CrossMark

Local influence for Liu estimators in semiparametric linear models

Hadi Emami[1]

Abstract Semiparameric linear regression models are extensions of linear models to include a nonparametric function of some covariate. They have been found to be useful in data modelling. This paper provides local influence analysis to the Liu penalized least squares estimators that uses a smoothing spline as a solution to its nonparametric component. The diagnostics under the perturbations of constant variance, individual explanatory variables and assessing the influence on the selection of the Liu penalized least squares estimators parameter are derived. The diagnostics are applied to a real data set with informative results.

Keywords Liu estimator · Local influence · Perturbation scheme · Multicollinearity · Smoothing spline

Mathematics Subject Classification 62J07 · 62J20 · 62J05

1 Introduction

Diagnostic techniques for the parametric regression model have received a great deal of attention in statistical literature since the seminal work of Cook (1977) and others including Cook and Weisberg (1982), Belsley et al. (1989) and Walker and Brich (1988). In semiparametric regression models (SPRMs), diagnostic results are quite rare; among them Eubank (1985), Thomas and Cook (1989) and Kim (1996) studied the basic diagnostic building blocks such as residuals and leverages. Kim et al. (2001,2002) and Fung et al. (2002) proposed some type of Cook's distances in SPRMs.

✉ Hadi Emami
h.emami@znu.ac.ir

[1] Department of Statistics, Faculty of Science, University of Zanjan, Zanjan, Iran

The existence of collinearity in the linear regression model can lead to a very sensitive least squares estimate, therefore mixed estimation and ridge regression are suggested to mitigate the effect of collinearity. However, as many authors noted, the influence of the observations on ridge regression is different from the corresponding least squares estimate, and collinearity can even disguise anomalous data (Belsley et al. 1989). Using case deletion method Walker and Brich (1988) and Jahufer and Chen (2009) studied the influence of observations in ordinary ridge regression (ORR) and modified ridge regression (MRR) respectively. They derived the dependence of several influence measures from case deletion on the estimates of ORR and MRR parameters. In SPRMs, Emami (2015) extended results from Walker and Brich (1988) and derived influence measures of ridge estimates using case deletion.

Instead of deleting cases one-by-one, the local influence approach considered in Cook (1986) assessed the influence by simultaneous perturbation of the assumed model, and the influence was measured by normal curvature of an influence graph based on likelihood displacement. This approach has received a lot of attention in the past. The local influence analysis does not involve recomputing the parameter estimates for every case deletion, so it is often computationally simpler. Furthermore, it permits perturbation of various aspects of the model to tell us more than what the case deletion approach is designed for. For example, it can help measure leverage of a design point and evaluate the impact of a small measurement error of x on our estimates. This approach has been extended to generalized models by Thomas and Cook (1989), to linear mixed models by Lesaffre and Verbeke (1998), to partial linear models by Zhu et al. (2003) and to linear measurement errors by Rasekh (2006). In multicollinearity problems, Shi and Wang (1999) and Jahufer and Chen (2012) studied the local influence of minor perturbations on the ridge estimate and Liu estimators in the ordinary regression, respectively. They derived the diagnostics under the perturbation of variance and explanatory variables.

In this paper we generalize the Shi and Wang (1999) results to the SPRMs and we assess the local influence of observations on the Liu estimates. We demonstrate that the local influence analysis of Cook (1986) can be extended to the Liu penalized least squares estimators (LPLSEs) in SPRMs and provide some insight into the interplay between the linear and the nonparametric components in the context of influence diagnostics.

The paper is organized as follows. In the next section, SPRMs are introduced, the relevant notations and some inferential results are also given. Section 3 derives the local influence diagnostics of LPLSEs under the perturbation of constant variance and explanatory variables. Section 4 provides a diagnostic for detecting the local influential observations of selecting the LPLSEs parameter. In Sect. 5 the proposed methods are illustrated through a simulation study and a real data set. A discussion is given in the last section.

2 Model and inference

Consider the semiparametric regression model

$$y_i = x_i'\beta + g(t_i) + \epsilon_i \quad 1 \le i \le n, \tag{1}$$

where y_i is the scalar response, β is a p-vector of regression coefficients, x_i is a p-vector of explanatory variables, t_i is a scalar ($a \leq t_i, \ldots t_n \leq b$), and $t_i's$ are not all identical, g is a twice differentiable unknown smooth function on some finite interval and the errors ϵ_i are uncorrelated with zero mean and unknown constant variance σ^2. This model is also called a partially linear model or a partial spline model. Model (1) has been used in discussion of many methods, e.g., penalized least square (see Fung et al. 2002; Chen and You 2005), smoothing spline (see Speckman 1988; Green and Silverman 1994). In this study we will focus our attentions on the local influence diagnostics for the penalized least square estimators as it is a well-studied method of estimation for such models.

2.1 Penalized least square estimators (PLSEs)

Let the ordered distinct values among $t_1, \ldots, t_n$ be denoted by $s_1, \ldots, s_q$. The connection between $t_1, \ldots, t_n$ and $s_1, \ldots, s_q$ is captured by means of $n \times q$ incidence matrix $\mathbf{N}$, with entries $N_{ij} = 1$ if $t_i = s_j$ and 0 otherwise. Let g be the vector of value $a_i = g(s_i)$. For model (1) the penalized sum of squares is

$$\|y - \mathbf{X}\beta - \mathbf{N}g\|^2 + \lambda \int g''(t)^2 dt, \tag{2}$$

where y is the vector of n response values and $\mathbf{X}$ is $n \times p$ design matrix. Minimizing (2) with respect to β and g, the PLSEs of β and g are

$$\hat{\beta} = \{\mathbf{X}'(\mathbf{I} - \mathbf{S})\mathbf{X}\}^{-1}\mathbf{X}'(\mathbf{I} - \mathbf{S})y, \tag{3}$$

and

$$\hat{g} = (\mathbf{N}'\mathbf{N} + \lambda\mathbf{K})^{-1}\mathbf{N}'(y - \mathbf{X}\hat{\beta}), \tag{4}$$

respectively, where $\mathbf{I}$ is the identity matrix of size n, $\mathbf{S} = \mathbf{N}(\mathbf{N}'\mathbf{N} + \lambda\mathbf{K})^{-1}\mathbf{N}'$ is a smoothing matrix, λ is a nonnegative tuning parameter and $\mathbf{K}$ is a $q \times q$ matrix whose entries only depend on the knots $\{s_j\}$ (see Speckman 1988).

