A NEW METHOD OF ESTIMATION IN A MOVING AVERAGE MODEL OF ORDER ONE

H.A. Niroomand Chapeh & M. Bhaskara Rao
Center for Multivariate Analysis
University of Pittsburgh

Center for Multivariate Analysis
University of Pittsburgh
A NEW METHOD OF ESTIMATION IN A MOVING AVERAGE MODEL OF ORDER ONE

H.A. Niroomand Chapeh & M. Bhaskara Rao
Center for Multivariate Analysis
University of Pittsburgh

Technical Report No. 86-46
December 1986

Center for Multivariate Analysis
Fifth Floor, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

This work is supported by Contract N00014-85-K-0292 of the Office of Naval Research and Contract F49620-85-C-0008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.
A NEW METHOD OF ESTIMATION IN A MOVING AVERAGE MODEL OF ORDER ONE

H.A. Niroomand Chapeh

and

M. Bhaskara Rao

Sheffield University

and

Center for Multivariate Analysis

University of Pittsburgh

Summary

The exact likelihood of the data coming from a moving average model of order is complicated. In this paper, we propose a method of estimation of the parameters of a moving average model of order one based on the approximate likelihood of the data and on the simulation of a pair of random variables. Some comparisons were made of this method with some well known methods for moderate sample sizes. A computer program is appended which is helpful in using this method.

Key words: Moving average model, Exact Likelihood, Approximate Likelihood, Simulation.
A NEW METHOD OF ESTIMATION IN A MOVING AVERAGE MODEL OF ORDER ONE

H.A. Niroomand Chapeh
and
M. Bhaskara Rao
Sheffield University
and
Center for Multivariate Analysis,
University of Pittsburgh

1. Introduction

Estimation of parameters of a moving average model is a delicate subject fraught with complications. There are several methods available in the literature for estimating the parameters of a moving average model. Many of these methods require an enormous amount of computational effort. Though most of these methods provide consistent estimators of the parameters, it is of some interest to compare their performance for moderate sample sizes. McClave (1974) compared the performance of some five methods of estimation (Durbin (1959), Walker (1961), Hannan (1969), Parzen and Clevenson (1970), and Anderson (1973)) by simulating a first-order moving average model

\[ Y_t = e_t + \beta e_{t-1}, \quad t = 0, 1, 2, \ldots, \]

where \( e_t, t = 0, 1, 2, \ldots \) is an independent identically distributed gaussian process with mean zero and variance \( \sigma^2 = 1 \), for \( \beta = 0.5 \) and 0.9, taking the sample size \( N = 100 \). McClave (1974) recommends the use of Parzen-Clevenson's, Hannan's and Walker's methods in that order for estimation. All these methods use approximations of one kind or another for the true expressions involved. As a result, these methods have inherently a certain
degree of complexity in their execution requiring a careful choice of trial approximations. Moreover, looking at the simulation studies of McClave (1974), one notices that $\theta$ being systematically under-estimated and this being more so in the case $\theta = 0.9$.

In this paper, we propose a new method of estimation of the parameter $\theta$ and $\sigma^2$ in a moving average model of order one, and examine its performance vis-a-vis with those of five methods considered by McClave (1974). This method is much simpler to use and is as good as any of these five methods. We apply this method to the time series data "IBM closing stock prices" and fit a moving average model of order one to these data. See Box and Jenkins (1976, p.526). We also give the computer program used in the estimation of $\theta$ and $\sigma^2$.

In an article under preparation, we show how this new method can be used to make inferences in the context of signal processing when the input forms a moving average model of order one.

2. Estimation

The basic problem is to estimate $\theta$ and $\sigma^2$ based on a single realization $y_i$ of $Y_i$, $i = 1,2,\ldots,N$. The random vector $Y = (Y_1,Y_2,\ldots,Y_N)$ has a $N$-variate normal distribution with mean vector $0$ and dispersion matrix

\[
\Sigma_N = \begin{bmatrix}
\sigma^2 & 1+\theta^2 & \theta & 0 & 0 & \cdots & 0 & 0 \\
\theta & 1+\theta^2 & \theta & 0 & \cdots & 0 & 0 \\
0 & \theta & 1+\theta^2 & \theta & \cdots & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & 0 & 0 & \cdots & 1+\theta^2 & \theta \\
0 & 0 & 0 & 0 & \cdots & \theta & 1+\theta^2 \\
\end{bmatrix}
\]
The likelihood of \( y_1, y_2, \ldots, y_N \) involves \( \Sigma_N^{-1} \). The inverse of \( \Sigma_N \) can be determined exactly but the exact likelihood of \( y_1, y_2, \ldots, y_N \) is too unwieldy to handle. See Shaman (1969). In many of the methods currently used in estimating \( \beta \) of the moving average model of order one, \( \Sigma_N \) is replaced by a matrix whose inverse has a simple form. See Shaman (1969), McClave (1974), Godolphin (1978) and Godolphin and de Gooijer (1982). The replacement considered is given by

