AD-A081 251   PITTSBURGH UNIV PA DEPT OF MATHEMATICS AND STATISTICS   F/G 12/1
MAXIMUM LIKELIHOOD ESTIMATION FOR AN AUTOREGRESSIVE PROCESS WIT--ETC(U)
AUG 79   S TAN   N00014-76-C-0839

UNCLASSIFIED   NL

1 OF 1
AD
A081251

END
DATE
FILMED
3-80
DDC

Maximum Likelihood Estimation For An Autoregressive Process With Missing Observations

by

Suan-Boon Tan
Department of Mathematics and Statistics
University of Pittsburgh

AUG 79

Abstract

Three methods are proposed for estimation of the parameters of an autoregressive process of order p with missing observations. These methods are based on the maximum likelihood approach and use the EM algorithm, the Newton-Raphson method and the method of scoring, which are applied to the likelihood equations. Finally, comparison on those methods is also discussed.

[1]Research has been supported in part by ONR Contract N00014-76-C-0839.

AMS 1970 Subject classifications. Primary 62M10; Secondary 62F10.

Key Words and Phrases: Maximum likelihood estimation, autoregressive processes, missing observations, EM algorithm, Newton-Raphson, scoring, lost information matrix.

1. Introduction

An autoregressive process $\{y_t, t = 0, \pm 1, \ldots\}$ of order p is defined by

$$(1.1) \qquad \sum_{i=0}^{p} \gamma_i \, y_{t-i} = \epsilon_t, \qquad t = 0, \pm 1, \ldots,$$

where $\gamma_0 = 1$ and $\{\epsilon_t\}$ is a sequence of uncorrelated random variables with mean 0 and common variance σ^2. We assume that the roots of $\sum_{i=0}^{p} \gamma_i Z^i = 0$ are outside the unit disc. The process (1.1) is completely specified by $\tilde{\phi} \equiv (\gamma_1, \ldots, \gamma_p, \sigma^2)'$ when the ϵ_t are assumed to be normally distributed. Throughout this paper we shall assume normality of ϵ_t.

Usually statistical inference is based on a set of T consecutive observations on y_t. Let

$$(1.2) \qquad \tilde{y} = (y_1, y_2, \ldots, y_T)',$$

and let $\tilde{P}$ be a permutation matrix such that $\tilde{P}\tilde{y} = (\tilde{s}', \tilde{m}')'$, where $\tilde{s}$ is a $(T-m) \times 1$ vector and $\tilde{m}$ is an $m{\times}1$ vector, with the ordering in $\tilde{s}$ and $\tilde{m}$ preserved. Suppose only observations in $\tilde{s}$ are available and those in $\tilde{m}$ are missing. Our goal here is to obrain maximum likelihood estimates of $\tilde{\phi}$.

For any T×T matrix $\tilde{C}$, let us define $\tilde{C}_{ss}$, $\tilde{C}_{sm}$, $\tilde{C}_{ms}$ and $\tilde{C}_{mm}$ to be the $(T-m) \times (T-m)$, $(T-m) \times m$, $m \times (T-m)$ and $m{\times}m$ matrices, respectively, satisfying

$$(1.3) \qquad \tilde{P}\,\tilde{C}\,\tilde{P}' = \begin{pmatrix} \tilde{C}_{ss} & \tilde{C}_{sm} \\ \tilde{C}_{ms} & \tilde{C}_{mm} \end{pmatrix} .$$

For the rest of this paper, let $f(\tilde{y}|\tilde{\phi})$ denote the probability density function of $\tilde{y}$, $f(\tilde{s}|\tilde{\phi})$ denote the probability density function of $\tilde{s}$, $f(\tilde{m}|\tilde{s}, \tilde{\phi})$ denote the conditional probability density function of $\tilde{m}$ given $\tilde{s}$, $\log f(\tilde{y}|\tilde{\phi})$ denote the log likelihood function based on $\underset{\sim}{y}$ and log $f(\underset{\sim}{s}|\underset{\sim}{\phi})$ denote the log likelihood function based on $\underset{\sim}{s}$. We assume that the maximum likelihood solutions satisfy

$$\frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}} = \underset{\sim}{0}. \tag{1.4}$$

