مجله‌ی پژوهش‌های آماری ایران
پژوهشکده‌ی آمار
دوره‌ی ۱۲، شماره‌ی ۱، بهار و تابستان ۱۳۹۴، صص ۳۹–۵۶
J. Statist. Res. Iran **12** (2015): 39–56

Influence Measures in Ridge Linear Measurement Error Models

Hadi Emami

University of Zanjan

Abstract. Usually the existence of influential observations is complicated by the presence of collinearity in linear measurement error models. However no method of influence measure available for the possible effect's that collinearity can have on the influence of an observation in such models. In this paper, a new type of ridge estimator based corrected likelihood function (REC) for linear measurement error models is defined. We show when this type of ridge estimator is used to mitigate the effects of collinearity the influence of some observations can be drastically modified. We propose a case deletion formula to detect influential points in REC. As an illustrative example two real data set are analysed.

Keywords. Corrected likelihood; diagnostics; leverage; measurement error models, shrinkage estimators.

MSC 2010: 62J05, 62J07, 62J20.

1 Introduction

In statistics, linear measurement error models are linear regression models that account for measurement errors in the independent variables. More statistical problems involve measurement error. In fact, measurement error occurs whenever we cannot exactly observe one or more of the variables that enter into a model of interest. There are many reasons such errors occur, the most common ones being instrument error and sampling error. Measurement error occurs in nearly every discipline such as epidemiology study, quality control, agricultural sciences, computer science, etc.

Diagnostic techniques for the regression model have received great attention in statistical literature. In linear measurement error models, only few works have been done that they are originally based on an unbiased function approach (see Stefanski and Caroll, 1989). Kelly (1984) derived the influence function under the assumption that the covariance matrix of the errors is known up to a multiple and with the aid of this influence, diagnostics statistic is defined by analogy with the corresponding ordinary linear model equivalents. Fuller (1987) defined the hat matrix using the estimated predictor variable values, Wellman and Gunst (1991) used the influence functions to show that extreme observations affect measurement error model estimates in directions orthogonal to and along the fitted plane, rather than vertically and horizontally as with least-squares estimates. Zhao and Lee (1994) derived influence function for generalized linear and non linear measurement error models. Rasekh and Feller (2003) gave an influence function in functional measurement error models with replicated observations. Gimenez and Galea (2013) and Castro et al. (2007) investigated influence and local influence measures in heteroscedastic and functional heteroscedastic measurement error models respectively.
In different way Zhong et al. (2000) also presented a unified diagnostic method in linear measurement error model based upon the corrected likelihood (CL) approach which is introduced by Nakamura (1990).
It is not unusual to have collinearity and influential cases simultaneously in a dataset. However, in ridge regression (RR), the RR estimators are sensitive to outlier thus, one or a subset of observations may have undue effects on them. Therefore mixed estimators and ridge type estimators are suggested to mitigate the effect of multicollinearity. However, as some authors have noted, that the influence of the observations on ridge type estimators are different from that of the corresponding least squares estimator and even multicollinearity can disguise anomalous data. In recent literature, Walker and Brich (1989) studied the influence of observations in ridge regression using global influence method. Shi (1997) proposed the local influence in principal component analysis by defining a generalized Cook statistic and showed that his method is equivalent to Cook's approach under the likelihood framework. Janufer and Jianbao (2012) studied global influential observations in modified ridge regression estimator. Janufer and Jianbao (2009) also studied identification of local influential observations in Liu type regression estimator. Influence analysis in measurement error models with ridge estimate rarely discussed. Rasekh and Feller (2003) derived influence

function of ridge estimate in measurement error models using case deletion. Rasekh (2006) studied the local influence of minor perturbations on the ridge estimate in the ordinary regression model. He derived the diagnostics under the perturbation of variance and explanatory variables. In this paper a ridge type estimator based upon the corrected likelihood function for linear measurement error model (REC) is derived. We generalize the Walker and Brich (1989) results to the linear measurement error models to assess the global influence of observations on the REC. We show that when REC is used to mitigate the effects of collinearity the influence of some observations can be drastically modified. The rest of paper is organized as follows. Section 2 introduces the linear measurement error models, the relevant notation, influence measures and some inferential results are also given. Section 3 ridge estimator based on corrected likelihood function (namely REC) is defined. The influence measures in REC, and case-deletion formulas for REC is also derived. Statistical properties and motivation of these measures are discussed. In Section 4 the proposed methods are applied through a real data set. Finally discussion is given in the Section 5.

