Inequalities in Statistics and Probability
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 165–172

MOMENT INEQUALITIES WITH APPLICATIONS TO REGRESSION AND TIME SERIES MODELS

BY TZE LEUNG LAI[1] and CHING ZONG WEI[2]
Columbia University and the University of Maryland

Herein we review several important moment inequalities in the literature and discuss their applications to strong (almost sure) limit theorems for linear processes and for least squares estimates in multiple regression models.

1. Introduction and Summary. A classical model for random noise in the regression and time series literature is that of *equinormed orthogonal* random variables ϵ_n, i.e.,

$$E(\epsilon_i \epsilon_j) = 0 \quad \text{for } i \neq j, \qquad = \sigma^2 \quad \text{for } i = j. \tag{1.1}$$

Such random variables have the important mean square property that for all constants c_i,

$$E(\Sigma_{i=m}^n c_i\epsilon_i)^2 = \sigma^2 \Sigma_{i=m}^n c_i^2 \qquad \text{for all } n \geq m. \tag{1.2}$$

For example, the so-called Gauss-Markov model in multiple regression theory is of the form

$$z_i = \beta_1 t_{i1} + \dots + \beta_k t_{ik} + \epsilon_i \, (i=1,2, \dots) \tag{1.3}$$

where t_{ij} are known constants, z_i are observed random variables, $\beta_i, \dots , \beta_k$ are unknown parameters, and ϵ_i are equinormed orthogonal random variables that represent unobservable random errors. Throughout the sequel we shall let $\mathbf{T}_n$ denote the design matrix $(t_{ij})_{1\leq i\leq n,\ 1\leq j\leq k}$, and let $\mathbf{Z}_n = (z_1, \dots , z_n)'$. For $n\geq k$, the least squares estimate $\mathbf{b}_n = (b_{n1}, \dots , b_{nk})'$ of $\beta = (\beta_1, \dots , \beta_k)'$ based on the design matrix $\mathbf{T}_n$ and the response vector $\mathbf{Z}_n$ is given by

$$\mathbf{b}_n = (\mathbf{T}'_n\mathbf{T}_n)^{-1}\mathbf{T}'_n\mathbf{Z}_n, \tag{1.4}$$

provided that $\mathbf{T}'_n\mathbf{T}_n$ is nonsingular. From (1.1), it follows easily that

$$\text{cov}(\mathbf{b}_n) = \sigma^2(\mathbf{T}'_n\mathbf{T}_n)^{-1}, \tag{1.5}$$

and therefore $\mathbf{b}_n$ is weakly consistent (i.e., $\mathbf{b}_n \xrightarrow{P} \mathbf{B}$) if

$$(\mathbf{T}'_n\mathbf{T}_n)^{-1} \to \mathbf{0} \text{ as } n\to\infty. \tag{1.6}$$

If $\sigma \neq 0$, the condition (1.6) is also necessary for the weak consistency of $\mathbf{b}_n$ (cf. Drygas (1976)).

In time series theory, it is well known that every wide-sense stationary sequence $\{y_n\}$ with zero means and an absolutely continuous spectral distribution can be represented as

$$y_n = \text{l.i.m.}_{N\to\infty}\Sigma_{i=-N}^N a_{n-i}\epsilon_i, \tag{1.7}$$

where $\{\epsilon_n\}$ is an orthonormal sequence (i.e., $\sigma=1$ in (1.1)), $\{a_n\}$ is a sequence of constants such that $\Sigma_{n=-\infty}^\infty a_n^2 < \infty$, and l.i.m. denotes limit in quadratic mean (cf. Doob (1953), page 499). From this representation, it follows that

---

[1,2] Research supported by the National Science Foundation.
AMS 1970 subject classifications: Primary 60F15; Secondary 60G35, 62J05, 62M10.
Key words and phrases. Least squares theory, linear processes, orthogonal random variables, moment inequalities, lacunary systems, almost sure convergence, law of the iterated logarithm.

$$E(\Sigma_{j=1}^n y_j)^2 = \Sigma_{i=-\infty}^{\infty}(\Sigma_{j=1}^n a_{j-i})^2 = o(n^2),$$

and therefore $\{y_n\}$ satisfies the weak law of large numbers:

(1.8) $$n^{-1}\Sigma_{j=1}^n y_j \xrightarrow{P} 0.$$

The representation (1.7) also provides an important stochastic model in the engineering literature, where the sequence $\{\epsilon_n\}$ is a white noise sequence and $\{y_n\}$ is the output sequence obtained by passing $\{\epsilon_n\}$ through a linear filter defined by $\{a_n\}$ (cf. Kailath (1974)). We shall call the sequence $\{y_n\}$ in (1.7) a *linear process generated by* $\{\epsilon_n\}$.

