Identifying Local Influence in Modified Ridge Regression Using Cook's Method

A JAHUFER[1] AND C. JIANBAO[2]

[1]Department of Mathematical Sciences, Faculty of Applied Sciences, South Eastern University of Sri Lanka, Sammanthurai 32200, Sri Lanka.

[2]Department of Planning and Statistics, School of Economics, Xiamen University, Xiamen 361005, P. R. China

ABSTRACT

In this paper, we study an assessment of the minor perturbation of the modified ridge regression estimator (MRRE), in the ridge type linear regression model to detect influential case using Cook's (1986) method when, σ^2 is known or unknown and perturbation of the individual explanatory variable. An example on the Longley data is analyzed for illustration. The influential observations identified by this method is the same as the influence diagnostic methods of Cook (1986), Walker and Birch (1988) and Shi and Wang (1999) methods, but the order of the magnitudes are changed.

Key words: Influential Observations; Modified Ridge Regression; Collinearity; Cook's Local Influence Analyze

INTRODUCTION

The main aim of this study is to identify the local influential observations in modified ridge regression estimator (MRRE) using Cook's (1986) minor perturbation method. The influential observations substantially affect the parametric estimation and statistical inferences. There are many relative works that have received great attention in the last two decades. For the identification of influential observations, the most popular methods are global and local influence analyzes. In the global influence analysis, case deletion (Belsley *et al.*, 1980; Cook and Weisberg, 1982) is very valuable to study the effects of single case and it is the sample version of influence function. But it can not obtain a satisfactory result for the complex problem or possible masking effects. Under a kind of model perturbation, Cook (1986) introduced local influence analysis, based on likelihood function, in which the curvature of influential graph has been used as the indicator of influential point. This method can not only identify the influential point effectively, but also is useful for masking effects to some extent. It has received extensive application for its flexibility (MaCulloch, 1989; Thomas and Cook, 1989; Beckman *et al.*, 1987; Schall and Dunne, 1992).

Much effort has been devoted to the detection of influential observations in ridge type biased estimators. Among others, Shi and Wang (1999) studied the local influence of minor perturbations on the ridge regression estimator. The diagnostics under the perturbation of variance and explanatory variables are derived. Also, they developed a new technique for the detection of influential observations on the ridge biasing parameter. Moreover, Shi (1997) used local influence in principal component analysis. Walker and Birch (1988) analyzed the influence of observations in ridge regression using the case deletion method.

This paper is composed of five sections: Section 1 describes the introduction, Section 2 gives the background, Section 3 derives the identifying local influential observations for MRRE estimator, and Section 4 provides numerical illustration. Conclusion is given in the last section.

BACKGROUND

A Multiple Linear Regression Model

A matrix multiple linear regression model is

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathbf{Y}$ is an n vector of observable random variables, $\mathbf{X}$ is a n × p known matrix, $\boldsymbol{\beta}$ is a p vector of unknown parameters, $\boldsymbol{\varepsilon}$ is an n vector of unobservable errors with $E(\varepsilon)=0$ and $Var(\varepsilon)=\sigma^2$ and $\mathbf{I}$ is an identity matrix of the order n. The ordinary least squares estimator (OLSE) of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$ and the estimator of σ^2 is $s^2 = \frac{\mathbf{e'e}}{(n-p)}$, where residual vector $\mathbf{e} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$.

When there exists a strong near-linear relationship among the columns of matrix $\mathbf{X}$, it can be said that collinearity exists in the data set (Belsley *et al.*, 1980). In this situation, the matrix $\mathbf{X'X}$ becomes near singular or ill-conditioned, and least squares estimator possess undue sensitivity to the data (Belsley *et al.*, 1980; Belsley, 1991).

Since a popular numerical technique to deal with near collinearity is that of ridge type regression technique, the ordinary ridge regression estimator (ORRE) introduced by Hoerl (1964) is given by

$$\hat{\boldsymbol{\beta}}_R = (\mathbf{X'X} + k\mathbf{I})^{-1}\mathbf{X'Y}, \qquad (2)$$

where k>0 is the ridge estimator biasing parameter. It is a biased estimator, however, the variances of its elements are less than the variance of the corresponding elements of the OLSE for suitable k. By accepting some bias to reduce variance, the mean squared error (MSE) might thus be improved.

