STATISTICAL REPORT

MULTIPLE BILINEAR TIME SERIES MODELS

by

Boonchai K. Stensholt and Dag Tjøstheim

Report No.11

June 1985

Department of Mathematics

UNIVERSITY OF BERGEN

Bergen, Norway

ISSN 0333 - 1865

Department of Mathematics
University of Bergen
5014 Bergen-U
NORWAY

MULTIPLE BILINEAR TIME SERIES MODELS

by

Boonchai K. Stensholt and Dag Tjøstheim

Report No.11 June 1985

Abstract.

A definition of multiple bilinear time series is given. Sufficient conditions are obtained for the existence of strictly stationary solutions conforming to the model, and a brief discussion of the first and second order structure is included.

AMS subject classification: 62M10 (62H99)

Keywords and phrases: Multiple bilinear time series models, strictly stationary processes, tensor products of matrices, spectral radius of a matrix.

1. Introduction.

The theory of bilinear scalar time series has been considered in a number of papers recently (cf. in particular Granger and Andersen (1978), Subba Rao (1981) and Bhaskara Rao et al (1983) and references therein). To our knowledge no theory, not even a definition, has been given in the multivariate case.

In the present paper we propose a definition of a multiple bilinear model. This definition has been motivated by three main concerns. Firstly, the definition should contain the most general scalar models as special cases. Secondly, one should be able to prove existence of multiple bilinear models having specified properties such as strict and/or second order stationarity, and in the second order case one would like to obtain information about the covariance structure. Finally, one should be able to obtain least squares or maximum likelihood estimates. In the following we will concentrate on the first two objectives. We plan to pursue the estimation problem in a separate publication.

2. A multiple bilinear model.

Let $\{X(t), t=0,\pm1,\ldots\}$ be a real scalar stochastic process defined on a probability space (Ω,F,P). The process $\{X(t)\}$ is said to be generated by a bilinear model (cf. Subba Rao 1981) if $X(t)$ satisfies the difference equation

$$X(t) + \sum_{i=1}^{p} a^{i} X(t-i) = e(t) + \sum_{i=1}^{q} b^{i} e(t-i) + \sum_{i=1}^{r} \sum_{j=1}^{s} c^{ij} X(t-i)e(t-j) \qquad (2.1)$$

for every $t=0,\pm1,\ldots$. Here $\{e(t), t=0,\pm1,\ldots\}$ is a sequence of

independent identically distributed random variables on $(\Omega, \mathcal{F}, P)$ with $E\{e(t)\} = 0$ and $E\{e^2(t)\} = \sigma^2 < \infty$, while $\{a^i, 1 \leq i \leq p\}$, $\{b^i, 1 \leq i \leq q\}$ and $\{c^{ij}, 1 \leq i \leq r, 1 \leq j \leq s\}$ are constants.

Now, let $\{X(t) = [X_1(t),\ldots,X_n(t)]^T, t = 0,\pm1,\ldots\}$ be an n-dimensional vector process. An n-dimensional generalization of (2.1) is obtained by requiring the kth component, $1 \leq k \leq n$, of $\{X(t)\}$ to be given for each t by

$$X_k(t) + \sum_{i=1}^{p} \sum_{u=1}^{n} a^i_{ku} X_u(t-i) = e_k(t) + \sum_{i=1}^{q} \sum_{u=1}^{n} b^i_{ku} e_u(t-i)$$
$$+ \sum_{i=1}^{r} \sum_{j=1}^{s} \sum_{u=1}^{n} \sum_{v=1}^{n} c^{ij}_{kuv} X_u(t-i) e_v(t-j) , \tag{2.2}$$

where $\{e(t) = [e_1(t),\ldots,e_n(t)]^T, t = 0,\pm1,\ldots\}$ is a sequence of independent identically distributed vector random variables with $E\{e(t)\} = 0$ and $E\{e(t)\, e^T(t)\} = G$, and where $\{a^i_{ku}\}$, $\{b^i_{ku}\}$ and $\{c^{ij}_{kuv}\}$ are constants, the range of indices being obvious from (2.2).

By introducing matrix notation and the tensor product denoted by $\otimes$ we can write (2.2) in vector form as

$$X(t) + \sum_{i=1}^{p} a^i X(t-i) = e(t) + \sum_{i=1}^{q} b^i e(t-i) + \sum_{i=1}^{r} \sum_{j=1}^{s} d^{ij} \{e(t-j) \otimes X(t-i)\} \tag{2.3}$$

Here $a^i = (a^i_{ku})$, $1 \leq i \leq p$, and $b^i = (b^i_{ku})$, $1 \leq i \leq q$, are $n \times n$ matrices, whereas d^{ij}, $1 \leq i \leq r$, $1 \leq j \leq s$, is an $n \times n^2$ matrix, where the kth row is obtained by vectorizing the $n \times n$ matrix $c^{ij}_k = (c^{ij}_{kuv}, 1 \leq u,v \leq n)$, where u is row index and v is column index; i.e.

$$d^{ij} = \begin{bmatrix} \{\text{vec}\ (c_1^{ij})\}^T \\ \cdot \\ \cdot \\ \cdot \\ \{\text{vec}\ (c_n^{ij})\}^T \end{bmatrix}, \tag{2.4}$$

where in general vec(a) is the column vector obtained by stacking the columns of a matrix a one on top of another in order from left to right.

