Local Influence in Constrained General Linear Models

Hadi Emami[1][*] and Mostafa Emami[1]

[1]Department of statistics, University of Zanjan

Abstract: Constrained general linear models (CGLMs) have wide applications in practice. Similar to other data analysis, the identification of influential observations that may be potential outliers is an important step beyond in CGLMs. We develop local influence approach for detecting influential observations in CGLMs. The procedure makes use of the normal curvature and the direction achieving the maximum curvature to assess the local influences of minor perturbation of CGLMs. An illustrative example with a real data set is also reported.

Key words: Constrained general linear models, Local influence, Perturbation scheme, Cook's distance

1. Introduction

Influence analysis and model diagnostics are well established for the linear regression model. The books of Belsley et al. (1989), Cook and Weis-berg (1982), Atkinson (1985) and Chatterjee and Hadi (1988) surveyed the field with applications to linear regression and other models. The study of influential observations has been extended to other statistical models using similar ideas to the ones developed in the linear regression. For example, Haslett and Haslet (2007) and Shi and Chen (2009) extended a generalized version of residuals with influential measures in general linear model, Banerjee and Frees (1997) and Tan et al. (2001) investigate the diagnostic problems in longitudinal models. Christensen et al (1992), Zewotir and Galpin (2005) and Norbe and Singer (2011) discussed the extension of influence diagnostic methods to mixed models and so on. Case deletion is a popular way to assess the individual impact of cases on the estimation process. Cook (1977) paper gave an impetus to the development of case deletion diagnostics for all sorts of statistical models. This approach is an example of a global influence analysis, namely, the effect of an observation is assessed by completely removing it. Cook (1986) paper dealt with a local approach, that is, each case is given a weight ω_i and the effect on the parameter estimation is measured by perturbing these weights around, say, $\omega_i = 1$. Choosing weights equal to zero or one corresponds to the global case-deletion approach. In our approach, a case will always correspond to an individual and hence to the whole vector of responses for that individual.

---

[*] Corresponding author.

Constrained general linear models (CGLMs) are used widely in the field of econometrics, for example in the estimation of Cobb-Douglas production functions Chipman and Rao (1964). However, diagnostics for assessing the influence of observations in CGLMs have not received the same attention. We found the method of local influence particularly useful for CGLMs. As in Cook (1986), the displacement in log-likelihood was taken as the metric to evaluate local influence. Kim (1998, 2007) worked out the local influence and case deletion diagnostic approach for linear hypothesis in multiple regression which is simpler in the sense that all random errors were assumed to be independent of each other. In this work we will suggest a method of detecting outliers, using local influence in linear model when there are some linear relationship among model coefficients and when random errors are not necessary independent. It measures the sensitivity of the analysis to a change in the model caused by a minor perturbation and has been known as a method of detecting outliers that avoids masking and swamping effect. The rest of paper is organized as follows. In section 2 constrained general linear models is described. In section 3 three necessary computational theorems for the local influence measure are provided. An illustrative example is given in section 4.

2. Constrained general linear models

Consider the general linear models

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

with

$$\mathbf{A}\boldsymbol{\beta} = c, \tag{2}$$

Where $y = (y_1, \ldots, y_n)'$ is an $n \times 1$ vector of observations, $\mathbf{X}$ is a $n \times p$ matrix of rank p whose ith row is $x_i' = (x_{i1}, \ldots, x_{ip})'$, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameter and ε is an $n \times 1$ vector of normality distributed random errors with $E(\varepsilon)=0$ and $Cov(\varepsilon)=\sigma^2\mathbf{V}$, where $\mathbf{V}$ is positive definite matrix, not necessary to be diagonal matrix and σ^2 is an unknown positive scalar. For the purpose of this article, similar to work of Shi and Chen (2009) we assume that $\mathbf{V}$ is known or can be replaced by a predetermined estimate.

