SENSITIVITY ANALYSIS IN LINEAR REGRESSION

José A. Díaz-García, Graciela González-Farías and Victor M. Alvarado-Castro

Comunicación Técnica No I-05-05/11-04-2005
(PE/CIMAT)

Sensitivity analysis in linear regression

José A. Díaz-García
Universidad Autónoma Agraria Antonio Narro
Department of Statistics and Computation
25315 Buenavista, Saltillo, Coahuila, México.
E-mail: jadiaz@uaaan.mx
Graciela González-Farías
Centro de Investigación en Matemáticas, A.C.
Department of Probability and Statistics
Jalisco S/N, Mineral de Valenciana
36240 Guanajuato, Guanajuato, México.
E-mail: farias@cimat.mx
and
Víctor M. Alvarado-Castro
Universidad Autónoma de Guerrero
Av. Lázaro Cárdenas s/n
39070 Chilpancingo, Gro., México.
E-mail: alvarado@cimat.mx

Abstract

Based on a multivariate linear regression model, we propose several generalizations to the multivariate classical and modified Cook's distances in order to detect one or more of influential observations including the case of linear transformations of the estimated regression parameter. For those distances, we derived the exact distributions and point out a method to extend the calculation of exact distributions for several other metrics available in the literature, for the univariate and multivariate cases. The results are extended to elliptical families not under the assumption of normality. An application is described in order to exemplify the methodology.

Key words: Diagnostic, outlier, Cook's distance, elliptical law, linear application, modified Cook's distance, exact distributions.

AMS 2000 Subject Classification: Primary 62J15
Secondary 62E10

1 Introduction

Sensitivity analysis in linear models, under normality assumptions has been deeply studied in the statistical literature since the seminal work of Cook (1977) and many more like Belsey *et al.* (1980), Cook and Weisberg (1982), and Chatterjee and Hadi (1988), among others. The research in this area has been extended to deal with several particular

regression models, like in Galea *et al.* (1997), Díaz-García *et al.* (2003) or Díaz-García and González- Farías (2004) but one common factor for most of the techniques dealing with detection of influential observation in linear regression models, is the so called Cook's distance. The work of Muller and Mok (1997) and Jensen and Ramirez (1997) have derived the exact distribution for the Cook's distance on univariate linear models either with fixed or random coefficients. For the multivariate case, there exists a proposed expression for the exact distribution of the Cook's distance, see Díaz-García *et al.* (2003).

In many instances when we are dealing with a regression model, it is not only of interest to study the sensitivity of the parameters $\boldsymbol{\beta}$, but some particular transformations that reflect hypothesis of interest for the researchers. Most of these transformations are basically expressed as linear transformations of the parameters of the form $\mathbf{N}\boldsymbol{\beta}\mathbf{M}$, where $\mathbf{N}$ and $\mathbf{M}$ are constant matrices of the appropriate orders, see Caroni (1987) and Díaz-García and González- Farías (2004). It would be beneficial to study the sensitivity of transformation of a most general kind, for example, linear applications of the form $\mathbf{N}\boldsymbol{\beta}\mathbf{M}$. One particular case would be if we consider the effect of outliers for the kr-th element of β_{kr} from $\boldsymbol{\beta}$. However, the most common cases we found on the literature are comparisons among the rows of $\boldsymbol{\beta}$, of the form $\mathbf{N}\boldsymbol{\beta}$; that is, we are interested on comparing the effects of the proposed model with each characteristic or linear applications in the form $\boldsymbol{\beta}\mathbf{M}$; for which we are making comparison among the columns of $\boldsymbol{\beta}$, and therefore comparison of the same effect for all the characteristics under the study.

In this paper we propose some generalizations in the multivariate context for the classical and the modified Cook's distances when we eliminate one or several observations, deriving the exact distributions. In the same way, we study the effect of eliminating one or several observations on the estimation of linear functions of the parameter regression matrix. In such case, we propose an extension of the classical and modified Cook's distance (Díaz-García *et al.*, 2003), as well as, the modified distances given in Díaz-García and González- Farías (2004). For all the cases we derived the exact distribution. We also extend all those results when the normality assumption is dropped considering instead, the family of elliptical distributions. We applied these results to a diet problem discussed in Srivastava and Carter (1983), and illustrate some of the linear transformations of the form $\boldsymbol{\beta}\mathbf{M}$.

2 Multivariate Elliptical Linear Regression Model

As we can see in the statistical literature, the elliptical family of distributions has received a lot of attention during the last 20 years, given the fact that many properties that belong to the multivariate normal distribution are either invariant or can be extended, in a very natural way, to the elliptical family. Some references in this line are Fang and Anderson (1990), Fang and Zhang (1990), Gupta and Varga (1993), Díaz-García *et al.* (2002), among many others.

In this section we provide some basic notation and results in the context of linear models that allow us to derive the corresponding distances and their distributions for sections 3 and 4.

We say that an $n \times p$ random matrix $\mathbf{Y} = (\mathbf{Y}_1, ..., \mathbf{Y}_p)'$ has an elliptical distribution with parameters $\boldsymbol{\mu} \in \Re^{n\times p}$ the location matrix and $\boldsymbol{\Sigma} \otimes \boldsymbol{\Phi} \in \Re^{np\times np}$ the scala matrix, $\boldsymbol{\Sigma} > \mathbf{0}$ and $\boldsymbol{\Phi} > \mathbf{0}$, with $\boldsymbol{\Sigma} \in \Re^{p\times p}$ and $\boldsymbol{\Phi} \in \Re^{n\times n}$, if its density function is given by

$$f_{\mathbf{Y}}(\mathbf{Y}) = |\boldsymbol{\Sigma}|^{-n/2}|\boldsymbol{\Phi}|^{-p/2} g\left\{\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}(\mathbf{Y}-\boldsymbol{\mu})'\boldsymbol{\Phi}^{-1}(\mathbf{Y}-\boldsymbol{\mu})\right)\right\}, \qquad (1)$$

where the function $g : \Re \to [0, \infty)$ is such that $\int_0^\infty u^{np/2-1} g(u)du < \infty$ and $\otimes$ denotes the usual Kronecker product. The function g is called the density generator and write $\mathbf{Y} \sim \mathcal{E}l_{n\times p}(\boldsymbol{\mu}, \boldsymbol{\Sigma}\otimes\boldsymbol{\Phi}; g)$. When it exists, we have that $E(\mathbf{Y}) = \boldsymbol{\mu}$ and $\mathrm{Var}(\mathbf{Y}) = c_g\boldsymbol{\Sigma}\otimes\boldsymbol{\Phi}$, where c_g is a constant positive. In the case where $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} \otimes \boldsymbol{\Phi} = \mathbf{I}_{np}$, we obtain the spherical family of densities. These class of distributions include Normal, t-Student, Contaminated Normal, Bessel and Kotz, among other distributions.

