A weighted pairwise likelihood approach to multivariate AR(1) models

M. R. Kazemi and A. R. Nematollahi*

Department of Statistics, College of Sciences, Shiraz University, Shiraz 71454, Iran
E-mails: kazemimr88@hotmail.com & nematollahi@susc.ac.ir

Abstract

In this paper, the use of weighted pairwise likelihood instead of the full likelihood in estimating the parameters of the multivariate AR(1) is investigated. A closed formula for typical elements of the Godambe information (sandwich information) is presented. Some efficiency calculations are also given to discuss the feasibility and computational advantages of the weighted pairwise likelihood approach relative to the full likelihood approach.

Keywords: Weighted pairwise likelihood; multivariate time series; Godambe information

1. Introduction

The multivariate autoregression model is an essential and informative tool for describing the dynamic behavior of the applied sciences, e.g. neurosciences, biomedical research, business, economic, environmental, industrial and financial time series, see Schelter et al. and therein references for more examples and applications [1]. In multivariate models, the standard likelihood-based inference often involves high-dimensional integrals which become more difficult to evaluate when the sample size increases, Xu [2] and Parner [3]. In the case that the likelihood function is difficult to evaluate or cannot be solved, certain methods based on modifications of the likelihood are used by several authors, e.g. Besag [4, 5], Cox [6] and Lindsay [7]. The conditional Gaussian likelihood procedure proposed by Hannan and Rissanen [8] has been complemented by Salau [9] to investigate the effects of different choices of order for autoregressive approximation on the fully efficient parameter estimates for ARMA models. Asymptotic quasi-likelihood has been advanced by Heyde and Lin [10], Heyde [11] and Jung et al. [12]. A quasi-maximum likelihood unit root test is studied by Rothenberg and Stock [13] and Carstensen [14], and the quasi-maximum likelihood theory to simulate the likelihood ratio test is given by Lee [15] and Ziegler [16]. Parameters estimation by quasimaximum likelihood through the optimization of a Gaussian log-likelihood function is discussed in Bollerslev and Wooldridge [17], Newey and Steigerwald [18] and Rodrigues and Rubia [19]. Modified versions of the likelihood function to find an approximate sufficient statistic for the parameters of the vector-valued (multivariate) ARMA models, in terms of the periodogram are applied by Kharrati-Kopaei et al. [20]. The pseudo-likelihood based estimation procedure is also used by many authors. The reader is referred to Hu and Zhang [21] and the references therein for a short discussion on this subject.

In recent years, the composite likelihood approach (based on the likelihoods of low-dimensional marginal distributions) has received much attention in statistical models. The pairwise likelihood, given by Cox and Reid [22], takes the bivariate margins to produce the pseudo-likelihood. An excellent review on the composite likelihoods with emphasis on some applications in genetic, longitudinal data, survival analysis and spatial statistics can be found in Varin [23]. Recently, Ng et al. [24] have established the consistency and asymptotic normality properties in a time series model with a latent Gaussian autoregressive process for various composite likelihood estimators. The asymptotic properties of pairwise likelihood estimation procedures for linear time series models are also studied by Davis and Yau [25].

Pairwise likelihood approach in estimating the parameters of a multivariate AR(1) is studied by Nematollahi and Kazemi [26]. They concentrate on the maximum pairwise likelihood estimators (MPLE) and specify its efficiency for the arbitrary matrix of parameters of multivariate AR(1) model, along the line of the equi-correlation normal matrix example in Cox and Reid [22]. They have used unweighted pairwise likelihood and by some
tedious computations, it is shown that the efficiency can be low, so that using weighted pairwise likelihood-based inferences is recommended. In this paper, we discuss the feasibility of the weighted pairwise likelihood approach (Joe and Lee [27]) and discuss the computational advantages of weighted pairwise likelihood relative to the full likelihood approach. Practical advantages of the weighted MPLE (WMPLE) to MLE are brought to light and some efficiency calculations are presented.

