Ann. Inst. Statist. Math.
Vol. 52, No. 2, 367–379 (2000)

INFLUENCE ANALYSIS FOR LINEAR MEASUREMENT ERROR MODELS*

XU-PING ZHONG[1], BO-CHENG WEI[1] AND WING-KAM FUNG[2]

[1] *Department of Mathematics, Southeast University, Nanjing, 210096, China*
[2] *Department of Statistics, University of Hong Kong, Hong Kong, China*

(Received March 9, 1998; revised June 2, 1998)

Abstract. In this paper, we present a unified diagnostic method for linear measurement error models based upon the corrected likelihood of Nakamura (1990, *Biometrika*, **77**, 127–137). Both global influence and local influence are discussed. The case-deletion model and mean-shift outlier model are considered, and they are shown to be approximately equivalent. Several diagnostic measures are derived and discussed. It is found that they can be written in terms of the residual and leverage measure. Some existing results are improved. Numerical example illustrates that our method is useful for diagnosing influential observations.

Key words and phrases: Corrected likelihood, diagnostics, global influence, local influence, measurement error models.

1. Introduction

This paper deals with the assessment of influence for linear measurement error models, based upon the corrected likelihood (CL) defined by Nakamura (1990). Some review of measurement error models can be found in Anderson (1984), and Fuller (1980, 1987) for linear models, and in Carroll *et al.* (1995) for nonlinear models. To deal with these models, there are two approaches shown in the literature as Hanfelt and Liang ((1997), p. 628) pointed out. One approach, the corrected likelihood (Nakamura (1990)), successfully corrects for measurement errors in normal, Poisson, gamma and inverse Gaussian regression models. The method is easy to compute in practice. Another approach given by Stefanski and Carroll (1987) is more direct on an unbiased estimating function rather than on an approximate likelihood. On regression diagnostics for measurement error models, only little work has been done and they are all basically based on the second approach mentioned above. Kelly (1984) gave an influence function for the structural models, Fuller (1987) defined the hat matrix using the estimated predictor variable values, and Wellman and Gunst (1991) proposed an one-step approximation to Cook's distance. Zhao *et al.* (1994), and Zhao and Lee (1995) derived the influence functions for generalized linear and non-linear measurement error models. However, the results stated above are quite complicated in computations.

In this paper, we present a new and unified diagnostic method for linear measurement error models based upon the corrected likelihood of Nakamura (1990). Both global and local influence diagnostics are derived. In Section 2, we review the basis of the corrected likelihood defined by Nakamura (1990), and some related properties are discussed. By using the corrected likelihood, Section 3 deals with two diagnostic models:

* The research is supported by NSFC, NSFJ of China and a RGC grant of Hong Kong.

case-deletion model and mean-shift outlier model. Our results show that the two diagnostic models are approximately equivalent. We can approach to global influence via deletion approach, and several diagnostics are derived. Section 4 covers local influence analysis via perturbations of model or data. The diagnostic procedure based upon Cook (1986), and the scaled curvature defined by Schall and Dunne (1992) are well applied to the corrected likelihood. Moreover, the corrected scaled curvature is found to connect well with the diagnostics mentioned in Section 3. The diagnostics given in this paper are easier to compute than those of Kelly (1984), and Wellman and Gunst (1991). In Section 5, an illustrative example is given. Section 6 gives a brief discussion on the generalization of the methods.

2. Corrected likelihood of measurement error models

The linear measurement error models can be written as

$$Y = Z\beta + \varepsilon, \qquad \varepsilon \sim N(0, \sigma^2 I_n), \tag{2.1a}$$

$$X = Z + \delta, \qquad \delta \sim N(0, I_n \otimes \Lambda), \tag{2.1b}$$

where Y is an $n \times 1$ vector of observations y_i, Z is an $n \times p$ matrix with z_i^T as its i-th row, I_n is an $n \times n$ identity matrix, and σ^2 is the unknown common variance. X is also an $n \times p$ matrix with x_i^T as its i-th row, ε and δ are independent, and Λ is a positive definite matrix.

The measurement error model begins with an underlying model for the response Y in terms of the covariates. We distinguish between two kinds of covariates: X represents those covariates which for all practical purposes are measured with errors, while Z cannot be observed exactly for all studied subjects.

We now first review the corrected likelihood of Nakamura (1990) and then apply it to model (2.1).

