

The Information Matrices of the Parameters of Multiple Mixed Time Series*

H. JOSEPH NEWTON

State University of New York at Buffalo

Communicated by E. J. Hannan

Closed form matrix equations are given for the information matrix of the parameters of the vector mixed autoregressive moving average time series model.

1. INTRODUCTION

Consider the d -dimensional mixed autoregressive moving average process $\{\mathbf{X}(t), t \in Z\}$, Z the set of integers, of order (p, q) ,

$$\sum_{j=0}^p A(j) X(t-j) = \sum_{k=0}^q B(k) \epsilon(t-k), \quad t \in Z$$

where $A(0) \equiv B(0) \equiv I_d$, the d -dimensional identity matrix, and $\{\epsilon(t), t \in Z\}$ is a collection of uncorrelated zero mean d -dimensional random variables each having positive definite covariance matrix Σ .

The estimation of the $(d \times d)$ matrices $A(1), \dots, A(p)$, $B(1), \dots, B(q)$, and Σ has received considerable attention in the recent literature (Akaike [1], Wilson [2], Dunsmuir and Hannan [3], for example). One of the most difficult computational problems involved in the estimation is that of determining the asymptotic covariance matrix V of the maximum likelihood estimators. Hannan [4, pp. 385 and 329], has given expressions for V for the case $p = 0$ or $q = 0$.

In this paper we derive closed form matrix expressions for V by using Whittle's [5] formula for the Fisher information matrix of the parameters of a Gaussian multiple time series.

Received August 29, 1977.

AMS 1970 subject classification: Primary.

Key words and phrases: Multiple mixed time series, information matrices.

* Research supported in part by the Office of Naval Research.

2. THE INFORMATION MATRIX OF THE MIXED PROCESS

Let $\{Y(t), t \in Z\}$ be a zero mean Gaussian time series whose distribution depends on parameters $\theta = (\theta_1, \dots, \theta_r)^T$, where A^T denotes the transpose of the matrix A . Let $f(\omega)$, $\omega \in [-\pi, \pi]$ be the spectral density matrix of $Y(\cdot)$. Then given a sample realization $y(1), \dots, y(t)$ of $Y(\cdot)$, Whittle [5] shows that the maximum likelihood estimators $\hat{\theta}_T$ of θ are such that $T^{1/2}(\hat{\theta}_T - \theta)$ is asymptotically r -dimensional normal with mean zero and covariance matrix $V(\theta) = I^{-1}(\theta)$ where the (j, k) th element of $I(\theta)$ is given by

$$I_{jk}(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \text{Tr} \left[\frac{\partial f(\omega)}{\partial \theta_j} f^{-1}(\omega) \frac{\partial f(\omega)}{\partial \theta_k} f^{-1}(\omega) \right] d\omega, \tag{2.1}$$

where $\partial A/\partial b$ denotes the matrix of scalar derivatives $(\partial A_{jk}/\partial b)$, and $\text{Tr } B$ denotes the trace of the matrix B .

Let $D = (D_{jk})$ be an $(n \times n)$ matrix and $U = U(D)$ and $V = V(D)$ be matrix functions of D . Then (Neudecker [6], for example)

$$\frac{\partial D}{\partial D_{jk}} = E_{jk}, \quad \frac{\partial D^T}{\partial D_{jk}} = E_{jk}^T = E_{kj}, \tag{2.2}$$

$$\frac{\partial UV}{\partial D_{jk}} = \frac{\partial U}{\partial D_{jk}} V + U \frac{\partial V}{\partial D_{jk}}, \tag{2.3}$$

$$\frac{\partial D^{-1}}{\partial a} = -D^{-1} \frac{\partial D}{\partial a} D^{-1}, \tag{2.4}$$

where E_{jk} is the zero matrix except the (j, k) th element which is one.

