Local influence analysis based on the perturbation manifold in functional measurement error models

Patricia Giménez

Universidad Nacional de Mar del Plata, pcgimene@mdp.edu.ar

Abstract: The assessment of the effects of minor perturbations of data on corrected score estimators in functional measurement error models is considered by using a differential-geometrical framework proposed by Zhu et al. [*Ann. Statist.* **35** (2007) 2565-2588]. An n-dimensional Riemannian manifold, called the perturbation manifold is defined. The metric tensor can be used to choose an appropriate perturbation vector. First and second-order terms on a covariant version of the Taylor's theorem, based on the Levi-Civita connection, are used to define influence measures for the corrected score estimator. To illustrate the calculation of the geometrical quantities of interest, the simple linear measurement error model is examined.

Keywords: *local influence, Riemannian manifold, measurement error models, corrected score*

1 Introduction

Local influence analysis is an important statistical tool because it can provide indication of bad model fitting or of influential observations that could somewhat distort the parameter estimates leading in some cases to erroneous inference. The study of influence diagnostics has been an active area of statistical research since the seminal work of Cook ([6]), where a perturbation scheme is introduced into the postulated model through a perturbation vector, and the influence is studied via the normal curvatures on the graph of the likelihood displacement versus the perturbation vector. A generalization of Cook's approach and the influence on the maximum likelihood estimate of any parameter in a regression model is presented by [22]. Recently, [26] developed a differential-geometrical framework of a perturbation model (called the perturbation manifold). This method extends Cook's approach in several aspects. First, it is showed that the metric tensor of the perturbation manifold provides important information about selecting an appropriate perturbation of a model. Second, new influence measures are defined for smooth objective functions, that avoid the scale dependence of normal curvature for objective functions at points with a nonzero first derivative ([8]). In addition, the proposed second-order influence measures reduces to normal curvature under an appropriate perturbation scheme for objective functions that have zero first derivative at the critical point.

Influence diagnostic for measurement error models have received attention in the literature. Most works derive influence functions or apply the local influence method of [6], that is, the so-called first order approach ([22]). [11] gave an influence function for the structural models. [7] defined the hat matrix using the estimated predictor variable values and [21] proposed a one-step approximation to Cook's distance. [23] and [24] derived the influence functions for generalized linear and non-linear measurement error models. [13] obtained some useful diagnostics based on the likelihood displacement functions for generalized linear measurement models. [25] presented a unified diagnostic method for linear measurement error models based upon the corrected likelihood of [15]. [9] considered influence and diagnostic methods in homoscedastic comparative calibration models in functional and structural versions using Cook's approach based on the likelihood displacement. [18] considered the construction and properties of influence functions in the context of functional measurement error models with replicated data.

In functional measurement error models we are typically concerned with structural parameter estimation in the presence of incidental parameters. The failure of the likelihood approach for some models in such situations ([16], [17], [20]) has motivated researchers to seek for alternative methods of estimation. One of these is the corrected score approach ([15], [10]), which yields unbiased estimating equations independent of the incidental parameters. Under convenient regularity conditions, corrected score estimators are consistent and asymptotically normally distributed.

The aim of this paper is to assess the effects of minor perturbations of data on corrected score estimators in functional measurement error models. Following the approach of [26] we obtain the perturbation manifold for these models and the geometrical quantities associated for checking appropriate choice of a perturbation vector and calculating influence measures.

The paper is organized as follows. Section 2 presents the functional measurement error model and review estimation by using the corrected score approach. Section 3 considers local influence analysis. Different perturbation schemes on the corrected score function are included. The density of the perturbed model which yields the perturbed corrected score and the statistical perturbation manifold is obtained. The associated metric tensor and affine connection are calculated. In Section 4 first and second-order influence measures for the corrected score estimator are defined. Section 5 illustrates the calculation of the geometrical quantities of interest in the simple linear regression model. Numerical computations from a small data set is also included. Concluding remarks are made in Section 6.

2 FUNCTIONAL MEASUREMENT ERROR MODELS

A measurement error model is a linear or non-linear regression model with (substantial) measurement error in the variables, above all in the explanatory variable. Disregarding these measurement errors in estimating the regression parameters results in asymptotically biased, i.e. inconsistent estimators. This is the motivation for investigating measurement error models.

On the other hand, most studies in the life sciences, biology, ecology and economics involve variables that cannot be recorded exactly. In engineering, the calibration of measuring instruments deals with measurement errors by definition ([3]). Recently measurement error methods have been applied in the masking of data to assure anonymity ([2]). Many more examples and contribution to this field can be found in the literature, in particular in [7], [4] and [5].

Suppose that we wish to estimate a $p \times 1$ vector of parameters $\boldsymbol{\theta}$ in an open subset Θ of $\mathbb{R}^p$, governing the density function $p(\mathbf{y}; \mathbf{z}, \boldsymbol{\theta})$ of a $r \times 1$ random vector of responses $\mathbf{y}$, depending on a $k \times 1$ vector of covariates $\mathbf{z}$, unobservable because it is measured with error. Instead, we observe a surrogate $\mathbf{x} = \mathbf{z} + \mathbf{u}$, independent of $\mathbf{y}$, where the measurement error $\mathbf{u}$ is normally distributed with mean zero and covariance matrix Σ_u^2, which we suppose known.

Inference is based on a sample of n independent observations $(\mathbf{y}_1, \mathbf{x}_1), \ldots, (\mathbf{y}_n, \mathbf{x}_n)$. If the unobserved covariates $\mathbf{z}_1, \ldots, \mathbf{z}_n$ are unknown constants, then the model is referred to as a functional model and $\mathbf{z}_1, \ldots, \mathbf{z}_n$ are nuisance parameters whose number increases with the sample size, called incidental parameters. If $\mathbf{z}_1, \ldots, \mathbf{z}_n$ are considered as a random sample from some distribution, then the model is referred to as a structural model. The terminology "functional" and "structural" is due to [12]. In practice it is hard to decide which of these models is more relevant. In this paper we consider functional measurement error models with normal measurement error. The parameter $\boldsymbol{\theta}$ is the parameter of interest or structural parameter and the unobserved covariates $\mathbf{z}_j$, $j = 1, \ldots, n$ are incidental parameters pertaining to the observation $\mathbf{x}_j$, $j = 1, \ldots, n$. Let $p(\mathbf{x}_j; \mathbf{z}_j, \boldsymbol{\theta})$ denote the density function of $\mathbf{x}_j$ depending on $\mathbf{z}_j$.

