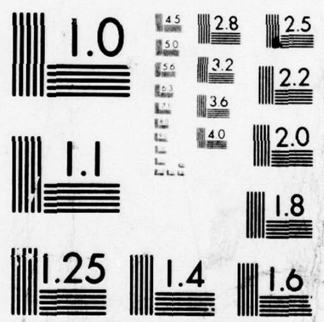
AD-A073 788 STANFORD UNIV CALIF DEPT OF STATISTICS F/G 12/1
MAXIMUM LIKELIHOOD ESTIMATION OF THE AUTOREGRESSIVE COEFFICIENT--ETC(U)
AUG 79 F AHRABI N00014-75-C-0442

UNCLASSIFIED TR-39 NL

1 OF 1

AD
A073788

END
DATE
FILMED
10-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

(12)

AD A073788

MAXIMUM LIKELIHOOD ESTIMATION OF THE AUTOREGRESSIVE COEFFICIENTS AND MOVING AVERAGE COVARIANCES OF VECTOR AUTOREGRESSIVE MOVING AVERAGE MODELS

BY

FEREYDOON AHRABI

DDC
RECEIVED
SEP 13 1979
C

TECHNICAL REPORT NO. 39
AUGUST 1979

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

DDC FILE COPY

This document has been approved for public release and sale; its distribution is unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

79 09 13 033

MAXIMUM LIKELIHOOD ESTIMATION OF THE AUTOREGRESSIVE COEFFICIENTS AND MOVING AVERAGE COVARIANCES OF VECTOR AUTOREGRESSIVE MOVING AVERAGE MODELS

DDC
RECEIVED
SEP 13 1979
C

by

FEREYDOON AHRABI

TECHNICAL REPORT NO. 39

AUGUST 1979

94p.

PREPARED UNDER CONTRACT N00014-75-C-0442

(NR-042-034)

OFFICE OF NAVAL RESEARCH

TR-39

Theodore W. Anderson, Project Director

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government.
Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

332 580

ACKNOWLEDGMENTS

I would like to express my thanks to my advisor Professor T. W. Anderson for his interest, encouragement and guidance in writing this dissertation.

I would also like to thank Professor Charles Stein and Professor Ray Faith for reading the manuscript and making helpful comments.

Finally, I would like to thank Carolyn Knutsen for a superb and accurate typing job.

Accession For
NTIS GRA&I
DDC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or special
A

Table of Contents

		Page
1.	Introduction	1
2.	The Model	4
3.	Estimation in the Time Domain	9
	3.1. The Likelihood Function	9
	3.2. The First Derivative of log $\ell(\underset{\sim}{\gamma})$	16
	3.3. The Numerical Approximations	21
	3.4. The Pure Moving Average Case	29
	3.5. The Problem of Computation	33
4.	Estimation in the Frequency Domain	41
	4.1. Preliminaries	41
	4.2. The First Derivative of Λ	44
	4.3. The Second Order Derivatives of Λ	53
	4.4. The Newton-Raphson Method	58
	4.5. The Scoring Method	59
	4.6. The Pure Moving Average Case	61
5.	Asymptotic Properties	62
6.	Estimation of the Coefficients and the Covariance Matrix of the $\underset{\sim}{\epsilon}_t$'s in the Time Domain (The Scoring Method)	67
	Appendix	77
	References	84

Maximum Likelihood Estimation of the Autoregressive Coefficients and Moving Average Covariances of Vector Autoregressive Moving Average Models

Fereydoon Ahrabi
Stanford University

1. Introduction.

The purpose of this paper is to derive asymptotically efficient estimates for the autoregressive matrix coefficients and moving average covariance matrices of the vector autoregressive moving average (VARMA) models in both time and frequency domains. To do this we shall apply the Newton-Raphson and scoring methods to the maximum likelihood equations derived from modified likelihood functions under the Gaussian assumption.

The parameterization in this paper differs from that of other works in the vector case, except Ahrabi (1978) which deals with the same estimation problems in the vector moving average case, and it follows that of Anderson (1975), Parzen (1971), and Clevenson (1970) in the scalar case. The usual parameterization of VARMA models is in terms of the autoregressive and moving average coefficients and the covariance matrix of the disturbance vector $\underset{\sim}{\epsilon}_t$. With this parameterization, Hannan (1969b,1970) has considered the problem in the pure moving average case in the frequency domain. Nicholls (1976) has extended this work to the estimation of full VARMA models which also contain exogenous variables. Reinsel (1976) has considered the problem in the time domain and has derived estimates using the Newton-Raphson method on the (modified) maximum likelihood equations. An iterative nonlinear least squares estimation procedure has been proposed by Tunnicliffe Wilson (1973). Other papers in this area include Akaike (1973), Kashyap (1970), Whittle (1963), and Osborn (1977).

As indicated above, there is an alternative parameterization which we will use in this paper. One advantage of this parameterization is that we do not need to assume that some past values of $\underset{\sim}{\epsilon}_t$'s are fixed. Also as Hannan (1975) has pointed out it is easy to recover the original parameters using the spectral density.

Newton (1975) considers, among other things, the estimation of moving average covariance matrices in the pure moving average case in the frequency domain. But his method is different from the methods used in this paper. He regresses the elements of the sample spectral density, evaluated at some equidistant points, on certain trigonometric functions using the method of weighted least squares.

To summarize, Chapter 2 describes the model and the parameters to be estimated. Chapter 3 deals with the estimation problem in the time domain. The modified likelihood function is derived under the assumption of normality, using the method developed in Anderson (1975). Then the Newton-Raphson and scoring methods are applied to the resulting maximum likelihood equations. The chapter closes with remarks about the computational problems. The estimation in the frequency domain is discussed in Chapter 4. The modified likelihood function used is similar to that of Whittle (1953,1961) and Dunsmuir and Hannan (1976). Again the Newton-Raphson and scoring methods are applied to the maximum likelihood equations.

The asymptotic properties are discussed in Chapter 5. The estimates are shown to be asymptotically efficient under suitable assumptions.

In Chapter 6 we return to the usual parameterization and derive estimates for the autoregressive and moving average matrix coefficients and the covariance matrix of $\underset{\sim}{\epsilon}_t$, using the scoring method in the time domain. Finally in the Appendix we present some of the mathematical results used in previous chapters.

2. The Model.

We have observations, $\underset{\sim}{y}_1, \underset{\sim}{y}_2, \ldots, \underset{\sim}{y}_T$, on the process $\{\underset{\sim}{y}_t\}$ generated by

$$\sum_{i=0}^{p} \underset{\sim}{B}_i \underset{\sim}{y}_{t-i} = \sum_{j=0}^{q} \underset{\sim}{A}_j \underset{\sim}{\epsilon}_{t-j} \equiv \underset{\sim}{u}_t, \quad t = 0, \pm 1, \ldots, \tag{2.1}$$

where $\underset{\sim}{y}_t$'s and $\underset{\sim}{\epsilon}_t$'s are $m \times 1$ vectors and $\underset{\sim}{B}_i$'s and $\underset{\sim}{A}_j$'s are $m \times m$ matrices and $\underset{\sim}{B}_0 = \underset{\sim}{A}_0 = \underset{\sim}{I}_m$. Let

$$\underset{\sim}{B}(z) = \sum_{i=0}^{p} \underset{\sim}{B}_i z^i , \tag{2.2}$$

$$\underset{\sim}{A}(z) = \sum_{j=0}^{q} \underset{\sim}{A}_j z^j . \tag{2.3}$$

Assumption 1. The $\underset{\sim}{\epsilon}_t$'s are independently identically distributed random vectors with mean zero and unknown covariance matrix $\underset{\sim}{V}$.

Assumption 2. The zeros of $|\underset{\sim}{B}(z)|$ lie outside the unit circle.

Assumption 3. The zeros of $|\underset{\sim}{A}(z)|$ lie outside the unit circle.

Assumption 4. A greatest common left divisor of $\underset{\sim}{A}(z)$ and $\underset{\sim}{B}(z)$ is $\underset{\sim}{I}_m$.

Assumption 5. The matrix $(\underset{\sim}{B}_p, \underset{\sim}{A}_q)$ is of full rank, i.e., $r(\underset{\sim}{B}_p, \underset{\sim}{A}_q) = m$.

Remarks.

(i) Assumption 2 ensures the stationarity of the process. It also makes $\underset{\sim}{y}_t$ independent of $\underset{\sim}{\epsilon}_{t+1}, \underset{\sim}{\epsilon}_{t+2}, \ldots$.

(ii) Assumptions 3, 4, 5 ensure that the system is identified. By this we mean that the autocovariances of the process defined by

$$(2.4) \qquad \underset{\sim}{V}_s = \mathcal{E}\, \underset{\sim}{y}_t \underset{\sim}{y}'_{t+s} \;, \qquad s = 0, \pm 1, \ldots,$$

determine $\underset{\sim}{A}(z)$ and $\underset{\sim}{B}(\underset{\sim}{z})$ uniquely. (See Hannan (1969a)). In particular we can get the moving average matrix coefficients from the moving average covariance matrices defined by

$$(2.5) \qquad \underset{\sim}{\Sigma}^{(s)} = \mathcal{E}\, \underset{\sim}{u}_t \underset{\sim}{u}'_{t+s} \;, \quad s = 0, 1, \ldots, q \;,$$

uniquely, by solving the following system of equations

$$\underset{\sim}{\Sigma}^{(s)} = \sum_{i=0}^{q-s} \underset{\sim}{A}_i \underset{\sim}{V} \underset{\sim}{A}'_{s+i} \;, \quad s = 0, 1, \ldots, q \;.$$

Finally, assumption 5 is not a necessary condition and it can be replaced by other conditions, see Hannan (1971) and Kashyap and Nasburg (1974).

The parameters of interest are

$$\underset{\sim}{\Sigma}^{(s)} = \mathcal{E}\, \underset{\sim}{u}_t \underset{\sim}{u}'_{t+s} \;, \quad s = 0,1, \ldots, q \;,$$

$$\underset{\sim}{B}_r \;, \quad r = 1,\ldots,p \;.$$

Since we will differentiate the log likelihood function of $(\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_T)$ with respect to the elements of the above matrix parameters, it is more convenient to vectorize them.

__Definition.__ __If__ $\underset{\sim}{C} = (c_1, \ldots, \underset{\sim}{c}_n)$, __where__ $\underset{\sim}{c}_i$'s __are column vectors,__

$$\operatorname{vec} \underset{\sim}{C} = \begin{pmatrix} \underset{\sim}{c}_1 \\ \vdots \\ \underset{\sim}{c}_n \end{pmatrix} .$$

Before we proceed with vectorization of the matrix parameters, we notice that $\underset{\sim}{\Sigma}^{(0)}$ is symmetric and hence should be treated differently from $\underset{\sim}{\Sigma}^{(s)}$, $s = 1, \ldots, q$. In vectorizing $\underset{\sim}{\Sigma}^{(0)}$ we only need to vectorize the diagonal and lower diagonal elements. So we let

$$\underset{\sim}{\theta}_0^{(1)} = \operatorname{dg} \underset{\sim}{\Sigma}^{(0)} \equiv \begin{pmatrix} \sigma_{11}^{(0)} \\ \vdots \\ \sigma_{mm}^{(0)} \end{pmatrix},$$

$$\underset{\sim}{\theta}_0^{(2)} = \widetilde{\operatorname{vec}}\, \underset{\sim}{\Sigma}^{(0)} ,$$

where $\widetilde{\operatorname{vec}}$ is an operator that vectorizes the elements of the matrix that it is applied to, ignoring the diagonal and upper diagonal elements, e.g.,

$$\widetilde{\operatorname{vec}} \begin{pmatrix} 1 & 3 & 4 \\ 0 & 5 & 2 \\ 4 & 8 & 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 8 \end{pmatrix} .$$

So the parameters are

$$\underset{\sim}{\theta}_0 = \begin{pmatrix} \underset{\sim}{\theta}_0^{(1)} \\ \underset{\sim}{\theta}_0^{(2)} \end{pmatrix} , \quad \underset{\sim}{\theta}_s = \operatorname{vec} \underset{\sim}{\Sigma}^{(s)} , \quad s = 1, \ldots, q .$$

$$\underset{\sim}{\beta}_r = \operatorname{vec} \underset{\sim}{B}_r , \quad r = 1, \ldots, p .$$

Finally we can state that the parameters of interest are the components of $\underset{\sim}{\rho}$ where

$$(2.6) \qquad \underset{\sim}{\rho}' = (\underset{\sim}{\beta}', \underset{\sim}{\theta}') = (\underset{\sim}{\beta}_1', \ldots, \underset{\sim}{\beta}_p', \underset{\sim}{\theta}_0', \underset{\sim}{\theta}_1', \ldots, \underset{\sim}{\theta}_q') .$$

We shall also find it useful to introduce another vector, $\tilde{\underset{\sim}{\theta}}$ where

$$(2.7) \qquad \tilde{\underset{\sim}{\theta}}' = (\text{vec}' \, \underset{\sim}{\Sigma}^{(0)}, \text{vec}' \, \underset{\sim}{\Sigma}^{(1)}, \ldots, \text{vec}' \, \underset{\sim}{\Sigma}^{(q)})$$

$$= (\tilde{\underset{\sim}{\theta}}_0', \tilde{\underset{\sim}{\theta}}_1', \ldots, \tilde{\underset{\sim}{\theta}}_q') .$$

<u>Remarks.</u>

(i) We can find a matrix $\underset{m \times m^2}{\underset{\sim}{C}}$ such that for any $m \times m$ matrix $\underset{\sim}{A}$

$$(2.8) \qquad \text{dg}(\underset{\sim}{A}) = \underset{\sim}{C} \, \text{vec} \, \underset{\sim}{A} .$$

It is easy to see that $\underset{\sim}{C}$ is obtained from the $m^2 \times m^2$ identity matrix by deleting all the rows except the 1st, m+2nd, 2m+3rd, ..., m^2th, i.e.,

$$\underset{\sim}{C} = \begin{pmatrix} \underset{\sim}{e}_1' \\ \underset{\sim}{e}_{m+2}' \\ \underset{\sim}{e}_{2m+3}' \\ \vdots \\ e_{m^2}' \end{pmatrix} ,$$

where

$$\underset{\sim}{I}_{m^2} = (\underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots, \underset{\sim}{e}_{m^2}) .$$

(ii) Similarly we can find an $\frac{m(m-1)}{2} \times m^2$ matrix $\underset{\sim}{D}$ such that for any $m \times m$ matrix $\underset{\sim}{A}$

(2.9) $$\widetilde{\text{vec}}\, \underset{\sim}{A} = \underset{\sim}{D}\, \text{vec}\, \underset{\sim}{A} .$$

It is easily verified that $\underset{\sim}{D}$ is obtained from $\underset{\sim}{I}_{m^2}$ by deleting the following rows

$$\begin{array}{rrrr} 1,\ m+1, & 2m+1, & \ldots, & (m-1)m+1 \\ m+2, & 2m+2, & \ldots, & (m-1)m+2 \\ & 2m+3, & \ldots, & (m-1)m+3 \\ & & & \vdots \\ & & & (m-1)m+m \end{array}$$

3. Estimation in the Time Domain.

3.1. The Likelihood Function.

We are concerned with maximum likelihood estimation, so the first step is to find the likelihood function. To find this we need to know the distribution of $\underset{\sim}{\epsilon}_t$. But in the previous chapter we assumed that $\underset{\sim}{\epsilon}_t$'s are i.i.d., without assuming any particular distribution. However in deriving the likelihood function we shall treat $\underset{\sim}{\epsilon}_t$'s as normal vectors and later we shall demonstrate that the resulting estimates have the same asymptotic covariance matrix irrespective of the distribution of $\underset{\sim}{\epsilon}_t$'s as long as the assumptions in Chapter 2 are satisfied.

Even with the assumption of normality we cannot find the exact likelihood function except in the pure moving average case. In order to find an approximate likelihood function, following Anderson (1975), we assume

$$\underset{\sim}{y}_0 = \underset{\sim}{y}_{-1} = \cdots = \underset{\sim}{y}_{1-p} = \underset{\sim}{0} .$$

The likelihood function we will derive is in fact the conditional likelihood of $\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_T$ given that $\underset{\sim}{y}_0, \underset{\sim}{y}_{-1}, \ldots, \underset{\sim}{y}_{1-p}$ are equal to their expected values. Now

$$\underset{\sim}{y}_t + \underset{\sim}{B}_1 \underset{\sim}{y}_{t-1} + \cdots + \underset{\sim}{B}_p \underset{\sim}{y}_{t-p} = \underset{\sim}{u}_t .$$

Transposing both sides yields

$$\underset{\sim}{y}'_t + \underset{\sim}{y}'_{t-1} \underset{\sim}{B}'_1 + \cdots + \underset{\sim}{y}'_{t-p} \underset{\sim}{B}'_p = \underset{\sim}{u}'_t .$$

Writing these equations for $t = 1, 2, \ldots, T$, we get

$$(3.1.1) \qquad \begin{pmatrix} \underset{\sim}{y}_1' \\ \vdots \\ \underset{\sim}{y}_T' \end{pmatrix} + \begin{pmatrix} \underset{\sim}{0} \\ \underset{\sim}{y}_1' \\ \vdots \\ \underset{\sim}{y}_{T-1}' \end{pmatrix} \underset{\sim}{B}_1' + \cdots + \begin{pmatrix} \underset{\sim}{0} \\ \vdots \\ \underset{\sim}{y}_1' \\ \vdots \\ \underset{\sim}{y}_{T-p}' \end{pmatrix} \underset{\sim}{B}_p' = \begin{pmatrix} \underset{\sim}{u}_1' \\ \vdots \\ \underset{\sim}{u}_T' \end{pmatrix}$$

As in Anderson (1975) we notice that

$$\begin{pmatrix} \underset{\sim}{0} \\ \underset{\sim}{y}_1' \\ \vdots \\ \underset{\sim}{y}_{T-1}' \end{pmatrix} = \underset{\sim}{L} \begin{pmatrix} \underset{\sim}{y}_1' \\ \vdots \\ \underset{\sim}{y}_T' \end{pmatrix} ,$$

where

$$\underset{\sim}{L} = \begin{pmatrix} \underset{\sim}{0} & 0 \\ \underset{\sim}{I}_{T-1} & \underset{\sim}{0} \end{pmatrix} .$$

Similarly

$$\begin{pmatrix} \underset{\sim}{0} \\ \underset{\sim}{0} \\ \underset{\sim}{y}_1 \\ \vdots \\ \underset{\sim}{y}_{T-2} \end{pmatrix} = L^2 \begin{pmatrix} \underset{\sim}{y}_1' \\ \vdots \\ y_T' \end{pmatrix}$$

and so on. This means that we can write (3.1.1) as

(3.1.2) $$\sum_{i=0}^{p} \underset{\sim}{L}^{i}\, \underset{\sim}{Y}\, \underset{\sim}{B}_i' = \underset{\sim}{U} ,$$

where

$$\underset{\sim}{Y}' = (\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_T) , \quad \underset{\sim}{U}' = (\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_T) .$$

Now we shall need the following lemma in vectorizing (3.1.2).

<u>Lemma 3.1</u>.

$$\text{vec } \underset{\sim}{A}\,\underset{\sim}{B}\,\underset{\sim}{C} = (\underset{\sim}{C}' \otimes \underset{\sim}{A})\text{vec } \underset{\sim}{B} .$$

See Minc and Marcus (1964).

Using this lemma on the left hand side of (3.1.2) we get

(3.1.3) $$\sum_{i=0}^{p} (\underset{\sim}{B}_i \otimes \underset{\sim}{L}^{i})\text{vec } \underset{\sim}{Y} = \text{vec } \underset{\sim}{U} .$$

Now let

$$\underset{\sim}{\mathcal{B}} = \sum_{i=0}^{p} (\underset{\sim}{B}_i \otimes \underset{\sim}{L}^{i}) , \quad \underset{\sim}{y} = \text{vec } \underset{\sim}{Y} , \quad \underset{\sim}{u} = \text{vec } \underset{\sim}{U} ,$$

then (3.1.3) can be written as

(3.1.4) $$\underset{\sim}{\mathcal{B}}\, \underset{\sim}{y} = \underset{\sim}{u} .$$

Because of the Gaussian assumption, to find the likelihood function (the density of $\underset{\sim}{y}$) we only need to find the covariance matrix of $\underset{\sim}{u}$. Now, it follows from the normality of $\underset{\sim}{\epsilon}_t$'s that

$$(3.1.5) \qquad \begin{pmatrix} \underset{\sim}{u}_1 \\ \vdots \\ \underset{\sim}{u}_T \end{pmatrix} \sim N(\underset{\sim}{0}, \underset{\sim}{\Sigma}) ,$$

where

$$(3.1.6)\ \underset{\sim}{\Sigma} = \begin{pmatrix} \underset{\sim}{\Sigma}^{(0)} & \underset{\sim}{\Sigma}^{(1)} & \cdots & \underset{\sim}{\Sigma}^{(q)} & 0 & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \underset{\sim}{\Sigma}'^{(1)} & \underset{\sim}{\Sigma}^{(0)} & \underset{\sim}{\Sigma}^{(1)} & \cdots & \underset{\sim}{\Sigma}^{(q)} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \vdots & \ddots & \ddots & \ddots & & \ddots & & \vdots \\ & & & & & & & \underset{\sim}{\Sigma}^{(q)} \\ \underset{\sim}{\Sigma}'^{(q)} & & \ddots & & \ddots & & \ddots & \\ \underset{\sim}{0} & \ddots & & \ddots & & \ddots & & \\ \vdots & & \ddots & & \ddots & & \ddots & \underset{\sim}{\Sigma}^{(1)} \\ \underset{\sim}{0} & \cdots & & \underset{\sim}{\Sigma}'^{(q)} & \cdots & \underset{\sim}{\Sigma}'^{(1)} & & \underset{\sim}{\Sigma}^{(0)} \end{pmatrix}$$

$$= \underset{\sim}{I}_T \otimes \underset{\sim}{\Sigma}^{(0)} + (\underset{\sim}{L}' \otimes \underset{\sim}{\Sigma}^{(1)} + \underset{\sim}{L} \otimes \underset{\sim}{\Sigma}^{(1)'}) + \cdots + (\underset{\sim}{L}'^{q} \otimes \underset{\sim}{\Sigma}^{(q)} + \underset{\sim}{L}^{q} \otimes \underset{\sim}{\Sigma}^{(q)'}) .$$

The following lemma will enable us to derive the distribution of $\underset{\sim}{u}$ from (3.1.5).