From (2.4) we can write (2.1) as

$$I_{jk} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \text{Tr} \left[\frac{\partial f^{-1}}{\partial \theta_j} f \frac{\partial f^{-1}}{\partial \theta_k} f \right] d\omega \tag{2.5}$$

$$= -\frac{1}{4\pi} \int_{-\pi}^{\pi} \text{Tr} \left[\frac{\partial f^{-1}}{\partial \theta_j} \frac{\partial f}{\partial \theta_k} \right] d\omega, \tag{2.6}$$

where we have now deleted the argument of $f(\cdot)$ for convenience.

The spectral density matrix $f(\cdot)$ of the mixed model can be written (see Hannan [4, p. 67])

$$f(\omega) = (1/2\pi) G^{-1}(e^{i\omega}) H(e^{i\omega}) \Sigma H^*(e^{i\omega}) G^{-*}(e^{i\omega}), \quad \omega \in [-\pi, \pi] \tag{2.7}$$

where the complex matrix polynomials $G(\cdot)$ and $H(\cdot)$ are given by $G(z) = \sum_{j=0}^p A(j)z^j$, $H(z) = \sum_{k=0}^q B(k)z^k$, and A^* is the complex conjugate transpose of the matrix A . We assume that the zeros of $\det(G(z))$ and $\det(H(z))$ are outside the unit circle so that the elements of $G^{-1}(z)$ and $H^{-1}(z)$ can be written as power series in z .

From (2.7) we write

$$2\pi f = G^{-1}H\Sigma H^*G^{-*} \tag{2.8}$$

$$\equiv G^{-1}QG^{-*},$$

$$(1/2\pi)f^{-1} = G^*Q^{-1}G, \tag{2.9}$$

where G does not involve the $B(\cdot)$, $Q = H\Sigma H^*$ is Hermitian and mathematically independent of the $A(\cdot)$, $Q = \Sigma$ if $q = 0$, and we have deleted the arguments of all functions.

Clearly we must order the elements of the $A(\cdot)$ and $B(\cdot)$ into a vector θ . However, we first find the element of the information matrix corresponding to $A_{jk}(v), A_{lm}(u)$ {denoted $I[A_{jk}(v), A_{lm}(u)]$ } by (2.5), the element corresponding to $B_{jk}(v), B_{lm}(u)$ {denoted $I[B_{jk}(v), B_{lm}(u)]$ } by (2.1), and the element corresponding to $A_{jk}(v), B_{lm}(u)$ {denoted $I[A_{jk}(v), B_{lm}(u)]$ } by (2.6). Then we consider various orderings of the elements of the $A(\cdot)$ and $B(\cdot)$ to find a convenient expression for $I(\theta)$.

2.1. $I[A_{jk}(v), A_{lm}(u)]$

From (2.2), (2.3), and (2.9) we obtain

$$\frac{\partial f^{-1}}{\partial A_{jk}(v)} f \frac{\partial f^{-1}}{\partial A_{lm}(u)} f$$

$$= E_{kj}G^{-*}E_{ml}G^{-*} \exp[-i(u + v)\omega] + 2\pi E_{kj}Q^{-1}E_{lm}f \exp[-i(v - u)\omega]$$

$$+ 2\pi G^*Q^{-1}E_{jk}fE_{ml}G^{-*} \exp[-i(u - v)\omega]$$

$$+ f^{-1}G^{-1}E_{jk}G^{-1}E_{lm}f \exp[i(u + v)\omega].$$

Denote by D^{rs} the matrix D replaced by zeros except for the r th row which is replaced by the s th row, i.e., $D^{rs} = E_{rs}D$. Then $\text{Tr}[E_{rs}DE_{tu}C] = \text{Tr}[D^{rs}D^{tu}] = D_{st}C_{ur}$. Thus

$$\text{Tr} \left[\frac{\partial f^{-1}}{\partial A_{jk}(v)} f \frac{\partial f^{-1}}{\partial A_{lm}(u)} f \right]$$

$$= G_{jm}^{-*}G_{lk}^{-*} \exp[-i(u + v)\omega] + 2\pi Q_{jl}^{-1}f_{mk} \exp[-i(v - u)\omega]$$

$$+ 2\pi Q_{ij}^{-1}f_{km} \exp[-i(u - v)\omega] + G_{mj}^{-1}G_{kl}^{-1} \exp[i(u + v)\omega].$$