2. Some basic results

Assume that $\underset{\sim}{y}$ is distributed as multivariate normal with mean $\underset{\sim}{0}$ and covariance matrix $\underset{\sim}{\Sigma}$, that is, $f(\underset{\sim}{y}|\underset{\sim}{\phi})$ is given by

$$f(\underset{\sim}{y}|\underset{\sim}{\phi}) = \frac{1}{\sqrt{(2\pi)^T \, |\underset{\sim}{\Sigma}|}} \exp \left\{-\frac{1}{2} \underset{\sim}{y}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{y}\right\}. \tag{2.1}$$

Then $\underset{\sim}{P}\underset{\sim}{y} = (\underset{\sim}{s}', \underset{\sim}{m}')'$ is distributed as multivariate normal with mean $\underset{\sim}{0}$ and covariance matrix $\underset{\sim}{P} \, \underset{\sim}{\Sigma} \, \underset{\sim}{P}'$. Since $\underset{\sim}{P} \, \underset{\sim}{P}' = \underset{\sim}{I}_T$, where $\underset{\sim}{I}_T$ is the T×T identity matrix,

$$\begin{aligned} (\underset{\sim}{P} \, \underset{\sim}{\Sigma} \, \underset{\sim}{P}')^{-1} &= \underset{\sim}{P} \, \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{P}' \\ &= \frac{1}{\sigma^2} \underset{\sim}{P} \, \underset{\sim}{M} \, \underset{\sim}{P}' \\ &= \frac{1}{\sigma^2} \begin{pmatrix} \underset{\sim}{M}_{ss} & \underset{\sim}{M}_{sm} \\ \underset{\sim}{M}_{ms} & \underset{\sim}{M}_{mm} \end{pmatrix}, \end{aligned} \tag{2.2}$$

where $\sigma^2 \underset{\sim}{\Sigma}^{-1} = \underset{\sim}{M}$, and $\underset{\sim}{M}_{ss}$, $\underset{\sim}{M}_{sm}$, $\underset{\sim}{M}_{ms}$ and $\underset{\sim}{M}_{mm}$ are as defined by (1.3). Also, by (1.3), we get

$$(2.3) \qquad \underset{\sim}{P}\,\underset{\sim}{\Sigma}\,\underset{\sim}{P}' = \begin{pmatrix} \underset{\sim}{\Sigma}_{ss} & \underset{\sim}{\Sigma}_{sm} \\ \underset{\sim}{\Sigma}_{ms} & \underset{\sim}{\Sigma}_{mm} \end{pmatrix}.$$

Therefore from (2.2) and (2.3), it follows that

$$(2.4) \qquad [\mathrm{Cov}\ (\underset{\sim}{s}|\underset{\sim}{\phi})]^{-1} = \underset{\sim}{\Sigma}_{ss}^{-1} = \frac{1}{\sigma^2}\ [\underset{\sim}{M}_{ss} - \underset{\sim}{M}_{sm}\,\underset{\sim}{M}_{mm}^{-1}\,\underset{\sim}{M}_{ms}],$$

$$(2.5) \qquad [\mathrm{Cov}\ (\underset{\sim}{m}|\underset{\sim}{s},\ \underset{\sim}{\phi})]^{-1} = [\underset{\sim}{\Sigma}_{mm} - \underset{\sim}{\Sigma}_{ms}\,\underset{\sim}{\Sigma}_{ss}^{-1}\,\underset{\sim}{\Sigma}_{sm}]^{-1}$$

$$= \frac{1}{\sigma^2}\ \underset{\sim}{M}_{mm}\ ,$$

$$(2.6) \qquad E[\underset{\sim}{m}|\underset{\sim}{s},\ \underset{\sim}{\phi}] = \underset{\sim}{\Sigma}_{ms}\,\underset{\sim}{\Sigma}_{ss}^{-1}\ \underset{\sim}{s}$$

$$= -\,\underset{\sim}{M}_{mm}^{-1}\ \underset{\sim}{M}_{ms}\ \underset{\sim}{s},$$

and

$$(2.7) \qquad |\underset{\sim}{M}| = |\underset{\sim}{P}\,\underset{\sim}{M}\,\underset{\sim}{P}'|$$

$$= |\underset{\sim}{M}_{mm}|\ |\underset{\sim}{M}_{ss} - \underset{\sim}{M}_{sm}\,\underset{\sim}{M}_{mm}^{-1}\ \underset{\sim}{M}_{ms}|.$$