Denote by $l(\beta, Z, Y)$, $U(\beta, Z, Y)$, and $J(\beta, Z, Y)$ the log-likelihood, score function, and observed information $-\partial U(\beta, Z, Y)/\partial\beta$ respectively, of β given Z and Y. The Fisher information of β given Z is denoted as $I(\beta, Z)$. Let β_0 be the true parameter value and E^+ denote the expectation with respect to the random vector Y, then

$$E^+\{U(\beta_0, Z, Y)\} = 0, \qquad E^+\{J(\beta, Z, Y)\} = I(\beta, Z). \tag{2.2}$$

When Z is subject to error and X is the observed value of Z, then $E^+\{U(\beta_0, X, Y)\} = 0$ does not necessarily hold and $\tilde{\beta}$ such that $U(\tilde{\beta}, X, Y) = 0$ is not necessarily consistent, in general. To correct this, Nakamura (1990) proposed a corrected likelihood $l^*(\beta, X, Y)$ which satisfies

$$E^*\{l^*(\beta, X, Y)\} = l(\beta, Z, Y), \tag{2.3}$$

where E^* denotes the conditional mean with respect to X given Z and Y, and parameter β is in an open convex subset $\mathcal{B}$. Let $U^*(\beta, X, Y) = \partial l^*(\beta, X, Y)/\partial\beta$ and $J^*(\beta, X, Y) = -\partial U^*(\beta, X, Y)/\partial\beta$ be the corrected score function and corrected observed information respectively. If E^* and $\partial\beta$ are interchangeable (Nakamura (1990)), then

$$E^*\{U^*(\beta, X, Y)\} = U(\beta, Z, Y), \qquad E^*\{J^*(\beta, X, Y)\} = J(\beta, Z, Y). \tag{2.4}$$

The value $\hat{\beta}$ such that $U^*(\hat{\beta}, X, Y) = 0$ with $J^*(\hat{\beta}, X, Y)$ being positive definite is called a corrected likelihood estimate (CLE) of β.

As pointed out by Nakamura (1990), the corrected likelihood l^* and the CL estimator $\hat{\beta}$ have nice properties (see also Hanfelt and Liang (1997)). Let $E = E^+E^*$ denote the global expectation, then it follows from (2.2) and (2.4) that

$$E\{U^*(\beta_0, X, Y)\} = E^+E^*\{U^*(\beta_0, X, Y)\} = E^+\{U(\beta_0, Z, Y)\} = 0.$$

This indicates that a corrected score function is an unbiased score function (Stefanski and Carroll (1987); Nakamura (1990)). In particular, as proved by Nakamura (1990), the estimator $\hat{\beta}$ has the properties of consistency and asymptotic normality, under certain regularity conditions, as expected. Moreover, from the asymptotic normality of $\hat{\beta}$, we have $LD(\beta) = 2\{l^*(\hat{\beta}) - l^*(\beta)\} \to \chi^2(p)$ (see, for example, Cox and Hinkley (1974)).

By the above description, our discussion on measurement error model (2.1) will be based on the CL $l^*(\beta, X, Y)$ and the corresponding CLE $\hat{\beta}$. For our model (2.1), it is easily seen that the log-likelihood $l(\beta, Z, Y)$ is

$$l(\beta, Z, Y) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum(y_i - z_i^T\beta)^2. \tag{2.5}$$

From this, we can get the corrected log-likelihood as

$$l^*(\beta, X, Y) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum\{(y_i - x_i^T\beta)^2 - \beta^T\Lambda\beta\}, \tag{2.6}$$

since $E^*\{l^*(\beta, X, Y)\} = l(\beta, Z, Y)$ (see also Nakamura (1990)).

Differentiating $l^*(\beta, X, Y)$ with respect to β, we have the corrected score function

$$\frac{\partial l^*}{\partial \beta} = \sigma^{-2}(X^TY - X^TX\beta + n\Lambda\beta). \tag{2.7}$$

Then $\hat{\beta}$, the CLE of β is obtained by solving equation $\partial l^*/\partial\beta = 0$, when the following matrix inversion is possible

$$\hat{\beta} = (X^TX - n\Lambda)^{-1}X^TY, \tag{2.8}$$

which coincides with the regression coefficient corrected for attenuation; see Fuller ((1987), p. 5).

From (2.8), the fitted value is $\hat{Y} = HY$, where $H = X(X^TX - n\Lambda)^{-1}X^T$ with entries h_{ij}, and the residual vector is $\hat{e} = Y - \hat{Y}$. The estimate of σ^2 is $\hat{\sigma}^2 = \|\hat{e}\|^2/(n-p)$. Further, from differentiating (2.7) we can get the corrected Fisher information matrix

$$J^*(\beta, X, Y) \triangleq \sigma^{-2}(X^TX - n\Lambda).$$

Let δ_k be a p-vector with 1 at the k-th position and zeros elsewhere, then $\hat{\beta}_k = \delta_k^T\hat{\beta}$ has the standard error s_k and the t-value $t_k = \hat{\beta}_k/s_k$, where $s_k^2 = \hat{\sigma}^2\delta_k^T(X^TX - n\Lambda)^{-1}\delta_k$. The joint $1-\alpha$ confidence region for parameter β is $\{\beta : (\hat{\beta}-\beta)^T(X^TX - n\Lambda)(\hat{\beta}-\beta) \le \hat{\sigma}^2F(p, n-p, \alpha)\}$, where $F(p, n-p, \alpha)$ denotes the upper α percentile of the Fisher's F-distribution with p and $n-p$ degrees of freedom.

3. Global influence analysis

This section will present several diagnostic measures based on corrected likelihood (2.6).