This paper is organized as follows: The pairwise likelihood function for m-dimensional random vector is introduced Section 2. In Section 3, we consider the multivariate AR(1) models and discuss the computational advantages of the composite score equation provided by n observations is given by

$$G(\theta) = \mathbf{J}_{wpl}(\theta) \mathbf{K}_{wpl}(\theta)^{-1} \mathbf{J}_{wpl}(\theta)^T,$$

where $\mathbf{J}_{wpl}(\theta) = E\left(-\frac{\partial^2 \mathbf{S}_{wpl}(\theta, y)}{\partial \theta^2}\right)$ and $\mathbf{K}_{wpl}(\theta)$ is the variance of the composite score matrix given by

$$\mathbf{K}_{wpl}(\theta) = E[\mathbf{S}_{wpl}(\theta, y) \mathbf{S}_{wpl}(\theta, y)^T].$$

The standard theory for inference functions [28] can be applied to derive the general asymptotic properties of $\tilde{\theta}$. Under the theory of estimating functions and the regularity conditions on the log-likelihood, it can be shown that $\tilde{\theta}$ is consistent and asymptotically normal distributed with asymptotic mean $\theta_0$ and variance matrix

$$\mathbf{G}^{-1}(\theta_0) = \mathbf{J}_{wpl}^{-1}(\theta_0) \mathbf{K}_{wpl}(\theta_0) \mathbf{J}_{wpl}^{-1}(\theta_0),$$

which follows from the Taylor expansion to second order. A proof can be found in Godambe [28]. See also Zhao [29], Zhao and Joe [30] and Joe and Lee [27] for more details.

In this paper, we consider a multivariate AR(1) model which satisfies the difference equation

$$x_i = \Phi x_{i-1} + z_i,$$  \hspace{1cm} (2)

where $\Phi$ is $m \times m$ matrix and $z_i$ is a sequence of independent multivariate normal with zero mean vector and covariance matrix $\Sigma$. We assume that the process is causal in the sense that all eigenvalues of $\Phi$ are less than 1 in absolute value. From the well-known Yule-Walker equations, we have

$$\Gamma(h) = \Phi \Gamma^{-1},$$

for $h = 1, 2, 3, \ldots$ and

$$\text{vec}(\Gamma(0)) = (I - (\Phi \otimes \Phi))^{-1} \text{vec}(\Sigma)$$

where the "vec" notation for a matrix $A = [a_1, a_2, \ldots, a_n]^{T}$ is defined by $\text{vec}(A) = [a_1^{T}, a_2^{T}, \ldots, a_n^{T}]^{T}$ and operator $\otimes$ is the Kronecker product.

Joe and Lee [27] studied the $m$-dimensional Gaussian AR(1) model with mean $\mu I_m$ and covariance matrix $\eta^2 \mathbf{R}(\rho)$, where $\mathbf{R}(\rho) = [\rho^{j-k}]_{j,k=1}^{m}$ with $-1 < \rho < 1$. They
obtained some efficiency results for the estimate of one of the parameters assuming the other two to be known, using the weighted pairwise likelihoods with weights depending on lag. In this paper, we study weighted pairwise likelihoods in a more general case with the weights constructed based on the autoregressive property of the time series and considering the serial dependence within each component series and interdependence between the different component series. Using this kind of weights, we will show that the loss of efficiency compared to maximum likelihood estimation is acceptable and balanced by the computational ease.

3. The weighted pairwise likelihood inferences

Let \{X_1, \ldots, X_n\} be a set of \(m\)-dimensional observations, where \(X_t = \{X_{1t}, X_{2t}, \ldots, X_{mt}\}^T\), \(t = 1, \ldots, n\). Following Nematollahi and Kazemi [26], consider a \(mn \times 1\) vector \(X\) constructed by

\[
X = \left( X_1^T, X_2^T, \ldots, X_n^T \right)^T
\]

\[
= (X_{11}, X_{12}, \ldots, X_{1m}, X_{21}, X_{22}, \ldots, X_{2m}, \ldots, X_{nm}, X_{n1}, X_{n2}, \ldots, X_{nm})^T. \tag{3}
\]

