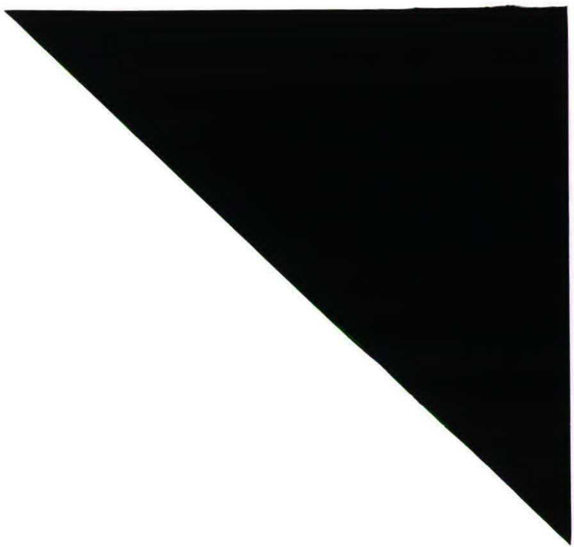
Research Memorandum

Faculty of Economics and
Business Administration

Tilburg University



THE SECOND DERIVATIVE OF THE LIKELIHOOD OF AN EXACT ARMA MODEL

Jan van der Leeuw

FEW 661

Communicated by Dr. H.H. Tigelaar

THE SECOND DERIVATIVE OF THE LIKELIHOOD OF AN EXACT ARMA MODEL

Jan van der Leeuw[1]
Dept. of Econometrics
Tilburg University
P.O.Box 90153
NL - 5000 LE Tilburg

Abstract

The likelihood of an exact ARMA model under the assumption of normality is investigated. Using a closed form expression of the covariance matrix the elements of the matrix of second derivatives of the concentrated likelihood function are derived. These elements consist in general of four or five terms. Two terms come from the determinant, one of them belonging to the information matrix. These terms are sums of the elements of the covariance matrix or its inverse. The next two terms are quadratic forms of the error vector. The last term is based on the matrix of independent variables, and thus only present in a regression model and not in the pure time series model.

The general form of second derivative does not permit conclusions about the existence of global maximum of the likelihood function.

JEL code: C22

Keywords : Autoregressive moving average process; exact ARMA covariance matrix; concentrated likelihood; second derivative.

1. Introduction

In Van der Leeuw (1993) it was shown how first order conditions for the parameters of a linear model with ARMA-errors can be derived and solved.

[1] I am indebted to H.H. Tigelaar for many suggestions and comments on an earlier draft.

Under the assumption of normality and using a closed form for the exact covariance matrix the (concentrated) likelihood function can be regarded as a function of the ARMA parameters and the error vector. The exact covariance is written in the form of lag matrices, which can simply be differentiated. This was done for the pure MA case, the pure AR case and the general ARMA model. The resulting first order conditions have at least one solution.

The solutions for the AR and MA parameters depend on the (computed) values of the error vector, which in turn is based upon the covariance matrix. Only in the pure MA and AR case of a time series model without explanatory variables direct solutions are found. In the general ARMA model the results for the MA part depend on the AR parameters and vice versa.

Supposing a linear model of the form $y=X\beta+\varepsilon$, with normally distributed errors, we maximize the likelihood function, which is equivalent to minimizing $S=|V|^{1/T}e'V^{-1}e$ (Judge *et al.*, p.284). Here $e=y-Xb$, b being the Aitken estimator of β, X a matrix of independent variables and $\sigma^2V=E\varepsilon\varepsilon'$. It is clear, that this model reduces to a pure time series model in case X is zero: e is identical to ε and y (see, *e.g.*, Anderson and Mentz, 1982). A convenient way to start with is to use T, the number of observations, times the logarithm of S : $S^{\bullet}=T\times\log S$. It consists *i.a.* of the covariance matrix and the error vector. The first derivative is simple enough to permit a useful expression for the second derivative.

First we will give a general expression of the second derivative, next we will discuss the details for the different cases: MA, AR and ARMA. It will become clear that conclusions about the existence of a global minimum cannot be based solely upon the behavioral of the second derivative.

Before we start, let us make clear what we mean when we use some expressions. Let A and B be two matrices. Then element (i,j) of matrix A is denoted as A[i,j]. If A[i,i+k]=A[i,i-k] for every i we will write A[k]. Furthermore **d**AB should be read as **d**(A)B and not **d**(AB) and **d**A+B as (**d**A)+B. When we use **tr**AB it should be understood as **tr**(AB), not (**tr**A)B.

2. First and second differential

As $S^{*}=T\times \log S$, its differential becomes

$$\mathbf{d}S^{*}=\mathbf{d}\{\log|V| + T\log(e'V^{-1}e)\}$$

or

$$\mathbf{d}S^{*}=\mathbf{tr}V^{-1}\mathbf{d}V + T(e'V^{-1}e)^{-1}e'\mathbf{d}V^{-1}e \tag{1}$$

because $e'V^{-1}\mathbf{d}e=(y-Xb)'V^{-1}\mathbf{d}(y-Xb)=-(y'V^{-1}X-b'X'V^{-1}X)\mathbf{d}b=0$. This is what Magnus (1978) called the ϑ-equation(s). This expression shall be our starting point. It has the advantage above using S, that the number of terms will be less, while it has the same stationary points.

Putting $e'V^{-1}e/T$ equal to s^2 we have $\mathbf{d}S^{*}=\mathbf{tr}V^{-1}\mathbf{d}V + e'\mathbf{d}V^{-1}e/s^2$.

In case an expression for V and its differential is available we rewrite $\mathbf{d}V^{-1}$ as $-V^{-1}\mathbf{d}VV^{-1}$ and have $s^2\mathbf{tr}V^{-1}\mathbf{d}V=e'V^{-1}\mathbf{d}VV^{-1}e$ as first order condition. When we have at our disposal the inverse of V - as in the AR case - we use $V\mathbf{d}V^{-1}=-V^{-1}\mathbf{d}V$ and get $s^2\mathbf{tr}V\mathbf{d}V^{-1}=e'\mathbf{d}V^{-1}e$.

Before we differentiate (1) again we first take the differential of $e=y-Xb$:

$$\begin{aligned}\mathbf{d}e &= -X\mathbf{d}b = -X\mathbf{d}((X'V^{-1}X)^{-1}X'V^{-1}y)=\\ &= X(X'V^{-1}X)^{-1}X'\mathbf{d}V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}y - X(X'V^{-1}X)^{-1}X'\mathbf{d}V^{-1}y=\\ &= -X(X'V^{-1}X)^{-1}X'\mathbf{d}V^{-1}(y-X(X'V^{-1}X)^{-1}X'V^{-1}y)\\ &= -X(X'V^{-1}X)^{-1}X'\mathbf{d}V^{-1}e.\end{aligned}$$

Hence $e'\mathbf{d}V^{-1}\mathbf{d}e$ is $-e'\mathbf{d}V^{-1}X(X'V^{-1}X)^{-1}X'\mathbf{d}V^{-1}e$.

Straightforward differentiation of (1) gives for the second differential:

$$\mathbf{d}^2S^{\bullet}= \mathbf{d}\{\mathbf{tr}V^{-1}\mathbf{d}V +T\ (e'V^{-1}e)^{-1}e'\mathbf{d}V^{-1}e\}$$

$$= \mathbf{tr}\mathbf{d}V^{-1}\mathbf{d}V +\mathbf{tr}V^{-1}\mathbf{d}^2V +T\mathbf{d}(e'V^{-1}e)^{-1}e'\mathbf{d}V^{-1}e +T(e'V^{-1}e)^{-1}(e'\mathbf{d}^2V^{-1}e +2e'\mathbf{d}V^{-1}\mathbf{d}e)$$

$$= \mathbf{tr}\mathbf{d}V^{-1}\mathbf{d}V +\mathbf{tr}V^{-1}\mathbf{d}^2V -T(e'V^{-1}e)^{-2}(e'\mathbf{d}V^{-1}e)\mathbf{d}(e'V^{-1}e)$$

$$+T(e'V^{-1}e)^{-1}e'\mathbf{d}^2V^{-1}e +2T(e'V^{-1}e)^{-1}e'\mathbf{d}V^{-1}\mathbf{d}e$$

or

$$\mathbf{d}^2S^{\bullet}= \mathbf{tr}\mathbf{d}V^{-1}\mathbf{d}V +\mathbf{tr}V^{-1}\mathbf{d}^2V -\frac{1}{T}\left(\frac{(e'\mathbf{d}V^{-1}e)}{s^2}\right)^2 +\frac{e'\mathbf{d}^2V^{-1}e}{s^2} -2\frac{e'\mathbf{d}V^{-1}X(X'V^{-1}X)^{-1}X'\mathbf{d}V^{-1}e}{s^2} \tag{2}$$

Of course the last term is not present in the pure time series model. The derivative corresponding to the first part of this expression, $\mathbf{tr}\mathbf{d}V^{-1}\mathbf{d}V$, is equal to minus the information matrix as shown by Magnus (1978).

