



HAL
open science

On Conditional Quantiles Approximation for Elliptical Distributions

Véronique Maume-Deschamps, Didier Rullière, Antoine Usseglio-Carleve

► **To cite this version:**

Véronique Maume-Deschamps, Didier Rullière, Antoine Usseglio-Carleve. On Conditional Quantiles Approximation for Elliptical Distributions. 2016. hal-01339520v1

HAL Id: hal-01339520

<https://hal.science/hal-01339520v1>

Preprint submitted on 29 Jun 2016 (v1), last revised 10 Jan 2017 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON CONDITIONAL QUANTILES APPROXIMATION FOR ELLIPTICAL DISTRIBUTIONS

V. MAUME-DESCHAMPS, D. RULLIÈRE, AND A. USSEGLIO-CARLEVE

ABSTRACT. In this work, we propose some approximations for the conditional quantile of one component of a random vector, given the other components. We focus on the case where the vector has an elliptical distribution. We first give exact expressions for conditional quantiles, and discuss problems that occur for computing these values. A first affine regression quantile estimator is detailed, its distribution is given, and direct simple expressions are derived for some particular elliptical distributions. The performance of this regression quantile is shown to be very poor for extremal quantile levels, so that a second approximation is proposed. We prove that this new extremal approximation is asymptotically equivalent to the true conditional quantile. Through numerical illustrations, the study shows that for usual techniques as Kriging, Quantile Regression may perform poorly when one leaves the usual Gaussian random field assumption, thus justifying the use of proposed extremal quantile approximations.

Keywords: *Elliptical distribution; quantile regression; extremal quantile; Kriging.*

1. INTRODUCTION

Kriging, introduced by Krige (1951), and formalized by Matheron (1963), aims at estimating the conditional mean of a process $(Z_t)_{t \in T}$ given the values Z_{t_1}, \dots, Z_{t_N} of the process at some points $t_1, \dots, t_N \in T$, where typically $T \subset \mathbb{R}^d$. When using the Kriging techniques, for any $x \in T$, the conditional mean of Z_x given Z_{t_1}, \dots, Z_{t_N} is approximated by a linear combination of Z_{t_1}, \dots, Z_{t_N} where the weight vector is the solution of a least square minimization problem (cf. Ligas and Kulczycki (2010)). It seems natural to approximate, in the same spirit as Kriging, other functionals by linear combinations. Our starting point is to apply this method to conditional quantiles.

In 1978, Koenker and Bassett proposed a conditional quantile estimation as an affine combination of Z_{t_1}, \dots, Z_{t_N} , called Quantile Regression (cf. Koenker and Bassett (1978)). The weight vector is the solution of a minimization problem, with an asymmetric loss function, different from the least squares.

Obviously, for a Gaussian process, $Z_x | (Z_{t_1}, \dots, Z_{t_N})$ is still Gaussian, and the conditional quantile of Z_x given $(Z_{t_1}, \dots, Z_{t_N})$ is easily computed. For non Gaussian processes, explicit formulas for conditional quantiles are more difficult to get. The Quantile Regression approach usually requires time consuming simulations to compute expectations. Moreover, the consistency of the estimation by quantile regression is not guaranteed.

In this paper, we focus on elliptical distributions. Elliptical distributions, formalized by Cambanis et al. (1981), have the advantage of being stable under affine transformations. Therefore, explicit formulas for the quantile regression may be obtained for consistent elliptical distributions (cf. Kano (1994)). Nevertheless, the

quantile regression is generally not equal to the conditional quantile and the difference may be large, especially for extreme levels of quantile. This is one reason why we propose an approximation that is adapted to conditional extremal quantiles.

The paper is organized as follows. In Section 2, we give some definitions, properties and examples of elliptical distributions satisfying the consistency property. For these models, we give formulas for conditional quantiles in Section 3. The Section 4 is devoted to quantile regression for consistently elliptical distributions: closed formulas are obtained. In Section 5, we prove asymptotic equivalences when the quantile level is close to 0 or 1. Section 6 provides a numerical study. In particular, we emphasize the fact that quantile regression is generally not consistent, especially for high level quantiles. We illustrate this point on several examples.

2. ELLIPTICAL DISTRIBUTIONS

We recall here the definition of an elliptical vector. We also give some useful properties and classical examples of elliptical distributions. Most results may be found, for instance, in Frahm (2004).

Definition 2.1. *Let X be a d -dimensional random vector. X is elliptical if and only if there exists a unique $\mu \in \mathbb{R}^d$, a semi-positive definite matrix $\Sigma \in \mathbb{R}^{d \times d}$, and a function $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}$ such that the characteristic function of $(X - \mu)$ is*

$$\mathbb{E}[\exp(it(X - \mu))] = \Phi(t'\Sigma t).$$

For such an elliptical random vector, we write $X \sim \mathcal{E}_d(\mu, \Sigma, \Phi)$

It seems important to note that Σ is not necessarily the covariance matrix of X . More precisely, Σ is proportional to the covariance matrix K of the random vector X , when it is defined, i.e there exists a positive coefficient τ such that:

$$(2.1) \quad \Sigma = \tau K$$

For example, for Gaussian distributions, $\tau = 1$, i.e $\Sigma = K$. But this is not always the case for all elliptical distributions: for Student distributions with ν degrees of freedom, $\tau = \frac{\nu-2}{\nu}$. Furthermore, K may not exist (e.g. for Cauchy distributions). In the present paper, we do only consider the case of non-degenerated distributions, i.e. we assume that the matrix Σ is invertible.

The following representation theorem is central in the theory of elliptical distributions. This result may be found in Cambanis et al. (1981).

Theorem 2.1 (Cambanis, Huang, Simons, 1981). *The random vector X is elliptical, $X \sim \mathcal{E}_d(\mu, \Sigma, \Phi)$, if and only if*

$$(2.2) \quad X = \mu + R\Lambda U^{(d)},$$

where $\Lambda\Lambda^T = \Sigma$, $U^{(d)}$ is a d -dimensional random vector uniformly distributed on \mathcal{S}^{d-1} (the unit disk of dimension d), and R is a non-negative random variable independent of $U^{(d)}$.

The representation of Theorem 2.1 is not unique (see Cambanis et al., 1981, for details). Given μ and Σ , the elliptic random vector X is characterized by the non-negative random variable R , called the radius of X . We now recall the consistency property of an elliptical distribution. The related definitions and properties may be found in Kano (1994).

Definition 2.2. *Let $X \sim \mathcal{E}_d(\mu, \Sigma, \Phi)$. X is said consistent if Φ is dimension-free, i.e. if Φ does not depend on d .*

Distribution	Coefficient c_d	Generator $g_d(t)$
Gaussian	$\frac{1}{(2\pi)^{\frac{d}{2}}}$	$\exp(-\frac{t}{2})$
Student, $\nu > 0$	$\frac{\Gamma(\frac{d+\nu}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{(\nu\pi)^{\frac{d}{2}}}$	$(1 + \frac{t}{\nu})^{-\frac{d+\nu}{2}}$
Logistic	$\frac{\Gamma(\frac{d}{2})}{(2\pi)^{\frac{d}{2}}} \left[\int_0^{+\infty} x^{\frac{d}{2}-1} \cdot \frac{e^{-x}}{(1+e^{-x})^2} dx \right]^{-1}$	$\frac{\exp(-\frac{1}{2}t)}{(1+\exp(-\frac{1}{2}t))^2}$
Kotz, $q, r, s > 0$	$\frac{s\Gamma(\frac{d}{2})}{\pi^{\frac{d}{2}} \Gamma(\frac{2q+d-2}{2s})} r^{\frac{2q+d-2}{2s}}$	$t^{q-1} \exp(-rt^s)$
Unimodal GM	$\frac{1}{(2\pi)^{\frac{d}{2}}}$	$\sum_{k=1}^n \pi_k \theta_k^d \exp\left(-\frac{\theta_k^2}{2} t\right)$
Laplace, $\lambda > 0$	$\frac{2}{\lambda(2\pi)^{\frac{d}{2}}}$	$\frac{K_{\frac{d}{2}-1}\left(\sqrt{\frac{\lambda}{2}} t\right)}{\left(\sqrt{\frac{\lambda}{2}} t\right)^{\frac{d}{2}-1}}$
Uniform GM	$\frac{\Gamma(\frac{d+1}{2})}{\sqrt{2}\pi^{\frac{d}{2}}}$	$\frac{\chi_{d+1}^2(t)}{t^{\frac{d+1}{2}}}$

TABLE 1. Some classical d -dimensional elliptical distributions with corresponding normalisation constants and generators

Kano (1994) established the following relation between Definition 2.2 and the radius R .

Proposition 2.2. *Let $X \sim \mathcal{E}_d(\mu, \Sigma, \Phi)$, and let R be the corresponding radius of X . X is consistent if and only if :*

$$(2.3) \quad R \stackrel{d}{=} \frac{\chi_d}{\epsilon},$$

where χ_d is the square root of a χ^2 distributed random variable with d degrees of freedom, ϵ is a non-negative random variable which does not depend on d , and χ_d , ϵ and $U^{(d)}$ are mutually independent.

Theorem 2.3 (Elliptical density). *Let $X \in \mathbb{R}^d$ be an elliptical random vector, $X \sim \mathcal{E}_d(\mu, \Sigma, \Phi)$, and let R be the corresponding radius of X .*

$$(2.4) \quad f_X(x) = \frac{c_d}{|\det(\Lambda)|} g_d\left((x - \mu)\Sigma^{-1}(x - \mu)\right)$$

where $g_d(t) = \frac{\Gamma(\frac{d}{2})}{2\pi^{\frac{d}{2}}} \sqrt{t}^{-(d-1)} f_R(\sqrt{t})$, and $f_R(t)$ is the p.d.f of R .

The coefficient c_d is called *the normalization constant* and the function g_d is called *the generator* of X . Table 1 provides some examples of elliptical distributions, associated with their coefficients and generators. Most of them may easily be found in the literature: Kotz distribution is introduced in Nadarajah (2003), Student in Nadarajah and Kotz (2004), and Laplace in Eltoft et al. (2006) and Kozubowski et al. (2013). We added two other distributions, obtained with Kano's representation : *Unimodal* Gaussian Mixture (Unimodal GM), and another one which will be called *Uniform* Gaussian Mixture (Uniform GM).

In Table 1, $K_m(x)$ denotes the modified Bessel function of the second kind and order m , and $\chi_m^2(x)$ denotes the c.d.f of the χ^2 distribution with m degrees of freedom, evaluated at x . Remark that the Cauchy distribution corresponds to a Student distribution with $\nu = 1$ degree of freedom.

In this paper, we focus on consistent elliptical distributions. Unfortunately, Logistic and Kotz distributions does not have this property (except the Kotz distribution with $s = q = 1$, and $r = \frac{1}{2}$, i.e the Gaussian distribution). Therefore, our examples

Distribution	ϵ
Gaussian	1
Student, $\nu > 0$	$\frac{X_\nu}{\sqrt{\nu}}$
Unimodal Gaussian Mixture	$\sum_{k=1}^n \pi_k \delta_{\theta_k}$
Laplace, $\lambda > 0$	$\frac{1}{\sqrt{\mathcal{E}(\frac{1}{\lambda})}}$
Uniform Gaussian Mixture	$\mathcal{U}([0, 1])$

TABLE 2. Some classical consistent d -dimensional elliptical distributions with corresponding random variable ϵ

will only refer to the distributions mentioned in Table 1. For these models, the non negative random variable ϵ is given in Table 2.

We have seen that an elliptical distribution is characterized by parameters μ , Σ , and by either the characteristic function Φ , the radius R or the generator g_d . For this reason, we define the distribution of an elliptical random vector by any of these three possible characterization, using indifferently the notation $X \sim \mathcal{E}_d(\mu, \Sigma, \Phi)$, $X \sim \mathcal{E}_d(\mu, \Sigma, R)$ or $X \sim \mathcal{E}_d(\mu, \Sigma, g_d)$. At last, in order to emphasis the role played by the radius and the dimension, we also use the denomination (R, d) -elliptical, as defined hereafter.