2.2 Liu penalized least squares estimators (LPLSEs)

The multicollinearity is a problem when the primary interest is in the estimation of the parameters in a regression model. In the case of multicollinearity we know that when the correlation matrix has one or more small eigenvalues, the estimates of the regression coefficients can be large in absolute value. The least squares estimator performs poorly in the presence of multicollinearity. Some biased estimators have been suggested as a means to improve the accuracy of the parameter estimate in the model when multicollinearity exists. There are a few studies that looked at overcoming the rank-deficient and ill-conditioned or multicollinearity problems in SPRMs (see Hu 2005; Tabakan and Akdeniz 2010; Roozbeh 2015; Roozbeh and Arashi 2015; Amini and Roozbeh 2015; Arashi and Valizadeh 2015). To overcome near multicollinearity,

Liu (1993) combined the Stein (1956) estimator with ordinary ridge estimator to obtain the Liu estimator. This approach extended in semiparametric linear regression models (see Akdeniz and Akdeniz Duran (2010); Akdeniz et al. (2015)). Here, we use Liu estimators which can be obtained by minimizing the term

$$\|y - \mathbf{X}\beta - \mathbf{N}g\|^2 + \lambda \int g''(t)^2 dt + \|d\hat{\beta} - \beta\|^2 \tag{5}$$

Following Green and Silverman (1994), minimization of (5) can be done in a two steps estimation process: first we minimize it subject to $g(s_j) = a_j, j = 1, \ldots, q$ and in the second step we minimize the result over the choice of g and β. The problem of minimizing $\int g''(t)^2 dt$ subject to g interpolating given points $g(s_j) = a_j$ is given by Green and Silverman (1994), and minimizing function g provides a cubic spline with knots $\{s_j\}$. There exists a matrix $\mathbf{K}$ only depending on the knots $\{s_j\}$, such that the minimized value of $\int g''(t)^2 dt$ is $g'\mathbf{K}g$. The equation in (5) is therefore of the form

$$\|y - \mathbf{X}\beta - \mathbf{N}g\|^2 + \lambda g'\mathbf{K}g + \|d\hat{\beta} - \beta\|^2. \tag{6}$$

Minimizing (6) subject to β and g, the LPLSEs, say, $\hat{\beta}_d$ and $\hat{g}_d$, solve

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} + \mathbf{I}_p & \mathbf{X}'\mathbf{N} \\ \mathbf{N}'\mathbf{X} & \mathbf{N}'\mathbf{N} + \lambda\mathbf{K} \end{pmatrix} \begin{pmatrix} \beta \\ g \end{pmatrix} = \begin{pmatrix} \mathbf{X}'y + d\hat{\beta} \\ \mathbf{N}'y \end{pmatrix}, \tag{7}$$

where $\mathbf{I}_p$ is the $p \times p$ identity matrix. From (7) by simple calculation the LPLSEs are defined as

$$\hat{\beta}_d = \left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1} \left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]y + d\hat{\beta}\right), \tag{8}$$

and

$$\hat{g}_d = (\mathbf{N}'\mathbf{N} + \lambda\mathbf{K})^{-1}\mathbf{N}'(y - \mathbf{X}\hat{\beta}_d). \tag{9}$$

Using (8) and (9), the vector of fitted values is

$$\hat{y} = \mathbf{X}\hat{\beta}_d + \mathbf{N}\hat{g}_d = (\mathbf{H}_\beta + \mathbf{H}_g)y$$
$$= \mathbf{H}_d y,$$

where $\mathbf{H}_\beta = \mathbf{X}\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + d\mathbf{I}_p)\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X}\right)\tilde{\mathbf{X}}'$, $\tilde{\mathbf{X}} = (\mathbf{I} - \mathbf{S})\mathbf{X}$, $\mathbf{H}_g = \mathbf{S}(\mathbf{I} - \mathbf{H}_\beta)$ and $\mathbf{H}_d$ is the hat matrix which in the expanded form is

$$\mathbf{H}_d = \mathbf{S} + \tilde{\mathbf{X}}\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + d\mathbf{I}_p)\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X}\right)\tilde{\mathbf{X}}'.$$

The first part, $\mathbf{S}$ is due to the nonparametric component of the model and the second part is due to the linear component of the model after adjusting for the former. The LPLSEs residual vector is evaluated as

$$e_d = y - \hat{y}_d = (\mathbf{I} - \mathbf{H}_d)y.$$

The small value d is called the Liu estimator biasing parameter or the shrinkage parameter. A prediction criteria suggested by Liu (1993), to choose biasing parameter d by minimizing Mallows (1973) statistic and it is can be generalized for PLSEs by

$$C_d = SSR_d/s^2 + 2tr(\mathbf{H}_d) - (n - 2). \tag{10}$$

where SSR_d is the sum of squares residuals, $s^2 = e_d'e_d/(n - p)$ is the estimator of σ^2 from LPLSEs in SPRM.

3 Local influence of LPLSEs

It is necessary to start by giving a brief sketch of the local influence approach suggested by Shi (1997) and Shi and Wang (1999) in which the generalized influence function and generalized Cook statistic are defined to assess the local change of small perturbation on some key issues. The generalized influence function of a concerned quantity $\mathbf{T} \in \mathbb{R}^n$ is given by

$$GIF(\mathbf{T}, l) = \lim_{a \to 0} \frac{\mathbf{T}(\omega_0 + l) - \mathbf{T}(\omega_0)}{a},$$

where $\omega = \omega_0 + al \in \mathbb{R}^n$ represents a perturbation, ω_0 is an null perturbation which satisfies $\mathbf{T}(\omega_0) = \mathbf{T}$ denotes an unit length vector. To assess the influence of the perturbations on $\mathbf{T}$, the generalized Cook's statistic is defined as

$$GC(\mathbf{T}, l) = \frac{[GIF(\mathbf{T}, l)]'\mathbf{M}[GIF(\mathbf{T}, l)]}{c},$$

where $\mathbf{M}$ is a $p \times p$ positive definite matrix, and c is a scalar. By maximizing the absolute value of $GC(\mathbf{T}; l)$ with respect to l, a direction $l_{max}(\mathbf{T})$ is obtained. This direction shows how to perturb the data to obtain the greatest local change in $\mathbf{T}$, and thus can be used as a main diagnostic. Maximum value $GC_{max}(\mathbf{T}) = GC(\mathbf{T}; l_{max})$ indicates the serious local influence. This method removes the need of likelihood. For a discussion on this method and its relationship with Cook (1986) approach, see Shi (1997).

3.1 Perturbing the variance

We first perturb the data by modifying the weight given to each case in the Liu penalized least squares criterion. This is equivalent to perturbing the variance of ϵ_i in the model. Using the perturbation of constant error variance the term in (6) becomes

$$\|y - \mathbf{X}'\beta_\omega - g_\omega(t_i))\|_{\mathbf{W}}^2 + \lambda g_\omega \mathbf{K} g_\omega' + \|d\hat{\beta} - \beta\|^2, \tag{11}$$

where $\mathbf{W} = diag(\omega)$ is a diagonal matrix with diagonal elements of $\omega = (\omega_1, \ldots, \omega_n)$ and $\|.\|_{\mathbf{W}}^2$ is weighted l^2-norm. Let $\omega = \omega_0 + al$, where $\omega_0 = 1$, the n vector of ones and $l = (l_1, \ldots, l_n)$.