In the second equation, $\mathbf{A}$ is a known $q \times p$ $(q \leq p)$ matrix of rank q, and c is a known $q \times 1$ vector. In fitting general linear model (1) without constraint (2) the least squares estimate of β can easily be obtained as $\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}\mathbf{V}^{-1}\mathbf{y}$. The $n \times 1$ vector of residuals is given by e $= \boldsymbol{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}$ and the residual sum of squares satisfies $s^2 = e'\mathbf{V}^{-1}e = y'(\mathbf{V}^{-1} - \mathbf{P})^{-1}y$, where $\mathbf{P} = \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$ The ML estimator of σ^2 can be obtained as $\sigma^2 = s^2/n$ (see Yan and Su (2009)). On the other hand the weighted least square of β under constraint (2) can be derived as

$$\widehat{\boldsymbol{\beta}}^* = \widehat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\widehat{\boldsymbol{\beta}} - \mathbf{c}), \tag{3}$$

which is also the maximum likelihood (ML) estimate of $\boldsymbol{\beta}$ under constraint (2) with normality assumption of errors.

Defining $e^* = (y - X\widehat{\beta}^*)$, we can write $s^{*2} = e^{*\prime}V^{-1}e^*$, therefore the maximum likelihood estimator of σ^2 under (2) will be $\hat{\sigma}^{*2} = s^{*2}/n$. The existence of the linear relationships among β can be checked by using the F-statistic given by $F = [(n-p)/q](\hat{\sigma}^{*2}/\hat{\sigma}^2 - 1)$ which is distributed as the F-distribution with q and $n-p$ degree of freedom when the linear-ship (2) hold. More detail can be found in chapter 7 of Yan and Su (2009).

3. Local influence procedure

In this section local influence measures for CGLMs will be derived. Let $\omega = (\omega_1, \ldots, \omega_n)'$ be an $n \times 1$ vector in Ω of R^n storing the case weights and Ω represents the set of relevant perturbations. For example, Ω may be the set defined by $0 \leq \omega_i \leq 1$ for $i = 1, \ldots, n$. We consider the perturbed model in which the $n \times 1$ vector ε is perturbed according to a normal distribution $N(0, \sigma^2 V_\omega)$, where V_ω is perturbed variance covariance matrix V such that lth diagonal element of V is perturbed according to $v_{ll}\omega$ for $l = 1, \ldots, n$. This perturbation scheme simultaneously perturbs all the cases. When the ω are set equal to one, the perturbed model reduces to unperturbed model.

Let Θ be the $(p+1) \times 1$ vector of parameters formed by stacking β and σ^2. The log-likelihoods for the unperturbed and perturbed models is denoted by $L(\Theta)$ and $L(\Theta \mid \omega)$ respectively. The classical log-likelihood $L(\Theta)$ is equal to $L(\Theta \mid \omega_0)$ with $\omega_0 = (1, \ldots, 1)^T$. Let $\widehat{\Theta}$ and Θ_ω be the maximum likelihood estimators of Θ under $L(\Theta)$ and $L(\Theta \mid \omega)$ respectively. Then their difference can be measured by the likelihood displacement function given by

$$LD(\omega) = 2[L(\widehat{\Theta} \mid \omega_0) - L(\widehat{\Theta} \mid \omega)].$$

This likelihood displacement, together with the case-weights perturbation scheme, can be considered as a natural generalization of Cook's distance (see Cook (1986) and Cook and Weisberg (1982)). At the point ω_0, achieves $LD(\omega)$ its minimum. A straight line in Ω passing through ω_0 is given by $\omega(l) = \omega_0 + al$, where a is scalar and l is a fixed column in vector in R^n. Cook (1986) suggested using the normal curvature C_l of the graph of the likelihood displacement function along a direction l at the optimal point ω_0 to examine the local behaviour of $LD(\omega)$ for assessing the local influence of the perturbation in the direction l. The direction vector $l$max associated with the largest curvature of curve at $a = 0$ provide information about outliers that cause a great change in the likelihood displacement. Observations corresponding to the component of l_{max} that has substantially larger magnitude than the others are potential outliers. Let

$$\dot{L}D = \frac{\partial LD(\omega)}{\partial \omega} \Big|_{\omega=0} \quad \text{and} \quad \ddot{L}D = \frac{\partial^2 LD(\omega)}{\partial \omega \partial \omega'} \Big|_{\omega=0},$$

for the CGLMs, $\partial L(\Theta)/\partial \Theta$ evaluated at $\Theta = \widehat{\Theta}$ is not general zero. However, it will be shown later that the first order derivative $\dot{L}D$ of likelihood displacements evaluated at $a = 0$ is zero. Thus the curvature is given by