In order to strengthen the weak consistency result on $\mathbf{b}_n$ into its strong consistency under the minimal assumption (1.6) on the design constants, or to strengthen the weak law (1.8) into the corresponding strong law, we have found it necessary to introduce additional structure into the noise sequence $\{\epsilon_n\}$. Indeed, Chen, Lai and Wei (1981) gave a counter-example to show that the condition (1.6) is not sufficient for the strong consistency of $\mathbf{b}_n$ in the Gauss-Markov model. A very useful additional assumption on $\{\epsilon_n\}$, which is satisfied by many important classes of random variables that are natural models for random noise, and which yields the desired strong limit theorems, takes the form of the following moment inequality, to be satisfied for some $p>2$ and all constants c_i:

(1.9) $$E|\Sigma_{i=m}^n c_i \epsilon_i|^p \leq K_p(\Sigma_{i=m}^n c_i^2)^{p/2} \qquad \text{for all } n \geq m.$$

Given $p>0$, a sequence of random variables $\{\epsilon_n\}$ is called a *lacunary system of order* p, or an S_p *system*, if there exists a positive constant K_p such that the moment inequality (1.9) is satisfied for all constants c_i. The concept of S_p systems was introduced by Banach (1930) and Sidon (1934). If $\{\epsilon_n\}$ is an S_p system for all $p>0$, then it is called an S_∞ system. In view of (1.2), an equinormed orthogonal system is an S_2 system, and the moment inequality (1.9) can be regarded as an L_p extension of the L_2 property (1.2). In Section 2, we give some basic properties and examples of S_p systems, and in this connection, review some important moment inequalities in the literature. In particular, we also discuss how the moment inequality (1.9) in the case $p>2$ is related to the almost sure limiting behavior of the sequence $\{\Sigma_{i=1}^n c_i \epsilon_i\}$.

While the moment restriction (1.9) appears more restrictive than the equinormed orthogonal situation (1.2) in the sense that it considers the p^{th} absolute moment with $p>2$, it is also less restrictive than (1.2) in the sense that it replaces equality in (1.2) by an upper inequality ($\leq$). If we replace equality in (1.2) by a lower inequality ($\geq$), then we get a Bessel-type inequality. A sequence of random variables $\{\epsilon_n\}$ is said to satisfy the *Bessel inequality* if there exists $K>0$ such that for all constants c_i,

(1.10) $$E(\Sigma_{i=m}^n c_i \epsilon_i)^2 \geq K\Sigma_{i=m}^n c_i^2 \qquad \text{for all } n \geq m$$

(cf. Gaposhkin (1966)). Since $E|Y|^p \geq (EY^2)^{p/2}$ for $p>2$, the inequality (1.10) in turn implies the existence of a positive constant $A_p>0$ such that for all constants c_i,

(1.11) $$E|\Sigma_{i=m}^n c_i \epsilon_i|^p \geq A_p(\Sigma_{i=m}^n c_i^2)^{p/2} \qquad \text{for all } n \geq m.$$

Gaposhkin (1966) showed that if $\{\epsilon_n\}$ is an S_p system for some $p>2$ and if it also satisfies the Bessel inequality, then it is a *Banach system*, i.e., there exists $A(=A_1)$ such that (1.11) holds with $p=1$ for all constants c_i. Clearly, if $\{\epsilon_n\}$ is a Banach system, then for every $p \geq 1$, there exists $A_p>0$ such that (1.11) holds for all constants c_i.

