ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 141, 1997

Pestano Gabino, Celina González Concepcion-Concepcion**

CHARACTERIZATION OF THE ORDERS IN VARMA MODELS

Abstract. In this paper we present a method which is able to identify the degrees of the matrix polynomials that are involved in the VARMA models. This method is based on the difference between the ranks of certain matrices defined from the sample covariance matrices of the process. The values of the mentioned difference are arranged in tabular from. The specific structure of this table lets us characterize a VARMA (p, q) model.

We study the relative significance of certain elements to confirm the used pattern. The proposed procedure is illustrated by data simulations.

Key words: VARMA models, matrix of plynomials, ranks method, rational representation, matrix Padé approximation.

1. INTRODUCTION

In (Pestano and González 1994b) we proposed a method to characterize a matrix rational function and to estimate the *minimum*[1] degrees of the polynomials that intervene in such function. This method is the result of the research that we are carrying out in the Numerical Analysis field and, specifically, in matrix Padé approximation. In order to illustrate an application, we have characterized a VARMA model, reduced in certain sense. Such method is based on the ranks of matrices built from the sample covariance matrices of VARMA process.

(Pestano and González 1994b) calculates the rank of a matrix by Gaussian elimination with partial pivoting (Atkinson 1989). In this procedure the rank of a matrix depends on the number of nonzero elements in the diagonal of a triangular matrix.

* University of La Laguna, Department of Applied Economics.
[1] Section 2 will introduce this concept.

Due to the rounding and estimation errors, the elements of such diagonal are not absolute zeros. For this reason, in the mentioned article we choose a number such that every value below it was considered null.

The results that we obtained with exact rational functions were satisfactory. In exchange, the ones obtained with simulated data were very sensitive to the number we used to decide if an element is zero or not, during the Gaussian elimination process. Therefore, the aim of this paper is to study the statistical significance of the elements that are candidates to constitute the diagonal of the mentioned triangular matrix.

In the literature, several researches have studied techniques to obtain the orders of a VARMA model. For instance (Francq 1989) proposes the matrix ε-algorithm as well as in (Tiao and Tsay 1989) is extensively studied the model specification stage by investigating linear combinations of the observed series – they belong to VARMA models.

In the following section we introduce the VARMA models, exposing only the properties that are going to be used. In the third section we present the method mentioning the steps that we have implemented in a FORTRAN program. Section 4 gives a criterion to calculate statistically the value of matrix's rank. Finally, section 5 illustrates the procedure through a simulation and shows the obtained results.

2. INTRODUCTION TO THE VARMA MODELS (Reinsel 1993)

One of the objectives of statistical analysis for multivariate time series is to understand the linear dynamic relationships among the variables. VARMA models have the ability to accommodate a variety of dynamic structures.

Let X_t be a k-dimensional, nondeterministic, stationary, with mean zero process. A multivariate generalization of Wold's Theorem states that X_t can be represented as an infinite vector moving average (MA) process,

$$X_t = W(L)\varepsilon_t \qquad W_0 = I$$

where

$$W(L) = \sum_{j=0}^{\infty} W_j L^j$$

is a $k \times k$ matrix in the backshift operator L, such that $L^j\varepsilon_t = \varepsilon_{t-j}$. The coefficients W_j are not necessarily absolutely summable but do satisfy the weaker condition $\sum_{j=0}^{\infty} \| W_j \|^2 < \infty$. ε_t is a vector white noise process such that $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_t^t) = \Sigma$ and $E(\varepsilon_t \varepsilon_{t+f}^t) = 0$ for $f \neq 0$.