Modified Ridge Regression

Swindel (1976) introduced one ridge type estimator based on prior information which is called MRRE. It is defined as

$$\hat{\boldsymbol{\beta}}_{(k,b)} = (\mathbf{X'X} + k\mathbf{I})^{-1}(\mathbf{X'Y} + k\mathbf{b}), \qquad (3)$$

where k is MRRE biasing parameter and **b** is an m×1 prior information vector. $\hat{\boldsymbol{\beta}}_{(k,b)}$ approaches **b** strictly monotonically as $k \to +\infty$ (whatever **X** and **Y**); $\hat{\boldsymbol{\beta}}_{(k,b)}$ is close to **b** in an equivalence class of estimators of $\boldsymbol{\beta}$ with the same residual sum of squares; and $\hat{\boldsymbol{\beta}}_{(k,b)}$ has minimum residual sum of squares in an equivalence class of estimator of $\boldsymbol{\beta}$ which has equal distance from **b**. With $\hat{\boldsymbol{\beta}}_{(k,b)}$ so inclined to **b**, in application, **b** might be chosen to reflect the prior information or hypothesis on $\boldsymbol{\beta}$. Formally then, call **b** the prior information on $\boldsymbol{\beta}$, and $\hat{\boldsymbol{\beta}}_{(k,b)}$ (0< k) the family of ridge estimators of $\boldsymbol{\beta}$ based on prior information **b**. Given this determination of **b**, the choice of a particular k, and hence $\hat{\boldsymbol{\beta}}_{(k,b)}$, is seen as one of the compromises between the best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$ based on the sample data and $\hat{\boldsymbol{\beta}}_{(+\infty,b)} \equiv \mathbf{b}$. The approach of $\hat{\boldsymbol{\beta}}_{(k,b)}$ to **b** requires no pleading, that is, the intuitive arguments for shrinking BLUE toward **b** are substantially stronger than those for shrinking BLUE toward **0**. Moreover, the MRRE introduced by Swindel is more stable than the OLSE.

A number of studies reveal the effect on MRRE. For instance, Pliskin (1987) compared the mean squared error matrix of the ORRE and MRRE (Kaçiranlar et al., 1998; Wijekoon, 1998) analyzed mean squared error comparisons of the restricted ridge regression estimator (RRRE) and MRRE, and Groß (2003) used this estimator to develop a new restricted ridge estimator.

The key problem in the MRRE is to choose the value of k. There are several methods for selecting the value of k, among which Hoerl and Kennard's an iterative procedure (Hoerl and Kennard, 1976), Mcdonald-Galarneau's method (McDonald and Galarneau, 1975), C_p statistic criterion (Mallows, 1973), GCV criterion (Wahba et al., 1979), PRESS procedure and VIF procedure (Marquardt, 1970) are popular.

Cook's Influential Measures

The local influence approach was proposed by Cook (1986) as a general method for assessing the influence of minor perturbations of a statistical model. The method makes use of differential geometry techniques to assess the change of the likelihood function due to perturbations of the model.

In his development of a general and unified approach to the diagnostic analysis of perturbation models, Cook (1986) modeled perturbation of a statistical model as a vector $\omega_{q\times 1}$ from a perturbation space $W \subset R^q$. When $L(\theta)$ is the log likelihood of the postulated model, the perturbation ω result in a perturbed log likelihood $L(\theta,\omega)$, where it is assumed that a `null' perturbation $\omega_0 \in W$ exists such that $L(\theta) = L(\theta,\omega_0)$, for all θ. Based on this approach, Cook (1986) proposed to assess the influence of perturbations by the likelihood displacement

$$LD(\omega) = 2[L(\hat{\theta}) - L(\hat{\theta}(\omega))]. \tag{4}$$

Here $\hat{\theta}(\omega)$ is the maximum likelihood estimate for θ under the log likelihood $L(\theta,\omega)$, for given ω, so that $\hat{\theta} = \hat{\theta}(\omega_0)$ is the maximum likelihood estimator of θ under the log likelihood of the postulated model $L(\theta) \doteq L(\theta,\omega_0)$. The associated influence graph is defined as the surface

$$\alpha(\omega) = \begin{pmatrix} \omega \\ LD(\omega) \end{pmatrix}.$$

For q=1, Cook (1986) proposed to assess the local influence of the perturbation ω by the curvature C_ω at C_0, the postulated model. He showed that

$$C_\omega = 2\left|\ddot{F}\right| = 2\left|\Delta' \ddot{L}^{-1} \Delta\right|,$$

where

$$\ddot{F} = \frac{\partial^2 (\hat{\theta}(\omega))}{\partial \omega^2}, \qquad \Delta_{p\times 1} = \frac{\partial^2 L(\theta,\omega)}{\partial \theta \partial \omega'} \qquad and \qquad \ddot{L}_{p\times p} = \frac{\partial^2 L(\theta)}{\partial \theta \partial \theta'}$$

Here and in the following, all derivatives of $L(\theta,\omega)$ are evaluated at the postulated model, that is, at $\hat{\theta}$ and ω_0. For q > 1, the curvature in direction $d \in R^q$ is given by

$$C(\ell'\omega) = 2\left|\ell' \ddot{F} \ell\right| = 2\left|\ell' \Delta' \ddot{L}^{-1} \Delta \ell\right|, \tag{5}$$

where $\|\ell\| = 1$, $\ddot{F}_{q\times q} = \dfrac{\partial^2 (\hat{\theta}(\omega))}{\partial \omega \partial \omega'}$, $\Delta_{p\times q} = \dfrac{\partial^2 L(\theta,\omega)}{\partial \theta \partial \omega'}$,

and $\ddot{L}$ is defined as before. A large curvature $C_{(\ell'\omega)}$ indicates a departure of $(\ell'\omega)$ from $(\ell'\omega_0)$ as influential.