\[
\Delta_N = \begin{bmatrix}
\sigma^2 & 1 & \beta & 0 & 0 & \cdots & 0 & 0 & 0 \\
1 & 1+\beta^2 & \beta & 0 & \cdots & 0 & 0 & 0 \\
0 & \beta & 1+\beta^2 & \beta & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 1+\beta^2 & \beta \\
0 & 0 & 0 & 0 & \cdots & 0 & 1
\end{bmatrix}.
\]

The inverse of \( \Delta_N \) is given by

\[
\Delta_N^{-1} = \left[ \frac{1}{\sigma^2(1-\beta^2)} \right] \begin{bmatrix}
1 & -\beta & \beta^2 & -\beta^3 & \cdots & (-\beta)^{N-2} & (-\beta)^{N-1} \\
-\beta & 1 & -\beta & \beta^2 & \cdots & (-\beta)^{N-3} & (-\beta)^{N-2} \\
\beta^2 & -\beta & 1 & -\beta & \cdots & (-\beta)^{N-4} & (-\beta)^{N-3} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
(-\beta)^{N-1} & (-\beta)^{N-2} & (-\beta)^{N-3} & (-\beta)^{N-4} & \cdots & -\beta & 1
\end{bmatrix}.
\]

We call the resultant likelihood of \( y_1, y_2, \ldots, y_N \) using \( \Delta_N \) as "approximate likelihood".

Whittle (1953) has shown that the maximum likelihood estimator (using exact likelihood) of \( \beta \) is consistent, efficient and asymptotically
normal. (In all these deliberations, we assume that \(|\beta| < 1\).) It can be shown that the estimator of \(\beta\) obtained using the approximate likelihood is consistent. See Niroomand Chapeh and Bhaskara Rao (1987). However, for moderate sample sizes the performance of the estimator based on approximate likelihood does not seem to be satisfactory for values of \(\beta\) close to zero and to one. We have simulated the moving average model of order one for \(\beta = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\) and \(\sigma^2 = 1\) taking \(N = 100\), and then used the approximate likelihood to estimate \(\beta\) and \(\sigma^2\). The symbols \(\bar{\beta}\) and \(\bar{\sigma}^2\) indicate average of estimates of \(\beta\) and \(\sigma^2\) (at 1), respectively, taken over 100 repetitions of the above procedure. The symbols \(\text{var}(\hat{\beta})\) and \(\text{var}(\hat{\sigma}^2)\) indicate the variances of these 100 estimates of \(\beta\) and \(\sigma^2\), respectively.

**TABLE 1: Mean and variances of the estimates using approximate likelihood**

<table>
<thead>
<tr>
<th>True Value of (\beta)</th>
<th>(\bar{\beta})</th>
<th>(\text{var}(\hat{\beta}))</th>
<th>(\bar{\sigma}^2)</th>
<th>(\text{var}(\hat{\sigma}^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.086446</td>
<td>0.013927</td>
<td>0.972431</td>
<td>0.021727</td>
</tr>
<tr>
<td>0.2</td>
<td>0.175466</td>
<td>0.009216</td>
<td>1.012012</td>
<td>0.022227</td>
</tr>
<tr>
<td>0.3</td>
<td>0.308813</td>
<td>0.010092</td>
<td>0.974329</td>
<td>0.016091</td>
</tr>
<tr>
<td>0.4</td>
<td>0.403180</td>
<td>0.008703</td>
<td>0.979329</td>
<td>0.017651</td>
</tr>
<tr>
<td>0.5</td>
<td>0.502262</td>
<td>0.009645</td>
<td>0.994639</td>
<td>0.014661</td>
</tr>
<tr>
<td>0.6</td>
<td>0.604383</td>
<td>0.006919</td>
<td>0.980099</td>
<td>0.021174</td>
</tr>
<tr>
<td>0.7</td>
<td>0.695617</td>
<td>0.004870</td>
<td>0.984986</td>
<td>0.020208</td>
</tr>
<tr>
<td>0.8</td>
<td>0.777597</td>
<td>0.004416</td>
<td>1.008638</td>
<td>0.019917</td>
</tr>
<tr>
<td>0.9</td>
<td>0.857607</td>
<td>0.003533</td>
<td>1.076121</td>
<td>0.030506</td>
</tr>
</tbody>
</table>

