A C T A   U N I V E R S I T A T I S   L O D Z I E N S I S

FOLIA OECONOMICA 54, 1986

Andrzej Tomaszewicz*, Abdul Majid Hamza Al-Nasir**

ON THE ESTIMATION IN SIMPLE LINEAR REGRESSION MODEL WITH AUTOREGRESSIVE MOVING AVERAGES (ARMA) ERROR

1. Introduction

Suppose that a response y_t follows the model:

$$y_t = B_o + B_1 x_t + e_t, \qquad t = 1, 2, \ldots \tag{1}$$

The simple linear equation (1) states that in period t, the value of y, the response, is determined by four factors: the population constant B_o, the population regression coefficient B_1, the level of x, and the level of e, the disturbance term.

The disturbance term is assumed to have certain properties in order to carry out statistical estimation and tests of significance. Departures from these assumptions bring some of the characteristic problems. For instance, usually, one assumes that all pairs of values of e_t, whether adjacent in time or not, are not correlated. A departure from this assumption gives rise to the problem of autocorrelation.

This problem has been studied by a number of statisticians, for instance, A n d e r s o n (1942), C o c h r a n (1949), Q u e n o u i l l e (1949), D u r b i n (1950), H a n n a n (1957), T h e i l and N a g a r (1961), and in recent years, B o x and P i e r c e (1970) studied the distribution of residuals which follow a mixed ARMA model, P i e r c e (1971) developed a method for estimating the parameters by using the first order terms in Taylor's expansion and applied it to an ARMA of the first order and N u-

*Lecturer, Institute of Econometrics and Statistics, University of Łódź.

**Lecturer, University of Baghdad.

r i (1979) proposed a method for finding approximate estimates for the parameters which is basically related to the least squares method with some modifications.

The purpose of this paper is to investigate the properties of the model (1), especially when the errors e_t follow a low order ARMA. That is because, in practice, it is frequently true that an adequate representation of actually occurring stationary time series can be obtained in mixed models, in which the order of autoregressive process p and the order of moving average q are not greater than 2 and often less than 2 (B o x and J e n k i n s 1970). So ARMA (3) is studied in some detail, and some new results are obtained. Attempts are made to construct suitable examples: artificial examples and economic data examples.

2. The Properties

In the model (1) suppose that

$$e_t = \Phi^{-1}(B)\Theta(B)a_t, \tag{2}$$

where a_t's are independently and normally distributed random variables with zero means and variance σ^2, Φ and Θ are polynomials such that:

$$\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2, \quad \Theta(B) = 1 - \Theta_1 B - \Theta_2 B^2$$

and B is the back-shift operator (lag operator) defined by $B^j f_t = f_{t-j}$ for any function f_t and for $j = 1, 2$. The following are some of the characterizations for the proposed model for e_t.

1. Using equation (2) the second order autoregressive, the second order moving average process ARMA (3) can be written as:

$$(1 - \Phi_1 B - \Phi_2 B^2)e_t = (1 - \Theta_1 B - \Theta_2 B^2)a_t,$$

or

$$e_t = \Phi_1 e_{t-1} + \Phi_2 e_{t-2} + a_t - \Theta_1 a_{t-1} - \Theta_2 a_{t-2}. \tag{3}$$

2. Multiplying equation (3) by e_{t-k} and taking expectations, it could be obtained that

$$\gamma_k = \Phi_1\gamma_{k-1} + \Phi_2\gamma_{k-2} + \gamma_{ae}(k) - \theta_1\ \gamma_{ae}(k-1) - \theta_2\gamma_{ae}(k-2), \tag{4}$$

where $\gamma_k = \mathrm{cov}(e_t, e_{t-k})$ and $\gamma_{ae}(k)$ is the cross covariance between a and e at lag difference k, defined by $E(a_t e_{t-k})$. Since e_{t-k} depends only on shocks which have occurred up to time t-k, we obtain

$$\gamma_{ae}(k) = \begin{cases} \sigma^2 f_1\ (\Phi_1, \theta_1) & k < 0 \\ \sigma^2 & k = 0 \\ 0 & k > 0 \end{cases} \tag{5}$$

3. It follows that

$$\gamma_o = \Phi_1\ \gamma_1 + \Phi_2\gamma_2 + \sigma^2 - \theta_1\gamma_{ae}(-1) - \theta_2\gamma_{ae}(-2), \tag{6}$$

$$\gamma_1 = \Phi_1\ \gamma_o + \Phi_2\gamma_1 - \theta_1\sigma^2 - \theta_2\gamma_{ae}(-1), \tag{7}$$

$$\gamma_2 = \Phi_1\ \gamma_1 + \Phi_2\gamma_o - \theta_2\sigma^2. \tag{8}$$

And for $k > 2$, $\gamma_k = \Phi_1\gamma_{k-1} + \Phi_2\gamma_{k-2}$, which does not involve the moving average parameters. Therefore, after lag 2, the autocovariances and consequently the autocorrelation coefficients behave as those for the AR process. And so we reach the same conclusion as Anderson (1976).

