Journal of Computational and Applied Mathematics 300 (2016) 97–102

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Restricted difference-based Liu estimator in partially linear model

Jibo Wu

Key Laboratory of Group & Graph Theories and Applications, Chongqing University of Arts and Sciences, Chongqing, 402160, China

ARTICLE INFO

Article history:
Received 14 September 2015
Received in revised form 14 December 2015

MSC:
primary 62G08
secondary 62J07

Keywords:
Linear restrictions
Difference-based estimator
Liu estimator
Restricted difference-based Liu estimator

ABSTRACT

Partially linear model is useful in statistical model as a multivariate nonparametric fitting method. This paper deals with statistical inference for the partially linear model in the presence of multicollinearity. When some additional linear restrictions are assumed to hold, the corresponding restricted difference-based Liu estimator for the parametric component is constructed. The asymptotically properties of the proposed estimators are discussed. Finally, a simulation study is presented to explain the performance of the estimators.

Contents

1. Introduction

Let us consider the following partially linear model

$$y_i = X_i'\beta + f(t_i) + \varepsilon_i, \quad i = 1, \dots, n \tag{1}$$

with y_i denotes a scalar response, $X_i = (X_{i1}, \dots, X_{ip})'$ denotes a $p \times 1$ independent vectors with a non-singular covariance matrix Σ_X, $\beta = (\beta_1, \dots, \beta_p)'$ denotes a p-vector of unknown parameters, $f(\cdot)$ is the unknown function, the model error ε_i is an independent random error with zero mean and variance σ^2.

Rewrite model (1) in matrix notation as

$$y = X\beta + f(t) + \varepsilon \tag{2}$$

where $y = (y_1, \dots, y_n)'$, $f(t) = (f(t_1), \dots, f(t_n))'$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$ and $X = (X_1, \dots, X_n)'$ is the $n \times p$ matrix.

E-mail address: linfen52@126.com.

http://dx.doi.org/10.1016/j.cam.2015.12.023
0377-0427/

Partial linear models are more flexible than standard linear models since they have a parametric and a nonparametric component. They can be a suitable choice when one suspects that the response y linearly depends on X, but that it is nonlinearly related to X.

The condition number is a measure of the presence of multicollinearity. The condition number of the matrix X presents some information about the existence of multicollinearity, however it does not illustrate the structure of the linear dependency among the column vectors $X_1, X_2, \ldots, X_n$. The best way of illustrating the existence and structure of multicollinearity is to see the eigenvalues of $X'X$. If $X'X$ is ill-conditioned with a large condition number a Liu regression estimator can be used to estimate β (see e.g. [1–7]). In this paper, we will examine a biased estimation techniques to be followed when the matrix $X'X$ appears to be ill-conditioned in the partial linear model. We suppose that the condition number of the parametric component is large explains that a biased estimation procedure is desirable.

In this paper, a restricted difference-based estimator is presented for the vector parameter β in the partially linear model when the linear nonstochastic constraint is assumed to hold. We also examine the properties of the proposed estimator.

The rest of the paper is organized as follows: the restricted difference-based Liu estimator is defined in Section 2 and the properties of the proposed estimator are discussed in Section 3. The performance of the new estimator is evaluated by a simulation study in Section 4 and some conclusions are given in Section 5.

2. Profile least-squares estimator

In this section we will propose the restricted difference-based Liu estimator in partially linear model.

2.1. Difference-based estimator

Let $d = (d_0, \ldots, d_m)$ be a $m + 1$ vector, where m is the order of differencing and $d_0, \ldots, d_m$ are differencing weights satisfying the conditions

$$\sum_{j=0}^{m} d_j = 0, \qquad \sum_{j=0}^{m} d_j^2 = 1. \tag{3}$$

Moreover, for $k = 1, \ldots, m$ let $c_k = \sum_{i=1}^{m+1-k} d_i d_{i+k}$. Now, we denote the $(n - m) \times n$ differencing matrix D whose elements satisfy Eq. (3) as follows:

$$D = \begin{pmatrix} d_0 & d_1 & \cdots & d_m & 0 & 0 & \cdots & 0 \\ 0 & d_0 & d_1 & \cdots & d_m & 0 & \cdots & 0 \\ \cdots & \cdots & & & & & & \\ \cdots & \cdots & & & & & & \\ \cdots & \cdots & & & & & & \\ 0 & 0 & \cdots & d_1 & \cdots & d_m & 0 & 0 \\ 0 & 0 & \cdots & d_0 & d_1 & \cdots & d_m & 0 \\ 0 & 0 & \cdots & 0 & d_0 & d_1 & \cdots & d_m \end{pmatrix}. \tag{4}$$

This and related matrices are given, for example, in [8]. Then we can use the differencing matrix to model (2), and this leads to direct estimation of the parametric effect. In particular, take

$$Dy = DX\beta + Df(t) + D\epsilon. \tag{5}$$

Since the data have been reordered so that the $X's$ are close, the application of the differencing matrix D in model (3) can remove the nonparametric effect in large samples [8]. This ignores the presence of $Df(t)$. Thus, we may write Eq. (7) as

$$Dy \doteq DX\beta + D\varepsilon \tag{6}$$

or

$$\widetilde{y} \doteq \widetilde{X}\beta + \widetilde{\varepsilon} \tag{7}$$

where $\widetilde{y} = Dy, \widetilde{X} = DX$ and $\widetilde{\varepsilon} = D\varepsilon$.

For arbitrary differencing coefficients satisfying Eq. (6), Yatchew [9] defines a simple differencing estimator of the parameter β in a partial linear model

$$\hat{\beta} = (\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}'\widetilde{y}. \tag{8}$$

In order to account for the parameter β in Eq. (3), we propose the modified estimator of σ^2, defined as

$$\hat{\sigma}^2 = \frac{\widetilde{y}'(I - P)\widetilde{y}}{tr(D'(I - P)D)} \tag{9}$$

where P is the projection matrix and defined as

$$P = \widetilde{X}(\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}'. \tag{10}$$