Definition 2.3. *An elliptical random vector of \mathbb{R}^d with radius R is called (R, d) -elliptical.*

The following proposition, from Hult and Lindskog (2002), is the basis of our study.

Proposition 2.4 (Affine transformation). *Let X a consistent (R, d) -elliptical random vector with parameters μ and Σ . Then for any $c \in \mathbb{R}^d$, $c^T X$ is $(R, 1)$ -elliptical with parameters $c^T \mu$ and $c^T \Sigma c$.*

Proposition 2.4 implies that an affine transformation of a (R, d) -elliptical random vector is a $(R, 1)$ -elliptical random variable. The proposition below is a direct consequence of this result (see Hult and Lindskog, 2002, for a proof).

Proposition 2.5 (Subvectors distributions). *Let $X = (X_1, X_2)^T$ be a consistent (R, d) -elliptical random vector with $X_1 \in \mathbb{R}^{d_1}$, $X_2 \in \mathbb{R}^{d_2}$, $d_1 + d_2 = d$ and parameters μ and Σ . Let us write:*

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}.$$

Then X_1 and X_2 are respectively (R, d_1) - and (R, d_2) -elliptical with parameters μ_1 , Σ_{11} and μ_2 , Σ_{22} , respectively.

Remark that a p -dimensional subvector of a consistent (R, d) -elliptical random vector is (R, p) -elliptical. As a consequence when $p = 1$, all marginals are $(R, 1)$ -elliptical. The following proposition gives some indications concerning the conditional distributions of elliptical vectors.

Proposition 2.6 (Conditional distribution). *Let $X = (X_1, X_2)^T$ be a consistent (R, d) -elliptical random vector with $X_1 \in \mathbb{R}^{d_1}$, $X_2 \in \mathbb{R}^{d_2}$, $d_1 + d_2 = d$ and parameters μ and Σ . Let us write:*

$$(2.5) \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}.$$

The conditional distribution $X_2|(X_1 = x_1)$ has parameters:

$$(2.6) \quad \begin{cases} \mu_{2|1} = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1) \\ \Sigma_{2|1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{cases}$$

Furthermore, $X_2|(X_1 = x_1)$ is elliptical, with generator R^* given by :

$$(2.7) \quad R^* \stackrel{d}{=} R\sqrt{1-\beta} \left| \left(R\sqrt{\beta}U^{(d)} = C_{11}^{-1}(x_1 - \mu_1) \right) \right.$$

where C_{11} is the Cholesky root of Σ_{11} , and $\beta \sim \text{Beta}(\frac{d_1}{2}, \frac{d_2}{2})$.

At last, the conditional density of $X_2|(X_1 = x_1)$ is given by :

$$(2.8) \quad f_{X_2|X_1}(x_2|x_1) = \frac{c_{2|1}}{|\Sigma_{2|1}|^{\frac{1}{2}}} g_d \left(q_1 + (x_2 - \mu_{2|1})^T \Sigma_{2|1}^{-1} (x_2 - \mu_{2|1}) \right)$$

with $c_{2|1} = \frac{c_d}{c_{d_1} g_{d_1}(q_1)}$, and $q_1 = (x_1 - \mu_1)^T \Sigma_{11}^{-1} (x_1 - \mu_1)$.

Proof. For Equations (2.6) and (2.7), a proof may be found in Frahm (2004). Concerning Equation (2.8), the proof is obvious, since $f_{X_2|X_1}(x_2|x_1) = \frac{f_X(x)}{f_{X_1}(x_1)}$. Then :

$$(2.9) \quad f_{X_2|X_1}(x_2|x_1) = \frac{\frac{c_d}{|\Sigma|^{\frac{1}{2}}} g_d \left((x - \mu)^T \Sigma^{-1} (x - \mu) \right)}{\frac{c_{d_1}}{|\Sigma_{11}|^{\frac{1}{2}}} g_{d_1}(q_1)}$$

Since $(x - \mu)^T \Sigma^{-1} (x - \mu) = q_1 + (x_2 - \mu_{2|1})^T \Sigma_{2|1}^{-1} (x_2 - \mu_{2|1})$ and $|\Sigma| = |\Sigma_{11}| |\Sigma_{2|1}|$, we get the expected result. \square

We have introduced the main definitions and properties of elliptical distributions. In the following section, we focus on conditional quantiles of $X_2|(X_1 = x_1)$, where $X_2 \in \mathbb{R}$ and $X_1 \in \mathbb{R}^N$.

3. CONDITIONAL QUANTILES

From now on, we consider the following context: $X = (X_1, X_2)^T$ is a consistent $(R, N+1)$ -elliptical vector, with $X_1 \in \mathbb{R}^N$ and $X_2 \in \mathbb{R}$. The aim is to estimate the α -quantile of $X_2|(X_1 = x_1)$.

Firstly, we give a general expression for $X_2|(X_1 = x_1)$. We then explain problems that occur for computing this value, and give two examples where we get explicit formulas: Gaussian and Student distributions.

3.1. General case. Let us firstly give an expression of the theoretical conditional quantile. Consider the respective distribution functions Φ_R and Φ_{R^*} ,

$$(3.1) \quad \begin{cases} \Phi_R(x) &= \mathbb{P}(RU^{(1)} \leq x), \\ \Phi_{R^*}(x) &= \mathbb{P}(R^*U^{(1)} \leq x), \end{cases}$$

where $U^{(1)}$ is 1 or -1 with probability $\frac{1}{2}$. Φ_R and Φ_{R^*} are respective distribution functions of the reduced centered $(R, 1)$ -elliptical random variable and $(R^*, 1)$ -elliptical random variable. With this notation and the conditional generator given in Equation (2.7), we get the following result.

Proposition 3.1 (Conditional elliptical quantile). *Let $X = (X_1, X_2)^T$ be a consistent $(R, N+1)$ -elliptical random vector with $X_1 \in \mathbb{R}^N$, $X_2 \in \mathbb{R}$ and parameters μ and Σ . Let us write Σ :*

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

Then the α -quantile of $X_2|(X_1 = x_1)$ is given by :

$$(3.2) \quad q_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sqrt{\Sigma_{2|1}} \Phi_{R^*}^{-1}(\alpha)$$

where $\mu_{2|1}$ and $\Sigma_{2|1}$ are given in Equation (2.6), and R^* in Equation (2.7).

Proof. We know that $X_2|(X_1 = x_1) \sim \mu_{2|1} + R^* \sqrt{\Sigma_{2|1}} U^{(1)}$. Then, our aim is to get q_α such as :

$$\mathbb{P}(\mu_{2|1} + R^* \sqrt{\Sigma_{2|1}} U^{(1)} \leq q_\alpha) = \alpha$$

Since $\mathbb{P}(\mu_{2|1} + R^* \sqrt{\Sigma_{2|1}} U^{(1)} \leq q_\alpha) = \mathbb{P}\left(R^* U^{(1)} \leq \frac{q_\alpha - \mu_{2|1}}{\sqrt{\Sigma_{2|1}}}\right) = \Phi_{R^*}\left(\frac{q_\alpha - \mu_{2|1}}{\sqrt{\Sigma_{2|1}}}\right)$, we easily get the expected result. \square

A general expression for conditional quantiles of consistent elliptical distributions is thus available. However, Equation (2.7) shows that the conditional radius R^* is difficult to exploit and thus, the calculation of the conditional quantile is not tractable in general. Indeed, the term $\Phi_{R^*}^{-1}(\alpha)$, is difficult to calculate, since the conditional density in Equation (2.8) leads to the inverse problem in x :

$$\int_{-\infty}^x c_{2|1} g_{N+1}(q_1 + t^2) dt = \alpha$$

where $c_{2|1}$ is given in Proposition 2.6. Computing and inverting the function Φ_{R^*} from the distribution of R^* is not easier, since the latter is hard to obtain. We thus have a general expression for the conditional quantile which is not easy to compute in the general case.

Fortunately, in several cases, we get an explicit formula for conditional quantiles. We propose two classical examples: Gaussian and Student distributions. We thus get explicit formulas for some simple cases, but it would not be possible in some other cases, such as Gaussian Mixture distribution.

3.2. Gaussian case. The Gaussian case is the most famous one and the simplest case of elliptical distribution. Since a conditional Gaussian distribution is still Gaussian, we have:

$$(3.3) \quad X_2|(X_1 = x_1) \sim \mathcal{N}(\mu_{2|1}, \Sigma_{2|1})$$

with $\mu_{2|1}$ and $\Sigma_{2|1}$ given in Equation (2.6). Then, the calculation of the conditional α -quantile of $X_2|(X_1 = x_1)$ is immediate, and gives:

$$(3.4) \quad q_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sqrt{\Sigma_{2|1}} \Phi^{-1}(\alpha)$$

where Φ is the standard Gaussian distribution function.

3.3. Student case. In the Student case also, we get explicit formulas for conditional quantiles. For that purpose, we need to introduce some properties of Student distributions. The lemma below and associated proof may be found in Nadarajah and Kotz (2004).

Lemma 3.2. *Let X be a d -dimensional Student distribution with ν degree of freedom and parameters $\mu \in \mathbb{R}^d$ and $\Sigma \in \mathbb{R}^{d \times d}$. Consider $X = (X_1, X_2)^T$ with $X_1 \in \mathbb{R}^{d_1}$, $X_2 \in \mathbb{R}^{d_2}$ and $d_1 + d_2 = d$. Then the density function of the conditional random variable $X_2|(X_1 = x_1)$ is given by the following equation.*

(3.5)

$$f_{X_2|X_1}(x_2|x_1) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{((\nu+d_1)\pi)^{\frac{d_2}{2}} \Gamma\left(\frac{\nu+d_1}{2}\right) |\Sigma_{2|1}|^{\frac{1}{2}}} \left[1 + \frac{1}{\nu} \frac{q_{2|1}(x_2)}{1 + \frac{1}{\nu} q_1}\right]^{-\frac{\nu+d}{2}} \times \left[\frac{\frac{\nu+d_1}{2}}{1 + \frac{1}{\nu} q_1}\right]^{\frac{d_2}{2}},$$

where $q_{2|1}(x_2)$ and q_1 are the Mahalanobis distances :

$$\begin{cases} q_{2|1}(x_2) = (x_2 - \mu_{2|1})^T \Sigma_{2|1}^{-1} (x_2 - \mu_{2|1}), \\ q_1 = (x_1 - \mu_1)^T \Sigma_{11}^{-1} (x_1 - \mu_1). \end{cases}$$

With the conditional density, we can deduce the conditional α -quantile, in the following proposition.