Replacing equality in (1.2) by the upper inequality (1.9) (with $p=2$) and the lower inequality (1.10) enables us to substantially enlarge the equinormed orthogonal model for random noise and include random errors that are correlated and have different variances. Assuming (1.9) for some $p>2$ in addition often enables us to extend the mean square convergence properties in classical regression and time series models with equinormed orthogonal errors to the corresponding almost sure convergence properties. For example, as shown by Lai and Wei (1983), if the random errors ϵ_n in the linear process (1.7) form an orthonormal S_p system with $p>2$, then the weak law (1.8) can indeed be strengthened into the strong law, i.e.,

$$n^{-1}\Sigma_{j=1}^{n} y_j \to 0 \quad \text{a.s.}$$

To establish the strong consistency of the least squares estimate $\mathbf{b}_n = (b_{n1}, \ldots, b_{nk})'$ in the multiple regression model (1.3) when the random errors ϵ_i form an S_p system with $p>2$, we fix $j=1, \ldots, k$ and note that b_{nj} can be represented for all large n as

$$b_{nj} - \beta_j = (\Sigma_{i=1}^{n} a_{ni}\epsilon_i)/(\Sigma_{i=1}^{n} a_{ni}^2), \tag{1.12}$$

where $\{a_{ni}: 1 \leq i \leq n, n=1, 2, \ldots\}$ is a triangular array of constants such that

$$\Sigma_{i=1}^{m} a_{ni}a_{mi} = \Sigma_{i=1}^{m} a_{mi}^2 \qquad \text{for } n \geq m \tag{1.13}$$

(cf. Lai and Wei (1982), Lemma 2). Thus, to study the limiting behavior of the least squares estimate b_{nj}, it is useful to consider more generally linear transformations of the form

$$x_n = \Sigma_{i=-\infty}^{\infty} a_{ni}\epsilon_i \tag{1.14}$$

where a_{ni} are constants such that $\Sigma_{i=-\infty}^{\infty} a_{ni}^2 < \infty$ for every n. Since $\{\epsilon_n\}$ is an S_p system with $p>2$, the series in (1.14) indeed converges a.s. (see Section 2). Partial sums $x_n = \Sigma_{i=1}^{n} y_i$ of the linear process $\{y_i\}$ defined in (1.7) can also be expressed in the form (1.14). In Section 3, we consider the almost sure limiting behavior of such linear transformations of S_p systems and discuss applications of the results to regression and time series models.

2. Lacunary systems, Banach systems, and related moment inequalities. We now give some examples of S_p systems and Banach systems, and in this connection, also review some important moment inequalities in the literature.

Example 1. If $\{\epsilon_n\}$ are i.i.d. standard normal random variables, then since $E|\Sigma_{i=m}^{n} c_i\epsilon_i|^p = (\Sigma_{i=m}^{n} c_i^2)^{p/2} E|N(0,1)|^p$, $\{\epsilon_n\}$ is an S_∞ system and a Banach system.

Example 2. Let $\{\epsilon_n\}$ be i.i.d. Bernoulli random variables such that $P\{\epsilon_n = 1\} = P\{\epsilon_n = -1\} = ½$. Then by an inequality of Khintchine (1924), for every $p>0$, there exist positive constants A_p and B_p such that

$$A_p(\Sigma_{i=m}^{n} c_i^2)^{p/2} \leq E|\Sigma_{i=m}^{n} c_i\epsilon_i|^p \leq B_p(\Sigma_{i=m}^{n} c_i^2)^{p/2} \tag{2.1}$$

for all $n \geq m$ and all constants c_i. Thus, Khintchine's inequality implies that $\{\epsilon_n\}$ is an S_∞ system and a Banach system.

Khintchine's inequality was generalized to general independent random variables by Marcinkiewicz and Zygmund (1937) who showed that if ϵ_n are independent random variables with zero means, then for every $p \geq 1$, there exist positive constants A_p and B_p depending only on p such that

$$A_p E\{(\Sigma_{i=m}^{n} \epsilon_i^2)^{p/2}\} \leq E|\Sigma_{i=m}^{n} \epsilon_i|^p \leq B_p E\{(\Sigma_{i=m}^{n} \epsilon_i^2)^{p/2}\} \tag{2.2}$$

for all $n \geq m$. In the case $p>1$, the moment inequality (2.2) was extended from independent random variables to martingale difference sequences $\{\epsilon_n\}$ by Burkholder (1966). Some other important martingale extensions of the Marcinkiewicz-Zygmund inequality can be found in Burkholder's survey paper (1973) and the references therein.