Therefore it has to be negative:

$$\mathbf{tr}\mathbf{d}V^{-1}\mathbf{d}V= -\mathbf{tr}V^{-1}\mathbf{d}VV^{-1}\mathbf{d}V= -\mathbf{vec}(V^{-1}\mathbf{d}VV^{-1})\mathbf{vec}(\mathbf{d}V)$$

$$= -\mathbf{vec}(\mathbf{d}V)'(V^{-1}\otimes V^{-1})\ \mathbf{vec}(\mathbf{d}V)$$

$$<0.$$

Furthermore it is obvious that the third term and the last one, if present, are always negative. This is a far from encouraging situation as we are looking for a minimum. On the other hand the sign of the second and fourth term are not clear without any information about the structure of V. We will show that at least in the MA case these expressions are always positive.

When V is known, rewrite $\mathbf{d}V^{-1}$ and $\mathbf{d}^2V^{-1}$:

$$\mathbf{d}V^{-1}= -V^{-1}\mathbf{d}VV^{-1}$$

$$\mathbf{d}^2V^{-1}= \mathbf{d}(\mathbf{d}V^{-1})= \mathbf{d}(\ -V^{-1}\mathbf{d}VV^{-1})= 2V^{-1}\mathbf{d}VV^{-1}\mathbf{d}VV^{-1} -V^{-1}\mathbf{d}^2VV^{-1}.$$

Substituting in (2) we get

$$d^2S^* = -trV^{-1}dVV^{-1}dV + trV^{-1}d^2V - \frac{1}{T}\left(\frac{e'V^{-1}dVV^{-1}e}{s^2}\right)^2 + \frac{e'V^{-1}dVV^{-1}dVV^{-1}e}{s^2}$$

$$-\frac{e'V^{-1}d^2VV^{-1}e}{s^2} - 2\frac{e'V^{-1}dVV^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}dVV^{-1}e}{s^2}$$

or

$$d^2S^* = -trV^{-1}dVV^{-1}dV + trV^{-1}d^2V - \frac{1}{T}\left(\frac{e'V^{-1}dVV^{-1}e}{s^2}\right)^2 - \frac{e'V^{-1}d^2VV^{-1}e}{s^2} +$$

$$2\frac{e'V^{-1}dV(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})dVV^{-1}e}{s^2}. \quad (3)$$

Here the last expression at the right hand side is positive, while the sign of the fourth one is not clear.

On the other hand, when V^{-1} is known we rewrite dV and d^2V:

$dV^{-1} = -V^{-1}dVV^{-1}$ or $dV = -VdV^{-1}V$,

$d^2V^{-1} = 2V^{-1}dVV^{-1}dVV^{-1} - V^{-1}d^2VV^{-1}$ or $d^2V = 2VdV^{-1}VdV^{-1}V - Vd^2V^{-1}V$.

Substitution in (2) gives

$$d^2S^* = -trdV^{-1}VdV^{-1}V + trV^{-1}(2VdV^{-1}VdV^{-1}V - Vd^2V^{-1}V) - \frac{1}{T}\left(\frac{e'dV^{-1}e}{s^2}\right)^2 +$$

$$\frac{e'd^2V^{-1}e}{s^2} - 2\frac{e'dV^{-1}X(X'V^{-1}X)^{-1}X'dV^{-1}e}{s^2}$$

or

$$d^2S^* = trdV^{-1}VdV^{-1}V - trVd^2V^{-1} - \frac{1}{T}\left(\frac{e'dV^{-1}e}{s^2}\right)^2 + \frac{e'd^2V^{-1}e}{s^2} - 2\frac{e'dV^{-1}X(X'V^{-1}X)^{-1}X'dV^{-1}e}{s^2}$$

(4)

Observe that the first term is positive and that the third and last one are always negative. The sign of the second and fourth one are unknown.

Next we will use these expressions to give a more detailed description of the second derivatives for the different ARMA cases.

3. The ARMA covariance matrix

The general form of ARMA distributed errors is given by

$$\varepsilon_t = -\sum_{i=1}^{p}\vartheta_i\varepsilon_{t-i}+v_t+\sum_{i=1}^{q}\alpha_i v_{t-i},\quad t=1,..,T,$$

where v_t is a sequence of independently and identically distributed random variables. ϑ denotes the vector $(\vartheta_1,\vartheta_2,..,\vartheta_p)'$ of AR-parameters, α is the vector $(\alpha_1,\alpha_2,..,\alpha_q)'$ of MA-parameters. We assume that the invertibility conditions are fulfilled. By definition ϑ_0 and α_0 are equal to 1. Use $\sigma^2 V$ to denote the covariance matrix of ε: $\sigma^2V=E\varepsilon\varepsilon'$.

Following Pagan (1974), we introduce two matrices for both the AR parameters and the MA parameters. We define a (square) lower band matrix P of dimensions TxT, and a Txp matrix Q as follows:

$$P=\begin{bmatrix} 1 & & & & & \\ \vartheta_1 & \cdot & & & & \\ \cdot & \cdot & & & & \\ \cdot & & & & & \\ \vartheta_p & & & & & \\ & \cdot & & \cdot & \cdot & \\ & & \vartheta_p \cdot & \cdot & \vartheta_1 & 1 \end{bmatrix} \qquad Q=\begin{bmatrix} \vartheta_p & \vartheta_{p-1} & \cdot & \vartheta_1 \\ 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \\ & & & \cdot \\ 0 & & & \vartheta_p \\ 0 & \cdot & \cdot & 0 \\ \cdot & & & \cdot \\ 0 & \cdot & \cdot & 0 \end{bmatrix}$$

The upper triangular part of a lower band matrix consists of zeros and the lower part has off-diagonals with the same elements. Q consists of an upper pxp part with an upper band matrix and a lower (T-p)xp part, which consists of only zeros. Like P and Q will be used to describe the AR part of the error vector, so are M and N defined for the MA part, where ϑ is replaced by α and p by q. As is proven elsewhere (Van der Leeuw, 1992) the exact covariance matrix for ARMA errors is equal to $V=[N\ M][\bar{P}'\bar{P}-\bar{Q}\bar{Q}']^{-1}[N\ M]'$, where $\bar{P}$ is like P, but of order (T+p)x(T+p) and $\bar{Q}$ like Q, but of order (T+p)xp. In the MA case this expression reduces to $V=[N\ M][N\ M]'$ and in the AR case it becomes $[P'P-QQ']^{-1}$.

For our purpose it is obvious to rewrite the matrices of which the covariance matrix consists in such a way that they can be differentiated easily. To do this we introduce the lagmatrix, as described in Van der Leeuw (1993). Define ι_h as the Tx1 vector of which all elements are zero, except element h, being 1. Then the lagmatrix, for which we shall use the symbol L, is $L_k(n,m)=\sum_{h=n+1}^{T-m+k}\iota_h\iota_{h-k}'$, $n,m\geq\max(0,k)$. If both n and m are zero we write L_k and if n is equal to m $L_k(n)$. L_0 is the unit matrix, L_T the null matrix. Its transpose, $L_k'(n,m)$, is equal to $L_{-k}(n-k,m-k)$.

Using lag matrices we write P as $\sum_{i=0}^{p}L_i(i)\vartheta_i$ and Q as $(\sum_{i=0}^{p}L_{i-p}\vartheta_i)[I_p\ 0]'$, where I_p is the pxp unit matrix: forms which are linear in the parameters and that can be differentiated easily. Of course M and N are rewritten in a similar form.

4. The Moving Average case

In the pure MA case we have $V=[N\ M][N\ M]'=NN'+MM'$. Its first differential is $\mathbf{d}V=\mathbf{d}NN'+N\mathbf{d}N'+\mathbf{d}MM'+M\mathbf{d}M'$. As M and N are linear functions of α $\mathbf{d}^2M=\mathbf{d}^2N=0$ and the second differential becomes $\mathbf{d}^2V=2\mathbf{d}N\mathbf{d}N'+2\mathbf{d}M\mathbf{d}M'$. First we will show that the parts containing the second differential of V (of which the sign was not clear) are positive. Define the matrices of derivatives to α_i:

$N_i:=\dfrac{\partial N}{\partial\alpha_i}$ and $M_i:=\dfrac{\partial M}{\partial\alpha_i}$. Then $\mathbf{d}N=\sum_i\dfrac{\partial N}{\partial\alpha_i}\mathbf{d}\alpha_i=\sum_i N_i\mathbf{d}\alpha_i$ and $\mathbf{d}M=\sum_i\dfrac{\partial M}{\partial\alpha_i}\mathbf{d}\alpha_i=\sum_i M_i\mathbf{d}\alpha_i$.