Theorem 1 *Under the perturbation of variance the influence function of $\hat{\beta}_d$ and $\hat{g}_d$ are:*

$$GIF(\hat{\beta}_d, l) = \left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1} \tilde{\mathbf{X}}'\mathbf{D}(l)\tilde{e}_d = \left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1} \tilde{\mathbf{X}}'\mathbf{D}(\tilde{e}_d)l,$$

$$GIF(\hat{g}_d, l) = \left[\mathbf{N}'\mathbf{N} + \lambda\mathbf{K}\right]^{-1} \mathbf{N}'\left(\mathbf{I} - \mathbf{X}\left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1} \tilde{\mathbf{X}}'\right)\mathbf{D}(\tilde{e}_d)l,$$

where $\tilde{e}_d = \tilde{y} - \tilde{\mathbf{X}}\hat{\beta}_d$ and $\tilde{y} = (\mathbf{I} - \mathbf{S})y$.

Proof Minimizing (11), the perturbed version of the LPLSEs $\hat{\beta}_d$ and $\hat{g}_d$ are

$$\hat{\beta}_{d,\omega} = \left(\mathbf{X}'\left[\mathbf{W} - \mathbf{W}\mathbf{S}_\omega\mathbf{W}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}\left(\mathbf{X}'\left[\mathbf{W} - \mathbf{W}\mathbf{S}_\omega\mathbf{W}\right]y + d\hat{\beta}\right), \qquad (12)$$

and

$$\hat{g}_{d,\omega} = (\mathbf{N}'\mathbf{W}\mathbf{N} + \lambda\mathbf{K})^{-1}\mathbf{N}'\mathbf{W}(y - \mathbf{X}\hat{\beta}_{d,\omega}), \qquad (13)$$

respectively.

Now, by differentiating and equating to the null matrix we obtain

$$\frac{\partial\left(X'\left[\mathbf{W} - \mathbf{W}\mathbf{S}_\omega\mathbf{W}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}}{\partial\omega}\bigg|_{a=0} = -\left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}$$
$$\times\tilde{\mathbf{X}}'\mathbf{D}(l)\,\tilde{X}\left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}, (14)$$

and

$$\frac{\partial\left(\mathbf{X}'\left[\mathbf{W} - \mathbf{W}\mathbf{S}_\omega\mathbf{W}\right]y\right)}{\partial\omega}\bigg|_{a=0} = \tilde{\mathbf{X}}'\mathbf{D}(l)\tilde{y}, \qquad (15)$$

Then using(12), (14) and (15) we have

$$\frac{\partial\hat{\beta}_{\omega,d}}{\partial\omega}\bigg|_{a=0} = -\left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}'\mathbf{D}(l)\tilde{\mathbf{X}}\left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}$$
$$\times\left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]y + d\hat{\beta}\right) + \left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}'\mathbf{D}(l)\tilde{y}$$
$$= \left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}'\mathbf{D}(l)(\tilde{y} - \tilde{\mathbf{X}}\hat{\beta}_d)$$
$$= \left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}'\mathbf{D}(l)\tilde{e}_d, \qquad (16)$$

Similary, from (13) and (16) we get

$$\frac{\partial\hat{g}_{\omega,d}}{\partial\omega}\bigg|_{a=0} = -\left[\mathbf{N}'\mathbf{N} + \lambda\mathbf{K}\right]^{-1}\mathbf{N}'\,\mathbf{D}(l)\mathbf{S}e_d + \left[\mathbf{N}'\mathbf{N} + \lambda\mathbf{K}\right]^{-1}\mathbf{N}'\mathbf{D}(l)e_d$$
$$- \left[\mathbf{N}'\mathbf{N} + \lambda\mathbf{K}\right]^{-1}\mathbf{N}'\mathbf{X}\left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}'\mathbf{D}(l)\tilde{e}_d$$
$$= \left[\mathbf{N}'\mathbf{N} + \lambda\mathbf{K}\right]^{-1}\mathbf{N}'\left(\mathbf{I} - \mathbf{X}\left(\mathbf{X}'\left[\mathbf{I} - \mathbf{S}\right]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}'\right)\mathbf{D}(l)\tilde{e}_d. \quad (17)$$

Therefore from (16) and (17) the proof is complete. Analogous to case deletion, two versions of the generalized Cook's statistic of β can be constructed as

$$GC_1(\hat{\beta}_d, l) = l'\mathbf{D}(\tilde{e}_d)\tilde{\mathbf{X}}\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1}\left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)$$
$$\times \left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}'\mathbf{D}(\tilde{e}_d)l/s^2 tr(\mathbf{H}_\beta)$$
$$= l'\mathbf{A}_\beta l/s^2 tr(\mathbf{H}_\beta), \tag{18}$$

and

$$GC_2(\hat{\beta}_d, l) = l'\mathbf{D}(\tilde{e}_d)\tilde{\mathbf{X}}\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + d\mathbf{I}_p\right)^{-1}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X})(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}$$
$$\times(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X})\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + d\mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}'\mathbf{D}(\tilde{e}_d)l/s^2 tr(\mathbf{H}_\beta)$$
$$= l'\mathbf{B}_\beta l/s^2 tr(\mathbf{H}_\beta), \tag{19}$$

In (18) generalized Cook's statistic is scaled by $\mathbf{M}$ in the LS regression framework and in (19) generalized Cook's statistic is scaled by $\mathbf{M}$ in the Liu version of SPRM framework using the fact that

$$cov(\hat{\beta}_d) = \sigma^2\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1}\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + d\mathbf{I}_p\right)(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}$$
$$(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X})^{-1}\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + d\mathbf{I}_p\right)\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1}.$$

The generalized Cooks statistic of $\hat{g}_d$ will be

$$GC_g(\hat{g}_d, l) = l'\mathbf{D}(\tilde{e}_d)(\mathbf{I} - \mathbf{X}\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}')\mathbf{N}\left[\mathbf{N}'\mathbf{N} + \lambda\mathbf{K}\right]^{-1}$$
$$(\mathbf{N}'\mathbf{N})\left[\mathbf{N}'\mathbf{N} + \lambda\mathbf{K}\right]^{-1}\mathbf{N}\left(\mathbf{I} - \mathbf{X}\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}'\right)\mathbf{D}(\tilde{e}_d)l/s^2 tr(\mathbf{H}_g)$$
$$= l'\mathbf{D}(\tilde{e}_d)\|(\mathbf{I} - \mathbf{X}\left(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p\right)^{-1}\tilde{\mathbf{X}}')\mathbf{S}\|^2\mathbf{D}(\tilde{e}_d)l/s^2 tr(\mathbf{H}_g)$$
$$= l'\mathbf{A}_g l/s^2 tr(\mathbf{H}_g). \tag{20}$$