Consider the multivariate linear regression model,

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2}$$

where $\mathbf{Y} \in \Re^{n\times p}$ is the observed matrix, $\mathbf{X} \in \Re^{n\times q}$ the regression matrix of rank q, $\boldsymbol{\beta} \in \Re^{q\times p}$ the parameter regression matrix and $\boldsymbol{\varepsilon} \in \Re^{n\times p}$ the error matrix with distribution given by $\mathcal{E}l_{n\times p}(\boldsymbol{\mu}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n; g)$, where $\boldsymbol{\Sigma} > \mathbf{0}$ is the scale matrix of dimension $p \times p$ and the density of $\mathbf{Y}$ is given in (1), with $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$. If g is a continuing and decreasing function the maximum likelihood estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ are given by,

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})\mathbf{X}'\mathbf{Y} = \mathbf{X}^-\mathbf{Y} \tag{3}$$

$$\widehat{\boldsymbol{\Sigma}} = u_0(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\beta}}) \tag{4}$$

where A^- is the Moore-Penrose inverse of A and u_0 maximize the function $h(u) = u^{-np/2}g(p/u)$, $u \geq 0$. More over, by Gupta and Varga (1993, Theorem 2.1.2 p. 20) we have,

$$\widehat{\boldsymbol{\beta}} \sim \mathcal{E}l_{q\times p}(\boldsymbol{\beta}, \boldsymbol{\Sigma} \otimes (\mathbf{X}'\mathbf{X})^{-1}; g). \tag{5}$$

Finally, if we defined $\mathbf{S} = \widehat{\boldsymbol{\Sigma}}/(u_0(n-q))$, we get $E(\mathbf{S}) = \boldsymbol{\Sigma}$.

3 Detecting an Influential Observation

We will establish all the results under the assumption that $\boldsymbol{\varepsilon}$ has a matrix normal distribution, noticing that, $u_0 = 1/n$. At the end of the section we will extend the result for the elliptical case.

3.1 Classical Cook's distance

Consider the general multivariate linear model which is obtained from (2) by deleting the i-th row of $\mathbf{Y}$, $\mathbf{X}$ and $\boldsymbol{\varepsilon}$. That is, by deleting i-th observation and getting the matrices $\mathbf{Y}_{(i)}$, $\mathbf{X}_{(i)}$ and $\boldsymbol{\varepsilon}_{(i)}$, respectively.

For the modified model, the maximum likelihood estimators for $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ are given by,

i) $\widehat{\boldsymbol{\beta}}_{(i)} = (\mathbf{X}'_{(i)}\mathbf{X}_{(i)})^{-1}\mathbf{X}'_{(i)}\mathbf{Y}_{(i)} = \mathbf{X}^-_{(i)}\mathbf{Y}_{(i)}$ and

ii) $\widehat{\boldsymbol{\Sigma}}_{(i)} = \dfrac{1}{n}(\mathbf{Y}_{(i)} - \mathbf{X}_{(i)}\widehat{\boldsymbol{\beta}}_{(i)})'(\mathbf{Y}_{(i)} - \mathbf{X}_{(i)}\widehat{\boldsymbol{\beta}}_{(i)})$

Díaz-García *et al.* (2003) proposed a multivariate version for the Cook distance to study the sensitivity of the regression parameters given by,

$$\mathcal{D}_i = \frac{1}{q}\,\mathrm{vec}'\left((\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\right)\widehat{\mathrm{Cov}}\left(\mathrm{vec}(\widehat{\boldsymbol{\beta}})\right)^{-1}\mathrm{vec}\left((\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\right). \tag{6}$$

Note that $\mathrm{Cov}\left(\mathrm{vec}(\widehat{\boldsymbol{\beta}})\right) = \boldsymbol{\Sigma} \otimes (\mathbf{X}'\mathbf{X})^{-1}$ and given that

$$\mathrm{tr}(\mathbf{B}\mathbf{X}'\mathbf{C}\mathbf{X}\mathbf{D}) = \mathrm{vec}'(\mathbf{X})(\mathbf{B}'\mathbf{D}' \otimes \mathbf{C})\,\mathrm{vec}(\mathbf{X}) \tag{7}$$

for matrices of appropriate orders, (6) can be written as

$$\mathcal{D}_i = \frac{1}{q}\,\mathrm{tr}\left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right)'(\mathbf{X}'\mathbf{X})\left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}\right)\mathbf{S}^{-1} \tag{8}$$

but

$$\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)} = \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i\widehat{\boldsymbol{\varepsilon}}_i'}{1-h_{ii}}, \tag{9}$$

where $\mathbf{X}_i$ is the i-th row of the matrix $\mathbf{X}$, $h_{ii} = \mathbf{X}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i$ is the i-th diagonal element of the orthogonal projector $\mathbf{H} = \mathbf{X}\mathbf{X}^{-} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ (also called the prediction matrix), $\widehat{\boldsymbol{\varepsilon}} = (\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}) = (\mathbf{I}-\mathbf{H})\mathbf{Y}$ and $\widehat{\boldsymbol{\varepsilon}}_i' = \mathbf{e}_i^{n\prime}(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}) = \mathbf{H}_i'\mathbf{Y}$, with $\mathbf{e}_i^n$ being the i-th canonical vector from $\Re^n$, see Chatterjee and Hadi (1988) and Díaz-García and González-Farías (2004). Therefore, substituting (9) in (8) and applying some trace properties

$$\begin{aligned}\mathcal{D}_i &= \frac{1}{q(1-h_{ii})^2}\,\widehat{\boldsymbol{\varepsilon}}_i'\mathbf{S}^{-1}\widehat{\boldsymbol{\varepsilon}}_i\mathbf{X}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i &(10)\\ &= \frac{h_{ii}}{q(1-h_{ii})}R_i^2 &(11)\end{aligned}$$

where $R_i^2 = \widehat{\boldsymbol{\varepsilon}}_i'\mathbf{S}^{-1}\widehat{\boldsymbol{\varepsilon}}_i/(1-h_{ii})$ is the square norm of the multivariate internal studentized residual. Now, since it is known (see Caroni (1987)) that $R_i^2/(n-q) \sim \boldsymbol{\beta}(q/2,(n-q-p)/2)$ where $\boldsymbol{\beta}(q/2,(n-q-p)/2)$ denote a central Beta distribution with parameters $q/2$ and $(n-q-p)/2$ it follows that,

$$\frac{q(1-h_{ii})\mathcal{D}_i}{h_{ii}(n-q)} = \frac{R_i^2}{(n-q)} \sim \boldsymbol{\beta}(p/2,(n-q-p)/2) \tag{12}$$

or, $\mathcal{D}_i \sim C_i\boldsymbol{\beta}(p/2,(n-q-p)/2)$, where

$$C_i = h_{ii}(n-q)/(q(1-h_{ii})) \tag{13}$$

The distribution of $\mathcal{D}_i$, in the univariate case has been studied by Muller and Mok (1997). More over, following the suggestion from Chatterjee and Hadi (1988, p. 124) and Díaz-García and González- Farías (2004), that is, if instead of $\mathbf{S}$ we use $\mathbf{S}_{(i)}$, obtain after eliminating the i-th observation and we denote $\mathcal{D}_i$ by $\mathcal{D}_i^*$, then,

$$\begin{aligned}\mathcal{D}_i^* &= \frac{1}{q(1-h_{ii})^2}\,\widehat{\boldsymbol{\varepsilon}}_i'\mathbf{S}_{(i)}^{-1}\widehat{\boldsymbol{\varepsilon}}_i\mathbf{X}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i &(14)\\ &= \frac{h_{ii}}{q(1-h_{ii})}T_i^2, &(15)\end{aligned}$$

where $T_i^2 = \widehat{\boldsymbol{\varepsilon}}_i'\mathbf{S}_{(i)}^{-1}\widehat{\boldsymbol{\varepsilon}}_i/(1-h_{ii})$ is the squared norm of the multivariate externally studentized residual, with

$$\frac{(n-q-p)T_i^2}{p(n-q-1)} \sim \mathcal{F}_{p,(n-q-p)}.$$

where $\mathcal{F}_{p,(n-q-p)}$ denote a central $\mathcal{F}$ distribution with parameters p and $(n-q-p)$, see (see Caroni, 1987). Therefore, $\mathcal{D}_i^* \sim E_i\mathcal{F}_{p,(n-q-p)}$, where

$$E_i = \frac{h_{ii}p(n-q-1)}{q(1-h_{ii})(n-q-p)}. \tag{16}$$

The same result but only for the univariate case, can be found in Jensen and Ramirez (1997), although their proof follows a rather different approach.

