J. Statist. Comput. Simul., 1985, Vol. 21, pp. 171–177
0094-9655/85/2102-0171 $18.50/0

A Priori Sample Size Evaluation and Information Matrix Computation for Time Series Models

BALA G. DHARAN
Assistant Professor of Accounting, Jones Graduate School of Administration, Rice University, Houston, TX 77251, U.S.A.

(*Received March 28, 1984; in final form December 11, 1984*)

This paper describes computational algorithms for determining the information matrix and evaluating the sample size needed for satisfactory estimation of a proposed time series model. These algorithms are potentially useful to many application studies where the researcher needs to know the necessary sample sizes for model estimation *prior to* committing costly resources for model development or estimation. They are also useful to simulation studies requiring *a priori* sample size decisions. The algorithms apply to any general univariate linear time series model with autoregressive and moving average coefficients.

KEY WORDS: Sample size evaluation, ARMA model, information matrix, time series model, algorithm.

1. INTRODUCTION

The autoregressive and moving average (ARMA) linear time series models of Box and Jenkins (1976) are often used in a wide variety of natural science and social science studies dealing with forecasting. A unique characteristic of this class of time series models—not shared by other econometric techniques such as linear regression models and simultaneous equation systems—is that if the researcher has

certain prior beliefs about the values of the ARMA model coefficients he (or she) wants to estimate, then he can analytically determine the sample size needed for satisfactory model estimation *prior to* doing the estimation. This is so because the information matrix and the related covariance matrix of the ARMA model coefficients do not depend on residual error variance and hence can be computed *a priori* (i.e., without having access to actual realizations of the model) solely as a function of the sample size and the values of the model coefficients postulated by the researcher. Using the computed theoretical standard errors and the postulated coefficient values, one can then determine the sample size needed to estimate the coefficients with acceptable magnitudes of standard errors. Alternatively, if the researcher has no control over the available sample size, he can evaluate whether a given sample size can give coefficient estimates with acceptable magnitudes of standard errors, and thus whether costly resources should be spent on further development and estimation of a proposed model.

This property of ARMA models has not been exploited in the past by the many application studies in economics and other areas. One likely reason is that computational algorithms for determining the information matrix of a *general* time series model are not available in the literature, though some theoretical discussion underlying such algorithms is available in Anderson (1971), Box and Jenkins (1976), and other places. The algorithms presented here fill this void. They apply to any general ARMA(p, q) model, where p refers to the number of autoregressive coefficients and q refers to the number of moving average coefficients. These algorithms are designed to evaluate the standard errors of the coefficients directly, rather than the standard errors of the corresponding characteristic roots. They are thus preferable to a narrower procedure discussed by Box and Jenkins (1976) for determining the covariance matrix of the characteristic roots of an ARMA model, which can be used only when the roots are real and different from one another. By contrast, the algorithms described here have a more general and unrestricted applicability.

2. INFORMATION MATRIX OF AN ARMA MODEL

Consider a stationary and invertible ARMA(p, q) model, written as

$$\phi(L) Z_t = \theta(L) e_t, \tag{2.1}$$

where $\phi(L)=1-\phi_1 L-\phi_2 L^2 \ldots -\phi_p L^p$, $\theta(L)=1-\theta_1 L-\theta_2 L^2 \ldots -\theta_q L^q$, Z_t is the stationary stochastic variable being studied, e_t is from a white noise process with zero mean and a variable of σ^2, and L is a lag operated such that $L^x Z_t = Z_{t-x}$ for $x \geqq 0$. Let $K=p+q$.

