

Detecting Influential observations in Two-Parameter Liu-Ridge Estimator

Adewale F. Lukman¹ and Kayode Ayinde²

¹*Department of Mathematics, Landmark University, Omu-Aran, Kwara State, Nigeria.*

²*Department of Statistics, Federal University of Technology, Akure, Ondo State, Nigeria.*

Abstract: Influential observations do posed a major threat on the performance of regression model. Different influential statistics including Cook's Distance and DFFITS have been introduced in literatures using Ordinary Least Squares (OLS). The efficiency of these measures will be affected with the presence of multicollinearity in linear regression. However, both problems can jointly exist in a regression model. New diagnostic measures based on the Two-Parameter Liu-Ridge Estimator (TPE) defined by Ozkale and Kaciranlar (2007) was proposed as alternatives to the existing ones. Approximate deletion formulas for the detection of influential cases for TPE are proposed. Finally, the diagnostic measures are illustrated with two real life dataset.

Key words: Influential Statistics, Multicollinearity, Diagnostic Measures, Approximate Deletion Formulas, Two-Parameter Liu-Ridge Estimator.

1. Introduction

In the application of regression analysis, there can be strong or nearly perfect relationship among the regressors. When this relationship exists, the regression model is said to suffer the problem of multicollinearity. The efficiency of Ordinary Least Square (OLS) estimator when applied to multicollinear data is seriously affected. The regression coefficients possess large standard errors and sometimes exhibit wrong sign (Johnston, 1972). In literature, some biased estimation techniques are introduced to solve this problem. Among them are the ridge regression estimator and Liu estimator introduced by Hoerl and Kennard (1970) and Liu (1993) respectively. The effect of influential observations on the OLS estimates has been thoroughly investigated over the years but little attention has been given to the effect of influential observations on biased estimation techniques. The problem of multicollinearity and influential observation on OLS has been addressed by different authors (Cook, 1977; Belsley *et al.*, 1980; Cook and Weisberg (1980); Cook and Weisberg (1982); Chattergee and Hadi, 1986; Cook, 1986). Belsley (1991) investigated the effect of leverage in ridge regression. Walker and Birch (1988) examined the effect of influential points on ridge regression using case deletion method. Local influence analysis in principal component analysis and ridge regression estimator was investigated by (Shi, 1997; Shi and Wang, 1999). Jahufer and Jianbao (2009) used the modified ridge regression to detect influential points. Local influential analysis in Liu estimator was studied by Jahufer and Chen (2011). Jahufer (2013) investigated the effect of influential points on Liu estimator. Yasin and Murat (2016) studied the impact of influential point on two-parameter ridge regression.

The objective of this paper is to introduce new influence diagnostics based on a Two-parameter estimator which combines ridge and Liu estimators defined by Ozkale and Kaciranlar (2007). Also, to obtain the generalized versions of Cook's D, DFFITS and approximate case deletion formulas for this estimator.

The organization of this paper is as follows: Background information and review of the influence measures in OLS is given in section 2. New diagnostic measures in two-parameter Liu-Ridge estimator is introduced in section 3. Section 4 covers application to two real life datasets. Discussion is provided in the last session.

2. Background and Definition

2.1. Background

Consider the linear regression model

$$y = X\beta + \varepsilon \quad (1)$$

where y is an $n \times 1$ vector of response variable, X is an $n \times p$ centered and standardized known matrix. β is $p \times 1$ vector of the unknown regression coefficients and ε is the $n \times 1$ vector of error terms with $E(\varepsilon) = 0$ and $V(\varepsilon) = \sigma^2 I_n$ and I_n is an $n \times n$ matrix of identity matrix.

The OLS estimator is defined as:

$$\hat{\beta} = (X'X)^{-1}X'y \quad (2)$$

The corresponding residual vector, e_i , is defined as

$$\begin{aligned} e_i &= y - \hat{y} = y - X\hat{\beta} \\ &= y - (X'X)^{-1}X'y = y(I - X(X'X)^{-1}X') \text{ since } \hat{\beta} = (X'X)^{-1}X'y \\ e_i &= (I - H)y \text{ where } H = X(X'X)^{-1}X' \text{ is the hat matrix} \end{aligned}$$