For subsequent discussions it will also be convenient to write (2.3) in state space form (cf. Priestley 1980). We will then assume that $r = p$ and $s = q$, and $q \leq p$. This is not an essential restriction, since it can be fulfilled by introducing a suitable number of zero matrices. We introduce the state vector $Y(t)$ defined by $Y^T(t) = [X^T(t),\ldots,X^T(t-p+1),e^T(t),\ldots,e^T(t-q+1)]$. In the following the symbol 0 will be used to denote a matrix whose elements are all zero. The dimensions of this zero-matrix will be clear from the context. The notation I_m will be used for the identity matrix of order m.

With this notation it is not difficult to verify using (2.3) that $\{Y(t)\}$ satisfies the equation

$$Y(t) = F\,e(t) + A\,Y(t-1) + \sum_{j=1}^{q} C_j\,\{e(t-j) \otimes I_{n(p+q)}\}Y(t-1)\,. \tag{2.5}$$

Here F and A are $n(p+q) \times n$ and $n(p+q) \times n(p+q)$ matrices given by

$$F = \begin{array}{c} \scriptstyle np \left\{ \right. \\ \scriptstyle nq \left\{ \right. \end{array} \overbrace{\left[\begin{array}{c} I_n \\ 0 \\ \hline I_n \\ 0 \end{array} \right]}^{n} \quad \text{and} \quad A = \begin{array}{c} \scriptstyle np \left\{ \right. \\ \scriptstyle nq \left\{ \right. \end{array} \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\overbrace{\qquad\qquad}^{np}} & \multicolumn{2}{c}{\overbrace{\qquad\qquad}^{nq}} \\ a_1 \cdots a_p & & b_1 \cdots b_q & \\ & & 0 \cdots 0 & \\ I_{n(p-1)} & 0 & \cdots\cdots & \\ & & 0 \cdots 0 & \\ \hline & & 0 \cdots 0 & \\ & 0 & & \\ & & I_{n(q-1)} & 0 \end{array} \right] \tag{2.6}$$

The matrices C_j , $1 \leqq j \leqq q$ have dimensions $n(p+q) \times n^2(p+q)$ and are given by

$$C_j = \begin{array}{r} n \left\{ \right. \\ n(p+q-1) \left\{ \right. \end{array} \overbrace{\left[\begin{array}{c} C_1^j \; \cdots \; C_n^j \\ \hline 0 \end{array} \right]}^{n^2(p+q)} \tag{2.7}$$

where C_v^j , $1 \leqq v \leqq n$, are the $n \times n(p+q)$ matrices given by

$$C_v^j = n \left\{ \left[\begin{array}{c|c} \overbrace{\qquad\qquad}^{np} & \overbrace{\quad}^{nq} \\ \{\text{vec } (c_{1v}^j)\}^T & \\ \vdots & 0 \\ \{\text{vec } (c_{nv}^j)\}^T & \end{array} \right] \right. \tag{2.8}$$

with c_{kv}^j , $1 \leqq k \leqq n$, being $n \times p$ matrices defined by

$c^j_{kv} = (c^{ij}_{kuv}\ , 1 \leqq u \leqq n, 1 \leqq i \leqq p)$, where u is row index and i is column index.

In the sequel we will make a distinction between superdiagonal models with $d^{ij}=0$ for $i > j$ in (2.3) and subdiagonal models with $d^{ij}=0$ for $i < j$. (Note that our definitions are different from Granger and Anderson 1978.) Subdiagonal models are easier to handle when it comes to questions of stationary solutions and computation of covariances, and we will be able to prove quite a general result for these models in Section 4. However, we start out by a somewhat special model that is neither sub- nor superdiagonal.

3. A simple bilinear model

If we look at the special case of (2.5) where $C_j = 0$ except for $j = k$, we obtain

$$Y(t) = Fe(t) + AY(t-1) + C\{e(t-k) \otimes I_{n(p+q)}\}Y(t-1) \tag{3.1}$$

where for ease of notation we have omitted the subscript on C_k. This corresponds to a model where d^{ij} in (2.3) is non-zero for $i = 1,\ldots,r$; $j = k$ only. We will look at the model (3.1) under the range of k-values $0 \leqq k \leqq q$, where it should be noted that $k = 0$ is not included in (2.5). We use $\rho(a)$ to denote the spectral radius of a matrix a and F^e_t to denote the σ-algebra generated by $\{e(s)\ ,\ s \leqq t\}$.

Theorem 3.1: Let $\{e(t)\ ,\ t = 0,\pm1,\ldots\}$ be a sequence of independent identically distributed vector random variables defined on the probability space (Ω,F,P) such that $E\{e(t)\} = 0$ and $E\{e(t)e^T(t)\} = G$. Let F , A and C be as in (2.6) and (2.7) and $H = E[\{e(t) \otimes I_{n(p+q)}\} \otimes \{e(t) \otimes I_{n(p+q)}\}]$.

$c^j_{kv} = (c^{ij}_{kuv}\,,\ 1 \le u \le m,\ 1 \le i \le p)$, where u is row index and i is column index.

In the sequel we will make a distinction between superdiagonal models with $d^{ij}=0$ for $i > j$ in (2.3) and subdiagonal models with $d^{ij}=0$ for $i < j$. (Note that our definitions are different from Granger and Andersen 1978.) Subdiagonal models are easier to handle when it comes to questions of stationary solutions and computation of covariances, and we will be able to prove quite a general result for these models in Section 4. However, we start out by a somewhat special model that is neither sub- nor superdiagonal.

