Journal of the Korean Statistical Society 43 (2014) 11–17

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

On partial linear additive isotonic regression

Kyusang Yu *

Konkuk University, Republic of Korea

ARTICLE INFO

Article history:
Received 31 March 2013
Accepted 15 July 2013
Available online 7 August 2013

AMS 2000 subject classifications:
62G08
62G20

Keywords:
Partially linear additive models
Tangent cone
Efficiency bound
Isotonic regression

ABSTRACT

In this paper we discuss semiparametric additive isotonic regression models. We discuss the efficiency bound of the model and the least squares estimator under this model. We show that the ordinary least square estimator studied by Huang (2002) and Cheng (2009) for the semiparametric isotonic regression achieves the efficiency bound for the regular estimator when the true parameter belongs to the interior of the parameter space. We also show that the result by Cheng (2009) can be generalized to the case that the covariates are dependent on each other.

1. Introduction

In this paper we discuss semiparametric additive isotonic regression models. We discuss the efficiency bound of the model and the least squares estimator under this model. We find that the monotonicity of infinite dimensional nuisance parameters is not helpful for estimating the linear part when we consider regular estimators. We also show that the ordinary least square estimator studied by Cheng (2009) and Huang (2002) for the semiparametric isotonic regression achieves the efficiency bound for the regular estimator when the true parameter belongs to the interior of the parameter space. We also show that the result by Cheng (2009) can be generalized to the case that the covariates are dependent on each other.

For multidimensional semiparametric regression, we often impose some restrictions on the regression function to avoid the curse of dimension. Additively separable models are popular among those structured models since they inherit the simple structure and easy interpretation from the linear model. Meanwhile, they allow flexible modeling for each component function. The semiparametric additively separable regression function has the following structure:

$$E(Y|\mathbf{W}) = f(\mathbf{X};\theta) + m(\mathbf{Z}) \quad (1)$$

where $\mathbf{W}^T = (\mathbf{X}^T, \mathbf{Z}^T)$ is the covariate vector with $\mathbf{X}^T \in \mathbb{R}^p$ and $\mathbf{Z}^T \in \mathbb{R}^d$, and Y is the response. Here, $f(\cdot;\theta)$ is a function parameterized by the finite dimensional parameter θ and m is a smooth function.

Model (1) is the partial linear model if $f(\mathbf{X};\theta) = \theta^T\mathbf{X}$. Furthermore, if we impose the additive structure on m, i.e., $m(\mathbf{Z}) = m_1(Z_1) + \cdots + m_d(Z_d)$, it becomes the partial linear additive model. For the estimation under these models, we refer to Bhattacharya and Zhao (1997), Schick (1993, 1996), Speckman (1988), and Yu, Mammen, and Park (2011) among others. As an extension of model (1), some researchers considered a transformation of the conditional mean instead of the conditional mean itself, that is $g(E(Y|\mathbf{W})) = f(\mathbf{X};\theta) + m(\mathbf{Z})$ for some function g. See, e.g., Hastie and Tibshirani (1990), Severini and Staniswalis (1994), and Yu and Lee (2010) for details. For the efficiency and efficient estimation under these

* Tel.: +82 2 2049 6121.
E-mail addresses: kyusangu@konkuk.ac.kr, kyusangu@hanmail.net.

1226-3192/$ – see front matter
http://dx.doi.org/10.1016/j.jkss.2013.07.003

models, we also refer to Cheng, Zhou, and Huang (in press), Cuzick (1992), Severini and Wong (1992), Yu and Lee (2010), and Yu et al. (2011).

In this paper we study the partial linear additive isotonic model, which is a subclass of model (1). We write the model as follows:

$$Y = \theta^T \mathbf{X} + m_1(Z_1) + \cdots + m_d(Z_d) + \epsilon \tag{2}$$

where $m_j, j = 1, \ldots, d$ are smooth monotone functions and ϵ is an error independent with the covariate $\mathbf{W}^T = (\mathbf{X}^T, \mathbf{Z}^T)$. Note that model (2) includes the usual isotonic regression model, additive isotonic regression model and partial linear isotonic regression model, for the case $p = 0,\ d = 1$, for the case $p = 0$ and $d > 1$, and for the case $p \geq 1$ and $d = 1$, respectively.

The isotonic regression is very useful for analyzing the monotone relationship between the input and the output, which are very common in economics, biosciences, and engineering. The usual isotonic regression is now a classical area in statistics (see, e.g., Robertson, Wright, & Dykstra, 1988). The additive isotonic regression model was discussed by Bacchetti (1989), De Boer, Besten, and Ter Braak (2002), and Morton-Jones, Diggle, Parker, Dickinson, and Binks (2000) from the applied side. Mammen and Yu (2007) obtained the asymptotic distribution of the least square estimator under this model and also showed that the least square problem can be solved by applying a cyclic pool adjacent violators algorithm. Huang (2002) obtained the asymptotic distribution of the least square estimator under the partial linear isotonic regression model when $d = 1$. He showed that the least square estimator $\hat{\theta}$ is $\sqrt{n}$-consistent and asymptotically normal and m_1 can be estimated asymptotically as if there does not exist the parametric part, $\theta^T \mathbf{X}$. Cheng (2009) extended this result for the case $d > 1$ with the pairwise independent covariates in the nonparametric part. Cheng, Zhao, and Li (2012) also studied an empirical likelihood method for the semiparametric additive isotonic regression to obtain a good confidence interval.