2.2 Influential Measures in Least Squares

2.2.1 DFFITS

DFFITS is the standardized change in the fitted value when a case is deleted. It is defined as:

$$DFFITS_i = \frac{x_i[\hat{\beta} - \hat{\beta}_{(i)}]}{S(x_i\hat{\beta})} \quad (3)$$

where $S(x_i\hat{\beta})$ is an estimator of standard error of the fitted values, x_i is the i th row of the X matrix, $\hat{\beta}_{(i)}$ is the least squares estimator of β when the i th case is omitted in fitting the regression function. It can also be expressed algebraically as:

$$(DFFITS)_i = \left(\frac{h_{ii}}{1-h_{ii}}\right)^{1/2} \frac{\varepsilon_i}{\sqrt{\hat{\sigma}^2(1-h_{ii})}} = \left(\frac{h_{ii}}{1-h_{ii}}\right)^{1/2} t_i \quad (4)$$

where $\hat{\sigma}^2$ is the estimate of σ^2 , h_{ii} is the diagonal elements of the hat matrix and $t_i = \frac{\varepsilon_i}{\sqrt{\hat{\sigma}^2(1-h_{ii})}}$ is the studentized residual (also called the external studentized residual). Jahufer (2013) suggested that for large data sets any observations for which the absolute value of $DFFITS$ exceeds $2\sqrt{\frac{p}{n}}$ warrants attention. Draper (1981) suggested that for small to medium data sets any observations for which the absolute value of $DFFITS$ exceeds 1 is influential.

2.2.2 Cook's Distance Measures

Cook's distance measure denoted by D_i , considers the influence of the i^{th} case on all n fitted values.

$$D_i = \frac{(\hat{\beta} - \hat{\beta}_{(i)})'(X'X)(\hat{\beta} - \hat{\beta}_{(i)})}{ps^2} \quad (5)$$

where $\hat{\beta}_{(i)}$ is the least squares estimator of β without the i^{th} case. An equivalent algebraic expression of Cook's D Measure is given by:

$$D_i = \frac{r_i^2}{p} \left(\frac{h_{ii}}{1-h_{ii}} \right) \quad (6)$$

where h_{ii} is the diagonal elements of the hat matrix and where r_i is i^{th} internally studentized residual. It was suggested that observations for which $D_i > 1$ warrants attention (Cook, 1977).

3. Influence Measure in Two-Parameter Ridge-Liu Estimator

3.1. Two-Parameter Liu-Ridge Estimator

Ridge estimator, $\hat{\beta}_R$, was introduced by Hoerl and Kennard (1970) and defined as:

$$\hat{\beta}_R = (X'X + kI)^{-1}X'y \quad (7)$$

where I is an identity matrix, k is the ridge parameter or biasing constant and often takes values between 0 and 1.

Liu (1993) introduced the Liu estimator, $\hat{\beta}_d$, which combines the ridge estimator by Hoerl and Kennard (1970) with the stein estimator by Stein (1956). It is defined as:

$$\hat{\beta}_d = (X'X + I)^{-1}(X'y + d\hat{\beta}) \quad (8)$$

where I is an identity matrix, $\hat{\beta}$ is the least square estimator of β , d is referred to as the Liu biasing parameter and do takes values between 0 and 1.

Ozkale and Kaciranlar (2007) introduced a Two-Parameter Ridge-Liu estimator (TPE) given as:

$$\hat{\beta}_{TP} = (X'X + kI)^{-1}(X'y + kd\hat{\beta}) \quad (9)$$

where I is an identity matrix, k is the ridge parameter and d is Liu biasing parameter. $\hat{\beta}$ is the least square estimator of β . TPE is introduced for handling the problem of multicollinaerity in the linear regression model.

3.2 Leverage and Residual Measures in Two-Parameter Ridge-Liu Estimator

The vector of fitted value of TPE is

$$\begin{aligned} \hat{y}_{TP} &= X\hat{\beta}_{TP} = X(X'X + kI)^{-1}(X'y + kd\hat{\beta}) \\ &= X(X'X + kI)^{-1}(X'y + kd(X'X)^{-1}X'y) \text{ where } \hat{\beta} = (X'X)^{-1}X'y \\ &= X(X'X + kI)^{-1}(I + kd(X'X)^{-1})X'y \\ \hat{y}_{TP} &= X(X'X + kI)^{-1}(X'X + kdI)(X'X)^{-1}X'y = H_{TP}y \end{aligned} \quad (10)$$

where $H_{TP} = X(X'X + kI)^{-1}(X'X + kdI)(X'X)^{-1}X'$ is Two-Parameter Ridge-Liu Estimator hat matrix and this performs the same role with the hat matrix (H) in Ordinary Least Square (OLS) estimator. The i^{th} fitted value can be written in terms of the elements of H_{TP} as $\hat{y}_{iTP} = \sum_{j=1}^n h_{TPji} y_j$.

