Elliptical multivariate semiparametric linear regression model with equicorrelated random errors: some theoretical aspects of the estimation process and diagnostic analysis

THESIS

To apply for the degree of

Master in Statistics

Institute of Statistics

University of Valparaíso

Chile

Presented by:

Margarita del Carmen Rivas Gómez

Thesis supervisor:

Dr. Germán Ibacache Pulgar

Valparaíso, 2019

Contents

List of Tables

Agradecimientos

Deseo expresar mi más sincero agradecimiento a todos quienes de alguna u otra forma me ayudaron en la realización de la presente tesis, en especial a mi profesor guía el Dr. Germán Ibacache Pulgar, por su orientación, motivación y por todo su apoyo incondicional que recibí durante todo mi proceso para lograr la obtención de mi grado académico.

No puedo dejar de agradecer a mis jefaturas y compañeras(os) de trabajo de la Vicerrectoría de Investigación, Postgrado e Innovación de la Universidad de Playa Ancha, por todos los consejos, ánimo, apoyo y comprensión que siempre me brindaron en todo mi proceso académico.

Y por supuesto un agradecimiento muy especial a toda mi familia que de alguna u otra forma me ayudaron para culminar mis estudios de postgrado, sin duda fue un camino difícil, pero que supimos llevar adelante.

A todos los que estuvieron y a los que dios ha llamado a su lado... Muchas gracias!!

Acknowledgements

I would like to express my sincere thanks to all those who in some way or another help me in the realization of this thesis, especially to my guide professor Dr. Germán Ibacache Thumb, for his guidance, motivation and for all his unconditional support.

I received throughout my process to obtain my academic degree. I cannot fail to thank my headquarters and co-workers of the Office of the Vice-Rector for Research, Postgraduate and Innovation of the University of Playa Ancha, for all the advice, encouragement, support and understanding that will always give me throughout my academic process.

And of course a very special thanks to all my family that in some way or another helps me to complete my postgraduate studies, it was certainly a difficult path, but we knew how to move forward.

To all who were and to those whom God has called by his side. . . . Thank you!!

Abstract

In this thesis, we will study elliptical multivariate semiparametric linear regression model with equicorrelated random errors, specifically, some theoretical aspects of the estimation process and diagnostic analysis. An estimation method based in penalized likelihood function is developed. In addition, the local influence curvature was developed to assess the sensitivity of the estimators to the observations that are influential.

Chapter 1

Introduction

The analysis of the linear regression is a statistical method, and is the most used for the linear relationships between a variable of interest or a response and a set of predictor variables. These analyzes can be seen more clearly through scatter charts, they can be seen in the set of data are in a linear form.

Statistical models of linear regression can be applied to different fields of research such as economics, administration, among many others. There are also situations in the areas of science and engineering where it is not correct to assume that there is a linear relationship. This would lead us to study non-linear regression models.

In this thesis work, only the mathematical analysis and the theoretical representation of the parametric description will be delivered in a class of known models "Elliptic multivariate semiparametric linear regression model with equicorrelated random errors: some theoretical aspects of the estimation process and diagnostic analysis", we will also introduce the function Score and the Hessian Matrix.

It will be possible to incorporate into the multivariate linear regression model with randomized equicorrelated errors, a nonparametric component?

Semiparametric regression models are a combination of models with parametric components and models with non-parametric components. An example of semi-parametric models can be analyzed in the studies conducted by Durban *et al* (1999) Approximate standard errors in semiparametric models and in Durban *et al* (2003) The practical use of semiparametric models in fiels trials, which shows examples of the structure of a semi-parametric model.

With the proposed model, the effect of small disturbances on the data and/or on the assumptions of the model, on the maximum likelihood estimators without eliminating observations, that is, evaluating the sensitivity of the estimators when there are small disturbances, will be evaluated. This method is called Local Influence and which is a method proposed by Cook (1986) and some examples of these analyzes and which were a basis for this thesis, were the studies conducted by Ibacache-Pulgar, G. and Paula, G. (2011) where he developed the study of Local Influence for Student-t partially linear models and in a later study in Ibacache *et al*. (2012) a Influence diagnostics for elliptical semiparametric mixed models.

It will be possible to develop the local influence method to assess the sensitivity of the estimators against data disturbances in Elliptical multivariate semiparametric linear regression model with equicorrelated random errors?

The objective of this work is to derive some theoretical aspects of the diagnostic estimation and analysis process in Elliptical multivariate semiparametric linear regression model with equicorrelated random errors. Also apply the Local influence method to the model under study. Based on the foregoing, the following specific objectives have been defined:
Therefore, the following specific objectives have been defined:

(a) Verify the proposed estimators for the model, calculate the Score function for the model parameters.

(b) Obtain the Hessian matrix (Matrix of second derivatives).

(c) Develop the theory of the method of local influence.

This thesis work will be organized in chapters as detailed below:

- Chapter 2, characteristics and analysis of The elliptical multivariate semi-parametric linear regression model with equicorrelated random errors: some theoretical aspects of the estimation process in addition to the presentation of diagnostic analysis and the penalty function.

- Chapter 3, the description of the fundamental elements will be carried out to adequately obtain the estimation parameters, for this, we derire the Score Function and Hessian Matrix associated with the parameter vector $\boldsymbol{\theta}$. Also, a procedure is proposed to estimate the parameters based on the Back-fitting algorithm.

- Chapter 4, the theoretical aspects of the Local Influence method were carried out for the proposed model. Through the studies carried out, we derive the normal curvature considering different disturbance schemes.

- Chapter 5, presents some conclusions obtained from the theoretical aspects and some possible future studies.

Chapter 2

Elliptical multivariate semiparametric linear regression model with equicorrelated random errors

In this chapter we will study the multivariate semiparametric linear regression model with equicorrelated random errors that follow a multivariate elliptical distribution.

