Journal of Multivariate Analysis 102 (2011) 1280–1292

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Estimates of MM type for the multivariate linear model

Nadia L. Kudraszow [a,*], Ricardo A. Maronna [b]

[a] CONICET, University of La Plata, Argentina

[b] C.I.C.P.B.A, University of La Plata, Argentina

ARTICLE INFO

Article history:
Received 10 June 2010
Available online 5 May 2011

AMS 2000 subject classifications:
62F35
62H12
62J05

Keywords:
Robust methods
MM-estimate
Multivariate linear model

ABSTRACT

We propose a class of robust estimates for multivariate linear models. Based on the approach of MM-estimation (Yohai 1987, [24]), we estimate the regression coefficients and the covariance matrix of the errors simultaneously. These estimates have both a high breakdown point and high asymptotic efficiency under Gaussian errors. We prove consistency and asymptotic normality assuming errors with an elliptical distribution. We describe an iterative algorithm for the numerical calculation of these estimates. The advantages of the proposed estimates over their competitors are demonstrated through both simulated and real data.

1. Introduction

Consider a multivariate linear model (MLM) with random predictors, i.e., we observe n independent identically distributed (i.i.d.) $(p+q)$-dimensional vectors, $\mathbf{z}_i = (\mathbf{y}_i', \mathbf{x}_i')$ with $1 \leq i \leq n$, where $\mathbf{y}_i = (y_{i1}, \ldots, y_{iq})' \in \mathbb{R}^q, \mathbf{x}_i = (x_{i1}, \ldots, x_{ip})' \in \mathbb{R}^p$, and $'$ denotes the transpose. The $\mathbf{y}_i$ are the response vectors and the $\mathbf{x}_i$ are the predictors, and both satisfy the equation

$$\mathbf{y}_i = \mathbf{B}_0'\mathbf{x}_i + \mathbf{u}_i \quad 1 \leq i \leq n, \tag{1.1}$$

where $\mathbf{B}_0 \in \mathbb{R}^{p\times q}$ is the matrix of the regression parameters and $\mathbf{u}_i$ is a q-dimensional vector independent of $\mathbf{x}_i$. If $\mathbf{x}_{ip} = 1$ for all $1 \leq i \leq n$, we obtain a regression model with intercept.

We denote the distributions of $\mathbf{x}_i$ and $\mathbf{u}_i$ by G_0 and F_0, respectively, and $\mathbf{\Sigma}_0$ is the covariance matrix of the $\mathbf{u}_i$. The p-multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$ is denoted by $N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$.

In the case of $\mathbf{u}_i$ with distribution $N_q(\mathbf{0}, \mathbf{\Sigma}_0)$, the maximum likelihood estimate (MLE) of $\mathbf{B}_0$ is the least squares estimate (LSE), and the MLE of $\mathbf{\Sigma}_0$ is the sample covariance matrix of the residuals. It is known that these estimates are not robust: a small fraction of outliers may have a large effect on their values.

Several approaches have been proposed to deal with this problem. The first proposal of a robust estimate for the MLM was given by Koenker and Portnoy [13]. They proposed to apply a regression M-estimator, based on a convex loss function, to each coordinate of the response vector. The problems with this estimate is lack of affine equivariance and zero breakdown point. Several other estimates without these problems were defined later. Rousseeuw et al. [20] proposed estimates for the MLM based on a robust estimate of the covariance matrix of $\mathbf{z} = (\mathbf{x}', \mathbf{y}')$. Bilodeau and Duchesne [4] extended the S-estimates introduced by Davies [6] for multivariate location and scatter; then Van Aelst and Willems [23] studied the robustness of

* Corresponding author.
E-mail addresses: nkudraszow@mate.unlp.edu.ar (N.L. Kudraszow), rmaronna@retina.ar (R.A. Maronna).

0047-259X/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2011.04.011

these estimators. Agulló et al. [1] extended the minimum covariance determinant estimate introduced by Rousseeuw [19] and Roelandt et al. [18] extended the definition of GS-estimates introduced by Croux et al. [5]. These estimates have a high breakdown point but are not highly efficient when the errors are Gaussian and q is small. In order to solve this, Agulló et al. [1] improved the efficiency of their estimates, maintaining their high breakdown point, by considering one-step reweighting and one-step Newton–Raphson GM-estimates. García Ben et al. [7] extended τ-estimates for multivariate regression, obtaining a estimate with high breakdown point and a high Gaussian efficiency. Another important approach to obtain robust and efficient estimates is constrained M (CM) estimation, proposed by Mendes and Tyler in [17] for regression and by Kent and Tyler in [12] for multivariate location and scatter. The bias of CM-estimates for regression was studied by Berrendero et al. in [3]. Following this approach, Bai et al. in [2] proposed CM-estimates for the multivariate linear model.

