Statistica Sinica 2(1992), 479-494

ON STATIONARITY AND ASYMPTOTIC INFERENCE OF BILINEAR TIME SERIES MODELS

Jian Liu

University of British Columbia

Abstract: One of the commonly used techniques in establishing ergodicity of a Markov chain has been developed in a series papers by Tweedie (1974, 1975) and his associates. The present paper intends to demonstrate a useful alternative technique originated by Benes (1967) in the context of continuous time Markov chains. This technique can be adapted to the case when a time series model observed at discrete time points is under consideration. One of the advantages of such a technique is that it enables us to drop off the crucial assumption of ϕ-irreducibility as required by Tweedie's technique. Examples showing how to obtain stationarity conditions for bilinear models are given under finite and infinite variance assumptions on the noise sequence. Existence of moments is examined and finally, a central limit theorem and a law of the iterated logarithm concerning sample moments of some bilinear time series models are established.

Key words and phrases: Bilinear model, stationarity, central limit theorem, law of the iterated logarithm, moments.

1. Introduction

Following Granger and Andersen (1978) and Subba Rao and Gabr (1984), a time series $\{X_t\}$ is said to be a bilinear time series, denoted by BL(p,q,m,l), if it satisfies the following equation,

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{j=0}^{q} \theta_j Z_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{l} b_{ij} X_{t-i} Z_{t-j}, \quad t \in \mathrm{Z}, \tag{1.1}$$

where $\{Z_t\}$ is an independent identically distributed (iid) sequence of random variables, $\mathrm{Z} = \{0, \pm 1, \dots\}$ and $\theta_0 = 1$. Clearly, a special case of the bilinear model is the well-known ARMA(p,q) model. As shown by many authors such as Subba Rao and Gabr (1984), the bilinear model is particularly attractive in modelling processes with sample paths of occasional sharp spikes. Such phenomena are often found in seismology, econometrics and control theory.

A bilinear time series defined by (1.1) is said to be *causal* if there exists a measurable function $g : \mathbf{R}^\infty \to \mathbf{R}$ such that

$$X_t = g(Z_t, Z_{t-1}, \ldots), \quad \forall t \in \mathbf{Z}.$$

The structure of the paper is as follows: Section 2 displays the usefulness of Benes' (Benes (1967)) technique in deriving stationarity conditions for bilinear time series models under both finite and infinite variance assumptions on the noise sequence; (It should be mentioned here that such a technique is also applicable to the so called threshold ARMA models proposed by Tong (1983).) Section 3 focuses on a recursive algorithm useful in deriving governing conditions for the existence of higher order moments of some bilinear models; the last section gives proofs of both the central limit theorem (CLT) and the law of the iterated logarithm (LIL) for the sample mean and sample covariances of a class of bilinear models.

2. Strict Stationarity

In this section, it is intended to demonstrate a useful alternative technique in establishing strict stationarity of a general Markov chain. This technique was first introduced by Benes (1967) for continuous time Markov chains and was adapted to the discrete time case by Liu and Susko (1992). For convenience, the technique is summarized in the following theorem.

First, let $\{X_t, t = 0, 1, \ldots\}$ be a discrete time Markov chain defined on a locally compact complete separable metric space $\mathcal{X}$ with homogeneous transition probabilities

$$P^n(x, A) = P(X_n \in A | X_0 = x), \quad x \in \mathcal{X}, A \in \mathcal{B},$$

where $\mathcal{B}$ is the Borel σ-field on $\mathcal{X}$. Assume further that $P^n(\cdot, A)$ is $\mathcal{B}$-measurable and for fixed $x \in \mathcal{X}$, $P^n(x, \cdot)$ is a probability measure on the σ-field $\mathcal{B}$.

Assumption. Assume that the open and closed sets of the metric space $(\mathcal{X}, \rho)$ are all $\mathcal{B}$-measurable and for every compact set K in $\mathcal{X}$,

$$\lim_{A \downarrow \phi} \sup_{x \in K} \{P^1(x, A)\} = 0. \tag{2.1}$$

Theorem 2.1. *Suppose that assumption* (2.1) *holds. Then there exists a strictly stationary Markov process* $\{X_t\}$ *with the above transition probabilities if and only if there exist a nonnegative measurable function* $g(\cdot)$ *satisfying*

$$\inf_{x \in K_n^c} g(x) \to \infty, \quad \textit{as some compact sets } K_n \uparrow \mathcal{X} \tag{2.2}$$

and an initial probability measure P_0 for X_0 such that

$$\sup_{t\geq 0}\left[\int_{\mathcal{X}}\int_{\mathcal{X}} g(y)P^t(x,dy)P_0(dx)\right] < \infty. \tag{2.3}$$

Remark. The advantage of this technique is that it does not require the assumption of ϕ-irreducibility as required by Tweedie (1974, 1975). It should also be mentioned here that Tweedie (1988) has extended his previous results to deal with second order stationarity by dropping off the ϕ-irreducibility requirement.

For a proof of this theorem, see Benes (1967) and Liu and Susko (1992).

Now let us apply the above theorem to the bilinear models. To do so, we need to write the model under consideration into the framework of a Markov chain or some sort of state space form. Without loss of generality, we assume here that $p = m$ and $q = l$. By introducing the state vector

$$\tilde{Y}_t = (X_t, \dots, X_{t-p+1}, Z_t, \dots, Z_{t-q+1})',$$

we can express (1.1) in the following equivalent form,

$$\tilde{Y}_t = \tilde{g}(\tilde{Y}_{t-1}) + (1,0,\dots,0,1,0,\dots,0)'Z_t, \tag{2.4}$$

for some appropriate measurable function $\tilde{g}(\cdot)$ mapping from $\mathbf{R}^{p+q}$ to itself. Hence $\{\tilde{Y}_t\}$ is evidently a Markov chain. Thus, to verify Theorem 2.1, we only need to examine the scalar case with some suitable scalar function $g(\cdot)$.

