Journal of Multivariate Analysis 150 (2016) 229–244

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Model identification using the Efficient Determination Criterion

Paulo Angelo Alves Resende *, Chang Chung Yu Dorea

Department of Mathematics, University of Brasília, Brasília 70910-900, Distrito Federal, Brazil

ARTICLE INFO

Article history:
Received 15 October 2014
Available online 28 June 2016

AMS 2010 subject classifications:
62M05
62F12
60J10

Keywords:
EDC
BIC
AIC
Order estimation
BEKK-GARCH

ABSTRACT

In the realm of the model selection context, Akaike's and Schwarz's information criteria, AIC and BIC, have been applied successfully for decades for model order identification. The Efficient Determination Criterion (EDC) is a generalization of these criteria, proposed originally to define a strongly consistent class of estimators for the dependency order of a multiple Markov chain. In this work, the EDC is generalized to partially nested models, which encompass many other order identification problems. Based on some assumptions, a class of strongly consistent estimators is established in this general environment. This framework is applied to BEKK multivariate GARCH models and, in particular, the strong consistency of the order estimator based on BIC is established for these models.

1. Introduction

Zhao et al. [24], on estimating the order of multiple Markov chains, introduced the Efficient Determination Criterion (EDC), which allows for adjustments on the penalty term used in the criteria AIC [1] and BIC [22]. Also, a class of strongly consistent estimators was established in the same work. Afterwards, Dorea [9] extended this class and proposed the asymptotic optimal order estimator, which had its better performance verified by the extensive use of numerical simulations [21].

In this work, the concept of "nested models" is generalized to a class of partially nested models and the EDC criterion is extended to this new context. Some results regarding the consistency of EDC order estimators are established based essentially on assumptions about the likelihood function. This approach is applied to state the consistency of the BIC order estimator for BEKK multivariate GARCH models, which encompass the univariate version GARCH as a particular case.

To exemplify the "partially nesting" concept, consider that $\mathbb{X} = \{X_t\}_{t\in\mathbb{N}}$ is an AR(k) process and $\mathbb{Y} = \{Y_t\}_{t\in\mathbb{N}}$ is an ARMA(k_1, k_2) process, i.e.,

$$X_t = \varepsilon_t + \sum_{i=1}^{k} \alpha_i X_{t-i}$$

and

$$Y_t = \varepsilon_t + \sum_{i=1}^{k_1} \alpha_i Y_{t-i} + \sum_{i=1}^{k_2} \beta_i \varepsilon_{t-i},$$

* Corresponding author.
E-mail addresses: pa@pauloangelo.com (P.A.A. Resende), changdorea@unb.br (C.C.Y. Dorea).

http://dx.doi.org/10.1016/j.jmva.2016.06.002

where $\{\varepsilon_t\}$ are i.i.d. $\mathcal{N}(0, 1)$ and $\alpha_i, \beta_i \in \mathbb{R}$. If $\mathbb{X}$ is an AR(1), it can be represented as an AR(2) taking $\alpha_2 = 0$. In this case, we say that the "AR(1) model is nested in AR(2)", denoted by AR(1) $\subseteq$ AR(2). Also, for any $k, p \in \mathbb{N}$, AR(k) $\subseteq$ AR(p) or AR(p) $\subseteq$ AR(k) and we may consider $\{\mathrm{AR}(k)\}_{k\in\mathbb{N}}$ as a sequence of nested models. Similarly, we can define ARMA(1, 1) $\subseteq$ ARMA(2, 2). However ARMA(1, 2) and ARMA(2, 1) are not "related". For such a case, we define the class of partially nested models $\{\mathrm{ARMA}(k_1, k_2)\}_{(k_1,k_2)\in\mathbb{N}^2}$, where we can define the order estimator based on the EDC criterion.

Section 2 provides the general results, that may be applied in a variety of models to establish the EDC order estimators. Section 3 presents the approach applied to BEKK multivariate GARCH models. The proofs of the stated results are in Appendices.

2. General framework

The essence of nested models have been used since the pioneer researches using hypothesis tests. However, practically all works focused on particular cases and the formal definition and treatment of the concept of nested models were unused. Nishii [20] firstly proposed a general estimator for the dimension of i.i.d. models. A relevant piece of Nishii's technique is adapted to our purposes.