A quadratic form like $\phi'\mathbf{d}^2V\phi$ is positive:

$$\phi'\mathbf{d}^2V\phi=2\phi'\mathbf{d}N\mathbf{d}N'\phi+2\phi'\mathbf{d}M\mathbf{d}M'\phi$$

$$=2\sum_i\phi'N_i\mathbf{d}\alpha_i\sum_j N_j'\phi\mathbf{d}\alpha_j+2\sum_i\phi'M_i\mathbf{d}\alpha_i\sum_j M_j'\phi\mathbf{d}\alpha_j$$

$$=2n'n+2m'm>0.$$

Here $n=\sum_j n_j$, $n_j= N_j'\phi d\alpha_j$ and $m=\sum_j m_j$, $m_j= M_j'\phi d\alpha_j$, all (Tx1)-vectors. The term containing the second differential is also positive:

$\mathbf{tr}V^{-1}\mathbf{d}^2V= \mathbf{tr}V^{-1}(2dNdN'+2dMdM')= 2\mathbf{tr}(dN'V^{-1}dN+dM'V^{-1}\mathbf{d}M)>0$, as both terms within parentheses are TxT positive definite matrices.

Using lagforms, we write the MA-covariance matrix as

$$V=\sum_{i=0}^{q}\Big(\sum_{j=0}^{i} L'_{j-i}\alpha_j+\sum_{j=i+1}^{q} L_{i-j}\alpha_j\Big)\alpha_i .$$

Its differential is

$$\mathbf{d}V=\sum_{i=0}^{q}\Big(\sum_{j=0}^{i}(L_{j-i}+L'_{j-i})\alpha_j+\sum_{j=i+1}^{q}(L_{i-j}+L'_{i-j})\alpha_j\Big)\mathbf{d}\alpha_i .$$

The corresponding derivatives are :

$$\frac{\partial V}{\partial\alpha_i}=\sum_{j=0}^{q}(L_{-|j-i|}+L'_{-|j-i|})\alpha_j,\quad i=1,..,q$$

and

$$\frac{\partial^2 V}{\partial\alpha_i\partial\alpha_j}=L_{-|j-i|}+L'_{-|j-i|},\quad i=1,..,q,\quad j=1,..,q.$$

From (3) we obtain the second derivative of the modified likelihood function:

Theorem 1

Second derivative MA-case.

$$\frac{\partial^2 S^{\bullet}}{\partial\alpha_i\partial\alpha_j}=-\mathbf{tr}V^{-1}\frac{\partial V}{\partial\alpha_i}V^{-1}\frac{\partial V}{\partial\alpha_j}+\mathbf{tr}V^{-1}\frac{\partial^2 V}{\partial\alpha_i\partial\alpha_j}-\frac{1}{T}\frac{1}{s^4}\phi'\frac{\partial V}{\partial\alpha_i}\phi\,\phi'\frac{\partial V}{\partial\alpha_j}\phi$$
$$-\frac{1}{s^2}\phi'\frac{\partial^2 V}{\partial\alpha_i\partial\alpha_j}\phi+\frac{2}{s^2}\phi'\frac{\partial V}{\partial\alpha_i}H\frac{\partial V}{\partial\alpha_j}\phi$$

where

1. $\mathrm{tr}V^{-1}\frac{\partial V}{\partial \alpha_i}V^{-1}\frac{\partial V}{\partial \alpha_j}= 2\sum_{k=0}^{q}\sum_{l=0}^{q}\sum_{h=1}^{T-|k-i|}\sum_{g=1}^{T-|l-j|}$

$(V^{-1}[g+|l-j|,h]V^{-1}[g,h+|k-i|]+V^{-1}[g,h]V^{-1}[g+|l-j|,h+|k-i|])\alpha_k\alpha_l$

2. $\mathrm{tr}V^{-1}\frac{\partial^2 V}{\partial \alpha_i \partial \alpha_j}= 2\sum_{h=1}^{T-|i-j|} V^{-1}[h+|i-j|,h]$

3. $\phi'\frac{\partial V}{\partial \alpha_i}\phi\ \phi'\frac{\partial V}{\partial \alpha_j}\phi= 4\{\sum_{k=0}^{q}\sum_{h=1}^{T-|k-i|}\phi_h\phi_{h+|k-i|}\alpha_k\}\{\sum_{l=0}^{q}\sum_{g=1}^{T-|l-j|}\phi_g\phi_{g+|l-j|}\alpha_l\}$

$\phi= V^{-1}e$

4. $\phi'\frac{\partial^2 V}{\partial \alpha_i \partial \alpha_j}\phi= 2\sum_{h=1}^{T-|j-i|}\phi_h\phi_{h+|j-i|}$

5. $\phi'\frac{\partial V}{\partial \alpha_i}H\frac{\partial V}{\partial \alpha_j}\phi= \sum_{k=0}^{q}(\phi(|k-i|)+\phi(-|k-i|))\alpha_k)'H(\sum_{l=0}^{q}\phi(|j-l|)+\phi(-|j-l|)\alpha_l).$

$\phi(k)= (\phi_{1+k}\ \ldots\ \phi_T\ \underbrace{0\ ..\ 0}_{\leftarrow k \rightarrow})'$

$\phi(-k)= (\underbrace{0\ ..\ 0}_{\leftarrow k \rightarrow}\ \phi_1\ \ldots\ \phi_{T-k}\ )'$

$H= V^{-1}\ -V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}.$

The proof is given in Appendix 1. Here we give a brief outline. Substitute the lagform of the covariance matrix in the expression of the derivative, next rewrite if necessary and use the properties of the trace operator. Eventually use the definition of the lagmatrix and the result follows. The first part and the most complicated one is $\mathrm{tr}V^{-1}\frac{\partial V}{\partial \alpha_i}V^{-1}\frac{\partial V}{\partial \alpha_j}$. It is a function of the elements of the inverse of the covariance matrix, with many elements if T is large. The next part, $\mathrm{tr}V^{-1}\frac{\partial^2 V}{\partial \alpha_i \partial \alpha_j}$, is rather simple and consists of the sum of the elements of the $|i-j|^{th}$ diagonal of the inverse of the dispersion matrix.

For the third and fourth part, $\phi'\frac{\partial V}{\partial\alpha_i}\phi\ \phi'\frac{\partial V}{\partial\alpha_j}\phi$ and $\phi'\frac{\partial^2 V}{\partial\alpha_i\partial\alpha_j}\phi$, we define $\phi=V^{-1}e$. The resulting expressions are simple sums and products of this vector.

For the last part, if present, we first define $\phi(k)=(\phi_{1+k}\ \dots\ \phi_T\ \underbrace{0\ ..\ 0}_{\leftarrow\ k\ \rightarrow})'$ to compute the vector $\frac{\partial V}{\partial\alpha_j}\phi$. Next we premultiply and postmultiply the quadratic form H with the appropriate vectors.

5. The Auto Regressive Case

The approach to the AR case is more or less as in the preceding section. However there are several differences. First, in this case we have an expression for the dispersion matrix in stead of the covariance matrix itself. As a consequence all elements of the off diagonals of several matrices we encounter are equal, which makes computations considerable easier. Second, both expressions, of which the sign is not clear, are not necessarily positive, be it that the quadratic form is almost always positive. Only if the T, the number of observations, is small compared to p, the number of parameters, it may become negative. Third, the determinant of the covariance matrix is equal to that of the submatrix consisting of the first p rows and columns.

In the AR-case we have $V^{-1}=P'P-QQ'$, with $\mathbf{d}V^{-1}=\mathbf{d}P'P+P'\mathbf{d}P-\mathbf{d}QQ'-Q\mathbf{d}Q'$ and $\mathbf{d}^2V^{-1}=2(\mathbf{d}P'\mathbf{d}P-\mathbf{d}Q\mathbf{d}Q')$, as P and Q are linear functions of the parameter vector. Write $\underline{V}$, $\underline{P}$ and $\underline{Q}$ for the pxp upper left submatrix of V, P and Q. Then $\underline{V}^{-1}=\underline{P}'\underline{P}-\underline{Q}\underline{Q}'$, and $|V|=|\underline{V}|$ (see Van der Leeuw, 1992). Therefore we have $\mathbf{tr}V\mathbf{d}^2V^{-1}=\mathbf{tr}\underline{V}\mathbf{d}^2\underline{V}^{-1}=2\mathbf{tr}\underline{V}(\mathbf{d}\underline{P}'\mathbf{d}\underline{P}-\mathbf{d}\underline{Q}\mathbf{d}\underline{Q}')=2\mathbf{trd}\underline{P}\underline{V}\mathbf{d}\underline{P}'-2\mathbf{trd}\underline{Q}'\underline{V}\mathbf{d}\underline{Q}$, the difference of the traces of two positive definite matrices. For the quadratic form

$e'd^2V^{-1}e$ we have $e'dP'dPe-e'dQdQ'e$, again the difference of two positive terms.