IDENTIFYING LOCAL INFLUENTIAL OBSERVATIONS FOR MRRE

Local Influence for MRRE

There are two types of perturbations that can be applied to detect influential observations: they are the multiplicative type and the additive type perturbations. In the multiplicative type perturbation reduces to the case-weights perturbation (Cook, 1986) which assigns a weight ω_{ij} to each observation and individual influential observations can be identified. Case-weights perturbation is a popular method in influence analysis literature because of its relationship to the intuitive diagnostic methods of case deletion (Cook, 1986; Lawrance, 1991).

In this section the assumption of constant variance in model Equation (3) is perturbed. This perturbation scheme is a better way to handle cases badly modeled (Lawrance, 1988). The distribution of under perturbation becomes

$$\varepsilon_{\omega} \sim N(\mathbf{0},\sigma^2\mathbf{W}^{-1}), \tag{6}$$

where $\mathbf{W} = diag(\omega)$ is a diagonal matrix with diagonal elements of $\omega' = (\omega_1, \omega_2, \ldots, \omega_n)$. Let $\omega = \omega_0 + al$, where $\omega_0 = 1$, ω denotes the n × 1 vector of case-weights for the regression model Equation (3), $a \in R^1$ and l is a fixed nonzero vector of unit length in R^q.

Suppose, it is assumed that σ^2 is known, then the relevant part of the multiplicative perturbed likelihood model is

$$L(\varepsilon_{\omega}) = \prod_{i=1}^{n} (2\pi\sigma^2)^{\frac{-n}{2}} |\mathbf{W}|^{\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}\mathbf{e}'_{(k,b)} W^{-1}\mathbf{e}_{(k,b)}\right], \tag{7}$$

where $\mathbf{e}_{(k,b)} = (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{(k,b)})$.

According to the discussion in section 1, influential patterns can be identified by studying the eigenvector $l^{m}_{\max}$ associated with maximum eigen value of the normal curvature matrix $C_{l^{\prime}\omega}$ in Equation (2). The component matrices Δ and $\ddot{L}$ are given below, as

$$\Delta = \frac{\partial^2 L(\varepsilon_{\omega})}{\partial\hat{\beta}_{(k,b)}\partial\omega}\Big|\hat{\beta}_{(k,b)}, \omega_0, \sigma^2 = \frac{1}{\sigma^2}\mathbf{X}'\mathbf{D}(\hat{e}_{(k,b)})$$

and

$$\ddot{L} = \frac{\partial^2 L(\varepsilon_{\omega})}{\partial\hat{\beta}^2_{(k,b)}}\Big|\hat{\beta}_{(k,b)}, \omega_0, \sigma^2 = -\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X}$$

where $D(\hat{e}_{(k,b)}) = diag(\hat{e}_{(k,b)1}, \ldots, \hat{e}_{(k,b)n})$.

Therefore, $C_{l'\omega}$ is

$$C_{l'\omega} = 2\left|l'\Delta'\ddot{L}^{-1}\Delta l\right| = \frac{2}{\sigma^2}\left|l'\mathbf{D}(\hat{e}_{(k,b)}\mathbf{P}_X\,\mathbf{D}(\hat{e}_{(k,b)})l\right|, \tag{8}$$

where

$\mathbf{P}_X = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, and $\|l\| = 1$. In what follows, $\mathbf{P}_M$ will be used to denote the projection operator for the column space of the matrix **M**, and $\mathbf{P}_M^{'} = \mathbf{I} - \mathbf{P}_M$.

When σ^2 is unknown, a similar calculation for $\theta' = (\beta_{(k,b)}^{'}, \sigma^2)$ yields

$$\Delta = \begin{pmatrix} \mathbf{X}'\mathbf{D}(e_{(k,b)}/\hat{\sigma}^2 \\ \mathbf{e}_{sq}^{'}/2\hat{\sigma}^4 \end{pmatrix} \tag{9}$$

where $\hat{\sigma}^2$ is the maximum likelihood estimator of σ^2 and $\mathbf{e}_{sq}$ is the n × 1 vector with elements $e_{(k,b)i}^2$. Since

$$\ddot{L} = -\begin{pmatrix} \mathbf{X}'\mathbf{X}/\hat{\sigma}^2 & 0 \\ 0 & n/2\hat{\sigma}^4 \end{pmatrix} \tag{10}$$

the analogous result for θ is

$$C_{l'\omega} = 2l'\left(\mathbf{D}(e_{k,b)})\mathbf{P}_X\,\mathbf{D}(e_{k,b)}) + e_{sq}e_{sq}^{'}/2n\hat{\sigma}^2\right)l/\hat{\sigma}^2\ , \tag{11}$$

General analytic expressions for l_{max} are not known for Equation (8) or Equation (11).