Therefore associated diagnostics, denoted by $l_{max}^1(\hat{\beta}_d)$, $l_{max}^2(\hat{\beta}_d)$ and $l_{max}(\hat{g}_d)$ are the eigenvectors corresponding to the largest absolute eigenvalues of matrices $\mathbf{A}_\beta$, $\mathbf{B}_\beta$ and $\mathbf{A}_g$ respectively. □

3.2 Perturbing the explanatory variables

It is known that the minor perturbation of the explanatory variables can seriously influence the least squares results when collinearity is present (Cook 1986, p. 147). This section considers the influence that perturbation of explanatory variables has on the LPLSEs. We define the matrix $\mathbf{X} = [x_1, \ldots, x_p]$ in which x_i, $i = 1, \ldots, p$ vectors of explanatory variables and we refer to $\mathbf{X}_\omega$ as the matrix $\mathbf{X}$ after the perturbation of ith column. Therefore,

$$\mathbf{X}_\omega = \mathbf{X} + as_i l\xi_i'$$

where ξ_i is a $p \times 1$ vector with a 1 in the ith position and zeroes elsewhere. s_i denotes the scale factor and accounts for the different measurement units associated with the columns of $\mathbf{X}$.

Theorem 2 *Under the perturbation of explanatory variables the influence function of $\hat{\beta}_d$ and $\hat{g}_d$ are:*

$$GIF(\hat{\beta}_d, l) = s_i \left(X' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{I}_p\right)^{-1} \left(\xi_i \tilde{e}'_d - \hat{\beta}_{d,i} \tilde{\mathbf{X}}'\right) l,$$

$$GIF(\hat{g}_d, l) = -s_i (\mathbf{N}'\mathbf{N} + \lambda\mathbf{K})^{-1} \mathbf{N}' \left[\hat{\beta}_{d,i}\mathbf{I} + \mathbf{X}(X' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{I}_p)^{-1}(\xi_i \tilde{e}'_d - \hat{\beta}_{d,i}\tilde{\mathbf{X}}')\right] l.$$

Proof Under the perturbation of ith column in (6) the LPLSEs will be

$$\hat{\beta}_{\omega,d} = \left(\mathbf{X}'_\omega \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X}_\omega + \mathbf{I}_p\right)^{-1} \left(\mathbf{X}'_\omega \left[\mathbf{I} - \mathbf{S}\right] y + d\hat{\beta}\right), \tag{21}$$

and

$$\hat{g}_{d,\omega} = \left(\mathbf{N}'\mathbf{N} + \lambda\mathbf{K}\right)^{-1} \mathbf{N}' \left(y - \mathbf{X}_\omega \hat{\beta}_{w,d}\right). \tag{22}$$

Since

$$\mathbf{X}'_\omega \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X}_\omega = \left(\mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + as_i l\xi'_i \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] as_i \xi_i l' \right.$$
$$\left. + as_i l\xi'_i \mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] as_i \xi_i l'\right)$$

and $\mathbf{X}'_\omega \left[\mathbf{I} - \mathbf{S}\right] y = \mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] y + as_i l\xi'_i \left[\mathbf{I} - \mathbf{S}\right] y$, then it is easy to obtain that

$$\left(\mathbf{X}'_\omega \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X}_\omega + \mathbf{I}_p\right)^{-1} = (\mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{I}_p)^{-1}$$
$$- as_i \left((\mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{I}_p)^{-1}\tilde{\mathbf{X}}' l\xi'_i (\mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{I}_p)^{-1}\right.$$
$$\left. + (\mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{I}_p)^{-1}\xi_i l'\tilde{\mathbf{X}}(\mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{I}_p)^{-1}\right) + o(a^2). \tag{23}$$

Hence, from (23) the (21) becomes

$$\hat{\beta}_{\omega,d} = \hat{\beta}_d + as_i(\mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{I}_p)^{-1}(\xi_i \tilde{e}' - \hat{\beta}_{d,i}\tilde{\mathbf{X}}')l + o(a^2). \tag{24}$$

From (24) by similar calculations for (22) we have

$$\hat{g}_{\omega,d} = \hat{g}_d - (\mathbf{N}'\mathbf{N} + \lambda\mathbf{K})^{-1}\mathbf{N}'as_i$$
$$\times \left[\hat{\beta}_{d,i}\mathbf{I} + \mathbf{X}(\mathbf{X}' \left[\mathbf{I} - \mathbf{S}\right] \mathbf{X} + \mathbf{I}_p)^{-1}(\xi_i \tilde{e}'_d - \hat{\beta}_{d,i}\tilde{\mathbf{X}}')\right] l - o(a^2) \tag{25}$$

Differentiating (24) and (25) with respect to a at $a = 0$ the proof will be complete. Analogous to the Sect. 3.1, the generalized Cook statistics can be written

$$GC_{\beta_\omega,d} = s_i^2 l'(\tilde{e}_d\xi_i' - \hat{\beta}_{d,i}\tilde{\mathbf{X}})(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\,\mathbf{X} + \mathbf{I}_p)^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\,\mathbf{X} + \mathbf{I}_p)^{-1}$$
$$\times(\xi_i\tilde{e}_d' - \hat{\beta}_{d,i}\tilde{\mathbf{X}}')l/s^2 tr(\mathbf{H}_\beta), \tag{26}$$

and

$$GC_{g_\omega,d} = s_i^2 l'(\tilde{e}_d\xi_i' - \hat{\beta}_{d,i}\tilde{\mathbf{X}})(\hat{\beta}_{d,i}\mathbf{I} + (\mathbf{X}'[\mathbf{I} - \mathbf{S}]\,\mathbf{X} + \mathbf{I}_p)^{-1}\mathbf{X}')$$
$$\times\mathbf{S}\left[\hat{\beta}_{d,i}\mathbf{I} + \mathbf{X}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\,\mathbf{X} + \mathbf{I}_p)^{-1}(\xi_i\tilde{e}_d' - \hat{\beta}_{d,i}\tilde{\mathbf{X}}')\right]l/s^2 tr(\mathbf{H}_g). \tag{27}$$

The diagnostic direction l_{max} can be obtained by finding the eigenvector corresponding to the largest absolute eigenvalue of matrices in (26) and (27) respectively. $\square$