Suppose the matrix $W(L)$ can be represented (at least approximately) as the product of two matrices in the from $A_p^{-1}(L)B_q(L)$ where $A_p(L) = I + A_{p-1}L + \dots + A_0L^p$ and $B_q(L) = I + B_{q-1}L + \dots + B_0L^q$ and the coefficients A_i and $B_j(i = 0, \dots p-1$ and $j = 0, \dots, q-1)$ are constant $k \times k$ matrices, then the linear model VARMA (p, q) is defined by the following relation:

$$X_t + A_{p-1}X_{t-1} + \dots + A_0X_{t-p} = \varepsilon_t + B_{q-1}\varepsilon_{t-1} + \dots + B_0\varepsilon_{t-q}$$

A VARMA process is stationary if the roots of $\det\{(A_p(z)\} = 0$ are all greater than one in absolute value. The process is invertible if the roots of $\det\{(B_q(z)\} = 0$ are all greater than one in absolute value.

Theorem 1. admits the VARMA (p, q) representation: $A_p(L)X_t = B_q(L)\varepsilon_t$, where the degrees of $A_p(L)$ and $B_q(L)$ and are *minimum*[2], if and only if, the sequence of covariance matrices, $(R(s)_{s \in Z})$, of the process X_t statisfies the equation in difference with constant matrix coefficients:

$$\sum_{i=0}^{p-1} A_iR(f+i) = -R(f+p) \quad \forall f \geqslant q-p+1 \tag{1}$$

and also

$$\sum_{i=0}^{p-1} A_iR(q-p+i) \neq -R(q) \tag{2}$$

p being as small as possible.

3. A METHOD TO ESTIMATE THE POSSIBLE MINIMUM ORDERS OF A VARMA MODEL

(Pestano and González 1994a) gives the associated proofs to the results of this section. They are based on the matrix Padé approximation theory (Draux 1987, Guo-liang and Bultheel 1988, Guo-liand 1990).

Proposition 1. Let $F_f(z)$ be the power series $\sum_{i=0}^{\infty} R(-f+i)z^i$. The following sentences are equivalents:

a) Some constant matrix coefficients $A_0, A_1, \dots, A_{p-1}$ exist such that the covariance matrices $(R(s)_{s \in Z})$, verify (1) and (2), being p as small as possible.

[2] We consider that the degrees of two matrix polynomials $A_p(L)$ and $B_q(L)$ are minimum when: If two matrix polynomials $D_g(L)$ and $N_h(L)$, exist with degrees g and h respectivity, and they verify that $D_g(0) = I$ and $A_p^{-1}(L)B_q(L) = D_g^{-1}(L)N_h(L)$ when $h < q$ implies $g > p$ and $g < p$ implies $h > q$.

b) $F_f(z) = P_p^{-1}(z)Q_{q+f}(z)$, $f \geqslant p-1$; where $P_p(z) = I + A_{p-1} + \ldots + A_0 z^p$ and $Q_{q+f}(z) = Q_{q+f} + Q_{q+f-1}z + \ldots + Q_0 z^{q+f}$, being p and q *minimum* degrees.

c) $W(L) = A_p^{-1}(L)B_q(L)$, that is, X_t is a VARMA (p, q) process with p and q *minimum*.

Ranks method (Pestano and González 1994b)

Step 1. Choose the degrees of a possible rational representation for $F_f(z)$.

Choose r and s such that the ranks of the matrices $(R(s-r+i+j-1))_{\substack{1\leqslant i\leqslant r\\ 1\leqslant j\leqslant r}}$, $(R(s-r+i+j-1))_{\substack{1\leqslant i\leqslant r+1\\ 1\leqslant j\leqslant r}}$ and $(R(s-r+i+j-1))_{\substack{1\leqslant i\leqslant r\\ 1\leqslant j\leqslant r+1}}$ are equal. Then $F_f(z)$ can be represented, at least, in the rational from $D_r^{-1}(z)N_{s+f}(z)$. Observe that r and s are not necessarily *minimum*.

Step 2. Look for rational representation of *minimum* degrees for $F_f(z)$.