If only $\boldsymbol{\beta}$ is of interest, the curvature is given by Equation (8) with σ^2 replaced with $\hat{\sigma}^2$.

Hence the i^{th} curvature is given by

$$C_{(l'\omega)i} = \frac{1}{\hat{\sigma}^2} e^2_{(k,b)i} \mathbf{X}_i' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_i + \frac{1}{2n\hat{\sigma}^4} e^4_{(k,b)i} \tag{12}$$

where x_i is the i[th] row of matrix **X**.

$C_{l'\omega}$ measures the influence of local changes in the case weight. Clearly, a case that is locally influential must be globally influential, but the reverse need not be true. Both types of information can be useful, depending on the concerns of the investigator. However, when considering many cases simultaneously, it will be easier to use Equation (8) to characterize the local behavior of an influence graph than to use multiple-case-deletion diagnostics to characterize global behavior by using the corners of **W**.

Perturbation of the Explanatory Variables

It is well known that minor perturbation of the explanatory variables in linear regression can seriously influence the results of a least squares analysis when collinearity is present. Such results may also be influenced by a few isolated errors in the values of the explanatory variables or by a few values that are widely separated from the remaining data. Here we assume that σ^2 is known. The following results can be easily adapted for the situation in which σ^2 is unknown and only $\beta_{(k,b)}$ is of interest by replacing σ^2 and $\hat{\sigma}^2$.

Suppose s_j, j=1, ... , p, denote scale factors to account for the different measurement units associated with the columns of **X**, then the perturbed log-likelihood $L(\beta_{(k,b)} \mid \omega)$ is constructed from Equation (3) with **X** replaced by

$$\mathbf{X}_\omega = \mathbf{X} + \mathbf{WS} \tag{13}$$

where $W = (\omega_{ij})$ is an n × p matrix of perturbations and $\mathbf{S} = diag(S_1, \ldots, S_n)$. The perturbed log-likelihood is given by

$$L(\boldsymbol{\beta}_{(k,b)}, \mathbf{W}) = -\frac{1}{2\sigma^2} \left[\mathbf{Y} - (\mathbf{X} + \mathbf{WS})\boldsymbol{\beta}_{(k,b)}\right]' \left[\mathbf{Y} - (\mathbf{X} + \mathbf{WS})\boldsymbol{\beta}_{(k,b)}\right] \tag{14}$$

The curvature C_{ij} associated with the perturbation ω_{ij} is given by

$$C_{ij} = \frac{2\Delta_{ij}^{'}(\mathbf{X}'\mathbf{X})^{-1}\Delta_{ij}}{\sigma^2}. \tag{15}$$

Here Δ_{ij} is the i^{th} column of the matrix $\Delta_j = s_j(\mathbf{d}_j\hat{\mathbf{e}}_{(k,b)}^{'} - \hat{\boldsymbol{\beta}}_{(k,b)j}\mathbf{X}_j)$, where s_j is the j^{th} scale factor, $\mathbf{d}_j$ is a p ×1 vector with a 1 in the j^{th} position and zeros elsewhere, $\hat{\mathbf{e}}_{(k,b)}$ is the vector of MRRE residuals, and $\hat{\boldsymbol{\beta}}_{(k,b)j}$ is the j^{th} MRRE parameter.

In this application, $\Delta_{ij}^{'}(\mathbf{X}'\mathbf{X})^{-1}\Delta_{ij}/\sigma^2$ is a potentially large np × np matrix and determining its eigenvalues may be an unpleasant task. However, it can be shown that the nonzero eigenvalues of $\Delta_{ij}^{'}(\mathbf{X}'\mathbf{X})^{-1}\Delta_{ij}/\sigma^2$ are

$$\frac{1}{\sigma^2}\left[\mathbf{e}_{(k,b)}^{'}\mathbf{e}_{(k,b)}\delta_i + \sum_j \hat{\beta}_{(k,b)j}^2 s_j^2\right], \tag{16}$$

where δ_i eigenvalue of $\mathrm{S}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{S}$, i=1, ..., p Thus

$$C_{\max} = \frac{2}{\sigma^2}\left[\mathbf{e}_{(k,b)}^{'}\mathbf{e}_{(k,b)}\delta_{\max} + \sum_j \hat{\beta}_{(k,b)j}^2 s_j^2\right] \tag{17}$$

Except in the special situation discussed below, an analytic form for $l_{\max}$ is unknown.

