

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I



www.sciencedirect.com

Statistics A multivariate generalized long memory model

Processus longue mémoire généralisés multivariés

Abdou Kâ Diongue

LERSTAD, université Gaston-Berger, UFR SAT, BP 234, Saint-Louis, Senegal

ARTICLE INFO

Article history: Received 2 March 2009 Accepted 2 February 2010 Available online 2 March 2010

Presented by Paul Deheuvels

ABSTRACT

In this Note, we propose a new flexible multivariate long memory process which is a selfsimilar model with the ability to capture short-range dependence, seasonality and longrange dependence characteristics. Specifically, we extend the multivariate ARFIMA model proposed by Sowell (1989) [8], and investigate some of its statistical properties.

© 2010 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

RÉSUMÉ

Dans cette Note, nous proposons une extension des processus longue mémoire multivariés VARFIMA permettant de modéliser à la fois la dépendence à mémoire courte, la saisonalité et la dependence à mémoire longue. Nous étudions quelques propriétés statistiques du modèle.

© 2010 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

In recent years, the importance of long memory processes has been realized in numerous applications, especially financial and electricity data (e.g., Norrbin and Smallwood [6], and Diongue et al. [2], among others). In these papers, the observations at each time point are assumed to be independent, and the dependence extends only along the time axis.

However, in some situations such as interconnected European electricity markets, when individuals at each point belong to the same family or share the same common environment, the observations on these individuals at any given time point would generally be correlated. In actual fact, there are a few studies in the integration or inter-relationship between time points as well as within the group of individuals at each point (e.g., Dueker and Startz [3], Diongue [1], Tsay [9], among others).

The aim of this Note is to introduce a new model that is the multivariate generalized long memory and to derive matrix expression for the theoretical spectral density function. Parameter estimation methods are also discussed.

2. Model and main results

The objective of this Note is to propose a general class of long memory processes in order to combine strands of long memory and seasonality literature. Let *k* be a nonnegative integer and $X_t = \{X_{ts}, t \in (0, \pm 1, \pm 2, ...)\}$ for s = 1, 2, ..., k be a multivariate generalized long memory process, denoted VGARMA, specified by

$$\Phi(L)\Delta^{d,\nu}(L)X_t = \Theta(L)\varepsilon_t,$$

(1)

E-mail address: abdou.diongue@gmail.com.

¹⁶³¹⁻⁰⁷³X/\$ – see front matter © 2010 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved. doi:10.1016/j.crma.2010.02.001

where $\Delta^{d,\nu}(L) = \text{diag}[(1 - 2\nu_1L + L^2)^{d_1}, \dots, (1 - 2\nu_kL + L^2)^{d_k}]$, a diagonal $(k \times k)$ matrix containing the Gegenbauer polynomial for each series with $d = (d_1, \dots, d_k)'$, $\nu = (\nu_1, \dots, \nu_k)'$, where c' being the transpose of the vector c, and L is the backwards shift operator. Here, $\Phi(L) = I_k - \Phi_1 L - \dots - \Phi_p L^p$ and $\Theta(L) = I_k - \Theta_1 L - \dots - \Theta_q L^q$ are matrix polynomials in the lag operator L of degrees p and q respectively, with Φ_i , $i = 1, \dots, p$, and Θ_j , $j = 1, \dots, q$, each $(k \times k)$ matrices, I_k is a $(k \times k)$ identity matrix, and $\varepsilon_t = (\varepsilon_{t1}, \dots, \varepsilon_{tk})$ is a k-dimensional independent and identically distributed white noise process with mean 0 and a nonsingular covariance matrix Σ . Notice that this model is a direct generalization of the VARFIMA process introduced by Sowell [8] when ignoring the presence of the constant $(k \times k)$ nonsingular matrix V which allows to address the estimation problems associated with the linear combination of nonstationary process.

In the following proposition, we provide conditions under which the multivariate generalized long memory process $\{X_{ts}, t \in (0, \pm 1, \pm 2, ...)\}$, for s = 1, 2, ..., k has a stationary and invertible solution:

Proposition 2.1. Assume that the polynomials det $[\Phi(L)]$ and det $[\Theta(L)]$ have all their roots outside the unit circle. Thus,

- (i) X_t is stationary if $d_i < 1/2$ when $v_i < 1$ or $d_i < 1/4$ when $v_i = 1$ for i = 1, ..., k;
- (ii) X_t possess an invertible moving average representation if $d_i > -1/2$ when $v_i < 1$ or $d_i > -1/4$ when $v_i = 1$ for i = 1, ..., k.

To proof Proposition 2.1, remark that the simple form of the Gegenbauer matrix, $\Delta^{d,\nu}(L)$, means that the stationarity and invertibility properties of the Gegenbauer X_t vector series can be obtained by the univariate proofs applied element by element.