<u>Lemma 3.2. If $\underset{\sim}{A}$ is any $r \times s$ matrix,</u>

$$\text{vec } \underset{\sim}{A}' = \underset{\sim}{K}_{r,s} \text{ vec } \underset{\sim}{A} ,$$

<u>where $\underset{\sim}{K}_{r,s}$ is a square $rs \times rs$ matrix partitioned into $r \times s$ submatrices such that the ij^{th} block has a 1 in the ji^{th} position and zeros elsewhere. (See MacRae (1974).)</u>

<u>Proof</u>. $\underset{\sim}{K}_{r,s}$ can be written as

$$\underset{\sim}{K}_{r,s} = \begin{pmatrix} \underset{\sim}{E}_{11} & \underset{\sim}{E}_{21} & \cdots & \underset{\sim}{E}_{s1} \\ \vdots & & & \\ \underset{\sim}{E}_{1r} & \underset{\sim}{E}_{2r} & \cdots & \underset{\sim}{E}_{sr} \end{pmatrix} ,$$

where $\underset{\sim}{E}_{ij}$ is an $s \times r$ matrix with 1 in the ij^{th} position and zeros elsewhere. Now

$$\underset{\sim}{K}_{r,s} \text{ vec } \underset{\sim}{A} = \underset{\sim}{K}_{r,s} \text{ vec}(\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_s) = \underset{\sim}{K}_{r,s} \begin{pmatrix} \underset{\sim}{a}_1 \\ \vdots \\ \underset{\sim}{a}_s \end{pmatrix} = \begin{pmatrix} \underset{\sim}{b}_1 \\ \vdots \\ \underset{\sim}{b}_r \end{pmatrix} , \text{ say .}$$

So

$$\underset{\sim}{b}_i = \sum_{\ell=1}^{s} \underset{\sim}{E}_{\ell i} \underset{\sim}{a}_\ell .$$

But $\underset{\sim}{E}_{\ell i} \underset{\sim}{a}_\ell$ is an $r \times 1$ vector with $a_{i\ell}$ in the i^{th} position and zeros elsewhere, hence

$$\underset{\sim}{b}_i = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{ir} \end{pmatrix} .$$

This means that

$$\begin{pmatrix} \underset{\sim}{b}_1 \\ \vdots \\ \underset{\sim}{b}_r \end{pmatrix} = \text{vec } \underset{\sim}{A}'$$

Q.E.D.

<u>Note</u>. We shall use the convention

$$\underset{\sim}{K}_n \equiv \underset{\sim}{K}_{n,n} \ .$$

Now we use the above lemma to derive the distribution of $\underset{\sim}{u}$. By definition

$$\underset{\sim}{u} = \text{vec } \underset{\sim}{U} = \text{vec}(\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_T)' \ .$$

Now, using lemma 3.2 we have

$$\begin{aligned} \text{vec}(\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_T)' &= \underset{\sim}{K}_{m,T} \text{ vec}(\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_T) \\ &= \underset{\sim}{K}_{m,T} \begin{pmatrix} \underset{\sim}{u}_1 \\ \vdots \\ \underset{\sim}{u}_T \end{pmatrix} \ . \end{aligned}$$

This, together with (3.1.5), yields

$$\underset{\sim}{u} \sim N(\underset{\sim}{0}, \underset{\sim}{\Gamma}) \ ,$$

where

$$\underset{\sim}{\Gamma} = \underset{\sim}{K}_{m,T} \underset{\sim}{\Sigma} \underset{\sim}{K}'_{m,T} .$$

Lemma 3.3.

(i) If $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are $m \times n$ and $r \times s$ matrices respectively, then

$$\underset{\sim}{K}_{m,r}(\underset{\sim}{A} \otimes \underset{\sim}{B})\underset{\sim}{K}_{s,n} = \underset{\sim}{B} \otimes \underset{\sim}{A} .$$

(ii) $\underset{\sim}{K}_{r,s} = \underset{\sim}{K}'_{s,r}$.

See MacRae (1974).

Using this lemma and noticing that $\underset{\sim}{\Sigma}$ is given by (3.1.6) we get

$$(3.1.7) \quad \underset{\sim}{\Gamma} = \underset{\sim}{\Sigma}^{(0)} \otimes \underset{\sim}{I}_T + (\underset{\sim}{\Sigma}^{(1)} \otimes \underset{\sim}{L}' + \underset{\sim}{\Sigma}^{(1)'} \otimes \underset{\sim}{L}) + \cdots + (\underset{\sim}{\Sigma}^{(q)} \otimes \underset{\sim}{L}'^{q} + \underset{\sim}{\Sigma}^{(q)'} \otimes \underset{\sim}{L}^{(q)}) .$$

We shall be using these lemmas in the later chapters as well.

Finally from (3.1.3) we get

$$\underset{\sim}{y} \sim N(\underset{\sim}{0}, \underset{\sim}{\mathfrak{B}}^{-1} \underset{\sim}{\Gamma} \underset{\sim}{\mathfrak{B}}'^{-1}) .$$

This gives us the (modified) log likelihood function

$$(3.1.8) \qquad \log \ell(\underset{\sim}{y}) = -\frac{1}{2} \underset{\sim}{y}' \underset{\sim}{\mathfrak{B}}' \underset{\sim}{\Gamma}^{-1} \underset{\sim}{\mathfrak{B}} \underset{\sim}{y} + \frac{1}{2} \log|\underset{\sim}{\Gamma}^{-1}| - \frac{Tm}{2} \log 2\pi ,$$

since $|\underset{\sim}{\mathfrak{B}}| = 1$, as we shall see in (3.5.2). The maximum likelihood estimates are a set of roots of

$$\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\rho}} = 0 \; .$$

So we proceed to derive the first derivative of $\log \ell(\underset{\sim}{y})$.

3.2. The First Derivative of $\log \ell(\underset{\sim}{y})$.

In differentiating the log likelihood function with respect to $\underset{\sim}{\theta}$ we only need to differentiate the first term and by using

$$\mathcal{E}\left(\frac{\partial \log \ell}{\partial \underset{\sim}{\theta}}\right) = 0$$

complete the derivative. That is

$$\text{(3.2.1)} \qquad \frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} = -\frac{1}{2}\left(\frac{\partial(\underset{\sim}{y}'\underset{\sim}{\mathcal{B}}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{\mathcal{B}}\,\underset{\sim}{y})}{\partial \underset{\sim}{\theta}} - \mathcal{E}\,\frac{\partial(\underset{\sim}{y}'\underset{\sim}{\mathcal{B}}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{\mathcal{B}}\,\underset{\sim}{y})}{\partial \underset{\sim}{\theta}} \right) .$$

We shall also use the following lemma

Lemma 3.4. For any two column vectors $\underset{\sim}{x}$ and $\underset{\sim}{z}$

$$\operatorname{vec}(\underset{\sim}{x}\,\underset{\sim}{z}') = \underset{\sim}{z} \otimes \underset{\sim}{x} \; .$$

Proof. It is easily verified by writing out the two sides.

The Derivative With Respect to Autoregressive Coefficients.

Let $b_{ij}^{(r)}$ denote the i,j^{th} element of $\underset{\sim}{B}_r$ then from (3.1.3)

$$\text{(3.2.2)} \qquad \frac{\partial \underset{\sim}{\mathcal{B}}}{\partial b_{ij}^{(r)}} = \underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^{r} \; , \quad i,j = 1, \ldots, m \; .$$

Using this we get

$$(3.2.3)\qquad \frac{\partial(\underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{B}\,\underset{\sim}{y})}{\partial b_{ij}^{(r)}} = \underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1}(\underset{\sim}{E}_{ij}\otimes\underset{\sim}{L}^{r})\underset{\sim}{y} + \underset{\sim}{y}'(\underset{\sim}{E}_{ji}\otimes\underset{\sim}{L}'^{r})\underset{\sim}{\Gamma}^{-1}\underset{\sim}{B}\,\underset{\sim}{y}$$

$$= 2\underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1}(\underset{\sim}{E}_{ij}\otimes\underset{\sim}{L}^{r})\underset{\sim}{y}\ .$$

Now using lemma 3.1 we have

$$\underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1}(\underset{\sim}{E}_{ij}\otimes\underset{\sim}{L}^{r})\underset{\sim}{y} = \mathrm{vec}[\underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1}(\underset{\sim}{E}_{ij}\otimes\underset{\sim}{L}^{r})\underset{\sim}{y}]$$

$$= (\underset{\sim}{y}'\otimes\underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1})\,\mathrm{vec}(\underset{\sim}{E}_{ij}\otimes\underset{\sim}{L}^{r})\ .$$

This together with (3.2.3), yields

$$(3.2.4)\qquad \frac{\partial(\underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{B}\,\underset{\sim}{y})}{\partial\,\underset{\sim}{\beta}_r'} = 2(\underset{\sim}{y}'\otimes\underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1})\underset{\sim}{E}_r\ ,$$

where

$$\underset{\sim}{E}_r = [\mathrm{vec}(\underset{\sim}{E}_{11}\otimes\underset{\sim}{L}^{r}),\ \mathrm{vec}(\underset{\sim}{E}_{21}\otimes\underset{\sim}{L}^{r}),\ \ldots,\ \mathrm{vec}(\underset{\sim}{E}_{mm}\otimes\underset{\sim}{L}^{r})],\ r = 1,\ \ldots,\ p.$$

Which in turn yields

$$(3.2.5)\qquad \frac{\partial(\underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1}\underset{\sim}{B}\,\underset{\sim}{y})}{\partial\,\underset{\sim}{\beta}'} = 2(\underset{\sim}{y}'\otimes\underset{\sim}{y}'\underset{\sim}{B}'\underset{\sim}{\Gamma}^{-1})\underset{\sim}{E}\ ,$$

where

$$(3.2.6)\qquad \underset{\sim}{E} = (\underset{\sim}{E}_1,\ \ldots,\ \underset{\sim}{E}_p)\ .$$

Finally using (3.1.8) we get

$$(3.2.7) \qquad \frac{\partial \log \ell}{\partial \underset{\sim}{\beta}} = -\underset{\sim}{E}'(\underset{\sim}{y} \otimes \underset{\sim}{\Gamma}^{-1} \underset{\sim}{\mathcal{B}} \underset{\sim}{Y}).$$

The Derivative with Respect to Moving Average Covariances.

As indicated by (3.2.1) we only need to find

$$\frac{\partial\, \underset{\sim}{Y}' \underset{\sim}{\mathcal{B}}' \underset{\sim}{\Gamma}^{-1} \underset{\sim}{\mathcal{B}}\, \underset{\sim}{Y}}{\partial\, \underset{\sim}{\theta}}.$$

However it is more convenient to find

$$\frac{\partial\, \underset{\sim}{Y}' \underset{\sim}{\mathcal{B}}' \underset{\sim}{\Gamma}^{-1} \underset{\sim}{\mathcal{B}}\, \underset{\sim}{Y}}{\partial\, \tilde{\underset{\sim}{\theta}}}$$

which is related to the former derivative. To find the relationship we note that

$$\frac{\partial \underset{\sim}{\Gamma}}{\partial \sigma_{ij}^{(0)}} = \left(\frac{\partial \underset{\sim}{\Gamma}}{\partial \sigma_{ji}^{(0)}}\right)',$$

where $\sigma_{ij}^{(0)}$ and $\sigma_{ji}^{(0)}$ are treated as different variables. This means

$$\frac{\partial(\underset{\sim}{Y}' \underset{\sim}{\mathcal{B}}' \underset{\sim}{\Gamma}^{-1} \underset{\sim}{\mathcal{B}}\, \underset{\sim}{Y})}{\partial \sigma_{ij}^{(0)}} = \frac{\partial(\underset{\sim}{Y}' \underset{\sim}{\mathcal{B}}' \underset{\sim}{\Gamma}^{-1} \underset{\sim}{\mathcal{B}}\, \underset{\sim}{Y})}{\partial \sigma_{ji}^{(0)}},$$

which in turn yields

$$(3.2.8) \qquad \frac{\partial\, \underset{\sim}{Y}' \underset{\sim}{\mathcal{B}}' \underset{\sim}{\Gamma}^{-1} \underset{\sim}{\mathcal{B}}\, \underset{\sim}{Y}}{\partial \bar{\sigma}_{ij}^{(0)}} = 2\, \frac{\partial\, \underset{\sim}{Y}' \underset{\sim}{\mathcal{B}}' \underset{\sim}{\Gamma}^{-1} \underset{\sim}{\mathcal{B}}\, \underset{\sim}{Y}}{\partial \sigma_{ij}^{(0)}},$$

where " — " indicates that we take the symmetry of $\underset{\sim}{\Sigma}^{(0)}$ into account. In view of (2.8) and (2.9) using (3.2.8) we get

(3.2.9) $$\frac{\partial \underset{\sim}{y}'\mathfrak{B}'\underset{\sim}{\Gamma}^{-1}\mathfrak{B}\,\underset{\sim}{y}}{\partial \underset{\sim}{\theta}} = \underset{\sim}{G}\,\frac{\partial \underset{\sim}{y}'\mathfrak{B}'\underset{\sim}{\Gamma}^{-1}\mathfrak{B}\,\underset{\sim}{y}}{\partial \underset{\sim}{\tilde{\theta}}},$$

where $\underset{\sim}{G}$ is a $[qm^2 + \frac{m(m+1)}{2}] \times (q+1)m^2$ matrix which can be written as

$$G = \begin{pmatrix} \underset{\sim}{G}_1 \\ \underset{\sim}{G}_2 \\ \underset{\sim}{G}_3 \end{pmatrix} \begin{matrix} m \\ \frac{m(m-1)}{2} \\ qm^2 \end{matrix},$$

and

$$\underset{\sim}{G}_1 = (\underset{\sim}{C}, \underset{\sim}{0}), \quad \underset{\sim}{G}_2 = (2\underset{\sim}{D}, \underset{\sim}{0}), \quad \underset{\sim}{G}_3 = (\underset{\sim}{0}, \underset{\sim}{I})$$

with $\underset{\sim}{C}$ and $\underset{\sim}{D}$ as in Chapter 2. It is obvious that from (3.2.9) we can conclude

(3.2.10) $$\frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} = \underset{\sim}{G}\,\frac{\partial \log \ell}{\partial \underset{\sim}{\tilde{\theta}}}.$$

We now proceed to derive $\frac{\partial \log \ell}{\partial \underset{\sim}{\tilde{\theta}}}$. Using lemma 3.1 we have

(3.2.11) $$\frac{\partial \underset{\sim}{y}'\mathfrak{B}'\underset{\sim}{\Gamma}^{-1}\mathfrak{B}\,\underset{\sim}{y}}{\partial \underset{\sim}{\tilde{\theta}}'} = \frac{\partial\, \mathrm{vec}(\underset{\sim}{y}'\mathfrak{B}'\underset{\sim}{\Gamma}^{-1}\mathfrak{B}\,\underset{\sim}{y})}{\partial \underset{\sim}{\tilde{\theta}}'}$$

$$= (\underset{\sim}{y}'\mathfrak{B}' \otimes \underset{\sim}{y}'\mathfrak{B}')\,\frac{\partial\, \mathrm{vec}\, \underset{\sim}{\Gamma}^{-1}}{\partial \underset{\sim}{\tilde{\theta}}'}$$

$$= -(\underset{\sim}{y}'\mathfrak{B}' \otimes \underset{\sim}{y}'\mathfrak{B}')\,\mathrm{vec}(\Gamma^{-1}\,\frac{\partial \underset{\sim}{\Gamma}}{\partial \underset{\sim}{\tilde{\theta}}'}\,\underset{\sim}{\Gamma}^{-1})$$

$$= -(\underset{\sim}{y}'\mathfrak{B}' \otimes \underset{\sim}{y}'\mathfrak{B}')(\underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{\Gamma}^{-1})\,\frac{\partial\, \mathrm{vec}\, \underset{\sim}{\Gamma}}{\partial \underset{\sim}{\tilde{\theta}}'}$$

$$= -(\underset{\sim}{y}'\mathfrak{B}'\underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{y}'\mathfrak{B}'\underset{\sim}{\Gamma}^{-1})\,\frac{\partial\, \mathrm{vec}\, \underset{\sim}{\Gamma}}{\partial \underset{\sim}{\tilde{\theta}}'}.$$

To find $\dfrac{\partial \text{ vec } \underset{\sim}{\Gamma}}{\partial \underset{\sim}{\tilde{\theta}}'}$ we need to find $\dfrac{\partial \underset{\sim}{\Gamma}}{\partial \sigma_{ij}^{(s)}}$, for $s = 0,1,\ldots,q$ and $i,j = 1,\ldots,m$. From (3.1.6) we get

$$(3.2.12) \qquad \frac{\partial \underset{\sim}{\Gamma}}{\partial \sigma_{ij}^{(0)}} = \underset{\sim}{E}_{ij} \otimes \underset{\sim}{I}_T ,$$

$$(3.2.13) \qquad \frac{\partial \underset{\sim}{\Gamma}}{\partial \sigma_{ij}^{(s)}} = \underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}'^{s} + \underset{\sim}{E}_{ji} \otimes \underset{\sim}{L}^{s} , \quad s = 1, \ldots, q,\ i,j = 1, \ldots, m,$$

where $\underset{\sim}{E}_{ij}$ is an $m \times m$ matrix with one in the ij^{th} position and zeros elsewhere. Now, vectorizing (3.2.12) yields

$$\frac{\partial \text{ vec } \underset{\sim}{\Gamma}}{\partial \sigma_{ij}^{(0)}} = \text{vec}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{I}_T) = \underset{\sim}{\alpha}_{ij}^{(0)}, \quad \text{say} ,$$

which yields

$$(3.2.14) \qquad \frac{\partial \text{ vec } \underset{\sim}{\Gamma}}{\partial \underset{\sim}{\tilde{\theta}}_0'} = (\underset{\sim}{\alpha}_{11}^{(0)}, \underset{\sim}{\alpha}_{21}^{(0)}, \ldots, \underset{\sim}{\alpha}_{mm}^{(0)}) = \underset{\sim}{F}_0, \quad \text{say} .$$

Similarly

$$(3.2.15) \qquad \frac{\partial \text{ vec } \underset{\sim}{\Gamma}}{\partial \underset{\sim}{\tilde{\theta}}_s'} = \underset{\sim}{F}_s = (\underset{\sim}{\alpha}_{11}^{(s)}, \ldots, \underset{\sim}{\alpha}_{mm}^{(s)}) ,$$

where

$$\underset{\sim}{\alpha}_{ij}^{(s)} = \text{vec}[\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}'^{s} + \underset{\sim}{E}_{ji} \otimes \underset{\sim}{L}^{s}] , \quad s = 1,\ldots,q .$$

So the derivative with respect to $\underset{\sim}{\tilde{\theta}}'$ is

$$(3.2.16) \qquad \frac{\partial \operatorname{vec} \boldsymbol{\Gamma}}{\partial \tilde{\boldsymbol{\theta}}'} = (\mathbf{F}_0, \mathbf{F}_1, \ldots, \mathbf{F}_q) = \mathbf{F}, \quad \text{say} .$$

Now using (3.2.11) we have

$$(3.2.17) \qquad \frac{\partial\, \mathbf{y}'\mathscr{B}'\boldsymbol{\Gamma}^{-1}\mathscr{B}\,\mathbf{y}}{\partial \tilde{\boldsymbol{\theta}}} = -\mathbf{F}'(\boldsymbol{\Gamma}^{-1}\mathscr{B}\,\mathbf{y} \otimes \boldsymbol{\Gamma}^{-1}\mathscr{B}\,\mathbf{y}) .$$

To complete the derivative of $\dfrac{\partial \log \ell}{\partial \tilde{\boldsymbol{\theta}}}$ we need to find the expected value of (3.2.16)

$$\mathcal{E}\left(\frac{\partial\, \mathbf{y}'\mathscr{B}'\boldsymbol{\Gamma}^{-1}\mathscr{B}\,\mathbf{y}}{\partial \tilde{\boldsymbol{\theta}}}\right) = -\mathbf{F}'\, \mathcal{E} \operatorname{vec}(\boldsymbol{\Gamma}^{-1}\mathscr{B}\,\mathbf{y}\mathbf{y}'\mathscr{B}'\boldsymbol{\Gamma}^{-1})$$

$$= -\mathbf{F}' \operatorname{vec}(\boldsymbol{\Gamma}^{-1}\mathscr{B}\mathscr{B}^{-1}\boldsymbol{\Gamma}\,\mathscr{B}'^{-1}\mathscr{B}'\boldsymbol{\Gamma}^{-1})$$

$$= -\mathbf{F}' \operatorname{vec} \boldsymbol{\Gamma}^{-1} .$$

Using (3.2.1) we get

$$(3.2.18) \qquad \frac{\partial \log \ell}{\partial \tilde{\boldsymbol{\theta}}} = \frac{1}{2}\mathbf{F}'[\boldsymbol{\Gamma}^{-1}\mathscr{B}\,\mathbf{y} \otimes \boldsymbol{\Gamma}^{-1}\mathscr{B}\,\mathbf{y} - \operatorname{vec} \boldsymbol{\Gamma}^{-1}] .$$

Finally using (3.2.10), we have

$$(3.2.19) \qquad \frac{\partial \log \ell}{\partial \boldsymbol{\theta}} = \frac{1}{2}\mathbf{G}\,\mathbf{F}'[\boldsymbol{\Gamma}^{-1}\mathscr{B}\,\mathbf{y} \otimes \boldsymbol{\Gamma}^{-1}\mathscr{B}\,\mathbf{y} - \operatorname{vec} \boldsymbol{\Gamma}^{-1}].$$

3.3. The Numerical Approximations.

The equation

$$\frac{\partial \log \ell(\mathbf{y})}{\partial \boldsymbol{\rho}} = 0$$

is nonlinear and cannot be solved explicitly. Therefore we will use numerical approximations that yield asymptotically efficient estimates.

These methods are the Newton-Raphson and Scoring methods. Both methods require that we start with an initial estimate that is consistent of order $T^{-1/2}$, call it $\hat{\underset{\sim}{\rho}}_{(0)}$. Then, the Newton-Raphson method consists of solving the following set of linear equations for $\hat{\underset{\sim}{\rho}}_{(1)}$.

$$(3.3.1) \qquad -\left.\frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\rho}\, \partial \underset{\sim}{\rho}'}\right|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}} (\underset{\sim}{\rho}_{(1)} - \hat{\underset{\sim}{\rho}}_{(0)}) = \left.\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\rho}}\right|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}} .$$

In the Scoring method,

$$\mathcal{E}\left(\frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\rho}\partial \underset{\sim}{\rho}'}\right)\Bigg|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}}$$

replaces

$$\left.\frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\rho}\, \partial \underset{\sim}{\rho}'}\right|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}} ,$$

i.e., we solve the following set of linear equations for $\hat{\underset{\sim}{\rho}}_{(1)}$

$$(3.3.2) \qquad -\mathcal{E}\left(\frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\rho}\, \partial \underset{\sim}{\rho}'}\right)\Bigg|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}} (\hat{\underset{\sim}{\rho}}_{(1)} - \hat{\underset{\sim}{\rho}}_{(0)}) = \left.\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\rho}}\right|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}} .$$

<u>Initial estimate</u> $\hat{\underset{\sim}{\rho}}_{(0)}$.

In the vector Yule-Walker equations

$$\sum_{r=1}^{p} \underset{\sim}{B}_r \underset{\sim}{V}_{r-s} = -\underset{\sim}{V}_{-s} , \quad s = q+1, \ldots, q+p ,$$

we can replace the autocovariances i.e., $\underset{\sim}{V}_s$'s by their sample analogues, $\underset{\sim}{C}_k$'s, and then solve the resulting equations to obtain initial estimates for $\underset{\sim}{B}_r$'s. So the equations are

$$(3.3.3) \qquad \sum_{r=1}^{p} \hat{\underset{\sim}{B}}_r(0)\underset{\sim}{C}_{r-s} = -\underset{\sim}{C}_{-s}\,, \quad s = q+1, \ldots, q+p\,,$$

where

$$(3.3.4) \qquad \underset{\sim}{C}_k = \frac{1}{T}\sum_{t=1}^{T-k} \underset{\sim}{y}_t\underset{\sim}{y}'_{t+k} = \underset{\sim}{C}'_{-k}\,.$$

Having obtained these estimates we can form

$$\hat{\underset{\sim}{u}}_t(0) = \sum_{r=0}^{p} \hat{\underset{\sim}{B}}_r(0)\underset{\sim}{y}_{t-r}\,, \quad t = 1, \ldots, T\,.$$

Bear in mind that $\hat{\underset{\sim}{B}}_0(0) = \underset{\sim}{I}$ and $\underset{\sim}{y}_0 = \cdots = \underset{\sim}{y}_{1-p} = \underset{\sim}{0}$. Now we can use the same autocovariances of $\hat{\underset{\sim}{u}}_t(0)$'s to estimate $\underset{\sim}{\Sigma}^{(s)}$, $s = 0,1, \ldots, q$. We estimate $\underset{\sim}{\Sigma}^{(s)}$ by

$$(3.3.5) \qquad \hat{\underset{\sim}{\Sigma}}^{(s)}_{(0)} = \frac{1}{T}\sum_{t=1}^{T-s} \hat{\underset{\sim}{u}}_t(0)\hat{\underset{\sim}{u}}'_{t+s}(0) = \hat{\underset{\sim}{\Sigma}}^{(-s)'}_{(0)}\,, \quad s = 0,1, \ldots, q\,.$$

Finally by vectorizing the initial estimates obtained in this manner we get an initial estimate of $\underset{\sim}{\rho}$, which we shall denote by $\hat{\underset{\sim}{\rho}}_{(0)}$.