The above results can be used in situations in which less than p explanatory variables are perturbed by setting $s_j = 0$ for the unperturbed variables. In particular, when only the i^{th} column of X is perturbed, $s_j = 0$ for $j \neq i$ and $\ddot{\mathbb{F}} = \Delta_i^{'}(\mathbf{X}'\mathbf{X})^{-1}\Delta_i\sigma^2$ where Δ_i is given in above with j = i.

Using this identity, $C_{\max}$ is obtained from Equation (17) as

$$C_{\max} = \frac{2s_i^2}{\sigma^2}\left[\|e_{(k,b)}\|\|r\|^{-2} + \hat{\beta}_{(k,b)i}^2\right], \tag{18}$$

where $\|r\|^{-2} = \mathbf{d}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{d}_i$.

Cook (1986) proposed that the curvature values depend on an essentially arbitrary choice of the matrix S. Thebsolution is to compute the scaled curvature that Schall and Dunneb (1992) suggested where by a treatment of local influence in terms of scaled curvature is equivalent to a treatment in terms of curvature Equation (8) with the canonical scaling matrix $\mathbf{S} = \sigma diag(\hat{\beta}_1^{-1},...,\hat{\beta}_p^{-1})$.

Hence, when the i[th] individual explanatory variable is perturbed, the l_{max} can be obtained by finding the eigenvector corresponding to the largest absolute eigenvalue of matrix

$$\frac{2s_i^2}{\sigma^2}\left[(\mathbf{d}_i\hat{\mathbf{e}}_{(k,b)}' - \hat{\boldsymbol{\beta}}_{(k,b)i}\mathbf{X}')'(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{d}_i\hat{\mathbf{e}}_{(k,b)}' - \hat{\boldsymbol{\beta}}_{(k,b)i}\mathbf{X}')\right]. \qquad (19)$$

NUMERICAL ILLUSTRATION

The Longley (1967) data is used for numerical illustration. This data consist of 6 explanatory variables and 16 observations. The scaled condition number for this data set is 43,275 (Walker and Birch, 1988), which suggests the presence of strong collinearity. Cook (1977) identified cases 5, 16, 4, 10, and 15 as the most influential observations by using the Cook statistic in least squares regression. Walker and Birch (1988) studied the influence of observations on the ridge regression by employing approximated case deletion formulae. They identified cases 16, 10, 4, 15, and 1 as the most influential observations by inspection of Cook statistic $\mathbf{D}_i^*$. Therefore the influence of case 5 is canceled in ridge regression. The effects of ridge parameter k on the influence of observations are also studied by plotting $\mathbf{D}_i^*$ or DFFITS$_i$ against k. The value of k that minimizes for these data set is 0.0002.

Shi and Wang (1999) studied the influence of observations on the ridge estimator by assessing influence on the selection of ridge parameter based on small perturbations. Using the variance perturbation method, they showed that cases 10, 4, 5 and 15 are the four most influential cases, and cases 6, 1 and 16 have a moderate influence in least squares regression. Cases 10, 4, 15, 16, and 1 in order are influential cases in the ridge regression. This is similar to that of case deletion, but the order of influence cases magnitude is changed.

In this paper, Cook's (1986) method is used to identify the local influential observations for biased MRRE in linear ridge type regression model when σ^2 is known or unknown and perturbation of individual explanatory variable. We have used the prior information vector and biasing parameter of MRRE is same as the ORRE (see Trenkler, 1988).

First, we consider the influence of observations on the MRRE based on the eigenvector l_{max} associated with maximum eigen value of the normal curvature matrix C_d in Equation (8). The plot l^m_{max} against observation number is given in Figure 1. When the i[th] element of l_{max} is found to be relatively large, this indicates that perturbations in the weight ω_i of the i[th] case may lead to substantial change in the results of the analysis and thus ω_i is relatively influential. In such situations, it will, of course, be important to investigate the i[th] case to find the specific cause of the sensitivity.

