Journal of the Korean Data & Information Science Society
2015, **26**(2), 465–474
http://dx.doi.org/10.7465/jkdi.2015.26.2.465
한국데이터정보과학회지

The local influence of LIU type estimator in linear mixed model†

Lili Zhang[1] · Jangsun Baek[2]

[12]Department of Statistics, Chonnam National University
Received 20 February 2015, revised 18 March 2015, accepted 23 March 2015

Abstract

In this paper, we study the local influence analysis of LIU type estimator in the linear mixed models. Using the method proposed by Shi (1997), the local influence of LIU type estimator in three disturbance models are investigated respectively. Furthermore, we give the generalized Cook's distance to assess the influence, and illustrate the efficiency of the proposed method by example.

Keywords: Generalized Cook's distance, linear mixed model, LIU type estimator, local influence.

1. Introduction

The linear mixed model is a statistical model containing both fixed effects and random effects. This model is frequently used in physical, biological and social science.

Detecting outliers and influential observations is an important step in the analysis of data. There exist many methods to assess the influence of data and model perturbations on the parameter estimates in linear mixed model, including case deletion, robust diagnostics and local influence. Diagnostic techniques for the linear regression model have received great attention in statistical literature. See Belsley *et al.* (1980), Cook and Weisberg (1982), Cook (1986), Rousseeuw and Leory (1987) and Chatterjee and Hadi (1988).

Case deletion is a method which is a global influence analysis to assess the individual impact of cases on the estimation process. We can find the method in Ronald and Larry (1992) and Shi and Chen (2008). On the other hand, for local influence considered in Cook (1986), instead of removing the case completely, giving each case a weight, the influence is assessed by perturbing these weights, and measured by normal curvature of an influence graph based on likelihood displacement. The application of this method can be found in Beckman *et al.* (1987) for linear mixed model, Lawrance (1988), Thomas and Cook (1990). In fact many statistical problems do not provide us the exact parametric distribution for the data. In this case, we cannot use Cook's method which is based on model likelihood replacement. Shi (1997) studied the local influence by defining a generalized Cook's statistic, and showed that his method is equivalent to Cook's approach under the likelihood framework.

† This study was financially supported by Chonnam National University, 2012.
[1] Ph.D. candidate, Department of Statistics, Chonnam National University, Gwangju 500-757, Korea.
[2] Corresponding author: Professor, Department of Statistics, Chonnam National University, Gwangju 500-757, Korea. E-mail: jbaek@jnu.ac.kr

In linear regression model, the existence of collinearity can lead to a very sensitive least square estimator. Statisticians proposed many biased estimators to improve least square estimator, such as the ridge estimator, the generalized ridge estimator, the principal component analysis estimator and the characteristic root estimator. Liu (1993) proposed a new biased estimator named LIU estimator. There are many good properties for LIU type estimator and its application has broad prospects. Many literatures are discussed for the LIU estimator (Liu, 2004; Liu, 2011, Zhang and Zhang, 2010; Li and Yang, 2012).

We extend the idea of the LIU type estimator of the linear regression model, and suggest a new LIU type estimator which can be applied to the linear mixed model. Then we investigate the local influence analysis of the LIU type estimator of fixed parameters in the linear mixed model under some perturbing assumptions.

In this paper, we propose a generalized LIU type estimator for the parameter in the linear mixed model, and then assess the local influence using the Shi's method with different perturbation structures. In Section 2, we give the LIU type estimator for linear mixed model. In Section 3, we apply Shi's method to derive the local influence measure of LIU type estimator in the variance disturbance model, independent variable perturbation model and dependent variable perturbation model, respectively. The methods are illustrated using examples in Section 4 and discussed in Section 5.

2. Linear mixed model and LIU type estimator

We consider the following linear mixed model

$$y = X\beta + Zu + e, \tag{2.1}$$

where β is a vector of fixed effects, u is a vector with mean $E(u) = 0$ and variance-covariance matrix $Var(u) = D$ which is a positive definite matrix, e is a vector with $E(e) = 0$ and positive definite variance-covariance matrix $Var(e) = R$, and also $Cov(u, e) = 0$. X is $n \times p$ design matrix with full column rank. y is a vector of observations with $E(y) = X\beta$ and $\Sigma = \mathrm{Cov}(y) = ZDZ' + R$.