Denote the $(K \times 1)$ vector of coefficients $(\phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q)^T$, where T represents transpose, as β. Then the $(K \times K)$ asymptotic theoretical covariance matrix of β is given by

$$V(\beta)=I_l^{-1}, \tag{2.2}$$

where I_l is the $(K \times K)$ *asymptotic information matrix* of (2.1). Hence, to determine the theoretical standard errors of the coefficients as a function of sample size, one must first compute the information matrix. To aid in this computation, Box and Jenkins (1976, pp. 240–242) suggest forming two variables, u_t and x_t, as

$$\phi(L)u_t = e_t, \tag{2.3}$$

and

$$\theta(L)x_t = -e_t. \tag{2.4}$$

The information matrix is composed of the autocovariances and cross-covariances of u_t and x_t. Let n be the sample size. Then the information matrix can be written as

$$I_l = n \cdot \begin{bmatrix} A & B \\ B^T & D \end{bmatrix} \cdot \sigma^{-2} = n\sigma^{-2} M, \tag{2.5}$$

where the submatrix A is of the order $(p \times p)$, B is $(p \times q)$, B^T is the transpose of B, and D is $(q \times q)$.

The submatrix A contains autocovariances of u_t, $\{A_{ij}\} = V_{uu}(i-j)$, where,

$$V_{uu}(k) = V_{uu}(-k) = E(u_t u_{t+k}) = E(u_t u_{t-k}), \quad (k=1,2,\ldots), \tag{2.6}$$

is the autocovariance of u_t at lag k, and $V_{uu}(0)$ is the variance of u_t. Similarly, the submatrix D contains $\{D_{ij}\} = V_{xx}(i-j)$. The submatrix B of the information matrix contains cross-covariances between u_t

and x_t, $\{B_{ij}\} = V_{ux}(i-j)$, where,

$$V_{ux}(k) = E(u_t x_{t+k}), \tag{2.7}$$

for all k. Note that $V_{ux}(k) = V_{xu}(-k)$.

3. ALGORITHMS FOR SAMPLE SIZE

Computing the covariance matrix $V(\beta)$ requires computing the variance, covariance and cross-covariance elements of u_t and x_t. Algorithms presented in this section do this by utilizing the fact that these elements can be expressed as functions of the coefficient vector, with the white noise variance acting only as a scale factor.

Determination of the A and D submatrices, which contain the autocovariances of the u_t and x_t processes respectively, can be done efficiently by using an algorithm given by McLeod (1975), as corrected by McLeod (1977), for deriving the theoretical auto-covariance function of an ARMA(p, q) model. Given the coefficient vector β, this algorithm computes A_{ij} and D_{ij} as functions of β and σ^2. As will be seen below, elements of submatrix B are also scaled by σ^2. However, the white noise variance is not a necessary information with respect to the information matrix since it can be seen from (2.5) that the submatrices are multiplied by σ^{-2}. In other words, elements of the information matrix are *independent of* σ^2. Hence in applying the algorithm of McLeod for A and D and the following algorithm for B, σ^2 can be set to 1 without loss of generality.

To compute the B submatrix (which needs to be done only when $p>0$ *and* $q>0$), it is clear from the definition of B_{ij} that one needs to compute $(p-1)+(q-1)+1 = p+q-1$ cross-covariances of u_t and x_t. Computational ease is obtained when these are placed in a vector z in the following order:

$$z = (V_{ux}(p-1)\, V_{ux}(p-2) \ldots V_{ux}(0)\, V_{ux}(-1) \ldots V_{ux}(1-q))^T. \tag{3.1}$$

Elements of z can be obtained by solving a system of $(p+q-1)$ linear equations of the form $Hz = y$, where H is a $(p+q-1)$ by $(p+q-1)$ matrix whose elements are functions of β. The vector y is zero except for the pth element which is σ^2, as shown below.