It should be noted here that though the handy Markovian representation (2.4) can be used directly in the application of Theorem 2.1, it is generally not suitable for Tweedie's (1974, 1975) results since ϕ-irreducibility is generally difficult to verify with such a representation due to possibly over-dimension of the state vector. Recently, in a series papers of Pham (1985, 1986), he derives the so called bilinear Markovian representation using the predictor space as a basis. He shows that the general bilinear model can be represented as

$$\begin{cases} X_t = H\tilde{Y}(t-1) + Z_t, \\ \tilde{Y}(t) = A(t)\tilde{Y}(t-1) + \tilde{\xi}(t), \end{cases} \tag{2.5}$$

where $A(t)$ and $\tilde{\xi}(t)$ are finite order polynomials in Z_t and H is a constant vector. Let $\|\cdot\|_v$ be a vector norm and define, for each $\gamma > 0$, the associated L^γ-norm (It is not a norm in the usual sense if $\gamma < 1$.),

$$\|A(t)\|_{v,\gamma} = \sup\left\{\frac{\{E[\|A(t)\tilde{z}\|_v]^\gamma\}^{1/\gamma}}{\|\tilde{z}\|_v}\right\}, \quad \tilde{z} \in \mathbf{R}^k,$$

provided that the right-hand-side is finite, where k is the dimension of the square matrix $A(t)$. Under the assumption of ϕ-irreducibility of the second equation of (2.5), which is clearly a Markov chain, Pham (1986) shows that

$$\|A(t)\|_{v,\gamma} < 1 \tag{2.6}$$

implies geometric ergodicity. If $\gamma \geq 1$, by iterating the second equation of (2.5) and using Minkowski's triangular inequality, one can easily see that the norm condition (2.6) is sufficient for the existence of a strictly stationary and ergodic $\{\tilde{Y}(t)\}$ and hence $\{X_t\}$.

As shown in Pham (1985, 1986), writing a general bilinear model into a bilinear Markovian form is not an easy task and the matrices involved in the representation (2.5) are in general quite complicated. In Example 2.1, we show how to derive stationarity condition using the simpler state space form (2.4) and in Example 2.2, we apply Theorem 2.1 to Pham's (Pham (1986)) bilinear Markovian representation (2.5) for a general $0 < \gamma \leq 1$.

Example 2.1. Consider the general subdiagonal BL(p,q,m,l) defined by

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + Z_t + \sum_{i=1}^{m}\sum_{j=1}^{l} b_{ij} X_{t-i} Z_{t-j}, \quad t \in \mathbf{Z}, \tag{2.7}$$

where $\{Z_t\}$ are iid with possibly infinite variance, $\theta_0 = 1$ and $b_{ij} = 0$ for $i < j$.

Case 1. Assume that $\{Z_t\}$ has finite fourth moment and $E(Z_t) = E(Z_t^3) = 0$. Write (2.7) in the following equivalent vector form,

$$\tilde{X}_t = CZ_t + A\tilde{X}_{t-1} + \sum_{j=1}^{l} B_j \tilde{X}_{t-j} Z_{t-j}, \tag{2.8}$$

where $\tilde{X}_t = (X_t, X_{t-1}, \dots, X_{t-p+1})'$ with p redefined as $p = \max(p,m)$, $C = (1,0,\dots,0)'$ is a $p \times 1$ vector, A and B_j are $p \times p$ matrices defined by

$$A = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \quad \text{and} \quad B_j = \begin{pmatrix} b_{jj} & \cdots & b_{mj} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

It follows from the above theorem that to derive a stationarity condition, one only needs to select a function $g(\cdot)$ so that (2.2) and (2.3) are satisfied. One of these choices is $g(\tilde{x}) = \|\tilde{x}\|^2$, the Euclidean norm of a $p \times 1$ vector. Set

$$\tilde{X}_n^* = \begin{cases} \tilde{0}, & \text{if } n < 0, \\ CZ_n + A\tilde{X}_{n-1}^* + \sum_{j=1}^{l} B_j \tilde{X}_{n-j}^* Z_{n-j}, & \text{if } n \geq 0, \end{cases}$$

for any $t \in \mathbf{Z}$,

$$\tilde{S}_n(t) = \begin{cases} \tilde{0}, & \text{if } n < 0, \\ CZ_t + A\tilde{S}_{n-1}(t-1) + \sum_{j=1}^{l} B_j \tilde{S}_{n-j}(t-j) Z_{t-j}, & \text{if } n \geq 0 \end{cases} \tag{2.9}$$

and

$$\tilde{\Delta}_n(t) = \tilde{S}_n(t) - \tilde{S}_{n-1}(t).$$

Clearly, for each fixed $n \geq 0$, there exists a measurable function $g_n(\cdot)$ such that $\tilde{S}_n(t) = g_n(Z_t, Z_{t-1}, \ldots)$. Noting that $\tilde{X}_n^* = \tilde{S}_n(n)$ for $n \geq 0$ and for each fixed n and arbitrary $t, t' \in \mathbf{Z}$, $\tilde{S}_n(t)$ and $\tilde{S}_n(t')$ have identical distributions, so $\tilde{X}_n^*$ and $\tilde{S}_n(0)$ have the same distribution. Though it is possible to derive a stationarity condition by using an argument similar to that of Liu (1989b) to show that

$$\sup_n E[\tilde{S}_n(0)'\tilde{S}_n(0)] < \infty,$$

we shall employ the results already established in Liu (1989b), namely,

$$E[\tilde{\Delta}_n(0)'\tilde{\Delta}_n(0)] \leq \text{const.}\lambda^{n/2}, \quad \forall n > 1,$$