2.1 The Model

Multivariate semiparametric linear regression models have become an important tool in economic and biometric applications. Here, we consider the following relationship between the vector of observed responses and the explanatory variables:

$$\mathbf{y}_i = \boldsymbol{\beta}_1 x_{i1} + ... + \boldsymbol{\beta}_k x_{ik} + \mathbf{N}_i \mathbf{f} + \boldsymbol{\epsilon}_i, \tag{2.1}$$

where $\mathbf{y}_i$ is a $(p \times 1)$ random vector of observed responses from the ith cluster $(i = 1, ...n)$, $\boldsymbol{\beta}_j$ is a $(p \times 1)$ fixed parameter vector $(j = 1, ..., k)$, x_{ij} is the jth explanatory variable value associated with the ith cluster, $\mathbf{N}_i$ is a $(p \times n)$ incidence matrix with a vector $\mathbf{1}$'s in the ith column , $\mathbf{f} = (f(t_1), ..., f(t_n))$ is a $(n \times 1)$ vector associated with nonparametric component and $\boldsymbol{\epsilon}_i$ is a $(p \times 1)$ vector of within-cluster errors such that $\boldsymbol{\epsilon}_i^{'}s$ are independent random vectors following an elliptical distribution with position vector $\mathbf{0}$ and scale matrix (or intraclass correlation matrix) given by

$$\boldsymbol{\Sigma}(\sigma^2, \rho) = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix} = \sigma^2[(1-\rho)\mathbf{I} + \rho\mathbf{J}],$$

where $\sigma^2 > 0$ and $-(p-1)^{-1} < \rho < 1$, $\mathbf{I}$ denotes a $(p \times p)$ identity matrix and $\mathbf{J}$ denotes a $(p \times p)$ matrix of ones. It is denoted $\boldsymbol{\epsilon}_i \sim \mathrm{El}_p(\mathbf{0}, \boldsymbol{\Sigma})$.

Note that the model (2.1) is an extension of other models proposed in the literature. For example,

(i) If the non-parametric component is not present in the model, then the model (2.1) is expressed as follows

$$\mathbf{y}_i = \boldsymbol{\beta}_1 x_{i1} + ... + \boldsymbol{\beta}_k x_{ik} + \boldsymbol{\epsilon}_i.$$

This model was proposed by Ibacache-Pulgar *et al.* (2014) to analyze data from the biomedical and financial area.

(ii) If $x_{i1} = 1$ and $x_{i3} = x_{i4} = ... = x_{in} = 0$, then model (2.1) is reduced to (model with intercept)

$$\mathbf{y}_i = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 x_{i2} + \mathbf{N}_i\mathbf{f} + \boldsymbol{\epsilon}_i$$

This model is called the multivariate partial linear regression model; see, for example, Ibacache-Pugar and Paula (2012).

(iii) If $x_{i1} = x_{i3} = ... = x_{ik} = 0$, then model (2.1) is reduced to (model without interception)

$$\mathbf{y}_i = \boldsymbol{\beta}_2 x_{i2} + \mathbf{N}_i\mathbf{f} + \boldsymbol{\epsilon}_i.$$

It is important designate that the elements of vector $\mathbf{f}$ correspond to the functional evaluation of function f in each value of nodes $t_1 < t_2 < ... < t_n$.

Considering the fact that $\boldsymbol{\epsilon}_i \sim \text{El}_p(\mathbf{0}, \boldsymbol{\Sigma})$, with $i = 1, ..., n$, we have that the observed response vector $\mathbf{y}_i$ follows an elliptical distribution with position vector $\boldsymbol{\mu}_i$ and scale matrix $\boldsymbol{\Sigma}$, named $\mathbf{y}_i \sim \text{El}_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$, whose density function takes the form

$$f_{\mathbf{y}}(\mathbf{y}_i) = |\boldsymbol{\Sigma}|^{-1/2} g(\delta_i), \tag{2.2}$$

where $\delta_i = \mathbf{r}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{r}_i = \text{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{r}_i\mathbf{r}_i^{\text{T}})$ is the square of the Mahalanobis distance, $\mathbf{r}_i = \mathbf{y}_i - \boldsymbol{\mu}_i$ is the shift vector, with $\boldsymbol{\mu}_i = \boldsymbol{\beta}_1 x_{i1} + ... + \boldsymbol{\beta}_k x_{ik} + \mathbf{N}_i\mathbf{f}$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\sigma^2, \rho)$ and $g(\cdot)$ is a function of $R \longrightarrow [0, \infty]$ know as the density generator function. Some examples of the function $g(\cdot)$, for different distributions that belong to the elliptical class, are shown in Table 2.1

Remark: When exists, the mean $\text{E}(\mathbf{y}_i) = \boldsymbol{\mu}_i$ y covariance-variance matrix $\text{Cov}(\mathbf{y}_i) = \zeta_i \boldsymbol{\Sigma}$. In particular, for the Student-t, we have $\zeta_i = \frac{\nu_i}{\nu_i - 2}(\nu_i > 2)$, where ν_i denotes the degrees of freedom.

Remark: The media vector $\boldsymbol{\mu}_i$ can be written as $\boldsymbol{\mu}_i = \mathbf{B}^T\mathbf{x}_i + \mathbf{N}_i\mathbf{f}$, where $\mathbf{x}_i = (x_{i1}, ..., x_{ik})^T$ denoting the ith row of the matrix $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_n)^T$, and $\mathbf{B} = (\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_k)^T$. This representation of $\boldsymbol{\mu}_i$ allows an algebraic facility of some matrix and vector operations.

Table 2.1: Some elliptical distribution and the respective density generator function.

Distribution	$g(\delta)$	
Normal	$c_1 \exp\{-\delta/2\}$	$\delta \geq 0$
Student-t	$c_2\{1+\delta/\nu\}^{-(\nu+1)/2}$	$\nu > 2$
Power Exponential	$c_3 \exp\{-\delta^\gamma/2\}$	$\gamma > 0$
Logistics	$c_4 \exp\{-u\}/(1{+}\exp\{-\delta\})^2$	$\delta \geq 0$
Laplace	$c_5 \exp\{-\sqrt{\delta}/2\}$	$\delta \geq 0$
Slash	$c_6\, \Gamma(\nu+1/2, \delta/2)$	$\delta \geq 0$
Pearson type VII	$c_7\{1+\delta/\nu\}^N$	$N > n/2, \nu > 0$
Kotz type	$c_8\delta^{N-1} \exp\{-r\delta^\nu\}$	$r, \nu > 0,\ 2N+n > 0$

δ is the Mahalanobis distance

c_l is constant of normalize

2.2 Penalized function

Considering that $\mathbf{y}_i \sim \mathrm{El}_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$, the log-likelihood function associated to the model 2.1 can be expressed as

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{n} L_i(\boldsymbol{\theta}), \tag{2.3}$$

where $L_i(\boldsymbol{\theta}) = -\frac{1}{2}\log|\boldsymbol{\Sigma}| + \log g(\delta_i)$, $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \mathbf{f}^{\mathrm{T}}, \sigma^2, \rho)^{\mathrm{T}} \in \boldsymbol{\Theta} \subseteq R^{p^*}$, $\boldsymbol{\beta} = \mathrm{vec}(\mathbf{B})$, $p^* = kp + n + 2$, and $|\boldsymbol{\Sigma}|{=}(\sigma^2)^p(1-\rho)^{p-1}[1+\rho(p-1)]$, with $vec(\cdot)$ denoting the vectorization of the matrix of parameters $\mathbf{B}$. It is important to note that the number of effective parameters of model (2.1) depends on the number of effective parameters associated with the nonparametric component. This will be considered in the following sections.