We summarize the above results in the following theorem.

Theorem 1. *Consider the general multivariate model (2) and definitions $\mathcal{D}_i$ and $\mathcal{D}_i^*$ given in (11) and (15), respectively. Suppose that $\boldsymbol{\varepsilon} \sim \mathcal{N}_{n\times p}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n)$, then,*

i)

$$\mathcal{D}_i \sim C_i \boldsymbol{\beta}(p/2, (n-q-p)/2), \tag{17}$$

where $\boldsymbol{\beta}(p/2, (n-q-p)/2)$ denote a central beta distribution with parameters $p/2$ and $(n-q-p)/2$ and C_i as given in (13).

ii)

$$\mathcal{D}_i^* \sim E_i \mathcal{F}_{p,(n-q-p)}, \tag{18}$$

where E_i defined in (16) and $\mathcal{F}(p, n-q-p)$ denote a central $\mathcal{F}$ distribution with p and $(n-q-p)$ degrees of freedom.

Note that all the multivariate Cook's distances defined here can be easily extended to study the sensitivity of linear functions of the parameters $\mathbf{N}\boldsymbol{\beta}\mathbf{M}$, in the following way. Let

$$\mathcal{D}_i = \frac{1}{l} \operatorname{vec}' \left(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M} \right) \widehat{\operatorname{Cov}} \left(\operatorname{vec}(\mathbf{N}\widehat{\boldsymbol{\beta}}\mathbf{M}) \right)^{-1} \operatorname{vec} \left(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M} \right) \tag{19}$$

where $\mathbf{N} \in \Re^{l\times q}$ and $\mathbf{M} \in \Re^{p\times s}$, are non random matrices (matrices of constants) of rank l and s, respectively. Observe that

$$\widehat{\operatorname{Cov}} \left(\operatorname{vec} \left(\mathbf{N}(\widehat{\boldsymbol{\beta}})\mathbf{M} \right) \right)^{-1} = \mathbf{M}^- \mathbf{S}^{-1} \mathbf{M}'^- \otimes \mathbf{N}'^- \mathbf{X}'\mathbf{X}\mathbf{N}^-,$$

and

$$\operatorname{vec} \left(\mathbf{N} \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)} \right) \mathbf{M} \right) = (\mathbf{M}' \otimes \mathbf{N}) \operatorname{vec} \left(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)} \right),$$

and following the same type of arguments than before, (see equation (10)),

$$\begin{aligned} \mathcal{D}_i &= \frac{1}{l} \operatorname{tr} \mathbf{M}\mathbf{M}^- \mathbf{S}^{-1} \mathbf{M}\mathbf{M}^- (\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})' \mathbf{N}^- \mathbf{N}\mathbf{X}'\mathbf{X}\mathbf{N}^- \mathbf{N} (\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}) && (20) \\ &= \frac{h_{ii}^*}{l(1-h_{ii})} R_i^2 && (21) \end{aligned}$$

where $h_{ii}^* = \mathbf{X}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{N}^-\mathbf{N}(\mathbf{X}'\mathbf{X})\mathbf{N}^-\mathbf{N}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i$, and

$$R_i^2 = \begin{cases} \widehat{\boldsymbol{\varepsilon}}_i' \mathbf{S}_1^- \widehat{\boldsymbol{\varepsilon}}_i / (1-h_{ii}), & \mathbf{S}_1^- = \mathbf{M}\mathbf{M}^- \mathbf{S}^{-1} \mathbf{M}\mathbf{M}^- \in \Re^{p\times p} \text{ of rank } s \le p; \\ \widehat{\boldsymbol{\varepsilon}}_i^{*'} \mathbf{S}^{-1} \widehat{\boldsymbol{\varepsilon}}_i^* / (1-h_{ii}), & \text{with } \widehat{\boldsymbol{\varepsilon}}_i^* = \mathbf{M}\mathbf{M}^- \widehat{\boldsymbol{\varepsilon}}_i. \end{cases}$$

Theses alternative expressions for R_i^2 allows us to establish the exact distribution under either one of the following assumptions, $\mathbf{S}_1$ has a singular central Wishart distribution or $\widehat{\boldsymbol{\varepsilon}}_i^*$ follows a singular multivariate normal distribution, see Eaton (1983) or Díaz-García and Gutiérrez-Jáimez (1997).

In this context we also may substitute, $\mathbf{S}$ by $\mathbf{S}_{(i)}$, and redefine $\mathcal{D}_i^*$. So, we get the following result.

Theorem 2. *Consider the general multivariate linear model (2) and definitions $\mathcal{D}_i$ and $\mathcal{D}_i^*$. Suppose that $\boldsymbol{\varepsilon} \sim \mathcal{N}_{n\times p}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n)$, and given the parametric linear functions $\mathbf{N}\boldsymbol{\beta}\mathbf{M}$ we have,*

i)

$$\mathcal{D}_i \sim C_i \boldsymbol{\beta}(s/2, (n-l-s)/2), \tag{22}$$

where $\boldsymbol{\beta}(s/2, (n-l-s)/2)$ denote a central beta distribution with parameters $s/2$ and $(n-l-s)/2$ and $C_i = h_{ii}^*(n-l)/(l(1-h_{ii}))$.

ii)

$$\mathcal{D}_i^* \sim E_i \mathcal{F}_{s,(n-l-s)}, \tag{23}$$

where $E_i = sh_{ii}^*(n-l-1)/(l(1-h_{ii})(n-l-s))$ *as defined in (16) and* $\mathcal{F}(s, n-l-s)$ *denote a central* $\mathcal{F}$ *distributions with* p *and* $(n-l-s)$ *degrees of freedom.*

3.2 Modified Cook's distance

The modified Cook's distance for sensitivity analysis of the matrix estimator of parameters $\boldsymbol{\beta}$, has been deeply studied in Díaz-García and González- Farías (2004). The objective of this section is to extend those results for sensitivity analysis of linear functions of the form $\mathbf{N}\boldsymbol{\beta}\mathbf{M}$. So, we first propose a similar modification to that given in Díaz-García and González- Farías (2004), called $\mathcal{AC}_i$and define it as,

$$\mathcal{AC}_i = \mathrm{vec}'(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M})\widehat{\mathrm{Cov}}\left(\mathrm{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M})\right)^{-} \mathrm{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M}). \tag{24}$$

In order to obtain an explicit expression for $\mathcal{AC}_i$, just note that

$$\widehat{\mathrm{Cov}}\left(\mathrm{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M})\right) = (\mathbf{M}' \otimes \mathbf{N})\widehat{\mathrm{Cov}}\left(\mathrm{vec}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\right)(\mathbf{M} \otimes \mathbf{N}')$$