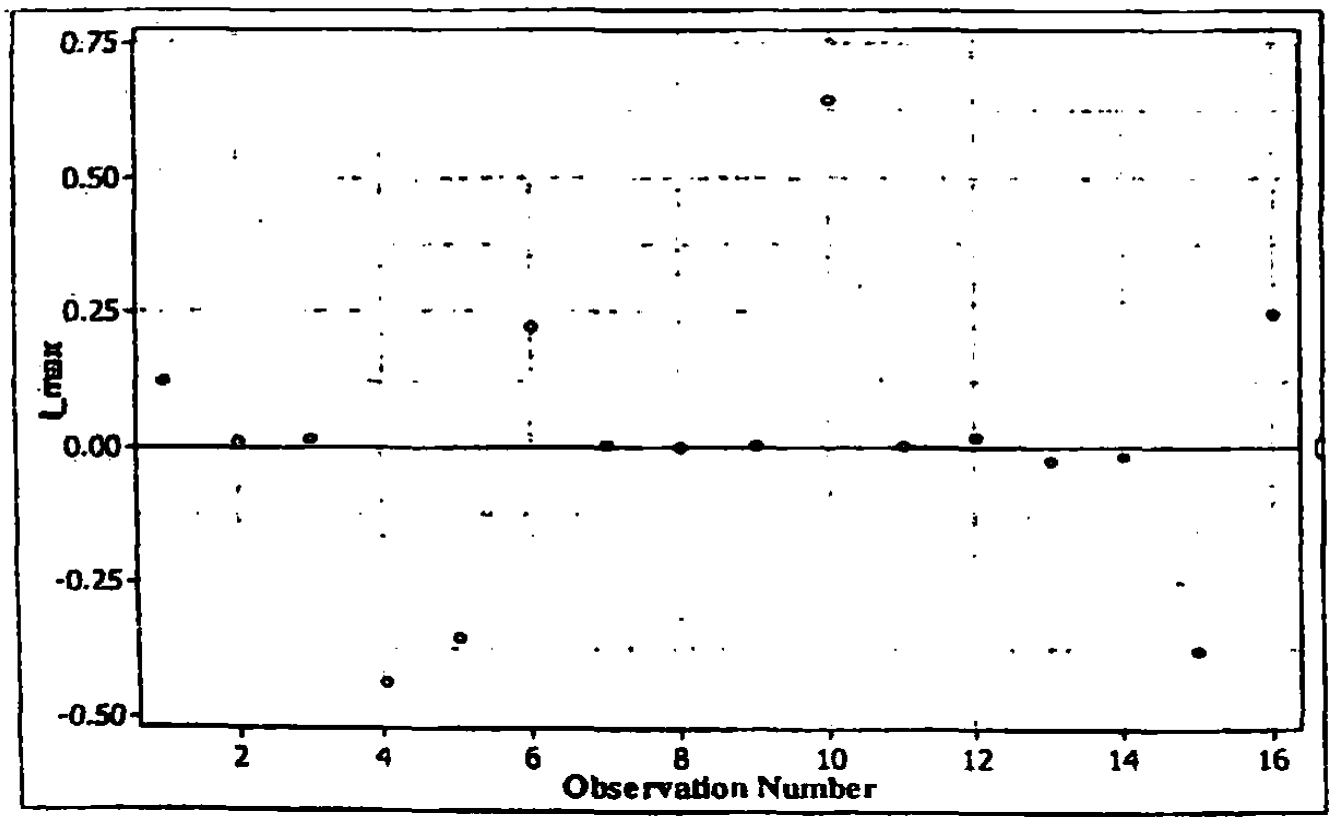

Figure 1: Index plot of l_{max} against Observation Number

Based on Figure 1, the most influential observations are cases 10, 4, 15, 5, 16 and 6 in this order. These influential observations are approximately the same for case deletion method but the order of magnitude is changed. Moreover, case 1 had been identified as high influential observation in Walker and Birch's (1988) and in Shi and Wang's (1999) methods. However, this does not occur in either Cook's (1977) or our method. Case 6 is not an influential point in Walker and Birch's and in Cook's method. However, it is a moderated influential observation in Shi and Wang's (1999) and in our method, respectively. The most five influential observations in ORRE, ORRE, and MRRE are given in table 1. From this table it can be observed that the influence observations are same but the order of the magnitudes is changed.

Table 1: The Most Influence Observations according to Cook's D_i: Longly Data

OLSE		ORRE		MRRE	
Case	Cook's D_i	Case	Cook's D_i	Case	Cook's D_i
5	0.614	16	0.582	10	0.526
16	0.467	10	0.251	4	0.229
4	0.244	4	0.219	15	0.216
10	0.235	15	0.145	5	0.142
15	0.170	5	0.142	16	0.121

Second, local influential observations for MRRE are analyzed using Equation (12), when σ^2 is unknown. The largest six $C_{(l'\omega)i}$ occurred for cases 10, 4, 5, 15, 16 and 1, in this order. Therefore, in this method, the most influential observations are cases 10, 4, 5, 15, 16 and 1. These influence observations are approximately the same as Cook's (1977), Walker and Birch's (1988) and Shi and Wang's (1999) methods, except that the order of magnitude is changed.