Making use of Burkholder's inequality (2.2) for martingale difference sequences and Minkowski's inequality, Lai and Wei (1983) obtained

Example 3. Let $p \geq 2$, and let $\{\epsilon_n\}$ be a martingale difference sequence (i.e., $E\{\epsilon_n | \epsilon_j, j \leq n-1\} = 0$ for all n) such that $\sup_n E|\epsilon_n|^p < \infty$. Then $\{\epsilon_n\}$ is an S_p system. If furthermore $\inf_n E|\epsilon_n| > 0$, then it follows from Lemma 4 of Burkholder (1968) that $\{\epsilon_n\}$ is also a Banach system.

Let r be a positive even integer. A sequence of random variables $\{\epsilon_n\}$ is said to be *multiplicative of order r* if

$$E(\epsilon_{i_1} \ldots \epsilon_{i_r}) = 0 \text{ for all } i_1 < i_2 < \ldots < i_r. \tag{2.3}$$

When $r=2$, this reduces to orthogonal random variables and therefore forms an S_2 system if $\sup_i E\epsilon_i^2 < \infty$. For $r \geq 4$, Komlós (1972) obtained the following

Example 4. Let $r \geq 4$ be an even integer, and let $\{\epsilon_n\}$ be a multiplicative sequence of order r such that $\sup_i E\epsilon_i^r < \infty$. Then as shown by Komlós (1972), $\{\epsilon_n\}$ is an S_r system. Obviously, if $\inf_n E\epsilon_n^2 > 0$, then $\{\epsilon_n\}$ satisfies the Bessel inequality, and this in turn implies that $\{\epsilon_n\}$ is a Banach system since it is an S_r system ($r>2$) satisfying the Bessel inequality. Longnecker and Serfling (1978) introduced three different ways to weaken the multiplicative condition (2.3) and showed that these three different classes of weakly multiplicative systems of order r are also S_r systems if $\sup_i E\epsilon_i^r < \infty$. They also showed that certain stationary mixing sequences and Gaussian sequences are special cases of these weakly multiplicative sequences.

The following maximal inequality plays an important role in the theory of S_p systems with $p>2$.

LEMMA 1. (Móricz (1976)). *Let $p>0$ and $\alpha>1$. Let $\{x_n\}$ be a sequence of random variables. Suppose that there exist nonnegative constants d_i such that*

$$E|x_n - x_m|^p \leq (\Sigma_{i=m+1}^n d_i)^\alpha \qquad \text{for } n>m\geq m_0. \tag{2.4}$$

Then there exists an absolute constant $C_{p,\alpha}$ such that

$$E(\max_{m\leq i\leq n} |x_i - x_m|^p) \leq C_{p,\alpha}(\Sigma_{i=m+1}^n d_i)^\alpha \qquad \text{for } n>m\geq m_0. \tag{2.5}$$

As a consequence of the maximal inequality (2.5), we obtain the following corollary on the almost sure convergence of $\{x_n\}$ and also its order of magnitude in case of divergence (cf. Lai and Wei (1983), Lemma 3.2).

COROLLARY 1. *With the same notation and assumptions as in Lemma 1, define*

$$D_n = \Sigma_{i=m_0+1}^n d_i. \tag{2.6}$$

(i) *If $\lim_{n\to\infty} D_n < \infty$, then x_n converges a.s. and in the L_p-norm.*

(ii) *If $\lim_{n\to\infty} D_n = \infty$, then for every $\delta > 0$,*

$$x_n = o(\{D_n^{\alpha/p}(\log D_n)^{1/p}(\log\log D_n)^{(1+\delta)/p}\}) \quad a.s. \tag{2.7}$$

Remark. Suppose that $\{\epsilon_n\}$ is an S_p system for some $p>2$ and $\{c_n\}$ is a sequence of constants. Let $x_n = \Sigma_{i=1}^n c_i\epsilon_i$. Then (1.9) implies that $\{x_n\}$ satisfies (2.4) with $\alpha = p/2>1$, $d_i = K_p^{2/p}c_i^2$ and $m_0=1$. Therefore Lemma 1 and Corollary 1 are applicable to $\{x_n\}$.

