A NOTE ON STRONG MIXING

By Soutir Bandyopadhyay

Department of Statistics

Iowa State University

April 21, 2006

Abstract

The strong mixing property for a sequence of random variables is interesting in its own right. It is discussed that under what conditions the strong mixing property holds for linear stochastic processes and in particular ARMA processes. Then an example of Non-Strong mixing Autoregressive Processes is discussed here.

Keywords: Harris Chain, Stationary Processes, Strong Mixing, ARMA Processes

1 Introduction

In order to be able to state our results precisely, let us start with a review of mixing conditions. Let $\{X_t : t \in \mathbb{Z}\}$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{F}_n^m = \sigma\{X_t : n \leq t \leq m\}$ be the σ-algebra generated by the random variables $\{X_n, \cdots, X_m\}$. Define

$$\alpha\,(m) = \sup |P(E \cap F) - P(E)P(F)|, \tag{1.1}$$

where the supremum is taken over all $E \in \mathcal{F}_{-\infty}^n$, $F \in \mathcal{F}_{n+m}^\infty$ and n. We say that $\{X_t\}$ is *strong mixing* if $\alpha\,(m)$ tends to zero as m increases to infinity.

The strong mixing condition was introduced by Rosenblatt in 1956 to prove the central limit theorem for 'weakly dependent' random variables. Since then it has assumed a position of considerable importance in probability theory. This is due to its tractability in the derivation of asymptotic properties of various functions of sequences of dependent random variables. Its areas of application are wide and include central limit theorems, strong laws of large numbers, laws of iterated logarithm, empirical processes, order statistics and robust estimators. Although the strong mixing condition has been widely adopted in the literature, a clear understanding of the condition itself is lacking. The categorization of well-known processes as strong mixing or non-strong mixing is still far from complete. However Chanda [1974] has shown that members of the important class of linear stochastic processes are strongly strong mixing, provided they are based on innovation random variables which have Lebesgue-integrable characteristic functions. The latter condition is not superfluous. Withers [1981] has given some conditions for a linear process to be strong mixing. Ibragimov and Linnik [1971] and Chernick [1981] give examples of first-order autoregressive (AR(1)) processes based on discrete-valued innovation random variables which are not strong mixing. Unfortunately, their proofs are by contradiction, and do not give much insight into the reason why the strong mixing condition fails. But Andrews [1984] shows that certain AR(1) processes are *not* strong mixing by explicitly constructing sequences of sets which violate the strong mixing condition. Again, Athreya and Pantula [1986a] have shown that a Harris-recurrent Markov chain on a general state space is strong mixing, provided there exists a stationary probability distribution for that Markov chain. They have also established that certain stationary autoregressive moving average(ARMA) processes are strong mixing (See Athreya and Pantula [1986b]).

2 Results

2.1 Strong Mixing Properties of Linear Stochastic Processes

2.1.1 Results for The Univariate Linear Stochastic Processes

There are a lot of articles on the strong mixing property of linear stochastic processes. Chanda [1974] has given rate of strong mixing of a linear stochastic process under some conditions. The main result of that article by Chanda is given in the following theorem.

Theorem :Let $\{Z_t : t \in \mathbb{Z}\}$ be a pure white noise process with $\gamma = E\left\{|Z_1|^\delta\right\} < \infty$ for some $\delta > 0$. Assume that the characteristic function of Z_1 is integrable, $(2\pi)^{-1}\int |\phi_0(u)|du \leq 1$, and let $\{X_t : t \in \mathbb{Z}\}$ be a linear process defined as $X_t = \sum_{j=0}^{\infty} g_v Z_{t-v}$, where $\sum_{v=0}^{\infty} v|g_v|^\lambda < \infty$, $\lambda = \delta(1+\delta)^{-1}$; then X_t is strongly mixing in the sense that

$$|P(A \cap B) - P(A)P(B)| \leq M\beta(k), \tag{2.2}$$

for all $A \in \mathcal{F}_{-\infty}^{0}$, $B \in \mathcal{F}_k^\infty$ where M is a finite positive constant depending only on ϕ_0 and

$$\beta(k) = \sum_{v=k}^{\infty} v|g_v|^\lambda,$$

By a pure white noise process the author means that the random variables Z_t are independent and identically distributed.