Compare the above with the simulations performed by McClave (1974) for the cases \(\beta = 0.5\) and \(0.9\). See also Godolphin and de Gooijer (1982).
The new method we want to propose has the following theoretical backing. Let \( X^T = (X_1, X_2, \ldots, X_N) \) be a random vector normally distributed with mean vector \( \vec{0} \) and dispersion matrix \( \Lambda_N \). Let \( Z^T = (Z_1, Z_2, \ldots, Z_N) \) be a random vector normally distributed with mean vector \( \vec{0} \) and dispersion matrix given by

\[
\Sigma_N = \begin{bmatrix}
\sigma^2 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \sigma^2
\end{bmatrix}
\]

The random variables \( Z_1, Z_2, \ldots, Z_N \) are independently normally distributed with \( Z_2, Z_3, \ldots, Z_{N-1} \) being degenerate, all with common mean zero and the variance of each of \( Z_1 \) and \( Z_N \) being equal to \( \beta^2 \sigma^2 \). Assume that \( X \) and \( Z \) are independent. It then follows that \( Y \overset{d}{=} X + Z \), where \( \overset{d}{=} \) indicates equality in distribution. If we have a realization \( z \) of \( Z \), then we could obtain a realization \( x \) of \( X \) using \( y \) (given data) and \( z \) by taking \( x = y - z \). Then we can work with the exact likelihood of \( x \) for estimating \( \beta \) and \( \sigma^2 \) instead of working with the approximate likelihood of \( y \) for the estimation of \( \beta \) and \( \sigma^2 \). One way of obtaining a realization of \( Z \) is by simulating \( Z \) once. The simulation of \( Z_1 \) and \( Z_N \) involves the unknown parameters \( \beta \) and \( \sigma^2 \). We can use some decent estimates of \( \beta \) and \( \sigma^2 \) to simulate \( Z_1 \) and \( Z_N \). This is the basic idea behind the new method. Schematically, we outline the procedure as follows.
Step 1. Estimate $\beta$ and $\sigma^2$ using the approximate likelihood of $y_1, y_2, \ldots, y_N$. Let $\hat{\beta}(1)$ and $\hat{\sigma}^2(1)$ be the respective estimates.

Step 2. Simulate $Z_1$ and $Z_N$ each once, where $Z_1$ and $Z_N$ are independently identically normally distributed with mean 0 and variance $\hat{\sigma}^2(1)$. Let the realizations be denoted by $z_1$ and $z_N$.

Step 3. Define $x_i = y_i$ for $i = 2, 3, \ldots, N-1$,
        $x_1 = y_1 - z_1$ for $i = 1, N$.

Step 1*. View $x^T = (x_1, x_2, \ldots, x_N)$ as a realization of $X^T = (X_1, X_2, \ldots, X_N)$.
Estimate $\beta$ and $\sigma^2$ using the likelihood of $x$. Let these estimates be denoted by $\hat{\beta}(2)$ and $\hat{\sigma}^2(2)$.

Keep repeating Steps 2 and 3 until two consecutive $\hat{\beta}$'s and $\hat{\sigma}^2$'s agree in a specified number of decimal places.

Some of the theoretical properties of these estimators are currently under investigation by the authors.

3. Simulation Studies

We simulated a moving average model of order one with $\beta = 0.1 \ (0.1)$, $\sigma^2 = 1$ taking $N = 100$, and used the above procedure to get estimates $\hat{\beta}$ and $\hat{\sigma}^2$ of $\beta$ and $\sigma^2$, respectively. This is repeated 100 times. The following table gives the averages and variances of these estimates over 100 repetitions.
Table 2: Mean and variances of the estimates obtained by using a combination of approximate likelihood and simulation of two dummy variables