From (2.4) and (2.7), we obtain

$$(2.8) \qquad f(\underset{\sim}{s}|\underset{\sim}{\phi}) = \left(\frac{1}{\sqrt{2\pi\ \sigma^2}}\right)^{T-m} \left(\frac{|\underset{\sim}{M}|}{|\underset{\sim}{M}_{mm}|}\right)^{\frac{1}{2}}\ .$$

$$\exp\{-\frac{1}{2\sigma^2}\underset{\sim}{s}'\ [\underset{\sim}{M}_{ss} - \underset{\sim}{M}_{sm}\,\underset{\sim}{M}_{mm}^{-1}\ \underset{\sim}{M}_{ms}]\ \underset{\sim}{s}\}.$$

Expressions (2.5) and (2.6) will be used in the following sections. Though (2.8) gives the expression for the probability density function of $\underset{\sim}{s}$, we will not use it to obtain the score function $\frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}$, due to the simplicity of $\frac{\partial \log f(\underset{\sim}{y}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}$ and Lemma 1 in section 3. We will use (2.8) in proving the asymptotic properties of the estimates in a subsequent paper. Under suitable conditions, the estimates of $\underset{\sim}{\phi}$ based on the Newton-Raphson method and the method of scoring are shown to be $\sqrt{T-m}$ - consistent, asymptotically normal and one-step asymptotically efficient if the initial estimates are $\sqrt{T-m}$ - consistent.

3. Estimation

Let

$$\underset{\sim}{\gamma} = (\gamma_1, \ldots, \gamma_p)' \tag{3.1}$$

and

$$\underset{\sim}{\gamma}_\mu = (1, \underset{\sim}{\gamma}')'. \tag{3.2}$$

Then (see Anderson (1971), sec 6.2, and Box and Jenkins (1976), sec 7.A.5)

$$\begin{aligned} \log f(\underset{\sim}{y}|\underset{\sim}{\phi}) &= -\frac{T}{2} \log (2\pi\sigma^2) + \frac{1}{2} \log |\underset{\sim}{M}| - \frac{1}{2\sigma^2} \underset{\sim}{y}' \underset{\sim}{M} \underset{\sim}{y} \\ &= -\frac{T}{2} \log (2\pi\sigma^2) + \frac{1}{2} \log |\underset{\sim}{M}| - \frac{1}{2\sigma^2} \underset{\sim}{\gamma}_\mu' \underset{\sim}{D} \underset{\sim}{\gamma}_\mu, \end{aligned} \tag{3.3}$$

where the elements m_{st} of the T×T matrix $\underset{\sim}{M}$ is given by

$$\begin{aligned} m_{st} &= m_{T+1-t,\, T+1-s} \\ &= \sum_{j=0}^{\min(s,t)-1} \gamma_j \gamma_{j+|s-t|}, && s, t = 1, \ldots, p, \\ &= \sum_{j=0}^{p-|s-t|} \gamma_j \gamma_{j+|s-t|}, && \max(s,t) \geq p+1, \\ & && \min(s,t) \leq T-p, \\ & && |s-t| = 0, 1, \ldots, p, \\ &= 0, && |s-t| = p+1, \ldots, \end{aligned} \tag{3.4}$$

and the element d_{ij} of the $(p+1) \times (p+1)$ matrix $\underset{\sim}{D}$ is given by

$$(3.5) \qquad d_{ij} = \underset{\sim}{y}' \underset{\sim}{A}_{ij} \underset{\sim}{y}$$

with the element a_{mn} of the $T \times T$ matrix $\underset{\sim}{A}_{ij}$ given

by

$$(3.6) \qquad \begin{aligned} a_{mn} &= 1, \qquad \text{if } (m,n) = (i+s,\ j+s), \\ &\qquad\qquad s = 0,1,\ldots,\ T+1-(i+j), \\ &= 0, \qquad \text{otherwise.} \end{aligned}$$

From (3.3) we obtain

$$(3.7) \qquad \frac{\partial \log f(\underset{\sim}{y}|\underset{\sim}{\phi})}{\partial \gamma_j} = \frac{1}{2}\frac{\partial \log |\underset{\sim}{M}|}{\partial \gamma_j} - \frac{1}{\sigma^2}\sum_{i=0}^{p} \gamma_i\, d_{i+1,j+1},$$

$$j = 1,\ldots,\ p,$$

and

$$(3.8) \qquad \frac{\partial \log f(\underset{\sim}{y}|\underset{\sim}{\phi})}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\underset{\sim}{\gamma}_\mu' \underset{\sim}{D}\, \underset{\sim}{\gamma}_\mu.$$