In order to prove this theorem the author first proves a lemma which is as follows:

lemma :Let

$$W_t = \sum_{v=0}^{t-1} g_v Z_{t-v}, \qquad k \leq t \leq k+m-1, (m \geq 1);$$

$W = (W_k, \cdots, W_{k+m-1})$. Assume that the characteristic function ϕ_0 of Z_1 is integrable. Then the d.f. of W_t admits a p.d.f. f_t which is bounded and continuous everywhere. Similarly, the d.f. of W admits a p.d.f. $f_{k,\cdots,k+m-1}$ which is bounded and continuous everywhere.

Now using the lemma the author has proved the result for any Borel set C in the space of $\{X_{-p}, \cdots, X_0\}$ and any arbitrary disjoint union D of intervals in the space of $\{X_k, X_{k+1}, \cdots\}$. Let $\mathcal{C}$ be the class of inverse images of all Borel sets in the space of $\{X_k, X_{k+1}, \cdots\}$ for which (2.2) holds. Then the following result holds. If $B \in \mathcal{C}, B^c \in \mathcal{C}$; also if $\{B_n\}$ is a monotone sequence of sets $\in \mathcal{C}$ then $\lim B_n = B \in \mathcal{C}$. The result follows from the property of probability measure because if we write $X^{-1}(C) = A, \;\; Y^{-1}(D_n) = B_n$ then

$$|P(A \cap B_n) - P(A)P(B_n)| \leq M\beta(k),$$

implies that

$$|P(A \cap B) - P(A)P(B)| \leq M\beta(k),$$

where $B = lim_{n\to\infty} B_n$. Further $\mathcal{C} \supset \mathcal{M}$ which is the field over the inverse images of all intervals in the space of $\{X_k, X_{k+1}, \cdots\}$. Therefore $\mathcal{C}$ is a monotone class over the field $\mathcal{M}$, and hence $\mathcal{C} = \mathcal{F}_k^\infty$. Since C is arbitrary so that $X^{-1}(C) \in \mathcal{F}_{-\infty}^0$, the result of the theorem follows immediately.

If $\{X_t\}$ is a stationary Gaussian (without being necessarily a linear process) then Rozanov [1967] has proved the process is strong mixing, provided the spectral density function $f(\mu)$ of the process exists everywhere and is continuous and non-vanishing over $[-\pi, \pi]$. If, additionally, $\log f(\mu)$ is integrable over $[-\pi, \pi]$ then by virtue of Wold's decomposition theorem (see Doob [1953]) $\{X_t\}$ is regular, without any singular component and hence is a linear process. But the result proved by Chanda [1974] holds for any linear process with or without any second order properties and covers both Gaussian and non-Gaussian processes.

After some years, Gorodetskii [1977] points out that Chanda's result is wrong and gives a counterexample. He then corrects the result of Chanda by adding more conditions (too complicated to give here) and manages to avoid the assumption that the $\{Z_t\}$ are identically distributed.

Withers [1981] gives an alternative set of conditions for linear processes to be strong mixing based on Section 7 of Withers [1978]. The theorem is as follows:

Theorem : Let $\{Z_j\}$ be independent r.v.s on $\Re$ with characteristic functions $\{\phi_j\}$ such that

$$K = (2\pi)^{-1} \max_j \int |\phi_j(t)| dt < \infty, \tag{2.3}$$

and for some $\delta > 0$

$$\gamma = \max_j E|Z_j|^\delta < \infty \tag{2.4}$$

Let $\{g_j\}$ be complex numbers such that

$$G_t = S_t(\min(1,\delta))^{\max(1,\delta)} \to 0 \quad \text{as} \quad t \to \infty,$$

where

$$S_t(\delta) = \sum_{v=t}^{\infty} |g_v|^\delta,$$

Then for all t, $X_{nt} = \sum_{j=0}^{n} g_j Z_{t-j}$ converges in probability to a r.v. X_t as $n \to \infty$. Suppose

$$M_0 = \sup_{m,s,k \geq 1} \sup_{\alpha,\beta,\nu} \max_t \left| \frac{\partial}{\partial \nu_t} \quad P\left(W + \nu \in \bigcup_{j=1}^{s} D_j \right) \right| < \infty, \tag{2.5}$$

where

$$D_j = \prod_{t=k}^{k+m+1} (\alpha_{jt}, \beta_{jt}), W_t = X_{t-1,t}, V_t = X_t - W_t, \nu = (\nu_k, \cdots, \nu_{k+m-1}), W = (W_k, \cdots, W_{k+m-1}).$$