where $\lambda = \rho(\Gamma)$, the spectral radius or the maximum eigenvalue in absolute value of Γ, is assumed to be less than 1 and Γ is defined by

$$\Gamma = \begin{pmatrix} \Gamma_1 & \Gamma_2 & \cdots & \Gamma_{l-1} & \Gamma_l \\ I_{p^2} & 0 & \cdots & 0 & 0 \\ 0 & I_{p^2} & \cdots & 0 & 0 \\ \cdots & & & & \cdots \\ 0 & 0 & \cdots & I_{p^2} & 0 \end{pmatrix},$$
$$\Gamma_1 = A \otimes A + \sigma^2(B_1 \otimes B_1),$$
$$\Gamma_j = \sigma^2\Big[B_j \otimes \Big(\sum_{i=1}^{j-1} A^i B_{j-i}\Big) + \Big(\sum_{i=1}^{j-1} A^i B_{j-i}\Big) \otimes B_j + B_j \otimes B_j\Big], \quad j = 2, \ldots, l \tag{2.10}$$

where I_k denotes the $k \times k$ identity matrix. Thus, if $\lambda < 1$, $E[\tilde{\Delta}_n(0)'\tilde{\Delta}_n(0)]$ decays to zero geometrically as n tends to infinity. Hence, it follows from the

Cauchy-Schwarz inequality that

$$\begin{aligned} E[g(\tilde{X}_t^*)] &= E[\tilde{S}_n(0)'\tilde{S}_n(0)] \\ &\le \text{const.} + \text{const.} E\Big[\Big\|\sum_{j=1}^{n}\tilde{\Delta}_j(0)\Big\|^2\Big] \\ &\le \text{const.} + \text{const.}\Big[\sum_{j=1}^{n}\sqrt{E(\|\tilde{\Delta}_j(0)\|^2)}\Big]^2 \\ &\le \text{const.} + \text{const.}\Big(\sum_{j=1}^{n}\lambda^{j/4}\Big)^2 \\ &< \infty. \end{aligned}$$

This shows the boundedness of (2.3). Hence, there exists a strictly stationary time series $\{X_t\}$ satisfying (2.7) and (2.8), provided that $\lambda < 1$, $E(Z_t^4) < \infty$ and $E(Z_t) = E(Z_t^3) = 0$.

Case 2. Suppose that $\{Z_t\}$ are iid with finite 2γth moments for some $\gamma \in (0,1]$. Define $g(x) = |x|^\gamma$ for any $x \in \mathbf{R}$ and set $X_t^* = Z_t + U_t$, where X_t^* is the first component of $\tilde{X}_t^*$. Then for any initial random variables $X_0^*, \ldots, X_{1-p}^*$ which are independent of $\{X_s^*, s \ge p\}$ (here p is assumed to be $\max(p,m)$ for notational convenience),

$$E[g(X_t^*)] \le \text{constant} + \sum_{i=1}^{p} E\Big\{\Big|\,\phi_i + \sum_{j\le i} b_{ij}Z_{t-j}\Big|^\gamma\Big\} Eg(U_{t-i}).$$

This implies that

$$E[g(U_t)] \le \text{constant} + \sum_{i=1}^{p} E\Big\{\Big|\phi_i + \sum_{j\le i} b_{ij}Z_{t-j}\Big|^\gamma\Big\} Eg(U_{t-i}).$$

So the sequence $\{E[g(U_t)]\}$ is bounded provided that

$$\sum_{i=1}^{p} E\Big\{\Big|\,\phi_i + \sum_{j\le i} b_{ij}Z_{t-j}\Big|^\gamma\Big\} < 1. \tag{2.11}$$

This in turn will be sufficient for the boundedness of $\{E[g(X_t^*)]\}$. Hence, there exists a strictly stationary bilinear time series $\{X_t\}$ satisfying (2.4).

Example 2.2. General Bilinear Models

Again, we consider the general bilinear model (1.1). As shown by Pham (1986), (1.1) has an equivalent bilinear Markovian representation (2.5). Define

$$\tilde{Y}^*(t) = (\tilde{Y}(t)', Z_t)', \quad \tilde{\xi}^*(t) = (\tilde{\xi}(t)', Z_t)',$$

$$A^*(t) = \begin{pmatrix} A(t) & 0 \\ 0 & 0 \end{pmatrix}, \quad H^* = (H', 1)'.$$

Then (2.5) can be rewritten as

$$\begin{cases} X_t = H^* \tilde{Y}^*(t-1), \\ \tilde{Y}^*(t) = A^*(t)\tilde{Y}^*(t-1) + \tilde{\xi}^*(t). \end{cases} \tag{2.12}$$

Let $\|\cdot\|_v$ be some vector norm and $\|\cdot\|_m$ be the induced matrix norm. Define $g(\cdot) = \|\cdot\|_v$. Assume that $E\{[\|\tilde{\xi}^*(t)\|_v]^\gamma\} < \infty$ for some $\gamma \in (0,1]$. Then a direct application of Theorem 2.1 yields strict stationarity of $\{\tilde{Y}^*(t)\}$ and hence that of $\{X_t\}$, provided that

$$E\{[\|A(t)\|_m]^\gamma\} = E\{[\|A^*(t)\|_m]^\gamma\} < 1.$$

3. Moments

For the general bilinear model, it has been shown by Liu and Brockwell (1988) that the second order moment of the observable time series $\{X_t\}$ is finite if the iid innovation sequence $\{Z_t\}$ has finite $2l$th moment, provided that the stated stationarity condition is satisfied. This requirement can be substantially simplified when the bilinear model reduces to the so called subdiagonal model (i.e. $b_{ij} = 0$ for all $i < j$). In this case, the existence of finite fourth moment of the innovation sequence $\{Z_t\}$ is sufficient for the existence of finite second order moment of the bilinear time series $\{X_t\}$ again under the stationarity assumption $\rho(\Gamma) < 1$, where Γ is defined in (2.10). A proof of this result in more general context of a subdiagonal multiple bilinear model is given in Liu (1989b). For the detailed causality conditions, see Liu and Brockwell (1988) and Liu (1989b).