<u>Note</u>. The initial estimate $\hat{\underset{\sim}{\Sigma}}_{(0)}$ may not satisfy the conditions for a moving average covariance matrix. For example when $m = q = 1$, $p = 0$ the constraint on $\underset{\sim}{\Sigma}$ is

$$\frac{\sigma_1^2}{\sigma_0^2} < \frac{1}{4}\,.$$

The Newton-Raphson Method.

This method consists of solving the following system of linear equations for $\hat{\underset{\sim}{\rho}}_{(1)}$

$$-\left.\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\rho}\, \partial \underset{\sim}{\rho}'}\right|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}} (\hat{\underset{\sim}{\rho}}_{(1)} - \hat{\underset{\sim}{\rho}}_{(0)}) = \left.\frac{\partial \log \ell}{\partial \underset{\sim}{\rho}}\right|_{\underset{\sim}{\rho}=\underset{\sim}{\rho}_{(0)}} .$$

So we need to find the second partial derivatives of $\log \ell$.

Derivation of $\dfrac{\partial^2 \log \ell}{\partial \underset{\sim}{\beta}\, \partial \underset{\sim}{\beta}'}$.

As in (3.2.7)

$$\frac{\partial \log \ell}{\partial \underset{\sim}{\beta}} = -\underset{\sim}{E}'(\underset{\sim}{y} \otimes \underset{\sim}{\Gamma}^{-1} \underset{\sim}{B}\, \underset{\sim}{y}) .$$

Differentiating this with respect to $b_{ij}^{(r)}$ yields

$$\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\beta}\, \partial b_{ij}^{(r)}} = -\underset{\sim}{E}'[\underset{\sim}{y} \otimes \underset{\sim}{\Gamma}^{-1} (\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^r)\underset{\sim}{y}] .$$

Now using lemma 3.4 we can rewrite the term inside the brackets as

$$\mathrm{vec}[\underset{\sim}{\Gamma}^{-1} (\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^r)\underset{\sim}{y}\underset{\sim}{y}'] = (\underset{\sim}{y}\underset{\sim}{y}' \otimes \underset{\sim}{\Gamma}^{-1})\mathrm{vec}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^r) ,$$

where we have used lemma 3.1. So finally we have

$$\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\beta}\, \partial \underset{\sim}{\beta}'} = -E'[\underset{\sim}{y}\underset{\sim}{y}' \otimes \underset{\sim}{\Gamma}^{-1}]\underset{\sim}{E} . \qquad (3.3.6)$$

<u>Derivation of</u> $\dfrac{\partial^2 \log \ell}{\partial \underset{\sim}{\beta}\, \partial \underset{\sim}{\theta}'}$.

Differentiating (3.2.7) with respect to $\sigma_{ij}^{(s)}$ we get

$$\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\beta}\, \partial \sigma_{ij}^{(s)}} = \underset{\sim}{E}'[\underset{\sim}{y} \otimes \underset{\sim}{\Gamma}^{-1}(\frac{\partial \underset{\sim}{\Gamma}}{\partial \sigma_{ij}^{(s)}})\, \underset{\sim}{\Gamma}^{-1} \underset{\sim}{B}\, \underset{\sim}{y}] .$$

Now using lemma 3.4 the right hand side can be rewritten as

$$\underset{\sim}{E}' \ \mathrm{vec}[\underset{\sim}{\Gamma}^{-1} \frac{\partial \underset{\sim}{\Gamma}}{\partial \sigma_{ij}^{(s)}} \underset{\sim}{\Gamma}^{-1} \underset{\sim}{B}\, \underset{\sim}{y}\, \underset{\sim}{y}'] ,$$

which in turn can be rewritten using lemma 3.1 as

$$\underset{\sim}{E}' (\underset{\sim}{y}\, \underset{\sim}{y}' \underset{\sim}{B}' \underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{\Gamma}^{-1}) \frac{\partial \ \mathrm{vec} \ \underset{\sim}{\Gamma}}{\partial \sigma_{ij}^{(s)}} .$$

This means

$$\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\beta}\, \partial \underset{\sim}{\tilde{\theta}}'} = \underset{\sim}{E}' (\underset{\sim}{y}\, \underset{\sim}{y}' \ \underset{\sim}{B}' \underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{\Gamma}^{-1}) \frac{\partial \ \mathrm{vec} \ \underset{\sim}{\Gamma}}{\partial \underset{\sim}{\tilde{\theta}}} .$$

But (3.2.15) states

$$\frac{\partial \ \mathrm{vec} \ \underset{\sim}{\Gamma}}{\partial \underset{\sim}{\tilde{\theta}}'} = \underset{\sim}{F} .$$

So

$$\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\beta}\, \partial \underset{\sim}{\tilde{\theta}}'} = \underset{\sim}{E}' (\underset{\sim}{y}\, \underset{\sim}{y}' \underset{\sim}{B}' \underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{\Gamma}^{-1}) \underset{\sim}{F} .$$

Finally we get

$$(3.3.7) \qquad \frac{\partial^2 \log \ell}{\partial \beta\, \partial \theta'} = \frac{\partial^2 \log \ell}{\partial \beta\, \partial \tilde{\theta}'} G' = E'(y y' B' \Gamma^{-1} \otimes \Gamma^{-1}) F G' .$$

<u>Derivation of</u> $\dfrac{\partial^2 \log \ell}{\partial \theta\, \partial \theta'}$.

As in (3.2.17) we have

$$\frac{\partial \log \ell}{\partial \tilde{\theta}} = \frac{1}{2} F' [\Gamma^{-1} B\, y \otimes \Gamma^{-1} B\, y - \operatorname{vec} \Gamma^{-1}] .$$

Differentiating this with respect to $\sigma_{ij}^{(s)}$ we get

$$(3.3.8) \qquad \frac{\partial^2 \log \ell}{\partial \tilde{\theta}\, \partial \sigma_{ij}^{(s)}} = \frac{1}{2} F' [-\Gamma^{-1} \frac{\partial \Gamma}{\partial \sigma_{ij}^{(s)}} \Gamma^{-1} B y \otimes \Gamma^{-1} B y$$

$$- \Gamma^{-1} B y \otimes \Gamma^{-1} \frac{\partial \Gamma}{\partial \sigma_{ij}^{(s)}} \Gamma^{-1} B y + \operatorname{vec} \Gamma^{-1} \frac{\partial \Gamma}{\partial \sigma_{ij}^{(s)}} \Gamma^{-1}] .$$

Now, using lemma 3.4 the right hand side of (3.3.8) can be rewritten as

$$\frac{1}{2} F' \operatorname{vec} [-\Gamma^{-1} B y y' B' \Gamma^{-1} \frac{\partial \Gamma}{\partial \sigma_{ij}^{(s)}} \Gamma^{-1} - \Gamma^{-1} \frac{\partial \Gamma}{\partial \sigma_{ij}^{(s)}} \Gamma^{-1} B y y' B' \Gamma^{-1}$$

$$+ \Gamma^{-1} \frac{\partial \Gamma}{\partial \sigma_{ij}^{(s)}} \Gamma^{-1}] .$$

And using lemma 3.1 this can be rewritten as

(3.3.9) $$-\frac{1}{2}\underset{\sim}{F}'[\underset{\sim}{\Gamma}^{-1}\otimes\underset{\sim}{\Gamma}^{-1}\underset{\sim}{\mathbf{z}}\underset{\sim}{y}\underset{\sim}{y}'\underset{\sim}{\mathbf{z}}'\underset{\sim}{\Gamma}^{-1}+\underset{\sim}{\Gamma}^{-1}\underset{\sim}{\mathbf{z}}\underset{\sim}{y}\underset{\sim}{y}'\underset{\sim}{\mathbf{z}}'\underset{\sim}{\Gamma}^{-1}\otimes\underset{\sim}{\Gamma}^{-1}-\underset{\sim}{\Gamma}^{-1}\otimes\underset{\sim}{\Gamma}^{-1}]\frac{\partial\,\mathrm{vec}\,\underset{\sim}{\Gamma}}{\partial\sigma_{ij}^{(s)}}$$

$$= -\frac{1}{2}\underset{\sim}{F}'\underset{\sim}{M}\frac{\partial\,\mathrm{vec}\,\underset{\sim}{\Gamma}}{\partial\sigma_{ij}^{(s)}}\ , \quad \text{say .}$$

Using this and (3.2.15) we get

$$\frac{\partial^2 \log \ell}{\partial\underset{\sim}{\tilde{\theta}}\,\partial\underset{\sim}{\tilde{\theta}}'} = -\frac{1}{2}\underset{\sim}{F}'\underset{\sim}{M}\frac{\partial\,\mathrm{vec}\,\underset{\sim}{\Gamma}}{\partial\underset{\sim}{\tilde{\theta}}'} = -\frac{1}{2}\underset{\sim}{F}'\underset{\sim}{M}\underset{\sim}{F}\ .$$

Finally

(3.3.10) $$\frac{\partial^2 \log \ell}{\partial\underset{\sim}{\theta}\,\partial\underset{\sim}{\theta}'} = \underset{\sim}{G}\frac{\partial^2 \log \ell}{\partial\underset{\sim}{\tilde{\theta}}\,\partial\underset{\sim}{\tilde{\theta}}'}\underset{\sim}{G}' = -\frac{1}{2}\underset{\sim}{G}\underset{\sim}{F}'\underset{\sim}{M}\underset{\sim}{F}\underset{\sim}{G}'\ .$$

Now, putting (3.3.6), (3.3.7) and (3.3.10) together we get

(3.3.11) $$\frac{\partial^2 \log \ell}{\partial\underset{\sim}{\rho}\,\partial\underset{\sim}{\rho}'} = -\begin{pmatrix}\underset{\sim}{E}' & \underset{\sim}{0}\\ \underset{\sim}{0} & \underset{\sim}{G}\underset{\sim}{F}'\end{pmatrix}\underset{\sim}{\Pi}\begin{pmatrix}\underset{\sim}{E} & \underset{\sim}{0}\\ \underset{\sim}{0} & \underset{\sim}{F}\underset{\sim}{G}'\end{pmatrix}\ ,$$

where

(3.3.12) $$\underset{\sim}{\Pi} = \begin{pmatrix}\underset{\sim}{y}\underset{\sim}{y}'\otimes\underset{\sim}{\Gamma}^{-1} & \underset{\sim}{y}\underset{\sim}{y}'\underset{\sim}{\mathbf{z}}'\underset{\sim}{\Gamma}^{-1}\otimes\underset{\sim}{\Gamma}^{-1}\\ \underset{\sim}{\Gamma}^{-1}\underset{\sim}{\mathbf{z}}\underset{\sim}{y}\underset{\sim}{y}'\otimes\underset{\sim}{\Gamma}^{-1} & \frac{1}{2}\underset{\sim}{M}\end{pmatrix}\ .$$

Now we are ready to write down the equations for the Newton-Raphson method. They are

$$(3.3.13) \qquad \begin{pmatrix} \underset{\sim}{E}' & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{G}\underset{\sim}{F}' \end{pmatrix} \hat{\underset{\sim}{\Pi}}_{(0)} \begin{pmatrix} \underset{\sim}{E} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{F}\underset{\sim}{G}' \end{pmatrix} (\hat{\underset{\sim}{\rho}}_{(1)} - \hat{\underset{\sim}{\rho}}_{(0)}) = \hat{\underset{\sim}{\delta}}_{(0)} ,$$

where $\underset{\sim}{\delta} = \frac{\partial \log \ell}{\partial \underset{\sim}{\rho}}$ which is given by (3.2.7) and (3.2.18). Once we have $\hat{\underset{\sim}{\rho}}_{(1)}$, we could carry out a second iteration by replacing $\hat{\underset{\sim}{\rho}}_{(1)}$ with $\hat{\underset{\sim}{\rho}}_{(2)}$, $\hat{\underset{\sim}{\rho}}_{(0)}$ with $\hat{\underset{\sim}{\rho}}_{(1)}$, $\hat{\underset{\sim}{\Pi}}_{(0)}$ with $\hat{\underset{\sim}{\Pi}}_{(1)}$ and $\hat{\underset{\sim}{\delta}}_{(0)}$ with $\hat{\underset{\sim}{\delta}}_{(1)}$ in (3.3.13) and solve for $\hat{\underset{\sim}{\rho}}_{(2)}$. But even for samples of moderate size this would be computationally very costly.

<u>The Scoring Method.</u>

The equation for this method is given by (3.3.2). We notice that we have to find

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\rho}\, \partial \underset{\sim}{\rho}'}\right) .$$

Taking expectations of both sides of (3.3.11) we get

$$(3.3.14) \qquad \mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\rho}\, \partial \underset{\sim}{\rho}'}\right) = - \begin{pmatrix} \underset{\sim}{E}' & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{G}\underset{\sim}{F}' \end{pmatrix} \mathcal{E}\, \underset{\sim}{\Pi} \begin{pmatrix} \underset{\sim}{E} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{F}\underset{\sim}{G}' \end{pmatrix} .$$

Now recall that

$$\mathcal{E}(\underset{\sim}{y}\underset{\sim}{y}') \doteq \underset{\sim}{B}^{-1} \underset{\sim}{\Gamma}\, \underset{\sim}{B}'^{-1} .$$

Applying this to (3.3.9) we get

$$\mathcal{E}(\underset{\sim}{M}) \doteq \underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{\Gamma}^{-1} .$$

We can also easily find the expectation of the other entries of $\underset{\sim}{\Pi}$. The end result is

$$\mathcal{E}\,\underset{\sim}{\Pi} = \begin{pmatrix} \underset{\sim}{S}^{-1}\underset{\sim}{\Gamma}\underset{\sim}{S}'^{-1}\otimes\underset{\sim}{\Gamma}^{-1} & \underset{\sim}{S}^{-1}\otimes\underset{\sim}{\Gamma}^{-1} \\ \underset{\sim}{S}'^{-1}\otimes\underset{\sim}{\Gamma}^{-1} & \frac{1}{2}(\underset{\sim}{\Gamma}^{-1}\otimes\underset{\sim}{\Gamma}^{-1}) \end{pmatrix}$$

$$= \underset{\sim}{\Delta}\ , \quad \text{say.}$$

Substituting this in (3.3.14) we get

$$(3.3.15) \qquad \mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\rho}\, \partial \underset{\sim}{\rho}'}\right) = -\begin{pmatrix} \underset{\sim}{E}' & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{G}\underset{\sim}{F}' \end{pmatrix} \underset{\sim}{\Delta} \begin{pmatrix} \underset{\sim}{E} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{F}\underset{\sim}{G}' \end{pmatrix} .$$

So finally the equations for the Scoring method are

$$(3.3.16) \qquad \begin{pmatrix} \underset{\sim}{E}' & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{G}\underset{\sim}{F}' \end{pmatrix} \hat{\underset{\sim}{\Delta}}_{(0)} \begin{pmatrix} \underset{\sim}{E} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{F}\underset{\sim}{G}' \end{pmatrix} (\hat{\underset{\sim}{\rho}}_{(1)} - \hat{\underset{\sim}{\rho}}_{(0)}) = \hat{\underset{\sim}{\delta}}_{(0)} .$$

3.4. The Pure Moving Average Case.

The pure moving average case corresponds to the model defined by (2.1) with $\underset{\sim}{\beta} = \underset{\sim}{0}$. This case was treated by Ahrabi (1978). But we can also derive the estimation equations from (3.3.13) and (3.3.16) by letting

$$\underset{\sim}{\beta} = 0\ , \quad \underset{\sim}{\rho} = \underset{\sim}{\theta}\ .$$

The resulting equations are in fact simplified versions of those of Ahrabi (1978). This is because in the latter, the second order derivatives of the log likelihood have a more complex representation and also that for the Scoring method instead of

$$- \mathcal{E} \frac{\partial^2 \log \ell}{\partial \underset{\sim}{\theta} \, \partial \underset{\sim}{\theta}'}$$

we had used

$$\mathcal{E}\left(\frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} \cdot \frac{\partial \log \ell}{\partial \underset{\sim}{\theta}'}\right) ,$$

which proved to be more cumbersome.

An important distinction between the general case and the case of the pure moving average model is that the log likelihood for the latter case, as derived from (3.1.8) by letting $\underset{\sim}{\beta} = \underset{\sim}{0}$, is the exsct log likelihood of the data. That is we do not need to assume that some past values of $\underset{\sim}{y}_t$ are fixed.

The model, as pointed out above, is

$$\underset{\sim}{y}_t = \underset{\sim}{\epsilon}_t + \underset{\sim}{A}_1 \underset{\sim}{\epsilon}_{t-1} + \cdots + \underset{\sim}{A}_q \underset{\sim}{\epsilon}_{t-q} .$$

The parameters to be estimated are the components of $\underset{\sim}{\theta}$ as defined in Chapter 2, with $\underset{\sim}{y}_t \equiv \underset{\sim}{u}_t$. And obviously in this case we only need assumptions 1 and 3.

The Newton Raphson Method.

We get the second order derivative of log ℓ with respect to $\underset{\sim}{\theta}$, by letting $\underset{\sim}{B} = \underset{\sim}{I}_m$ in (3.3.10) which yields

$$\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\theta} \, \partial \underset{\sim}{\theta}'} = - \frac{1}{2} \underset{\sim}{G} \underset{\sim}{F}' \underset{\sim}{M} \underset{\sim}{F} \underset{\sim}{G}' , \tag{3.4.1}$$

where

$$(3.4.2)\qquad \underset{\sim}{M} = \underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{\Gamma}^{-1} \underset{\sim}{y}\underset{\sim}{y}' \underset{\sim}{\Gamma}^{-1} + \underset{\sim}{\Gamma}^{-1} \underset{\sim}{y}\underset{\sim}{y}' \underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{\Gamma}^{-1} - \underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{\Gamma}^{-1} .$$

Now let

$$(3.4.3)\qquad \tilde{\underset{\sim}{y}} = \begin{pmatrix} \underset{\sim}{y}_1 \\ \vdots \\ \underset{\sim}{y}_T \end{pmatrix} = \text{vec } \underset{\sim}{Y}' = K'_{m,T} \text{ vec } \underset{\sim}{Y} = \underset{\sim}{K}'_{m,T} \underset{\sim}{y} ,$$

where $\underset{\sim}{Y}$ was introduced in (3.1.2). We also recall that

$$(3.4.4)\qquad \underset{\sim}{\Gamma} = \underset{\sim}{K}_{m,T} \underset{\sim}{\Sigma} \underset{\sim}{K}'_{m,T} ,$$

where $\underset{\sim}{\Sigma}$ was defined by (3.1.6). From (3.4.3) we get

$$\underset{\sim}{y} = \underset{\sim}{K}'^{-1}_{m,T} \tilde{\underset{\sim}{y}} = \underset{\sim}{K}_{m,T} \tilde{\underset{\sim}{y}} .$$

Substituting this and $\underset{\sim}{\Gamma}$ as in (3.4.4) in (3.4.2) we get

$$\begin{aligned}(3.4.5)\qquad \underset{\sim}{M} &= (\underset{\sim}{K}'_{m,T} \otimes \underset{\sim}{K}'_{m,T})(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1} \underset{\sim}{y}\underset{\sim}{y}' \underset{\sim}{\Sigma}^{-1} + \underset{\sim}{\Sigma}^{-1} \tilde{\underset{\sim}{y}}\tilde{\underset{\sim}{y}}' \underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1} \\ &\qquad - \underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1})(\underset{\sim}{K}_{m,T} \otimes \underset{\sim}{K}_{m,T}) \\ &= (\underset{\sim}{K}_{m,T} \otimes \underset{\sim}{K}_{m,T})\, \underset{\sim}{N} (\underset{\sim}{K}'_{m,T} \otimes \underset{\sim}{K}'_{m,T}) ,\end{aligned}$$

say. Now if we let

$$(3.4.6) \qquad (\underset{\sim}{K}'_{m,T} \otimes \underset{\sim}{K}'_{m,T})\underset{\sim}{F} = \underset{\sim}{\tilde{F}} \; ,$$

then (3.4.1) can be rewritten as

$$(3.4.7) \qquad \frac{\partial^2 \log \ell}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'} = -\frac{1}{2}\, \underset{\sim}{G}\underset{\sim}{\tilde{F}}'\, \underset{\sim}{N}\underset{\sim}{\tilde{F}}\underset{\sim}{G}' \; .$$

We note that

$$(3.4.8) \qquad (\underset{\sim}{K}_{m,T} \otimes \underset{\sim}{K}_{m,T}) \operatorname{vec}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^{s}) = \operatorname{vec}[\underset{\sim}{K}_{m,T}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^{s})\underset{\sim}{K}'_{m,T}] = \operatorname{vec}(\underset{\sim}{L}^{s} \otimes \underset{\sim}{E}_{ij}) \; ,$$

using Lemmas 3.1 and 3.3. It is now clear that $\underset{\sim}{\tilde{F}}$ is what was called $\underset{\sim}{E}$ and $\underset{\sim}{\tilde{y}}$ is what was called $\underset{\sim}{y}$ in Ahrabi (1978). The first order derivative, in the same manner, is derived from (3.2.19) which yields

$$\frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} = \frac{1}{2}\, \underset{\sim}{G}\underset{\sim}{\tilde{F}}'[\underset{\sim}{\Sigma}^{-1}\underset{\sim}{\tilde{y}} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{\tilde{y}} - \operatorname{vec} \underset{\sim}{\Sigma}^{-1}] \; ,$$

which is identical to (3.2.11) of Ahrabi (1978). We need an initial estimate for $\underset{\sim}{\theta}$ which is derived from (3.3.5) if we replace $\hat{\underset{\sim}{u}}_t$ by $\underset{\sim}{y}_t$. That is

$$\hat{\underset{\sim}{\Sigma}}^{(s)}_{(0)} = \frac{1}{T}\sum_{t=1}^{T-s} \underset{\sim}{y}_t \underset{\sim}{y}'_{t+s} \; , \quad s = 0,1, \ldots, q \; .$$

So finally the Newton-Raphson equations are

$$\underset{\sim}{G}\underset{\sim}{\tilde{F}}'\, \hat{\underset{\sim}{N}}_{(0)}\underset{\sim}{\tilde{F}}\underset{\sim}{G}'\, (\hat{\underset{\sim}{\theta}}_{(1)} - \hat{\underset{\sim}{\theta}}_{(0)}) = \underset{\sim}{G}\underset{\sim}{\tilde{F}}'\, [\hat{\underset{\sim}{\Sigma}}^{-1}_{(0)}\underset{\sim}{\tilde{y}} \otimes \hat{\underset{\sim}{\Sigma}}^{-1}_{(0)}\underset{\sim}{\tilde{y}} - \operatorname{vec} \hat{\underset{\sim}{\Sigma}}^{-1}_{(0)}] \; .$$

The Scoring Method.