We argue that the first and last terms integrate to zero, as follows: because the roots of $\det(G(x)) = 0$ are assumed strictly outside the unit circle, G_{jk}^{-1} , the

(j, k) th element of G^{-1} , can be written as a power series $G_{jk}^{-1}(z) = \sum_{l=0}^{\infty} C_{jk}(l)z^l$, and thus

$$\int_{-\pi}^{\pi} G_{mj}^{-1}G_{kl}^{-1} \exp[i(u + v)\omega] d\omega = \sum_{t=0}^{\infty} \sum_{r=0}^{\infty} C_{mj}(t) C_{kl}(r) \int_{-\pi}^{\pi} \exp[i(u + v + t + r)\omega] d\omega, \tag{2.10}$$

and the integral is always zero since $(u + v + t + r) > 0$.

Thus defining $C = 2\pi E_{kj}Q^{-1}E_{lm}f \exp[-i(v - u)\omega]$, we have

$$I[A_{jk}(v), A_{lm}(u)] = \frac{1}{4\pi} \int_{-\pi}^{\pi} \text{Tr}[C + C^*] d\omega = \int_{-\pi}^{\pi} Q_{jl}^{-1}f_{mk} \exp[-i(v - u)\omega] d\omega. \tag{2.11}$$

2.2. $I[B_{jk}(v), B_{lm}(u)]$

From (2.8) it can be shown in a similar fashion:

$$\text{Tr} \left[\frac{\partial f}{\partial B_{jk}(v)} f^{-1} \frac{\partial f}{\partial B_{lm}(u)} f^{-1} \right] = \text{Tr}[C + C^*] + \text{Tr}[D + D^*],$$

where $C = E_{lm}H^{-1}E_{jk}H^{-1} \exp[i(u+v)\omega]$ and $D = E_{mi}Q^{-1}E_{jk}\mathcal{Z} \exp[-i(u-v)\omega]$. The integral of the first term again vanishes, and we have

$$I[B_{jk}(v), B_{lm}(u)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} Q_{il}^{-1}\mathcal{Z}_{mk} \exp[-i(v - u)\omega] d\omega. \tag{2.12}$$

2.3. $I[A_{jk}(v), B_{lm}(u)]$

To find the off block diagonal elements of I we obtain, using arguments similar to the ones above, that

$$\text{Tr} \left[\frac{\partial f^{-1}}{\partial A_{jk}(v)} \frac{\partial f}{\partial B_{lm}(u)} \right] = \text{Tr}[C + C^*] + \text{Tr}[D + D^*],$$

where $C = E_{lm}H^{-1}E_{jk}G^{-1} \exp[i(u + v)\omega]$ and $D = E_{kj}Q^{-1}E_{lm}\mathcal{Z}H^*G^{-*} \times \exp[-i(v - u)\omega]$. Thus

$$I[A_{jk}(v), B_{lm}(u)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} Q_{jl}^{-1}[\mathcal{Z}H^*G^{-*}]_{mk} \exp[-i(v - u)\omega] d\omega, \tag{2.13}$$

where the (m, k) th element of the product $\mathcal{Z}H^*G^{-*}$ is denoted $[\mathcal{Z}H^*G^{-*}]_{mk}$.

So the integrands in the expressions for $I[A_{jk}(v), A_{lm}(u)]$, $I[B_{jk}(v), B_{lm}(u)]$, and $I[A_{jk}(v), B_{lm}(u)]$ are all of the form $C_{ji}^{-1}D_{mk} \exp[-i(v - u)\omega]$ for some matrices C and D , i.e., we have the transformation of indices denoted

$$(j, k, v), (l, m, u) \rightarrow (j, l, -v), (m, k, u). \tag{2.14}$$

This transformation allows us to find general expressions for the information matrix for various orderings of the elements of the $A(\cdot)$ and $B(\cdot)$ matrices.