4. Multiplying equation (3) first by a_{t-1} and secondly by a_{t-2} and taking expectations we find

$$\gamma_{ae}(-1) = (\Phi_1 - \theta_1)\,\sigma^2 \tag{9}$$

and

$$\gamma_{ae}(-2) = (\Phi_1^2 - \Phi_1\theta_1 + \Phi_2 - \theta_2)\,\sigma^2. \tag{10}$$

Substituting these in equations (6) and (7) leads to

$$\gamma_o = \frac{(1-\Phi_2)\left[1+\theta_1^2+\theta_2{}^2 - 2\theta_2\Phi_2\right] - 2\Phi_1\left[\theta_1+\theta_2\Phi_1 - \theta_1\theta_2\right]}{(1+\Phi_2)\left[(1-\Phi_2)^2 - \Phi_1^2\right]}\sigma^2, \tag{11}$$

$$\gamma_1 = \frac{\Phi_1}{1 - \Phi_2}\gamma_0 - \frac{\theta_1 + \theta_2\Phi_1 - \theta_2\theta_1}{1 - \Phi_2}\sigma^2 \tag{12'}$$

and

$$\gamma_2 = \frac{\Phi_2 + \Phi_1^2 - \Phi_2^2}{1 - \Phi_2}\gamma_0 - \frac{\theta_2 + \theta_1\Phi_1 - \theta_2\Phi_2 + \Phi_1^2\theta_2 - \Phi_1\theta_1\theta_2}{(1 - \Phi_2)}\sigma^2. \tag{13}$$

3. Least Squares Estimators

To obtain the least squares estimators of the coefficients, we first rewrite equation (1) as follows

$$\frac{\Phi(B)}{\theta(B)}\, y_t = \frac{\Phi(B)}{\theta(B)}\left(B_0 + B_1 x_t\right) + a_t. \tag{14}$$

Then the problem is carried out by minimizing

$$v = \sum_{t=1}^{n}\left[\frac{\Phi(B)}{\theta(B)}\left(y_t - B_0 - B_1 x_t\right)\right]^2. \tag{15}$$

We find that

$$\frac{\partial v}{\partial B_0} = -2\sum_t \left\{\frac{\Phi(B)}{\theta(B)}\right\}^2 \left(y_t - B_0 - B_1 x_t\right),$$

$$\frac{\partial v}{\partial B_1} = -2\sum_t \left\{\frac{\Phi(B)}{\theta(B)}\right\}^2 \left(y_t - B_0 - B_1 x_t\right)x_t,$$

$$\frac{\partial v}{\partial \Phi_r} = -2\sum_t \frac{\Phi(B)B^r}{\left\{\theta(B)\right\}^2}\left(y_t - B_0 - B_1 x_t\right)^2, \quad r = 1, 2$$

and

$$\frac{\partial v}{\partial \theta_s} = 2\sum_t \frac{\left\{\Phi(B)\right\}^2 B^s}{\left\{\theta(B)\right\}^3}\left(y_t - B_0 - B_1 x_t\right)^2, \quad s = 1, 2. \tag{16}$$

Following the approximations proposed by N u r i (1979) we obtain:

$$\left\{\frac{\Phi(B)}{\theta(B)}\right\}^2 \simeq 1 - 2\Phi_1 B - 2\Phi_2 B^2 + 2\theta_1 B + 2\theta_2 B^2 = A(B),$$

$$\frac{\Phi(B)B^r}{\{\theta(B)\}^2} \simeq B^r - \Phi_1 B^{r+1} - \Phi_2 B^{r+2} + 2\theta_1 B^{r+1} + 2\theta_2 B^{r+2}, \ r = 1, 2$$

and

$$\frac{\{\Phi(B)\}^2 B^s}{\{\theta(B)\}^3} \simeq B^s - 2\Phi_1 B^{s+1} - 2\Phi_2 B^{s+2} + 3\theta_1 B^{s+1} + 3\theta_2 B^{s+2}. \quad (17)$$

1. An initial estimator for $\underline{B} = \begin{pmatrix} B_o \\ B_1 \end{pmatrix}$, is obtained by the least squares method for the model $y_t = B_0 + B_1 x_t + e_t$.

