Inference of Multivariate Continuous-time Long Memory Processes

Henghsiu Tsai[a], Heiko Rachinger[b], and Kung-Sik Chan[c]
[a]Academia Sinica, [b]University of Vienna, and [c]University of Iowa

Abstract

Continuous-time long-memory models have found diverse applications in many fields, including option pricing, volatility modeling, environmental study, and annual tree-ring measurements, among many others. Marquardt (2007) introduces a class of multivariate continuous-time autoregressive fractionally integrated moving average (MCARFIMA) models without discussing the estimation of the model parameters and thus without rendering it applicable to real data. Alternative continous time models exhibiting long range dependence either do not model the long-range dependence explicitly or assume all processes to have a common Hurst parameter.

In this paper, we develop the missing estimation theory of the MCARFIMA models with different long memory parameters. The MCARFIMA models are useful for analyzing multivariate discrete-time long memory data sampled regularly or irregularly. In contrast to a discrete-time ARFIMA process being defined as a solution of a difference equation, a d-dimensional MCARFIMA process is defined as the solution of a p-th order stochastic equation with suitable initial condition and driven by a vector of independent fractional Brownian motions with different Hurst parameters. We derive the spectral density matrix function of the model and show that the continuous-time model given discrete-time regularly spaced data is identifiable under certain regularity conditions. The spectral density function enables us to consider the estimation of the MCARFIMA Models with discrete-time regularly spaced data by maximizing the Whittle likelihood. We show that the spectral maximum Whittle likelihood estimator (SMLE) is asymptotically normal and efficient. Finite-sample properties of the SMLE are studied by simulations. We illustrate the method with a real application.

Keywords: Asymptotic normality; Multivariate time series; Spectral maximum likelihood estimator; Whittle likelihood

Address for correspondence: Henghsiu Tsai, Institute of Statistical Science, Academia Sinica, Taipei, Taiwan 115, Republic of China.

E-mail: htsai@stat.sinica.edu.tw

1 Introduction

Continuous-time long-memory models have found diverse applications in many fields, including option pricing (Comte and Renault, 1998), volatility modeling (Casas and Gao, 2008), environmental study (Tsai and Chan, 2005a), and annual tree-ring measurements (Tsai and Chan, 2005b), among many others. For further developments of univariate continuous-time long-memory models, see, for example, Chambers (1996), Comte (1996), Comte and Renault (1996), and Brockwell and Marquardt (2005). In order to study interactions and comovements among a group of time series variables, one needs to consider multivariate time series models. There are three extensions of univariate continuous-time long-memory models to multivariate continuous-time long-memory models. Marquardt (2007) introduces a class of multivariate fractionally integrated CARMA processes and studies their probabilistic properties, however, without discussing the estimation. Barndorff-Nielsen and Stelzer (2011) propose multivariate supOU (superpositions of Ornstein-Uhlenbeck-type) processes which can exhibit long-range dependence and suggest moments based estimation methods for estimating the parameters. The finite- and large- sample properties of the estimator are, however, unknown. Asai and McAleer (2013) propose a fractionally integrated Wishart stochastic volatility model with a common long memory parameter for multivariate stochastic volatility modeling. In this paper, we develop the missing estimation theory of the MCARFIMA models with different Hurst parameters. The MCARFIMA models are useful for analyzing multivariate discrete-time long memory data sampled regularly or irregularly.

The rest of the paper is organized as follows. The multivariate continuous-time fractionally integrated ARMA processes are described in Section 2. Spectral maximum likelihood estimator (SMLE) and its large sample properties are discussed in Section 3. In Section 4, we report some empirical performance of the SMLE. We illustrate the use of the MCARFIMA model with a real application in Section 5. Section 6 concludes.

2 Multivariate continuous-time fractionally integrated ARMA processes

We first introduce some notations. Let the indicator function of a set B, denoted by $I_B(\cdot)$, defined to be one if the argument lies in B and zero otherwise. The real numbers and the integers are denoted by $\mathbb{R}$ and $\mathbb{Z}$, respectively. The ring of polynomial expressions in z over a ring $\mathbb{K}$ is denoted by $\mathbb{K}[z]$. The symbols $M_{m,n}(\mathbb{K})$, or $M_n(\mathbb{K})$ if $m = n$, stand for the space of $m \times n$ matrices with entries in $\mathbb{K}$.

Heuristically, a d-dimensional MFCARFIMA(p, H, q) process $\{Y(t)\}$ is defined as the solution of a p-th order stochastic differential equation with suitable initial condition and driven by a vector of independent standard fractional Brownian motions with Hurst parameters $H = (H_1, \cdots, H_m)'$, where m is a positive integer.