In this paper, we extend Cheng's (2009) result to the general case of dependent covariate vector and show that the least square estimator attains the efficient asymptotic variance under the Gaussian error model. The semiparametric Fisher information bound for model (2) when $d = 1$ was first studied by Tripathi (2000). We also extend that result to the general dimension in this paper.

2. Main result

In this section we discuss semiparametric efficiency for estimating the parameter θ under the partial linear additive isotonic model (2) with a Gaussian error. Suppose we have a random sample of size n, $(Y^1, \mathbf{W}^1), \ldots, (Y^n, \mathbf{W}^n)$, which obeys the semiparametric regression model (2). Here, we further assume that the errors ϵ^i have the normal distribution with mean zero and variance σ^2. Note that the additive function $m(z_1, \ldots, z_d) = m_1(z_1) + \cdots + m_d(z_d)$ is not identifiable up to constants. We impose $Em_j(Z_j) = 0$ for the identification and we also assume w.l.o.g. that $Em(Z_1, \ldots, Z_d) = 0$ for the simplicity. We also assume that the covariate vector $\mathbf{W}^T = (\mathbf{X}^T, \mathbf{Z}^T)$ has a joint density q with respect to $\nu = \nu_1 \times \nu_2$ where ν_1 is a σ-finite measure and ν_2 is the Lebesgue measure on each support of $\mathbf{X}$ and $\mathbf{Z}$. We only require that ν_1 is a σ-finite measure, which allows discrete random variables as well as continuous ones. Then, the random vector $(Y, \mathbf{W}^T)^T$ has the density:

$$p(y, \mathbf{x}, \mathbf{z}; \theta, m, \sigma^2, q) = q(\mathbf{x}, \mathbf{z}) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(y - \theta^T \mathbf{x} - m_1(z_1) - \cdots - m_d(z_d))^2}{2\sigma^2} \right\} \tag{3}$$

where m_j's are increasing functions. Note that, by the symmetry, increasing functions can be replaced by decreasing functions with the reverse ordering in the covariate.

We will take a framework to Tripathi (2000) and van der Vaart (1989). Let $C_1(A)$ be the set of all continuously differentiable functions defined on a set $A \subset \mathbb{R}$. Define $\mathcal{F}_j = \{f \in C_1(\mathcal{S}_j) : f'(s) \geq 0,\ Ef(Z_j) = 0\}$ where $\mathcal{S}_j$ is the support of Z_j. Then the sets $\mathcal{F}_j$ are closed convex cones in $L_2(q_{Z_j})$ where q_{Z_j} is the density of Z_j if $\mathcal{S}_j$ are compact intervals. Let $\mathcal{F}_{\text{add}} = \mathcal{F}_1 + \cdots + \mathcal{F}_d = \{f(\mathbf{z}) = f_1(z_1) + \cdots, f_d(z_d); f_j \in \mathcal{F}_j \text{ for } j = 1, \ldots, d\}$. Then it is also clear that $\mathcal{F}_{\text{add}}$ is a closed convex cone in $L_2(q_{\mathbf{Z}})$ where $q_{\mathbf{Z}}$ is the density of the random vector $\mathbf{Z}$. The difficulty in calculating the Fisher information under model (3) arises since $\mathcal{F}_{\text{add}}$ is not a linear space but a proper convex cone and thus it has the boundary. We need a special treatment at the boundary of $\mathcal{F}_{\text{add}}$. Note that the boundary of $\mathcal{F}_{\text{add}}$ is given as $\partial(\mathcal{F}_{\text{add}}) = \{f \in \mathcal{F}_{\text{add}} : f_j'(x_j) = 0 \text{ for at least a } x_j \text{ and a } j\}$.

Under model (3), the estimation problem of (θ, m) is orthogonal to the parameter σ^2 and q, and thus we may regard σ^2 and q are known for the moment w.l.o.g. Then we have the log-likelihood of (θ, m), $l_n(\theta, m; \{(Y^i, \mathbf{W}^i)\}_{i=1,\ldots,n})$ so that

$$l_n(\theta, m; \{(Y^i, \mathbf{W}^i)\}_{i=1,\ldots,n}) \propto -\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(Y^i - \theta^T \mathbf{X}^i - m(Z_1^i, \ldots, Z_d^i)\right)^2 \tag{4}$$

$$\equiv \sum_{i=1}^{n} \ell(\theta, m; (Y^i, \mathbf{W}^i)) \tag{5}$$

where $\theta \in \Theta \subset \mathbb{R}$ and $m \in \mathcal{F}_{\text{add}}$. For the simplicity, hereafter, we suppress the notation for data in l_n and ℓ.