The lagform of the dispersion matrix is

$$V^{-1}=\sum_{i=0}^{p}\sum_{j=0}^{p}L_{j-i}(j)\vartheta_i\vartheta_j,$$

with as differential

$$dV^{-1}=\sum_{i=0}^{p}\sum_{j=0}^{p}(L_{j-i}(j)+L'_{j-i}(j))\vartheta_j d\vartheta_i.$$

The representation of $\underline{V}$ in lagform is

$$\underline{V}^{-1}=\sum_{i=0}^{p}(\sum_{j=0}^{p-i-1}L_{j-i}(j)\vartheta_j-\sum_{j=p-i+1}^{p}L_{j-i}(p-i)\vartheta_j)\vartheta_i$$

with differential

$$d\underline{V}^{-1}=\sum_{i=0}^{p}(\sum_{j=0}^{p-i-1}\{L_{j-i}(j)+L'_{j-i}(j)\}\vartheta_j-\sum_{j=p-i+1}^{p}\{L_{j-i}(p-i)+L'_{j-i}(p-i)\}\vartheta_j)d\vartheta_i.$$

Using these expressions we give the second derivative of the modified likelihood function.

Theorem 2

Second derivative AR-case.

$$\frac{\partial^2 S^*}{\partial\vartheta_i\partial\vartheta_j}=\operatorname{tr}\underline{V}\frac{\partial\underline{V}^{-1}}{\partial\vartheta_i}\underline{V}\frac{\partial\underline{V}^{-1}}{\partial\vartheta_j}-\operatorname{tr}\underline{V}\frac{\partial^2\underline{V}^{-1}}{\partial\vartheta_i\partial\vartheta_j}-\frac{1}{T}\frac{1}{s^4}e'\frac{\partial V^{-1}}{\partial\vartheta_i}e\;e'\frac{\partial V^{-1}}{\partial\vartheta_j}e+\frac{1}{s^2}e'\frac{\partial^2V^{-1}}{\partial\vartheta_i\partial\vartheta_j}e$$
$$-\frac{2}{s^2}e'\frac{\partial V^{-1}}{\partial\vartheta_i}\{X(X'V^{-1}X)^{-1}X'\}\frac{\partial V^{-1}}{\partial\vartheta_j}e,$$

where

1. $\mathbf{tr}\underline{V}\frac{\partial \underline{V}^{-1}}{\partial\vartheta_i}\underline{V}\frac{\partial \underline{V}^{-1}}{\partial\vartheta_j}=$

$$2\{\sum_{k=0}^{p-i-1}\sum_{l=0}^{p-j-1}\sum_{g=1+k}^{p-i}\sum_{h=1+l}^{p-j} - \sum_{k=0}^{p-i-1}\sum_{l=p-j+1}^{p}\sum_{g=1+k}^{p-i}\sum_{h=1+p-j}^{l}$$

$$-\sum_{k=p-i+1}^{p}\sum_{l=0}^{p-j-1}\sum_{g=1+p-i}^{k}\sum_{h=1+l}^{p-j} + \sum_{k=p-i+1}^{p}\sum_{l=p-j+1}^{p}\sum_{g=1+p-i}^{k}\sum_{nnh+p-j}^{l}\}$$

$$\{\underline{V}[h-l+j-g]\underline{V}[g-k+i-h] + \underline{V}[h-g]\ \underline{V}[g-k+i-h+l-j]\}\vartheta_k\vartheta_l$$

2. $\mathbf{tr}\underline{V}\frac{\partial^2\underline{V}^{-1}}{\partial\vartheta_i\partial\vartheta_j} = 2|p-i-j|\underline{V}[j-i]$

3. $e'\frac{\partial V^{-1}}{\partial\vartheta_i}e\ e'\frac{\partial V^{-1}}{\partial\vartheta_j}e = 4\{\sum_{k=0}^{p}(\sum_{h=1}^{T-i-k} e_{h+k}e_{h+i})\vartheta_k\}\{\sum_{k=0}^{p}(\sum_{h=1}^{T-j-k} e_{h+k}e_{h+j})\vartheta_k\}$

4. $e'\frac{\partial^2 V^{-1}}{\partial\vartheta_i\partial\vartheta_j}e = 2\sum_{h=1}^{T-i-j} e_{h+j}e_{h+i}$

5. $e'\frac{\partial V^{-1}}{\partial\vartheta_i}\{X(X'V^{-1}X)^{-1}X'\}\frac{\partial V^{-1}}{\partial\vartheta_j}e=$

$$\{\sum_{k=0}^{p}(e(i,k)+e(k,i))'\vartheta_k\}\{X(X'V^{-1}X)^{-1}X'\}\{\sum_{k=0}^{p}(e(j,k)+e(k,j))\vartheta_k\}$$

$$e(i,j) := (\underbrace{0\ ..\ 0}_{\leftarrow i \rightarrow}\ e_{1+j}\ \cdots\ e_{T-i}\ \underbrace{0\ ..\ 0}_{\leftarrow j \rightarrow})'.$$

The derivation of this formula can be found in Appendix 2. As in the MA case the information matrix gives most complications, be it that the number of computations is relatively small, because of the structure of the determinant. The second part is very simple: the second derivative is here equal to one of the elements of $\underline{V}$ times a scalar. At the same time it makes clear, that the sign is not certain.

The differential corresponding to this term is $\mathbf{tr}Vd^2V^{-1}$ or

$$\sum_{i=1}^{p}\sum_{j=1}^{p} 2|p-i-j|\underline{V}[j-i]\mathbf{d}\vartheta_i\mathbf{d}\vartheta_j.$$

If p=1 we get $\mathbf{tr}V\mathbf{d}^2V^{-1}=2\underline{V}[0]$, which is clearly positive. For p=2 it becomes $\mathbf{tr}V\mathbf{d}^2V^{-1}=\begin{bmatrix}\mathbf{d}\vartheta_1 & \mathbf{d}\vartheta_2\end{bmatrix}\begin{bmatrix}0 & 2\underline{V}[1]\\ 2\underline{V}[1] & 4\underline{V}[0]\end{bmatrix}\begin{bmatrix}\mathbf{d}\vartheta_1\\ \mathbf{d}\vartheta_2\end{bmatrix}$, a matrix with one positive and one negative root.

The fourth term, containing the second derivative of V^{-1}, is

$e'\frac{\partial^2 V^{-1}}{\partial\vartheta_i\partial\vartheta_j}e=2\sum_{h=1}^{T-i-j}e_{h+j}e_{h+i}$ or $2\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{h=1+i+j}^{T}e_{h-i}e_{h-j}$. The corresponding matrix of derivatives is almost sure positive. The differential $e'\mathbf{d}^2V^{-1}e$ can be split up in two parts, of which the larger one is always positive.

$$
\begin{aligned}
e'\mathbf{d}^2V^{-1}e&=\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{h=1+i+j}^{T}e_{h-i}e_{h-j}\mathbf{d}\vartheta_i\mathbf{d}\vartheta_j\\
&=\sum_{i=1}^{p}\sum_{j=1}^{p}\Big\{\sum_{h=1+i+j}^{2p}e_{h-i}e_{h-j}\mathbf{d}\vartheta_i\mathbf{d}\vartheta_j+\sum_{h=2p+1}^{T}e_{h-i}e_{h-j}\mathbf{d}\vartheta_i\mathbf{d}\vartheta_j\Big\}
\end{aligned}
$$

The former trem in this expression contains only $p^2(p-1)$, the latter one $p^2(T-2p)$ terms and is positive:

$$\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{h=2p+1}^{T}e_{h-i}e_{h-j}\mathbf{d}\vartheta_i\mathbf{d}\vartheta_j=\sum_{h=2p+1}^{T}\Big(\sum_{k=1}^{p}e_{h-k}\mathbf{d}\vartheta_k\Big)^2.$$

The sign of the former term is indeterminate. It can be split up again, such that the first part has the same structure as the complete expression, while the second part can be positive, zero or negative.

$$
\begin{aligned}
&\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{h=1+i+j}^{2p}e_{h-i}e_{h-j}\mathbf{d}\vartheta_i\mathbf{d}\vartheta_j=\\
&\quad=\sum_{i=1}^{p-1}\sum_{j=1}^{p-1}\sum_{h=1+i+j}^{2p}e_{h-i}e_{h-j}\mathbf{d}\vartheta_i\mathbf{d}\vartheta_j+2\sum_{k=1}^{p}\sum_{h=1+p+k}^{2p}e_{h-p}e_{h-k}\mathbf{d}\vartheta_h\mathbf{d}\vartheta_k
\end{aligned}
$$

For the last term of the second derivative, if present, we first compute a vector containing the derivative. It is obvious, that the corresponding matrix of second derivatives is positive definite.