3. A simple bilinear model

If we look at the special case of (2.5) where $C_j = 0$ except for $j = k$, we obtain

$$Y(t) = Fe(t) + AY(t-1) + C\{e(t-k) \otimes I_{n(p+q)}\}Y(t-1) \qquad (3.1)$$

where for ease of notation we have omitted the subscript on C_k. This corresponds to a model where d^{ij} in (2.3) is non-zero for $i = 1,\ldots,r$; $j = k$ only. We will look at the model (3.1) under the range of k-values $0 \le k \le q$, where it should be noted that $k = 0$ is not included in (2.5). We use $\rho(a)$ to denote the spectral radius of a matrix a and F^e_t to denote the σ-algebra generated by $\{e(s)\,,\ s \le t\}$.

<u>Theorem 3.1</u>: Let $\{e(t)\,,\ t = 0, \pm 1, \ldots\}$ be a sequence of independent identically distributed vector random variables defined on the probability space (Ω, F, P) such that $E\{e(t)\} = 0$ and $E\{e(t)e^T(t)\} = G$. Let F, A and C be as in (2.6) and (2.7) and $H = E[\{e(t) \otimes I_{n(p+q)}\} \otimes \{e(t) \otimes I_{n(p+q)}\}]$.

Then if

$$\rho\{(A \otimes A) + (C \otimes C)H\} = \lambda < 1 \ , \tag{3.2}$$

there exists a unique (in almost sure sense) strictly stationary F_t^e-measurable solution to (3.1). This solution is given by

$$Y(t) = Fe(t) + \sum_{j=1}^{\infty} \prod_{r=1}^{j} [A + C\{e(t-k-r+1) \otimes I_{n(p+q)}\}] Fe(t-j) \tag{3.3}$$

where the expression on the right of(3.3) converges absolutely almost surely as well as in the mean for every fixed t.

Proof: The proof is patterned after a similar proof in the scalar case (cf. Bhaskara Rao et al 1983, pp 99-102) and therefore only the main steps will be indicated. The crucial point is the use of a well-known result in probability theory (Chung 1974, p.42) according to which $\sum_j X_j$ converges absolutely almost surely if $\sum_j E(|X_j|) < \infty$ for a sequence of random variables $\{X_n\}$.

We denote by $(a)_i$ the ith component of a vector a and by $(b)_{ij}$ the element in row i and column j of a matrix b. Moreover we omit the subscript $n(p+q)$ on the identity matrix. We have, using independence of the $e(t)$'s, that for $j \geq k$

$$E\Big(\Big| \Big(\prod_{r=1}^{j} [A + C\{e(t-k-r+1) \otimes I\}]Fe(t-j)\Big)_i \Big| \Big)$$

$$= E\Big\{ \Big| \sum_{s=1}^{n(p+q)} \Big(\prod_{r=1}^{j-k} [A + C\{e(t-k-r+1) \otimes I\}]\Big)_{is} \cdot$$

$$\Big(\prod_{r=j-k+1}^{j} [A + C\{e(t-k-r+1) \otimes I\}]Fe(t-j)\Big)_s \Big| \Big\}$$

$$\leq \sum_{s=1}^{n(p+q)} E\Big\{ \Big| \Big(\prod_{r=1}^{j-k} [A + C\{e(t-k-r+1) \otimes I\}]\Big)_{is} \Big| \Big\} \cdot$$

$$E\Big\{ \Big| \Big(\prod_{r=j-k+1}^{j} [A + C\{e(t-k-r+1) \otimes I\}]Fe(t-j)\Big)_s \Big| \Big\} \tag{3.4}$$

Multiplying out in the last expectation and using the fact that only the first and the last factor are dependent, it is seen that typically we get terms of type

$$E\left[\,|e_{i_1}(t-j)\prod_{r=2}^{k}\{e_{i_r}(t-j+1-r)\}e_{i_{k+1}}(t-j)|\,\right]$$

$$\leq [\max_{1\leq i\leq n} E\{|e_i(t)|\}]^{k-1}\cdot E\{|e_{i_1}(t-j)e_{i_{k+1}}(t-j)|\}\ . \tag{3.5}$$

Since second moments exist and since k is fixed, it follows that there is a constant K such that

$$E\left\{|(\sum_{r=j-k+1}^{j}[A+C\{e(t-k-r+1)\otimes I\}]Fe(t-j))_s|\right\}\leq K \tag{3.6}$$

The first expectation in the last expression of (3.4) is treated exactly as in the corresponding scalar case (Bhaskara Rao et al pp. 100-101) and we obtain

$$E\left(|(\prod_{r=1}^{j}[A+C\{e(t-k-r+1)\otimes I\}]Fe(t-j))_i|\right)$$

$$\leq K'n(p+q)[\rho\{(A\otimes A)+(C\otimes C)H\}]^{(j-k)/2} = K'n(p+q)\lambda^{(j-k)/2} \tag{3.7}$$

where K' is a constant. Since $\lambda < 1$, the absolute almost sure convergence and the convergence in the mean of the infinite series in (3.3) follows from the quoted result in Chungs book, and thus $Y(t)$ defined by (3.3) is well-defined. It is easy to check that $\{Y(t), t=0,\pm1,\dots\}$ is strictly stationary and that it defines a solution of (3.1).

