



Research article

A generalized Liu-type estimator for logistic partial linear regression model with multicollinearity

Dayang Dai and Dabuxilatu Wang*

School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China

* **Correspondence:** Email: wangdabu@gzhu.edu.cn.

Abstract: This paper is concerned with proposing a generalized Liu-type estimator (GLTE) to address the multicollinearity problem of explanatory variable of the linear part in the logistic partially linear regression model. Using the profile likelihood method, we propose the GLTE as a general class of Liu-type estimator, which includes the profile likelihood estimator, the ridge estimator, the Liu estimator and the Liu-type estimator as special cases. The conditional superiority of the proposed GLTE over the other estimators is derived under the asymptotic mean square error matrix (MSEM) criterion. Moreover, the optimal choices of biasing parameters and function of biasing parameter are given. Numerical simulations demonstrate that the proposed GLTE performs better than the existing estimators. An application on a set of real data arising from the study of Indian Liver Patient is shown for illustrating our theoretical results.

Keywords: Liu estimator; logistic partial linear model; multicollinearity; profile likelihood; ridge estimator

Mathematics Subject Classification: 62J07, 62G05, 62F10

1. Introduction

A generalized partial linear regression model (GPLRM) is a semiparametric extension of a generalized linear regression model. In the study of GPLRM, the topics such as parameters estimation, the effectiveness of estimators have been paid much attention by researchers [1–5]. Here the famous profile likelihood method (PLM) came into being. As a special case of GPLRM, the logistic partial linear regression model (LPLRM) is widely used to study the relationship between binomial response variable and explanatory variables with linear part and nonparametric part in fields of medical studies, social studies and biological studies, and its parameters estimation as well as the asymptotic properties of the estimators can be obtained from the framework of GPLRM. As known, the LPLRM is assumed that the explanatory variables in the linear part are mutually independent.

However, in application, the assumption of independence between explanatory variables in the linear part seldom holds, and a violation of the assumption causes the multicollinearity problem. It is well known that multicollinearity may cause large variance of the estimators, which in turn leads to wide confidence intervals and erroneous sign estimates.

In the study of the multicollinearity problem, to our best knowledge, the multicollinearity of linear regression models, generalized linear regression models (logistic, poisson and gamma), and partial linear regression models has been well investigated, respectively. (a) Most proposals of powerful novel methods for dealing with multicollinearity have appeared in the study of the linear regression models. Hoerl and Kennard [6] firstly proposed a ridge estimation program of unknown parameters to combat multicollinearity of a linear regression model. Liu [7] presented a class of biased Liu estimators of the unknown parameters that are superior to ridge estimators with ease of selecting bias parameters. Liu [8] introduced a two-parameter Liu-type estimator of unknown parameters to deal with the serious multicollinearity problem that the ridge estimator failed to combat. Kurnaz and Akay [9] proposed a general Liu-type estimator of unknown parameters based on the Liu-type estimator. Moreover, Zeinal [10] contributed an extension of the two-parameter estimator presented by Özkale and Kaçiranlar [11]. (b) In the case of a generalized linear regression model, for dealing with multicollinearity, some modified estimators are established. The estimators to combat multicollinearity in the logistic linear regression models (LLRM) are listed below. Kibria, Mnsson and Shukur [12] generalized and compared logistic ridge estimators with different ridge parameters. Inan and Erdogan [13] introduced a Liu-type estimator that had a smaller total mean squared error (MSE) than the ridge estimator under certain conditions. Asar and Genç [14]) constructed a two-parameter ridge estimator for the LLRM. Varathan and Wijekoon [15] proposed estimator called Modified almost unbiased logistic Liu estimator (MAULLE). Ertan and Akay [16] modified the general Liu-type estimator with reconstructing the form of the biasing parameter function. Jadhav [17] proposed a new estimator designated as linearized ridge logistic estimator. And also there are new poisson ridge estimator by Rashad et al. [18], and gamma modified ridge-type estimator by Lukman et al. [19]. (c) In order to eliminate the multicollinearity of a partial linear regression model (PLRM), some biased estimators of unknown parameters are constructed by considering additive errors, correlated errors or linear constraint, etc. For example, Roozbeh and Arashi [20] proposed difference-based ridge type estimators combining the restricted least squares method in the seemingly unrelated semiparametric models. Wu [21] proposed a difference-based almost unbiased Liu estimator in PLRM. Emami and Aghamohammadi [22] constructed the difference-based ridge and Liu type estimators in PLRM when the covariates are measured with additive errors. Akdeniz and Roozbeh [23] introduced a generalized difference-based almost unbiased ridge estimator in PLRM when the errors are correlated. Wu and Kibria [24] considered the generalized difference-based mixed two-parameter estimator in PLRM.

Theoretically, statistical regression models (except for univariate regression model) including linear part may face to a multicollinearity problem. Now we consider the case of a LPLRM, to our best knowledge, there is no literature reported on combating multicollinearity problem in study of a LPLRM. Apparently, the existing methods mentioned above combating multicollinearity in LLRM and PLRM can be recommended to construct a suitable biased estimator for our purpose. However, it is difficult to properly coordinate the existing methods for LLRM and PLRM, i.e. the combination of the methods should generate an alternative (estimator) which optimally reduces the impact of

multicollinearity on LPLRM, meanwhile the linear part and the nonparametric part of the LPLRM are estimated in accordance with some standard statistical criteria. We attempt to deal such issues in section 3 where the PLM is employed to construct a more generalized Liu-type estimator, and the optimal choices of the biasing parameters and the function of the biasing parameter as well as the superiority conditions of the proposed estimator over other estimators are given.

This paper is organized as follows. Section 2 states the LPLRM, its estimation and evaluation criteria as well as the PLM. Several biased estimators and the proposed generalized Liu-type estimator (GLTE) are given in section 3. In section 4, theoretical conditions are derived to study the superiority of the GLTE over the other estimators under MSEM criterion. In section 5, the optimal choices of the biasing parameters and the function of biasing parameter are determined. In section 6, Monte Carlo simulations are given to evaluate the performance of the proposed GLTE. Section 7 presents a real data application. Finally, a brief summary and conclusions are given in section 8.

2. Statistical methodology

2.1. Model and estimations

The dependent variable $y_i \in \{0, 1\}$ is the binary response variable. \mathbf{x}_i^T is the i th row of $n \times p$ explanatory variables matrix \mathbf{X} , which may take any form of continuous, discrete or mixture of discrete and continuous. $\mathbf{t}_i \in \mathcal{R}^q$ is i th sample of a q -variate random vector of continuous explanatory variables. We consider the LPLRM as follows:

$$\log\left(\frac{\pi_i}{1 - \pi_i}\right) = \mathbf{x}_i^T \boldsymbol{\beta} + m(\mathbf{t}_i), i = 1, 2, \dots, n, \quad (2.1)$$

where $\pi_i = \text{prob}(y_i = 1 | \mathbf{x}_i, \mathbf{t}_i)$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^T$ is $p \times 1$ parametric vector, $m(\cdot)$ is a nonparametric function.

The PLM is often used to obtain estimators of model (2.1). In PLM, the smoothed or local log-likelihood for the nonparametric function $m_{\boldsymbol{\beta}}(\mathbf{t})$ at point \mathbf{t} is given by

$$L^{\mathbf{H}}(m_{\boldsymbol{\beta}}(\mathbf{t})) = \sum_{j=1}^n \kappa_{\mathbf{H}}(\mathbf{t} - \mathbf{t}_j) \left(y_j \log \frac{\pi_j(\boldsymbol{\beta}, m_{\boldsymbol{\beta}}(\mathbf{t}))}{1 - \pi_j(\boldsymbol{\beta}, m_{\boldsymbol{\beta}}(\mathbf{t}))} + \log(1 - \pi_j(\boldsymbol{\beta}, m_{\boldsymbol{\beta}}(\mathbf{t}))) \right), \quad (2.2)$$

with $\pi_j(\boldsymbol{\beta}, m_{\boldsymbol{\beta}}(\mathbf{t})) = \frac{\exp(\mathbf{x}_j^T \boldsymbol{\beta} + m_{\boldsymbol{\beta}}(\mathbf{t}))}{1 + \exp(\mathbf{x}_j^T \boldsymbol{\beta} + m_{\boldsymbol{\beta}}(\mathbf{t}))}$. Here, $\kappa_{\mathbf{H}}(\mathbf{t} - \mathbf{t}_j)$ denote local kernel weights with a (multidimensional) kernel function κ and a bandwidth matrix \mathbf{H} , $m_{\boldsymbol{\beta}}(\mathbf{t})$ is a differentiable function with respect to $\boldsymbol{\beta}$ for each \mathbf{t} . The Logarithmic profile likelihood for $\boldsymbol{\beta}$ can be written as

$$L(\boldsymbol{\beta}) = \sum_{i=1}^n \left(y_i \log \frac{\pi_i(\boldsymbol{\beta}, m_{\boldsymbol{\beta}}(\mathbf{t}_i))}{1 - \pi_i(\boldsymbol{\beta}, m_{\boldsymbol{\beta}}(\mathbf{t}_i))} + \log(1 - \pi_i(\boldsymbol{\beta}, m_{\boldsymbol{\beta}}(\mathbf{t}_i))) \right), \quad (2.3)$$

with $\pi_i(\boldsymbol{\beta}, m_{\boldsymbol{\beta}}(\mathbf{t}_i)) = \frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta} + m_{\boldsymbol{\beta}}(\mathbf{t}_i))}{1 + \exp(\mathbf{x}_i^T \boldsymbol{\beta} + m_{\boldsymbol{\beta}}(\mathbf{t}_i))}$. Abbreviate $m_i = m_{\boldsymbol{\beta}}(\mathbf{t}_i)$, the likelihood equations are obtained from Eqs (2.2) and (2.3) as follows

$$\frac{\partial L^{\mathbf{H}}(m_{\boldsymbol{\beta}}(\mathbf{t}_i))}{\partial m_i} = \sum_{j=1}^n (y_j - \pi_j(\boldsymbol{\beta}, m_i)) \kappa_{\mathbf{H}}(\mathbf{t}_i - \mathbf{t}_j) = \mathbf{0}, \quad (2.4)$$

and

$$\frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n (y_i - \pi_i(\boldsymbol{\beta}, m_i))(x_i + \mathbf{m}'_i) = \mathbf{0}, \quad (2.5)$$

where \mathbf{m}'_i is the partial derivative vector of $m_{\beta}(t_i)$ with respect to $\boldsymbol{\beta}$. Taking the partial derivative of Eq (2.4) with respect to $\boldsymbol{\beta}$, we have

$$\mathbf{m}'_i = -\frac{\sum_{j=1}^n \pi_j(\boldsymbol{\beta}, m_j)(1 - \pi_j(\boldsymbol{\beta}, m_j))\kappa_{\mathbf{H}}(t_i - t_j)\mathbf{x}_j}{\sum_{j=1}^n \pi_j(\boldsymbol{\beta}, m_j)(1 - \pi_j(\boldsymbol{\beta}, m_j))\kappa_{\mathbf{H}}(t_i - t_j)}.$$

The second partial derivative of Eq (2.3) with respect to $\boldsymbol{\beta}$ is given by

$$\frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} = -\sum_{i=1}^n \pi_i(\boldsymbol{\beta}, m_i)(1 - \pi_i(\boldsymbol{\beta}, m_i))(x_i + \mathbf{m}'_i)(x_i + \mathbf{m}'_i)^T. \quad (2.6)$$

By the iterative Newton-Raphson algorithm [25] (see Eqs (2.5) and (2.6)), the iterative formula of $\boldsymbol{\beta}$ is given by

$$\begin{aligned} \boldsymbol{\beta}^{new} &= \boldsymbol{\beta} - (\tilde{\mathbf{X}}^T \mathbf{W} \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^T (\mathbf{y} - \boldsymbol{\pi}) \\ &= (\tilde{\mathbf{X}}^T \mathbf{W} \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^T \mathbf{W} (\tilde{\mathbf{X}} \boldsymbol{\beta} + \mathbf{W}^{-1}(\mathbf{y} - \boldsymbol{\pi})), \end{aligned}$$

where $\tilde{\mathbf{X}} = \mathbf{X} - \mathbf{S}\mathbf{X}$, \mathbf{S} is the smoothing matrix with the following elements

$$S_{ij} = \frac{\pi_i(\boldsymbol{\beta}, m_j)(1 - \pi_i(\boldsymbol{\beta}, m_j))\kappa_{\mathbf{H}}(t_i - t_j)}{\sum_{i=1}^n \pi_i(\boldsymbol{\beta}, m_j)(1 - \pi_i(\boldsymbol{\beta}, m_j))\kappa_{\mathbf{H}}(t_i - t_j)},$$

$\tilde{\mathbf{X}}$ is the matrix with rows $\tilde{\mathbf{x}}_i^T$, $\tilde{\mathbf{x}}_i = \mathbf{x}_i + \mathbf{m}'_i$, the diagonal matrix $\mathbf{W} = \text{diag}(\pi_i(\boldsymbol{\beta}, m_i)(1 - \pi_i(\boldsymbol{\beta}, m_i)))_{i=1, \dots, n}$, $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$, $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_n)^T$. We will be able to get the iterative weighted least squares algorithm [25] for the model (2.1). Define an adjusted dependent variable $\mathbf{Z} = \mathbf{X}\boldsymbol{\beta} + \mathbf{m} + \mathbf{W}^{-1}(\mathbf{y} - \boldsymbol{\pi})$ to obtain $\tilde{\mathbf{Z}} = \mathbf{Z} - \mathbf{S}\mathbf{Z} = \tilde{\mathbf{X}}\boldsymbol{\beta} + \mathbf{W}^{-1}(\mathbf{y} - \boldsymbol{\pi})$, where $\mathbf{m} = (m_1, m_2, \dots, m_n)^T$.