6. The ARMA case

As can be expected the ARMA case is the most complicated one. The differential of the MA part is quite easy to find. For the AR part we only have the inverse in a form which can be rewritten in lagmatrices. Therefore formulas become longer and more complicated, but essential technical problems do not arise. In the sequel symbols containing a bar denote 'enlarged' matrices or vectors of order T+p in stead of T for expressions without a bar. Moreover we will use p for the number of parameters. This give no loss of generality as we can fill up the shorter vector with zeros. First we will treat the second differential in the direction of the MA parameter, next in the AR direction and eventually the mixed case.
The expression for V we will use is $[N\ M][\bar{P}'\bar{P}-\bar{Q}\bar{Q}']^{-1}[N\ M]'$, and thus we use (3) as the equation for the second differential. To facilitate notations we will use Δ for the (inverted) AR part, $\bar{P}'\bar{P}-\bar{Q}\bar{Q}'$. As Δ^{-1} is a covariance matrix, every diagonal has the same elements. From this expression it is clear, that the MA-differential of V will always contain the AR covariance matrix. This results in sums over all the elements of Δ. On the other hand the AR differentials of V suffer from the fact that only the inverse of a differentiable form is available. The consequence is second differential of two parts.

6.1 The ARMA case: MA-part

What we need is an expression for the covariance matrix that can simply be differentiated to α_1. To do this we first rewrite those parts of the covariance matrix containing MA parameters in such a way, that the parameter vector becomes explicit. Observe that [N M] can be written as $[0\ I]\bar{M}$, where $[0\ I]=\sum_{h=1}^{T}\iota_h\bar{\iota}'_{h+p}$ and $\bar{M}=\sum_{i=0}^{p}\bar{L}_i(i,i)\alpha_i$. Then $[0\ I]\bar{M}$ becomes:

$$[0\ I]\bar{M}=\sum_{h=1}^{T}\iota_h\bar{\iota}'_{h+p}\sum_{i=0}^{p}\sum_{g=1+i}^{T+p}\bar{\iota}_g\bar{\iota}'_{g-i}\alpha_i$$

$$=\sum_{i=0}^{p}\sum_{h=1}^{T}\iota_h\bar{\iota}'_{h+p}\bar{\iota}_{h+p}\bar{\iota}'_{h+p-i}\alpha_i$$

$$=\sum_{i=0}^{p}\sum_{h=1}^{T}\iota_h\bar{\iota}'_{h+p-i}\alpha_i$$

The covariance matrix is

$$V=\Big(\sum_{i=0}^{p}\sum_{h=1}^{T}\iota_h\bar{\iota}'_{h+p-i}\alpha_i\Big)\Delta^{-1}\Big(\sum_{j=0}^{p}\sum_{g=1}^{T}\iota_g\bar{\iota}'_{g+p-j}\alpha_j\Big)'$$

$$=\sum_{i=0}^{p}\sum_{j=0}^{p}\sum_{h=1}^{T}\sum_{g=1}^{T}\Delta^{-1}[h-i-g+j]\iota_h\iota'_g\alpha_i\alpha_j.$$

Its differential is

$$\mathbf{d}V=\sum_{i=1}^{p}\sum_{j=0}^{p}\sum_{h=1}^{T}\sum_{g=1}^{T}(\Delta^{-1}[h-g-i+j]+\Delta^{-1}[h-g-j+i])\iota_h\iota'_g\alpha_j\mathbf{d}\alpha_i$$

or

$$=\sum_{i=1}^{p}\sum_{j=0}^{p}\sum_{h=1}^{T}\sum_{g=1}^{T}(\Delta^{-1}[h-g-i+j]+\Delta^{-1}[g-h-i+j])\iota_h\iota'_g\alpha_j\mathbf{d}\alpha_i$$

as $\Delta^{-1}[k]=\Delta^{-1}[-k]$.

Theorem 3.1

Second derivative ARMA-case; MA-part.

$$\frac{\partial^2 S^*}{\partial\alpha_i\partial\alpha_j}=-\mathbf{tr}V^{-1}\frac{\partial V}{\partial\alpha_i}V^{-1}\frac{\partial V}{\partial\alpha_j}+\mathbf{tr}V^{-1}\frac{\partial^2 V}{\partial\alpha_i\partial\alpha_j}-\frac{1}{T}\frac{1}{s^4}\phi'\frac{\partial V}{\partial\alpha_i}\phi\,'\frac{\partial V}{\partial\alpha_j}\phi-\frac{2}{s^2}\phi'\frac{\partial^2 V}{\partial\alpha_i\partial\alpha_j}\phi+$$

$$+2\phi'\frac{\partial V}{\partial\alpha_i}H\frac{\partial V}{\partial\alpha_j}\phi$$

with

1. $\mathbf{tr}V^{-1}\frac{\partial V}{\partial\alpha_i}V^{-1}\frac{\partial V}{\partial\alpha_j}=\sum_{h_1=1}^{T}\sum_{h_2=1}^{T}\sum_{g_1=1}^{T}\sum_{g_2=1}^{T}$

$$\{\sum_{k_1=0}^{p}(\Delta^{-1}[h_1-g_1-i+k_1]+\Delta^{-1}[g_1-h_1-i+k_1])\alpha_{k_1}\}$$

$$\{\sum_{k_2=0}^{p}(\Delta^{-1}[h_2-g_2-k_2+j]+\Delta^{-1}[g_2-h_2-k_2+j])\alpha_{k_2}\}V^{-1}[g_2,h_1]V^{-1}[g_1,h_2]$$

2. $\mathbf{tr}V^{-1}\frac{\partial^2 V}{\partial\alpha_i\partial\alpha_j}=2\sum_{h=1}^{T}\sum_{g=1}^{T}\Delta^{-1}[h-g-i+j]V^{-1}[h,g]$

3. $\phi'\frac{\partial V}{\partial\alpha_i}\phi\phi'\frac{\partial V}{\partial\alpha_j}\phi=$

$$4\{\sum_{h=1}^{T}\sum_{g=1}^{T}(\sum_{k=0}^{p}\Delta^{-1}[h-g-i+k]\alpha_k)\phi_h\phi_g\}\{\sum_{h=1}^{T}\sum_{g=1}^{T}(\sum_{k=0}^{p}\Delta^{-1}[h-g-j+k]\alpha_k)\phi_h\phi_g\}$$

$\phi= V^{-1}e$

4. $\phi'\frac{\partial^2 V}{\partial\alpha_i\partial\alpha_j}\phi=2\sum_{h=1}^{T}\sum_{g=1}^{T}\Delta^{-1}[h-g-i+j]\phi_h\phi_g$

5. $\phi'\frac{\partial V}{\partial\alpha_i}H\frac{\partial V}{\partial\alpha_j}\phi=\{\sum_{h_1=1}^{T}\sum_{g_1=1}^{T}\sum_{h_2=1}^{T}\sum_{g_2=1}^{T}$

$$\{\sum_{k_1=0}^{p}(\Delta^{-1}[h_1-g_1-i+k_1]+\Delta^{-1}[g_1-h_1-i+k_1])\alpha_{k_1}\}$$

$$\{\sum_{k_2=0}^{p}(\Delta^{-1}[h_2-g_2-j+k_2]+\Delta^{-1}[g_2-h_2-j+k_2])\alpha_{k_2}\}\ H[h_1,h_2]\phi_{g_1}\phi_{g_2}\}$$

$H=(V^{-1}-V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})$.

The proof can be found in Appendix 3.1. The derivations are similar to the those in the MA and AR part. The first part, minus the information matrix has a simple structure but contains for large T many terms, consisting of elements of the (enlarged) AR covariance matrix and the dispersion matrix.

The second term looks like the first one, but in a simplified form. The quadratic parts are similar to the MA case, but now weighted with the elements of the AR covariance matrix. The last term with the quadratic form is again a complicated expression.

6.2 The ARMA case: AR-part

From $V = [N\ M]\Delta^{-1}[N\ M]'$ with $\Delta = \bar{P}'\bar{P} - \bar{Q}\bar{Q}'$ we conclude that we need the differential of Δ or $\mathbf{d}V = -[N\ M]\Delta^{-1}\mathbf{d}\Delta\Delta^{-1}[N\ M]'$. Hence the second differential is a less friendly expression with two terms: $\mathbf{d}^2V = 2[N\ M]\Delta^{-1}\mathbf{d}\Delta\Delta^{-1}\mathbf{d}\Delta\Delta^{-1}[N\ M]' - [N\ M]\Delta^{-1}\mathbf{d}^2\Delta\Delta^{-1}[N\ M]'$. It implicates that the resulting second derivatives become correspondingly longer. Otherwise the derivation brings no specific problems. Furthermore, contrary to the pure AR case, the determinant is not equal to the upper left submatrix of the covariance matrix.

