$

when $r = 1$,

or

(ii) $\left[(\mathbf{X}_n' \mathbf{X}_n)^{-1}\right]_{jj} = O\left(n^{-1/r} \operatorname{Log}^{-2(r-1)/r} n\right)$ *as* $n \to \infty$ *for all* j *and*

$$\sum_{1 \leqslant k < \ell \leqslant \infty} |x_{ki} x_{\ell j}| \int_{\frac{\ell^{1/r}}{\operatorname{Log}^{2/r} \ell}}^{\infty} \frac{G_{\varepsilon_k, \varepsilon_\ell}(t)}{t^3 \operatorname{Log}^2 t}\, \mathrm{d}t < \infty \qquad (i, j = 1, \ldots, p)$$

whenever $1 < r < 2$,

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

Proof. From the expression of $\widehat{\boldsymbol{\beta}}_n$, it follows that the strong consistency of the least-squares estimate is equivalent to

$$(\mathbf{X}_n' \mathbf{X}_n)^{-1} \sum_{k=1}^{n} \mathbf{x}_k \varepsilon_k \xrightarrow{\text{a.s.}} \mathbf{0}$$

where $\mathbf{x}_k = (x_{k1}, \dots, x_{kp})'$. Since $(\mathbf{X}_n'\mathbf{X}_n)^{-1}$, $n \geqslant n_0$ is symmetric positive-definite, we have

$$\left|\left[(\mathbf{X}_n'\mathbf{X}_n)^{-1}\right]_{ij}\right| \leqslant \left[(\mathbf{X}_n'\mathbf{X}_n)^{-1}\right]_{ii}^{1/2} \left[(\mathbf{X}_n'\mathbf{X}_n)^{-1}\right]_{jj}^{1/2}$$

(see Harville 1997, page 280). Hence,

$$\left|\left[(\mathbf{X}_n'\mathbf{X}_n)^{-1}\right]_{ij} \sum_{k=1}^{n} x_{kj}\varepsilon_k\right| \leqslant$$

$$\left[(\mathbf{X}_n'\mathbf{X}_n)^{-1}\right]_{ii}^{1/2} \left[(\mathbf{X}_n'\mathbf{X}_n)^{-1}\right]_{jj}^{1/2} \left|\sum_{k=1}^{n} x_{kj}\varepsilon_k\right| \leqslant C \left|\frac{1}{n}\sum_{k=1}^{n} x_{kj}\varepsilon_k\right| \xrightarrow{\text{a.s.}} 0$$

when $r = 1$ by Theorem 1; from Theorem 2, we get

$$\left|\left[(\mathbf{X}_n'\mathbf{X}_n)^{-1}\right]_{ij} \sum_{k=1}^{n} x_{kj}\varepsilon_k\right| \leqslant$$

$$\left[(\mathbf{X}_n'\mathbf{X}_n)^{-1}\right]_{ii}^{1/2} \left[(\mathbf{X}_n'\mathbf{X}_n)^{-1}\right]_{jj}^{1/2} \left|\sum_{k=1}^{n} x_{kj}\varepsilon_k\right| \leqslant C \left|\frac{1}{n^{1/r}\operatorname{Log}^{2(r-1)/r} n}\sum_{k=1}^{n} x_{kj}\varepsilon_k\right| \xrightarrow{\text{a.s.}} 0$$

whenever $1 < r < 2$ establishing the thesis. □

3.2.2 Stochastic regressors

In model (3.15), let us assume that the design levels $\{x_{ij}, 1 \leqslant j \leqslant p, i \geqslant 1\}$ are random variables. If the errors $\varepsilon_1, \varepsilon_2, \dots$ are pairwise PQD and identically distributed random variables then we can use Theorem 1 to prove the strong consistency of $\widehat{\boldsymbol{\beta}}_n$.