We need to find $\mathcal{E}\,\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}$. We get this from (3.3.15) which yields

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}\right) = -\frac{1}{2}\, \underset{\sim}{G}\underset{\sim}{F}'(\underset{\sim}{\Gamma}^{-1} \otimes \underset{\sim}{\Gamma}^{-1})\underset{\sim}{F}'\underset{\sim}{G} = -\frac{1}{2}\, \underset{\sim}{G}\underset{\sim}{\tilde{F}}'(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1})\underset{\sim}{\tilde{F}}'\underset{\sim}{G} .$$

So the equations for this method are

$$\underset{\sim}{G}\underset{\sim}{F}'(\hat{\underset{\sim}{\Sigma}}_{(0)}^{-1} \otimes \hat{\underset{\sim}{\Sigma}}_{(0)}^{-1})\underset{\sim}{F}\underset{\sim}{G}'(\hat{\underset{\sim}{\theta}}_{(1)} - \underset{\sim}{\theta}_{(0)}) = \underset{\sim}{G}\underset{\sim}{\tilde{F}}'(\hat{\underset{\sim}{\Sigma}}_{(0)}^{-1}\underset{\sim}{V} \otimes \hat{\underset{\sim}{\Sigma}}_{(0)}^{-1}\underset{\sim}{V} - \operatorname{vec} \hat{\underset{\sim}{\Sigma}}_{(0)}^{-1}) .$$

3.5. The Problem of Computation.

If we look at equation (3.3.13) and (3.3.16), which are the equations for the Newton-Raphson and Scoring methods respectively, we realize that to get $\hat{\underset{\sim}{\theta}}_{(1)}$ we have to tackle two computational problems:

(i) The problem of computing $\hat{\underset{\sim}{\Pi}}_{(0)}$, $\hat{\underset{\sim}{\Delta}}_{(0)}$, $\hat{\underset{\sim}{\delta}}_{(0)}$.

(ii) The problem of solving the resulting equations.

The second problem is the less serious one, because it involves solving a set of $r = (p+q)m^2 + m(m+1)/2$ linear equations. Even though r can be large it remains fixed as the number of observations (T) increases. By comparison in the first, as we shall see, the matrices that are to be inverted have dimensions of order T. So we shall concentrate on (i).

If we look at (3.3.12) and (3.3.15) it becomes apparent that the major computational problem for computing $\hat{\underset{\sim}{\Pi}}_{(0)}$ and $\hat{\underset{\sim}{\Delta}}_{(0)}$ is the problem of inversion of $\underset{\sim}{\mathfrak{B}}$ and $\underset{\sim}{\Gamma}$. This is also the case for $\hat{\underset{\sim}{\delta}}_{(0)}$ which is derived from (3.2.7) and (3.2.19).

Inversion of $\underset{\sim}{\mathfrak{B}}$.

We recall that as in (3.1.4) $\underset{\sim}{\mathfrak{B}}$ is given by

$$\underset{\sim}{\mathfrak{B}} = \sum_{i=0}^{p} (\underset{\sim}{B}_i \otimes \underset{\sim}{L}^i) ,$$

where

$$\underset{\sim}{L} = \begin{pmatrix} \underset{\sim}{0}' & 0 \\ \underset{\sim}{I}_{T-1} & \underset{\sim}{0} \end{pmatrix}.$$

Now, using lemma (3.3) we have

$$\underset{\sim}{K}_{m,T} \underset{\sim}{\mathfrak{B}} \underset{\sim}{K}'_{m,T} = \sum_{i=0}^{p} (\underset{\sim}{L}^i \otimes \underset{\sim}{B}_i) = \underset{\sim}{\tilde{\mathfrak{B}}} , \tag{3.5.1}$$

say. We notice that since $\underset{\sim}{L}^i$ is lower triangular, so is $\underset{\sim}{\tilde{\mathfrak{B}}}$. This makes it possible to find $\underset{\sim}{\tilde{\mathfrak{B}}}^{-1}$ via some recursive equations. It is clear from (3.5.1) that $\underset{\sim}{\tilde{\mathfrak{B}}}$ has 1's for the diagonal elements. This means

$$|\underset{\sim}{\mathfrak{B}}| = |\underset{\sim}{\tilde{\mathfrak{B}}}| = 1 , \tag{3.5.2}$$

which was used in deriving (3.1.8). We also notice that $\underset{\sim}{\tilde{\mathfrak{B}}}$ is block Toeplitz.

Lemma 3.5. For $\underset{\sim}{\tilde{\mathfrak{B}}}$ defined by (3.5.1)

$$\underset{\sim}{\tilde{\mathfrak{B}}}^{-1} = \sum_{j=0}^{T-1} (\underset{\sim}{L}^j \otimes \underset{\sim}{B}^{(j)}), \tag{3.5.3}$$

where $\underset{\sim}{B}^{(0)} = \underset{\sim}{I}_m$ and $\underset{\sim}{B}^{(j)}$, $j = 1, \ldots, T-1$ are given by the recursive equations

$$(3.5.4) \qquad \sum_{i+j=r} \underset{\sim}{B}_i \underset{\sim}{B}^{(j)} = \underset{\sim}{0}\ , \quad r = 1,2,\ldots, T-1\ .$$

Proof. Multiplying $\underset{\sim}{\tilde{\mathfrak{B}}}$ by the right hand side of (3.5.3) yields

$$(3.5.5) \qquad \underset{\sim}{\tilde{\mathfrak{B}}} \sum_{j=0}^{T-1} (\underset{\sim}{L}^j \otimes \underset{\sim}{B}^{(j)}) = \sum_{i=0}^{p} \sum_{j=0}^{T-1} (\underset{\sim}{L}^{i+j} \otimes \underset{\sim}{B}_i \underset{\sim}{B}^{(j)})$$

$$= \sum_{r=0}^{T-1} [\underset{\sim}{L}^r \otimes \sum_{i+j=r} \underset{\sim}{B}_i \underset{\sim}{B}^{(j)}]\ ,$$

since $\underset{\sim}{L}^{T+h} = 0$, $h = 0,1,\ldots$. Now the first term of (3.3.5) is

$$\underset{\sim}{L}^0 \otimes \underset{\sim}{B}_0 \underset{\sim}{B}^{(0)} = \underset{\sim}{I}_{mT}\ ,$$

so letting the left hand side of (3.3.5) be equal to $\underset{\sim}{I}_{mT}$ we get

$$\sum_{i+j=r} \underset{\sim}{B}_i \underset{\sim}{B}^{(j)} = \underset{\sim}{0}\ , \quad r = 1,2,\ldots, T-1\ .$$

If we write out the first few equations it becomes clear that these can easily be solved recursively. The first equation is

$$\underset{\sim}{B}_0 \underset{\sim}{B}^{(1)} + \underset{\sim}{B}_1 \underset{\sim}{B}^{(0)} = \underset{\sim}{0}\ ,$$

which yields

$$(3.5.6) \qquad \underset{\sim}{B}^{(1)} = -\underset{\sim}{B}_1\ .$$

The second equation is

$$\underset{\sim}{B}_0\underset{\sim}{B}^{(2)} + \underset{\sim}{B}_1\underset{\sim}{B}^{(1)} + \underset{\sim}{B}_2\underset{\sim}{B}^{(0)} = \underset{\sim}{0} ,$$

which yields

$$\underset{\sim}{B}^{(2)} = -\underset{\sim}{B}_1\underset{\sim}{B}^{(1)} - \underset{\sim}{B}_2 = \underset{\sim}{B}_1^2 - \underset{\sim}{B}_2 ,$$

using (3.5.6). So at stage n

$$\underset{\sim}{B}^{(n)} = -\underset{\sim}{B}_1\underset{\sim}{B}^{(n-1)} - \underset{\sim}{B}_2\underset{\sim}{B}^{(n-2)} \dots - \underset{\sim}{B}_n .$$

Notice that there are n terms here, but we know that for $n > p$, $\underset{\sim}{B}_n = \underset{\sim}{0}$. This means that for any n we shall at most have p terms. That is, the equations (3.5.4) are recursive of order p. This makes the computation task much easier. Q.E.D.

From (3.5.1) we get

$$\tilde{\underset{\sim}{\mathfrak{B}}}^{-1} = \underset{\sim}{K}_{m,T}\underset{\sim}{\mathfrak{B}}^{-1}\underset{\sim}{K}'_{m,T} , \tag{3.5.7}$$

where we have used the fact that $\underset{\sim}{K}_{m,T}$ is orthogonal. From (3.5.7) we get

$$\underset{\sim}{\mathfrak{B}}^{-1} = \underset{\sim}{K}'_{m,T}\tilde{\underset{\sim}{\mathfrak{B}}}^{-1}\underset{\sim}{K}_{m,T} ,$$

which together with (3.5.3) and lemma (3.3) yields -

(3.5.8)
$$\underset{\sim}{B}^{-1} = \sum_{j=0}^{T-1} (\underset{\sim}{B}^{(j)} \otimes \underset{\sim}{L}^{j}) .$$

The Problem of Inversion of $\underset{\sim}{\Gamma}$.

We recall that

(3.5.9)
$$\underset{\sim}{\Gamma} = \underset{\sim}{K}_{m,T} \underset{\sim}{\Sigma} \underset{\sim}{K}'_{m,T} ,$$

where $\underset{\sim}{\Sigma}$ is defined by (3.1.6). This means

$$\underset{\sim}{\Gamma}^{-1} = \underset{\sim}{K}_{m,T} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{K}'_{m,T} .$$

So the problem reduces to the problem of inversion of $\underset{\sim}{\Sigma}$. Now from (3.1.6) it is clear that $\underset{\sim}{\Sigma}$ is a symmetric, banded block-Toeplitz matrix, which makes it easier to compute $\underset{\sim}{\Sigma}^{-1}$. There are efficient algorithms for inversion of symmetric block-Toeplitz matrices, e.g. see Friedlander, Morf, Kailath and Ljung (1978). The idea is that for an $N \times N$ toeplitz matrix $\underset{\sim}{R}$ the inverse can be represented by

$$\underset{\sim}{R}^{-1} = \frac{1}{r_N} [\underset{\sim}{L}_1 \underset{\sim}{U}_1 - \underset{\sim}{L}_2 \underset{\sim}{U}_2] ;$$

where

$$\underset{\sim}{L}_1 = \begin{pmatrix} 1 & & & 0 \\ b_{1N} & \ddots & & \\ \vdots & \ddots & \ddots & \\ b_{NN} & \cdots & b_{1N} & 1 \end{pmatrix} ,$$

$$\underset{\sim}{L}_2 = \begin{pmatrix} 0 & & & 0 \\ a_{NN} & \ddots & & \\ \vdots & \ddots & \ddots & \\ a_{1N} & \cdots & a_{NN} & 0 \end{pmatrix},$$

$$\underset{\sim}{U}_1' = \begin{pmatrix} 1 & & & 0 \\ a_{1N} & \ddots & & \\ \vdots & \ddots & \ddots & \\ a_{NN} & \cdots & a_{1N} & 1 \end{pmatrix},$$

$$\underset{\sim}{U}_2' = \begin{pmatrix} 0 & & & 0 \\ b_{NN} & \ddots & & \\ \vdots & \ddots & \ddots & \\ b_{1N} & \cdots & b_{NN} & 0 \end{pmatrix}.$$

Now the sequences $\underset{\sim}{a}$ and $\underset{\sim}{b}$ can be found recursively. This has also been extended to the block-toeplitz case. Now, when we have a banded toeplitz matrix then there are simplifications in the recursive formulae (similar to the simplifications that arose in solving (3.5.4)) and this has been done, in the scalar case by Dickinson (1978). It appears that the method could be extended to the case of symmetric banded block-toeplitz matrices.

Additional Computing Hints.

After computing $\hat{\underset{\sim}{\Gamma}}^{-1}_{(0)}$ and $\hat{\underset{\sim}{\mathfrak{B}}}^{-1}_{(0)}$, there is one more tedious computation in the Scoring method and that is the computation of

$$\underset{\sim}{F}'(\hat{\underset{\sim}{\Gamma}}^{-1}_{(0)} \otimes \hat{\underset{\sim}{\Gamma}}^{-1}_{(0)})\underset{\sim}{F} . \tag{3.5.10}$$

If we look at the definition of $\underset{\sim}{F}$ in (3.2.14) and (3.2.16) we notice that to compute (3.5.10) we have to compute terms like

$$[\mathrm{vec}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^r)]'(\hat{\underset{\sim}{\Gamma}}^{-1}_{(0)} \otimes \hat{\underset{\sim}{\Gamma}}^{-1}_{(0)})\ \mathrm{vec}(\underset{\sim}{E}_{uv} \otimes \underset{\sim}{L}^s) , \tag{3.5.11}$$

$$i,j,u,v = 1, \ldots, m, \quad r,s = 0,1, \ldots, q .$$

And also terms that are essentially of the form (3.5.11) except that one or both of the $\underset{\sim}{L}$'s might be replaced by $\underset{\sim}{L}'$. Now using lemma (3.1), (3.5.11) can be rewritten as

$$[\mathrm{vec}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^r)]'\ \mathrm{vec}[\hat{\underset{\sim}{\Gamma}}^{-1}_{(0)}(\underset{\sim}{E}_{uv} \otimes \underset{\sim}{L}^s)\hat{\underset{\sim}{\Gamma}}^{-1}_{(0)}] ,$$

which, using lemma 1(i) of the Appendix, is equal to

$$\mathrm{tr}[\hat{\underset{\sim}{\Gamma}}^{-1}_{(0)}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^r)\hat{\underset{\sim}{\Gamma}}^{-1}_{(0)}(\underset{\sim}{E}_{vu} \otimes \underset{\sim}{L}'^s)] . \tag{3.5.12}$$

Now let

$$\hat{\underset{\sim}{\Gamma}}^{-1}_{(0)} = \begin{pmatrix} \underset{\sim}{\Gamma}^{11} & \cdots & \underset{\sim}{\Gamma}^{1m} \\ \vdots & & \\ \underset{\sim}{\Gamma}^{m1} & \cdots & \underset{\sim}{\Gamma}^{mm} \end{pmatrix} ,$$

then by just carrying out the multiplications in (3.5.12) we see that it is equal to

$$\mathrm{tr}(\underset{\sim}{\Gamma}^{ui}\underset{\sim}{L}^{r}\underset{\sim}{\Gamma}^{jv}\underset{\sim}{L}'^{s}) .$$

Note. Throughout this paper we have assumed that $\mathcal{E}\underset{\sim}{y}_t = 0$. However in practice the mean of $\underset{\sim}{y}_t$ is unknown and will be estimated by

$$\bar{\underset{\sim}{y}} = \frac{1}{T}\sum_{t=1}^{T}\underset{\sim}{y}_t .$$

Then the estimating equations will be the same as in this paper except that $\underset{\sim}{y}_t$ will be replaced by $\underset{\sim}{y}_t - \bar{\underset{\sim}{y}}$.

4. Estimation in the Frequency Domain.

4.1. Preliminaries.

For a stationary process $\{\underset{\sim}{z}_t, t=0,\pm 1, \ldots\}$ with mean zero and covariances $\underset{\sim}{D}_s = \mathcal{E}(\underset{\sim}{z}_t \underset{\sim}{z}'_{t+s})$, $s = 0, \pm 1, \ldots,$ the spectral density matrix $\underset{\sim}{f}$ is defined by

$$(4.1.1) \qquad \underset{\sim}{f}(\lambda) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \underset{\sim}{D}_s e^{-is\lambda} ,$$

if the sum converges. The covariances can be recovered from $\underset{\sim}{f}(\cdot)$ via

$$\underset{\sim}{D}_s = \int_{-\pi}^{\pi} \underset{\sim}{f}(\lambda) e^{is\lambda} d\lambda .$$

The sample analogue of the spectral density, the periodogram, is defined by

$$(4.1.2) \qquad \underset{\sim}{I}(\lambda) = \frac{1}{2\pi} \sum_{-(T-1)}^{T-1} \underset{\sim}{C}_s e^{-is\lambda} ,$$

where

$$\underset{\sim}{C}_s = \frac{1}{T} \sum_{t=1}^{T-s} \underset{\sim}{z}_t \underset{\sim}{z}'_{t+s} = \underset{\sim}{C}'_{-s}, \quad s = 0, 1, \ldots, T-1 .$$

We can also represent $\underset{\sim}{I}(\lambda)$ in terms of the discrete Fourier transforms

$$(4.1.3) \qquad \underset{\sim}{I}(\lambda) = \underset{\sim}{w}(\lambda) \underset{\sim}{w}^*(\lambda) ,$$

where

$$(4.1.4) \qquad \underset{\sim}{w}(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} \underset{\sim}{z}_t e^{it\lambda} ,$$

and "*" indicates "conjugate transpose". For a fuller treatment see Anderson (1971), Chapter 7.

If the process $\{\underset{\sim}{z}_t\}$ is Gaussian the log likelihood is

$$(4.1.5) \qquad \log \ell = -\frac{1}{2} |\underset{\sim}{D}| - \frac{1}{2} \underset{\sim}{z}'\underset{\sim}{D}^{-1}\underset{\sim}{z} - \frac{Tm}{2} \log 2\pi ,$$

where

$$\underset{\sim}{z} = (\underset{\sim}{z}_1', \ldots, \underset{\sim}{z}_T')' ,$$

$$\underset{\sim}{D} = \mathcal{E}\, \underset{\sim}{z}\underset{\sim}{z}' .$$

Following Whittle (1953,1961) and Dunsmuir and Hannon (1976), we will approximate the second term in (4.1.5) by

$$-\frac{1}{2} \sum_t \mathrm{tr}[\underset{\sim}{f}^{-1}(\lambda_t)\underset{\sim}{I}(\lambda_t)] ,$$

where $\underset{\sim}{f}(\lambda)$ and $\underset{\sim}{I}(\lambda)$ were defined by (4.1.1) and (4.1.2) and

$$\lambda_t = \frac{2\pi t}{T} , \quad t = 0,1, \ldots, T-1 .$$

We shall also approximate the remaining terms in (4.1.5) by

$$-\frac{1}{2} \sum_t \log|\underset{\sim}{f}(\lambda_t)| .$$

We shall show the relation of this approximation to that proposed by Dunsmuir and Hannan (1976), in the Appendix.

For our problem

$$\underset{\sim}{z}_t \equiv \underset{\sim}{y}_t \, , \quad \underset{\sim}{D}_s \equiv \underset{\sim}{V}_s \, .$$

It is well known that the spectral density for the process defined by (2.1) is given by

$$(4.1.6) \qquad \underset{\sim}{f}(\lambda) = \frac{1}{2\pi} \underset{\sim}{B}^{-1}(e^{i\lambda})\underset{\sim}{A}(e^{i\lambda})\underset{\sim}{V}\,\underset{\sim}{A}^{*}(e^{i\lambda})\underset{\sim}{B}^{*-1}(e^{i\lambda}) \, ,$$

where $\underset{\sim}{B}(\cdot)$, $\underset{\sim}{A}(\cdot)$ and $\underset{\sim}{V}$ were introduced in Chapter 2. (See Hannan (1970)). It follows that the spectral density of the moving average part is given by

$$(4.1.7) \qquad \underset{\sim}{G}(\lambda) = \frac{1}{2\pi}\, \underset{\sim}{A}(e^{i\lambda})\underset{\sim}{V}\,\underset{\sim}{A}^{*}(e^{i\lambda}) = \frac{1}{2\pi} \sum_{-q}^{q} \underset{\sim}{\Sigma}^{(s)} e^{-i\lambda s} \, ,$$

where we have used the definition given in (4.1.1). Now using (4.1.6) and (4.1.7) we get

$$(4.1.8) \qquad \underset{\sim}{f}(\lambda) = \underset{\sim}{B}^{-1}(e^{i\lambda})\underset{\sim}{G}(\lambda)\underset{\sim}{B}^{*-1}(e^{i\lambda}) \, .$$

Finally, we approximate the log likelihood by

$$(4.1.9) \qquad \Lambda = -\frac{1}{2}\sum_{t=0}^{T-1} \log|\underset{\sim}{f}_t| - \frac{1}{2}\sum_{t=0}^{T-1} \mathrm{tr}(\underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t) \, ,$$

where

$$\underset{\sim}{f}_t \equiv \underset{\sim}{f}(\lambda_t) \ , \quad \underset{\sim}{I}_t \equiv \underset{\sim}{I}(\lambda_t) \ .$$

As in the time domain, we shall use the Newton-Raphson and Scoring methods to arrive at approximate maximum likelihood estimates that are asymptotically efficient.

Note: For any square matrix A we have

$$\frac{\partial \log |\underset{\sim}{A}|}{\partial a_{ij}} = \frac{1}{|\underset{\sim}{A}|} \operatorname{cof} a_{ij} = (\underset{\sim}{A}^{-1})_{ji} \ .$$

Using this we get

$$\frac{\partial \log |\underset{\sim}{A}|}{\partial x} = \sum_{i,j} (\underset{\sim}{A}^{-1})_{ji} \left(\frac{\partial \underset{\sim}{A}}{\partial x}\right)_{ij} = \operatorname{tr}(\underset{\sim}{A}^{-1} \frac{\partial \underset{\sim}{A}}{\partial x}) \ .$$

Using this on (4.1.9) we get

$$(4.1.10) \qquad \frac{\partial \Lambda}{\partial x} = -\frac{1}{2} \sum_t \operatorname{tr}(\underset{\sim}{f}_t^{-1} \frac{\partial \underset{\sim}{f}_t}{\partial x}) + \frac{1}{2} \sum_t \operatorname{tr}(\underset{\sim}{f}_t^{-1} \frac{\partial \underset{\sim}{f}_t}{\partial x} \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)$$

$$= -\frac{1}{2} \frac{\partial \sum_t \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial x} + \frac{1}{2} \left. \frac{\partial \sum_t \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial x} \right|_{\underset{\sim}{I}_t = \underset{\sim}{f}_t} \ .$$

4.2. The First Derivative of Λ.

Derivation of $\frac{\partial \Lambda}{\partial \underset{\sim}{\beta}}$.

