We have applied the trapezium method to approximate integrals in an implementation of the EM algorithm proposed by Tsai and Chan (1999b) for estimating continuous-time autoregressive models, whose original implementation was based on Euler’s method for approximating integrals. It is well known that the trapezium method generally provides a second order approximation to an integral of a well-behaved functional of Wiener process, whereas the Euler method is generally of first order. Simulation results confirm that with increasing discretization frequency, the EM estimators based on the trapezium method converge to the (conditional) ML estimator at a faster rate than the EM estimators based on Euler’s method. However, with an appropriate choice of discretization frequency, the EM estimator based on Euler’s method outperforms both the EM estimator based on the trapezium method and the ML estimator in terms of biases and standard deviations of the estimates. An invariance property of the EM estimator based on the trapezium method is briefly discussed.

Some key words: Trapezium method, Girsanov formula, Maximum likelihood estimation, Stochastic differential equations, irregularly sampled time series, Kalman filter.

1 Introduction

Owing to the sampling procedure or the presence of missing data, many time series, say, \{Y_t\}_{t=0,...,N}, are sampled with irregular time intervals. Irregularly sampled time series data are often analyzed by assuming that these data are sampled from an underlying continuous-time process. The underlying continuous-time process may be modeled as driven by some stochastic differential equations, for example, the linear continuous-time autoregressive moving average (ARMA) models. This linear specification results in a tractable likelihood for the observed discrete-time data. Hence this method is rather popular for analyzing irregularly sampled time-series data. See, e.g., Harvey (1989), Bergstrom (1990), Tong (1990) and Jones (1980, 1993). See Belcher et at.
(1994) for some discussions of parameterizing continuous-time autoregressive models. We also note that in many applications, for example, economics, the main interest may consist of drawing inference on the underlying stochastic differential equation even with equally spaced data; see Bergstrom (1990).

The likelihood function of a CAR\((p)\) model can be computed via Kalman filters. Maximum likelihood estimation can be done by means of some nonlinear optimization algorithm such as the simplex method. Tsai and Chan (1999b) proposed the use of the Expectation-Maximization (EM) algorithm to derive approximate ML estimators of the CAR\((p)\) models. The EM algorithm is based on an integral representation of the likelihood function and an approximation of the integrals by Euler’s method. Simulation results reported by Tsai and Chan (1999b) suggested that with suitable choices of the discretization frequency, the EM estimators are comparable to the ML estimator in terms of bias, but with smaller standard errors than those of the ML estimator. Tsai and Chan (1999b) proved in the first order case that as the discretization frequency increases to infinity, the EM estimators converge to the ML estimator in probability. Tsai and Chan (1999b) conjectured that in the limit of no discretization error, the EM estimator becomes the ML estimator.

An alternative to Euler’s method is the trapezium approximation. It is well known that the trapezium approximation converges to its limit with a faster convergence rate than Euler’s method. We now briefly recall the trapezium method for approximating an integral, and some of its properties without proofs; see Milstein (1995, pp.6, 135-141) for the exact statements and their proofs. Let \(\{Z(s), a \leq s \leq b\}\) be a well-behaved functional of a Wiener process; more specifically, \(Z(s)\) is a smooth functional of \(\{W(t), a \leq t \leq s\}\), with polynomial growth in the tail. Consider integrals of the form \(\int_a^b Z(s)ds\). Let \(a = t_0 < t_1 < \cdots < t_N = b\) be an equally spaced partition of the interval \([a, b]\). Denote \(h = (b - a)/N\). Write \(\int_a^b Z(s)ds = \sum_1^N \int_{t_{i-1}}^{t_i} Z(s)ds\). The trapezium method approximates \(\int_{t_{i-1}}^{t_i} Z(s)ds\) by \((Z(t_{i-1}) + Z(t_i))h/2\) whereas the Euler method approximates the integrals by \(Z(t_{i-1})h\). The trapezium method can be shown to be of order two, that is, the mean error is \(O(h^2)\), but the Euler method is of order one. Moreover, it holds generically that there exists no ‘simple’ approximating scheme for \(\int_a^b Z(s)ds\) which is of order higher than two.

In this paper, we compare the finite sample performance of the (conditional) ML estimator with those of the EM estimators when the integrals are approximated by Euler’s method and the trapezium method. Simulation results confirm that the EM estimators based on the trapezium method converge to the (conditional) ML estimator at a faster rate than the EM estimators based on Euler’s method. In a first order case, the EM estimators based on
the trapezium method converge to a limit very close to the conditional ML estimator. In a second order case, the EM estimators based on the trapezium method seem to converge to the ML estimator generated by the simplex method in terms of biases. The standard deviations of the EM estimators based on the trapezium method are also close to those of the ML estimator, although a bit larger. It is also interesting to note that with appropriate choices of discretization frequencies, the EM estimators based on Euler’s method outperform the EM estimators based on the trapezium method and the ML estimator in terms of biases and standard deviations of the estimates.

The organization of this paper is as follows. Section 2 provides a review of ML estimation of CAR(p) models. An integral representation of the likelihood, its approximation by the trapezium method and the EM algorithm are described in section 3. Simulation results comparing the finite sample performance of the EM estimators based on Euler’s approximation method and the trapezium method, and that of the (conditional) ML estimator are reported in section 4. An invariance property of the EM estimator based on the trapezium method is briefly discussed. We conclude in section 5.