Besides, and using similar arguments as those given in Díaz-García and González-Farías (2004)

$$\begin{aligned} \mathrm{vec}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}) &= \frac{(\mathbf{I}_p \otimes (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i\mathbf{H}_i')}{1-h_{ii}}\,\mathrm{vec}(\mathbf{Y}), & (25)\\ \mathrm{Cov}\left(\mathrm{vec}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\right) &= \frac{\mathbf{S} \otimes (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i\mathbf{X}_i'(\mathbf{X}'\mathbf{X})^{-1}}{(1-h_{ii})}, & (26) \end{aligned}$$

where $\mathbf{H}_i' = \mathbf{e}_i^{n\prime}(I - \mathbf{H})$ is the i-th row of the matrix $(I - \mathbf{H})$ and $\mathbf{e}_i^n$ being the i-th vector for a canonical base in $\Re^n$. Then,

$$\mathrm{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M}) = \frac{\left(\mathbf{M}' \otimes \mathbf{N}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i\mathbf{H}_i'\right)}{1-h_{ii}}\,\mathrm{vec}(\mathbf{Y}), \tag{27}$$

and

$$\widehat{\mathrm{Cov}}\left(\mathrm{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M})\right) = \frac{\mathbf{M}'\mathbf{S}\mathbf{M} \otimes \mathbf{N}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i\mathbf{X}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{N}'}{1-h_{ii}}. \tag{28}$$

Now, let us denote $\mathbf{v}_i = \mathbf{N}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i$, $\mathbf{v}_i \in \Re^l$, then

$$\begin{aligned} \left(\widehat{\mathrm{Cov}}\left(\mathrm{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M})\right)\right)^{-} &= \left(\frac{\mathbf{M}'\mathbf{S}\mathbf{M} \otimes \mathbf{v}_i\mathbf{v}_i'}{1-h_{ii}}\right)^{-}\\ &= \frac{1-h_{ii}}{\|\mathbf{v}_i\|^4}((\mathbf{M}'\mathbf{S}\mathbf{M})^{-1} \otimes \mathbf{v}_i\mathbf{v}_i'). \end{aligned}$$

Therefore the modified Cook's distance, $\mathcal{AC}_i$, can be written as,

$$\begin{aligned} \mathcal{AC}_i &= \mathrm{vec}'(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M})\left(\widehat{\mathrm{Cov}}\left(\mathrm{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M})\right)\right)^{-}\mathrm{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\mathbf{M})\\ &= \left(\frac{(\mathbf{M}' \otimes \mathbf{v}_i\mathbf{H}_i')}{1-h_{ii}}\,\mathrm{vec}(\mathbf{Y})\right)'\frac{(1-h_{ii})((\mathbf{M}'\mathbf{S}\mathbf{M})^{-1} \otimes \mathbf{v}_i\mathbf{v}_i')}{\|\mathbf{v}_i\|^4}\\ &\qquad\qquad\left(\frac{(\mathbf{M}' \otimes \mathbf{v}_i\mathbf{H}_i')}{1-h_{ii}}\,\mathrm{vec}(\mathbf{Y})\right)\\ &= (1-h_{ii})^{-1}\,\mathrm{vec}'(\mathbf{Y})(\mathbf{M}(\mathbf{M}'\mathbf{S}\mathbf{M})^{-1}\mathbf{M}' \otimes \mathbf{H}_i\mathbf{H}_i')\,\mathrm{vec}(\mathbf{Y}). \end{aligned} \tag{29}$$

Given (7), $\mathcal{AC}_i$ may be written as

$$\mathcal{AC}_i = (1-h_{ii})^{-1}\,\mathrm{tr}(\mathbf{M}(\mathbf{M}'\mathbf{S}\mathbf{M})^{-1}\mathbf{M}'\mathbf{Y}'\mathbf{H}_i\mathbf{H}'_i\mathbf{Y}).$$

But, $\widehat{\boldsymbol{\varepsilon}}'_i = \mathbf{e}_i^{n\prime}(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}}) = \mathbf{P}'_i\mathbf{Y}$, then,

$$\begin{aligned}\mathcal{AC}_i &= (1-h_{ii})^{-1}\,\mathrm{tr}(\mathbf{M}(\mathbf{M}'\mathbf{S}\mathbf{M})^{-1}\mathbf{M}'\widehat{\boldsymbol{\varepsilon}}_i\widehat{\boldsymbol{\varepsilon}}'_i)\\ &= (1-h_{ii})^{-1}(\mathbf{M}'\widehat{\boldsymbol{\varepsilon}}_i)'(\mathbf{M}'\mathbf{S}\mathbf{M})^{-1}(\mathbf{M}'\widehat{\boldsymbol{\varepsilon}}_i).\end{aligned}$$

As for the classical Cook's distance case, it is also possible to replace the estimator $\mathbf{S}$ of $\boldsymbol{\Sigma}$ by the estimator $\mathbf{S}_{(i)}$, obtained from the reduced sample, and get a modified Cook's distance that will denote as $\mathcal{AC}_i^*$.

The exact distributions for $\mathcal{AC}_i$ and $\mathcal{AC}_i^*$, are given on the following result.

Theorem 3. *Consider the general linear multivariate] model (2) and the definitions of $\mathcal{AC}_i$ and $\mathcal{AC}_i^*$. Suppose that, $\boldsymbol{\varepsilon} \sim \mathcal{N}_{n\times p}(\mathbf{0}, \boldsymbol{\Sigma}\otimes\mathbf{I}_n)$ and given the matrices $\mathbf{N}\in\Re^{l\times q}$ and $\mathbf{M}\in\Re^{p\times s}$, of rank l and s respectively, we have*

i)

$$\frac{\mathcal{AC}_i}{n-q} \sim \boldsymbol{\beta}(s/2,(n-q-s)/2), \tag{30}$$

where $\boldsymbol{\beta}(s/2,(n-q-s)/2)$ denote a central beta distribution with parameters $s/2$ and $(n-q-s)/2$.

ii)

$$\frac{(n-q-s)\mathcal{AC}_i^*}{s(n-q-1)} \sim \mathcal{F}(s,n-q-s), \tag{31}$$

where $\mathcal{F}(s,n-q-s)$ denote a central $\mathcal{F}$ distribution with s and $(n-q-s)$ degrees of freedom.

Proof. It follows immediately from Caroni (1987) and Díaz-García and González- Farías (2004). ∎

4 Detecting a set of influential observations

Let $I=\{i_1,i_2,...,i_k\}$ a set of size k of $\{1,2,..n\}$, such that $(n-k)\geq q$. Now, with respect to the model (2) denote $\mathbf{X}_{(I)}$, $\mathbf{Y}_{(I)}$ and $\boldsymbol{\varepsilon}_{(I)}$ the regression, data and error matrices respectively, after deleting the corresponding observations in accordance with the subindexes given in I. Let $\widehat{\boldsymbol{\beta}}_{(I)}$ and $\widehat{\boldsymbol{\Sigma}}_{(I)}$ be the corresponding maximum likelihood estimators for the general multivariate linear model after eliminating the set of observation in I.