2. Defining $z_t = (y_t - b_0 - b_1 x_t)^2$, the normal equations (16), take the form

$$nb_0 A(B) + b_1 \sum_t A(B)x_t = \sum_t A(B)y_t, \quad (18)$$

$$b_0 \sum_t A(B)x_t + b_1 \sum_t A(B)x_t^2 = \sum_t A(B)x_t y_t, \quad (19)$$

$$\Phi_1 \sum_t z_{t-r-1} + \Phi_2 \sum_t z_{t-r-2} - 2\theta_1 \sum_t z_{t-r-1} - 2\theta_2 \sum_t z_{t-r-2} = \sum_t z_{t-r}, \quad (20)$$

$$2\Phi_1 \sum_t z_{t-s-1} + 2\Phi_2 \sum_t z_{t-s-2} - 3\theta_1 \sum_t z_{t-s-1} - 3\theta_2 \sum_t z_{t-s-2} = \sum_t z_{t-s}, \quad s = 1, 2. \quad (21)$$

3. Solving equations (20) and (21) an initial estimator of $\underline{\alpha} = (\Phi_1 \Phi_2 \theta_1 \theta_2)^T$ is obtained, namely

$$\hat{\underline{\alpha}}^{(0)} = \left(\hat{\Phi}_1^{(0)} \hat{\Phi}_2^{(0)} \ \hat{\theta}_1^{(0)} \hat{\theta}_2^{(0)}\right)^T.$$

4. After obtaining the initial estimate of $\underline{\alpha}$, the sets of equations in (18) and (19) are employed to obtain a first approximate estimator $\underline{b}^{(1)}$ and so the first approximate estimator of $\underline{\alpha}^{(1)}$ is then obtained. And so on iteratively these steps are repeated until we have $\left|b^{(m)} - b^{(m-1)}\right| < \delta_1$ and $\left|\alpha^{(m)} - \alpha^{(m-1)}\right| < \delta_2$ for some specified numbers δ_1 and δ_2.

4. The Distribution of the Estimators

Define a random variable u_t such that $\Phi(B)u_t = a_t$, where $\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2$ and a_t is NID$(0, \sigma)^2$.

1. Following W o l d and M a n n (1943), $(\hat{\Phi}_1\ \hat{\Phi}_2)^T$ is asymptotically normal with mean$(\Phi_1\ \Phi_2)^T$ and variance-covariance matrix $\frac{\sigma}{n}^2 \Gamma^{-1}$, where $\Gamma = \begin{pmatrix} \Gamma_o & \Gamma_1 \\ \Gamma_1 & \Gamma_o \end{pmatrix}$ and $\Gamma_r = \mathrm{cov}(u_t, u_{t+r}) = E\, u_t u_{t+r}$, $r = 0, 1$.

2. Γ can be obtained alternatively as follows

$$u_t = \frac{1}{1 - \Phi_1 B - \Phi_2 B^2} a_t = \frac{1}{(1 - T_1 B)(1 - T_2 B)} a_t,$$

where $T_1 + T_2 = \Phi_1$ and $T_1 T_2 = -\Phi_2$, or

$$u_t = \frac{1}{T_2 - T_1}\left(\frac{T_2}{1 - T_2 B} - \frac{T_1}{1 - T_1 B}\right) a_t,$$

or

$$u_t = \frac{1}{T_2 - T_1} \sum_{j=0}^{\infty}\left(T_2^{j+1} - T_1^{j+1}\right) a_{t-j}, \tag{22}$$

Therefore $\Gamma_r = Eu_t u_{t+r}$

$$u_t = \frac{1}{\left(T_2 - T_1\right)^2} \sum_{j=0}^{\infty}\left(T_2^{j+1} - T_1^{j+1}\right)\left(T_2^{j+1} - T_1^{j+1}\right)^T Ea_{t-j} a_{t+r-j},$$

or

$$u_t = \frac{\sigma^2}{(T_2 - T_1)^2} \sum_{j=0}^{\infty}\left(T_2^{j+1} - T_1^{j+1}\right)\left(T_2^{j+r+1} - T_1^{j+r+1}\right) \tag{23}$$

substituting $r = 0, 1$, and after few steps we obtain

$$\Gamma_o = \frac{(1 - \Phi_2)\sigma^2}{(1 + \Phi_2)\left[(1 - \Phi_2)^2 - \Phi_1^2\right]} \tag{24}$$

and

$$\Gamma_1 = \frac{\Phi_1\sigma^2}{(1 + \Phi_2)\left[(1 - \Phi_2)^2 - \Phi_1^2\right]} \cdot \tag{25}$$