We will now investigate the analytical expression of the spectral density function of X_t . For this purpose, let U_t be the process defined by $U_t = \Delta^{d,\nu}(L)X_t$. Thus, the vector time series U_t is generated by the ARMA(p,q) process $\Phi(L)U_t = \Theta(L)\varepsilon_t$. The Wold representation of U_t is

$$U_t = \Phi(L)^{-1} \Theta(L) \varepsilon_t = \frac{B(L)}{A(L)} \varepsilon_t$$

where B(L) is a matrix polynomial of order $M \leq (k-1)p+q$ and A(L) is a scalar polynomial of order $H \leq kp$. If we assume that the roots of $A(\xi)$ are outside the unit circle, then $A(\xi)$ can be written as

$$A(\xi) = \prod_{n=1}^{H} (1 - \rho_n \xi)^{-1}, \text{ where } |\rho_n| < 1 \text{ for } n = 1, \dots, H.$$

Lemma 2.2. Assume that the roots of $A(\xi)$ are all unique. The spectral density function of U_t , denoted $f_U(\lambda)$, can be written with representative (i, j) element notation $f_U(\lambda) = [f_U(\lambda)_{i,j}]$, where

$$f_{U}(\lambda)_{i,j} = \sum_{l=-M}^{M} \sum_{m=1}^{H} \psi_{i,j}(l) \omega^{H+l} \zeta_m \bigg[\frac{\rho_m^{2H}}{(1-\rho_m \omega)} - \frac{1}{(1-\rho_m^{-1} \omega)} \bigg],$$
(2)

with $\omega = e^{-i\lambda}$, and the coefficients ζ_i and $\psi_{i,j}(l)$ are given by

$$\zeta_{j} = \frac{1}{[\rho_{j} \prod_{i=1}^{H} (1 - \rho_{i} \rho_{j}) \prod_{k=1, k \neq j}^{H} (\rho_{j} - \rho_{k})]},$$

and

$$\psi_{i,j}(l) = \sum_{h=1}^{k} \sum_{t=1}^{k} \sum_{s=n_1}^{n_2} \sigma_{ht} B_{i,h}(s) B_{j,t}(s-l),$$

with $n_1 = \max(0, l)$, and $n_2 = \min(M, M - l)$, if σ_{ij} and $B_{i,j}(\omega) = \sum_{n=0}^{M} B_{i,j}(n)\omega^n$ are the (i, j) element of Σ and $B(\omega)$, respectively.

Proof. The spectral density of U_t is given by

$$f_U(\lambda) = B(\omega) \Sigma B(\omega^{-1})' \prod_{n=1}^{H} (1 - \rho_n \omega)^{-1} (1 - \rho_n \omega^{-1})^{-1}.$$

Thus, considering the hypothesis and using the partial decomposition of the product, it is straightforward to obtain the result. For more details, we refer to Sowell [8]. \Box

Theorem 2.3. If we assume the hypothesis in Lemma 2.2 is hold, then the spectral density function of X_t , denoted by $f_X(\lambda)$, can be expressed with representative (r, s) element as $f_X(\lambda) = [f_X(\lambda)_{r,s}]$ where

$$f_X(\lambda)_{r,s} = 2^{-d_r - d_s} (-1)^{d_r + d_s} (\cos \lambda_r - \cos \lambda)^{-d_r} (\cos \lambda_s - \cos \lambda)^{-d_s} \times \exp(i\lambda(d_r - d_s)) f_U(\lambda)_{r,s}, \quad \text{with } \lambda_r = \arccos v_r.$$
(3)

Sketch of proof. The spectral density of X_t can be written as

$$f_X(\lambda) = \Delta^{d,\nu}(\omega)^{-1} f_U(\lambda) \left[\Delta^{d,\nu} \left(\omega^{-1} \right)^{-1} \right]'.$$

$$\tag{4}$$

Thus, the (r, s) element of $f_X(\lambda)$ is given by $f_X(\lambda)_{r,s} = f_U(\lambda)_{r,s}(1 - 2\nu_r\omega + \omega^2)^{-d_r}(1 - 2\nu_s\omega^{-1} + \omega^{-2})^{-d_s}$. Finally, using the facts that:

a)
$$(1 - 2\nu z + z^2) = (1 - e^{i\lambda}z)(1 - e^{-i\lambda}z)$$
 with $\nu = \cos(\lambda)$; and
b) $1 - e^{i\lambda} = 2\sin(\frac{\lambda}{2})\exp(i\frac{\lambda - \pi}{2})$,

the result is obtained.

We propose in this study two approaches for parameter estimation: the Whittle-type quasi-maximum likelihood method and the fast likelihood approximation method of Luceño [5]. Given data $X_1, ..., X_n$, along with parameters $\Psi = (\Phi, \Theta, d, \nu, \Sigma)$, the Whittle quasi-log-likelihood function for the VGARMA(p, d, ν, q) model can be written as

$$\ell_f(X,\Psi) = -\frac{1}{2} \sum_{j=1}^{n-1} \log \det[f_X(\lambda_j)] - \pi \operatorname{tr}\left[\sum_{j=1}^{n-1} f_X(\lambda_j)^{-1} I(\lambda_j)\right],\tag{5}$$

where $f_X(\lambda)$ is the spectral density function of X_t , and $I(\lambda)$ is the sample spectrum of X_t defined by

$$I(\lambda) = \frac{1}{2n\pi} \left[\sum_{t=1}^{n} X_t e^{-it\lambda} \right] \left[\sum_{t=1}^{n} X'_t e^{it\lambda} \right].$$

Notice that the highest dimension of matrices involved in the spectral log-likelihood calculation is $(k \times k)$ which makes the optimization practically manageable. Following Hosoya [4, Theorems 1 and 2], we can obtain the theoretical properties (consistency and asymptotic normality) of Whittle-type quasi-maximum likelihood estimators for multivariate GARMA processes.