<table>
<thead>
<tr>
<th>True value of ( \beta )</th>
<th>( \bar{\beta} )</th>
<th>var(( \bar{\beta} ))</th>
<th>( \bar{\sigma}^2 )</th>
<th>var(( \bar{\sigma}^2 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.0934</td>
<td>0.0139</td>
<td>0.9935</td>
<td>0.0227</td>
</tr>
<tr>
<td>0.2</td>
<td>0.1908</td>
<td>0.0084</td>
<td>0.9916</td>
<td>0.0191</td>
</tr>
<tr>
<td>0.3</td>
<td>0.3016</td>
<td>0.0076</td>
<td>1.0216</td>
<td>0.0226</td>
</tr>
<tr>
<td>0.4</td>
<td>0.3992</td>
<td>0.0098</td>
<td>1.0096</td>
<td>0.0130</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5012</td>
<td>0.0085</td>
<td>0.9916</td>
<td>0.0183</td>
</tr>
<tr>
<td>0.6</td>
<td>0.6044</td>
<td>0.0081</td>
<td>1.0013</td>
<td>0.0201</td>
</tr>
<tr>
<td>0.7</td>
<td>0.6928</td>
<td>0.0072</td>
<td>1.0204</td>
<td>0.0271</td>
</tr>
<tr>
<td>0.8</td>
<td>0.7706</td>
<td>0.0042</td>
<td>1.0182</td>
<td>0.0203</td>
</tr>
<tr>
<td>0.9</td>
<td>0.8589</td>
<td>0.0034</td>
<td>1.0499</td>
<td>0.0240</td>
</tr>
</tbody>
</table>

The computer program for the above is available from the authors on request.

McClave (1974) recommends Parzen-Clevenson's method for estimating \( \beta \) with samples of moderate size. His simulation studies cover \( \beta = 0.5 \) and \( \beta = 0.9 \) only. The following table gives his results (with the best in each category) along with the ones obtained above.
Table 3: Mean and variances of estimates of $\beta$

<table>
<thead>
<tr>
<th>Method</th>
<th>True value of $\beta$</th>
<th>0.5</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Variance</td>
</tr>
<tr>
<td>Durbin's</td>
<td></td>
<td>0.4991</td>
<td>0.0104</td>
</tr>
<tr>
<td>Walker's</td>
<td></td>
<td>0.4920</td>
<td>0.0072</td>
</tr>
<tr>
<td>Hannan's</td>
<td></td>
<td>0.5025</td>
<td>0.0072</td>
</tr>
<tr>
<td>Parzen-Clevenson's</td>
<td></td>
<td>0.4912</td>
<td>0.0070</td>
</tr>
<tr>
<td>Anderson's</td>
<td></td>
<td>0.4997</td>
<td>0.0075</td>
</tr>
<tr>
<td>Approximate likelihood</td>
<td></td>
<td>0.5023</td>
<td>0.0096</td>
</tr>
<tr>
<td>Chapeh-Rao's</td>
<td></td>
<td>0.5012</td>
<td>0.0085</td>
</tr>
</tbody>
</table>

Note: In the first five methods, the averages were taken over 200 repetitions. In the last two methods, the averages were taken over 100 repetitions.

In conclusion, we remark that the execution of the new method is direct and simpler than the five methods examined by McClave (1974). There seems to be some improvement in the accuracy of the estimates obtained by using the new method than by using approximate likelihood especially in the lower reaches of $\beta$. 
4. A Case Study

Box and Jenkins (1976, p.239 and p.526) fitted a moving average model of order one for the time series data "IBM Common Stock Closing Prices - Daily, 17th May 1961 to 2nd November 1962" after obtaining the first order differences of these series. The method they have used is the iterative least squares procedure and the model works out to be

$$7 Y_t = e_t + 0.09 e_{t-1}$$

with residual variance equal to 52.2, where $7 Y_t = Y_t - Y_{t-1}$ and $Y_t$'s are the original series. We have applied the method based on a combination of approximate likelihood and the simulation of two dummy variables to the first order differences of the above time series data and obtained

$$\hat{\sigma} = 0.0867 \quad \text{and} \quad \hat{\sigma}^2 = 52.219.$$  

In Section 5, we give the computer program we have used in this connection, and this program can be used to fit moving average models of order one by the new method. This program gives both the estimates of $\hat{\sigma}$ and $\hat{\sigma}^2$. 
REFERENCES


5. COMPUTER PROGRAM

Language
FORTRAN 77

Description and purpose
Estimate the parameters of a moving average model of order one using a combination of approximate likelihood and simulation of two independent normal random variables.

Structure
SUBROUTINE GO5DDF(XM, SIG1) : gives a random sample from a normal distribution with mean XM and variance SIG1.