2 Review of Maximum Likelihood Estimation of CAR Models

Nonlinear optimization is the most popular method in the literature for doing maximum likelihood estimation of the CAR(p) models, which we briefly review below. For further discussions of the CAR(p) processes, see, e.g., Brockwell (1993) and Brockwell and Stramer (1995). A CAR(p) process is defined to be a solution of the p-th order differential equation:

\[ dY_t^{(p-1)} = (\alpha_0 + \alpha_1 Y_t + \cdots + \alpha_p Y_t^{(p-1)})dt + \sigma dW_t, \]  

where the superscript \((j)\) denotes \(j\)-fold differentiation with respect to \(t\); \(\{W_t, t \geq 0\}\) is the standard Brownian motion, and \(\alpha_0, \ldots, \alpha_p\) and \(\sigma\) are constants. We assume that \(\sigma > 0\) and \(\alpha_1 \neq 0\). The Brownian motion \(\{W_t, t \geq 0\}\) is nowhere differentiable. So, the solution of equation (1) is interpreted as satisfying the integral equation:

\[ Y_t^{(p-1)} - Y_0^{(p-1)} = \int_0^t (\alpha_0 + \alpha_1 Y_s + \cdots + \alpha_p Y_s^{(p-1)})ds + \sigma W_t. \]

The term \((\alpha_0 + \alpha_1 Y_s + \cdots + \alpha_p Y_s^{(p-1)})\) is referred to as the instantaneous mean of the process, and \(\sigma\) the instantaneous standard deviation. Equation (1) can also be equivalently cast in terms of the observation and state equations:

\[ Y_t = H'X_t, \quad t \geq 0, \]

\[ dX_t = (AX_t dt + \alpha_0 \delta_p)dW_t, \]
where

\[
A = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix},
X_t = \begin{bmatrix}
X_t^{(0)} \\
X_t^{(1)} \\
\vdots \\
X_t^{(p-2)} \\
X_t^{(p-1)}
\end{bmatrix},
\delta_p = \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
1
\end{bmatrix},
H = \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
0
\end{bmatrix}
\]

and the superscript $'$ denotes the transpose of a vector, and $X_t^{(0)} = Y_t$. Equation (3) is an Ito differential equation for the state vector $X_t$. We assume that $X_0$ is independent of \{W_t, t \geq 0\} and $X_0$ is determined by initial conditions that could be random or deterministic. The process \{Y_t, t \geq 0\} is thus said to be a CAR(p) process with parameter $(\theta, \sigma^2) = (\alpha_0, ..., \alpha_p, \sigma^2)$ if $Y_t = H'X_t$, where \{X_t\} is a solution of (3). The solution of (3) can be written as

\[
X_t = e^{At}X_0 + \alpha_0 \int_0^t e^{A(t-u)}\delta_p du + \sigma \int_0^t e^{A(t-u)}\delta_p dW_u,
\]

where $e^{At} = I + \sum_{n=1}^{\infty} [(At)^n(n!)^{-1}]$, and $I$ is the identity matrix.

Let the mean vector of $X_t$ be denoted by $\mu_t$. It satisfies the equation:

\[
\mu_t = \frac{\alpha_0}{\alpha_1} (e^{At} - I) H + e^{At} \mu_0.
\]

The covariance kernel of \{X_t\}, denoted by $\gamma_{s,t}$, equals

\[
\gamma_{s,t} = E[(X_s - \mu_s)(X_t - \mu_t)'] = e^{As}V_0e^{A't} + \sigma^2 \int_0^{t \wedge s} e^{A(s-u)}\delta_p\delta_p' e^{A'(t-u)} du
\]

\[
= \begin{cases}
e^{A(s-t)}V_t, & 0 \leq t \leq s < \infty \\
V_t e^{A'(t-s)}, & 0 \leq s \leq t < \infty,
\end{cases}
\]

where $t \wedge s = \min(s, t)$ and

\[
V_t = \gamma_{t,t} = e^{At}V_0e^{A't} + \sigma^2 \int_0^{t} e^{A(t-u)}\delta_p\delta_p' e^{A'(t-u)} du.
\]

It follows from the above equations on $\mu_t$ and $V_t$ that the states and the observations, $X_t$ and $Y_t$, at the sampling times $t_0, t_1, ..., t_i$, satisfy the discrete-time state and observation equations:

\[
X_{t_{i+1}} = \mu_{t_{i+1}} + e^{A(t_{i+1}-t_i)}(X_{t_i} - \mu_{t_i}) + Z_{t_i}, \quad i = 0, 1, ..., \tag{4}
\]
\[ Y_t = H'X_{t_i}, \quad i = 0, 1, \ldots, \] (5)

where \( Z_{t_i} \) is independent of \( X_{t_i} \) and \( \{Z_{t_i}, i = 0, 1, \ldots\} \) is an independent sequence of Gaussian random vectors with zero mean and covariance matrices

\[ \Sigma_i = E(Z_t, Z_t') = \sigma^2 \int_{t_i}^{t_{i+1}} e^{A(t_{i+1}-u)} \delta_{p} \delta_{u} e^{A'(t_{i+1}-u)} du. \] (6)

These equations are needed for applications of the Kalman recursions (see, e.g., Chapter 12 of Brockwell and Davis, 1991). Let capital letters be used for random vectors and corresponding small letters for observed vectors. Let \( Y = \{Y_{t_i}\}_{i=0}^{N} \) and define \( \hat{X}_{t_i} \) as the conditional expectation of \( X_t \) based on the observations up to time \( s \) and \( P_{t_i} \) the corresponding covariance matrix, i.e., \( \hat{X}_{t_i} = E(X_t | y_j, j \leq s) = (\hat{X}_{t_i}^{(0)}, \ldots, \hat{X}_{t_i}^{(p-1)}) \) and \( P_{t_i} = \text{var}(X_t | y_j, j \leq s) \).

Then we can compute recursively \( \tilde{y}_{t_i} = y_{t_i} - \hat{X}_{t_i|t_{i-1}} \), the predictive residuals, and \( p^{(1,1)}_{t_i|t_{i-1}} \), the \((1,1)\)th element of \( P_{t_i|t_{i-1}} \), \( i \geq 1 \), which are handy for computing minus twice the log-likelihood function,

\[
-2l_v(\theta, \sigma^2) = \sum_{i=0}^{N} \left[ \frac{\tilde{y}_{t_i}^2}{p^{(1,1)}_{t_i|t_{i-1}}} + \log p^{(1,1)}_{t_i|t_{i-1}} \right] + (N + 1) \log(2\pi). \] (7)

Here, we start with a diffuse initial condition as we do not assume stationarity, i.e., let

\[
\hat{X}_{t-1|t_{i-1}} = [\bar{y}, 0, \ldots, 0]^T, \quad P_{t-1|t_{i-1}} = \delta s^2 I,
\]

where \( t_{i-1} \) is some arbitrarily chosen time point, \( \delta \) is some positive number, \( \bar{y} \) and \( s^2 \) are the sample mean and sample variance of \( y \), respectively. A reasonable choice of \( \delta \) is, e.g., 5.