It is $\dfrac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}$ that is of interest to us and not $\dfrac{\partial \log f(\underset{\sim}{y}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}$, since observations in $\underset{\sim}{m}$ are missing. However, $\dfrac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}$ can be derived from $\dfrac{\partial \log f(\underset{\sim}{y}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}$ as indicated in the following lemma.

<u>LEMMA 1</u>. (Orchard and Woodbury (1972)).

$$(3.9) \qquad \frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}} = E\left[\frac{\partial \log f(\underset{\sim}{y}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}} \;\middle|\; \underset{\sim}{s}, \underset{\sim}{\phi}\right].$$

Proof. The result follows immediately from

$$E\left[\frac{\partial \log f(\underset{\sim}{m}|\underset{\sim}{s},\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}} \mid \underset{\sim}{s},\underset{\sim}{\phi}\right] = \underset{\sim}{0}. \tag{3.10}$$

It is clear from (3.7) - (3.9) that

$$\frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \gamma_j} = \frac{1}{2}\frac{\partial \log |\underset{\sim}{M}|}{\partial \gamma_j} - \frac{1}{\sigma^2}\sum_{i=0}^{p} \gamma_i \, E\left[d_{i+1,j+1} \mid \underset{\sim}{s}, \underset{\sim}{\phi}\right] \tag{3.11}$$

$$j = 1,\dots, p,$$

and

$$\frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\underset{\sim}{\gamma}_\mu' \, E\left[\underset{\sim}{D} \mid \underset{\sim}{s}, \underset{\sim}{\phi}\right] \underset{\sim}{\gamma}_\mu . \tag{3.12}$$

The term $\log |\underset{\sim}{M}|$ is $O(1)$ (see Hannan (1973), e.g.) while d_{ij} is $O_p(T)$. The effect of neglecting $\log |\underset{\sim}{M}|$ is negligible for moderate or large T, and we shall neglect $\log |\underset{\sim}{M}|$ and other negligible terms henceforth. From (3.11) and (3.12), it follows that the likelihood equations are given by

$$\sum_{g=0}^{p} \gamma_g \, E\left[d_{g+1,j+1} \mid \underset{\sim}{s},\underset{\sim}{\phi}\right] = 0, \qquad j = 1,\dots, p, \tag{3.13}$$

and

$$\sigma^2 = \frac{1}{T}\underset{\sim}{\gamma}_\mu' \, E\left[\underset{\sim}{D} \mid \underset{\sim}{s},\underset{\sim}{\phi}\right]\underset{\sim}{\gamma}_\mu . \tag{3.14}$$

When there are no missing observations, $E\left[d_{g+1,j+1} \mid \underset{\sim}{s},\underset{\sim}{\phi}\right] = d_{g+1,j+1}$ does not involve unknown parameters. Then the equations are linear in γ_i, $i = 1,\dots, p$, and are the Yule-Walker equations. When missing observations do occur, $E\left[d_{g+1,j+1} \mid \underset{\sim}{s}, \underset{\sim}{\phi}\right]$ involves unknown parameters and (3.13) and (3.14) are highly non-linear in the unknown parameters. In fact, from (2.5) and (2.6),

$$\begin{aligned}
&E\left[d_{g+1,j+1} \mid \underset{\sim}{s}, \underset{\sim}{\phi}\right] \\
&= \underset{\sim}{s}' \Big[(\underset{\sim}{A}_{g+1,j+1})_{ss} - (\underset{\sim}{A}_{g+1,j+1})_{sm} \underset{\sim}{K} \\
&\quad - \underset{\sim}{K}' (\underset{\sim}{A}_{g+1,j+1})_{ms} + \underset{\sim}{K}' (\underset{\sim}{A}_{g+1,j+1})_{mm} \underset{\sim}{K}\Big] \underset{\sim}{s} \\
&\quad + \sigma^2 \, tr\, (A \qquad) \quad M^{-1} .
\end{aligned} \tag{3.15}$$

where $\underset{\sim}{K} = \underset{\sim}{M}_{mm}^{-1} \underset{\sim}{M}_{ms}$, and the matrices involved are as defined in (1.3), (3.4) and (3.6). Therefore solutions of (3.13) and (3.14) are not straightforward and iterative procedures have to be used.