Using the canonical equation $X'\Sigma^{-1}X\beta = X'\Sigma^{-1}y$, we can get the generalized least square (GLS) estimator

$$\hat{\beta}_{GLS} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y \tag{2.2}$$

for parameter β in model (2.1). In fact, D and R are unknown, we use the estimator $\tilde{D}$ and $\tilde{R}$ instead, that is $\tilde{\Sigma} = Z\tilde{D}Z' + \tilde{R}$, then we get the two-stage estimator $\tilde{\beta}_{GLS} = (X'\tilde{\Sigma}^{-1}X)^{-1}X'\tilde{\Sigma}^{-1}y$.

In linear regression model

$$y = X\beta + e,\ e \sim (0, \sigma^2 I),$$

the least square (LS) estimator $\hat{\beta} = (X'X)^{-1}X'y$ performs very poor when the design matrix X is close to singular. Liu (1993) proposed a new biased estimator named LIU estimator $\hat{\beta}(k) = (X'X + I)^{-1}(X'X + kI)(X'X)^{-1}X'y$.

Now we propose the LIU type estimator

$$\hat{\beta}_{LIU}(k) = (X'\Sigma^{-1}X + I)^{-1}(X'\Sigma^{-1}X + kI)(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y \tag{2.3}$$

for the linear mixed model, which is a biased estimator of β.

3. Local influence of the LIU type estimator

Shi (1997) studied the local influence in principal components analysis, the general influence function of a concerned quantity $T \in R^p$ is given by

$$GIF(T, l) = \lim_{a \to 0} \frac{T(\omega) - T(\omega_0)}{a}, \tag{3.1}$$

where $\omega = \omega_0 + al \in R^n$ represents a perturbation, ω_0 is a null perturbation which satisfies $T(\omega_0) = T$ and $l \in R^n$ denotes unit-length vector. The generalized Cook's distance in Shi (1997) is defined as

$$GC(T, l) = GIF(T, l)' M GIF(T, l)/c, \tag{3.2}$$

where M is positive definite matrix, c is a positive number.

By maximizing the absolute value of $GC(T, l)$ with respect to l, a direction $l_{\max}(T)$ is obtained. This direction shows how to perturb the data to obtain the greatest local change in T and maximum value $GC_{\max}(T, l) = GC(T, l_{\max})$ indicates the serous local influence.

In the linear mixed model, we are interested in estimating β, so the above T is the LIU type estimator $\hat{\beta}_{LIU}(k)$. In this paper we will get the LIU type estimator of the fixed effects β under three perturbation models and consider the local influence of the LIU type estimator using the Shi's method.

For the calculation of $\hat{\beta}_{LIU}(k)$, Σ should be replaced by its estimate $\tilde{\Sigma}$. Therefore the expression Σ in both the LIU type estimator and the generalized Cook's distance is conceived to be $\tilde{\Sigma}$ hereafter.

3.1. Perturbation of covariance matrix

We consider the perturbation with disturbing the ith row and ith column of the covariance matrix in model (2.1) as following

$$\Sigma^* = \Sigma + a(l{d_i}' + d_i l' - d_i {d_i}' D(l)), \tag{3.3}$$

where $i = 1, ..., n$, and d_i is an $n \times 1$ vector with an 1 in the ith position and zeros elsewhere, $D(l) = diag(l)$ is a diagonal matrix with diagonal elements of $l = (l_1, ..., l_n)'$. This perturbation scheme is a better way to handle cases badly modeled (Lawrance (1988)).

Theorem 3.1 For perturbation (3.3), the generalized Cook's distance of LIU estimator for β is as following:

$$\begin{aligned} GC(\hat{\beta}_{LIU}(k), l) = l'[&((X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1} d_i + \Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}) d_i' \\ &- D(\Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})) d_i d_i')\Sigma^{-1} X - ((y - X\hat{\beta}_{GLS})'\Sigma^{-1} d_i \\ &+ \Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}) d_i' - D(\Sigma^{-1}(y - X\hat{\beta}_{GLS})) d_i d_i')\Sigma^{-1} X (X'\Sigma^{-1} X)^{-1} \\ &(X'\Sigma^{-1} X + kI)](X'\Sigma^{-1} X + kI)^{-1}(X'\Sigma^{-1} X)(X'\Sigma^{-1} X + kI)^{-1} \\ &[X'\Sigma^{-1}(d_i'\Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}) + d_i(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1} \\ &- d_i d_i' D(\Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}))) - (X'\Sigma^{-1} X + kI)(X'\Sigma^{-1} X)^{-1} X'\Sigma^{-1} \\ &(d_i'\Sigma^{-1}(y - X\hat{\beta}_{GLS}) + d_i(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1} - d_i d_i' D(\Sigma^{-1}(y - X\hat{\beta}_{GLS})))]l/p. \end{aligned}$$

