

Bias correction in ARMA models

Gauss M. Cordeiro

Universidade Federal de Pernambuco, Departamento de Economia, Cidade Universitária, Recife, PE, Brazil

Ruben Klein

Laboratório Nacional de Computação Científica, Botafogo, Rio de Janeiro, RJ, Brazil

Received May 1992

Revised May 1993

Abstract: We give a general matrix formula for computing the bias of the exact unconditional maximum likelihood estimate in ARMA models, with known and unknown mean, up to order $1/n$, where n is the length of the series. Some illustrative examples are presented.

Keywords: ARMA model; bias correction; maximum likelihood estimate.

1. Introduction

In this paper we derive a general matrix formula for computing second-order biases of the exact unconditional maximum likelihood (MLEs) in ARMA models with known and unknown mean. The formula depends only on the covariance matrix structure V , its inverse V^{-1} and the first and second partial derivatives of either V or V^{-1} . Expressions for V or V^{-1} for general ARMA models are given, for instance, in Shaman (1973), Galbraith and Galbraith (1974) and Ljung and Box (1979).

It is possible, but algebraically hard, to find explicit expressions for the second-order biases in terms of the parameters in particular cases. The use of an algebraic computation package like REDUCE (Hearn, 1984) or MAPLE (see Char et al., 1988) is very helpful. We have found *such* expressions for the AR(1), AR(2), and MA(1) processes and they appear to work properly. But for particular values of the parameters, it is quite easy to compute the biases numerically via a language supporting numerical linear algebra such as GAUSS, APL or Splus. For doing this, it is sufficient to find V , V^{-1} and the first two partial derivatives of either one with respect to the parameters of the model. We also shown that the bias of the MLE up to order $1/n$ for the unknown mean case decomposes in the sum of the bias up to order $1/n$ of the MLE for the known mean case plus a term that also depends only on V , V^{-1} and the first two partial derivatives of either one with respect to the parameters. It is shown that the MLE for the unknown mean is unbiased up to order $1/n$. In this way it becomes clear how the bias depends on whether or not the mean is estimated.

Tanaka (1984) computed these biases for the cases of known and unknown mean, through the use of asymptotic expansions associated with the MLEs in ARMA models. He has no such decomposition of

Correspondence to: Dr. R. Klein, Laboratorio Nacional de Computação Científica, Rua Lauro Muller 455, 22290-160 Rio de Janeiro, RJ, Brazil.

the bias. We believe our method is much simpler than his and has great advantages for numerical purposes.

Bias for other estimators have been studied in the literature. In particular, Shaman and Stine (1988) studied the bias of least-squares and Yule-Walker estimators for autoregressive processes, and Marriot and Pope (1954) had already shown for the AR(1) case, that the bias of the least squares estimates depends on whether or not the mean is estimated.

Godolphin and de Gooijer (1982) and Magee (1989) present small simulation studies for the MA(1) and AR(1) models respectively, where the bias for the MLE is estimated from the simulation.

2. A general matrix expression for the bias

Let $\ell(\beta)$ be the total log likelihood for some vector β of unknown parameters of dimension p . We assume that $\ell(\beta)$ is regular with respect to all β derivatives up to and including those of third order. We define joint cumulants of the log likelihood derivatives by $k_{ij} = E(\partial^2 \ell / \partial \beta_i \partial \beta_j)$, $k_{ijl} = E(\partial^3 \ell / \partial \beta_i \partial \beta_j \partial \beta_l)$ and $k_{ijl}^{(l)} = E\{(\partial^2 \ell / \partial \beta_i \partial \beta_j)(\partial \ell / \partial \beta_l)\}$ for $i, j, l = 1, \dots, p$. Furthermore, we define the following cumulant derivative $k_{ij}^{(l)} = \partial k_{ij} / \partial \beta_l$ for $i, j, l = 1, \dots, p$. The total Fisher Information matrix of order p for β is $K = \{-k_{ij}\}$ and let $K^{-1} = \{-k^{ij}\}$ be its inverse. All k 's are assumed to be $O(n)$, where n is the sample size.