Conversely if $\{Y(t), t=0,\pm1,\dots\}$ is a F_t^e-measurable solution of (3.1), then by repeated application of (3.1) we have

$$Y(t) = Fe(t) + \sum_{j=1}^{m-1} \prod_{r=1}^{j} [A + C\{e(t-k-r+1) \otimes I\}] Fe(t-j)$$

$$+ \sum_{r=1}^{m} [A + C\{e(t-k-r+1) \otimes I\}] Y(t-m)$$

from which it can be shown that almost surely $Y(t)$ must be given by (3.3) using identical arguments to those of Bhaskara Rao et al (1983). ||

A corresponding solution to (2.3) is obtained from $Y(t)$ of (3.3) by taking the n first components. In the case where $n = k = 1$, $q = 0$ our result degenerates to the theorem of Section 3 of Bhaskara Rao et al (1983).

4. A general subdiagonal model

We choose to work with the state space representation (2.5). Alternatively we could have started with a direct multivariate generalization of the equation (1.3) of Bhaskara Rao et al (1981) extended with a linear MA component. Working with (2.5) has the advantage that the same technique can be used as in the preceding theorem. Moreover it enables us to state the solution in explicit form and to prove its uniqueness. However, the condition for existence of a stationary solution is most easily stated in terms of a representation generalizing (1.3) of Bhaskara Rao et al. At the end of this section we will comment more closely on the connection between the two representations and show how the result of Bhaskara Rao et al comes out as a special case

The model we will consider is the general subdiagonal case with $c^{ij}_{kuv} = 0$ for $i < j$ in (2.2). This means that the block matrices C^j_v, $1 \leq v \leq n$, composing the upper block row of C_j, $1 \leq j \leq q$, have the structure

$$C_v^j = n \left\{ \begin{bmatrix} \overbrace{0}^{n(j-1)} & \overbrace{R_v^j}^{n(p-j+1)} & \overbrace{0}^{nq} \end{bmatrix} \right. \tag{4.1}$$

where R_v^j is a $n \times (p-j+1)n$ matrix whose detailed structure need not be specified. The following lemma is used in the proof of the main theorem.

Lemma 4.1: Let the matrices A and C_j be as defined in (2.6) and (2.7), and let $I = I_{n(p+q)}$. Then for a subdiagonal model with $c_{kuv}^{ij}=0$ for $i<j$ in (2.2) we have for arbitrary integers m and r

$$C_i\{e(t-m) \otimes I\}A^k C_j\{e(t-r) \otimes I\} = 0 \tag{4.2}$$

for $k+2 \leqq i \leqq q$, $1 \leqq j \leqq q$ and $0 \leqq k \leqq q-2$.

Proof: Using the definition of C_i and the subdiagonality property (4.1) it is not difficult to show that

$$C_i\{e(t-m) \otimes I\} = \begin{matrix} n\{ \\ n(p+q-1)\{ \end{matrix} \overbrace{\begin{bmatrix} S_i \\ 0 \end{bmatrix}}^{n(p+q)} \tag{4.3}$$

where

$$S_i = n \left\{ \begin{bmatrix} \overbrace{0}^{n(i-1)} & \overbrace{\sum_{v=1}^{n} R_v^i \otimes e_v(t-m)}^{n(p-i+1)} & \overbrace{0}^{nq} \end{bmatrix} \right. \tag{4.4}$$

and it follows at once that (4.2) holds for $k=0$.

Using (4.3), (4.4) and the definition of A it is not difficult to show that $A^k C_j\{e(t-r) \otimes I\}$ has non-zero elements only in its first

$(k+1)n$ rows. On the other hand for $i \geq k+2$, the matrix $C_i\{e(t-r) \otimes I\}$ has zeros in its first $(k+1)n$ columns and (4.2) follows for $1 \leq k \leq q-2$. ||

We can now state the main theorem.

Theorem 4.1: Let $\{e(t), t = 0, \pm 1, \ldots\}$ be a sequence of independent identically distributed random variables defined on the probability space $(\Omega, \mathcal{F}, P)$ such that $E\{e(t)\} = 0$ and $E\{e(t)e^T(t)\} = G$. Let F, A and C_j, $1 \leq j \leq q$, be as in (2.6), (2.7) and (2.8) with $c^{ij}_{kuv} = 0$ for $i < j$, and let $H = E[\{e(t) \otimes I_{n(p+q)}\} \otimes \{e(t) \otimes I_{n(p+q)}\}]$. Moreover, let Γ_j, $1 \leq j \leq q$ be the $n^2(p+q)^2 \times n^2(p+q)^2$ matrices defined by

$$\Gamma_1 = A \otimes A + (C_1 \otimes C_1)H$$

$$\Gamma_j = \sum_{i=1}^{j-1} \{(A^{j-i}C_i) \otimes C_j\}H(A^{i-1} \otimes A^{j-1}) + (C_j \otimes C_j)H(A^{j-1} \otimes A^{j-1})$$

$$+ \sum_{i=1}^{j-1} \{C_j \otimes (A^{j-i}C_i)\}H(A^{j-1} \otimes A^{i-1}) \qquad 2 \leq j \leq q \qquad (4.5)$$

where $A^0 = I_{n(p+q)}$, and let L be the $qn^2(p+q)^2 \times qn^2(p+q)^2$ matrix defined by

$$L = \begin{bmatrix} \Gamma_1 \; \Gamma_2 & \cdots\cdots & \Gamma_q \\ I_{(q-1)n^2(p+q)^2} & & 0 \end{bmatrix} \qquad (4.6)$$

Then if $\rho(L) = \lambda < 1$, there exists a unique (in almost sure sense) $\mathcal{F}^e_t$-measurable solution to (2.5). This solution is given by

$$Y(t) = Fe(t) + \sum_{j=1}^{\infty} \prod_{r=1}^{j} [A + \sum_{i=1}^{q} C_i\{e(t-i-r+1) \otimes I_{n(p+q)}\}]Fe(t-j) \qquad (4.7)$$

where the expression on the right of (4.7) converges absolutely almost surely as well as in the mean for every fixed t.