Let $\alpha_1 = \text{vec}(A_1)$, $\alpha_2 = \text{vec}(A_1^T)$, $\alpha_3 = \text{vec}(A_2)$, $\beta_1 = \text{vec}(B_1^T)$, $\beta_2 = \text{vec}(B_2^T)$, $\beta_3 = \text{vec}(B_1)$, where $A_1 = (A(1), \dots, A(p))$, $A_2 = (A^T(1), \dots, A^T(p))$, $B_1 = (B(1), \dots, B(q))$, $B_2 = (B^T(1), \dots, B^T(q))$, and $\text{vec}(A) = (\mathbf{a}_1^T, \dots, \mathbf{a}_n^T)^T$, where $\mathbf{a}_1, \dots, \mathbf{a}_n$ are the columns of A .

Let $J_{nm}(\omega) = (J_{jk}(\omega)) = \exp[i(j - k)\omega]$, $j = 1, \dots, n$, $k = 1, \dots, m$.

Denote the information matrix of parameters θ_1, θ_2 by $I^{\theta_1\theta_2}$. Note that $I^{\theta_2\theta_1} = (I^{\theta_1\theta_2})^T$, and that $(C \otimes D)^T = C^T \otimes D^T$, where $C \otimes D$ denotes the Kronecker product $(C_{jk}D)$ of C and D . Then we have from (2.11)–(2.13),

$$\begin{aligned} I^{\alpha_1\alpha_1} &= \int_{-\pi}^{\pi} J_{pp}^T \otimes f^T \otimes Q^{-1} d\omega, & I^{\alpha_2\alpha_2} &= \int_{-\pi}^{\pi} Q^{-T} \otimes J_{pp} \otimes f d\omega, \\ I^{\alpha_3\alpha_3} &= \int_{-\pi}^{\pi} J_{pp} \otimes Q^{-T} \otimes f d\omega, & I^{\beta_1\beta_1} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} J_{qq} \otimes Q^{-T} \otimes \mathcal{Z}^* d\omega, \\ I^{\beta_2\beta_2} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{Z} \otimes J_{qq} \otimes Q^{-T} d\omega, & I^{\beta_3\beta_3} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} J_{qq}^T \otimes \mathcal{Z} \otimes Q^{-1} d\omega, \\ I^{\alpha_1\beta_3} &= -\frac{1}{2\pi} \int_{-\pi}^{\pi} J_{qp}^T \otimes (\mathcal{Z}H^*G^{-*})^T \otimes Q^{-1} d\omega. \end{aligned}$$

2.4. Summary of Results

$$(a) \quad I = \begin{bmatrix} I^{\alpha_1\alpha_1} & I^{\alpha_1\beta_3} \\ \text{sym} & I^{\beta_3\beta_3} \end{bmatrix} = \int_{-\pi}^{\pi} \begin{pmatrix} X \\ Y \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}^* d\omega,$$

where $X = (2\pi)^{-1/2} \mathbf{J}_p \otimes (\mathcal{Z}^{1/2}H^*G^{-*})^T \otimes H^{-*}\mathcal{Z}^{-1/2}$, $Y = -(2\pi)^{-1/2} \mathbf{J}_q \otimes \mathcal{Z}^{1/2} \otimes H^{-*}\mathcal{Z}^{-1/2}$, $\mathcal{Z}^{1/2}$ denotes a positive definite square root of the positive definite matrix \mathcal{Z} , and \mathbf{J}_r is an r -dimensional vector whose k th element is $\exp[-ik\omega]$.

