Statistical Methodology 16 (2014) 14–31

Contents lists available at ScienceDirect

Statistical Methodology

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

On influence diagnostics in elliptical multivariate regression models with equicorrelated random errors

Germán Ibacache-Pulgar [a], Gilberto A. Paula [b], Manuel Galea [c,*]

[a] *Departamento de Estadística, Universidad de Valparaíso, Chile*
[b] *Departamento de Estatística, Universidade de São Paulo, Brazil*
[c] *Pontificia Universidad Católica de Chile, Chile*

A R T I C L E I N F O

Article history:
Received 12 June 2012
Received in revised form
14 June 2013
Accepted 15 June 2013

Keywords:
Multivariate model
Elliptical distribution
Equicorrelation matrix
Local influence

A B S T R A C T

In this paper we discuss estimation and diagnostic procedures in elliptical multivariate regression models with equicorrelated random errors. Two procedures are proposed for the parameter estimation and the local influence curvatures are derived under some usual perturbation schemes to assess the sensitivity of the maximum likelihood estimates (MLEs). Two motivating examples preliminarily analyzed under normal errors are reanalyzed considering appropriate elliptical distributions. The local influence approach is used to compare the sensitivity of the model estimates.

1. Introduction

Diagnostic methods for regression models under normal errors have been largely investigated in the statistical literature. The majority of the works have given emphasis on studying the effect of eliminating observations on the results from the fitted model, particularly on the parameter estimates; see, for instance [2–5].

Since case deletion methods do not directly reflect the impact of other perturbations in the model, [6] proposed an interesting method, named local influence, to assess the effect of small perturbations in the model (or data) on the parameter estimates. The local influence analysis does not involve recomputing the parameter estimates for every case deletion, so it is often computationally

* Corresponding author. Tel.: +56 02 2354 5973.
E-mail address: mgalea@mat.puc.cl (M. Galea).

1572-3127/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.stamet.2013.06.001

simpler. Several authors have applied the local influence method to various regression models. For example, [10,18] extended the local influence methodology to univariate linear regression models with symmetric errors, whereas [8,19] discussed the local influence methodology in multivariate linear regression models with elliptical errors. [12] derived local influence curvatures in symmetric nonlinear models and, more recently, [14,15] derived the local influence curvatures in semiparametric elliptical models, whereas [17] presented the curvature calculations for a class of multivariate elliptical models with a general parametrization. Even though the equicorrelation structure may be derived from this general parametrization, sensitivity studies are not known for equicorrelated errors.

The aim of this paper is to apply the approach of local influence in elliptical multivariate regression models with equicorrelated random errors. The paper is presented as follows. Section 2 contains two motivating examples analyzed under normal multivariate regression models. In Section 3 we introduce the elliptical multivariate regression models with equicorrelated random errors. Section 4 contains the estimation and inference procedures for the model parameters. The local influence methodology is discussed in Section 5 under some usual perturbation schemes. In Section 6, the two motivating examples are reanalyzed under elliptical multivariate regression models with equicorrelated random errors, specifically, under Student-t errors. The local influence diagnostics developed in the paper are applied for the purpose of comparison between the fitted models under normal and heavy-tailed errors. The final section presents some concluding remarks.

2. Motivating examples

In this section two real data sets with multivariate structures and presence of outlying observations are analyzed under normal multivariate regression models with equicorrelated random errors.

2.1. Chilean Pension Funds

As a first illustration, we will consider a data set on the Chilean Pension Funds. The data correspond to monthly returns of pension funds of four Pension Fund Administrators (PFAs), namely Habitat, Planvital, Provida and Santa Maria and correspond to the period from January 1999 to December 2003. The monthly return of the Selective Index of Share Prices (IPSA) was used as explanatory variable. Fig. 1 displays the relationship between the monthly return of each pension fund and the IPSA return.

In this descriptive analysis we notice linear tendencies between the return of each pension fund and IPSA. Additionally, we see in Table 1 that the sample correlations (also the covariances) among the pension funds are similar, suggesting a covariance structure proportional to the equicorrelation matrix among the pension fund returns in the period. The likelihood ratio test to assess absence of equicorrelation under a normal multivariate model with equicorrelated random errors (see, for instance, [23]) leads to a p-value of 0.054. Thus, there is indication that a covariance structure proportional to the equicorrelation matrix seems to be appropriate.

In order to identify influential observations under the normal multivariate regression model, some influence graphs are presented in Fig. 2. Fig. 2(a) exhibits the index plot of Cook's distance and we notice observations $\{11, 12, 45\}$ as potentially influential on the regression coefficient estimates. Fig. 2(b) exhibits the index plot of the scale ratio measure indicating observations $\{6, 33, 52\}$ as potentially influential on the estimate of the variance–covariance matrix for the regression coefficient estimates. Details on the calculation of these diagnostic measures under multivariate regression models with equicorrelated random errors are given, for instance, in [13]. In Section 6 we reanalyze this example under heavy-tailed error models.

2.2. Testicular volume data

The second illustration corresponds to a data set in which measurements of the testicular volume of 42 adolescents were obtained through five different techniques: ultrasound (US), graphic method (I), dimensional measurement (II), prader orchidometer (III) and ring orchidometer (IV) (see [11]).

From the sample correlations (covariances), given in Table 2, a covariance structure proportional to the equicorrelation matrix among the five techniques of the testicular volume seems to be