$$C_l = \mid l^T \ddot{L}Dl \mid.$$

Then l_{max} is the eigenvector corresponding to the largest absolute eigen-value of $\ddot{L}D$. For deriving $\dot{L}D$ and $\ddot{L}D$, first we note that

$$\dot{L}D = -2\{\frac{\partial\hat{\Theta}_\omega}{\partial\omega}\mid_{a=0}\}\dot{L}, \tag{4}$$

$$\ddot{L}D = -2\{\frac{\partial\hat{\Theta}_\omega}{\partial\omega}\mid_{a=0}\}\ddot{L}\{\frac{\partial\hat{\Theta}_\omega}{\partial\omega}\mid_{a=0}\} - 2\sum_{i=1}^{p+1}\dot{L}_i\{\frac{\partial^2\hat{\Theta}_{i\omega}}{\partial\omega\partial\omega^T}\mid_{a=0}\}, \tag{5}$$

where $\dot{L} = \partial L(\Theta)/\partial\Theta$ evaluated at $\Theta = \hat{\Theta}$, $\ddot{L} = \partial^2 L(\Theta)/\partial\Theta\,\partial\Theta^T$ evaluated at $\Theta = \hat{\Theta}$, $\dot{L}_i$ is ith element of $\dot{L}$ and $\hat{\Theta}_{i\omega}$ is ith element of $\hat{\Theta}_\omega$.

The infinitesimal perturbation for the likelihood estimator of β under constraint (2) can be considered as

$$\hat{\beta}_\omega^* = \hat{\beta}_\omega - (X'V_\omega^{-1}X)^{-1}A'[A(X'V_\omega^{-1}X)^{-1}A']^{-1}(A\hat{\beta}_\omega - c), \tag{6}$$

where $\hat{\beta}_\omega = (X'V_\omega^{-1}X)^{-1}XV_\omega^{-1}y$. The following lemma is useful to computing the next theorems.

Lemma 3.1. *Consider the matrix of partial derivatives of* $(X'V_\omega^{-1}X)^{-1}$ *and denote by* ξ_r *the r-th column of identity matrix* I_n *then*

$$\frac{\partial(X'V_\omega^{-1}X)^{-1}}{\partial\omega_r}\mid_{a=0} = (X'V^{-1}X)^{-1}X'V^{-1}\xi_r\xi_r'V^{-1}X(X'V^{-1}X).$$

Proof. By perturbing rth diagonal elements of covariance matrix V we can write $V_\omega = V - \xi_r\left(v_{rr} - \frac{v_{rr}}{\omega_r}\right)\xi_r'$ so

$$V_\omega^{-1} = V^{-1} + \frac{V^{-1}\xi_r\xi_r'V^{-1}v_{rr}(\omega_r - 1)}{\omega_r - v_{rr}(\omega_r - 1)v^{rr}},$$

where $v^{rr} = \xi_r'V^{-1}\xi_r$ is rth diagonal element of V^{-1}. Therefore $\frac{\partial V^{-1}}{\partial\omega_r}\mid_{a=0} = V^{-1}\xi_r\xi_r'V^{-1}$. Using the equality $\partial F(\omega)^{-1}/\partial\omega_r = -F(\omega)^{-1}[\partial F(\omega)/\partial\omega_r]F(\omega)$ for the matrix $F(\omega)$ the proof will be complete.

Using lemma (3.1), after some algebraic calculation we easily get

$$\frac{\partial\hat{\beta}_\omega}{\partial\omega_r}\mid_{a=0} = (X'V^{-1}X)^{-1}X'V^{-1}W_rV^{-1}e, \tag{7}$$

$$\frac{\partial^2\hat{\beta}_\omega}{\partial\omega_r\omega_s}\mid_{a=0} = -(X'V^{-1}X)^{-1}X'V^{-1}W_rPW_sV^{-1}e$$
$$-(X'V^{-1}X)^{-1}X'V^{-1}W_sPW_rV^{-1}e, \tag{8}$$

where $W_i = \xi_i\xi_i'$, in the other words W_i is a diagonal matrix in which the i th element is one and the other elements is zero.