Proposition 3.3 (Conditional Student quantile). *Let X a $N + 1$ -dimensional Student distribution with ν degree of freedom and parameters $\mu \in \mathbb{R}^{N+1}$ and $\Sigma \in \mathbb{R}^{(N+1) \times (N+1)}$. Consider $X = (X_1, X_2)^T$ with $X_1 \in \mathbb{R}^N$ and $X_2 \in \mathbb{R}^1$. The conditional α -quantile of $X_2|(X_1 = x_1)$ has the following expression*

$$(3.6) \quad q_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sqrt{\Sigma_{2|1}} \sqrt{\frac{\nu}{\nu + N}} \sqrt{1 + \frac{1}{\nu} q_1} \Phi_{\nu+N}^{-1}(\alpha).$$

Proof. We consider the density (3.5), with $d_1 = N$, $d_2 = 1$. We have $q_{2|1}(x_2) = \frac{(x_2 - \mu_{2|1})^2}{\Sigma_{2|1}} \in \mathbb{R}$. Let q_α be the conditional quantile of $X_2|(X_1 = x_1)$. It satisfies:

$$\int_{-\infty}^{q_\alpha} f_{X_2|X_1}(x_2|x_1) dx_2 = \alpha.$$

Hence :

$$\frac{\Gamma\left(\frac{\nu+N+1}{2}\right)}{((\nu+N)\pi)^{\frac{1}{2}} \Gamma\left(\frac{\nu+N}{2}\right) |\Sigma_{2|1}|^{\frac{1}{2}}} \left[\frac{\frac{\nu+N}{\nu}}{1 + \frac{1}{\nu} q_1} \right]^{\frac{1}{2}} \times \int_{-\infty}^{q_\alpha} \left[1 + \frac{1}{\nu+N} \frac{\frac{\nu+N}{\nu} q_{2|1}(x_2)}{1 + \frac{1}{\nu} q_1} \right]^{-\frac{\nu+N+1}{2}} dx_2 = \alpha$$

Considering $z = \frac{\sqrt{\frac{\nu+N}{\nu}}(x_2 - \mu_{2|1})}{\sqrt{\Sigma_{2|1}} \sqrt{1 + \frac{1}{\nu} q_1}}$ we obtain:

$$\begin{aligned} \frac{\Gamma\left(\frac{\nu+N+1}{2}\right)}{((\nu+N)\pi)^{\frac{1}{2}} \Gamma\left(\frac{\nu+N}{2}\right)} \times \int_{-\infty}^{z_\alpha} \left[1 + \frac{z^2}{\nu+N} \right]^{-\frac{\nu+N+1}{2}} dz = \alpha, \text{ with} \\ z_\alpha = \frac{\sqrt{\frac{\nu+N}{\nu}}(q_\alpha - \mu_{2|1})}{\sqrt{\Sigma_{2|1}} \sqrt{1 + \frac{1}{\nu} q_1}}. \end{aligned}$$

Let $\Phi_{\nu+N}$ be the one dimensional Student distribution with $N + \nu$ degrees of freedom. It is obvious that

$$\Phi_{\nu+N} \left(\frac{\sqrt{\frac{\nu+N}{\nu}}(q_\alpha - \mu_{2|1})}{\sqrt{\Sigma_{2|1}} \sqrt{1 + \frac{1}{\nu} q_1}} \right) = \alpha.$$

As a consequence,

$$q_\alpha = \mu_{2|1} + \sqrt{\frac{\nu}{\nu+N}} \sqrt{\Sigma_{2|1}} \sqrt{1 + \frac{1}{\nu} q_1} \Phi_{\nu+N}^{-1}(\alpha).$$

□

We did not obtain such simple results for other elliptical distributions. In order to approximate conditional quantiles for other elliptical distributions, we propose, in what follows, two approaches. In the following section, we will apply quantile regression in the case of consistent elliptical distributions. In Section 5, we will focus on extreme quantile levels, i.e when the quantile level α is close to 0 or 1.

4. QUANTILE REGRESSION

Quantile regression, introduced by Koenker and Bassett (1978), is a classical way to estimate conditional quantiles of a distribution. If $X_1 \in \mathbb{R}^N$ and $X_2 \in \mathbb{R}$, the α -quantile of $X_2|(X_1 = x_1)$ is approximated by

$$(4.1) \quad \hat{q}_\alpha(X_2|X_1 = x_1) = \beta^{*T} x_1 + \beta_0^*,$$

where β^* and β_0^* are solutions of the following minimization problem

$$(4.2) \quad (\beta^*, \beta_0^*) = \arg \min_{\beta \in \mathbb{R}^N, \beta_0 \in \mathbb{R}} \mathbb{E} [\phi_\alpha(X_2 - \beta^T X_1 - \beta_0)].$$

and where the scoring function ϕ_α Ziegel (2015) is

$$(4.3) \quad \phi_\alpha(x) = (\alpha - 1)x\mathbf{1}_{\{x < 0\}} + \alpha x\mathbf{1}_{\{x > 0\}} = \alpha x - x\mathbf{1}_{\{x < 0\}}.$$

Obviously, given Equation (4.2), we need to assume the existence of moments $\mathbb{E}[X_1]$ and $\mathbb{E}[X_2]$. It excludes some elliptical distributions with the consistency property (such as the Cauchy distribution). Most of the time, we need simulations to find the solution, and an appropriate algorithm, for example MM algorithm (see Hunter and Lange, 2000) or stochastic gradient (see Zheng, 2011). As an example, the package `quantreg` of R performs simulations of (X_1, X_2) and solves Equation (4.2) by simplex or interior point algorithm. These simulations are computationally expensive. Furthermore, it is difficult to get the distribution of the estimator. It is a reason why we propose another approach for consistent elliptical distributions.

In the following, we denote by $\mathcal{E}_c^1(N+1)$ the set of consistent elliptical random vectors X in \mathbb{R}^{N+1} , with $X = (X_1, X_2)^T$, $X_1 \in \mathbb{R}^N$ and $X_2 \in \mathbb{R}$, having an order one moment. We shall get an explicit formula for β^* and β_0^* . Let us recall the distribution function introduced in Equation (3.1),

$$\Phi_R(x) = \mathbb{P}(RU^{(1)} \leq x).$$

From Equation (2.8), we get

$$(4.4) \quad \Phi_R(x) = \int_{-\infty}^x c_1 g_1(x^2) dx.$$

We will use this notation later in our formulas.

The next lemma is an explicit formula of a truncated moment for a bivariate elliptical distribution with the consistency property.

Lemma 4.1. *Let (X, Y) be a consistent $(R, 2)$ -elliptical vector, which admits an order one moment and with parameters $\mu = (\mu_X, \mu_Y)^T$ and $\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$.*

Then

$$(4.5) \quad \mathbb{E}[X\mathbf{1}_{\{Y > 0\}}] = \mu_X \Phi_R\left(\frac{\mu_Y}{\sigma_Y}\right) + \rho\sigma_X \int_{\frac{\mu_Y}{\sigma_Y}}^{+\infty} c_1 z g_1(z^2) dz.$$

Proof. Let $f_{(X,Y)}$ be the density of the random vector (X, Y) . Then

$$(4.6) \quad \mathbb{E}[X\mathbf{1}_{\{Y > 0\}}] = \int_{-\infty}^{+\infty} \int_0^{+\infty} x f_{X,Y}(x, y) dx dy.$$

We get:

$$\int_{-\infty}^{+\infty} x f_{X,Y}(x, y) dx = f_Y(y) \int_{-\infty}^{+\infty} x \frac{f_{X,Y}(x, y)}{f_Y(y)} dx = f_Y(y) \int_{-\infty}^{+\infty} x f_{X|Y}(x|y) dx .$$

We recognize the integral expression of the conditional mean of $X|Y$, which is given in Equation (2.6) :

$$\int_{-\infty}^{+\infty} x f_{X,Y}(x, y) dx = f_Y(y) \mathbb{E}[X|Y = y] = f_Y(y) \left(\mu_X + \rho \frac{\sigma_X}{\sigma_Y} (y - \mu_Y) \right).$$

Integrating with respect to the second variable y ,

$$(4.7) \quad \mathbb{E}[X \mathbf{1}_{\{Y > 0\}}] = \int_0^{+\infty} f_Y(y) \left(\mu_X + \rho \frac{\sigma_X}{\sigma_Y} (y - \mu_Y) \right) dy.$$

Recall that $f_Y(y) = \frac{c_1}{\sigma_Y} g_1 \left(\frac{(y - \mu_Y)^2}{\sigma_Y^2} \right)$, this leads to

$$\mathbb{E}[X \mathbf{1}_{\{Y > 0\}}] = \mu_X \int_0^{+\infty} f_Y(y) dy + \rho \frac{\sigma_X}{\sigma_Y} c_1 \int_0^{+\infty} \frac{y - \mu_Y}{\sigma_Y} g_1 \left(\frac{(y - \mu_Y)^2}{\sigma_Y^2} \right) dy.$$

Using the change of variable $z = \frac{y - \mu_Y}{\sigma_Y}$, we get

$$\mathbb{E}[X \mathbf{1}_{\{Y > 0\}}] = \mu_X \Phi_R \left(\frac{\mu_Y}{\sigma_Y} \right) + \rho \sigma_X \int_{-\frac{\mu_Y}{\sigma_Y}}^{+\infty} c_1 z g_1(z^2) dz.$$

Hence the result. \square

Previous lemma is a keystone to find β^* and β_0^* . Let ρ_j be the correlation coefficient between X_{1j} and the random variable $X_2 - \beta^{*T} X_1 - \beta_0^*$. Indeed, X_{1j} is $(R, 1)$ -elliptical with parameters μ_{1j} and σ_{1j}^2 . Furthermore, $X_2 - \beta^{*T} X_1 - \beta_0^*$ is $(R, 1)$ -elliptical too, with parameters $\mu_2 - \beta^{*T} \mu_1 - \beta_0^*$ and $(-\beta^*, 1)^T \Sigma (-\beta^{*T}, 1)$. Hence the $(R, 2)$ -elliptical vector $(X_{1j}, X_2 - \beta^{*T} X_1 - \beta_0^*)$ admits as second parameter the matrix :

$$\begin{pmatrix} \sigma_{1j}^2 & \rho_j \sigma_{1j} \sqrt{(-\beta^*, 1)^T \Sigma (-\beta^{*T}, 1)} \\ \rho_j \sigma_{1j} \sqrt{(-\beta^*, 1)^T \Sigma (-\beta^{*T}, 1)} & (-\beta^*, 1)^T \Sigma (-\beta^{*T}, 1) \end{pmatrix}$$

Proposition 4.2 (Explicit form of β^*). *Let $X = (X_1, X_2)^T \in \mathcal{E}_c^1(N + 1)$. The optimal β^* is given by :*

$$(4.8) \quad \beta^* = \Sigma_{11}^{-1} \Sigma_{12}.$$

Proof.

$$(4.9) \quad (\beta^*, \beta_0^*) = \arg \min_{\beta \in \mathbb{R}^N, \beta_0 \in \mathbb{R}} \mathbb{E} [\phi_\alpha (X_2 - \beta^T X_1 - \beta_0)]$$

Equation (4.9) is equivalent to :

$$(4.10) \quad \begin{cases} \mathbb{E} [-X_1 \phi'_\alpha (X_2 - \beta^{*T} X_1 - \beta_0^*)] = 0 \\ \mathbb{E} [-\phi'_\alpha (X_2 - \beta^{*T} X_1 - \beta_0^*)] = 0 \end{cases}$$

with the scoring function derivative :

$$(4.11) \quad \phi'_\alpha(x) = (\alpha - 1) \mathbf{1}_{\{x < 0\}} + \alpha \mathbf{1}_{\{x > 0\}} = (\alpha - 1) + \mathbf{1}_{\{x > 0\}}$$

We obtain

$$(4.12) \quad \begin{cases} (1 - \alpha)\mathbb{E}[X_1] - \mathbb{E}[X_1 \mathbf{1}_{\{X_2 - \beta^{*T} X_1 - \beta_0^* > 0\}}] & = 0 \\ (1 - \alpha) - \mathbb{E}[\mathbf{1}_{\{X_2 - \beta^{*T} X_1 - \beta_0^* > 0\}}] & = 0 \end{cases}$$

Since $X_2 - \beta^{*T} X_1 - \beta_0^*$ is $(R, 1)$ -elliptical, with parameters $\bar{\mu} = \mu_2 - \beta^{*T} \mu_1 - \beta_0^*$ and $\bar{\Sigma} = (-\beta^*, 1)^T \Sigma (-\beta^*, 1)$. Then $\mathbb{E}[\mathbf{1}_{\{X_2 - \beta^{*T} X_1 - \beta_0^* > 0\}}] = \mathbb{P}(X_2 - \beta^{*T} X_1 - \beta_0^* > 0)$, i.e, if we denote $\bar{\sigma} = \bar{\Sigma}^{\frac{1}{2}}$:

$$(4.13) \quad \mathbb{E}[\mathbf{1}_{\{X_2 - \beta^{*T} X_1 - \beta_0^* > 0\}}] = 1 - \Phi_R\left(\frac{-\bar{\mu}}{\bar{\sigma}}\right) = \Phi_R\left(\frac{\bar{\mu}}{\bar{\sigma}}\right)$$