The vector-valued (multivariate) time series \(X_i\) having not only serial dependence within each component series \(X_i\), but also interdependence between the different component series \(X_i\) and \(X_j\), \(i \neq j\). According to the autoregressive property, each \(X_i\) is also dependent on \(X_{i-1}\). So, we will consider the weighted pairwise likelihood for \(X\) as

\[
L_{wpl}(x, \theta) = \prod_{i=1}^{n} \prod_{j=1}^{m} f_{ij}(x_{ij}, x_{ij, \theta}) \prod_{i=1}^{n} \prod_{j=1}^{m} f_{ij}(x_{ij}, x_{ij, \theta})
\]

\[
= \left[ \prod_{i=1}^{n} \prod_{j=1}^{m} f_{ij}(x_{ij}, x_{ij, \theta}) \prod_{i=1}^{n} \prod_{j=1}^{m} f_{ij}(x_{ij}, x_{ij, \theta}) \right] \tag{4}
\]

where

\[
\pi_{ijc} = \begin{cases} 1, & (i = j, 1 \leq b < c \leq m) \text{or} (j = i+1, 1 \leq b, c \leq m) \\ 0, & \text{otherwise}. \end{cases}
\]

**Remark 1.** Note that in order to use all of the information in the time series, we have to write \(X\) and \(L_{wpl}(x, \theta)\) as above, since in the multivariate time series the vector of observations at different times are dependent and it is also dependent with other observations. In (4), the first term is the likelihood function of the observations within the \(X_n\) and the second term is the counterpart between groups \(\{X_n\}\) and \(\{X_{n+1}\}\), this is one justification for considering the weights \(\pi_{ijc}\) as given above. Another justification is the autoregressive property.

**Remark 2.** The pairwise likelihood function considered by Nematollahi and Kazemi [26] can be easily derived by setting \(\pi_{ijc} = 1\), for all \(i, j, b\) and \(c\).

Now, suppose that \(\{X_i, i = 1, 2, \ldots, n\}\) is a set of \(n\) mean-zero \(m\)-dimensional Gaussian time series with covariance matrix \(K(i, j) = E(X_iX_j^T)\), then for a \(mm\) -dimensional Gaussian vector \(X\), we can say

\[
\begin{pmatrix} X_{ib} \\ X_{jc} \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \gamma_{bb}(0) & \gamma_{bc}(i-j) \\ \gamma_{cb}(i-j) & \gamma_{cc}(0) \end{pmatrix} \right). \tag{5}
\]

Let

\[
M_{bc}(h) = \begin{pmatrix} \gamma_{bb}(0) & \gamma_{bc}(h) \\ \gamma_{cb}(h) & \gamma_{cc}(0) \end{pmatrix}
\]

then (5) can be written as

\[
\begin{pmatrix} X_{ib} \\ X_{jc} \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, M_{bc}(i-j) \right). \tag{6}
\]

Now, using (6) the logarithm of the weighted pairwise likelihood function \(L_{wpl}(x, \theta)\) reduces to

\[
\ln L_{wpl}(\theta) = \log(\ln L_{wpl}(x, \theta))
\]

\[
= -\frac{1}{2} \sum_{a=1}^{n} \sum_{b=1}^{m} \log[\mathbf{M}_{bb}(0)\mathbf{X}_{bb}^t(\mathbf{X}_{bb}(a))\mathbf{X}_{bb}(a)]
\]

\[
-\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} \log[\mathbf{M}_{bc}(-1)\mathbf{X}_{bc}(i)\mathbf{M}_{bc}(-1)\mathbf{X}_{bc}(i)] \tag{7}
\]

where \(\theta = (\Phi, \Sigma)\) is the unknown parameter of the model, \(X_{bc}(a) = \left( X_{ab}, X_{ac} \right)^T\) and \(X_{bc}(i) = \left( X_{ib}, X_{ic} \right)^T\).