Specifically, for $t \geq 0$,

$$P(D)Y(t) = Q(D)D\bar{B}^H(t), \tag{1}$$

where $D = d/dt$, $\bar{B}^H(t) = [B^{H_1}(t), \cdots, B^{H_m}(t)]'$, the superscript $'$ denotes the transpose, and for $1 \leq k \leq m$, $\{B^{H_k}(t), t \geq 0\}$ is a standard Brownian motion with Hurst parameter $0 < H_k < 1$, and $\{B^{H_k}(t), t \geq 0\}$, $k = 1, ..., m$, are m independent stochastic processes; $P(z) = z^p + A_1 z^{p-1} + \cdots + A_p \in M_d(\mathbb{R}[z])$, $Q(z) = B_0 + B_1 z + \cdots + B_q z^q \in M_{d,m}(\mathbb{R}[z])$. The fractional Brownian motion is nowhere differentiable (Mandelbrot and Van Ness, 1968), so the stochastic equation (1) has to be appropriately interpreted as some integral equation as explained below. Analogous to the case of univariate continuous-time ARMA processes (see, e.g., Brockwell, 1993), equation (1) can be equivalently cast in terms of the *state* and *observation* equations:

$$dX(t) = \bar{A}X(t) + \beta d\bar{B}^H(t), \quad (2)$$

$$Y(t) = CX(t), \quad (3)$$

where

$$\bar{A} = \begin{pmatrix} 0 & I_d & 0 & \cdots & 0 \\ 0 & 0 & I_d & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_d \\ -A_p & -A_{p-1} & -A_{p-2} & \cdots & -A_1 \end{pmatrix} \in M_{pd}(\mathbb{R}),$$

where I_d is the $d \times d$ identity matrix, $\beta = (\beta_1', \cdots, \beta_p') \in M_{pd,m}(\mathbb{R})$, $\beta_{p-j} = -I_{0,\dots,q}(j)\left[\sum_{i=1}^{p-j-1} A_i\beta_{p-j-i} - B_j\right]$, and $C = (I_d, 0, ..., 0) \in M_{d,pd}(\mathbb{R})$. Under the condition that all the eigenvalues of A have strictly negative real parts, the solution of (2) can be written as

$$X_t = e^{\bar{A}t}X_0 + \int_0^t e^{\bar{A}(t-u)}\beta dB^H(t),$$

where $e^{\bar{A}t} = I_{dp} + \sum_{n=1}^{\infty}\left\{(\bar{A}t)^n(n!)^{-1}\right\}$.

The spectral density matrix of $\{Y(t)\}$ is given in Theorem 1.

Theorem 1 *The spectral density matrix of $\{Y(t), t \geq 0\}$ equals*

$$f_Y(\omega) = \frac{1}{2\pi}P^{-1}(i\omega)Q(i\omega)\Sigma\ diag(D_1(\omega), \cdots, D_m(\omega))Q(-i\omega)'\{P^{-1}(-iw)\}' \quad (4)$$

where $D_j(\omega) = \Gamma(2H_j + 1)\sin(\pi H_j)|w|^{1-2H_j}$, for $j = 1, \cdots, m$, and $\Gamma()$ is the Gamma function.

Remark 1: note that in Theorem 1, $D_j(\omega) = 1$ if $H_j = 1/2$.
Remark 2: the multivariate CARFIMA(p, H, q) defined for $\{t \geq 0\}$ can be extended to be well-defined for $\{t \in \mathbb{R}\}$. The extension is similar to that of Tsai (2009) in the univariate case.

One major problem with continuous-time modeling is the identifiability of the continuous-time model, given discrete-time data. Let $\theta = (H_1, .., H_m, A_1, .., A_p, B_0.., B_q)$, and $\{Y(ih)\}_{i=1,\dots,N}$ be the observations sampled from a stationary MCARFIMA(p, H, q)

process, where h is the step size. By the aliasing formula (Priestley, 1981), the spectral density matrix of $\{Y(ih)\}_{i=1,\ldots,N}$ equals

$$f_h(\omega;\theta) = \frac{1}{h}\sum_{k\in Z} f_Y\left(\frac{\omega+2k\pi}{h}\right), \qquad \omega \in [-\pi,\pi],$$

where $f_Y(\cdot)$ is as defined in Equation (4). Using the frequency domain method, Tsai and Chan (2005b) showed that the univariate CARFIMA(p,H,q) model with $1/2 < H < 1$ is identifiable. For the multivariate case, we have similar identification results in the following theorem.

Theorem 2 *Let $Y = \{Y(ih)\}_{i=1}^N$ be sampled from a stationary (Gaussian) MCARFIMA(p,H,q) process given by Equation (1),* $\det P(z)$ *and* $\det Q(z)$ *have no common zeros, all roots of* $\det P(z) = 0$*, and the roots of* $\det Q(z) = 0$ *have strictly negative real parts. If $h > 0$, then for $\theta_1 \neq \theta_2$, the set $\{\omega | f_h(\omega;\theta_1) \neq f_h(\omega;\theta_2)\}$ has positive Lebesgue measure.*

We note that the roots of the determinant of the polynomial $P(z)$ are the same as the eigenvalues of the matrix $\bar{A}$, and the condition on the roots of the determinant of the polynomial $P(z)$ is necessary for the stationarity of the process, whereas the condition on the determinant of the polynomial $P(z)$ is akin to the invertibility condition for discrete-time processes.