Next, thanks to Equation (4.5), we have, for $j \in \{1, \dots, N\}$:

$$(4.14) \quad \mathbb{E}[X_{1j} \mathbf{1}_{\{X_2 - \beta^{*T} X_1 - \beta_0^* > 0\}}] = \mu_{1j} \Phi_R\left(\frac{\bar{\mu}}{\bar{\sigma}}\right) + \rho_j \sigma_{1j} \int_{\frac{\bar{\mu}}{\bar{\sigma}}}^{+\infty} c_1 z g_1(z^2) dz$$

Finally, Equation (4.12) may be written as :

$$(4.15) \quad \begin{cases} \mu_{1j} \left[(1 - \alpha) - \Phi_R\left(\frac{\bar{\mu}}{\bar{\sigma}}\right) \right] - \rho_j \sigma_{1j} \int_{\frac{\bar{\mu}}{\bar{\sigma}}}^{+\infty} c_1 z g_1(z^2) dz & = 0 \\ (1 - \alpha) - \Phi_R\left(\frac{\bar{\mu}}{\bar{\sigma}}\right) & = 0 \end{cases}, \forall j \in \{1, \dots, N\}$$

Therefore :

$$(4.16) \quad \rho_j \sigma_{1j} \int_{\frac{\bar{\mu}}{\bar{\sigma}}}^{+\infty} c_1 z g_1(z^2) dz = 0, \forall j \in \{1, \dots, N\}$$

Then $\rho_j = 0, \forall j \in \{1, \dots, N\}$. Now, it is not difficult to see that $\beta^* = \Sigma_{11}^{-1} \Sigma_{12}$. Indeed, the second parameter of the $(R, N+1)$ -elliptical vector $(X_1, X_2 - \beta^{*T} X_1 - \beta_0^*)$ is :

$$\begin{pmatrix} \Sigma_{11} & 0_{\mathbb{R}^N} \\ 0_{\mathbb{R}^N}^T & (-\beta^*, 1)^T \Sigma (-\beta^{*T}, 1) \end{pmatrix}$$

Hence its determinant is $(-\beta^*, 1)^T \Sigma (-\beta^{*T}, 1) |\Sigma_{11}|$. Using the conditional moments of $X_2 - \beta^{*T} X_1 - \beta_0^* | X_1$ given in Proposition 2.6, we get the following equation :

$$\Sigma_{2|1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} = (-\beta^*, 1)^T \Sigma (-\beta^{*T}, 1)$$

Since $(-\beta^*, 1)^T \Sigma (-\beta^{*T}, 1) = \Sigma_{22} + \beta^{*T} \Sigma_{11} \beta^* - 2\beta^{*T} \Sigma_{12}$, the previous equation is equivalent to :

$$2\beta^{*T} \Sigma_{12} - \beta^{*T} \Sigma_{11} \beta^* - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} = 0$$

The convex function $2\beta^{*T} \Sigma_{12} - \beta^{*T} \Sigma_{11} \beta^* - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$ reaches its minimum 0 at $\beta^* = \Sigma_{11}^{-1} \Sigma_{12}$. \square

This result shows that a quantile regression through a linear model is meaningless. Indeed, the vector $\Sigma_{11}^{-1} \Sigma_{12}$ does not depend on α and corresponds to the vector of weights in the Kriging mean Cressie (1988). Using a linear estimator, the result of the quantile regression, whatever the quantile level $\alpha \in [0, 1]$, would be the conditional mean $\mu_{2|1}$ given in Equation (2.6). This is why the addition of the affine term β_0^* is required. Let us now calculate this term β_0^* . As mentioned in the last proof that, obviously, $X_2 - \beta^{*T} X_1 - \beta_0^*$ is elliptical with the same radius as

(X_1, X_2) . With Equation (4.8), we are now able to prove that its second parameter is $\Sigma_{2|1}$, given in Equation (2.6).

Lemma 4.3. *Let $X = (X_1, X_2)^T \in \mathcal{E}_c^1(N+1)$. $X_2 - \beta^{*T} X_1$ is $(R, 1)$ -elliptical with parameters $\mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1$ and $\Sigma_{2|1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.*

Proof. The random variable $X_2 - \beta^{*T} X_1$, as an affine transformation of a consistent (R, d) -elliptical random vector, is obviously $(R, 1)$ -elliptical. Furthermore, with the expression of β^* given in Proposition 4.2, $\mathbb{E}[X_2 - \beta^{*T} X_1] = \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1$. Concerning the second parameter, it is equal to $(-\beta^*, 1)^T \Sigma(-\beta^*, 1)$. Thus,

$$(-\beta^*, 1)^T \Sigma(-\beta^*, 1) = \sigma_2^2 - 2 \sum_{i=1}^n \beta_i^* \rho_{ix} \sigma_{1i} \sigma_2 + \sum_{j=1}^n \beta_j^* \sum_{i=1}^n \beta_i^* \rho_{ij} \sigma_{1i} \sigma_{1j}.$$

It may be rewritten in matrix form

$$(-\beta^*, 1)^T \Sigma(-\beta^*, 1) = \sigma_2^2 - 2\beta^{*T} \Sigma_{12} + \beta^{*T} \Sigma_{11} \beta^*.$$

Using that $\beta^* = \Sigma_{11}^{-1} \Sigma_{12}$, we prove the following equality.

$$(-\beta^*, 1)^T \Sigma(-\beta^*, 1) = \sigma_2^2 - 2\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{11}\Sigma_{11}^{-1}\Sigma_{12} = \sigma_2^2 - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12},$$

which concludes the proof. \square

We have seen that the conditional second parameter of our $(R, 1)$ -elliptical random variable is exactly the conditional second parameter $\Sigma_{2|1}$ of Equation (2.6). We are led to our main result of this section on quantile regression for elliptical distributions.

Theorem 4.4 (Quantile regression approximation). *Let $X = (X_1, X_2)^T \in \mathcal{E}_c^1(N+1)$. The quantile regression vector (β^*, β_0^*) of $X_2|X_1 = x_1$, satisfying Equation (4.2), is given by*

$$(4.17) \quad \begin{cases} \beta^* = & \Sigma_{11}^{-1} \Sigma_{12} \\ \beta_0^* = & \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1 + \sigma_{2|1}\Phi_R^{-1}(\alpha) \end{cases}$$

The quantile regression approximation with level $\alpha \in [0, 1]$ is given by:

$$(4.18) \quad \hat{q}_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\Phi_R^{-1}(\alpha)$$

Furthermore, its distribution is

$$(4.19) \quad \hat{q}_\alpha(X_2|X_1) \sim \mathcal{E}_1(\mu_2 + \sigma_{2|1}\Phi_R^{-1}(\alpha), \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}, g_1)$$

Proof. β^* is given in Equation (4.8). Concerning β_0^* , it satisfies:

$$\beta_0^* = \arg \min_{\beta_0 \in \mathbb{R}} \mathbb{E}[\phi_\alpha(X_2 - \beta^{*T} X_1 - \beta_0)].$$

Finally, β_0^* is the α -quantile of the random variable $X_2 - \beta^{*T} X_1$. We have seen in Lemma 4.3, that $X_2 - \beta^{*T} X_1$ is $(R, 1)$ -elliptical with parameters $\mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1$ and $\Sigma_{2|1}$. Then, using the quantile formula of an elliptical distribution, and denoting $\sigma_{2|1} = \sqrt{\Sigma_{2|1}}$, we get:

$$\beta_0^* = \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1 + \sigma_{2|1}\Phi_R^{-1}(\alpha).$$

With our optimal parameters β^* and β_0^* , we can now express our quantile regression estimator of X_2 given $X_1 = x_1$, $\hat{q}_\alpha(X_2|X_1 = x_1)$, using Equation (4.1):

$$\hat{q}_\alpha(X_2|X_1 = x_1) = \beta^{*T} x_1 + \beta_0^* = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1) + \sigma_{2|1}\Phi_R^{-1}(\alpha).$$

We recognize, on the left, the expression of $\mu_{2|1}$ given in Equation (2.6).

Concerning its distribution, we know that

$$\hat{q}_\alpha(X_2|X_1) = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) + \sigma_{2|1}\Phi_R^{-1}(\alpha),$$

with $X_1 \sim \mathcal{E}_N(\mu_1, \Sigma_{11}, g_N)$. As an affine combination of X_1 , using Proposition 2.4, $\hat{q}_\alpha(X_2|X_1) \sim \mathcal{E}_1(\mu_2 + \sigma_{2|1}\Phi_R^{-1}(\alpha), (\Sigma_{21}\Sigma_{11}^{-1})\Sigma_{11}(\Sigma_{11}^{-1}\Sigma_{12}), g_1)$. Hence the result. \square

As an illustration of Theorem 4.4, we now propose several examples, such as Gaussian and Student cases, where the theoretical conditional quantiles are known (see Section 3). We also consider the Unimodal Gaussian Mixture and Laplace cases, where we do not have any explicit formula for theoretical quantiles. Numerical applications are proposed in the last section.

Example 1 : Gaussian case. We recall the theoretical formula of the conditional quantile of a Gaussian distribution :

$$(4.20) \quad q_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\Phi^{-1}(\alpha).$$

In that case, the quantile regression estimator $\hat{q}_\alpha(X_2|X_1 = x_1)$ is exactly the same :

$$(4.21) \quad \hat{q}_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\Phi^{-1}(\alpha).$$

For Gaussian processes, the quantile regression and the theoretical conditional quantile coincide.

Example 2 : Student case. The Gaussian case is the most simple one. Unfortunately, for Student distributions, quantile regression and conditional quantile do not coincide. Indeed, the conditional quantile of a multivariate Student distribution is given by Equation (3.6):

$$q_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\sqrt{\frac{\nu}{\nu + N}}\sqrt{1 + \frac{1}{\nu}(x_1 - \mu_1)^T\Sigma_{11}^{-1}(x_1 - \mu_1)}\Phi_{\nu+N}^{-1}(\alpha).$$

The quantile regression is

$$\hat{q}_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\Phi_\nu^{-1}(\alpha).$$

The error $q_\alpha(X_2|X_1 = x_1) - \hat{q}_\alpha(X_2|X_1 = x_1)$ that is done when using the quantile regression depends on the Mahalanobis distance $(x_1 - \mu_1)^T\Sigma_{11}^{-1}(x_1 - \mu_1)$ which may be large.

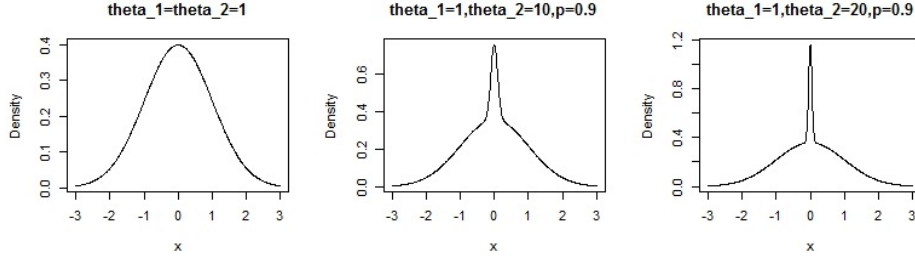
Example 3 : Unimodal Gaussian Mixture case. We consider here a mixture of 2 centered Gaussian distributions, i.e a radius R such that

$$R = \frac{\chi_d}{p\theta_1 + (1-p)\theta_2}.$$

Immediately, we get the following multivariate p.d.f :

$$(4.22) \quad f_X(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \left[p\theta_1^d e^{-\frac{\theta_1^2}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} + (1-p)\theta_2^d e^{-\frac{\theta_2^2}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} \right].$$

As an illustration, we propose some examples of univariate densities (4.22), with different values of p , θ_1 and θ_2 : For these models, we do not have any explicit formula for $q_\alpha(X_2|X_1 = x_1)$. On the other hand, using the conditional density expressed in Equation (2.8), we may get a numerical approximation of this value, and compare it with the Quantile Regression Estimator.