Here, the estimation of the LPLRM (2.1) is called LPLE. In summary, the estimation algorithm for the model (2.1) is given as follows

Step 1: Give suitable initial values $\boldsymbol{\beta}^{(0)}$, $\mathbf{m}^{(0)}$ for $\boldsymbol{\beta}$ and \mathbf{m} .

Step 2: Repeat step (a), (b) and (c) until convergence.

(a) Calculate $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Z}}$;

(b) Updating step for $\boldsymbol{\beta}$: $\boldsymbol{\beta}^{new} = (\tilde{\mathbf{X}}^T \mathbf{W} \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^T \mathbf{W} \tilde{\mathbf{Z}}$;

(c) Updating step for \mathbf{m} : $\mathbf{m}^{new} = \mathbf{S}(\mathbf{Z} - \mathbf{X}\boldsymbol{\beta}^{new})$.

Step 3: Obtain the final estimates $\hat{\boldsymbol{\beta}}$, $\hat{\mathbf{m}}_{\hat{\boldsymbol{\beta}}}$ of $\boldsymbol{\beta}$ and \mathbf{m} .

2.2. Evaluation criteria for related estimators of unknown parameters

There are two common evaluation criteria for estimators, namely, the asymptotic scalar mean squared error (SMSE) and the asymptotic mean square error matrix (MSEM). The SMSE are traces of the MSEM of an estimator $\tilde{\boldsymbol{\beta}}$, and they are defined as

$$\begin{aligned} \text{MSEM}(\tilde{\boldsymbol{\beta}}) &= E(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})^T = \text{Cov}(\tilde{\boldsymbol{\beta}}) + \text{Bias}(\tilde{\boldsymbol{\beta}})(\text{Bias}(\tilde{\boldsymbol{\beta}}))^T, \\ \text{SMSE}(\tilde{\boldsymbol{\beta}}) &= \text{tr}(\text{MSEM}(\tilde{\boldsymbol{\beta}})) = \text{tr}(\text{Cov}(\tilde{\boldsymbol{\beta}})) + (\text{Bias}(\tilde{\boldsymbol{\beta}}))^T (\text{Bias}(\tilde{\boldsymbol{\beta}})), \end{aligned}$$

where $\text{Cov}(\cdot)$ is the covariance matrix, $\text{Bias}(\cdot)$ is the biasing vector, $\text{tr}(\cdot)$ denote trace operation. We denote $\alpha = \mathbf{Q}^T \boldsymbol{\beta}$, $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p) = \mathbf{Q}^T (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}}) \mathbf{Q}$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ are the ordered eigenvalues of $\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}}$, $\hat{\mathbf{W}}$ is the final estimate of \mathbf{W} in the above algorithm, \mathbf{Q} is the orthogonal matrix consisting of the eigenvectors of $\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}}$ in columns, α_j denote the j th element of $\mathbf{Q}^T \boldsymbol{\beta}$, $j = 1, 2, \dots, p$. Since the LPLE is asymptotically unbiased [1], the MSEM and the SMSE of $\hat{\boldsymbol{\beta}}$ are given by

$$\text{MSEM}(\hat{\boldsymbol{\beta}}) = \text{Cov}(\hat{\boldsymbol{\beta}}) = (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}})^{-1} = \mathbf{Q} \boldsymbol{\Lambda}^{-1} \mathbf{Q}^T, \quad (2.7)$$

$$\text{SMSE}(\hat{\boldsymbol{\beta}}) = \text{tr}(\text{Cov}(\hat{\boldsymbol{\beta}})) = \text{tr}((\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}})^{-1}) = \sum_{j=1}^p \frac{1}{\lambda_j}.$$

3. Construction of the proposed estimator

In this section, using the PLM, we introduce a GLTE which includes the LPLE, the ridge estimator, the Liu estimator and the Liu-type estimator. We still use some denotation same as that used in Section 2 such as \mathbf{Q} , $\boldsymbol{\Lambda}$, $\hat{\mathbf{W}}$, \mathbf{Z} , $\tilde{\mathbf{Z}}$, $\tilde{\mathbf{X}}$, \mathbf{X} without change of meanings. We begin to consider how to combat a multicollinearity of the LPLRM, since in the presence of multicollinearity, the matrix $\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}}$ becomes ill-conditioned, which leads to a large variance and instability of LPLE though the LPLE is asymptotically unbiased [1]. Inspired by the famous ridge estimation method (a popular penalty estimation used for dealing with multicollinearity), we attempt to construct an appropriate biased estimation to eliminate the multicollinearity in LPLRM and meanwhile the LPLRM are well estimated in accordance with some standard statistical criteria. Using the PLM, in first step, we will simply extend the existing biased estimators such as the ridge estimator, the Liu estimator and the Liu-type estimator to the case of the LPLRM. In second step, we take a complex procedure in which we modify the biasing parameters and the function of biasing parameter, and introduce an objective function of $\boldsymbol{\beta}$ to construct the GLTE.

3.1. Logistic partial linear ridge estimator

Now, using LPLE, we adopt the ridge estimator [6] to construct logistic partial linear ridge estimator (LPLRE) as follows

$$\hat{\boldsymbol{\beta}}_R(k) = (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + k\mathbf{I})^{-1} \tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{Z}} = (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + k\mathbf{I})^{-1} (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}}) \hat{\boldsymbol{\beta}},$$

where $k \geq 0$ is biasing parameter, \mathbf{I} is the identity matrix with dimension $p \times p$. The estimates of the nonparametric functions \mathbf{m} are given by

$$\hat{\mathbf{m}}_{\hat{\boldsymbol{\beta}}_R}(k) = \mathbf{S}(\mathbf{Z} - \mathbf{X} \hat{\boldsymbol{\beta}}_R(k)).$$

The Bias, Cov, MSEM and SMSE of $\hat{\boldsymbol{\beta}}_R(k)$ are given by

$$\begin{aligned} \text{Bias}(\hat{\boldsymbol{\beta}}_R(k)) &= E(\hat{\boldsymbol{\beta}}_R(k)) - \boldsymbol{\beta} = -k(\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + k\mathbf{I})^{-1} \boldsymbol{\beta}, \\ \text{Cov}(\hat{\boldsymbol{\beta}}_R(k)) &= \text{Cov}((\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + k\mathbf{I})^{-1} (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}}) \hat{\boldsymbol{\beta}}) \\ &= (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + k\mathbf{I})^{-1} \tilde{\mathbf{X}} \hat{\mathbf{W}} \tilde{\mathbf{X}}^T ((\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + k\mathbf{I})^{-1})^T, \end{aligned}$$

$$\begin{aligned} \text{MSEM}(\hat{\beta}_R(k)) &= \text{Cov}(\hat{\beta}_R(k)) + (\text{Bias}(\hat{\beta}_R(k)))(\text{Bias}(\hat{\beta}_R(k)))^T \\ &= \mathbf{Q}((\mathbf{\Lambda} + k\mathbf{I})^{-1}\mathbf{\Lambda}(\mathbf{\Lambda} + k\mathbf{I})^{-1} + k^2(\mathbf{\Lambda} + k\mathbf{I})^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}^T(\mathbf{\Lambda} + k\mathbf{I})^{-1})\mathbf{Q}^T, \end{aligned} \quad (3.1)$$

$$\begin{aligned} \text{SMSE}(\hat{\beta}_R(k)) &= \text{tr}(\text{Cov}(\hat{\beta}_R(k))) + (\text{Bias}(\hat{\beta}_R(k)))^T (\text{Bias}(\hat{\beta}_R(k))) \\ &= \sum_{j=1}^p \frac{\lambda_j}{(\lambda_j + k)^2} + \sum_{j=1}^p \frac{k^2 \alpha_j^2}{(\lambda_j + k)^2}. \end{aligned}$$

3.2. Logistic partial linear Liu estimator

Liu estimator was firstly proposed for dealing with a multicollinearity problem in a linear regression model by Liu [7]. Here, using LPLE, we adopt the Liu estimator to construct logistic partial linear Liu estimator (LPLLE) as follows

$$\hat{\beta}_L(d) = (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + \mathbf{I})^{-1} (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{Z}} + d\hat{\beta}) = (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + \mathbf{I})^{-1} (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + d\mathbf{I})\hat{\beta},$$

where $0 < d < 1$ is a biasing parameter. In addition, the estimates of nonparametric functions m are given by

$$\hat{m}_{\hat{\beta}_L}(d) = S(\mathbf{Z} - \mathbf{X}\hat{\beta}_L(d)).$$

The Bias, Cov, MSEM and SMSE of $\hat{\beta}_L(d)$ are given by

$$\begin{aligned} \text{Bias}(\hat{\beta}_L(d)) &= E(\hat{\beta}_L(d)) - \beta \\ &= (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + \mathbf{I})^{-1} (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + d\mathbf{I})\beta - (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + \mathbf{I})^{-1} (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + \mathbf{I})\beta \\ &= (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + \mathbf{I})^{-1} ((d-1)\mathbf{I})\beta, \end{aligned}$$

$$\begin{aligned} \text{Cov}(\hat{\beta}_L(d)) &= \text{Cov}((\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + \mathbf{I})^{-1} (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + d\mathbf{I})\hat{\beta}) \\ &= (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + \mathbf{I})^{-1} (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + d\mathbf{I})(\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}})^{-1} (\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + d\mathbf{I})^T \\ &\quad ((\tilde{\mathbf{X}}^T \hat{\mathbf{W}} \tilde{\mathbf{X}} + \mathbf{I})^{-1})^T, \end{aligned}$$

$$\begin{aligned} \text{MSEM}(\hat{\beta}_L(d)) &= \text{Cov}(\hat{\beta}_L(d)) + (\text{Bias}(\hat{\beta}_L(d)))(\text{Bias}(\hat{\beta}_L(d)))^T \\ &= \mathbf{Q}((\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{\Lambda} + d\mathbf{I})\mathbf{\Lambda}^{-1}(\mathbf{\Lambda} + d\mathbf{I})(\mathbf{\Lambda} + \mathbf{I})^{-1} + (d-1)^2(\mathbf{\Lambda} + \mathbf{I})^{-1} \\ &\quad \boldsymbol{\alpha}\boldsymbol{\alpha}^T(\mathbf{\Lambda} + \mathbf{I})^{-1})\mathbf{Q}^T, \end{aligned} \quad (3.2)$$

$$\begin{aligned} \text{SMSE}(\hat{\beta}_L(d)) &= \text{tr}(\text{Cov}(\hat{\beta}_L(d))) + (\text{Bias}(\hat{\beta}_L(d)))^T (\text{Bias}(\hat{\beta}_L(d))) \\ &= \sum_{j=1}^p \frac{(\lambda_j + d)^2}{\lambda_j(\lambda_j + 1)^2} + \sum_{j=1}^p \frac{(d-1)^2 \alpha_j^2}{(\lambda_j + 1)^2}. \end{aligned} \quad (3.3)$$