Theorem 3.1. *The partial derivatives of $\hat{\beta}_\omega^*$ can be derived as*

$$\frac{\partial \hat{\beta}_\omega^*}{\partial \omega_r} \mid_{a=0} = (\mathbf{I} - \mathbf{R})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{W}_r\mathbf{V}^{-1}e^*, \tag{9}$$

$$\frac{\partial^2 \hat{\beta}_\omega^*}{\partial \omega_r \omega_s} \mid_{a=0} = -(\mathbf{I} - \mathbf{R})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{W}_r(\mathbf{P} - \mathbf{P}_A)\mathbf{W}_s\mathbf{V}^{-1}e^*$$
$$-(\mathbf{I} - \mathbf{R})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{W}_s(\mathbf{P} - \mathbf{P}_A)\mathbf{W}_r\mathbf{V}^{-1}e^*, \tag{10}$$

where $\mathbf{R} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{A}]^{-1}\mathbf{A}$ and
$\mathbf{P}_A = \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{A}'$

Proof. Using lemma(3.1) and a little complicated algebra with relations (7) and (8) the proof will be complete.

Theorem 3.2. *The partial derivatives of $\hat{\sigma}_\omega^*$ respect to ω_r and ω_s is*

$$\frac{\partial \hat{\sigma}_\omega^*}{\partial \omega_r} \mid_{a=0} = \frac{1}{n} e^{*\prime} \mathbf{V}^{-1}\mathbf{W}_r\mathbf{V}^{-1}e^*, \tag{11}$$

$$\frac{\partial^2 \hat{\sigma}_\omega^*}{\partial \omega_r \, \partial \omega_s} \mid_{a=0} = \frac{1}{n} e^{*\prime} \mathbf{V}^{-1}[\mathbf{W}_r(\mathbf{P}_A - \mathbf{P})\mathbf{W}_s + \mathbf{W}_s(\mathbf{P}_A - \mathbf{P})\mathbf{W}_r]\mathbf{V}^{-1}e^*. \tag{12}$$

Proof. For the perturbed model, the maximum likelihood estimator of σ^2 is

$$\hat{\sigma_\omega^{*2}} = \frac{1}{n}(y - \mathbf{X}\hat{\beta}_\omega^*)'\mathbf{V}_\omega^{-1}(y - \mathbf{X}\hat{\beta}_\omega^*)$$
$$= \frac{1}{n}(y - \mathbf{X}\hat{\beta}_\omega)'\mathbf{V}_\omega^{-1}(y - \mathbf{X}\hat{\beta}_\omega)$$
$$+ \frac{1}{n}(\mathbf{A}\hat{\beta}_\omega - c)[\mathbf{A}(\mathbf{X}'\mathbf{V}_\omega^{-1}\mathbf{X})^{-1}\mathbf{A}]^{-1}(\mathbf{A}\hat{\beta}_\omega - c),$$

and using the previous relations (9) and (10) with further computations the proof will be complete.

The first order derivatives of the log-likelihood evaluated at $\Theta = \hat{\Theta}$ are easily obtained as

$$\frac{\partial L(\Theta)}{\partial \beta} \mid_{\Theta = \hat{\Theta}} = \frac{1}{\hat{\sigma}^2}\mathbf{A}'[\mathbf{A}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\beta} - c), \tag{13}$$

$$\frac{\partial L(\Theta)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{(y - \mathbf{X}\beta)'\mathbf{V}^{-1}(y - \mathbf{X}\beta)}{2\sigma^4} \mid_{\Theta = \hat{\Theta}} = 0. \tag{14}$$

Using (9), (11), (13) and (14) it follows from relation in (4) that $\dot{L}D = 0$

Theorem 3.3. *The (r, s) th element of $\ddot{L}D$ is computed as*

$$\ddot{L}D_{(r,s)} = \frac{2}{\hat{\sigma}^2}e^{*\prime}\mathbf{V}^{-1}\mathbf{W}_r\mathbf{V}^{-1}\mathbf{X}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}(\mathbf{I} - \mathbf{R})\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$$

$$(\mathbf{I} - \mathbf{R})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}\mathbf{V}^{-1}\mathbf{W}_s\mathbf{V}^{-1}e^* + \frac{[e^{*\prime}\mathbf{V}^{-1}\mathbf{W}_r\mathbf{V}^{-1}e^*]^2[e^{*\prime}\mathbf{V}^{-1}\mathbf{W}_s\mathbf{V}^{-1}e^*]^2}{n\hat{\sigma}^4}. \tag{15}$$