Cox and Snell (1968) showed that for independent, but not necessarily identically distributed observations, the bias b of the MLE $\hat{\beta}$ of β reduces to

$$b_s = E(\hat{\beta}_s - \beta_s) = \sum_{i,j,l=1}^p k^{si} k^{jl} \left\{ \frac{1}{2} k_{ijl} + k_{ij,l} \right\} + O(n^{-2}) \quad (1)$$

for $s = 1, \dots, p$. Recently, Cordeiro and McCullagh (1991) using (1) derived general formulae for second-order biases of MLEs of all parameters in generalized linear models. We now apply formula (1) for ARMA models. It is important to point out that (1) continues to be valid for non-independent observations provided that all k 's are $O(n)$, as is the case for ARMA models. This follows by examining the orders of the terms in the Taylor series expansions and their expectations in Cox and Snell's (1968) proof.

We can rewrite (1) up to order n^{-1} as

$$b_s = \sum_{i=1}^p k^{si} \sum_{j,l=1}^p \left\{ k_{ij}^{(l)} - \frac{1}{2} k_{ijl} \right\} k^{jl},$$

for $s = 1, \dots, p$.

In matrix notation, the vector b is then

$$b = E(\hat{\beta} - \beta) = K^{-1} A \text{vec}(K^{-1}) + O(n^{-2}) \quad (2)$$

where $\text{vec}(K^{-1})$ is the vector obtained by stacking the columns of K^{-1} and where $A = \{A^{(1)} | \dots | A^{(p)}\}$, with $A^{(l)} = \{a_{ij}^{(l)}\}$ having its (i, j) th element defined by $a_{ij}^{(l)} = k_{ij}^{(l)} - \frac{1}{2} k_{ijl}$ for $l = 1, \dots, p$. The bias corrected MLE is given by $\tilde{\beta} = \hat{\beta} - \hat{K}^{-1} \hat{A} \text{vec}(\hat{K}^{-1})$, where \hat{K}^{-1} and \hat{A} are the matrices K^{-1} and A computed at $\hat{\beta}$, and would be expected to have better sampling properties than the uncorrected $\hat{\beta}$.

3. Bias of the MLE in ARMA models

An ARMA(p, q) model is defined by

$$y_t - \mu - \alpha_1(y_{t-1} - \mu) - \dots - \alpha_p(y_{t-p} - \mu) = u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q},$$

where the u_i 's are independent random variables with mean μ and variance σ^2 and $y = (y_1 \dots y_n)^T$ is the observed time series of length n . Let $\beta = (\alpha_1, \dots, \alpha_p, \theta_1, \dots, \theta_q, \sigma^2)^T$ be the $(p+q+1)$ vector of unknown parameters in the known mean case and let $\beta^* = (\beta^T, \mu)^T$ be the $(p+q+2)$ vector of unknown parameters in the unknown mean case.

The log likelihood $\ell(\beta)$ for β given y is

$$\ell(\beta) = -\frac{1}{2}n \log(2\pi) - \frac{1}{2}n \log(\sigma^2) - \frac{1}{2} \log(|V|) - ((y - \mu 1_n)^T V^{-1} (y - \mu 1_n)) / 2\sigma^2, \quad (3)$$

where $\sigma^2 V$ is the covariance matrix of y ,

$$V = V(\alpha_1, \dots, \alpha_p, \theta_1, \dots, \theta_q)$$

and 1_n is the n -dimensional vector with all entries equal to 1.

We define the derivatives, for $i, j = 1, \dots, p+q$, $V_i = \partial V / \partial \beta_i$, $V^i = \partial V^{-1} / \partial \beta_i$, $V_{ij} = \partial^2 V / \partial \beta_i \partial \beta_j$ and $V^{ij} = \partial^2 V^{-1} / \partial \beta_i \partial \beta_j$ and adopt the following notation:

$$\begin{aligned} \tilde{V}_i &= V^{-1} V_i, & \tilde{V}^i &= V^i V = -\tilde{V}_i, & \tilde{V}_{ij} &= V^{-1} V_{ij}, & \tilde{V}^{ij} &= V^{ij} V, \\ m_i &= \text{tr}(\tilde{V}_i), & m^i &= \text{tr}(\tilde{V}^i) = -m_i, & m_{ij} &= \text{tr}(\tilde{V}_{ij}), & m^{ij} &= \text{tr}(\tilde{V}^{ij}), \\ m_{i,j} &= \text{tr}(\tilde{V}_i \tilde{V}_j), & m^{i,j} &= \text{tr}(\tilde{V}^i \tilde{V}^j), \\ m_{ij,l} &= \text{tr}(\tilde{V}_{ij} \tilde{V}_l), & m^{ij,l} &= \text{tr}(\tilde{V}^{ij} \tilde{V}^l), & m_{i,j,l} &= \text{tr}(\tilde{V}_i \tilde{V}_j \tilde{V}_l), & m^{i,j,l} &= \text{tr}(\tilde{V}^i \tilde{V}^j \tilde{V}^l). \end{aligned}$$

The m 's defined above satisfy certain equations which facilitate their calculations. For example, $m_{ij} = 2m^{i,j} - m^{ij}$, $m^{ij} = 2m_{i,j} - m_{ij}$, $m_{ri,j} = m^{r,i,j} - m^{r,i,j} - m^{r,j,i}$, $m_{i,j,r} = -m^{i,j,r}$ and so on.

Differentiating (3) and making use of the relations $V^i = -V^{-1} V_i V^{-1}$, $\partial \log |V| / \partial \beta_i = \text{tr}(V^{-1} V_i)$, together with the multiplicative rule for differentiation and the equation $E((y - \mu 1_n)^T G (y - \mu 1_n)) = \sigma^2 \text{tr}(VG)$, for any positive definite matrix G (see Press, 1982 and Mardia and Marshall, 1984), we find after some algebra the following.

Case 1. Known mean.

$$\begin{aligned} K &= \frac{1}{2} \begin{pmatrix} (m_{i,j}) & (m_i/\sigma^2) \\ (m_j/\sigma^2) & n/\sigma^4 \end{pmatrix}_{1 \leq i,j \leq p+q} \\ &= \frac{1}{2} \begin{pmatrix} (m^{i,j}) & (-m^i/\sigma^2) \\ (-m^j/\sigma^2) & n/\sigma^4 \end{pmatrix}_{1 \leq i,j \leq p+q}, \end{aligned} \quad (4)$$

$$\begin{aligned} A^{(l)} &= -\frac{1}{4} \begin{pmatrix} (m_{li,j} + m_{lj,i} - m_{ij,l}) & m_{li}/\sigma^2 \\ (m_{lj}/\sigma^2) & 2m_l/\sigma^4 \end{pmatrix}_{1 \leq i,j \leq p+q} \\ &= -\frac{1}{4} \begin{pmatrix} (m^{li,j} + m^{lj,i} + m^{ij,l} - 2m^{i,j,l}) & ((2m^{l,i} - m^{li})/\sigma^2) \\ ((2m^{l,j} - m^{lj})/\sigma^2) & -2m^l/\sigma^4 \end{pmatrix}_{1 \leq i,j \leq p+q} \end{aligned} \quad (5)$$

for $l = 1, \dots, p+q$, and

$$\begin{aligned} A^{(\sigma^2)} &= A^{(p+q+1)} = -\frac{1}{4} \begin{pmatrix} ((2m_{i,j} - m_{ij})/\sigma^2) & 0 \\ 0 & 0 \end{pmatrix}_{1 \leq i,j \leq p+q} \\ &= -\frac{1}{4} \begin{pmatrix} (m^{ij}/\sigma^2) & 0 \\ 0 & 0 \end{pmatrix}_{1 \leq i,j \leq p+q} \end{aligned} \quad (6)$$