Formal parameters
XM INTEGER INPUT : ZERO
SIG1 REAL INPUT : ONE

SUBROUTINE CO5NBF(FUN, 2, X, F, TOL, W, 20, IFAIL)

Formal parameters
X Real array of dimension 2 Input : Initial guesses of parameters Output : gives estimates of the parameters
TOL The accuracy
W The work space
IFAIL INTEGER Output : a fault indicator equal to
IFAIL = 1 On entry, N ≤ 0, or XTOL < 0.0, or LWA ≤ N(3N + 4)
IFAIL = 2 There have been at least 200(N+1) evaluations.
IFAIL = 3 No further improvement in the solution is possible.
IFAIL = 4 The iteration is not making good progress.

AUXILIARY ALGORITHM
CO5NBF uses the subroutine FUN(L, X, F, IFAIL) : it finds the zeros of two nonlinear equations F(1) and F(2) in the unknown parameters \( \hat{\beta} \) and \( \sigma^2 \).

N = the total sample size.
Estimation of parameters of a moving average model of order one

IMPLICIT REAL*8 (A-H,O-Z)
IMPLICIT INTEGER*2 (I-N)
DIMENSION XV(N=), S(N=), X(2), W(30), F(2), YV(N=), U(2), T(N=)
COMMON S,N
EXTERNAL FUN

IFAIL = 0
X1 = 0.0D0
X2 = 0.0D0
XM = 0.0D0
TOL = DSQRT(X02AAF(0.0D0))
open(5, file = 'data')
OPEN(6, FILE = 'RESULTS')
read(5,*) XV
WRITE(6,201)
N = 
DO 4 J = 1,N
4  S(J) = 0.0
DO 5 I = 1,N
  L = N-I+1
  DO 5 K = 1,L
    J = K+I-1
5  S(I) = S(I)+XV(K)*XV(J)
X(1) = 1D-1
X(2) = 1.0D0
CALL COSNB(FUN,2,X,F,TOL,W,20,IFAIL)
IF(ABS(X(1)-X1).LT.0.1D-2.AND.ABS(X(2)-X2).LT.0.1D-2) GO TO 90
X1 = X(1)
X2 = X(2)
100 SIGI = X1*X1*X2
DO 6 I = 1,2
C Simulation of two independent random variables
U(I) = GO5DDF(XM,SIG1)
6 CONTINUE
YV(1) = XV(1) - U(1)
YV(N) = XV(N) - U(2)
DO 7 I = 2, N - 1
YV(I) = XV(I)
7 CONTINUE
DO 9 J = 1, N
T(J) = 0.0
DO 10 I = 1, N
L = N - I + 1
DO 10 K = 1, L
J = K + I - 1
10 T(I) = T(I) + YV(K) * YV(J)
X(1) = 1D-1
X(2) = 1.0D0
CALL CO5NBF(FUN, 2, X, F, TOL, W, 20, IFAIL)
IF (ABS(X(1) - X1) .LT. 0.1D-2 .AND. ABS(X(2) - X2) .LT. 0.1D-2) GO TO 90
X1 = X(1)
X2 = X(2)
GO TO 100
90 WRITE(6, 202) X1, X2
201 FORMAT("")
202 FORMAT('OFINAL RESULTS ARE :- B hat', F10.6, 'SIG.SQ', F10.6)
STOP
END
SUBROUTINE FUN(L, X, F, IFAIL)
IMPLICIT REAL*8 (A-H, O-Z)
IMPLICIT INTEGER*2 (I-N)
DIMENSION F(2), S(N=), X(2)
COMMON S, N
F(1) = - N * X(2) * (1.0D0 - X(1) * X(1)) + S(1)
F(2) = X(1) * X(2) * (1.0D0 - X(1) * X(1)) - X(1) * S(1)
DO 1 I = 2, N
F(1) = F(1) + 2.0D0 * (-1.0D0) ** (I + 1) * (X(1) ** (I - 1)) * S(I)
F(2) = F(2) + (-1.0D0) ** I * (X(1) ** (I - 2)) * (I + 3 - I) * X(1) * X(1) * S(I)
1 CONTINUE
RETURN
END
A new method of estimation in a moving average model of order one

The exact likelihood of the data coming from a moving average model of order is complicated. In this paper, we propose a method of estimation of the parameters of a moving average model of order one based on the approximate likelihood of the data on the simulation of a pair of random variables. Some comparisons were made of this method with some well known methods for moderate sample sizes. A computer program is appended which is helpful in using this method.
END
DATE FILMED
DEC. 1987