3.3. Logistic partial linear Liu-type estimator

Liu [8], Inan and Erdogon [13] introduced a Liu-type estimator to modify Liu estimator when they deal with the multicollinearity problem in the LLRM. Using LPLE, we adopt the Liu-type estimator to construct the logistic partial linear Liu-type estimator (LPLLTE) as follows

$$\begin{aligned}\hat{\beta}_{LT}(d, k) &= (\tilde{X}^T \hat{W} \tilde{X} + kI)^{-1} (\tilde{X}^T \hat{W} \tilde{Z} - d\hat{\beta}) \\ &= (\tilde{X}^T \hat{W} \tilde{X} + kI)^{-1} (\tilde{X}^T \hat{W} \tilde{X} - dI)\hat{\beta},\end{aligned}$$

where $k > 0$ and $d \in \mathcal{R}$ are biasing parameters. The estimates of the nonparametric functions m are given by

$$\hat{m}_{\hat{\beta}_{LT}}(d, k) = S(Z - X\hat{\beta}_{LT}(d, k)).$$

The MSEM and SMSE of $\hat{\beta}_{LT}(d, k)$ are given as follows

$$\begin{aligned}\text{MSEM}(\hat{\beta}_{LT}(d, k)) &= Q((\Lambda + kI)^{-1}(\Lambda - dI)\Lambda^{-1}(\Lambda - dI)(\Lambda + kI)^{-1} \\ &\quad + (d + k)^2(\Lambda + kI)^{-1}\alpha\alpha^T(\Lambda + kI)^{-1})Q^T,\end{aligned}\quad (3.4)$$

$$\text{SMSE}(\hat{\beta}_{LT}(d, k)) = \sum_{j=1}^p \frac{(\lambda_j - d)^2}{\lambda_j(\lambda_j + k)^2} + \sum_{j=1}^p \frac{(d + k)^2 \alpha_j^2}{(\lambda_j + k)^2}.\quad (3.5)$$

3.4. Proposed generalized Liu-type estimator

Kurnaz and Akay [9], Ertan and Akay [16] introduced some new Liu-type estimators when they deal with multicollinearity of linear regression model and LLRM. The new Liu-type estimators are defined as

$$\hat{\beta}_{E\&A} = (X^T \hat{W} X + kI)^{-1} (X^T \hat{W} X + f(k)I)\hat{\beta}^*, \quad k > 0,$$

and

$$\hat{\beta}_{K\&A} = (X^T X + kI)^{-1} (X^T X + f(k)I)\hat{\beta}^*, \quad k > 0,$$

where $\hat{\beta}^*$ is any estimator of the regression coefficient vector β , k is a biasing parameter and $f(k)$ is a continuous function of biasing parameter k .

Using LPLE, we modify the constructions of the new Liu-type estimators $\hat{\beta}_{E\&A}$, $\hat{\beta}_{K\&A}$ to propose a new estimator which is called a generalized Liu-type estimator (GLTE). Setting the following objective function

$$\begin{aligned}F(\beta) &= (\tilde{Z} - \tilde{X}\beta)^T \hat{W} (\tilde{Z} - \tilde{X}\beta) + ((QKQ^T)^{-1}(QFQ^T)\hat{\beta} - \beta)^T QKQ^T \\ &\quad ((QKQ^T)^{-1}(QFQ^T)\hat{\beta} - \beta),\end{aligned}\quad (3.6)$$

where $K = \text{diag}(k_1, k_2, \dots, k_p)$ is the matrix of biasing parameters, $F = \text{diag}(f(k_1), f(k_2), \dots, f(k_p))$ is the matrix of function $f(k_j)$ of biasing parameter, $f(k_j) = ak_j + b$, $k_j > 0$, $j = 1, \dots, p$, a and b are

constants (refer to section 5). Minimize the function (3.6) with respect to β , we can obtain the GLTE as follows

$$\begin{aligned}\hat{\beta}_G(\mathbf{K}) &= (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1} (\tilde{X}^T \hat{W} \tilde{Z} + \mathbf{QFQ}^T \hat{\beta}) \\ &= (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1} (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QFQ}^T) \hat{\beta}.\end{aligned}\quad (3.7)$$

In addition, the estimates of the nonparametric functions m are given by

$$\hat{m}_{\hat{\beta}_G}(\mathbf{K}) = S(\mathbf{Z} - \mathbf{X}\hat{\beta}_G(\mathbf{K})).$$

Remark 1. In our proposed construction (3.7) of GLTE, a group of biasing parameters is considered, which is the key difference from the constructions of the new Liu-type estimators by Kurnaz and Akay [9], Ertan and Akay [16]. We add the matrix of biasing parameters and the matrix of function of biasing parameters in order to ensure that the adjustment of the estimate of each component of the vector β can be different.

Remark 2. The GLTE is a more generalized estimator which includes the other estimators as special cases:

- (i) $\hat{\beta}_G(\mathbf{K}) = \hat{\beta}$, for $f(k_j) = k_j$, where $a = 1$ and $b = 0$;
- (ii) $\hat{\beta}_G(\mathbf{K}) = \hat{\beta}_R(k)$, for $\mathbf{K} = k\mathbf{I}$, $f(k_j) = 0$, where $a = 0$ and $b = 0$;
- (iii) $\hat{\beta}_G(\mathbf{K}) = \hat{\beta}_L(d)$, for $\mathbf{K} = \mathbf{I}$, $f(k_j) = d$, where $a = 0$ and $b = d$;
- (iv) $\hat{\beta}_G(\mathbf{K}) = \hat{\beta}_{LT}(d, k)$, for $\mathbf{K} = k\mathbf{I}$, $f(k_j) = -d$, where $a = 0$ and $b = -d$.

The Bias, Cov, MSEM and SMSE of $\hat{\beta}_G(\mathbf{K})$ are given by

$$\begin{aligned}\text{Bias}(\hat{\beta}_G(\mathbf{K})) &= E(\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1} (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QFQ}^T) \hat{\beta}_M - \beta \\ &= E(\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1} (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QFQ}^T) (\tilde{X}^T \hat{W} \tilde{X})^{-1} \tilde{X}^T \hat{W} \tilde{Z} - \beta \\ &= (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1} (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QFQ}^T) (\tilde{X}^T \hat{W} \tilde{X})^{-1} \tilde{X}^T \hat{W} \tilde{X} \beta - \beta \\ &= (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1} (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QFQ}^T) \beta - \beta \\ &= (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1} \mathbf{Q}(\mathbf{F} - \mathbf{K})\mathbf{Q}^T \beta, \\ \text{Cov}(\hat{\beta}_G(\mathbf{K})) &= \text{Cov}((\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1} (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QFQ}^T) \hat{\beta}_M) \\ &= (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1} (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QFQ}^T) (\tilde{X}^T \hat{W} \tilde{X})^{-1} \\ &\quad (\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QFQ}^T)^T ((\tilde{X}^T \hat{W} \tilde{X} + \mathbf{QKQ}^T)^{-1})^T, \\ \text{MSEM}(\hat{\beta}_G(\mathbf{K})) &= \text{Cov}(\hat{\beta}_G(\mathbf{K})) + (\text{Bias}(\hat{\beta}_G(\mathbf{K}))) (\text{Bias}(\hat{\beta}_G(\mathbf{K})))^T \\ &= \mathbf{Q}((\Lambda + \mathbf{K})^{-1} (\Lambda + \mathbf{F}) \Lambda^{-1} (\Lambda + \mathbf{F}) (\Lambda + \mathbf{K})^{-1} + (\Lambda + \mathbf{K})^{-1} \\ &\quad (\mathbf{F} - \mathbf{K}) \alpha \alpha^T (\mathbf{F} - \mathbf{K}) (\Lambda + \mathbf{K})^{-1}) \mathbf{Q}^T, \\ \text{SMSE}(\hat{\beta}_G(\mathbf{K})) &= \text{tr}(\text{Cov}(\hat{\beta}_G(\mathbf{K}))) + (\text{Bias}(\hat{\beta}_G(\mathbf{K})))^T (\text{Bias}(\hat{\beta}_G(\mathbf{K}))) \\ &= \sum_{j=1}^p \frac{(\lambda_j + f(k_j))^2}{\lambda_j(\lambda_j + k_j)^2} + \sum_{j=1}^p \frac{(f(k_j) - k_j)^2 \alpha_j^2}{(\lambda_j + k_j)^2}.\end{aligned}\quad (3.8)$$

4. The superiority of the proposed GLTE

In this section, we compare the performance of the proposed GLTE with that of the LPLE, the ridge estimator, the Liu estimator and the Liu-type estimator under MSEM criterion.

Let $\hat{\beta}_1$ and $\hat{\beta}_2$ be any two estimators of the vector β . From literature [26], we know that $\hat{\beta}_2$ is superior to $\hat{\beta}_1$ with respect to the MSEM criteria if and only if (iff) $\text{MSEM}(\hat{\beta}_1) - \text{MSEM}(\hat{\beta}_2)$ is a positive definite (*p.d.*) matrix. If $\text{MSEM}(\hat{\beta}_1) - \text{MSEM}(\hat{\beta}_2)$ is a non-negative definite matrix, then $\text{SMSE}(\hat{\beta}_1) - \text{SMSE}(\hat{\beta}_2) \geq 0$, but the inverse is not true. To compare the superiority of GLTE $\hat{\beta}_G(\mathbf{K})$, we will use the following lemma.

Lemma 1. (Farebrother [27]) Let \mathbf{A} be a *p.d.* matrix, namely $\mathbf{A} > 0$, and \mathbf{c} be nonzero vector. Then, $\mathbf{A} - \mathbf{c}\mathbf{c}^T$ is *p.d.* matrix, iff $\mathbf{c}^T \mathbf{A}^{-1} \mathbf{c} \leq 1$.

For clarity, we use the following abbreviations, $\Lambda_1 = \Lambda + \mathbf{I}$, $\Lambda_d = \Lambda + d\mathbf{I}$, $\Lambda_{-d} = \Lambda - d\mathbf{I}$, $\Lambda_k = \Lambda + k\mathbf{I}$, $\Lambda_K = \Lambda + \mathbf{K}$, $\Lambda_F = \Lambda + \mathbf{F}$. We give the following theorems to show the superiority of $\hat{\beta}_G(\mathbf{K})$ over $\hat{\beta}$, $\hat{\beta}_R(k)$, $\hat{\beta}_L(d)$, $\hat{\beta}_{LT}(d, k)$.

Theorem 1. Let be $k_j > 0$ and $-2\lambda_j - k_j < f(k_j) < k_j$, $j = 1, \dots, p$. Then $\text{MSEM}(\hat{\beta}) - \text{MSEM}(\hat{\beta}_G(\mathbf{K})) > 0$ iff

$$\left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right)^T \mathbf{Q}(\Lambda^{-1} - \Lambda_K^{-1} \Lambda_F \Lambda^{-1} \Lambda_F \Lambda_K^{-1}) \mathbf{Q}^T \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right) < 1,$$

where $\text{Bias}(\hat{\beta}_G(\mathbf{K})) = \mathbf{Q}\Lambda_K^{-1}(\mathbf{F} - \mathbf{K})\alpha$.

Proof. From Eqs (2.7) and (3.8), we can immediately obtain the difference between the MSEM of $\hat{\beta}$ and $\hat{\beta}_G(\mathbf{K})$ as follows

$$\begin{aligned} \Delta_1 &= \text{MSEM}(\hat{\beta}) - \text{MSEM}(\hat{\beta}_G(\mathbf{K})) \\ &= \mathbf{Q}(\Lambda^{-1} - \Lambda_K^{-1} \Lambda_F \Lambda^{-1} \Lambda_F \Lambda_K^{-1}) \mathbf{Q}^T - \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right) \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right)^T \\ &= \mathbf{Q} \text{diag} \left\{ \frac{1}{\lambda_j} - \frac{(\lambda_j + f(k_j))^2}{\lambda_j(\lambda_j + k_j)^2} \right\}_{j=1}^p \mathbf{Q}^T - \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right) \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right)^T. \end{aligned}$$

The $\Lambda^{-1} - \Lambda_K^{-1} \Lambda_F \Lambda^{-1} \Lambda_F \Lambda_K^{-1}$ is the *p.d.* matrix if $(2\lambda_j + k_j + f(k_j))(k_j - f(k_j)) > 0$, $j = 1, \dots, p$. Since $k_j > 0$, $j = 1, \dots, p$, this condition becomes $-2\lambda_j - k_j < f(k_j) < k_j$. By Lemma 1, the Theorem 1 is proved. \square

Theorem 2. Let be $k > 0$, $k_j > 0$ and $-\lambda_j - \frac{\lambda_j(\lambda_j + k_j)}{\lambda_j + k} < f(k_j) < -\lambda_j + \frac{\lambda_j(\lambda_j + k_j)}{\lambda_j + k}$, $j = 1, \dots, p$. Then $\text{MSEM}(\hat{\beta}_R(k)) - \text{MSEM}(\hat{\beta}_G(\mathbf{K})) > 0$ iff

$$\left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right)^T \mathbf{Q}(\Lambda_k^{-1} \Lambda \Lambda_k^{-1} - \Lambda_K^{-1} \Lambda_F \Lambda^{-1} \Lambda_F \Lambda_K^{-1}) \mathbf{Q}^T \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right) < 1,$$

where $\text{Bias}(\hat{\beta}_G(\mathbf{K})) = \mathbf{Q}\Lambda_K^{-1}(\mathbf{F} - \mathbf{K})\alpha$.