We propose the following three methods of solving (3.13) and (3.14): the EM algorithm, the Newton-Raphson method and the method of scoring.

a. <u>The EM algorithm</u>. Since $\underset{\sim}{\phi}$ is to be estimated, it is natural that one replace $E[\cdot \mid \underset{\sim}{s}, \underset{\sim}{\phi}]$ in (3.13) and (3.14) by $E[\cdot \mid \underset{\sim}{s}, \underset{\sim}{\phi}_i]$, where $\underset{\sim}{\phi}_i$ is some estimated value of $\underset{\sim}{\phi}$, and obtain $\underset{\sim}{\phi}_{i+1}$ iteratively by solving

$$\text{(3.16)} \qquad \sum_{g=0}^{p} (\gamma_g)_{i+1} \, E\left[d_{g+1,j+1} \mid \underset{\sim}{s}, \underset{\sim}{\phi}_i\right] = 0. \qquad j = 1, \ldots, p,$$

and

$$\text{(3.17)} \qquad \sigma_{i+1}^2 = \frac{1}{T} (\underset{\sim}{\gamma}_\mu)_i' \, E\left[\underset{\sim}{D} \mid \underset{\sim}{s}, \underset{\sim}{\phi}_i\right] (\underset{\sim}{\gamma}_\mu)_i.$$

Here $(\gamma_g)_j$ and $(\underset{\sim}{\gamma}_\mu)_j$ denote the estimates of γ_g and $\underset{\sim}{\gamma}_\mu$, respectively, at the j-th iteration. As shown in Tan (1979), the above method gives the same solutions as the EM algorithm proposed by Dempster, Laird and Rubin (1977).

b. <u>The Newton-Raphson Method</u>. From (3.11) and (3.12), we obtain

$$\text{(3.18)} \qquad \Theta_{kj} \equiv -\frac{\partial^2 \log f(\underset{\sim}{s} \mid \underset{\sim}{\phi})}{\partial\gamma_k \, \partial\gamma_j} = \frac{1}{\sigma^2} \left\{ \sum_{g=0}^{p} \gamma_g \frac{\partial}{\partial\gamma_k} E\left[d_{g+1,j+1} \mid \underset{\sim}{s}, \underset{\sim}{\phi}\right] + E\left[d_{k+1,j+1} \mid \underset{\sim}{s}, \underset{\sim}{\phi}\right]\right\},$$

$$j, k = 1, \ldots, p.$$

$$(3.19)\qquad \Theta_{j,p+1} \equiv -\frac{\partial^2 \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial\gamma_j\, \partial\sigma^2} = \sum_{g=0}^{p} \gamma_g \{\frac{1}{\sigma^2} \mathrm{tr}\, (\underset{\sim}{A}_{g+1,j+1})_{mm} \underset{\sim}{M}_{mm}^{-1}$$

$$- \frac{1}{\sigma^4} E\left[d_{g+1,j+1} \mid \underset{\sim}{s},\underset{\sim}{\phi}\right]\},$$

$$j = 1,\ldots, p,$$

and

$$(3.20)\qquad \Theta_{p+1,p+1} \equiv -\frac{\partial^2 \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial(\sigma^2)^2} = -\left[\frac{T}{2\sigma^4} + \frac{1}{2}\, \underset{\sim}{\gamma}_\mu' \,\underset{\sim}{B}\, \underset{\sim}{\gamma}_\mu\right],$$

where the element b_{ij} of the $(p+1) \times (p+1)$ matrix $\underset{\sim}{B}$ is given by

$$(3.21)\qquad b_{ij} = -\frac{2}{\sigma^6} E\left[d_{ij} \mid \underset{\sim}{s},\underset{\sim}{\phi}\right] + \frac{1}{\sigma^4} \mathrm{tr}\, (\underset{\sim}{A}_{i+1,j+1})_{mm} \underset{\sim}{M}_{mm}^{-1}.$$