Proof: The LIU estimator under the perturbation (3.3) of β is given as

$$\hat{\beta^*}_{LIU}(k) = (X'\Sigma^{*-1}X + I)^{-1}(X'\Sigma^{*-1}X + kI)(X'\Sigma^{*-1}X)^{-1}X'\Sigma^{*-1}y.$$

By using matrix inversion formula and some technical matrix computing methods, we get

$$\begin{aligned}\hat{\beta^*}_{LIU}(k) = \hat{\beta}_{LIU}(k) &+ a[(X'\Sigma^{-1}X + I)^{-1}X'\Sigma^{-1}(d_i'\Sigma^{-1}(X\hat{\beta}_{LIU}(k) \\ &- X\hat{\beta}_{GLS}) + d_i(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1} - d_id_i'D(\Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}))) \\ &- (X'\Sigma^{-1}X + I)^{-1}(X'\Sigma^{-1}X + kI)(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}(d_i'\Sigma^{-1}(y - X\hat{\beta}_{GLS}) \\ &+ d_i(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1} - d_id_i'D(\Sigma^{-1}(y - X\hat{\beta}_{GLS})))]l + o(a^2),\end{aligned}$$

where $\hat{\beta}_{LIU}(k)$ is the LIU type estimator of the linear model (2.1) as given in (2.3). According to (3.1) with $T(\omega = \Sigma^*) = \hat{\beta^*}_{LIU}(k)$ and $T(\omega_0 = \Sigma) = \hat{\beta}_{LIU}(k)$, the generalized influence function is obtained as

$$\begin{aligned}GIF(\hat{\beta}_{LIU}(k), l) = &[(X'\Sigma^{-1}X + I)^{-1}X'\Sigma^{-1}(d_i'\Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}) \\ &+ d_i(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1} - d_id_i'D(\Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}))) \\ &- (X'\Sigma^{-1}X + I)^{-1}(X'\Sigma^{-1}X + kI)(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}(d_i'\Sigma^{-1}(y - X\hat{\beta}_{GLS}) \\ &+ d_i(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1} - d_id_i'D(\Sigma^{-1}(y - X\hat{\beta}_{GLS})))]l.\end{aligned}$$

The result can be easily obtained, if we take $c = p$ and $M = \mathrm{Cov}(\hat{\beta}_{LIU}(k)) = (X'\Sigma^{-1}X + I)(X'\Sigma^{-1}X + kI)^{-1}(X'\Sigma^{-1}X)^{-1}(X'\Sigma^{-1}X + kI)^{-1}(X'\Sigma^{-1}X + I)$ in (3.2). □

3.2. Perturbation of independent variables

It is known that the minor perturbation of the explanatory variables can seriously influence the least square regression results when collinearity is present (Cook (1986)). We only consider the individual perturbation of the jth column of X as did in Shi (1999),

$$X^* = X + as_jlb_j{}', \tag{3.4}$$

where $j = 1, ..., p$, and b_j is a $p \times 1$ vector with an 1 in the jth position and zeros elsewhere, s_j denotes the scale factor of the jth regression variable (that is the mode of the jth regression variable,$s_j = \|X_j\|$).

Theorem 3.2 For perturbation (3.4), the generalized Cook's distance of LIU estimator is:

$$\begin{aligned}GC(\hat{\beta}_{LIU}(k), l) = &s_j^2l'[(\Sigma^{-1}X(\hat{\beta}_{GLS} - \hat{\beta}_{LIU}(k))'b_j + b_j'(\hat{\beta}_{GLS} - \hat{\beta}_{LIU}(k))\Sigma^{-1}X) \\ &+ (\Sigma^{-1}(y - X\hat{\beta}_{GLS})b_j' - \Sigma^{-1}X\hat{\beta}_{GLS}'b_j)(X'\Sigma^{-1}X)^{-1}(X'\Sigma^{-1}X + kI)] \\ &(X'\Sigma^{-1}X + kI)^{-1}(X'\Sigma^{-1}X)(X'\Sigma^{-1}X + kI)^{-1}[(b_j'(\hat{\beta}_{GLS} - \hat{\beta}_{LIU}(k))X'\Sigma^{-1} \\ &+ X'\Sigma^{-1}(\hat{\beta}_{GLS} - \hat{\beta}_{LIU}(k))'b_j) + (X'\Sigma^{-1}X + kI)(X'\Sigma^{-1}X)^{-1}(b_j(y - X\hat{\beta}_{GLS})'\Sigma^{-1} \\ &- b_j'\hat{\beta}_{GLS}X'\Sigma^{-1})]l/p.\end{aligned}$$