3. Similarly define the random variable v_t such that $(1 - \theta_1 B - \theta_2 B^2)v_t = a_t$, where $\hat{\underline{\theta}} = (\hat{\theta}_1 \quad \hat{\theta}_2)^T$ is also asymptotic normal with mean $\underline{\theta}$ and variance-covariance matrix $\frac{1}{2}\sigma^2\Omega^{-1}$, where $\Omega = \begin{pmatrix} \Omega_0 & \Omega_1 \\ \Omega_1 & \Omega_0 \end{pmatrix}$ such that $\Omega_s = Ev_t v_{t+s}$, $s = 0,1$; using the same way as for Γ, we obtain:

$$v_t = \frac{1}{\lambda_2 - \lambda_1} \sum_{j=0}^{\infty} \left(\lambda_2^{j+1} - \lambda_1^{j+1}\right) a_{t-j} \tag{26}$$

where $\lambda_1 + \lambda_2 = \theta_1$ and $\lambda_1 \lambda_2 = -\theta_2$

and

$$\Omega_s = \frac{\sigma^2}{(\lambda_2 - \lambda_1)^2} \sum_{j=0}^{\infty} \left(\lambda_2^{j+1} - \lambda_1^{j+1}\right)\left(\lambda_2^{j+s+1} - \lambda_1^{j+s+1}\right), \tag{27}$$

substituting $s = 0, 1$ we obtain

$$\Omega_0 = \frac{(1 - \theta_2)\sigma^2}{(1 + \theta_2)\left[(1 - \theta_2)^2 - \theta_1^2\right]} \tag{28}$$

and

$$\Omega_1 = \frac{\theta_1 \sigma^2}{(1 + \theta_2)\left[(1 - \theta_2)^2 - \theta_1^2\right]} \cdot \tag{29}$$

4. Define a matrix $w = \begin{pmatrix} w_0 & w_{-1} \\ w_1 & w_0 \end{pmatrix}$ such that

$w_k = \mathrm{cov}(u_t, v_{t+k}) = E\, u_t v_{t+k}$.

Now, for $k = 0$

$$w_0 = \frac{\sigma^2}{(T_2 - T_1)(\lambda_2 - \lambda_1)} \sum_{j=0}^{\infty} \left(T_2^{j+1} - T_1^{j+1}\right)\left(\lambda_2^{j+1} - \lambda_1^{j+1}\right).$$

And after few steps we obtain

$$w_0 = \frac{(1 - \theta_2\Phi_2)\,\sigma^2}{(1 - \Phi_2\theta_2)^2 - (\theta_1 + \Phi_1\theta_2)(\Phi_1 + \theta_1\Phi_2)}, \tag{30}$$

where $k = 1$,

$$w_1 = E\,u_t v_{t+1} = \frac{\sigma^2}{(T_2 - T_1)(\lambda_2 - \lambda_1)} \sum_{j=0}^{\infty} \left(T_2^{j+1} - T_1^{j+1}\right)\left(\lambda_2^{j+2} - \lambda_1^{j+2}\right), \tag{31}$$

or

$$w_1 = \frac{(\theta_1 + \Phi_1\theta_2)\sigma^2}{(1 - \Phi_2\theta_2)^2 - (\theta_1 + \Phi_1\theta_2)(\Phi_1 + \theta_1\Phi_2)}. \tag{32}$$

But, where $k = -1$, we have

$$w_{-1} = E u_t v_{t-1} = \frac{\sigma^2}{(T_2 - T_1)(\lambda_2 - \lambda_1)} \sum_{j=0}^{\infty} \left(T_2^{j+1} - T_1^{j+1}\right)\left(\lambda_2^{j} - \lambda_1^{j}\right),$$

or

$$w_{-1} = \frac{(\Phi_1 + \theta_1\Phi_2)\,\sigma^2}{(1 - \Phi_2\theta_2)^2 - (\theta_1 + \Phi_1\theta_2)(\Phi_1 + \theta_1\Phi_2)}, \tag{33}$$

comparing (32) with (33) it appears that $w_1 \neq w_{-1}$.

