On Estimation and Influence Diagnostics for the Grubbs' model under Heavy-tailed Distributions

Felipe Osorio [a,1,*], Gilberto A. Paula [b], Manuel Galea [a]

[a] *Departamento de Estadística, CIMFAV, Universidad de Valparaíso, Chile*

[b] *Instituto de Matemática e Estatística, Universidade de São Paulo, Brazil*

Abstract

The Grubbs' measurement model is frequently used to comparing several measuring devices. It is common to assume that the random terms have a normal distribution. However, such assumption makes the inference vulnerable to outlying observations whereas scale mixtures of normal distributions have been an interesting alternative to produce robust estimates keeping the elegancy and simplicity of the maximum likelihood theory. The aim of this paper is to develop an EM-type algorithm for the parameter estimation and to use the local influence method to assessing the robustness aspects of these parameter estimates under some usual perturbation schemes. In order to identify outliers and to criticize the model building we use the local influence procedure in a study to compare the precision of several thermocouples.

Key words: Grubbs' model; Heavy-tailed distributions; Outliers; Q-displacement; Regression diagnostics.

---

1 Introduction

The problem of comparing the precision and accuracy of different measuring instruments may appear in various scientific applications like engineering (Grubbs, 1948, 1973) and medicine (Barnett, 1969). Taking measurements of

* Corresponding author.
Email addresses: felipe.osorio@uv.cl (Felipe Osorio), giapaula@ime.usp.br (Gilberto A. Paula), manuel.galea@uv.cl (Manuel Galea).
[1] Address for correspondence: Departamento de Estadística, CIMFAV, Universidad de Valparaíso, Casilla 5030, Valparaíso, Chile.

Preprint submitted to Elsevier *9 September 2008*

the same unknown characteristic x from different individuals or experimental units has been the usual way for comparing the instruments. Wich may differ in some aspects such as cost, speed and convenience. The relative quality in the measurements is evaluated considering the precision and bias of the different instruments.

The assessment of robustness aspects of the parameter estimates in statistical models has been an important concerning of various researchers in the last decades. The deletion methodology, which consists in studying the impact on the parameter estimates after dropping individual observations, is probably the most employed technique to detect influential observations (see, for example, Cook and Weisberg, 1982 and Chatterjee and Hadi, 1988). Nevertheless, the local influence procedure (Cook, 1986), that investigates the influence of small perturbations in the model/data on the parameter estimates, has received an increasing attention in the last 20 years, mainly due to its flexibility in constructing different kinds of graphics and its applicability in various statistical models (see discussion in Cook, 1997). In particular, Galea, Bolfarine and de Castro (2002) and Lachos, Vilca and Galea (2007) applied the methodology in normal comparative calibration and Grubbs' models, notifying under some usual perturbation schemes the well known lack of robustness of the least-squares estimates against outlying observations.

Several methodologies have been proposed to attenuate the influence of outlying observations on the parameter estimates under normality, such as modifications of the least-squares methodology (see, for instance, Huber, 1981). Other approaches that assume heavy-tailed error distributions for which the maximum likelihood estimates appear to be robust against extreme observations have been proposed (see, for example, Galea, Bolfarine and Vilca, 2005). In this work, we will assume scale mixtures of normal distributions (Andrews and Mallows, 1974) for the accommodation of extreme and outlying observations in the Grubbs' model. Properties of distributions in this class, such as Student-t, power exponential and contaminated normal may be found in Andrews and Mallows (1974) and Lange and Sinsheimer (1993). In this paper scale mixtures of normal distributions are assumed for the Grubbs' model and the hierarchical representation proposed by Pinheiro, Liu and Wu (2001) is considered. Our aim is to apply the local influence method in the Grubbs' model under heavy-tailed distributions in order to assess the influence of minor perturbations on the model/data, our results are generalizations of the results obtained by Lachos, Vilca and Galea (2007). The rest of the paper is organized as follows. In Section 2 some inferential aspects are discussed and an EM-type algorithm is developed for the parameter estimation. Section 3 introduces the local influence methodology (Cook, 1986 and Zhu and Lee, 2001). The normal curvature for some usual perturbation schemes is derived in Section 4. The methodology is illustrated in Section 5 in which Grubbs' models under normal and scale mixtures of normal distributions are compared according to the robustness aspects of the maximum likelihood estimates. Finally, some concluding remarks are given in Section 6.

2 Model description

Grubbs (1948, 1973, 1983) proposes a linear model for comparing p different instruments, in which the characteristic x_i of the ith experimental unit is measured once by all the p instruments. The model assumes the form

$$Y_{ij} = \alpha_j + x_i + \epsilon_{ij}, \qquad i = 1, \ldots, M \text{ and } j = 1, \ldots, p, \tag{1}$$

where Y_{ij} denotes the measurement of the jth instrument for the ith experimental unit and α_j is called additive bias. The measurement errors ϵ_{ij} are assumed to be independent of the random variables $x_1, \ldots, x_M$. In addition, one has that x_i and ϵ_{ij} are mutually independent and distributed according to $x_i \sim N(\mu_x, \phi_x)$ and $\epsilon_{ij} \sim N(0, \phi_j)$.

In order to allow model (1) be identifiable we may consider $\alpha_1 = 0$ (see, for example, Shyr and Gleser, 1986; Bedrick, 2001 and Christensen and Blackwood, 1993). However, in this work, we will assume the transformation $z_i = x_i - \mu_x$, $i = 1, \ldots, M$ (Theobald and Mallison, 1978) so that the Grubbs' model is expressed in the alternative form

$$\boldsymbol{Y}_i = \boldsymbol{\mu} + \mathbf{1} z_i + \boldsymbol{\epsilon}_i, \qquad i = 1, \ldots, M, \tag{2}$$

where $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{ip})^T$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)^T$ and $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{ip})^T$ with z_i denoting a random variable normally distributed of mean zero and dispersion parameter ϕ_x.

Christensen and Blackwood (1993) reported 64 sets of simultaneous measurements for temperature obtained by five thermocouples previously used, with the aim of examining their precision and exactitude after one or more thermocouples had sustain certain damage. Let Y_{ij} represent the measurement for the temperature by the jth thermocouple for the ith item, $i = 1, \ldots, 64$; $j = 1, \ldots, 5$, and these data are plotted in Figure 1. The plot shows several measurements as potential outliers.

It is well known that models developed under the assumption of normality are susceptible to outlying observations. The Grubbs model given in (2) considers two sources of variation, which may generate outliers in the error component as well as in the latent variable component (see, for instance, Pinheiro, Liu and

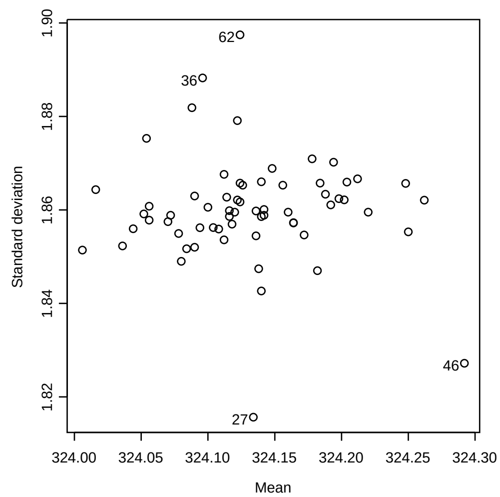

Fig. 1. Means and standard deviations of the measurements of the thermocouples.

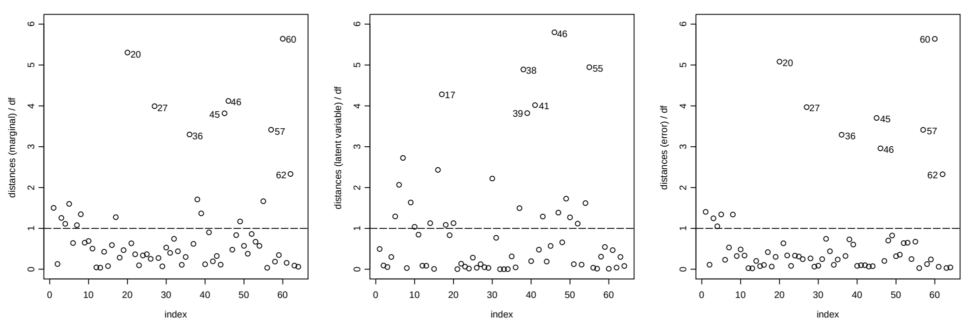

Fig. 2. Index plots of the Mahalanobis distances under normality.

Wu, 2001). In order to identify such observations we consider the distances

$$U_i = (\boldsymbol{Y}_i - \boldsymbol{\mu})^T (D(\boldsymbol{\phi}) + \phi_x \mathbf{1}\mathbf{1}^T)^{-1} (\boldsymbol{Y}_i - \boldsymbol{\mu}),$$
$$U_{\epsilon_i} = \boldsymbol{\epsilon}_i^T D^{-1}(\boldsymbol{\phi}) \boldsymbol{\epsilon}_i, \quad \text{and} \quad U_{z_i} = z_i^2/\phi_x,$$

where $\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - \boldsymbol{\mu} - z_i \mathbf{1}$ for $i = 1, \ldots, 64$ and $D(\boldsymbol{\phi}) = \text{diag}(\phi_1, \ldots, \phi_p)$. Under the assumption of normality one has that $U_i \sim \chi^2_p$, $U_{\epsilon_i} \sim \chi^2_p$ and $U_{z_i} \sim \chi^2_1$. Since $\text{E}(U_i) = p$, $\text{E}(U_{\epsilon_i}) = p$ and $\text{E}(U_{z_i}) = 1$, Pinheiro, Liu and Wu (2001) proposed the quantities $\widehat{U}_i/p$, $\widehat{U}_{\epsilon_i}/p$ and $\widehat{U}_{z_i}$ to identify outlying observations. These statistics have expected value equals to one. The parameter estimates under normal error (standard errors in parenthesis) are given in the first column of Table 1. From Figure 2 we can notice observations 20, 27, 36, 45, 46, 57, 60 and 62 as possible outliers. In the sequel we introduce the class of scale mixtures of normal error models to accommodate outlying observations.