It is known fact that maximizing the log-likelihood function without imposing restrictions over the nonparametric function may cause over-fitting and non identification of $\boldsymbol{\beta}$. A well known procedure that can solve this problem is based on the idea of log-likelihood penalization and consists in incorporating a penalty function over the function $f(t)$ such that

$$L_p(\boldsymbol{\theta}, \lambda) = L(\boldsymbol{\theta}) + \lambda^* J(f),$$

where $J(f)$ denotes the penalty function over f and $\lambda^*(\lambda)$ is a constant that to depend of the smoothing parameter $\lambda > 0$. In this thesis, we will consider

$$J(f) = \int_a^b \left[f^2(t)\right]^2 \mathrm{d}t,$$

where $f^{(2)}(t) = \frac{d^2}{dt^2} f(t)$, with $t \in [a, b]$, and the function f belongs to the Sobolev space defined as

$$W_2^{(2)} = \left\{ f : f, f^{(1)}, \text{abs.cont.}, f^{(2)} \in \mathcal{L}^2[a, b] \right\}.$$

In this case, the estimatin of f leads to a smooth cubic spline with knots at the points $t_1, ...t_n$. According to Green and Siverman (1994), we may express the penalty function as quadratic form given by

$$J(f) = \mathbf{f}^T \mathbf{K} \mathbf{f},$$

where $\mathbf{K}$ is a $(n \times n)$ nonnegative definitive smoothing marix associated with the explonatory variable t and that depends only on the knots. Then, if we consider $\lambda^*(\lambda) = -\frac{1}{2}\lambda$, the penalized log-likehood function can be expressed as

$$L_p(\boldsymbol{\theta}, \lambda) = L(\boldsymbol{\theta}) - \frac{\lambda}{2} \mathbf{f}^T \mathbf{K} \mathbf{f}, \tag{2.4}$$

where λ controls the trade of between goodness of fit and the smothness estimated function.

In the next chapter we consider some theoretical aspects concerning the estimation of the parameters involved in model (2.1).

Chapter 3

Parameters estimation

In this chapter we derive the score function and hessian matrix associated with the parameter vector $\boldsymbol{\theta}$. Also, we propose a procedures for estimating the parameters based on the Backfitting algorithm. The observed matrix will be used to calculate the standard errors of the estimate $\widehat{\boldsymbol{\theta}}$.

3.1 Score function

Let

$$L_p(\boldsymbol{\theta}, \lambda) = \sum_{i=1}^{n} L_{p_i}(\boldsymbol{\theta}, \lambda)$$

where

$$L_{p_i}(\boldsymbol{\theta}, \lambda) = -\frac{1}{2}\log|\boldsymbol{\Sigma}| + \log g(\delta_i) - \frac{\lambda}{2n}\mathbf{f}^T\mathbf{K}\mathbf{f} \tag{3.1}$$

with $\delta_i = \mathbf{r}_i^T\boldsymbol{\Sigma}^{-1}\mathbf{r}_i$, $\mathbf{r}_i = \mathbf{y}_i - \boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_i = \mathbf{B}^T\mathbf{x}_i + \mathbf{N}_i\mathbf{f}$. Assuming that (3.1) is regular with respect to each element of $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \mathbf{f}^{\mathrm{T}}, \sigma^2, \rho)^{\mathrm{T}}$, the score function are given by

$$\mathbf{U}_{\boldsymbol{\theta}} = \sum_{i=1}^{n} \frac{\partial L_{p_i}(\boldsymbol{\theta}, \lambda)}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \mathbf{U}_{\boldsymbol{\beta}} \\ \mathbf{U}_{\mathbf{f}} \\ U_{\sigma^2} \\ U_{\rho} \end{pmatrix}.$$

To obtain the score function associated to $\boldsymbol{\beta}$, we must derive $L_{p_i}(\boldsymbol{\theta})$ with respect to each element of $\boldsymbol{\beta}$. For $\boldsymbol{\beta}_s$, we have

$$\frac{\partial L_{p_i}(\boldsymbol{\theta}, \lambda)}{\partial \boldsymbol{\beta}_s} = v_i x_{is} \boldsymbol{\Sigma}^{-1}(\mathbf{y}_i - \boldsymbol{\mu}_i) \tag{3.2}$$

and, consequently,

$$\frac{\partial L_{p_i}(\boldsymbol{\theta}, \lambda)}{\partial \boldsymbol{\beta}} = v_i(\mathbf{x}_i \otimes \boldsymbol{\Sigma}^{-1}(\mathbf{y}_i - \boldsymbol{\mu}_i)),$$

with $\otimes$ denoting the kronecker product. Thus,

$$\mathbf{U}_{\boldsymbol{\beta}} = \sum_{i=1}^{n} v_i(\mathbf{x}_i \otimes \boldsymbol{\Sigma}^{-1}(\mathbf{y}_i - \boldsymbol{\mu}_i)), \tag{3.3}$$

where $v_i = -2\zeta_g(\delta_i)$, with $\zeta_g(\delta_i) = \frac{d\log(\delta_i)}{d\delta_i}$ $i = 1, ..., n$. Then, after some algebraic operations, the score function for $\mathbf{B}$ takes the form

$$\mathbf{U}_{\mathbf{B}} \sum_{i=1}^{n} v_i \mathbf{x}_i (\mathbf{y}_i - \mathbf{B}^T \mathbf{x}_i - \mathbf{N}_i \mathbf{f})^T \boldsymbol{\Sigma}^{-1}$$

Analogously, to obtain the score function associated to vector $\mathbf{f}$, we must derive the function $L_{p_i}(\boldsymbol{\theta}, \lambda)$ with respect to $\mathbf{f}$. In effect,

$$\frac{\partial L_{p_i}(\boldsymbol{\theta}, \lambda)}{\partial \mathbf{f}} = v_i \mathbf{N}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{r}_i - \frac{\lambda}{n} \mathbf{K}\mathbf{f}$$

and, consequently,

$$\mathbf{U_f} = \sum_{i=1}^{n} v_i \mathbf{N}_i^T \mathbf{\Sigma}^{-1}\mathbf{r}_i - \lambda \mathbf{K}\mathbf{f}. \tag{3.4}$$