 FIGURE 1. Univariate mixture distributions, with $\mu = 0$ and $\Sigma = 1$

Example 4 : Laplace case. The multivariate Laplace distribution, introduced in Kozubowski et al. (2013) and Eltoft et al. (2006), has the following density:

$$(4.23) \quad f(x) = \frac{1}{|\Sigma|^{\frac{1}{2}}(2\pi)^{\frac{d}{2}}} \frac{2 K_{\frac{d}{2}-1} \left(\sqrt{\frac{2}{\lambda}} (x - \mu)^T \Sigma^{-1} (x - \mu) \right)}{\left(\sqrt{\frac{\lambda}{2}} (x - \mu)^T \Sigma^{-1} (x - \mu) \right)^{\frac{d}{2}-1}}, \forall x \in \mathbb{R}^d \setminus \{\mu\}.$$

where $K_m(x)$ denotes the modified Bessel function of the second kind and order m , evaluated at x . In this case, we can easily give an expression for the quantile regression $\hat{q}_\alpha(X_2|X_1 = x_1)$. This is the aim of the following proposition.

Proposition 4.5. *In the Laplace case, the quantile regression estimator is given by:*

$$(4.24) \quad \begin{cases} \hat{q}_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sqrt{\frac{\lambda}{2}} \sigma_{2|1} \ln(2\alpha), & \text{for } \alpha \leq \frac{1}{2} \\ \hat{q}_\alpha(X_2|X_1 = x_1) = \mu_{2|1} - \sqrt{\frac{\lambda}{2}} \sigma_{2|1} \ln(2(1-\alpha)), & \text{for } \alpha > \frac{1}{2}. \end{cases}$$

Proof. From (4.18), we just have to calculate $\Phi_R^{-1}(\alpha)$, where Φ_R is the distribution function of the univariate reduced and centered Laplace distribution, with density function (4.23), and with $d = 1$. Then, $\Phi_R^{-1}(\alpha)$ equals $x \in \mathbb{R}$ so that

$$\int_{-\infty}^x \frac{1}{\sqrt{2\lambda}} e^{-\sqrt{\frac{2}{\lambda}}|t|} dt = \alpha.$$

Since the univariate reduced centered Laplace distribution is symmetric, we have the equivalence $\alpha \leq \frac{1}{2} \Leftrightarrow x \leq 0$. Then, we will separate the cases $\alpha \leq \frac{1}{2}$ and $\alpha > \frac{1}{2}$. If $\alpha \leq \frac{1}{2}$, we have to solve

$$\int_{-\infty}^x \frac{1}{\sqrt{2\lambda}} e^{\sqrt{\frac{2}{\lambda}}t} dt = \alpha \Rightarrow x = \sqrt{\frac{\lambda}{2}} \ln(2\alpha).$$

If $\alpha > \frac{1}{2}$, we write :

$$\int_{-\infty}^0 \frac{1}{\sqrt{2\lambda}} e^{\sqrt{\frac{2}{\lambda}}t} dt + \int_0^x \frac{1}{\sqrt{2\lambda}} e^{-\sqrt{\frac{2}{\lambda}}t} dt = \alpha.$$

Since $\int_{-\infty}^0 \frac{1}{\sqrt{2\lambda}} e^{\sqrt{\frac{x}{\lambda}}t} dt = \frac{1}{2}$ and $\int_0^x \frac{1}{\sqrt{2\lambda}} e^{-\sqrt{\frac{x}{\lambda}}t} dt = \frac{1}{2} - \frac{1}{2} e^{-\sqrt{\frac{x}{\lambda}}x}$, we quickly get $x = -\ln(2(1-\alpha))$. Finally,

$$\begin{cases} \hat{q}_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sqrt{\frac{\lambda}{2}}\sigma_{2|1} \ln(2\alpha) & , \alpha \leq \frac{1}{2} \\ \hat{q}_\alpha(X_2|X_1 = x_1) = \mu_{2|1} - \sqrt{\frac{\lambda}{2}}\sigma_{2|1} \ln(2(1-\alpha)) & , \alpha > \frac{1}{2} \end{cases}$$

which is the announced result. \square

5. EXTREMAL QUANTILES

In this part, in order to simplify the notations, we denote by q_1 the Mahalanobis distance $(x_1 - \mu_1)^T \Sigma_{11}^{-1} (x_1 - \mu_1)$. We have previously defined an estimator \hat{q}_α based on quantile regression. We have seen that the best affine estimator of conditional quantiles of an R -elliptical distribution with the consistency property is given by:

$$(5.1) \quad \hat{q}_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1} \Phi_R^{-1}(\alpha)$$

where $\Phi_R(x) = \int_{-\infty}^x c_1 g_1(t^2) dt$. The conditional quantile is:

$$(5.2) \quad q_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1} \Phi_{R^*}^{-1}(\alpha)$$

with the conditional radius R^* defined in Equation (2.7). In this section, the aim is to establish a relation between Φ_R^{-1} and $\Phi_{R^*}^{-1}$ for extremal values of α , i.e for $\alpha \rightarrow 0$ or $\alpha \rightarrow 1$. To this aim, we need the following hypothesis.

Assumption 1. *Their exist $0 < \ell < +\infty$ and $\gamma \in \mathbb{R}$ such as :*

$$(5.3) \quad \lim_{x \rightarrow +\infty} \frac{\bar{\Phi}_{R^*}(x)}{\bar{\Phi}_R(x^\gamma)} = \ell$$

where $\bar{\Phi} = 1 - \Phi$ is the survival function associated to Φ .

We also recall that the distribution functions are $\Phi_R(x) = \int_{-\infty}^x c_1 g_1(t^2) dt$ and

$$\Phi_{R^*}(x) = \int_{-\infty}^x c_1^* g_{N+1}(q_1 + t^2) dt, \text{ with } c_1^* \text{ given by Equation (2.8).}$$

$$(5.4) \quad c_1^* = \frac{c_{N+1}}{c_N g_N(q_1)},$$

where c_{N+1} and c_N are the normalization coefficients of the elliptical distributions with radius R in dimensions $N+1$ and N . Then, coefficients γ and ℓ may also be obtained as

$$(5.5) \quad \lim_{x \rightarrow +\infty} \frac{c_1^* g_{N+1}(q_1 + x^2)}{c_1 \gamma x^{\gamma-1} g_1(x^{2\gamma})} = \ell$$

Table 3 gives values of γ and ℓ for the examples introduced in Section 2. The calculations will be detailed later on. Unfortunately, the Laplace distribution does not satisfy Assumption 1, this is why we will not consider this example in what follows.

Thanks to Equation (5.3), we have, under Assumption 1, the following equivalence when x goes to ∞ :

Distribution	γ	ℓ
Gaussian	1	1
Student, $\nu > 0$	$\frac{N+\nu}{\nu}$	$\frac{\Gamma(\frac{\nu+N+1}{2})\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+N}{2})\Gamma(\frac{\nu+1}{2})} \left(1 + \frac{q_1}{\nu}\right)^{\frac{N+\nu}{2}} \frac{\nu^{\frac{N}{2}+1}}{\nu+N}$
Unimodal Gaussian Mixture	1	$\frac{\min(\theta_1, \dots, \theta_n)^N \exp\left(-\frac{\min(\theta_1, \dots, \theta_n)^2}{2} q_1\right)}{\sum_{k=1}^n \pi_k \theta_k^N \exp\left(-\frac{\theta_k^2}{2} q_1\right)}$
Uniform Gaussian Mixture	$N+1$	$\frac{\Gamma(\frac{N+2}{2}) q_1^{\frac{N+1}{2}} \sqrt{2}}{\Gamma(\frac{N+1}{2}) (N+1) \chi_{N+1}^2(q_1)}$

TABLE 3. Coefficients γ and ℓ for classical distributions, where $q_1 = (x_1 - \mu_1)^T \Sigma_{11}^{-1} (x_1 - \mu_1)$ is a function of x_1 . (see Proposition 2.6).

$$(5.6) \quad \bar{\Phi}_{R^*}(x) \underset{x \rightarrow \infty}{\sim} \ell \bar{\Phi}_R(x^\gamma)$$

Our aim is now to get an equivalence relationship between the quantile function $\bar{\Phi}_{R^*}^{-1}$ and $\bar{\Phi}_R^{-1}$. For that, we refer to the paper of Djurčić and Torgasev, which gives some conditions to deduce the equivalence of inverse functions if these functions are equivalent Djurčić and Torgasev (2001).

Definition 5.1. *A function f is a φ -function if $f : [0, +\infty[\rightarrow [0, +\infty[$, $f(0) = 0$, f is continuous, non decreasing on $[0, +\infty[$, and $f \rightarrow +\infty$ when $x \rightarrow +\infty$.*

Clearly, our two equivalent functions $\bar{\Phi}_{R^*}(x)$ and $\ell \bar{\Phi}_R(x^\gamma)$ are not φ -functions for several reasons: $\bar{\Phi}_{R^*}(0) = \frac{1}{2}$, $\ell \bar{\Phi}_R(0) = \frac{1}{2} \ell$, $\lim_{x \rightarrow +\infty} \bar{\Phi}_{R^*}(x) = 1$ and $\lim_{x \rightarrow +\infty} \ell \bar{\Phi}_R(x^\gamma) = \ell < +\infty$. Then, we have to transform these functions in order to find an equivalence. But before, we need more definitions and properties. The following is the definition of a general class of functions K_c , which contains in particular Regularly Varying functions..

Definition 5.2. *K_c is the set of all φ -functions f with the property :*

$$(5.7) \quad \lim_{\lambda \rightarrow 1} \frac{f(\lambda x)}{f(x)} = 1$$

In fact, the K_c is the set of φ -functions for which the limit $\lim_{\lambda \rightarrow 1} \frac{f(\lambda x)}{f(x)}$ exists,

because if it exists, then automatically we can swap the limits, and $\lim_{\lambda \rightarrow 1} \frac{f(\lambda x)}{f(x)} =$

$\lim_{x \rightarrow +\infty} \left(\lim_{\lambda \rightarrow 1} \frac{f(\lambda x)}{f(x)} \right) = 1$. We can also notice that the condition in Equation (5.7) is more general than the regular variation functions in the sense of Karamata.

These two last definitions are very important for the following lemma, which establishes the relation between the equivalence of two functions and equivalence of their inverses.

Lemma 5.1. *[Djurčić and Torgasev (2001)] Suppose that f and g are two strictly increasing φ -functions, and that at least one of the functions f^{-1} , g^{-1} belongs to the class K_c , and $f(x) \underset{x \rightarrow \infty}{\sim} g(x)$. Then $f^{-1}(x) \underset{x \rightarrow \infty}{\sim} g^{-1}(x)$*

Now, as we said previously, the first step is to transform our functions $\bar{\Phi}_{R^*}(x)$ and $\ell \bar{\Phi}_R(x^\gamma)$ into φ -functions.

Definition 5.3. Let Ψ and Ψ_* be

$$(5.8) \quad \begin{cases} \Psi(x) &= \frac{1}{\ell \bar{\Phi}_R(x^\gamma)} - \frac{2}{\ell} \\ \Psi_*(x) &= \frac{1}{\bar{\Phi}_{R^*}(x)} - 2 \end{cases}$$

Lemma 5.2. If $\gamma > 0$, then Ψ and Ψ_* are φ -functions.