The special case $p=\alpha=4$ in Lemma 1 was first established by Erdös (1943) for lacunary trigonometric series. The result of Erdös was subsequently extended by several authors (cf. Móricz (1976) and the references therein), and Móricz (1976) considered in addition to the case $\alpha>1$ in Lemma 1 also the cases $\alpha=1$ and $0<\alpha<1$. The latter two cases are quite different from the case $\alpha>1$; instead of the absolute constant $C_{p,\alpha}$ in (2.5), the corresponding maximal inequalities in these two cases involve constants of the form $C_{p,\alpha}(m,n)$. These results generalize the classical Rademacher-Mensov inequality for orthogonal random variables (cf. Doob (1953), page 156): If $\epsilon_1, \ldots, \epsilon_n$ are orthogonal random variables with finite variances $\sigma_1^2, \ldots, \sigma_n^2$, then

$$E\{\max_{1\leq j\leq n}(\Sigma_{i=1}^{j}\epsilon_i)^2\} \leq \left(\frac{\log 4n}{\log 2}\right)^2 \Sigma_{i=1}^{n}\sigma_i^2. \tag{2.8}$$

The following recent generalization of this kind of maximal inequalities is due to Móricz, Serfling and Stout (1982).

LEMMA 2. *Let* $g:\{1, \ldots, n\} \times \{1, \ldots, n\} \to [0,\infty)$ *such that for some* $Q \geq 1$

$$g(i,j) + g(j+1,k) \leq Qg(i,k) \quad \text{for } i \leq j < k, \tag{2.9}$$

$$g(i,j) \leq g(i,j+1) \quad \text{for } i \leq j. \tag{2.10}$$

Let $\epsilon_1, \ldots, \epsilon_n$ *be random variables such that for some* $p \geq 1$ *and* $\alpha \geq 1$,

$$E|\Sigma_{t=1}^{j}\epsilon_t|^p \leq g^{\alpha}(i,j) \quad \text{for all } 1 \leq i \leq j \leq n. \tag{2.11}$$

(i) *If* $\alpha > 1$ *and* $Q < 2^{(\alpha-1)/\alpha}$, *then there exists an absolute constant* $C_{p,\alpha,Q}$ *such that*

$$E(\max_{1\leq j\leq n}|\Sigma_{t=1}^{j}\epsilon_t|^p) \leq C_{p,\alpha,Q}\, g^{\alpha}(1,n). \tag{2.12}$$

(ii) *In the case* $\alpha = 1$, *we have the maximal inequality*

$$E(\max_{1\leq j\leq n}|\Sigma_{t=1}^{j}\epsilon_t|^p) \leq (\Sigma_{t=0}^{[\log n/\log 2]} Q^{t/p})^p\, g(1,n). \tag{2.13}$$

For the case $Q=1$, inequality (2.9) says that g is superadditive and implies (2.10). In this case, as pointed out by Longnecker and Serfling (1977), there exist nonnegative constants $d_1, \ldots, d_n$ such that

$$g(1,n) = \Sigma_{t=1}^{n}u_t \quad \text{and} \quad g(i,j) \leq \Sigma_{t=i}^{j}u_t \quad \text{for} \quad 1 \leq i \leq j \leq n, \tag{2.14}$$

and therefore the maximal inequality (2.12) reduces to that of Lemma 1.

For the case where $g(i,j)$ takes the form $g(i,j) = g(j-i+1)$, (2.9) becomes

$$g(i) + g(j-i) \leq Q\, g(j) \quad \text{for } i \leq j. \tag{2.15}$$

Another maximal inequality of this nature but under an assumption different from (2.15) is

LEMMA 3. (Lai and Stout (1980)). *Let* $g:\{1, 2, \ldots\} \to (0, \infty)$ *be a function satisfying*

$$\liminf_{n\to\infty} g(Kn)/g(n) > K \quad \text{for some integer } K \geq 2, \tag{2.16}$$

and

$$\text{for all } \delta > 0, \text{ there exists } \rho = \rho(\delta) < 1 \text{ for which } \limsup_{n\to\infty}\{\max_{\rho n\leq i\leq n} g(i)/g(n)\} < 1 + d. \tag{2.17}$$

Let $\epsilon_1, \epsilon_2, \ldots$ *be random variables such that for some* $p > 0$,

$$E|\Sigma_{t=v+1}^{v+n}\epsilon_t|^p \leq g(n) \quad \text{for all } v \geq 0 \text{ and } n \geq 1. \tag{2.18}$$