Now let \(\Phi = \phi_{ij, 1 \leq j \leq m}\) and \(\Sigma = \sigma_{ij, 1 \leq j \leq m}\) and take

\[
w_1 = \phi_{11}, w_2 = \phi_{21}, \ldots, w_m = \phi_{m1}, \ldots, w_q = \sigma_{mm},
\]

where \(q = mm + (m(m+1)/2)\) is the number of unknown parameters. Then using (7), we have
S_{npl}(x,0) = \left[ \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_q \partial w_t} \right]_{\theta = 1}^{q},

J_{npl}(\theta) = \left[ E \left( -\frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} \right) \right]_{k,l=1}^{q},

and finally

K_{npl}(\theta) = \left[ E \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} \right) \right]_{k,l=1}^{q}.

To compute the above quantities, first note that

\frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} = \frac{1}{2} \sum_{l = 1}^{q} \sum_{x = 1}^{n} \left[ \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} \right) + X_{n}(a) \left[ \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} \right] X_{n}(a) \right],

\frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} = \frac{1}{2} \sum_{l = 1}^{q} \sum_{x = 1}^{n} \left[ \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} \right) + X_{n}(a) \left[ \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} \right] X_{n}(a) \right],

Now we can state the following theorem which plays the main role in computing the asymptotic variance-covariance matrix of the WMLE.

**Theorem 1.** The typical (k,l)th element of the matrices J_{npl}(\theta) and K_{npl}(\theta) have the form

J_{npl}(\theta)_{kl} = E \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_l} \right) = \frac{1}{2} \sum_{j = 1}^{q} \sum_{x = 1}^{n} \left[ \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_l} \right) + tr \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_l} \right) M_{n}(\theta) \right],

\frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} = \frac{1}{2} \sum_{l = 1}^{q} \sum_{x = 1}^{n} \left[ \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} \right) + tr \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_t} \right) M_{n}(\theta) \right],

and

K_{npl}(\theta)_{kl} = E \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_l} \right) = \frac{1}{2} \sum_{j = 1}^{q} \sum_{x = 1}^{n} \left[ \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_l} \right) + tr \left( \frac{\partial^2 \log \gamma_{npl}(\theta)}{\partial w_k \partial w_l} \right) M_{n}(\theta) \right],

respectively, where

N_{bc,i}(h) = \partial M_{bc}^{-1}(h) \partial w_i,

M_{bc}^{(0)}(h,h') = \begin{bmatrix} \gamma_{bb}(h-h') & \gamma_{bc}(h-h') \\ \gamma_{cb}(h-h') & \gamma_{cc}(h-h') \end{bmatrix},

M_{bc}^{(1)}(h,h') = \begin{bmatrix} \gamma_{bb}(h-h') & \gamma_{bc}(h-h'-1) \\ \gamma_{cb}(h-h') & \gamma_{cc}(h-h'-1) \end{bmatrix},

M_{bc}^{(2)}(h,h') = \begin{bmatrix} \gamma_{bc}(h+h') & \gamma_{bc}(h-h') \\ \gamma_{cb}(h+h') & \gamma_{cc}(h-h') \end{bmatrix},

M_{bc}^{(3)}(h,h') = \begin{bmatrix} \gamma_{bc}(h+h'-1) & \gamma_{bc}(h-h') \\ \gamma_{cb}(h+h'-1) & \gamma_{cc}(h-h') \end{bmatrix}.

**Proof:** The proof of (10) is simply done by using the following fact. If E(X) = 0, then

E(X'AX) = tr(\Omega), where \Omega = E(XX').

To prove (11), at first note that,

E(\partial^2 \sum_{i = 1}^{\varphi} \partial w_i \partial w_i) = \pi_1 + \pi_2 + \pi_3 + \cdots + \pi_{16},

where the equations \pi_1 up to \pi_{16} can be divided to 4 groups as below. Note that for the four following groups (q,r) varies in \{(0,0),(0,-1),(-1,0),(-1,-1)\} according to s, varying from 1 up to 4 for the first group, from 5 up to 8 for the second group, from 9 up to 12 for the third group and from 13 up to 16 for the fourth group, respectively.

