Journal of the Korean Statistical Society 44 (2015) 58–67

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Convex analysis in the semiparametric model with Bernstein polynomials

Jianhua Ding[a,b], Zhongzhan Zhang[a,*]

[a] College of Applied Sciences, Beijing University of Technology, Beijing, 100124, China

[b] Department of Mathematics, Shanxi Datong University, Datong, 037009, China

ARTICLE INFO

Article history:
Received 26 October 2013
Accepted 15 May 2014
Available online 6 June 2014

AMS 2000 subject classifications:
primary 62J05
secondary 62G08

Keywords:
Bernstein polynomials
Convexity
Empirical process
Asymptotic distribution

ABSTRACT

In this paper, we propose Bernstein polynomial estimation for the partially linear model when the nonparametric component is subject to convex (or concave) constraint. We employ a nested sequence of Bernstein polynomials to approximate the convex (or concave) nonparametric function. Bernstein polynomial estimation can be obtained as a solution of a constrained least squares method and hence we use a quadratic programming algorithm to compute efficiently the estimator. We show that the estimator of the parametric part is asymptotically normal. The rate of convergence of the nonparametric function estimator is established under very mild conditions. The small sample properties of our estimation are provided via simulation study and compared with regression splines method. A real data analysis is conducted to illustrate the application of the proposed method.

1. Introduction

We consider the following semiparametric partially linear model:

$$Y = \mathbf{X}^T\boldsymbol{\beta} + \psi(Z) + \varepsilon, \tag{1}$$

where $\mathbf{X} = (X_1, \ldots, X_d)^T$ and Z are d and 1 dimensional explanatory variables respectively, $\boldsymbol{\beta}$ is a $d \times 1$ vector of the unknown regression parameters, ψ is an unknown smooth function which subjects to convex (or concave) constraint of the form $\psi^{(2)}(z) \geq 0$ (or $\psi^{(2)}(z) \leq 0$), the error term ε has mean 0 and finite variance σ^2, $(\mathbf{X}, Z)$ and ε are independent. Convex restriction arises in many applications, for example, convexity for production or Engle curves (Hildreth, 1954; Matzkin, 1991). The model (1) without any restriction has been extensively studied by many authors, see, for example, Bianco and Boente (2004), Engle, Granger, Rice, and Weiss (1986), Green, Jennison, and Seheult (1985), Green and Silverman (1994), Robinson (1988), and Schimek (2009) among many others. They investigated the asymptotic behaviors of the estimates of the regression parameters and the smooth nonparametric component using smoothing spline, kernel smoothing, or local linear smoother method, etc. With monotone restriction the model (1) has also been studied in recent years, see, for example, Huang (2002), Sun, Zhang, and Du (2012), Wright (1981), etc. However, for the model (1) with convex (or concave) constraint on the nonparametric component $\psi(Z)$, it appears that no systematic study has been done on the asymptotic behavior of the estimator of $(\boldsymbol{\beta}, \psi)$.

* Corresponding author. Tel.: +86 10 67392067; fax: +86 10 67391459.
E-mail address: zzhang@bjut.edu.cn (Z. Zhang).

http://dx.doi.org/10.1016/j.jkss.2014.05.003
1226-3192/

In the nonparametric model $Y = \psi(Z) + \varepsilon$ with ψ being convex (or concave), estimation of ψ is a basic problem in the convex regression literature. Hildreth (1954) pioneered the constrained least squares method to estimate a concave function, and Fraser and Massam (1989) and Wu (1982) proposed efficient algorithms to compute the estimator. Aït-Sahalia and Duarte (2003) have also considered a decreasing and convex function with certain bounds on the derivative. Groeneboom, Jongbloed, and Wellner (2001), Hanson and Pledger (1976) and Mammen (1991) derived several asymptotic rates of convergence. Alternative approaches which combine convex restriction and smoothing process have also been proposed, for example, kernel-based estimator (Birke & Dette, 2007), spline-based estimator (Meyer, 2008) and Bernstein-polynomial-based estimator (Chang, Chien, Hsiung, Wen, & Wu, 2007; Wang & Ghosh, 2012).

In the present paper, we extend nonparametric model to the semiparametric partially linear model (1) with ψ being convex (or concave). We make use of Bernstein polynomial to approximate the nonparametric component. The desired convex (or concave) constraint can easily be imposed by imposing suitable linear constraints on the parameters of the basis functions. Thus, Bernstein polynomial estimation can be obtained as a solution of a constrained least squares method and hence we employ a quadratic programming algorithm to compute efficiently the estimator. The convex (or concave) constraint is maintained for any finite sample size and satisfied over the entire domain of the nonparametric function.