Then for $\{X_t\}$,

$$\alpha(k) \leq 2(4M_0 + \gamma)\alpha_0(k),$$

where,

$$\alpha_0(k) = \sum_{t=k}^{\infty} G_t$$

Corollary : Suppose that the conditions (2.3),(2.4),(2.5) hold and that

$$g_k = O\left(k^{-\nu}\right) \quad \text{where} \quad \nu > 1 + \delta^{-1} + \max(1, \delta^{-1}).$$

Then for $\{X_t\}, \alpha(k) = O\left(k^{-\varepsilon}\right)$ where

$$\varepsilon = (\nu\delta - \max(1,\delta))\,(1+\delta)^{-1} - 1 > 0.$$

Note : Theorem 2.1 yields (for the i.i.d. case) the weaker result $\varepsilon = \nu\delta\,(1+\delta)^{-1} - 2$

2.1.2 Results for The Multivariate Linear Stochastic Processes

Chanda, Gorodetskii and Withers consider only the univariate case while Pham and Tran [1985] give a result for the multivariate case.

Let $X(t), t = \cdots, -1, 0, 1, \cdots$ be a p-variate random process, $\mathcal{F}_n^m = \sigma\{X(t) : n \le t \le m\}$ be the σ-algebra generated by $\{X(n), \cdots, X(m)\}$. Let $X = (\cdots, X(-1), X(0))$, $Y = (X(n), X(n+1), \cdots)$ and P_{XY}, P_X, P_Y be respectively the joint distribution Of X, Y and the marginal distributions of X, Y.

Define a function $\Delta_n(x)$ by the condition that for any measurable set A of $\Re^p \times \Re^p \times \cdots$

$$\int \Delta_n(x) P_X(dx) = \sup_{|h|\le 1} \left\{ \int \int_{A}, \; h(x,y)[P_X(dx)P_Y(dy) - P_{X,Y}(dx,dy)] \right\}, \tag{2.6}$$

Now, Δ_n always exists since the right-hand side of (2.6) defines a measure absolutely continuous with respect to $P_X(dx)$. If the conditional distribution of $P_Y^X(dy|x)$ of Y given X exists, then Δ_n is just the total variation of $P_Y^X(dy|x) - P_Y(dy)$.

Let $\|\Delta_n\|_s$ be the L^s norm of Δ_n. Then $X(t)$ is said to satisfy the Gastwirth and Rubin condition if $\|\Delta_n\|_s \to 0$ as $n \to \infty$ for some $0 < s < \infty$.

When $\|\Delta_n\|_1 \to 0$ as $n \to \infty$, the process is often referred to as absolutely regular, weakly Bernoulli or completely regular.

Let us define $\alpha(n)$ as (1.1) and say $X(t)$ satisfies the strong mixing condition if $\alpha(n) \to 0$ as $n \to \infty$.

It can be shown that $\alpha(n) \le 4\,\|\Delta_n\|_1$ and hence the condition that $\|\Delta_n\|_1 \to 0$ as $n \to \infty$ is stronger than strong mixing. Some results for absolutely regular process do not hold just under strong mixing (See, for example, Berbee [1979] or Bradley [1983]). Volkonskii and Rozanov [9, p.187] have pointed out that the condition of absolute regularity is also more suitable for research. It is thus of interest to determine whether a process is absolutely regular.