A non-linear optimization algorithm can then be used in conjunction with the expression for \(-2l_v(\theta, \sigma^2)\) to find the maximum likelihood estimate of the parameter \((\theta, \sigma^2)\). The calculations of \( e^{A't} \) are most readily performed by first block-diagonalizing \( A \) and then applying a Padé approximation on each block. See, e.g., Ward (1977). For a faster alternative for computing \( \Sigma_i \), see Tsai and Chan (1999a).

The parameter \( \sigma^2 \) can be concentrated out of the likelihood (see Jones, 1980). We only need to replace \( P_{t_i} \) by \( P^*_t = P_{t_i}/\sigma^2 \) and equation (7) becomes

\[
-2l_v(\theta, \sigma^2) = \sum_{i=0}^{N} \left[ \frac{\tilde{y}_{t_i}^2}{\sigma^2 p^{(1,1)}_{t_i|t_{i-1}}} + \log \left( \sigma^2 p^{*(1,1)}_{t_i|t_{i-1}} \right) \right] + (N + 1) \log(2\pi). \] (8)
Differentiating (8) with respect to \( \sigma^2 \) and equating to zero gives

\[
\sigma^2 = \frac{1}{N+1} \sum_{i=0}^{N} \frac{\hat{y}_{t_i}^2}{p_{t_i|t_{i-1}}} \tag{9}
\]

and substituting into (8), the objective function becomes

\[
-2I_Y(\theta) = (N+1) \log \left( \sum_{i=0}^{N} \frac{\hat{y}_{t_i}^2}{p_{t_i|t_{i-1}}} \right) + \sum_{i=0}^{N} \log p_{t_i|t_{i-1}} + C, \tag{10}
\]

where \( C = (N+1)(1 - \log(N+1) + \log(2\pi)) \). This function is then minimized with respect to \( \theta \) to get the maximum likelihood estimate \( \hat{\theta} \). The parameter estimate \( \hat{\sigma}^2 \) is then calculated from (9).

### 3 An Integral Representation of the Likelihood, Its Approximation by the Trapezium Method and the EM Algorithm

Tsai and Chan (1999b) applied the Cameron-Martin-Girsanov formula (see, e.g., Corollary 8.23 of Øksendal, 1995) to derive an integral representation of the pdf of \( Y = \{Y_{t_0}, Y_{t_1}, \ldots, Y_{t_N}\} \) with respect to the Lebesgue measure. This representation will be useful for applying an Expectation-Maximization (EM) algorithm to estimate the parameters. The notations \( E_\theta(\cdot|y) \), \( \text{var}_\theta(\cdot|y) \) and \( \text{cov}_\theta(\cdot|y) \) denote the conditional expectation, variance and covariance of the enclosed expression given \( Y = y \), respectively, where \( \theta \) is the true parameter. Also, the parameter \( \theta \) is omitted if no confusion would occur. The cumulative distribution function of \( Y \) is

\[
P_{\theta,\sigma^2}(Y_{t_0} \leq y_{t_0}, Y_{t_1} \leq y_{t_1}, \ldots, Y_{t_N} \leq y_{t_N})
= E_\theta,\sigma^2[I(Y_{t_0} \leq y_{t_0}, Y_{t_1} \leq y_{t_1}, \ldots, Y_{t_N} \leq y_{t_N})]
= E_{0,\sigma^2}[I(Y_{t_0} \leq y_{t_0}, Y_{t_1} \leq y_{t_1}, \ldots, Y_{t_N} \leq y_{t_N}) \frac{dP_{\theta,\sigma^2}}{dP_{0,\sigma^2}}(X_s^{(p-1)}, t_0 \leq s \leq t_N)]
= E_{0,\sigma^2}\left[E_{0,\sigma^2}[I(Y_{t_0} \leq y_{t_0}, \ldots, Y_{t_N} \leq y_{t_N}) \frac{dP_{\theta,\sigma^2}}{dP_{0,\sigma^2}}(X_s^{(p-1)}, t_0 \leq s \leq t_N) | Y]\right]
= E_{0,\sigma^2}\left[I(Y_{t_0} \leq y_{t_0}, \ldots, Y_{t_N} \leq y_{t_N}) E_{0,\sigma^2}\left[\frac{dP_{\theta,\sigma^2}}{dP_{0,\sigma^2}}(X_s^{(p-1)}, t_0 \leq s \leq t_N) | Y\right]\right]
= \int_{-\infty}^{y_{t_0}} \cdots \int_{-\infty}^{y_{t_N}} E_{0,\sigma^2}\left[\frac{dP_{\theta,\sigma^2}}{dP_{0,\sigma^2}}(X_s^{(p-1)}, t_0 \leq s \leq t_N) | Y\right] f_{Y_{0,\sigma^2}}(y) dy_{t_0} \cdots dy_{t_N}
\]
where
\[
\frac{dP_{\theta, \sigma^2}}{dP_{0, \sigma^2}} \left\{ X_s^{(p-1)}; 0 \leq s \leq t_N \right\} = \exp \left\{ \frac{1}{\sigma^2} \int_0^{t_N} (\alpha_0 + \alpha' X_s) dX_s^{(p-1)} - \frac{1}{2\sigma^2} \int_0^{t_N} (\alpha_0 + \alpha' X_s)^2 ds \right\}.
\]
Thus,
\[
f_{Y; (\theta, \sigma^2)}(y) = E_{0, \sigma^2} \left[ \left. \frac{dP_{\theta, \sigma^2}}{dP_{0, \sigma^2}} \left\{ X_s^{(p-1)}; 0 \leq s \leq t \right\} \right| \right] f_{Y; (0, \sigma^2)}(y),
\]
where \( f_{Y; (\theta, \sigma^2)}(y) \) is the joint pdf of \( Y = y \) under the CAR model with \( (\theta, \sigma^2) \) as its true parameter. Therefore, \( l_Y(\theta, \sigma^2) \), the log-likelihood function of \( Y \) is
\[
l_Y(\theta, \sigma^2) = \log E_{0, \sigma^2} \left[ \left. \frac{dP_{\theta, \sigma^2}}{dP_{0, \sigma^2}} \left\{ X_s^{(p-1)}; 0 \leq s \leq t \right\} \right| \right] + l_Y(0, \sigma^2).
\]
Let \( R = R(\theta, \sigma^2) = 1/\sigma^2 \int_0^{t_N} (\alpha_0 + \alpha' X_s) dX_s^{(p-1)} - 1/(2\sigma^2) \int_0^{t_N} (\alpha_0 + \alpha' X_s)^2 ds \). Note that for fixed \( \sigma^2 \), \( R \) is the complete-data log-likelihood for the process \( \{X_s, 0 \leq s \leq t_N\} \). Let \( D \) be the differentiation operator with respect to \( \theta \), then
\[
Dl_Y(\theta, \sigma^2) = \frac{DE_{0, \sigma^2} [\exp(R) | y]}{E_{0, \sigma^2} [\exp(R) | y]} = \frac{E_{0, \sigma^2} [\exp(R) (DR) | y]}{E_{0, \sigma^2} [\exp(R) | y]} = E_{\theta, \sigma^2} [(DR) | y] (11)
\]
where the last equality follows from the change of measure formula (see, e.g., Lemma 3.5.3 of Karatzas and Shreve, 1991). We remark that (11) is similar to a result due to Louis (1982) for Euclidean sample spaces. Thus,
\[
\frac{\partial l_Y}{\partial \alpha_0} = \frac{1}{\sigma^2} E_{\theta, \sigma^2} \left[ \int_{t_0}^{t_N} dX_s^{(p-1)} - \int_{t_0}^{t_N} (\alpha_0 + \alpha' X_s) ds \right] y, \tag{12}
\]
\[
\frac{\partial l_Y}{\partial \alpha_r} = \frac{1}{\sigma^2} E_{\theta, \sigma^2} \left[ \int_{t_0}^{t_N} X_s^{(r-1)} dX_s^{(p-1)} - \int_{t_0}^{t_N} (\alpha_0 + \alpha' X_s) X_s^{(r-1)} ds \right] y, \tag{13}
\]
\( r = 1, 2, \ldots, p \).