The class of scale mixtures of normal distributions (Andrews and Mallows, 1974) has been applied in the context of regression models (see, for instance, Lange and Sinsheimer, 1993 and Liu, 1996) as well as in linear mixed models (Rosa, Padovani and Gianola, 2003, 2004) for obtaining robust estimates against outlying observations. To define this class, suppose the m-dimensional vector $\boldsymbol{Y}$ written as

$$\boldsymbol{Y} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + \kappa^{1/2}(V)\boldsymbol{W}, \tag{3}$$

then we say that $\boldsymbol{Y}$ follows a scale mixtures of normal distributions of position parameter $\boldsymbol{\mu} \in \mathbb{R}^m$ and definite positive matrix $\boldsymbol{\Lambda}$, where $\boldsymbol{W} \sim N_m(\boldsymbol{0}, \boldsymbol{\Lambda})$, V is a positive random variable, named mixture variable, with distribution $H(v; \boldsymbol{\nu})$ independent of $\boldsymbol{W}$ indexed by the parameter vector $\boldsymbol{\nu}$, whereas $\kappa(\cdot)$ is a strictly positive function which is associated to the mixture variable V. Here $\stackrel{\mathrm{d}}{=}$ denotes equivalence of distributions. It is easy to see that the conditional distribution $(\boldsymbol{Y}|V = v) \sim N_m(\boldsymbol{\mu}, \kappa(v)\boldsymbol{\Lambda})$. In addition, the marginal density of $\boldsymbol{Y}$ takes the form

$$f(\boldsymbol{y}) = |2\pi\boldsymbol{\Lambda}|^{-1/2} \int_0^\infty \{\kappa(v)\}^{-m/2} \exp\{-\tfrac{1}{2}\kappa^{-1}(v)u\} \,\mathrm{d}H, \tag{4}$$

where $u = (\boldsymbol{Y} - \boldsymbol{\mu})^T \boldsymbol{\Lambda}^{-1} (\boldsymbol{Y} - \boldsymbol{\mu})$. When the density of $\boldsymbol{Y}$ assumes the form (4) we will denote $\boldsymbol{Y} \sim SMN_m(\boldsymbol{\mu}, \boldsymbol{\Lambda}; H)$. The class of scale mixtures of normal distributions presents similar properties of the normal distribution, is simple to work with it and allows the development of robust procedures for the parameter estimation. Examples of distributions in this class may be found, for instance, in Lange and Sinsheimer (1993).

Based on the suggestion of Pinheiro, Liu and Wu (2001), we will introduce scale mixtures of normal distributions in the Grubbs' model defined in (2), by considering the following hierarchical structure:

$$\begin{aligned} \boldsymbol{Y}_i|z_i &\stackrel{\text{ind}}{\sim} SMN_p(\boldsymbol{\mu} + \boldsymbol{1}z_i, D(\boldsymbol{\phi}); H), \\ z_i &\stackrel{\text{ind}}{\sim} SMN(0, \phi_x; H), \qquad i = 1, \dots, M, \end{aligned}$$

or, equivalently, as

$$\begin{aligned} \boldsymbol{Y}_i|z_i, v_i &\stackrel{\text{ind}}{\sim} N_p(\boldsymbol{\mu} + \boldsymbol{1}z_i, \kappa(v_i)D(\boldsymbol{\phi})), \quad z_i|v_i \stackrel{\text{ind}}{\sim} N(0, \kappa(v_i)\phi_x), \\ v_i &\stackrel{\text{ind}}{\sim} H(v_i; \boldsymbol{\nu}), \qquad i = 1, \dots, M, \end{aligned} \tag{5}$$

where $D(\boldsymbol{\phi}) = \mathrm{diag}(\phi_1, \dots, \phi_p)$.

Let $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T, \phi_x)^T$ be the parameter vector of interest. We will apply an EM-type algorithm (Meng and Rubin, 1993; see also McLachlan and Krishnan, 1997) for the parameter estimation by assuming that $(\boldsymbol{z}^T, \boldsymbol{v}^T)^T$ with $\boldsymbol{z} = (z_1, \dots, z_M)^T$ and $\boldsymbol{v} = (v_1, \dots, v_M)^T$ are not observable. Thus, the vector of complete data will be given by $\boldsymbol{Y}_c = (\boldsymbol{Y}^T, \boldsymbol{z}^T, \boldsymbol{v}^T)^T$, where $\boldsymbol{Y} = (\boldsymbol{Y}_1^T, \dots, \boldsymbol{Y}_M^T)^T$ corresponds to the vector of observable responses for the M

individuals. The log-likelihood function for the complete data will be denoted by $L(\boldsymbol{\theta}|\boldsymbol{Y}_c) = \sum_{i=1}^M L_i(\boldsymbol{\theta}|\boldsymbol{Y}_c)$, where

$$\begin{aligned} L_i(\boldsymbol{\theta}|\boldsymbol{Y}_c) = &-\tfrac{1}{2}\log|D(\boldsymbol{\phi})| - \tfrac{\kappa^{-1}(v_i)}{2}(\boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}z_i)^T D^{-1}(\boldsymbol{\phi})(\boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}z_i) \\ &- \tfrac{1}{2}\log\phi_x - \tfrac{\kappa^{-1}(v_i)}{2\phi_x} z_i^2 + \log h(v_i;\boldsymbol{\nu}) + C, \end{aligned} \tag{6}$$

with $h(v_i;\boldsymbol{\nu})$ being the density function of the mixture variable V_i and C a constant. For Grubbs' model in (5) it is possible to show that the expected complete data log-likelihood function, called Q-function in Dempster, Laird and Rubin (1977) evaluated in the current estimate $\widehat{\boldsymbol{\theta}}^{(k)} = (\boldsymbol{\mu}^{(k)^T}, \boldsymbol{\phi}^{(k)^T}, \phi_x^{(k)^T})^T$, may be expressed as $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) = \sum_{i=1}^M Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)})$ with

$$\begin{aligned} Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)}) = &-\tfrac{1}{2}\log|D(\boldsymbol{\phi})| - \tfrac{1}{2}\widehat{\tau}^{(k)}\mathbf{1}^T D^{-1}(\boldsymbol{\phi})\mathbf{1} \\ &- \tfrac{1}{2}\widehat{\kappa}_i^{(k)}(\boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}\widehat{z}_i^{(k)})^T D^{-1}(\boldsymbol{\phi})(\boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}\widehat{z}_i^{(k)}) \\ &- \tfrac{1}{2}\log\phi_x - \tfrac{1}{2\phi_x}(\widehat{\kappa}_i^{(k)}\{\widehat{z}_i^{(k)}\}^2 + \widehat{\tau}^{(k)}), \end{aligned} \tag{7}$$

where

$$\widehat{z}_i^{(k)} = \widehat{\tau}^{(k)}\, \mathbf{1}^T D^{-1}(\widehat{\boldsymbol{\phi}}^{(k)})(\boldsymbol{Y}_i - \widehat{\boldsymbol{\mu}}^{(k)}), \qquad \widehat{\tau}^{(k)} = \widehat{\phi}_x^{(k)}/\widehat{s}^{(k)}, \tag{8}$$

with $\widehat{s}^{(k)} = 1 + \widehat{\phi}_x^{(k)}\mathbf{1}^T D^{-1}(\widehat{\boldsymbol{\phi}}^{(k)})\mathbf{1}$ and $\widehat{\kappa}_i^{(k)} = \mathrm{E}(\kappa_i^{-1}(V_i)|\boldsymbol{Y}_i, \widehat{\boldsymbol{\theta}}^{(k)})$, for $i = 1,\ldots,M$.

In general, the conditional expectation $\mathrm{E}(\kappa^{-1}(V)|\boldsymbol{Y})$ is given by

$$\mathrm{E}(\kappa^{-1}(V)|\boldsymbol{Y}) = \frac{\int_0^\infty \{\kappa(v)\}^{-(p/2+1)} \exp\{-\kappa^{-1}(v)u/2\}\,\mathrm{d}H}{\int_0^\infty \{\kappa(v)\}^{-p/2} \exp\{-\kappa^{-1}(v)u/2\}\,\mathrm{d}H}, \tag{9}$$

which may be evaluated by using, for instance, the Laplace method (see Kass, 1997). However, for the most popular members of the scale mixture of normal distributions, the expectation in (9) can be easily computed, some examples are given below (see also Lange and Sinsheimer, 1993):

Student-t distribution. The Student-t distribution, denoted $t_p(\boldsymbol{\mu}, \boldsymbol{\Lambda}, \nu)$, with $\nu > 0$ degrees of freedom belongs to the class of scale mixtures of normal distributions with $\kappa(v) = 1/v$ and $V \sim Gamma(\nu/2, \nu/2)$, so that

$$\mathrm{E}(\kappa^{-1}(V)|\boldsymbol{Y}) = \frac{\nu + p}{\nu + u}.$$

The Cauchy distribution is obtained when $\nu = 1$.

Slash distribution. For the Slash distribution one has $\kappa(v) = 1/v$ and the mixture variable has density function

$$h(v;\nu) = \nu v^{\nu-1}, \qquad 0 < v \leq 1 \text{ and } \nu > 0,$$

with conditional expectation

$$\mathrm{E}(\kappa^{-1}(V)|\boldsymbol{Y}) = \Big(\frac{p+2\nu}{u}\Big)\frac{P_1(p/2+\nu+1,u/2)}{P_1(p/2+\nu,u/2)},$$

where $P_x(a,b)$ denotes the cumulative distribution function of a random variable $Gamma(a,b)$, that is

$$P_x(a,b) = \frac{b^a}{\Gamma(a)}\int_0^x r^{a-1}e^{-br}\,\mathrm{d}r.$$

Contaminated normal distribution. For the contaminated normal distribution, $CN(\boldsymbol{\mu},\boldsymbol{\Lambda},\epsilon,\gamma)$, $0\le\epsilon\le 1$ and $0<\gamma<1$, one has $\kappa(v)=1/v$ and the mixture variable follows the discrete probability function

$$h(v;\boldsymbol{\nu}) = \begin{cases} \epsilon, & \text{if } v=\gamma,\\ 1-\epsilon, & \text{if } v=1,\end{cases}$$

where $\boldsymbol{\nu}=(\epsilon,\gamma)^T$. The expectation in (9) reduces to

$$\mathrm{E}(\kappa^{-1}(V)|\boldsymbol{Y}) = \frac{1-\epsilon+\epsilon\gamma^{p/2+1}e^{(1-\gamma)u/2}}{1-\epsilon+\epsilon\gamma^{p/2}e^{(1-\gamma)u/2}}.$$