Proof. We need to compute the second order derivative of the log-likelihood evaluated at Θ = $\widehat{\Theta}$ that are given by

$$\frac{\partial^2 L(\Theta)}{\partial \beta \partial \beta} \Big|_{\Theta=\hat{\Theta}} = -\frac{1}{\hat{\sigma}^2} \mathbf{X}'\mathbf{V}^{-1}\mathbf{X},$$

$$\frac{\partial^2 L(\Theta)}{\partial \beta \partial \sigma^2} \Big|_{\Theta=\hat{\Theta}} = \frac{1}{\hat{\sigma}^4} \mathbf{A}'[\mathbf{A}(\mathbf{X}\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{A}]^{-1}(\mathbf{A}\hat{\beta} - c),$$

$$\frac{\partial^2 L(\Theta)}{\partial (\sigma^2)^2} \Big|_{\Theta=\hat{\Theta}} = -\frac{n}{2\hat{\sigma}^4}.$$

Using relations (10), (12), (13) and (14) the second term in (5) turns out to be zero and proof will be complete.

4. PBC data

The performance of the proposed method in previous sections is illustrated with an example of the real data set. We consider primary biliary cirrhosis (PBC) data for 312 patients. These data are described in Fleming and Harrington (1999), pp.359-371. The data consist of survival time, an indicator for censoring and 17 covariate variables. Fleming and Harrington (1991, pp.160-161) suggest that the proportional hazard model with covariates log(Bilirubin), age, log(Albumin), log(protime), log(Protime) and Edma is biologicaly reasonable for this data. The linear predictor

$$\beta_1 age + \beta_2 \log(Bilirubin) + \beta_3 \log(Albumin) + \beta_4 Edma + \beta_5 \log(Protime),$$

for the risk score is now considered. The normality assumption of residuals of the above logarithm transformation regression can be checked by Q.Q plot which presented in Figure1.(a). It is seen that the assumption of normality of residuals is not violated. Because the effect of Bilirubin and Albumin on risk score can be similar, so the following linear-ship for this linear predictor can be considered.

$$\beta_2 - \beta_3 = 0 \tag{16}$$

Since the log likelihood function $L(y, \eta)$ is available, where $\eta = (\eta_1, \ldots, \eta_N)$, and where $\eta_i = x_i \beta$. The Newton-Raphson iteration leads to the sequence of estimates $\hat{\beta}_{t+1} = \hat{\beta}_t - (\mathbf{X}'\mathbf{V}(\hat{\beta}_t)^{-1} \mathbf{X})^{-1}\mathbf{X}'\mathbf{S}(\hat{\beta}_t)$, where $\mathbf{V}(\hat{\beta}_t) = [\partial^2 \log\{L(y, \eta)\}/\partial \eta' \partial \eta]^{-1}$ and $\mathbf{S}(\beta_t) = \{\partial L(y, \eta)\}/\partial \eta$. Further simplification gives

$$\hat{\beta}_{t+1} = (\mathbf{X}'\mathbf{V}(\hat{\beta}_t)^{-1} \mathbf{X})^{-1} \mathbf{X})^{-1} \mathbf{V}(\hat{\beta}_t)^{-1} z(\hat{\beta}_t),$$

where $z(\beta_t) = \mathbf{X}\beta - \mathbf{V}(\beta_t)\mathbf{S}(\beta_t)$. The process is iterated until convergence at the maximum likelihood estimate $\hat{\beta}$, that is

$$\hat{\beta} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})\mathbf{X}'\mathbf{V}^{-1}z,$$

in which $\mathbf{V} = \mathbf{V}(\widehat{\boldsymbol{\beta}})$ and $z = z(\widehat{\boldsymbol{\beta}}_t)$. Based these informations the one step approximation for $\widehat{\boldsymbol{\beta}}_\omega$ is'

$$\widehat{\boldsymbol{\beta}}_\omega = \left(\mathbf{X}'\mathbf{V}_\omega^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}_\omega^{-1}\mathbf{z}. \tag{17}$$