Finally, to obtain score functions of σ^2 and ρ we must calculate

$$\frac{\partial L_{p_i}(\boldsymbol{\theta}, \lambda)}{\partial \sigma^2} = -\frac{1}{2\sigma^2}\left\{p - v_i \delta_i\right\}$$

and

$$\frac{\partial L_{p_i}(\boldsymbol{\theta}.\lambda)}{\partial \rho} = -\frac{\sigma^2}{2}\left\{\gamma p(p-1) - v_i \mathbf{r}_i^T \mathbf{V}\mathbf{r}_i\right\}.$$

Applying the sum to the previous equations you get

$$\begin{aligned}\mathbf{U}_{\sigma^2} &= \sum_{i=1}^{n} -\frac{1}{2\sigma^2}\left\{p - v_i\delta_i\right\} \\ &= -\frac{n}{2\sigma^2}\{p - \operatorname{tr}(\mathbf{R}\mathbf{\Sigma}^{-1})\}\end{aligned}$$

and

$$\begin{aligned}\mathbf{U}_{\rho} &= \sum_{i=1}^{n} -\frac{\sigma^2}{2}\left\{\gamma p(p-1) - v_i \mathbf{r}_i^T \mathbf{V}\mathbf{r}_i\right\} \\ &= -\frac{n\sigma^2}{2}\{\gamma p(p-1) - \operatorname{tr}(\mathbf{R}\mathbf{V})\},\end{aligned}$$

where $\mathbf{V} = \mathbf{\Sigma}^{-1}(\mathbf{1}\mathbf{1}^T - \mathbf{I})\mathbf{\Sigma}^{-1}$, $\mathbf{1}$ denote a $(p \times 1)$ vector of 1's, $\gamma = \frac{\rho}{\sigma^2}(\rho - 1)[\rho(p-1) + 1]$ and $\mathbf{R} = \sum_{i=1}^{n} v_i \mathbf{r}_i \mathbf{r}_i^T$. The expressions of v_i for some distributions that belong to the class of elliptical distributions are presented in Table 3.1

Table 3.1: Quantities v_i, $v_i^{'}$, d_{g_i} and f_{g_i} for some distributions that belong to the elliptical class.

Distribution	v_i	$v_i^{'}$	d_{g_i}	f_{g_i}
Normal	1	0	$\frac{p}{4}$	$\frac{p(p+2)}{4}$
Student-t	$\frac{\nu+p}{\nu+\delta_i}$	$\frac{\nu+p}{(\nu+\delta_i)^2}$	$\frac{p}{4}\left(\frac{\nu+p}{\nu+p+2}\right)$	$\frac{p(p+2)}{4}\left(\frac{\nu+p}{\nu+p+2}\right)$
Power Exponential	$\gamma\delta_i^{\gamma-1}$	$\gamma(\gamma-1)\delta_i^{\gamma-2}$	$\frac{\frac{\gamma^2}{2^{1/\gamma}}\Gamma\left(\frac{p-2}{2\gamma}+2\right)}{\Gamma(\frac{p}{2\gamma})}$	$\frac{p(p+2\gamma)}{4}$
Cauchy	$\frac{1+p}{1+\delta_i}$	$\frac{1+p}{(1+\delta)^2}$	$\frac{p}{4}\left(\frac{1+p}{3+p}\right)$	$\frac{p(p+2)}{4}\left(\frac{1+p}{3+p}\right)$

p is the dimension of the response variables vector.

3.2 Hessian matrix

In this section the Hessian matrix (derivative second matrix) is determined under the elliptical multivariate semiparametric lineal regresision model with equicorrelated random errors. The Hessian matrix of dimension $(p^* \times p^*)$, denoted by $\mathbf{L}_{\mathbf{p}}^{\boldsymbol{\theta\theta}}$, is defined as follows:

$$\mathbf{L}_p^{\boldsymbol{\theta\theta}} = \frac{\partial^2 L_{p_i}(\boldsymbol{\theta},\lambda)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T} = \begin{pmatrix} \mathbf{L}_p^{\boldsymbol{\beta\beta}} & \mathbf{L}_p^{\boldsymbol{\beta}\mathbf{f}} & \mathbf{L}_p^{\boldsymbol{\beta}\sigma^2} & \mathbf{L}_p^{\boldsymbol{\beta}\rho} \\ \mathbf{L}_p^{\mathbf{f}\boldsymbol{\beta}} & \mathbf{L}_p^{\mathbf{ff}} & \mathbf{L}_p^{\mathbf{f}\sigma^2} & \mathbf{L}_p^{\mathbf{f}\rho} \\ \mathbf{L}_p^{\sigma^2\boldsymbol{\beta}} & \mathbf{L}_p^{\sigma^2\mathbf{f}} & \mathbf{L}_p^{\sigma^2\sigma^2} & \mathbf{L}_p^{\sigma^2\rho} \\ \mathbf{L}_p^{\rho\boldsymbol{\beta}} & \mathbf{L}_p^{\rho\mathrm{f}} & \mathbf{L}_p^{\rho\sigma^2} & \mathbf{L}_p^{\rho\rho} \end{pmatrix}$$

The elements of the matrix $\mathbf{L}_p^{\boldsymbol{\theta\theta}}$ are given by the following equations:

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^T} = \begin{cases} -\boldsymbol{\Sigma}^{-1}(2v_i^{'}\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}x_{is}^2 & \text{s=s'} \\ -\boldsymbol{\Sigma}^{-1}(2v_i^{'}\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}x_{is}x_{is'} & \text{s} \neq \text{s'}, \end{cases}$$

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\mathbf{f}^{\mathrm{T}}} = -x_i^T\boldsymbol{\Sigma}^{-1}\{2v_i^{'}\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma}\}\boldsymbol{\Sigma}^{-1}\mathbf{N}_i \ ,$$

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\sigma^2} = -\frac{x_{ij}}{\sigma^2}\boldsymbol{\Sigma}^{-1}\{v_i^{'}\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma}^{-1}\}\boldsymbol{\Sigma}^{-1}\mathbf{r}_i \ ,$$

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\rho} = -x_{ij}\sigma^2\boldsymbol{\Sigma}^{-1}\{v_i^{'}\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma}^{-1}\}\mathbf{V}\mathbf{r}_i \ ,$$

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\mathbf{f}\partial\mathbf{f}^{\mathrm{T}}} = -\mathbf{N}_i^T\boldsymbol{\Sigma}^{-1}\{2v_i^{'}\mathbf{r}_i\mathbf{r}_i^T + v_i\Sigma\}\Sigma^{-1}\mathbf{N}_i - \frac{\lambda}{n}\mathbf{K} \ ,$$

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\mathbf{f}\partial\sigma^2} = -\frac{1}{\sigma^2}\mathbf{N}_i^T\boldsymbol{\Sigma}^{-1}\{v_i^{'}\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma}^{-1}\}\boldsymbol{\Sigma}^{-1}\mathbf{r}_i \ ,$$