From $\Delta = \sum_{i=0}^{q} \sum_{j=0}^{q} \bar{L}_{j-i}(j)\vartheta_j\vartheta_i$

we get

$$\mathbf{d}\Delta = \sum_{i=0}^{q} \sum_{j=0}^{q} (\bar{L}_{j-i}(j) + \bar{L}'_{j-i}(j))\vartheta_j \mathbf{d}\vartheta_i$$

and as derivative

$$\frac{\partial\Delta}{\partial\vartheta_i} = \sum_{k=0}^{q} (\bar{L}_{k-i}(k) + \bar{L}'_{k-i}(k))\vartheta_k \quad i = 1,\ldots,q.$$

Theorem 3.2

Second derivative ARMA-case; AR-part.

$$\frac{\partial^2 S^*}{\partial\vartheta_i\partial\vartheta_j} = -\mathbf{tr}V^{-1}\frac{\partial V}{\partial\vartheta_i}V^{-1}\frac{\partial V}{\partial\vartheta_j} + \mathbf{tr}V^{-1}\frac{\partial^2 V}{\partial\vartheta_i\partial\vartheta_j} - \frac{1}{T}\frac{1}{s^4}\phi'\frac{\partial V}{\partial\vartheta_i}\phi\ \phi'\frac{\partial V}{\partial\vartheta_j}\phi - \frac{2}{s^2}\phi'\frac{\partial^2 V}{\partial\vartheta_i\partial\vartheta_j}\phi$$
$$+ 2\phi'\frac{\partial V}{\partial\vartheta_i}H\frac{\partial V}{\partial\vartheta_j}\phi$$

with

$$1.\quad \mathbf{tr}U\frac{\partial\Delta}{\partial\vartheta_i}U\frac{\partial\Delta}{\partial\vartheta_j}= \sum_{k=0}^{p}\sum_{l=0}^{p}$$

$$\{\sum_{g=1+k}^{T+p-i}\sum_{h=1+l}^{T+p-j}U[h-l+j,g]U[g-k+i,h]+\sum_{g=1+k}^{T+p-i}\sum_{h=1+j}^{T+p-l}U[h-j+l,g]U[g-k+i,h]$$

$$+\sum_{g=1+i}^{T+p-k}\sum_{h=1+l}^{T+p-j}U[h-l+j,g]U[g-i+k,h]+\sum_{g=1+i}^{T+p-k}\sum_{h=1+j}^{T+p-l}U[h-j+l,g]U[g-i+k,h]\}\vartheta_k\vartheta_l$$

$$U= \Delta^{-1}[N\ M]'V^{-1}[N\ M]\Delta^{-1}$$

$$2.\quad \mathbf{tr}V^{-1}\frac{\partial^2 V}{\partial\vartheta_i\partial\vartheta_j}= 2\mathbf{tr}U\frac{\partial\Delta}{\partial\vartheta_i}\Delta^{-1}\frac{\partial\Delta}{\partial\vartheta_j}-\mathbf{tr}U\frac{\partial^2\Delta}{\partial\vartheta_i\partial\vartheta_j}$$

$$2.1\quad \mathbf{tr}U\frac{\partial\Delta}{\partial\vartheta_i}\Delta^{-1}\frac{\partial\Delta}{\partial\vartheta_j}=\sum_{k=0}^{p}\sum_{l=0}^{p}$$

$$\{\sum_{g=1+k}^{T+p-i}\sum_{h=1+l}^{T+p-j}U[h-l+j,g]\Delta^{-1}[g-k+i-h]+\sum_{g=1+k}^{T+p-i}\sum_{h=1+j}^{T+p-l}U[h-j+l,g]\Delta^{-1}[g-k+i-h]$$

$$+\sum_{g=1+i}^{T+p-k}\sum_{h=1+l}^{T+p-j}U[h-l+j,g]\Delta^{-1}[g-i+k-h]+\sum_{g=1+i}^{T+p-k}\sum_{h=1+j}^{T+p-l}U[h-j+l,g]\Delta^{-1}[g-i+k-h]\}\vartheta_k\vartheta_l$$

$$2.2\quad \mathbf{tr}U\frac{\partial^2\Delta}{\partial\vartheta_i\partial\vartheta_j}= 2\sum_{h=1+j}^{T+p-i}U[h-j+i,h]$$

$$3.\quad \zeta'\frac{\partial\Delta}{\partial\vartheta_i}\zeta\ \zeta'\frac{\partial\Delta}{\partial\vartheta_j}\zeta= 4\{\sum_{k=0}^{p}\sum_{g=1+k}^{T+p-i}\zeta_g\zeta_{g-k+i}\vartheta_k\}\{\sum_{l=0}^{p}\sum_{h=1+l}^{T+p-j}\zeta_h\zeta_{h-l+j}\vartheta_l\}$$

$$\zeta= \Delta^{-1}[N\ M]'V^{-1}e$$

$$4.\quad e'V^{-1}\frac{\partial^2 V}{\partial\vartheta_i\partial\vartheta_j}V^{-1}e= 2\zeta'\frac{\partial\Delta}{\partial\vartheta_i}\Delta^{-1}\frac{\partial\Delta}{\partial\vartheta_j}\zeta-\zeta'\frac{\partial^2\Delta}{\partial\vartheta_i\partial\vartheta_j}\zeta$$

$$\zeta(i,j)= (\underbrace{0\ ..\ 0}_{\leftarrow\ i\ \rightarrow}\ \zeta_{1+j}\ \cdots\ \zeta_{T+p-i}\ \underbrace{0\ ..\ 0}_{\leftarrow\ j\ \rightarrow})'$$

$$4.1\quad \zeta'\frac{\partial\Delta}{\partial\vartheta_i}\Delta^{-1}\frac{\partial\Delta}{\partial\vartheta_j}\zeta= \{\sum_{k=0}^{p}\{\zeta(i,k)+\zeta(k,i)\}\vartheta_k\}'\Delta^{-1}\{\sum_{l=0}^{p}\{\zeta(j,l)+\zeta(l,j)\}\vartheta_l\}$$

$$4.2\ \zeta'\frac{\partial^2\Delta}{\partial\vartheta_i\partial\vartheta_j}\zeta=\sum_{h=1}^{T+p-i-j}\zeta_{h+j}\zeta_{h+i}$$

$$5.\ \zeta'\frac{\partial\Delta}{\partial\vartheta_i}G\frac{\partial\Delta}{\partial\vartheta_j}\zeta=\{\sum_{k=0}^{p}(\zeta(k,i)+\zeta(i,k))\vartheta_k\}'G\{\sum_{k=0}^{p}(\zeta(k,i)+\zeta(i,k))\vartheta_k\zeta'\}$$

$$G=\Delta^{-1}[N\ M]'(V^{-1}-V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})[N\ M]\Delta^{-1}$$

For the proof see Appendix 3.2. To describe the information matrix we introduce U, a mixture of the AR covariance matrix, a matrix with MA parameters and the dispersion matrix. We get an expression as in the pure AR case, but with much more elements. For the second and fourth part (containing the second derivative of V) we need two terms. For the quadratic form we introduce ζ, a transformation of the error vector, like ϕ in the preceding sections.

6.3 The ARMA case: mixed part

In the preceding sections the direction of the differential was clear. Here however, we have to make clear which differential is meant. Therefore we introduce $\mathbf{d}_\alpha V=\frac{\partial V}{\partial\alpha}\mathbf{d}\alpha$ and $\mathbf{d}_\vartheta V=\frac{\partial V}{\partial\vartheta}\mathbf{d}\vartheta$.

For the MA- direction we use as before

$$\mathbf{d}_\alpha V=\sum_{i=0}^{p}\sum_{s=0}^{p}\sum_{h=1}^{T}\sum_{g=1}^{T}(\bar\iota'_{h+p-i}\Delta^{-1}\bar\iota_{g+p-s}\iota'_g+\bar\iota'_{h+p-s}\Delta^{-1}\bar\iota_{g+p-i})\iota_h\iota'_g\alpha_s\mathbf{d}\alpha_i$$

or

$$=\sum_{i=0}^{p}\sum_{s=0}^{p}\sum_{h=1}^{T}\sum_{g=1}^{T}\{\Delta^{-1}[h-g-i+s]+\Delta^{-1}[g-h-i+s]\}\iota_h\iota'_g\alpha_s\mathbf{d}\alpha_i.$$

For the ϑ direction we have $\mathbf{d}_{\vartheta}V = -[N\ M]\Delta^{-1}\mathbf{d}_{\vartheta}\Delta\Delta^{-1}[N\ M]'$ and

$$\mathbf{d}_{\vartheta}\Delta = \sum_{j=0}^{p}\sum_{k=0}^{p}(\bar{L}_{k-j}(k)+\bar{L}'_{k-j}(k))\vartheta_k \mathbf{d}\vartheta_j.$$

Taking the differential of $\mathbf{d}_{\alpha}V$ in the ϑ-direction gives, as shown in appendix 3.3,

$$\mathbf{d}_{\vartheta}\mathbf{d}_{\alpha}V = -\sum_{i=0}^{p}\sum_{j=0}^{p}\sum_{s=0}^{p}\sum_{t=0}^{p}\sum_{h=1}^{T}\sum_{g=1}^{T}\sum_{r=1+t}^{T+p-j}$$

$$\{\Delta^{-1}[h+p-i-r]\Delta^{-1}[r-t+j-g-p+s]+\Delta^{-1}[h+p-i-r+t-j]\Delta^{-1}[r-g-p+s]+$$

$$\Delta^{-1}[h+p-s-r]\Delta^{-1}[r-t+j-g-p+i]+\Delta^{-1}[h+p-s-r+t-j]\Delta^{-1}[r-g-p+i]\}\iota_h\iota'_g\alpha_s\vartheta_t\mathbf{d}\vartheta_j\mathbf{d}\alpha_i,$$

which shows that we may expect complicated expressions.