Group1: for \pi_s, s = 1,2,3,4 we have,

\pi_s = \frac{1}{4} \sum_{q,b,c} \sum_{r,b,c} \left[ \frac{\partial \log M_{bc}(q)}{\partial w_k} \right] \left[ \frac{\partial \log M_{bc}(r)}{\partial w_l} \right].

Group2: for \pi_s, s = 5,6,7,8 we have,

\pi_s = \frac{1}{4} \sum_{q,b,c} \sum_{r,b,c} \left[ \frac{\partial \log M_{bc}(q)}{\partial w_k} \right] \left[ \frac{\partial \log M_{bc}(r)}{\partial w_l} \right].

Group3: for \pi_s, s = 9,10,11,12 we have,

\pi_s = \frac{1}{4} \sum_{q,b,c} \sum_{r,b,c} \left[ \frac{\partial \log M_{bc}(q)}{\partial w_k} \right] \left[ \frac{\partial \log M_{bc}(r)}{\partial w_l} \right].

Group4: for \pi_s, s = 13,14,15,16 we have,
\[
\pi_s = \frac{1}{4} \sum_{q=1}^{n} \sum_{b,c} \sum_{b',c'} \frac{1}{\sqrt{2\pi}} e^{-\frac{(X_k)_b^{(q)} - \frac{1}{2} tr \left( \frac{M_{b,c}^{(q)}}{\delta w_q} \right) M_{b,c}^{(q)(r)}}{2}}.
\]

Note that in the all of the above groups, if \( i_q \) varies from 1 up to \( n-1 \) then both \( b,c \) vary from 1 up to \( m \) and if \( i_q \) varies from 1 to \( n \) then \( b,c \) vary from 1 to \( b < c \leq m \).

For simplifying these expressions, we have to use the following useful facts.

(I) If \( E(X) = 0 \) then \( E(X^T A X) = tr (A \Omega) \)

(II) \( \frac{\partial \log |V|}{\partial w} = tr \left( V \left( \frac{\partial V}{\partial w} \right)^{-1} \right) \) and so \( \frac{\partial \log |V|}{\partial w} = - \frac{\partial \log |V|}{\partial w} \).

(III) For simplifying \( \pi_{13}, \ldots, \pi_{16} \) we have used the following well-known result (Searle, [31]). Let \( X_i \sim N(0, C_{ij}) \) and \( C_{ij} = \Sigma (X_i X_i^T) \) then,

\[
\det \Sigma (a)_{ij} = \det \Sigma (X_i)_{ij} \Sigma (a)_{ij} = \det (A_{ij} C_{ij} A_{ij}^T) + A_{ij} C_{ij} A_{ij}^T).
\]

Using (I) and (II) it can be easily shown that,

\[
\pi_1 + \pi_2 = \pi_3 + \pi_4 = \pi_5 + \pi_6 = \pi_7 = \pi_8 = 0.
\]

And also by using (I) and (II) we can have,

\[
\pi_s = -\frac{f(n)}{4} \sum_{b,c} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} tr \left( \frac{M_{b,c}^{(q)}}{\delta w_q} \right) M_{b,c}^{(q)(r)}}.
\]

for \( s = 9,10,11,12 \). If \( s = 9 \) and \( (q,r) = (0,0) \) then \( f(n) = n^3 \), and if \( s = 11 \) and \( (q,r) e \{0,-1,(-1,0)\} \) then \( f(n) = n(n-1)^2 \), and if \( s = 12 \) and \( (q,r) = (-1,-1) \) then \( f(n) = (n-1)^3 \).

Recall \( N_{h,a}(h) \cdot M_{b,c}^{(q)}(h,a) \cdot M_{b,c}^{(r)}(h,a) \) and \( M_{b,c}^{(q)}(h,a) \) as in Theorem 1 and using results (I), (II) and (III) we have,

\[
\pi_{13} = \sum_{a=1}^{n} \sum_{b=1}^{m} \sum_{c=1}^{m} \sum_{b'=1}^{m} \sum_{c'=1}^{m} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} tr \left( \frac{M_{b,c}^{(q)}}{\delta w_q} \right) M_{b,c}^{(q)(r)}}.
\]

Remark 3. The asymptotic variance-covariance matrix can now be derived, by using the relations (10) and (11) via (1).