Power exponential distribution. It is possible to express the power exponential distribution ($0<\nu\le 2$) in the class of scale mixtures of normal distributions (West, 1987; Lange and Sinsheimer, 1993), however the evaluation of $\mathrm{E}(\kappa^{-1}(V)|\boldsymbol{Y})$ in (9) may be hard. In this work we apply the approach proposed by (Lange and Sinsheimer, 1993, sec. 3), in which the power exponential density takes the form

$$|2\pi\boldsymbol{\Lambda}|^{-1/2}\exp\{-\eta(u)/2\},$$

where $\eta(u) = u^{\nu/2}+p\log c_n$ and c_n is the normalizing constant. Here, the quantity $\eta'(U)$ corresponds to $\mathrm{E}(\kappa^{-1}(V)|\boldsymbol{Y})$, for which we obtain

$$\eta'(u) = (\nu/2)u^{\nu/2-1}, \quad \text{for } u\neq 0 \text{ and } \nu\neq 1.$$

Typically, the conditional expectation $\widehat{\kappa}_i^{(k)} = \mathrm{E}(\kappa_i^{-1}(V_i)|\boldsymbol{Y}_i,\widehat{\boldsymbol{\theta}}^{(k)})$ required in (7) depends on the distance $\widehat{u}_i^{(k)} = u_i(\widehat{\boldsymbol{\theta}}^{(k)})$, where

$$\begin{aligned}\widehat{u}_i^{(k)} &= (\boldsymbol{Y}_i-\widehat{\boldsymbol{\mu}}^{(k)})^T(D(\widehat{\boldsymbol{\phi}}^{(k)})+\widehat{\phi}_x^{(k)}\mathbf{1}\mathbf{1}^T)^{-1}(\boldsymbol{Y}_i-\widehat{\boldsymbol{\mu}}^{(k)})\\ &= \boldsymbol{e}_i^{(k)T}D^{-1}(\widehat{\boldsymbol{\phi}}^{(k)})\boldsymbol{e}_i^{(k)}+\{\widehat{z}_i^{(k)}\}^2/\widehat{\phi}_x^{(k)},\end{aligned}$$

with $\boldsymbol{e}_i^{(k)} = \boldsymbol{Y}_i-\widehat{\boldsymbol{\mu}}^{(k)}-\mathbf{1}\widehat{z}_i^{(k)}$, for $i=1,\dots,M$.

2.1 ECM Algorithm

Similarly to the work by Pinheiro, Liu and Wu (2001), in this section we propose an ECM algorithm that is a simple extension of the popular EM algorithm. The M-step is replaced by a constrained maximization (CM-step), where the maximization of $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(k)})$ is made with some function of $\boldsymbol{\theta}$ held fixed.

E-step: Given $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}^{(k)}$, compute $\widehat{\tau}^{(k)}$, $\widehat{z}_i^{(k)}$ and $\widehat{\kappa}_i^{(k)}$ for a particular scale mixture of normal distribution using equations (8)-(9).

CM-step 1: Update $\widehat{\boldsymbol{\mu}}^{(k+1)}$ by maximizing (7) with respect to $\boldsymbol{\mu}$, using

$$\widehat{\boldsymbol{\mu}}^{(k+1)} = \frac{1}{\sum_{i=1}^{M} \widehat{\kappa}_i^{(k)}} \sum_{i=1}^{M} \widehat{\kappa}_i^{(k)} (\boldsymbol{Y}_i - \mathbf{1}\widehat{z}_i^{(k)}).$$

CM-step 2: Fix $\boldsymbol{\mu} = \widehat{\boldsymbol{\mu}}^{(k+1)}$ and update $\widehat{\boldsymbol{\phi}}^{(k)}$ by maximizing (7) over $\boldsymbol{\phi}$, obtaining

$$\widehat{\boldsymbol{\phi}}^{(k+1)} = \widehat{\tau}^{(k)}\,\mathbf{1} + \frac{1}{M}\sum_{i=1}^{M} \widehat{\kappa}_i^{(k)} D(\boldsymbol{e}_i^{(k)})\boldsymbol{e}_i^{(k)},$$

where $\boldsymbol{e}_i^{(k)} = \boldsymbol{Y}_i - \widehat{\boldsymbol{\mu}}^{(k+1)} - \mathbf{1}\widehat{z}_i^{(k)}$, for $i = 1, \ldots, M$.

CM-step 3: Update $\phi_x^{(k)}$ by maximizing (7) with respect to ϕ_x, which gives

$$\widehat{\phi}_x^{(k+1)} = \widehat{\tau}^{(k)} + \frac{1}{M}\sum_{i=1}^{M} \widehat{\kappa}_i^{(k)} \widehat{z}_i^{(k)2}.$$

The algorithm iterates between the E and CM steps until reach convergence. In fact, under some mild conditions the sequence $\{\widehat{\boldsymbol{\theta}}^{(k)}\}$ converges to the maximum likelihood estimate $\widehat{\boldsymbol{\theta}}$.

In this work we suppose that the parameters associated to the mixture variable V are known. In particular, for the Student-t distribution, Lucas (1997) notice that the parameter estimates are robust against extreme observation only in the case that the degrees of freedom are kept fixed. Whereas Fernández and Steel (1999), alert on the estimation of ν and they notice that in this case the function of log-likelihood is unbounded and that indeed it corresponds to an nonregular estimation problem. Thus, to estimate $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T, \phi_x)^T$ we consider a set of acceptable values for the parameters related to the mixture variable and we choose the one that maximizes the log-likelihood function (see Lange, Little and Taylor, 1989).

We notice that the proposed algorithm shares the simplicity and stability of the EM algorithms for maximum likelihood estimation introduced by Bolfarine and Galea (1995, 1996) in comparative calibration models. In addition, it is computationally no expensive and guarantees no negative scale parameter estimates.

3 Local Influence

The aim of local influence Cook (1986) is to investigate the behavior of some influence measure $T(\boldsymbol{\omega})$ when small perturbations are made into the model/data, where $\boldsymbol{\omega}$ is a q-dimensional vector of perturbations restricted to some open subset $\Omega \subset \mathbb{R}^q$. In this work we assess local influence by using an appropriate measure based on the complete log-likelihood function and particularly recommended for incomplete data.

Let $L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y})$ and $L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)$ be the perturbed log-likelihood functions for observed and complete data, respectively. We will assume that the no perturbed model is nested into the perturbed model, that is, there exists $\boldsymbol{\omega}_0 \in \Omega$ such that $L(\boldsymbol{\theta}, \boldsymbol{\omega}_0|\boldsymbol{Y}) = L(\boldsymbol{\theta}|\boldsymbol{Y})$ and $L(\boldsymbol{\theta}, \boldsymbol{\omega}_0|\boldsymbol{Y}_c) = L(\boldsymbol{\theta}, |\boldsymbol{Y}_c)$ for all $\boldsymbol{\theta}$. The influence assessment of a particular perturbation scheme on the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})$ will be evaluated by the influence measure $T(\boldsymbol{\omega})$ as $\boldsymbol{\omega}$ varies in Ω.

Zhu and Lee (2001) propose an approach to perform influence diagnostic in models with incomplete data based on the Q-displacement

$$f_Q(\boldsymbol{\omega}) = 2\{Q(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) - Q(\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})|\widehat{\boldsymbol{\theta}})\}, \tag{10}$$

where $\widehat{\boldsymbol{\theta}}(\boldsymbol{\omega})$ denotes the solution for $\boldsymbol{\theta}$ by maximizing $Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}}) = \mathrm{E}\{L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)$ $|\boldsymbol{Y}, \widehat{\boldsymbol{\theta}}\}$.

Similarly to Cook (1986), Zhu and Lee (2001) study the behavior of the surface $\boldsymbol{\gamma}(\boldsymbol{\omega}) = (\boldsymbol{\omega}^T, f_Q(\boldsymbol{\omega}))^T$ and calculate the normal curvature $C_{f_Q,h}$ at the unitary direction $\boldsymbol{h} \in \mathbb{R}^q$, given by

$$C_{f_Q,h}(\boldsymbol{\theta}) = 2\, \boldsymbol{h}^T \boldsymbol{\Delta}^T \{-\ddot{\boldsymbol{Q}}(\boldsymbol{\theta})\}^{-1} \boldsymbol{\Delta} \boldsymbol{h}, \tag{11}$$

where

$$\ddot{\boldsymbol{Q}}(\boldsymbol{\theta}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \qquad \text{and} \qquad \boldsymbol{\Delta} = \frac{\partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T} \tag{12}$$

which are evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$. As the case of normal curvature given in Cook (1986) the suggestion here to examine the elements of the eigenvector associated with the largest eigenvalue of the matrix $\ddot{\boldsymbol{T}} = \boldsymbol{\Delta}^T \{-\ddot{\boldsymbol{Q}}\}^{-1} \boldsymbol{\Delta}$. Alternatively, one may also examine the total local influence $C_i = C_{f_Q,h_i}(\boldsymbol{\theta})$, where $\boldsymbol{h}_i$ is an $q \times 1$ vector of zeros with one at the ith position.

In some occasions, we may be interested in to assess the influence on a subset $\boldsymbol{\theta}_1$ of the $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)^T$. In this case, the Q-displacement is given by

$$f_Q(\boldsymbol{\omega}) = 2\{Q(\widehat{\boldsymbol{\theta}}|\widehat{\boldsymbol{\theta}}) - Q(\widehat{\boldsymbol{\theta}}_1(\boldsymbol{\omega}), \widehat{\boldsymbol{\theta}}_2(\widehat{\boldsymbol{\theta}}_1(\boldsymbol{\omega}))|\widehat{\boldsymbol{\theta}})\},$$

where $\widehat{\boldsymbol{\theta}}_2(\widehat{\boldsymbol{\theta}}_1(\boldsymbol{\omega}))$ is the maximum likelihood estimate of $\boldsymbol{\theta}_2$ in the perturbed model, for $\boldsymbol{\theta}_1$ fixed. Consider the following matrices in partitioned form

$$\ddot{\boldsymbol{Q}}(\boldsymbol{\theta}) = \begin{pmatrix} \ddot{\boldsymbol{Q}}_{11} & \ddot{\boldsymbol{Q}}_{12} \\ \ddot{\boldsymbol{Q}}_{21} & \ddot{\boldsymbol{Q}}_{22} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{B}_{22} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{\boldsymbol{Q}}_{22}^{-1} \end{pmatrix}.$$

Then, the normal curvature at the direction $\boldsymbol{h}$ for $\boldsymbol{\theta}_1$ assumes the form

$$C_{f_Q,h}(\boldsymbol{\theta}_1) = -2\boldsymbol{h}^T\boldsymbol{\Delta}^T\{\ddot{\boldsymbol{Q}}(\boldsymbol{\theta})^{-1} - \boldsymbol{B}_{22}\}\boldsymbol{\Delta h}.$$

Since $C_{f_Q,h}$ is not invariant under uniform change of scale, Poon and Poon (1999) proposed the conformal normal curvature $B_{f_Q,h}(\boldsymbol{\theta}) = C_{f_Q,h}(\boldsymbol{\theta})/||2\ddot{\boldsymbol{T}}||$, where $\|\cdot\|$ denotes some matricial norm. For example, Zhu and Lee (2001) consider the norm $\mathrm{tr}(2\ddot{\boldsymbol{T}})$. An interesting property of the conformal normal curvature is that for any unitary direction $\boldsymbol{h}$ one has $0 \leq B_{f_Q,h}(\boldsymbol{\theta}) \leq 1$, which allows comparison of curvatures among different scale mixtures of normal models.