To find this we need to find the first order derivatives of $\underset{\sim}{f}_t$. Now (4.1.8) can be written as

$$f_t = B_t^{-1} Q_t B_t^{*-1}, \quad t = 0,1, \ldots, T-1, \tag{4.2.1}$$

and $B_t \equiv B(e^{i\lambda_t})$, $Q_t \equiv Q(\lambda_t)$. **Differentiating** (4.2.1) **with respect** to $b_{rs}^{(h)}$ we get

$$\frac{\partial f_t}{\partial b_{rs}^{(h)}} = -B_t^{-1} \frac{\partial B_t}{\partial b_{rs}^{(h)}} B_t^{-1} Q_t B_t^{*-1} - B_t^{-1} Q_t B_t^{*-1} \frac{\partial B_t^*}{\partial b_{rs}^{(h)}} B_t^{*-1} \tag{4.2.2}$$

$$= -B_t^{-1} E_{rs} f_t e^{i\lambda_t h} - f_t E_{sr} B_t^{*-1} e^{-i\lambda_t h},$$

where we have differentiated (2.2) to get

$$\frac{\partial B_t}{\partial b_{rs}^{(h)}} = E_{rs} e^{i\lambda_t h}. \tag{4.2.3}$$

Using (4.2.2) we can get the derivative of $\mathrm{tr}(f_t^{-1} I_t)$ in the following manner

$$\frac{\partial\, \mathrm{tr}(f_t^{-1} I_t)}{\partial b_{rs}^{(h)}} = -\mathrm{tr}\left(f_t^{-1} \frac{\partial f_t}{\partial b_{rs}^{(h)}} f_t^{-1} I_t\right) \tag{4.2.4}$$

$$= \mathrm{tr}(f_t^{-1} B_t^{-1} E_{rs} f_t e^{i\lambda_t h} f_t^{-1} I_t)$$

$$+ \mathrm{tr}(f_t^{-1} f_t E_{sr} B_t^{*-1} e^{-i\lambda_t h} f_t^{-1} I_t)$$

$$= e^{i\lambda_t h}\, \mathrm{tr}(I_t f_t^{-1} B_t^{-1} E_{rs}) + e^{-i\lambda_t h}\, \mathrm{tr}(B_t^{*-1} f_t^{-1} I_t E_{sr})$$

$$= e^{-i\lambda_t h} (B_t^{*-1} f_t^{-1} I_t)_{rs} + e^{i\lambda_t h} (I_t f_t^{-1} B_t^{-1})_{sr},$$

$$r,s = 1,\ldots,m, \; h = 1,\ldots,p.$$

Notice that in the above expression the two terms are conjugates. To see this note that

$$e^{i\lambda_t h}(\underset{\sim}{I}_t \underset{\sim}{f}_t^{-1} \underset{\sim}{B}_t^{-1})_{sr} = e^{i\lambda_t h}(\underset{\sim}{I}_t \underset{\sim}{f}_t^{-1} \underset{\sim}{B}_t^{-1})'_{rs}$$

$$= e^{i\lambda_t h}(\underset{\sim}{B}_t' \underset{\sim}{f}_t'^{-1} \underset{\sim}{I}_t')_{rs} = e^{i\lambda_t h}\,\overline{(\underset{\sim}{B}_t^{*} \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}_{rs}$$

$$= \overline{e^{-i\lambda_t h}(\underset{\sim}{B}_t^{*} \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)_{rs}} \,.$$

Now using (4.1.9) and (4.1.10) we have

$$(4.2.5) \qquad \frac{\partial \Lambda}{\partial b_{rs}^{(h)}} = -\frac{1}{2} \sum_t \frac{\partial \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial b_{rs}^{(h)}} + \frac{1}{2} \sum_t \frac{\partial \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial b_{rs}^{(h)}}\Bigg|_{\underset{\sim}{I}_t = \underset{\sim}{f}_t} .$$

From (4.2.4)

$$\sum_t \frac{\partial \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial b_{rs}^{(h)}} = \sum_t e^{-i\lambda_t h}(\underset{\sim}{B}_t^{*-1} \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)_{rs}$$

$$+ \sum_t e^{i\lambda_t h}(\underset{\sim}{I}_t \underset{\sim}{f}_t^{-1} \underset{\sim}{B}_t^{-1})_{sr} = 2 \sum_t e^{-i\lambda_t h}(\underset{\sim}{B}_t^{*-1} \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)_{rs} \,,$$

since the sums are real and the summands are conjugates. The sums are real because the summands are functions of $e^{i\lambda_t}$ and because $e^{i\lambda_t} = e^{-i\lambda_{T-t}}$. So (4.2.5) can be written as

$$\frac{\partial \Lambda}{\partial b_{rs}^{(h)}} = -\sum_t e^{-i\lambda_t h}(\underset{\sim}{B}_t^{*-1}\underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t)_{rs} + \sum_t e^{-i\lambda_t h}(\underset{\sim}{B}_t^{*-1})_{rs} .$$

But we notice that the second term on the right hand side is $o(T)$. This is because

$$\lim_{T \to \infty} \frac{1}{T} \sum_t e^{-i\lambda_t h}(\underset{\sim}{B}_t^{*-1})_{rs} = \frac{1}{2\pi}\int_0^{2\pi} e^{-i\lambda h}(B^{*-1})_{rs} d\lambda = 0 ,$$

Since

$$e^{-i\lambda h}(B^{*-1})_{rs} = \sum_{j=0}^{\infty} d_j \, e^{-i\lambda(j+h)} ,$$

and

$$\int_0^{2\pi} e^{-i\lambda(j+h)} d\lambda = 0 \quad \text{for} \quad h \neq 0 \quad j = 0,1,\ldots .$$

<u>Note</u>: We can omit terms that are $o(T)$ because of the forms of the Newton-Raphson and Scoring method equations and the fact that

$$\frac{1}{T}\frac{\partial^2 \Lambda}{\partial \underset{\sim}{\rho}\, \partial \underset{\sim}{\rho}'} \to -\underset{\sim}{\vartheta}(\underset{\sim}{\rho}) .$$

So we finally get

$$\frac{\partial \Lambda}{\partial \underset{\sim}{\beta}_h} \doteq -\sum_t e^{-i\lambda_t h} \operatorname{vec}(\underset{\sim}{B}_t^{*-1}\underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t)$$

$$= -\sum_t e^{-i\lambda_t h} \operatorname{vec}(\underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t) , \quad h = 1,\ldots,p .$$

Putting these together for $h = 1, \ldots, p$ we get

$$(4.2.6) \qquad \frac{\partial\Lambda}{\partial \underset{\sim}{\beta}} \doteq - \sum_t \underset{\sim}{R}_t \; \mathrm{vec}(\underset{\sim}{Q}_t^{-1} \underset{\sim}{B}_t \underset{\sim}{I}_t) ,$$

where

$$(4.2.7) \qquad \underset{\sim}{R}_t' \equiv \underset{\sim}{R}'(\lambda_t) = (e^{-i\lambda_t}, e^{-2i\lambda_t}, \ldots, e^{-pi\lambda_t}) \otimes \underset{\sim}{I}_{m^2} .$$

Derivation of $\frac{\partial\Lambda}{\partial \underset{\sim}{\theta}}$.

We shall need the derivatives of $\underset{\sim}{f}_t$ which follow from (4.1.7) and (4.2.1)

$$(4.2.8) \qquad \frac{\partial \underset{\sim}{f}_t}{\partial \sigma_{k\ell}^{(0)}} = \underset{\sim}{B}_t^{-1} \frac{\partial \underset{\sim}{Q}_t}{\partial \sigma_{k\ell}^{(0)}} \underset{\sim}{B}_t^{*-1} = \frac{1}{2\pi} \underset{\sim}{B}_t^{-1} \underset{\sim}{E}_{k\ell} \underset{\sim}{B}_t^{*-1} ,$$

$$(4.2.9) \qquad \frac{\partial \underset{\sim}{f}_t}{\partial \sigma_{k\ell}^{(s)}} = \frac{1}{2\pi} \underset{\sim}{B}_t^{-1} (\underset{\sim}{E}_{k\ell} \, e^{-i\lambda_t s} + \underset{\sim}{E}_{\ell k} \, e^{i\lambda_t s}) \underset{\sim}{B}_t^{*-1} ,$$

$$k,\ell = 1,\ldots,m, \quad s = 1,\ldots,q .$$

Using these we get

$$\frac{\partial \, \mathrm{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial \sigma_{k\ell}^{(0)}} = -\mathrm{tr}[\underset{\sim}{f}_t^{-1} (\frac{1}{2\pi} \underset{\sim}{B}_t^{-1} \underset{\sim}{E}_{k\ell} \underset{\sim}{B}_t^{*-1}) \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t]$$

$$= - \frac{1}{2\pi} \mathrm{tr}(\underset{\sim}{B}_t^{*-1} \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t \underset{\sim}{f}_t^{-1} \underset{\sim}{B}_t^{-1} \underset{\sim}{E}_{k\ell})$$

$$= - \frac{1}{2\pi} \mathrm{tr}(\underset{\sim}{Q}_t^{-1} \underset{\sim}{B}_t \underset{\sim}{I}_t \underset{\sim}{B}_t^{*-1} \underset{\sim}{Q}_t^{-1} \underset{\sim}{E}_{k\ell})$$

$$= -\frac{1}{2\pi}\,(\underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t\underset{\sim}{B}_t^*\underset{\sim}{Q}_t^{-1})_{\ell k}$$

$$= -\frac{1}{2\pi}\,(\underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t\underset{\sim}{B}_t^*\underset{\sim}{Q}_t^{-1})'_{k\ell}$$

$$= -\frac{1}{2\pi}\,(\overline{\underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t\underset{\sim}{B}_t^*\underset{\sim}{Q}_t^{-1}})_{k\ell}\ ,$$

since the matrix in brackets is Hermitian. This yields

$$\frac{\partial\,\mathrm{tr}(\underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t)}{\partial\underset{\sim}{\tilde{\theta}}_0} = -\frac{1}{2\pi}\,\mathrm{vec}(\overline{\underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t\underset{\sim}{B}_t^*\underset{\sim}{Q}_t^{-1}})\ ,$$

which in turn yields

$$\text{(4.2.10)}\qquad \frac{\partial\sum_t\mathrm{tr}(\underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t)}{\partial\underset{\sim}{\tilde{\theta}}_0} = -\frac{1}{2\pi}\sum_t\mathrm{vec}(\overline{\underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t\underset{\sim}{B}_t^*\underset{\sim}{Q}_t^{-1}})$$

$$= -\frac{1}{2\pi}\sum_t\mathrm{vec}(\underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t\underset{\sim}{B}_t^*\underset{\sim}{Q}_t^{-1})\ ,$$

since the sum is real. Now, using (4.1.10)

$$\frac{\partial\Lambda}{\partial\underset{\sim}{\tilde{\theta}}_0} = -\frac{1}{2}\,\frac{\partial}{\partial\underset{\sim}{\tilde{\theta}}_0}\sum_t\mathrm{tr}(\underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t) + \frac{1}{2}\,\frac{\partial}{\partial\underset{\sim}{\tilde{\theta}}_0}\sum_t\mathrm{tr}(\underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t)\Big|_{\underset{\sim}{I}_t=\underset{\sim}{f}_t}\ .$$

Now

$$\text{(4.2.11)}\qquad \frac{\partial}{\partial\underset{\sim}{\tilde{\theta}}_0}\sum_t\mathrm{tr}(\underset{\sim}{f}^{-1}\underset{\sim}{I}_t)\Big|_{\underset{\sim}{I}_t=\underset{\sim}{f}_t} = -\frac{1}{2\pi}\sum_t\mathrm{vec}\ \underset{\sim}{Q}_t^{-1}\ .$$

This together with (4.2.10) yields

$$(4.2.12) \qquad \frac{\partial \Lambda}{\partial \underset{\sim}{\tilde{\theta}}_0} = \frac{1}{4\pi} \sum_t \operatorname{vec}(\underset{\sim}{Q}_t^{-1} \underset{\sim}{B}_t \underset{\sim}{I}_t \underset{\sim}{B}_t^* \underset{\sim}{Q}_t^{-1} - \underset{\sim}{Q}_t^{-1}) \,.$$

We proceed to derive $\dfrac{\partial \Lambda}{\partial \underset{\sim}{\tilde{\theta}}_s}$. Using (4.2.9) we have

$$(4.2.13) \qquad \frac{\partial \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial \sigma_{k\ell}^{(s)}} = -\frac{1}{2\pi} \operatorname{tr}[\underset{\sim}{f}_t^{-1} \underset{\sim}{B}_t^{-1} (\underset{\sim}{E}_{k\ell} e^{-i\lambda_t s} + \underset{\sim}{E}_{\ell k} e^{i\lambda_t s}) \underset{\sim}{B}_t^{*-1} \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t]$$

$$= -\frac{1}{2\pi} e^{i\lambda_t s} (\underset{\sim}{Q}_t^{-1} \underset{\sim}{B}_t \underset{\sim}{I}_t \underset{\sim}{B}_t^* \underset{\sim}{Q}_t^{-1})_{k\ell}$$

$$-\frac{1}{2\pi} e^{-i\lambda_t s} \overline{(\underset{\sim}{Q}_t^{-1} \underset{\sim}{B}_t \underset{\sim}{I}_t \underset{\sim}{B}_t^* \underset{\sim}{Q}_t^{-1})_{k\ell}} \,.$$

From this we get

$$\frac{\partial \sum_t \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial \sigma_{k\ell}^{(s)}} = -\frac{1}{\pi} \sum_t e^{i\lambda_t s} (\underset{\sim}{Q}_t^{-1} \underset{\sim}{B}_t \underset{\sim}{I}_t \underset{\sim}{B}_t^* \underset{\sim}{Q}_t^{-1})_{k\ell} \,.$$

This means

$$\frac{\partial \sum_t \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial \underset{\sim}{\tilde{\theta}}_s} = -\frac{1}{\pi} \sum_t e^{i\lambda_t s} \operatorname{vec}(\underset{\sim}{Q}_t^{-1} \underset{\sim}{B}_t \underset{\sim}{I}_t \underset{\sim}{B}_t^* \underset{\sim}{Q}_t^{-1}) \,,$$

which finally gives us

$$(4.2.14) \qquad \frac{\partial \Lambda}{\partial \underset{\sim}{\theta}_s} \doteq \frac{1}{2\pi} \sum_t e^{i\lambda_t s} \operatorname{vec}(\underset{\sim}{Q}_t^{-1} \underset{\sim}{B}_t \underset{\sim}{I}_t \underset{\sim}{B}_t^* \underset{\sim}{Q}_t^{-1} - \underset{\sim}{Q}_t^{-1}) \,, \quad s = 1, \ldots, q \,,$$

using (4.1.10). Putting (4.2.12) and (4.2.13) together, we get

$$(4.2.15)\qquad \frac{\partial\Lambda}{\partial\tilde{\theta}} \doteq \frac{1}{2\pi}\sum_t \underset{\sim}{J}_t \operatorname{vec}(\underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t\underset{\sim}{B}_t^*\underset{\sim}{Q}_t^{-1}-\underset{\sim}{Q}_t^{-1}) ,$$

where

$$(4.2.16)\qquad \underset{\sim}{J}_t' \equiv \underset{\sim}{J}'(\lambda_t) = (\tfrac{1}{2}, e^{i\lambda_t}, \ldots, e^{qi\lambda_t}) \otimes \underset{\sim}{I}_{m^2} .$$

Now, as in the time domain

$$\frac{\partial\Lambda}{\partial\underset{\sim}{\theta}} = \underset{\sim}{G}\,\frac{\partial\Lambda}{\partial\tilde{\theta}} .$$

So

$$(4.2.17)\qquad \frac{\partial\Lambda}{\partial\underset{\sim}{\theta}} = \frac{1}{2\pi}\sum_t \underset{\sim}{G}\,\underset{\sim}{J}_t \operatorname{vec}(\underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t\underset{\sim}{B}_t^*\underset{\sim}{Q}_t^{-1}-\underset{\sim}{Q}_t^{-1}) .$$

<u>Note</u>. There is an alternative form for (4.2.16) which we shall find more useful in deriving the second derivatives of Λ. To derive this alternative form we note that (4.2.13) can be rewritten as

$$(4.2.18)\qquad \frac{\partial\, \mathrm{tr}(\underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t)}{\partial\sigma_{k\ell}^{(s)}} = -\frac{1}{2\pi}\,[e^{i\lambda_t s}\,\underset{\sim}{g}_t + e^{-i\lambda_t s}\,\underset{\sim}{g}_t']_{k\ell} ,$$

where

$$\underset{\sim}{g}_t = \underset{\sim}{Q}_t^{-1}\underset{\sim}{B}_t\underset{\sim}{I}_t\underset{\sim}{B}_t^*\underset{\sim}{Q}_t^{-1} ,$$

and we have used the fact that $\underset{\sim}{g}_t$ is a Hermitian matrix. Now, from (4.2.18) we get

$$\frac{\partial\, \mathrm{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial \underset{\sim}{\theta}_s} = -\frac{1}{2\pi}[e^{i\lambda_t s} \mathrm{vec}\, \underset{\sim}{g}_t + e^{-i\lambda_t s} \mathrm{vec}\, \underset{\sim}{g}_t'] .$$

But using lemma (3.2) we have

$$\mathrm{vec}\, \underset{\sim}{g}_t' = \underset{\sim}{K}_{m,m}\, \mathrm{vec}\, \underset{\sim}{g}_t \equiv \underset{\sim}{K}_m\, \mathrm{vec}\, \underset{\sim}{g}_t .$$

So

$$\frac{\partial\, \mathrm{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)}{\partial \underset{\sim}{\theta}_s} = -\frac{1}{2\pi}[e^{i\lambda_t s} \underset{\sim}{I}_{m^2} + e^{-i\lambda_t s} \underset{\sim}{K}_m] \mathrm{vec}\, \underset{\sim}{g}_t ,$$

which gives us

$$\frac{\partial \Lambda}{\partial \underset{\sim}{\theta}_s} = \frac{1}{4\pi} \sum_t (e^{i\lambda_t s} \underset{\sim}{I}_{m^2} + e^{-i\lambda_t s} \underset{\sim}{K}_m)\, \mathrm{vec}(\underset{\sim}{g}_t - \underset{\sim}{Q}_t^{-1}) .$$

Using this and (4.2.12) we get

(4.2.19) $$\frac{\partial \Lambda}{\partial \underset{\sim}{\tilde{\theta}}} = \frac{1}{4\pi} \sum_t \underset{\sim}{H}_t\, \mathrm{vec}(\underset{\sim}{g}_t - \underset{\sim}{Q}_t^{-1}) ,$$

where

(4.2.20) $$\underset{\sim}{H}_t' \equiv \underset{\sim}{H}'(\lambda_t) = (1, e^{i\lambda_t}, \ldots, e^{qi\lambda_t}) \otimes \underset{\sim}{I}_{m^2} + (1, e^{-i\lambda_t}, \ldots, e^{-qi\lambda_t}) \otimes \underset{\sim}{K}_m .$$

Finally for $\frac{\partial \Lambda}{\partial \underset{\sim}{\theta}}$ we have the alternative form

(4.2.21) $$\frac{\partial \Lambda}{\partial \underset{\sim}{\theta}} = \frac{1}{4\pi} \sum_t \underset{\sim}{G}\, \underset{\sim}{H}_t\, \mathrm{vec}(\underset{\sim}{g}_t - \underset{\sim}{Q}_t^{-1}) .$$

4.3. Second Order Derivatives of Λ.

Derivation of $\frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \partial \underset{\sim}{\beta}'}$.

Using lemma (3.1) we can rewrite (4.2.6) as

$$\text{(4.3.1)} \qquad \frac{\partial \Lambda}{\partial \underset{\sim}{\beta}} = - \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \otimes \underset{\sim}{Q}_t^{-1}) \operatorname{vec} \underset{\sim}{B}_t .$$

Differentiating this with respect to $\underset{\sim}{\beta}'$ we get

$$\text{(4.3.2)} \qquad \frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \partial \underset{\sim}{\beta}'} = - \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \otimes \underset{\sim}{Q}_t^{-1}) \frac{\partial \operatorname{vec} \underset{\sim}{B}_t}{\partial \underset{\sim}{\beta}'} .$$

Now, from (2.2)

$$\operatorname{vec} \underset{\sim}{B}_t = \sum_{j=0}^{p} \underset{\sim}{\beta}_j \, e^{i\lambda_t j} ,$$

which gives us

$$\frac{\partial \operatorname{vec} \underset{\sim}{B}_t}{\partial \underset{\sim}{\beta}_j'} = e^{i\lambda_t j} \underset{\sim}{I}_{m^2} , \quad j = 1, \ldots, p .$$

Or

$$\frac{\partial \operatorname{vec} \underset{\sim}{B}_t}{\partial \underset{\sim}{\beta}'} = (e^{i\lambda_t}, \ldots, e^{pi\lambda_t}) \otimes \underset{\sim}{I}_{m^2} = \underset{\sim}{R}_t^* .$$

Substituting this in (4.3.2) we get

$$\text{(4.3.3)} \qquad \frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \partial \underset{\sim}{\beta}'} = - \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \otimes \underset{\sim}{Q}_t^{-1}) \underset{\sim}{R}_t^* .$$

Derivation of $\frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \partial \underset{\sim}{\theta}'}$.

We shall first find $\frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \partial \underset{\sim}{\tilde{\theta}}'}$ and then use

$$\frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \partial \underset{\sim}{\theta}'} = \frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \partial \underset{\sim}{\tilde{\theta}}'} \underset{\sim}{G}' \ . \tag{4.3.4}$$

Now, (4.2.6) can be rewritten as

$$\frac{\partial \Lambda}{\partial \underset{\sim}{\beta}} = - \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \underset{\sim}{B}_t' \otimes \underset{\sim}{I}_m) \text{vec } \underset{\sim}{Q}_t^{-1} \ .$$

Differentiating with respect to x we get

$$\frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \partial x} = - \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \underset{\sim}{B}_t' \otimes \underset{\sim}{I}_m) \frac{\partial \text{ vec } \underset{\sim}{Q}_t^{-1}}{\partial x}$$

$$= \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \underset{\sim}{B}_t' \otimes \underset{\sim}{I}_m) \text{ vec} (\underset{\sim}{Q}_t^{-1} \frac{\partial \underset{\sim}{Q}_t}{\partial x} \underset{\sim}{Q}_t^{-1})$$

$$= \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \underset{\sim}{B}_t' \otimes \underset{\sim}{I}_m)(\underset{\sim}{Q}_t'^{-1} \otimes \underset{\sim}{Q}_t^{-1}) \frac{\partial \text{ vec } \underset{\sim}{Q}_t}{\partial x}$$

$$= \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \underset{\sim}{B}_t' \underset{\sim}{Q}_t'^{-1} \otimes \underset{\sim}{Q}_t^{-1}) \frac{\partial \text{ vec } \underset{\sim}{Q}_t}{\partial x} \ .$$

By letting x be the components of $\underset{\sim}{\tilde{\theta}}'$ we get

$$\frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \partial \underset{\sim}{\tilde{\theta}}'} = \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \underset{\sim}{B}_t' \underset{\sim}{Q}_t'^{-1} \otimes \underset{\sim}{Q}_t^{-1}) \frac{\partial \text{ vec } \underset{\sim}{Q}_t}{\partial \underset{\sim}{\tilde{\theta}}'} \ . \tag{4.3.5}$$

Now, as in (4.2.8) and (4.2.9)

$$\frac{\partial \underset{\sim}{Q}_t}{\partial \sigma_{k\ell}^{(0)}} = \frac{1}{2\pi} \underset{\sim}{E}_{k\ell},$$

$$\frac{\partial \underset{\sim}{Q}_t}{\partial \sigma_{k\ell}^{(s)}} = \frac{1}{2\pi} [e^{-i\lambda_t s} \underset{\sim}{E}_{k\ell} + \underset{\sim}{E}_{\ell k} e^{i\lambda_t s}] .$$

Vectorizing these results in

$$(4.3.6) \qquad \frac{\partial \text{ vec } \underset{\sim}{Q}_t}{\partial \sigma_{k\ell}^{(0)}} = \frac{1}{2\pi} \text{ vec}(\underset{\sim}{E}_{k\ell}) = \frac{1}{2\pi} \underset{\sim}{e}_{k\ell}, \text{ say },$$

$$(4.3.7) \qquad \frac{\partial \text{ vec } \underset{\sim}{Q}_t}{\partial \sigma_{k\ell}^{(s)}} = \frac{1}{2\pi} [e^{-i\lambda_t s} \underset{\sim}{e}_{k\ell} + e^{i\lambda_t s} \underset{\sim}{e}_{\ell k}] .$$

But $\underset{\sim}{e}_{\ell k} = \text{vec } \underset{\sim}{E}_{\ell k} = \text{vec } \underset{\sim}{E}_{k\ell}' = \underset{\sim}{K}_m \underset{\sim}{e}_{k\ell}$, using lemma (3.2). So

$$(4.3.8) \qquad \frac{\partial \text{ vec } \underset{\sim}{Q}_t}{\partial \sigma_{k\ell}^{(s)}} = \frac{1}{2\pi} [e^{-i\lambda_t s} \underset{\sim}{I}_{m^2} + e^{i\lambda_t s} \underset{\sim}{K}_m] \underset{\sim}{e}_{k\ell} .$$

Now (4.3.6) and (4.3.8) yield

$$(4.3.9) \qquad \frac{\partial \text{ vec } \underset{\sim}{Q}_t}{\partial \underset{\sim}{\tilde{\theta}}_0'} = \frac{1}{2\pi} [\underset{\sim}{e}_{11}, \underset{\sim}{e}_{21}, \ldots, \underset{\sim}{e}_{mm}] = \frac{1}{2\pi} \underset{\sim}{I}_{m^2} ,$$

$$(4.3.10) \qquad \frac{\partial \text{ vec } \underset{\sim}{Q}_t}{\partial \underset{\sim}{\tilde{\theta}}_s'} = \frac{1}{2\pi} (e^{i\lambda_t s} \underset{\sim}{I}_{m^2} + e^{i\lambda_t s} \underset{\sim}{K}_m) .$$

Finally putting these together we have

$$(4.3.11) \qquad \frac{\partial \text{ vec } \underset{\sim}{Q}_t}{\partial \underset{\sim}{\tilde{\theta}}'} = \frac{1}{2\pi} \underset{\sim}{H}_t^* .$$

Substituting this in (4.3.5) we have

$$\frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \, \partial \underset{\sim}{\tilde{\theta}}'} = \frac{1}{2\pi} \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \underset{\sim}{B}_t' \underset{\sim}{Q}_t'^{-1} \otimes \underset{\sim}{Q}_t^{-1}) \underset{\sim}{H}_t^* .$$

Substituting this in (4.3.4) yields

$$(4.3.12) \qquad \frac{\partial^2 \Lambda}{\partial \underset{\sim}{\beta} \, \partial \underset{\sim}{\theta}'} = \frac{1}{2\pi} \sum_t \underset{\sim}{R}_t (\underset{\sim}{I}_t' \underset{\sim}{B}_t' \underset{\sim}{Q}_t'^{-1} \otimes \underset{\sim}{Q}_t^{-1}) \underset{\sim}{H}_t^* \underset{\sim}{G}' .$$

<u>Derivation of</u> $\dfrac{\partial^2 \Lambda}{\partial \underset{\sim}{\theta} \, \partial \underset{\sim}{\theta}'}$.