As for the other case, it can be verified, see Chatterjee and Hadi (1988) and Díaz-García and González- Farías (2004), that

$$\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_I(\mathbf{I}_k-\mathbf{H}_I)^{-1}\widehat{\boldsymbol{\varepsilon}}_I. \tag{32}$$

where $\mathbf{H}_I = \mathbf{X}'_I(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_I$, with $\mathbf{X}_I$ as the regression matrix , and $\widehat{\boldsymbol{\varepsilon}}_I = \mathbf{U}'_I(\mathbf{I}-\mathbf{H})\mathbf{Y}$, with $\mathbf{U}_I = (\mathbf{e}^n_{i_1},\mathbf{e}^n_{i_2},...,\mathbf{e}^n_{i_2})$.

4.1 Classical Cook's distance

In this case, a generalization of the classical Cook's distance for a multivariate linear model can be written as

$$\mathcal{D}_I = \frac{1}{k}\,\mathrm{vec}'\left((\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\right)\widehat{\mathrm{Cov}}\left(\mathrm{vec}(\widehat{\boldsymbol{\beta}})\right)^{-1}\mathrm{vec}\left((\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\right). \tag{33}$$

Given (32) and applying properties for the trace operator vec can be re-expressed (33) as

$$\mathcal{D}_I = \frac{1}{k}\,\mathrm{tr}\,\widehat{\boldsymbol{\varepsilon}}_I\mathbf{S}^{-1}\widehat{\boldsymbol{\varepsilon}}_I'(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{H}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1} \tag{34}$$

Again, (34), $\mathbf{S}$ may be substitute by $\mathbf{S}_{(I)}$, calculated from the reduce sample, to get the modified Cook's distance that it is denoted by D_I^*.

Assuming that $k < p$ (this will be enough since the distributions when $k \geq p$, can be derived from the case $k < p$, see Muirhead (1982, p. 455)), of Díaz-García and González-Farías (2004) it is known that,

$$\begin{aligned}\mathbf{B} &= (\mathbf{I}_k - \mathbf{H}_I)^{-1/2}\widehat{\boldsymbol{\varepsilon}}_I((n-q)\mathbf{S})^{-1}\widehat{\boldsymbol{\varepsilon}}_I'(\mathbf{I}_k - \mathbf{H}_I)^{-1/2}\\ \mathbf{F} &= (\mathbf{I}_k - \mathbf{H}_I)^{-1/2}\widehat{\boldsymbol{\varepsilon}}_I((n-q-k)\mathbf{S}_{(I)})^{-1}\widehat{\boldsymbol{\varepsilon}}_I'(\mathbf{I}_k - \mathbf{H}_I)^{-1/2}\end{aligned}$$

has a matric variate type Beta and matric variate $\mathcal{F}$ distributions respectively; also called matrix variate beta type I and matrix variate beta type II distribution, respectively, see Gupta and Nagar (2000, pp. 165-166). Let $\mathbf{A}^{1/2} = \mathbf{H}_I^{1/2}(\mathbf{I}_k - \mathbf{H}_I)^{-1/2}$ then the matrices

$$\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}'^{1/2}$$
$$\mathbf{A}^{1/2}\mathbf{F}\mathbf{A}'^{1/2}$$

have a generalized matrix variate beta type I and a generalized matrix variate beta type II distribution, respectively, see Gupta and Nagar (2000, p. 175). Then,

$$\frac{k\ \mathcal{D}_I}{(n-q)} = \mathrm{tr}\,\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}'^{1/2} \quad \text{and} \quad \frac{k\ \mathcal{D}_I^*}{(n-q-k)} = \mathrm{tr}\,\mathbf{A}^{1/2}\mathbf{F}\mathbf{A}'^{1/2}$$

We have summarized those results in Theorem 4.

Theorem 4. *Consider the general multivariate linear model (2) and definitions $\mathcal{D}_I$ and $\mathcal{D}_I^*$. Suppose that $\boldsymbol{\varepsilon} \sim \mathcal{N}_{n\times p}(\mathbf{0}, \boldsymbol{\Sigma}\otimes\mathbf{I}_n)$, then,*

$$\frac{k\ \mathcal{D}_I}{(n-q)} \quad \textit{and} \quad \frac{k\ \mathcal{D}_I^*}{(n-q-k)},$$

have the distribution of the trace of a generalized matrix variate beta type I and a generalized matrix variate beta type II distribution, respectively. That is, they have the distributions of the Pillai and Lawley-Hotelling statistics base on the generalized versions of the betas type I and II, respectively.

Unfortunately, distributions given in Theorem 4 have not yet been tabulated.

One way to circumvent this fact and give a solution to our problem, is proposing an alternative metric for (34), similar to those given in Díaz-García and González- Farías (2004, see Table 1). In such a way that the distributions of $\mathcal{D}_I$ and $\mathcal{D}_I^*$ can be expressed as functions of the distribution for the matrices $\mathbf{B}$ and $\mathbf{F}$.

Let us delete the denominator k and consider the metric base on determinants instead of traces in (34), we get the following variant of the Cook distance

$$\begin{aligned} \mathcal{DM}_I &= \frac{|\mathbf{H}_I|}{|\mathbf{I}_k - \mathbf{H}_I|} \left| \frac{1}{(n-q)} \mathbf{S}^{-1} \widehat{\boldsymbol{\varepsilon}}_I' (\mathbf{I}_k - \mathbf{H}_I)^{-1} \widehat{\boldsymbol{\varepsilon}}_I \right|^{-1} \\ &= \frac{|\mathbf{H}_I|}{|\mathbf{I}_k - \mathbf{H}_I|} \; S - \text{criterion.} \end{aligned} \tag{35}$$

Where the S-criterion is due to Ch. L. Olson, see Kres (1983, p. 8).

Similarly, deleting the denominator k, taking as a metric the inverse of the determinant in (34) considering $\mathbf{S}_{(I)}$, we get this other variant for the Cook distance

$$\begin{aligned} \mathcal{DM}_I^* &= \frac{|\mathbf{I}_k - \mathbf{H}_I|}{|\mathbf{H}_I|} \left| \frac{1}{(n-q-k)} \mathbf{S}_{(I)}^{-1} \widehat{\boldsymbol{\varepsilon}}_I' (\mathbf{I}_k - \mathbf{H}_I)^{-1} \widehat{\boldsymbol{\varepsilon}}_I \right| \\ &= \frac{|\mathbf{I}_k - \mathbf{H}_I|}{|\mathbf{H}_I|} \; U - \text{criterion.} \end{aligned} \tag{36}$$

The U-criterion has been credited to Gnanadesikan in Kres (1983, p. 8) and Olson (1974), perhaps because it was thought to have been introduced in Roy *et al.* (1971, p. 72). However this statistic is none other than the U-statistic of Wilks given in Wilks (1932) and in Hsu (1940). Hsu even gives the distribution of the latter for $p = 2$. Unfortunately, in Kres (1983, p. 6), the expression for the U-statistic of Wilks as a function of the eigenvalues was given incorrectly, and perhaps this explains why it was not clear that this statistic and the one of Gnanadesikan, (presented also in Kres (1983, p. 8)), are in fact the same. Here we provide the correct representation for the U-statistic of Wilks and show that it is equivalent to the U-criterion of Gnanadesikan.