The partial derivative of \hat{y}_{iTP} with respect to y_i is given as $\frac{\partial \hat{y}_{iTP}}{\partial y_i} = h_{TPii}$, which is the i th diagonal

element of H_{TP} . It should be noted that H_{TP} is not idempotent. It is therefore a quasi-projection matrix (Walker and Birch, 1988). However, the canonical reduction can be applied by applying singular value decomposition (SVD) (Mandel, 1982). Design matrix X can be decomposed as $X=U\Lambda V'$ where U and V are orthogonal matrices of order $n \times p$ and $p \times p$, respectively; and Λ is a diagonal matrix of order $p \times p$ containing the singular values of X . Matrix V are eigenvectors of X such that $X'X = V\Lambda V'$. The ij th element of $n \times p$ matrix U is such that $u_{ij}\sqrt{\lambda_j}$ is the projection of the i th row, x_i onto the j th eigenvector of X . The i th leverage of TPE can be written as follows:

$$\begin{aligned} h_{TPii} &= x(X'X+kI)^{-1}(X'X+kI)(X'X)^{-1}x' \\ &= \sqrt{\lambda_j}u_{ij}v_i'(V\Lambda V'+kI)^{-1}(V\Lambda V')^{-1}(V\Lambda V'+kI)v_iu_{ij}\sqrt{\lambda_j} \\ &= \sum_{j=1}^p \frac{(\lambda_j+kd)}{(\lambda_j+kI)}u_{ij}^2 \end{aligned} \quad (12)$$

It is observed that h_{TPii} values approaches the OLS leverages, h_{ii} , as k and d approaches 0 and 1 respectively.

The i th residual of TPE is defined as:

$$e_{iTP} = y_i - \hat{y}_{iTP} = (1 - h_{TPii})y_i.$$

3.3 DFFITS and Cook's D Measures in Two-Parameter Ridge-Liu Estimator

Following equation (3), *DFFITS* for TPE is defined as:

$$DFFITS_{iTP} = \frac{x_i[\hat{\beta}_{TP} - \hat{\beta}_{(i)TP}]}{S(x_i\hat{\beta}_{TP})} = \frac{x_i[\hat{\beta}_{TP} - \hat{\beta}_{(i)TP}]}{s(i)\sqrt{\sum_{j=1}^n h_{TPij}^2}} \quad (13)$$

where $\hat{\beta}_{(i)TP}$ is the TPE in equation (3) without the i th case and the denominator is the estimator of the standard error of the TPE fitted value such that $S(x_i\hat{\beta}_{TP}) = s(i)\sqrt{\sum_{j=1}^n h_{TPij}^2}$ such that k and d are assumed to be non-scholastic and $s(i) = \sqrt{\frac{(n-p)s^2 - e_i^2/(1-h_{ii})}{n-p-1}}$ is the OLS estimator s of σ without the i th case. OLS estimators of s and $s(i)$ will be used as measure of scales because they affected by multicollinearity.

Cook's D version for TPE is defined as in two forms. A direct version of Cook's D in equation (5) is defined as:

$$D_{iTP}^* = \frac{1}{ps^2} [\hat{\beta}_{TP} - \hat{\beta}_{(i)TP}]' (X'X) [\hat{\beta}_{TP} - \hat{\beta}_{(i)TP}] \quad (14)$$

while the other version is defined as

$$D_{iTP}^{**} = \frac{1}{ps^2} [\hat{\beta}_{TP} - \hat{\beta}_{(i)TP}]' (X'X + kI)(X'X + kdI)^{-1}(X'X)(X'X + kdI)^{-1}(X'X + kI) [\hat{\beta}_{TP} - \hat{\beta}_{(i)TP}] \quad (15)$$

Based on the fact that $\text{var}(\hat{\beta}_{TP}) = \sigma^2(X'X + kI)^{-1}(X'X + kdI)(X'X)^{-1}(X'X + kI)^{-1}(X'X + kdI)$.

It is not possible to write the measures obtained in equation (13) to (15) as functions of leverage and residual because of the scale dependency of TPE. The estimator is not scale invariant, thus, the design matrix X with the i th row deleted has to be rescaled before $\hat{\beta}_{(i)TP}$ is computed. The approximate versions of these measures are provided through approximate case deletion formulas.

3.4. Approximate Case Deletion Formulas for Two-Parameter Ridge-Liu Estimator

Deleting i th row from $\hat{\beta}_{TP}$, $\hat{\beta}_{(i)TP}$ can be written as $\hat{\beta}_{(i)TP} = (X'_{(i)}X_{(i)} + kI)^{-1}(X'_{(i)}y_i + kd\hat{\beta}_{(i)})$. where $X_{(i)}$ is the matrix X without the i th row and y_i is the vector of response variable without the i th entry. $X_{(i)}$ is scaled so that $X'_{(i)}X_{(i)}$ is in correlation form. Applying Sherman-Morrison- Woodbury (SMW) theorem (Rao, 1973), $\hat{\beta}_{(i)TP}$ can be approximated as:

$$\begin{aligned} \hat{\beta}_{(i)TP} &= (X'X - x'_i x_i + kI)^{-1}(X'y - x'_i y_i + kd\hat{\beta}_{(i)}) \text{ such that } X'X + kI = C_k. \\ \hat{\beta}_{(i)TP} &= (C_k - x'_i x_i)^{-1}(X'y - x'_i y_i + kd\hat{\beta}_{(i)}) \\ \text{where } &= (C_k - x'_i x_i)^{-1} = C_k^{-1} + \frac{C_k^{-1} x'_i x_i C_k^{-1}}{1 - x_i C_k^{-1} x'_i}. \\ \hat{\beta}_{(i)TP} &= (C_k^{-1} + \frac{C_k^{-1} x'_i x_i C_k^{-1}}{1 - x_i C_k^{-1} x'_i})(X'y + kd\hat{\beta}_i - x'_i y_i) \text{ such that } m_{ii} = x_i C_k^{-1} x'_i \\ &= C_k^{-1}(X'y + kd\hat{\beta}_i) - C_k^{-1} x'_i y_i + \frac{C_k^{-1} x'_i x_i C_k^{-1}}{1 - m_{ii}}(X'y + kd\hat{\beta}_i) - \frac{C_k^{-1} x'_i x_i C_k^{-1}}{1 - m_{ii}} x'_i y_i \\ &= \hat{\beta}_{TP} + \frac{C_k^{-1} x'_i}{1 - m_{ii}}(x_i C_k^{-1}(X'y + kd\hat{\beta}_i) - x_i C_k^{-1} x'_i y_i - (1 - m_{ii}) y_i) \end{aligned}$$

Recall that

$$\begin{aligned} m_{ii} &= x_i C_k^{-1} x'_i, \hat{y}_{TP} = x_i C_k^{-1}(X'y + kd\hat{\beta}_i) \\ &= \hat{\beta}_{TP} + \frac{C_k^{-1} x'_i}{1 - m_{ii}}(\hat{y}_{TP} - m_{ii} y_i - (1 - m_{ii}) y_i) \\ &= \hat{\beta}_{TP} + \frac{C_k^{-1} x'_i}{1 - m_{ii}}(\hat{y}_{TP} - y_i) \\ &= \hat{\beta}_{TP} - \frac{C_k^{-1} x'_i}{1 - m_{ii}}(y_i - \hat{y}_{TP}) \text{ where } e_{iTP} = y_i - \hat{y}_{TP} \\ \hat{\beta}_{(i)TP} &\cong \hat{\beta}_{TP} - \frac{e_{iTP} C_k^{-1} x'_i}{1 - m_{ii}} \\ \hat{\beta}_{TP} - \hat{\beta}_{(i)TP} &\cong \frac{e_{iTP} C_k^{-1} x'_i}{1 - m_{ii}} \end{aligned} \tag{16}$$

Based on these result, the approximate version of (13) can be written as:

$$DFFITs_{iTP} = \left[\frac{m_{ii}}{1 - m_{ii}} \right] \frac{e_{iTP}}{S(x_i \hat{\beta}_{TP})} = \left[\frac{m_{ii}}{1 - m_{ii}} \right] \frac{e_{iTP}}{s(i) \sqrt{\sum_{j=1}^n h_{TPij}^2}} \tag{17}$$

$$\text{where } s(i) = \sqrt{\frac{(n-p)s^2 - e_i^2 / (1 - h_{ii})}{n - p - 1}}, m_{ii} = x_i C_k^{-1} x'_i, C_k^{-1} = (X'X + kI)^{-1}$$

The approximate version of (14) can be written as follows:

$$\begin{aligned} D_{iTP}^* &= \frac{1}{ps^2} \left[\frac{e_{iTP} C_k^{-1} x'_i}{1 - m_{ii}} \right]' (X'X) \left[\frac{e_{iTP} C_k^{-1} x'_i}{1 - m_{ii}} \right] \\ &= \frac{1}{ps^2} \frac{x'_i}{(1 - m_{ii})^2} C_k^{-1} e_{iTP}^2 (X'X) C_k^{-1} x'_i \text{ where } C_k^{-1} = (X'X + kI)^{-1} \\ D_{iTP}^* &= D_{iTP}^{**} = \frac{1}{ps^2} \left[\frac{e_{iTP}}{1 - m_{ii}} \right]^2 x'_i (X'X + kI)^{-1} (X'X) (X'X + kI)^{-1} x'_i \\ &= \frac{1}{ps^2} \left[\frac{e_{iTP}}{1 - m_{ii}} \right]^2 \sum_{j=1}^n m_{ij}^2 \end{aligned} \tag{18}$$

The approximate version of (15) can be written as follows:

$$\begin{aligned}
 D_{iTP}^{**} &= \frac{1}{ps^2} \left[\frac{e_{iTP}}{1-m_{ii}} \right]^2 x_i (X'X + kI)^{-1} (X'X + kdI) (X'X)^{-1} (X'X + kdI) (X'X + kI)^{-1} x_i' \\
 &= \frac{1}{ps^2} \left[\frac{e_{iTP}}{1-m_{ii}} \right]^2 \sum_{j=1}^n h_{TPij}^2
 \end{aligned}
 \tag{19}$$

4. Application to Real life Dataset

Real life data sets are used to illustrate the performance of the influential statistics. The results are as follows.