As shown by Liu (1989a, 1990), the existence of a causal and ergodic solution does not necessarily require the existence of higher order moments of the innovation sequence. In particular, as demonstrated above, for any $\gamma \in (0,1]$, the existence of finite 2γth moment of the iid innovation sequence will be sufficient for the existence of finite γth moment of the subdiagonal bilinear time series, provided that the stated stationarity condition in Example 2.1 is met.

Another approach to study the existence of higher order moments is given by Pham (1985, 1986) using the bilinear Markovian representation (2.5). Though his approach often leads to a more general (implicit) sufficient condition for the existence of higher order moments, in view of difficulties involved in obtaining (2.5) for a general bilinear model, we shall not proceed further along this direction. Instead, our attention will be focused on deriving some simpler sufficient conditions under which higher order moments exist.

As is well known, the existence of higher order moments of a bilinear time series normally requires additional restrictions on model parameters as well as possibly even higher order moments of the innovations. As an example, we shall examine the fourth moment of the vector-wise subdiagonal model,

$$\tilde{X}_t = CZ_t + (A + BZ_{t-1})\tilde{X}_{t-1}, \tag{3.1}$$

where A and B are arbitrary $p \times p$ matrices, C is a $p \times 1$ constant vector and $\{Z_t\}$ are iid. Assume that

$$E(Z_t) = E(Z_t^3) = 0, \quad \sigma^2 = E(Z_t^2), \quad \text{and} \quad \gamma^4 = E(Z_t^4) < \infty. \tag{3.2}$$

Also assume that the causality (or ergodicity) condition

$$\lambda = \rho[(A \otimes A) + \sigma^2(B \otimes B)] < 1 \tag{3.3}$$

is satisfied. To establish existence of the fourth order moment, we introduce $\tilde{S}_n(t)$ and $\tilde{\Delta}_n(t)$ as in (2.9) and the subsequent equation. Clearly, $\tilde{\Delta}_n(t)$ is measurable with respect to the σ-field $\sigma(Z_s, s < t)$ and satisfies the equations,

$$\tilde{\Delta}_n(t) = (A + BZ_{t-1})\tilde{\Delta}_{n-1}(t-1).$$

By the L^p-theory, $p > 1$, the problem of existence of the fourth order moment reduces to the convergence of $\{\tilde{S}_n(t), n \geq 0\}$ in L^4. The quantity of interest in determining the L^4-convergence is

$$V_n = E[(\tilde{\Delta}_n(t)\tilde{\Delta}_n(t)') \otimes (\tilde{\Delta}_n(t)\tilde{\Delta}_n(t)')].$$

It is not difficult to show that

$$\begin{aligned}
V_n = {} & (A \otimes A)V_{n-1}(A' \otimes A') \\
& + \sigma^2[(A \otimes B)V_{n-1}(A' \otimes B') + (B \otimes A)V_{n-1}(B' \otimes A')] \\
& + \sigma^2[(A \otimes A)V_{n-1}(B' \otimes B')] \\
& + \sigma^2[(A \otimes B)V_{n-1}(B' \otimes A') + (B \otimes A)V_{n-1}(A' \otimes B')] \\
& + \sigma^2(B \otimes B)V_{n-1}(A' \otimes A') + \gamma^4(B \otimes B)V_{n-1}(B' \otimes B'),
\end{aligned}$$

which can be written in the vector form,

$$\begin{aligned}\vec{V}_n = \{&[(A\otimes A)\otimes(A\otimes A)] + \sigma^2[(A\otimes B)\otimes(A\otimes B)]\\ &+\sigma^2[(B\otimes A)\otimes(B\otimes A)] + \sigma^2[(B\otimes B)\otimes(A\otimes A)]\\ &+\sigma^2[(B\otimes A)\otimes(A\otimes B)] + \sigma^2[(A\otimes B)\otimes(B\otimes A)]\\ &+\sigma^2[(A\otimes A)\otimes(B\otimes B)] + \gamma^4[(B\otimes B)\otimes(B\otimes B)]\}\vec{V}_{n-1}\\ = \Lambda&\vec{V}_{n-1},\end{aligned}$$

where, for a matrix M, $\vec{M}$ denotes the vector obtained by stacking the columns of M one on top of the another. Thus,

$$\rho(\Lambda) < 1 \tag{3.4}$$

is sufficient for $\{\vec{V}_n\}$ converging to zero geometrically. It then follows from the following relation,

$$E\Big[\Big\|\sum_n \tilde{\Delta}_n(t)\Big\|^4\Big] \le \Big[\sum_n (E\|\tilde{\Delta}_n(t)\|^4)^{1/4}\Big]^4,$$

(3.3) and (3.4) that $\{\tilde{S}_n(t)\}$ is a Cauchy sequence in L^4 and hence its limit $\tilde{X}_t$ is also in L^4, i.e., $\tilde{X}_t$ has finite fourth order moment.

Remark. The condition (3.4) is usually different from (3.3), since at least (3.3) does not contain γ^4, the fourth order moment of $\{Z_t\}$. But condition (3.4) implies (3.3).