Proof. From Eqs (3.1) and (3.8), we can immediately obtain the difference between the MSEM of

$\hat{\beta}_R(k)$ and $\hat{\beta}_G(\mathbf{K})$ as follows

$$\begin{aligned}\Delta_2 &= \text{MSEM}(\hat{\beta}_R(k)) - \text{MSEM}(\hat{\beta}_G(\mathbf{K})) \\ &= \mathbf{Q}(\Lambda_k^{-1}\Lambda\Lambda_k^{-1} - \Lambda_K^{-1}\Lambda_F\Lambda^{-1}\Lambda_F\Lambda_K^{-1})\mathbf{Q}^T + (\text{Bias}(\hat{\beta}_R(k)))(\text{Bias}(\hat{\beta}_R(k)))^T \\ &\quad - (\text{Bias}(\hat{\beta}_G(\mathbf{K})))(\text{Bias}(\hat{\beta}_G(\mathbf{K})))^T \\ &= \mathbf{Q}\text{diag}\left\{\frac{\lambda_j}{(\lambda_j+k)^2} - \frac{(\lambda_j+f(k_j))^2}{\lambda_j(\lambda_j+k_j)^2}\right\}_{j=1}^p \mathbf{Q}^T + (\text{Bias}(\hat{\beta}_R(k)))(\text{Bias}(\hat{\beta}_R(k)))^T \\ &\quad - (\text{Bias}(\hat{\beta}_G(\mathbf{K})))(\text{Bias}(\hat{\beta}_G(\mathbf{K})))^T.\end{aligned}$$

Since $(\text{Bias}(\hat{\beta}_R(k)))(\text{Bias}(\hat{\beta}_R(k)))^T$ is the *p.d.* matrix, we just need to prove that $\mathbf{Q}(\Lambda_k^{-1}\Lambda\Lambda_k^{-1} - \Lambda_K^{-1}\Lambda_F\Lambda^{-1}\Lambda_F\Lambda_K^{-1})\mathbf{Q}^T - (\text{Bias}(\hat{\beta}_G(\mathbf{K})))(\text{Bias}(\hat{\beta}_G(\mathbf{K})))^T$ is the *p.d.* matrix. The $\Lambda_k^{-1}\Lambda\Lambda_k^{-1} - \Lambda_K^{-1}\Lambda_F\Lambda^{-1}\Lambda_F\Lambda_K^{-1}$ is the *p.d.* matrix if $\left(\frac{\lambda_j}{\lambda_j+k} + \frac{\lambda_j+f(k_j)}{\lambda_j+k_j}\right)\left(\frac{\lambda_j}{\lambda_j+k} - \frac{\lambda_j+f(k_j)}{\lambda_j+k_j}\right) > 0$. Since $k > 0, k_j > 0, j = 1, \dots, p$, this condition becomes $-\lambda_j - \frac{\lambda_j(\lambda_j+k_j)}{\lambda_j+k} < f(k_j) < -\lambda_j + \frac{\lambda_j(\lambda_j+k_j)}{\lambda_j+k}$. By Lemma 1, the Theorem 2 is proved. \square

Theorem 3. Let be $0 < d < 1, k_j > 0$ and $-\lambda_j - \frac{(\lambda_j+d)(\lambda_j+k_j)}{\lambda_j+1} < f(k_j) < -\lambda_j + \frac{(\lambda_j+d)(\lambda_j+k_j)}{\lambda_j+1}, j = 1, \dots, p$. Then $\text{MSEM}(\hat{\beta}_L(d)) - \text{MSEM}(\hat{\beta}_G(\mathbf{K})) > 0$ iff

$$(\text{Bias}(\hat{\beta}_G(\mathbf{K})))^T \mathbf{Q}(\Lambda_1^{-1}\Lambda_d\Lambda^{-1}\Lambda_d\Lambda_1^{-1} - \Lambda_K^{-1}\Lambda_F\Lambda^{-1}\Lambda_F\Lambda_K^{-1})\mathbf{Q}^T (\text{Bias}(\hat{\beta}_G(\mathbf{K}))) < 1,$$

where $\text{Bias}(\hat{\beta}_G(\mathbf{K})) = \mathbf{Q}\Lambda_K^{-1}(\mathbf{F} - \mathbf{K})\alpha$.

Proof. From Eqs (3.2) and (3.8), we can immediately obtain the difference between the MSEM of $\hat{\beta}_L(d)$ and $\hat{\beta}_G(\mathbf{K})$ as follows

$$\begin{aligned}\Delta_3 &= \text{MSEM}(\hat{\beta}_L(d)) - \text{MSEM}(\hat{\beta}_G(\mathbf{K})) \\ &= \mathbf{Q}(\Lambda_1^{-1}\Lambda_d\Lambda^{-1}\Lambda_d\Lambda_1^{-1} - \Lambda_K^{-1}\Lambda_F\Lambda^{-1}\Lambda_F\Lambda_K^{-1})\mathbf{Q}^T + (\text{Bias}(\hat{\beta}_L(d)))(\text{Bias}(\hat{\beta}_L(d)))^T \\ &\quad - (\text{Bias}(\hat{\beta}_G(\mathbf{K})))(\text{Bias}(\hat{\beta}_G(\mathbf{K})))^T \\ &= \mathbf{Q}\text{diag}\left\{\frac{(\lambda_j+d)^2}{\lambda_j(\lambda_j+1)^2} - \frac{(\lambda_j+f(k_j))^2}{\lambda_j(\lambda_j+k_j)^2}\right\}_{j=1}^p \mathbf{Q}^T + (\text{Bias}(\hat{\beta}_L(d)))(\text{Bias}(\hat{\beta}_L(d)))^T \\ &\quad - (\text{Bias}(\hat{\beta}_G(\mathbf{K})))(\text{Bias}(\hat{\beta}_G(\mathbf{K})))^T.\end{aligned}$$

Since $(\text{Bias}(\hat{\beta}_L(d)))(\text{Bias}(\hat{\beta}_L(d)))^T$ is the *p.d.* matrix, we just need to prove that $\mathbf{Q}(\Lambda_1^{-1}\Lambda_d\Lambda^{-1}\Lambda_d\Lambda_1^{-1} - \Lambda_K^{-1}\Lambda_F\Lambda^{-1}\Lambda_F\Lambda_K^{-1})\mathbf{Q}^T - (\text{Bias}(\hat{\beta}_G(\mathbf{K})))(\text{Bias}(\hat{\beta}_G(\mathbf{K})))^T$ is the *p.d.* matrix. The $\Lambda_1^{-1}\Lambda_d\Lambda^{-1}\Lambda_d\Lambda_1^{-1} - \Lambda_K^{-1}\Lambda_F\Lambda^{-1}\Lambda_F\Lambda_K^{-1}$ is the *p.d.* matrix if $\left(\frac{\lambda_j+d}{\lambda_j+1} + \frac{\lambda_j+f(k_j)}{\lambda_j+k_j}\right)\left(\frac{\lambda_j+d}{\lambda_j+1} - \frac{\lambda_j+f(k_j)}{\lambda_j+k_j}\right) > 0$. Since $0 < d < 1, k_j > 0, j = 1, \dots, p$, this condition becomes $-\lambda_j - \frac{(\lambda_j+d)(\lambda_j+k_j)}{\lambda_j+1} < f(k_j) < -\lambda_j + \frac{(\lambda_j+d)(\lambda_j+k_j)}{\lambda_j+1}, j = 1, \dots, p$. By Lemma 1, the Theorem 3 is proved. \square

Theorem 4. Let be $k > 0$, $d \in R$, $k_j > 0$ and $-\lambda_j - \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k} < f(k_j) < -\lambda_j + \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k}$, or $-\lambda_j + \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k} < f(k_j) < -\lambda_j - \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k}$, $j = 1, \dots, p$. Then $\text{MSEM}(\hat{\beta}_{LT}(d, k)) - \text{MSEM}(\hat{\beta}_G(\mathbf{K})) > 0$ iff

$$\left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right)^T \mathbf{Q}(\Lambda_k^{-1} \Lambda_{-d} \Lambda^{-1} \Lambda_{-d} \Lambda_k^{-1} - \Lambda_K^{-1} \Lambda_F \Lambda^{-1} \Lambda_F \Lambda_K^{-1}) \mathbf{Q}^T \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right) < 1,$$

where $\text{Bias}(\hat{\beta}_G(\mathbf{K})) = \mathbf{Q} \Lambda_K^{-1} (\mathbf{F} - \mathbf{K}) \alpha$.

Proof. From Eqs (3.4) and (3.8), we can immediately obtain the difference between the MSEM of $\hat{\beta}_{LT}(d, k)$ and $\hat{\beta}_G(\mathbf{K})$ as follows

$$\begin{aligned} \Delta_4 &= \text{MSEM}(\hat{\beta}_{LT}(d, k)) - \text{MSEM}(\hat{\beta}_G(\mathbf{K})) \\ &= \mathbf{Q}(\Lambda_k^{-1} \Lambda_{-d} \Lambda^{-1} \Lambda_{-d} \Lambda_k^{-1} - \Lambda_K^{-1} \Lambda_F \Lambda^{-1} \Lambda_F \Lambda_K^{-1}) \mathbf{Q}^T + \left(\text{Bias}(\hat{\beta}_{LL}(d, k))\right) \left(\text{Bias}(\hat{\beta}_{LL}(d, k))\right)^T \\ &\quad - \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right) \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right)^T \\ &= \mathbf{Q} \text{diag} \left\{ \frac{(\lambda_j - d)^2}{\lambda_j(\lambda_j + k)^2} - \frac{(\lambda_j + f(k_j))^2}{\lambda_j(\lambda_j + k_j)^2} \right\}_{j=1}^p \mathbf{Q}^T + \left(\text{Bias}(\hat{\beta}_{LL}(d, k))\right) \left(\text{Bias}(\hat{\beta}_{LL}(d, k))\right)^T \\ &\quad - \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right) \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right)^T. \end{aligned}$$

Since $\left(\text{Bias}(\hat{\beta}_{LL}(d, k))\right) \left(\text{Bias}(\hat{\beta}_{LL}(d, k))\right)^T$ is the *p.d.* matrix, we just need to prove that $\mathbf{Q}(\Lambda_k^{-1} \Lambda_{-d} \Lambda^{-1} \Lambda_{-d} \Lambda_k^{-1} - \Lambda_K^{-1} \Lambda_F \Lambda^{-1} \Lambda_F \Lambda_K^{-1}) \mathbf{Q}^T - \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right) \left(\text{Bias}(\hat{\beta}_G(\mathbf{K}))\right)^T$ is the *p.d.* matrix. The $\Lambda_k^{-1} \Lambda_{-d} \Lambda^{-1} \Lambda_{-d} \Lambda_k^{-1} - \Lambda_K^{-1} \Lambda_F \Lambda^{-1} \Lambda_F \Lambda_K^{-1}$ is the *p.d.* matrix if $\left(\frac{\lambda_j-d}{\lambda_j+k} + \frac{\lambda_j+f(k_j)}{\lambda_j+k_j}\right) \left(\frac{\lambda_j-d}{\lambda_j+k} - \frac{\lambda_j+f(k_j)}{\lambda_j+k_j}\right) > 0$. Since $k > 0, d \in R, k_j > 0, j = 1, \dots, p$, this condition becomes $-\lambda_j - \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k} < f(k_j) < -\lambda_j + \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k}$, or $-\lambda_j + \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k} < f(k_j) < -\lambda_j - \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k}$, $j = 1, \dots, p$. By Lemma 1, the Theorem 4 is proved. \square

5. Determinations of biasing parameters and function of the biasing parameter

5.1. Estimators of function $f(k_j)$ of the biasing parameter and the biasing parameter k_j in GLTE ($j = 1, \dots, p$)

To make GLTE more effective, it is important to make the appropriate choices of function $f(k_j)$ of the biasing parameter and the biasing parameter k_j . Inspired by Kurnaz and Akay [9], Ertan and Akay [16], we use the minimum SMSE criterion to select the best function of $f(k_j)$ and the optimal values of k_j , $j = 1, \dots, p$.