Check if the ranks of:

$$(R(i-j+k+m-1))_{\substack{1\leqslant k\leqslant j\\ 1\leqslant m\leqslant r+s-i}} \qquad (*)$$

and

$$(R(i-j+k+m-1))_{\substack{1\leqslant k\leqslant j+1\\ 1\leqslant m\leqslant r+s-i}} \qquad (**)$$

are equal for $0\leqslant j\leqslant r$ and $0\leqslant i\leqslant s$.

Step 3. Build a table

Build a table with $s+1$ columns (from the 0-th to the s-th) and $r+1$ rows (from the 0-th to the r-th). Place a "0" in the intersection of the column i with the row j if the ranks of (*) and (**) are equal, and place a "1" if the mentioned ranks are different.

It is necessary to comment that if an intersection (g, h) of the table has a "0" then all the intersections (a, b) with $g\leqslant a\leqslant s$ and $b\leqslant h\leqslant r$, have a zero too. Therefore, theoretically, it is not necessary to build the whole table. However, we have built it in order to make firm the results.

Step 4. Interpretation of the table elements.

To interpret the table we give the following theorem:

Theorem 2. follows a VARMA (p, q) model where the orders p and q are *minimum*, if and only if, the table of the Step 3 has special pattern. This pattern contains a right lower rectangle, with null elements, which left upper corner is (p, q). This corner is well delimited.

Note that Theorem 2 ensures that if X_t admits a VARMA (p, q) model and a VARMA (e, d) model, where p and q and e and d are *minimum* orders, then the corresponding table presents this pattern:

	0	1	...	.	q	.	.	.	.	d	.	.	.	s
0	1	1	...	.	1	.	.	.	.	1	.	.	.	1
1	1	1	...	.	1	.	.	.	.	1	.	.	.	1
.	.	.	...	.	.	.	.	.	.	.	.	.	.	.
.	1	1	...	.	1	.	.	.	1	1	.	.	.	1
e	1	1	...	.	1	.	.	.	1	0	0	0	...	0
.	1	1	...	.	1	.	.	.	1	0	0	0	...	0
.	1	1	...	1	1	.	.	.	1	0	0	0	...	0
p	1	1	...	1	0	0	0	...	0	0	0	0	...	0
.	1	1	...	1	0	0	0	...	0	0	0	0	...	0
.	1	1	...	1	0	0	0	...	0	0	0	0	...	0
r	1	1	...	1	0	0	0	...	0	0	0	0	...	0

4. A STATISTICAL PROCEDURE TO DETERMINE THE RANK OF A MATRIX

We have made an algorithm that takes the necessary decisions and carries out all the operations involved in the table building.

As we have seen, during the procedure it is essential to calculate the rank, or more specifically, the difference between the ranks of certain matrices defined from the sample covariance matrices of the process. To do it we have used the well known method of Gaussian elimination with partial pivoting[3]. Remember that in this method the rank of a matrix is exactly the number of nonzero elements in the diagonal of a transformed triangular matrix.

In each stage of the trangularization we must decide which will be the nonzero pivot element. Due to the estimation[4] and rounding errors[5], certain elements should be, theoretically, null but they are not.

[3] Often an element would be zero except for rounding errors that have ocurred in calculating. Using such an element as pivot element will result in gross errors in the further calculations in the matrix. To guard against this, and for other reasons involving the propagation of rounding errors, we introduce partial pivoting (Atkinson 1989).

[4] The covariance matrix at $\log n$, $n \in Z$, $R(n) = E(X_t X'_{t+n})$ is estimated by the sample covariance matrix $C(n) = N^{-1} \sum_{t=1}^{N-n} (X_t - \bar{X})(X_{t+n} - \bar{X})$ where $\bar{X} = N^{-1} \sum_{t=1}^{N} X_t$ is the vector sample mean and N is the sample size.

[5] If the coefficients of the matrices $C(n)$ vary greatly in size, then it is likey that large loss of significance errors will be introduced and the propagation of rounding errors will be introduced worse. To avoid this problem, it is usually scaling the $C(n)$ so that the elements vary less (Atkinson 1989).