Let $\mathbf{Y}$ be the $n \times r$ matrix with $\mathbf{y}_j^T$ as its j-th row, $\mathbf{X}$ and $\mathbf{Z}$ the $n \times k$ matrices with $\mathbf{x}_j^T$ and $\mathbf{z}_j^T$ as its j-th rows, respectively. The density of the postulated model is given by

$$p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}) = p(\mathbf{Y}; \mathbf{Z}, \boldsymbol{\theta}) p(\mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}) = \prod_{j=1}^{n} p(\mathbf{y}_j; \mathbf{z}_j, \boldsymbol{\theta}) \prod_{j=1}^{n} p(\mathbf{x}_j; \mathbf{z}_j, \boldsymbol{\theta}),$$

where $\mathbf{Z}$ is part of the parameters.

The log-likelihood, given $\mathbf{Y}$ and $\mathbf{X}$ is

$$\ell(\boldsymbol{\theta}, \mathbf{Z}; \mathbf{Y}, \mathbf{X}) = \sum_{j=1}^{n} \ell(\boldsymbol{\theta}, \mathbf{z}_j; \mathbf{y}_j) + \sum_{j=1}^{n} \ell(\boldsymbol{\theta}, \mathbf{z}_j; \mathbf{x}_j), \tag{1}$$

where $\ell(\boldsymbol{\theta}, \mathbf{z}_j; \mathbf{y}_j) = \log p(\mathbf{y}_j; \mathbf{z}_j, \boldsymbol{\theta})$ and $\ell(\boldsymbol{\theta}, \mathbf{z}_j; \mathbf{x}_j) = \log p(\mathbf{x}_j; \mathbf{z}_j, \boldsymbol{\theta})$.

2.1 ESTIMATION BY THE CORRECTED SCORE APPROACH

It is not generally true that maximizing (1) produces consistent estimators of $\boldsymbol{\theta}$ ([20]). The problem is due to the large number of nuisance parameters. The unwieldy functional likelihood and its failure to produce consistent estimators has motivated the search of alternative methods of estimation.

[15], [19] and [10] consider the use of corrected score functions in measurement error models. The approach depends on the existence of a function $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X})$, called a corrected score function, such that

$$\mathrm{E}[\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X})|\mathbf{Y},\mathbf{Z}] = \mathbf{U}(\boldsymbol{\theta};\mathbf{Y},\mathbf{Z}) \tag{2}$$

for all $\mathbf{Y},\ \mathbf{Z}$ and $\boldsymbol{\theta}$, where $\mathbf{U}(\boldsymbol{\theta};\mathbf{Y},\mathbf{Z}) = \dfrac{\partial}{\partial\boldsymbol{\theta}}\log p(\mathbf{Y};\mathbf{Z},\boldsymbol{\theta})$ is the unobserved score function, that is, the usual score if there were no measurement error. From (2), with the help of the iterative expectation principle and the fact that the unobserved score is unbiased, $\mathbf{U}^*$ can be seen as an unbiased estimating function, and so, under appropriate regularity conditions ([10]), it exists $\hat{\boldsymbol{\theta}}$ solving

$$\mathbf{U}^*(\hat{\boldsymbol{\theta}};\mathbf{Y},\mathbf{X}) = \sum_{j=1}^{n}\mathbf{U}^*(\hat{\boldsymbol{\theta}};\mathbf{y}_j,\mathbf{x}_j) = \mathbf{0}$$

which is a consistent and asymptotically normal estimator, called the corrected score estimator.

The corrected score method effectively estimates the estimator one would use if there were no measurement error. The corrected score function is independent of the incidental parameters $\mathbf{Z}$, so one can directly find estimators of the parameters of interest $\boldsymbol{\theta}$ avoiding the problem of estimating the incidental parameters $\mathbf{Z}$. However, corrected score function do not always exists. Existence depends critically on the assumed normality of the measurement error. [15] derived corrected score for some common generalized linear models.

3 Local influence analysis

We are interested on the assessment of effects of minor perturbations of data on the corrected score estimator of $\boldsymbol{\theta}$.

Let $\boldsymbol{\omega} = (\omega_1,\ldots,\omega_n)^T$ be a perturbation vector, $\boldsymbol{\omega}\in\Omega\subset\mathbb{R}^n$, which is introduced to perturb $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X})$. If $\boldsymbol{\omega}$ has a large effect, then it is important to know the cause (e.g. influential observations or invalid model assumptions) of such large effect.

Let $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega})$ denote the perturbed corrected score and $\boldsymbol{\omega}^0$ a vector representing no perturbation, that is, $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}^0) = \mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X})$. The perturbed corrected score estimator $\hat{\boldsymbol{\theta}}(\boldsymbol{\omega})$ solves the equation $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}) = \mathbf{0}$. We study each component of the vector $\hat{\boldsymbol{\theta}}(\boldsymbol{\omega})$ separately. As pointed out by [22], this approach is sometimes more informative than studying mixed effects, which, from different sources, may cancel out each other. In the following, $\hat{\theta}(\boldsymbol{\omega})$ denotes a particular component of the vector $\hat{\boldsymbol{\theta}}(\boldsymbol{\omega})$.

We need to know how the perturbation $\boldsymbol{\omega}$ affects the postulated model. This implies to find the density of the perturbed model $p(\mathbf{Y},\mathbf{X};\mathbf{Z},\boldsymbol{\theta},\boldsymbol{\omega})$ such that $\int p(\mathbf{Y},\mathbf{X};\mathbf{Z},\boldsymbol{\theta},\boldsymbol{\omega})\ d\mathbf{Y}\ d\mathbf{X} = 1$, from which we can obtain $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega})$, having $p(\mathbf{Y},\mathbf{X};\mathbf{Z},\boldsymbol{\theta},\boldsymbol{\omega}^0) = p(\mathbf{Y},\mathbf{X};\mathbf{Z},\boldsymbol{\theta})$.

Note that we begin perturbing the corrected score, which is independent of the incidental parameters $\mathbf{Z}$, while the density of the postulated model includes them. If we were interested in the assessment of the local influence on some objective function based on maximum likelihood approach (e.g. maximum likelihood estimator, likelihood displacement, etc.) we first would consider perturbation of the log-likelihood function of the postulated model (1) and then obtaining the density of the perturbated model would be direct.

