Available online at www.sciencedirect.com
ScienceDirect
Statistics and Probability Letters 78 (2008) 1430–1433
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

The asymptotic and exact Fisher information matrices of a vector ARMA process

André Klein[a,*], Guy Mélard[b], Abdessamad Saidi[c]

[a] *Department of Quantitative Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands*
[b] *ECARES CP 114, Université Libre de Bruxelles, Avenue Franklin Roosevelt 50, B-1050 Bruxelles, Belgium*
[c] *Département de mathématiques et de statistique, Université de Montréal, CP 6128, Succursale A, Montréal, Québec, H3C 3J7, Canada*

Received 12 July 2006; accepted 11 December 2007
Available online 28 December 2007

Abstract

The exact Fisher information matrix of a Gaussian vector autoregressive-moving average (VARMA) process has been considered for a time series of length N in relation to the exact maximum likelihood estimation method. In this paper it is shown that the Gaussian exact Fisher information matrix converges to the asymptotic Fisher information matrix when N goes to infinity.
© 2008 Published by Elsevier B.V.

MSC: 62M10; 62F12; 62B10

1. Introduction

The exact Fisher information matrix of a Gaussian vector autoregressive-moving average (VARMA) process has been considered for a time series of length N in relation to the exact maximum likelihood estimation method. In this paper it is shown that the Gaussian exact Fisher information matrix converges to the asymptotic Fisher information matrix when N goes to infinity.

Several recent papers have discussed either the asymptotic Fisher information matrix (e.g. Godolphin and Bane (2005)) or the exact Fisher information matrix (e.g. Terceiro (2000)) but we have seen no indication of the result mentioned in the previous paragraph. Only Zadrozny (1989, 1992) mentions the two information matrices, exact and asymptotic, but we could not see a convergence between the two expressions. On the contrary, the asymptotic Fisher information is defined as the limit of the exact Fisher information.

Consider $\{y_t, t \in \mathbb{Z}\}$, $\mathbb{Z}$ the set of integers, a Gaussian vector autoregressive-moving average (VARMA) process of order (p, q) in dimension n, which satisfies the vector difference equation

$$\sum_{j=0}^{p} \alpha_j y_{t-j} = \sum_{j=0}^{q} \beta_j \varepsilon_{t-j}, \quad t \in \mathbb{Z} \tag{1}$$

* Corresponding author. Tel.: +31 20 5254245; fax: +31 205254349.
E-mail addresses: A.A.B.Klein@uva.nl (A. Klein), gmelard@ulb.ac.be (G. Mélard), saidiabd@dms.umontreal.ca (A. Saidi).

0167-7152/$ - see front matter © 2008 Published by Elsevier B.V.
doi:10.1016/j.spl.2007.12.013

where $\{\varepsilon_t, t \in \mathbb{Z}\}$ is the innovation process, a sequence of independent zero-mean n-dimensional random variables each having positive definite covariance matrix Σ, and where $\alpha_j, \beta_j \in \mathbb{R}^{n\times n}$ are the parameter matrices, and $\alpha_0 \equiv \beta_0 \equiv I_n$.

We use L to denote the backward shift operator on $\mathbb{Z}$, for example $Ly_t = y_{t-1}$; then (1) can be written as

$$\alpha(L)y_t = \beta(L)\varepsilon_t \tag{2}$$

where

$$\alpha(z) = \sum_{j=0}^{p} \alpha_j z^j, \qquad \beta(z) = \sum_{j=0}^{q} \beta_j z^j$$

are the associated matrix polynomials. We further assume the eigenvalues of the matrix polynomials $\alpha(z)$ and $\beta(z)$ to be outside the unit circle so the elements of $\alpha^{-1}(z)$ and $\beta^{-1}(z)$ can be written as power series in z. These eigenvalues are obtained by solving the scalar polynomials $\det(\alpha(z)) = 0$ and $\det(\beta(z)) = 0$, where $\det(X)$ is the determinant of X. We assume that the matrix polynomials $\alpha(z)$ and $\beta(z)$ have no common eigenvalues so that non-singularity of Fisher's information matrix is guaranteed (e.g. Klein et al. (2005)).