Consider the expression for the Λ of Wilks statistic and the U-criterion as functions of the eigenvalues like given in Seber (1984, pp. 412 and 413, respectively)

$$\Lambda = \prod_{i=1}^{w} (1 - \theta_i) \quad \text{and} \quad U = \prod_{i=1}^{w} \theta_i,$$

where $\mathbf{H}$ and $\mathbf{E}$ are the sum of squared and product matrices for the hypothesis and the error respectively, w being the rank and $\boldsymbol{\Theta} = (\theta_1, \ldots, \theta_w)'$ the no-zero eigenvalues of the matrix $\mathbf{H}(\mathbf{H} + \mathbf{E})^{-1}$, such that $1 > \theta_1 \geq \cdots \geq \theta_w > 0$. Denote the density of $\boldsymbol{\Theta}$ by $p(\boldsymbol{\Theta}, w, m_1, m_2)$, where m_1 and m_2 are functions of n, p, q, defined in Section 2, see Muirhead (1982, pp. 451 and 454-455) or Díaz-García and Gutiérrez-Jáimez (1997).

Now, it is known that $\Lambda \sim$ Wilks's Λ. If $\boldsymbol{\Theta}^* = ((1 - \theta_1), \ldots, (1 - \theta_w))' = (\theta_1^*, \ldots, \theta_w^*)'$, the distribution of $\boldsymbol{\Theta}^*$is the same as that of $\boldsymbol{\Theta}$, interchanging m_1 and m_2. Then,

$$\Lambda^* = \prod_{i=1}^{w} \theta_i^* \sim \text{Wilks's } \Lambda, \quad \text{with } m_1 \text{ and } m_2 \text{ interchanged}$$

but note that $\Lambda^* = U$. Therefore,

$$U \sim \text{Wilks's } \Lambda, \quad \text{with } m_1 \text{ and } m_2 \text{ interchanged}$$

In summary, the critical value for the U-criterion, can be obtained from the tables for the statistics Λ of Wilks interchanging m_1 and m_2.

Theorem 6 summarizes the above results.

Theorem 5. *Consider the general multivariate linear model (2) and definitions $\mathcal{DM}_I$ and $\mathcal{DM}_I^*$. Suppose that $\boldsymbol{\varepsilon} \sim \mathcal{N}_{n\times p}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n)$,*

$$\frac{|\mathbf{I}_k - \mathbf{H}_I|}{|\mathbf{H}_I|}\ \mathcal{DM}_I \quad \textit{and} \quad \frac{|\mathbf{H}_I|}{|\mathbf{I}_k - \mathbf{H}_I|}\ \mathcal{DM}_I^*,$$

have the S-criterion distribution and U-criterion distribution respectively.

For the linear application case, $\mathbf{N}\boldsymbol{\beta}\mathbf{M}$ the distances $\mathcal{DM}_I$ and $\mathcal{DM}_I^*$ are obtained simply making the following changes

$$\begin{aligned} \mathbf{H}_I &\rightarrow \mathbf{H}_I^* = \mathbf{X}_I'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{N}^-\mathbf{N}(\mathbf{X}'\mathbf{X})\mathbf{N}^-\mathbf{N}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_I \\ \mathbf{S}^{-1} &\rightarrow \mathbf{S}_1^- = \mathbf{M}\mathbf{M}^-\mathbf{S}^{-1}\mathbf{M}\mathbf{M}^- \end{aligned}$$

and their corresponding distributions, change the parameters

$$q \rightarrow l \quad \text{and} \quad p \rightarrow s$$

to obtain the correct percentile.

4.2 The modified Cook's distance

The modified Cook distance for several influential observations for the estimated regression parameter matrix $\boldsymbol{\beta}$ was also studied in Díaz-García and González- Farías (2004). Here again, we present the direct extension for linear transformations of the form $\mathbf{N}\boldsymbol{\beta}\mathbf{M}$. Denote the modified Cook's distance as $\mathcal{AC}_I$ and define it in the following way,

$$\mathcal{AC}_I = \text{vec}'(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M})\widehat{\text{Cov}}\left(\text{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M})\right)^- \text{vec}(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M}), \tag{37}$$

where$\mathbf{N} \in \Re^{l\times q}$ and $\mathbf{M} \in \Re^{p\times s}$, with rank l and s, respectively. Working in the same way as before and following Díaz-García and González- Farías (2004) we get that $\mathcal{AC}_I$ can be written as

$$\mathcal{AC}_I = \text{vec}'(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\left(\mathbf{M}((\mathbf{M}'\mathbf{S}\mathbf{M}))^{-1}\mathbf{M}' \otimes \mathbf{N}'(\mathbf{N}\mathbf{R}\mathbf{N}')^-\mathbf{N}\right)\text{vec}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)}),$$

where $\mathbf{R} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_I(I_k - \mathbf{H}_I)^{-1}\mathbf{X}_I'(\mathbf{X}'\mathbf{X})^{-1}$.

Due to (7), $\mathcal{AC}_i$ we get

$$\mathcal{AC}_I = \text{tr}(\mathbf{M}'\mathbf{S}\mathbf{M})^{-1}\left(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M}\right)'(\mathbf{N}\mathbf{R}\mathbf{N}')^-\left(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M}\right). \tag{38}$$

But $\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M} = \mathbf{N}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\widehat{\boldsymbol{\varepsilon}}_I\mathbf{M}$, therefore

$$\begin{aligned} &\left(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M}\right)'(\mathbf{N}\mathbf{R}\mathbf{N}')^-\left(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M}\right) \\ &\qquad = \mathbf{M}'\widehat{\boldsymbol{\varepsilon}}_I'(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{X}_I'\mathbf{X}_I'^{-}(\mathbf{I}_k - \mathbf{H}_I)\mathbf{X}_I^-\mathbf{X}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\widehat{\boldsymbol{\varepsilon}}_I\mathbf{M}. \end{aligned}$$

Note that if $\mathbf{N} \in \Re^{l\times q}$, of rank l and

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{N}'\mathbf{N}'^{-}(\mathbf{X}'\mathbf{X}) = \left(\mathbf{N}'^{-}(\mathbf{X}'\mathbf{X})\right)^-\left(\mathbf{N}'^{-}(\mathbf{X}'\mathbf{X})\right) = I_l$$

then,

$$\left(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M}\right)'(\mathbf{N}\mathbf{R}\mathbf{N}')^-\left(\mathbf{N}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(I)})\mathbf{M}\right) = \mathbf{M}'\widehat{\boldsymbol{\varepsilon}}_I'(\mathbf{I}_k - \mathbf{H}_I)^{-1}\widehat{\boldsymbol{\varepsilon}}_I\mathbf{M}.$$

and $\mathcal{AC}_I$ can be finally written as

$$\mathcal{AC}_I = \text{tr}((\mathbf{M}'\mathbf{S}\mathbf{M})^{-1}\mathbf{M}'\widehat{\boldsymbol{\varepsilon}}_I'(\mathbf{I}_k - \mathbf{H}_I)^{-1}\widehat{\boldsymbol{\varepsilon}}_I\mathbf{M}. \tag{39}$$

Of course we can replace $\mathbf{S}$, the estimator of $\boldsymbol{\Sigma}$, by $\mathbf{S}_{(I)}$, obtain from the reduced sample after deleting the k observations. In that case we denote the modified Cook's distance as $\mathcal{AC}_I^*$.

For the case of multiple observations, the exact distributions of $\mathcal{AC}_I$ and $\mathcal{AC}_I^*$, are given in the following result.