2 Back Ground of Model

2.1 Linear Measurement Error Models

The linear measurement error models is defined by

$$y = \mathbf{Z}\beta + \epsilon, \qquad \epsilon \sim N(0, \sigma^2 \mathbf{I}_n),$$

$$\mathbf{X} = \mathbf{Z} + \delta, \qquad \delta \sim N(0, \mathbf{I}_n \otimes \Lambda), \tag{1}$$

where y is the $n \times 1$ vector of observations y_i, $\mathbf{Z}$ is $n \times p$ matrix whose ith row is z_i', β is vector of parameters, $\mathbf{I}_n$ is $n \times n$ identity matrix, and σ^2 is the unknown common variance. $\mathbf{X}$ is $n \times p$ matrix with x_i as its row, ϵ and δ are independent, and Λ is positive definite matrix. This is underlying model for response vector y in term of covariates. Covariate $\mathbf{Z}$ is unobservable for all study subjects which can be observed from random matrix $\mathbf{X}$.
For model (1) the log likelihood function $l(\beta, \mathbf{Z}, y)$ is

$$l(\beta, \mathbf{Z}, y) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - z_i'\beta)^2. \tag{2}$$

The following equation is useful to find the corrected log-likelihood function

$$E^*(\mathbf{X}'\mathbf{X}) = \mathbf{Z}'\mathbf{Z} + n\Lambda,$$

where E^* denotes the conditional mean respect to $\mathbf{X}$ given $\mathbf{Z}$ and y. From this the corrected likelihood can be obtained as

$$l^*(\beta, \mathbf{X}, y) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\{(y_i - x_i'\beta)^2 - \beta'\Lambda\beta\}, \tag{3}$$

for more details see Zhong et al. (2000). By solving $\frac{\partial l^*}{\partial \beta} = 0$, the corrected likelihood estimator (CLE) of β obtained as:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X} - n\Lambda)^{-1}\mathbf{X}'y. \tag{4}$$

Using estimator (4), Zhong et al. (2000) defined the vector of fitted values as

$$\hat{y} = \mathbf{X}\hat{\beta} = \mathbf{H}y$$

where $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X} - n\Lambda)^{-1}\mathbf{X}'$ with entries h_{ij}.
Therefore, the hat matrix $\mathbf{H}$ plays the same roles as the hat matrix in LS. The ith hat diagonal h_i can be interpreted as leverage in the same sense as the hat diagonals in LS. In continue the vector of residual is $e = (\mathbf{I} - \mathbf{H})y$ and the estimate of σ^2 will be $\hat{\sigma}^2 = \frac{e'e}{n-p}$.

2.2 Influential Measures in Linear Measurement Error

The general purpose of influence analysis is to measure the changes induced in a given aspect of the analysis when the data are perturbed. particularly, appealing perturbation scheme is case deletion. This scheme will be used throughout this article. In general, the influence of a case can be viewed as the product of two factors: the first a function of the residual and the second a function of the position of the point in the $\mathbf{X}$ space. The position or leverage of the ith point is measured by h_i. Among the most popular single-case influential measure is the difference in fit standardized DFFITS Belsley et al. (1989), which evaluated at the ith case, for linear measurement error can be given by

$$\begin{aligned} DFFITS(i) &= \frac{x_i(\hat{\beta}_{(i)} - \hat{\beta})}{\hat{\sigma}_{\hat{y}}} \\ &= \frac{h_{ii}e_i}{p\hat{\sigma}_{\hat{y}}(1 - h_{ii})}, \end{aligned} \tag{5}$$

where $\hat{\beta}_{(i)}$ is the CLE of β with out the ith case and $\hat{\sigma}_{\hat{y}}$ is an estimator of the standard error of the fitted values. $DFFITS(i)$ is the standardized change in the fitted value of a case when it is deleted. Thus it can be considered a measure of influence on individual fitted values. Another useful measure of influence is Cook's distance (Cook, 1977), which generalized in linear measurement error model as

$$\begin{aligned} D_i &= (\hat{\beta}_{(i)} - \hat{\beta})'\mathbf{M}(\hat{\beta}_{(i)} - \hat{\beta}) \\ &= \frac{h_{ii}e_i^2}{p\sigma^2(1-h_{ii})^2}. \end{aligned} \tag{6}$$

Where $\mathbf{M} = \sigma^{-2}(\mathbf{X}'\mathbf{X} - n\Lambda)$ is the corrected fisher information matrix of y for β. D_i is a measure of the change in all of the fitted values when a case is deleted. Even though D_i is based on different theoretical consideration, it is closely related to DFFITS. Observations with large values of D_i have considerable influence on the CLE of β. In general, observations for which $D_i \geqslant F_{(\alpha,p,n-p)}$ to be influential (see Zhao et al., 2000). In DFFITS measures any observation for which $|DFFITS_i| \geqslant 2\sqrt{\frac{p}{n}}$ warrants attention. It is important to mention that these measures are useful for detecting single cases with an unduly high influence.