Finally, we consider the perturbation of individual explanatory variables. We use the scale factors of explanatory variables as suggested by Schall and Dunne (1992). The maximum value of C_{max} in Equation (19) for separately perturbing explanatory variables x_i, i = 1, ..., 6 are 569.222, 16.971, 2.861, 2.683, 356.403, and 2.930, respectively. Hence x_1, x_5, and x_2 are the largest values (in this order) among the others on $\hat{\beta}_{(k,b)}$. The index plots of l_{max} based on perturbation of x_1, x_2, and x_5 are given in figure 2. It is obvious that most influential cases in figure 2 are cases 10, 4, 15, 5, 1, and 6 in (A), and cases 10, 4, 15, 1, 6, and 5 in (C), respectively. But in (B) the most influential cases are 10, 5, 4, 15 1, and 7. These imply that the MRRE $\hat{\beta}_{(k,b)}$ is sensitive for values of x_1 and x_5 at cases 10, 4 and 15, and values of x_2 at cases 10 and 5.

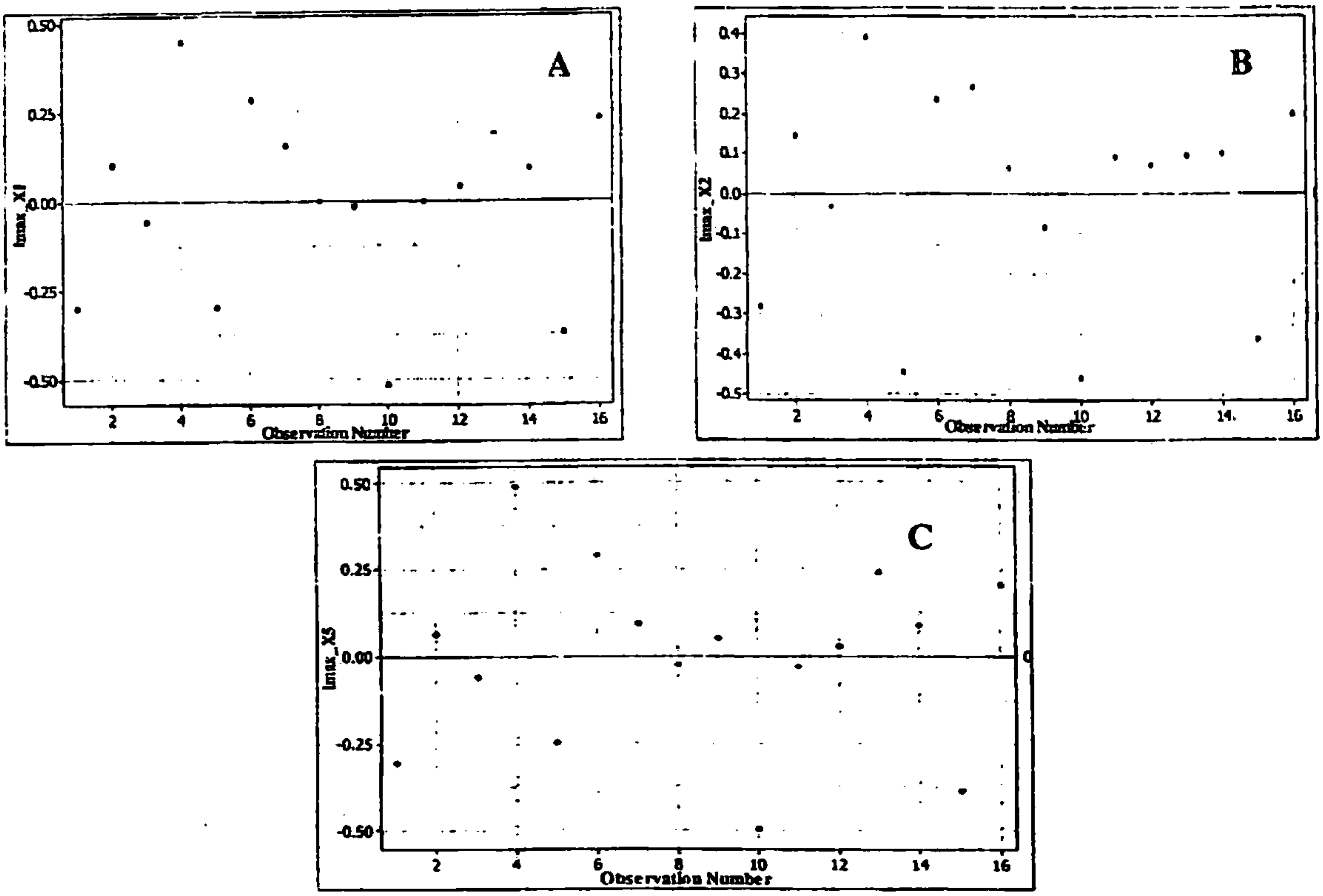

Figure 2: Index plot of $l_{max}X_1$, $l_{max}X_2$ and $l_{max}X_5$ against Observation Number respectively.

CONCLUSION

According to Belsey *et al.*, (1980), when biased estimation techniques are used to reduce the effect of collinearity, the influence of some cases can be modified. Based on this fact, Belsey *et al.*, (1980) suggested that collinearity can be controlled before attempting to measure influence. It is well known that a large number of biased estimators have been proposed to fit the regression model when collinearity presents among the regressors. But few methods were constructed for detecting influential observations for biased estimators.