Differentiating (4.2.21) with respect to $\underset{\sim}{\theta}'$ we get

$$(4.3.13) \qquad \frac{\partial^2 \Lambda}{\partial \underset{\sim}{\theta} \, \partial \underset{\sim}{\theta}'} = \frac{1}{4\pi} \sum_t \underset{\sim}{G} \underset{\sim}{H}_t^* \frac{\partial \text{ vec } \underset{\sim}{P}_t}{\partial \underset{\sim}{\theta}'} ,$$

where

$$\underset{\sim}{P}_t = \underset{\sim}{Q}_t^{-1} \underset{\sim}{B}_t \underset{\sim}{I}_t \underset{\sim}{B}_t^* \underset{\sim}{Q}_t^{-1} - \underset{\sim}{Q}_t^{-1} = \underset{\sim}{Q}_t^{-1} (\underset{\sim}{B}_t \underset{\sim}{I}_t \underset{\sim}{B}_t^* - \underset{\sim}{Q}_t) \underset{\sim}{Q}_t^{-1} .$$

So we need to find the first order derivatives of $\underset{\sim}{P}_t$. Differentiating $\underset{\sim}{P}_t$ with respect to x we get

$$\frac{\partial \underset{\sim}{P}_t}{\partial x} = -\underset{\sim}{Q}_t^{-1} \frac{\partial \underset{\sim}{Q}_t}{\partial x} \underset{\sim}{P}_t - \underset{\sim}{P}_t \frac{\partial \underset{\sim}{Q}_t}{\partial x} \underset{\sim}{Q}_t^{-1} - \underset{\sim}{Q}_t^{-1} \frac{\partial \underset{\sim}{Q}}{\partial x} \underset{\sim}{Q}_t^{-1} .$$

Vectorizing both sides yields

$$(4.3.14) \qquad \frac{\partial \text{ vec } \underset{\sim}{P}_t}{\partial x} = -[(\underset{\sim}{P}_t' \otimes \underset{\sim}{Q}_t^{-1}) + (\underset{\sim}{Q}_t'^{-1} \otimes \underset{\sim}{P}_t) + (\underset{\sim}{Q}_t'^{-1} \otimes \underset{\sim}{Q}_t^{-1})] \text{ vec } \frac{\partial \underset{\sim}{Q}_t}{\partial x}$$

$$= - \underset{\sim}{M}_t \text{ vec } \frac{\partial \underset{\sim}{Q}_t}{\partial x} \text{ , say .}$$

Now, for $x = \sigma_{k\ell}^{(0)}$ we get

$$\frac{\partial \text{ vec } \underset{\sim}{P}_t}{\partial \sigma_{k\ell}^{(0)}} = -\underset{\sim}{M}_t \text{ vec}(\frac{1}{2\pi} \underset{\sim}{E}_{k\ell}) = - \frac{1}{2\pi} \underset{\sim}{M}_t \underset{\sim}{e}_{k\ell} ,$$

which gives us

$$(4.3.15) \qquad \frac{\partial \text{ vec } \underset{\sim}{P}_t}{\partial \underset{\sim}{\tilde{\theta}}_0'} = - \frac{1}{2\pi} \underset{\sim}{M}_t .$$

Similarly for $x = \sigma_{k\ell}^{(s)}$ we get

$$\frac{\partial \text{ vec } \underset{\sim}{P}_t}{\partial \sigma_{k\ell}^{(s)}} = - \frac{1}{2\pi} \underset{\sim}{M}_t (e^{-i\lambda_t s} \underset{\sim}{I}_{m^2} + e^{i\lambda_t s} \underset{\sim}{K}_m) \underset{\sim}{e}_{k\ell} .$$

Or

$$(4.3.16) \qquad \frac{\partial \text{ vec } \underset{\sim}{P}_t}{\partial \underset{\sim}{\tilde{\theta}}_s'} = - \frac{1}{2\pi} \underset{\sim}{M}_t (e^{-i\lambda_t s} \underset{\sim}{I}_{m^2} + e^{i\lambda_t s} \underset{\sim}{K}_m) \text{ , } s = 1,\dots,q .$$

Putting (4.3.15) and (4.3.16) together we get

$$\frac{\partial \text{ vec } \underset{\sim}{P}_t}{\partial \underset{\sim}{\tilde{\theta}}'} = - \frac{1}{2\pi} \underset{\sim}{M}_t \underset{\sim}{H}_t^* ,$$

which means

$$(4.3.17) \qquad \frac{\partial \operatorname{vec} P_t}{\partial \theta'} = -\frac{1}{2\pi} M_t H_t^* G' .$$

Substituting this in (4.3.13) we get

$$(4.3.18) \qquad \frac{\partial^2 \Lambda}{\partial \theta \, \partial \theta'} = -\frac{1}{8\pi^2} \sum_t G H_t M_t H_t^* G' .$$

This completes the derivation of the second order derivatives of Λ.

4.4. The Newton-Raphson Method.

This method consists of solving the following system of linear equations for $\hat{\rho}_{(1)}$

$$(4.4.1) \qquad -\left.\frac{\partial^2 \Lambda}{\partial \rho \, \partial \rho'}\right|_{\rho=\hat{\rho}_{(0)}} (\hat{\rho}_{(1)} - \hat{\rho}_{(0)}) = \left.\frac{\partial \Lambda}{\partial \rho}\right|_{\rho=\hat{\rho}_{(0)}} .$$

We get the matrix of second order derivatives of Λ, evaluated at $\rho = \hat{\rho}_{(0)}$, from (4.3.3), (4.3.12) and (4.3.18)

$$(4.4.2) \qquad -\left.\frac{\partial^2 \Lambda}{\partial \rho \, \partial \rho'}\right|_{\rho=\hat{\rho}_{(0)}} =$$

$$\begin{pmatrix} \sum_t R_t (I_t' \otimes \hat{Q}_t^{-1}(0)) R_t^* & -\frac{1}{2\pi} \sum_t R_t (I_t' \hat{B}_t'(0) \hat{Q}_t'^{-1}(0) \otimes \hat{Q}_t^{-1}) H_t^* G' \\ -\frac{1}{2\pi} \sum_t G \bar{H}_t (\hat{Q}_t^{-1}(0) \hat{B}_t(0) I_t \otimes \hat{Q}_t'^{-1}(0)) R_t' & \frac{1}{8\pi^2} \sum_t G H_t \hat{M}_t(0) H_t^* G' \end{pmatrix} .$$

We get the first derivative of Λ, evaluated at $\underset{\sim}{\rho} = \hat{\underset{\sim}{\rho}}_{(0)}$, from (4.2.6) and (4.2.17)

$$\left.\frac{\partial\Lambda}{\partial\underset{\sim}{\rho}}\right|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}} = \begin{pmatrix} -\sum_t \underset{\sim}{R}_t \ \mathrm{vec}(\hat{\underset{\sim}{Q}}{}^{-1}_t(0)\hat{\underset{\sim}{B}}_t(0)\underset{\sim}{I}_t) \\ \\ \frac{1}{2\pi}\sum_t \underset{\sim}{G}\underset{\sim}{J}_t \ \mathrm{vec} \ \hat{\underset{\sim}{P}}_t(0) \end{pmatrix} . \tag{4.4.3}$$

We substitute these in (4.4.1) and solve for $\hat{\underset{\sim}{\rho}}_{(1)}$.

4.5. The Scoring Method.

This method consists of solving the following system of linear equations for $\hat{\underset{\sim}{\rho}}_{(1)}$.

$$-\mathcal{E}\left(\frac{\partial^2\Lambda}{\partial\underset{\sim}{\rho}\partial\underset{\sim}{\rho}'}\right)\bigg|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}}(\hat{\underset{\sim}{\rho}}_{(1)}-\hat{\underset{\sim}{\rho}}_{(0)}) = \left.\frac{\partial\Lambda}{\partial\underset{\sim}{\rho}}\right|_{\underset{\sim}{\rho}=\hat{\underset{\sim}{\rho}}_{(0)}} . \tag{4.5.1}$$

Now, we have seen in section (4.3) that $\frac{\partial^2\Lambda}{\partial\underset{\sim}{\rho}\partial\underset{\sim}{\rho}'}$, depends on the observations only through $\underset{\sim}{I}_t$, $t = 0, \ldots, T-1$. Lemma (6) of the Appendix allows us to replace $\underset{\sim}{I}_t$ by $\underset{\sim}{f}_t$ when taking expectations of (4.3.3), (4.3.12) and (4.3.18). So we get

$$\mathcal{E}\left(\frac{\partial^2\Lambda}{\partial\underset{\sim}{\beta}\,\partial\underset{\sim}{\beta}'}\right) = \mathcal{E}\left(-\sum_t \underset{\sim}{R}_t(\underset{\sim}{I}_t' \otimes \underset{\sim}{Q}_t^{-1})\underset{\sim}{R}_t^*\right) \tag{4.5.2}$$

$$\doteq -\sum_t \underset{\sim}{R}_t(\underset{\sim}{f}_t' \otimes \underset{\sim}{Q}_t^{-1})\underset{\sim}{R}_t^* ,$$

$$(4.5.3) \qquad \mathcal{E}\left(\frac{\partial^2 \Lambda}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right) = \mathcal{E}\left(\frac{1}{2\pi} \sum_t \mathbf{R}_t (\mathbf{I}_t' \mathbf{B}_t' \mathbf{Q}_t'^{-1} \otimes \mathbf{Q}_t^{-1}) \mathbf{H}_t^* \mathbf{G}'\right)$$

$$\doteq \frac{1}{2\pi} \sum_t \mathbf{R}_t (\mathbf{f}_t' \mathbf{B}_t' \mathbf{Q}_t'^{-1} \otimes \mathbf{Q}_t^{-1}) \mathbf{H}_t^* \mathbf{G}'$$

$$= \frac{1}{2\pi} \sum_t \mathbf{R}_t (\overline{\mathbf{B}}_t^{-1} \otimes \mathbf{Q}_t^{-1}) \mathbf{H}_t^* \mathbf{G}' \ .$$

Finally

$$\mathcal{E}\left(\frac{\partial^2 \Lambda}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right) = -\frac{1}{8\pi^2} \mathcal{E}\left(\sum_t \mathbf{G} \mathbf{H}_t \mathbf{M}_t \mathbf{H}_t^* \mathbf{G}'\right) .$$

Now

$$\mathbf{M}_t = \mathbf{P}_t' \otimes \mathbf{Q}_t^{-1} + \mathbf{Q}_t'^{-1} \otimes \mathbf{P}_t + \mathbf{Q}_t'^{-1} \otimes \mathbf{Q}_t^{-1} ,$$

where

$$\mathbf{P}_t = \mathbf{Q}_t^{-1} \mathbf{B}_t \mathbf{I}_t \mathbf{B}_t^* \mathbf{Q}_t^{-1} - \mathbf{Q}_t^{-1} \ .$$

So

$$\mathcal{E}\ \mathbf{P}_t \doteq \mathbf{Q}_t^{-1} \mathbf{B}_t \mathbf{f}_t \mathbf{B}_t^* \mathbf{Q}_t^{-1} - \mathbf{Q}_t^{-1} = 0 ,$$

which means

$$\mathcal{E}\ \mathbf{M}_t \doteq \mathbf{Q}_t'^{-1} \otimes \mathbf{Q}_t^{-1} \ .$$

Hence

$$(4.5.4) \qquad \mathcal{E}\left(\frac{\partial^2 \Lambda}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right) \doteq -\frac{1}{8\pi^2} \sum_t \mathbf{G} \mathbf{H}_t (\mathbf{Q}_t'^{-1} \otimes \mathbf{Q}_t^{-1}) \mathbf{H}_t^* \mathbf{G}' \ .$$

So we evaluate these expressions at $\boldsymbol{\rho} = \hat{\boldsymbol{\rho}}_{(0)}$ and then substitute them in (4.5.1) and solve the resulting system of linear equations for $\hat{\boldsymbol{\rho}}_{(1)}$.

4.6. The Pure Moving Average Case.

In this special case we have

$$\underset{\sim}{B}_t \equiv \underset{\sim}{I}_m , \quad \underset{\sim}{f}_t \equiv \underset{\sim}{Q}_t ,$$

which means

$$\underset{\sim}{g}_t = \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t \underset{\sim}{f}_t^{-1} ,$$

$$\underset{\sim}{P}_t = \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t \underset{\sim}{f}_t^{-1} - \underset{\sim}{f}_t^{-1} .$$

Using these and (4.3.18) we can get the Newton-Raphson equations

$$-\frac{1}{8\pi^2} [\sum_t \underset{\sim}{G} \underset{\sim}{H}_t \hat{\underset{\sim}{M}}_t(0) \underset{\sim}{H}_t^* \underset{\sim}{G}'] (\hat{\underset{\sim}{\theta}}(1) - \hat{\underset{\sim}{\theta}}(0)) = \frac{1}{2\pi} \sum_t \underset{\sim}{G} \underset{\sim}{J}_t \operatorname{vec} \hat{\underset{\sim}{P}}_t(0) .$$

This is identical to the equation derived by Ahrabi (1978). Similarly the equation for the scoring method can be derived using (4.5.4)

$$\frac{1}{8\pi^2} [\sum_t \underset{\sim}{G} \underset{\sim}{H}_t (\hat{\underset{\sim}{f}}{}'^{-1}_t(0) \otimes \hat{\underset{\sim}{f}}{}^{-1}_t(0)) \underset{\sim}{H}_t^* \underset{\sim}{G}'] (\hat{\underset{\sim}{\theta}}(1) - \hat{\underset{\sim}{\theta}}(0)) = \frac{1}{2\pi} \sum_t \underset{\sim}{G} \breve{\underset{\sim}{J}}_t \operatorname{vec} \hat{\underset{\sim}{P}}_t(0) ,$$

which is identical to the equation derived by Ahrabi (1978).

5. Asymptotic Properties.

The four estimates proposed in the preceding chapters are asymptotically equivalent and we shall show that they are asymptotically efficient, i.e.,

$$\sqrt{T}\,(\hat{\rho}_{(1)} - \rho) \xrightarrow{d} N(0, \vartheta^{-1}(\rho)) ,$$

where $\vartheta(\rho)$ is the limiting average information matrix and "$\xrightarrow{d}$" indicates convergence in distribution.

To find $\vartheta(\rho)$, by definition we have

$$\vartheta(\rho) = \lim_{T \to \infty} -\frac{1}{T}\, \mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \rho \partial \rho'}\right)$$

$$= \lim_{T \to \infty} -\frac{1}{T}\, \mathcal{E}\left(\frac{\partial^2 \Delta}{\partial \rho \partial \rho'}\right) .$$

Now let

(5.1)
$$\vartheta(\rho) = \begin{pmatrix} \varphi & \Omega' \\ \Omega & \psi \end{pmatrix} .$$

Then from (4.5.2), (4.5.3) and (4.5.4) we have

(5.2)
$$\varphi = \lim_{T \to \infty} \frac{1}{T} \sum_t R_t (f_t' \otimes Q_t^{-1}) R_t^*$$

$$= \frac{1}{2\pi} \int_0^{2\pi} R (f' \otimes Q^{-1}) R^* d\lambda ,$$

(5.3) $$\underset{\sim}{\Omega} = \lim_{T \to \infty} -\frac{1}{T}\left(\frac{1}{2\pi} \sum_t \underset{\sim}{\mathfrak{R}}_t (\underset{\sim}{\bar{B}}_t^{-1} \otimes \underset{\sim}{Q}_t^{-1}) \underset{\sim}{H}_t^* \underset{\sim}{G}\right)$$

$$= -\frac{1}{(2\pi)^2} \int_0^{2\pi} \underset{\sim}{\mathfrak{R}} (\underset{\sim}{\bar{B}}^{-1} \otimes \underset{\sim}{Q}^{-1}) \underset{\sim}{H}^* \underset{\sim}{G}' \, d\lambda \, .$$

(5.4) $$\underset{\sim}{\Psi} = \lim_{T \to \infty} \frac{1}{T}\left(\frac{1}{8\pi^2} \sum_t \underset{\sim}{G} \underset{\sim}{H}_t (\underset{\sim}{Q}_t'^{-1} \otimes \underset{\sim}{Q}_t^{-1}) \underset{\sim}{H}_t^* \underset{\sim}{G}'\right)$$

$$= \frac{1}{16\pi^3} \int_0^{2\pi} \underset{\sim}{G} \underset{\sim}{H} (\underset{\sim}{Q}'^{-1} \otimes \underset{\sim}{Q}^{-1}) \underset{\sim}{H}^* \underset{\sim}{G}' \, d\lambda \, .$$

The four estimates are obtained from equations like

(5.5) $$\hat{\underset{\sim}{\vartheta}}_{(0)} (\hat{\underset{\sim}{\rho}}_{(1)} - \hat{\underset{\sim}{\rho}}_{(0)}) = \frac{1}{T} \left. \frac{\partial \log \ell}{\partial \underset{\sim}{\rho}} \right|_{\underset{\sim}{\rho} = \hat{\underset{\sim}{\rho}}_{(0)}} ,$$

where $\hat{\underset{\sim}{\vartheta}}_{(0)}$ is an initial estimate of $\underset{\sim}{\vartheta}(\underset{\sim}{\rho})$ and is consistent. We can write (5.5) as

(5.6) $$\hat{\underset{\sim}{\vartheta}}_{(0)} \sqrt{T} \, (\hat{\underset{\sim}{\rho}}_{(1)} - \underset{\sim}{\rho}) = \hat{\underset{\sim}{\vartheta}}_{(0)} \sqrt{T} \, (\hat{\underset{\sim}{\rho}}_{(0)} - \underset{\sim}{\rho}) + \frac{1}{\sqrt{T}} \left. \frac{\partial \log \ell}{\partial \underset{\sim}{\rho}} \right|_{\underset{\sim}{\rho} = \hat{\underset{\sim}{\rho}}_{(0)}} ,$$

where $\underset{\sim}{\rho}$ is the true parameter value. Now

(5.7) $$\frac{1}{\sqrt{T}} \frac{\partial \log \ell}{\partial \underset{\sim}{\rho}} = \frac{1}{\sqrt{T}} \left. \frac{\partial \log \ell}{\partial \underset{\sim}{\rho}} \right|_{\underset{\sim}{\rho} = \hat{\underset{\sim}{\rho}}_{(0)}} + \frac{1}{\sqrt{T}} \left. \frac{\partial^2 \log \ell}{\partial \underset{\sim}{\rho} \, \partial \underset{\sim}{\rho}'} \right|_{\underset{\sim}{\rho} = \underset{\sim}{\rho}^+} (\underset{\sim}{\rho} - \hat{\underset{\sim}{\rho}}_{(0)}) ,$$

where $|\underset{\sim}{\rho} - \underset{\sim}{\rho}^+| \leq |\underset{\sim}{\rho} - \hat{\underset{\sim}{\rho}}_{(0)}|$. Now (5.6) can be rewritten using (5.7)

(5.8) $$\hat{\mathfrak{g}}_{(0)}\sqrt{T}\,(\hat{\rho}_{(1)}-\rho) = [\hat{\mathfrak{g}}_{(0)} + \frac{1}{T}\frac{\partial^2 \log \ell}{\partial \rho\,\partial \rho'}\Big|_{\rho=\rho^+}]$$

$$\times \sqrt{T}\,(\hat{\rho}_{(0)}-\rho) + \frac{1}{\sqrt{T}}\;\frac{\partial \log \ell}{\partial \rho} .$$

Now noticing that

$$\frac{1}{T}\;\frac{\partial^2 \log \ell}{\partial \rho\,\partial \rho'}\Big|_{\rho=\rho^+} \xrightarrow{\mathcal{L}} -\,\mathfrak{g}(\rho)$$

and that $\sqrt{T}\,(\hat{\rho}_{(0)}-\rho)$ is bounded in probability, we see that (5.8) is (asymptotically) equivalent to

(5.9) $$\sqrt{T}\,(\hat{\rho}_{(1)}-\rho) = \mathfrak{g}^{-1}(\rho)\;\frac{1}{\sqrt{T}}\;\frac{\partial \log \ell}{\partial \rho} .$$

Theorem. If in addition to Assumptions 1-5 of Chapter 2 we assume that the ϵ_t's have finite fourth order moments, then

(5.10) $$\sqrt{T}\,(\hat{\rho}_{(1)}-\rho) \xrightarrow{d} N(0,\mathfrak{g}^{-1}(\rho)) ,$$

where $\hat{\rho}_{(1)}$ is any one of the four estimates proposed in the previous chapters.