4.1. Application to Longley Data

This was adopted from the study of Longley (1967). The regression model is defined as:

$$y = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_5 X_5 + \beta_6 X_6 + e \tag{20}$$

where y is the total derived employment, X_1 is the gross national product implicit price deflator,

X_2 is the gross national product, X_3 is unemployment, X_4 is the size of armed forces, X_5 is the non-institutional population 14 years of age and over and X_6 is the time. The scaled condition number of this data is 43,275 (Walker and Birch, 1988). Several authors had used this data to identify influential observations (Cook, 1977; Walker and Birch, 1988; Jahufer and Jianbao, 2009; Jahufer, 2013; Ullah *et al.*, 2013; Yasin and Murat, 2016). The results are summarized in Table 1. Ridge parameter suggested by Hoerl *et al.* (1975) and classified in the study of Lukman and Ayinde (2015) as Fixed Maximum Original is used. It is defined as $\hat{K}_{HK}^{FMO} = \frac{\hat{\sigma}^2}{\max(\hat{\alpha}_i^2)}$. The value of K and d in this study are computed to be 5.36488D-08 and 0.9 respectively. The value of Liu ridge parameter, d , used in this study was the one used in the study of Ullah *et al.* (2013) when applied to the same data set.

Table 1: The most five influential observations using leverage, residual, DFFITS and Two versions of Cook's Distance

Leverage _{OLS}		Leverage _{iTP}		DFFITS _{iTP}		D_{iTP}^*		D_{iTP}^{**}	
Case	Value	Case	Value	Case	Value	Case	Value	Case	Value
16	0.6261	16	0.6582	16	1.8237	4	0.7003	16	0.4013
5	0.5530	2	0.5645	10	1.6448	10	0.6623	5	0.2493
2	0.5025	8	0.4851	4	1.6406	16	0.6592	4	0.2020
8	0.4422	12	0.4828	5	1.3824	5	0.4862	15	0.1882
7	0.4290	1	0.4245	15	1.1696	1	0.3327	1	0.1530

Table 2: Summary of most influential diagnostics on Longley datasets compare with propose

Authors	Year	Influential points in order	Method
Cook	1977	5, 16, 4, 10, 15	Cooks distance in OLS
Walker and Birch	1988	16, 10, 4, 15, 1	Cooks distance in ridge regression
Shi and Wang	1999	10, 4, 15, 16 and 1	Local influence in ridge regression
Jahufer and Jianbao	2009	16, 4, 1, 10, 15	Modified Ridge regression of the usual diagnostics
Jahufer	2013	15,4 1, 6, 16	Cooks distance and DFFITs in in Liu regression
Ullah et al.	2013	16, 10, 4, 6, 1 16, 10,4,6, 5 and 16, 5, 4, 10, 15	Influential points in Liu regression for different d=0.1, 0.5 and 0.9 respectively
Yasin and Murat	2016	16, 10, 6, 1 and 4	Influential points in Two-parameter ridge
$DFFITS_{ITP}$	Proposed	16, 10, 4, 5 and 15	Influential points in Two-parameter Ridge-Liu
D_{ITP}^*	Proposed	4, 10, 16, 5, 1	Influential points in Two-parameter Ridge-Liu
D_{ITP}^{**}	Proposed	16, 5, 4, 15, 1	Influential points in Two-parameter Ridge-Liu

Table 1 shows that the leverage based on OLS and the proposed estimator identified observations 16, 2 and 8 as leverages in common but different in order of ranking. This also agrees with the study of Yasar and Murat (2016) where the first three leverages are case number 16, 2 and 8. This is further illustrated in Figure 1. Figure 1 is the plot of the hat diagonal matrix against the observations. The summary from to Table 2 shows that the proposed statistics $DFFITS_{ITP}$ identified the same value that Cook (1997) and Ullah (2013) identified as influential points, though, in a different order. D_{ITP}^* and D_{ITP}^{**} identified the same cases as identified by other authors except case 1. The result is also validated in Figure 2.