3.1 *On diagnostic models*

A fundamental approach of diagnostic analysis is based on the comparison of parameter estimates $\hat{\beta}$, $\hat{\sigma}^2$ with parameter estimates $\hat{\beta}_{(i)}$, $\hat{\sigma}^2_{(i)}$ that correspond to the so-called case deletion model (CDM) with the i-th case deleted:

$$y_j = z_j^T \beta + \epsilon_j, \quad j \neq i, \quad j = 1, \ldots, n. \tag{3.1}$$

LEMMA 3.1. *The estimates $\hat{\beta}_{(i)}$ in case deletion model* (3.1) *can be expressed as*

$$\hat{\beta} - \hat{\beta}_{(i)} \approx \frac{(X^T X - n\Lambda)^{-1} x_i \hat{e}_i}{1 - h_{ii}}, \tag{3.2}$$

where h_{ii} is the i-th diagonal entry of H, measuring the leverage for case i, and $\hat{e}_i = y_i - \hat{y}_i$ is the i-th residual.

This lemma is easily obtained by a few calculations from (2.6) (2.8) and (3.1). A similar result is given for linear models without measurement error by Cook and Weisberg ((1982), p. 110).

The case deletion model is the basis for constructing effective diagnostic statistics, and it is the most important one in practice because it is straightforward and easy to compute. Another commonly used diagnostic model is the mean-shift outlier model (MSOM, Cook and Weisberg (1982), p. 20). MSOM can be represented as

$$\begin{aligned} y_j &= z_j^T \beta + \epsilon_j, \quad \text{for} \quad j \neq i, \quad j = 1, \ldots, n, \\ y_i &= z_i^T \beta + \gamma + \epsilon_i, \end{aligned} \tag{3.3}$$

where γ is an extra parameter to indicate the presence of an outlier. It is easily seen that the nonzero value of γ implies that the i-th case may be an outlier because the case (x_i, z_i, y_i) no longer comes from the original model. This model is usually easier to formulate than the case deletion model. To detect outliers, one may either estimate the parameter γ or make a test of hypothesis $H_0 : \gamma = 0$, using MSOM which will be discussed later. The CLE of β, γ and σ^2 in (3.3) are denoted by $\hat{\beta}_{mi}$, $\hat{\gamma}_{mi}$ and $\hat{\sigma}^2_{mi}$, respectively.

CDM and MSOM are central models frequently used in regression diagnostics. It is well known that in linear regression models CDM and MSOM are equivalent in the following sense: the least square estimates (LSE) of parameters are equal under CDM and MSOM (Cook and Weisberg (1982)). However, the equivalence of CDM and MSOM on our models (2.1) have not been studied in the literature. It was Storer and Crowley (1985) who conjectured that $\hat{\beta}_{mi} = \hat{\beta}_{(i)}$ may hold in a broader class of models. Williams (1987) used the result $\hat{\beta}_{mi} = \hat{\beta}_{(i)}$ for generalized linear models, but he did not give the proof. Ross (1987) and Dzieciolowski and Ross (1990) used the result $\hat{\beta}_{mi} = \hat{\beta}_{(i)}$ implicitly for nonlinear regression models. Wei and Shih (1994) have solved this problem for many commonly encountered models, but not including measurement error models. Now we shall show that $\hat{\beta}_{mi} \approx \hat{\beta}_{(i)}$ holds for our models.

THEOREM 3.1. *Under the notation and definitions stated above, we have $\hat{\beta}_{mi} \approx \hat{\beta}_{(i)}$.*

PROOF. It follows from (2.6) that the corrected log-likelihood of MSOM is

$$l_{mi}^*(\beta, X, Y) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\left\{\sum_{j\neq i}(y_j - x_j^T\beta)^2 - (y_i - x_i^T\beta - \gamma)^2 - n\beta^T\Lambda\beta\right\}.$$

The CLE $\hat{\beta}_{mi}$ and $\hat{\gamma}_{mi}$ satisfy

$$\frac{\partial l_{mi}^*(\beta, X, Y)}{\partial\beta} = 0, \qquad \frac{\partial l_{mi}^*(\beta, X, Y)}{\partial\gamma} = 0,$$

which results in $\hat{\gamma}_{mi} = y_i - x_i^T\hat{\beta}_{mi}$ and

$$\hat{\beta}_{mi} = (X^TX - n\Lambda - x_ix_i^T)^{-1}(X^TY - x_iy_i),$$

$$\hat{\beta}_{mi} = \hat{\beta} - \frac{(X^TX - n\Lambda)^{-1}x_i\hat{e}_i}{1 - h_{ii}}. \tag{3.4}$$

From Lemma 3.1, we have $\hat{\beta}_{mi} \approx \hat{\beta}_{(i)}$.

3.2 *Influence diagnostics*

Once we get the corrected likelihood and its corresponding estimates, many diagnostics are immediate consequences, now we list them below.

3.2.1 *Basic diagnostics*

As in linear regression, the residual and the internally studentized residual are defined as

$$\hat{e}_i = y_i - x_i^T\hat{\beta}, \qquad \text{and}$$
$$r_i = \frac{\hat{e}_i}{\hat{\sigma}\sqrt{1 - h_{ii}}}.$$

It is easily seen that $h_{ii} = \partial\hat{y}_i/\partial y_i$. This means that h_{ii} measures the sensitivity of $\hat{y}_i$ with respect to y_i.