Remark 4. For \( \pi_{ij} = 1 \), for all \( i,j \) and \( b \), (10) and (11) reduce to the formulas obtained by Nematollahi and Kazemi [26].

4. Simulation studies

In the following section we have conducted some simulation studies to study the asymptotic relative efficiency of the WMPLE with respect to the MLE (ARE). The optimization method for the weighted pairwise likelihood function is the quasi-Newton method, with a relative tolerance of \( 10^{-8} \), implemented by the \textit{fmincon} function of the software package Matlab 2009a, v7.8. The \textit{fmincon} function is a numerical optimization routine, including sequential quadratic programming algorithm to solve for constrained optima.

This section consists of three parts. In the first part, we assume that \( m = 2 \) and the variance-covariance matrix \( \Sigma \) is known and is given by

\[
\Sigma = \begin{bmatrix}
0.71 & 0.01 \\
0.01 & 0.71
\end{bmatrix}
\]

The model (2) is then written as

\[
\begin{bmatrix}
X_{11} \\
X_{12}
\end{bmatrix} = \begin{bmatrix}
\phi_1 & \phi_2 \\
\phi_1 & \phi_2
\end{bmatrix} \begin{bmatrix}
X_{1,-1} \\
X_{1,-2}
\end{bmatrix} + \begin{bmatrix}
Z_{11} \\
Z_{12}
\end{bmatrix},
\]

We compare the variance of WMPLE computed by (3.8) and (11) with respect to MLE. We will
study the relative efficiency of WPMLE with respect to MLE in terms of $\theta_1$ by letting $\phi_{12} = \phi_{21} = 0, \phi_{22} = 0.9$ and $\phi_{11}$ varying in the set $0.1, 0.2, ..., 0.9$.

The condition required for the existence of a causal stationary solution of the multivariate AR(1) model is all the eigenvalues of $\Phi$ are less than 1 in absolute value, i.e., provided that

$$\det(I - z\Phi) \neq 0,$$

(12)

for all complex values of $z$, such that $|z| \leq 1$. For the case $m = 2$, this causality condition, which ensure us the existence of a stationary solution of the considered model can be written as

$$\det\begin{pmatrix} 1 - z\phi_{11} & -z\phi_{12} \\ -z\phi_{21} & 1 - z\phi_{22} \end{pmatrix} \neq 0,$$

for all complex values of $z$, such that $|z| \leq 1$. In the first example, this condition reduces to $|\phi_{11}| < 1$ and $|\phi_{22}| < 1$, which are satisfied for $\phi_{12} = \phi_{21} = 0, \phi_{22} = 0.9$ and $\phi_{11}$ varying in the set $0.1, 0.2, ..., 0.9$.

To compute the asymptotic variance of the MLE of matrix of coefficients $\Phi$, we use the matrix $1/n(\Sigma \otimes \Gamma(\theta_{11}))$, see Reinsel [32] and Shumway and Stoffer [33] for more details.

In Table 1, for various values of $\phi_{11}$, the asymptotic variance of the WMPLE (avar(\(\tilde{\phi}_{11}\))), MLE (avar(\(\hat{\phi}_{11}\))), and ARE of the WMPLE with respect to the MLE for $n=100$ are reported.