Note that, the SMSE of $\hat{\beta}_G(\mathbf{K})$ is a nonlinear function of k_j , $j = 1, \dots, p$ and denoted by

$$g(k_1, k_2, \dots, k_p) = \sum_{j=1}^p \frac{(\lambda_j + f(k_j))^2}{\lambda_j(\lambda_j + k_j)^2} + \sum_{j=1}^p \frac{(f(k_j) - k_j)^2 \alpha_j^2}{(\lambda_j + k_j)^2}.$$

We obtain partial derivative of $g(k_1, k_2, \dots, k_p)$ with respect to $k_j, j = 1, \dots, p$ as follow

$$\frac{\partial g(k_1, k_2, \dots, k_p)}{\partial k_j} = \frac{2((\lambda_j + f(k_j)) + \lambda_j \alpha_j^2 (f(k_j) - k_j))(f'(k_j)(\lambda_j + k_j) - (\lambda_j + f(k_j)))}{\lambda_j (\lambda_j + k_j)^3}, \quad (5.1)$$

after equating the Eq (5.1) to 0, we obtain

$$f(k_j) = \frac{\lambda_j \alpha_j^2}{1 + \lambda_j \alpha_j^2} k_j - \frac{\lambda_j}{1 + \lambda_j \alpha_j^2}, \quad (5.2)$$

or

$$f(k_j) = f'(k_j) k_j + \lambda_j (f'(k_j) - 1), \quad (5.3)$$

where $f'(k_j)$ is the first derivative of $f(k_j)$ with respect to $k_j, j = 1, \dots, p$.

Remark 3. Our construction Eq (3.7) is able to carry out the general form of $f(k_j)$ by minimizing SMSE of $\hat{\beta}_G(\mathbf{K})$, instead of the two special cases Fact1, Fact2 in Kurnaz and Akay [9], Ertan and Akay [16]. This is due to the fact that there is no summation after taking the partial derivative, see Eq (5.1).

In Eq (5.2), we know that the first derivative $f'(k_j) = \frac{\lambda_j \alpha_j^2}{1 + \lambda_j \alpha_j^2}$, and both $\frac{\lambda_j \alpha_j^2}{1 + \lambda_j \alpha_j^2}$ and $-\frac{\lambda_j}{1 + \lambda_j \alpha_j^2}$ are constants, $f(k_j)$ is a linear function. In Eq (5.3), when $f'(k_j) = \frac{\lambda_j \alpha_j^2}{1 + \lambda_j \alpha_j^2}$, $f(k_j)$ is also a linear function. Thus, Eqs (5.2) and (5.3) can be unified as a linear function $f(k_j) = ak_j + b, j = 1, \dots, p$. Substituting $f(k_j) = ak_j + b$ into Eq (5.1) and making it equal to 0, we can obtain the optimal value of k_j as follows,

$$\hat{k}_j = \frac{-\lambda_j - b(1 + \lambda_j \hat{\alpha}_j^2)}{a + (a - 1)\lambda_j \hat{\alpha}_j^2}, j = 1, \dots, p, \quad (5.4)$$

where a and b are constants need to be specified, $\hat{\alpha}_j$ denote the j th element of $\mathbf{Q}^T \hat{\beta}$. From Eq (5.2), we know that $0 < a < 1$ and $b < 0$. From Eq (5.3), we know that $b = \lambda_j(a - 1)$. In practice, we might as well take $a = \tau$ and $b = \lambda_{\min}(\tau - 1)$, $\tau \in (0, 1)$. However, in Eq (5.4), \hat{k}_j may be negative. Thus, it is better to use the following estimator

$$\hat{k}_{G_j} = \max\{0, \hat{k}_j\}, j = 1, \dots, p, \quad (5.5)$$

the estimator of $f(k_j)$ is $\hat{f}(k_j) = \tau + \lambda_{\min}(\tau - 1)\hat{k}_{G_j}$. A more detailed discussion of constant τ is given in section 6.

5.2. Estimator of the biasing parameter k in LPLRE

In the LPLRE method, we use three biasing parameter estimators (namely, $\hat{k}_{R1}, \hat{k}_{R2}, \hat{k}_{R3}$) recommended by Kibria et al. [12], which have been verified by simulation study to have better performance than other biasing parameter estimators. These parameter estimators are given as follows

$$\hat{k}_{R1} = \max\left(\frac{1}{m_j}\right), \quad \hat{k}_{R2} = \prod_{j=1}^p \left(\frac{1}{q_j}\right)^{\frac{1}{p}}, \quad \hat{k}_{R3} = \text{median}\left(\frac{1}{q_j}\right),$$

where $\hat{\sigma}^2 = \frac{1}{n-p}(y_i - \hat{\pi}_i)^2, m_j = \sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}_j^2}}, q_j = \frac{\lambda_{\max}}{(n-p)\hat{\sigma}^2 + \lambda_{\max}\hat{\alpha}_j^2}$.

5.3. Estimator of the biasing parameter d in LPLLE

By Liu [7], the optimal estimator of d is obtained by minimizing Eq (3.3) as follows

$$\hat{d}_{opt} = \sum_{j=1}^p \frac{\hat{\alpha}_j^2 - 1}{(\lambda_j + 1)^2} / \sum_{j=1}^p \frac{\hat{\alpha}_j^2 + \frac{1}{\lambda_j}}{(\lambda_j + 1)^2},$$

It is obvious that $\hat{d}_{opt} < 1$, but \hat{d}_{opt} may be negative. Refer to [17], we use the following estimator

$$\hat{d}_L = \max\{0, \hat{d}_{opt}\}.$$

5.4. Estimators of biasing parameters d and k in LPLLTE

Based on the method in [13], for the biasing parameters k and d in $\hat{\beta}_{LT}$, we fix k and find the optimal value of d by minimizing Eq (3.5), which is as follows

$$\hat{d}_{LT} = \sum_{j=1}^p \frac{1 - \hat{\alpha}_j^2 \hat{k}_{LT}}{(\lambda_j + \hat{k}_{LT})^2} / \sum_{j=1}^p \frac{\hat{\alpha}_j^2 + \frac{1}{\lambda_j}}{(\lambda_j + \hat{k}_{LT})^2},$$

with $\hat{k}_{LT} = \frac{p}{\hat{\alpha}'\hat{\alpha}}$.

Remark 4. The above mentioned biasing parameter k_j and the function $f(k_j)$ of the biasing parameter in GLTE, as well as the biasing parameter k in ridge estimator, the biasing parameter d in Liu estimator and the biasing parameters k, d in Liu-type estimator, can also be determined by the generalized cross validation (GCV) criterion (see, [28, 29]). Based on the GCV, we are able to simultaneously select the optimal biasing parameters in the biased estimators and the bandwidth of kernel smoother, and obtain the biased estimators with good performance.

6. Monte Carlo simulations

In this section, we present some simulations to compare the performance of the proposed GLTE with that of the other estimators in the LPLRM. From the sample size (n), the degree of collinearity (r) and the number of explanatory variables in linear part (p), we discuss the performance of different estimators of parameters of linear part, as well as the performance of different estimators of nonparametric function in LPLRM. We generate the explanatory variables with following formula given by Zeinal [10], Varathan and Wijekoon [15], Ertan and Akay [16] as

$$x_{ij} = (1 - r^2)^{\frac{1}{2}} \xi_{ij} + r \xi_{ip}, i = 1, 2, \dots, n, j = 1, 2, \dots, p,$$

where r^2 denotes the correlation between any two design variables, which is specified by $r = 0.70, 0.80, 0.85$, and 0.90 , ξ_{ij} are independent standard normal pseudo-random numbers. The binary response variable y_i is generated from the Bernoulli(π_i) distribution with

$$\pi_i = \frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta} + m(t_i))}{1 + \exp(\mathbf{x}_i^T \boldsymbol{\beta} + m(t_i))} = \frac{\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + m(t_i))}{1 + \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + m(t_i))}, i = 1, \dots, n,$$

where $\beta = (0.1, 0.6, 0.4, 0.7)^T$ or $\beta = (0.2, 0.3, 0.5, 0.4, 0.5, 0.2, 0.4)^T$ so that $\beta^T \beta = 1$. The nonparametric functions are generated by using the design given in [23] as

$$m(t_i) = \sin(t_i)\cos(8t_i), \quad t_i \sim U(0, 1).$$

The sample sizes are taken as $n = 200, 300, 400$, and 500 . The bandwidth vector is computed by Scott's rule of thumb [30] for the Gaussian kernel. The simulation is repeated $M = 2,000$ times with the above setup and the simulated average scalar mean square error (ASMSE) of estimator $(\hat{\beta}^*)$ for linear parametric part $(x_i^T \beta)$ is given by

$$\text{ASMSE}(\hat{\beta}^*) = \frac{1}{M} \sum_{l=1}^M (\hat{\beta}_l^* - \beta)^T (\hat{\beta}_l^* - \beta),$$

where the subscript l indicates the number of repetition, $\hat{\beta}^*$ represents the estimator of the various methods. The simulated mean squared error (mse) of the estimator vector $\hat{m}(t)$ of the nonparametric function $m(t)$ is obtained by using the following equation,

$$\text{mse}(\hat{m}(t), m(t)) = \frac{1}{nM} \sum_{l=1}^M \|\hat{m}_l(t) - m(t)\|_2^2,$$

where $\|v\|_2^2 = \sum_{i=1}^n v_i^2$ for $v = (v_1, \dots, v_n)'$.

6.1. Performance of the LPLE, LPLRE, LPLLE, LPLLTE and GLTE

First, we will evaluate the performance of the LPLE, LPLRE, LPLLE, LPLLTE and GLTE. For different combinations of p, r , and n , the ASMSE of the estimator $(\hat{\beta}^*)$ for each method is shown in Table 1, the mse of the estimator $(\hat{m}(t))$ for each method is shown in Table 2.

Table 1 clearly reveals that the sample size (n), the degree of collinearity (r) and the number of explanatory variables in linear part (p) influence the ASMSE of estimator $(\hat{\beta}^*)$ for each method. When any other two parameters in p, r, n are fixed, the ASMSE of each estimator $(\hat{\beta}^*)$ decreases with the increase of the sample size n , increases with the increase of the degree of collinearity (r), and increases with the increase of the number of explanatory variables for linear part (p). It is observed that the ASMSE of the estimator $(\hat{\beta}_G)$ is smaller than that of the other estimators for all combinations of p, r, n . It implies that the performance of the estimator $(\hat{\beta}_G)$ better than that of the other estimators.

Table 2 indicate that the sample size (n) and the number of explanatory variables in the linear part (p) influence the mse of estimator $\hat{m}(t)$ of nonparametric function for each method, and the concrete situation of the influence is similar to that in Table 1, and the influence degree is much smaller than that in Table 1. However, the mse of the estimators of the nonparametric functions of the proposed method is still smaller than that of other methods, only less obvious than that of the parametric part. The degree of collinearity (r) has some slight effect on the mse of estimator $\hat{m}(t)$ of nonparametric function of each method, but the mse of estimator $\hat{m}(t)$ of nonparametric function under the proposed method is generally at the lowest level. These results reflect that the proposed method also has some superior performance in the estimator of nonparametric functions, but it is not obvious in the parameter part.

It is noted that in Tables 1 and 2, τ values are close to 0, which is not arbitrary. Simulation results and simple theoretical derivation will be explained below.