3.2.2 *Score test statistic of outlier*

We can get a score test statistic to detect the outlier based on the mean-shift outlier model. In fact, for model (3.3), one can make a test of hypotheses:

$$H_0 : \gamma = 0; \qquad H_1 : \gamma \neq 0.$$

If H_0 is rejected, then the i-th case is a possible outlier because this case may not come from the original model. We now derive a score test statistic for H_0.

THEOREM 3.2. *For MSOM, the score test statistic for the hypothesis $H_0 : \gamma = 0$ is given by*

$$SC_i = \frac{\hat{e}_i^2}{\hat{\sigma}^2(1 - h_{ii})} = r_i^2. \tag{3.5}$$

PROOF. Since CLE is asymptotically normal, the score test can be used (Cox and Hinkley (1974)). Let the corrected Fisher information matrix of Y for β and γ be $J_{(\beta,\gamma)}$, then the score statistic for H_0 is

$$SC_i = \left\{ \left(\frac{\partial l^*_{mi}}{\partial \gamma} \right)^T J^{\gamma\gamma} \left(\frac{\partial l^*_{mi}}{\partial \gamma} \right) \right\} \Bigg|_{(\hat{\beta},\hat{\sigma^2})},$$

where $J^{\gamma\gamma}$ is the lower right corner of $J^{-1}_{(\beta,\gamma)}$. It is easily seen that

$$\frac{\partial l^*_{mi}(\beta,\sigma)}{\partial \gamma} = \frac{1}{\sigma^2}(y_i - x_i^T \beta - \gamma),$$
$$J_{(\beta,\gamma)} = \frac{1}{\sigma^2} \begin{pmatrix} X^T X - n\Lambda & x_i \\ x_i^T & 1 \end{pmatrix},$$
$$J^{\gamma\gamma} = \sigma^2/(1 - h_{ii}),$$

then we can get the formula (3.5) from the above equations.

This theorem shows that the score statistic SC_i is just the square of studentized residual that is an adequate diagnostic statistic as often used in linear regression diagnostics.

3.2.3 *Generalized Cook distance*

The generalized Cook (1977) distance is the norm of $\hat{\beta} - \hat{\beta}_{(i)}$ with respect to certain weight matrix $M > 0$, i.e.

$$CD_i = \|\hat{\beta} - \hat{\beta}_{(i)}\|^2_M = (\hat{\beta} - \hat{\beta}_{(i)})^T M (\hat{\beta} - \hat{\beta}_{(i)}).$$

Choosing $M = J^*(\beta, X, Y) = \hat{\sigma}^{-2}(X^T X - n\Lambda)$, the corrected Fisher information matrix of Y for β, yields

$$CD_i = \frac{(\hat{\beta} - \hat{\beta}_{(i)})^T (X^T X - n\Lambda)(\hat{\beta} - \hat{\beta}_{(i)})}{\hat{\sigma}^2}.$$

By Lemma 3.1 we can get, approximately,

$$CD_i = \frac{h_{ii}}{1 - h_{ii}} r_i^2. \tag{3.6}$$

Rio (1988) defined $G_i^2(u)$ where

$$G_i(u) = \frac{x_i^T (X^T X - n\Lambda)^{-1} u}{\sqrt{u^T (X^T X - n\Lambda)^{-1} u} \cdot \sqrt{1 - h_{ii}}},$$

for some $p \times 1$ nonzero vector u, as the so-called u potential. He argued that $G_i^2(u)$ can be used to measure the influence of case i on the precision of the estimation of $u^T\beta$. Let $G_{ki} = G_i(\delta_k)$, and

$$CD_i^{(k)} = r_i^2 G_{ki}^2, \tag{3.7}$$

then case i can be considered a highly influential point if $CD_i^{(k)} > F(1, n-1, 0.1)$, since this case, if deleted, would move the estimate of $\hat{\beta}_k$ to the edge of the 90% confidence region, or if $CD_i > F(p, n-p, 0.1)$ for a similar reason. In the $CD_i^{(k)}$ measure, the factors G_{ki} is the δ_k-potential, which indicates the influence of case i on the precision of the estimate $\hat{\beta}_k$; whereas in the CD_i measure, the term $h_{ii}/(1-h_{ii})$, which is a monotone function of h_{ii}, measures the relative sensitivity at each data point.

As in linear regression, other diagnostics which have similar rationale as the Cook distance can be defined. One of such diagnostics is the Welsch and Kuh (1977) statistic given as

$$WK_i = \left(\frac{h_{ii}}{1-h_{ii}}\right)^{1/2} t_i,$$

where $t_i = \hat{e}_i/(\hat{\sigma}_{(i)}\sqrt{1-h_{ii}})$ is called the externally studenized residual.

