Journal of Multivariate Analysis 145 (2016) 132–147

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

On oracle property and asymptotic validity of Bayesian generalized method of moments

Cheng Li [a,*], Wenxin Jiang [b]

[a] Department of Statistical Science, Duke University, Box 90251, Durham NC 27708, United States

[b] Department of Statistics, Northwestern University, 2006 Sheridan Road, Evanston IL 60208, United States

ARTICLE INFO

Article history:
Received 30 January 2015
Available online 29 December 2015

AMS subject classifications:
62F15
62F12

Keywords:
Bayesian
GEE (generalized estimating equations)
GMM (generalized method of moments)
MCMC
Model selection
Moment condition
Oracle property
Posterior validity

ABSTRACT

Statistical inference based on moment conditions and estimating equations is of substantial interest when it is difficult to specify a full probabilistic model. We propose a Bayesian flavored model selection framework based on (quasi-)posterior probabilities from the Bayesian Generalized Method of Moments (BGMM), which allows us to incorporate two important advantages of a Bayesian approach: the expressiveness of posterior distributions and the convenient computational method of Markov Chain Monte Carlo (MCMC). Theoretically we show that BGMM can achieve the posterior consistency for selecting the unknown true model, and that it possesses a Bayesian version of the oracle property, i.e. the posterior distribution for the parameter of interest is asymptotically normal and is as informative as if the true model were known. In addition, we show that the proposed quasi-posterior is valid to be interpreted as an approximate posterior distribution given a data summary. Our applications include modeling of correlated data, quantile regression, and graphical models based on partial correlations. We demonstrate the implementation of the BGMM model selection through numerical examples.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

We consider the estimation problem based on the following unconditional moment restrictions

$$\mathrm{E}\{g(D,\theta)\}=0 \tag{1}$$

where D is a set of random variables with domain $\mathcal{D}$, θ is a p-dimensional vector of parameters to be estimated, and g is a m-dimensional mapping from $\mathcal{D}\times\mathbb{R}^p$ to $\mathbb{R}^m$. Typically it is necessary to have $m\ge p$ for the point identification of θ. Given an i.i.d. or stationary realization $\boldsymbol{D}=\{D_1,\ldots,D_n\}$ of D, one can estimate θ directly from such a set of m moment functions, without needing to fully specify the underlying data generating process of D. In this paper, we consider the case where in (1), the true parameter θ_0 could possibly lie in a lower dimensional subspace. Our goal is to consistently select the relevant variables and estimate their effects, namely the nonzero components of θ_0, when the specification of full probabilistic model is unavailable but a sufficient number of moment conditions are present.

We consider a Bayesian-flavored approach, where a quasi-posterior can be derived from a prior distribution and a quadratic form of moment restrictions. This enables us to accommodate two important advantages of the Bayesian approach: the expressiveness of the posterior distributions and the convenient computational method of MCMC. These are particularly

* Corresponding author.
E-mail addresses: cl332@stat.duke.edu (C. Li), wjiang@northwestern.edu (W. Jiang).

http://dx.doi.org/10.1016/j.jmva.2015.12.009
0047-259X/© 2015 Elsevier Inc. All rights reserved.

useful for the model selection problem that we study. We are able to report the most probable model, the second most probable model and so on, together with their quasi-posterior probabilities, which are shown to be asymptotically valid in large samples. We can also use the reversible jump MCMC algorithm [15,9] to traverse the space of different models and simulate the quasi-posterior probabilities.

For this framework of moment-based Bayesian method of model selection and model averaging, our paper will prove several appealing fundamental theorems. They will address model selection consistency, oracle property, and valid interpretation of the quasi-posterior distribution. In the following, we will first review the related works and then describe in detail the contributions of our current paper.

1.1. GMM and BGMM

The moment based estimation problem (1) is important and has been extensively studied in econometrics and statistics. Well known methods include the generalized method of moments (GMM, [17,18,41]), the empirical likelihood (EL, [43,44]), the exponential tilting (ET, [30]), the exponential tilted empirical likelihood (ETEL, [46,47]) and the generalized empirical likelihood (GEL, [42]). Essentially they all share the same first order efficiency of optimally weighted GMM estimator, and have been applied to independent data, time series data and panel data in econometrics. On the other hand, researchers in statistics also use the moment based methods for constructing efficient estimators, especially for clustered and correlated longitudinal data. For example, Qu et al. [45] proposed a GMM type estimator to avoid the inefficiency from misspecified working correlation matrices in generalized estimating equations (GEE) for longitudinal data. Wang et al. [50] considered the EL approach to address the within-subject correlation structure. Recently frequentist penalization methods have been proposed to accommodate increasing dimension p. See for example [51,34,8,4], etc. In general, the moment based estimation methods only require information on the low order moments of D and are therefore more flexible, efficient and robust to model misspecification, as long as the moment conditions are correctly specified.