Following [26], the perturbed model $p(\mathbf{Y},\mathbf{X};\mathbf{Z},\boldsymbol{\theta},\boldsymbol{\omega})$, characterized by a set of perturbations $\boldsymbol{\omega}$ can be regarded as an n-dimensional manifold M. Considering $\hat{\theta}(\boldsymbol{\omega})\colon\mathbb{R}^n\to\mathbb{R}$ as the objective function and $\boldsymbol{\omega}(t)$ a smooth curve on M, first and second-order terms from a Taylor expansion of $\hat{\boldsymbol{\theta}}(\boldsymbol{\omega}(t))$ are used to define influence measures.

3.1 Perturbed corrected score and perturbed log- likelihood of the model

Let $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega})$ be the perturbed corrected score and

$$\ell(\boldsymbol{\theta},\mathbf{Z};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}) = \ell(\boldsymbol{\theta},\mathbf{Z};\mathbf{Y},\boldsymbol{\omega}) + \ell(\boldsymbol{\theta},\mathbf{Z};\mathbf{X},\boldsymbol{\omega}) \tag{3}$$

be the perturbed log-likelihood of the model.

Perturbation $\boldsymbol{\omega}$ introduced in $\mathbf{U}^*$ primarily affects the first term on the right hand of (3) since

$$\mathrm{E}[\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega})|\mathbf{Y},\mathbf{Z},\boldsymbol{\omega}] = \mathbf{U}(\boldsymbol{\theta};\mathbf{Y},\mathbf{Z},\boldsymbol{\omega}) = \frac{\partial}{\partial\boldsymbol{\theta}}\ell(\boldsymbol{\theta},\mathbf{Z};\mathbf{Y},\boldsymbol{\omega}),$$

where $\mathbf{U}(\boldsymbol{\theta};\mathbf{Y},\mathbf{Z},\boldsymbol{\omega})$ is the perturbed unobserved score.

We consider some commonly used perturbation schemes on corrected score and then find the perturbed log-likelihood of the model.

1. *Case weights perturbation*

   The case weights are often the basis for the study of influence. We define an $n \times 1$ vector of weights $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^T$ to perturb the contribution of each case to the corrected score $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X})$, resulting in the perturbed corrected score

$$\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}) = \sum_{j=1}^{n} \omega_j \mathbf{U}^*(\boldsymbol{\theta};\mathbf{y}_j,\mathbf{x}_j), \tag{4}$$

   which generalizes the inclusion ($\omega_j = 1$) or the exclusion ($\omega_j = 0$) of an observation from the estimation of $\boldsymbol{\theta}$, so that this device enables us to learn about the relative importance of the observation to the estimation process.

   In this case $\boldsymbol{\omega}^0 = \mathbf{1}_n$, where $\mathbf{1}_n$ is the $n \times 1$ vector of ones.

   It can be seen easily that (4) can be obtained from

$$\ell(\boldsymbol{\theta},\mathbf{Z};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}) = \sum_{j=1}^{n} \omega_j \ell(\boldsymbol{\theta},\mathbf{z}_j;\mathbf{y}_j) + \sum_{j=1}^{n} \omega_j \ell(\boldsymbol{\theta},\mathbf{z}_j;\mathbf{x}_j).$$

2. *Perturbation of the observed covariate* $\mathbf{x}$

   Consider perturbing the data for the i-th observed covariate, by modifying $\mathbf{x}_j$ as

$$\mathbf{x}_j(\omega_j) = \mathbf{x}_j + \omega_j \boldsymbol{\delta}_i, \ \ j = 1, \ldots n,$$

   where $\boldsymbol{\delta}_i$ is an $k \times 1$ vector with 1 at the i-th position and zeros elsewhere.

   The perturbed corrected score can be written as

$$\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}) = \sum_{j=1}^{n} \mathbf{U}^*(\boldsymbol{\theta};\mathbf{y}_j,\mathbf{x}_j + \omega_j\boldsymbol{\delta}_i).$$

   Here $\boldsymbol{\omega}^0 = \mathbf{0}_n$, where $\mathbf{0}_n$ is the $n \times 1$ vector of zeros.

   In this case

$$\ell(\boldsymbol{\theta},\mathbf{Z};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}) = \sum_{j=1}^{n} \ell(\boldsymbol{\theta},\mathbf{z}_j + \omega_j\boldsymbol{\delta}_i;\mathbf{y}_j) + \sum_{j=1}^{n} \ell(\boldsymbol{\theta},\mathbf{z}_j + \omega_j\boldsymbol{\delta}_i;\mathbf{x}_j).$$

   Note that perturbing the observed covariate $\mathbf{x}_j$ as $\mathbf{x}_j(\omega_j) = \mathbf{x}_j + \omega_j\boldsymbol{\delta}_i$ on the corrected score is equivalent to perturbing the unobserved covariate $\mathbf{z}_j$ as $\mathbf{z}_j(\omega_j) = \mathbf{z}_j + \omega_j\boldsymbol{\delta}_i$ on the log-likelihood since $\mathbf{x}_j(\omega_j) = \mathbf{z}_j(\omega_j) + \mathbf{u}_j$ and

$$\mathrm{E}[\mathbf{U}^*(\boldsymbol{\theta};\mathbf{y}_j,\mathbf{x}_j(\omega_j))|\mathbf{y}_j,\mathbf{z}_j,\omega_j] = \mathbf{U}(\boldsymbol{\theta};\mathbf{y}_j,\mathbf{z}_j(\omega_j)) = \frac{\partial}{\partial\boldsymbol{\theta}}\ell(\boldsymbol{\theta},\mathbf{z}_j(\omega_j);\mathbf{y}_j).$$

3. *Perturbation of the response variable* $\mathbf{y}$

   We perturb the data for the i-th response variable, leading to

$$\mathbf{y}_j(\omega_j) = \mathbf{y}_j + \omega_j \boldsymbol{\delta}_i, \quad j = 1, \ldots n,$$

   where $\boldsymbol{\delta}_i$ is an $r \times 1$ vector with 1 at the i-th position and zeros elsewhere. The perturbed corrected score is given by

$$\mathbf{U}^*(\boldsymbol{\theta}; \mathbf{Y}, \mathbf{X}, \boldsymbol{\omega}) = \sum_{j=1}^{n} \mathbf{U}^*(\boldsymbol{\theta}; \mathbf{y}_j + \omega_j \boldsymbol{\delta}_i, \mathbf{x}_j)$$

   and the perturbed log-likelihood can be written as

$$\ell(\boldsymbol{\theta}, \mathbf{Z}; \mathbf{Y}, \mathbf{X}, \boldsymbol{\omega}) = \sum_{j=1}^{n} \ell(\boldsymbol{\theta}, \mathbf{z}_j; \mathbf{y}_j + \omega_j \boldsymbol{\delta}_i) + \sum_{j=1}^{n} \ell(\boldsymbol{\theta}, \mathbf{z}_j, \mathbf{x}_j),$$

   where $\boldsymbol{\omega}^0 = \mathbf{0}_n$.