Immediately afterwards we expose the ideas of the statistic procedure that we have made in order to know if an element (possible pivot) is zero or not.

We denote by $M^0 = (m_{ij}^0)$, $1 \leqslant i \leqslant u$ and $1 \leqslant j \leqslant v$, to the original matrix. We would like to know the value of its rank. Let $M^k = (m_{ij}^k)$ be the matrix that stems from the k-th stage of the triangularization. Supposing that the pivot element, in the $k+1$ stage, belongs to the n-th row of M^k, the matrix $M^{*k} = (m_{ij}^{*k})$ is built from the rows of M^k in the following way:

i) $m_{k+1j}^k = m_{nj}^k$ and $m_{nj}^{*k} = m_{k+1j}^k$ for $j = 1, 2, ..., v$[6].

ii) $m_{pj}^{*k} = m_{pj}^k$ otherwise.

By and large and supposing that the pivot element in the stage $k+1$ is m_{k+1h}^{*k}, where $h \in \{k+1, k+2, ..., v\}$, the Gaussian elimination calculates m_{ij}^{k+1} as follows:

$$m_{ij}^{k+1} = \begin{cases} m_{ij}^{*k} - \dfrac{m_{ih}^{*k}}{m_{k+1h}^{*k}} m_{kj}^{*k} & k \leqslant i \leqslant u, \ h \leqslant j \leqslant v \\ m_{ij}^{*k} & \text{otherwise} \end{cases}$$

To decide which one is the pivot element in the $k+2$ stage we must study statistically if certain elements of M^{k+1} are zero or not. Given the structure of m_{ij}^{k+1} the problem is that of testing the hypothesis: $H_0: f(B) = 0$ that involves a nonlinear function of the sample covariance matrices' coefficients[7]. The statistic to this test would be (Greene 1991):

$$z = \frac{f(\hat{B})}{\text{stimated standard error}}$$

which is distributed as the standard normal distribution.

The discrepancy in the numerator presents no difficulty. However, obtaining an estimate of the sampling variance of $f(\hat{B})$ involves the variance of a nonlinear function of $\hat{B}$. An approach that uses the large-sample properties of the estimates and provides an approximation to the variances we need is based on a linear Taylor series approximation. A linear Taylor series approximation to $f(\hat{B})$ around the true parameter vector B, is:

$$f(\hat{B}) \cong f(B) + \left(\frac{df(B)}{dB}\right)^T (\hat{B} - B)^8.$$

[6] Note that the pivot element in the $k+1$ stage can be an element of the column h of M^k, with $h \in (k+1, ..., v)$; it is due to the fact that can happen $\underset{k+1 \leqslant i \leqslant u}{Max} |m_{ik+1}^k| = 0$ in the partial pivoting, then we choose the pivot element in the following column.

[7] To generalize we denoting by $m_{ij}^{k+1} \equiv f(\hat{B})$ where B is the vector $(R_0^1 ... R_0^r R_1^1 ... R_1^r R_2^1 ... R_2^r ... R_E^1 ... R_E^r)$ being R_i^j the j-th row of $R(i)$ and $E = r + s + 1$.

[8] T denotes transpose.

In general, the expected value of a nonlinear function is not equal to the function of the expected value, so, we must rely on consistency rather than unbiasedness here. Thus, in view of the approximation, and assuming that $p\lim(\hat{B}) = B$, we are justified in using $f(\hat{B})$ as an estimate of $f(B)$ (the relevant result is the Slutsky theorem). Assuming that our use this approximation is appropriate, the variance of the nonlinear function is approximately equal to the variance of the right-hand side, which is, then,

$$Var[f(\hat{B})] \cong g^T Var[\hat{B} - B] g$$

$$g = \frac{df(B)}{dB}$$

The derivatives[9] in the expression for the variance are functions of the unknown parameters. Since these are being estimated, we use our sample estimates in computing the derivatives. To estimate the variance of the estimator, we use an approximation given by (H a n n a n 1976):