To this end, observe that in the definition of $\tilde{S}_n(t)$, the initial vector $\tilde{S}_0(t)$ could be chosen as any arbitrary random vector $\tilde{S}_0^*$ which is independent of $\{Z_s, s\in \mathbf{Z}\}$ and has finite fourth order moment. By proceeding exactly the same as above and letting $C = \tilde{0}$, one can derive the same relation for V_n. This shows the L^4-convergence of $\tilde{S}_n(t)$ for any initial $\tilde{S}_0(t)$ independent of $\{Z_s, s\in\mathbf{Z}\}$ and having finite fourth order moment provided $\rho(\Lambda) < 1$. This in turn shows that $\tilde{S}_n(t)$ is also L^2-convergent for any initial $\tilde{S}_0(t)$ independent of $\{Z_s, s\in\mathbf{Z}\}$ and having finite fourth order moment. Also, for any two initial vectors $\tilde{S}_{0,1}^* = \tilde{d}_1$ and $\tilde{S}_{0,2}^* = \tilde{d}_2$, denote the corresponding $\{\tilde{S}_n(t)\}$ sequences by $\{\tilde{S}_{n,1}(t)\}$ and $\{\tilde{S}_{n,2}(t)\}$. Then it is easy to show that $W_n = E[\tilde{S}_{n,1}(t)\tilde{S}_{n,2}(t)']$ satisfies the equations,

$$\vec{W}_n = (A\otimes A + \sigma^2 B\otimes B)\vec{W}_{n-1} = \Gamma\vec{W}_{n-1} = \Gamma^{n-1}\vec{W}_1 = \Gamma^n\vec{W}_0, \tag{3.5}$$

where $W_0 = \tilde{d}_1\tilde{d}_2'$. Since (3.5) holds for any $\tilde{d}_1, \tilde{d}_2 \in \mathbf{R}^p$ and hence for arbitrary W_0 and

$$\lim_{n\to\infty} W_n = 0,$$

we conclude by Jordan decomposing Γ that $\rho(\Gamma)$ has to be less than unity.

4. Limit Theorems for Subdiagonal Bilinear Time Series

We shall consider the simple subdiagonal bilinear model,

$$\tilde{X}_t = CZ_t + A\tilde{X}_{t-1} + B\tilde{X}_{t-1}Z_{t-1}, \tag{4.1}$$

where $\{Z_t\}$ are iid and A, B and C are arbitrary nonrandom matrices. Obviously, (4.1) is a special case of (2.7) with $l = 1$. In this section, it will be shown that under appropriate conditions, the sample mean and the sample covariances satisfy the central limit theorem (CLT) and the law of the iterated logarithm (LIL).

Let the iid innovation sequence $\{Z_t\}$ be defined on the probability space $(\Omega, \mathcal{F}, P)$ and set $\mathcal{F}_t = \sigma\{Z_s, s \le t\}$, the σ-field generated by $\{Z_s, s \le t\}$. Define T to be the shift operator in $\mathbf{R}^\infty$, i.e. for $\tilde{x} = (\dots, x_{-1}, x_0, x_1, \dots) \in \mathbf{R}^\infty$, $T(\tilde{x}) = (\dots, x_0, x_1, x_2, \dots)$.

Throughout this section, it is assumed that

$$E(Z_t^4) < \infty, \quad E(Z_t^m) = 0, \quad \text{for } m = 1 \text{ and } 3, \tag{4.2}$$

and

$$\lambda = \rho(A \otimes A + \sigma^2 B \otimes B) < 1. \tag{4.3}$$

Then there exists a measurable function $g(\cdot)$ mapping from $\mathbf{R}^\infty$ to $\mathbf{R}$ such that

$$X_t = g(Z_t, Z_{t-1}, \dots), \quad \text{a.s.}, \quad \forall t \in \mathbf{Z},$$

i.e. $\{X_t\}$ is *causal*. Furthermore, $\{\tilde{X}_t\}$ can be written in the form of an almost surely convergent as well as L^2-convergent infinite series,

$$\tilde{X}_t = CZ_t + \sum_{n=1}^{\infty}\prod_{j=0}^{n-1}(A + BZ_{t-j})CZ_{t-n}. \tag{4.4}$$

For a proof of this, see either Bhaskara Rao et al. (1983) or Liu and Brockwell (1988).

Lemma 4.1. *For any real $p \times p$ matrices A and B, if*

$$\rho(A \otimes A + B \otimes B) < 1,$$

then $\rho(A) < 1$.

Proof. Consider the subdiagonal bilinear model (4.1) with $\{Z_t\}$ iid $N(0,1)$ and $C = 0$. Then the condition $\rho(A \otimes A + B \otimes B) < 1$ implies that (4.1) is ergodic and causal. Furthermore, if $\{\tilde{S}_n(t), n \geq 0\}$ and $\tilde{\Delta}_n(t)$ are defined as in Section 3 except that the initial vector $\tilde{S}_0(t)$ is now redefined as $Z_{t-1}^2 \tilde{d}$ for arbitrary $\tilde{d} \in \mathbf{R}^p$, then it is easy to show that for $n > 1$,

$$D_n = E[\tilde{\Delta}_n(t)\tilde{\Delta}_n(t)'] = AD_{n-1}A' + BD_{n-1}B',$$

or equivalently,

$$\overrightarrow{D_n} = (A \otimes A + B \otimes B)\overrightarrow{D_{n-1}}.$$

Hence $\tilde{\Delta}_n(t)$ converges to $\tilde{0}$ geometrically in L^2 as well as in L^1. So $\{\tilde{S}_n(t), n \geq 0\}$ converges to $\tilde{X}_t = \tilde{0}$ both in L^2 and in L^1 since $\tilde{X}_t = \tilde{0}$ is the unique stationary solution of (4.1) (Bhaskara Rao et al. (1983) and Liu (1989b)). Now, observe that

$$\tilde{\eta}_n = E[\tilde{S}_n(t)] = A\tilde{\eta}_{n-1} = \cdots = A^n \tilde{d}.$$

This together with the fact that $\tilde{X}_t = \tilde{0}$ is the unique stationary solution of (4.1) and hence $\lim_{n\to\infty} \tilde{\eta}_n = \tilde{0}$ for all $\tilde{d} \in \mathbf{R}^p$ implies that

$$\lim_{n\to\infty} A^n \tilde{d} = \tilde{0}, \quad \forall \tilde{d} \in \mathbf{R}^p.$$

Finally, we conclude by Jordan decomposing A and using the above relation that $\rho(A) < 1$.