Assume that there exists a sequence of independent random vectors $e(t)$, and matrices $A(j)$ such that

$$X(t) = \sum_{j=0}^{\infty} A(j)e(t-j), \quad A(0) = I, \tag{2.7}$$

where I is the identity matrix. We further assume that the $e(t)$ admit a density, say g_t. Pham and Tran [1985] studies the convergence of $\|\Delta_n\|_1$ to zero for linear processes and prove the following theorem.

Theorem: Suppose that the following conditions hold:

(a)$\int |g_t(v-u) - g_t(v)|dv < K_1 \|u\|$ for all t;

(b)$\sum_{j=0}^{\infty} \|A(j)\| < \infty$ and $\sum_{j=0}^{\infty} A(j)z^j \neq 0$ for all z with $|z| \leq 1$.

(c)$E \|e(t)\|^{\delta} < K_2$, for some $\delta > 0$ and for all t;

If

$$\sum_{j=1}^{\infty} \alpha(j)^{\frac{\delta}{1+\delta}} < \infty \quad \text{where} \quad \alpha(j) = \sum_{k \geq j} \|A_k\|,$$

Then

$$\|\Delta_n\|_1 \leq K_3 \sum_{j=n}^{\infty} \alpha(j)^{\frac{\delta}{1+\delta}} \quad \text{and} \quad X_t \text{ is absolutely regular.}$$

where K_1, K_2, K_3 are constants whose values are unimportant.

To prove the theorem they first prove the following two lemmas.

lemma 1: Let $r(n) = \sum_{j=n}^{\infty} A(j)e(n-j)$ and $\xi(n) = X(n) - r(n)$. Then $(\xi(n), \cdots, \xi(n+m))'$ admits a density, say, $f_{n,m}$ and

$$\Delta_n(X) \leq \sup_{m \geq 0} \{E[\delta_{n,m}(R_{n,m})|X] + E[\delta_{n,m}(R_{n,m})]\} \quad a.s.$$

where $R_{n,m} = (r(n), \cdots, r(n+m)), X = (\cdots, X(-1), X(0))$ and $\delta_{n,m}(u) = \int |f_{n,m}(z-u) - f_{n,m}(z)|dz$.

lemma 2: Suppose that

(a)$\int |g_t(v-u) - g_t(v)|dv < K \|u\|$ for all t;

(b)$\sum_{j=0}^{\infty} \|A(j)\| < \infty$ and $\sum_{j=0}^{\infty} A(j)z^j \neq 0$ for all z with $|z| \leq 1$.Then

$$\sup_{m \geq 0} \delta_{n,m}(R_{n,m}) \leq K \sum_{j=0}^{\infty} \alpha(j+n) \|e(-j)\|$$

where $\alpha(j) = \sum_{k \geq j} \|A_k\|$ and K is a constant.

Proof of the Theorem: Let $\{c_j\}$ be a sequence of positive numbers. Since $\Delta_n \leq 2$ a.s.,we have from lemma 1 and 2, $\|\Delta_n\|_1 \leq K \sum_{j=n}^{\infty} \alpha(j)c_j + 2 \sum_{j=n}^{\infty} P\{\|e(j)\| > c_j\}$. By Schwartz inequality, $P\{\|e(j)\| > c_j\} \leq K/c_j^{\delta}$. The theorem then follows by putting $c_j = \alpha(j)^{-1/1+\delta}$.

2.2 Strong Mixing Properties of Markov chains and ARMA Processes

2.2.1 Results for The Univariate ARMA Processes

In applications of time series like econometrics there is an increasing interest in establishing the asymptotic normality of various estimators of the model parameters assuming the strong mixing property for the basic model. It is, therefore, of interest to see for what classes of time series strong mixing holds, and in particular to see if it is true for stationary ARMA processes.

Athreya and Pantula [1986a] and Athreya and Pantula [1986b] discuss the mixing properties of autoregressive processes. They first establish the strong mixing property for a wide class of Harris-recurrent markov chains. Using this result, they derive a set of sufficient conditions to guarantee the strong mixing property for autoregressive processes.