Theorem 6. *Consider the general multivariate linear model (2) and the definitions $\mathcal{AC}_I$ and $\mathcal{AC}_I^*$. Suppose that $\boldsymbol{\varepsilon} \sim \mathcal{N}_{n\times p}(\mathbf{0}, \boldsymbol{\Sigma}\otimes\mathbf{I}_n)$ and considering the matrices $\mathbf{N} \in \Re^{l\times q}$ and $\mathbf{M} \in \Re^{p\times s}$, of rank l and s, respectively, we have*

i)

$$\frac{\mathcal{AC}_I}{n-q} \sim \mathcal{P}(w, m, h), \tag{40}$$

where $\mathcal{P}(w, m, h)$ denote the central distribution for the Pillai's statistics, with parameters w, m, and h, see Seber (1984) or Rencher (1995).

ii)

$$\frac{\mathcal{AC}_I^*}{n-q-k} \sim \mathcal{LH}(w, m, h), \tag{41}$$

where$\mathcal{LH}(w, m, h)$ denote the central distribution of the Lawley-Hotelling's statistic with parameters w, m, and h, see see Seber (1984) or Rencher (1995).

In both cases the parameters are defined as $w = \min(s, k)$, $m = (|s-k|-1)/2$ and $h = (n-q-s-1)/2$.

Proof. It follows directly from Díaz-García and González- Farías (2004). ∎

A very important observation here is that if we look at several of the metrics that have been proposed in the study of sensitivity of the parameter estimates of the linear model $\boldsymbol{\beta}$ or $\boldsymbol{\Sigma}$, such as: the Cook's distance, Welsh's distance, methods based on volumes, and methods based on the likelihood function, among many others, (see Chatterjee and Hadi (1988)), those can be written as a function of the internally studentized residual or externally studentized residual, either in the univariate or the multivariate case. Moreover, we can write all those statistics as

$$\text{Metric} = GR$$

where R denote the squared norm of some of those type of the residuals and G is a constant. Then Metric/G has the same distribution as R. Therefore all the distributions of the metrics, in general, coincide with some of the distributions for the metrics given in Díaz-García and González- Farías (2004). Which means that, we are able to establish the exact distribution, for the univariate as well as for the multivariate case, of all those metrics.

If we now assume that $\boldsymbol{\varepsilon} \sim \mathcal{E}l_{n\times p}(\mathbf{0}, \boldsymbol{\Sigma}\otimes\mathbf{I}_n; g)$ as an immediate consequence of Theorem 5.3.1 in Gupta and Varga (1993, p. 182) the distributions associated with each of the distances defined in Theorem 1 through 8, are invariant under the family of elliptical distributions. Therefore, the extension of the sensitivity analysis of linear applications of the form $\mathbf{N}\boldsymbol{\beta}\mathbf{M}$ and $\boldsymbol{\beta}$ follow the general multivariate linear model with elliptical errors.

Now, if we look at the exact distribution and the proposed metrics for the modified Cook's distance, we note that neither one depend on the matrix $\mathbf{N}$, not even any of

its properties like its rank or its dimensions. The researchers should be aware that the distances proposed to study the effect of the estimation of those linear applications ($\mathbf{N}\boldsymbol{\beta}\mathbf{M}$), are not taking into account the effect of $\mathbf{N}$. Therefore, those distances are in principle measuring the sensitivity for the estimation of linear applications of the form $\boldsymbol{\beta}\mathbf{M}$.

The same comments extend to the classical Cook distance since even when the metrics depend on $\mathbf{N}$ through de h_{ii}^* or $\mathbf{H}_I^*$, for the case of one or several influential observations, when we calculate their exact distributions, such effect vanishes so at the end it is only possible to detect effects for the sensitivity of liner applications of the form $\boldsymbol{\beta}\mathbf{M}$. We illustrate this type of linear transformation in the following section.

5 Application

Consider the quantity of the food ingested by 32 rats, previously assigned into groups of 4, for about 12 days. The diet consists of a treatment containing different levels of phosphate, see Srivastava and Carter (1983) or Srivastava (2002).

The model to consider is

$$y_i = \beta_{0i} + \beta_{1i}x + \beta_{2i}x^2 + \beta_{3i}x^3 + \beta_{4i}w + \varepsilon_i \quad i = 1, ..., 12$$

where x represents the quantity of phosphate given and w is the covariable initial weight of the rats.

In matrix notation,

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where $\mathbf{Y} \in \Re^{32\times 12}$, is the response matrix, with $n = 32$ and y_{ij} = the quantity of food consumed by the rat i on the day j; $\mathbf{X} \in \Re^{n\times 5}$ is the matrix of covariables, the first column is formed by corresponding intercepts, columns 2, 3 and 4 correspond to x, x^2 and x^3, respectively and the last column corresponds to w; and finally, $\boldsymbol{\beta} \in \Re^{5\times 12}$,

$$\beta = \begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{012} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{112} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{41} & \beta_{42} & \cdots & \beta_{412} \end{pmatrix},$$

is the matrix of parameters with vectors corresponding to the intercepts, x, x^2, x^3 and w respectively.

Suppose you are interested in the following linear applications,

$$\begin{pmatrix} \beta_{01} - (\beta_{03} + \beta_{04}) & \beta_{02} - \beta_{04} & \beta_{03} - \beta_{05} \\ \beta_{11} - (\beta_{13} + \beta_{14}) & \beta_{12} - \beta_{14} & \beta_{13} - \beta_{15} \\ \beta_{21} - (\beta_{23} + \beta_{24}) & \beta_{22} - \beta_{24} & \beta_{23} - \beta_{25} \\ \beta_{31} - (\beta_{33} + \beta_{34}) & \beta_{32} - \beta_{34} & \beta_{33} - \beta_{35} \\ \beta_{41} - (\beta_{43} + \beta_{44}) & \beta_{42} - \beta_{44} & \beta_{43} - \beta_{45} \end{pmatrix}, \tag{42}$$

Where, for example, for column 1, we have the following interpretation: in the first function one is interested in contrasting the quantity of food consumed in day 1 and the total food consumed in days 3 and 4; the second function establishes the effect in the quantity of phosphate in day 1, being the same as the sum for the linear effect on days 3 and 4; the third function establishes that the quadratic effect of the amount of phosphate in day 1 is the same as the quadratic effect of days 3 and 4, the fourth function

establishes the the cubic effect of the amount of phosphate on day one is the same as the sum of the cubic effects on days 3 and 4; and finally the fifth linear function establishes the effect of the initial weight on day one as being the same as the sum of the effects of the initial weight of days 3 and 4. Columns 2 and 3 in (42) are interpreted in a similar fashion.

Note that (42)can be expressed βM, with

$$\mathbf{M}' = \begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then, it would be of interest to see if any individual or group of observations is influential on the estimator of the linear functions given in (42).

The figure 1 show the distances $\mathcal{AC}_i$ and $\mathcal{AC}_i^*$, which are useful in detecting the influence of the i-th observation in the linear combination $\widehat{\boldsymbol{\beta}}\mathbf{M}$. The results of the tests allow us to identify the observations 16, 17, 20 and 26 with as having a strong individual influence over $\widehat{\boldsymbol{\beta}}\mathbf{M}$.