Theorem 3.3

Second derivative ARMA-case: AR/MA-part

$$\frac{\partial^2 S^*}{\partial\alpha_i\partial\vartheta_j} = -\mathbf{tr}Z'\frac{\partial\Delta}{\partial\vartheta_i}Z\frac{\partial V}{\partial\alpha_j} + \mathbf{tr}V^{-1}\frac{\partial^2 V}{\partial\alpha_i\partial\vartheta_j} - \frac{1}{T}\frac{1}{s^4}\phi'\frac{\partial V}{\partial\alpha_i}\phi\ \zeta'\frac{\partial V}{\partial\vartheta_j}\zeta - \frac{1}{s^2}\phi'\frac{\partial^2 V}{\partial\alpha_i\partial\vartheta_j}\phi + 2\phi'\frac{\partial V}{\partial\alpha_i}H\frac{\partial\Delta}{\partial\vartheta_j}\zeta$$

with

1. $$\mathbf{tr}Z'\frac{\partial\Delta}{\partial\vartheta_i}Z\frac{\partial V}{\partial\alpha_j} = \sum_{h=1}^{T}\sum_{g=1}^{T}\{\sum_{s=0}^{p}(\Delta^{-1}[h-g-j+s]+\Delta^{-1}[g-h-j+s])\alpha_s\}$$

$$\{\sum_{k=0}^{p}\sum_{n=1+k}^{T+p-i}(Z[g,n]Z[n-k+i,h]+Z[h,n]Z[n-k+i,g])\vartheta_k\}$$

$Z = \Delta^{-1}[N\ M]'V^{-1}$

2. $$\mathbf{tr}V^{-1}\frac{\partial^2 V}{\partial\alpha_i\partial\vartheta_j} = -\sum_{s=0}^{p}\sum_{t=0}^{p}\sum_{h=1}^{T}\sum_{g=1}^{T}\sum_{r=1+t}^{T+p-s}$$

$$\{\Delta^{-1}[h+p-i-r]\Delta^{-1}[r-t+j-g-p+s]+\Delta^{-1}[h+p-i-r+t-j]\Delta^{-1}[r-g-p+s]+$$

$$\Delta[h+p-s-r]\Delta^{-1}[r-t+j-g-p+i]+\Delta^{-1}[h+p-s-r+t-j]\Delta^{-1}[r-g-p+i]\}V^{-1}[h,g]\alpha_s\vartheta_t$$

3. $\phi' \frac{\partial V}{\partial \alpha_i} \phi \zeta' \frac{\partial \Delta}{\partial \vartheta_j} \zeta =$

$$2\{\sum_{s=0}^{p} \sum_{h=1}^{T} \sum_{g=1}^{T} \{\Delta^{-1}[h-g-i+s]+\Delta^{-1}[h-g-s+i]\}\phi_h \phi_g \alpha_s\}\{\sum_{k=0}^{p} \sum_{g=1}^{T+p-i-k} \zeta_{g+k}\zeta_{g+i}\vartheta_k\}$$

$\phi = V^{-1}e$

$\zeta = \Delta^{-1}[N\ M]'V^{-1}e$

$$4.\ \phi' \frac{\partial^2 V}{\partial \alpha_i \partial \vartheta_j}\phi = -\sum_{s=0}^{p} \sum_{t=0}^{p} \sum_{h=1}^{T} \sum_{g=1}^{T} \sum_{r=1+t}^{T+p-j}$$

$$\{\Delta^{-1}[h+p-i-r]\Delta^{-1}[r-t+j-g-p+s]+\Delta^{-1}[h+p-i-r+t-j]\Delta^{-1}[r-g-p+s]+$$

$$\Delta^{-1}[h+p-s-r]\Delta^{-1}[r-t+j-g-p+i]+\Delta^{-1}[h+p-s-r+t-j]\Delta^{-1}[r-g-p+i]\}\phi_h\phi_g\alpha_s\vartheta_t$$

5. $\phi' \frac{\partial V}{\partial \alpha_i} H \frac{\partial \Delta}{\partial \vartheta_j} \zeta =$

$$\sum_{h_1=1}^{T} \sum_{g_2=1}^{T} \{ \sum_{g_1=1}^{T} \sum_{s=0}^{p} \{\Delta^{-1}[h_1-g_1-i+s]+\Delta^{-1}[g_1-h_1-i+s]\}\phi_{g_1}\alpha_s\}\ H[h_1,g_2]$$

$$\{\sum_{k=0}^{p} \sum_{l=0}^{p} \sum_{h_2=1}^{T+p-j-l} \Delta^{-1}[g_2+p-k-h_2-l]\zeta_{h_2+j}+\Delta^{-1}[g_2+p-k-h_2-j]\zeta_{h_2+l}\}\vartheta_l$$

$H = (V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})$.

For the proof see appendix 3.3. In line with the foregoing cases, the result is as may be expected: rather complicated expressions. The information matrix part is maybe less complicated than expected, a consequence of the relatively simple form of the first derivative in the ϑ-direction.

7. Conclusion

The differential of the logarithm of the concentrated likelihood consists of two parts: the differential of the determinant and the differential of the quadratic form of the errors. Another differentiation gives five terms, two of them coming from the determinant, the other three from the quadratic form. Some of these are positive definite, some negative definite. In several cases the sign is not clear or can be either positive or negative. All terms can be expressed as function of the covariance matrix or matrices of which the covariance matrix is composed. The resulting algorithms are in several cases very computer time consuming because of the number of summations. The second derivative cannot assure us whether a stationary point is a unique global optimum.

Appendix

The appendix is available upon request.

References

Anderson, T.W. and R.P. Mentz, 1982, Maximum Likelihood Estimation in autoregressive and Moving Average Models, Time Series Analysis: Theory and Practice, Amsterdam, 23-29.

Judge, G.C. et al., 1985, The Theory and Practice of Econometrics, New York.

Kohn,R. and C.F. Ansley, 1985, Computing the Likelihood and its Derivatives for a Gaussian ARMA Model, Journal of Statistical Computation and Simulation, 22, 229-263.

Van der Leeuw, J.L., forthcoming, The Covariance Matrix of ARMA-errors in Closed Form, Journal of Econometrics.

Van der Leeuw, J.L., 1993, First Order Conditions for the Maximum Likelihood Estimation of an Exact ARMA-model, Research Memorandum FEW 611, Tilburg University.

Magnus, J.R., 1978, Maximum Likelihood Estimation of the GLS Model with Unknown Parameters in the Disturbance Covariance Matrix, Journal of Econometrics, 7, 281-312.

Pagan, A. 1974, A Generalised Approach to the Treatment of Autocorrelation, Australian Economic Papers, 13, 267-280.

IN 1993 REEDS VERSCHENEN

588 Rob de Groof and Martin van Tuijl
The Twin-Debt Problem in an Interdependent World
Communicated by Prof.dr. Th. van de Klundert

589 Harry H. Tigelaar
A useful fourth moment matrix of a random vector
Communicated by Prof.dr. B.B. van der Genugten

590 Niels G. Noorderhaven
Trust and transactions; transaction cost analysis with a differential behavioral assumption
Communicated by Prof.dr. S.W. Douma

591 Henk Roest and Kitty Koelemeijer
Framing perceived service quality and related constructs A multilevel approach
Communicated by Prof.dr. Th.M.M. Verhallen

592 Jacob C. Engwerda
The Square Indefinite LQ-Problem: Existence of a Unique Solution
Communicated by Prof.dr. J. Schumacher

593 Jacob C. Engwerda
Output Deadbeat Control of Discrete-Time Multivariable Systems
Communicated by Prof.dr. J. Schumacher

594 Chris Veld and Adri Verboven
An Empirical Analysis of Warrant Prices versus Long Term Call Option Prices
Communicated by Prof.dr. P.W. Moerland

595 A.A. Jeunink en M.R. Kabir
De relatie tussen aandeelhoudersstructuur en beschermingsconstructies
Communicated by Prof.dr. P.W. Moerland

596 M.J. Coster and W.H. Haemers
Quasi-symmetric designs related to the triangular graph
Communicated by Prof.dr. M.H.C. Paardekooper

597 Noud Gruijters
De liberalisering van het internationale kapitaalverkeer in historisch-institutioneel perspectief
Communicated by Dr. H.G. van Gemert

598 John Görtzen en Remco Zwetheul
Weekend-effect en dag-van-de-week-effect op de Amsterdamse effectenbeurs?
Communicated by Prof.dr. P.W. Moerland