(b) To find the information matrix of the elements of \mathcal{Z} , and the cross information, in turn, between the $A(\cdot)$ and \mathcal{Z} , and the $B(\cdot)$ and \mathcal{Z} , we note that since \mathcal{Z} is symmetric, $\partial f / \partial \mathcal{Z}_{jk} = (1/2\pi) CD_{jk}C^*$, where $C = G^{-1}H$, $D_{jk} = E_{jk} + (1 - \delta_{jk})E_{kj}$, and $\delta_{jk} = 1$ if $j = k$ and 0 otherwise. Thus

$$\text{Tr} \left[\frac{\partial f}{\partial \mathcal{Z}_{jk}} f^{-1} \frac{\partial f}{\partial \mathcal{Z}_{lm}} f^{-1} \right] = \text{Tr} [D_{jk} \mathcal{Z}^{-1} D_{lm} \mathcal{Z}^{-1}],$$

which is independent of ω . Thus the element of the information matrix corresponding to $\mathcal{Z}'_{jk}, \mathcal{Z}'_{lm}$, denoted $I[\mathcal{Z}'_{jk}, \mathcal{Z}'_{lm}]$, is given by

$$I[\mathcal{Z}'_{jk}, \mathcal{Z}'_{lm}] = (1/2)[(2 - \delta_{jk} - \delta_{lm} + \delta_{jk}\delta_{lm}) \mathcal{Z}'^{-1}_{jl} \mathcal{Z}'^{-1}_{mk} + (2 - \delta_{jk} - \delta_{lm}) \mathcal{Z}'^{-1}_{kl} \mathcal{Z}'^{-1}_{mj}], \quad (2.15)$$

where we consider $\mathcal{Z}'_{u,v}$, $u \geq v = 1, \dots, d$ as the distinct elements of \mathcal{Z}' . To write this in matrix form, define $\sigma_1 = \text{vec}(\mathcal{Z}')$, $\sigma_2 = l \text{vec}(\mathcal{Z}')$, where $l \text{vec}(A)$ is the vec operation on the lower triangular portion of the matrix A . Thus σ_2 contains the distinct elements of \mathcal{Z}' .

If we treat the d^2 elements of \mathcal{Z}' as being distinct (i.e., ignore the symmetry of \mathcal{Z}'), then $D_{jk} = E_{jk}$, and it is easy to show that $I^{\sigma_1 \sigma_1} = \frac{1}{2} \mathcal{Z}'^{-1} \otimes \mathcal{Z}'^{-1}$. Note that we can write $\sigma_1 = A \sigma_2$, where A is a $d^2 \times \frac{1}{2}d(d+1)$ matrix of zeros and ones. Then (Pagano [7])

$$\begin{aligned} I^{\sigma_2 \sigma_2} &= A^T I^{\sigma_1 \sigma_1} A \\ &= \frac{1}{2} A^T (\mathcal{Z}'^{-1} \otimes \mathcal{Z}'^{-1}) A. \end{aligned} \quad (2.16)$$

Thus the information matrix of σ_1 can be found from either (2.15) or (2.16).

(c) One can also show that $I[A_{jk}(v), \mathcal{Z}'_{lm}] = 0$ and $I[B_{jk}(v), \mathcal{Z}'_{lm}] = 0$ for the same reason that the integrals vanish in the derivation of the information matrices of the $A(\cdot)$ and the $B(\cdot)$.

(d) If $q = 0$ we have $I^{\alpha_1 \alpha_1} = \Gamma_p \otimes \mathcal{Z}'^{-1}$ and $I^{\alpha_2 \alpha_2} = \mathcal{Z}'^{-1} \otimes \Gamma_p$, since

$$\begin{aligned} \Gamma_p &\equiv \text{BTOEPL}(R(0), R(-1), \dots, R(1-p)) \\ &= \int_{-\pi}^{\pi} J_{pp} \otimes f \, d\omega, \end{aligned}$$

where $\text{BTOEPL}(R(0), \dots, R(1-p))$ is a block Toeplitz matrix having $R(j-k)$ in the j th row and k th column of blocks, $j, k = 1, \dots, p$. Note that $I^{\alpha_2 \alpha_2}$ is the result given by Hannan [4, p. 329].