Let $\{y_t, t = 1,\ldots,N\}$ be a time series generated by the VARMA process (2) and let the set of parameters $\vartheta = (\vartheta_1,\ldots,\vartheta_\ell)^\top$, where $\top$ denotes transposition and $\ell = n^2(p+q)$. The following definition of the parameter vector ϑ is introduced: $\vartheta = \operatorname{vec}\left\{\alpha_1, \alpha_2, \ldots, \alpha_p, \beta_1, \beta_2, \ldots, \beta_q\right\}$, where vec X as usual stands for the vector resulting from stacking the columns of a matrix X on top of each other. We assume, as usual, that the nuisance parameters included in Σ are functionally independent from the parameters of interest included in ϑ.

2. State space form and Fisher information matrices

Although it is not strictly needed, the exact Fisher information of a Gaussian process is often introduced using a state space representation (e.g. Hannan and Deistler (1988)) using a vector of the state variables $x_t \in \mathbb{R}^m$, $t \in \mathbb{N}$. Among other possibilities, using a specific basis in the state space, the following state space structure is considered:

$$x_{t+1} = \phi x_t + F\varepsilon_t \tag{3}$$

$$y_t = Hx_t + \varepsilon_t, \tag{4}$$

where $\varepsilon_t \in \mathbb{R}^n$ is a Gaussian white noise process with $\mathbb{E}(\varepsilon_t) = 0$, $\mathbb{E}\left(\varepsilon_t\varepsilon_t^\top\right) = \Sigma > 0$, and

$$\phi = \begin{pmatrix} -\alpha_1 & I_n & & 0_n \\ -\alpha_2 & 0_n & \ddots & \\ \vdots & & \ddots & I_n \\ -\alpha_h & 0_n & \cdots & 0_n \end{pmatrix}, \qquad F = \begin{pmatrix} \beta_1 - \alpha_1 \\ \beta_2 - \alpha_2 \\ \vdots \\ \beta_h - \alpha_h \end{pmatrix}, \quad \text{and} \quad H^\top = \begin{pmatrix} I_n \\ 0_n \\ \vdots \\ 0_n \end{pmatrix}, \tag{5}$$

and $h = \max(p,q)$, $\alpha_i = 0_n$, $i > p$, $\beta_i = 0_n$, $i > q$, and consequently $m = hn$.

In Klein and Neudecker (2000) an appropriate representation at the vector–matrix level for the exact Fisher information matrix $\mathcal{J}_N(\vartheta)$ is set out. It is based on the multivariate version of minus the logarithm of the likelihood of the system described by (3) and (4) and is given by

$$l(\vartheta) = -\log L(\vartheta) = \sum_{t=1}^{N} \left\{ \frac{n}{2}\log 2\pi + \frac{1}{2}\log\det B_t + \frac{1}{2}\widetilde{y}_t^\top B_t^{-1}\widetilde{y}_t \right\},$$

where $\widetilde{y}_t$ and B_t are defined below. The exact information matrix is then

$$N\mathcal{J}_N(\vartheta) = \mathbb{E}\frac{\partial^2 l(\vartheta)}{\partial\vartheta\,\partial\vartheta^\top},$$

for obtaining

$$N\mathcal{J}_N(\vartheta) = \sum_{t=1}^{N} \left[\frac{1}{2}\left(\frac{\partial \operatorname{vec} B_t}{\partial\vartheta^\top}\right)^\top (B_t \otimes B_t)^{-1} \left(\frac{\partial \operatorname{vec} B_t}{\partial\vartheta^\top}\right) + \mathbb{E}\left\{ \left(\frac{\partial\widetilde{y}_t}{\partial\vartheta^\top}\right)^\top B_t^{-1} \left(\frac{\partial\widetilde{y}_t}{\partial\vartheta^\top}\right) \right\} \right]. \tag{6}$$