Journal of Multivariate Analysis 101 (2010) 1898–1901

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Note(s)

On an asymptotically more efficient estimation of the single-index model

Ziqing Chang [a], Liugen Xue [b], Lixing Zhu [a,*]

[a] *Hong Kong Baptist University, China*
[b] *Beijing University of Technology, China*

ARTICLE INFO

Article history:
Received 9 November 2008
Available online 11 March 2010

AMS 2001 subject classifications:
62J05
62G08
62G20

Keywords:
The single-index model
Asymptotical efficiency
Least squares estimation

ABSTRACT

In this note, we revisit the single-index model with heteroscedastic error, and recommend an estimating equation method in terms of transferring restricted least squares to unrestricted least squares: the estimator of the index parameter is asymptotically more efficient than existing estimators in the literature in the sense that it is of a smaller limiting variance.

1. Introduction

The single-index model has received much attention recently because of its usefulness in several areas such as econometrics. The single-index model is of the following form:

$$Y = g(X^T\beta_0) + \varepsilon, \tag{1}$$

where Y is the univariate response and X is a p-dimensional covariable, β_0 is an unknown index of interest, the function g is an unknown link, and $E(\varepsilon|X) = 0$. Furthermore, for the sake of identifiability, we assume with no loss of generality that $\|\beta_0\| = 1$ and the first component of β_0 is positive, where $\|\cdot\|$ denotes the Euclidean norm. Such a model with Gaussian predictors was formulated by Brillinger [1]. It is of dimension-reduction structure in the sense that, if we can estimate the index β_0 efficiently, we can use the univariate $X^T\hat{\beta}_0$ as the covariable and Y to estimate the nonparametric link g, and thus avoid the curse of dimensionality when nonparametric smoothing is employed. Thus, much effort has been devoted to estimating the index β_0 efficiently. Hall [2] investigated a locally defined estimator, and Zhu and Fang [3] considered a globally defined estimator. Both were done under a projection pursuit framework. Ichimura [4] showed that the parameter β_0 can be estimated root-n consistently.

When ε is homoscedastic, Härdle et al. [5] obtained the asymptotic normality of an estimator of β_0 in a least squares sense that is similar to projection pursuit regression method in [2] and then [3]. In a special case of the model investigated by Carroll et al. [6], the limiting variance is the same as that of Härdle et al. [5]. In the heteroscedastic case, when the model studied by Xia and Härdle [7] is reduced to the single-index model, their estimator is of the following variance:

$$Q_0^- Q_2 Q_0^-, \tag{2}$$

where $Q_k = E\{|\varepsilon|^k g'(X^T\beta_0)^2[X - E(X|X^T\beta_0)][X - E(X|X^T\beta_0)]^T\}$ with $k = 0, 2$, and Q_0^- is a generalized inverse of Q_0. It is clear that when ε is homoscedastic, this variance is equal to $\sigma^2 Q_0^-$, identical to Härdle, Hall and Ichimura's [5], where

* Corresponding address: Department of Mathematics, Hong Kong Baptist University, Hong Kong, China.
E-mail address: lzhu@hkbu.edu.hk (L. Zhu).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2010.02.005

σ^2 is the variance of ε. A parallel estimation method to nonlinear parametric regression is that of [8]. Xia [9] proposed a refined minimum average conditional variance estimation (rMAVE) method that is a modification of MAVE [10]. In the homoscedastic case, the limiting variance of rMAVE is also similar to the one derived by Härdle et al. [5]. This kind of variance has become a standard result in the literature. It is noted that Xia [9] also obtained a refined outer product of gradients (rOPG) method. The limiting variance is of a very complicated structure and is difficult to compute. Also the rOPG method is asymptotically less efficient than rMAVE (see [9]). A relevant reference is [11].