Case 2. Unknown mean. In this case $\beta^* = (\beta^T, \mu)^T$ is the $(p + q + 2)$ vector of unknown parameters and

$$b^* = (b_{\beta}^{*T}, b_{\mu}^*)^T = E(\hat{\beta}^* - \beta^*) = K^{*-1} A^* \text{vec}(K^{*-1}) + O(n^{-2}),$$

$$K^* = \begin{pmatrix} K & O \\ O & 1_n^T V^{-1} 1_n / \sigma^2 \end{pmatrix} \quad (7)$$

$$A^* = (A^{*(1)} | \dots | A^{*(p+q)} | A^{*(\sigma^2)} | A^{*(\mu)}),$$

$$A^{*(l)} = \begin{pmatrix} A^{(l)} & O \\ O & 1_n^T V^{-1} 1_n / 2\sigma^2 \end{pmatrix}, \quad l = 1, \dots, p + q, \quad (8)$$

$$A^{*(\sigma^2)} = \begin{pmatrix} A^{(\sigma^2)} & O \\ O & -1_n^T V^{-1} 1_n / 2\sigma^4 \end{pmatrix} \quad (9)$$

and

$$A^{*(\mu)} = \begin{pmatrix} O & a^* \\ a^{*T} & O \end{pmatrix} \quad (10)$$

where

$$a^* = (1_n^T V^{-1} 1_n / 2\sigma^2, \dots, 1_n^T V^{p+q} 1_n / 2\sigma^2, -1_n^T V^{-1} 1_n / 2\sigma^4)^T \quad (11)$$

is a $(p + q + 1)$ -dimensional vector, and K , A^l , $l = 1, \dots, p + q$, $A^{(\sigma^2)}$ are as in Case 1. Then

$$b_{\mu}^* = 0 + O(n^{-2}) \quad (12)$$

and

$$b_{\beta}^* = b + c + O(n^{-2}) \quad (13)$$

where

$$b = K^{-1} A \text{vec}(K^{-1}) \quad (14)$$

is the bias in the known mean case and

$$c = \sigma^2 (1_n^T V^{-1} 1_n)^{-1} K^{-1} a^* \quad (15)$$

is the added term for the bias in the unknown mean case.

4. Computation of the biases

Formula (2) in conjunction with (4)–(15) may be implemented in a computer algebra system to obtain bias-corrected MLEs in ARMA models to order n^{-1} with minimal effort. In fact, for an ARMA(p , q) model, we know V and V^{-1} , and we can find V^i , V_i , V_{ij} or V , V^1 , V^{ij} . Furthermore, we can evaluate these matrices at a particular value of β , for example, at $\hat{\beta}$. Thus, we can calculate, at least numerically, all the \tilde{V} 's and m 's given in Section 3. Therefore, in the known mean case, we can obtain K and A from (4)–(6). Inverting K and using (2), we obtain the biases by simple matrix multiplication. For the unknown mean case, we just have to add to b the second term, c , in (13).

5. Application to special models

To apply algebraically formula (2) to a specific example involves a lot of work. In the examples below, we have used the algebraic computing packages REDUCE and MAPLE to verify some computations and to do some others. Let us consider the cases with known and unknown mean.

Example 1. A stationary AR(1). The model is given by

$$y_t - \mu - \alpha(y_{t-1} - \mu) = u_t, \quad |\alpha| < 1, \quad t = 1, \dots, n.$$

The matrix V has a very simple form

$$V = (1 - \alpha^2)^{-1} \begin{pmatrix} 1 & \alpha \cdots \alpha^{n-1} \\ \alpha^{n-1} & \alpha^{n-2} \cdots 1 \end{pmatrix}.$$

and, therefore, we can compute the biases of $\hat{\alpha}$ and $\hat{\sigma}^2$ to order n^{-1} from (2) and (4)–(15).

