Journal of Computational and Applied Mathematics 302 (2016) 24–37

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Smooth-threshold estimating equations for varying coefficient partially nonlinear models based on orthogonality-projection method

Jing Yang *, Hu Yang

College of Mathematics and Statistics, Chongqing University, Chongqing 401331, China

ARTICLE INFO

Article history:
Received 27 July 2015
Received in revised form 8 December 2015

Keywords:
Estimating equations
Orthogonality-projection
Varying coefficient partially nonlinear models
Variable selection
Oracle property

ABSTRACT

In this paper, a two-stage estimation procedure is proposed for varying coefficient partially nonlinear models, in which the estimates of parametric vector and coefficient functions do not affect each other. Specifically, we first employ an orthogonality-projection-based method for the estimation of parametric coefficient by removing out the varying coefficient parts. In the second stage, we approximate each coefficient function via B-spline basis functions and develop a novel variable selection procedure based on smooth-threshold estimating equations. The proposed procedure can automatically eliminate the irrelevant covariates by setting the corresponding coefficient functions as zero, and simultaneously estimate the nonzero regression components. Besides, this approach not only avoids solving a convex optimization problem that is required in previous variable selection procedures, but also is flexible and easy to implement. Under some mild conditions, the asymptotic theories of the generated estimates are established, including model selection consistency, rate of convergence as well as asymptotic distribution. Finally, some numerical simulations are conducted to examine the finite sample performance of the proposed methodologies, and a real data analysis is followed to further illustrate the application of the methods.

1. Introduction

As one of the most important semiparametric models, varying coefficient partially linear models (VCPLMs) have been widely studied by many statisticians. It can be seen as a linear regression model that some coefficients of the predictor variables are assumed to be constant while others are assumed to vary with another element. Related literature can be referred to [1–5]. However, this linear relationship between the response variable and covariate variable may not be sufficient in many practical situations, and some more general underlying relationship such as nonlinear dependencies should be taken into consideration. Therefore, in this paper, we focus our attention on varying coefficient partially nonlinear models (VCPNMs), which was first introduced by Li and Mei [6]. Specifically, let Y be the univariate response variable, $\mathbf{X} = (X_1, X_2, \ldots, X_p)^T, \mathbf{Z} = (Z_1, Z_2, \ldots, Z_q)^T$ and U be the corresponding covariates, a varying coefficient partially nonlinear model is of the following form

$$Y = \mathbf{X}^T\boldsymbol{\alpha}(U) + g(\mathbf{Z}, \boldsymbol{\beta}) + \varepsilon, \quad (1)$$

* Corresponding author.
E-mail address: yang2009jing@163.com (J. Yang).

http://dx.doi.org/10.1016/j.cam.2016.01.038

where "T" denotes the transpose of a vector or matrix throughout this paper, $\boldsymbol{\alpha}(\cdot) = (\alpha_1(\cdot), \alpha_2(\cdot), \ldots, \alpha_p(\cdot))^T$ is a p-dimensional vector consisting of unknown coefficient functions, $g(\cdot, \cdot)$ is a known nonlinear function, and $\boldsymbol{\beta}$ is a vector of unknown coefficients that do not necessarily have the same dimension with $\mathbf{Z}$. The random error term ε is independent of $(\mathbf{X}, \mathbf{Z}, U)$ with mean zero and finite variance σ^2. Obviously, this model preserves the flexibility of varying coefficient models and easy interpretation of nonlinear model.

In the last couple of decades, variable selection has been an important topic in all regression analysis and many procedures have been developed for this. Generally speaking, most of the variable selection procedures are based on penalized estimation using penalty functions. For instance, Frank and Friedman [7] considered the L_q penalty, which yields a "Bridge regression". Tibshirani [8] proposed the Lasso penalty, which can be viewed as a solution to the penalized least squares with the L_1 penalty. Fan and Li [9] developed the SCAD penalty and proved that this variable selection procedure enjoys the oracle property. Zou [10] proposed the adaptive Lasso penalty and also demonstrated its oracle property. Wang et al. [11] considered one-step estimator for ultrahigh dimensionality linear regression model with various penalty functions. Note that all above methods are based on penalized estimation procedures by using a penalty function that has a singularity at zero. Therefore, these variable selection procedures require convex optimization, which will incur a computational burden. To this end, Ueki [12] developed a new variable selection procedure named the smooth-threshold estimating equations that can automatically eliminate irrelevant parameters by setting them as zero and simultaneously estimate the nonzero regression components in linear regression model. Lai et al. [13] and Lai et al. [14] further extended this procedure to single-index models and partially linear single-index model, respectively. Li et al. [15] developed a smooth-threshold generalized estimating equations in generalized linear models with longitudinal data. Tian et al. [16] applied this technique for estimation and variable selection in VCPLM.