The above equations can be used to estimate the parameters by an EM algorithm. The M-step is done by equating the above equations to zero and then solving the resulting linear equations. The E-step involves the computation of the conditional expectation of integrals, which is complex. Tsai and Chan (1999b) applied Euler’s method to approximate the integrals in these
equations. The conditional expectations are then computed by Kalman filters. An alternative to Euler’s method is the trapezium method which has a faster convergence rate than the former.

The EM algorithm is a data augmentation method which augments the observed data $Y$ with some latent data $Z$ so that $l_{Y,Z}(\theta)$, the log-likelihood function of $X = (Y, Z)$, is tractable. Here $\theta$ is an arbitrary element of the parameter space $\Omega$. Let $Y = y$ be the observed incomplete data. Let $X = \{X_0, X_{1/m}, \ldots, X_{kN/m}\}$ be the unobserved complete data of which $Y$ is a measurable function, where $Y = \{Y_{t_j}\}_{0,\ldots,N}$ and $m$ is chosen to be some moderately large integer such that, for each $0 \leq j \leq N$, $t_j = k_j/m$ (approximately) for some integer $k_j$. Note that the preceding condition that the observed times $t_j = k_j/m$ (approximately) can be lifted by employing irregularly spaced partitions at the expense of more complex notations. To simplify notations, we write $X_k$ for $X_{{k}/m}$, $Y_k$ for $Y_{{k}/m}$ and $q$ for $k_N$ in this section.

The EM algorithm consists of two steps:

- **E step:** form $Q(\theta'|\theta) = E_{\theta}(l_X(\theta')|y)$;
- **M step:** maximize $Q(\cdot|\theta)$.

Suppose that for all $\theta$, $Q(\cdot|\theta)$ has a unique global maximizer at $M(\theta)$ and that $M(\theta)$ is continuous in $\theta$. Then an EM sequence $\{\theta_k\}$ is obtained from $\{\theta_{k+1}\} = M(\theta_k)$, and $\{\theta_k\}$ is a Markov chain which converges to a stationary point of $l_{Y}(\theta)$. An important property of the EM algorithm is that the likelihood of the observed data always increases along an EM sequence. See, e.g., Tanner (1991). Dempster, Laird and Rubin (1977) showed that the EM algorithm converges at a linear rate, with the rate depending on the proportion of information about $\theta$ in $l_{Y}(\theta)$ which is observed. For other convergence properties of the EM algorithm, see Dempster, Laird and Rubin (1977), and Wu (1983).

We now describe some formulas useful for the E-step. The conditional distribution of $X$ given $Y = y$ is Gaussian. The computation of the means and the variances of the conditional distribution of $X$ given $Y = y$ can be carried out by a forward Kalman filtering sequence, followed by backward iterations. See, e.g., page 189 of Anderson and Moore (1979). We now outline the Kalman computation below. For $0 \leq k \leq q$, $\hat{X}_{k|k}$, $\hat{X}_{k+1|k}$, $P_{k|k}$, $P_{k+1|k}$, can be computed via a forward Kalman filter as follows.

First, let $\hat{X}_{-1|-1} = [\hat{y}_0, 0, \ldots, 0]'$ and $P_{-1|-1} = \delta s_N^2 I$ as in section 2. Then, for $0 \leq k \leq q$, compute $P_{k|k-1}$ and $P_{k|k}$ iteratively via equations (14) and (15):

$$P_{k|k-1} = e^{\Delta k}P_{k-1|k-1}e^{\Delta k'} + \Sigma,$$

(14)
\[ P_{k|k} = \begin{cases} 
 P_{k|k-1} - \frac{1}{p_{k|k-1}} P_{k|k-1} H H' P_{k|k-1}, & \text{if } k \in \{k_0, \ldots, k_N\} \\
 P_{k|k-1}, & \text{if } k \notin \{k_0, \ldots, k_N\}, 
\end{cases} \tag{15} \]

where \( p_{k|k-1}^{(i,j)} \) is the \((i, j)\)th element of \( P_{k|k-1} \); \( \Sigma = V - e^{\hat{A}_V} V e^{\hat{A}_V'} \) and \( V \) is the solution of the matrix equation \( AV + V A' = -\sigma^2 \delta \delta' \). The preceding result on \( \Sigma \) is well-known for the stationary case but it also holds for the non-stationary case (see Tsai and Chan, 1999a, for a proof).