In what follows, we shall define $\rho(\mathbf{A}) = \sup\{|\lambda| : \lambda \in \operatorname{Spec}(\mathbf{A})\}$ where $\operatorname{Spec}(\mathbf{A})$ is the spectrum of the matrix $\mathbf{A} = \left(a_{ij}\right)_{1\leqslant i,j\leqslant p}$. The column space of the matrix $\mathbf{M} = \left(m_{ij}\right)_{1\leqslant i\leqslant n, 1\leqslant j\leqslant p}$ will be indicated by $\operatorname{Col}(\mathbf{M})$. Given a n-dimensional vector $\mathbf{a}$ we shall use $\|\mathbf{a}\|$ to denote the *Euclidean* vector norm, that is, $\|\mathbf{a}\| = \sqrt{\mathbf{a}'\mathbf{a}}$.

Theorem 5 *Suppose that in model* (3.15), $\{\varepsilon_n,\ n \geqslant 1\}$ *is a sequence of pairwise PQD random variables stochastically dominated by a random variable* $\varepsilon \in \mathscr{L}_2$ *satisfying* (3.4). *If* $\{x_{ij}\}$ $(i = 1, 2, \dots; j = 1, \dots, p)$ *is an arbitrary double array of random variables such that* $\mathbf{X}_n'\mathbf{X}_n$ *is non-singular a.s. for some* $n \geqslant p$ *and* $n\rho\left((\mathbf{X}_n'\mathbf{X}_n)^{-1}\right) \xrightarrow{\text{a.s.}} 0$, *then* $\widehat{\boldsymbol{\beta}}_n \xrightarrow{\text{a.s.}} \boldsymbol{\beta}$.

Proof. From Proposition 1 of Lita da Silva 2014, we have

$$\|\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}\|^2 \leqslant \rho\left((\mathbf{X}_n'\mathbf{X}_n)^{-1}\right) \|\mathbf{P}_{\operatorname{Col}(\mathbf{X}_n)}\boldsymbol{\varepsilon}_n\|^2 \quad \text{a.s.} \tag{3.16}$$

where $\mathbf{P}_{\operatorname{Col}(\mathbf{X}_n)}\boldsymbol{\varepsilon}_n$ is the orthogonal projection of $\boldsymbol{\varepsilon}_n$ on $\operatorname{Col}(\mathbf{X}_n)$. Using Gram-Schmidt process we can construct an orthonormal basis $\{\mathbf{w}_{n,1}, \dots, \mathbf{w}_{n,p}\}$ of $\operatorname{Col}(\mathbf{X}_n)$ such that

$$\|\mathbf{P}_{\operatorname{Col}(\mathbf{X}_n)}\boldsymbol{\varepsilon}_n\|^2 = \langle\mathbf{w}_{n,1}, \boldsymbol{\varepsilon}_n\rangle^2 + \dots + \langle\mathbf{w}_{n,p}, \boldsymbol{\varepsilon}_n\rangle^2 \tag{3.17}$$

where $\langle \cdot\,, \cdot \rangle$ denotes the usual inner product in $\mathbb{R}^n$. From Cauchy-Schwarz inequality we have, for each $j = 1, \ldots, p$,

$$\rho\left((\mathbf{X}_n'\mathbf{X}_n)^{-1}\right)\langle \mathbf{w}_{n,j}, \boldsymbol{\varepsilon}_n\rangle^2 \leqslant \rho\left((\mathbf{X}_n'\mathbf{X}_n)^{-1}\right)\|\mathbf{w}_{n,j}\|^2\|\boldsymbol{\varepsilon}_n\|^2 = n\rho\left((\mathbf{X}_n'\mathbf{X}_n)^{-1}\right)\cdot\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2 \xrightarrow{\text{a.s.}} 0 \tag{3.18}$$

via (3.5). By (3.16), (3.17) and (3.18), it follows

$$\|\mathbf{b}_n - \boldsymbol{\beta}\|^2 \leqslant \rho\left((\mathbf{X}_n'\mathbf{X}_n)^{-1}\right)\left[\langle \mathbf{w}_{n,1}, \boldsymbol{\varepsilon}_n\rangle^2 + \ldots + \langle \mathbf{w}_{n,p}, \boldsymbol{\varepsilon}_n\rangle^2\right] \xrightarrow{\text{a.s.}} 0.$$