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\mathbf{f}\partial\rho} = -\mathbf{N}_i^T\sigma^2\boldsymbol{\Sigma}^{-1}\{v_i^{'}\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma}^{-1}\}\mathbf{V}\mathbf{r}_i \ ,$$

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\sigma^2\partial\sigma^2} = \frac{p}{2\sigma^4} - \frac{1}{2\sigma^4}\mathbf{r}_i^T\boldsymbol{\Sigma}^{-1}(v_i^{'}\mathbf{r}_i\mathbf{r}_i^T\boldsymbol{\Sigma}^{-1} + 2v_i\mathbf{I})\mathbf{r}_i \ ,$$

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\sigma^2\partial\rho} = -\frac{1}{2}\mathbf{r}_i^T\{\mathbf{V}(v_i^{'}\mathbf{r}_i\mathbf{r}_i^T\boldsymbol{\Sigma}^{-1} + v_i\mathbf{I})\}\mathbf{r}_i \qquad \text{and}$$

$$\frac{\partial^2 L_{p_i}(\boldsymbol{\theta})}{\partial\rho\partial\rho} = -\frac{\sigma^2\gamma_\rho p(p-1)}{2} - \frac{\sigma^4}{2}\mathbf{r}_i^T \ \mathbf{V}\{v_i^{'}(\mathbf{r}_i^T\mathbf{V}\mathbf{r}_i)\mathbf{I} + 2v_i\boldsymbol{\Sigma}\mathbf{V}\}\mathbf{r}_i \quad ,$$

with $\gamma_\rho = -\gamma\left\{\sigma^2\gamma\left(\frac{(p-1)(2\rho-1)+1}{p}\right)+1\right\}$. The expressions of $\nu_i^{'}$ for some distributions that belong to the class of elliptical distributions are presented in table [3.1]

3.3 Estimation procedure

Here we present the method for the estimation of parameter vector $\boldsymbol{\theta}$ under the elliptical multivariate semiparametric linear regression model with equicorrelated random errors. For simplicity, we consider a particular case of model (2.1); specifically, the model

$$\mathbf{y}_i = \boldsymbol{\beta} x_i + \mathbf{N}_i \mathbf{f} + \boldsymbol{\epsilon}_i,$$

where x_i is a scalar explanatory variable. The proposed estimation method is based on the Backfitting algorithm, which allows solving the estimation equations

$$\mathbf{U}_{\boldsymbol{\beta}} = 0 \quad , \quad \mathbf{U}_{\mathbf{f}} = 0 \quad , \quad \mathbf{U}_{\sigma^2} = 0 \quad \text{and} \quad \mathbf{U}_{\rho} = 0, \tag{3.5}$$

and leads to the following two-step iterative process:

Step 1 Back-fitting algorithm.

Denoted by $\boldsymbol{\theta}^{(0)}$ and $\boldsymbol{\theta}^{(s+1)}$ the starting and current values of $\boldsymbol{\theta}$, respectively, for $s = 0, 1, \ldots$. In this stage, the updating equations for $\boldsymbol{\beta}$ and $\mathbf{f}$, which are obtained from (3.5), induce an inner back-fitting algorithm, where the solution is the limit of the iteration obtained by repeatedly cycling between the following two equations:

$$\widehat{\boldsymbol{\beta}} = \left[\sum_{i=1}^{n} v_i x_i^2 (\boldsymbol{\Sigma}^{(s)})^{-1}\right]^{-1} \left[\sum_{i=1}^{n} v_i x_i (\boldsymbol{\Sigma}^{s})^{-1} (\mathbf{y}_i - \mathbf{N}_i \mathbf{f}^{(\mathrm{s+1,s^*})})\right] \tag{3.6}$$

and

$$\widehat{\mathbf{f}} = \left[\sum_{i=1}^{n}(v_i\mathbf{N}_i^T(\mathbf{\Sigma}^{(s)})^{-1}\mathbf{N}_i - \lambda\mathbf{K})\right]^{-1}\left[\sum_{i=1}^{n}v_i\mathbf{N}_i^T(\mathbf{\Sigma}^{s})^{-1}(\mathbf{y}_i - x_i\boldsymbol{\beta}^{s+1,s^*+1})\right] \tag{3.7}$$

for $s^* = 0, 1, ...,$ where $\mathbf{y} = (\mathbf{y}_1^T, ..., \mathbf{y}_n^T)^T$, with $\mathbf{X}$ and $\mathbf{N}$ being denoted in the same way. Here, we can use as statting value for the back-fitting algorithm $\mathbf{f}^{(s+1,0)} = \mathbf{f}^{(0)}$ and the MPLEs from elliptical multivariate semiparametric linear regression model with equicorrelated random errors can be used as starting values $\boldsymbol{\beta}^{(0)}$, $\mathbf{f}^{(0)}$, $\sigma^{2(0)}$ and $\rho^{(0)}$ Equations (3.6) and (3.7) are iterated until convergence.

Step 2 Concentrated penalized log-likelihood.

To update σ^2 and ρ by using the following equations:

$$\sigma^{2^{(m+1)}} = \frac{1}{p}\mathrm{tr}(\mathbf{R}^{(m)}) \quad \text{and} \quad \rho^{(m+1)} = \frac{1}{(p-1)}\left\{\frac{\mathbf{1}^T\mathbf{R}^{(m)}\mathbf{1}}{\mathrm{tr}(\mathbf{R}^{(m)})} - 1\right\},$$

where $\mathbf{R}^{(m)} = \mathbf{R}\,|_{\theta}^{(m)}$.

Thus, alternating between step 1 and 2, this iterative process leads to the MLE of $\boldsymbol{\theta}$. We suggest to use as starting values $\boldsymbol{\theta}^0 = \widehat{\boldsymbol{\theta}}_N$, the MLE for $\boldsymbol{\theta}$ under the normal model. Note that the v_i's can be interpreted as weights and since $g(\delta_i)$ is for the majarity of the elliptical a positive decreasing function, one has in general $v_i > 0$. Exceptions are Kotz, generalizd Kotz and double exponential distributions.