In order to determine if the ith-observation is possible influential some authors (see, for instance, Zhu and Lee, 2001; Poon and Poon, 1999) have proposed a benchmark value for

$$B_{f_Q,h_j}(\boldsymbol{\theta}) = \sum_{i=1}^{r} \widetilde{\lambda}_i u_{ij}^2$$

where $\{(\lambda_i, \boldsymbol{u}_i)\}_{i=1}^q$ corresponds to the eigenvalues and eigenvectors of the matrix

$$-2\ddot{\boldsymbol{T}} = \sum_{i=1}^{q} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^T$$

with $\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_q = 0$, $\widetilde{\lambda}_i = \lambda_i / \sum_{k=1}^r \lambda_k$ and $\boldsymbol{u}_i = (u_{i1}, \ldots, u_{iq})^T$. Zhu and Lee (2001) defined $M(0)_j = B_{f_Q,h_i}$ and by noting that $\overline{M}(0) = 1/q$ they proposed classify the ith-observation as possible influential if B_{f_Q,h_i} is greather than the benchmark

$$\overline{M}(0) + 2SM(0)$$

where $SM(0)$ is the sample standard error of $\{M(0)_k, k = 1, \ldots, q\}$ (see also, Lee and Xu, 2004).

4 Curvature Derivation

In this section we will derive the normal curvature for the Grubbs' model by considering the hierarchical formulation given in (5). We will compute $\ddot{\boldsymbol{Q}}(\boldsymbol{\theta}) = \partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T$ and $\boldsymbol{\Delta} = \partial^2 Q(\boldsymbol{\theta},\boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^T$ by using results of matrix differentiation described in Magnus and Neudecker (1988). Details on the differential calculations for the matrices $\ddot{\boldsymbol{Q}}$ and $\boldsymbol{\Delta}$ under different perturbation schemes are given in Appendix.

4.1 Hessian Matrix, $\ddot{\boldsymbol{Q}}$

Let $\boldsymbol{\theta} = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T, \phi_x)^T$ be the parameter vector of interest. The hessian matrix $\ddot{\boldsymbol{Q}}$ evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ is given by $\ddot{\boldsymbol{Q}}(\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{M} \ddot{\boldsymbol{Q}}_i(\widehat{\boldsymbol{\theta}})$ with

$$\ddot{\boldsymbol{Q}}_i(\widehat{\boldsymbol{\theta}}) = \frac{\partial^2 Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\bigg|_{\theta=\widehat{\theta}} = \begin{pmatrix} \ddot{\boldsymbol{Q}}_{11,i} & \ddot{\boldsymbol{Q}}_{12,i} & \mathbf{0} \\ \ddot{\boldsymbol{Q}}_{12,i} & \ddot{\boldsymbol{Q}}_{22,i} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddot{Q}_{33,i} \end{pmatrix}, \tag{13}$$

with

$$\begin{aligned} \ddot{\boldsymbol{Q}}_{11,i} &= -\widehat{\kappa}_i D^{-1}(\widehat{\boldsymbol{\phi}}), \qquad \ddot{\boldsymbol{Q}}_{12,i} = -\widehat{\kappa}_i D^{-2}(\widehat{\boldsymbol{\phi}}) D(\boldsymbol{e}_i), \\ \ddot{\boldsymbol{Q}}_{22,i} &= \tfrac{1}{2} D^{-2}(\widehat{\boldsymbol{\phi}}) - \widehat{\tau} D^{-3}(\widehat{\boldsymbol{\phi}}) - \widehat{\kappa}_i D(\boldsymbol{e}_i) D^{-3}(\widehat{\boldsymbol{\phi}}) D(\boldsymbol{e}_i) \qquad \text{and} \\ \ddot{Q}_{33,i} &= \tfrac{1}{2\widehat{\phi}_x^2} - \tfrac{1}{\widehat{\phi}_x^3}(\widehat{\kappa}_i \widehat{z}_i^2 + \widehat{\tau}), \end{aligned}$$

where $\boldsymbol{e}_i = \boldsymbol{Y}_i - \widehat{\boldsymbol{\mu}} - \mathbf{1}\widehat{z}_i$, $i = 1, \ldots, M$ and $D^{-m}(\boldsymbol{a}) = \text{diag}(a_1^{-m}, \ldots, a_p^{-m})$, for $\boldsymbol{a}$ being a p-dimensional vector and $m > 0$.

4.2 Perturbation Schemes

We will evaluate in the sequel the matrix $\boldsymbol{\Delta}$ under the following perturbation schemes for the Grubbs' model given in (5): *case-weight* for detecting observations with outstanding contribution on the log-likelihood function and that may exercise high influence on the maximum likelihood estimates; *response perturbation* made on the observed values from the instruments used in the study, which may indicate observations with large influence on their own predicted values and *multiplicative bias perturbation* that may indicate if the relationship between the observed measurements from the p instruments and the true characteristic values is adequate. For each perturbation scheme one has the partitioned form

$$\boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_1 \\ \boldsymbol{\Delta}_2 \\ \boldsymbol{\Delta}_3 \end{pmatrix},$$

where $\boldsymbol{\Delta}_1 = \partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})/\partial\boldsymbol{\mu}\partial\boldsymbol{\omega}^T \in \mathbb{R}^{p\times q}$, $\boldsymbol{\Delta}_2 = \partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})/\partial\boldsymbol{\phi}\partial\boldsymbol{\omega}^T \in \mathbb{R}^{p\times q}$ and $\boldsymbol{\Delta}_3 = \partial^2 Q(\boldsymbol{\theta}, \boldsymbol{\omega}|\widehat{\boldsymbol{\theta}})/\partial\phi_x\partial\boldsymbol{\omega}^T \in \mathbb{R}^{1\times q}$ with q being the dimension of the perturbation vector $\boldsymbol{\omega}$. We will assume that integration and differentiation operations may be exchanged.

Case-weight

First, consider the following arbitrary atribution of weights for the experimental units in the log-likelihood function for complete data

$$L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) = \sum_{i=1}^{M} \omega_i L_i(\boldsymbol{\theta}|\boldsymbol{Y}_c),$$

with $L_i(\boldsymbol{\theta}|\boldsymbol{Y}_c)$ given in (6). Here $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_M)^T$ where $0 \le \omega_i \le 1$ for $i = 1, \ldots, M$ and $\boldsymbol{\omega}_0 = \mathbf{1}_M$. Note that, for $\omega_i = 0$ and $\omega_j = 1$, $j \neq i$, the ith experimental unity is dropped from the log-likelihood function for complete data. In addition, it is possible to show that the local influence for this perturbation scheme is equivalent to the deletion method (Osorio, 2006).

For this perturbation scheme we find

$$\begin{aligned}
\frac{\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\boldsymbol{\mu}\partial\omega_i}\bigg|_{\omega=\omega_0} &= \kappa^{-1}(v_i)D^{-1}(\boldsymbol{\phi})(\boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}z_i),\\
\frac{\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\boldsymbol{\phi}\partial\omega_i}\bigg|_{\omega=\omega_0} &= -\tfrac{1}{2}\{D^{-1}(\boldsymbol{\phi})\mathbf{1} - \kappa^{-1}(v_i)D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\epsilon}_i)\boldsymbol{\epsilon}_i\} \quad \text{and}\\
\frac{\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\phi_x\partial\omega_i}\bigg|_{\omega=\omega_0} &= -\tfrac{1}{2}\Big\{\tfrac{1}{\phi_x} - \tfrac{\kappa^{-1}(v_i)}{\phi_x^2}\, z_i^2\Big\},
\end{aligned}$$

where $\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}z_i$, for $i = 1, \ldots, M$.

Measurement Perturbation

We will consider additive perturbations made on the measurements obtained by the p instruments under study. Let $\boldsymbol{Y}_{i\omega}$ denote the perturbed measurements for the ith experimental unity. The following perturbation schemes will be evaluated:

Joint perturbation on the measurements obtained for the p instruments: here we replace $\boldsymbol{Y}_i$ by $\boldsymbol{Y}_{i\omega} = \boldsymbol{Y}_i + \boldsymbol{\omega}_i$, where $\boldsymbol{\omega}_i \in \mathbb{R}^p$, $i = 1, \ldots, M$, in which one has $\boldsymbol{\omega}_0 = \mathbf{0} \in \mathbb{R}^{Mp}$. The perturbed log-likelihood function for the complete data is given by $L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) = \sum_{i=1}^{M} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)$, where

$$\begin{aligned} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) = &-\tfrac{1}{2}\log|D(\boldsymbol{\phi})| - \tfrac{\kappa^{-1}(v_i)}{2}(\boldsymbol{\epsilon}_i + \boldsymbol{\omega}_i)^T D^{-1}(\boldsymbol{\phi})(\boldsymbol{\epsilon}_i + \boldsymbol{\omega}_i) \\ &- \tfrac{1}{2}\log\phi_x - \tfrac{\kappa^{-1}(v_i)}{2\phi_x} z_i^2 + \log h(v_i; \boldsymbol{\nu}) + C, \end{aligned}$$

with $\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}z_i$, for $i = 1, \ldots, M$. We obtain

$$\begin{aligned} \frac{\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\boldsymbol{\mu}\partial\boldsymbol{\omega}_i^T}\bigg|_{\omega=\omega_0} &= \kappa^{-1}(v_i)D^{-1}(\boldsymbol{\phi}), \\ \frac{\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\omega}_i^T}\bigg|_{\omega=\omega_0} &= \kappa^{-1}(v_i)D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\epsilon}_i) \qquad \text{and} \\ \frac{\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\phi_x\partial\boldsymbol{\omega}_i^T}\bigg|_{\omega=\omega_0} &= \mathbf{0}. \end{aligned}$$