The proof is complete. □

3.3 Simple ridge regression model

In model (3.15), suppose $p = 1$ and the ridge estimator

$$\widehat{\gamma}_n = \left(\sum_{j=1}^{n} x_j^2 + \kappa\right)^{-1}\mathbf{x}_n'\mathbf{y}_n$$

where $\kappa = \widehat{\sigma}_n^2/\widehat{\beta}_n^2$, $\mathbf{x}_n = (x_1, \ldots, x_n)'$ and $\widehat{\sigma}_n^2 = (\mathbf{y}_n - \mathbf{x}_n\widehat{\beta}_n)'(\mathbf{y}_n - \mathbf{x}_n\widehat{\beta}_n)/(n-1)$ (see Groß 2003, page 9).

Theorem 6 *Suppose model* (3.15) *with* $p = 1$ *and* $\{\varepsilon_n,\, n \geqslant 1\}$ *a sequence of pairwise PQD random variables stochastically dominated by a random variable* $\varepsilon \in \mathscr{L}_2$ *satisfying* (3.4)*. If* $\{x_n,\, n \geqslant 1\}$ *is an arbitrary sequence of random variables such that* $\sum_{j=1}^n x_j^2 \neq 0$ *a.s. for some* $n \geqslant 1$ *and* $n/\sum_{j=1}^n x_j^2 \xrightarrow{\text{a.s.}} 0$*, then* $\widehat{\gamma}_n \xrightarrow{\text{a.s.}} \beta$.

Proof. We have

$$\left(\sum_{j=1}^{n} x_j^2 + \kappa\right)^{-1} = \left(1 - \frac{\kappa/\sum_{j=1}^n x_j^2}{1 + \kappa/\sum_{j=1}^n x_j^2}\right)\left(\sum_{j=1}^{n} x_j^2\right)^{-1}.$$

which yields

$$\widehat{\gamma}_n = \left(1 - \frac{\kappa/\sum_{j=1}^n x_j^2}{1 + \kappa/\sum_{j=1}^n x_j^2}\right)\beta + \left(1 - \frac{\kappa/\sum_{j=1}^n x_j^2}{1 + \kappa/\sum_{j=1}^n x_j^2}\right)\cdot\frac{\sum_{j=1}^n x_j\varepsilon_j}{\sum_{j=1}^n x_j^2}.$$

Since,

$$\widehat{\sigma}_n^2 = \frac{\left\|[\mathbf{I}_n - \mathbf{x}_n(\mathbf{x}_n'\mathbf{x}_n)^{-1}\mathbf{x}_n']\boldsymbol{\varepsilon}_n\right\|^2}{n-1} \leqslant \frac{\|\boldsymbol{\varepsilon}_n\|^2}{n-1}$$

it follows

$$\frac{\widehat{\sigma}_n^2}{\sum_{j=1}^n x_j^2} \leqslant \frac{\sum_{j=1}^n \varepsilon_j^2}{(n-1)\sum_{j=1}^n x_j^2} = \frac{n}{n-1}\cdot\frac{\sum_{j=1}^n \varepsilon_j^2}{n^2}\cdot\frac{n}{\sum_{j=1}^n x_j^2}$$

Recall that $\sum_{j=1}^{n} \varepsilon_j^2/n^2 \xrightarrow{\text{a.s.}} 0$ via Kronecker's lemma provided that $\sum_{n=1}^{\infty} \mathbb{E}\,\varepsilon_n^2/n^2 \leqslant C\,\mathbb{E}\,\varepsilon^2 < \infty$. From Theorem 5, we obtain strong consistency of $\widehat{\beta}_n$ i.e. $\widehat{\beta}_n \xrightarrow{\text{a.s.}} \beta$. Thus,