3.4 Effective degress of freddom

In the elliptical multivariate semiparametric linear regression model with equicorrelated random errors, the degree of freedom, which measure the individual effect contribution of the nonparametric component, is given by (see, for instance, Hastie and Tibshirani (1990).

$$\mathrm{df}(\lambda) = \sum_{i=1}^{n} \mathbf{N}_i \left[v_i \mathbf{N}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{N}_i + \lambda \mathbf{K} \right] \mathbf{N}_i^T \boldsymbol{\Sigma}^{-1}$$

3.5 Selecting an appropriate model

Under the elliptical multivariate semiparametric linear regression model with equicorrelated random errors one has a total of $kp + 2 + \mathrm{df}(\lambda)$ parameters to be estimated, with $\mathrm{df}(\lambda)$ denoting approximately the number of effective parameters involved in modelling of the nonparametric component. In this case, the Akaike information criterion (AIC) (Akaike, 1973) or the Bayesian information criterion (BIC) (Schwarz *et al.*, 1978) can be used for selecting an appropriate model. The idea is to minimize the function

$$\mathrm{AIC}(\lambda) = -2L_p(\widehat{\boldsymbol{\theta}}, \lambda) + 2[kp + 2 + \mathrm{df}(\lambda)] \, ,$$

where $L_p(\widehat{\boldsymbol{\theta}}, \lambda)$ denotes the penalized log-likelihood function available at $\widehat{\boldsymbol{\theta}}$ for a fixed λ. It is important mention that AIC is based on information theory and is useful for selecting an appropriate model given data with adequate sample size. An alternative version of the AIC criterion, denoted by AIC_c, was proposed by Hurvich *et al.* (1998) in the context of parametric linear regression and autoregressive time series. Recently, Relvas (2016) adapted this criterion for the partially linear model with first-order autoregressive symmetric errors. Considering such proposals, we propose the AIC_c criterion as an alternative for the selection of models under the elliptical multivariate semiparametric linear regression model with equicorrelated random errors, which is given by

$$\mathrm{GCV}_c(\lambda) = \log\left\{\frac{\|\sqrt{\widehat{\mathbf{W}}}(\mathbf{y}-\widehat{\mathbf{y}})\|^2}{n}\right\} + \frac{2\left[\mathrm{tr}(\widehat{\mathbf{H}(\lambda)})+1\right]}{n-\mathrm{tr}(\widehat{\mathbf{H}(\lambda)})-2} + 1 \ ,$$

3.6 Smoothing parameters

The determination of the parameter is a crucial part in the estimation process and different choice methods are available in the literature. For example, is usual to consider the cross-validation method or the generalized cross-validation method (Craven and Wahda, 1978). Following Relvas (2016), an alternative to select smoothing parameters under the elliptical multivariate semiparametric linear regression model with equicorrelated random errors, is to consider a generalized cross-validation method defined by

$$\mathrm{GCV}(\lambda) = \frac{\|\sqrt{\widehat{\mathbf{W}}}(\mathbf{y}-\widehat{\mathbf{y}})\|^2}{[1-n^{-1}\mathrm{tr}(\widehat{\mathbf{H}(\lambda)})]} \ ,$$

where $\widehat{\mathbf{H}(\lambda)}$ is commonly called smoother matrix and is equivalent to the *hat* matrix defined in the class of parametric regression models which satisfies $\widehat{\mathbf{y}} = \widehat{\mathbf{H}(\lambda)}\mathbf{y}$.

In this case, λ should be obtained by minimizing $\mathrm{GCV}(\lambda)$ for a grid of λ values. Alternatively, these parameters may be selected by applying the Akaike information criterion. In particular, we can consider the $\mathrm{AIC}(\lambda)$ or $\mathrm{AIC}_c(\lambda)$ criteria defined in the previous section, and use the effective degress of freedom involved in nonparametric modeling to select appropriate smoothing parameters.

In the next chapter we will consider some theoretical aspects of the Local Influence Method for model 2.1.

Chapter 4

Local Influence

In this section we derive the normal curvatures of local influence for the model (2.1) under some usual perturbation schemes. We will consider the case-weight, scale parameter and explanatory variable perturbation schemes.

4.1 The method

Let $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^T$ be and $(n \times 1)$ vector of perturbations restricted to some open subser $\boldsymbol{\Omega} \in R^n$ and the logarithm of the perturbed liklihood function denoted by $L(\boldsymbol{\theta}, \lambda|\boldsymbol{\omega})$. Suppose that there is a point $\omega_0 \in \Omega$ that represents no perturbation of the data so that $L(\boldsymbol{\theta}, \lambda|\boldsymbol{\omega}_0) = L(\boldsymbol{\theta}, \lambda)$ To assess the influence of minor perturbations on $\widehat{\boldsymbol{\theta}}$, we consider the likelihood displacement

$$LD_\omega = 2\left[L(\widehat{\boldsymbol{\theta}}, \lambda) - L(\widehat{\boldsymbol{\theta}}_w, \lambda)\right] \geq 0, \tag{4.1}$$

where $\widehat{\boldsymbol{\theta}}_w$ is the MLE under $L(\boldsymbol{\theta}, \lambda|\omega)$. This measure $LD(\omega)$ is useful for assessing the distance between $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}_w$. Cook (1986) suggested studying the local behavior of $LD(\omega)$ around ω_0. The procedure consists in selecting a unit direction $\boldsymbol{\ell} \in \Omega$ ($\|\boldsymbol{\ell}\| = 1$), and then to considering the plot of $LD(\boldsymbol{\omega}_0 + a\boldsymbol{\ell})$ against a, where $a \in \mathcal{R}$. This plot is called lifted line. Each lifted line can be characterized by considering the normal curvature $C_\ell(\boldsymbol{\theta})$ around $a = 0$. The suggestion is to consider the direction $\boldsymbol{\ell} = \boldsymbol{\ell}_{max}$ corresponding to the largest curvature $C_{\ell_{max}}(\boldsymbol{\theta})$. The index plot of $\boldsymbol{\ell}_{max}$ may reveal those observations that under small perturbations exercise influence on $LD(\boldsymbol{\omega})$.

According to Cook (1986), the normal curvature in the unitary direction $\boldsymbol{\ell}$ is given by $C_\ell(\boldsymbol{\theta}) = -2\{\boldsymbol{\ell}^T\boldsymbol{\Delta}^T\ddot{\mathbf{L}}^{-1}\boldsymbol{\Delta}\boldsymbol{\ell}\}$, where

$$\ddot{\mathbf{L}} = \left.\frac{\partial^2 L(\boldsymbol{\theta}, \lambda)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right|_{\widehat{\boldsymbol{\theta}}} \quad \text{and}$$

$$\boldsymbol{\Delta} = \left.\frac{\partial^2 L(\boldsymbol{\theta}, \lambda\,|\,\boldsymbol{\omega})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^T}\right|_{\widehat{\boldsymbol{\theta}},\, \boldsymbol{\omega}_0}.$$