Our work focuses on the Bayesian inference of θ under the moment constraint (1). Compared to the abundance of frequentist literature, the development of Bayesian methods on this problem still remains limited. One difficulty that hinders the fully probabilistic Bayesian modeling is that some prior distribution on both the distribution of D (denoted as P_D) and the parameter θ needs to be specified, such that the pair (P_D, θ) satisfies the set of restrictions (1). Recent progress in this direction includes Kitamura and Otsu [29] and Florens and Simoni [13]. Kitamura and Otsu [29] tried to minimize the Kullback–Leibler divergence of P_D to a Dirichlet process, which leads to an ET type likelihood function that computationally requires optimizations within each MCMC iteration step. Florens and Simoni [13] exploited the Gaussian process prior and required a functional transformation of the data that is only asymptotically Gaussian, which still leads to a misspecified likelihood function in finite samples. Besides, both methods have only been tested on simple examples that involve a few parameters and moments. Instead, another analytically simpler Bayesian way of modeling (1) is *the Bayesian generalized method of moments* (BGMM), first proposed and studied by Kim [27] and Chernozhukov and Hong [6], which constructs the simple quasi-likelihood function

$$q(\boldsymbol{D}|\theta) = \frac{1}{\det(2\pi \boldsymbol{V}_n/n)^{\frac{1}{2}}} \exp\left\{-\frac{n}{2}\bar{g}(\boldsymbol{D},\theta)^\top \boldsymbol{V}_n^{-1}\bar{g}(\boldsymbol{D},\theta)\right\}, \tag{2}$$

where $\bar{g}(\boldsymbol{D},\theta)$ is the sample average of $g(D_i,\theta)$, $i = 1,\ldots,n$, $\boldsymbol{V}_n$ is a $m \times m$ positive definite matrix that could possibly depend on the data $\boldsymbol{D}$, and $\det(\boldsymbol{A})$ denotes the determinant of a matrix $\boldsymbol{A}$. Hereafter we use the symbol "q" to denote the quasi-likelihood function and the quasi-posterior. This quasi-likelihood function has been studied under a Bayesian framework in [27] and is named *the limited information likelihood (LIL)*, which minimizes the Kullback–Leibler divergence of the true data generating process P_D to the set of all distributions satisfying the less restrictive asymptotic constraint $\lim_{n\to\infty} \mathrm{E}\left\{n\bar{g}(\boldsymbol{D},\theta_0)^\top \boldsymbol{V}_n^{-1}\bar{g}(\boldsymbol{D},\theta_0)\right\}/m = 1$. This relation holds when we choose $\boldsymbol{V}_n$ to be a consistent estimator of the covariance matrix $\mathrm{Var}(g(D,\theta_0))$. Given a prior distribution $\pi(\theta)$, the quasi-posterior takes the form

$$q(\theta|\boldsymbol{D}) \propto \frac{1}{\det(2\pi \boldsymbol{V}_n/n)^{\frac{1}{2}}} \exp\left\{-\frac{n}{2}\bar{g}(\boldsymbol{D},\theta)^\top \boldsymbol{V}_n^{-1}\bar{g}(\boldsymbol{D},\theta)\right\}\pi(\theta). \tag{3}$$

By using $q(\boldsymbol{D}|\theta)$ in the Bayesian model, we only need to specify a prior on θ and thus circumvent the difficulty of directly assigning a prior on the pair (P_D,θ) with constraints (1). In the computational aspect, $q(\boldsymbol{D}|\theta)$ takes an explicit analytical form that allows straightforward MCMC updating for the corresponding Bayesian posterior without any iterative optimization steps [6]. Furthermore, when $\boldsymbol{V}_n$ is chosen as a consistent estimator of $\mathrm{Var}(g(D,\theta_0))$, the exponential part of $q(\boldsymbol{D}|\theta)$ resembles the optimally weighted GMM criterion function [17], which in large samples can be viewed as a second order approximation to the true negative log-likelihood function that follows a chi-square distribution with p degrees of freedom if $m = p$ and both are fixed [52].

The theoretical properties of BGMM have been investigated extensively in Chernozhukov and Hong [6] and Belloni and Chernozhukov [2], who show that a Bernstein–von Mises theorem holds, i.e. the posterior distribution converges asymptotically to normal. The computational aspects of BGMM with no model selection have been investigated in [52,53]. Kim [28] has established the pairwise consistency theoretically when each candidate model is compared to the true model separately, and has used MCMC in simulations for such model comparison. Hong and Preston [19] have discussed a more

Download English Version:

https://daneshyari.com/en/article/1145194

Download Persian Version:

https://daneshyari.com/article/1145194

Daneshyari.com