There was a number of results in the literature related to the strong mixing of Markov chains. Strong mixing was introduced by Rosenblatt in 1956. Rosenblatt [1971] gives necessary and sufficient conditions for a process to be strong mixing. He established that a stationary markov process is strong mixing if and only if it is uniformly pure non-deterministic. He also gives equivalent conditions for strong mixing in terms of the transition operator and the invariant probability measure. Ibragimov and Linnik [1971] established that a stationary gaussian sequence is strong mixing if it has a continuous spectral density that is bounded away from 0.

Chanda [1974] and Withers [1981] have considered strong mixing properties of the process $Y_n = \sum_{j=0}^{\infty} w_j e_{n-j}$. They assume that $\{e_n\}$ are i.i.d. random variables with an integrable characteristic function and with a density function that is Lipschitz in L^1. The condition on the smoothness is used to obtain a bound on $\alpha(m)$. But Athreya and Pantula consider processes of the form $Y_n = \sum_{j=0}^{n-1} w_j e_{n-j} + Z_n$ where $\{w_j\}$ decay exponentially and Z_n converges to 0 in probability. Athreya and Pantula [1986b] derive the strong mixing of such processes and their results are not derivable from that of Withers [1981].

Let $(A, \mathcal{A})$ be a measurable space and $P(\cdot, \cdot) : A \times \mathcal{A} \to [0, 1]$ be a transition function, i.e., $P(x, \cdot)$ is a probability measure on $\mathcal{A}$ for each fixed x in A, and $P(\cdot, E)$ is a measurable function on A for each fixed E in $\mathcal{A}$. Let $\{Y_n : n \geq 0\}$ be a Markov chain with $(A, \mathcal{A})$ as its state space and $P(\cdot, \cdot)$ as its transition function. Thus

$$P[Y_{n+1} \in \cdot \ \ |\mathcal{F}_n] = P(Y_n, \cdot) \ \ a.s.$$

where $\mathcal{F}_n = \mathcal{F}_0^n$

We say that a Markov chain Y_n is $Harris-recurrent$ if there exists a non-trivial σ-finite measure $\psi(\cdot)$ on $(A, \mathcal{A})$ such that,

$$\psi(E) > 0 \Rightarrow P_x[Y_n \in E, \quad for \quad some \quad n \geq 1] = 1,$$

for all x in A where P_x refers to the probability measure corresonding to the initial condition $Y_0 = x$.

A σ-finite measure $\pi(\cdot)$ on $(A, \mathcal{A})$ is called an *invariant measure* for the chain $\{Y_n\}$ if

$$\pi(\cdot) = \int_A P(x, \cdot)\pi(dx)$$

An invariant probability measure $\pi(\cdot)$ is also called a stationary probability distribution forY_n.

Athreya and Pantula [1986b] prove the following theorem:

Theorem: Let Y_t be an autoregressive process given by $Y_t = \rho Y_{t-1} + e_t, t = 1, 2, \cdots$, where $|\rho| \leq 1$ and $|e_t|$ are i.i.d. random variables independent of Y_0.Assume that

(a) $E[\{log|e_1|\}^+]$ is finite, and

(b) e_1 has a non-trivial absolutely continuos component.

Then, for any initial distribution Λ of Y_0, $\{Y_n\}$ is strong mixing.

To prove the theorem they first prove the following three lemmas.

lemma 1: Let $\{Y_t\}$ be a Harris-recurrent Markov chain on a state space $(A, \mathcal{A})$ and with transition function $P(\cdot, \cdot)$. Suppose $\pi(\cdot)$ is stationary probability distribution for $\{Y_t\}$. The $\{Y_t\}$ is strong mixing. (Proved in Athreya and Pantula [1986a])

lemma 2: Let $\{Y_t\}$ be and $AR(1)$ process given by $Y_t = \rho Y_{t-1} + e_t, t = 1, 2, \cdots$,. Assume that, $E[\{log|e_1|\}^+]$ is finite. Then, for any initial distribution of Y_0, Y_n converges in distribution to $U = \sum_{j=1}^{\infty} \rho^j e_j$. Also, $\pi(\cdot) = P[U \in \cdot]$, is a stationary probability measure for the Markov chain $\{Y_t\}$ and it is absolutely continuous. (Proved in Athreya and Pantula [1986a] and Lai and Wei [1982])

lemma 3: Under the hypothesis of the theorem, there exist an integer k, a finite interval $I = (a, b)$, a number $\delta > 0$ and a multiple m of k such that,