For \( 0 \leq k \leq q \), compute \( \hat{X}_{k|k-1} \) and \( \dot{X}_{k|k} \) iteratively via equations (16) and (17).

\[ \dot{X}_{k|k-1} = \mu + e^{\hat{A}} (\hat{X}_{k-1|k-1} - \mu) \tag{16} \]

\[ \dot{X}_{k|k} = \begin{cases} 
 \hat{X}_{k|k-1} + \frac{1}{p_{k|k-1}} P_{k|k-1} H (y_k - \hat{X}_{k|k-1}), & \text{if } k \in \{k_0, \ldots, k_N\} \\
 \hat{X}_{k|k-1}, & \text{if } k \notin \{k_0, \ldots, k_N\}, 
\end{cases} \tag{17} \]

where \( \mu = -\alpha H / \alpha_1 \). Next, go through the Kalman filter backward, we can compute the conditional means and the variances of \( X_q \)'s given all observed data via the following recursive equations:

\[ \hat{X}_{k|q} = \hat{X}_{k|k} + B_k (\hat{X}_{k+1|q} - \hat{X}_{k+1|k}), \tag{18} \]

\[ P_{k|q} = P_{k|k} + B_k (P_{k+1|q} - P_{k+1|k}) B_k', \tag{19} \]

where \( B_k = P_{k|k} e^{\hat{A}_P} P^{-1}_{k+1|k} \), \( k = q - 1, \ldots, 0 \).

The M-step can be carried out by solving the equations obtained by equating the right hand sides of equations (12) and (13) to zero. The integrals in equations (12) and (13) can be approximated by Euler’s method (see Tsai and Chan, 1999b). Alternatively, we may approximate the integrals by the trapezium method. Using the trapezium method and the fact that for any \( a < b \)

\[ \int_a^b X_a^{(r-1)} dX_a^{(p-1)} = X_b^{(r-1)} X_b^{(p-1)} - X_a^{(r-1)} X_a^{(p-1)} - \int_a^b X_a^{(p-1)} dX_a^{(r-1)} \]

\[ = X_b^{(r-1)} X_b^{(p-1)} - X_a^{(r-1)} X_a^{(p-1)} - \int_a^b X_a^{(p-1)} X_s^{(r)} d s, \]

\[ r = 1, \ldots, p-1, \]

\[ \int_a^b X_s^{(p-1)} dX_a^{(p-1)} = \frac{1}{2} \left[ (X_b^{(p-1)})^2 - (X_a^{(p-1)})^2 - \sigma^2 (b - a) \right], \]

\[ \int_a^b \left( X_a^{(p-1)} \right)^2 dX_a^{(p-1)} = \frac{1}{3} \left[ (X_b^{(p-1)})^3 - (X_a^{(p-1)})^3 - 3 \sigma^2 \int_a^b X_a^{(p-1)} dX_a^{(p-1)} \right]. \]
equations (12) and (13) become

\[
\frac{\partial l_v}{\partial \alpha_0} = \frac{1}{\sigma^2} \left[ \delta'_{p}(\hat{X}_{q|q} - \hat{X}_{0|q}) - \frac{1}{2m} \sum_{j=1}^{q} \left\{ (\alpha_0 + \alpha' X_{j-1}) + (\alpha_0 + \alpha' X_j) \right\} y \right] \\
\frac{\partial l_v}{\partial \alpha_r} = \frac{1}{\sigma^2} \delta'_{p} \left[ E(X_q X'_q | y) - E(X_0 X'_0 | y) \right] \delta_p + \frac{1}{2m} \sum_{j=1}^{q} \left[ (\alpha_0 + \alpha' X_{j-1}) X'_{j-1} \delta_p - (\alpha_0 + \alpha' X_j X'_j) \delta_{j=1} \right] \\
\frac{\partial l_v}{\partial \alpha_p} = \frac{1}{2m} \delta'_{p} \left[ E(X_q X'_q | y) - E(X_0 X'_0 | y) \right] \delta_p - \frac{1}{2} (t_n - t_0) \\
\frac{\partial l_v}{\partial \alpha_0} = \frac{1}{2m} \delta'_{p} \left[ (\alpha_0 + \alpha' X_{j-1}) X'_{j-1} \delta_p + (\alpha_0 + \alpha' X_j X'_j) \delta_{j=1} \right].
\]

Setting equations (20), (21) and (22) to zero yields the approximate maximum likelihood estimate \( \hat{\theta} = (\hat{\alpha}_0, ..., \hat{\alpha}_p) = (\hat{\alpha}_0, \hat{\alpha}) \) as the solution of the linear equations:

\[
\begin{bmatrix}
2q \\
\sum_{j=1}^{q} (\hat{X}'_{j-1|q} + \hat{X}'_{j|q}) \\
\sum_{j=1}^{q} (\hat{X}'_{j-1|q} + \hat{X}'_{j|q})
\end{bmatrix}
\begin{bmatrix}
\hat{\alpha}_0 \\
\hat{\alpha}
\end{bmatrix}
= \left[ 2m(\hat{X}_{q|q} - \hat{X}_{0|q}) \delta_p b_1 \cdots b_p \right]'.
\]

where

\[
\Gamma_{1,j} = E(X_{j-1} X'_{j-1} | y) + E(X_j X'_j | y) \\
= P_{j-1|q} + \hat{X}_{j-1|q} \hat{X}'_{j-1|q} + P_{j|q} + \hat{X}_{j|q} \hat{X}'_{j|q},
\]

and

\[
b_r = 2m \delta'_{p} \left[ E(X_q X'_q | y) - E(X_0 X'_0 | y) \right] \delta_p - \delta'_{r+1} \sum_{j=1}^{q} \Gamma_{1,j} \delta_p,
\]