   Moreover, some perturbation on model assumptions could be introduced, for instance, to consider an heterogeneous variance of $\mathbf{Y}$.

3.2 PERTURBATION MANIFOLD

To assess the local influence of a data perturbation, we are primarily interested in the behavior of the density of the perturbed model $p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\omega})$ as a function of $\boldsymbol{\omega}$ around $\boldsymbol{\omega}^0$. Here the parameters $\boldsymbol{\theta}$ and $\mathbf{Z}$ are assumed to be known or be fixed at a given value.

Given $\ell(\boldsymbol{\theta}, \mathbf{Z}; \mathbf{Y}, \mathbf{X}, \boldsymbol{\omega})$, obtained in Section 3.1, the density of the perturbed model $p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\omega})$ can be written as

$$p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\omega}) = \prod_{j=1}^{n} \{\exp\{\ell(\boldsymbol{\theta}, \mathbf{z}_j; \mathbf{y}_j, \mathbf{x}_j, \omega_j)\}[c_j(\boldsymbol{\theta}, \mathbf{z}_j, \omega_j)]^{-1}\},$$

where

$$c_j(\boldsymbol{\theta}, \mathbf{z}_j, \omega_j) = \int \exp\{\ell(\boldsymbol{\theta}, \mathbf{z}_j; \mathbf{y}_j, \mathbf{x}_j, \omega_j)\} d\mathbf{y}_j \, d\mathbf{x}_j.$$

Moreover, we assume that $p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\omega})$ satisfies the following regularity conditions, considered on page 16 of [1]:

1. All the $p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\omega})$'s have a common support, so that $p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\omega}) > 0$ for all $(\mathbf{Y}, \mathbf{X})$ in the support.

2. Let $\ell(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta}) = \log p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\omega})$. For every fixed $\boldsymbol{\omega}$, n functions in $(\mathbf{Y}, \mathbf{X})$

$$\frac{\partial}{\partial \omega_j} \ell(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta}), \quad j = 1, \ldots, n$$

   are linearly independent.

3. The moments of random variables $\dfrac{\partial}{\partial \omega_j} \ell(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta})$ exist up to necessary orders.

4. The partial derivatives $\frac{\partial}{\partial \omega_j}$ and the integration with respect to the Lebesgue measure λ can always be interchanged as

$$\frac{\partial}{\partial \omega_j} \int f(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta}) \, d\lambda = \int \frac{\partial}{\partial \omega_j} f(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta}) \, d\lambda$$

   for any function $f(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta})$ that we treat in the following.

The perturbed model

$$M = \{p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\omega}) : \ \boldsymbol{\omega} \in \Omega\}$$

can be regarded as an n-dimensional manifold (see [1]).

When a coordinate system $\boldsymbol{\omega}$ is given, $\partial_j = \partial/\partial\omega_j,\ (j = 1, \ldots, n)$ are the natural basis of the tangent space T_ω at point $\boldsymbol{\omega}$ of the manifold M, associated with the coordinate system. But, there is a more familiar representation of the tangent space in the case of the manifold M of a statistical model, that is, $T_\omega^{(1)}$, the so called 1-representation of the tangent space of M at $\boldsymbol{\omega}$, which is spanned by n functions $\partial_j \ell(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta})$. We can identify T_ω with $T_\omega^{(1)}$ considering that

$$\mathbf{h} = \sum_{j=1}^{n} h^j \partial_j \in T_\omega \ \leftrightarrow \ \mathbf{h}(\mathbf{Y}, \mathbf{X}) = \sum_{j=1}^{n} h^j \partial_j \ell(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta}) \in T_\omega^{(1)}$$

(see [1]).

3.2.1 Metric tensor and appropriate perturbation

The inner product of two basis operators ∂_i and ∂_j is

$$g_{ij}(\boldsymbol{\omega}) = \mathrm{E}_\omega[\partial_i \ell(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta}) \partial_j \ell(\boldsymbol{\omega}; \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \boldsymbol{\theta})],$$

where E_ω denotes the expectation taken with respect to $p(\mathbf{Y}, \mathbf{X}; \mathbf{Z}, \boldsymbol{\theta}, \boldsymbol{\omega})$. The n^2 quantities $g_{ij}(\boldsymbol{\omega}),\ i, j = 1 \ldots n$, form the metric tensor ([26]).

The metric matrix $G(\boldsymbol{\omega}) = (g_{ij}(\boldsymbol{\omega}))$ is the Fisher information matrix with respect to the perturbation vector $\boldsymbol{\omega}$. The element $g_{jj}(\boldsymbol{\omega})$ indicates the amount of perturbation introduced by the j-th component of $\boldsymbol{\omega}$. The elements $g_{ij}(\boldsymbol{\omega}),\ i \neq j$ represent the association between different components of $\boldsymbol{\omega}$.

The authors of [26] define, based on these observations, an appropriate perturbation as that satisfying that $G(\boldsymbol{\omega}^0) = \mathrm{diag}(g_{11}(\boldsymbol{\omega}^0), \ldots, g_{nn}(\boldsymbol{\omega}^0))$. This condition avoids any redundant components of $\boldsymbol{\omega}$ and determines the ortogonality between the different components of $\boldsymbol{\omega}$, to ensure that we can easily pinpoint the cause of a large effect. Moreover, we can always choose a new perturbation vector $\tilde{\boldsymbol{\omega}}$ such that $G(\tilde{\boldsymbol{\omega}})$ evaluated at $\boldsymbol{\omega}^0$ equals $c\mathbf{I}_n$, where $c > 0$.

(M, G) defines a Riemannian manifold, called a statistical perturbation manifold.