For the univariate bilinear time series generated by the first components of $\{\tilde{X}_t\}$ which satisfies (4.1), set $Y_t = g(Z_t, Z_{t-1}, \ldots) - \mu$, here $\mu = E(X_0)$ and $g(\cdot)$ is the measurable function determined by (4.4) so that $X_t = g(Z_t, Z_{t-1}, \ldots)$. Then $E(Y_0) = 0$, $E(Y_0^2) < \infty$ and Y_0 is $\mathcal{F}_0$-measurable. Also, $Y_t = Y_0(T^t(\tilde{Z}))$, where $\tilde{Z} = (\ldots, Z_{-1}, Z_0, Z_1, \ldots)$ is strictly stationary and ergodic. The first partial sum of interest is

$$S_n = \sum_{t=1}^{n} Y_t = \sum_{t=1}^{n} X_t - n\mu.$$

Theorem 4.1. *Consider the bilinear model* (4.1) *with* $\{X_t\}$ *defined to be the first component of* $\{\tilde{X}_t\}$. *Under the causality assumptions* (4.2) *and* (4.3), *the following limit laws hold:*

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0, \sigma^{*2}), \quad \text{as } n \to \infty, \tag{4.5}$$

where $\sigma^* = \sqrt{\lim_{n\to\infty} ES_n^2/n} < \infty$; *if, in addition,* $\sigma^* > 0$, *then*

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = \sigma^*, \quad \text{a.s.,} \tag{4.6}$$

and

$$\liminf_{n\to\infty} \frac{S_n}{\sqrt{2n\ln\ln n}} = -\sigma^*, \quad \text{a.s..} \tag{4.7}$$

Proof. After some algebraic computation, we have, for each $t > 0$,

$$E(\tilde{X}_t|\mathcal{F}_0) = \sigma^2 \sum_{j=0}^{t-2} A^j BC + A^{t-1}(A + BZ_0)\tilde{X}_0. \tag{4.8}$$

To establish the above theorem, we only need to verify the conditions of Theorem 5.5 of Hall and Heyde (1980). Set

$$y_t = E(Y_t|\mathcal{F}_0) - E(Y_t|\mathcal{F}_{-1}).$$

Then $y_t = 0$ a.s. for $t < 0$ and for $t \geq 0$,

$$y_t = \tilde{h}' A^{t-1}[\sigma^2 BC + (A + BZ_0)\tilde{X}_0 + A(A + BZ_{-1})\tilde{X}_{-1}],$$

where $\tilde{h} = (1, 0, \dots, 0)' \in \mathbf{R}^p$. Noting that $(\rho(A))^2 = \rho(A \otimes A) < 1$, we have, by Proposition 2.1 of Liu and Brockwell (1988), that

$$\begin{aligned} Ey_t^2 &= \tilde{h}' A^{t-1} E\{[\sigma^2 BC + (A + BZ_0)\tilde{X}_0 + A(A + BZ_{-1})\tilde{X}_{-1}] \\ &\quad \cdot [\sigma^2 BC + (A + BZ_0)\tilde{X}_0 + A(A + BZ_{-1})\tilde{X}_{-1}]'\} A^{t-1'}\tilde{h} \\ &\leq \text{const.}\lambda^t \end{aligned}$$

for $t > 0$. It then follows from above and the Cauchy-Schwarz inequality that

$$E\Big[\sum_{j=m}^{n} y_j\Big]^2 \leq \text{const.}\lambda^m, \quad \text{for some } 0 < \lambda < 1 \text{ and } \forall\, 0 < m < n.$$

This establishes (5.24) of Hall and Heyde (1980), namely,

$$\sum_{m=1}^{\infty} \Big\{ \limsup_n E\Big[\sum_{j=m}^{n} y_j\Big]^2 + \limsup_n E\Big[\sum_{j=m}^{n} y_{-j}\Big]^2 \Big\} < \infty.$$

Finally, we need to show that

$$E(Y_0|\mathcal{F}_\infty) = Y_0, \quad \text{a.s., and} \quad E(Y_0|\mathcal{F}_{-\infty}) = 0, \quad \text{a.s.,}$$

where $\mathcal{F}_\infty = \sigma\{\cup_t \mathcal{F}_t\}$ and $\mathcal{F}_{-\infty} = \cap_t \mathcal{F}_t$. The first relation is obvious since Y_0 is $\mathcal{F}_0$-measurable. To show the second relation, observe that for $\tilde{Y}_t = \tilde{X}_t - \tilde{\mu}$,

$$\begin{aligned} E(\tilde{Y}_0|\mathcal{F}_{-m}) &= E[\tilde{X}_0 - \tilde{\mu}|\mathcal{F}_{-m}] \\ &= \sigma^2 \sum_{j=0}^{m-2} A^j BC + A^{m-1}(A + BZ_{-m})\tilde{X}_{-m} - \tilde{\mu}, \quad \forall m > 0, \end{aligned}$$

and

$$\tilde{\mu} = \sigma^2 \sum_{j=0}^{\infty} A^j BC.$$

These, together with the Cauchy-Schwarz inequality, imply that

$$\begin{aligned} E\{E[Y_0|\mathcal{F}_{-\infty}]\}^2 &= E\{E[E(Y_0|\mathcal{F}_{-m})|\mathcal{F}_{-\infty}]\}^2 \\ &\le E\{E[[E(Y_0|\mathcal{F}_{-m})]^2|\mathcal{F}_{-\infty}]\} \\ &\to 0 \end{aligned}$$

as $m \to \infty$. This completes the proof.