3 Influence in Ridge Estimator

3.1 Ridge Estimator Based on Corrected Likelihood Function (REC)

When the ordinary least squares method of estimation is applied for multicollinearity data, it produces large variances for the estimated regression coefficients. To overcome this problem the restricted least squares (RLS), mixed estimators and the ridge type estimators methods are used in the literature. In ordinary and generalized linear model the idea of ridge estimation is to fix the length of the regression coefficient vector β to say H, and to minimize the sum of square (maximize log-likelihood function) subject to this constraint. We examine this argument in order to derive a ridge estimate of parameter β in model (1).
According to Zhong et al. (2000) the CLE of β can be attained using (3). To derive a ridge estimate of parameter β, instead of maximizing the restricted

log-likelihood

$$l_k(\beta, \mathbf{Z}, y) = l(\beta, \mathbf{Z}, y) - \frac{k}{2}(\beta'\beta - H),$$

we can maximize the restricted corrected log-likelihood

$$l_k^*(\beta, \mathbf{X}, y) = l^*(\beta, \mathbf{X}, y) - \frac{k}{2}(\beta'\beta - H),$$

in which $\frac{k}{2}$ is the Lagrange multiple. Since $E^*\{l_k^*(\beta, \mathbf{X}, y)\} = l_k(\beta, \mathbf{Z}, y)$ we can say that the restricted corrected log-likelihoods is corrected the restricted log-likelihood.
Solving $\frac{\partial l_k^*(\beta, \mathbf{X}, y)}{\partial \beta} = 0$ the REC will be

$$\hat{\beta}_{(k)} = (\mathbf{X}'\mathbf{X} - n\Lambda + kI_p)^{-1}\mathbf{X}'y, \tag{7}$$

where k refereed to as shrinkage parameter and I_p is an identity matrix. It is obvious that for k equal to zero, the REC defined in relation (7) is exactly the same as CLE of β. The REC defined in relation (7) can also be written as

$$\hat{\beta}_{(k)} = R_{(k)}\hat{\beta},$$

in which $R(k)$ is the ridge factor, namely $R_{(k)} = [\mathbf{I}_p + k(\mathbf{X}'\mathbf{X} - n\Lambda)^{-1}]^{-1}$. The key question in the REC is to choose the value of the k. The choice of this parameter in ordinary least square is still unsolved and because of this problem several approaches have been developed the guide to data analyst in the selection of shrinkage parameter. However, there are some popular methods for selecting the value of k in RR, for example Hoerl and Kennard's iterative procedure Horel and Kennard (1976), Mc Donald Galarneau's method (Mc Donald and Galarneau, 1975), C_p statistic criterion (Mallows, 1973) and so on. Some of these methods do not have a strong theoretical justification and are subjective in nature. We suggest the C_p criteria technique to selecting the ridge parameter in linear measurement error models. It consists of selecting the value of k that minimizes

$$C_k = \frac{SSR_k}{\hat{\sigma}^2} - n + 2tr(H^*).$$

In which SSR_k is the sum of squares of residuals from linear measurement error, $H^* = \mathbf{X}(\mathbf{X}'\mathbf{X} - n\Lambda + kI_p)^{-1}\mathbf{X}'$ and tr stands for the trace. (Note that when $k = 0$ this expression reduces to Mallow's C_p statistic.

3.2 Leverages and Residuals in REC

Using the estimator in (7), The vector of fitted values is

$$\hat{y}^* = \mathbf{X}\hat{\beta}_{(k)} = \mathbf{X}(\mathbf{X}'\mathbf{X} - n\Lambda + kI_p)^{-1}\mathbf{X}'y.$$

Therefore, the matrix $H^* = \mathbf{X}(\mathbf{X}'\mathbf{X} - n\Lambda + kI_p)^{-1}\mathbf{X}'$ plays the same role as the hat matrix in linear measurement error model. The ith fitted value can be written in terms of elements of H^* as

$$\hat{y}_i^* = \sum_{j=1}^{n} h_{ij}^* y_j.$$

consequently $\frac{\partial \hat{y}_i^*}{\partial y_i} = h_{ii}^*$. The ridge hat diagonals h_{ii}^* can be interpreted as leverage in the same sense as the hat diagonals in linear measurement error. It is important to note, however, that the matrix H^* is not a projection matrix because it is not idempotent.

***Lemma* 1.** *Let h_{ii} be the ith diagonal element of linear measurement error leverage matrix H then*

$$\frac{\gamma_1}{\gamma_1 + k} h_{ii} \leqslant h_{ii}^* \leqslant \frac{\gamma_p}{\gamma_p + k} h_{ii},$$

where γ_1 and γ_p are the minimum and maximum eigenvalue of matrix $(\mathbf{X}'\mathbf{X} - n\Lambda)$ respectively.

Proof. Let $\mathbf{B} = (\mathbf{X}'\mathbf{X} - n\Lambda)^{-1}$ and $\mathbf{A} = (\mathbf{X}'\mathbf{X} - n\Lambda + k\mathbf{I}_p)^{-1}$ for the formula in Appendix 1 and proof will be done. □