4.2. Application to Hald Data

This has been previously adopted by Cook (1977); Yasin and Murat (2016) in the investigation of influential observations. Four regressors were used with thirteen (13) observations. The condition number of matrix X is computed to be 249.578 which show that the model suffers the problem of strong multicollinearity. Cook (1977) identified observations 8, 3, 11, 6 and 13 as influential points in this order. Yasin and Murat (2016) detect the influential points to be 8, 11, 10, 6 and 13 when Cooks Distance based on Two-parameter ridge was used. Observations 8, 11, 10, 3 and 6 were identified when generalized cooks D was used while observations 8, 11, 6, 10 and 13 were identified with the use of DFFITS based on Two-parameter ridge. The value of k and d in this study are computed to be 0.0076761 and 1.18495 respectively. The results obtained in this study are summarized in Table 3. It was observed that the Cooks D obtained in this study and previous study performs similarly by identifying cases 8, 13, 6 and 11 in common but the order of appearance differs. DFFITS based on OLS identified cases 8, 11, 7, 2 and 4 as influential observations. The method of DFFITS suggested by Yasin and Murat (2016) identified the cases 8, 11, 6, 10 and 13 as influential while the method proposed in this study identified cases 8, 6, 13, 7 and 11 as influential points. Comparing the three methods, observation 8 is the most influential. However, the method proposed in this study and that of Yasin and Murat (2016) identifies four cases in common, though, the order differs. The new diagnostics defined based on Two-parameter Liu ridge estimator competes favorably with the existing ones.

Table 3: The most five influential observations using leverage, residual, DFFITS and Two versions of Cook's Distance

Leverage _{OLS}		Leverage _{iTP}		DFFITS _{iTP}		D_{iTP}^*		D_{iTP}^{**}	
Case	Value	Case	Value	Case	Value	Case	Value	Case	Value
10	0.7004	10	0.6986	8	1.6854	8	0.4105	8	0.4619
3	0.5769	1	0.5444	6	0.7431	13	0.1500	13	0.1631
1	0.5503	3	0.5379	13	0.6606	6	0.0837	11	0.0992
11	0.4255	7	0.4229	7	0.5858	11	0.0794	2	0.0922
8	0.4085	11	0.4210	11	0.5614	2	0.0704	7	0.0745