Proof. Since $\bar{\Phi}_R$ and $\bar{\Phi}_{R^*}$ are clearly strictly decreasing, Ψ and Ψ_* are obviously strictly increasing. Furthermore, $\bar{\Phi}_R(0) = \bar{\Phi}_{R^*}(0) = \frac{1}{2}$, thus

$$\begin{cases} \Psi(0) &= 0 \\ \Psi_*(0) &= 0 \end{cases}.$$

Finally, since $\lim_{x \rightarrow +\infty} \bar{\Phi}_R(x^\gamma) = \lim_{x \rightarrow +\infty} \bar{\Phi}_{R^*}(x) = 0$, functions Ψ and Ψ_* are such that $\lim_{x \rightarrow +\infty} \Psi(x) = \lim_{x \rightarrow +\infty} \Psi_*(x) = +\infty$. \square

Now, given that $\bar{\Phi}_{R^*}(x) \underset{x \rightarrow \infty}{\sim} \ell \bar{\Phi}_R(x^\gamma)$, we have to check whether $\Psi(x) \underset{x \rightarrow \infty}{\sim} \Psi_*(x)$.

Lemma 5.3. Under Assumption 1, we have

$$(5.9) \quad \Psi_*(x) \underset{x \rightarrow \infty}{\sim} \Psi(x)$$

Proof. Let us focus on the limit

$$\lim_{x \rightarrow +\infty} \frac{\Psi_*(x)}{\Psi(x)} = \lim_{x \rightarrow +\infty} \frac{\frac{1}{\bar{\Phi}_{R^*}(x)} - 2}{\frac{1}{\ell \bar{\Phi}_R(x^\gamma)} - \frac{2}{\ell}}.$$

This limit is equal to

$$\lim_{x \rightarrow +\infty} \frac{\ell \bar{\Phi}_R(x^\gamma)}{\bar{\Phi}_{R^*}(x)} \cdot \frac{1 - 2\bar{\Phi}_{R^*}(x)}{1 - 2\bar{\Phi}_R(x^\gamma)}.$$

Thanks to (5.3), $\lim_{x \rightarrow +\infty} \frac{\ell \bar{\Phi}_R(x^\gamma)}{\bar{\Phi}_{R^*}(x)} = 1$. Furthermore, $\lim_{x \rightarrow +\infty} \bar{\Phi}_{R^*}(x) = \lim_{x \rightarrow +\infty} \bar{\Phi}_R(x^\gamma) = 0$. Then :

$$\lim_{x \rightarrow +\infty} \frac{\Psi_*(x)}{\Psi(x)} = \lim_{x \rightarrow +\infty} \frac{\ell \bar{\Phi}_R(x^\gamma)}{\bar{\Phi}_{R^*}(x)} \cdot \frac{1 - 2\bar{\Phi}_{R^*}(x)}{1 - 2\bar{\Phi}_R(x^\gamma)} = 1.$$

In other words, $\Psi_*(x) \underset{x \rightarrow \infty}{\sim} \Psi(x)$. \square

In order to use Lemma 5.1, we shall need an additional assumption on Ψ_*^{-1} .

Assumption 2. Ψ_*^{-1} belongs to the class K_c .

The following proposition is the key step to approximate conditional quantiles.

Proposition 5.4. Under Assumptions 1 and 2, we have

$$(5.10) \quad \Phi_{R^*}^{-1}(\alpha) \underset{\alpha \rightarrow 1}{\sim} \left[\Phi_R^{-1} \left(1 - \frac{1}{\frac{\ell}{1-\alpha} + 2(1-\ell)} \right) \right]^{\frac{1}{\gamma}}$$

Proof. Since $\Psi_*^{-1}(x)$ belongs to the K_c class, and $\Psi_*(x)$ and $\Psi(x)$ are two strictly increasing φ -functions, and $\Psi_*(x) \underset{x \rightarrow +\infty}{\sim} \Psi(x)$, then Lemma 5.1 gives

$$\Psi_*^{-1}(x) \underset{x \rightarrow +\infty}{\sim} \Psi^{-1}(x)$$

In other words

$$(5.11) \quad \Phi_{R^*}^{-1} \left(1 - \frac{1}{x+2} \right) \underset{x \rightarrow +\infty}{\sim} \left[\Phi_R^{-1} \left(1 - \frac{1}{\ell x + 2} \right) \right]^{\frac{1}{\gamma}}.$$

This may also be rewritten as

$$\Phi_{R^*}^{-1}(\alpha) \underset{\alpha \rightarrow 1}{\sim} \left[\Phi_R^{-1} \left(1 - \frac{1}{\frac{\ell}{1-\alpha} + 2(1-\ell)} \right) \right]^{\frac{1}{\gamma}}$$

with $\alpha = 1 - \frac{1}{x+2}$, hence the result. \square

This result leads us to the following conditional extremal quantile approximation.

Definition 5.4 (Extreme conditional quantiles approximations). *Define*

$$(5.12) \quad \begin{cases} \hat{q}_{\alpha\uparrow}(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1} \left[\Phi_R^{-1} \left(1 - \frac{1}{\frac{\ell}{1-\alpha} + 2(1-\ell)} \right) \right]^{\frac{1}{\gamma}} \\ \hat{q}_{\alpha\downarrow}(X_2|X_1 = x_1) = \mu_{2|1} - \sigma_{2|1} \left[\Phi_R^{-1} \left(1 - \frac{1}{\frac{\ell}{\alpha} + 2(1-\ell)} \right) \right]^{\frac{1}{\gamma}} \end{cases}$$

\hat{q} will be called ECQA.

The following proposition shows that ECQA $\hat{q}_{\alpha\uparrow}(X_2|X_1 = x_1)$ and $\hat{q}_{\alpha\downarrow}(X_2|X_1 = x_1)$ are asymptotically equivalent to the theoretical quantile $q_\alpha(X_2|X_1 = x_1)$, respectively for $\alpha \rightarrow 1$ and $\alpha \rightarrow 0$.

Theorem 5.5 (Equivalences of extreme conditional quantiles approximations). *Let $(X_1, X_2)^T \in \mathcal{E}_c^1(N+1)$. Under Assumptions 1 and 2,*

$$(5.13) \quad \begin{cases} q_\alpha(X_2|X_1 = x_1) \underset{\alpha \rightarrow 1}{\sim} \hat{q}_{\alpha\uparrow}(X_2|X_1 = x_1) \\ q_\alpha(X_2|X_1 = x_1) \underset{\alpha \rightarrow 0}{\sim} \hat{q}_{\alpha\downarrow}(X_2|X_1 = x_1) \end{cases}$$

Proof. Recall Equation (3.2):

$$q_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1} \Phi_{R^*}^{-1}(\alpha)$$

From (5.10) and (5.12), we immediatly prove the first half of the result:

$$q_\alpha(X_2|X_1 = x_1) \underset{\alpha \rightarrow 1}{\sim} \mu_{2|1} + \sigma_{2|1} \left[\Phi_R^{-1} \left(1 - \frac{1}{\frac{\ell}{1-\alpha} + 2(1-\ell)} \right) \right]^{\frac{1}{\gamma}} = \hat{q}_{\alpha\uparrow}(X_2|X_1 = x_1)$$

Now, let us prove the equivalence when $\alpha \rightarrow 0$. By the symmetry properties of elliptical distributions, we have $\Phi_{R^*}^{-1}(\alpha) = -\Phi_{R^*}^{-1}(1-\alpha), \forall \alpha \in [0, 1]$. Then,

$$q_\alpha(X_2|X_1 = x_1) = \mu_{2|1} - \sigma_{2|1} \Phi_{R^*}^{-1}(1-\alpha).$$

Using the equivalence (5.10) and Equation (5.12), we finally get:

$$q_\alpha(X_2|X_1 = x_1) \underset{\alpha \rightarrow 0}{\sim} \mu_{2|1} - \sigma_{2|1} \left[\Phi_R^{-1} \left(1 - \frac{1}{\frac{\ell}{\alpha} + 2(1-\ell)} \right) \right]^{\frac{1}{\gamma}} = \hat{q}_{\alpha\downarrow}(X_2|X_1 = x_1)$$

\square

We finish this section by checking Assumption 2 is satisfied for examples that are given in Table 3.

Example 1: Gaussian case.

Lemma 5.6. *Gaussian distribution satisfies Assumptions 1 and 2, with $\ell = 1$ and $\gamma = 1$.*

Proof. Firstly, let us calculate ℓ and γ . Take $\gamma = 1$, it remains the limit of Assumption 1 :

$$\lim_{x \rightarrow +\infty} \frac{c_1^* g_{N+1}(q_1 + x^2)}{c_1 g_1(x^2)}$$

The values c_1 , $g_{N+1}(q_1 + x^2)$ and $g_1(x^2)$ may be easily deduced from Table 1. Let $c_1^* = \frac{c_{N+1}}{c_N g_N(q_1)}$,

$$\ell = \lim_{x \rightarrow +\infty} \frac{\exp\left(-\frac{q_1 + x^2}{2}\right)}{\exp\left(-\frac{q_1}{2}\right) \exp\left(-\frac{x^2}{2}\right)} = 1.$$

It remains to check whether the limit

$$\epsilon_\ell = \lim_{\substack{x \rightarrow +\infty \\ \lambda \rightarrow 1}} \frac{\Psi^{-1}(\lambda x)}{\Psi^{-1}(x)} = \lim_{\substack{x \rightarrow +\infty \\ \lambda \rightarrow 1}} \frac{\Phi^{-1}\left(1 - \frac{1}{\lambda x + 2}\right)}{\Phi^{-1}\left(1 - \frac{1}{x + 2}\right)}$$

exists, or equivalently, whether the limit

$$\epsilon_\ell = \lim_{(\delta, y) \rightarrow (0, 0)} \frac{\Phi^{-1}\left(1 - \frac{y}{2y + \delta + 1}\right)}{\Phi^{-1}\left(1 - \frac{y}{2y + 1}\right)}$$

exists.

Now, we move to polar coordinates, i.e take $\delta = r \cos(\theta)$ and $y = r \sin(\theta)$, and calculate the limit when $r \rightarrow 0$. If this limit do not depend on θ , then the limit when $(\delta, y) \rightarrow (0, 0)$ exists. Consider

$$\lim_{r \rightarrow 0} \frac{\Phi^{-1}\left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1}\right)}{\Phi^{-1}\left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1}\right)} = f(\theta).$$

Clearly, the numerator and denominator both tend to $+\infty$, then we use the L'Hôpital's rule to get

$$f(\theta) = \lim_{r \rightarrow 0} \frac{\left(-\frac{\sin(\theta)(2r \sin(\theta) + r \cos(\theta) + 1) - r \sin(\theta)(2 \sin(\theta) + \cos(\theta))}{(2r \sin(\theta) + r \cos(\theta) + 1)^2}\right) \frac{1}{\phi\left(\Phi^{-1}\left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1}\right)\right)}}{\left(-\frac{\sin(\theta)(2r \sin(\theta) + 1) - r \sin(\theta) 2 \sin(\theta)}{(2r \sin(\theta) + 1)^2}\right) \frac{1}{\phi\left(\Phi^{-1}\left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1}\right)\right)}}$$

On the left, the bulky ratio clearly tends to 1 when $r \rightarrow 0$. On the right, we recall that ϕ is the density function of the standard normal distribution, i.e $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Then :

$$\epsilon_\ell = \lim_{r \rightarrow 0} \frac{\phi\left(\Phi^{-1}\left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1}\right)\right)}{\phi\left(\Phi^{-1}\left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1}\right)\right)}$$

Now, the numerator and denominator both tend to 0. Once again, we use the L'Hôpital's rule and then,

$$f(\theta) = \lim_{r \rightarrow 0} \frac{\frac{\left(-\frac{\sin(\theta)(2r \sin(\theta) + 1) - r \sin(\theta)2 \sin(\theta)}{(2r \sin(\theta) + 1)^2} \right)}{\phi \left(\Phi^{-1} \left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1} \right) \right)}{\frac{\left(-\frac{\sin(\theta)(2r \sin(\theta) + r \cos(\theta) + 1) - r \sin(\theta)(2 \sin(\theta) + \cos(\theta))}{(2r \sin(\theta) + r \cos(\theta) + 1)^2} \right)}{\phi \left(\Phi^{-1} \left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1} \right) \right)}}{\phi' \left(\Phi^{-1} \left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1} \right) \right)}$$

We have seen that the ratio on the left tends to 1. Furthermore, we have the relationship $\phi'(x) = -x\phi(x)$. Finally,

$$f(\theta) = \lim_{r \rightarrow 0} \frac{\Phi^{-1} \left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1} \right)}{\Phi^{-1} \left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1} \right)} = \frac{1}{f(\theta)},$$

thus, $f(\theta) = 1$ hence the result. \square

Since $\Psi^{-1}(x)$ belongs to the K_c class and $\ell = \gamma = 1$, we have the following formulas for the extremal estimators $\hat{q}_{\alpha\uparrow}$ and $\hat{q}_{\alpha\downarrow}$:

$$\begin{cases} \hat{q}_{\alpha\uparrow}(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\Phi_R^{-1}(\alpha) \\ \hat{q}_{\alpha\downarrow}(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\Phi_R^{-1}(\alpha) \end{cases}$$

Remark that $\hat{q}_{\alpha\uparrow}$ and $\hat{q}_{\alpha\downarrow}$ equal to the quantile regression estimator \hat{q}_α , which equals the theoretical quantile q_α .