3.2.4 *Likelihood distance*

The likelihood distance is defined as (Cook and Weisberg (1982), p. 183)

$$LD_i(\beta) = 2\{l^*(\hat{\beta}) - l^*(\hat{\beta}_{(i)})\}.$$

Taylor expansion of $l^*(\hat{\beta}_{(i)})$ at $\hat{\beta}$ gives

$$LD_i = 2\left[\dot{l}^{*T}(\hat{\beta})(\hat{\beta}-\hat{\beta}_{(i)}) + \frac{1}{2}(\hat{\beta}-\hat{\beta}_{(i)})^T\{-\ddot{l}^*(\hat{\beta})\}(\hat{\beta}-\hat{\beta}_{(i)})\right],$$

where $\dot{l}^*(\hat{\beta}) = U^*(\hat{\beta}, X, Y)$ and $-\ddot{l}^*(\hat{\beta}) = J^*(\hat{\beta}, X, Y)$. This result is exact because the third derivative is zero. Since $\dot{l}^*(\hat{\beta}) = 0$, we have

$$LD_i(\beta) = CD_i,$$

and the two diagnostic measures are identical.

We have introduced some basic diagnostics, the generalized Cook distance CD_i, the likelihood distance LD_i and score statistic SC_i based upon the corrected likelihood. It is observed that the diagnostic measures can be written in terms of the basic diagnostics. Assessment of the adequacy of model fit can be done by plotting residuals r_i versus various fitted values $\hat{y}_i$, or versus the values of some of the above diagnostics and so on. Any patterns in such plots tend to indicate possible model inadequacy.

4. Local influence measure

The local influence approach was presented by Cook (1986) and developed further by several authors (Thomas and Cook (1989), Escobar and Meeker (1992), Wu and Luo (1993), Wu and Wan (1994), Fung and Kwan (1997), and so on). In this section we first review the basic idea and formulas of local influence approach, and then apply them to linear measurement error models. Notice that our work is based upon the corrected likelihood (2.3).

4.1 *Cook's general procedure and its modification*

To study the sensitivity of uncertainties in the data or model, one can proceed by specifying a perturbation scheme through an $n \times 1$ vector ω with components ω_i attaching to case i. Here ω is admitted to vary over a neighborhood of ω_0, a null point at which the perturbed log-likelihood satisfies $l(\theta, \omega_0) = l(\theta)$. Assuming that given ω, $l(\theta, \omega)$ is maximized at $\hat{\theta}(\omega)$, while $l(\theta)$ is maximized at $\hat{\theta}$, then $\hat{\theta}(\omega_0) = \hat{\theta}$. From arguments of Cook (1986), the likelihood displacement surface is given in the form as

$$\alpha(\omega) = \begin{pmatrix} \omega \\ LD(\omega) \end{pmatrix}, \tag{4.1}$$

with $LD(\omega) = 2\{l(\hat{\theta}) - l(\hat{\theta}(\omega))\}$ as the solution locus. The normal curvature of $\alpha(\omega)$ at ω_0 along the line $\omega = \omega_0 + \tau d$ takes the simple form

$$C_d = 2d^T \ddot{F} d \quad \text{with} \quad \ddot{F} = \dot{T}^T(-\ddot{l})\dot{T}, \tag{4.2}$$

where d is an $n \times 1$ unit vector, and τ is a scale parameter, while $\dot{T}$ is a $p \times n$ matrix of first derivatives of $\hat{\theta}(\omega)$ with respect to ω, evaluated at ω_0. Cook (1986) argued that the maximum curvature direction can provide important diagnostic information.

A modification of Cook's local influence approach which is invariant over reparametrizations of the perturbation scheme was introduced by Schall and Dunne (1992). They suggested the scaled curvature which is defined as

$$C_d^s = \frac{d^T \ddot{F} d}{d^T G d}, \tag{4.3}$$

where $G = E(-\partial^2 l(\theta, \omega)/\partial\omega\partial\omega^T)$ is evaluated at ω_0. Based on a linear transformation to the scaled curvature, we have

$$C_d^I = d^T \ddot{A} d \quad \text{with} \quad \ddot{A} = G^{-1/2}\ddot{F}G^{-1/2}. \tag{4.4}$$

It is easily seen that C_d^s and C_d^I have the same maximum curvature direction.

Let $G^c = E\{-\partial^2 l(\hat{\theta}(\omega))/\partial\omega\partial\omega^T\}$ be evaluated at ω_0, then the corrected scaled curvature might be given as

$$C_d^{II} = \frac{d^T \ddot{F} d}{d^T G^c d}, \tag{4.5}$$

and it well connects local influence of commonly used perturbations with the influence diagnostics given in Section 3.

By the description stated in Section 2, we can naturally apply the corrected likelihood (2.6) to the Cook's (1986) theory with $LD^*(\omega) = 2\{l^*(\hat{\theta}) - l^*(\hat{\theta}(\omega))\}$.