Thus, the Newton-Raphson method leads to the following set of equations:

$$(3.22)\qquad \underset{\sim}{\Theta}\Big|_{\underset{\sim}{\phi}_i} (\underset{\sim}{\phi}_{i+1} - \underset{\sim}{\phi}_i) = \frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial\underset{\sim}{\phi}}\Big|_{\underset{\sim}{\phi}_i} ,$$

where the element Θ_{ij} of $\underset{\sim}{\Theta}$ is as given by (3.18) - (3.21) and $\frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial\underset{\sim}{\phi}}$ is as given by (3.11) (without the first term on the right-hand side) and (3.12).

c. <u>The Method of Scoring</u>. From (3.15), (3.18) - (3.21), we obtain

$$(3.23)\qquad \Phi_{ij} \equiv E\left[\Theta_{ij}\right] = \frac{1}{\sigma^2} \{T\, \sigma(i-j) + \sum_{g=0}^{p} \gamma_g\, E\, \frac{\partial}{\partial\gamma_i}\, E\left[d_{g+1,j+1} \mid \underset{\sim}{s},\underset{\sim}{\phi}\right]\},$$

$$i,\ j = 1,\ldots, p,$$

$$(3.24)\qquad \Phi_{j,p+1} \equiv E\left[\Theta_{j,p+1}\right] = \frac{1}{\sigma^2} \sum_{g=0}^{p} \gamma_g\, \mathrm{tr}\, (\underset{\sim}{A}_{g+1,j+1})_{mm} \underset{\sim}{M}_{mm}^{-1},$$

$$j = 1,\ldots, p,$$

and

$$(3.25) \qquad \Phi_{p+1,p+1} \equiv E\left[\Theta_{p+1,p+1}\right] = -\frac{T}{2\sigma^4} + \frac{1}{\sigma^6} T \sum_{i,j=0}^{p} \gamma_i \gamma_j \sigma(i-j)$$

$$- \frac{1}{2\sigma^4} \sum_{i,j=0}^{p} \gamma_i \gamma_j \operatorname{tr} (A_{i+1,j+1})_{mm} M_{mm}^{-1},$$

where $\sigma(k) = E\left[y_{t+k}\, y_t\right]$. Thus, the method of scoring leads to the following set of equations:

$$(3.26) \qquad \underset{\sim}{\Phi}\Big|_{\underset{\sim}{\phi}_i} (\underset{\sim}{\phi}_{i+1} - \underset{\sim}{\phi}_i) = \frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}\Big|_{\underset{\sim}{\phi}_i},$$

where the elements Φ_{ij} of $\underset{\sim}{\Phi}$ are given by (3.23) - (3.25), and $\frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}$ is given by (3.11) (without the first term on the right-hand side) and (3.12).

We have used the fact that $E\left[d_{i+1,j+1}\right] = \left[T - (i+j)\right] \sigma(i-j)$, which can be approximated by $T\, \sigma(i-j)$ for moderate or large T.

4. Comparison of the methods of estimation

The estimates of $\underset{\sim}{\phi}$ based on the Newton-Raphson method, the method of scoring and the EM algorithm can be expressed in the following form

$$\underset{\sim}{H}\Big|_{\underset{\sim}{\phi}_i} (\underset{\sim}{\phi}_{i+1} - \underset{\sim}{\phi}_i) = \frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}\Big|_{\underset{\sim}{\phi}_i},$$

where $\frac{\partial \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}}$ is given by (3.11) (without the first term on the right-hand side) and (3.12). In the Newton-Raphson method, we have

$$\underset{\sim}{H} = \underset{\sim}{\Theta},$$

where $\underset{\sim}{\Theta}$ is given by (3.18) - (3.21). In the method of scoring, we have

$$\underset{\sim}{H} = \underset{\sim}{\Phi},$$

where $\underset{\sim}{\Phi}$ is given by (3.23) - (3.25). In the EM algorithm, we have

$$\underset{\sim}{H} = \underset{\sim}{\omega},$$

where elements ω_{ij} of $\underset{\sim}{\omega}$ are given by

$$\begin{aligned}
\omega_{gj} &= \omega_{jg} \\
&= \frac{1}{\sigma^2} \, E\left[d_{g+1,j+1} \mid \underset{\sim}{s}, \underset{\sim}{\phi}\right], && g,j = 1,\ldots, p, \\
&= 0, && g = 1,\ldots, p,\ j = p+1 \\
&= \frac{T}{2\sigma^4}, && g = p+1,\ j = p+1.
\end{aligned}$$