Proof: The LIU type estimator for β under perturbation (3.4) is given as

$$\hat{\beta^*}_{LIU}(k) = (X^{*\prime}\Sigma^{-1}X^* + I)^{-1}(X^{*\prime}\Sigma^{-1}X^* + kI)(X^{*\prime}\Sigma^{-1}X^*)^{-1}X^{*\prime}\Sigma^{-1}y.$$

By using matrix inversion formula and some technical matrix computing methods, we get

$$\begin{aligned}
&\hat{\beta^*}_{LIU}(k)\\
&\quad= \hat{\beta}_{LIU}(k) + as_j[(X'\Sigma^{-1}X+I)^{-1}(b_j'(\hat{\beta}_{GLS}-\hat{\beta}_{LIU}(k))X'\Sigma^{-1}+X'\Sigma^{-1}(\hat{\beta}_{GLS}-\hat{\beta}_{LIU}(k))'b_j)\\
&\quad+(X'\Sigma^{-1}X+I)^{-1}(X'\Sigma^{-1}X+kI)(X'\Sigma^{-1}X)^{-1}(b_j(y-X\hat{\beta}_{GLS})'\Sigma^{-1}-b_j'\hat{\beta}_{GLS}X'\Sigma^{-1})]l\\
&\quad+o(a^2).
\end{aligned}$$

Similarly, according to (3.1), the generalized influence function is obtained as

$$\begin{aligned}
&GIF(\hat{\beta}_{LIU}(k), l)\\
&\quad= s_j[(X'\Sigma^{-1}X + I)^{-1}(b_j'(\hat{\beta}_{GLS} - \hat{\beta}_{LIU}(k))X'\Sigma^{-1} + X'\Sigma^{-1}(\hat{\beta}_{GLS} - \hat{\beta}_{LIU}(k))'b_j)\\
&\quad+(X'\Sigma^{-1}X+I)^{-1}(X'\Sigma^{-1}X+kI)(X'\Sigma^{-1}X)^{-1}(b_j(y-X\hat{\beta}_{GLS})'\Sigma^{-1}-b_j'\hat{\beta}_{GLS}X'\Sigma^{-1})]l.
\end{aligned}$$

The result can be easily proved, if we take $c = p$ and $M = \mathrm{Cov}(\hat{\beta}_{LIU}(k)) = (X'\Sigma^{-1}X + I)(X'\Sigma^{-1}X + kI)^{-1}(X'\Sigma^{-1}X)^{-1}(X'\Sigma^{-1}X + kI)^{-1}(X'\Sigma^{-1}X + I)$ in (3.2). □

3.3. Perturbation of dependent variables

The perturbation of dependent variables can also influence the estimation results. Consider the dependent variable perturbation as the following

$$y^* = y + al. \tag{3.5}$$

Theorem 3.3 For the perturbation (3.5), the generalized Cook's distance of LIU estimator is

$$GC(\hat{\beta}_{LIU}(k), l) = l'\Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}l/p.$$

Proof:

The LIU type estimator for β under (3.5) is

$$\hat{\beta^*}_{LIU}(k) = (X'\Sigma^{-1}X + I)^{-1}(X'\Sigma^{-1}X + kI)(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y^*.$$

Expanding the right of the formula by using some technical computing methods, we have

$$\hat{\beta^*}_{LIU}(k) = \hat{\beta}_{LIU}(k) + a(X'\Sigma^{-1}X + I)^{-1}(X'\Sigma^{-1}X + kI)(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}l.$$

So the generalized influence function is

$$GIF(\hat{\beta}_{LIU}(k), l) = (X'\Sigma^{-1}X + I)^{-1}(X'\Sigma^{-1}X + kI)(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}l.$$

The result can be easily obtained, if we take $c = p$ and $M = \mathrm{Cov}(\hat{\beta}_{LIU}(k)) = (X'\Sigma^{-1}X + I)(X'\Sigma^{-1}X + kI)^{-1}(X'\Sigma^{-1}X)^{-1}(X'\Sigma^{-1}X + kI)^{-1}(X'\Sigma^{-1}X + I)$ in (3.2). □