4.2 *Mean shift perturbation*

One common way to describe the uncertainties in the mean is to perturb $X\beta$ to $X\beta + \omega$, which is identical with perturbing the vector of the observed responses in normal case. Assuming for simplicity that σ^2 is known or replaced by $\hat{\sigma}^2$, the relevant part of the perturbed corrected log-likelihood is

$$l^*(\beta, \omega) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}(\|Y - X\beta - \omega\|^2 - n\beta^T\Lambda\beta), \tag{4.6}$$

where $\omega_0 = 0$ yields the non-perturbed corrected log-likelihood. In the direction of the i-th unit vector of R^n, the curvature of the influence graph of ω derived from (4.2) is given by $C_i = h_{ii}/\hat{\sigma}^2$. The expected information is $G = \hat{\sigma}^{-2}I$, then the scaled curvature is given by $C_i^I = h_{ii}$. By a little calculation from (4.6), we obtain

$$l^*(\hat{\beta}(\omega)) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}(Y - \omega)^T(I - X(X^TX - n\Lambda)^{-1}X^T)(Y - \omega), \tag{4.7}$$

which results in $G^c = (I - H)/\hat{\sigma}^2$, then the corrected scaled curvature is

$$C_i^{II} = \frac{h_{ii}}{1 - h_{ii}}. \tag{4.8}$$

The scaled curvature associated with a mean-shift in the i-th observation turns out to be the leverage of the observation y_i, and the corrected scaled curvature, is expressed as a monotone function of h_{ii}. The diagnostics here emphasize the effect of case leverage.

4.3 *Case weights perturbation*

The case weights are often the basis for the study of influence; deleting a case is identical with attaching a zero weight to that case. Let ω denote an $n \times 1$ vector of case weights, then the perturbed corrected log-likelihood can be denoted as

$$l^*(\beta, \omega) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum\{\omega_i(y_i - x_i^T\beta)^2 - \beta^T\Lambda\beta\} + \frac{1}{2}\sum \log \omega_i, \tag{4.9}$$

where $\omega_0 = (1, 1, \ldots, 1)^T$ yields the non-perturbed corrected log-likelihood. In the direction of the i-th unit vector of R^n, the curvature of the influence graph of ω derived from (4.2) is given by $C_i = 2h_{ii}\hat{e}_i{}^2/\hat{\sigma}^2$. The expected information is $G = I_n/2$, then the scaled curvature is given by $C_i^I = 2C_i$. By a little calculation from equation (4.9), we get

$$l^*(\hat{\beta}(\omega)) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}Y^T\{W - WX(X^TWX - n\Lambda)^{-1}X^TW\}Y, \tag{4.10}$$

where $W = \text{diag}(\omega_1, \omega_2, \ldots, \omega_n)$, then the diagonal entry of G^c is $G_{ii}^c = 2h_{ii}(1 - h_{ii})$, and the corrected scaled curvature is

$$C_i^{II} = r_i^2. \tag{4.11}$$

This is just the square of the studentized residual. This result illustrates that the studentized residual and the case weights perturbation here have the same effect to detect the model fit.

4.4 *Perturbation of explanatory variables*

Consider perturbing the data for the k-th explanatory variable, by modifying the data matrix X as $X_\omega = X + \omega\delta_k^T$. In this situation, the perturbed corrected log-likelihood can be written as

$$l^*(\beta, \omega) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\{\|Y - X_\omega\beta\|^2 - n\beta^T\Lambda\beta\}, \tag{4.12}$$

where $\omega_0 = 0$ yields the non-perturbed corrected log-likelihood. In the direction of the i-th unit vector of R^n, the curvature of the influence graph of ω derived from (4.2) is given as

$$C_i = \frac{1}{\hat{\sigma}^2}\{\delta_k^T(X^TX - n\Lambda)^{-1}\delta_k\hat{e}_i^2 - 2\delta_k^T(X^TX - n\Lambda)^{-1}x_i\hat{\beta}_k\hat{e}_i + \hat{\beta}_k{}^2h_{ii}\}. \tag{4.13}$$

The expected information is $G = \hat{\sigma}^{-2}\hat{\beta}_k{}^2 I_n$, then the scaled curvature is given by $C_i^I = \hat{\beta}_k^{-2}C_i$. By a little calculation from equation (4.12), we get

$$l^*(\hat{\beta}(\omega)) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}Y^T\{I - X_\omega(X_\omega^TX_\omega - n\Lambda)^{-1}X_\omega^T\}Y. \tag{4.14}$$

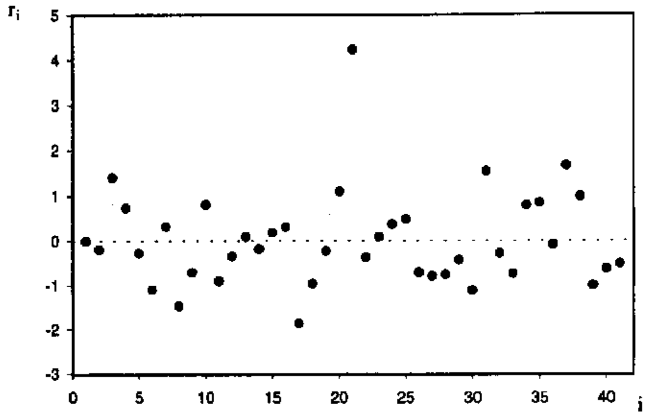

Fig. 1. Index plot of studentized residual r_i.