3.2.2 Affine connection

If we want to talk about the straightness (and hence curvature) of a smooth curve $\boldsymbol{\omega}(t)$ in M, an affine connection must be introduced. The metric tensor defines the Levi-Civita connection by its Christoffel symbols

$$\Gamma_{ijk}(\boldsymbol{\omega}) = \frac{1}{2}[\partial_i g_{jk}(\boldsymbol{\omega}) + \partial_j g_{ik}(\boldsymbol{\omega}) - \partial_k g_{ij}(\boldsymbol{\omega})].$$

It has the property that the geodesics of the Levi-Civita connection are curves of minimum length among those paths that lie in the manifold.

Of course there are connections which are not metrics or derived from metrics in this way. [26] introduce a covariant 3-tensor, symmetric in all indices, and a related family of affine connections Γ^α for any $\alpha \in \mathbb{R}$ are defined therefrom. We restrict attention to the Levi-Civita connection Γ $(\alpha = 0)$.

4 Influence measures for the corrected score estimator

Let $\hat{\theta}(\boldsymbol{\omega})\colon \mathbb{R}^q \to \mathbb{R}$ be a particular component of the perturbed corrected score estimator $\hat{\boldsymbol{\theta}}(\boldsymbol{\omega})$ and $\boldsymbol{\omega}(t)$ the geodesic, which is unique and defined in an interval containing 0 such that $\boldsymbol{\omega}(0) = \boldsymbol{\omega}^0$ and $d\boldsymbol{\omega}(t)/dt|_{t=0} = \mathbf{h} \in T_{\omega^0}$. [14] state a covariant version of the Taylor theorem, which has the advantage that each term in the series is a tensor, thus invariant to reparametrization, unlike the standard Taylor's series expansion, as follows

$$\hat{\theta}(\boldsymbol{\omega}(t)) = \hat{\theta}(\boldsymbol{\omega}^0) + t\nabla_{\hat{\theta}}^T \mathbf{h} + \frac{1}{2} t^2 \mathbf{h}^T \tilde{H}_{\hat{\theta}} \mathbf{h} + o(t^2),$$

with $\nabla_{\hat{\theta}} = \frac{\partial\hat{\theta}(\boldsymbol{\omega}^0)}{\partial\boldsymbol{\omega}}$ and $\tilde{H}_{\hat{\theta}} = \tilde{H}_{\hat{\theta}(\omega^0)}$, where the (i,j)-th element of $\tilde{H}_{\hat{\theta}(\omega)}$ is given by

$$[\tilde{H}_{\hat{\theta}(\omega)}]_{(i,j)} = \partial_i\partial_j\hat{\theta}(\boldsymbol{\omega}) - \sum_{s,r} g^{s,r}(\boldsymbol{\omega})\Gamma_{ijs}(\boldsymbol{\omega})\partial_r\hat{\theta}(\boldsymbol{\omega}),$$

where $g^{s,r}(\boldsymbol{\omega})$ is the (s,r)-th element of $G(\boldsymbol{\omega})^{-1}$. The matrix $\tilde{H}_{\hat{\theta}(\omega)}$ is called the covariant Hessian of $\hat{\theta}(\boldsymbol{\omega})$.

First and second derivatives of $\hat{\theta}(\boldsymbol{\omega}(t))$ on M, at $t = 0$ can be used to construct influence measures (see [26]).

A first order (FI) influence measure in the direction $\mathbf{h} \in T_{\omega^0}$ is given by

$$\mathrm{FI}_{\hat{\theta},h} = \frac{\mathbf{h}^T\nabla_{\hat{\theta}}\nabla_{\hat{\theta}}^T\mathbf{h}}{\mathbf{h}^T G\mathbf{h}}, \tag{5}$$

where $G = G(\boldsymbol{\omega}^0)$.

A second order (SI) influence measure in the direction $\mathbf{h} \in T_{\omega^0}$ is defined as

$$\mathrm{SI}_{\hat{\theta},h} = \frac{\mathbf{h}^T\tilde{H}_{\hat{\theta}}\mathbf{h}}{\mathbf{h}^T G\mathbf{h}}. \tag{6}$$

[26] showed that if $\boldsymbol{\omega}$ is an appropriate perturbation and $\nabla_{\hat{\theta}} = \mathbf{0}$, then $\mathrm{SI}_{\hat{\theta},h}$ coincides with the normal curvature defined in [6] in the direction $\mathbf{h}$, which would be calculated in this case from the surface formed by the corrected score estimator. Moreover, $\mathrm{SI}_{\hat{\theta},h}$ is scale invariant even when $\nabla_{\hat{\theta}} \neq \mathbf{0}$, whereas the normal curvature of Cook is not ([8]).

Maximum values of $\mathrm{FI}_{\hat{\theta},h}$ and $\mathrm{SI}_{\hat{\theta},h}$ quantify the degree of local influence of $\boldsymbol{\omega}$ to a statistical model, while the associated directions can be used for identifying influential observations. Also, the absolute values of $\mathrm{FI}_j = \mathrm{FI}_{\hat{\theta},e_j}$ and $\mathrm{SI}_j = \mathrm{SI}_{\hat{\theta},e_j}$, where $\mathbf{e_j}$ is an $n \times 1$ vector with j-th element one and zero otherwise, for $j = 1, \ldots, n$ can be used for this purpose. If $\nabla_{\hat{\theta}} \neq \mathbf{0}$, then $\mathrm{FI}_{\hat{\theta},h}$ and $\mathrm{SI}_{\hat{\theta},h}$ are used together.

5 Application to the simple linear regression model

We consider the simple linear functional measurement error model to illustrate how to calculate geometrical quantities for a perturbation manifold. We check whether the perturbation $\boldsymbol{\omega}$ is appropriate and calculate influence measures FI and SI associated with the corrected score estimator of a particular component of $\boldsymbol{\theta}$, to assess local influence of the perturbation.

The simple linear functional measurement error model can be represented by the equations

$$\begin{aligned} y_j &= \alpha + \beta z_j + \varepsilon_j, \\ x_j &= z_j + u_j, \qquad j = 1\ldots n, \end{aligned} \tag{7}$$

with $\varepsilon_j \sim N(0,\sigma_\varepsilon^2)$ independent of $u_j \sim N(0,\sigma_u^2)$. To make the model identifiable, the variance σ_u^2 is taken as known.

