Journal of the Korean Statistical Society 45 (2016) 114–122

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

CGMM LASSO-type estimator for the process of Ornstein–Uhlenbeck type

Yinfeng Wang [a], Yanlin Tang [b], Xinsheng Zhang [c,*]

[a] School of Business, Shanghai Dianji University, Shanghai, 201306, China
[b] Department of Mathematics, Tongji University, Shanghai, 200092, China
[c] Department of Statistics, Fudan University, Shanghai, 200433, China

ARTICLE INFO

Article history:
Received 27 January 2015
Accepted 13 August 2015
Available online 3 September 2015

AMS 2000 subject classifications:
62G20
62M05

Keywords:
LASSO-type estimator
GMM with continuum of moment method (CGMM) estimator
The process of Ornstein–Uhlenbeck type

ABSTRACT

In this paper, we study the LASSO-type penalized CGMM (GMM with continuum of moment method) estimator for the process of Ornstein–Uhlenbeck type. This LASSO-type estimator is obtained by minimizing the summation of the CGMM object function and a LASSO-type penalty, which is included for model selection. In the proposed method, model selection and estimation are done simultaneously. Under some regularity conditions, the proposed estimator asymptotically follows a non-standard normal distribution (Caner, 2009). Simulation study shows that the proposed estimator correctly selects the true model much more frequently than the commonly used Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC).

1. Introduction

Model selection is an important issue in applied econometric analysis, because correct model selection is crucial in the subsequent step of estimation. In Caner (2009), model selection can rise to non-nested models with two subsets of parameters and two sets of orthogonality restrictions, then one can construct a single large parametric model and select one of the two models. Among the various methods for variable selection, LASSO (least absolute shrinkage and selection operator, Tibshirani, 1996) is well studied and very easy to implement, see for examples, Efron, Hastie, Johnstone, and Tibshirani (2004), Leng, Lin, and Wahba (2006), Yuan and Lin (2007), etc. In recent years, the LASSO procedure has also been studied in time series analysis by several authors, mainly in the case of autoregressive models. For example, Wang, Li, and Tsai (2007) considered the problem of shrinkage estimation of regressive and autoregressive coefficients, Hsu, Hung, and Chang (2008) studied vector autoregression (VAR) case. However, it is known that LASSO leads to asymptotic bias.

Knight and Fu (2000) proposed a LASSO-type penalized estimator under the framework of ordinary least squares regression for linear model, named L_q LASSO, in which they penalize $|\theta|^q$ with $0 < q < 1$. Under some regularity conditions, Knight and Fu (2000) showed that L_q LASSO is consistent in parameter estimation; furthermore, under some additional conditions, L_q LASSO enjoys oracle property (see Fan & Li, 2001). Caner (2009) extended the L_q LASSO to a generalized method of moments-based (GMM-based) LASSO estimator (GMM L_q LASSO for short) for time series data, and also verified the oracle property of this estimator. Nardi and Rinaldo (2011) considered penalized order selection in an AR(p) model.

* Corresponding author.
E-mail address: xszhang@fudan.edu.cn (X. Zhang).

http://dx.doi.org/10.1016/j.jkss.2015.08.002
1226-3192/

More recently, Gregorio and Iacus (2012) considered quasi-likelihood-based adaptive LASSO-type estimator for multivariate diffusion processes.

Recently, the processes of Ornstein–Uhlenbeck (OU for short) type is becoming more and more popular. The processes of OU type are constructed by a given correlation structure and a stationary marginal distribution, and has wide applications in finance and econometrics, see for examples, Barndorff-Nielsen (1998, 2001), Barndorff-Nielsen, Jensen, and Sørensen (1998) and Barndorff-Nielsen and Shephard (2001, 2003). In the context of the processes of OU type, the density or distribution function is usually unknown. Carrasco, Chernov, Florens, and Ghysels (2007) proposed CGMM (GMM with Continuum of Moment Conditions) method for parameter estimation in Markov processes. In this paper, we propose L_q LASSO penalized estimator for the multidimensional processes of OU type, to select the correct structure, under the CGMM framework. Under some regularity conditions, we prove that the proposed penalized estimator is consistent in estimation and enjoys the oracle property.