Finally the one step approximation for $\widehat{\boldsymbol{\beta}}_\omega^*$ will be achieved substituting (17) in (6). By applying theorems of (3.1 -3.3) the matrix of $\ddot{LD}$ in expression (15) can be obtained by replacing y with z. In order to check the adequacy of the linear relationship in (16) for the PBC data, we use the extended usual F-test for hypothesis given in chapter 7 of (Yan and Su (2009)). The value of F-test statistic is 0.468 and the associated p-value is 0.5. Hence it is reasonable to conclude that this linear relationship is not rejected in the linear predictor model for this dataset. An index plot of the direction cosines in l_{max} for the linear model with $\beta_2 - \beta_3 = 0$ is provided in Figure 1.(b) from which it is seen that cases 81, 293, 107 and 253 are stand out. Hence these observations can be considered as outliers for the GLM model with constraint $\beta_2 - \beta_3 = 0$. The direction cosines for other observations are small enough. In the overall analyses of PBC data these observations have an influence on the model fit. Fleming and Harrington (1991) indicated that there was a data bases error in the pro-time of observation 107. This result can be supported by investigating the CGLM residuals, observations 253, 293, 81 and 107 have largest absolute residuals $|e_i^*|$, 7.62, 5.43, 4.19 and 1.86 respectively and the other observations have small residuals relative to this cases. Plot of standardized CGLM residuals is presented in Figure 2 and like Figure 1.(b) the results of this measure is similar.
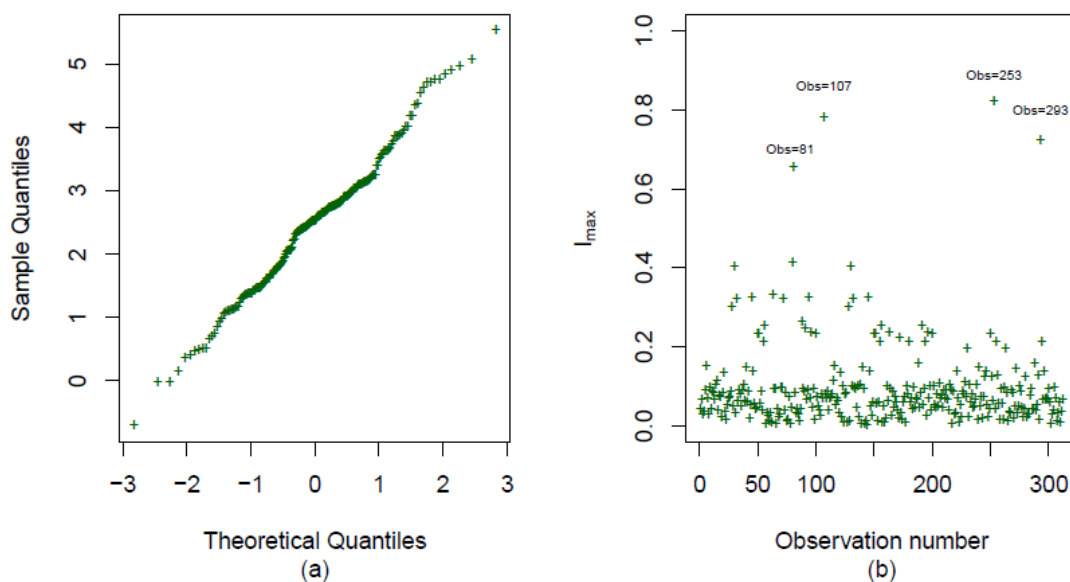

Figure 1: (a):Q-Q plot of residuals (b):Index plot of the direction cosines in l_{max} for PBC data
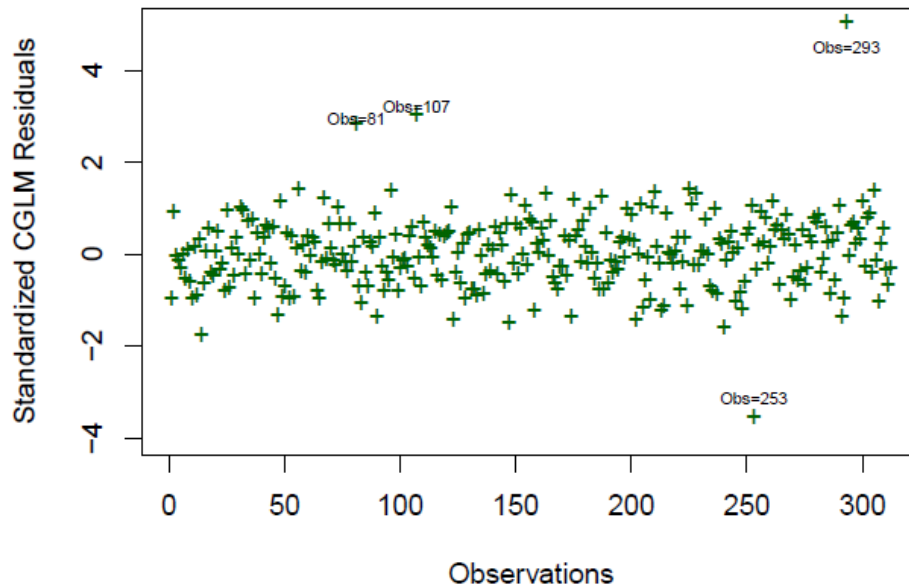
Figure 2: Plot of standardized CGLM residuals