In this case, $\boldsymbol{\theta} = (\alpha, \beta, \sigma_\varepsilon^2)^T$ is the structural parameter vector and $\mathbf{Z} = (z_1, \ldots, z_n)^T$ is the corresponding vector of incidental parameters.

The log-likelihood of the model is given by

$$\sum_{j=1}^n \ell(\boldsymbol{\theta}, z_j; y_j, x_j) = \sum_{j=1}^n \ell(\boldsymbol{\theta}, z_j; y_j) + \sum_{j=1}^n \ell(\boldsymbol{\theta}, z_j; x_j),$$

where

$$\ell(\boldsymbol{\theta}, z_j; y_j) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\sigma_\varepsilon^2 - \frac{1}{2\sigma_\varepsilon^2}(y_j - \alpha - \beta z_j)^2$$

corresponds to the unobserved log-likelihood and

$$\ell(\boldsymbol{\theta}, z_j; x_j) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\sigma_u^2 - \frac{1}{2\sigma_u^2}(x_j - z_j)^2$$

incorporates the relation between x_j and z_j.

The corrected score function can be written ([10]) as

$$\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X}) = \sum_{j=1}^{n} \mathbf{U}^*(\boldsymbol{\theta}; y_j, x_j),$$

where

$$\mathbf{U}^*(\boldsymbol{\theta}; y_j, x_j) = \frac{1}{\sigma_\varepsilon^2}\begin{pmatrix} y_j - \alpha - \beta z_j \\ (y_j - \alpha - \beta x_j)x_j + \beta\sigma_u^2 \\ -\frac{1}{2} + \frac{1}{2\sigma_\varepsilon^2}[(y_j - \alpha - \beta x_j)^2 - \beta^2\sigma_u^2] \end{pmatrix}.$$

The solution to the equation $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X}) = \mathbf{0}$ leads to the estimators

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}, \quad \hat{\beta} = \frac{S_{xy}}{S_{xx} - \sigma_u^2} \quad \text{and} \quad \hat{\sigma}_\varepsilon^2 = S_{yy} - \hat{\beta}S_{xy},$$

with $S_{xx} > \sigma_u^2$ and $S_{yy} > \frac{S_{xy}^2}{S_{xx}-\sigma_u^2}$, where $\bar{x} = \frac{1}{n}\sum_{j=1}^n x_j$, $\bar{y} = \frac{1}{n}\sum_{j=1}^n y_j$, $S_{xx} = \frac{1}{n}\sum_{j=1}^n (x_j - \bar{x})^2$, $S_{yy} = \frac{1}{n}\sum_{j=1}^n (y_j - \bar{y})^2$ and $S_{xy} = \frac{1}{n}\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})$.

The estimators $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\sigma}_\varepsilon^2$ differ from the maximum likelihood estimators, which are inconsistent.

For brevity we include the calculations only for case weights perturbation scheme, so that

$$\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}) = \sum_{j=1}^{n} \omega_j \mathbf{U}^*(\boldsymbol{\theta}; y_j, x_j), \text{ with } \boldsymbol{\omega}^0 = \mathbf{1}_n. \tag{8}$$

Solving $\mathbf{U}^*(\boldsymbol{\theta};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}) = \mathbf{0}$, we obtain the perturbed corrected score estimators

$$\hat{\alpha}(\boldsymbol{\omega}) = \bar{y}_\omega - \hat{\beta}(\boldsymbol{\omega})\bar{x}_\omega, \quad \hat{\beta}(\boldsymbol{\omega}) = \frac{S_{xy\omega}}{S_{xx\omega} - \sigma_u^2} \quad \text{and} \quad \hat{\sigma}_\varepsilon^2(\boldsymbol{\omega}) = S_{yy\omega} - \hat{\beta}(\boldsymbol{\omega})S_{xy\omega},$$

where $\bar{x}_\omega = \sum_{j=1}^n \gamma_{j\omega} x_j$, $\bar{y}_\omega = \sum_{j=1}^n \gamma_{j\omega} y_j$, $S_{xx\omega} = \sum_{j=1}^n \gamma_{j\omega}(x_j - \bar{x}_\omega)^2$, $S_{yy\omega} = \sum_{j=1}^n \gamma_{j\omega}(y_j - \bar{y}_\omega)^2$ and $S_{xy\omega} = \sum_{j=1}^n \gamma_{j\omega}(x_j - \bar{x}_\omega)(y_j - \bar{y}_\omega)$, with $\gamma_{j\omega} = \omega_j/(\sum_{j=1}^n \omega_j)$.

The perturbed log-likelihood of the model can be written as

$$\ell(\boldsymbol{\theta},\mathbf{Z};\mathbf{Y},\mathbf{X},\boldsymbol{\omega}) = \sum_{j=1}^{n} \omega_j\{-\log(2\pi\sigma_u\sigma_\varepsilon) - \frac{1}{2\sigma_\varepsilon^2}(y_j - \alpha - \beta z_j)^2 - \frac{1}{2\sigma_u^2}(x_j - z_j)^2\}.$$

Thus, the density of the perturbed model is given by

$$p(\mathbf{Y},\mathbf{X};\mathbf{Z},\boldsymbol{\theta},\boldsymbol{\omega}) = \prod_{j=1}^{n}\{\frac{1}{2\pi\sigma_\varepsilon\sigma_u}\omega_j \exp\{-\omega_j[\frac{1}{2\sigma_\varepsilon^2}(y_j - \alpha - \beta z_j)^2 + \frac{1}{2\sigma_u^2}(x_j - z_j)^2]\}\}.$$

Then

$$\begin{aligned}&\ell(\boldsymbol{\omega};\mathbf{Y},\mathbf{X},\mathbf{Z},\boldsymbol{\theta}) = \log p(\mathbf{Y},\mathbf{X};\mathbf{Z},\boldsymbol{\theta},\boldsymbol{\omega}) \\ &= \sum_{j=1}^{n}\{-\log(2\pi\sigma_u\sigma_\varepsilon) + \log\omega_j - \omega_j[\frac{1}{2\sigma_\varepsilon^2}(y_j - \alpha - \beta z_j)^2 + \frac{1}{2\sigma_u^2}(x_j - z_j)^2]\}.\end{aligned}$$

After some calculations, we have

$$g_{ij}(\boldsymbol{\omega}) = \frac{1}{\omega_i^2}\delta_{ij} \quad \text{and} \quad \Gamma_{ijk}(\boldsymbol{\omega}) = -\frac{1}{\omega_i^3}\delta_{ij}\delta_{ik}, \quad i,j,k = 1\ldots n.$$

Thus, $G(\boldsymbol{\omega}^0) = I_n$ and the perturbation in (8) is an appropriate one.