(e) If $p = 0$, then $Q^{-T} = (1/2\pi) f^{-T}$ which is of the form of an autoregressive spectral density matrix. Thus $I^{\beta_1 \beta_1}$ should be of the same form as $I^{\alpha_1 \alpha_1}$ of the previous note with the autocovariances corresponding to the spectral density Q^{-T} replacing $R(\cdot)$ and $2\pi \mathcal{Z}'^{-1}$ replacing \mathcal{Z}' . Thus $I^{\beta_1 \beta_1} = \Gamma_i \otimes \mathcal{Z}'$ where $\Gamma_i = \text{BTOEPL}(Ri(0), \dots, Ri(1-q))$, and $Ri(v) = (1/4\pi^2) \int_{-\pi}^{\pi} f^{-T}(\omega) \exp(iv\omega) \, d\omega$, $v \in \mathbb{Z}$. See Newton [8], p. 11, *et seq.* for a discussion of the inverse autocovariances $Ri(\cdot)$. Note that $I^{\beta_1 \beta_1}$ is the result stated by Hannan [4], p. 385.

(f) $I^{\alpha_2 \alpha_2}$ and $I^{\beta_1 \beta_1}$ also provide an interpretation of information matrices of random matrices (transformed into a random vector by the vec operator) as the covariance matrix of some random vector. We can write

$$I^{\alpha_1 \alpha_1} = \{R(j-k) \otimes \mathcal{Z}'^{-1}\}_{p \times p}, \quad I^{\beta_1 \beta_1} = \{Ri(j-k) \otimes \mathcal{Z}'^{-1}\}_{q \times q},$$

i.e., $I^{\alpha_1\alpha_1}$ is a $(p \times p)$ block matrix of $(d^2 \times d^2)$ blocks, while $I^{\beta_1\beta_1}$ is a $(q \times q)$ block matrix of $(d^2 \times d^2)$ blocks. Thus there exist d^2 -dimensional random vectors $\mathbf{z}_1, \dots, \mathbf{z}_p$ and $\mathbf{y}_1, \dots, \mathbf{y}_q$ such that $\mathbf{z} = (\mathbf{z}_1^T, \dots, \mathbf{z}_p^T)^T$ has a covariance matrix given by $I^{\alpha_1\alpha_1}$ and $\mathbf{y} = (\mathbf{y}_1^T, \dots, \mathbf{y}_q^T)^T$ has a covariance matrix given by $I^{\beta_1\beta_1}$.

(g) Akaike [1] derives approximate expressions for the elements of the Hessian of the log likelihood function. From formulas (2.11)–(2.13), it is clear that Akaike's formulas for the Hessian are the sample analogues of the information matrix of the mixed scheme parameters. Thus one can use the block Toeplitz matrix inversion techniques developed by Akaike [9] to find the asymptotic covariance matrix of the maximum likelihood estimators.

ACKNOWLEDGMENTS

The author would like to thank Professor E. Parzen and Professor M. Pagano for many helpful discussions.

REFERENCES

- [1] AKAIKE, H. (1973a). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika* **60** 255–265.
- [2] WILSON, G. T. (1973). The estimation of parameters in multivariate time series models. *J. Royal Statist. Soc. B.* **35** 76–85.
- [3] DUNSMIUR, W. AND HANNAN, E. J. (1976). Vector linear time series models. *Adv. Appl. Prob.* **8** 339–364.
- [4] HANNAN, E. J. (1970). *Multiple Time Series*. John Wiley, New York.
- [5] WHITTLE, P. (1953). The analysis of multiple stationary time series. *J. Royal Statist. Soc. B.* **15** 125–139.
- [6] NEUDECKER, H. (1969). Some theorems on matrix differentiation with special reference to Kronecker matrix products. *J. Amer. Statist. Assoc.* **64** 953–964.
- [7] PAGANO, M. (1974). *Modification of parameter estimates under linear constraints*. Technical Report Ho. 12, Statistical Science Division, SUNY at Buffalo.
- [8] NEWTON, H. J. (1975). *The efficient estimation of stationary multiple time series mixed models: theory and algorithms*. Technical Report No. 33, Statistical Science Division, SUNY at Buffalo.
- [9] AKAIKE, H. (1973b). Block Toeplitz matrix inversion. *SIAM J. Appl. Math.* **24** 234–241.
- [10] RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, (2nd ed.). John Wiley, New York.