Perturbation on the measurements obtained for a particular instrument: suppose the interest is on perturbing the measurements obtained for the tth instrument, $t = 1, \ldots, p$. In this case one has $\boldsymbol{Y}_{i\omega} = \boldsymbol{Y}_i + \omega_i\boldsymbol{c}_t$, where $\boldsymbol{c}_t$ denotes a p-dimensional vector of zeros with one at the tth position. The perturbation vector is given by $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_M)^T$, $\boldsymbol{\omega}_0 = \mathbf{0} \in \mathbb{R}^M$ and the log-likelihood function for complete data assumes the form $L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) = \sum_{i=1}^{M} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)$, where

$$\begin{aligned} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) = &-\tfrac{1}{2}\log|D(\boldsymbol{\phi})| - \tfrac{\kappa^{-1}(v_i)}{2}(\boldsymbol{\epsilon}_i + \omega_i\boldsymbol{c}_t)^T D^{-1}(\boldsymbol{\phi})(\boldsymbol{\epsilon}_i + \omega_i\boldsymbol{c}_t) \\ &- \tfrac{1}{2}\log\phi_x - \tfrac{\kappa^{-1}(v_i)}{2\phi_x} z_i^2 + \log h(v_i; \boldsymbol{\nu}) + C, \end{aligned}$$

with $\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}z_i$, for $i = 1, \ldots, M$. Applying the differentials given in Appendix we obtain

$$\begin{aligned} \frac{\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\boldsymbol{\mu}\partial\omega_i}\bigg|_{\omega=\omega_0} &= \kappa^{-1}(v_i)D^{-1}(\boldsymbol{\phi})\boldsymbol{c}_t, \\ \frac{\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\boldsymbol{\phi}\partial\omega_i}\bigg|_{\omega=\omega_0} &= \kappa^{-1}(v_i)D^{-2}(\boldsymbol{\phi})D(\boldsymbol{\epsilon}_i)\boldsymbol{c}_t \qquad \text{and} \\ \frac{\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\phi_x\partial\omega_i}\bigg|_{\omega=\omega_0} &= \mathbf{0}. \end{aligned}$$

Multiplicative Bias Perturbation

Consider the perturbed model

$$\boldsymbol{Y}_i = \boldsymbol{\mu} + \boldsymbol{\omega}z_i + \boldsymbol{\epsilon}_i, \qquad i = 1, \ldots, M,$$

where $\boldsymbol{\omega} \in \mathbb{R}^p$. Influence diagnostics for this kind of model under Student-t errors have been studied by Galea, Bolfarine and Vilca (2005) with an application to the data set described by Barnett (1969) on lung vital capacity. Here $\boldsymbol{\omega}_0 = \mathbf{1}_p$ and the log-likelihood function for complete data is given by $L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) = \sum_{i=1}^{M} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)$, with

$$\begin{aligned} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) = & -\tfrac{1}{2}\log|D(\boldsymbol{\phi})| - \tfrac{\kappa^{-1}(v_i)}{2}(\boldsymbol{Y}_i - \boldsymbol{\mu} - \boldsymbol{\omega} z_i)^T D^{-1}(\boldsymbol{\phi})(\boldsymbol{Y}_i - \boldsymbol{\mu} - \boldsymbol{\omega} z_i) \\ & - \tfrac{1}{2}\log\phi_x - \tfrac{\kappa^{-1}(v_i)}{2\phi_x} z_i^2 + \log h(v_i; \boldsymbol{\nu}) + C, \end{aligned}$$

for $i = 1, \ldots, M$. One may show that $\partial^2 L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)/\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^T = \sum_{i=1}^{M} \partial^2 L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)/\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^T$, where

$$\begin{aligned} \frac{\partial^2 L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\boldsymbol{\mu}\partial\boldsymbol{\omega}^T} &= -\kappa^{-1}(v_i) z_i D^{-1}(\boldsymbol{\phi}), \\ \frac{\partial^2 L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\omega}^T} &= -\kappa^{-1}(v_i) z_i D^{-2}(\boldsymbol{\phi}) D(\boldsymbol{\epsilon}_i(\boldsymbol{\omega})) \qquad \text{and} \\ \frac{\partial^2 L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)}{\partial\phi_x\partial\boldsymbol{\omega}^T} &= \mathbf{0}, \end{aligned}$$

with $\boldsymbol{\epsilon}_i(\boldsymbol{\omega}) = \boldsymbol{Y}_i - \boldsymbol{\mu} - \boldsymbol{\omega} z_i$ for $i = 1, \ldots, M$, which must be evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$.

5 Application

We consider Grubbs' model given in (2) with the following hierarchical formulation:

$$\begin{aligned} \boldsymbol{Y}_i|z_i &\overset{\text{ind}}{\sim} SMN_5(\boldsymbol{\mu} + \mathbf{1} z_i, D(\boldsymbol{\phi}); H) \qquad \text{and} \\ z_i &\overset{\text{ind}}{\sim} SMN(0, \phi_x; H), \qquad i = 1, \ldots, 64, \end{aligned}$$

where $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{i5})^T$, $D(\boldsymbol{\phi}) = \text{diag}(\phi_1, \ldots, \phi_5)$ and H denote the distribution function for the mixture variable V_i, $i = 1, \ldots, 64$.

In our analyzes we suppose that the mixture variables follows a Gamma distribution, Beta, discrete and point mass in V_i, that is, the marginal response $\boldsymbol{Y}_i$ follows a Student-t, slash, contaminated normal and normal distribution, respectively. We set $\nu = 2.3$ and $\nu = 0.8$ for the degrees of freedom in the Student-t distribution and the slash distribution, respectively and $\epsilon = 0.15$ and $\gamma = 0.05$ for the parameters of the contaminated normal, such parameters were chosen a set of acceptable values. The plot of the profile log-likelihood function for these models is given in Figure 3.

Table 1
Parameter estimates of the fitted models to the thermocouples data.

Parameter	Normal	Student-t	Slash	Contaminated normal
μ_1	32608.3 (5.832)	32608.3 (5.262)	32608.2 (4.447)	32608.2 (5.318)
μ_2	32198.4 (6.650)	32198.6 (5.353)	32198.5 (4.504)	32198.5 (5.355)
μ_3	32604.8 (5.864)	32605.2 (5.295)	32605.1 (4.474)	32605.1 (5.350)
μ_4	32363.8 (5.739)	32363.8 (5.272)	32363.8 (4.462)	32363.8 (5.340)
μ_5	32290.7 (5.948)	32290.6 (5.310)	32290.5 (4.489)	32290.5 (5.348)
ϕ_1	1.8945 (0.322)	0.5190 (0.771)	0.2543 (1.688)	0.6029 (0.809)
ϕ_2	12.0975 (0.061)	1.3062 (0.412)	0.5549 (1.027)	1.0017 (0.589)
ϕ_3	2.2657 (0.281)	0.8023 (0.605)	0.3943 (1.334)	0.9515 (0.612)
ϕ_4	0.8156 (0.396)	0.6075 (0.720)	0.3320 (1.491)	0.8377 (0.670)
ϕ_5	3.2543 (0.207)	0.9291 (0.543)	0.4777 (1.158)	0.9273 (0.624)
log-likelihood	-753.495	-699.891	-699.740	-696.143

In Table 1 are presented the maximum likelihood estimates for the parameter vector $\boldsymbol{\theta}$ (standard errors in parenthesis) for the normal, Student-t, slash and contaminated normal models. The asymptotic variances of the parameter estimates were obtained from the Fisher information matrix presented in Appendix A. For the slash distribution we use the expressions given in Lange and Sinsheimer (1993) for obtaining the constants d_g and f_g, whereas for the contaminated normal distribution these expressions were obtained using a Laplace approximation.

The estimates of ϕ_x are 32.123; 21.137 and 11.578 and 27.673 for the normal, Student-t, slash and contaminated normal models, respectively. We see that the inference for the four fitted models is quite similar, but the estimates for the scale parameters are not comparable because there are in different scales. The power exponential model was also fitted to the data, but the shape parameter estimate does not seem to be satisfactory so the estimates were omitted. The likelihood ratio statistics for the Student-t, slash and contaminated normal models against the normal model, corresponding to $LR = 107.208$, $LR = 107.510$ and $LR = 114.704$, respectively, indicate that models based in heavy-tailed distributions provides a better fit than the normal model.

Note from Figure 4 that when we use distributions with tails heavier than the normal ones the ECM algorithm allows to accommodate such observations attributing to them small weights in the estimation procedure. The weights for the normal distribution ($\widehat{\kappa}_i = 1, \forall i$) are indicated in Figure 4 as a segmented line. All our results have been obtained using a library for S-Plus, available upon request from the first author. Next we identify influential observations for the thermocouples data set using the conformal curvature B_i. The perturbation schemes described in the previous section were considered. In all the influence graphics the benchmark proposed by Zhu and Lee (2001) was considered as cut-off value.

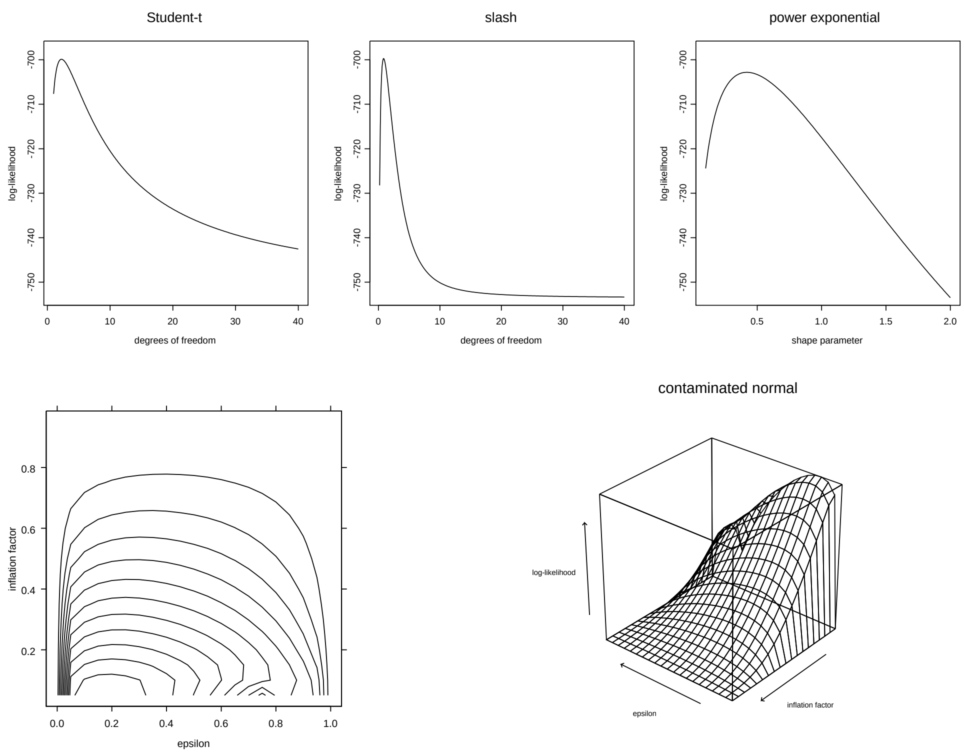

Fig. 3. Plot of the profile log-likelihood for the fitted models.