Table 1. The simulated ASMSE values of $\hat{\beta}, \hat{\beta}_{R_1}, \hat{\beta}_{R_2}, \hat{\beta}_{R_3}, \hat{\beta}_L, \hat{\beta}_{LT}, \hat{\beta}_G$.

r	$p = 4$ & $\tau = 0.03$				$p = 7$ & $\tau = 0.01$			
	0.70	0.80	0.85	0.90	0.70	0.80	0.85	0.90
$n = 200$								
$\hat{\beta}$	0.3174	0.4692	0.6252	0.9403	0.6911	1.0891	1.5213	2.3644
$\hat{\beta}_{R_1}$	0.2207	0.2750	0.3115	0.3464	0.4393	0.5442	0.6092	0.6359
$\hat{\beta}_{R_2}$	0.2074	0.2592	0.2963	0.3387	0.4258	0.5357	0.6081	0.6680
$\hat{\beta}_{R_3}$	0.2086	0.2613	0.3000	0.3458	0.4390	0.5598	0.6434	0.7203
$\hat{\beta}_L$	0.2522	0.3424	0.4246	0.5711	0.5211	0.7344	0.9413	1.3091
$\hat{\beta}_{LT}$	0.1901	0.2545	0.3240	0.4606	0.3572	0.5177	0.6965	1.0615
$\hat{\beta}_G$	0.0978	0.1143	0.1409	0.1940	0.2009	0.2863	0.4055	0.6063
$n = 300$								
$\hat{\beta}$	0.1979	0.2949	0.3890	0.5911	0.4205	0.6654	0.9254	1.4517
$\hat{\beta}_{R_1}$	0.1578	0.2120	0.2519	0.3091	0.3185	0.4331	0.5119	0.5967
$\hat{\beta}_{R_2}$	0.1519	0.2032	0.2418	0.2988	0.3140	0.4257	0.5075	0.6050
$\hat{\beta}_{R_3}$	0.1526	0.2042	0.2433	0.3012	0.3204	0.4385	0.5283	0.6394
$\hat{\beta}_L$	0.1707	0.2389	0.2968	0.4066	0.3506	0.5070	0.6527	0.9115
$\hat{\beta}_{LT}$	0.1303	0.1736	0.2133	0.3040	0.2334	0.3278	0.4362	0.6611
$\hat{\beta}_G$	0.0838	0.0908	0.1052	0.1420	0.1433	0.1910	0.2485	0.3731
$n = 400$								
$\hat{\beta}$	0.1370	0.2026	0.2665	0.4102	0.2990	0.4752	0.6649	1.0494
$\hat{\beta}_{R_1}$	0.1160	0.1590	0.1931	0.2533	0.2457	0.3487	0.4312	0.5344
$\hat{\beta}_{R_2}$	0.1134	0.1548	0.1877	0.2470	0.2428	0.3438	0.4264	0.5374
$\hat{\beta}_{R_3}$	0.1140	0.1557	0.1890	0.2489	0.2461	0.3515	0.4396	0.5624
$\hat{\beta}_L$	0.1228	0.1730	0.2169	0.3055	0.2616	0.3875	0.5086	0.7219
$\hat{\beta}_{LT}$	0.0963	0.1244	0.1521	0.2152	0.1743	0.2429	0.3254	0.4962
$\hat{\beta}_G$	0.0750	0.0743	0.0841	0.1110	0.1196	0.1485	0.1896	0.2727
$n = 500$								
$\hat{\beta}$	0.1137	0.1668	0.2223	0.3326	0.2334	0.3684	0.5082	0.7911
$\hat{\beta}_{R_1}$	0.0997	0.1374	0.1721	0.2255	0.2006	0.2893	0.3635	0.4706
$\hat{\beta}_{R_2}$	0.0981	0.1348	0.1682	0.2199	0.1991	0.2863	0.3592	0.4685
$\hat{\beta}_{R_3}$	0.0986	0.1357	0.1694	0.2213	0.2013	0.2913	0.3676	0.4852
$\hat{\beta}_L$	0.1043	0.1469	0.1880	0.2597	0.2102	0.3131	0.4077	0.5750
$\hat{\beta}_{LT}$	0.0821	0.1062	0.1314	0.1827	0.1459	0.1975	0.2550	0.3716
$\hat{\beta}_G$	0.0710	0.0678	0.0756	0.0952	0.1075	0.1261	0.1550	0.2207

Table 2. The simulated mse values of $\hat{m}_{\hat{\beta}}, \hat{m}_{\hat{\beta}_{R_1}}, \hat{m}_{\hat{\beta}_{R_2}}, \hat{m}_{\hat{\beta}_{R_3}}, \hat{m}_{\hat{\beta}_L}, \hat{m}_{\hat{\beta}_{LT}}$ and $\hat{m}_{\hat{\beta}_G}$.

r	$p = 4$ & $\tau = 0.03$				$p = 7$ & $\tau = 0.01$			
	0.70	0.80	0.85	0.90	0.70	0.80	0.85	0.90
$n = 200$								
$\hat{m}_{\hat{\beta}}$	0.3140	0.3163	0.3185	0.3192	0.3481	0.3543	0.3587	0.3611
$\hat{m}_{\hat{\beta}_{R_1}}$	0.3015	0.3022	0.3032	0.3021	0.3298	0.3312	0.3311	0.3283
$\hat{m}_{\hat{\beta}_{R_2}}$	0.2991	0.3000	0.3011	0.3004	0.3275	0.3287	0.3290	0.3271
$\hat{m}_{\hat{\beta}_{R_3}}$	0.2995	0.3004	0.3015	0.3008	0.3287	0.3300	0.3305	0.3287
$\hat{m}_{\hat{\beta}_L}$	0.3060	0.3076	0.3092	0.3093	0.3360	0.3398	0.3422	0.3428
$\hat{m}_{\hat{\beta}_{LT}}$	0.2881	0.2919	0.2958	0.2992	0.3126	0.3196	0.3254	0.3311
$\hat{m}_{\hat{\beta}_G}$	0.2723	0.2846	0.2923	0.2987	0.3051	0.3187	0.3231	0.3273
$n = 300$								
$\hat{m}_{\hat{\beta}}$	0.2886	0.2892	0.2890	0.2886	0.2950	0.3015	0.3038	0.3062
$\hat{m}_{\hat{\beta}_{R_1}}$	0.2827	0.2829	0.2823	0.2812	0.2871	0.2917	0.2924	0.2924
$\hat{m}_{\hat{\beta}_{R_2}}$	0.2819	0.2821	0.2815	0.2805	0.2864	0.2908	0.2915	0.2916
$\hat{m}_{\hat{\beta}_{R_3}}$	0.2822	0.2823	0.2817	0.2807	0.2869	0.2915	0.2922	0.2924
$\hat{m}_{\hat{\beta}_L}$	0.2848	0.2852	0.2847	0.2841	0.2897	0.2951	0.2965	0.2979
$\hat{m}_{\hat{\beta}_{LT}}$	0.2709	0.2726	0.2739	0.2759	0.2715	0.2790	0.2826	0.2871
$\hat{m}_{\hat{\beta}_G}$	0.2574	0.2664	0.2709	0.2754	0.2648	0.2778	0.2785	0.2809
$n = 400$								
$\hat{m}_{\hat{\beta}}$	0.2667	0.2679	0.2686	0.2694	0.2775	0.2815	0.2848	0.2855
$\hat{m}_{\hat{\beta}_{R_1}}$	0.2632	0.2641	0.2646	0.2649	0.2730	0.2761	0.2786	0.2781
$\hat{m}_{\hat{\beta}_{R_2}}$	0.2629	0.2637	0.2642	0.2646	0.2727	0.2756	0.2780	0.2776
$\hat{m}_{\hat{\beta}_{R_3}}$	0.2631	0.2639	0.2644	0.2647	0.2730	0.2760	0.2785	0.2781
$\hat{m}_{\hat{\beta}_L}$	0.2645	0.2655	0.2660	0.2666	0.2745	0.2779	0.2807	0.2809
$\hat{m}_{\hat{\beta}_{LT}}$	0.2533	0.2553	0.2571	0.2598	0.2600	0.2648	0.2693	0.2721
$\hat{m}_{\hat{\beta}_G}$	0.2409	0.2492	0.2538	0.2587	0.2539	0.2638	0.2680	0.2711
$n = 500$								
$\hat{m}_{\hat{\beta}}$	0.2530	0.2557	0.2561	0.2561	0.2703	0.2743	0.2762	0.2783
$\hat{m}_{\hat{\beta}_{R_1}}$	0.2505	0.2523	0.2533	0.2538	0.2674	0.2709	0.2723	0.2736
$\hat{m}_{\hat{\beta}_{R_2}}$	0.2502	0.2528	0.2529	0.2529	0.2673	0.2707	0.2720	0.2734
$\hat{m}_{\hat{\beta}_{R_3}}$	0.2504	0.2529	0.2531	0.2531	0.2675	0.2709	0.2723	0.2737
$\hat{m}_{\hat{\beta}_L}$	0.2513	0.2540	0.2542	0.2550	0.2684	0.2720	0.2736	0.2753
$\hat{m}_{\hat{\beta}_{LT}}$	0.2419	0.2451	0.2464	0.2489	0.2562	0.2608	0.2638	0.2675
$\hat{m}_{\hat{\beta}_G}$	0.2292	0.2385	0.2426	0.2475	0.2503	0.2598	0.2622	0.2651

In order to study the influence of τ value on the ASMSE of $\hat{\beta}_G$, the curves of τ ($\tau \in (0, 1)$) and the ASMSE of $\hat{\beta}_G$ at $p = 4, n = 200, r = 0.70$ or 0.90 ; $p = 7, n = 200, r = 0.70$ or 0.90 and $p = 7, n = 300, r = 0.70$ or 0.90 are plotted in Figure 1. The Figure 1 clearly indicates that the ASMSE of $\hat{\beta}_G$ for various combinations is in a small state when τ value tends to 0, and reaches a maximum value when τ value tends to 1, which is close to the ASMSE of $\hat{\beta}$ (refer to Table 1). The situations of

the two endpoints can also be derived theoretically. From $\hat{f}(k_j) = \tau k_j + \lambda_{\min}(\tau - 1)$, Eqs (5.4) and (5.5), it follows that $\hat{f}(k_j) \rightarrow -\lambda_{\min}$ and $k_j \rightarrow \frac{-\lambda_j + \lambda_{\min}(1 + \lambda_j \hat{\alpha}_j^2)}{-\lambda_j \hat{\alpha}_j^2}$ as $\tau \rightarrow 0$, and $\hat{f}(k_j) \rightarrow 0$ and $k_j \rightarrow 0$ as $\tau \rightarrow 1$, thus $\hat{\beta}_G \rightarrow e$ (e is the minimum value of $\hat{\beta}_G$) as $\tau \rightarrow 0$, and $\hat{\beta}_G \rightarrow \hat{\beta}$ as $\tau \rightarrow 1$. In addition, Figure 1 also implies that the ASMSE of $\hat{\beta}_G$ varies greatly with large p and large r . Large p and large r are more likely to have multicollinearity, it indirectly reflects that the proposed GLTE $\hat{\beta}_G$ has significant ability to eliminate multicollinearity.

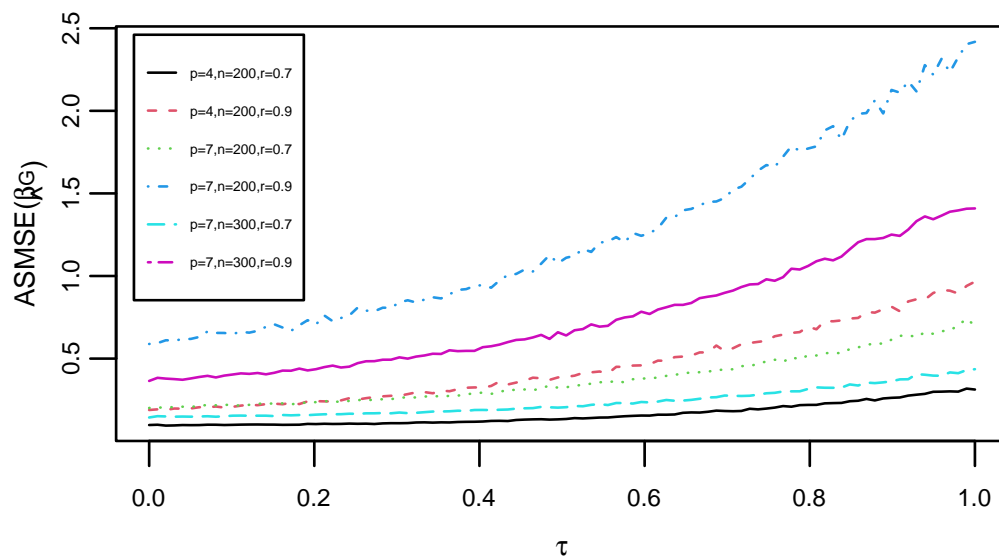


Figure 1. ASMSE of the estimators $\hat{\beta}_G$ versus τ at some combinations of p , r , and n .

6.2. Performance of GLTE and unmodified methods

In this subsection, we directly apply the estimator in [16] to the LPLRM for simulation and obtain estimators of unmodified method (UM) as follows

$$\hat{\beta}_{UM}(k) = (\tilde{X}^T \hat{W} \tilde{X} + kI)^{-1} (\tilde{X}^T \hat{W} \tilde{X} + f(k)I) \hat{\beta}, \quad k > 0,$$

$$\hat{m}_{\hat{\beta}_{UM}}(k) = S(Z - X \hat{\beta}_{UM}(k)).$$

We chose $k = \hat{k}_{R_1}$ (case II) and the seven $f(k)$ functions (I, II, III, IV, V, VI, VII) in [16] and simulate with the setting in the beginning of section 6. The simulation results are presented in Table 3, Table 4 together with those of GLTE and LPLE in subsection 6.1.