In order to obtain the l_{max}, we proceeded as following:

For all three perturbation above, we can express $GC(\hat{\beta}_{LIU}(k), l)$ as $GC(\hat{\beta}_{LIU}(k), l) = l'Fl$, here for example we take

$$\begin{aligned}
F = {} & [((X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1}d_i + \Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})d_i' \\
& - D(\Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}))d_id_i')\Sigma^{-1}X - ((y - X\hat{\beta}_{GLS})'\Sigma^{-1}d_i \\
& + \Sigma^{-1}(\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})d_i' - D(\Sigma^{-1}(y - X\hat{\beta}_{GLS}))d_id_i')\Sigma^{-1}X(X'\Sigma^{-1}X)^{-1} \\
& (X'\Sigma^{-1}X + kI)](X'\Sigma^{-1}X + kI)^{-1}(X'\Sigma^{-1}X)(X'\Sigma^{-1}X + kI)^{-1} \\
& [X'\Sigma^{-1}(d_i'\Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}) + d_i(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1} \\
& - d_id_i'D(\Sigma^{-1}(X\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS}))) - (X'\Sigma^{-1}X + kI)(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1} \\
&](d_i'\Sigma^{-1}(y - X\hat{\beta}_{GLS}) + d_i(\hat{\beta}_{LIU}(k) - X\hat{\beta}_{GLS})'\Sigma^{-1} - d_id_i'D(\Sigma^{-1}(y - X\hat{\beta}_{GLS})))]/p,
\end{aligned}$$

when the covariance is perturbed as in section 3.1. By maximizing the absolute value of $GC(\hat{\beta}_{LIU}(k), l)$ with respect to l, we get a direction l_{max} which corresponds to the eigenvector associated with the largest absolute eigenvalue of F. We can obtain l_{max} similarly for the other two perturbation models, that is, F is the matrix in the middle of $GC(\hat{\beta}_{LIU}(k), l))$ for each model.

4. Illustrative example

Harrison and Rubinfeld (1978) reported the housing prices in the Boston Standard Metropolitan Statistical Area. We choose a part of the data to analyze as in Shi (1996) (see Table 4.1). The sample was taken from the five districts (Auston-Brighton, Bock Bay, Charistrun, East Boston and Hyde Park) of Boston. We use the linear mixed model which includes 8 fixed effects and one random effect to fit the housing data. The fixed effect variables are $X1 - X8$: CRIM (Crime rate by town), CHAS (Charles River dummy; =1 if tract bounds the Charles River; =0 if otherwise), NOX (Nitrogen oxide concentrations in pphm), RM (Average number of rooms in owner units), AGE (Proportion of owner units built prior to 1940), DIS (Weighted distances to five employment centers in the Boston region), B (Black proportion of population), and LSTAT (Proportion of population that is lower status=1/2), respectively. The dependent variable Y in the model is the median value of the owner-occupied homes in the census tract. The fitting model is $Y_i = X_i\beta + 1_{t_i}u_i + e_i$, where Y_i and e_i are $t_i \times 1$ random vectors, X_i is a $t_i \times 8$ design matrix, t_i denotes the number of observations in the ith district, u_i is a random effect variable which follows $N(0, \sigma_i^2)$, $i = 1, ..., 5$. After introducing some notations, the model can be written as model (2.1). We use restricted maximum likelihood estimator to estimate the variance R and D, and then we get the two-stage estimator of parameter β. There are many methods for selecting the value of k. We set the parameter $k = 0.0002$ in the LIU estimator following a prediction criterion proposed by Myers (1986, p.249). The GLS estimates and LIU type estimates of the model are obtained and shown in Table 4.2.

Table 4.1 Boston housing data.