For an arbitrary time discrete stochastic process $\mathbb{X} = \{X_t\}_{t\in\mathbb{N}}$, $E \subseteq \mathbb{R}^p$ the set of possible values of X_t and ν a fixed measure on E, we define a family of statistical models for $\mathbb{X}$ as

$$M = \left\{f(x_1^n, \theta, n) : \ \theta \in \Theta, \ n \geq 1\right\}$$

where $f(x_1^n, \theta, n)$ represents the set of possible densities for $X_1^n := (X_1, \ldots, X_n)$ with respect to the product measure on E^n, which depends on the parameter $\theta \in \Theta \subseteq \mathbb{R}^d$. We may denote $f(x_1^n, \theta) = f(x_1^n, \theta, n)$ to simplify the notation.

Two statistical models

$$M_k = \left\{f(x_1^n, \theta, n) : \ \theta \in \Theta_k, \ n \geq 1\right\} \quad \text{and}$$

$$M_p = \left\{f(x_1^n, \theta, n) : \ \theta \in \Theta_p, \ n \geq 1\right\}$$

are nested, denoted by $M_k \subseteq M_p$, if $\Theta_k \subseteq \Theta_p$ and, for all $\theta \in \Theta_k$, $x_1^\infty \in E^\infty$, there exists $c \in (0, \infty)$ such that

$$\lim_{n\to\infty} \frac{f_k(x_1^n, \theta)}{f_p(x_1^n, \theta)} = c.$$

For $q \in \mathbb{N}$, $p = (p_1, \ldots, p_q) \in \mathbb{N}^q$ and $k = (k_1, \ldots, k_q) \in \mathbb{N}^q$, we define the usual order relation $p \geq k$ iff $p_i \geq k_i$ for $i = 1 \ldots q$, which makes $(\mathbb{N}, \geq)$ a partially ordered set. By $p \ngeq k$ we mean that $p < k$ or p and k are not related. The set $\mathbb{M} = \{M_k\}_{k\in\mathbb{N}^q}$ is a class of partially nested models if

$$M_k \subseteq M_p \Leftrightarrow k \leq p.$$

We say that an element $m_r \in \bigcup_{k\in\mathbb{N}^q} M_k$ has order $r \in \mathbb{N}^q$ if $m_r \in M_r$ and $m_r \in M_k$ implies that $M_r \subseteq M_k$. Analogously, a process $\mathbb{X}$ has order r if, $\forall n > 0$, the densities of X_1^n are elements in M_r and if the densities of X_1^n are represented by elements in M_k, then $M_r \subseteq M_k$. In this context, we denote a sample by x_1^n and the maximum likelihood estimator (MLE) of θ supposing the order k by $\hat{\theta}_k$. The EDC criterion was originally established using the likelihood functions. However, we observed that the estimator's consistency depends just on a few properties of the likelihood functions. Therefore, in what follows we consider $\left\{L_{n,k}(x_1^n, \theta)\right\}_{(n,k)\in\mathbb{N}\times\mathbb{N}^q}$ as a class of functions $L_{n,k} : E^n \times \Theta_k \to \mathbb{R}$ that satisfies

$$L_{n,k}(x_1^n, \hat{\theta}_k) = \sup_{\theta\in\Theta_k} \left\{L_{n,k}(x_1^n, \theta)\right\} \tag{1}$$

and for $\theta \in \Theta_k$ and $p \geq k$,

$$L_{n,p}(x_1^n, \theta) \geq L_{n,k}(x_1^n, \theta) \quad \text{and} \quad \lim_{n\to\infty} \frac{L_{n,p}(x_1^n, \theta)}{L_{n,k}(x_1^n, \theta)} < \infty. \tag{2}$$

To simplify notation, we shall denote $L_{n,k}(\theta) = L_{n,k}(x_1^n, \theta)$. In most situation, the $L_{n,k}$ functions are merely the likelihood for each n and k. However, this generalization enables the use of the EDC in atypical situations. Now, we define the EDC estimator for a class of partially nested models.

Definition 1. Let $\mathbb{X}$ be a process of order r and $\mathbb{M}$ its respective class of partially nested models. The EDC estimator for r is defined by

$$\hat{r} = \underset{k\leq K}{\operatorname{argmin}} \ \{\mathrm{EDC}(k)\} \tag{3}$$

for

$$\mathrm{EDC}(k) = -\ln L_{n,k}(\hat{\theta}_k) + c_n\gamma(k),$$

c_n a sequence of positive numbers, $\gamma(k) = \dim(\Theta_k)$ and K a known constant greater than r.