From this lemma first we can see for $k > 0$, $h_{ii}^* \leqslant h_{ii}$ for $i = 1, \ldots, n$; that is, for every observation the REC leverage is smaller than the corresponding linear measurement error leverage. Second the leverage decreases monotonically as k increases. Although the leverage of every point decreases monotonically as k increases, the effect of this increment on the residuals is far less clear.
The ith REC residual is given by:

$$e_i^* = y_i - \hat{y}_i^* = y_i - x_i\hat{\beta}_{(k)}.$$

3.3 Measuring Influence in REC

A version of DFFITS for REC is

$$DFFITS_i^* = \frac{x_i\{\hat{\beta}_{(k)} - \hat{\beta}_{(k)}(i)\}}{SE(x_i\hat{\beta}_{(k)})},$$

where $\hat{\beta}_{(k)}(i)$ is the estimator (7) computed with the ith case deletion and the denominator is an estimator of the standard error of the REC fitted values. If k is assumed non stochastic then

$$\begin{aligned} SE(x_i\hat{\beta}_{(k)}) &= \hat{\sigma}[x_i(\mathbf{X}'\mathbf{X} - n\Lambda + kI_p)^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X} - n\Lambda + kI_p)^{-1}x_i']^{\frac{1}{2}} \\ &= \hat{\sigma}[\sum_{j=1}^{n} h_{ij}^{*2}]. \end{aligned}$$

Therefore DFFITS can be written as

$$DFFITS_i^* = \frac{x_i(\hat{\beta}_{(k)} - \hat{\beta}_{(k)}(i))}{\hat{\sigma}\left[\sum_{j=1}^{n} h_{ij}^{*2}\right]}. \tag{8}$$

Hence, the mean squared error is a function of the fitted values and the response, neither of which depends on individual eigenvalues of $\mathbf{X}'\mathbf{X}$, it is not affected by multicollinearity. For this reason, the least squares estimators of $[\hat{\sigma}$ and $\hat{\sigma}(i)]$ will be used as measures of scale.
At least two versions of D_i can be constructed for REC, namely

$$D_i^* = \frac{1}{p\hat{\sigma}^2}\{\hat{\beta}_{(k)} - \hat{\beta}_{(k)}(i)\}'\mathbf{M}\{\hat{\beta}_{(k)} - \hat{\beta}_{(k)}(i)\}, \tag{9}$$

and

$$D_i^{**} = \frac{1}{p\hat{\sigma}^2}\{\hat{\beta}_{(k)} - \hat{\beta}_{(k)}(i)\}'\mathbf{A}^{-1}(\mathbf{X}'\mathbf{X} - n\Lambda)^{-1}\mathbf{A}^{-1}\{\hat{\beta}_{(k)} - \hat{\beta}_{(k)}(i)\}, \tag{10}$$

where D_i^* is the direct generalization of Cook's D (6) and D_i^{**} based on the fact that $var(\hat{\beta}_{(k)}) = \mathbf{A}(\mathbf{X}'\mathbf{X} - n\Lambda)\mathbf{A}\sigma^2$. Note that both D_i^* and D_i^{**} simplify to D_i when $k = 0$. It would be desirable to be able to write these measures as functions of leverage and residual, as was done in (5) and (6). This is not possible, however, because of the scale dependency of the ridge estimator. Since the ridge REC is not scale invariant, $\mathbf{X}_{(i)}$ has to be rescaled to unit column length before computing $\hat{\beta}_{(k)}(i)$. In Section 4, some approximated deletion formulas are presented.

3.4 Case Deletion for REC

When ith row is deleted from $\hat{\beta}_{(k)}$ then $\hat{\beta}_{(k)}(i)$ can be obtained by $\hat{\beta}_{(k)}(i) = (\mathbf{X}'_{(i)}\mathbf{X}_{(i)} - n\Lambda + kI_p)^{-1}\mathbf{X}'_{(i)}y_{(i)}$, where $\hat{\beta}_{(k)}(i)$ is the REC with out the i-row in $\hat{\beta}_{(k)}$, $\mathbf{X}_{(i)}$ is the matrix $\mathbf{X}$ when ith row is deleted and $y_{(i)}$ is the vector of response with out the ith entry. So the $\hat{\beta}_{(k)}(i)$ can be rewrite as $\hat{\beta}_{(k)}(i) = (\mathbf{X}'\mathbf{X} - x'_i x_i - n\Lambda + kI_p)^{-1}(\mathbf{X}'y - x'_i y_i)$ which, using the SMW theorem can be expended as;

$$\begin{aligned}\hat{\beta}_{(k)}(i) &= \left[\mathbf{A} + \frac{\mathbf{A}x'_i x_i \mathbf{A}}{1 - x_i \mathbf{A} x'_i}\right][\mathbf{X}'y - x'_i y_i] \\ &= \hat{\beta}_{(k)} - \frac{\mathbf{A}x'_i}{1 - h^*_i}(y_i - \hat{y}^*_i) \\ &= \hat{\beta}_{(k)} - \frac{\mathbf{A}x'_i}{1 - h^*_i} e^*_i.\end{aligned}$$

Based the above results the relations in (8)-(10) can be written in the form of residuals and leverage's as bellow

$$DFFITS^*_i = \frac{e^*_i}{p\hat{\sigma}_{(i)}} \cdot \frac{h^*_i}{1 - h^*_i} \left[\frac{1}{\sum_{j=1}^{n} h^{*2}_{ij}}\right]^{\frac{1}{2}}, \tag{11}$$

$$D^*_i = \frac{e^{*2}_i}{p\hat{\sigma}^2} \cdot \frac{\sum_{j=1}^{n} h^{*2}_{ij}}{(1 - h^*_i)^2}, \tag{12}$$

$$D^{**}_i = \frac{e^{*2}_i}{p\hat{\sigma}^2} \cdot \frac{h_i}{(1 - h^*_i)^2}. \tag{13}$$

The role of both the residual and leverage can be clearly seen in the preceding formulas in Section 3.2. Actually, these roles are almost identical to their roles in (6) and (7).