This rest of the paper is organized as follows. Section 2 introduces the d-dimensional process of OU type; Section 3 introduces the problem of CGMM-LASSO type estimation for the process of OU type; Section 4 gives the consistency and the oracle property of L_q LASSO estimator. Simulation study is carried out in Section 5 for the process of *IG*-OU type and d-dimensional process of *NIG*-OU type. Proofs are collected in the Appendix.

2. Preliminaries of the d-dimensional process of OU type

Given a d-dimensional time-homogeneous Lévy process $Z = \{Z_t,\ t \geq 0\}$ starting from the origin and a $d \times d$ matrix $\mathcal{Q}$, the d-dimensional process of OU type $X = \{X_t,\ t \geq 0\}$ driven by Z is defined by

$$X_t = e^{-t\mathcal{Q}}X_0 + \int_0^t e^{-(t-s)\mathcal{Q}}dZ_s, \quad t \in \mathbb{R}_+, \tag{2.1}$$

where X_0 is independent from Z. The process of OU type is equivalently defined as the unique strong solution of the stochastic differential equation

$$dX_t = -\mathcal{Q}X_t dt + dZ_t. \tag{2.2}$$

Under some regularity conditions on $\mathcal{Q}$ and the Lévy measure of Z, X admits a unique invariant distribution F, and the class of all possible Fs forms the class of all $\mathcal{Q}$-self-decomposable distributions, see Jacod (1985) and Wolfe (1982) for the one-dimensional case and Masuda (2004) for multidimensional case.

In this paper, we focus on the process of OU type which satisfies the following assumptions.

- (a) The existence of a smooth transition density follows from two assumptions of the Lévy measure ν of Z near the origin, even if the Gaussian covariance matrix A of Z degenerates. Assumptions of ν are: 1. $E[\log\{\max(1, |Z_1|)\} < \infty]$; 2. for any $u \in \mathbb{R}^d$ satisfying $|u| \geq 1$, there exist constants $\alpha > 2$ and $c > 0$ such that

  $$\int_{\{z:|u'z|\leq 1\}} |u'z|^2 \nu(dz) \geq c|u|^{2-\alpha},$$

  respectively. Here, $|A| = \{trace(A'A)\}^{1/2}$ for any matrix (or vector) A, with $'$ denoting transpose.

  Due to the Markovian nature of X, the β-mixing coefficient $\beta_X(t)$ of X is given by

$$\beta_X(t) = \int \|P(t, x, \cdot) - F(\cdot)\|_{TV} F(dx),$$

where $\|\cdot\|_{TV}$ stands for the total variation norm, $P(t, x, \cdot)$ is the transaction probability of X and $F(\cdot)$ is the marginal selfcomposable distribution of X (see, Page 107 in Masuda (2004)).

The following lemma is the Theorem 4.3 of Masuda (2004), which gives the exponential β-mixing bound condition of X.

Lemma 1. *Let X be the strictly stationary process of OU type given by (2.2) with a selfdecomposable marginal distribution F. If we have*

$$\int |x|^p F(dx) < \infty$$

for some $p > 0$, then there exists a constant $a > 0$, such that $\beta_X(t) = O(e^{-at})$ as $t \to \infty$. In particular, X is ergodic.

- (b) In the strictly stationary case, we assume that the process of OU type X has the exponential β-mixing bound.
  In Definition 11.9, Sato (1999), $\{Z_t, t \geq 0\}$ is of type A, if the triplet of Z_t, (A, ν, γ) satisfies: $A = 0$ and $\nu(\mathbb{R}) < \infty$.
- (c) Lévy process $\{Z_t,\ t \geq 0\}$ is of type A, and Lévy measure ν has continuous derivatives at least up to the second order with respect to θ in $\aleph_1$, where θ is the parameter vector of $X = \{X_t,\ t \geq 0\}$, $\aleph_1$ is a compact subset of Θ and contains the true parameter vector θ_0. $\int_{\mathbf{R}^d} \nu''(x)dx < \infty$, where ν'' is the second derivative of ν with respect to θ.