Denoting $\tau_{ab}(n) = C_{ab}(n) - R_{ab}(n)$, then $Cov(\tau_{ab}(m)\tau_{cd}(n)) \cong$

$$\sum_{j=-N+1}^{N-1} \frac{1}{N}(1 - \frac{|j|}{N})[R_{ac}(j)R_{bd}(j+n-m) + R_{ad}(j+n)R_{bc}(j-m)]$$ [10]

assuming that the process X_t has zero fourth-order cumulates as in the case of a Gaussian process.

5. ILLUSTRATIVE EXAMPLE

To illustrate the behavior of the proposed procedure, we conducted a simulation study. The VARMA (1, 0) model:

$$(I + A_0 L)X_t = \varepsilon_t$$

with

$$A_0 = \begin{bmatrix} -0.2 & -0.3 \\ 0.6 & 1.1 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix}$$

was employed in the simulation.

The results are based on 100 replications each with 350 observations, and the 5% level was used.

[9] The recursive structure of the m_{ij}^{k+1} is very useful to calculate the derivatives g.

[10] To estimate them we replace $R(n)$ by $C(N)$ (F r a n c q 1989).

The expected table is:

	0	1	2	3	4	5
0	1	1	1	1	1	1
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0

We obtained the results below:

- 94 tables show that the degree of the numerator is zero.
- 90 tables show that the degree of the denominator is one.

And more specifically:

- 74 tables show a clear pattern that identifies the VARMA(1,0) model.

REFERENCES

Atkinson K. E. (1989): *An Introduction to Numerical Analysis*, John Wiley & Sons, New York.

Draux A. (1987): *On the Non-Normal Padé Table in a Non-Commutative Algebra*, Publication ANO 175, Lille.

Franq C. (1989): *Identification et minimalité dans les séries chronologiques*, These Université de Montpellier.

Greene W. H. (1991): *Econometric Analysis*, Maxwell Macmillan International Editions.

Guo-liang Xu and Bultheel A. (1988): *The Problem of Matrix Padé Approximation*, Report TW 116, Department of Computer Science, K. U, Leuven.

Guo-liang Xu (1990): *Existence and Uniqueness of Matrix Padé Approximation*, J. Comp. Math. Vol. 8, No 1, p. 65–74.

Hannan E. J. (1976): *The Asymptotic Distribution of Serial Covariances*, The Annals of Statistics, Vol. 4, No 2, p. 396–399.

Pestano C. and González C. (1994a): *Funciones racionales matriciales: Una aplicación al estudio de modelos VARMA*, (pending publication).

Pestano C. and González C. (1994b): *Un método de rangos para estimar los órdenes de un modelo VARMA en su forma reducida*, VIII Estudios de Economía Aplicada, Vol II, p. 103–110.

Reinsel G. C. (1993): *Elements of Multivariate Times Series Analysis*, Springer Verlag, New York.

Tiao G. C. and Tsay R. S. (1989): *Model Specification in Multivariate Time Series*, J. R. Statist. Soc. B, 51, No 2, p. 157–213.

Pestano Gabino, Celina González Concepcion-Concepcion

CHARAKTERYSTYKA UPORZĄDKOWANIA W MODELACH VARMA

W artykule prezentujemy metodę stosowaną do określenia stopnia macierzy wielomianów występującą w procesach stochastycznych typu VARMA.

Proponowana metoda oparta jest na różnicy pomiędzy rzędami pewnych macierzy uzyskanych z próbkowych macierzy kowariancji tego procesu. Wartości rozważanych różnic są podane w formie tablicowej. Specyficzna struktura tej tablicy pozwala nam scharakteryzować model VARNA (p, q).

Przedstawiona procedura jest zilustrowana w punkcie 5 symulacyjnym przykładem.