599 Philip Hans Franses and H. Peter Boswijk
Temporal aggregration in a periodically integrated autoregressive process
Communicated by Prof.dr. Th.E. Nijman

600 René Peeters
On the p-ranks of Latin Square Graphs
Communicated by Prof.dr. M.H.C. Paardekooper

601 Peter E.M. Borm, Ricardo Cao, Ignacio García-Jurado
Maximum Likelihood Equilibria of Random Games
Communicated by Prof.dr. B.B. van der Genugten

602 Prof.dr. Robert Bannink
Size and timing of profits for insurance companies. Cost assignment for products with multiple deliveries.
Communicated by Prof.dr. W. van Hulst

603 M.J. Coster
An Algorithm on Addition Chains with Restricted Memory
Communicated by Prof.dr. M.H.C. Paardekooper

604 Ton Geerts
Coordinate-free interpretations of the optimal costs for LQ-problems subject to implicit systems
Communicated by Prof.dr. J.M. Schumacher

605 B.B. van der Genugten
Beat the Dealer in Holland Casino's Black Jack
Communicated by Dr. P.E.M. Borm

606 Gert Nieuwenhuis
Uniform Limit Theorems for Marked Point Processes
Communicated by Dr. M.R. Jaïbi

607 Dr. G.P.L. van Roij
Effectisering op internationale financiële markten en enkele gevolgen voor banken
Communicated by Prof.dr. J. Sijben

608 R.A.M.G. Joosten, A.J.J. Talman
A simplicial variable dimension restart algorithm to find economic equilibria on the unit simplex using $n(n+1)$ rays
Communicated by Prof.Dr. P.H.M. Ruys

609 Dr. A.J.W. van de Gevel
The Elimination of Technical Barriers to Trade in the European Community
Communicated by Prof.dr. H. Huizinga

610 Dr. A.J.W. van de Gevel
Effective Protection: a Survey
Communicated by Prof.dr. H. Huizinga

611 Jan van der Leeuw
First order conditions for the maximum likelihood estimation of an exact ARMA model
Communicated by Prof.dr. B.B. van der Genugten

612 Tom P. Faith
Bertrand-Edgeworth Competition with Sequential Capacity Choice
Communicated by Prof.Dr. S.W. Douma

613 Ton Geerts
The algebraic Riccati equation and singular optimal control: The discrete-time case
Communicated by Prof.dr. J.M. Schumacher

614 Ton Geerts
Output consistency and weak output consistency for continuous-time implicit systems
Communicated by Prof.dr. J.M. Schumacher

615 Stef Tijs, Gert-Jan Otten
Compromise Values in Cooperative Game Theory
Communicated by Dr. P.E.M. Borm

616 Dr. Pieter J.F.G. Meulendijks and Prof.Dr. Dick B.J. Schouten
Exchange Rates and the European Business Cycle: an application of a 'quasi-empirical' two-country model
Communicated by Prof.Dr. A.H.J.J. Kolnaar

617 Niels G. Noorderhaven
The argumentational texture of transaction cost economics
Communicated by Prof.Dr. S.W. Douma

618 Dr. M.R. Jaïbi
Frequent Sampling in Discrete Choice
Communicated by Dr. M.H. ten Raa

619 Dr. M.R. Jaïbi
A Qualification of the Dependence in the Generalized Extreme Value Choice Model
Communicated by Dr. M.H. ten Raa

620 J.J.A. Moors, V.M.J. Coenen, R.M.J. Heuts
Limiting distributions of moment- and quantile-based measures for skewness and kurtosis
Communicated by Prof.Dr. B.B. van der Genugten

621 Job de Haan, Jos Benders, David Bennett
Symbiotic approaches to work and technology
Communicated by Prof.dr. S.W. Douma

622 René Peeters
Orthogonal representations over finite fields and the chromatic number of graphs
Communicated by Dr.ir. W.H. Haemers

623 W.H. Haemers, E. Spence
Graphs Cospectral with Distance-Regular Graphs
Communicated by Prof.dr. M.H.C. Paardekooper

624 Bas van Aarle
The target zone model and its applicability to the recent EMS crisis
Communicated by Prof.dr. H. Huizinga

625 René Peeters
Strongly regular graphs that are locally a disjoint union of hexagons
Communicated by Dr.ir. W.H. Haemers

626 René Peeters
Uniqueness of strongly regular graphs having minimal p-rank
Communicated by Dr.ir. W.H. Haemers

627 Freek Aertsen, Jos Benders
Tricks and Trucks: Ten years of organizational renewal at DAF?
Communicated by Prof.dr. S.W. Douma

628 Jan de Klein, Jacques Roemen
Optimal Delivery Strategies for Heterogeneous Groups of Porkers
Communicated by Prof.dr. F.A. van der Duyn Schouten

629 Imma Curiel, Herbert Hamers, Jos Potters, Stef Tijs
The equal gain splitting rule for sequencing situations and the general nucleolus
Communicated by Dr. P.E.M. Borm

630 A.L. Hempenius
Een statische theorie van de keuze van bankrekening
Communicated by Prof.Dr.Ir. A. Kapteyn

631 Cok Vrooman, Piet van Wijngaarden, Frans van den Heuvel
Prevention in Social Security: Theory and Policy Consequences
Communicated by Prof.Dr. A. Kolnaar

IN 1994 REEDS VERSCHENEN

632 B.B. van der Genugten
Identification, estimating and testing in the restricted linear model
Communicated by Dr. A.H.O. van Soest

633 George W.J. Hendrikse
Screening, Competition and (De)Centralization
Communicated by Prof.dr. S.W. Douma

634 A.J.T.M. Weeren, J.M. Schumacher, and J.C. Engwerda
Asymptotic Analysis of Nash Equilibria in Nonzero-sum Linear-Quadratic Differential Games. The Two-Player case
Communicated by Prof.dr. S.H. Tijs

635 M.J. Coster
Quadratic forms in Design Theory
Communicated by Dr.ir. W.H. Haemers

636 Drs. Erwin van der Krabben, Prof.dr. Jan G. Lambooy
An institutional economic approach to land and property markets - urban dynamics and institutional change
Communicated by Dr. F.W.M. Boekema

637 Bas van Aarle
Currency substitution and currency controls: the Polish experience of 1990
Communicated by Prof.dr. H. Huizinga

638 J. Bell
Joint Ventures en Ondernemerschap: Interpreneurship
Communicated by Prof.dr. S.W. Douma

639 Frans de Roon and Chris Veld
Put-call parities and the value of early exercise for put options on a performance index
Communicated by Prof.dr. Th.E. Nijman

640 Willem J.H. Van Groenendaal
Assessing demand when introducing a new fuel: natural gas on Java
Communicated by Prof.dr. J.P.C. Kleijnen

641 Henk van Gemert & Noud Gruijters
Patterns of Financial Change in the OECD area
Communicated by Prof.dr. J.J Sijben

642 Drs. M.R.R. van Bremen, Drs. T.A. Marra en Drs. A.H.F. Verboven
Aardappelen, varkens en de termijnhandel: de reële optietheorie toegepast
Communicated by Prof.dr. P.W. Moerland

643 W.J.H. Van Groenendaal en F. De Gram
The generalization of netback value calculations for the determination of industrial demand for natural gas
Communicated by Prof.dr. J.P.C. Kleijnen

644 Karen Aardal, Yves Pochet and Laurence A. Wolsey
Capacitated Facility Location: Valid Inequalities and Facets
Communicated by Dr.ir. W.H. Haemers

645 Jan J.G. Lemmen
An Introduction to the Diamond-Dybvig Model (1983)
Communicated by Dr. S. Eijffinger

646 Hans J. Gremmen and Eva van Deurzen-Mankova
Reconsidering the Future of Eastern Europe: The Case of Czecho-Slovakia
Communicated by Prof.dr. H.P. Huizinga

647 H.M. Webers
Non-uniformities in spatial location models
Communicated by Prof.dr. A.J.J. Talman

648 Bas van Aarle
Social welfare effects of a common currency
Communicated by Prof.dr. H. Huizinga

649 Laurence A.G.M. van Lent
De winst is absoluut belangrijk!
Communicated by Prof.drs. G.G.M. Bak

650 Bert Hamminga
Jager over de theorie van de internationale handel
Communicated by Prof.dr. H. Huizinga

651 J.Ch. Caanen and E.N. Kertzman
A comparison of two methods of inflation adjustment
Communicated by Prof.dr. J.A.G. van der Geld

652 René van den Brink
A Note on the τ-Value and τ-Related Solution Concepts
Communicated by Prof.dr. P.H.M. Ruys

653 J. Engwerda and G. van Willigenburg
Optimal sampling-rates of digital LQ and LQG tracking controllers with costs associated to sampling
Communicated by Prof.dr. J.M. Schumacher

654 J.C. de Vos
A Thousand Golden Ten Orbits
Communicated by Prof.dr. B.B. van der Genugten

Katholieke Universiteit Brabant
PO Box 90153
5000 LE Tilburg
The Netherlands