Similarly, we may consider the conventional moment estimates. For example, one of the frequently used estimates for $\gamma(k)$, the covariance at lag k, is

$$\hat{\gamma}(k) = \frac{1}{n} \sum_{j=k+1}^{n} (X_j - \bar{X})(X_{j-k} - \bar{X}).$$

This motivates the consideration of the following partial sums,

$$S_n(k) = \sum_{j=k+1}^{n} X_j X_{j-k} - (n-k)\mu(k), \quad k \ge 0,$$

where $\mu(k) = E[X_k X_0]$. In addition to assumptions (4.2) and (4.3), if condition (3.4) is satisfied, similar limit theorems hold.

Theorem 4.2. *Let $\{X_t\}$ be the first component series of $\{\tilde{X}_t\}$ defined by* (4.1). *Assume that* (4.2) *and* (3.4) *are satisfied. Then the following limit laws hold:*

$$\frac{S_n(k)}{\sqrt{n}} \xrightarrow{d} N(0, \sigma(k)^2), \quad \text{as} \quad n \to \infty, \tag{4.9}$$

where $\sigma(k) = \sqrt{\lim_{n\to\infty} ES_n(k)^2/n} < \infty$; if, in addition, $\sigma(k) > 0$, then

$$\limsup_{n\to\infty} \frac{S_n(k)}{\sqrt{2n \ln\ln n}} = \sigma(k), \quad \text{a.s.}, \tag{4.10}$$

and

$$\liminf_{n\to\infty} \frac{S_n(k)}{\sqrt{2n \ln\ln n}} = -\sigma(k), \quad \text{a.s.} \tag{4.11}$$

Proof. Similar to the proof of Theorem 4.1, it suffices to show that

$$\left|E[\tilde{X}_{t+k}\tilde{X}_t'|\mathcal{F}_0] - E[\tilde{X}_{t+k}\tilde{X}_t']\right| \le \delta^{t-1}C(\mathcal{F}_0), \quad \forall t > 1, \tag{4.12}$$

where the matrix $C(\mathcal{F}_0)$ is $\mathcal{F}_0$-measurable and in L^2, the nonrandom constant $0 < \delta < 1$ and the matrix relation $\le$ is defined to be elementwise.

Without loss of generality, we consider only $k = 1$. Since

$$\begin{aligned}
E[Z_t\tilde{X}_t\tilde{X}_t'|\mathcal{F}_0] &= \sigma^2 C E[(A\tilde{X}_{t-1} + B\tilde{X}_{t-1}Z_{t-1})'|\mathcal{F}_0] \\
&\quad + \sigma^2 E[(A\tilde{X}_{t-1} + B\tilde{X}_{t-1}Z_{t-1})|\mathcal{F}_0]C \\
&= \sigma^2 C\Big\{\Big[\sigma^2 \sum_{j=0}^{t-3} A^{j+1}BC + A^{t-1}(A + BZ_0)\tilde{X}_0\Big]' + \sigma^2 C'B'\Big\} \\
&\quad + \sigma^2\Big\{\Big[\sigma^2 \sum_{j=0}^{t-3} A^{j+1}BC + A^{t-1}(A + BZ_0)\tilde{X}_0\Big] + \sigma^2 BC\Big\}C' \\
&= \text{const.}(t) + (A^{t-1}U_1) + (A^{t-1}U_1)',
\end{aligned} \tag{4.13}$$

we have

$$\begin{aligned}
E[\tilde{X}_{t+1}\tilde{X}_t'|\mathcal{F}_0] &= AE[\tilde{X}_t\tilde{X}_t'|\mathcal{F}_0] + BE[Z_t\tilde{X}_t\tilde{X}_t'|\mathcal{F}_0] \\
&= AE[\tilde{X}_t\tilde{X}_t'|\mathcal{F}_0] + \text{const.}(t) + (A^{t-1}U_1) + (A^{t-1}U_1)',
\end{aligned}$$

here U_1 is $\mathcal{F}_0$-measurable and in L^2, const.(t) is nonrandom bounded constant sequence, and $\rho(A) < 1$. It thus suffices to show that

$$\left|E[\tilde{X}_t\tilde{X}_t'|\mathcal{F}_0] - E[\tilde{X}_t\tilde{X}_t']\right| \le \delta^{t-1}C(\mathcal{F}_0), \quad \forall t > 1, \tag{4.14}$$

for some (other) nonrandom constant $0 < \delta < 1$ and some matrix $C(\mathcal{F}_0)$ measurable with respect to $\mathcal{F}_0$ and in L^2.

Set

$$M_1 = \begin{pmatrix} A \otimes A & B \otimes B \\ \sigma^2 A \otimes A & \sigma^2 B \otimes B \end{pmatrix}.$$

Then by Lemma 4.1, condition (3.4) implies the ergodicity condition

$$\rho(\Gamma) = \rho(A \otimes A + \sigma^2 B \otimes B) < 1,$$

which in turn implies that $\rho(M_1) < 1$, since the characteristic polynomial of M_1 is

$$|xI - M_1| = x^{p^2}|xI - \Gamma|.$$

Define

$$F_t = E[\tilde{X}_t\tilde{X}_t'|\mathcal{F}_0], \quad \text{and} \quad W_t = E[Z_t^2\tilde{X}_t\tilde{X}_t'|\mathcal{F}_0].$$