We consider the corrected score estimator $\hat{\beta}(\boldsymbol{\omega})$ as our objective function. After some algebraic derivations, first and second derivatives of $\hat{\beta}(\boldsymbol{\omega})$ at $\boldsymbol{\omega}^0 = \mathbf{1}_n$ are obtained as follows

$$\nabla_{\hat{\beta}} = \frac{\partial \hat{\beta}(\boldsymbol{\omega}^0)}{\partial \boldsymbol{\omega}} = \frac{1}{n(S_{xx} - \sigma_u^2)} \{\mathbf{d}_x \odot \mathbf{d}_y - \hat{\beta}\mathbf{s}_x\}$$

and

$$H_{\hat{\beta}} = \frac{\partial^2 \hat{\beta}(\boldsymbol{\omega}^0)}{\partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^T} = A + A^T,$$

with

$$A = -\frac{1}{n^2(S_{xx} - \sigma_u^2)^2} \{(S_{xx} - \sigma_u^2)(\mathbf{d}_y - \hat{\beta}\mathbf{d}_x)\mathbf{d}_x^T + (\mathbf{d}_x \odot \mathbf{d}_y - \hat{\beta}\mathbf{s}_x)\mathbf{s}_x^T\},$$

where $\mathbf{d}_x = (x_1 - \bar{x}, \ldots, x_n - \bar{x})^T$, $\mathbf{d}_y = (y_1 - \bar{y}, \ldots, y_n - \bar{y})^T$, $\mathbf{s}_x = \mathbf{d}_x \odot \mathbf{d}_x - \sigma_u^2 \mathbf{1}_n$ and $\odot$ denotes the element-wise product.

Direct calculations leads to the covariant Hessian matrix

$$\tilde{H}_{\hat{\beta}} = H_{\hat{\beta}} + \mathrm{diag}(\nabla_{\hat{\beta}}),$$

where $\mathrm{diag}(\nabla_{\hat{\beta}})$ represents the diagonal matrix with the elements of the vector $\nabla_{\hat{\beta}}$ in the diagonal.

Then, first and second-order influence measures, $\mathrm{FI}_{\hat{\beta},h}$ and $\mathrm{SI}_{\hat{\beta},h}$, can be calculated from formulas (5) and (6), respectively.

Perturbations of the observed covariable and response variable can be analyzed similarly. They are also appropriate and in these cases $\Gamma_{ijk}(\boldsymbol{\omega}) = 0$ for all $i, j, k = 1 \ldots n$, so the straight line $\boldsymbol{\omega}(t) = \boldsymbol{\omega}^0 + t\mathbf{h}$ is a geodesic and $\tilde{H}_{\hat{\beta}} = H_{\hat{\beta}}$.

5.1 Numerical illustration

To illustrate the computation of influence measures, we reanalyze the data set considered in [11]. The data given in Table 1 are measurement of serum kanamycin levels in blood samples drawn from twenty babies. One of the measurements was obtained by a heelstick method (x), the other using an umbilical catheter (y). Determinations of serum kanamycin contain measurement error in both methods. In [11] it is assumed a structural model (that is, with x a random variable) under the assumption $\sigma_u^2 = \sigma_\varepsilon^2$, to assure identifiability. Maximum likelihood estimates are calculated and the influence function is used for detection of influential observations. For the purpose of the example, here we assume that the functional model (7) holds, with σ_u^2 known for making the model identifiable. In [11], the estimates from the analysis were $\hat{\alpha} = -1.16$, $\hat{\beta} = 1.07$ and $\hat{\sigma}_u^2 = \hat{\sigma}_\varepsilon^2 = 4.60$. We choose $\sigma_u^2 = 4.40$, as the known value of the variance, with the aim of making both analysis comparable.

Table 1: Serum kanamycin levels in blood samples

Baby	*Heelstick*	*Catheter*	*Baby*	*Heelstick*	*Catheter*
1	23.0	25.2	11	26.4	24.8
2	33.2	26.0	12	21.8	26.8
3	16.6	16.3	13	14.9	15.4
4	26.3	27.2	14	17.4	14.9
5	20.0	23.2	15	20.0	18.1
6	20.0	18.1	16	13.2	16.3
7	20.6	22.2	17	28.4	31.3
8	18.9	17.2	18	25.9	31.2
9	17.8	18.8	19	18.9	18.0
10	20.0	16.4	20	13.8	15.6

Under these assumptions the estimates from the analysis by the corrected score approach are

$$\hat{\alpha} = -1.19, \quad \hat{\beta} = 1.07, \quad \hat{\sigma}_\varepsilon^2 = 4.34.$$

The slope and intercept estimates are similar to those obtained by the analysis of [11].

In the perturbation of case weights scheme the maximum absolute values of $\mathrm{FI}_{\hat{\beta},h}$ and $\mathrm{SI}_{\hat{\beta},h}$ are 0.074011 and 0.45629, respectively.

Figure 1 gives the index plot of $|\mathrm{FI}_j|$ and $|\mathrm{SI}_j|$, $j = 1, \ldots 20$. First and second order influence measures reveal that case 2 is the most influential on the corrected score estimate of the slope. This result coincides with that obtained by [11].