No.	LMV Y	CRIM X1	CHAS X2	NOX X3	RM X4	AGE X5	DIS X6	B X7	LSTAT X8
1	9.78695	8.98396	1	59.2899	38.5889	97.4	0.75245	0.37779	-1.7375
2	9.98507	3.8497	1	59.2899	40.896	91	0.91837	0.39134	-2.01956
3	10.0301	5.20177	0	59.3899	37.5401	83.4	1.00162	0.39543	-2.16439
4	10.0257	4.26131	0	59.3899	37.3565	81.3	0.91992	0.39037	-3.06577
5	10.1266	4.54192	0	59.2899	40.9344	88	0.92354	0.37456	-3.55194
6	9.89848	3.83584	0	59.2899	39.075	91.1	0.83095	0.35065	-1.95298
7	9.94271	3.67822	0	59.2899	28.751	96.2	0.74365	0.38097	-2.28416
8	9.73913	4.32339	1	59.2899	33.6748	89.0	0.64432	0.35304	-1.92114
9	9.99124	3.47429	1	51.5523	37.0884	82.9	0.64432	0.35455	-2.93973
10	10.2219	4.55587	0	51.5533	12.6807	87.9	0.47832	0.3547	-3.6434
11	9.99424	3.69695	0	51.5533	24.6314	91.4	0.56093	0.31604	-1.96597
12	10.0476	13.5222	0	39.816	14.9328	100	0.41251	0.13142	-2.01545
13	10.8198	4.89822	0	39.816	34.7099	100	0.28706	0.37552	-3.42353
14	10.8198	5.66998	1	39.816	44.66 25	96.8	0.30505	0.37533	-3.2893
15	9.61581	19.6091	0	45.024	53.4799	97.9	0.37482	0.3969	-2.00723
16	9.53864	15.288	0	45.024	44.2092	93.3	0.39632	0.36302	-1.45925
17	9.49552	9.82349	0	45.024	46.1584	98.8	0.30601	0.3969	-1.54934
18	9.48037	23.6482	0	45.024	40.7044	96.2	0.32649	0.3969	-1.44016
19	9.23011	17.8667	0	45.024	38.7257	100	0.32649	0.39374	-1.75974
20	9.24956	88.9762	0	45.024	48.553	91.9	0.34819	0.3969	-1.75974
21	9.29653	15.8744	0	45.024	42.837	99.1	0.41818	0.3969	-1.55689
22	9.33256	9.18702	0	48.9399	30.6473	100	0.4568	0.3969	-1.44401
23	9.41735	7.99248	0	48.9999	30.4704	100	0.42729	0.3969	-1.40393
24	9.08251	20.0843	0	48.9999	19.0794	91.2	0.36492	0.28583	-1.18312
25	8.88184	16.8118	0	48.9999	27.8467	98.1	0.35494	0.3969	-1.17723
26	9.25913	24.3938	0	48.9999	21.6411	100	0.38336	0.3969	-1.26291
27	8.90924	22.5971	0	48.9999	25	89.5	0.41766	0.3969	-1.13968
28	0.23014	14.3337	0	48.9999	23.8144	100	0.46342	0.37292	-1.18345
29	9.3501	8.15174	0	38.9999	29.0521	98.9	0.54702	0.39443	-1.76562
30	9.62245	6.96215	0	48.9999	32.6384	97	0.6557	0.39443	-1.76562
31	10.0519	5.29305	0	48.9999	36.6146	82.5	0.77371	0.37838	-1.67323
32	9.17988	11.5779	0	48.9999	25.3613	97	0.57098	0.3969	-1.35934
33	9.98957	2.81838	0	28.3024	33.2006	40.3	1.41057	0.39292	-3.26133
34	9.93305	2.37857	0	33.9889	34.4686	41.9	1.3148	0.37073	-3.01463
35	9.96176	3.67367	0	33.9889	39.8413	51.9	1.38432	0.38862	-2.2463
36	9.85741	5.69175	0	33.9889	37.381	79.8	1.26579	0.39268	-1.89872
37	9.93305	4.83567	0	33.9889	34.869	53.2	1.14813	0.38822	-2.16727

We plot the l_{max} against the index or individual. The plot shows that the jth individual will be more affected after perturbation if the jth element of l_{max} is far away from 0, $j = 1, \cdots, 37$. In that case the jth individual is more likely to be detected as an influential observation. That influential observation produces the greatest local change in the fixed parameter. In Figure 4.1, we detect the observation making the largest influence when perturbing each jth column and jth row of the covariance. Here we just show only three figures among 37 results for illustration. The dashed lines denote the value of $\pm \sum_{j=1}^{37} |l_{\max}(j)|/37$, which may serve as possible threshold points to decide the influential observation, and then we choose the largest and the second largest values which are larger or smaller than the two lines, respectively. In Shi *et al.* (1996), using the data deletion method, they detected points 9, 10, 12, 20, 24, 31, 36 as the strong influential data. Using the method proposed in our paper, when we perturb the covariance matrix we can detect the observations 20 (when perturbing 18, 22-24, 28-29, 31-32, 36-37 column and row), 28 (when perturbing 1-17, 19-21, 25-27, 30, 33-35 column and row) are the strongest influential points, and the point 28 can not be detected when using the data deletion method, but we can see the point 28 is a singular point which has very small y (LMV) value compared to the rest of the data as shown in table 4.1. Then individuals 24 (when perturbing 1-4, 6-14, 16-17, 19-21 column and row), $v12$ (when perturbing 15 column and row), 9 (when perturbing 34-36 column and row), 14 (when perturbing 28, 32-33 column and row), and 31 (when perturbing 20-23 column and row) are detected to be the second strongest influential points. We did not detect the individuals 10 and 36 as influential data whereas did in Shi *et al.* (1996). As shown in Figure 4.1, the individuals 10 and 36 did not have large values for each plot.