Example 2 : Student case. We have seen in part 2 that the quantile regression estimator \hat{q}_α was not good in the Student case. Then, for extremal levels of quantile, we can guess that the estimator \hat{q}_α will be worse and worse if α is close to 0 or 1. It is a reason why, in this case, our extremal estimators $\hat{q}_{\alpha\uparrow}$ and $\hat{q}_{\alpha\downarrow}$ may potentially be especially interesting, and improve the estimation. Firstly, we prove that Assumptions 1 and 2 are filled.

Lemma 5.7. *Student distributions satisfies Assumptions 1 and 2, with $\gamma = \frac{N+\nu}{\nu}$ and ℓ given in Table 3.*

Proof. Concerning Assumption 1, we have to calculate the limit :

$$\lim_{x \rightarrow +\infty} \frac{c_1^* g_{N+1}(q_1 + x^2)}{c_1 \gamma x^{\gamma-1} g_1(x^{2\gamma})}$$

From Table 1 and (5.4), we deduce $c_1 = \frac{\Gamma(\frac{1+\nu}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}$, $c_1^* = \frac{\Gamma(\frac{N+1+\nu}{2})}{\sqrt{\nu\pi}\Gamma(\frac{N+\nu}{2})} \left(1 + \frac{q_1}{\nu}\right)^{\frac{N+\nu}{2}}$, $g_{N+1}(q_1 + x^2) = \left(1 + \frac{q_1 + x^2}{\nu}\right)^{-\frac{N+1+\nu}{2}}$ and $g_1(x^{2\gamma}) = \left(1 + \frac{x^{2\gamma}}{\nu}\right)^{-\frac{1+\nu}{2}}$. Then the previous limit is the following :

$$\lim_{x \rightarrow +\infty} \frac{\Gamma(\frac{N+1+\nu}{2})\Gamma(\frac{\nu}{2})}{\Gamma(\frac{1+\nu}{2})\Gamma(\frac{N+\nu}{2})} \left(1 + \frac{q_1}{\nu}\right)^{\frac{N+\nu}{2}} \frac{\left(1 + \frac{q_1 + x^2}{\nu}\right)^{-\frac{N+1+\nu}{2}}}{\gamma x^{\gamma-1} \left(1 + \frac{x^{2\gamma}}{\nu}\right)^{-\frac{1+\nu}{2}}}$$

From now, we consider $\gamma = \frac{N+\nu}{\nu}$ and we focus on the ratio on the right. The leading monomials are $\nu^{\frac{N+1+\nu}{2}} x^{-N-1-\nu}$ (numerator) and $\nu^{\frac{1+\nu}{2}} x^{-N-1-\nu}$ (denominator). Then the limit is simply $\nu^{\frac{N}{2}}$. Finally :

$$\ell = \frac{\Gamma(\frac{N+1+\nu}{2})\Gamma(\frac{\nu}{2})}{\Gamma(\frac{1+\nu}{2})\Gamma(\frac{N+\nu}{2})} \left(1 + \frac{q_1}{\nu}\right)^{\frac{N+\nu}{2}} \frac{\nu^{\frac{N}{2}+1}}{N+\nu}$$

Concerning Assumption 2, the proof is similar to the Gaussian case, but we have $\gamma \neq 1$ and $\ell \neq 1$ here. We are dealing with the existence and value of the limit:

$$\lim_{\substack{x \rightarrow +\infty \\ \lambda \rightarrow 1}} \frac{\Phi_R^{-1} \left(1 - \frac{1}{\ell \lambda x + 2} \right)^{\frac{1}{\gamma}}}{\Phi_R^{-1} \left(1 - \frac{1}{\ell x + 2} \right)^{\frac{1}{\gamma}}}.$$

As in the Gaussian case, we use the polar coordinates and the L'Hôpital's rule to get the result. \square

In the case of Student distributions, we shall apply the approximations, with the coefficients ℓ and γ given in Table 3 :

$$\begin{cases} \ell = \frac{\Gamma(\frac{\nu+N+1}{2})\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+N}{2})\Gamma(\frac{\nu+1}{2})} \left(1 + \frac{q_1}{\nu} \right)^{\frac{N+\nu}{2}} \frac{\nu \frac{N}{2} + 1}{\nu + N} \\ \gamma = \frac{N+\nu}{\nu} \end{cases}$$

Example 3 : Unimodal Gaussian Mixture case.

Lemma 5.8. *Unimodal Gaussian Mixture distribution satisfies Assumptions 1 and 2, with $\gamma = 1$ and ℓ given in Table 3.*

Proof. Firstly, to prove that Assumption 1 is satisfied, we take $\gamma = 1$ and calculate the limit :

$$\ell = \lim_{x \rightarrow +\infty} \frac{c_1^* g_{N+1}(q_1 + x^2)}{c_1 g_1(x^2)}$$

According to Table 1 and Equation (5.4), $g_d(t) = \sum_{k=1}^n \pi_k \theta_k^d e^{-\frac{\theta_k^2}{2} t}$, $c_1 = \frac{1}{\sqrt{2\pi}}$ and $c_1^* = \frac{1}{\sqrt{2\pi} \sum_{k=1}^n \pi_k \theta_k^N \exp\left(-\frac{\theta_k^2}{2} q_1\right)}$. Hence the limit :

$$\ell = \lim_{x \rightarrow +\infty} \frac{\sum_{k=1}^n \pi_k \theta_k^{N+1} \exp\left(-\frac{\theta_k^2}{2} q_1\right) \exp\left(-\frac{\theta_k^2}{2} x^2\right)}{\left[\sum_{k=1}^n \pi_k \theta_k^N \exp\left(-\frac{\theta_k^2}{2} q_1\right) \right] \left[\sum_{k=1}^n \pi_k \theta_k \exp\left(-\frac{\theta_k^2}{2} x^2\right) \right]}.$$

Let k^* such as $\theta_{k^*} = \min(\theta_1, \dots, \theta_n)$. After simplifications, it remains :

$$\ell = \frac{\theta_{k^*}^N \exp\left(-\frac{\theta_{k^*}^2}{2} q_1\right)}{\sum_{k=1}^n \pi_k \theta_k^N \exp\left(-\frac{\theta_k^2}{2} q_1\right)}$$

Now we have to check Assumption 2, i.e prove that :

$$\lim_{\substack{x \rightarrow +\infty \\ \lambda \rightarrow 1}} \frac{\Phi_R^{-1} \left(1 - \frac{1}{\ell \lambda x + 2} \right)}{\Phi_R^{-1} \left(1 - \frac{1}{\ell x + 2} \right)} = 1$$

Like the Gaussian case, we do the same changes of variables, and the limit becomes

$$\epsilon_\ell = \lim_{r \rightarrow 0} \frac{\phi_R \left(\Phi_R^{-1} \left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1} \right) \right)}{\phi_R \left(\Phi_R^{-1} \left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1} \right) \right)} = f(\theta)$$

Here, $\phi_R(t) = \sum_{k=1}^n \pi_k \theta_k \phi(\theta_k t)$, where ϕ is the normal p.d.f, i.e $\phi(t) = e^{-\frac{t^2}{2}}$. Since $\Phi_R^{-1} \left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1} \right)$ and $\Phi_R^{-1} \left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1} \right)$ tend to $+\infty$ when $r \rightarrow 0$, asymptotically, we only have to consider the terms of the sum with the largest coefficient $-\frac{\theta_k^2}{2}$, i.e the smallest θ_k . Let k^* such as $\theta_{k^*} = \min(\theta_1, \dots, \theta_n)$. We get :

$$\epsilon_\ell = \lim_{r \rightarrow 0} \frac{\pi_k^* \theta_k^* \phi(\theta_k^* \Phi_R^{-1}(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1}))}{\pi_k^* \theta_k^* \phi(\theta_k^* \Phi_R^{-1}(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1}))} = \lim_{r \rightarrow 0} \frac{\phi(\theta_k^* \Phi_R^{-1}(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1}))}{\phi(\theta_k^* \Phi_R^{-1}(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1}))}$$

Finally, we have already calculated this kind of limit in the Gaussian case, and proved that it was equal to 1. \square

Example 4 : Uniform Mixture case. In this case, the radius $R \stackrel{d}{=} \frac{\chi_d}{\mathcal{U}(0,1]}$. The density obtained is then :

$$(5.14) \quad f(x) = \frac{1}{|\Sigma|^{\frac{1}{2}}} \frac{\Gamma(\frac{d+1}{2})}{\sqrt{2\pi}^{\frac{d}{2}}} \frac{\chi_{d+1}^2((x-\mu)^T \Sigma^{-1}(x-\mu))}{[(x-\mu)^T \Sigma^{-1}(x-\mu)]^{\frac{d+1}{2}}}, \forall x \in \mathbb{R}^d \setminus \{\mu\}$$

where $\chi_m^2(x)$ denotes the c.d.f of the χ^2 distribution, with m degrees of freedom, evaluated at x .

Lemma 5.9. *Uniform Mixture distribution satisfies Assumptions 1 and 2, with $\gamma = N + 1$ and ℓ given in Table 3.*

Proof. Concerning Assumption 1, we consider $\gamma = N + 1$, and calculate the limit :

$$\ell = \lim_{x \rightarrow +\infty} \frac{c_1^* g_{N+1}(q_1 + x^2)}{c_1 \gamma x^{\gamma-1} g_1(x^2)}$$

According to Table 1 and Equation (5.4), $g_d(t) = \frac{\chi_{d+1}^2(t)}{t^{\frac{d+1}{2}}}$, $c_1 = \frac{1}{\sqrt{2\pi}}$ and $c_1^* = \frac{\Gamma(\frac{N+2}{2}) q_1^{\frac{N+1}{2}}}{\Gamma(\frac{N+1}{2}) \sqrt{\pi} \chi_{N+1}^2(q_1)}$. Hence the limit

$$\ell = \lim_{x \rightarrow +\infty} \sqrt{2\pi} \frac{\Gamma(\frac{N+2}{2}) q_1^{\frac{N+1}{2}}}{\Gamma(\frac{N+1}{2}) \sqrt{\pi} \chi_{N+1}^2(q_1)} \frac{\chi_{N+2}^2(q_1 + x^2)}{\chi_2^2(x^{2N+2})} \frac{x^{2N+2}}{(q_1 + x^2)^{\frac{N+2}{2}}} \frac{1}{(N+1)x^N},$$

which leads to

$$\ell = \frac{\Gamma(\frac{N+2}{2}) q_1^{\frac{N+1}{2}} \sqrt{2}}{\Gamma(\frac{N+1}{2}) (N+1) \chi_{N+1}^2(q_1)}.$$

To prove that Assumption 2 is filled, we have, as in the other cases, to calculate the limit :

$$\lim_{\substack{x \rightarrow +\infty \\ \lambda \rightarrow 1}} \frac{\Phi_R^{-1}\left(1 - \frac{1}{\ell \lambda x + 2}\right)}{\Phi_R^{-1}\left(1 - \frac{1}{\ell x + 2}\right)} = \lim_{r \rightarrow 0} \frac{\phi_R\left(\Phi_R^{-1}\left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + 1}\right)\right)}{\phi_R\left(\Phi_R^{-1}\left(1 - \frac{r \sin(\theta)}{2r \sin(\theta) + r \cos(\theta) + 1}\right)\right)} = f(\theta).$$

We get once more that $f(\theta) = 1$. \square

In this section, we introduced asymptotic approximations based solely on the function Φ_R , that is assumed to be known. These approximations are asymptotically equivalent to the theoretical conditional quantiles, for levels of quantiles close to 0 or 1. Thus, they are expected to improve the quantile regression. In the following section, we perform some numerical applications to study the approximation quality.