Case 1. Known mean case. We have

$$b_\alpha = E(\hat{\alpha} - \alpha) = \frac{-2\alpha(1 - \alpha^2)[n^3(1 - \alpha^2) - 2n^2(1 - 2\alpha^2) + n(1 - 5\alpha^2) + 2\alpha^2]}{[n^2(1 - \alpha^2) - n(1 - 3\alpha^2) - 2\alpha^2]^2}$$

and

$$b_{\sigma^2} = E(\hat{\sigma}^2 - \sigma^2) = \frac{-n(1 - \alpha^2)\sigma^2}{n^2(1 - \alpha^2) - n(1 - 3\alpha^2) - 2\alpha^2}.$$

We can approximate further, up to $O(n^{-1})$, and the last two equations become

$$b_\alpha = E(\hat{\alpha} - \alpha) = -2\alpha/n \quad \text{and} \quad b_{\sigma^2} = E(\hat{\sigma}^2 - \sigma^2) = -\sigma^2/n.$$

Case 2. Unknown mean case. We have after approximations, up to $O(n^{-1})$,

$$c_\alpha = \frac{-(1 + \alpha)}{n} \quad \text{and} \quad c_{\sigma^2} = \frac{-\sigma^2}{n}$$

and the biases up to $O(n^{-1})$ are given by

$$b_\alpha^* = E(\hat{\alpha} - \alpha) = -\frac{2\alpha}{n} - \frac{1 + \alpha}{n} = -\frac{1 + 3\alpha}{n}$$

and

$$b_{\sigma^2}^* = E(\hat{\sigma}^2 - \sigma^2) = -\frac{\sigma^2}{n} - \frac{\sigma^2}{n} = -\frac{2\sigma^2}{n}.$$

These results are in agreement with the results obtained by Tanaka (1984, p. 66).

Example 2. An invertible MA(1). The model is

$$y_t - \mu = u_t - \theta u_{t-1}, \quad |\theta| < 1, \quad t = 1, \dots, n.$$

Shaman (1973) gave a closed-form expression for V^{-1} for a moving average process of order q . Using (2) and (4)–(15), we can compute the biases of $\hat{\theta}$ and $\hat{\sigma}^2$.

Case 1. Known mean case. We have

$$b_\theta = E(\hat{\theta} - \theta) = \theta(1 - \theta^2)(1 - \theta^{2n+2})A/B^2$$

and

$$b_{\sigma^2} = E(\hat{\sigma}^2 - \sigma^2) = -\sigma^2 n(1 - \theta^2)(1 - \theta^{2n+2})^2 / B,$$

where

$$\begin{aligned} A = n^4 & \left[2\theta^{2n}(1 - \theta^2)^2(\theta^{2n+2} - 2\theta^{2n} + 1) \right] \\ & - n^3 \left[(1 - \theta^2)(\theta^{6n+4} - 2\theta^{6n+2} + 3\theta^{4n+4} - 12\theta^{4n+2} + 10\theta^{4n} + 6\theta^{2n+2} - 5\theta^{2n} - 1) \right] \\ & - n^2 \left[(1 - \theta^{2n})(1 - \theta^{2n+2})(\theta^{2n+4} + 3\theta^{2n+2} - 8\theta^{2n} + 5\theta^2 - 1) \right] \\ & + n \left[2(1 - \theta^{2n})^2(5\theta^{2n+2} - 4\theta^2 - 1) \right] - 4\theta^2(1 - \theta^{2n})^3 \end{aligned}$$

and

$$\begin{aligned} B = n^3 & \left[2\theta^{2n}(1 - \theta^2)^2 \right] - n^2 \left[(1 - \theta^2)(\theta^{4n+2} + 3\theta^{2n+2} - 3\theta^{2n} - 1) \right] \\ & + n \left[(1 - \theta^{2n})(\theta^{2n+4} + 3\theta^{2n+2} - 3\theta^2 - 1) \right] - 2\theta^2(1 - \theta^{2n})^2. \end{aligned}$$

If $\theta^{2n} \cong O$, we conclude, ignoring terms of order less than n^{-1} , that

$$b_{\theta} = E(\hat{\theta} - \theta) = \frac{\theta(1 - \theta^2)[n^3(1 - \theta^2) + n^2(1 - 5\theta^2) - 2n(1 + 4\theta^2) - 4\theta^2]}{[n^2(1 - \theta^2) - n(1 + 3\theta^2) - 2\theta^2]^2}$$

and

$$b_{\sigma^2} = E(\hat{\sigma}^2 - \sigma^2) = \frac{-\sigma^2 n(1 - \theta^2)}{n^2(1 - \theta^2) - n(1 + 3\theta^2) - 2\theta^2}.$$