In this paper, we propose robust estimates for the linear model based on the MM approach, first proposed by Yohai [24] for the univariate linear model, and later by Lopuhaä [15], Tatsuoka and Tyler [22], and Salibián-Barrera et al. [21] for multivariate location and scatter. We show that our estimates have both a high breakdown point and high normal efficiency.

In Section 2, we define MM-estimates for the MLM and prove some properties. In Sections 3 and 4, we study their breakdown point and influence function. In Sections 5 and 6, we study the asymptotic properties (consistency and asymptotic normality) of the MM-estimates assuming random predictors and errors with an elliptical unimodal distribution. In Section 7, we describe a computing algorithm based on an iterative weighted MLE. In Section 8, we present the results of a simulation study, and we present a real example in Section 9. All the proofs can be found in [14].

2. Definition and properties

Before defining our class of robust estimates for the MLM, we will define a robust estimate of scale.

Definition 1. Given a sample of size n, $\mathbf{v} = (v_1, \ldots, v_n)$, an *M-estimate of scale* $s(\mathbf{v})$ is defined as the value of s that is solution of

$$\frac{1}{n}\sum_{i=1}^{n}\rho_0\left(\frac{v_i}{s}\right) = b, \tag{2.1}$$

where $b \in (0, 1)$, or $s = 0$ if $\sharp(v_i = 0) \geq n(1 - b)$, where $\sharp$ is the symbol for cardinality.

In this paper, we use $b = 0.5$, which ensures the maximal asymptotic breakdown point (see [10]).

The function ρ_0 should satisfy the following definition.

Definition 2. A *ρ-function* will denote a function $\rho(u)$ which is a continuous nondecreasing function of $|u|$ such that $\rho(0) = 0$, $\sup_u \rho(u) = 1$, and $\rho(u)$ is increasing for nonnegative u such that $\rho(u) < 1$.

Note that according to the terminology of Maronna et al. [16] this would be a "bounded ρ-function". A popular ρ-function is the *bisquare function*:

$$\rho_B(u) = 1 - (1 - u^2)^3 I(|u| \leq 1), \tag{2.2}$$

where $I(\cdot)$ is the indicator function.

Definition 3. Given a vector $\mathbf{u}$ and a positive definite matrix $\mathbf{V}$, the *Mahalanobis norm of* $\mathbf{u}$ *with respect to* $\mathbf{V}$ is defined as

$$d(\mathbf{u}, \mathbf{V}) = (\mathbf{u}'\mathbf{V}^{-1}\mathbf{u})^{1/2}.$$

For particular given $\mathbf{B} \in \mathbb{R}^{p\times q}$ and $\mathbf{\Sigma} \in \mathbb{R}^{q\times q}$, we denote by $d_i(\mathbf{B}, \mathbf{\Sigma})$ $(i = 1, \ldots, n)$ the Mahalanobis norms of the residuals with respect to the matrix $\mathbf{\Sigma}$; that is,

$$d_i(\mathbf{B}, \mathbf{\Sigma}) = (\widehat{\mathbf{u}}_i(\mathbf{B})'\mathbf{\Sigma}^{-1}\widehat{\mathbf{u}}_i(\mathbf{B}))^{1/2},$$

with $\widehat{\mathbf{u}}_i(\mathbf{B}) = \mathbf{y}_i - \mathbf{B}'\mathbf{x}_i$.

Using the concepts defined above, we can describe an *MM-estimate for the MLM* by the following procedure.

Let $(\widetilde{\mathbf{B}}_n, \widetilde{\mathbf{\Sigma}}_n)$ be an initial estimate of $(\mathbf{B}_0, \mathbf{\Sigma}_0)$, with a high breakdown point and such that $|\widetilde{\mathbf{\Sigma}}_n| = 1$, where $|\widetilde{\mathbf{\Sigma}}_n|$ is the determinant of $\widetilde{\mathbf{\Sigma}}_n$. Compute the Mahalanobis norms of the residuals using $(\widetilde{\mathbf{B}}_n, \widetilde{\mathbf{\Sigma}}_n)$,

$$d_i(\widetilde{\mathbf{B}}_n, \widetilde{\mathbf{\Sigma}}_n) = (\widehat{\mathbf{u}}_i'(\widetilde{\mathbf{B}}_n)\widetilde{\mathbf{\Sigma}}_n^{-1}\widehat{\mathbf{u}}_i(\widetilde{\mathbf{B}}_n))^{1/2} \qquad 1 \leq i \leq n. \tag{2.3}$$

Then, compute the M-estimate of scale $\hat{\sigma}_n := s(\mathbf{d}(\widetilde{\mathbf{B}}_n, \widetilde{\mathbf{\Sigma}}_n))$ of the above norms, defined by (2.1), using the function ρ_0 as specified in Definition 2 and $b = 0.5$.

Let ρ_1 be another ρ-function such that

$$\rho_1 \leq \rho_0, \tag{2.4}$$