In order to have a curvature invariant under uniform change of scale Poon and Poon (1999) proposed the conformal normal curvature defined as

$$B_\ell(\boldsymbol{\theta}) = -\frac{\boldsymbol{\ell}^T\boldsymbol{\Delta}^T\ddot{\mathbf{L}}^{-1}\boldsymbol{\Delta}\boldsymbol{\ell}}{\sqrt{\mathrm{tr}(\boldsymbol{\Delta}^{\mathrm{T}}\ddot{\mathbf{L}}^{-1}\boldsymbol{\Delta})^2}}.$$

This curvature is characterized to allow for any unitary direction $\boldsymbol{\ell}$ that $0 \leq B_\ell(\boldsymbol{\theta}) \leq 1$. A suggestion is to consider the direction $\boldsymbol{\ell} = \boldsymbol{\ell}_{max}$ corresponding to the largest curvature $B_{\ell_{max}}(\boldsymbol{\theta})$ or, alternatively, evaluating the normal curvature at the direction $\boldsymbol{\ell} = \mathbf{e}_i$, where $\mathbf{e}_i$ is an n-dimensional vector with 1 in the ith position and zeros in the remaining positions, and observing the index plot of $B_{\mathbf{e}_i}(\boldsymbol{\theta})$. We suggest using $B_i > \mathcal{B} + c^*\mathrm{SE}(\mathcal{B})$ to discriminate if an observation is influential or not, where $\mathcal{B}$ is the mean of $\mathcal{B} = \{B_i = B_{\mathbf{e}_i}(\boldsymbol{\theta}) : i = 1, \ldots, n\}$, $\mathrm{SE}(\mathcal{B})$ denotes the standard error of $\mathcal{B}$ and c^* is a constant selected appropriately.

4.2 Normal curvature derivation

In this section we present the expressions of the elements of the matrix $\boldsymbol{\Delta}$ for case-weight, scale matrix and explanatory variable perturbation schemes.

4.2.1 Case-weight perturbation

Let us consider the attributed weights for the observations in the log-likelihood function as

$$L(\boldsymbol{\theta}, \lambda|\boldsymbol{\omega}) = \sum_{i=1}^{n} \omega_i \, L_i(\boldsymbol{\theta}) \; - \frac{\lambda}{2}\mathbf{f}^T\mathbf{K}\mathbf{f}, \tag{4.2}$$

where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^T$ is the vector of weights, with $0 \leq \omega_i \leq 1$ $(i = 1, \ldots, n)$. In this case, the vector of no perturbation is given by $\boldsymbol{\omega}_0 = (1, \ldots, 1)^T$. Differentiating (4.2) with respect to the elements of $\boldsymbol{\theta}$ and ω_i we obtain, the expressions

$$\begin{aligned} \frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial \boldsymbol{\beta}_s \partial \omega_i} &= v_i x_{is} \, \boldsymbol{\Sigma}^{-1}\mathbf{r}_i\Big|_{\widehat{\boldsymbol{\theta}}} \\ \frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial \mathbf{f} \partial \omega_i} &= v_i \mathbf{N}_i^T \boldsymbol{\Sigma}^{-1}\mathbf{r}_i\Big|_{\widehat{\boldsymbol{\theta}}} \end{aligned}$$

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\sigma^2\partial\omega_i} = -\frac{1}{2\sigma^2}\{p - v_i\delta_i\}\Big|_{\widehat{\boldsymbol{\theta}}} \qquad \text{and}$$

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\rho\partial\omega_i} = -\frac{\sigma^2}{2}\{\gamma p(p-1) - v_i\mathbf{r}_i^T\mathbf{V}\mathbf{r}_i\}\Big|_{\widehat{\boldsymbol{\theta}}}.$$

4.2.2 Scale perturbation

Under the scale perturbation scheme we assume that

$$\mathbf{y}_i \sim \mathrm{El_p}(\boldsymbol{\mu}_i, \omega_i^{-1}\boldsymbol{\Sigma}), \tag{4.3}$$

where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^T$ is the vector of perturbations, with $\omega_i > 0$, for $i = 1, \ldots, n$. In this case, the vector of no perturbation is given by $\boldsymbol{\omega}_0 = (1, \ldots, 1)^T$ such that $L(\boldsymbol{\theta}, \lambda|\boldsymbol{\omega}) = L(\boldsymbol{\theta}, \lambda)$. Taking differentials of $L(\boldsymbol{\theta}, \lambda|\boldsymbol{\omega})$ with respect to the elements of $\boldsymbol{\theta}$ and ω_i, simple algebra yields

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\boldsymbol{\beta}_s\partial\omega_i} = \{v_i^{'}\,\delta_i + v_i\}\ x_{is}\boldsymbol{\Sigma}^{-1}\mathbf{r}_i\Big|_{\widehat{\boldsymbol{\theta}}}\ ,$$

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\mathbf{f}\partial\omega_i} = \{v_i^{'}\delta_i + v_i\}\mathbf{N}_i^T\boldsymbol{\Sigma}^{-1}\mathbf{r}_i\Big|_{\widehat{\boldsymbol{\theta}}}\ ,$$

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\sigma^2\partial\omega_i} = \frac{\delta_i}{2\sigma^2}\{v_i'\delta_i + v_i\}\Big|_{\widehat{\boldsymbol{\theta}}} \qquad \text{and}$$

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\rho\partial\omega_i} = \frac{\sigma^2}{2}\{v_i'\delta_i + v_i\}\mathbf{r}_i^T\mathbf{V}\mathbf{r}_i\Big|_{\widehat{\boldsymbol{\theta}}}$$

4.2.3 Explanatory variable perturbation

Here the sth explanatory variable (assumed continuous) is perturbed by considering additive perturbation scheme, namely $x_{is\omega} = x_{is} + \omega_i$ $(i = 1, \ldots, n)$, where x_{is} denotes the sth explanatory variable value for the ith cluster and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^T$ is the vector of perturbations. In this case, the desing matrix $\mathbf{X}_i$ associated to the parametric components of the model is replaced by $\mathbf{X}_{i\omega} = (x_{i1}...x_{ik\omega}...x_{ip})$. Thus, perturbed penalized log-likelihood function is constructed from (2.4) with $\mathbf{X}_{is}$ replaced by $\mathbf{X}_{i\omega}$, that is,

$$L(\boldsymbol{\theta}, \lambda|\boldsymbol{\omega}) = L(\boldsymbol{\theta}|\boldsymbol{\omega}) - \frac{\lambda}{2}\mathbf{f}^T\mathbf{K}\mathbf{f}, \tag{4.4}$$

where $L_i(\boldsymbol{\theta}|\boldsymbol{\omega}) = -\frac{1}{2}\log|\Sigma| + \log g(\delta_{iw})$ with $\delta_{i\omega} = \mathbf{r}_{i\omega}^T\boldsymbol{\Sigma}^{-1}\mathbf{r}_{i\omega}$, $\mathbf{r}_{i\omega} = \mathbf{y}_i - \boldsymbol{\mu}_{i\omega}$ and $\boldsymbol{\mu}_{i\omega} = \boldsymbol{\beta}_1 x_{i1} + ... + \boldsymbol{\beta}_s x_{isw} + ... + \boldsymbol{\beta}_k x_{ik} + \mathbf{N}_i^T\mathbf{f}$. In this case, the vector of no perturbation is given by $\boldsymbol{\omega}_0 = (0, ...0)^T$, since $x_{isw_0} = x_{is}$, for $i = 1, ..., n$ and $s = 1, ..., k$. Differentiating (4.4) with respect to $\boldsymbol{\theta}$ and w_i, we have, for $s = 1, ..., k$, that