Then there exists a positive constant C *such that*

$$E(\max_{1\leq j\leq n}|\Sigma_{t=v+1}^{v+j}\epsilon_t|^p) \leq C\, g(n) \quad \text{for all } v \geq 0 \text{ and } n \geq 1. \tag{2.19}$$

3. Linear transformations of S_p systems and their applications. In this section we first consider the multiple regression model (1.3) where the random errors ϵ_i form an S_p system with $p > 2$, and apply Corollary 1 to establish the strong consistency of the least squares estimate $\mathbf{b}_n = (b_{n1}, \ldots, b_{nk})'$ under the assumption (1.6) on the design constants. This is the content of

COROLLARY 2. *Suppose that in the multiple regression model* (1.3) *the random variables* ϵ_i *form an* S_p *system for some* $p>2$. *Let* $\mathbf{V}_n = (v_{ij}^{(n)})_{1\leq i,j\leq k} = (\mathbf{T}'_n\mathbf{T}_n)^{-1}$. *Fix* $j=1, \ldots, k$. *If* $\lim_{n\to\infty} v_{jj}^{(n)} = 0$, *then for every* $\delta>1/p$,

$$b_{nj}-\beta_j = o(\{(v_{jj}^{(n)})^{1/2}|\log v_{jj}^{(n)}|^{1/p}(\log|\log v_{jj}^{(n)}|)^{\delta}\}) \quad \text{a.s.} \tag{3.1}$$

Proof. By (1.12), $b_{nj}-\beta_j = (\sum_{i=1}^n a_{ni}\,\epsilon_i)/(\sum_{i=1}^n a_{ni}^2)$ for all large n, where a_{ni} are constants satisfying (1.13). Let $x_n = \sum_{i=1}^n a_{ni}\epsilon_i$. *Since* $\{\epsilon_n\}$ is an S_p system, for $n>m$,

$$\begin{aligned} E|x_n-x_m|^p &= E|\textstyle\sum_{i=1}^m(a_{ni}-a_{mi})\,\epsilon_i + \sum_{i=m+1}^n a_{ni}\epsilon_i|^p \\ &\leq K_p\{\textstyle\sum_{i=1}^m(a_{ni}-a_{mi})^2 + \sum_{i=m+1}^m a_{ni}^2\}^{p/2}. \end{aligned} \tag{3.2}$$

Let $D_n = \sum_{i=1}^n a_{ni}^2$, $d_n = D_n-D_{n-1}(D_0=0)$. It follows from (1.13) that for $n>m$

$$\textstyle\sum_{i=1}^m(a_{ni}-a_{mi})^2 + \sum_{i=m+1}^n a_{ni}^2 = \sum_{i=1}^n a_{ni}^2 - \sum_{i=1}^m a_{mi}^2 = D_n-D_m,$$

and therefore by (3.2), for $n>m$

$$E|x_n-x_m|^p \leq K_p(D_n-D_m)^{p/2} = K_p(\textstyle\sum_{i=m+1}^n d_i)^{p/2}.$$

As $n\to\infty$, $D_n = 1/v_{jj}^{(n)} \to \infty$ (cf. Lai, Robbins and Wei (1978)), and therefore we can apply Corollary 1 (ii) to obtain that for every $\delta>1/p$,

$$x_n = o(\{D_n^{1/2}(\log D_n)^{1/p}(\log\log D_n)^{\delta}\}) \quad \text{a.s.} \tag{3.3}$$

proving the desired conclusion (3.1). □

Corollary 2 extends the result of Lai, Robbins and Wei (1978) who considered the special case $p=4$. The above proof also shows that the linear transformation $x_n=\sum_{i=1}^n a_{ni}\,\epsilon_i$ of an S_p system $\{\epsilon_i\}$ has the asymptotic behavior (3.3) if $D_n = \sum_{i=1}^n a_{ni}^2 \to \infty$ and if the constants a_{ni} satisfy (1.13).