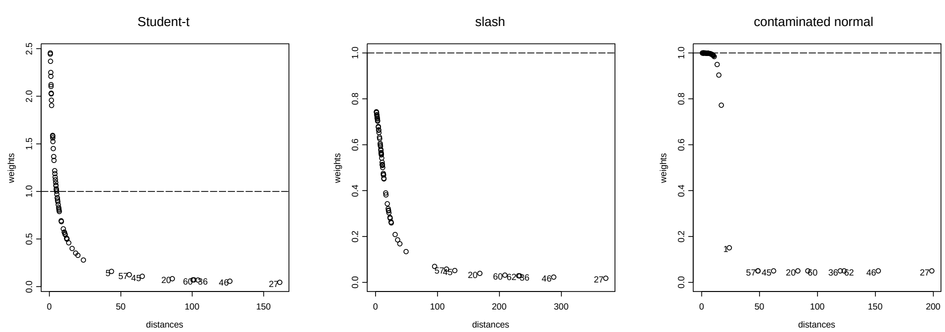

Fig. 4. Estimated weights versus distances for the Student-t, slash and contaminated normal models.

Case-weight perturbation: from Figures 5-7 is noted that under normal errors, the observations detected as outliers in Figure 1 are identified as influential on $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\mu}}^T, \widehat{\boldsymbol{\phi}}^T, \widehat{\phi}_x)^T$ and, in particular, observations 20 and 60 are influential on $\widehat{\boldsymbol{\mu}}$ as well as on $\widehat{\boldsymbol{\phi}}$. As espected, the influence of such observations is reduced when we consider distributions with heavier tails than the normal ones. For this data set the slash model with small degrees of freedom accommodates slightly better the influential observations.

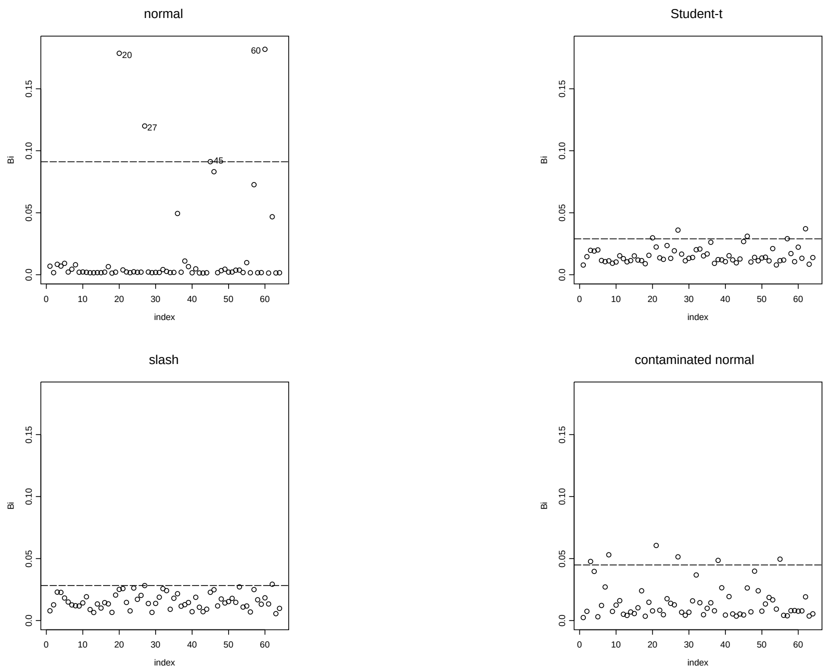

Fig. 5. Index plots of B_i for $\widehat{\boldsymbol{\theta}}$ under case-weight perturbation.

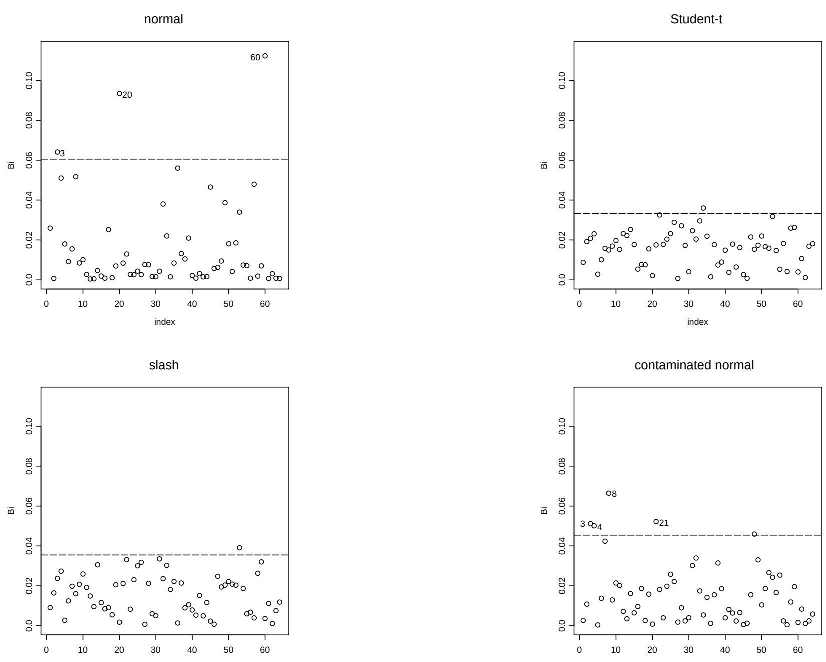

Fig. 6. Index plots of B_i for $\widehat{\boldsymbol{\mu}}$ under case-weight perturbation.

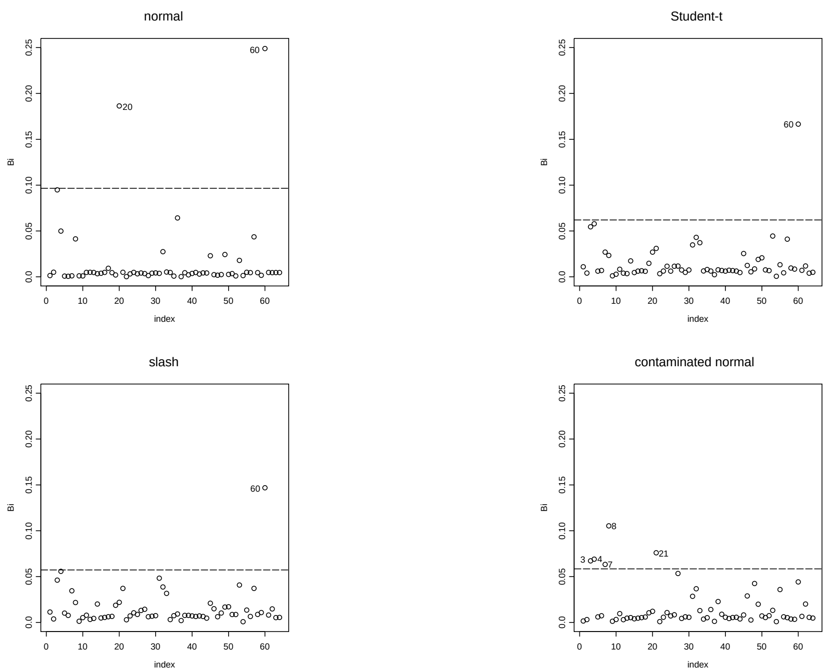

Fig. 7. Index plots of B_i for $\widehat{\boldsymbol{\phi}}$ under case-weight perturbation.

Joint measurement perturbation: the conformal curvatures $B_i(\boldsymbol{\theta}_1)$ with $\boldsymbol{\theta}_1 = (\boldsymbol{\mu}^T, \boldsymbol{\phi}^T)^T$ for those individuals detected as outliers for the Mahalanobis distance (see Figure 2) are presented in Figure 8 for the normal $(--)$, Student-t $(\cdots)$, slash (—) and contaminated normal $(\cdot - \cdot)$ models. Using this perturbation scheme we can examine the influence on the measurements *within* each items. In addition, we can show that exist a connection between the generalized leverage and local influence under this perturbation scheme (see, for example, Osorio, 2006 and Osorio, Paula and Galea, 2007). In Figure 8 is appreciated some influence when the measurements of items 20, 36, 45, 57 and 60 are perturbed under normal errors. This influence is reduced when we use distributions with heavier tails than the normal ones.

Multiplicative bias perturbation: Figure 9 presents the conformal curvatures for the three fitted models. We note that for the considered models, the values of B_i are quite different, suggesting that the assumption of equality of multiplicative biases is not supported by the data. This conclusion agrees with the results of the hypothesis test described in Christensen and Blackwood (1993). This result indicates that the comparative calibration model (Barnett, 1969) may be more appropriate to modelling this data set. Therefore, the importance of this perturbation scheme is that of allowing to criticize the model building.