Thus, to $O(n^{-1})$, it follows

$$b_{\theta} = E(\hat{\theta} - \theta) = \theta/n \quad \text{and} \quad b_{\sigma^2} = E(\hat{\sigma}^2 - \sigma^2) = -\sigma^2/n.$$

Case 2. Unknown mean case. We have after approximations, up to $O(n^{-1})$,

$$c_{\theta} = \frac{1 + \theta}{n} \quad \text{and} \quad c_{\sigma^2} = \frac{-\sigma^2}{n}.$$

Then the biases, up to $O(n^{-1})$, are given by

$$b_{\theta}^* = E(\hat{\theta} - \theta) = \frac{\theta}{n} + \frac{1 + \theta}{n} = \frac{1 + 2\theta}{n}$$

and

$$b_{\sigma^2}^* = E(\hat{\sigma}^2 - \sigma^2) = -\frac{\sigma^2}{n} - \frac{\sigma^2}{n} = -\frac{2\sigma^2}{n}.$$

These results coincide with the results of Tanaka (1984, p. 66).

Example 3. A stationary AR(2). The model is given by

$$y_t - \mu - \alpha_1(y_{t-1} - \mu) - \alpha_2(y_{t-2} - \mu) = u_t, \quad t = 1, \dots, n,$$

where the roots of $\alpha(z) = 1 - \alpha_1 z - \alpha_2 z^2$ are such that $|z| > 1$. Hence using (2) and (4)–(15), we can compute up to $O(n^{-1})$, the biases of $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\sigma}^2$.

Case 1. Known mean case. We have

$$\begin{aligned} b_{\alpha_1} &= E(\hat{\alpha}_1 - \alpha_1) = -\alpha_1(\alpha_2 + 1) \left[n((\alpha_2 - 1)^2 - \alpha_1^2) - 2(3\alpha_2^2 + 1 - \alpha_1^2) \right] / D_1, \\ b_{\alpha_2} &= E(\hat{\alpha}_2 - \alpha_2) \\ &= -(\alpha_2 + 1) \left[n(3\alpha_2 + 1)((\alpha_2 - 1)^2 - \alpha_1^2) - 2\alpha_2(5\alpha_2^2 - 6\alpha_2 - 3\alpha_1^2 + 1) \right] / D_1, \\ b_{\sigma^2} &= E(\hat{\sigma}^2 - \sigma^2) = -\sigma^2 \left[2n^2((\alpha_2 - 1)^2 - \alpha_1^2)(\alpha_2 + 1) \right. \\ &\quad \left. - n((9\alpha_2^2 + 2\alpha_2 - 3)(\alpha_2 - 1) - (7\alpha_2 + 5)\alpha_1^2) + 8\alpha_2^2(\alpha_2 - 1) \right. \\ &\quad \left. - 2\alpha_1^2(3\alpha_2 + 1) \right] / D_2, \end{aligned}$$

where

$$\begin{aligned} D_1 &= n^2((\alpha_2 - 1)^2 - \alpha_1^2)(\alpha_2 + 1) - n(\alpha_2^2(7\alpha_2 - 5) - \alpha_1^2(5\alpha_2 + 3) + 1 - 3\alpha_2) \\ &\quad + 4\alpha_2^2(3\alpha_2 - 1) - 2\alpha_1^2(3\alpha_2 + 1) \end{aligned}$$

and

$$\begin{aligned} D_2 &= n^3((\alpha_2 - 1)^2 - \alpha_1^2)(\alpha_2 + 1) - n^2((9\alpha_2^2 + 2\alpha_2 - 3)(\alpha_2 - 1) - (7(\alpha_2 + 5)\alpha_1^2) \\ &\quad + 2n(13\alpha_2^3 - 7\alpha_2^2 - 3\alpha_2 + 1 - 4\alpha_1^2(2\alpha_2 + 1)) - 8\alpha_2^2(3\alpha_2 - 1) + 4\alpha_1^2(3\alpha_2 + 1)). \end{aligned}$$