More generally, let $\{a_{ni} : n\geq 1, -\infty<i<\infty\}$ be a double array of constants such that

$$\textstyle\sum_{i=-\infty}^{\infty} a_{ni}^2 < \infty \text{ for every } n. \tag{3.4}$$

Thus, $\mathbf{a}_n = (a_{ni})_{-\infty<i<\infty}\in\ell^2$, and we shall let $\|\mathbf{a}_n\| = (\sum_{i=-\infty}^{\infty} a_{ni}^2)^{1/2}$ denote the ℓ^2 norm of $\mathbf{a}_n$. Let $\{\epsilon_n\}_{-\infty<n<\infty}$ be an S_p system with $p>2$. Define

$$x_n = \textstyle\sum_{i=-\infty}^{\infty} a_{ni}\epsilon_i, \tag{3.5}$$

noting that the series in (3.5) converges a.s. and in the L_p norm in view of Corollary 1(i) and (3.4). By (1.9),

$$E|x_n-x_m|^p \leq K_p\{\textstyle\sum_{i=-\infty}^{\infty}(a_{ni}-a_{mi})^2\}^{p/2} = K_p\|\mathbf{a}_n-\mathbf{a}_m\|^p. \tag{3.6}$$

If furthermore $\{\epsilon_n\}$ satisfies the Bessel inequality, then $E(x_n-x_m)^2 \geq K\sum_{i=-\infty}^{\infty}(a_{ni}-a_{mi})^2$ by (1.10), and it then follows from (3.6) that

$$E|x_n-x_n|^p \leq K'_p\{E(x_n-x_m)^2\}^{p/2}. \tag{3.7}$$

This inequality in turn enables us to relate the L_p properties of $\{x_n\}$ to its L_2 and spectral properties. Making use of this observation, Lai and Wei (1983) obtained the following

THEOREM 1. *Consider the linear process* y_n *defined in* (1.7) *where the random errors* ϵ_n *form an orthonormal* S_p *system with* $p>2$. *Let* f *be the spectral density of* $\{y_n\}$. *If* ess $\sup_{0\leq\theta\leq 2\pi} f(\theta) < \infty$, *then* $\{y_n\}$ *is an* S_p *system. Consequently,* $\sum_{i=1}^{\infty} c_i y_i$ *converges a.s. and in the* L_p *norm for all constants* c_i *such that* $\sum_1^{\infty} c_i^2 < \infty$.

In view of the inequality (3.6), the random variables $x_n = \sum_{i=-\infty}^{\infty} a_{ni}\epsilon_i$ satisfy moment inequalities of the type in Lemma 1 or 2 or 3 if the function $h(m,n) = \|\mathbf{a}_n - \mathbf{a}_m\|$ satisfies corresponding conditions of the type $h(m,n) \leqslant (\sum_{i=m+1}^{\infty} d_i)^{\alpha/p}$, or $h(m,n) \leqslant g^{\alpha/p}(m+1,n)$, or $h(m, n) \leqslant g^{1/p}(n-m)$ for $n>m$. Under such assumptions on $\|\mathbf{a}_n - \mathbf{a}_m\|$, we can therefore apply the maximal inequalities in these lemmas to obtain almost sure limit theorems of the type in Corollary 1 above or in Corollary 3.3 of Lai and Wei (1983) for linear transformations $x_n = \sum_{i=-\infty}^{\infty} a_{ni}\epsilon_i$ of S_p systems $\{\epsilon_n\}$ satisfying the Bessel inequality. Such maximal inequalities can also be applied in conjunction with exponential bound of the type

$$P\{|x_n| > \tau(\theta)(D_n \log\log D_n)^{1/2}\} = 0(\exp(-\theta \log\log D_n)), \tag{3.8}$$

where $D_n = \sum_{i=-\infty}^{\infty} a_{ni}^2$, $\theta>1$ and $\tau(\theta)>0$, to establish laws of the iterated logarithm for x_n (cf. Lai and Wei (1982), Theorem 4). Using this approach and certain truncation techniques, Lai and Wei (1982) obtained the following law of the iterated logarithm for double arrays of independent random variables and applied the result to regression and time series problems.

THEOREM 2. *Let* $\ldots, \epsilon_{-1}, \epsilon_0, \epsilon_1, \ldots$ *be independent random variables such that*

$$E\epsilon_n = 0 \text{ and } E\epsilon_n^2 = \sigma^2 \text{ for all } n, \text{ and } \sup_n E|\epsilon_n|^p < \infty \text{ for some } p > 2, \tag{3.9}$$

and let $\{a_{ni} : n \geq 1, -\infty<i<\infty\}$ *be a double array of constants such that* (3.4) *holds and*

$$D_n = \sum_{i=-\infty}^{\infty} a_{ni}^2 \to \infty, \tag{3.10}$$

$$\sup_n a_{ni}^2 = o(D_n(\log D_n)^{-r}) \quad \text{for all } r>0. \tag{3.11}$$

Let $x_n = \sum_{i=\infty}^{\infty} a_{ni}\epsilon_i$.