4 Assessing influence on the selection of Liu parameter d

In this section, using local influence analysis, a method is given to study the detection of the possible influential observations in the data which may have a serious influence on the estimation of d. The selection criterion we used is given in (10), and the perturbation scheme is (11). Let $C_{d,\omega}$, $SSR_{\omega,d}$ and $\mathbf{H}_\omega$ denote the perturbed versions of C_d, SSR_d and $\mathbf{H}_d$ respectively. Let d_ω denote the estimator of d by minimizing

$$C_{d,\omega} = SSR_\omega/s^2 + 2tr(\mathbf{H}_{d,\omega}) - (n - 2).$$

Then the $l_{max}(\hat{d})$ which is the main diagnostic direction of local influence for $\hat{d}$ has the form

$$l_{max}(\hat{d}) \propto \frac{\partial \hat{d}_\omega}{\partial \omega}\,|_{\omega=\omega_0}\,.$$

Since $C_{d,\omega}$ achieves a local minimum at $\hat{d}_\omega$, we have

$$\frac{\partial C_{d,\omega}}{\partial d}\,|_{\hat{d}=\hat{d}_\omega} = 0. \tag{28}$$

Differentiating both sides of (28) with respect to ω and evaluating at ω_0, we obtain

$$\frac{\partial^2 C_{d,\omega}}{\partial d^2}\frac{\partial \hat{d}_\omega}{\partial \omega}\,|_{\omega=\omega_0,d=\hat{d}} + \frac{\partial^2 C_{d,\omega}}{\partial \omega \partial d}\,|_{\omega=\omega_0,d=\hat{d}} = 0.$$

We can get the following relation

$$\frac{\partial \hat{d}_\omega}{\partial \omega}\,|_{\omega=\omega_0} = -\Delta/\ddot{C}_d,$$

where $\Delta = \partial^2 C_{\omega,d} / \partial d \partial \omega \mid_{\omega=\omega_0, d=\hat{d}}$ and $\ddot{C}_d = \partial^2 C_{\omega,d} / \partial d^2 \mid_{\omega=\omega_0, d=\hat{d}}$. Under perturbation of variance, the sum of the squares of the residual $SSR_{d,\omega}$ and the hat matrix $\mathbf{H}_{d,\omega}$ in LPLSEs become $SSR_{d,\omega} = (y - \mathbf{X}\hat{\beta}_\omega - \mathbf{N}\hat{g}_\omega)'\mathbf{W}(y - \mathbf{X}\hat{\beta}_\omega - \mathbf{N}\hat{g}_\omega)$ and

$$\mathbf{H}_{d,\omega} = \mathbf{S}_\omega \mathbf{W} + (\mathbf{I} - \mathbf{S}_\omega \mathbf{W})\mathbf{X}(\mathbf{X}'[\mathbf{W} - \mathbf{W}\mathbf{S}_\omega \mathbf{W}]\mathbf{X} + \mathbf{I}_p)^{-1}$$
$$\times (\mathbf{X}'[\mathbf{W} - \mathbf{W}\mathbf{S}_\omega \mathbf{W}] + d(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X})^{-1}\mathbf{X}'[\mathbf{I} - \mathbf{S}]) \quad .$$

By the known matrix theory, we get

$$\frac{\partial SSR_{d,\omega}}{\partial \omega} \mid_{\omega=\omega_0} \approx e_{i,d}^2,$$

and

$$\frac{\partial e_{d,\omega}}{\partial d} = -\tilde{\mathbf{X}}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p)^{-1}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X})^{-1}\mathbf{X}'[\mathbf{I} - \mathbf{S}]y$$
$$= -\tilde{\mathbf{X}}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + I)^{-1}\hat{\beta},$$

where $\hat{\beta}$ is PLSE of β.

$$\frac{\partial^2 SSR_{d,\omega}}{\partial \omega \partial d} \mid_{\omega=\omega_0, d=\hat{d}} = -2e_{i,d}\tilde{x}_i'(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p)^{-1}\hat{\beta}.$$

A similar matrix partial differentiation for $tr(\mathbf{H}_{d,\omega})$ gives that

$$\frac{\partial^2 tr(\mathbf{H}_{d,\omega})}{\partial \omega \partial d} \mid_{\omega=\omega_0, d=\hat{d}} = -\left(S_i' + \tilde{x}_i'(\mathbf{X}'[\mathbf{I}-\mathbf{S}]\mathbf{X} + \mathbf{I}_p)^{-1}\tilde{\mathbf{X}}' \right)\tilde{\mathbf{X}}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p)^{-1}$$
$$\times (\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X})^{-1}\tilde{x}_i,$$

where S_i' is ith row of $\mathbf{S}$. Therefore, the ith element of $l_{max}(\hat{d})$ is given by

$$l_{max}^{(i)}(\hat{d}) \propto \Delta = -e_{i,d}\tilde{x}_i'(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p)^{-1}\hat{\beta}/s^2$$
$$- (S_i' + \tilde{x}_i'(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + \mathbf{I}_p)^{-1}\tilde{\mathbf{X}}')\tilde{\mathbf{X}}(\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X} + I)^{-1}$$
$$\times (\mathbf{X}'[\mathbf{I} - \mathbf{S}]\mathbf{X})^{-1}\tilde{x}_i. \tag{29}$$

5 Numerical illustration

5.1 Simulation study

A simulation study has been carried out in order to evaluate the performances of the proposed method in different situation. To achieve different degrees of collinearity,

following McDonald and Galarneau (1975), the explanatory were generated using the following device:

$$x_{ij} = (1 - \gamma)^{1/2} z_{ij} + \gamma z_{i3} \quad i = 1, \ldots, n \quad j = 1, \ldots, 3$$

where z_{ij} are independent standard normal pseudo-random numbers, and is specified so that the correlation between any two explanatory variables is given by γ^2. Three different sets of correlations corresponding to $\gamma = 0.80, 0.90,$ and 0.99 are considered. Then, n observations for the dependent variable are determined by

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + g(t_i) + \epsilon_i \quad i = 1, \ldots, n.$$

with $g(t_i) = cos(2\pi t_i)$, $t_i \sim U(0, 1)$ in which $U(0, 1)$ denotes the uniform distribution in interval $(0, 1)$. We vary the sample size with $n = 15, 30$ and $n = 50$. In ridge regression Newhouse and Oman (1971) stated that if the mean squared error is a function of β, σ, and ridge parameter and, if the explanatory variables are fixed, then the mean squared error is minimized when β is the normalized eigenvector corresponding to the largest eigenvalue of $\mathbf{X}'\mathbf{X}$ matrix subject to constraint that $\beta'\beta = 1$. Here we can selected the coefficients β_1, β_2 and β_3 as normalized eigenvectors corresponding to the largest eigenvalues of $\mathbf{X}'(\mathbf{I} - \mathbf{S})\mathbf{X}$ matrix so that $\beta'\beta = 1$. An outlier is created by adding ν to the response y_{10}, i.e., $y_{10} = y_{10} + \nu$, where ν corresponds to the standard deviation of response y. We calculate the diagnostic measures of $l_{max}^1(\hat{\beta}_d)$, $l_{max}^2(\hat{\beta}_d)$ and $l_{max}(\hat{g}_d)$ in different datasets. The results are shown in Table 1. It is easily seen from Table 1 that case 10 is the most influential observation. For example for $n = 15$ with $\gamma = 0.8$, the $l_{max}^1(\hat{\beta}_d)$, $l_{max}^2(\hat{\beta}_d)$ and $l_{max}(\hat{g}_d)$ have maximum values for case 10 compared to any other observation (which have values less than 0.104, 0.153 and 0.096 respectively). These results imply that our proposed diagnostic measures can identify the potential outlier.