(a) for all x in (a, b) and $E \subset (a, b)$, $P_x[Y_{nk} \in E] \geq [\delta\lambda(E)]^n$ and

(b) for any y_0, $P_{y_0}[Y_{jm} \in I \quad for \quad some \quad j \geq 1] = 1$ where $\lambda(\cdot)$ is the Lebesgue measure. (Proved in Athreya and Pantula [1986b])

Proof of the Theorem: By lemma 3, $X_t = Y_{tm}$ is Harris recurrent. Also, from lemma 2, we know that X_t has an invariant probability distribution $\pi(\cdot)$. Therefore, from lemma 1, it follows that X_t is strong mixing. Finally,since $\{Y_t\}$ is a markov chain, Y_t is also strong mixing.

They also generalize the result for the p-th order autoregressive process $\{Y_t\}$ given by $Y_t = \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + e_t$.The generalized result is as follows:

Theorem(Generalized Version): $\{Y_t\}$ is strong mixing provided

(a) $E[\{log|e_1|\}^+]$ is finite,

(b) the distribution of e_1 has a non-trivial absolutely continuos component,

(c) $Y_0 = (Y_0, Y_{-1}, \cdots Y_{1-p})$ is independent of $\{e_j\}$,

(d) $\{e_j\}$ are i.i.d. random variables, and

(e) the roots of the characteristic equation $z^p - \alpha^1 z^{p-1} - \cdots - \alpha_p = 0$, are less than one in modulus.

Again, White and Domowitz [1984] has shown that a finite order moving average of a strong mixing process is itself a strong mixing. The two results together imply that finite order ARMA processes satisfying the above (a)-(e) are strong mixing.

2.2.2 Results for The Multivariate ARMA Processes

Let us define the p-variate random process as before and Δ_n as (2.6). Now let us assume that $X(t)$ is an autoregressive moving average (ARMA) process with values in $\Re^p$. Pham and Tran [1985] studies the convergence of $\|\Delta_n\|_s$ for ARMA processes.

Now $X(t)$ admits a Markovian representation

$$X(t) = HZ(t), \qquad Z(t) = FZ(t-1) + Ge(t), \tag{2.8}$$

where $Z(t)$ are random vectors,H, F, G are appropriate matrices and $e(t)$ are i.i.d. random vectors. Assume that the $e(t)$ have a density g. Here $A(j) = HF^jG$ and hence $r(n)$ and $\xi(n)$ of lemma 1 of section 2.1.2 equals $HF^nZ(0)$; and $\sum_{j=1}^{n} F^{n-j}Ge(j) = H\xi(n)$, say. Now for $u = (u_{n+1}, \cdots, u_{n+m})$,

$$\delta_{n,m}(u) = \sup_{|h|\leq 1} E[h(\xi_{n,m} - u) - h(\xi_{n,m})]$$

where $\xi_{n,m} = (\xi(n), \cdots, \xi(n+m))$. Take $u_j = HF^jz$, then

$$\xi(j) - u_j = HF^{j-n}[\xi(n) - F^n z] + \sum_{i=n}^{j} HF^{i-n}e(i) \quad for \quad j \geq n,$$

Thus,

$$h(\xi_{n,m} - u) = \widetilde{h}[\xi(n) - F^n z, e(n+1), \cdots, e(n+m)],$$

$$h(\xi_{n,m}) = \widetilde{h}[\xi(n) - F^n z, e(n+1), \cdots, e(n+m)] \quad for \quad some \quad function \quad \widetilde{h}.$$

Hence,

$$\delta_{n,m}(u) \leq \sup_{|\widetilde{h}| \leq 1} E\left\{\widetilde{h}[\xi(n) - F^n z, e(n+1), \cdots, e(n+m)] - \widetilde{h}[\xi(n), e(n+1), \cdots, e(n+m)]\right\}.$$