From the results shown in Tables 3 and 4 for the estimation of parameter and nonparametric function, we can see that the performance of GLTE method is better than the UM method under the seven function types. Both GLTE and UM have a large improvement in the estimation of the parameter, and these have a slightly smaller improvement in the estimation of the nonparametric function, comparing with LPLE. Also, as a result of increasing sample size n while keeping r and p fixed, the ASMSE and mse values in LPLE, GLTE and UM methods under various function types are generally decreases. Similarly, when n and p are fixed and r is increased, it is seen that the ASMSE and mse values commonly increased in various methods.

Table 3. The simulated ASMSE values of parameter estimator of LPLE, GLTE and UM.

	n	r	LPLE	UM(I)	UM(II)	UM(III)	UM(IV)	UM(V)	UM(VI)	UM(VII)	GLTE
$p = 4$	200	0.70	0.32	0.10	0.20	0.10	0.10	0.11	0.15	0.18	0.10
	200	0.80	0.47	0.11	0.24	0.11	0.11	0.12	0.17	0.22	0.11
	200	0.85	0.63	0.14	0.30	0.14	0.15	0.15	0.21	0.26	0.14
	200	0.90	0.94	0.19	0.36	0.19	0.19	0.20	0.25	0.30	0.19
	300	0.70	0.20	0.09	0.15	0.09	0.08	0.09	0.12	0.14	0.08
	300	0.80	0.29	0.09	0.19	0.09	0.09	0.10	0.14	0.18	0.09
	300	0.85	0.39	0.11	0.23	0.11	0.11	0.12	0.17	0.21	0.11
	300	0.90	0.59	0.14	0.29	0.14	0.15	0.15	0.21	0.26	0.14
	400	0.70	0.14	0.08	0.11	0.08	0.07	0.08	0.09	0.10	0.07
	400	0.80	0.20	0.08	0.14	0.08	0.07	0.08	0.11	0.14	0.07
	400	0.85	0.27	0.08	0.17	0.08	0.08	0.09	0.13	0.16	0.08
	400	0.90	0.41	0.11	0.23	0.11	0.11	0.12	0.17	0.21	0.11
	500	0.70	0.11	0.07	0.09	0.08	0.07	0.07	0.08	0.09	0.07
	500	0.80	0.17	0.07	0.13	0.07	0.07	0.07	0.10	0.12	0.07
	500	0.85	0.22	0.08	0.17	0.09	0.08	0.09	0.13	0.16	0.08
$p = 7$	200	0.70	0.69	0.20	0.38	0.20	0.21	0.20	0.25	0.37	0.20
	200	0.80	1.09	0.29	0.49	0.29	0.30	0.30	0.35	0.46	0.29
	200	0.85	1.52	0.41	0.61	0.41	0.42	0.42	0.47	0.60	0.41
	200	0.90	2.36	0.71	0.79	0.61	0.62	0.61	0.65	0.76	0.61
	300	0.70	0.42	0.14	0.27	0.14	0.14	0.14	0.17	0.27	0.14
	300	0.80	0.67	0.19	0.36	0.18	0.19	0.19	0.23	0.36	0.19
	300	0.85	0.93	0.26	0.44	0.25	0.26	0.26	0.31	0.44	0.25
	300	0.90	1.45	0.37	0.55	0.37	0.38	0.37	0.42	0.55	0.37
	400	0.70	0.30	0.12	0.22	0.12	0.12	0.12	0.14	0.21	0.12
	400	0.80	0.48	0.15	0.30	0.15	0.15	0.15	0.19	0.30	0.15
	400	0.85	0.66	0.19	0.36	0.18	0.19	0.19	0.23	0.36	0.19
	400	0.90	1.05	0.28	0.46	0.28	0.28	0.28	0.33	0.46	0.27
	500	0.70	0.23	0.11	0.18	0.11	0.11	0.11	0.12	0.18	0.11
	500	0.80	0.37	0.13	0.26	0.13	0.13	0.13	0.16	0.26	0.13
	500	0.85	0.51	0.15	0.30	0.15	0.15	0.15	0.19	0.30	0.16
500	0.90	0.79	0.22	0.39	0.22	0.23	0.22	0.27	0.39	0.22	

Note: $\tau = 0.03$ while $p = 4$; $\tau = 0.01$ while $p = 7$.

Table 4. The simulated mse values of nonparametric function estimator of LPLE, GLTE and UM.

	n	r	LPLE	UM(I)	UM(II)	UM(III)	UM(IV)	UM(V)	UM(VI)	UM(VII)	GLTE
$p = 4$	200	0.70	0.31	0.27	0.30	0.27	0.27	0.27	0.29	0.30	0.27
	200	0.80	0.32	0.28	0.30	0.28	0.28	0.28	0.29	0.30	0.28
	200	0.85	0.32	0.29	0.31	0.29	0.30	0.30	0.30	0.31	0.29
	200	0.90	0.32	0.30	0.31	0.30	0.30	0.30	0.30	0.31	0.30
	300	0.70	0.29	0.26	0.29	0.26	0.27	0.27	0.28	0.29	0.26
	300	0.80	0.29	0.27	0.29	0.27	0.27	0.28	0.28	0.29	0.27
	300	0.85	0.29	0.27	0.28	0.27	0.27	0.27	0.27	0.28	0.27
	300	0.90	0.29	0.28	0.29	0.28	0.28	0.28	0.29	0.29	0.28
	400	0.70	0.27	0.24	0.26	0.24	0.24	0.24	0.25	0.26	0.24
	400	0.80	0.27	0.25	0.26	0.25	0.25	0.25	0.26	0.26	0.25
	400	0.85	0.27	0.25	0.26	0.25	0.25	0.25	0.26	0.26	0.25
	400	0.90	0.27	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26
	500	0.70	0.25	0.23	0.25	0.23	0.23	0.23	0.24	0.25	0.23
	500	0.80	0.26	0.24	0.25	0.24	0.24	0.24	0.25	0.25	0.24
	500	0.85	0.26	0.24	0.25	0.24	0.24	0.24	0.25	0.25	0.24
$p = 7$	200	0.70	0.35	0.31	0.33	0.31	0.31	0.31	0.32	0.33	0.31
	200	0.80	0.35	0.32	0.33	0.32	0.32	0.32	0.32	0.33	0.32
	200	0.85	0.36	0.32	0.32	0.32	0.32	0.32	0.32	0.33	0.32
	200	0.90	0.36	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33
	300	0.70	0.29	0.26	0.28	0.26	0.26	0.26	0.27	0.28	0.26
	300	0.80	0.30	0.28	0.30	0.28	0.28	0.28	0.29	0.30	0.28
	300	0.85	0.30	0.28	0.29	0.28	0.28	0.28	0.28	0.29	0.28
	300	0.90	0.31	0.28	0.29	0.28	0.28	0.28	0.29	0.29	0.28
	400	0.70	0.28	0.25	0.27	0.25	0.25	0.25	0.26	0.27	0.25
	400	0.80	0.28	0.26	0.27	0.26	0.26	0.26	0.27	0.27	0.26
	400	0.85	0.28	0.27	0.28	0.27	0.27	0.27	0.27	0.28	0.27
	400	0.90	0.29	0.27	0.28	0.27	0.27	0.27	0.27	0.28	0.27
	500	0.70	0.27	0.25	0.27	0.25	0.25	0.25	0.26	0.27	0.25
	500	0.80	0.27	0.26	0.27	0.26	0.26	0.26	0.26	0.27	0.26
	500	0.85	0.28	0.26	0.27	0.26	0.26	0.26	0.26	0.27	0.26
500	0.90	0.28	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	

Note: $\tau = 0.03$ while $p = 4$; $\tau = 0.01$ while $p = 7$.

7. An application to the dataset of Indian Liver Patient

To motivate the multicollinearity problem in LPLRM, we consider the Indian Liver Patient Dataset from the UCI Repository of Machine Learning Databases [31]. The dataset contains 416 records of liver patients and 167 records of non-liver patients from north east of Andhra Pradesh, India. We consider the following eight variables: the binomial response variable y , taking values 1 if this patient

has liver disease and 0 otherwise, the explanatory variables include age of the patient (Age), total bilirubin (TB), direct bilirubin (DB), serum glutamate-pyruvate transaminase (SGPT), serum glutamate-oxaloacetate transaminase (SGOT), total proteins (TP), albumin (ALB). Based on the study of Hartatik et al. [32], there are strong correlations between TB and DB, SGPT and SGOT, but a weak correlation between Age and other variables. However, we are more interested in the curve of the effect of Age on the logarithmic odds of liver disease, so we set Age as a nonparametric variable and the other six explanatory variables are included in the linear part to establish the following LPLRM:

$$\log \frac{p_i}{1-p_i} = f(\text{Age}_i) + \beta_1 TB_i + \beta_2 DB_i + \beta_3 SGPT_i + \beta_4 SGOT_i + \beta_5 TP_i + \beta_6 ALB_i, \quad (7.1)$$

where $p_i = P(y_i = 1)$ is the probability of the i th individual having liver disease. After the final iteration process in section 2, the eigenvalues of the matrix $\tilde{\mathbf{X}}^T \tilde{\mathbf{W}} \tilde{\mathbf{X}}$ are $\lambda_1 = 189342.42$, $\lambda_2 = 32497.93$, $\lambda_3 = 276.57$, $\lambda_4 = 153.35$, $\lambda_5 = 16.40$, $\lambda_6 = 11.18$. Thus the condition number is $\kappa = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} = 130.1194$ showing that there is a multicollinearity problem. For this real data, the parameter estimates and SMSE values of parametric estimator of the model (7.1) under various methods are given in Table 5. The SMSE is used instead of ASMSE because the actual parameter values (β) are unknown during real data modeling. The results from the Table 5 reveal that the presence of multicollinearity affects the parameter estimates and that the SMSE of the proposed estimator $\hat{\beta}_G$ is smaller than that of the other estimators. Also, all the theoretical conditions in section 4 can be verified in this dataset when τ takes a very small value ($\tau = 0.01$) (see Table 6). The estimates of nonparametric functions under various methods are plotted in Figure 2. The Figure 2 shows that the difference between the nonparametric curve ($\hat{m}_{\beta_G}(\text{Age})$) of the proposed method and the nonparametric curve ($\hat{m}_{\beta}(\text{Age})$) of the PLM is the largest, but their shapes are basically similar, while the nonparametric curves of other methods are very similar to the nonparametric curve ($\hat{m}_{\beta}(\text{Age})$) of the PLM. Combined with the results in Table 2, this implies that the nonparametric function estimator of the proposed GLTE method outperform those of other methods.

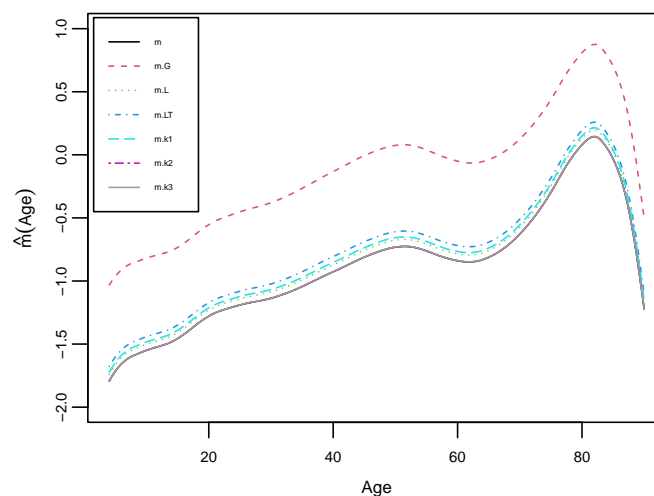


Figure 2. The estimations of nonparametric part of the model (7.1).