The rest of the paper is organized as follows: The convex (or concave) restricted Bernstein polynomial estimator $(\hat{\boldsymbol{\beta}}_n, \hat{\psi}_n)$ is presented in Section 2. Asymptotic results are given in Section 3. The small sample properties of our estimation are provided via simulation study and compared with convex (or concave) restricted regression splines, and a real data analysis is conducted to illustrate the application of the proposed method in Section 4. The proofs of asymptotic results are sketched in the Appendix.

2. Convex-restricted Bernstein polynomial estimator

Without the loss of generality, we assume that the domain of ψ lies in the unit [0, 1]. Let

$$\mathscr{F} = \{\psi : \psi \text{ is convex function on } [0, 1]\}.$$

Let $(Y_1, \mathbf{X}_1, Z_1), \ldots, (Y_n, \mathbf{X}_n, Z_n)$ be a set of observations which are assumed to be independently and identically distributed (i.i.d) as $(Y, \mathbf{X}, Z)$ satisfying model (1). Let $(\boldsymbol{\beta}_0, \psi_0)$ be the true value of the parameter, which minimizes the L_2 risk (Gyorfi, Kohler, Krzyzak, & Walk, 2002):

$$M(\boldsymbol{\beta}, \psi) = E(Y - \mathbf{X}^T\boldsymbol{\beta} - \psi(Z))^2,$$

where $\boldsymbol{\beta}$ belongs to a convex and compact subset $\Theta \subset \mathbb{R}^d$ and $\psi \in \mathscr{F}$ satisfies convex restriction. The L_2 empirical risk is

$$M_n(\boldsymbol{\beta}, \psi) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \mathbf{X}_i^T\boldsymbol{\beta} - \psi(Z_i))^2, \tag{2}$$

subject to $\boldsymbol{\beta} \in \Theta$ and $\psi \in \mathscr{F}$.

For a continuous function such as $\psi(Z)$ on [0, 1], the approximating Bernstein polynomial of order N is given by

$$B(Z; N, \psi) = \sum_{j=0}^{N} \psi\left(\frac{j}{N}\right) C_N^j (Z)^j (1-Z)^{N-j} = \sum_{j=0}^{N} \alpha_j b_j(Z, N),$$

where $b_j(Z, N) = \binom{N}{j} Z^j(1-Z)^{N-j}$ are the Bernstein basis polynomials and $\alpha_j = \psi(\frac{j}{N})$ are corresponding coefficients, $j = 0, 1, \ldots, N$. By the Weierstrass theorem, $B(\cdot; N, \psi) \to \psi(\cdot)$ uniformly over [0, 1] as $N \to \infty$ (Lorentz, 1986). Since the second derivative of $B(Z; N, \psi)$ can be written as

$$B''(Z; N, \psi) = N(N-1)\sum_{j=0}^{N-2}(\alpha_{j+2} - 2\alpha_{j+1} + \alpha_j) b_j(Z, N-2),$$

the convex restriction is satisfied provided that $\alpha_{j+2} + \alpha_j \geq 2\alpha_{j+1}, j = 0, \ldots, N-2$. Adopting the method of sieves, we consider the constrained Bernstein polynomial sieve $\mathscr{F}_N$ as follows:

$$\mathscr{F}_N = \left\{ B_N(Z) = \sum_{j=0}^{N} \alpha_j \cdot b_j(Z, N) : \alpha_{j+2} + \alpha_j \geq 2\alpha_{j+1}, j = 0, \ldots, N-2, \ \sum_{j=0}^{N} |\alpha_j| \leq L_N \right\},$$

for $N = 1, 2, \ldots$. Each element in the sieve $\mathscr{F}_N$ preserves the desired convex restrictions. The sequence of sieves $\mathscr{F}_N$ is nested in $\mathscr{F}$, *i.e.*, $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \cdots \subset \mathscr{F}_N \subset \mathscr{F} \subset L_2[0, 1]$ and $\bigcup_{N=1}^{\infty} \mathscr{F}_N$ is dense in $\mathscr{F}$ with respect to sup-norm (Wang & Ghosh, 2012).

By replacing $\psi(Z)$ by $\sum_{j=0}^{N} \alpha_j \cdot b_j(Z, N)$ in the L_2 empirical risk (2), we obtain the Bernstein polynomial L_2 empirical risk function,

$$M_n(\alpha, \boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \mathbf{X}_i^T\boldsymbol{\beta} - \sum_{j=0}^{N}\alpha_j b_j(Z_i, N)\right)^2, \tag{3}$$