Since the $e(t), t > n$ are independent of $\xi(n)$, if $\xi(n)$ admits a density ϕ_n, then

$$\delta_{n,m}(u) \leq \int |\phi_n(z - F^n u) - \phi_n(z)| dz.$$

Thus, lemma 1 in section 2.1.2 becomes

lemma 1: Let $X(t)$ be the ARMA process as in (2.8). Suppose that $\xi(n) = \sum_{j=1}^{n} F^{n-j} G e(j)$ admits a density ϕ_n. Then

$$\Delta_n \leq E\left\{\delta_n[F^n Z(0)]|X\right\} + E\delta_n[F^n Z(0)],$$

where $\delta_n(z) = \int |\phi_n(v - z) - \phi_n(v)| dv$ and X is as in lemma 1 in section 2.1.2.

lemma 2: Let g be an integrable and f be a bounded function on $\Re^d$. Assume that for some $\gamma > 0$,

$$\int \|x\|^\gamma |g(x)| dx < \infty \quad and \quad f(x) = O(\|x\|^\gamma), \quad x \to 0.$$

Then, as the $d \times d$ matrix u tends to the zero matrix,

$$\int f(ux) g(x) dx = O(\|u\|^\gamma),$$

where $\|u\| = \sup_{\|x\| \leq 1} \|ux\|$.

With the help of these two lemmas Pham and Tran [1985] prove the following theorem regarding the rate of convergence of $\|\Delta_n\|_s$ for ARMA process. The theorem is as follows.

Theorem: Suppose that the eigenvalues of F are of modulus stricly less than 1, then $\|\Delta_n\|_s \to 0$ as $n \to \infty$ for all s. If moreover $\int \|x\|^\delta g(x) dx < \infty$ $\int |g(x) - g(x - \theta)| dx = O(|\theta|^\gamma)$ for some $\delta > 0$ and $\gamma > 0$, then $\|\Delta_n\|_s \to 0$ at an exponential rate for all $0 < s < \infty$.

2.2.3 Non-Strong Mixing Autoregressive Processes

Andrews [1984] shows that certain first-order autoregressive processes are not strong mixing. A direct proof is given. The author first demonstrates that Markov processes do not necessarily satisfy Rosenblatt's strong mixing condition. In particular, he exhibits a class of first-order autoregressive processes given by $X_t = \sum_{l=0}^{\infty} \rho^l \varepsilon_{t-l}$, where $\{\varepsilon_t\}$ is a doubly infinite sequence of independent Bernoulli random variables and $0 < \rho \leq \frac{1}{2}$. His proof is constructive in nature and sheds new light on the characteristics of strong mixing.

Before going into the statement of the results a brief description of the direct proof the certain $AR(1)$ processes are non-strong mixing may be helpful. Suppose $\langle X_t \rangle$ is an $AR(1)$ process based on Bernoulli(q) innovation random variables, and $X_{t,s}$ is equal to X_{t+s} minus its component which depends on $X_t, X_{t-1}, \cdots$. If we know X_t is small, then we know that with probability 1 X_{t+s} must fall in a set which is a small neighbourhood of the support of $X_{t,s}$. A sequence of such small neighbourhoods can be constructed for $s = 1, 2, \cdots$ which have unconditional probability bounded away from 1. Hence, knowledge that X_t is small increases the probability of certain sets which are determined by the 'future' of the process, no matter how far in the future, by a non-negligible amount. This implies $\langle X_t \rangle$ is non-strong mixing.

First he proves two lemmas which are as follows:

lemma 1: If there exists a set A in $\mathcal{F}_{-\infty}^{t}$ with $P(A) > 0$ and sets B_s in $\mathcal{F}_{t+s}^{\infty}$ with $P(B_s) \leq k \;\; \forall s = 1, 2, \cdots$, for some constant $k < 1$, such that

$$P(B_s|A) = 1 \;\; \forall \;\; s,$$

then $\langle X_t \rangle$ is not strong mixing.