Proof: The same basic principle is used as in the proof of Theorem 3.1. Again we write I for $I_{n(p+q)}$.

The main task is thus to establish the convergence of

$$\sum_{j=1}^{\infty} E(|(\prod_{r=1}^{j} [A + \sum_{i=1}^{q} C_i\{e(t-i-r+1) \otimes I\}]Fe(t-j))_k|) \tag{4.8}$$

for an arbitrary component k, $1 \leqq k \leqq n(p+q)$, under the stated conditions. To this end let $g_{j,t}$ be the $n(p+q)\times n$ matrix defined by

$$g_{j,t} = \prod_{r=1}^{j} [A + \sum_{i=1}^{q} C_i\{e(t-i-r+1) \otimes I\}]F \tag{4.9}$$

Then using Lemma 4.1 successively we have for $j \geqq q+1$

$$\begin{aligned} g_{j,t} &= [A + \sum_{i=1}^{q} C_i\{e(t-i) \otimes I\}]\, g_{j-1,t-1} \\ &= [A + C_1\{e(t-1) \otimes I\}]g_{j-1,t-1} + \sum_{i=2}^{q} C_i\{e(t-i) \otimes I\}\, Ag_{j-2,t-2} \\ &= [A + C_1\{e(t-1) \otimes I\}]g_{j-1,t-1} + \sum_{i=2}^{q} C_i\{e(t-i) \otimes I\}\, A^{i-1} g_{j-i,t-i} \end{aligned} \tag{4.10}$$

From this it is seen that $g_{j,t} = g_{j,t}\{e(t-1),\ldots,e(t-j-q+1)\}$ contains at most first order powers of any $e(t-k)$, $1 \leqq k \leqq j+q-1$, and due to the independence of the e(t)'s it follows that $g_{j,t}$ is square integrable. Let $g_{j,t}' = g_{j,t}e(t-j)$. Then it is easily seen that

$$E\{|(g_{j,t}')_k|\} \leqq (\max_{1\leqq s\leqq n} E[\{(e(t-j))_s\}^2]\cdot \sum_{s=1}^{n} E[\{(g_{j,t})_{ks}\}^2])^{\frac{1}{2}}\,. \tag{4.11}$$

Since the first factor on the right hand side is bounded by a constant, and since $\{(g_{j,t})_{ks}\}^2 = (g_{j,t} \otimes g_{j,t})_{ks;ks}$, it suffices to evaluate

$$E(g_{j,t} \otimes g_{j,t}) = M_j\,, \tag{4.12}$$

where we do not get dependence on t due to the strict stationarity of $\{e(t)\}$.

We introduce

$$D_1 = A + C_1\{e(t-1) \otimes I\}g_{j-1,t-1} \text{ and } D_i = C_i\{e(t-i) \otimes I\}A^{i-1} g_{j-i,t-i} \tag{4.13}$$

for $2 \leqq i \leqq q$, such that $g_{j,t} \otimes g_{j,t} = (\sum_{i=1}^{q} D_i) \otimes (\sum_{i=1}^{q} D_i)$ by (4.10). We can

now use the same technique as in Bhaskara Rao et al (1983, pp.108-109) and obtain after some computation

$$Z_j = LZ_{j-1} , \tag{4.14}$$

where Z_j is the $qn^2(p+q)^2 \times n^2$ matrix given by

$$Z_j = \begin{bmatrix} M_j \\ \vdots \\ M_{j-q+1} \end{bmatrix} , \tag{4.15}$$

and where L is defined in (4.6). Combining (4.11), (4.12), (4.14) and (4.15), it follows that there exists a constant K such that $E\{|(g_{j,t}')_k|\} \leqq K\{\rho(L)\}^{(j-q)/2}$, and since $\rho(L) = \lambda < 1$ this guarantees the convergence of the infinite series in (4.8) and it follows that the expression on the right hand side of (4.7) is well-defined, the convergence being absolute almost sure and in the mean. As for Theorem 3.1 it is easy to check that $\{Y(t), t= 0, \pm1,\ldots\}$ is **strictly stationary** and that it defines a solution of (2.5), and that the first n components constitute a solution of (2.3), both being valid under the subdiagonality assumption.

Conversely if $\{Y(t), t= 0, \pm1,\ldots\}$ is an F_t^e-measurable solution of (2.5), then using the same technique as in the proof of Theorem 3.1, it can be shown that almost surely Y(t) must be given by (4.7).||

There is an alternative formulation of the spectral radius condition of Theorem 4.1 which is easier to compare to the one given in the scalar case in Bhaskara Rao et al (1983, Theorem p.106).