Table 5. The estimated parameter values and SMSE values.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	SMSE
$\hat{\beta}$	0.0121	0.4752	0.0120	0.0039	0.4358	-0.6615	0.1606
$\hat{\beta}_{R_1}(\hat{k}_{R_1} = 1.4202)$	0.0251	0.4421	0.0119	0.0040	0.3905	-0.5938	0.1599
$\hat{\beta}_{R_2}(\hat{k}_{R_2} = 0.0118)$	0.0122	0.4749	0.0120	0.0039	0.4354	-0.6608	0.1604
$\hat{\beta}_{R_3}(\hat{k}_{R_3} = 0.0148)$	0.0122	0.4749	0.0120	0.0039	0.4353	-0.6607	0.1604
$\hat{\beta}_L(\hat{d}_L = 0)$	0.0206	0.4528	0.0119	0.0040	0.4028	-0.6122	0.1575
$\hat{\beta}_{LT}(\hat{k}_{LT} = 7.0296, \hat{d}_{LT} = -3.7394)$	0.0354	0.4182	0.0118	0.0041	0.3610	-0.5497	0.1302
$\hat{\beta}_G(\tau = 0.01)$	0.1305	0.2079	0.0087	0.0055	0.0496	-0.1113	0.0284

Table 6. Verification of the superiority conditions for the proposed GLTE in Theorems 1–4.

Theorem 1: $k_j = \hat{k}_{G_j}, j = 1, \dots, 6$			Theorem 2: $k_j = \hat{k}_{G_j} (j = 1, \dots, 6), k = \hat{k}_{R_1} = 1.4202$			
j	$-2\lambda_j - k_j$	$f(k_j)$	k_j	$-\lambda_j - \frac{\lambda_j(\lambda_j+k_j)}{\lambda_j+k}$	$f(k_j)$	$-\lambda_j + \frac{\lambda_j(\lambda_j+k_j)}{\lambda_j+k}$
1	-387029.6000	72.3767	8344.8020	-387028.2000	72.3767	8343.3190
2	-85304.9900	192.0201	20309.1400	-85302.6800	192.0201	20306.8300
3	-576.5383	-10.8374	23.3910	-575.0059	-10.8374	21.8586
4	-10945.8700	95.3205	10639.1800	-10846.8400	95.3205	10540.1400
5	-32.7933	-11.0713	0.0000	-31.4863	-11.0713	-1.3070
6	-22.3663	-11.0713	0.0000	-21.1061	-11.0713	-1.2601
Theorem 3: $k_j = \hat{k}_{G_j} (j = 1, \dots, 6), d = \hat{d}_L = 0$			Theorem 2: $k_j = \hat{k}_{G_j} (j = 1, \dots, 6), k = \hat{k}_{R_2} = 0.0118$			
j	$-\lambda_j - \frac{(\lambda_j+d)(\lambda_j+k_j)}{\lambda_j+1}$	$f(k_j)$	$-\lambda_j + \frac{(\lambda_j+d)(\lambda_j+k_j)}{\lambda_j+1}$	$-\lambda_j - \frac{\lambda_j(\lambda_j+k_j)}{\lambda_j+k}$	$f(k_j)$	$-\lambda_j + \frac{\lambda_j(\lambda_j+k_j)}{\lambda_j+k}$
1	-387028.6000	72.3767	8343.7580	-387029.6000	72.3767	8344.7890
2	-85303.3600	192.0201	20307.5100	-85304.9700	192.0201	20309.1200
3	-575.4576	-10.8374	22.3103	-576.5255	-10.8374	23.3782
4	-10875.9500	95.3205	10569.2500	-10945.0400	95.3205	10638.3500
5	-31.8508	-11.0713	-0.9425	-32.7815	-11.0713	-0.0118
6	-21.4484	-11.0713	-0.9179	-22.3545	-11.0713	-0.0118
Theorem 4: $k_j = \hat{k}_{G_j}, k = \hat{k}_{LT} = 7.0296, d = \hat{d}_{LT} = -3.7394$			Theorem 2: $k_j = \hat{k}_{G_j} (j = 1, \dots, 6), k = \hat{k}_{R_3} = 0.0148$			
j	$-\lambda_j - \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k}$	$f(k_j)$	$-\lambda_j + \frac{(\lambda_j-d)(\lambda_j+k_j)}{\lambda_j+k}$	$-\lambda_j - \frac{\lambda_j(\lambda_j+k_j)}{\lambda_j+k}$	$f(k_j)$	$-\lambda_j + \frac{\lambda_j(\lambda_j+k_j)}{\lambda_j+k}$
1	-387026.2000	72.3767	192.0201	-387029.6000	72.3767	8344.7860
2	-85299.6400	192.0201	20303.7900	-85304.9600	192.0201	20309.1100
3	-573.0583	-10.8374	19.9110	-576.5222	-10.8374	23.3749
4	-10724.4600	95.3205	10417.7700	-10944.8300	95.3205	10638.1300
5	-30.4904	-11.0713	-2.3029	-32.7785	-11.0713	-0.0148
6	-20.3460	-11.0713	-2.0203	-22.3515	-11.0713	-0.0148

8. Conclusions

In this paper, we propose a GLTE to combat the multicollinearity of the linear part in a LPLRM. The GLTE is a more general estimator which includes the other estimators as special cases. Under some theoretical conditions, the performance of GLTE is superior to the LPLE, the ridge estimator, the Liu estimator and the Liu-type estimator. Also, the optimal choices of biasing parameters and function of biasing parameter are carried out, and their empirical choices are suggested. Monte Carlo simulations show that the finite sample performances of the proposed GLTE are better than those of the other estimators. The superior performance of GLTE is obtained when the empirical parameter τ takes a very small value ($\tau \rightarrow 0$ & $\tau \neq 0$). Also, the estimators are applied to the dataset of Indian Liver Patient, and the obtained results are consistent with the simulation study.

Acknowledgments

The research in this article is supported by the National Natural Science Foundation of China under grant number No.61973096. Their financial aids are greatly appreciated.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. T. A. Severini, J. G. Staniswalis, Quasi-likelihood estimation in semiparametric models, *J. Am. Stat. Assoc.*, **89** (1994), 501–511. <https://doi.org/10.1080/01621459.1994.10476774>
2. H. Sally, Semiparametric regression in likelihood-based models, *J. Am. Stat. Assoc.*, **89** (1994), 1354–1365. <https://doi.org/10.1080/01621459.1994.10476874>
3. H. Liang, Y. S. Qin, X. Y. Zhang, D. Ruppert, Empirical likelihood-based inferences for generalized partially linear models, *Scand. J. Stat.*, **36** (2009), 433–443. <https://doi.org/10.1111/j.1467-9469.2008.00632.x>
4. G. Boente, D. Rodriguez, Robust inference in generalized partially linear models, *Comput. Stat. Data An.*, **54** (2010), 2942–2966. <https://doi.org/10.1016/j.csda.2010.05.025>
5. J. Rahman, S. H. Luo, Y. W. Fan, X. H. Liu, Semiparametric efficient inferences for generalised partially linear models, *J. Nonparametr. Stat.*, **32** (2020), 704–724. <https://doi.org/10.1080/10485252.2020.1790557>
6. A. E. Hoerl, R. W. Kennard, Ridge regression: biased estimation for nonorthogonal problems, *Technometrics*, **12** (1970), 55–67. <https://doi.org/10.1080/00401706.1970.10488634>
7. K. J. Liu, A new class of biased estimate in linear regression, *Commun. Stat. Theor. M.*, **22** (1993), 393–402. <https://doi.org/10.1080/03610929308831027>
8. K. J. Liu, Using Liu-type estimator to combat collinearity, *Commun. Stat. Theor. M.*, **32** (2003), 1009–1020. <https://doi.org/10.1081/STA-120019959>

9. F. S. Kurnaz, K. U. Akay, A new Liu-type estimator, *Stat. Pap.*, **56** (2015), 495–517. <https://doi.org/10.1007/s00362-014-0594-6>
10. A. Zeinal, The extended two-type parameter estimator in linear regression model, *Commun. Stat. Theor. M.*, **1** (2021), 1. <https://doi.org/10.1080/03610926.2021.1916528>
11. M. R. Özkale, S. Kaçiranlar, The restricted and unrestricted two-parameter estimators, *Commun. Stat. Theor. M.*, **36** (2007), 2707–2725. <https://doi.org/10.1080/03610920701386877>
12. B. M. G. Kibria, K. Mnsson, G. Shukur, Performance of some logistic ridge regression estimators, *Comput. Ecom.*, **40** (2012), 401–414. <https://doi.org/10.1007/s10614-011-9275-x>
13. D. Inan, B. E. Erdogan, Liu-type logistic estimator, *Commun. Stat. Simul. C.*, **42** (2013), 1578–1586. <https://doi.org/10.1080/03610918.2012.667480>
14. Y. Asar, A. Genç, Two-parameter ridge estimator in the binary logistic regression, *Commun. Stat. Simul. C.*, **46** (2017), 7088–7099. <https://doi.org/10.1080/03610918.2016.1224348>
15. N. Varathan, P. Wijekoon, Modified almost unbiased Liu estimator in logistic regression, *Commun. Stat. Simul. C.*, **50** (2021), 1009–1020. <https://doi.org/10.1080/03610918.2019.1626888>
16. E. Ertan, K. U. Akay, A new Liu-type estimator in binary logistic regression models, *Commun. Stat. Theor. M.*, **51** (2022), 4370–4394. <https://doi.org/10.1080/03610926.2020.1813777>
17. N. H. Jadhav, On linearized ridge logistic estimator in the presence of multicollinearity, *Computation. Stat.*, **35** (2020), 667–687. <https://doi.org/10.1007/s00180-019-00935-6>
18. N. K. Rashad, Z. Y. Algamal, A new ridge estimator for the poisson regression model, *Iran. J. Sci. Technol. Tran. Sci.*, **43** (2019), 2921–2928. <https://doi.org/10.1007/s40995-019-00769-3>
19. A. F. Lukman, K. Ayinde, B. M. G. Kibria, E. T. Adewuyi, Modified ridge-type estimator for the gamma regression model, *Commun. Stat. Simul. C.*, **51** (2022), 5009–5023. <https://doi.org/10.1080/03610918.2020.1752720>
20. M. Roozbeh, M. Arashi, Feasible ridge estimator in partially linear models, *J. Multivariate Anal.*, **116** (2013), 35–44. <https://doi.org/10.1016/j.jmva.2012.11.006>
21. J. B. Wu, Performance of the difference-based almost unbiased Liu estimator in partial linear model, *J. Stat. Comput. Sim.*, **86** (2016), 2874–2887. <https://doi.org/10.1080/00949655.2015.1136628>
22. H. Emami, A. Aghamohammadi, Elliptical difference based ridge and Liu type estimators in partial linear measurement error models, *Commun. Stat. Theor. M.*, **50** (2021), 4913–4933. <https://doi.org/10.1080/03610926.2018.1472793>
23. F. Akdeniz, M. Roozbeh, Generalized difference-based weighted mixed almost unbiased ridge estimator in partially linear models, *Stat. Pap.*, **60** (2020), 1717–1739. <https://doi.org/10.1007/s00362-017-0893-9>
24. J. Wu, B. M. G. Kibria, A generalized difference-based mixed two-parameter estimator in partially linear models, *Commun. Stat. Theor. M.*, 2022. <https://doi.org/10.1080/03610926.2021.2024234>
25. P. McCullagh, J. A. Nelder, *Generalized linear models*, 2nd edition, New York: Chapman & Hall, 1989.

26. C. M. Theobald, Generalizations of mean square error applied to ridge regression, *J. R. Stat. Soc. B.*, **36** (1974), 103–106. <https://doi.org/10.1111/j.2517-6161.1974.tb00990.x>
27. R. W. Farebrother, Further results on the mean square error of ridge regression, *J. R. Stat. Soc. B.*, **38** (1976), 248–250. <https://doi.org/10.1111/j.2517-6161.1976.tb01588.x>
28. M. Amini, M. Roozbeh, Optimal partial ridge estimation in restricted semiparametric regression models, *J. Multivariate Anal.*, **136** (2015), 26–40. <https://doi.org/10.1016/j.jmva.2015.01.005>
29. M. Roozbeh, Optimal QR-based estimation in partially linear regression models with correlated errors using GCV criterion, *Comput. Stat. Data An.*, **117** (2018), 45–61. <https://doi.org/10.1016/j.csda.2017.08.002>
30. D. W. Scott, *Multivariate density estimation: theory, practice, and visualization*, Hoboken: Wiley, 1992.
31. ILPD (Indian Liver Patient Dataset) Data Set, Machine learning repository, 2012. Available from: <http://archive.ics.uci.edu/ml>
32. H. Hartatik, M. B. Tamam, A. Setyanto, Prediction for diagnosing liver disease in patients using KNN and Naive Bayes algorithms, *ICORIS*, 2020, 20288260. <https://doi.org/10.1109/ICORIS50180.2020.9320797>



AIMS Press

© 2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)