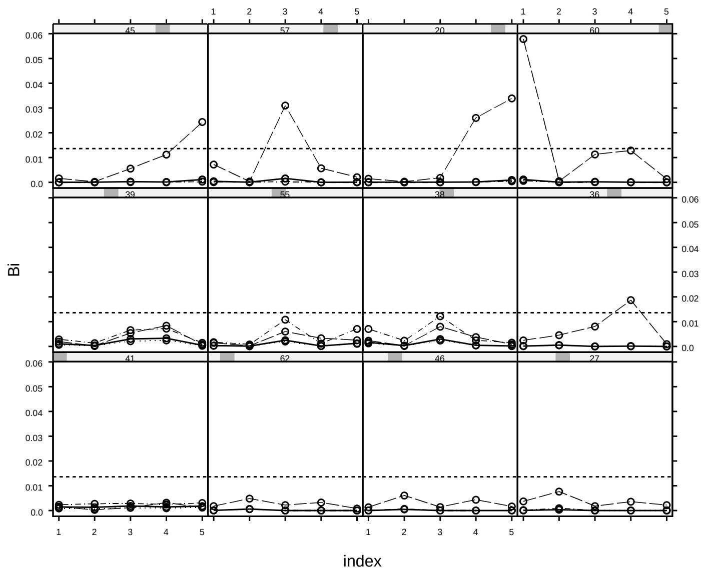

Fig. 8. Index plots of B_i for $\widehat{\boldsymbol{\theta}}_1$ under joint measurement perturbation.

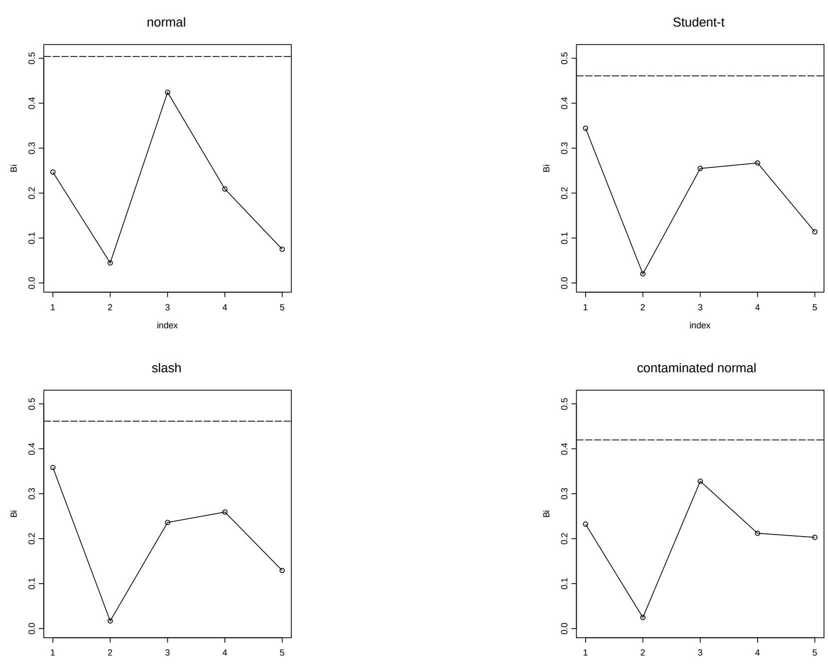

Fig. 9. Index plots of B_i for $\widehat{\boldsymbol{\theta}}$ under multiplicative bias perturbation.

6 Concluding Remarks

In this work we have discussed the parameter estimation in the Grubbs' model under a class of distributions that presents heavier tails than the normal ones. Through a local influence study some aspects of robustness of the maximum likelihood estimators under the scale mixture of normal distributions were noted. Explicit expressions are obtained for matrix $\boldsymbol{\Delta}$ under different perturbation schemes considered. It is noted, however that other perturbation schemes can be considered in analogous way. The results presented in this work represent an extension of the work by Lachos, Vilca and Galea (2007). It is important to emphasize the capacity of such models to attenuate outlying observations, by means of attributing to these observations a small weight in the estimation process. The results derived in this work agree with the considerations that in this respect are presented in Lange and Sinsheimer (1993) as well as with the comments of Pinheiro, Liu and Wu (2001) and Rosa, Padovani and Gianola (2003, 2004) for the linear mixed-effects model and Galea, Bolfarine and Vilca (2005) for comparative calibration models.

Acknowledgements

This work was partially supported by the FONDECYT grants 11075071 and 1070919, Chile and CNPq and FAPESP, Brazil. The authors thank the Associate Editor and a referee for valuable suggestions.

A Fisher Information Matrix

From the equation (5) it is possible to show that the marginal log-likelihood function for $\boldsymbol{Y}$ is given by $L(\boldsymbol{\theta}) = \sum_{i=1}^{M} L_i(\boldsymbol{\theta})$, with

$$L_i(\boldsymbol{\theta}) = -\frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\kappa^{-1}(v_i)(\boldsymbol{Y}_i - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}_i - \boldsymbol{\mu}) + C,$$

where $\boldsymbol{\Sigma} = D(\boldsymbol{\phi}) + \phi_x\mathbf{1}\mathbf{1}^T$ and C is a constant. Thus, following Galea (1995), the Fisher information matrix for $\boldsymbol{\theta}$ assumes the block diagonal form

$$\boldsymbol{K}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{K}(\boldsymbol{\mu}) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{K}(\boldsymbol{\psi}) \end{pmatrix},$$

where

$$\boldsymbol{K}(\boldsymbol{\psi}) = \begin{pmatrix} \boldsymbol{K}(\boldsymbol{\phi}) & \boldsymbol{K}(\boldsymbol{\phi}, \phi_x) \\ \boldsymbol{K}^T(\boldsymbol{\phi}, \phi_x) & K(\phi_x) \end{pmatrix},$$

with

$$\begin{aligned} \boldsymbol{K}(\boldsymbol{\mu}) &= \frac{4d_g}{p}\boldsymbol{\Sigma}^{-1}, \quad K(\phi_x) = \left(\frac{s-1}{s}\right)^2 \frac{c_1 + c_2}{\phi_x^2}, \\ \boldsymbol{K}(\boldsymbol{\phi}) &= c_1 \boldsymbol{h}\boldsymbol{h}^T + c_2\{\boldsymbol{b}\boldsymbol{b}^T + D(\boldsymbol{d})\}, \\ \boldsymbol{K}(\boldsymbol{\phi}, \phi_x) &= \left(\frac{s-1}{s}\right)\frac{c_1}{\phi_x} D^{-1}(\boldsymbol{\phi}) + \frac{c_2 - (s-1)c_1}{s^2} D^{-2}(\boldsymbol{\phi}) \end{aligned}$$

and

$$\boldsymbol{h} = D^{-1}(\boldsymbol{\phi})\mathbf{1} - \tau D^{-2}(\boldsymbol{\phi})\mathbf{1}, \quad \boldsymbol{b} = \tau D^{-1}(\boldsymbol{\phi})\mathbf{1}, \quad \boldsymbol{d} = D^{-2}(\boldsymbol{\phi})\mathbf{1} - 2\tau D^{-3}(\boldsymbol{\phi})\mathbf{1},$$

where $c_1 = -\frac{1}{4} + \frac{f_g}{p(p+2)}$, $c_2 = \frac{2f_g}{p(p+2)}$, with $d_g = \mathrm{E}\{W_g^2(U)U\}$, $f_g = \mathrm{E}\{W_g^2(U)U^2\}$, $U = ||\boldsymbol{Z}||^2$ and $\boldsymbol{Z}$ has a multivariate spherical distribution, i.e. (see Fang, Kotz and Ng, 1990) $\boldsymbol{Z} \sim EC_p(\mathbf{0}, \boldsymbol{I}; g)$.

B Differential Calculations

In this appendix the differentials $\mathrm{d}_\theta^2\, Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})$ and $\mathrm{d}_{\theta\omega}^2\, L(\boldsymbol{\theta}|\boldsymbol{Y}_c)$ will be derived for various perturbation schemes. The necessary matrices $\ddot{\boldsymbol{Q}}$ and $\boldsymbol{\Delta}$ for the normal curvature evaluation for model (5) may be obtained from the differentials by applying some theorems of identification given in Magnus and Neudecker (1988).

B.1 Hessian Matrix, $\ddot{\boldsymbol{Q}}$

Using results of matrix differentiation one has that the second differential of $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})$ with respect to $\boldsymbol{\theta}$ is given by $\mathrm{d}_\theta^2\, Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^M \mathrm{d}_\theta^2\, Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})$, with

$$\begin{aligned} \mathrm{d}_\mu^2\, Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) &= -\widehat{\kappa}_i(\mathrm{d}\,\boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi})\,\mathrm{d}\,\boldsymbol{\mu}, \qquad \text{and} \\ \mathrm{d}_{\mu\phi}^2\, Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) &= -\widehat{\kappa}_i(\mathrm{d}\,\boldsymbol{\mu})^T D^{-2}(\boldsymbol{\phi}) D(\widehat{\boldsymbol{\epsilon}}_i)\,\mathrm{d}\,\boldsymbol{\phi}, \end{aligned}$$

where $\widehat{\boldsymbol{\epsilon}}_i = \boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}\widehat{z}_i$, $i = 1, \ldots, M$. The differentials of $Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})$ with respect to $\boldsymbol{\phi}$ lead to

$$\begin{aligned} \mathrm{d}_\phi^2\, Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) = \tfrac{1}{2}(\mathrm{d}\,\boldsymbol{\phi})^T D^{-2}(\boldsymbol{\phi})\,\mathrm{d}\,\boldsymbol{\phi} - \widehat{\tau}(\mathrm{d}\,\boldsymbol{\phi})^T D^{-3}(\boldsymbol{\phi})\,\mathrm{d}\,\boldsymbol{\phi} \\ - \widehat{\kappa}_i(\mathrm{d}\,\boldsymbol{\phi})^T D(\widehat{\boldsymbol{\epsilon}}_i) D^{-3}(\boldsymbol{\phi}) D(\widehat{\boldsymbol{\epsilon}}_i)\,\mathrm{d}\,\boldsymbol{\phi}. \end{aligned}$$

Finally, taking differencials of $Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})$ with respect to ϕ_x we find

$$\mathrm{d}^2_{\phi_x} Q_i(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) = \tfrac{1}{2\phi_x^2}\,\mathrm{d}^2\phi_x - \tfrac{1}{\phi_x^3}(\widehat{\kappa}_i\widehat{z}_i^2 + \widehat{\tau})\,\mathrm{d}^2\phi_x.$$

The hessian matrix $\ddot{\boldsymbol{Q}}$ is obtained by evaluating the differentials above at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$.

B.2 Perturbation Schemes

The differential $\mathrm{d}^2_{\theta\omega} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)$ will be derived for each perturbation scheme. Evaluating these differentials at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ one may obtain $\partial^2 Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^T$ for the perturbation scheme under study. It is assumed in the calculations below that one may exchange integration and differentiation operators.