Table 4.2 The GLS and LIU type estimates.

Estimate	β_1	β_2	β_3	β_4
${\beta_L}^*$	-0.0005601803	-0.2252245002	0.0549869203	0.0563565346
${\beta_k}^*$	-0.0008131812	-0.2644397566	0.0543315471	0.0547935343
Estimate	β_5	β_6	β_7	β_8
${\beta_L}^*$	0.0304399772	1.1547986265	-1.8377645709	-1.8377645709
${\beta_k}^*$	0.0281791580	0.8867741245	-0.3680949177	-0.6522629361

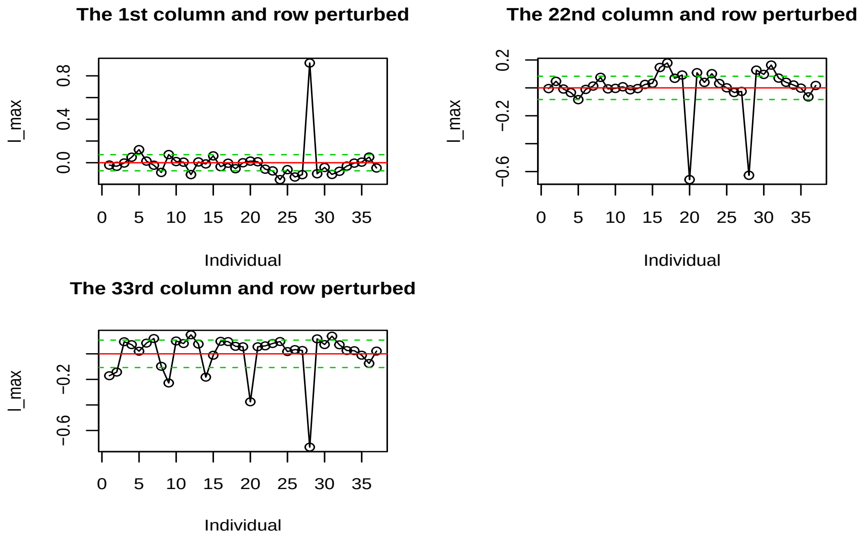

Figure 4.1 Index plot of l_{max} for independent variables disturbance

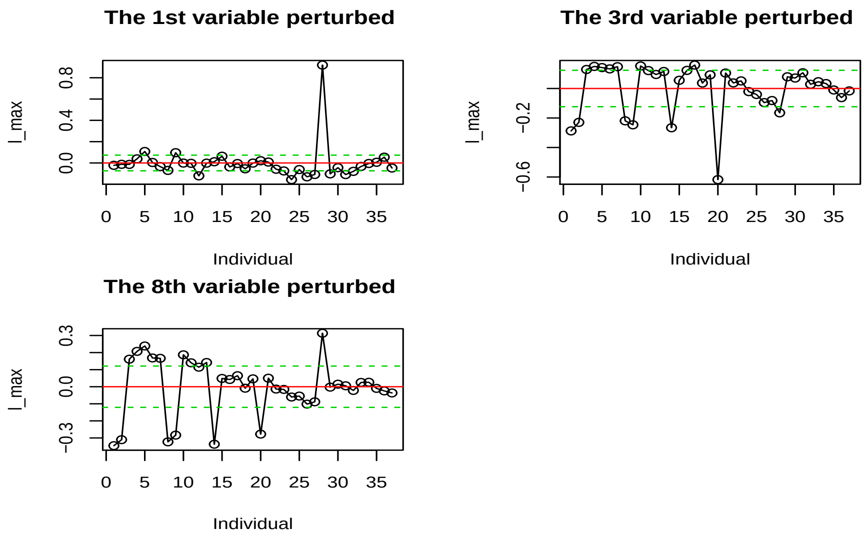

Figure 4.2 Index plot of l_{max} for independent variables disturbance

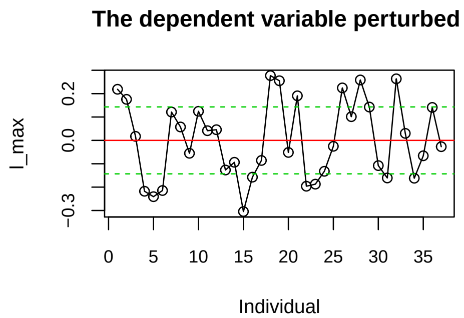

Figure 4.3 Index plot of l_{max} for dependent variables disturbance

Figure 4.2 shows the largest influential direction when perturbing the 1st, the 3rd and the 8th column of the independent variables, for illustration. We can see, the individuals 1 (when we disturbed X_8), 20 (when we disturbed X_3-X_6) and 28 (when disturbing X_1 and X_2) are detected as the strongest influential points. The observations 14 (when we disturbed X_4 and X_8)and 24 (when disturbing X_1,X_2 and X_7) are detected as the second strongest influential points. Figure 4.3 shows that the largest influential direction when perturbing the dependent variable. The observations 15 and 18 are detected as the largest and the second largest influential points.

5. Conclusions

In this paper we consider three different perturbations in linear mixed models, and propose the LIU estimator for the fixed effects parameter in each model. And then we study the local influence. Cook (1986) proposed the method of local influence, and Beckman *et al.* (1987) detected influential observations in a linear mixed model using this method. In our paper, we use the Shi's (1997) method which are not based on model likelihood replacement to study the local influence. We validate the results by numerical examples and it is easy to find the strong influential point of the model. The future study will be how to determine the local influence of the random effect parameter.

References

Beckman, R. J., Nachtsheim, C. J. and Cook, R. D. (1987). Diagnostics for mixed-Model of variance. *Journal of Technometrics*, **29**, 413-426.