4 Numerical Example

4.1 Egyptian Pottery Data

As a numerical illustration, we consider a set of data, which refereed as the Egyptian pottery data. The data consist of measurements of chemical

contents (mineral elements) made on many samples of pottery using two different techniques which are known as neutron activation analysis (NAA) and inductively coupled plasma (ICP) spectrometry (Smith, et al., 1988). The set of pottery has been collected from different locations around the city. In general, two types of clay are known that have been used to make the ancient Egyptian pottery-Silt and Marl. In addition, archaeologists have classified some sherds as imports from north African countries, to distinguish them from known Egyptian Silt and Marlwares. The group structure among the objects arises from two main sources, fabric code and location of pottery. Both of these subdivisions are important to the archaeologists. Consequently, according to this group structure the selected vessels have been divided into 27 groups. In each group there are different numbers of vessels from the same fabric code and provenance which can essentially be regarded as replicated observations.

Among all mineral elements Rasekh and Feller (2003) have concentrated on the relation between the Na values measured by NAA as the dependent variable versus four mineral elements Na, Al, K and Ti measured by ICP technique as the independent variables. They fitted a functional measurement error model to this data set. By applying empirical influence function and Cook's statistic, they found that group number 5, 3, 13, and 19 were the most influence groups in model.

Rasekh (2006) analysed the same data to detect influential observations on the ridge regression estimator using the local influence approach with minor perturbation on variance, six explanatory variables (Na, Al, K,V ,Cr and Mn). By the variance perturbation, He detected that groups 18, 23, 15 and 6 (imported vessels groups) are most influential groups in ridge measurement error regression. He applied Cook's statistic and found groups 18, 6 and 12 have more influence receptively. Therefore, local influential groups 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 case's influence. In this paper we used the same data set to assess the influential observations in REC. Since we have replicated observations, we can derive an estimate of the covariance matrix of errors in this case.

Minimizing (Mallows, 1973) statistic for this dataset we obtain $k = 0.1$. Using the method of case deletion influential measures such as DFFITS, Cook's D_i^*, Leverage, and Residual. Influential measures DFFITS and Cook's D_i^*, were computed for various values of k and the results, for the four most influential groups, are presented in Figure 1 and Table 1. The effect of k

on the influence of each case is apparent. Influential groups are the same in these two plots but the order of influence magnitude is changed. In particular, we note that in these two plots the influence of cases decrease to a fix point as k grows. In contrast, the influence of group 18 decreases faster than others as k increases. Leverage and residuals are plotted against k in Figure 2. In leverage plot the influential groups are the same as those in the above DFFITS and Cook's D_i^* influential measures, but the order of magnitude is changed. Leverage plot shows that group 18 decreases faster than the rest, whereas case 15 as the highest leverage group decreases very slowly. From the residual plot, it can be observed that case 18 is the highest residual group and the influence of all groupss is relatively constant as k increases from 0.1.

Table 1. The four most influential groups according to D_i^* for Egyptian pottery data.

CLE ($k = 0$)		REC ($k = 0.1$)	
Group	D_i^*	*Group*	D_i^*
18	0.641	18	0.385
15	0.291	15	0.322
6	0.244	12	0.123
12	0.167	23	0.077

4.2 Soil Data

The data used here were obtained subsurface soil samples of representative soils from 48 corn plants in Zanjan, Iran. An investigation of the source from which corn farms obtain their phosphorus was carried out. Although the laboratory determination of phosphor is simple, in soil science it is interesting to evaluate the possibility of estimating accurate value of it from more easily and routinely determined soil properties. This approach can make the phosphor determination faster and cheaper, which is desirable when the each number of soil sample need to be analysed. Therefore most research evaluate the possibility of estimating phosphor through a multiple linear regression model analysis using routinely determined soil chemical data. Here, the continuous characteristics associated with the phosphorus content of corn grown in the soil are X_1 =*inorganic phosphorus in soil*, X_2 = *organic phosphorus in soil*, X_3 = *percentage of lime* ($Caco_3$) *in soil*, X_4 = *percentage of clay in soil* and X_1 = *PH of soil*. The contious response variable is y =*the phosphorus content of corn grown in the soil*. All above mentioned determination