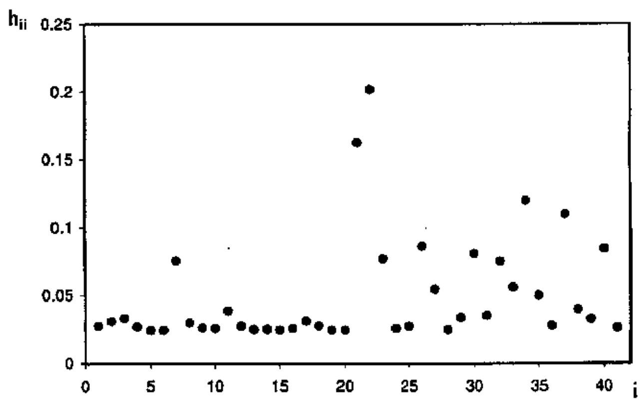

Fig. 2. Index plot of leverage measure h_{ii}.

From (4.14), we can obtain that the i-th diagonal entry of G^c is $G^c_{ii} = \hat{\sigma}^{-2}\hat{\beta}_k^2(1 - h_{ii})$ and then the corrected scaled curvature is

$$C_i^{II} = \frac{h_{ii}}{1 - h_{ii}} - 2t_k^{-1} r_i G_{ki} + t_k^{-2} r_i^2, \tag{4.15}$$

where t_i is defined in Section 2, and the result here associates with the studentized residual, the leverage and G_{ki} defined in Section 3, the potential measure.

5. Example

Concrete data. These data were given by Wellman and Gunst ((1991), p. 378), and we call them the concrete data for short. The data set contains comprehensive strength measurements of 41 samples of concrete. It was desired to use a linear regression model to predict comprehensive strength of concrete 28 days after pouring from the strength measurements taken two days after pouring. Wellman and Gunst (1991) used the unbiased score to discuss the linear measurement error models. Here we use the corrected likelihood to analyze the data.

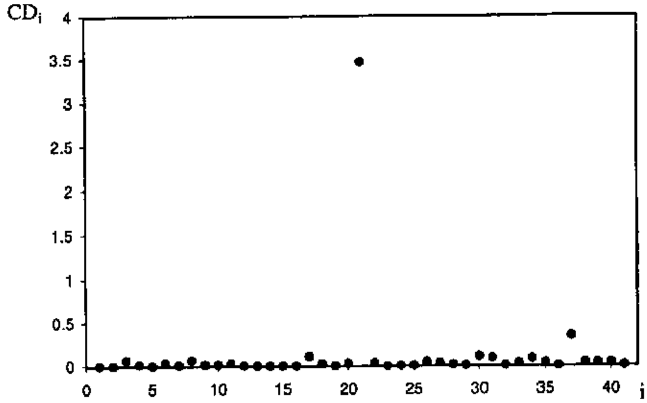

Fig. 3. Index plot of generalized Cook distance CD_i.

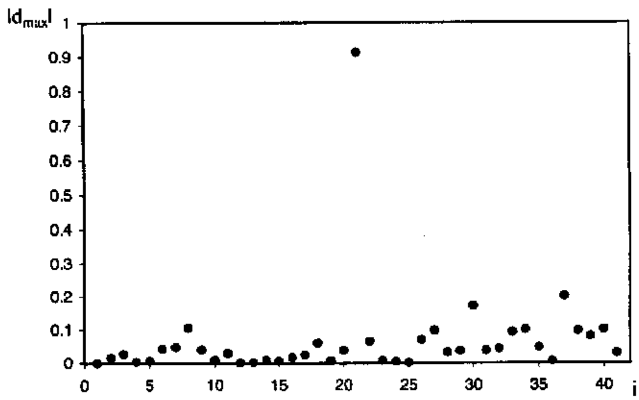

Fig. 4. Absolute value of direction $|d_{\max}|$ of maximum curvature under the case weight perturbation.

The statistical diagnostics of Sections 3 and 4 are constructed for the data set. Since many of the diagnostics can be written in terms of the basic statistics, studentized residual r_i and leverage measure h_{ii}, so they give similar influence information about the data set. For brevity, only the popular and more informative diagnostics are presented below. Figure 1 gives the index plot for the studentized residual r_i. Case 21 has a large r_i which indicates that it may be an outlying observation.

Figure 2 plots the leverage values of the data. The figure shows that cases 22 and 21 have the largest leverage. However, case 22 does not have a large residual as observed from Fig. 1. The effect of the observations to the estimates of regression coefficients is investigated by the generalized Cook distance which is given in Fig. 3. It is clear that case 21 has a much larger influence than the other cases.

The local influence of the observations is investigated. Figure 4 gives the absolute value of the direction $d_{\max}$ of the maximum curvature under the common case weight perturbation scheme. Again, case 21 is identified to be the most influential. The local influence of other perturbation schemes are also investigated, and they give very similar results to Fig. 4. They are omitted here.