5.2 Real data

The Longley (1967) data consisting of 7 economical variables, $x_1 = GNP$ implicit price deflator, $x_2 = Gross$ National Product, $x_3 = number$ of people in the armed forces, $x_4 = number$ of unemployed, $x_5 = Population$, $x_6 = Year$ and $y = number$ of people employed. This data has been used to explain the effect of extreme multicollinearity on the ordinary least squares estimators. The scaled condition number (see Walker and Brich 1988) of this data set is 43,275. This large value suggests the presence of an unusually high level of collinearity. Cook (1977) applied Cook's distance to this data and found that cases 5, 16, 4, 10, and 15 (in this order) were the most influential observations in OLS. Walker and Brich (1988) analysed the same data to find anomalous observations in ORR using the method of case deletion influential measures. They found that cases 16, 10, 4, 15 and 5 (in this order) were the most influential observations in Cook's and *DFFITS* measures. In local influence approach Shi and Wang (1999) find the cases 10, 4, 5 and 15 were the most influential observation for ridge estimation and Jahufer and Chen (2012) find the cases 4, 10, 1, 5 and 6 in

Table 1 Influential analysis of simulated data

n	γ	β	observation	$l^1_{max}(\hat{\beta}_d)$	$l^2_{max}(\hat{\beta}_d)$	$l_{max}(\hat{g}_d)$
15	0.8	$(.816, .419, .398)'$	10	.834	.819	.659
			Other	$\leq .104$	$\leq .153$	$\leq .096$
	0.9	$(.511, .699, .508)'$	10	.817	.847	0.701
			Other	$\leq .192$	$\leq .080$	$\leq .150$
	0.99	$(.301, .892, .338)'$	10	.884	.750	.865
			Other	$\leq .058$	$\leq .166$	$\leq .160$
30	0.8	$(.916, .325, .236)'$	10	.814	.790	.758
			Other	$\leq .111$	$\leq .099$	$\leq .063$
	0.9	$(.727, .428, .536)'$	10	.890	.721	.629
			Other	$\leq .101$	$\leq .164$	$\leq .051$
	0.99	$(.688, .611, .398)'$	10	.791	.801	.606
			Other	$\leq .097$	$\leq .131$	$\leq .091$
50	0.8	$(.303, .861, .410)'$	10	.981	.881	.792
			Other	$\leq .083$	$\leq .100$	$\leq .045$
	0.9	$(.399, .169, .902)'$	10	.916	.841	.761
			Other	$\leq .103$	$\leq .095$	$\leq .083$
	0.99	$(.526, .571, .631)'$	10	.971	.931	.806
			Other	$\leq .099$	$\leq .110$	$\leq .094$

this order were the five most influential observations for Liu estimator in ordinary regression. Recently, Emami (2015) used the same data to identify influential cases in ridge semiparametric regression model. By case deletion method he identified 12, 16, 2 and 5 were the most influential cases.

In this section, we use this data set to illustrate the method suggested in this article. The influence of observations on the LPLSEs of SPRM are studied based on small perturbations. Therefore, the influence of the different aspects of the model can be well approached. Here, we fit model (1) to the data, which $\mathbf{X} = [x_1, x_2, x_4, x_5, x_6]'$ and $g(t_i) = g(x_3)$. The parameter λ in this model is 0.007, which is obtained by minimizing GCV criterion. Estimate of nonparametric function for Longley data for $d = 0.985$ is shown in Fig. 1. First, we consider, the variance perturbation. The index plots of $l^1_{max}(\hat{\beta}_d)$ and $l_{max}(\hat{g}_d)$ are shown in Fig. 2, respectively (index plot of $l^2_{max}(\hat{\beta}_d)$ for $d = 0.985$ has a similar structure as $l^1_{max}(\hat{\beta}_d)$. In Fig. 2a, cases 16, 10, 15 and 5 are four most influential cases in $\hat{\beta}_d$. However, the largest absolute component of $l_{max}(\hat{g}_d)$ in Fig. 2b directs attention to cases 10, 16, 2, and 12 in order for $\hat{g}_d$. Therefore, local influential observations are slightly different from those in case deletion. This is partly due to the fact that local influence considers the joint influence instead of individual cases influence. Second, we considered the perturbation of individual explanatory variables. The maximum values of $l^1_{max}(\hat{\beta}_d)$ for separately perturbing explanatory variables x_j, ($j = 1, 2, 4, 5$ and 6) are 10.28, 0.761 ,5.43 ,2.11 and 1.21 respectively, and also the maximum values of $l_{max}(\hat{g}_d)$ are 8.91, 0.44, 3.28,
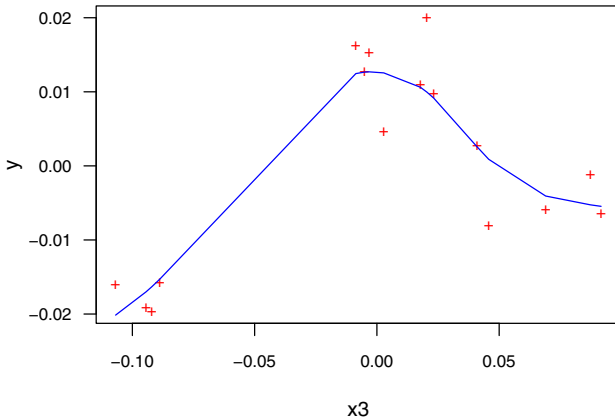
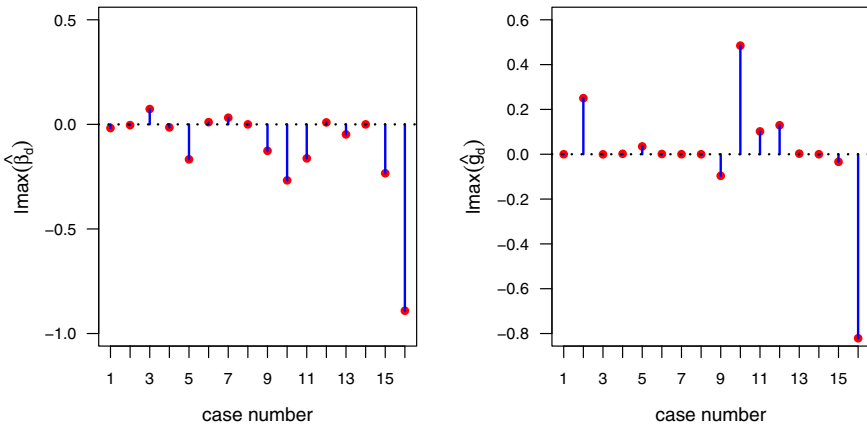**Fig. 1** Estimate of nonparametric function for Longley data with $d = 0.985$