10
\[ r = 1, \ldots, p - 1, \]
\[ b_p = m \delta_p \left[ E(X_{q|q}X^\prime |y) - E(X_0X^\prime |y) \right] \delta_p - ma^2(t_N - t_0). \]

Given \( \hat{\theta} = (\hat{\alpha}_0, \alpha) \), the (approximate) maximum likelihood estimate \( \hat{\sigma}^2 \) is computed by (9). Thus, a sequence of EM iterates of \((\theta, \sigma^2)\) can now be constructed as follows. Given the \( j \)th iterate \((\hat{\theta}_j, \hat{\sigma}_j^2)\), compute \( \hat{X}_{k|q} \) and \( P_{q|q^*} \) for \( 0 \leq k \leq q \), using equations (14) to (19). Then solve equation (23) to get a new set of estimates \( \hat{\theta}_{j+1} \). Let \( P_{t-1|t-1}^* = P_{t-1|t-1}/\hat{\sigma}_j^2 = \delta\tilde{s}_q I/\hat{\sigma}_j^2 \). Then with \( \hat{\theta}_{j+1} \) and making use of equation (9), we can get \( \hat{\sigma}_{j+1}^2 \). It remains to specify the initial estimates, which will be described below.

- **Initial estimates**

Here, we describe how to get a set of initial estimates of the parameters.

1. Note that \( X_{t_j}^{(0)} = Y_{t_j} \), for all \( j \).
2. If \( p > 1 \), then for \( i \) from 1 to \( p - 1 \), estimate \( X_{t_j}^{(i)} \) by \( (X_{t_j}^{(i-1)} - X_{t_{j-1}}^{(i-1)})/\Delta t_j \), for \( i \leq j \leq N \), iteratively, where \( \Delta t_j = t_j - t_{j-1} \).

Now, use the trapezium method again, but set the discretization interval to be \( \Delta t_i \) between \( t_{j-1} \) and \( t_j \), and then we can approximate equations (12) and (13) by the following equations.

\[
\frac{\partial l}{\partial \alpha_0} = \frac{1}{\sigma^2} \left[ X_{t_N}^{(p-1)} - X_{t_{p-1}}^{(p-1)} - \frac{1}{2} \sum_{j=p}^{N} \left\{ (\alpha_0 + \alpha X_{t_{j-1}}) + (\alpha_0 + \alpha X_{t_j}) \right\} \Delta t_j \right]
\]

\[
\frac{\partial l}{\partial \alpha_r} = \frac{1}{\sigma^2} \left( X_{t_N}^{(r-1)} X_{t_N}^{(p-1)} - X_{t_{p-1}}^{(r-1)} X_{t_{p-1}}^{(p-1)} \right)
\]

\[
- \frac{1}{2\sigma^2} \sum_{j=p}^{N} \left( X_{t_{j-1}}^{(p-1)} X_{t_{j-1}}^{(r)} + X_{t_{j}}^{(p-1)} X_{t_{j}}^{(r)} \right) \Delta t_j
\]

\[
- \frac{1}{2\sigma^2} \sum_{j=p}^{N} \left\{ (\alpha_0 + \alpha X_{t_{j-1}}) X_{t_{j-1}}^{(r-1)} + (\alpha_0 + \alpha X_{t_{j}}) X_{t_{j}}^{(r-1)} \right\} \Delta t_j,
\]

\( r = 1, 2, \ldots, p - 1, \)

\[
\frac{\partial l}{\partial \alpha_p} = \frac{1}{2\sigma^2} \left[ (X_{t_N}^{(p-1)})^2 - (X_{t_{p-1}}^{(p-1)})^2 - \sigma^2(t_N - t_{p-1}) \right]
\]

\[
- \frac{1}{2\sigma^2} \sum_{j=p}^{N} \left\{ (\alpha_0 + \alpha X_{t_{j-1}}) X_{t_{j-1}}^{(p-1)} + (\alpha_0 + \alpha X_{t_{j}}) X_{t_{j}}^{(p-1)} \right\} \Delta t_j.
\]
Setting the above equations to zero yields the initial estimate \( \tilde{\theta}_0 = (\tilde{\alpha}_0, \tilde{\alpha}) \) as the solution of the linear equations:

\[
\begin{bmatrix}
2 \sum_{j=p}^{N} \Delta t_j \\
\sum_{j=p}^{N} (X_{t_{j-1}} + X_{t_j}) \Delta t_j \\
\sum_{j=p}^{N} (X_{t_{j-1}} X_{t_{j-1}}' + X_{t_j} X_{t_j}') \Delta t_j
\end{bmatrix}
\begin{bmatrix}
\tilde{\alpha}_0 \\
\tilde{\alpha}
\end{bmatrix}
= \left[2(X^{(p-1)}_{t_N} - X^{(p-1)}_{t_{p-1}}) d_1 \cdots d_p \right] \text{,}
\tag{25}
\]

where

\[
d_r = 2 \left(X^{(r-1)}_{t_N} X^{(r-1)}_{t_{N-1}} - X^{(r-1)}_{t_{p-1}} X^{(r-1)}_{t_{p-1}} \right)
- \sum_{j=p}^{N} \left(X^{(r-1)}_{t_{j-1}} X^{(r)}_{t_{j-1}} + X^{(r-1)}_{t_j} X^{(r)}_{t_j} \right) \Delta t_j, \quad r = 1, \ldots, p - 1,
\]

\[
d_p = \left(X^{(p-1)}_{t_N}\right)^2 - \left(X^{(p-1)}_{t_{p-1}}\right)^2 - \sigma^2(t_N - t_{p-1}),
\]

and the initial estimate of \( \sigma^2 \) is again given by \( \tilde{\sigma}^2 = (N+1)^{-1} \sum_{i=0}^{N} \tilde{y}_i^2 / p_{*,t_{i-1}}^{(1,1)} \), where \( p_{*,t_{i-1}}^{(1,1)} \) is defined as before, while \( P_{*,t_{i-1}}^{*} = \delta s^2 t / \tilde{\sigma}^2 \) and \( \tilde{\sigma}^2 = N^{-1} \sum_{i=1}^{N} (X_{t_i} - X_{t_{i-1}})^2 / (t_i - t_{i-1}) \).