<table>
<thead>
<tr>
<th>$\phi_{11}$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>avar((\tilde{\phi}_{11}))</td>
<td>0.0175</td>
<td>0.0171</td>
<td>0.0165</td>
<td>0.0155</td>
<td>0.0141</td>
<td>0.0124</td>
<td>0.0103</td>
<td>0.0076</td>
<td>0.0042</td>
</tr>
<tr>
<td>avar((\hat{\phi}_{11}))</td>
<td>0.0226</td>
<td>0.0312</td>
<td>0.0339</td>
<td>0.0321</td>
<td>0.0268</td>
<td>0.0195</td>
<td>0.0143</td>
<td>0.0093</td>
<td>0.0044</td>
</tr>
<tr>
<td>ARE</td>
<td>0.7716</td>
<td>0.5494</td>
<td>0.4864</td>
<td>0.4827</td>
<td>0.5291</td>
<td>0.6372</td>
<td>0.7205</td>
<td>0.8233</td>
<td>0.9613</td>
</tr>
</tbody>
</table>

The asymptotic variances of the WMPLE (\(\tilde{\phi}_{11}\)), MLE (\(\hat{\phi}_{11}\)), and ARE of the WMPLE with respect to the MLE for different values of $\phi_{11}$ are plotted in Fig. 1.

![Fig. 1. The asymptotic variances of the WMPLE (\(\tilde{\phi}_{11}\)), MLE (\(\hat{\phi}_{11}\)) (left) and the relative efficiency of the WMPLE with respect to the MLE (ARE) (right) for different values of $\phi_{11}$](image-url)
As seen, the loss of efficiency is acceptable and balanced by the computational ease, however, the role of the values of $n, m$ and $\phi_{11}$ in the value of efficiency is crucial.

In the second part, given $m = 2$ and the variance-covariance matrix $\Sigma$ considered in the first part, we present a preliminary study on performance of WMPLE in which all parameters in the matrix of coefficients are unknown but we limit attention to the case $\phi_{11} = \phi_{22}$ and $\phi_{12} = \phi_{21}$. The causality condition given by (4.1) of the above model, in this case, reduces to the following constraints

$$|\phi_{11} - \phi_{12}| < 1, |\phi_{11} + \phi_{12}| < 1.$$  

We simulate 500 multivariate time series of the above AR(1) model with length $n = 1000$. For evaluating our estimator, we have reported the asymptotic variance of WMPLE ($\text{avar}(\tilde{\phi}_{11})$), and the asymptotic variance of the MLE ($\text{avar}(\tilde{\phi}_{11})$).

In Table 2, for various values of $\phi_{11}$ and $\phi_{12}$, the sample means and standard errors of the WMPLE and MLE are reported.

Table 2. Sample means and standard errors of the WMPLE and MLE of the parameters $\phi_{11}$ and $\phi_{12}$

<table>
<thead>
<tr>
<th>Parameters</th>
<th>True values</th>
<th>Mean of WMPLE</th>
<th>Mean of MLE</th>
<th>SE of WMPLE</th>
<th>SE of MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_{11}, \phi_{12}$</td>
<td>(0,0)</td>
<td>(.0063,0114)</td>
<td>(.0091,.0088)</td>
<td>(.0352,.03)</td>
<td>(.0346,.0283)</td>
</tr>
<tr>
<td>$\phi_{11}, \phi_{12}$</td>
<td>(.75,.05)</td>
<td>(.7325,.0624)</td>
<td>(.7417,.0503)</td>
<td>(.0279,.02)</td>
<td>(.02,.02)</td>
</tr>
<tr>
<td>$\phi_{11}, \phi_{12}$</td>
<td>(.75,.15)</td>
<td>(.7250,.1661)</td>
<td>(.7391,.1503)</td>
<td>(.03,.02)</td>
<td>(.0173,.0173)</td>
</tr>
<tr>
<td>$\phi_{11}, \phi_{12}$</td>
<td>(.95,0)</td>
<td>(.9417,.0004)</td>
<td>(.9415,.0014)</td>
<td>(.0084,.01)</td>
<td>(.0082,.01)</td>
</tr>
<tr>
<td>$\phi_{11}, \phi_{12}$</td>
<td>(-.95,0)</td>
<td>(-.9386,.0012)</td>
<td>(-.9388,.0013)</td>
<td>(.0173,.0079)</td>
<td>(.0173,.0079)</td>
</tr>
</tbody>
</table>

As shown in Table 2, the WMPLE performs very well.