Let $adj(\Theta(L))$ be the adjoint of a matrix $\Theta(L)$ and $\widehat{\Phi(L)} = adj\{\Theta(L)\}\Phi(L)$ be an autoregressive operator of order p^* . Following Luceño [5], the quasi-maximum likelihood function for the VGARMA(p, d, ν, q) may be approximated by

$$L(X,\Psi) = (2\pi)^{-\frac{nk}{2}} \det(\Omega)^{-1} \times \exp\left[\frac{1}{2} \operatorname{tr}(\delta_0^{(p^*,d,\nu,q)} P_0) + \sum_{i=1}^N \operatorname{tr}(\delta_i^{(p^*,d,\nu,q)} P_i)\right],\tag{6}$$

where $\Omega = \operatorname{cov}(X')$. In (6), the P_i 's satisfy

$$P_i = \sum_{t=1-M}^{M-1+M} \hat{X}_{t+i} \hat{X}$$

with $N \leq n - 1 + 2M$, $M \leq n$, and \hat{X}'_t is the observed series for t = 1, ..., n and the forecasts for t < 1 or t > n. The coefficients $\delta_i^{(p^*,d,\nu,q)}$'s are given by

$$\delta_i^{(p^*,d,\nu,q)} = \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \alpha_m \alpha_j \delta_{i+j-m}^{(p^*,d,\nu,0)}, \quad \text{where the } \alpha_j \text{'s are such that } \sum_{j=0}^{\infty} \alpha_j L^j = \frac{1}{\det[\Theta(L)]}.$$
(7)

The coefficients $\delta_i^{(p^*,d,\nu,0)}$ verify

$$\delta_{i}^{(p^{*},d,\nu,0)} = \sum_{j=0}^{p^{*}} \sum_{m=0}^{p^{*}} \delta_{i+j-m}^{(d,\nu)} \big(\tilde{\Phi}_{m}^{\prime} \Sigma^{-1} \tilde{\Phi}_{j} \big),$$

with

$$\begin{split} \delta_i^{(d,\nu)}(A) &= \sum_{j=0}^{\infty} \left(\pi_{i+j}^{(d,\nu)} \right)' A \pi_j^{(d,\nu)}, \\ \pi_j^{(d,\nu)} &= \text{diag} \Big[\pi_j^{(d_1,\nu_1)}, \dots, \pi_j^{(d_k,\nu_k)} \Big], \\ \pi_j^{(d_i,\nu_i)} &= \sum_{m=0}^{\left\lfloor \frac{j}{2} \right\rfloor} \frac{(-1)^m \Gamma(-d_i+j-m)(2\nu_i)^{j-2m}}{\Gamma(-d)\Gamma(m+1)\Gamma(j-2m+1)} \end{split}$$

and [j/2] is the integer part of j/2.

Remark that, for the evaluation of the time domain quasi-log-likelihood function, we need to obtain the determinant and the inverse of the covariance matrix by crude numerical methods which cost much computation time (see, So and Kwok, [7]). \Box

References

- A.K. Diongue, Modélisation longue mémoire multivariée : applications aux problématiques du producteur d'EDF dans le cadre de la libéralisation du marché européen de l'électricité, Ph.D. thesis, ENS Cachan, France, 2005.
- [2] A.K. Diongue, D. Guégan, B. Vignal, Forecasting electricity spot market prices with a k-factor GIGARCH process, Applied Energy 86 (2009) 505-510.
- [3] M. Dueker, R. Startz, Maximum likelihood estimation of fractionally cointegration with an application to U.S. and Canadian Bond Rates, Review of Economics and Statistics 80 (1998) 420–426.
- [4] Y. Hosoya, The quasi-maximum likelihood approach to statistical inference on multiple time series with long range dependence, Journal of Econometrics 71 (1996) 217–236.
- [5] A. Luceňo, A fast likelihood approximation for vector general linear processes with long series: Application to fractional differencing, Biometrika 83 (3) (1996) 603–614.
- [6] S.C. Norrbin, A.D. Smallwood, Generalized long memory and mean reversion of the real exchange rate, Applied Economics (2008) 1-10, iFirst.
- [7] M.K.P. So, S.W.Y. Kwok, A multivariate long memory stochastic volatility model, Physica A 362 (2006) 450-464.
- [8] F. Sowell, Maximum likelihood estimation of fractionally integrated time series models, Working Paper, Carnegie-Mellon University, 1989.
- [9] W.-J. Tsay, Maximum likelihood estimation of stationary multivariate ARFIMA processes, Journal of Statistical Computation and Simulation (2009), forthcoming.