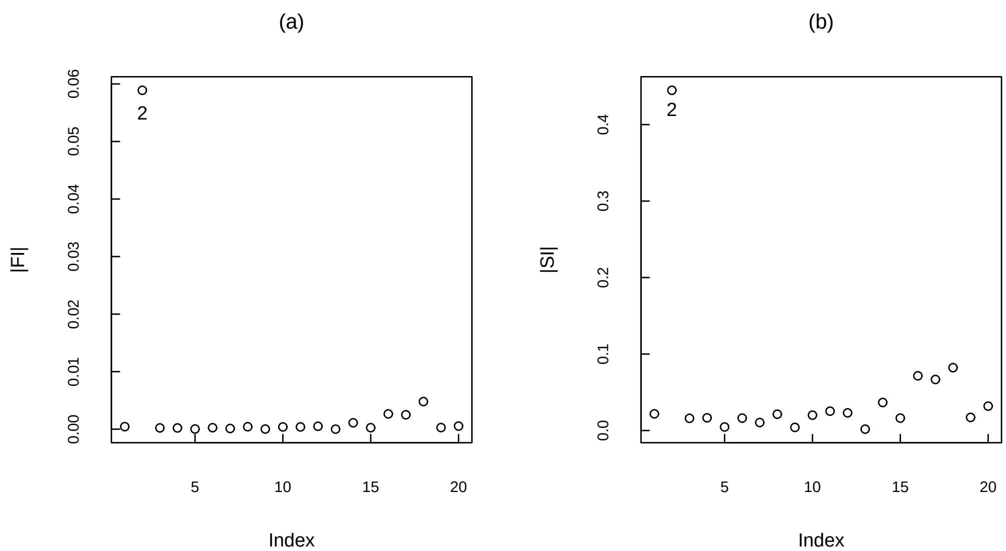

Figure 1: Index plot of (a): $|\mathrm{FI}_j|$ and (b): $|\mathrm{SI}_j|$

6 CONCLUSIONS

We have defined a Riemannian manifold, called the perturbation manifold, for assessment of local influence on corrected score estimators for functional measurement error models. A perturbation vector is introduced on the corrected score function (which is independent of the incidental parameters) and then the density of the corresponding perturbed model (including structural and incidental parameters) is found. The perturbed model, as a function of the perturbation vector, defines the manifold. The metric tensor permits to select an appropriate perturbation. First and second-order influence measures are defined based on a covariant version of the Taylor's theorem.

We consider the simple linear functional measurement error model with one of the variances known, to illustrate the calculation of the geometrical quantities of interest in the case weight perturbation scheme. We selected as our objective function the corrected score estimator of the slope and obtained simple formulae for first and second-order influence measures. The influence analysis on the corrected score estimators of the others parameters in the model can be performed separately.

The calculations of influence measures based on perturbation manifolds in more complex measurement error models, both functional and structural, by using different estimation approaches merits further research. Relationships between the influence measures here defined an other influence diagnostics in measurement error models could be also studied.

ACKNOWLEDGEMENTS

The author would like to thank the referee for helpful suggestions.

REFERENCES

[1] S. Amari. *Differential-Geometrical Methods in Statistics.*, Lecture Notes in Statistics, 28. Springer, Berlin. (1985).

[2] R. BRAND. *Microdata protection through noise addition* in Inference Control in Statistical Databases - From Theory to Practice, Lecture Notes in Computer Science 2316, Springer, Berlin. (2002).

[3] P.J. BROWN, *Multivariate calibration*, Journal of the Royal Statistical Society - Series B, 44 (1982), pp.287-321.

[4] C.-L Cheng, and J.W. Van Ness. *Statistical Regression with Measurement Error.*, Arnold, London. (1999).

[5] R.J. Carroll, D. Ruppert, L.A. Stefanski, and C.M. Crainiceanu, *Measurement Error in Nonlinear Models: A Modern Perspective.*, Second Edition. Chapman & Hall, London. (2006).

[6] R.D. COOK, *Assessment of Local Influence (with Discussion)*, Journal of the Royal Statistical Society - Series B, 48 (1986), pp.133-169.

[7] W. Fuller, *Measurement error models.*, Wiley, N.Y. (1987).

[8] W. FUNG, AND C. KWAN, *A note on local influence based on normal curvature*, Journal of the Royal Statistical Society - Series B, 59 (1997), pp.839-843.

[9] M. GALEA-ROJAS, H. BOLFARINE, AND M. DE CASTRO, *Local influence in comparative calibration models*, Biometrical Journal, 1 (2002), pp.59-81.

[10] P. GIMÉNEZ, AND H. BOLFARINE, *Corrected score functions in classical error-in-variables and incidental parameters models*, The Australian Journal of Statistics, 39 (1997), pp.325-344.

[11] G.E. KELLY, *The influence function in the error in variables problem*, The Annals of Statistics, 12 (1984), pp.87-100.

[12] M.G. Kendall, and A. Stuart. *The Advanced Theory of Statistics.*, Vol.2, 4th ed. Hafner, New York. (1979).

[13] A.H. LEE, AND Y. ZHAO, *Assessing local influence in measurement error models*, Biometrical Journal, 38 (1996), pp.829-841.

[14] M.K. Murray, and J.W. Rice. *Differential Geometry and Statistics.*, Chapman and Hall, London. (1993).

[15] T. NAKAMURA, *Corrected score function for errors-in-variables models: Methodology and application to generalized linear models*, Biometrika, 77 (1990), pp.127-137.

[16] J. NEYMANN, AND E. SCOTT, *Consistent estimates based on partially consistent observations*, Econometrica, 16 (1948), pp.1-32.

[17] W.M. PATEFIELD, *On the information matrix in the linear functional relationship problem*, Applied Statistics, 26 (1977), pp.69-70.

[18] A.R. RASEKH, AND N.R.J. FIELLER, *Influence functions in functional measurement error models with replicated data*, Statistics, 37 (2003), pp.169-178.

[19] L.A. STEFANSKI, *Unbiased estimation of a linear function of a normal mean with application to measurement error models*, Communications in Statistics - Series A, 18 (1989), pp.4335-4358.

[20] L.A. STEFANSKI, AND R.J. CARROLL, *Covariate measurement error in logistic regression*, Annals of Statistics, 13 (1985), pp.1335-1351.

[21] J.M. WELLMAN, AND R.F. GUNST, *Influence diagnostics for linear measurement error models*, Biometrika, 78 (1991), pp.373-380.

[22] X. WU, AND Z. LUO, *Second-order approach to local influence*, Journal of the Royal Statistical Society - Series B, 55 (1993), pp.929-936.

[23] Y. ZHAO, A.H. LEE, AND Y.V. HUI, *Influence diagnostics for generalized linear measurement error models*, Biometrics, 50 (1994), pp.1117-1128.

[24] Y. ZHAO, AND A.H. LEE, *Assessment of influence in nonlinear measurement error models*, Journal of Applied Statistics, 22 (1995), pp.215-225.

[25] X.P. ZHONG, B.C. WEI, AND W.K. FUNG, *Influence analysis for linear measurement error models*, Annals of Institute of Statistical Mathematics, 52 (2000), pp.367-379.

[26] H. ZHU, J.G. IBRAHIM, S. LEE, AND H. ZHANG, *Perturbation selection and influence measures in local influence analysis*, Annals of Statistics, 35 (2007), pp.2565-2588.