In the third part, we simulate 500 multivariate time series of the considered AR(1) model with $m = 3$ and length $n = 500$. We assume that the variance-covariance matrix $\Sigma$ is known and is given by

$$\Sigma = \begin{bmatrix} 1 & 0.7 & 0.6 \\ 0.7 & 1 & 0.5 \\ 0.6 & 0.5 & 1 \end{bmatrix}.$$  

The ARE is studied by letting the matrix of parameters be

$$\Phi = \begin{bmatrix} 0.500 & 0.030 & 0.070 \\ 0.030 & 0.600 & 0.030 \\ 0.070 & 0.030 & 0.500 \end{bmatrix}.$$  

Note that the condition required for the existence of a causal stationary solution of the multivariate AR(1) given by (4.1) is again satisfied for this proposed $\Phi$. In the following table, the sample means and standard errors of the WMPLE and MLE for all parameters in $\Phi$ including the ARE of the WMPLE with respect to the MLE are reported.

As shown in Tables 1, 4.2 and 4.3 the WMPLE has a good practical performance. It is consistent, sufficiently accurate and acts as good as the MLE for the large sample size.

5. Discussion

In this paper, we have shown how weighting can be considered for inferences about the parameters in the multivariate AR(1) time series. By using the analytical and numerical computations, we showed that where the unweighted pairwise likelihood does poorly in efficiency and may be computationally very slow, difficult and complicated, the weighted pairwise likelihood is suggested as a practically good choice.

It also seems that dimension parameter $m$ is a key variable here, which should be brought into play, as $m$ is the key variable in increasing the computational complexity that might encourage one to use pairwise.
Table 3. Sample means, standard errors and ARE of the WMPLE with respect to the MLE of the parameters of the matrix $\Phi$

<table>
<thead>
<tr>
<th>Parameters</th>
<th>True values</th>
<th>Mean of WMPLE</th>
<th>Mean of MLE</th>
<th>SE of WMPLE</th>
<th>SE of MLE</th>
<th>ARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_{1}$</td>
<td>0.500</td>
<td>0.415</td>
<td>0.474</td>
<td>0.2079</td>
<td>0.1400</td>
<td>0.4386</td>
</tr>
<tr>
<td>$\phi_{2}$</td>
<td>0.030</td>
<td>0.053</td>
<td>0.025</td>
<td>0.1810</td>
<td>0.1237</td>
<td>0.6830</td>
</tr>
<tr>
<td>$\phi_{3}$</td>
<td>0.070</td>
<td>0.106</td>
<td>0.069</td>
<td>0.1799</td>
<td>0.1164</td>
<td>0.6540</td>
</tr>
<tr>
<td>$\phi_{21}$</td>
<td>0.030</td>
<td>0.051</td>
<td>0.020</td>
<td>0.2145</td>
<td>0.1444</td>
<td>0.659</td>
</tr>
<tr>
<td>$\phi_{22}$</td>
<td>0.600</td>
<td>0.551</td>
<td>0.584</td>
<td>0.2331</td>
<td>0.1217</td>
<td>0.5220</td>
</tr>
<tr>
<td>$\phi_{23}$</td>
<td>0.030</td>
<td>0.054</td>
<td>0.025</td>
<td>0.2154</td>
<td>0.1118</td>
<td>0.5190</td>
</tr>
<tr>
<td>$\phi_{31}$</td>
<td>0.070</td>
<td>0.084</td>
<td>0.078</td>
<td>0.2006</td>
<td>0.1382</td>
<td>0.6889</td>
</tr>
<tr>
<td>$\phi_{32}$</td>
<td>0.030</td>
<td>0.015</td>
<td>0.019</td>
<td>0.1756</td>
<td>0.1248</td>
<td>0.7107</td>
</tr>
<tr>
<td>$\phi_{33}$</td>
<td>0.500</td>
<td>0.404</td>
<td>0.459</td>
<td>0.1917</td>
<td>0.1200</td>
<td>0.6259</td>
</tr>
</tbody>
</table>

References


pairwise likelihood in time series models, Statistica Sinica, 21, 255-277.