Let

$$B_V^j = n\{\Big[\overbrace{R_V^j}^{n(p-j+1)} \quad \overbrace{0}^{n(j-1)} \quad \overbrace{0}^{nq}\Big] \tag{4.16}$$

with R_V^j as in (4.1) and let

$$B_j = \begin{array}{c} n \\ n(p+q-1) \end{array} \left\{ \overbrace{\left[\begin{array}{ccc} B_1^j & \cdots & B_n^j \\ \hline & 0 & \end{array} \right]}^{n^2(p+q)} \right. \tag{4.17}$$

In terms of the B_j, $1 \leqq j \leqq q$, the defining equation (2.5) can be written as

$$Y(t) = Fe(t) + AY(t-1) + \sum_{j=1}^{q} B_j\{e(t-j) \otimes Y(t-j)\} \tag{4.18}$$

whose scalar version is similar but not quite the same as equation (1.3) in Bhaskara Rao et al (1983). We have the following lemma.

Lemma 4.2 : Let A, C_i and B_i be as defined in (2.6),(2.7) and (4.17) and let $I_{n(p+q)} = I$. Then for arbitrary integers t and k, and $1 \leqq i \leqq q$

$$C_i\{e(t-k) \otimes I\}A^{i-1} = B_i\{e(t-k) \otimes I\} \tag{4.19}$$

Proof: The lemma is trivial for i=1 since $C_1 = B_1$. From the definition of A in (2.6) it follows that A^{i-1}, $2 \leqq i \leqq q$, has the form

$$A^{i-1} = \begin{array}{c} n(i-1) \\ n(p-i+1) \\ n(i-1) \\ n(q-i+1) \end{array} \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\overbrace{\qquad U \qquad}^{np}} & \multicolumn{2}{c}{\overbrace{\qquad V \qquad}^{nq}} \\ I_{n(p-i+1)} & 0 & \multicolumn{2}{c}{0} \\ \hline \multicolumn{2}{c|}{0} & \multicolumn{2}{c}{0} \\ \multicolumn{2}{c|}{0} & I_{n(q-i+1)} & 0 \end{array} \right] \tag{4.20}$$

where U and V are $n(i-1)\times np$ and $n(i-1)\times nq$ matrices whose detailed structure need not be specified. The matrix $C_i\{e(t-k) \otimes I\}$ is depicted in (4.3) and (4.4). Similarly we get

$$B_i\{e(t-k) \otimes I\} = \begin{array}{c} n \\ n(p+q-1) \end{array} \left\{ \overbrace{\left[\begin{array}{c} T_i \\ 0 \end{array} \right]}^{n(p+q)} \right. \tag{4.21}$$

where

$$T_i = n\Big\{ \Big[\overbrace{\sum_{v=1}^{n} R_v^i \otimes e_v(t-k)}^{n(p-i+1)} \quad \overbrace{0}^{n(i-1)} \quad \overbrace{0}^{nq} \Big] \tag{4.22}$$

with R_v^i as in (4.1). Combining (4.3), (4.4) and (4.20)-(4.22) yields the conclusion of the lemma. ||

Theorem 4.2 : The spectral radius condition $\rho(L) < 1$ for existence of a solution stated in Theorem 4.1 can be rephrased in terms of the matrices B_j, $1 \leq j \leq q$, defined in (4.16) and (4.17) since the matrices Γ_j, $1 \leq j \leq q$, given in (4.5) can be expressed as

$$\Gamma_1 = (A \otimes A) + (B_1 \otimes B_1)H$$

$$\Gamma_j = [\{(\sum_{i=1}^{j-1} A^{j-i}B_i) \otimes B_j\} + \{B_j \otimes B_j\} + \{B_j \otimes \sum_{i=1}^{j-1} A^{j-i}B_i\}]H, \quad 2 \leqq j \leqq q \tag{4.23}$$

Proof: For j=1 there is nothing to prove. For $j \geqq 2$ we look at a typical term of Γ_j as given in (4.5), namely

$$\{(A^{j-i}C_i) \otimes C_j\}H(A^{i-1} \otimes A^{j-1})$$

$$= E[\{(A^{j-i}C_i) \otimes C_j\}[\{e(t) \otimes I\} \otimes \{e(t) \otimes I\}](A^{i-1} \otimes A^{j-1})] \tag{4.24}$$

Using well known properties of the tensor product this equals

$$E[[[A^{j-i}C_i\{e(t) \otimes I\}] \otimes [C_j\{e(t) \otimes I\}]](A^{i-1} \otimes A^{j-1})]$$

$$= E[[A^{j-i}C_i\{e(t) \otimes I\}A^{i-1}] \otimes [C_j\{e(t) \otimes I\}A^{j-1}]] , \tag{4.25}$$

which by Lemma 4.2 and properties of the tensor product, reduces to

$$E[[A^{j-i}B_i\{e(t) \otimes I\}] \otimes [B_j\{e(t) \otimes I\}]] = \{(A^{j-i}B_i) \otimes B_j\}H. \tag{4.26}$$

This is the corresponding typical term of (4.23). The other terms are treated similarly. ||

In the subdiagonal case the bilinear term of (2.1) is given by

$\sum_{j=1}^{q} \sum_{i=j}^{p} c^{ij}X(t-i)e(t-j)$ and it follows from (2.7), (2.8), (4.1) and (4.16)

that

$$B_j = \frac{1}{p+q-1}\left\{\left[\begin{array}{c} \overbrace{c^{jj} \quad c^{j+1,j} \quad \cdots \quad c^{pj}}^{p-j+1} \quad \overbrace{0}^{j-1} \quad \overbrace{0}^{q} \\ \hline 0 \end{array}\right]\right. . \tag{4.27}$$

Bhaskara Rao et al (1983) have treated the scalar case, where there is no linear MA part; i.e. $b^1 = \cdots = b^q = 0$ in (2.1), and where the bilinear term is given (cf. their equation (1.3)) as $\sum_{j=1}^{q} \sum_{i=1}^{p} b^{ij} X(t-i-j+1)e(t-j)$.