Proof. Using (5.9) it suffices to show that

(5.11) $$\frac{1}{\sqrt{T}}\;\frac{\partial \log \ell}{\partial \rho} \xrightarrow{d} N(0,\mathfrak{g}(\rho)) .$$

Let

(5.12)
$$\underset{\sim}{\xi} = \begin{pmatrix} \underset{\sim}{\beta} \\ \mathrm{vec}(\underset{\sim}{A}_1,\ldots,\underset{\sim}{A}_q) \\ \mathrm{dg}\ \underset{\sim}{V} \\ \widetilde{\mathrm{vec}}\ \underset{\sim}{V} \end{pmatrix},$$

where $\underset{\sim}{A}_i$'s and $\underset{\sim}{V}$ were introduced in Chapter 2. Now

$$\frac{\partial \log \ell}{\partial \xi_i} = \sum_j \frac{\partial \log \ell}{\partial \rho_j} \cdot \frac{\partial \rho_j}{\partial \xi_i},$$

which means

(5.13)
$$\frac{\partial \log \ell}{\partial \underset{\sim}{\xi}} = \frac{\partial \underset{\sim}{\rho}'}{\partial \underset{\sim}{\xi}} \cdot \frac{\partial \log \ell}{\partial \underset{\sim}{\rho}}.$$

It follows from Assumptions 3-5 of Chapter 2 that $\frac{\partial \underset{\sim}{\rho}'}{\partial \underset{\sim}{\xi}}$ is nonsingular, which means

(5.14)
$$\frac{1}{\sqrt{T}} \frac{\partial \log \ell}{\partial \underset{\sim}{\rho}} = \left(\frac{\partial \underset{\sim}{\rho}'}{\partial \underset{\sim}{\xi}}\right)^{-1} \frac{1}{\sqrt{T}} \frac{\partial \log \ell}{\partial \underset{\sim}{\xi}}.$$

Now, it has been shown by Nicholls (1976) and Reinsel (1976) that

(5.15)
$$\sqrt{T}\,(\hat{\underset{\sim}{\xi}}_{(1)} - \hat{\underset{\sim}{\xi}}) \xrightarrow{\mathcal{L}} N(\underset{\sim}{0},\cdot).$$

where $\hat{\underset{\sim}{\xi}}_{(1)}$ is the estimate obtained by solving equations of the form

(5.16)
$$\hat{\underset{\sim}{\mathfrak{J}}}(\hat{\underset{\sim}{\xi}}_{(0)})(\hat{\underset{\sim}{\xi}}_{(1)} - \hat{\underset{\sim}{\xi}}_{(0)}) = \frac{1}{T} \left.\frac{\partial \log \ell}{\partial \underset{\sim}{\xi}}\right|_{\underset{\sim}{\xi} = \hat{\underset{\sim}{\xi}}_{(0)}},$$

where $\hat{\underset{\sim}{\xi}}_{(0)}$ is an initial estimate of $\underset{\sim}{\xi}$ and $\hat{\underset{\sim}{\vartheta}}(\hat{\underset{\sim}{\xi}}_{(0)})$ an initial estimate of the limiting average information matrix of $\underset{\sim}{\xi}$. Now applying the same argument as we did for $\underset{\sim}{\rho}$ we see that (5.16) is equivalent to

$$(5.17) \qquad \sqrt{T}\,(\hat{\underset{\sim}{\xi}}_{(1)}-\underset{\sim}{\xi}) = \underset{\sim}{\vartheta}^{-1}(\underset{\sim}{\xi})\,\frac{1}{\sqrt{T}}\,\frac{\partial \log \ell}{\partial \underset{\sim}{\xi}} .$$

Now (5.15) and (5.17) imply

$$(5.18) \qquad \frac{1}{\sqrt{T}}\,\frac{\partial \log \ell}{\partial \underset{\sim}{\xi}} \xrightarrow{\mathcal{L}} N(\underset{\sim}{0},\cdot) .$$

Finally from (5.14), using (5.18), we get

$$(5.19) \qquad \frac{1}{\sqrt{T}}\,\frac{\partial \log \ell}{\partial \underset{\sim}{\rho}} \to N(\underset{\sim}{0},\underset{\sim}{\vartheta}(\underset{\sim}{\rho})) .$$

The desired result is obtained from (5.9), which together with (5.19) gives us

$$(5.20) \qquad \sqrt{T}\,(\hat{\underset{\sim}{\rho}}_{(1)}-\underset{\sim}{\rho}) \to N(\underset{\sim}{0},\underset{\sim}{\vartheta}^{-1}(\underset{\sim}{\rho})) .$$

Q.E.D.

Note. The approximation to the log likelihood used by Nicholls (1976) is not identical to ours, i.e., (4.1.9). But as we shall demonstrate in the Appendix it is asymptotically equivalent to it. This means that if Nicholls had used (4.1.9) he would still have obtained asymptotically efficient estimates, as we assumed in the proof above.

6. Estimation of the Coefficients and the Covariance Matrix of the $\underset{\sim}{\epsilon}_t$'s in the Time Domain (The Scoring Method).

For the model defined by (2.1), Reinsel(1976) gives equations for the estimates of $\underset{\sim}{A}_1,\dots,\underset{\sim}{A}_q$, $\underset{\sim}{B}_1,\dots,\underset{\sim}{B}_p$, $\underset{\sim}{V}$, using Newton-Raphson method on the (modified) log likelihood of the data. In this chapter we shall use the techniques developed in the preceding chapters to arrive at the equations for the estimatesof these parameters using the scoring method.

The Likelihood Function.

Assuming that $\underset{\sim}{\epsilon}_0 = \underset{\sim}{\epsilon}_{-1} = \cdots = \underset{\sim}{\epsilon}_{1-q} = \underset{\sim}{0}$, and using the same method as in section 3.1 we have

$$\underset{\sim}{\mathscr{B}}\underset{\sim}{y} = \underset{\sim}{\mathscr{A}}\underset{\sim}{e}\ , \tag{6.1}$$

where

$$\underset{\sim}{\mathscr{A}} = \sum_{i=0}^{q}(\underset{\sim}{A}_i \otimes \underset{\sim}{L}^i), \tag{6.2}$$

$$\underset{\sim}{e} = \text{vec}(\underset{\sim}{\epsilon}_1,\dots,\underset{\sim}{\epsilon}_T)'\ , \tag{6.3}$$

and $\underset{\sim}{\mathscr{B}}$ and $\underset{\sim}{y}$ were introduced in section 3.1. Now, to derive the likelihood function we need the covariance matrix of $\underset{\sim}{e}$. Using lemma 3.2 we have

$$\underset{\sim}{e} = \underset{\sim}{K}_{m,T}\ \text{vec}(\underset{\sim}{\epsilon}_1,\dots,\underset{\sim}{\epsilon}_T) = \underset{\sim}{K}_{m,T}\begin{pmatrix}\underset{\sim}{\epsilon}_1\\ \vdots \\ \underset{\sim}{\epsilon}_T\end{pmatrix}. \tag{6.4}$$

But

$$\begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_T \end{pmatrix} \sim N(0, I_T \otimes V) .$$

So

$$(6.5) \qquad \mathcal{E}(e e') = K_{m,T}(I_T \otimes V)K'_{m,T} = V \otimes I_T = \mathcal{J} ,$$

say, where we have used lemma 3.3. Using (6.1) and (6.5) we have

$$y \sim N(0, \mathcal{B}^{-1} \mathcal{Q} \mathcal{J} \mathcal{Q}' \mathcal{B}'^{-1}) .$$

So finally the modified log likelihood is

$$(6.6) \qquad \log \ell(y) = -\frac{Tm}{2} \log 2\pi - \frac{1}{2} \log|\mathcal{B}^{-1}\mathcal{Q}\mathcal{J}\mathcal{Q}'\mathcal{B}'^{-1}|$$

$$- \frac{1}{2} y'\mathcal{B}'\mathcal{Q}'^{-1}\mathcal{J}^{-1}\mathcal{Q}^{-1}\mathcal{B}y .$$

We can further simplify this by noticing that as in (3.5.2)

$$|\mathcal{B}| = |\mathcal{Q}| = 1 ,$$

and

$$|\mathcal{J}| = |V \otimes I_T| = |V|^T .$$

So (6.6) becomes

$$\log \ell(\underset{\sim}{y}) = -\frac{Tm}{2}\log 2\pi - \frac{T}{2}\log|\underset{\sim}{V}| - \frac{1}{2}\underset{\sim}{y}'\underset{\sim}{\mathcal{B}}'\underset{\sim}{\mathcal{A}}'^{-1}\mathcal{V}^{-1}\underset{\sim}{\mathcal{A}}'^{-1}\underset{\sim}{\mathcal{B}}\underset{\sim}{y}\,. \tag{6.7}$$

The Estimation Method.

The parameters to be estimated are

$$\underset{\sim}{\alpha}_i = \text{vec } \underset{\sim}{A}_i \;,\; i = 1,\ldots,q \;,$$

$$\underset{\sim}{\beta}_j = \text{vec } \underset{\sim}{B}_j \;,\; j = 1,\ldots,p \;,$$

$$\underset{\sim}{V}\;.$$

We let

$$\underset{\sim}{\eta} = \begin{pmatrix}\underset{\sim}{\alpha}\\ \underset{\sim}{\beta}\end{pmatrix} \;, \tag{6.8}$$

where $\underset{\sim}{\alpha}' = (\underset{\sim}{\alpha}_1', \ldots, \underset{\sim}{\alpha}_q')$, $\underset{\sim}{\beta}' = (\underset{\sim}{\beta}_1', \ldots, \underset{\sim}{\beta}_p')$. We are going to apply the scoring method to arrive at approximate maximum likelihood estimates that are asymptotically efficient. It is well known (See Dunsmuir and Hannan (1976).) that

$$\lim_{T\to\infty} \frac{1}{T}\,\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\eta}\,\partial v_{rs}} = 0\;. \tag{6.9}$$

This means that the limiting average information matrix is block diagonal. So we can write separate equations for estimates of $\underset{\sim}{\eta}$ and $\underset{\sim}{V}$ without violating asymptotic efficiency of the estimates. We shall first derive the quations for $\hat{\eta}_{(1)}$, the estimate of $\underset{\sim}{\eta}$ by the scoring method. These equations are

$$(6.10) \qquad -\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \boldsymbol{\eta}\, \partial \boldsymbol{\eta}'}\right)\Bigg|_{\substack{\boldsymbol{\eta}=\hat{\boldsymbol{\eta}}_{(0)} \\ \mathbf{V}=\hat{\mathbf{V}}_{(0)}}} (\hat{\boldsymbol{\eta}}_{(1)}-\hat{\boldsymbol{\eta}}_{(0)}) = \frac{\partial \log \ell}{\partial \boldsymbol{\eta}}\Bigg|_{\substack{\boldsymbol{\eta}=\hat{\boldsymbol{\eta}}_{(0)} \\ \mathbf{V}=\hat{\mathbf{V}}_{(0)}}} ,$$

where $\hat{\boldsymbol{\eta}}_{(0)}$ and $\hat{\mathbf{V}}_{(0)}$ are initial estimates of $\boldsymbol{\eta}$ and $\mathbf{V}$ that are consistent of order $T^{-1/2}$, as given by Reinsel (1976). We proceed to find the first and second order derivatives of $\log \ell$ with respect to $\boldsymbol{\eta}$.

<u>The First Order Derivatives.</u>

<u>Derivation of</u> $\dfrac{\partial \log \ell}{\partial \boldsymbol{\alpha}}$.

Differentiating (6.7) with respect to $a_{ij}^{(h)}$ we get

$$\frac{\partial \log \ell}{\partial a_{ij}^{(h)}} = \mathbf{y}'\mathfrak{B}'\mathfrak{A}'^{-1}\mathfrak{S}^{-1}\mathfrak{A}^{-1}(\mathbf{E}_{ij} \otimes \mathbf{L}^h)\mathfrak{A}^{-1}\mathfrak{B}\mathbf{y}$$

$$= (\mathbf{y}'\mathfrak{B}'\mathfrak{A}'^{-1} \otimes \mathbf{y}'\mathfrak{B}'\mathfrak{A}'^{-1}\mathfrak{S}^{-1}\mathfrak{A}^{-1})\ \mathrm{vec}(\mathbf{E}_{ij} \otimes \mathbf{L}^h) ,$$

using lemma 3.1. Using the same method as used in deriving (3.2.5) we get

$$(6.11) \qquad \frac{\partial \log \ell}{\partial \boldsymbol{\alpha}} = \tilde{\mathbf{E}}'(\mathfrak{A}^{-1}\mathfrak{B}\mathbf{y} \otimes \mathfrak{A}'^{-1}\mathfrak{S}^{-1}\mathfrak{A}^{-1}\mathfrak{B}\mathbf{y}) ,$$

where

(6.12) $$\tilde{E} = (E_1, \ldots, E_q) ,$$

and E_r was defined in (3.2.4).

<u>Derivation of</u> $\frac{\partial \log \ell}{\partial \beta}$.

This was derived in Chapter 3 and is given by (3.2.7), which states

(6.13) $$\frac{\partial \log \ell}{\partial \beta} = -E'(y \otimes \Gamma^{-1} B y) .$$

Now we need to express Γ^{-1} in terms of the parameters in this chapter. We recall that

(6.14) $$\Gamma = \mathcal{E}(u u') ,$$

where

$$u = B y ,$$

as given by (3.1.4). Now using (6.1), (6.4), (6.14) we have

$$\Gamma = \mathcal{E}(C e e' C') = C \Delta C' .$$

Substituting this in (6.13) we get

(6.15) $$\frac{\partial \log \ell}{\partial \beta} = -E'(y \otimes C'^{-1} \Delta^{-1} C^{-1} B y) .$$

The Second Order Derivatives.

Derivation of $\mathcal{E}\,\frac{\partial^2 \log \ell}{\partial \alpha\, \partial \alpha'}$.

Differentiating (6.11) with respect to $a_{ij}^{(h)}$ we get

$$
(6.16) \qquad \frac{\partial^2 \log \ell}{\partial \alpha\, \partial a_{ij}^{(h)}} = -\tilde{E}'[\mathcal{A}^{-1}(E_{ij} \otimes L^h)\mathcal{A}^{-1}\mathcal{B}y \otimes \mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1}\mathcal{B}y
$$

$$
+ \mathcal{A}^{-1}\mathcal{B}y \otimes \mathcal{A}'^{-1}(E'_{ij} \otimes L'^h)\mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1}\mathcal{B}y + \mathcal{A}^{-1}\mathcal{B}y \otimes \mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1}
$$

$$
(E_{ij} \otimes L^h)\mathcal{A}^{-1}\mathcal{B}y] \;.
$$

Now, using lemmas 3.4, (6.16) can be rewritten as

$$
(6.17) \qquad \frac{\partial^2 \log \ell}{\partial \alpha\, \partial a_{ij}^{(h)}} = -\tilde{E}' \operatorname{vec}[\mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1}\mathcal{B}yy'\mathcal{B}'\mathcal{A}'^{-1}(E'_{ij} \otimes L'^h)\mathcal{A}'^{-1}
$$

$$
+ \mathcal{A}'^{-1}(E'_{ij} \otimes L'^h)\mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1}\mathcal{B}yy'\mathcal{B}'\mathcal{A}'^{-1} + \mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1}
$$

$$
(E_{ij} \otimes L^h)\mathcal{A}^{-1}\mathcal{B}yy'\mathcal{B}'\mathcal{A}'^{-1}].
$$

Taking expectations of both sides of (6.17) and using (6.5) we have

$$
(6.18) \qquad \mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \alpha\, \partial a_{ij}^{(h)}}\right) = -2\tilde{E}' \operatorname{vec}[\mathcal{A}'^{-1}(E'_{ij} \otimes L'^h)\mathcal{A}'^{-1}
$$

$$
-\tilde{E}' \operatorname{vec}[\mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1}(E_{ij} \otimes L^h)\mathcal{S}] \;.
$$

As we shall now show in the Appendix (lemma 3) the first term on the right hand side of (6.18) is equal to zero. Using this and lemma 3.1, (6.18) can be rewritten as

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\alpha}\, \partial a_{ij}^{(h)}}\right) = -\underset{\sim}{\tilde{E}}'\,(\boldsymbol{\mathcal{S}} \otimes \boldsymbol{\mathcal{A}}'^{-1}\boldsymbol{\mathcal{S}}^{-1}\boldsymbol{\mathcal{A}}^{-1})\ \mathrm{vec}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^{h}) .$$

From this we finally get

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\alpha}\, \partial \underset{\sim}{\alpha}'}\right) = -\underset{\sim}{\tilde{E}}'\,(\boldsymbol{\mathcal{S}} \otimes \boldsymbol{\mathcal{A}}'^{-1}\boldsymbol{\mathcal{S}}^{-1}\boldsymbol{\mathcal{A}}^{-1})\underset{\sim}{\tilde{E}} . \tag{6.19}$$

<u>Derivation of</u> $\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\alpha}\, \partial \underset{\sim}{\beta}'}\right)$.

Differentiating (6.11) with respect to $b_{ij}^{(h)}$ we get

$$\begin{aligned}\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\alpha}\, \partial b_{ij}^{(h)}} &= \underset{\sim}{\tilde{E}}'\,[\boldsymbol{\mathcal{A}}^{-1}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^{h})\underset{\sim}{y} \otimes \boldsymbol{\mathcal{A}}'^{-1}\boldsymbol{\mathcal{S}}^{-1}\boldsymbol{\mathcal{A}}^{-1}\boldsymbol{\mathcal{B}}\underset{\sim}{y} \\ &\quad + \boldsymbol{\mathcal{A}}^{-1}\boldsymbol{\mathcal{B}}\underset{\sim}{y} \otimes \boldsymbol{\mathcal{A}}'^{-1}\boldsymbol{\mathcal{S}}^{-1}\boldsymbol{\mathcal{A}}^{-1}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^{h})\underset{\sim}{y}] \\ &= \underset{\sim}{\tilde{E}}'\ \mathrm{vec}[\boldsymbol{\mathcal{A}}'^{-1}\boldsymbol{\mathcal{S}}^{-1}\boldsymbol{\mathcal{A}}^{-1}\boldsymbol{\mathcal{B}}\underset{\sim}{y}\,\underset{\sim}{y}'\,(\underset{\sim}{E}_{ij}' \otimes \underset{\sim}{L}'^{h})\boldsymbol{\mathcal{A}}'^{-1} \\ &\quad + \boldsymbol{\mathcal{A}}'^{-1}\boldsymbol{\mathcal{S}}^{-1}\boldsymbol{\mathcal{A}}^{-1}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^{h})\underset{\sim}{y}\,\underset{\sim}{y}'\boldsymbol{\mathcal{B}}'\boldsymbol{\mathcal{A}}'^{-1}] ,\end{aligned} \tag{6.20}$$

using lemma 3.4. Taking expectations we have

$$\begin{aligned}\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\alpha}\, \partial b_{ij}^{(h)}}\right) &= \underset{\sim}{\tilde{E}}'\ \mathrm{vec}[\boldsymbol{\mathcal{B}}'^{-1}(\underset{\sim}{E}_{ij}' \otimes \underset{\sim}{L}'^{h})\boldsymbol{\mathcal{A}}'^{-1}] \\ &\quad + \underset{\sim}{\tilde{E}}'\ \mathrm{vec}[\boldsymbol{\mathcal{A}}'^{-1}\boldsymbol{\mathcal{S}}^{-1}\boldsymbol{\mathcal{A}}^{-1}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}^{h})\boldsymbol{\mathcal{B}}^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{S}}] .\end{aligned} \tag{6.21}$$

The first term on the right hand side of (6.21) is equal to zero. This follows from lemma 2 of the Appendix. Now using lemma 3.1, (6.21) can be rewritten as

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \alpha\, \partial b_{ij}^{(h)}}\right) = \tilde{E}'(\mathcal{S}'\mathcal{A}'\mathcal{B}'^{-1} \otimes \mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1})\ \mathrm{vec}(E_{ij} \otimes L^h) .$$

From this we finally get

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \alpha\, \partial \beta'}\right) = \tilde{E}'(\mathcal{S}'\mathcal{A}'\mathcal{B}'^{-1} \otimes \mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1})E \quad . \tag{6.22}$$

<u>Derivation of</u> $\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \beta\, \partial \beta'}\right)$.

This has already been derived in (3.3.14) and (3.3.15) which together yield

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \beta\, \partial \beta'}\right) = -E'(\mathcal{B}^{-1}\Gamma\mathcal{B}'^{-1} \otimes \Gamma^{-1})E \quad .$$

But

$$\Gamma = \mathcal{A}\,\mathcal{S}\,\mathcal{A}' \quad .$$

So

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \beta\, \partial \beta'}\right) = -E'(\mathcal{B}^{-1}\mathcal{A}\mathcal{S}\mathcal{A}'\mathcal{B}'^{-1} \otimes \mathcal{A}'^{-1}\mathcal{S}^{-1}\mathcal{A}^{-1})E \quad . \tag{6.23}$$

Putting (6.19), (6.22), and (6.23) together we get

$$(6.24) \qquad -\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \boldsymbol{\eta}\, \partial \boldsymbol{\eta}'}\right) = \begin{pmatrix} \tilde{\mathbf{E}}' \\ \mathbf{E}' \end{pmatrix} \left[\begin{pmatrix} \boldsymbol{\mathcal{S}} & \boldsymbol{\mathcal{S}}'\boldsymbol{\mathcal{A}}'\boldsymbol{\mathcal{B}}'^{-1} \\ \boldsymbol{\mathcal{B}}^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{S}} & -\boldsymbol{\mathcal{B}}'^{-1}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{S}}\boldsymbol{\mathcal{A}}'\boldsymbol{\mathcal{B}}'^{-1} \end{pmatrix} \otimes \boldsymbol{\mathcal{A}}'^{-1}\boldsymbol{\mathcal{S}}^{-1}\boldsymbol{\mathcal{A}}^{-1} \right]$$

$$(\tilde{\mathbf{E}}, \mathbf{E}) = \mathbf{W}, \text{ say} .$$

<u>The Scoring Method Equation for</u> $\boldsymbol{\eta}_{(1)}$.

Substituting the expressions derived for the first and second order derivatives of $\log \ell$ in (6.10), we get the desired equation, which is

$$(6.25) \qquad \hat{\mathbf{W}}_{(0)}(\hat{\boldsymbol{\eta}}_{(1)} - \hat{\boldsymbol{\eta}}_{(0)}) = \hat{\mathbf{w}}_{(0)} ,$$

where

$$(6.26) \qquad \hat{\mathbf{w}}_{(0)} = \left.\frac{\partial \log \ell}{\partial \boldsymbol{\eta}}\right|_{\substack{\boldsymbol{\eta} = \hat{\boldsymbol{\eta}}_{(0)} \\ \mathbf{V} = \hat{\mathbf{V}}_{(0)}}} ,$$

and $\dfrac{\partial \log \ell}{\partial \boldsymbol{\eta}}$ is given by (6.11) and (6.13).

<u>Estimation of</u> $\mathbf{V}$.