6. Discussion

We have discussed the linear measurement error models and obtained several diagnostic measures based on the corrected likelihood. As Nakamura (1990) successfully corrected for measurement errors in Poisson, Gamma and Inverse Gaussian regression models, the method used in this paper can be also applied to these generalized linear models. By similar derivations from the corrected likelihoods given by Nakamura ((1990), pp. 131–133), both global and local influence diagnostics can be obtained for these models. For example, the score statistic for a Gamma regression model can be constructed as in Subsection 3.2, and it is given as

$$SC_i^G = \frac{(\hat{e}_i^G)^2}{\hat{\kappa}_i^G(1 - h_{ii}^G)}$$

where $\hat{\kappa}_i^G = y_i \exp\{x_i^T\hat{\beta} - (\hat{\beta}^T\Lambda\hat{\beta})/2\}$, $h_{ii}^G = y_i\hat{\kappa}_i^G(x_i - \Lambda\hat{\beta})^T I_G^*(x_i - \Lambda\hat{\beta})$, $\hat{e}_i^G = \hat{\phi} - y_i\hat{\kappa}_i^G$, and I_G^* is the corrected observed information matrix for the Gamma model. We shall discuss the diagnostics for these other models in detail in a separate paper.

Acknowledgement

The authors are grateful to a referee for comments which improved the presentation of the paper.

REFERENCES

Anderson, T. W. (1984). The 1982 Wald memorial lectures: Estimating linear statistical relation, *Ann. Statist.*, **12**, 1–45.

Carroll, R. J., Ruppert, D. and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models*, Chapman and Hall, London.

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

Cook, R. D. (1986). Assessment of local influence (with discussion), *J. Roy. Statist. Soc. Ser. B*, **48**, 133–169.

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

Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics*, Chapman and Hall, London.

Dzieciolowski, K. and Ross, W. H. (1990). Assessing case influence on confidence intervals in nonlinear regression, *Canad. J. Statist.*, **18**, 127–139.

Escobar, L. A and Meeker, W. Q. (1992). Assessing influence in regression analysis with censored data, *Biometrics*, **48**, 507–528.

Fuller, W. A. (1980). Properities of some estimators for the errors-in-variables models, *Ann. Statist.*, **8**, 407–422.

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

Fung, W. K. and Kwan, C. W. (1997). A note on local influence based on normal curvature, *J. Roy. Statist. Soc. Ser. B*, **59**, 839–843.

Hanfelt J. J. and Liang, K. Y. (1997). Approximate likelihood for generalized linear errors-in-variables models, *J. Roy. Statist. Soc. Ser. B*, **59**, 627–637.

Kelly, G. E. (1984). The influence function in the errors in variables problems, *Ann. Statist.*, **12**, 87–100.

Nakamura, T. (1990). Corrected score function for errors-in-variables models: Methodology and application to generalized linear models, *Biometrika*, **77**, 127–137.

Rio, M. (1988). On the potential in the estimation of linear functions in regression, *Commun. Statist. Theory Methods*, **17**, 729–738.

Ross, W. H. (1987). The geometry of case deletion and the assessment of influence in nonlinear regression, *Canad. J. Statist.*, **15**, 91–103.

Schall, R. and Dunne, T. T. (1992). A connection between local influence analysis and residual diagnostics, *Technometrics*, **33**, 103–104.

Stefanski, L. A. and Carroll, R. J. (1987). Conditional scores and optimal scores for generalized linear measurement error models, *Biometrika*, **74**, 703–716.

Storer, B. E. and Crowley, J. (1985). A diagnostic for Cox regression and general conditional likelihoods, *J. Amer. Statist. Assoc.*, **80**, 139–147.

Thomas, W. and Cook, R. D. (1989). Assessing influence on regression coefficients in generalized linear models, *Biometrika*, **76**, 741–749.

Wei, B. C. and Shi, J. Q. (1994). On statistical models in regression diagnostics, *Ann. Inst. Statist. Math.*, **46**, 267–278.

Wellman, J. M. and Gunst, R. F. (1991). Influence diagnostics for linear measurement errors models, *Biometrika*, **78**, 373–380.

Welsch, R. E. and Kuh, E. (1977). Linear regression diagnostics, Tech. Report, 923-77, Sloan School of Management, Massachusetts Institute of Technology.

Williams, D. A. (1987). Generalized linear model diagnostics using the deviance with single case deletion, *Applied Statistics*, **36**, 181–191.

Wu, X. and Luo, Z. (1993). Second-order approach to local influence, *J. Roy. Statist. Soc. Ser. B*, **55**, 929–936.

Wu, X. and Wan, F. H. (1994). A perturbation scheme for nonlinear models, *Statist. Probab. Lett.*, **20**, 197–202.

Zhao, Y. and Lee, A. H. (1995). Assessment of influence in non-linear measurement error models, *J. Appl. Statist.*, **22**, 215–225.

Zhao, Y., Lee, A. H. and Hui, Y. V. (1994). Influence diagnostics for generalized linear measurement error models, *Biometrics*, **50**, 1117–1128.