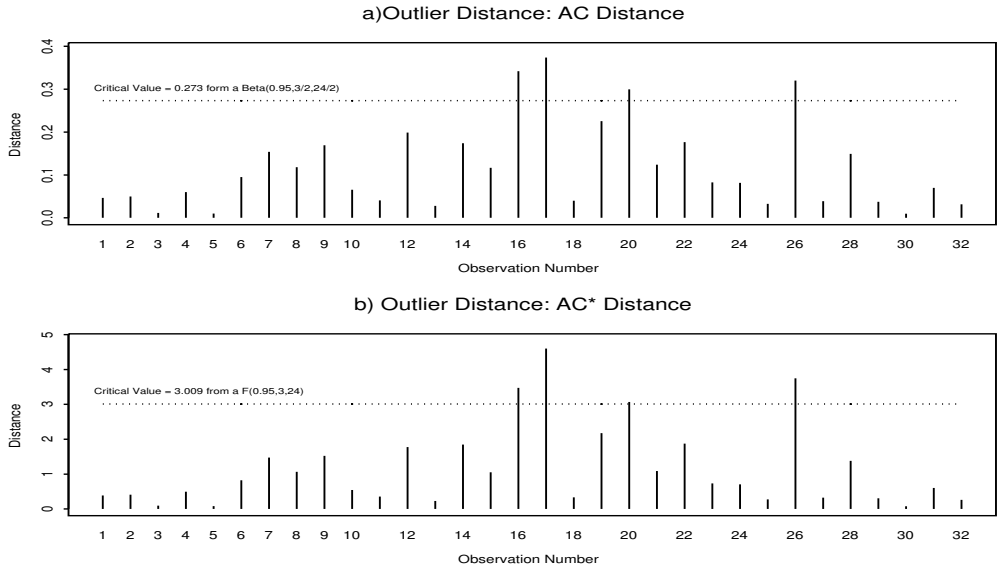

Figure 1: Identification of influential observations, based on a) the distance $\mathcal{AC}_i$ and b) the distance $\mathcal{AC}_i^*$.

Now, it is necessary to evaluate if the observations 16, 17, 20 and 26, as a group, have influence in the parametric functions $\widehat{\boldsymbol{\beta}}\mathbf{M}$. Then, using the metrics proposed in the theorem 6 we obtained the results described in Table 1.

In the four cases the test statistics are greater than the corresponding critical value α, consequently, both tests can identify the observations 16, 17, 20 and 26 as jointly influential in the linear combination $\widehat{\beta}M$.

Acknowledgements

This work was supported partially by the research project 39017E of CONACYT-México.

Table 1: Four metrics to detect an influential set of observations in linear combinations $\widehat{\beta}M$.

Metric	Statistic[a]	α Critical value
$\mathcal{AC}_I = 1.2532$	4.8430 [b]	1.8737
$\mathcal{AC}_I^* = 2.5209$	4.9718 [c]	1.8912
$\frac{\lvert\mathbf{I}_k-\mathbf{H}_I\rvert}{\lvert\mathbf{H}_I\rvert}\ \mathcal{DM}_I = 33.8781$	33.8781	not available
$\frac{\lvert\mathbf{H}_I\rvert}{\lvert\mathbf{I}_k-\mathbf{H}_I\rvert}\ \mathcal{DM}_I^* = 0.0477$	0.0477	0.000587[d]

[a] Observe that for four tests, the decision rule is: statistics $\geq$ critical value
[b] Using an approximate F-statistics, see equation (6.20) in Rencher (1995, p.185, 1995)
[c] Using an F approximation, see equation (6.24) in Rencher (1995, p.185, 1995)
[d] In our example, $p \to s$ then $m_1 = (|s-k|-1)/2$ and $m_2 = (n-q-s-1)/2$.

References

Besley, D., Kuh, E. and Welsch, R. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, New York.

Caroni, C. (1987). Residuals and influence in the multivariate linear model, *The Statistician* **36**: 365-370.

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

Cook, R. D. and Weisberg, S. (1982). *Residual and Influence in Regression*. Chapman and Hall, London.

Chatterjee, S. and Hadi, A. S. (1988). *Sensitivity Analysis in Linear Regression*. John Wiley & Sons. New York.

Díaz-García, J.A. and González- Farías, G. (2004). A note on the Cook's distance, *J. Statist. Plan. Inference* **120**: 119-136.

Díaz-García, J.A., Galea, M. and Leiva-Sánchez, V. (2001). Influence diagnostics for elliptical regression linear models, *Comm. Ststist. T. M.* **32**(2): 625-641.

Díaz-García, J.A., Leiva-Sánchez, V. and Galea, M. (2002). Singular elliptic distribution: density and applications, *Comm. Ststist. T. M.* **31**(5): 661-682.

Díaz-García J. A. and Gutiérrez-Jáimez, R. (1997). Proof of conjectures of H. Uhlig on the singular multivariate Beta and the Jacobian of a certain matrix transformation, *Ann. of Statist.* **25**: 2018-2023.

Draper, N. and Smith, H. (1981). *Applied Regression Analysis*. (2nd ed.), John Wiley & Sons, New York.

Eaton, M. L. (1983). *Multivariate Statistics. A Vector Space Approach*. John Wiley & Sons.

Fang, K. T. and Anderson T. W. (1990). *Statistical Inference in Elliptically Contoured and Related Distributions*. Allerton Press Inc., New York.

Fang, K. T. and Zhang, Y. T. (1990). *Generalized Multivariate Analysis*. Science Press, Beijing, Springer-Verlang.

Galea, M., Paula, G. and Bolfarine, H. (1997). Local influence in elliptical linear regression models, *The Statistician* **46**: 71-79.

Gupta, A. K. and Nagar, T. (2000). *Matrix variate distributions*. Chapman & Hall/CR, Boca Raton.

Gupta, A. K. and Varga, T. (1993). *Elliptically Contoured Models in Statistics*. Kluwer Academic Publishers, Dordrecht.

Jensen, D. R. and Ramirez, D. E. (1997). Some exact properties of Cook's D_I; In *Handbook of Statistics-16: Order Statistics and Their Applications*, Rao, C. R. and Balakrishnan, N., eds., North-Holland, Amsterdam.

Hsu, P. L. (1940). On generalization analysis of variance, *Biometrika* **31**: 221-237.

Kres, H. (1983). *Statistical Tables for Multivariate Analysis*, Springer- Verlag, New York.

Olson, Ch. L. (1974). Comparative robustness of six test in multivariate analysis of variance, *JASA* **69**: 894-908.

Muirhead, R. J. (1982). *Aspects of Multivariated Statistical Theory*, John Wiley & Sons, New York.

Muller, K. E. and Mok, M. Ch. (1997). The distribution of Cook's distance, *Commun. Satatist. - Theory Meth.* **26**(3): 525-546.

Rencher, A. C. (1995). *Methods of Multivariate Analysis*. John Wiley & Sons, New York.

Roy, J.; Gnanadesikan, R., and Srivastava, J. N. (1971). *Analysis and Design of Certain Quantitative Multiresponse Experiments*,Pergamon Pres Ltd., Oxford.

Seber, G. A. F. (1984). *Multivariate Observations*. John Wiley & Sons, New York.

Srivastava, M. S. (2002). *Methods of Multivariate Statistics*. John Wiley & Sons, New York.

Srivastava, M. S. and Carter, E. M. (1983). *An Introduction to Applied Multivariate Statistics*. North-Holland Publ., New York.

Wilks, S. S. (1932). Certain generalizations in the analysis of variance, *Biometrika* **24**: 471-494.