4 Simulations

Regularly spaced time series data, \( Y_{t_i} = X_i, i = 0, 1, 2, \ldots, 100 \), are simulated from the following stationary CAR(\( p \)) processes via equations (4) and (5):

Model 1, \( dX_t = -X_t dt + dW_t \),

Model 2, \( dX_t^{(1)} = (-2X_t - X^{(1)}_t) dt + dW_t \).

For model 1, the conditional maximum likelihood estimator has a closed form solution. See Tsai and Chan (1999b) for details. The averages, the standard deviations and the mean squared errors (MSEs) of 1000 replications of the conditional maximum likelihood estimators, the EM estimators with both Euler’s method and the trapezium approximation method are given in Table 1.

From Table 1, we see that with increasing discretization frequency, the EM estimators based on the trapezium approximation method converge to the conditional maximum likelihood estimator much faster than the EM estimators based on Euler’s approximation method. In fact, with the trapezium method, the EM estimators remain the same for the discretization frequency.
Table 1: Averages, (standard deviations) and (MSEs) of 1000 simulations of the conditional ML and EM estimators of Model 1 (m = 1, 5, 10, 20, 40 and 80)

<table>
<thead>
<tr>
<th>value of m</th>
<th>True Value</th>
<th>Euler’s method</th>
<th>trapezium method</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial estimate</td>
<td>α₀</td>
<td>0</td>
<td>-0.0001(0.0690)(0.0048)</td>
</tr>
<tr>
<td></td>
<td>α₁</td>
<td>-1</td>
<td>-0.6515(0.0886)(0.1293)</td>
</tr>
<tr>
<td></td>
<td>σ²</td>
<td>1</td>
<td>0.7825(0.1317)(0.0647)</td>
</tr>
<tr>
<td>EM (m=1)</td>
<td>α₀</td>
<td>0</td>
<td>-0.0001(0.0690)(0.0048)</td>
</tr>
<tr>
<td></td>
<td>α₁</td>
<td>-1</td>
<td>-0.6515(0.0886)(0.1293)</td>
</tr>
<tr>
<td></td>
<td>σ²</td>
<td>1</td>
<td>0.7830(0.1318)(0.0645)</td>
</tr>
<tr>
<td>EM (m=5)</td>
<td>α₀</td>
<td>0</td>
<td>-0.0002(0.0906)(0.0082)</td>
</tr>
<tr>
<td></td>
<td>α₁</td>
<td>-1</td>
<td>-0.8519(0.1421)(0.0421)</td>
</tr>
<tr>
<td></td>
<td>σ²</td>
<td>1</td>
<td>0.8889(0.1620)(0.0386)</td>
</tr>
<tr>
<td>EM (m=10)</td>
<td>α₀</td>
<td>0</td>
<td>-0.0003(0.0989)(0.0098)</td>
</tr>
<tr>
<td></td>
<td>α₁</td>
<td>-1</td>
<td>-0.9268(0.1741)(0.0357)</td>
</tr>
<tr>
<td></td>
<td>σ²</td>
<td>1</td>
<td>0.9346(0.1806)(0.0369)</td>
</tr>
<tr>
<td>EM (m=20)</td>
<td>α₀</td>
<td>0</td>
<td>-0.0003(0.1055)(0.0111)</td>
</tr>
<tr>
<td></td>
<td>α₁</td>
<td>-1</td>
<td>-0.9849(0.2062)(0.0427)</td>
</tr>
<tr>
<td></td>
<td>σ²</td>
<td>1</td>
<td>0.9723(0.2093)(0.0409)</td>
</tr>
<tr>
<td>EM (m=40)</td>
<td>α₀</td>
<td>0</td>
<td>-0.0004(0.1103)(0.0122)</td>
</tr>
<tr>
<td></td>
<td>α₁</td>
<td>-1</td>
<td>-1.0257(0.2353)(0.0560)</td>
</tr>
<tr>
<td></td>
<td>σ²</td>
<td>1</td>
<td>1.0001(0.2191)(0.0480)</td>
</tr>
<tr>
<td>EM (m=80)</td>
<td>α₀</td>
<td>0</td>
<td>-0.0005(0.1135)(0.0129)</td>
</tr>
<tr>
<td></td>
<td>α₁</td>
<td>-1</td>
<td>-1.0523(0.2594)(0.0700)</td>
</tr>
<tr>
<td></td>
<td>σ²</td>
<td>1</td>
<td>1.0188(0.2354)(0.0558)</td>
</tr>
<tr>
<td>conditional MLE</td>
<td>α₀</td>
<td>0</td>
<td>-0.0005(0.1185)(0.0140)</td>
</tr>
<tr>
<td></td>
<td>α₁</td>
<td>-1</td>
<td>-1.0936(0.3050)(0.1018)</td>
</tr>
<tr>
<td></td>
<td>σ²</td>
<td>1</td>
<td>1.0509(0.2639)(0.0722)</td>
</tr>
</tbody>
</table>
$m$ larger than 10 in terms of the averages and the standard deviations of 1000 replications. For $m$ larger than 10, there is a minor difference between the EM estimators based on the trapezium method and the conditional maximum likelihood estimator. This can be explained as follows: The objective function, $-2l_Y(\theta, \sigma^2)$, is the same for these two methods up to contributions from the initial conditions.

We repeat the same simulation study for model 2. The results are given in Table 2. The conditional maximum likelihood estimator of model 2 does not have a closed form solution. So, the conditional maximum likelihood estimators are replaced by the maximum likelihood estimators obtained from the simplex nonlinear optimization method. The simplex method is executed by applying the DUMPOL subroutine of the Microsoft Fortran PowerStation IMSL package to minimize the right hand side of equation (10). For a detailed description of the implementation of the DUMPOL subroutine for estimating continuous-time autoregressive models, see Tsai and Chan (1999b).