Fig. 2 **a** (*left panel*) Index plot of $l_{max}^1(\hat{\beta}_d)$ with $d = 0.985$. **b** (*right panel*) Index plot of $l_{max}(\hat{g}_d)$ with $d = 0.985$

0.87 and 0.96 respectively. Hence local change caused by perturbing x_1 is the largest among the others and local changes by perturbing the other explanatory variables are almost the same. The index plots of $l_{max}^1(\hat{\beta}_d)$ and $l_{max}(\hat{g}_d)$ based on perturbation of x_j, $j = 1, 4, 5$ and 6 are listed in Figs. 3 and 4 respectively. Note that in these figures vertical scales have been chosen identically (except for the sign). From Fig. 3 it is obvious that the LPLSE $\hat{\beta}_d$ is sensitive for the values of x_1, x_4 and x_5 at cases 1 and 11 and 16 and values of x_5 at cases 5, 10 and 14. Also, from Fig. 4 LPLSE $\hat{g}_d$ is sensitive for the values of x_1, x_4, x_5 and x_6 at cases 2, 6 and 12. Finally we estimated the $l_{max}(\hat{d})$ values using (29). It is observed from the index plot of $l_{max}(\hat{d})$ shown in Fig. 5, cases 2, 10, 5, 15 and 16 in this order are the five most influential observations on LPLSEs parameter.
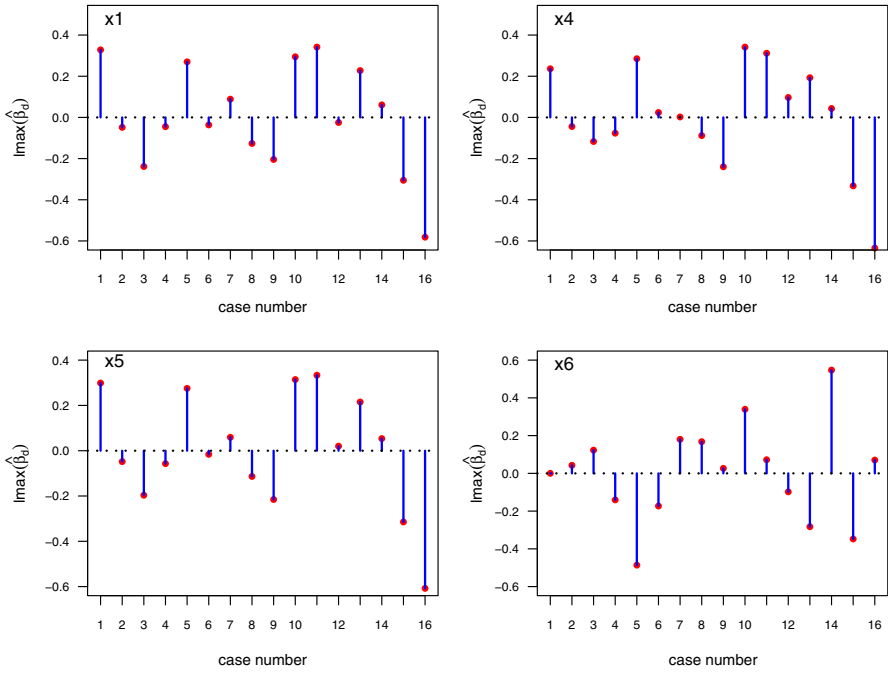
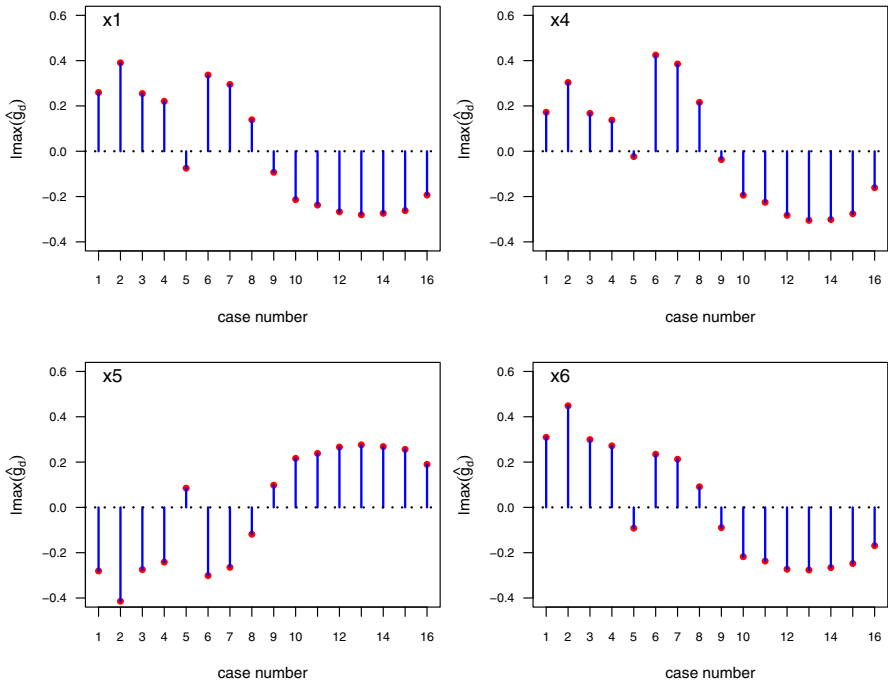**Fig. 3** Index plot of $l_{max}(\hat{\beta}_d)$ for separately perturbing x_1, x_4, x_5 and x_6