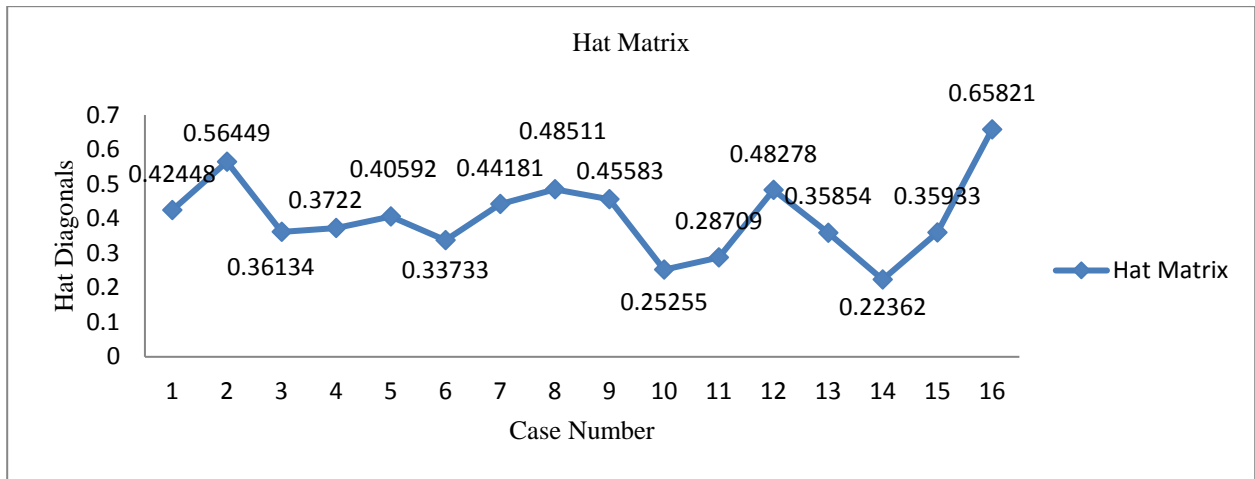


Figure 1: Plot of hat diagonals based on TPE against its observations

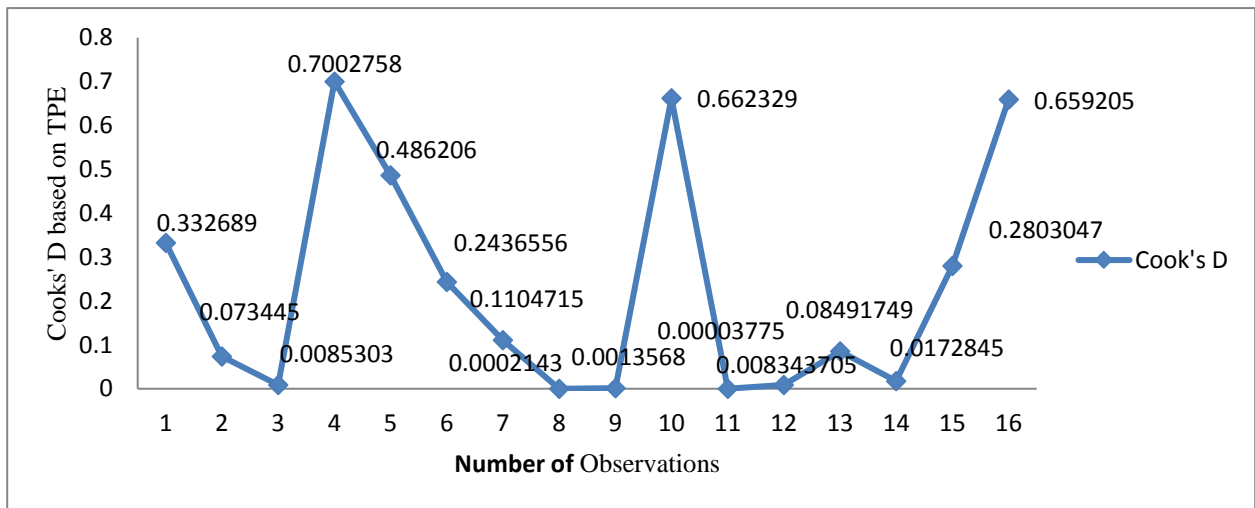


Figure 2: Plot of Cook's D based on TPE against its observations

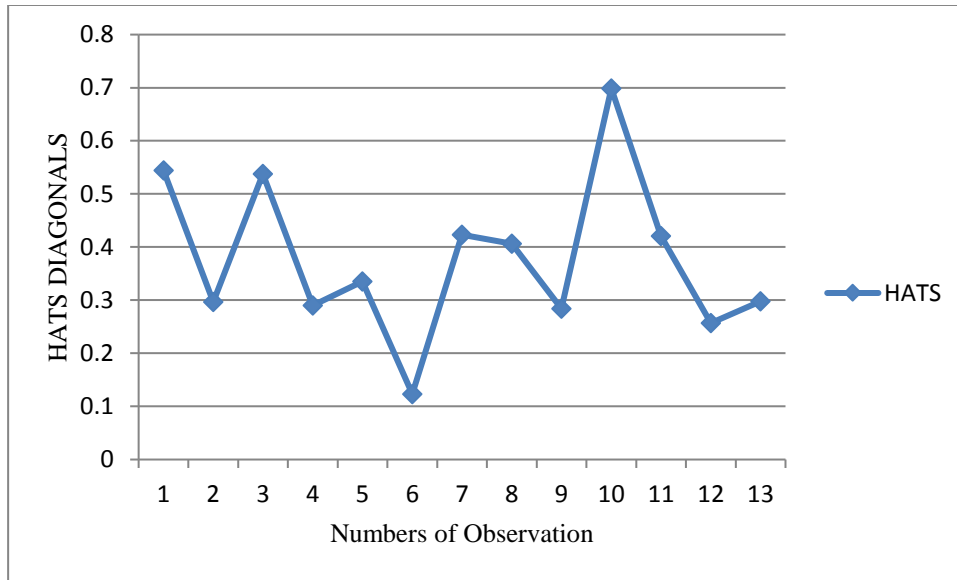


Figure 3: Plot of hat diagonals based on TPE against its observations

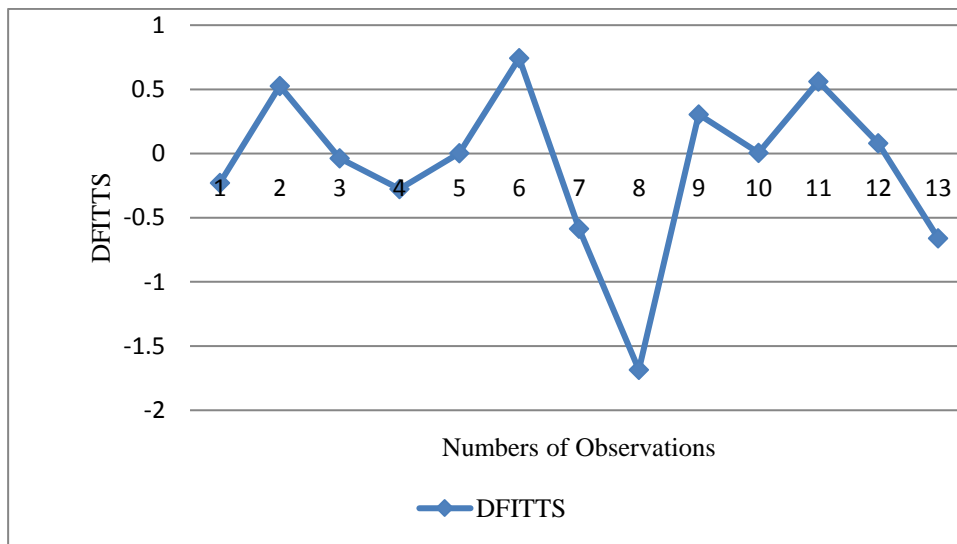


Figure 4: Plot of DFFITS based on TPE against its observations

5. Conclusion

The problems of multicollinearity and influential points have been jointly considered in this paper. A new diagnostic measures using Two-parameter Liu-Ridge Estimator (TPE) was proposed. The approximate case deletion formulas in TPE using Sherman-Morrison-Woodbury theorem by Rao (1973) were used to obtain the approximate versions of DFFITS, the two versions of COOK distance. The performance of these measures was illustrated with two real life dataset. The results show that the proposed measures compete favorably with the existing ones in identifying influential observations. Index plot adopted in this study is a conventional procedure to identified influential cases, even though, no conventional cut off points are introduced or developed for the TPE influence diagnostics. These measures will assist practitioners to decide whether to retain, remove or down sized influential points using robust estimators when identified in a study.

References

- [1] Belsley, D. A., Kuh, E. and Welsch, R.E. (1980). *Regression Diagnostics; Identifying Influence Data and Source of Collinearity*. Wiley, New York.
- [2] Belsley, D.A. (1991). *Conditioning Diagnostics: Collinearity and Weak Data in Regression*. New York: Wiley.
- [3] Cook, R.D. (1977) Detection of influential observations in linear regression. *Technometrics* 19: 15-18.
- [4] Cook, R. D. (1982) and Weisberg, S. (1980). Characterization of an empirical influence function for detecting influential cases in regression. *Technometrics*, 22, 495-508.
- [5] Cook, R. D. (1982) and Weisberg, S. (1982). *Residuals and Influence in Regression*. Chapman and Hall, New York.
- [6] Cook, R. D. (1986). Assessment of local influence (with discussion). *Journal of Royal Statistical Society, B*, 48, 133-69.
- [7] Chatterjee, S., Hadi, A. S. (1986). Influential observations, high leverage points, and outliers in linear regression. *Statistical Science*, 1, 379-416.