Belsley, D. A., Kuh, E. and Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of colinearity*, John Wiley, New York.

Chatterjee, S. and Hadi, A. S. (1988). *Sensitivity analysis in linear regression*, John Wiley, New York.

Cook, R. D. (1986). Access of local influence. *Journal of the royal statistical society*, **48**, 133-169.

Cook, R. D. and Weisberg, S. (1982). *Residuals and influence in regression*, John Wiley, New York.

Lawrance, A. J. (1988). Regression transformation diagnostics using local influence. *J. Am. Statist. Assoc.*, **83**, 1067-1072.

Li, Y. L. and Yang, H. (2012). A new Liu-type estimator in linear regression model. *Springer-Verlag*, **53**, 427-437.

Liu, K. J. (1993). A new class of biased estimate in linear regression. *Journal of Commun. Statist. Theo. and Meth.*, **22**, 393-402.

Liu, K. J. (2004). More on Liu-type in estimator in linear regression. *Communications in Statistics-Theory and Methods*, **33**, 223-2733.
Liu, X. Q. (2011). Improved Liu estimator in a linear regression model. *Journal of Statistical Planning and Inference*, **141**, 189-196.
Myers, R. H. (1986). *Classical and modern regression with applications*, Duxbury Press, Boston.
Ronald, C. and Larry, M. P. (1992). Case deletion diagnostics for mixed models. *Technometrics*, **34**, 38-45.
Rousseeuw, P. J. and Leroy, A. M. (1987). *Robust regression and outlier detection*, John Wiley and Sons, New York.
Shi, L. (1997). Local influence in principal components analysis. *Biometrika*, **84**, 175-186.
Shi, L. (1999). Local influence in ridge regression. *Computational Statistics and Data Analysis*, **31**, 341-353.
Shi, L. and Chen, G. M. (2008). Case deletion diagnostics in multilevel models. *Journal of multivariate Analysis*, **99**, 1860-1877.
Shi, L., Xiang, L. M. and Wang, X. R. (1996). Local influence of effect parameter in linear mixed model. *Acta Mathematica Scientia*, **16**, 316-323.
Thomas, W. and Cook, R. D. (1990). Assessing influence on predictions from generalized linear models. *Technometrics*, **32**, 59-65.
Zhang, L. L. and Zhang, S. L. (2010). Local influence analysis of Liu estimate in linear model. *Journal of information and computational science 7*, **13**, 2731-2737.