$$\frac{\kappa}{\sum_{j=1}^{n} x_j^2} = \frac{1}{\widehat{\beta}_n^2} \cdot \frac{\widehat{\sigma}_n^2}{\sum_{j=1}^{n} x_j^2} \xrightarrow{\text{a.s.}} 0 \tag{3.19}$$

provided that $\beta \neq 0$; obviously, for $\beta = 0$ one have

$$|\widehat{\gamma}_n| = \left(\sum_{j=1}^{n} x_j^2 + \kappa \right)^{-1} \left| \sum_{j=1}^{n} x_j \varepsilon_j \right| \leqslant \frac{\left| \sum_{j=1}^{n} x_j \varepsilon_j \right|}{\sum_{j=1}^{n} x_j^2}. \tag{3.20}$$

On the other hand,

$$\left(\frac{\sum_{j=1}^{n} x_j \varepsilon_j}{\sum_{j=1}^{n} x_j^2} \right)^2 \leqslant \frac{\sum_{j=1}^{n} \varepsilon_j^2}{\sum_{j=1}^{n} x_j^2} = \frac{\sum_{j=1}^{n} \varepsilon_j^2}{n} \cdot \frac{n}{\sum_{j=1}^{n} x_j^2} \tag{3.21}$$

and (3.5) ensures $\sum_{j=1}^{n} x_j \varepsilon_j / \sum_{j=1}^{n} x_j^2 \xrightarrow{\text{a.s.}} 0$ because $n/\sum_{j=1}^{n} x_j^2 \xrightarrow{\text{a.s.}} 0$. Hence, (3.19) (or (3.20)) and (3.21) lead to $\widehat{\gamma}_n \xrightarrow{\text{a.s.}} \beta$ and the strong consistency of the ridge estimator $\widehat{\gamma}_n$ is established. □

Remark 4 Similarly, under the assumptions of Theorem 6 we can conclude also the strong consistency of the shrinkage estimator $\widehat{\theta}_n = \widehat{\beta}_n/(1+\varrho)$, where $\varrho = \left(\sum_{j=1}^{n} x_j^2 \right)^{-1} \widehat{\sigma}_n^2/\widehat{\beta}_n^2$ (see Groß 2003, page 9).

Acknowledgements. The author is grateful to Editor-in-Chief Prof. Narayanaswamy Balakrishnan and Editors who handled the submitted manuscript. The author also wishes to express his gratitude to both Referees for the very careful reading of the manuscript, which resulted in elimination of several typos and improved readability of this final version.

This work is a contribution to the Project UIDB/04035/2020, funded by FCT - Fundação para a Ciência e a Tecnologia, Portugal.

References

Bingham, N. H., Goldie, C. M., and J. L. Teugels. 1987. *Regular Variation.* Cambridge: Cambridge University Press.

Birkel, T. 1989. A note on the strong law of large numbers for positively dependent random variables. *Statist. Probab. Lett.* 7:17–20. doi:10.1016/0167-7152(88)90080-6.

Birkel, T. 1989. A functional central limit theorem for positively dependent random variables. *J. Multivariate Anal.* 44:314–320. doi:10.1006/jmva.1993.1018.

Bulinski, A., and A. Shashkin. 2007 *Limit Theorems for Associated Random Fields and Related Systems.* Singapore: Advanced Series on Statistical Science & Applied Probability, Vol. 10, World Scientific. doi:10.1142/6555.

Chen, P., and S. H. Sung. 2019. Strong laws of large numbers for positively dependent random variables. *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Math. RACSAM* 113:3089–3100. doi:10.1007/s13398-019-00679-4.

Drygas, H. 1976. Weak and strong consistency of the least squares estimators in regression model. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* 34:119–127. doi:10.1007/BF00535679.

Esary, J., F. Proschan, and D. Walkup. 1967. Association of random variables with applications. *Ann. Math. Statist.* 38:1466–1474. doi:10.1214/aoms/1177698701.

Gui-Jing, C., Lai, T. L., and C. Z. Wei. 1981. Convergence systems and strong consistency of least squares estimates in regression models. *J. Multivariate Anal.* 11:319–333. doi:10.1016/0047-259X(81)90078-6.