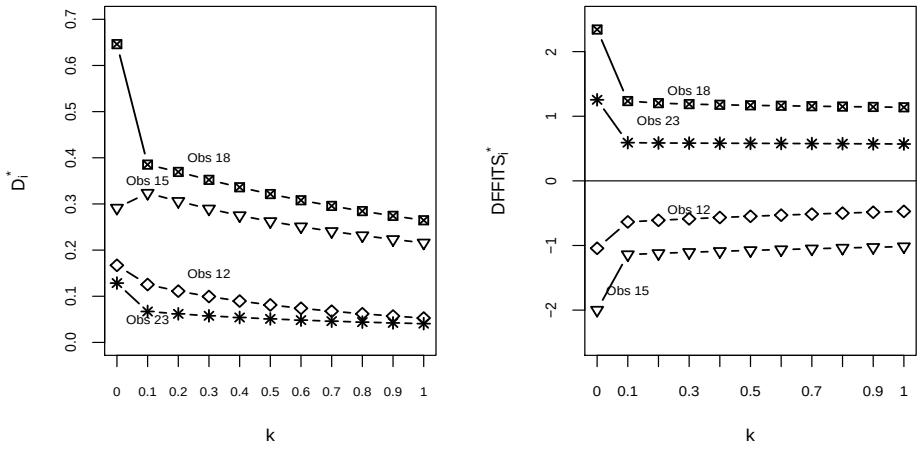

Figure 1. Plot of Cook's statistic D_i^* and $DFFITS_i$ against k for Egyptian pottery data.

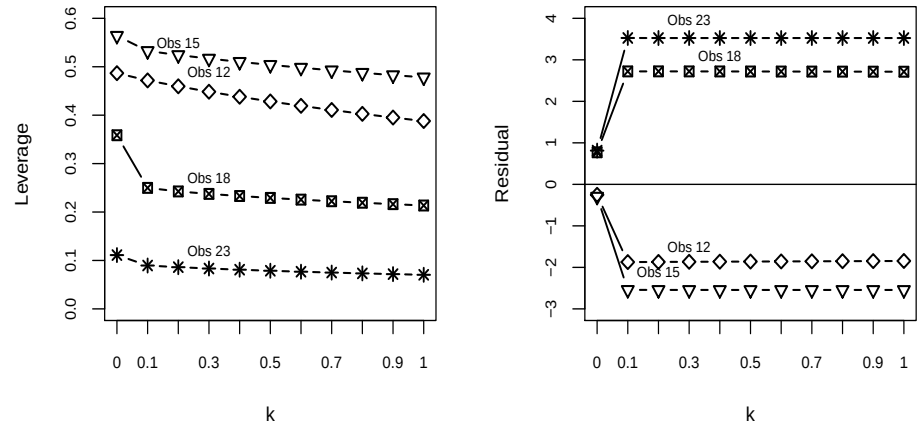

Figure 2. Plot of leverage and residuals against k for Egyptian pottery data.

were carried out on a volume basis dm^{-3} as done in most Iranian laboratories whose analyses are directed to soil fertility evaluation. However, the estimates of $X_1 - X_5$ contain measurement error arising from two sources. First, only a small sample of soil is selected from each farm and, as a result, there is the sampling error associated with the use of the sample to represent the whole. Second, there is a measurement error associated with the chemical analysis used to determine the values of $X_1 - X_5$ the soil sample. Also, the condition number of X is 244.130 indicating collinearly problem in X. From the scatter plot of Figure 3, it can be seen that there are moderate correlations between $X_1 - X_5$. Thus the linear measurement error model (1) with ridge estimation is used to mitigate the effect of collinearity in this dataset.

Table 2. The estimates of ridge measurement error regression model (with $k = 0.02$) for Soil data.

Parameter	β_1	β_2	β_3	β_4	β_5
Estiate	0.276	0.077	-0.106	0.441	0.131
Standard error	0.072	0.0703	0.108	0.067	0.100

Table 2 shows the estimates and standard error of estimates. From the left panel of Figure 4 the ridge parameter k=0.02 was selected by minimizing the cross validation criterion (GCV).

Since, our interest lies on the sensitivity analysis of the fitted model model, Figure 4 and Figure 5 shows the diagnostic plots according to proposed influence measures. From the right panel of Figure 4 we see that cases 19, 21, 13 and 24 (in this order) are the most high leverage points in the fitted ridge measurement error regression model.

Figure 5 shows the Cook's distance and standardized ridge residuals in the left and right panel respectively. From Cook's distance plot it appears that cases 13, 31, 29 and 21 are the most influential points in the mentioned fitted model.The joint effect of leverage and standardized residuals is reflected in the change of influence observed in Cook's distance plot.