However, note that β_0 is not an inner point in R^p. Thus any estimator of β_0 is not regular and then the estimator cannot be globally asymptotically efficient although it can be locally asymptotically efficient; see [6]. Note that all existing methods are restricted least squares type with the constraint $\|\beta\| = 1$, although this is not explicitly clarified. As is well known, in regular cases, the least squares method is of efficiency properties; see [12]. Hence, if we can transfer the restricted least squares to unrestricted least squares and then an estimating equation, we may be able to obtain an asymptotically more efficient estimation than the ones in the literature. In homoscedastic cases, Wang et al. [13] suggested an algorithm to implement this idea, and confirmed our expectation. In this note, we consider the heteroscedastic case and obtain an estimator whose limiting variance is as follows:

$$J_{\beta_0^{(r)}} V^{-1} Q V^{-1} J^T_{\beta_0^{(r)}}, \tag{3}$$

where

$$Q = E\{|\varepsilon|^2 g'(X^T\beta_0)^2 J^T_{\beta_0^{(r)}}[X - E(X|X^T\beta_0)][X - E(X|X^T\beta_0)]^T J_{\beta_0^{(r)}}\},$$
$$V = E[g'(X^T\beta_0)^2 J^T_{\beta_0^{(r)}} XX^T J_{\beta_0^{(r)}}],$$

and $J_{\beta_0^{(r)}}$ is defined in (10), which is the Jacobian matrix of β with respect to $\beta^{(r)}$ evaluated at $\beta_0^{(r)}$, where $\beta^{(r)}$ is defined in (9) in Section 2. We can prove that this variance is smaller than that of (2). That is, our estimator is asymptotically more efficient than those of Xia [9] and Xia and Härdle [7] in heteroscedastic cases.

In the next section, we describe the estimation procedure and main results. We do not include the details of the proof in this note. The full version of the paper can be obtained from Lixing Zhu at lzhu@hkbu.edu.hk upon request.

2. Estimation procedure and main results

Suppose that $(X_i, Y_i), 1 \le i \le n$ are independent identically distributed (iid) from model (1). For the nonparametric estimator of the link function g, the local linear smoother with the linear weighted functions $\{W_{ni}(z, \beta, h); 1 \le i \le n\}$ is defined in the following. For the true parameter β_0 in case it were given, the local linear estimators of g and its derivative g' are the pair (a, b) minimizing the weighted sum of squares (see [14])

$$\sum_{i=1}^n \left[Y_i - a - b(X_i^T\beta_0 - z)\right]^2 K_h(X_i^T\beta_0 - z), \tag{4}$$

where $K_h(\cdot) = K(\cdot/h)/h$, with K being a symmetric kernel function on $\mathbf{R}^1$ and $h = h_n$ being the bandwidth.

Let $\hat{g}(z;\beta_0,h) = \hat{a}$ and $\hat{g'}(z;\beta_0,h) = \hat{b}$ be the resulting estimators. Via a simple calculation, we have

$$\begin{cases} \hat{g}(z;\beta_0,h) = \sum_{i=1}^n W_{ni}(z;\beta_0,h)Y_i, \\ \hat{g'}(z;\beta_0,h) = \sum_{i=1}^n \widetilde{W}_{ni}(z;\beta_0,h)Y_i, \end{cases} \tag{5}$$

where, for $1 \le i \le n$,

$$W_{ni}(z;\beta_0,h) = \frac{U_{ni}(z;\beta_0,h)}{\sum_{i=1}^n U_{ni}(z;\beta_0,h)}, \tag{6}$$
$$\widetilde{W}_{ni}(z;\beta_0,h) = \frac{\widetilde{U}_{ni}(z;\beta_0,h)}{\sum_{i=1}^n U_{ni}(z;\beta_0,h)},$$

with

$$U_{ni}(z;\beta_0,h) = [S_{n,2}(z;\beta_0,h) - (X_i^T\beta_0 - z)S_{n,1}(z;\beta_0,h)]K_h(X_i^T\beta_0 - z),$$
$$\widetilde{U}_{ni}(z;\beta_0,h) = [(X_i^T\beta_0 - z)S_{n,0}(z;\beta_0,h) - S_{n,1}(z;\beta_0,h)]K_h(X_i^T\beta_0 - z),$$
$$S_{n,r}(z;\beta_0,h) = \frac{1}{n}\sum_{i=1}^n (X_i^T\beta_0 - z)^r K_h(X_i^T\beta_0 - z), \quad r = 0, 1, 2.$$