To order n^{-1} we hence have

$$\begin{aligned} b_{\alpha_1} &= E(\hat{\alpha}_1 - \alpha_1) = -\alpha_1/n, \quad b_{\alpha_2} = E(\hat{\alpha}_2 - \alpha_2) = -(3\alpha_2 + 1)/n \quad \text{and} \\ b_{\sigma^2} &= E(\hat{\sigma}^2 - \sigma^2) = -2\sigma^2/n, \end{aligned}$$

Case 2. Unknown mean case. We have, after approximations, up to $O(n^{-1})$,

$$c_{\alpha_1} = \frac{-(1 + \alpha_2)}{n}, \quad c_{\alpha_2} = \frac{-(1 + \alpha_2)}{n} \quad \text{and} \quad c_{\sigma^2} = \frac{-\sigma^2}{n}.$$

Then the biases, up to $O(n^{-1})$, are given by

$$\begin{aligned} b_{\alpha_1}^* &= E(\hat{\alpha}_1 - \alpha_1) = -\frac{\alpha_1}{n} - \frac{1 + \alpha_2}{n} = -\frac{1 + \alpha_1 + \alpha_2}{n}, \\ b_{\alpha_2}^* &= E(\hat{\alpha}_2 - \alpha_2) = -\frac{1 + 3\alpha_2}{n} - \frac{1 + \alpha_2}{n} = -\frac{2(1 + 2\alpha_2)}{n} \end{aligned}$$

and

$$b_{\sigma^2}^* = E(\hat{\sigma}^2 - \sigma^2) = -\frac{2\sigma^2}{n} - \frac{\sigma^2}{n} = -\frac{3\sigma^2}{n}.$$

These results again agree with the results of Tanaka (1984, p. 66).

Acknowledgment

The authors thank Telma Suaiden Klein for the computations in MAPLE.

References

- Char, B.W., K.O. Geddes, G.H. Gonnet, M.B. Monagan and S.M. Watt (1988), *Maple Reference Manual* (WATCOM Publications Ltd., Waterloo, Canada, 5th ed.).
- Cordeiro, G.M. and P. McCullagh (1991), Bias correction in generalized linear models, *J.R. Statist. Soc. B* **53**, 629–643.
- Cox, D.R. and E.J. Snell (1968), A general definition of residuals, *J.R. Statist. Soc. B* **30**, 248–275.
- Galbraith, R.F. and J.I. Galbraith (1974), On the inverses of some patterned matrices arising in the theory of stationary time series, *J. Appl. Prob.* **11**, 63–71.
- Godolphin, E.J. and S.G. de Gooijer (1982), On the maximum likelihood estimation of the parameters of a Gaussian moving average process, *Biometrika* **69**, 443–451.
- Hearn, A.C. (ed.) (1984), *Reduce User's Manual*, version 3.1 (The Rand Corporation).
- Ljung, G.M. and G.E.P. Box (1979), The likelihood function of stationary autoregressive-moving average models, *Biometrika* **66**, 265–270.
- Magee, L. (1989), Bias approximation for covariance parameter estimators in the linear model with AR(1) errors, *Commun. Statist-Theory Meth.* **18**, 395–422.
- Mardia, K.V. and R.J. Marshall (1984), Maximum likelihood estimation of models for residual covariance in spatial regression, *Biometrika* **71**, 135–146.
- Marriot, F.H.C. and J.A. Pope (1954), Bias in the estimation of autocorrelations, *Biometrika* **41**, 390–402.
- Press, S.J. (1982), *Applied Multivariate Analysis Including Bayesian and Frequentist Methods of Inference* (Krieger Publishing Company, Melbourne, Florida).
- Shaman, P. (1973), On the inverse of the covariance matrix for an autoregressive-moving average process, *Biometrika* **60**, 193–196.
- Shaman, P. and R.A. Stine (1988), The bias of autoregressive coefficient estimators, *J. Amer. Statist. Assoc.* **83**, 842–848.
- Tanaka, K. (1984), An asymptotic expansion associated with the maximum likelihood estimators in ARMA models, *J.R. Statist. Soc. B* **46**, 58–67.