3 Spectral maximum likelihood estimator and its large sample properties

Let $I_Y(\omega) = J_Y(\omega)J_Y(\omega)^*/(2\pi N)$, where $J_Y(\omega) = \sum_{t=1}^N Y_t e^{it\omega}$, $J(\omega)^*$ denotes the conjugate transpose of $J(\omega)$. Let $\mathrm{tr}(A)$ be the trace of the matrix A, $\omega_j := 2\pi j/N \in (0,\pi)$ the Fourier frequencies, and T be the largest integer $\leq (N-1)/2$. Then the (negative) log-likelihood function of $\{Y(ih)\}$ can be approximated, up to a multiplicative constant, by the (negative) Whittle log-likelihood function (see Hosoya, 1996)

$$-\tilde{l}(\theta) \quad = \quad \sum_{i=1}^{T}\left[\log\det f_h(\omega_i;\theta) + \mathrm{tr}\{f_h(\omega_i;\theta)^{-1} I_Y(\omega_i)\}\right] \qquad (5)$$

The objective function (5) is minimized with respect to θ to get the spectral maximum likelihood estimator (SMLE) $\hat{\theta}$.

Theorem 3 *Let the data $Y = \{Y(ih)\}_{i=1}^N$ be sampled from a stationary Gaussian long-memory process given by (1), where all roots of* $\det P(z) = 0$, *and the roots of* $\det Q(z) = 0$ *have strictly negative real parts. Let the spectral maximum likelihood estimator $\hat{\theta} \in \Theta$, a compact parameter space, and the true parameter θ_0 be in the interior of the parameter space. Then $\sqrt{N}(\hat{\theta} - \theta_0)$ converges in distribution to a normal random vector with mean 0 and covariance matrix $\Gamma(\theta_0)^{-1}$, where the (i, j)-th element of $\Gamma(\theta)$ is given by*

$$\Gamma_{ij}(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} tr \left[f_h(\omega;\theta)^{-1} \frac{\partial f_h(\omega;\theta)}{\partial \theta_i} f_h(\omega;\theta)^{-1} \frac{\partial f_h(\omega;\theta)}{\partial \theta_j} \right] d\omega.$$

4 Simulation

5 Real Application

6 Conclusion

Acknowledgements

Henghsiu Tsai thanks Academia Sinica, the National Science Council (NSC 98-2118-M-001-023-MY2), R.O.C., and Kung-Sik Chan thanks U.S. National Science Foundation (DMS-1021292), for partial support. Part of this paper was written while Heiko Rachinger was visiting Academia Sinica.

References

[1] Asai, M. and McAleer, M. (2013). A fractionally integrated Wishart stochastic volatility model. Available at `http://eprints.ucm.es/18068/`

[2] Barndorff-Nielsen, O. E. and Stelzer, R. (2011). Multivariate supOU processes. *The Annals of Applied Probability* **21**, 140V182.

[3] Brockwell, P. J. and Marquardt, T. (2005). Lévy-driven and fractionally integrated ARMA processes with continuous time parameter. *Statist. Sinica* **15** 477-494.

[4] Casas, I. and Gao, J. (2008). Econometric estimation in long-range dependent volatility models: theory and practice. *Journal of Econometrics* **147** 72-83.

[5] Chambers, M. J. (1996). The estimation of continuous parameter long-memory time series models. *Econometric Theory* **12** 374-390.

[6] Comte, F. (1996). Simulation and estimation of long memory continuous time models. *Journal of Time Series Analysis* **17**, 19-36.

[7] Comte, F. and Renault, E. (1996). Long memory continuous time models. *Journal of Econometrics* **73**, 101-149.

[8] Comte, F. and Renault, E. (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance* **8** 291-323.

[9] Hosoya, Y. (1996). The Quasi-likelihood approach to statistical inference on multiple time-series with long-range dependence. *Journal of Econometrics* 73 217-236.

[10] Mandelbrot, B.B. and Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review* **10** 422-437.

[11] Marquardt, T. (2007). Multivariate fractionally integrated CARMA processes. *Journal of Multivariate Analysis*. **98** 1705V1725.

[12] Tsai, H, (2009). On continuous-time autoregressive fractionally integrated moving average processes. *Bernoulli* **15** 178-194.

[13] Tsai, H. and Chan, K. S. (2005a). Maximum likelihood estimation of linear continuous time long memory processes with discrete time data. *J. Roy. Statist. Soc. Ser. B* **67** 703-716.

[14] Tsai, H. and Chan, K. S. (2005b). Quasi-maximum likelihood estimation for a class of continuous-time long-memory processes. *Journal of Time Series Analysis* **26** 691-713.