5. Conclusions

No method of influence measures is available for constrained linear models with general covariance matrix of errors. Instead of using case deletion, this paper uses the local influence approach to study the detection of influential observations in CGLMs. We have shown how to extend and use likelihood placement and local influence to detect influential cases when one is faced with CGLMs. Local influence is a useful tool for identifying importantly influential cases and for assessing the effects that perturbations to the assumed data/model will have on inferences. The methods are computationally simple and the results are easy to displaying graphically. They give useful information about model and the data assumptions that needs extra scrutiny. A new feature of our approach is that we extended local influence diagnostics for parameters in CGLMs and derived a closed formulae which is able to detect anomalous observations in the data.

Acknowledgment

References

[1] Atkinson, A.C. (1985). *Plots, Transformations and Regression.* Oxford, U.K. Clarendon Press.

[2] Banerjee, M., and Frees, E.W. (1997). Influence diagnostics for linear longitudinal models. *J. American Statistical Association*., **92**, 999-1005.

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

[4] Chatterjee, S., and Hadi, A.S. (1988). *Sensitivity Analysis in Linear Regression.* New York: Wiley.

[5] Chipman, J. S. and Rao, M. M. (1964). The treatment of linear restriction in regression analysis. *Econometrica*., **32**, 198-209.

[6] Christensen, R., Pearson, L.M., and Johnson, W. (1992). Case deletion diagnostics for mixed models. *Technometrics*., **34**, 38-45.

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

[8] Cook, R. D. (1986). Assessment of local influence (with discussion). *J. Royal Statistical Society.*, **B48**, 133-169.

[9] Cook, R.D., and Weisberg, S. (1982). *Residuals and Influence in Regression.* New York: Chapman & Hall.

[10] Fleming, T.R., Harrington, D.P. (1991). *Counting Processes and Survival Analysis*, New York: Wiley.

[11] Haslett, J., and Haslett, S.J. (2007). The three basic types of residuals for a linear model. *J. International Statistical Review.*, **75**, 124

[12] Kim, M.G. (1998). Local influence on a test of linear hypothesis in multiple regression.models. *J. Applied Statistics*., **25,** 145152.

[13] Kim, M.G. (2007). Influence analysis of constrained linear models. *J. The Korean communication in statistics*., **14**, 281-286.

[14] Nobre, J.S., and Singer, J.M. (2011). Leverage analysis for linear mixed models. *J. Applied Statistics.*, **38**, 10631072.

[15] Shi, L., and Chen.G.(2009). Influence Measure for General Linear Models With Correlated Errors. *J. The American Statistician*., **63**(**1**),40-42.

[16] Tan, F.E.S., Ouwens, M.J.N., and Berger, M.P.F. (2001). Detection of influential observations in longitudinal mixed effects regression models. *J. Statistician*., **50**, 271-284.

[17] Yan, X., and Su, G.X. (2009). *Linear Regression Analysis.* World Scientific Publishing Co.Pte. Ltd.

[18] Zewotir,T., and Galpin, J.S. (2005). Influence diagnostics for linear mixed models. *J. Data Science* ., **3**,153-177.

Hadi Emami
Department of statistics, University of Zanjan
Third floor, University Blvd., 45371-38791, Zanjan, I. R. Iran
h.emami@znu.ac.ir

Mostafa Emami
Department of statistics, University of Zanjan
University Blvd., 45371-38791, Zanjan, I. R. Iran
Emamemostafa@yahoo.com