Once we have $\hat{\boldsymbol{\eta}}_{(1)}$, we can replace $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ with $\hat{\boldsymbol{\mathcal{A}}}_{(1)}$, $\hat{\boldsymbol{\mathcal{B}}}_{(1)}$ in (6.7) and maximize the resulting function, which we denote by $\log \tilde{\ell}$. So we will maximize

$$(6.27) \qquad \log \tilde{\ell} = -\frac{Tm}{2} \log 2\pi - \frac{T}{2} \log|\mathbf{V}| - \frac{1}{2} \mathbf{y}'\hat{\boldsymbol{\mathcal{B}}}'_{(1)}\hat{\boldsymbol{\mathcal{A}}}'^{-1}_{(1)}\boldsymbol{\mathcal{S}}^{-1}\hat{\boldsymbol{\mathcal{A}}}^{-1}_{(1)}\hat{\boldsymbol{\mathcal{B}}}_{(1)}\mathbf{y} .$$

Now the last term can be rewritten as

$$(6.28) \qquad -\frac{1}{2}\, \underset{\sim}{y}'\hat{\underset{\sim}{\mathcal{B}}}'_{(1)}\hat{\underset{\sim}{\mathcal{Q}}}'^{-1}_{(1)}\underset{\sim}{K}_{m,T}(\underset{\sim}{I}_T \otimes \underset{\sim}{V}^{-1})\underset{\sim}{K}'_{m,T}\hat{\underset{\sim}{\mathcal{Q}}}^{-1}_{(1)}\hat{\underset{\sim}{\mathcal{B}}}_{(1)}\underset{\sim}{y} = -\frac{1}{2}\sum_{t=1}^{T}\hat{\underset{\sim}{\epsilon}}'_{t(1)}\underset{\sim}{V}^{-1}\hat{\underset{\sim}{\epsilon}}_{t(1)} \, .$$

This follows from (6.1) and (6.4). Now it is well known that the value of $\underset{\sim}{V}$ which maximizes (6.27) is given by

$$(6.29) \qquad \hat{\underset{\sim}{V}}_{(1)} = \frac{1}{T}\sum_{t=1}^{T}\hat{\underset{\sim}{\epsilon}}_{t(1)}\hat{\underset{\sim}{\epsilon}}'_{t(1)} \, .$$

See Anderson (1958), Chapter 3. We can express $\hat{\underset{\sim}{V}}_{(1)}$ in terms of $\hat{\underset{\sim}{\mathcal{Q}}}_{(1)}$, $\hat{\underset{\sim}{\mathcal{B}}}_{(1)}$, $\underset{\sim}{y}$ using the same argument as in (6.28). So

$$(6.30) \qquad \hat{\underset{\sim}{V}}_{(1)} = \frac{1}{T}\underset{\sim}{K}'_{m,T}\hat{\underset{\sim}{\mathcal{Q}}}^{-1}_{(1)}\hat{\underset{\sim}{\mathcal{B}}}_{(1)}\underset{\sim}{y}\underset{\sim}{y}'\,\hat{\underset{\sim}{\mathcal{B}}}'_{(1)}\hat{\underset{\sim}{\mathcal{Q}}}'^{-1}_{(1)}\underset{\sim}{K}_{m,T} \, .$$

We could theoretically carry out further iterations, but this would be computationaly costly. The estimates given above are asymptotically equivalent to the estimates derived via Newton-Raphson method and hence are asymptotically efficient as demonstrated by Reinsel (1976).

Appendix.

We shall now derive some of the results that we have used in the previous chapters.

The Time Domain.

Lemma 1. (i) For any two matrices $\underset{r\times s}{\underset{\sim}{A}}$ and $\underset{s\times r}{\underset{\sim}{B}}$

$$\mathrm{tr}(\underset{\sim}{A}\underset{\sim}{B}) = (\mathrm{vec}\ \underset{\sim}{A})'\ \mathrm{vec}\ \underset{\sim}{B}' . \tag{1}$$

(ii) For square matrices $\underset{\sim}{A}$ and $\underset{\sim}{B}$

$$\mathrm{tr}(\underset{\sim}{A}\otimes \underset{\sim}{B}) = \mathrm{tr}(\underset{\sim}{A})\cdot \mathrm{tr}(\underset{\sim}{B}) . \tag{2}$$

Proof. (1) and (2) are easily verified by writing out the two sides.

Lemma 2. For $\underset{\sim}{\mathfrak{A}}$, $\underset{\sim}{\mathfrak{B}}$, $\underset{\sim}{E}_{uv}$, $\underset{\sim}{E}_{rs}$, and $\underset{\sim}{L}$ as defined in the time domain

$$\mathrm{tr}[\underset{\sim}{\mathfrak{A}}^{-1}(\underset{\sim}{E}_{uv}\otimes \underset{\sim}{L}^h)\underset{\sim}{\mathfrak{B}}^{-1}(\underset{\sim}{E}_{rs}\otimes \underset{\sim}{L}^k)] = 0 , \tag{3}$$

for positive integers h,k.

Proof. From (3.5.3) we have

$$\underset{\sim}{\mathfrak{B}}^{-1} = \sum_{i=0}^{T-1} (\underset{\sim}{B}^{(i)}\otimes \underset{\sim}{L}^i) .$$

Similarly

$$\underset{\sim}{\mathfrak{A}}^{-1} = \sum_{j=0}^{T-1} (\underset{\sim}{A}^{(j)}\otimes \underset{\sim}{L}^j) .$$

So the left hand side of (3) can be rewritten as

$$(4) \qquad \sum_{j=0}^{T-1} \sum_{i=0}^{T-1} \operatorname{tr}(\underset{\sim}{A}^{(j)} \underset{\sim}{E}_{uv} \underset{\sim}{B}^{(i)} \underset{\sim}{E}_{rs} \otimes L^{j+h+i+k}) .$$

Now, Lemma 1(ii) applied to the summand in (4) yields

$$(5) \qquad \operatorname{tr}(\underset{\sim}{A}^{(j)} \underset{\sim}{E}_{uv} \underset{\sim}{B}^{(i)} \underset{\sim}{E}_{rs}) \cdot \operatorname{tr}(L^{j+h+i+k}) = 0 \quad ,$$

since $i+j+h+k$ is a positive integer and

$$\operatorname{tr}(L^r) = 0 \ , \quad r = 1,2, \ldots \quad .$$

This means that (4) is identically zero which proves the lemma.

Lemma 3.

$$(6) \qquad \underset{\sim}{\tilde{E}}' \operatorname{vec}[\underset{\sim}{\alpha}'^{-1}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}'^{h})\underset{\sim}{\beta}'^{-1}] = \underset{\sim}{0} .$$

Proof. The left hand side of (6) is a column vector, a typical element of which is

$$(7) \qquad \operatorname{vec}(\underset{\sim}{E}_{uv} \otimes \underset{\sim}{L}^{k}) \operatorname{vec}[\underset{\sim}{\alpha}'^{-1}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}'^{h})\underset{\sim}{\beta}'^{-1}] .$$

This follows from the definition of $\underset{\sim}{\tilde{E}}$ as given by (6.12). Now applying Lemma 1(i) and Lemma 2 to (7) we show that it is identically zero.

Q.E.D.

Note: It is obvious that by the same reasoning as in Lemma 3

(8) $$\tilde{\underset{\sim}{E}}' \operatorname{vec}[\underset{\sim}{\mathfrak{a}}'^{-1}(\underset{\sim}{E}_{ij} \otimes \underset{\sim}{L}'^{h})\underset{\sim}{\mathfrak{B}}'^{-1}] = \underset{\sim}{0} ,$$

which was used in (6.18).

The Frequency Domain.

Lemma 4. For $\underset{\sim}{A}(\cdot)$ and $\underset{\sim}{B}(\cdot)$ as in Chapter 2

(9) $$\int_0^{2\pi} \log|\underset{\sim}{A}\underset{\sim}{A}^*|\, d\lambda = \int_0^{2\pi} \log|\underset{\sim}{B}\underset{\sim}{B}^*|\, d\lambda = 0 ,$$

where we have omitted the argument $e^{i\lambda}$.

Proof. We shall prove this Lemma for $\underset{\sim}{A}$ and the argument for $\underset{\sim}{B}$ is identical. We shall show that

(10) $$\frac{\partial}{\partial \underset{\sim}{A}_h} \int_0^{2\pi} \log|\underset{\sim}{A}\underset{\sim}{A}^*|\, d\lambda = \underset{\sim}{0} , \quad h = 1, \ldots, q .$$

Then since for $\underset{\sim}{\alpha} = \underset{\sim}{0}$

$$\int_0^{2\pi} \log|\underset{\sim}{A}\underset{\sim}{A}^*|\, d\lambda = \int_0^{2\pi} \log|\underset{\sim}{I}_m|\, d\lambda = 0 ,$$

the desired result will follow. To derive the left hand side of (10) we have

$$\frac{\partial}{\partial a_{rs}^{(h)}} \log|\underset{\sim}{A}\underset{\sim}{A}^*| = tr(\underset{\sim}{A}^{*-1}\underset{\sim}{A}^{-1}\underset{\sim}{E}_{rs}e^{i\lambda h}\underset{\sim}{A}^*) + tr(\underset{\sim}{A}^{*-1}\underset{\sim}{A}^{-1}\underset{\sim}{A}\underset{\sim}{E}_{sr}e^{-ih\lambda})$$

$$= tr(\underset{\sim}{A}^{-1}\underset{\sim}{E}_{rs}e^{i\lambda h}) + tr(\underset{\sim}{A}^{*-1}\underset{\sim}{E}_{sr}e^{-i\lambda h})$$

$$= e^{i\lambda h}(\underset{\sim}{A}^{-1})_{sr} + e^{-i\lambda h}(\underset{\sim}{A}^{*-1})_{rs} .$$

From this we get

$$\frac{\partial \log|\underset{\sim}{A}\underset{\sim}{A}^*|}{\partial \underset{\sim}{A}_h} = e^{i\lambda h}\underset{\sim}{A}'^{-1} + e^{-i\lambda h}A^{*-1} .$$

So finally

$$\frac{\partial}{\partial \underset{\sim}{A}_h} = \int_0^{2\pi} \log|\underset{\sim}{A}\underset{\sim}{A}^*|d\lambda = \int_0^{2\pi} e^{i\lambda h}\underset{\sim}{A}'^{-1}d\lambda + \int_0^{2\pi} e^{-i\lambda h}\underset{\sim}{A}^{*-1}d\lambda = \underset{\sim}{0} ,$$

since $\underset{\sim}{A}'^{-1}$ is a power series in $e^{i\lambda}$ and $\underset{\sim}{A}^{*-1}$ a power series in $e^{-i\lambda}$ and

$$\int_0^{2\pi} e^{ik\lambda}d\lambda = \int_0^{2\pi} e^{-ik\lambda}d\lambda = 0 .$$

Q.E.D.

Lemma 5. For $\underset{\sim}{f}_t$, $\underset{\sim}{V}$ as in Chapter 4.

(11) $$\frac{1}{T}\sum_{t=0}^{T-1} \log|\underset{\sim}{f}_t| \to -m \log 2\pi + \log|\underset{\sim}{V}| .$$

Proof. The left hand side, as $T \to \infty$, tends to

(12) $$\frac{1}{2\pi}\int_0^{2\pi} \log|\underset{\sim}{f}(\lambda)|d\lambda \ .$$

But

$$|\underset{\sim}{f}(\lambda)| = (2\pi)^{-m}|\underset{\sim}{B}^{-1}\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{A}^*\underset{\sim}{B}^{*-1}| = (2\pi)^{-m}|\underset{\sim}{V}|\,|\underset{\sim}{B}\underset{\sim}{B}^*|^{-1}\,|\underset{\sim}{A}\underset{\sim}{A}^*| \ ,$$

which yields

$$\log|\underset{\sim}{f}(\lambda)| = -m \log 2\pi + \log|\underset{\sim}{V}| - \log|\underset{\sim}{B}\underset{\sim}{B}^*| + \log|\underset{\sim}{A}\underset{\sim}{A}^*| \ .$$

So (12) becomes

$$\begin{aligned} -m \log 2\pi + \log|\underset{\sim}{V}| &+ \frac{1}{2\pi}\int_0^{2\pi} (\log|\underset{\sim}{A}\underset{\sim}{A}^*| - \log|\underset{\sim}{B}\underset{\sim}{B}^*|)\, d\lambda \\ &= -m \log 2\pi + \log|\underset{\sim}{V}| \ , \end{aligned}$$

using Lemma 4.

Q.E.D.

Note. It follows from Lemma 5 that the modified log likelihood used in Chapter 4, given by (4.1.9) is asymptotically equivalent to the one used by Dunsmuir and Hannan (1976) and Nicholls (1976). One consequence of this is that maximizing (4.1.9) with respect to $\underset{\sim}{A}_1, \ldots, \underset{\sim}{A}_q, \underset{\sim}{B}_1, \ldots, \underset{\sim}{B}_p, \underset{\sim}{V}$ leads to asymptotically efficient estimates for these parameters. We have used this in the proof of the theorem in Chapter 5.

Lemma 6. For $\underset{\sim}{I}_t$ and $\underset{\sim}{f}_t$ as in Chapter 4

(13) $$\|\mathcal{E}\, \underset{\sim}{I}_t - \underset{\sim}{f}_t\| \le g_T \;, \quad t = 0,1, \ldots, T-1 \;,$$

where

$$\|\underset{\sim}{A}\| = \operatorname{tr} \underset{\sim}{A}\underset{\sim}{A}^* ,$$

for any matrix $\underset{\sim}{A}$ and

$$g_T = o(1) \;.$$

Proof. Using (4.1.2) we get

$$\mathcal{E}\, \underset{\sim}{I}_t = \frac{1}{2\pi} \sum_{|s| \le T-1} \underset{\sim}{V}_s\, e^{-is\lambda_t} \;.$$

So

(14) $$\mathcal{E}\, \underset{\sim}{I}_t - \underset{\sim}{f}_t = -\frac{1}{2\pi} \sum_{|s| > T-1} \underset{\sim}{V}_s\, e^{-is\lambda_t} - \frac{1}{2\pi T} \sum_{|s| \le T-1} |s| \underset{\sim}{V}_s e^{-is\lambda_t} \;.$$

Using triangle inequality on (14) yields

(15) $$\|\mathcal{E}\, \underset{\sim}{I}_t - \underset{\sim}{f}_t\| \le \frac{1}{2\pi} \sum_{|s| > T-1} \|\underset{\sim}{V}_s\| + \frac{1}{2\pi T} \sum_{|s| \le T-1} \|s\underset{\sim}{V}_s\| = g_T \;,$$

say. Now the first term of g_T is $o(1)$. This follows from

$$\frac{1}{2\pi} \sum_{|s| \leq T} \underset{\sim}{V}_s \to \underset{\sim}{f}_0 ,$$

which means

$$\sum_{|s| > T} \|\underset{\sim}{V}_s\| \to 0 \quad \text{as} \quad T \to \infty ,$$

The second term of g_T is also $o(1)$ because

$$\frac{\partial f}{\partial \lambda}\Big|_{\lambda=0} = \frac{i}{2\pi} \sum_{s=-\infty}^{\infty} s\underset{\sim}{V}_s \, e^{-is\lambda} , \tag{16}$$

this follows because $\underset{\sim}{f}$ is a rational function of $e^{i\lambda}$. Now (16) implies that

$$\sum_{|s| \leq T-1} \|s\underset{\sim}{V}_s\| \leq \sum_{s=-\infty}^{\infty} \|s\underset{\sim}{V}_s\| < \infty .$$

Q.E.D.

<u>Note</u>. Lemma 6 says that

$$\lim_{T \to \infty} \mathcal{E}\, \underset{\sim}{I}_t = \underset{\sim}{f}_t ,$$

and the convergence is uniform in t. This enabled us to derive a suitable approximation to

$$\mathcal{E}\left(\frac{\partial^2 \Lambda}{\partial \underset{\sim}{\rho}\, \partial \underset{\sim}{\rho}'}\right)$$

in section 4.5.

References

Ahrabi, Fereydoon (1978), "Maximum likelihood estimation of the covariances of the vector moving average models in the time and frequency domains," Technical Report No. 36, Department of Statistics, Stanford University, Stanford, CA.

Akaike, Hirotugu (1973), "Maximum likelihood identification of Gaussian autoregressive moving average model," Biometrika, Vol. 60, pp. 225-265.

Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, Inc., New York.

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

Anderson, T. W. (1975), "Maximum likelihood estimation of parameters of autoregressive processes with moving average residuals and other covariance matrices with linear structure," Annals of Statistics, Vol. 3, pp. 1283-1304.

Clevenson, M. Lawrence (1970), "Asymptotically efficient estimates of the parameters of a moving average time series," Technical Report No. 15, Department of Statistics, Stanford University, Stanford, CA.

Dickinson, Bradley W. (1978), "Efficient solution of linear equations with banded toeplitz matrices," To be published.

Dunsmuir, W. and Hannan, E. J. (1976), "Vector linear time series models," Advances in Applied Probability, Vol. 8, pp. 339-364.

Friedlander, B., Morf, M., Kailath, T., and Ljung, L. (1978), "New inversion formulas for matrices classified in terms of their distance from toeplitz matrices," To be published.

Hannan, E. J. (1969a), "The identification of vector mixed autoregressive moving average systems," Biometrika, Vol. 56, pp. 223-225.

Hannan, E. J. (1969b), "The estimation of mixed moving average autoregressive systems," Biometrika, Vol. 56, pp. 579-593.

Hannan, E. J. (1970), Multiple Time Series, John Wiley & Sons, Inc., New York.

Hannan, E. J. (1971), "The identification problem for multiple equation systems with moving average errors," Econometrika, Vol. 39, pp. 751-765.

Hannan, E. J. (1975), "The estimation of ARMA models," Annals of Statistics, Vol. 3, pp. 975-981.

Kashyap, R. L. (1970), "Maximum likelihood identification of stochastic linear systems," I.E.E.E. Transactions on Automatic Control, Vol. AC-15, pp. 25-34.

Kashyap, R. L., and Nasburg, Robert E. (1974), "Parameter estimation in multivariate stochastic difference equations," Transactions on Automatic Control, Vol. AC-19, pp. 784-797.

MacRae, Elizabeth Chase (1974), "Matrix derivatives with an application to an adaptive linear decision problem," The Annals of Statistics, Vol. 2, No. 2, pp. 337-346.

Minc, Henryk and Marcus, Marvin (1964), A Survey of Matrix Theory and Matrix Inequalities, Prindle, Weber & Schmidt, Boston.

Newton, Howard Joseph (1975), "The efficient estimation of stationary multiple time series mixed models: Theory and algorithms," Technical Report No. 33, Statistical Science Division, State University of New York at Buffalo.

Nicholls, D. F. (1976), "The efficient estimation of vector linear time series models," Biometrika, Vol. 64, pp. 85-90.

Osborn, Denise, R. (1977), "Exact and approximate maximum likelihood estimators for vector moving average processes," Journal of the Royal Statistical Society, Series B, Vol. 39, pp. 114-118.

Parzen, Emanuel (1971), "Efficient estimation of stationary time series mixed schemes," Bulletin of the International Statistical Institute, Vol. 44, pp. 315-319.

Reinsel, Gregory C. (1976), "Maximum likelihood estimation of vector autoregressive moving average models," Technical Report No. 117, Department of Statistics, Carnegie-Mellon University, Pittsburgh, PA.

Tunnicliffe, Wilson, G. (1973), "The estimation of parameters in multivariance time series models," Journal of the Royal Statistical Society, Series B, Vol. 35, pp. 76-85.

Whittle, P. (1953), "Estimation and information in stationary time series," Arkiv för Matematik, Vol. 2, pp. 423-434.

Whittle, P. (1961), "Gaussian estimation in stationary time series," Bulletin of the International Statistical Institute, Vol. 33, pp. 1-26.

Whittle, P. (1963), "On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix," Biometrika, Vol. 50, pp. 129-134.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.

2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.

3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.

4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.

5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.

6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.

7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.

8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.

9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.

10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.

11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.

12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.

13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.

14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.

15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.

16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations," Steffen L. Lauritzen, May, 1974.

18. "On the Interrelationships among Sufficiency, Total Sufficiency and Some Related Concepts," Steffen L. Lauritzen, June, 1974.

19. "Statistical Inference for Multiply Truncated Power Series Distributions," T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models in the Time and Frequency Domains," T. W. Anderson, June 1975.

21. "Asymptotic Properties of Some Estimators in Moving Average Models," Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting," Mituaki Huzii, February 1976.

23. "Estimating Means when Some Observations are Classified by Linear Discriminant Function," Chien-Pai Han, April 1976.

24. "Panels and Time Series Analysis: Markov Chains and Autoregressive Processes," T. W. Anderson, July 1976.

25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson, September 1976.

26. "The Recurrence Classification of Risk and Storage Processes," J. Michael Harrison and Sidney I. Resnick, September 1976.

27. "The Generalized Variance of a Stationary Autoregressive Process," T. W. Anderson and Raul P.Mentz, October 1976.

28. "Estimation of the Parameters of Finite Location and Scale Mixtures," Javad Behboodian, October 1976.

29. "Identification of Parameters by the Distribution of a Maximum Random Variable," T. W. Anderson and S.G. Ghurye, November 1976.

30. "Discrimination Between Stationary Guassian Processes, Large Sample Results," Will Gersch, January 1977.

31. "Principal Components in the Nonnormal Case: The Test for Sphericity," Christine M. Waternaux, October 1977.

32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a Multivariate Linear Model," F. Pukelsheim and George P.H. Styan, May 1978.

33. "Canonical Correlations with Respect to a Complex Structure," Steen A. Andersson, July 1978.

34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and I. Olkin, July 1978.

35. " Maximum likelihood Estimation for Vector Autoregressive Moving Average Models," T. W. Anderson, July 1978.

36. "Maximum likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.

37. "Efficient Estimation of a Model with an Autoregressive Signal with White Noise," Y. Hosoya, March 1979.

38. "Maximum Likelihood Estimation of the Parameters of a Multivariate Normal Distribution,"T.W. Anderson and I. Olkin, July 1979.

39. "Maximum Likelihood Estimation of the Autoregressive Coefficients and Moving Average Covariances of Vector Autoregressive Moving Average Models," Fereydoon Ahrabi, August 1979.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 39	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MAXIMUM LIKELIHOOD ESTIMATION OF THE AUTO-REGRESSIVE COEFFICIENTS AND MOVING AVERAGE COVARIANCES OF VECTOR AUTOREGRESSIVE MOVING AVERAGE MODELS		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) FEREYDOON AHRABI		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE AUGUST 1979 13. NUMBER OF PAGES 84
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Maximum likelihood estimation; vector autoregressive moving average model; Newton-Raphson; scoring method; time domain; frequency domain.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE REVERSE SIDE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

The vector autoregressive moving average process is a stationary stochastic process $\{\underset{\sim}{y}_t\}$ satisfying $\sum_{i=0}^{p} \underset{\sim}{B}_i \underset{\sim}{y}_{t-i} = \sum_{j=0}^{q} \underset{\sim}{A}_j \underset{\sim}{\epsilon}_{t-j} \equiv \underset{\sim}{u}_t$, where the unobservable vector process $\{\underset{\sim}{\epsilon}_t\}$ consists of independently identically distributed random variables. The matrix parameters $\underset{\sim}{B}_i$, $i = 1, \ldots, p$, $\underset{\sim}{\Sigma}^{(s)} = \mathcal{E}\underset{\sim}{u}_t \underset{\sim}{u}'_{t+s}$, $s = 0, 1, \ldots, q$ are estimated using the observations $\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_T$. The (modified) likelihood function is derived under the assumption of normality and to solve the maximum likelihood equations numerically, the Newton-Raphson and Scoring methods are used. The estimation problem is considered in the time and frequency domains. Asymptotic efficiency of the estimates is established. Finally estimates for $\underset{\sim}{B}_i$, $i = 1, \ldots, p$, $\underset{\sim}{A}_j$, $j = 1, \ldots, q$, $\underset{\sim}{V} = \mathcal{E}\underset{\sim}{\epsilon}_t \underset{\sim}{\epsilon}'_t$ are derived using the scoring method on the maximum likelihood equations in the time domain.