Case-weight Perturbation

For this perturbation scheme we obtain the differentials

$$\begin{aligned}
\mathrm{d}^2_{\mu\omega_i} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= \kappa^{-1}(v_i)(\mathrm{d}\,\boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi})(\boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}z_i)\,\mathrm{d}\,\omega_i,\\
\mathrm{d}^2_{\phi\omega_i} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= -\tfrac{1}{2}(\mathrm{d}\,\boldsymbol{\phi})^T\{D^{-1}(\boldsymbol{\phi})\mathbf{1} - \kappa^{-1}(v_i)D(\boldsymbol{\epsilon}_i)D^{-2}(\boldsymbol{\phi})\boldsymbol{\epsilon}_i\}\,\mathrm{d}\,\omega_i \quad \text{and}\\
\mathrm{d}^2_{\phi_x\omega_i} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= -\tfrac{1}{2}\Big\{\tfrac{1}{\phi_x} + \tfrac{\kappa^{-1}(v_i)}{\phi_x^2}\,z_i^2\Big\}\,\mathrm{d}\,\phi_x\,\mathrm{d}\,\omega_i,
\end{aligned}$$

where $\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}z_i$, for $i = 1, \ldots, M$.

Joint perturbation on the measurements obtained for the p instruments

Using the differential methodology, we find

$$\begin{aligned}
\mathrm{d}^2_{\mu\omega_i} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= \kappa^{-1}(v_i)(\mathrm{d}\,\boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi})\,\mathrm{d}\,\boldsymbol{\omega}_i,\\
\mathrm{d}^2_{\phi\omega_i} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= \kappa^{-1}(v_i)(\mathrm{d}\,\boldsymbol{\phi})^T D(\boldsymbol{\epsilon}_i + \boldsymbol{\omega}_i)D^{-2}(\boldsymbol{\phi})\,\mathrm{d}\,\boldsymbol{\omega}_i \qquad \text{and}\\
\mathrm{d}^2_{\phi_x\omega_i} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= \mathbf{0}.
\end{aligned}$$

with $\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1}z_i$, for $i = 1, \ldots, M$.

Perturbation on the masurements obtained for a particular instrument

Differentiating $L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)$ with respect to $\boldsymbol{\theta}$ and ω_i, we obtain

$$\begin{aligned}
\mathrm{d}^2_{\mu\omega_i} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= \kappa^{-1}(v_i)(\mathrm{d}\,\boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi})\boldsymbol{c}_t\,\mathrm{d}\,\omega_i,\\
\mathrm{d}^2_{\phi\omega_i} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= \kappa^{-1}(v_i)(\mathrm{d}\,\boldsymbol{\phi})^T D(\boldsymbol{\epsilon}_i + \omega_i\boldsymbol{c}_t)D^{-2}(\boldsymbol{\phi})\boldsymbol{c}_t\,\mathrm{d}\,\omega_i \qquad \text{and}\\
\mathrm{d}^2_{\phi_x\omega_i} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= 0.
\end{aligned}$$

with $\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - \boldsymbol{\mu} - \mathbf{1} z_i$, for $i = 1, \ldots, M$.

Perturbation on the multiplicative biases

In this case we obtain $\mathrm{d}^2_{\theta\omega} L(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) = \sum_{i=1}^{M} \mathrm{d}^2_{\theta\omega} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c)$, where

$$\begin{aligned}
\mathrm{d}^2_{\mu\omega} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= -\kappa^{-1}(v_i) z_i (\mathrm{d}\,\boldsymbol{\mu})^T D^{-1}(\boldsymbol{\phi})\,\mathrm{d}\,\boldsymbol{\omega}, \\
\mathrm{d}^2_{\phi\omega} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= -\kappa^{-1}(v_i) z_i (\mathrm{d}\,\boldsymbol{\phi})^T D(\boldsymbol{\epsilon}_{i\omega}) D^{-2}(\boldsymbol{\phi})\,\mathrm{d}\,\boldsymbol{\omega} \qquad \text{and} \\
\mathrm{d}^2_{\phi_x\omega} L_i(\boldsymbol{\theta}, \boldsymbol{\omega}|\boldsymbol{Y}_c) &= \mathbf{0}
\end{aligned}$$

with $\boldsymbol{\epsilon}_{i\omega} = \boldsymbol{Y}_i - \boldsymbol{\mu} - \boldsymbol{\omega} z_i$, for $i = 1, \ldots, M$.

References

Andrews, D.F., Mallows, C.L., 1974. Scale mixtures of normal distributions. Journal of the Royal Statistical Society, Ser. B 36, 99-102.

Barnett, V.D., 1969. Simultaneous pairwise linear structural relationships. Biometrics 25, 129-142.

Bedrick, E.J., 2001. An efficient scores test for comparing several measuring devices. Journal of Quality Technology 33, 96-103.

Bolfarine, H., Galea, M., 1995. Structural comparative calibration using the EM algorithm. Journal of Applied Statistics 2, 277-292.

Bolfarine, H., Galea, M., 1996. One structural comparative calibration under a t-model. Computational Statistics 11, 63-85.

Chatterjee, S., Hadi, A.S., 1988. Sensitivity Analysis in Linear Regression. Wiley, New York.

Christensen, R., Blackwood, L.G., 1993. Test for precision and accuracy of multiple measuring devices. Technometrics 35, 411-420.

Cook, R.D., 1986. Assessment of local influence" (with discussion). Journal of the Royal Statistical Society, Ser. B 48, 133-169.

Cook, R.D., 1997. Local influence. In S. Kotz, C.B. Read and D.L. Banks (Eds.), Encyclopedia of Statistical Sciences, Update, Vol. 1, Wiley, New York, 380-385.

Cook, R.D., Weisberg, S., 1982. Residuals and Influence in Regression. Chapman & Hall, London.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm (with discussion). Journal of the Royal Statistical Society, Ser. B 39, 1-38.

Fang, K.T., Kotz, S., Ng, K.W., 1990. Symmetric Multivariate and Related Distributions. Chapman & Hall, London.

Fernández, C. and Steel, M.F.J., 1999. Multivariate Student-t regression models: pitfalls and inference. Biometrika 86, 153-167.

Galea, M. Calibração Comparativa Estrutural e Funcional. Unpublished Ph.D. dissertation. (Dept. of Statistics, University of São Paulo, Brazil, 1995).

Galea, M., Bolfarine, H., de Castro, M., 2002. Local influence in comparative calibration models. Biometrical Journal 44, 59-81.

Galea, M., Bolfarine, H., Vilca, F., 2005. Local influence in comparative calibration models under elliptical t-distributions. Biometrical Journal 47, 691-706.

Grubbs, F.E., 1948. On estimating precision of measuring instruments and product variability. Journal of the American Statistical Association 43, 243-264.

Grubbs, F.E., 1973. Errors of measurement, precision, accuracy, and the statistical comparison of measuring instruments. Technometrics 15, 53-66.

Grubbs, F.E., 1983. Grubbs' estimators (precision and accuracy of measurement. In S. Kotz and N.L. Johnson (Eds.), Encyclopedia of Statistical Sciences, Vol. 3, Wiley, New York, 542-549.

Huber, P.J., 1981. Robust Statistics. Wiley, New York.

Kass, R.E., 1997. Laplace's method. In S. Kotz, C.B. Read and D.L. Banks (Eds.), (Eds.), Encyclopedia of Statistical Sciences, Update, Vol. 1, Wiley, New York, 347-354.

Lachos, V.H., Vilca, F., Galea, M., 2007. Influence diagnostics for the Grubbs' model. Statistical Papers 48, 419-436.

Lange, K.L., Little, R.J.A. and Taylor, J.M.G., 1989. Robust statistical modeling using the t distribution. Journal of the American Statistical Association 84, 881-896.

Lange, K., Sinsheimer, J.S., 1993. Normal/Independent distributions and their applications in robust regression. Journal of Computational and Graphical Statistics 2, 175-198.

Lee, S.Y., Xu, L., 2004. Influence analysis of nonlinear mixed-effects models. Computational Statistics & Data Analysis 45, 321-341.

Liu, C., 1996. Bayesian robust multivariate linear regression with incomplete data. Journal of the American Statistical Association 91, 1219-1227.

Lucas, A., 1997. Robustness of the Student t based M-estimator. Communications in Statistics - Theory and Methods 26, 1165-1182.

Magnus, J.R., Neudecker, H., 1988. Matrix Differential Calculus with Applications in Statistics and Econometrics. Wiley, Chichester.

McLachlan, G.L., Krishnan, T., 1997. The EM Algorithm and Extensions. Wiley, New York.

Meng, X., Rubin, D.B., 1993. Maximum likelihood estimation via the ECM algorithm: a general framework. Biometrika 80, 267-278.

Osorio, F., Diagnóstico de Influência em Modelos Elípticos com Efeitos Mistos. Unpublished Ph.D. dissertation. (Dept. of Statistics, University of São Paulo, Brazil, 2006).

Osorio, F., Paula, G.A., Galea, M., 2007. Assessment of local influence in elliptical linear models with longitudinal structure. Computational Statistics & Data Analysis 51, 4354-4368.

Pinheiro, J.C., Liu, C., Wu, Y.N., 2001. Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. Journal of Computational and Graphical Statistics 10, 249-276.

Poon, W., Poon, Y.S., 1999. Conformal normal curvature and assessment of local influence. Journal of the Royal Statistical Society, Ser. B 61, 51-61.

Rosa, G.J.M., Padovani, C.R., Gianola, D., 2003. Robust linear mixed models with Normal/Independent distributions and Bayesian MCMC implementation. Biometrical Journal 45, 573-590.

Rosa, G.J.M., Padovani, C.R., Gianola, D., 2004. Bayesian longitudinal data analysis with mixed models and thick-tailed distributions using MCMC. Journal of Applied Statistics 31, 855-873.

Shyr, J.Y., Gleser, L.J., 1986. Inference about comparative precision in linear structural relationships. Journal of Statistical Planning and Inference 14, 339-358.

Theobald, C.M., Mallison, J.R., 1978. Comparative calibration, linear structural relationships and congeneric measurements. Biometrics 34, 39-45.

West, M., 1987. On scale mixture of normal distributions. Biometrika 74, 646-648.

Zhu, H., Lee, S.Y., 2001. Local influence for incomplete-data models. Journal of the Royal Statistical Society, Ser. B 63, 111-126.