Equations in the above equation system can be derived from exploiting the following relationship obtained from (2.3) and (2.4):

$$\phi(L)u_t = -\theta(L)x_t. \tag{3.2}$$

Multiplying both sides by u_0 and taking expectations,

$$E\phi(L)u_t u_0 = 0 = -E\theta(L)x_t u_0$$

$$= \sum_{j=0}^{q} \theta_j V_{ux}(t-j), \tag{3.3}$$

where $\theta_0 = -1$ and E is the expectation operator. The first $p-1$ equations are obtained by taking $t = p-1, p-2, \ldots, 1$. Similarly, the last $q-1$ equations arise from

$$E\theta(L)x_t x_0 = 0 = -E\phi(L)u_t x_0$$

$$= \sum_{j=0}^{p} \phi_j V_{ux}(j-t), \tag{3.4}$$

where $\phi_0 = -1$, by taking $t = 1, 2, \ldots, q-1$. Finally, the middle (i.e., pth) equation is obtained from (3.3) by taking $t = p$, solving for $V_{ux}(p)$, and substituting into the following expression:

$$Ee_t x_t = \sigma^2 = E\phi(L)u_t x_t$$

$$= \sum_{j=0}^{p} \phi_{p-j} V_{ux}(p-j). \tag{3.5}$$

The resulting middle equation has the following format:

$$\sigma^2 = \sum_{j=1}^{p} \phi_{p-j} V_{ux}(p-j) + \sum_{j=1}^{q} \phi_p \theta_j V_{ux}(p-j). \tag{3.6}$$

Based on the above analysis, the algorithm to compute the submatrices B and B^T for an ARMA(p, q) model is:

ALGORITHM I

1) Start with postulated values for the coefficient vector.

2) Form the equation system $Hz = y$ described in (3.3), (3.4) and (3.6).
3) Set $\sigma^2 = 1$.
4) Compute H^{-1} and form the vector z in (3.1) as $H^{-1}y$.
5) Form submatrices B and B^T.

As noted earlier, the elements of A, B and D are scaled by σ^2, and the elements of I_l in (2.5) have a factor σ^{-2}. Hence, by arbitrarily setting $\sigma^2 = 1$, the above algorithm and the one by McLeod have computed the values of the elements of I_l except for the sample size factor n in (2.5). To estimate n, assume that the researcher is interested in a sample size that would yield sufficiently small standard errors for each coefficient estimate. Let μ be the minimum desired value for the ratio of each coefficient estimate to its standard error. Then the desired minimum sample size is given by the following algorithm:

ALGORITHM II

1) Start with matrix M defined in (2.5) whose submatrices are computed using Algorithm I and McLeod (1975).
2) Compute $M^{-1} = \{\alpha_{ij}\}$.
3) Compute the desired minimum sample size as

$$n = \min_{i=1,\ldots,K} \left(\frac{\alpha_{ii}\mu^2}{\beta_i^2} \right).$$

To summarize, the standard errors of an ARMA(p, q) model coefficients are obtained by first computing the model's information matrix whose elements are functions of the postulated coefficients. The standard errors are then obtained by assuming a desired ratio of coefficient value to standard error. Alternatively, a procedure similar to Algorithm II can be written which will compute the expected standard errors if a sample size is assumed. When the residual errors are assumed to come from a normal distribution, the ratio, μ, of a coefficient value to its standard error has a t-distribution. Hence the above algorithms can be used to obtain univariate measures of the *expected* statistical significance of the postulated coefficients, for assumed sample sizes.

4. CONCLUSION

This paper has presented the algorithms for the computation of the information matrix of an ARMA model. The algorithms have been used successfully by Dharan (1983) to evaluate the sufficiency of available sample sizes for the identification of an ARMA model. The general applicability of the algorithms to any ARMA(p,q) model makes them a robust and powerful tool for other researchers to evaluate the sample size needs for satisfactory estimation of a time series model whose coefficients are specified by theory or are initialized by prior belief.

The comments of the reviewer, which greatly simplified the presentation of the analysis in the paper, are gratefully acknowledged.

References

Anderson, T. W. (1971). *The Statistical Analysis of Time Series.* John Wiley and Sons, New York.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco, Revised Edition.

Dharan, B. G. (1983). Identification and estimation issues for a causal earnings model. *Journal of Accounting Research*, **21,** 18–41.

McLeod, I. (1975). Derivation of the theoretical autocovariance function of autoregressive-moving average time series. *Applied Statistics*, **24,** 255–256.

McLeod, I. (1977). Correction. *Applied Statistics*, **26,** 194.