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\boldsymbol{\beta_s}\partial\omega_i} = -2v_i' x_{is}\boldsymbol{\Sigma}^{-1}\mathbf{r}_i\mathbf{r}_i^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}_s - v_i\boldsymbol{\Sigma}^{-1}\big\{x_{is}\boldsymbol{\beta_s} - \mathbf{r}_i\big\}\Big|_{\widehat{\boldsymbol{\theta}}},$$

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\mathbf{f}\partial\omega_i} = -\mathbf{N}_i^T\boldsymbol{\Sigma}^{-1}\{2v_i'\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma}^{-1}\}\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta}_s\Big|_{\widehat{\boldsymbol{\theta}}},$$

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\sigma^2\partial\omega_i} = -\frac{1}{\sigma^2}\mathbf{r}_i^T\boldsymbol{\Sigma}^{-1}\{v_i'\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma}\}\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta_s}\Big\}\Big|_{\widehat{\boldsymbol{\theta}}} \quad \text{and}$$

$$\frac{\partial^2 L_i(\widehat{\boldsymbol{\theta}}|\boldsymbol{\omega}_0)}{\partial\rho\partial\omega_i} = -\sigma^2\mathbf{r}_i^T\mathbf{V}\{v_i'\mathbf{r}_i\mathbf{r}_i^T + v_i\boldsymbol{\Sigma}\}\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta_s}\Big\}\Big|_{\widehat{\boldsymbol{\theta}}}$$

for $i = 1, ..., n$.

Chapter 5

Conclusions and future research

In this thesis we consider the study of the elliptical multivariate semiparametric linear regression model with equicorrelated random errors, which is an extension of the models proposed by Ibacache *et al*. (2012, 2014). In this context, we study some theoretical aspects associated with the process of parameter estimation. In addition, we developed the technique of local influence for different perturbation. Among the results that stand out the most are the calculation of the score functions, the matrix of derivatives second and the derivation of the normal curvature.

Among the future works to develop we have the following:

(i) The application of the proposed model and the local influence technique to a set of real data.

(ii) Incorporate the nonlinear effect of at least two explanatory variables. This proposal would give rise to the elliptic multivariate additive semiparametric linear regression model with equicorrelated random errors.

(iii) Consider new structures for the scale matrix associated with the elliptical distribution.

Bibliography

[1] Akaike, H.(1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, **60(2)**: 255–265.

[2] Cook, D. (1986). Assessment of Local Influence. *Journal of the Royal Statistical Society. Series B (Methodological)*, **48**, 133-168

[3] Christensen, R. (1997). Linear Models for Multivariate, Time Series, and Spatial Data. *Springer Texts in Statistics.*

[4] Craven, P. and Wahba, G. Numer. Math. (1978). Smoothing noisy data with spline functions. 31: 377.

[5] Dabson, A. (2002). An Introduction to Generalized Linear Models. Second Edition. *Chapman and Hall/CRC.*

[6] Durban, M: Hackett, C. and Currie, I. (1999) Approximate standard errors in semiparametric models. *Biometric*, **55**: 699-703.

[7] Durban, M.; Hackett, C., McNicol, J.; Newton, A.; Thomas, W. and Currie, I. (2003) The praticaluse of semiparametric models in field trials. *Journal of agricultural, biological and Environmental Statistics*, **8(1)**: 48-66.

[8] Green, P. J. (1987). Penalized likelihood for general semi-parametric regression models. *International Statistical Review*, **55**, 245-259.

[9] Green, P. J. and Silverman, B. W. (1994). Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach. *Chapman and Hall*.

[10] Hastie, T., Tibshirani, R. (1986). Generalized additive models. *Statistical Science*, **1**, 297-318.

[11] Hastie, T. and Tibshirani, R. (1990), Generalized Additive Models, Chapman and Hall.

[12] Hurvich, Clifford M., Simonoff, Jeffrey S. and Tsai, Chih-Ling. (1998). Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, **60(2)**: 271-293

[13] Ibacache-Pulgar, G. and Gilberto A. Paula. (2011). Local Influence for Student-t partially linear models. *Computational Satatistis and Data Analysis*, **55 (2011**): 1462-1478.

[14] Ibacache-Pulgar, G., Paula, G.A. and Galea, M. (2012). Influence diagnostic for ellipticl semiparametric mixed model. *Statistical Modelling*, **12(2)**: 165-193.

[15] Ibacache-Pulgar, G., Paula, G. A. and Galea, M. (2014). On influence diagnostic in elliptical multivariate regression models with equicorrelated random errors. *Statistical Methodology*, **16**,14-31.

[16] McCullagh, P. and Nelder, J. A. (1989). Generalized Linear Models. *Chapman and Hall/CRC*.

[17] Poon W.-Y. and Poon S. (1999).Conformal normal curvature and assessment of local influence. *Royal Statistical Science*. **61**: 51-61.

[18] Relvas, C.E.M. and Paula, G.A. (2016). Partially linear models with first-order autoregressive symmetric errors. *Stat Papers* **57**: 795. https://doi.org/10.1007/s00362-015-0680-4

[19] Seber,G.a. (2004) Multivariate observations. *John Wiley*, New York.

[20] Schwarz, G., et al. (1978) Estimating the Dimension of a Model. *The Annals of Statistics* **6**: 461-464.

[21] Seber, G.A. Multivariate Observations. *John Wiley*. New York.

[22] Reyes, S. (2016). Local influence for elliptical partially varying-coefficient model. En: Tesis de Magíster. Departamento de Estadística. Univeridad de Valparaíso.

[23] Yandell & Green(1986). Yandell, B. S and Green, P. J. (1986). Semi-parametric generalized linear model diagnostics. *Technical Report, Dept. of Statistics, U. of Wisconsin.*