Fan, G. L., Liang, H. Y., Wang, J. F., and H. X. Xu. 2010. Asymptotic properties for LS estimators in EV regression model with dependent errors, *AStA Adv. Stat. Anal.* 94:89–103. doi: 10.1007/s10182-010-0124-3.

Fuller, W. A. 1987. *Measurement Error Models.* New York: John Wiley & Sons. doi:10.1002/9780470316665.

Groß, J. 2003. *Linear Regression.* Heidelberg: Springer-Verlag Berlin. doi:10.1007/978-3-642-55864-1.

Harville, D. A. 1997. *Matrix Algebra From a Statiscian's Perspective.* New York: Springer-Verlag. doi:10.1007/b98818.

Hu, D., Chen, P., and S. H. Sung. 2017. Strong laws for weighted sums of ψ-mixing random variables and applications in errors-in-variables regression models. *TEST* 26:600–617. doi:10.1007/s11749-017-0526-6.

Hutchinson, T. P., and C. D. Lai. 1990. *Continuous Bivariate Distributions, Emphasising Applications.* Adelaide: Rumsby Scientific Publishing.

Kolmogorov, A. N. 1933. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Berlin: Springer.

Lai, T. L., Robbins, H., and C. Z. Wei. 1979. Strong consistency of least squares estimates in multiple regression II. *J. Multivariate Anal.* 9:343–362. doi:10.1016/0047-259X(79)90093-9.

Lai, C. D., and X. Xie. 2000. A new family of positive quadrant dependent bivariate distributions. *Statist. Probab. Lett.* 46:359–364. doi:10.1016/S0167-7152(99)00122-4.

Lehmann, E. 1966. Some concepts of dependence. *Ann. Math. Stat.* 37:1137–1153. doi:10.1214/aoms/1177699260.

Louhichi, S. 2000. Convergence rates in the strong law for associated random variables. *Probab. Math. Statist.* 20:203–214.

Lita da Silva, J., and J. T. Mexia. 2013. Strong consistency of least squares estimates with i.i.d. errors with mean values not necessarily defined. *Statistics* 47:707–714. doi:10.1080/02331888.2011.581759.

Lita da Silva, J. 2014. Some strong consistency results in stochastic regression. *J. Multivariate Anal.* 129:220–226. doi:10.1016/j.jmva.2014.04.022.

Lita da Silva, J. 2015. Almost sure convergence for weighted sums of extended negatively dependent random variables. *Acta Math. Hungar.* 146:56–70. doi:10.1007/s10474-015-0502-0.

Lita da Silva, J. 2018. Strong laws of large numbers for pairwise quadrant dependent random variables. *Statist. Probab. Lett.* 137:349–358. doi:10.1016/j.spl.2018.01.031.

Liu, J., and X. Chen. 2005. Consistency of LS estimator in simple linear EV regression models. *Acta Math. Sci. Ser. B Engl. Ed.* 25:50–58. doi:10.1016/S0252-9602(17)30260-6.

Matuła, P. 2005. On almost sure limit theorems for positively dependent random variables. *Statist. Probab. Lett.* 74:59–66. doi:10.1016/j.spl.2005.04.032.

Newman, C. M. 1984. Asymptotic independence and limit theorems for positively and negatively dependent random variables. In: Y. L. Tong (Ed.), Institute of Mathematical Statistics Lecture Notes – Monograph Series 5:127–140. doi: 10.1214/lnms/1215465639.

Tong, Y. L. 1980. *Probability Inequalities in Multivariate Distributions*. New York: Academic Press.

Walk, H. 2005. Strong laws of large numbers by elementary tauberian arguments. *Monatsh. Math.* 144:329–346. doi:10.1007/s00605-004-0284-x.

Yang, S., Su, C., and K. Yu. 2008. A general method to the strong law of large numbers and its applications. *Statist. Probab. Lett.* 78:794–803. doi:10.1016/j.spl.2007.09.046.