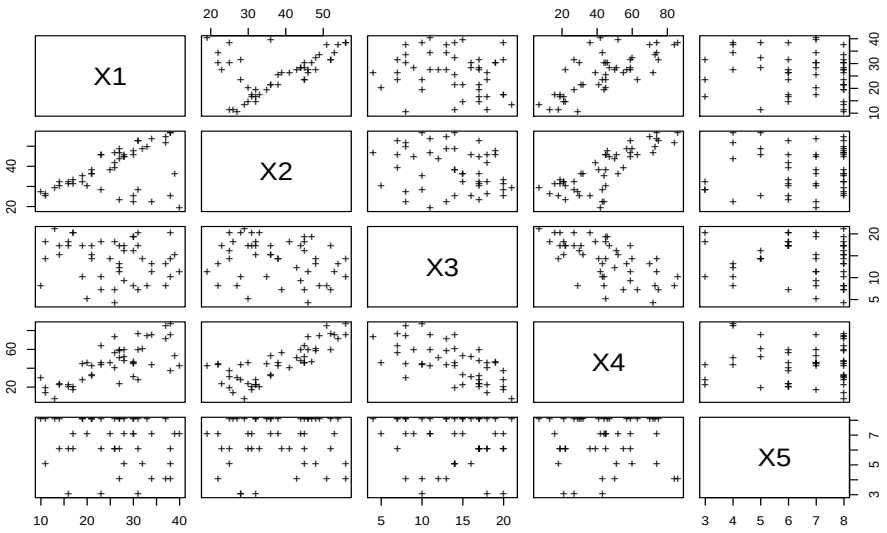

Figure 3. Scatter plot of $X_1 - X_5$ for Soil data.

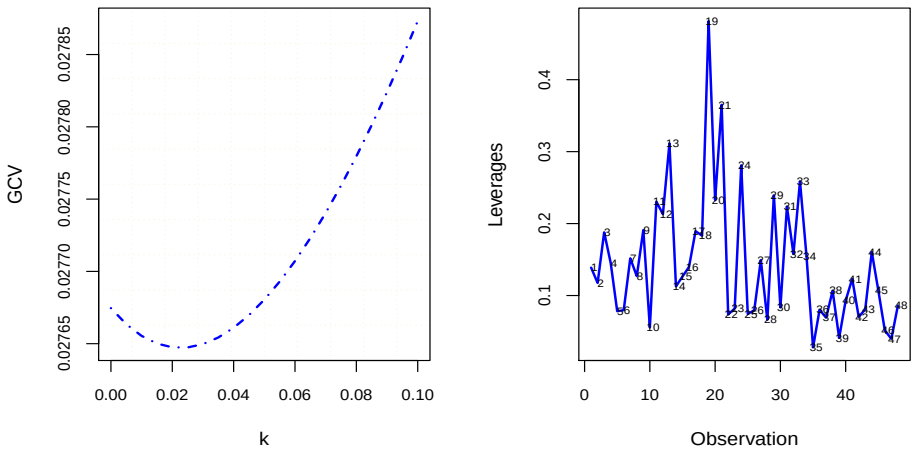

Figure 4. left panel: Plot of GCV against k, right panel: Leverages plot for Soil data.

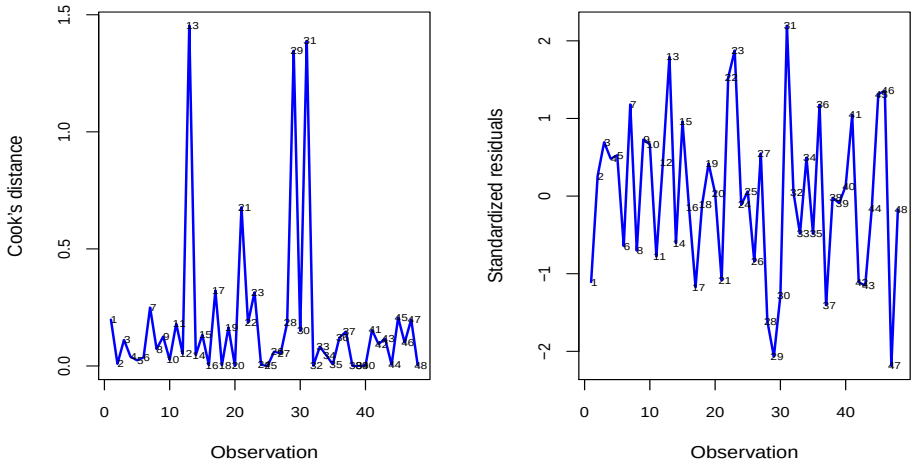

Figure 5. left panel: Plot of Cook's distance, right panel: Plot of standardized residuals for Soil data.

5 Conclusion

There is no method available for the possible effect's that collinearity can have on the influence of an observation in linear measurement error. Belsley et al. (1989) noted that biased estimators are used to reduce the affect of multicollinearity and that the influence of some cases can be modified in linear regression. Based on this fact, Belsley et al. (1989) suggested that multicollinearity should be controlled before attempting to measure influence. In this article, we show that using REC we can not rely on influence measures obtained for CLE. Once the value of k is determined, influence measures should be computed for that k. If, after analysing these indexes, it is decided to delete one or more cases from the analysis, the whole situation has to be reassessed in terms of both influence and multicollinearity parameter k is estimated first and for that k value the REC coefficients are estimated. Using these parameter quantities the influential observations are identified. But, the value of shrinkage parameter k depends on the every observation. Hence, for every influential case the value of k is estimated. The main advantage of the deletion formulas in Section 3 is that, as in least squares, the estimator does not have to be computed every time a case is

deleted. For a value of k all of the elements in (11), (12) and (13) are readily available from a single run of REC. Moreover, these measures, based on deletion formulas are particularly helpful for large data sets. Furthermore, the deletion formulas provide computationally inexpensive approximate influence measures for REC. We also note that, although the issue of detecting influential for high dimensional data has been studied in linear regression (see Pena, 2005), but dealing with high dimensional covariates no conventional diagnostics methods such as Pena's statistic is introduced or developed for linear measurement error models and ridge linear measurement errors. These are additional active issues for future research study.