Fig. 4 Index plot of $l_{max}(\hat{g}_d)$ for separately perturbing x_1, x_4, x_5 and x_6
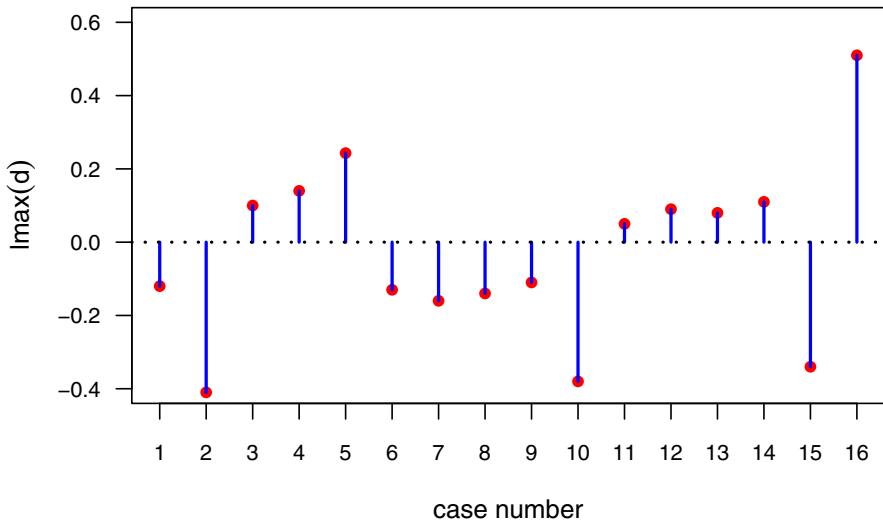
Fig. 5 Index plot of $l_{max}(\hat{d})$ with $d = 0.985$

6 Conclusion

Local influence diagnostics consider the joint influence of the data set, therefore it is useful to identify some influential patterns appearing in the data set. In this paper, we have studied several local influence diagnostic measures that seem practical and can play a considerable part in LPLSEs of SPRMs. Instead of using case deletion, we use the local influence method to study the detection of influential observations. By perturbing different aspects of the model, the influence compact of the data on the LPLSEs of SPRMs can be studied. The proposed techniques provide to the practitioner numerical and pictorial results that complement the analysis. We believe that the local influence diagnostics we derive here can be useful as part of any serious data analysis. All the proposed measures are the function of residuals, leverage points and LPLSEs. Furthermore, we study the influence of observations on selection of d, which is also important in Liu type regression models. Although no conventional cutoff points are introduced or developed for the Liu estimator local influence diagnostic quantities, it seems that index plot is an optimistic and conventional procedure to disclose influential cases.

References

Akdeniz F, Akdeniz Duran E (2010) Liu-type estimator in semiparametric segression models. J Stat Comput Simul 80:853–871

Akdeniz F, Akdeniz Duran E, Roozbeh M, Arashi M (2015) Effciency of the generalized difference-based Liu estimators in semiparametric segression models with correlated errors. J Stat Comput Simul 85:147–165

Amini M, Roozbeh M (2015) Optimal partial ridge estimation in restricted semiparametric regression models. J Multivar Anal 136:26–40

Arashi M, Valizadeh T (2015) Performance of Kibiria's methods in partial linear ridge regression models. Stat Pap 56:231–246

Belsley DA, Kuh E, Welsch RE (1989) Regression diagnostics: identifying influential data and sources of collinearity. Wiley, New York

Chen G, You J (2005) An asymptotic theory for semiparametric generalized least squares estimation in partially linear regression models. Stat Pap 46:173–193

Cook RD (1977) Detection of influential observations in linear regression. Technometrics 19:15–18

Cook RD (1986) Assessment of local influence (with discussion). J R Stat Soc B 48:133–169

Cook RD, Weisberg S (1982) Residuals and influence in regression. Chapman & Hall, New York

Emami H (2015) Influence diagnostics in ridge semiparametric regression models. Stat Probab Lett 105:106–115

Eubank RL (1985) Diagnostics for smoothing splines. J R Stat Soc B47:332–341

Fung WK, Zhu ZY, Wei BC, He X (2002) Influence diagnostics and outlier tests for semiparametric mixed models. J R Stat Soc B 47:332–341

Green PJ, Silverman BW (1994) Nonparametric regression and generalized linear models. Chapman and Hall, London

Hu H (2005) Ridge estimation of a semiparametric regression model. J Comput Appl Math B 176:215–222

Jahufer A, Chen J (2009) Assessing global influential observations in modified ridge regression. Stat Probab Lett 79:513–518

Jahufer A, Chen J (2012) Identifying local influential observations in Liu estimator. Metrika 75:425–438

Kim C (1996) Cook's distance in spline smoothing. Stat Probab Lett 31:139–144

Kim C, Lee Y, Park BU (2001) Cook's distance in local polynomial regression. Stat Probab Lett 54:33–40

Kim C, Lee Y, Park BU (2002) Influence diagnostics in semiparametric regression models. Stat Probab Lett 60:49–58

Lesaffre E, Verbeke G (1998) Local influence in linear mixed models. Biometrics 54:570–582

Liu K (1993) A new class of biased estimate in linear regression. Commun Stat Theory 22:393–402

Longley JW (1967) An appraisal of least squares programs for electronic computer from the point of view of the user. J Am Stat Assoc 62:819–841

Mallows CL (1973) Some comments on Cp. Technometrics 15:661–675

McDonald GC, Galarneau DI (1975) A Monte Carlo evaluation of some ridge-type estimators. J Am Stat Assoc 70:407–416

Newhouse JP, Oman SD (1971) An evaluation of ridge estimators. http://www.rand.org/pubs/reports/2007/R716

Rasekh A (2006) Local influence in measurement error models with ridge estimate. Comput Stat Data Anal 50:2822–2834

Roozbeh M (2015) Shrinkage ridge estimators in semiparametric regression models. J Multivar Anal 136:56–74

Roozbeh M, Arashi M (2015) Feasible ridge estimator in partially linear models. J Multivar Anal 116:35–44

Shi L (1997) Local influence in principal component analysis. Biometrika 87:175–186

Shi L, Wang X (1999) Local influence in ridge regression. Comput Stat Data Anal 31:341–353

Speckman PE (1988) Regression analysis for partially linear models. J R Stat Soc B 50:413–436

Stein C (1956) Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In: Proceedings of the third Berkeley symposium on mathematical statistics and probability vol 1, pp 197–206

Tabakan G, Akdeniz F (2010) Difference-based ridge estimator of parameters in partial linear model. Stat Pap 51:357–368

Thomas W, Cook RD (1989) Assessing influence on regression coefficients in generalized linear models. Biometrika 76:741–749

Walker E, Birch JB (1988) Influence measures in ridge regression. Technometrics 30:221–227

Zhu Z, He X, Fung W (2003) Local influence analysis for penalized Gaussian likelihood estimators in partially linear models. Scan J Stat 30:767–780