We can accomodate this model into our general subdiagonal framework by introducing a new AR order p'= p+q-1 and a bilinear term $\sum_{j=1}^{q} \sum_{i=1}^{p+q-1} c^{ij} X(t-i)e(t-j)$ where $c^{ij} = b^{i-j+1,j}$ for $j \leqq i \leqq j+p-1$ and $c^{ij} = 0$ for $j+p \leqq i \leqq p+q-1$. Analogously to (4.27) we obtain

$$B_j = \frac{1}{p+2q-2}\left\{\left[\begin{array}{c} \overbrace{b^{1j} \quad b^{2j} \quad \cdots \quad b^{pj}}^{p} \quad \overbrace{0}^{q-j} \quad \overbrace{0}^{j-1} \quad \overbrace{0}^{q} \\ \hline 0 \end{array}\right]\right. \tag{4.28}$$

Under the added assumption that the linear MA part is zero the Bhaskara Rao et al (1983) scalar model can be represented in our framework (4.18) with matrices

$$A = \begin{array}{c} p \\ 2q-1 \end{array}\left\{\begin{bmatrix} \overbrace{A'}^{p} & \overbrace{0}^{2q-1} \\ 0 & 0 \end{bmatrix}\right. \qquad B_j = \begin{array}{c} p \\ 2q-1 \end{array}\left\{\begin{bmatrix} \overbrace{B_j'}^{p} & \overbrace{0}^{2q-1} \\ 0 & 0 \end{bmatrix}\right.$$

where A' and B_j', $1 \leqq j \leqq q$, correspond to the matrices A and B_j defined in Bhaskara Rao et al (1983, bottom of p.96). Then it is an easy task in linear algebra to show that the general spectral radius condition $\rho(L)<1$ given in Theorems 4.1 and 4.2 reduces to the condition given in Bhaskara Rao et al (1983, Theorem p.106) for this particular scalar case.

5. First and second order structure.

We only treat the n-dimensional subdiagonal model of the preceding section and we assume that the conditions of Theorems 4.1 and 4.2 are fulfilled so that a strictly stationary solution Y(t) exists as given in (4.7). The L^1-convergence of the series expansion for Y(t) guarantees the existence of $\mu_Y = E\{Y(t)\}$.

Theorem 5.1 : Let the matrices A, B_j and F be given by (2.6) and (4.17). Assume that $\rho(A) \neq 1$ and that the conditions of Theorems 4.1 and 4.2 are fulfilled. Then

$$\mu_Y = (I-A)^{-1} \sum_{j=1}^{q} B_j F_j \tag{5.1}$$

with $F_j = E[\{e(t-j) \otimes I_{n(p+q)}\}Fe(t-j)]$.

Proof: As before we let $I = I_{n(p+q)}$. By taking expectations in (2.5) and using $E\{e(t)\} = 0$, we have that μ_Y must satisfy

$$\mu_Y = A\mu_Y + \sum_{j=1}^{q} E[C_j\{e(t-j) \otimes I\}Y(t-1)] \tag{5.2}$$

Inserting for Y(t-1) from (2.5) and using independence of the e(t)'s it follows that

$$\sum_{j=1}^{q} E[C_j\{e(t-j) \otimes I\} Y(t-1)]$$

$$= E[C_1\{e(t-1) \otimes I\}Fe(t-1)] + \sum_{j=2}^{q} E[C_j\{e(t-j) \otimes I\}AY(t-2)] \tag{5.3}$$

Using the same technique as when proving Lemma 4.1 it is not difficult to show that in the superdiagonal case

$$C_i\{e(t-m) \otimes I\}A^kF = 0 \tag{5.4}$$

for an arbitrary integer m and for $k+2 \leqq i \leqq q$, $1 \leqq k \leqq q-2$. Using (5.3), (5.4), Lemma 4.1 and the independence of the e(t)'s, it is shown by inserting successively from (2.5) that

$$\sum_{j=1}^{q} E[C_j\{e(t-j) \otimes I\} Y(t-1)] = \sum_{j=1}^{q} E[C_j\{e(t-j) \otimes I\} A^{j-1}Fe(t-j)] \tag{5.5}$$

5. First and second order structure.

We only treat the n-dimensional subdiagonal model of the preceding section and we assume that the conditions of Theorems 4.1 and 4.2 are fulfilled so that a strictly stationary solution $Y(t)$ exists as given in (4.7). The L^1-convergence of the series expansion for $Y(t)$ guarantees the existence of $\mu_Y = E\{Y(t)\}$.

Theorem 5.1 : Let the matrices A, B_j and F be given by (2.6) and (4.17). Assume that $\rho(A) \neq 1$ and that the conditions of Theorems 4.1 and 4.2 are fulfilled. Then

$$\mu_Y = (I-A)^{-1} \sum_{j=1}^{q} B_j F_j \qquad (5.1)$$

with $F_j = E[\{e(t-j) \otimes I_{n(p+q)}\}Fe(t-j)]$.