(i) *If there exist constants* $d_i \geqslant 0$ *and* $\lambda > 1/p$ *such that*

$$\|\mathbf{a}_n - \mathbf{a}_m\| < (\sum_{i=m+1}^{\infty} d_i)^{\lambda} \quad \text{for } n>m\geqslant m_0, \text{ and} \tag{3.12}$$

$$(\sum_{i=1}^{n} d_i)^{\lambda} = 0(D_n^{1/2}) \quad \text{as } n\to\infty, \tag{3.13}$$

then

$$\limsup_{n\to\infty} |x_n|/(2D_n \log\log D_n)^{1/2} \leqslant \sigma \quad a.s. \tag{3.14}$$

(ii) *If* $\|\mathbf{a}_n - \mathbf{a}_m\| \leqslant g^{1/p}(n-m)$, *where* $g : \{1, 2, \ldots\} \to (0,\infty)$ *satisfies conditions* (2.16) *and* (2.17) *and* $g(n) = 0(D_n^{p/2})$, *then* (3.14) *still holds.*

REFERENCES

BANACH, S. (1930). Über einige Eigenschaften der lacunaren trigonometrische Reihen. *Studia Math.* *2* 207–220.

BURKHOLDER, D. L. (1966). Martingale transforms. *Ann. Math. Statist.* *37* 1494–1504.

BURKHOLDER, D. L. (1968). Independent sequences with the Stein property. *Ann. Math. Statist.* *39* 1282–1288.

BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probability* *1* 19–42.

CHEN, G. C., LAI, T. L., and WEI, C. Z. (1981). Convergence systems and strong consistency of least squares estimates in regression models. *J. Multivariate Anal.* *11* 319–333.

DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

DRYGAS, H. (1976). Weak and strong consistency of the least squares estimators in regression models. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* *34* 119–127.

ERDÖS, P. (1943). On the convergence of trigonometric series. *J. Math. Phys. (M.I.T.)* *22* 37–39.

GAPOSHKIN, V. F. (1966). Lacunary series and independent functions. *Russian Math. Surveys* *21*, No. 6, 1–82.

KAILATH, T. (1974). A view of three decades of linear filtering theory. *IEEE Trans. Inform. Theory IT-20* 146–181.

KHINTCHINE, A. (1924). Über einen Satz der Wahrscheinlichkeitsrechnung. *Fund. Math. 6* 9–20.

KOMLÓS, J. (1972). On the series $\Sigma c_k \varphi_k$. *Studia Sci. Math. Hungar. 7* 451–458.

LAI, T. L., ROBBINS, H., and WEI, C. Z. (1978). Strong consistency of least squares estimates in multiple regression. *Proc. Nat. Acad. Sci. USA 75* 3034–3036.

LAI, T. L., and STOUT, W. (1980). Limit theorems for sums of dependent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 51* 1–14.

LAI, T. L., and WEI, C. Z. (1982). A law of the iterated logarithm for double arrays of independent random variables with applications to regression and time series models. *Ann. Probability 10* 320–335.

LAI, T. L., and WEI, C. Z. (1983). Lacunary systems and generalized linear processes. *Stochastic Processes Appl. 14* 187–199.

LONGNECKER, M. and SERFLING, R. J. (1977). General moment and probability inequalities for the maximum partial sum. *Acta Math. Acad. Sci. Hungar. 30* 129–133.

LONGNECKER, M. and SERFLING, R. J. (1978). Moment inequalities for S_n under general dependence restrictions with applications. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 43* 1–21.

MARCINKIEWICZ, J. et ZYGMUND, A. (1937). Sur les fonctions independantes. *Fund. Math. 29* 60–90.

MÓRICZ, F. (1976). Moment inequalities and the strong laws of large numbers. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 35* 299–314.

MÓRICZ, F., SERFLING, R., and STOUT, W. (1982). Moment and probability bounds with quasi-superadditive structure for the maximum partial sum. *Ann. Probability 10* 1032–1040.

SIDON, S. (1934). Ein Satz über Fourierische Reihen mit Lücken. *Math. Z. 32* 481–482.