In this paper, we propose Cook's (1986) statistical influence diagnostic method to detect influential observations for MRRE when σ^2 is known or unknown, and perturbation of individual explanatory variable. The influential observations identified by this method is the same as the influence diagnostic methods of Cook's, Walker and Birch's and Shi and Wang's methods, but the order of the magnitudes are changed. A case-by-case deletion method often suffers from masking effects. So, instead of deleting influential cases one by one, the method proposed in this paper is an alternative method to detect the anomalous observations for MRRE.

ACKNOWLEDGEMENTS

Much of this paper is taken from the first author's Ph.D. thesis research work, under the supervision of the second author at the Xiamen University of China. This work was supported by the Chinese Scholarship Council (CSC) government of P.R. of China–2006. The authors are thankful to Prof. P. Wijekoon of University of Peradeniya, Sri Lanka for her constructive comments.

REFERENCES

Beckman, R.J., C.J. Nachtsheim and R.D. Cook (1987). Diagnostics for mixed-model analysis of variance. *Technometrics,* 27(4): 413-426.

Belsley, D.A. (1991). *Conditioning Diagnostics: Collinearity and Weak Data in Regression.* Wiley New York.

Belsley, D.A., E. Kuh and R.E. Welsch (1980). *Regression Diagnostics: Identifying Influence data and Source of Collinearity.* Wiley New York.

Cook, R.D. (1977). Detection of influential observations in linear regression. *Technometrics,* 19: 15-18.

Cook, R.D. (1986). Assessment of local influence. *Journal of Royal Statistical Association* Series, B48: 133-169.

Cook, R.D. and S. Weisberg (1982). *Residuals and Influence in Regression.* Chapman & Hall, London.

Groß, J. (2003). Restricted ridge estimation. *Statistics & Probability Letter,s* 65: 57-64.

Hoerl, A.E. (1964). Ridge analysis. *Chemical Engineering Process Symposium,* Series60: 67-77.

Hoerl, A.E. and R.W. Kennard (1976). Ridge regression: Iterative estimation of the biasing parameter. *Communications in Statistics - Theory and Methods,* A5: 77-88.

Kaçiranlar, S., S. Sakallioğlu and F. Akdeniz (1998). Mean Squared Error comparisons of the modified ridge regression estimator and the restricted ridge regression estimator. *Communications in Statistics - Theory and Methods*, 27(1): 131-138.

Lawrance, A.J. (1988). Regression transformation diagnostics using local Influence. *Journal of American Statistical Association*, 83: 1067-1072.

Lawrance, A.J. (1991). Local and deletion influence. (In) *Directions in Robust Statistics and Diagnostics, Part 1.* W. Stahel and S. Weisberg (Eds.). Pp. 141-157. Springer, Berlin.

Longley, J.W. (1967). An Appraisal of least squares programs for electronic computer from the point of view of the user. *Journal of American Statistical Association*, 62: 819-841.

Marquardt, D.W. (1970). Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, 12: 591-612.

Mallows, C.L. (1973). Some comments on Cp. *Technometrics*, 15: 661-675.

Mcdonald, G.C. and D.I. Galarneau (1975). A Monte Carlo evaluation of some ridge-type estimators. *Journal of American Statistical Association*, 70: 407-416.

MaCulloch, R.E. (1989). Local model influence. *Journal of the American Statistical Association*, 84: 473-478.

Pliskin, J.L. (1987). A ridge type estimator and good prior means. *Communications in Statistics - Theory and Methods*, 16(12): 3429-3437.

Schall, R. and T.T. Dunne (1992). A note on the relationship between parameter collinearity and local influence. *Biometrika*, 79: 399-404.

Shi, L. (1997). Local influence in principal component analysis. *Biometrika*, 84(1): 175-186.

Shi, L. and X. Wang (1999). Local influence in ridge regression. *Computational Statistics & Data Analysis*, 31: 341-353.

Swindel, B.F. (1976). Good ridge estimators based on prior information. *Communications in Statistics - Theory and Methods*, A5(11): 1065-1075.

Thomas, W. and R.D. Cook (1989). Assessment on Regression Coefficients in GLM. *Biometrika*, 76(4): 741-749.

Trenkler, G. (1988). Some remarks on a ridge-type-estimator and good prior means. *Communications in Statistics - Theory and Methods*,17(12): 4251-4256.

Wahba, G., G.H. Golub and C.G. Heath (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21: 215-223.

Walker, E. and J.B. Birch (1988). Influence measures in ridge regression. *Technometrics*, 30: 221-227.

Wijekoon, P. (1998). A comparison of the restricted ridge regression and the modified ridge regression estimators. *Ceylon Journal of Science: Physical Sciences*, 5(1): 75-79.