Proof: As before we let $I = I_{n(p+q)}$. By taking expectations in (2.5) and using $E\{e(t)\} = 0$, we have that μ_Y must satisfy

$$\mu_Y = A\mu_Y + \sum_{j=1}^{q} E[C_j\{e(t-j) \otimes I\}Y(t-1)] \qquad (5.2)$$

Inserting for $Y(t-1)$ from (2.5) and using independence of the $e(t)$'s it follows that

$$\sum_{j=1}^{q} E[C_j\{e(t-j) \otimes I\} Y(t-1)]$$

$$= E[C_1\{e(t-1) \otimes I\}Fe(t-1)] + \sum_{j=2}^{q} E[C_j\{e(t-j) \otimes I\}AY(t-2)] \qquad (5.3)$$

Using the same technique as when proving Lemma 4.1 it is not difficult to show that in the superdiagonal case

$$C_j\{e(t-m) \otimes I\}A^k F = 0 \qquad (5.4)$$

for an arbitrary integer m and for $k+2 \leq j \leq q$, $1 \leq k \leq q-2$. Using (5.3), (5.4), Lemma 4.1 and the independence of the $e(t)$'s, it is shown by inserting successively from (2.5) that

$$\sum_{j=1}^{q} E[C_j\{e(t-j) \otimes I\} Y(t-1)] = \sum_{j=1}^{q} E[C_j\{e(t-j) \otimes I\} A^{j-1} Fe(t-j)] \qquad (5.5)$$

From (5.2), (5.5) and Lemma 4.2 we have

$$\mu_Y = A\mu_Y + \sum_{j=1}^{q} B_j\ E[\{e(t-j) \otimes I\}\ Fe(t-j)] \tag{5.6}$$

and the conclusion of the theorem follows. ||

It is easily checked using the definition of A, F and B_j, $1 \leqq j \leqq q$, that the last nq components of μ_Y are zero, which is consistent with $E\{e(t)\} = 0$, and $E\{X(t)\}$, where X(t) is given by (2.3), is obtained by taking the first n components of (5.1). In the case where n=1 (scalar case) q=1 and with no linear MA part, the expression (5.1) degenerates into the corresponding expression (3.3) of Subba Rao (1981).

To ensure the mean square convergence of the expansion (4.7) for Y(t) we now assume the existence of fourth moments for {e(t)} (cf. corresponding assumption in Bhaskara Rao et al 1983, remark 3 p.103).

Theorem 5.2 : Assume that the conditions of Theorem 5.1 are fulfilled and that in addition the fourth moments of {e(t)} exist. Let X(t) be as in (2.3) (with $d^{ij}=0$ for $i \leqq j$). Then for s>q we have

$$\text{Cov}\ \{X(t)\ ,\ X(t-s)\} = \sum_{i=1}^{p} a^i\ \text{Cov}\ \{X(t-i)\ ,\ X(t-s)\}\ ; \tag{5.7}$$

i.e. {X(t)} has the same covariance structure as a vector ARMA(p,q) process.

Proof: Since our assumptions guarantee the existence of a second order stationary solution as defined by (4.7), we can multiply (2.3) with X(t-s) and take expectations. Using independence of the e(t)'s and subdiagonality we have for s>q (with $I_{n(p+q)} = I$)

$$E\{X(t)\ X^T(t-s)\} = \sum_{j=1}^{p} a^j\ E\{X(t-j)\ X^T(t-s)\} + \sum_{j=1}^{q} d^{jj}\ E[\{e(t-j) \otimes I\}\ X(t-j)\ X^T(t-s)] \tag{5.8}$$

On the other hand, inserting from (2.3) and using stationarity and the

fact that $E\{e(t)\} = 0$, we have that

$$E\{X(t)\}\, E\{X^T(t-s)\} = \sum_{j=1}^{p} a^j\, E\{X(t-j)\}\, E\{X^T(t-s)\}$$
$$+ \sum_{j=1}^{q} d^{jj}\, E[\{e(t-j) \otimes I\}\, X(t-j)]\, E\{X^T(t-s)\} \qquad (5.9)$$

Moreover, inserting for $X(t-j)$ from (2.3) and using independence of the $e(t)$'s, it is not difficult to show that for $1 \leqq j \leqq q$ and $s > q$, we have

$$E[\{e(t-j) \otimes I\}\, X(t-j)\, X^T(t-s)] = E[\{e(t-j) \otimes I\}\, e(t-j)\, X^T(t-s)]$$
$$= E[\{e(t-j) \otimes I\}\, e(t-j)]\, E\{X^T(t-s)\} = E[\{e(t-j) \otimes I\}\, X(t-j)]\, E\{X^T(t-s)\} \qquad (5.10)$$

The equation (5.7) follows by combining (5.8)-(5.10). ||

Again this reduces to the case treated by Subba Rao (1981, formula (3.18) for $n = q = 1$.

An important task is to try to fit multiple bilinear models to data. In principle this can be done by adapting the procedure of Subba Rao (1981) to the multivariate case. The second order structure can then be used as a way of obtaining preliminary estimates. The estimation and fitting problem will be the subject of a subsequent publication.

References

Bhaskara Rao, M., Subba Rao, T. and Walker, A.M. (1983). On the existence of some bilinear time series models, Journal of Time Series Analysis, 4, 95-110

Chung, K.L. (1974). A Course in Probability Theory. 2nd. ed., Academic Press, London

Granger, C.W.J. and Andersen, A. (1978). An Introduction to Bilinear Time Series Models. Vanderhoeck and Ruprecht, Göttingen.

Priestley, M.B. (1980). State dependent models: A general approach to time series analysis, Journal of Time Series Analysis,1, 47-71.

Subba Rao, T. (1981). On the theory of bilinear time series models, Journal of the Royal Statistical Society, Ser.B, 43, 244-255.