- [8] Draper, N.R. and John, J.A. (1981). Influential observations and outliers in regression. *Technometrics*, 23, 21-26.
- [9] Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: biased estimation for non-orthogonal problems. *Technometrics*, 12, 55-67.
- [10] Hoerl, A. E., Kennard, R. W. and Baldwin, K. F. (1975). Ridge regression: Some simulation. *Communications in Statistics*, 4 (2), 105–123.
- [11] Jahufer, A., and Jianbao, C. (2009). Assessing global influential observations in modified ridge regression. *Statistics & Probability Letters*, 79(4), 513-518.
- [12] Jahufer, A. and Chen, J. (2008). Identifying Local Influence in Modified Ridge Regression Using Cook's Method. *Sri Lankan Journal of Applied Statistics*, 9, 93-108.
- [13] Jahufer, A. (2013). Detecting Global Influential Observations in Liu Regression Model. *Open Journal of Statistics*, 3(1), 5-11.
- [14] Johnston, J. (1972). *Econometric Methods*, 2nd Ed. McGraw-Hill Book Co., Inc., New York.
- [15] Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics*, 22(2), 393-402.
- [16] 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.

-
- [17] Lukman, A. F. and Ayinde, K. (2015). Review and classification of the Ridge Parameter Estimation Techniques. Hacettepe Journal of Mathematics and Statistics. Accepted for Publication.
- [18] Mandel, J. (1982). Use of the singular value decomposition in regression analysis. *The American Statistician*, 36(1), 15-24.
- [19] Ozkale, M. R. and Kaciranlar, S., (2007). The restricted and unrestricted two-parameter estimators. *Commun. Statist. Theor. Meth.* 36, 2707–2725.
- [20] Rao, C. R. (1973). *Linear statistical inference and its applications*, 22: John Wiley & Sons.
- [21] Shi, L. (1997). Local influence in principal components analysis. *Biometrika*, 84(1), 175-186.
- [22] Shi, L., and Wang, X. (1999). Local influence in ridge regression. *Computational Statistics and Data Analysis*, 31(3), 341-353.
- [23] Ullah, M. A., Pasha, G. R., and Aslam, M. (2013). Assessing Influence on the Liu Estimates in Linear Regression Models. *Communications in Statistics - Theory and Methods*, 42(17), 3100-3116.
- [24] Yasin, A. and Murat, E. (2016). Influence Diagnostics in Two-Parameter Ridge Regression. *Journal of Data Science*, 14, 33-52.
- [25] Walker, E. and Birch, J. B. (1988). Influence Measures in Ridge Regression. *Technometrics*, 30(2), 221- 227.