Acknowledgement

The authors would like to thank the Editor and anonymous referees for several helpful comments and suggestions, which resulted in a significant improvement in the presentation of this article.

References

Belsley, D.A., Kuh, E. and Welsch, R.E. (1989). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* Wiley, New York.

Castro, M., Rojas, M., and Bolfarine, H., (2007). Local Influence Assessment in Heteroscedastic Measurement Error Models. *Computational Statistics and Data Analysis* , **52**, 1132-1142.

Cook, R.D. (1977). Detection of Influential Observations in Linear Regression. *Technometrics*, **19**, 15-18.

Fuller, W.A. (1987). *Measurements Error Models.* Wiley, New York.

Gimenez, P. and Galea, M. (2013). Influence Measures on Corrected Score Estimators in Functional Heteroscedastic Measurement Error Models. *Multivariate Analysis*, **114**, 1-15.

Hoerl, A.E. and Kennard, R.W. (1976). Ridge Regression: Iterative Estimation of the Biasing Parameter. *Communications in Statistics - Theory and Methods* ,**5**, 77-88.

Janufer, A. and Jianbao, Ch. (2009). Assessing Global Influential Observations in Modified Ridge Regression. *Statistics and Probability Letters*, **79**, 513-518.

Janufer, A. and Jianbao, Ch. (2012). Identifying Local Influential Observation in Liu Estimator. *Metrika*, **75**, 425-438.

Kelly, G. (1984). The Influence Function in the Errors in Variables Problem. *Annals of Statistics*, **12**, 87-100.

Mallows, C.L. (1973). Some Comments on C_p. *Technometrics*, **15**, 661-675.

Mc Donald, G.C. and Galarneau, D.I. (1975). A Monte Carlo Evaluation of Some Ridge-type Estimators. *American Statistical Association*, **70**, 407-416.

Nakamura, T. (1990). Corrected Score Functions for Error in Variables Models: Methodology and Application to Generalized Linear Models. *Biometrika*, **77**, 127-137.

Pena, D.A. (2005). New Statistic for Influence in Linear Regression. *J. American Statistical Association*, **47**, 1-13.

Rasekh, A.R. (2006). Local Influence in Measurement Error Models with Ridge Estimate.*Computational Statistic and Data Analysis*, **50**, 2822-2834.

Rasekh, A.R. and Feller, N.R.J. (2003). Influence Functions in Functional Measurement Error Models with Replicated data. *Statistics*, **37**, 169-178.

Seber, G.A.F. (1982). *A Matrix Handbook for Statisticians*. Wiley, New York.

Shi, L. (1997) Local Influence in Principal Component Analysis. *Biometrika*, **84**, 175-186.

Smith, D.M., Hart, F.A., Symond, R.D. and Walsh, J.N. (1988). Analysis of Roman Pottery from Colchester by Inductively Coupled Plasma Spectrometry. In: Slater, E.A., Tate, J.O. (Eds.), Science and Archaeology Glasgow 1987, vol. 196(I). B.A.R., Oxford, 41–55

Stefanski, L.A. and Carroll, R.J. (1989). Conditional Scores and Optimal Scores for Generalized Measurement Error Models. *Biometrica*, **74**, 703-716.

Walker, E. and Birch, J. (1989). Influence Measures in Ridge Regression. *Technometrics*, **30**, –221-227.

Wellman, J.M. and Gunst, R.F. (1991). Influence Diagnostics for Linear Measurement Error Models *Biometrika*, **78**, 373-387.

Zhao, Y. and Lee, A.H. (1994). Influence Diagnostics for Generalized Measurement Error Models. *Biometrics*, **50**, 1117-1128.

Zhong, X., Wei, B. and Fung, W. (2000). Influence Analysis for Measurement Error Models. *Ann. Inst. Statist. Math.*, **52**, 367-379.

Appendix 1

Let $\mathbf{A}$ be $n \times n$ symmetric matrix and $\mathbf{B}$ be $n \times n$ positive definite matrix. Let $\gamma_1 \geqslant, \gamma_2 \geqslant \ldots \geqslant \gamma_n$ be the eigenvalues of $\mathbf{B}^{-1}\mathbf{A}$. Then $(x \neq 0)$ $\max \frac{x'\mathbf{A}x}{x'\mathbf{B}x} = \gamma_1$

and $\min \frac{x'\mathbf{A}x}{x'\mathbf{B}x} = \gamma_n$ (see Seber, 1982).

Hadi Emami
Department of Statistics,
University of Zanjan,
Zanjan, Iran.
email: *h.emami@znu.ac.ir*