5. We obtain that

$$\underline{\hat{B}} = \underline{b} = (b_0 b_1)^T$$

is normally distributed (bivariate) with mean $\underline{B}$ and variance-covariance matrix $(\sigma^2/n)\,B^{-1}$, where

$$B = \left(\left(\frac{1-\Phi_1-\Phi_2}{1-\theta_1-\theta_2}\right)^2 \frac{1-\Phi_1-\Phi_2}{n(1-\theta_1-\theta_2)} \sum b_{it}\right) \frac{1}{n} \sum b_{it}^2$$

$\hat{\underline{\alpha}} = (\hat{\underline{\Phi}}\hat{\underline{\theta}})^T$ is asymptotic normal (4-dimensional) with mean $\underline{\alpha}$ and variance-covariance matrix $\frac{\sigma^2}{n}\begin{pmatrix} \Gamma & w \\ w^T & \Omega \end{pmatrix}^{-1}$ and consequently $(\hat{\underline{B}}\hat{\underline{\alpha}}\hat{6}^2)^T$ is asymptotic normal (7-dimen.) with mean $(\underline{B}\ \underline{\alpha}\ 6^2)^T$ and variance--covariance matrix:

$$\frac{2}{n}\begin{pmatrix} B & 0 & 0 & 0 \\ 0 & \Gamma & w & 0 \\ 0 & w^T & \Omega & 0 \\ 0 & 0 & 0 & \frac{1}{26^4} \end{pmatrix}^{-1}. \qquad (34)$$

References

Anderson R. L., 1942, Distribution of the Serial Correlation Coefficients, Ann. Math. Statist., vol. 13, p. 1-13.

Anderson R. L., 1954, The Problem of Autocorrelation in Regression Analysis, JASA, vol. 49, p. 113-129.

Box G. E. P. and Jenkins G. M., 1970, Time-Series Analysis Forecasting and Control, Holden-Day, London.

Box G. E. P. and Pierce D. A., 1970, Distribution of Residuals Autocorrelation in ARMA Time Series Model, JASA, vol. 65, p. 1509-1526.

Butter W. F., Kavesh R. A. and Platt R. B., 1974, Methods and Techniques of Business Forecasting, Prentice--Hall, Inc., New Jersey.

Cochran D. and Orcutt G. H., 1949, Application of Least Squares Regression to Relationships Counting Autocorrelation Error Terms, JASA, vol. 44, p. 32-61.

Durbin J. and Watson G. S., 1950, 1951, Testing for Serial Correlation in Least Squares Regression, Pts I and II, Biometrica, p. 409-428 and 159-177.

D u r b i n J., 1960, Estimation of Parameters in Time Series Regression Models, JRSS, vol. 22, p. 139-153.

D u r b i n J., 1970, Testing for Serial Correlation in Least Squares Regression when Some of the Regressors Are Lagged Dependent Variables, "Econometrica", vol. 38, p. 410-421.

G r e n a n d e r U., 1954, On the Estimation of Regression Coefficients in the Case of Autocorrelated Disturbance, Ann. Math. Statist., vol. 25, p. 252-272.

G u r l a n d J., 1954, An Example of Auto-Correlated Disturbances in Linear Regression, "Econometrica", vol. 22, p. 218-227.

H a n n a n E. J., 1957, Testing for Serial Correlation in Least Squares Regression, "Biometrica", vol. 44, p. 57-66.

N u r i W. A., 1979, Estimation in Regression Models with Errors Following an ARMA Scheme, International Time Series Meeting at Nottingham University.

P i e r c e D. A., 1971, Distribution of Residuals Autocorrelation in the Regression Model with ARMA Error, JRSS, B. 33, pp. 140-146.

Q u e n o u i l l e M. H., 1949, Approximate Tests of Correlation in Time Series, JRSS, B. 11, p. 68.

T h e i l H. and N a g a r A. L., 1961, Testing Independence of Regression Disturbances, JASA, p. 793-806.

W a t s o n G. S. and H a n n a n E. J., 1955, Serial Correlation in Regression Analysis, "Biometrica", vol. 43, p. 436-448.

W i s e J., 1956, Regression Analysis of Relationships Between Autocorrelated Time Series, JRSS, B. 18, p. 240-256.

W o l d H. and M a n n H. B., 1943, On the Statistical Treatment of Linear Stochastic Difference Equations, "Econometrica", vol. 11, p. 173-220.

Andrzej Tomaszewicz, Abdul Majid Hamza Al-Nasir

O ESTYMACJI PARAMETRÓW LINIOWYCH MODELI ZE SKŁADNIKAMI LOSOWYMI TYPU ARMA NISKICH RZĘDÓW

W badaniach empirycznych istnieją zwykle podstawy do założenia, że badany szereg czasowy generowany jest przez "mieszany" proces stochastyczny będący sumą procesu autoregresyjnego i procesu średnich ruchomych ARMA (B o x, J e n k i n s 1970); w niniejszej pracy zanalizowano niektóre własności modeli typu ARMA niskich rzędów, szczególnie procesu ARMA (2,2).