From Table 2, we also see that the EM estimators based on the trapezium approximation method converge to the maximum likelihood estimator much faster than the EM estimators based on Euler’s method. With the trapezium method, the EM estimators seem to converge to a limit soon after $m = 20$, but this is not the case with Euler’s method. The EM estimators based on the trapezium method seem to converge to the ML estimator generated by the simplex method in terms of biases. The standard deviations of the EM estimators based on the trapezium method are also close to those of the ML estimator, although the standard deviations of the former seem to be larger than those of the latter a bit. It is also interesting to note that with an appropriate discretization frequency, the EM estimator based on Euler’s method outperforms the EM estimator based on the trapezium method and the ML estimator in terms of biases and standard deviations of the estimates.

Tsai and Chan (1999b) proved an invariance property of the EM estimators under linear transformations when Euler’s approximation method is used. It can be shown that the invariance property also holds when the trapezium approximation method is used. The proof is similar to that of Tsai and Chan (1999b), although it is more tedious. Now, we state the results in the following theorem without proof.

**Theorem 4.1 (Invariance property of the EM estimators under linear transformations)**

Let $\{Y_{t_j}\}_{j=0,1,...,N}$ be a series of discrete-time data sampled from a CAR($p$) model defined by

$$dX_t^{(p-1)} = (\alpha_0 + \alpha' X_t)dt + \sigma dW_t.$$
Table 2: Averages, (standard deviations) and (MSEs) of 1000 simulations of the ML and EM estimators of Model 2 (m = 1, 2, 5, 10 and 20)

<table>
<thead>
<tr>
<th>value of m</th>
<th>True Value</th>
<th>Euler’s method</th>
<th>trapezium method</th>
</tr>
</thead>
<tbody>
<tr>
<td>initial estimate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_0$</td>
<td>0</td>
<td>0.0009(0.0475)(0.0023)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-2</td>
<td>-0.9314(0.1108)(1.1542)</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2$</td>
<td>1</td>
<td>0.7549(0.1385)(0.0793)</td>
</tr>
<tr>
<td>EM (m=1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_0$</td>
<td>0</td>
<td>0.0009(0.0448)(0.0020)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-2</td>
<td>-0.8807(0.0806)(1.2593)</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2$</td>
<td>1</td>
<td>1.0717(0.1829)(0.0386)</td>
</tr>
<tr>
<td>EM (m=2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_0$</td>
<td>0</td>
<td>0.0015(0.0737)(0.0054)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-2</td>
<td>-1.4462(0.1467)(0.3282)</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2$</td>
<td>1</td>
<td>-1.1348(0.0929)(0.0268)</td>
</tr>
<tr>
<td>EM (m=5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_0$</td>
<td>0</td>
<td>0.0019(0.0923)(0.0085)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-2</td>
<td>-1.8090(0.1971)(0.0753)</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2$</td>
<td>1</td>
<td>-1.1050(0.1475)(0.0328)</td>
</tr>
<tr>
<td>EM (m=10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_0$</td>
<td>0</td>
<td>0.0020(0.0984)(0.0097)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-2</td>
<td>-1.9277(0.2164)(0.0521)</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2$</td>
<td>1</td>
<td>-1.0704(0.1861)(0.0396)</td>
</tr>
<tr>
<td>EM (m=20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_0$</td>
<td>0</td>
<td>0.0021(0.1013)(0.0103)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-2</td>
<td>-1.9848(0.2276)(0.0520)</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2$</td>
<td>1</td>
<td>-1.0452(0.2153)(0.0484)</td>
</tr>
<tr>
<td>simplex method</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_0$</td>
<td>0</td>
<td>0.0021(0.1041)(0.0108)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>-2</td>
<td>-2.0425(0.2409)(0.0598)</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2$</td>
<td>1</td>
<td>-1.0194(0.2557)(0.0658)</td>
</tr>
</tbody>
</table>
Consider \( \{Y_t^*\}_{j=0,\ldots,N} \), where \( Y_t^* = aY_t + b, \forall j \), and \( a \neq 0 \), a series of discrete-time data sampled from the corresponding CAR\((p)\) model defined by

\[
dX_t^{*\,(p-1)} = (\alpha_0^* + \alpha^* X_t^*)dt + \sigma^* dW_t.
\]

Then, the two CAR models are related as follows: \( (\alpha_0^*, \alpha^*, \sigma^*) = (a\alpha_0 - b\alpha_1, \alpha, a\sigma) \). Furthermore, for a fixed \( m \), let \( \{\hat{\alpha}_j\}_{j=0,\ldots,p} \) and \( \hat{\sigma}^2 \) be the EM estimators of \( \{\alpha_j\}_{j=0,\ldots,p} \) and \( \sigma^2 \) based on \( \{Y_t\} \); similarly, \( \{\hat{\alpha}_j^*\}_{j=0,\ldots,p} \) and \( \hat{\sigma}^{*2} \) denote the EM estimators of \( \{\alpha_j^*\}_{j=0,\ldots,p} \) and \( \sigma^{*2} \) based on \( \{Y_t^*\} \). Then, \( (\hat{\alpha}_0^*, \hat{\alpha}^*, \hat{\sigma}^*) = (a\hat{\alpha}_0 - b\hat{\alpha}_1, \hat{\alpha}, a\hat{\sigma}) \).

The above invariance properties imply that the choice of \( m \) is independent of the scale of the data.

5 Conclusion

We have applied the trapezium approximation method in an implementation of the EM algorithm proposed by Tsai and Chan (1999b) for estimating continuous-time autoregressive models. Simulation studies confirm that the EM estimators converge much faster to the ML estimator with the trapezium method than with Euler’s method. Tsai and Chan (1999c) proposed a likelihood based approach for testing for nonlinearity with irregularly sampled data. Tsai and Chan (1999c) used Euler’s method for approximating the integrals involved in calculating the score and the information matrix. It is interesting to investigate whether or not the power of their test may be enhanced by an alternative implementation of the test via the trapezium method.

Acknowledgments

The work of Tsai was partially supported by the Institute of Statistical Science, Academia Sinica, Taipei, Taiwan. Part of this work was done during Tsai’s visit at the Institute of Statistical Science, Academia Sinica, Taipei, Taiwan and he gratefully thanks the institute for providing a very productive environment and computer facilities.

References


18. Tsai, H. and Chan, K. S. (1999a), A note on the covariance structure of