In view of (4.13), it is easy to see that, for $t > 3$,

$$\begin{aligned} F_t &= \sigma^2 CC' + AF_{t-1}A' + BW_{t-1}B' \\ &\quad + AE[Z_{t-1}\tilde{X}_{t-1}\tilde{X}_{t-1}'|\mathcal{F}_0]B' + BE[Z_{t-1}\tilde{X}_{t-1}\tilde{X}_{t-1}'|\mathcal{F}_0]A' \\ &= \text{const.}(2,t) + AF_{t-1}A' + BW_{t-1}B' + (A^{t-1}U_2) + (A^{t-1}U_2)' \end{aligned}$$

and

$$\begin{aligned} W_t &= \gamma^4 CC' + \sigma^2 E[(A\tilde{X}_{t-1} + B\tilde{X}_{t-1}Z_{t-1})(A\tilde{X}_{t-1} + B\tilde{X}_{t-1}Z_{t-1})'|\mathcal{F}_0] \\ &= \text{const.}(3,t) + \sigma^2[AF_{t-1}A' + BW_{t-1}B'] + (A^{t-1}U_3) + (A^{t-1}U_3)', \end{aligned}$$

and their vector forms are

$$\overrightarrow{F_t} = (A\otimes A)\overrightarrow{F}_{t-1} + (B\otimes B)\overrightarrow{W}_{t-1} + (I\otimes A^{t-1})\overrightarrow{U}_2 + (A^{t-1'}\otimes I)\overrightarrow{U}_2' + \text{const.}(2,t)\mathbf{1},$$
$$\overrightarrow{W_t} = \sigma^2(A\otimes A)\overrightarrow{F}_{t-1} + \sigma^2(B\otimes B)\overrightarrow{W}_{t-1} + (I\otimes A^{t-1})\overrightarrow{U}_3 + (A^{t-1'}\otimes I)\overrightarrow{U}_3' + \text{const.}(3,t)\mathbf{1},$$

where both U_2 and U_3 are $\mathcal{F}_0$-measurable and in L^2, const.$(2,t)$ and const.$(3,t)$ are constant sequences bounded above and $\mathbf{1}$ is a column vector with all elements being 1 and dimension compatible with F_t and W_t. Hence if $\overrightarrow{D}_t = [\overrightarrow{F}_t', \overrightarrow{W}_t']'$, then

$$\overrightarrow{D}_t = M_1\overrightarrow{D}_{t-1} + M_2^{t-1}\overrightarrow{m}_0 + (\text{const.}(2,t)\mathbf{1}', \text{const.}(3,t)\mathbf{1}')', \tag{4.15}$$

here $\overrightarrow{m}_0$ is $\mathcal{F}_0$ measurable and in L^2 and $\rho(M_2) < 1$. Finally, we have

$$\overrightarrow{D}_t - E(\overrightarrow{D}_t) = M_1^{t-1}\overrightarrow{D}_1 + \sum_{i=0}^{t-1} M_1^i M_2^{t-i-1}\overrightarrow{m}_0 + M_1^{t-1}\tilde{c}_0$$

for some constant vector $\tilde{c}_0$. This clearly implies (4.14) with δ chosen to be $\delta < 1$ and $\delta > \max(\rho(M_1), \rho(M_2))$, completing the proof.

Acknowledgement

The research is partially supported by NSERC grant No. 80490. The author would like to thank the Edinburgh Nonlinear Time Series Workshop and the

British NSERC for providing partial support and the three referees for their valuable comments.

References

Benes, V. E. (1967). Existence of finite invariant measures for Markov processes. *Proc. Amer. Math. Soc.* **18**, 1058–1061.

Bhaskara Rao, M., Subba Rao, T. and Walker, A. M. (1983). On the existence of some bilinear time series models. *J. Time Ser. Anal.* **4**, 95–110.

Granger, C. W. and Andersen, A. (1978). *An Introduction to Bilinear Time Series Models.* Vandenhoeck and Ruprecht, Gottingen.

Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and Its Application.* Academic Press.

Liu, J. and Brockwell, P. J. (1988). On the general bilinear time series model. *J. Appl. Probab.* **25**, 553–564.

Liu, J. (1989a). A simple condition for the existence of some stationary bilinear time series. *J. Time Ser. Anal.* **10**, 33–40.

Liu, J. (1989b). On the existence of a general multiple bilinear time series. *J. Time Ser. Anal.* **10**, 341–356.

Liu, J. and Susko, E. (1992). On strict stationarity and ergodicity of a nonlinear ARMA model. To appear in *J. Appl. Probab.*

Liu, J. (1990). A note on causality and invertibility of a general bilinear time series model. *Adv. Appl. Probab.* **22**, 247–250.

Pham, D. T. (1985). Bilinear Markovian representation and bilinear models. *Stoch. Proc. Appl.* **20**, 295–306.

Pham, D. T. (1986). The mixing property of bilinear and generalized random coefficient autoregressive models. *Stoch. Proc. Appl.* **23**, 291–300.

Subba Rao, T. and Gabr, M. M. (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models.* Lecture Notes in Statistics **24**, Springer-Verlag, New York.

Tong, H. (1983). *Threshold Models in Non-Linear Time Series Analysis.* Lecture Notes in Statistics **21**, Springer-Verlag, New York.

Tweedie, R. L. (1974). R-theory for Markov chains on a general state space I: solidarity properties and R-recurrent chains. *Ann. Probab.* **2**, 840–864.

Tweedie, R. L. (1975). Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stoch. Proc. Appl.* **3**, 385–403.

Tweedie, R. L. (1988). Invariant measures for Markov chains with no irreducibility assumptions. *Appl. Probab.* **25**A, 275–285.

Statistics Department, University of British Columbia, Vancouver, BC V6T 1Z2, Canada.

(Received November 1989; accepted August 1991)