To find sets A and B_s as in lemma 1 write $X_{t+s} = X_{t,s} + \rho^s X_t$, where $X_{t,s} = \sum_{l=0}^{s-1} \rho^l \varepsilon_{t+s-l}$. $X_{t,s}$ is independent of X_t. Let W_s be the support of $X_{t,s}$, and $w_j, j = 1, \cdots, J$, be the elements of W_s, where $w_j < w_{j+1}$, and $J \leq 2^s$. If X_t takes a value x_t in $(0, \rho)$, then X_{t+s} must fall in the set $\bigcup_{j=1}^{J} (w_j, w_j + \rho^{s+1})$, since $0 < \rho^s x_t < \rho^{s+1}$. Define,

$$A = \{X_t \in (0, \rho)\}\,, \quad and \quad B_s = \left\{ X_{t+s} \in \bigcup_{j=1}^{J} (w_j, w_j + \rho^{s+1}) \right\}. \tag{2.9}$$

Then $P(A) > 0$. This is actually the proof of the second lemma which is as follows:

lemma 2: Let A and B_s be defined as in (2.9).If A occurs, then B_s occurs, $\forall \;\; s$. That is, $P(B_s|A) = 1 \;\; \forall \;\; s$. Further, $P(A) > 0$.

To get the desired non-strong mixing result it remains to show $P(B_s) \leq k < 1, \quad \forall s$.The upper bound, $w_j + \rho^{s+1}$, of the intervals in B_s has been chosen sufficiently small so that the following result holds.

Theorem: For $A \in \mathcal{F}_{-\infty}^{t}$ *and* $B_s \in \mathcal{F}_{t+s}^{\infty}, \forall \;\; s = 1, 2, \cdots$, as defined in (2.9),

$$|P(A \cap B_s) - P(A)P(B_s)| > P(A)[1-k] > 0 \;\; \forall s, \tag{2.10}$$

where $k \in (0,1)$ is independent of s. That is, the sets A and B_s violate the strong mixing condition (1.1), and the AR(1) process $\langle X_t \rangle$ based on Bernoulli(q) innovation random variables and AR parameter $\rho \in (0, \frac{1}{2}]$ is not strong mixing.

References

Andrews, D. W. K. (1984). Non-strong mixing autoregressive processes. *Journal of Applied Probability 21*, 930–934.

Athreya, K. B. and S. G. Pantula (1986a). Mixing properties of Harris chains and autoregressive processes. *Journal of Applied Probability 23*, 880–892.

Athreya, K. B. and S. G. Pantula (1986b). A note on strong mixing of ARMA processes. *Statistics Probability Letters 4*, 187–190.

Berbee, H. C. P. (1979). *Random walks with stationary increments and renewal theory.* CWI, Math. Centrum [Centrum voor Wiskunde en Informatica].

Bradley, R. C. (1983). Absolute regularity and functions of Markov chains. *Stoch. Proc. Appl. 14*, 67–77.

Chanda, K. C. (1974). Strong mixing properties of linear stochastic processes. *Journal of Applied Probability 11*, 401–408.

Chernick, M. R. (1981). A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. *Ann. Prob. 9*, 145–149.

Doob, J. L. (1953). Stochastic processes. *John Wiley, New York.*

Gorodetskii, V. V. (1977). On the strong mixing property for linear sequences. *Theory Probab. Appl. 22.*

Ibragimov, I. A. and Y. V. Linnik (1971). Independent and stationary sequences of random variables. *Wolters-Nordhoff,Groningen.*

Lai, T. L. and C. Z. Wei (1982). Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *The Annals of Statistics 10*, 154–166.

Pham, D. T. and L. T. Tran (1985). Some mixing properties of time series models. *Stochastic Processes and their Applications 19*, 297–303.

Rosenblatt, M. (1971). *Markov processes,Structure and Asymptotic Behavior.* Springer-Verlag Inc.

Rozanov, Y. A. (1967). Stationary Random Processes. *Holden-Day, San Francisco*.

White, H. and I. Domowitz (1984). Nonlinear regression with dependent observations. *Econometrica 52*, 143–162.

Withers, C. S. (1978). Central limit theorems for dependent variables. *Technical Report, Applied Mathematics Division, D. S. I. R., Wellington 76.*

Withers, C. S. (1981). Conditions for linear processes to be strong-mixing. *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete [Became: @J(ProbTher)] 57*, 477–480.