Let

$$\underset{\sim}{A} = \underset{\sim}{\Theta} - \underset{\sim}{\Phi},$$

$$\underset{\sim}{B} = \underset{\sim}{\omega} - \underset{\sim}{\Theta},$$

$$\underset{\sim}{C} = \underset{\sim}{\omega} - \underset{\sim}{\Phi},$$

and

$$\begin{aligned}
\underset{\sim}{I}_m(\underset{\sim}{\phi}) &= E\left[-\frac{\partial^2 \log f(\underset{\sim}{y}|\underset{\sim}{\phi})}{\partial\underset{\sim}{\phi}\,\partial\underset{\sim}{\phi}'}\right] \\
&\quad - E\left[-\frac{\partial^2 \log f(\underset{\sim}{s}|\underset{\sim}{\phi})}{\partial\underset{\sim}{\phi}\,\partial\underset{\sim}{\phi}'}\right] \\
&= E\left[-\frac{\partial^2 \log f(\underset{\sim}{y}|\underset{\sim}{\phi})}{\partial\underset{\sim}{\phi}\,\partial\underset{\sim}{\phi}'}\right] - E[\underset{\sim}{\Theta}].
\end{aligned}$$

$\underset{\sim}{I}_m(\underset{\sim}{\phi})$ is referred to as the lost information matrix in Orchard and Woodbury (1972). It follows that

$$E(\underset{\sim}{A}) = \underset{\sim}{0},$$

and also

$$\lim_{T-m\to\infty} \frac{E(\underset{\sim}{B})}{T-m} = \lim_{T-m\to\infty} \frac{E(\underset{\sim}{C})}{T-m} = \lim_{T-m\to\infty} \frac{\underset{\sim}{I}_m(\underset{\sim}{\phi})}{T-m},$$

since $\lim_{T-m\to\infty} \frac{E(\underset{\sim}{\omega})}{T-m} = \lim_{T-m\to\infty} - E\, \frac{1}{T-m} \frac{\partial^2 \log f(\underset{\sim}{y}|\underset{\sim}{\phi})}{\partial \underset{\sim}{\phi}\, \partial \underset{\sim}{\phi}'}$ (see Box and Jenkins (1976) section 7.A.5). In general, $\lim_{T-m\to\infty} \frac{\underset{\sim}{I}_m(\underset{\sim}{\phi})}{T-m}$ is not negligible. For example, when p = 1 in (1.1) and the process $\{y_t\}$ is periodically observed for α time points and then not observed for two time points, it can be shown that (see Tan (1979))

$$\lim_{T-m\to\infty} \frac{\underset{\sim}{I}_m(\underset{\sim}{\phi})}{T-m}$$

$$= \frac{1}{\alpha\sigma^2} \begin{pmatrix} \dfrac{\sigma^2(1+2\gamma_1^2)\,(3+\gamma_1^2-\gamma_1^4)}{\lambda^2} & \dfrac{-\gamma_1(1+2\gamma_1^2)}{\lambda} \\[2ex] \dfrac{-\gamma_1\,(1+2\gamma_1^2)}{\lambda} & \dfrac{1}{\sigma^2} \end{pmatrix}$$

where $\lambda = (1+\gamma_1^2+\gamma_1^4)$. It is easy to see that the above matrix is positive definite.

Acknowledgement. I am indebted to Paul Shaman for his suggestions and advice on this topic.

REFERENCES

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

2. Box, G.E.P., and Jenkins, G.M. (1976). Time Series Analysis Forecasting and Control 2nd edition. Holden-Day, Inc., San Francisco.

3. Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). "Maximum likelihood from incomplete data via the EM algorithm." (With Discussion). Journal of the Royal Statistical Society, Series B. 39, 1-38.

4. Hannan, E.J. (1973). "The asymptotic theory of linear time-series models." Journal of Applied Probability 10, 130-145.

5. Orchard, T., and Woodbury, M.A. (1972). "A missing information principle: theory and applications." Proceeding of 6th Berkeley Symposium on Mathematical Statistics and Probability 1, 697-715.

6. Tan, Suan-Boon (1979). "Maximum likelihood estimation in autoregressive processes with missing observations." Ph.D. Thesis, University of Pittsburgh.