6. NUMERICAL STUDY

In order to give a visual overview of the estimators we have defined, we have plotted in Figure 2, the conditional quantiles of an elliptical process observed at $N = 5$ points. We call $X_1 \in \mathbb{R}^5$ the covariates vector. For $x \in \mathbb{R}$, X_2 denotes the process at x and the aim is to estimate / approximate the quantile of $X_2 | X_1 = x_1$. For simplicity, we assume that the process is centered, and stationary (matrices Σ and Σ_{11} are obtained through an exponential kernel). But our results even work without these assumptions. We also add that parameters for the Gaussian Mixture

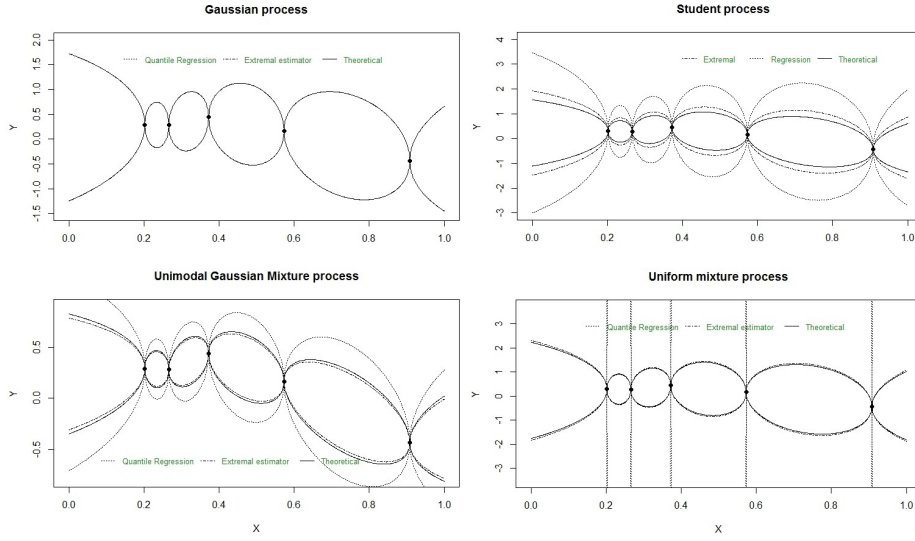


FIGURE 2. Theoretical quantiles, quantile regression and extremal approximation for Gaussian, Student, Gaussian Mixture and Uniform Mixture processes, with levels of quantile $\alpha = 0.995$ and $\alpha = 0.005$. In solid lines, the theoretical quantiles (estimated most of the time numerically), dotted, the Quantile Regression Estimator, and dashed, the Extremal Estimator.

example are $\theta_1 = 1$, $\theta_2 = 3$ and $p = 0.9$.

Of course, for the Gaussian process, the curves coincide. For the other examples, Quantile Regression Estimators seem very far from the theoretical curves, especially in the Uniform Mixture case. On the other hand, extremal approximations seem significantly better; they look closer to the target conditional quantiles here. We propose to use the following *RMSE* in order to quantify the error.

$$(6.1) \quad RMSE(\hat{q}_\alpha) = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(q_\alpha(X_2^{(i)} | X_1 = x_1) - \hat{q}_\alpha(X_2^{(i)} | X_1 = x_1) \right)^2}.$$

In the same way, we define the *RMSE* for Extremal Estimator :

$$(6.2) \quad RMSE(\hat{q}_\alpha) = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(q_\alpha(X_2^{(i)} | X_1 = x_1) - \hat{q}_\alpha(X_2^{(i)} | X_1 = x_1) \right)^2}.$$

The RMSE measures the average error in the approximation of the conditional quantiles. Table 4 is a summary of the RMSE for all treated examples, and different levels of α . Obviously, we only consider the cases $\alpha \geq \frac{1}{2}$ because elliptical distributions are symmetric.

For non Gaussian distributions, the $RMSE(\hat{q}_\alpha)$ decreases when α is close to 1, while the $RMSE(\hat{q}_\alpha)$ increases, which is the expected behavior. We also remark that for the median ($\alpha = 0.5$), the RMSE are always equal to 0. Indeed, if we replace α by $\frac{1}{2}$ in Equation (5.12) and Equation (4.18), our two approximations are equal to the conditional mean $\mu_{2|1}$, defined in Proposition 2.6. Since median and

α	Gaussian		Student		Unimodal GM		Uniform GM	
	$r(\hat{q}_\alpha)$	$r(\hat{\hat{q}}_\alpha)$	$r(\hat{q}_\alpha)$	$r(\hat{\hat{q}}_\alpha)$	$r(\hat{q}_\alpha)$	$r(\hat{\hat{q}}_\alpha)$	$r(\hat{q}_\alpha)$	$r(\hat{\hat{q}}_\alpha)$
0.5	0	0	0	0	0	0	0	0
0.6	0	0	0.027	0.372	0.001	0.019	0.067	0.334
0.7	0	0	0.059	0.376	0.003	0.036	0.151	0.301
0.8	0	0	0.104	0.358	0.006	0.051	0.297	0.249
0.9	0	0	0.195	0.322	0.012	0.061	0.770	0.177
0.95	0	0	0.314	0.289	0.024	0.061	1.880	0.123
0.9995	0	0	2.880	0.148	0.163	0.000	250.172	0.020
0.999995	0	0	16.546	0.081	0.109	$7.130 \cdot 10^{-06}$	25 178.530	0.008

TABLE 4. $r(\hat{\hat{q}}_\alpha) = RMSE(\hat{\hat{q}}_\alpha)$ and $r(\hat{q}_\alpha) = RMSE(\hat{q}_\alpha)$, for different levels of α , and different consistent elliptical distributions.

mean are coincident for elliptical distributions, we exactly estimate the conditional α -quantile for $\alpha = \frac{1}{2}$.

CONCLUSION

In this paper, we focused on conditional quantiles approximation, for elliptical processes with the consistency property. We proposed two different methods. The first one is to use quantile regression, i.e express the conditional quantile as affine transformation of the observed values. This approach is widely used in the literature but it often requires a large number of simulations, especially for extreme levels of quantile ($\alpha \rightarrow 0$ or $\alpha \rightarrow 1$). We have seen, in a first time, that we can obtain some explicit formulas in our case of consistent elliptical processes (4.1). Furthermore, we have given the distribution of the quantile regression (Theorem 4.4). We have seen that regression quantile is not adapted for non Gaussian distributions. It is a reason why we have proposed an extremal approximation (5.12).

Our approximation methods require the knowledge of the X_1 distribution. We have not explored the estimation procedure when the X_1 's distribution is estimated (parametrically e.g.). This is an interesting perspective for future work. Finally, we would like to emphasize that we gave examples in dimension $d = 1$, but all the results may be used in higher dimensions (see Figure 3).

REFERENCES

- Cambanis, S., Huang, S., and Simons, G. (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, (11):368–385.
- Cressie, N. (1988). Spatial prediction and ordinary kriging. *Mathematical Geology*, 20(4).
- Djurčić, D. and Torgasev, A. (2001). Strong asymptotic equivalence and inversion of functions in the class \mathcal{K}_c . *Journal of Mathematical Analysis and Applications*, 255:383–390.
- Eltoft, T., Kim, T., and Lee, T.-W. (2006). On the multivariate Laplace distribution. *IEEE Signal Processing Letters*, 13(5).
- Frahm, G. (2004). *Generalized Elliptical Distributions: Theory and Applications*. PhD thesis, Universität zu Köln.
- Hult, H. and Lindskog, F. (2002). Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability*, 34(3):587–608.
- Hunter, D. and Lange, K. (2000). Quantile regression via an mm algorithm. *Journal of Computational and Graphical Statistics*.

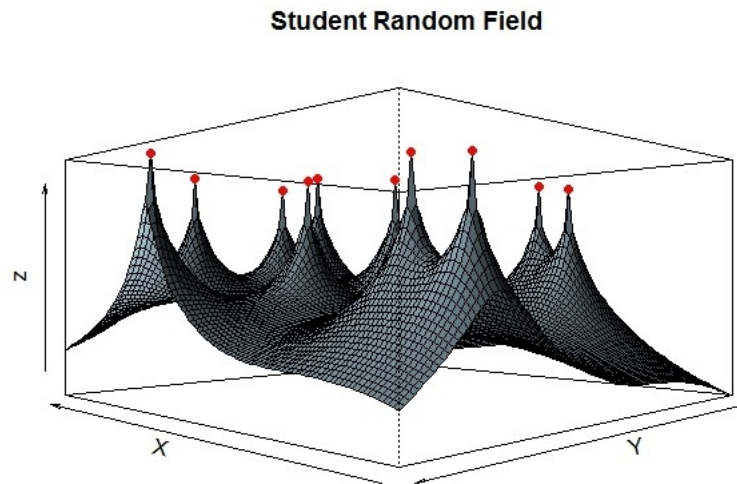


FIGURE 3. Quantile regression for a Student random field, with $\alpha = 0.0005$, in dimension $d = 2$

- Kano, Y. (1994). Consistency property of elliptical probability density functions. *Journal of Multivariate Analysis*, 51:139–147.
- Koenker, R. and Bassett, G. J. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.
- Kozubowski, T., Podgórski, K., and Rychlik, I. (2013). Multivariate generalized Laplace distribution and related random fields. *Journal of Multivariate Analysis*, 113:59–72.
- Krige, D. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society*, 52:119–139.
- Ligas, M. and Kulczycki, M. (2010). Simple spatial prediction by least squares prediction, simple kriging, and conditional expectation of normal vector. *Geodasy and Cartography*, 59(2):69–81.
- Matheron, G. (1963). *Traité de géostatistique appliquée*. Bureau de recherches géologiques et minières (France).
- Nadarajah, S. (2003). The katz type distribution with applications. *Statistics*, 37(4):341–358.
- Nadarajah, S. and Kotz, S. (2004). *Multivariate T-Distributions and Their Applications*. Cambridge University Press.
- Zheng, S. (2011). Gradient descent algorithms for quantile regression with smooth approximation. *International Journal of Machine Learning and Cybernetics*.
- Ziegel, J. (2015). Coherence and elicibility. *Mathematical Finance*.

UNIVERSITÉ DE LYON, UNIVERSITÉ LYON 1, INSTITUT CAMILLE JORDAN ICJ UMR
5208 CNRS

E-mail address: `veronique.maume@univ-lyon1.fr`

UNIVERSITÉ DE LYON, UNIVERSITÉ LYON 1, LABORATOIRE SAF EA2429

E-mail address: `didier.rulliere@univ-lyon1.fr`

UNIVERSITÉ DE LYON, UNIVERSITÉ LYON 1, INSTITUT CAMILLE JORDAN ICJ UMR
5208 CNRS

E-mail address: `usseglie@math.univ-lyon1.fr`