Note that for model (1), there exists little literature on its statistical inference up to now. Li and Mei [6] proposed a profile nonlinear least squares estimation approach for the parameter vector $\boldsymbol{\beta}$ and coefficient function vector $\boldsymbol{\alpha}(\cdot)$ and established the asymptotic properties of the corresponding estimates. Due to the mechanism of profile method, the estimation accuracies of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}(\cdot)$ are mutually affected. Therefore, motivated by the similar idea of [17–19], here we firstly construct an efficient estimate of $\boldsymbol{\beta}$, which does not depend on the coefficient functions, based on orthogonality projection of $\mathbf{Z}$ via a simple linear regression model. Then, we approximate each coefficient function via B-spline basis functions and develop a novel variable selection procedure based on smooth-threshold estimating equations from [12]. The proposed variable selection procedure can automatically eliminate the irrelevant variables by setting the corresponding coefficient functions as zero, and simultaneously estimate the nonzero coefficient functions. In summary, this paper mainly offer the following three contributions. Firstly, compared with the profile nonlinear least squares estimation method of Li and Mei [6], we separately estimate the parametric vector and coefficient functions, and the corresponding estimates do not affect each other. Besides, we do not bring in any extra parameter in this process such as bandwidth involved in [6], which will reduce the computation burden. Secondly, we extend the smooth-threshold estimating equations approach to model (1) for variable selection purpose that was not considered in [6]. Finally, compared with the existing variable selection procedure, our approach can be easily implemented without solving any convex optimization problem and possess the oracle property.

The rest of this paper is organized as follows. In Section 2, we present the details for the estimate of parametric vector based on orthogonality-projection method, and derive the corresponding asymptotic property under suitable conditions. In Section 3, we give the variable selection procedure by using smooth-threshold estimating equation, and establish the theoretical properties including consistency and asymptotic normality of the proposed procedure. Moreover, the choices of tuning parameters are also discussed in this part. Some numerical examples are conducted in Section 4 to examine the finite sample performance of the proposed methodologies. We further illustrate the methods via a real data analysis in Section 5, and a brief conclusion is followed in Section 6. All technical proofs of the main theoretical results are provided in the Appendix.

2. Orthogonality-projection-based methodology and results

Suppose that $\{\mathbf{X}_i, \mathbf{Z}_i, U_i, Y_i\}_{i=1}^n$ is an independent and identically distributed (i.i.d.) random sample from model (1), that is,

$$Y_i = \mathbf{X}_i^T \boldsymbol{\alpha}(U_i) + g(\mathbf{Z}_i, \boldsymbol{\beta}) + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{2}$$

where $\{\varepsilon_i\}_{i=1}^n$ are i.i.d. random errors satisfying the conditions $E(\varepsilon_i) = 0$ and $\mathrm{Var}(\varepsilon_i) = \sigma^2 < \infty$. To explore the ideal of orthogonality-projection-based estimation method, we first introduce some notations for convenience. Let

$$\tilde{\mathbf{X}} = \begin{pmatrix} \mathbf{X}_1^T & & \\ & \ddots & \\ & & \mathbf{X}_n^T \end{pmatrix} = \begin{pmatrix} X_{11} & \cdots & X_{1p} & & & & \\ & & & \ddots & & & \\ & & & & X_{n1} & \cdots & X_{np} \end{pmatrix},$$

$$\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)^T, \qquad \mathbf{U} = (U_1, U_2, \ldots, U_n)^T, \qquad \tilde{\mathbf{Z}} = \left(\mathbf{Z}_1^T, \mathbf{Z}_2^T, \ldots, \mathbf{Z}_n^T\right)^T,$$

$$\mathbf{g}(\tilde{\mathbf{Z}}, \boldsymbol{\beta}) = \left(g(\mathbf{Z}_1, \boldsymbol{\beta}), g(\mathbf{Z}_2, \boldsymbol{\beta}), \ldots, g(\mathbf{Z}_n, \boldsymbol{\beta})\right)^T,$$