Spatial Expectile Predictions for Elliptical Random Fields

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

SPATIAL EXPECTILE PREDICTIONS FOR ELLIPTICAL RANDOM FIELDS

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

Abstract. In this work, we consider an elliptical random field. We propose some spatial expectile predictions at one site given observations of the field at some other locations. To this aim, we first give exact expressions for conditional expectiles, and discuss problems that occur for computing these values. A first affine expectile regression predictor is detailed, an explicit iterative algorithm is obtained, and its distribution is given. Direct simple expressions are derived for some particular elliptical random fields. The performance of this expectile regression is shown to be very poor for extremal expectile levels, so that a second predictor is proposed. We prove that this new extremal prediction is asymptotically equivalent to the true conditional expectile. We also provide some numerical illustrations, and conclude that Expectile Regression may perform poorly when one leaves the Gaussian random field setting.

Keywords: *Elliptical distribution; Expectile regression; Extremal expectile; Spatial prediction, Kriging.*
AMS Classification : 60G15; 60G60; 62H11; 62M30

1. Introduction

Kriging, introduced by Krige (1951), and formalized by Matheron (1963), aims at predicting the conditional mean of a random field $(Z_t)_{t\in T}$ given the values $Z_{t_1},...,Z_{t_N}$ of the field at some points $t_1,...,t_N \in T$, where typically $T \subset \mathbb{R}^d$. When using Kriging techniques, for any $x \in T$, the conditional mean of Z_x given $Z_{t_1},...,Z_{t_N}$ is approximated by a linear combination of $Z_{t_1},...,Z_{t_N}$ where the weight vector is the solution of a least square minimization problem (see Ligas and Kulczycki, 2010, for example). It seems natural to predict, in the same spirit as Kriging, other functionals by linear combinations. In a previous work (see Maume-Deschamps et al., 2016), we focused on quantiles. In this paper, we apply the same methodology to conditional expectiles in order to get spatial expectile predictions. However, as we will see, more technical details are involved, and some numerical algorithms will be required.

In 1978, Koenker and Bassett proposed a conditional quantile estimation as an affine combination of $Z_{t_1},...,Z_{t_N}$, called Quantile Regression (cf. Koenker and Bassett, 1978). More recently, some papers propose an Expectile Regression, using the same approach (see Yang et al. (2015) or Sobotka and Kneib (2012), for example). The weight vector is the solution of a minimization problem, with an asymmetric loss function. In the case where $\alpha = \frac{1}{2}$, it corresponds exactly to the conditional mean regression, or Kriging. Otherwise, it is more difficult to get explicit formulas. The Expectile Regression approach usually requires time consuming simulations to compute expectations. Moreover, in a non-gaussian setting, the conditional expectile may not be expressed as a linear combination of the covariates, thus the consistency of the estimation by expectile regression is not guaranteed.
In this paper, we focus on elliptical random fields. Elliptical distributions, formalized by Cambanis et al. (1981), have the advantage of being stable under affine transformations. Therefore, explicit iterative algorithms for the expectile regression may be obtained for consistent elliptical distributions (cf. Kano (1994)). Nevertheless, the expectile regression is generally not equal to the conditional expectile and the difference may be large, especially for extreme levels of expectile. This is why we propose a new dedicated expectile prediction that is adapted to extremal levels.

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 expectiles in Section 3. Section 4 is devoted to expectile regression for consistent elliptical random fields: iterative algorithms are obtained. In Section 5, we propose some extremal predictions and prove asymptotic equivalences when the expectile level is close to 0 or 1. Section 6 provides a numerical study. In particular, we emphasize the fact that expectile regression is generally not appropriate, especially for high level expectiles. We illustrate this point on several examples.

This work was supported by the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR)..

2. Elliptical distributions

In this section, we recall some useful properties and classical examples of elliptical distributions. Most results may be found, for instance, in Frahm (2004). As these results are classical, we use here almost the same formulation as in our previous article Maume-Deschamps et al. (2016).

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 $\Phi : \mathbb{R}^+ \to \mathbb{R}$ such that the characteristic function of $(X-\mu)$ is*

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

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:

$$\Sigma = \tau K \tag{2.1}$$

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 $\nu > 2$ 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 et al. (1981))**.** *The random vector X is elliptical, $X \sim \mathcal{E}_d(\mu, \Sigma, \Phi)$, if and only if*

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

where $\Lambda\Lambda^T = \Sigma$, $U^{(d)}$ is a $d-$dimensional random vector uniformly distributed on $\mathcal{S}^{d-1}$ (the unit sphere 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 to be consistent if Φ is dimension-free, i.e. if Φ does not depend on d.*

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 denote its radius. X is consistent if and only if:*

$$R \overset{d}{=} \frac{\chi_d}{\epsilon}\,, \tag{2.3}$$

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 whose law 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.*

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

where $c_d g_d(t) = \frac{\Gamma(\frac{d}{2})}{2\pi^{\frac{d}{2}}}\sqrt{t}^{-(d-1)} f_R(\sqrt{t})$, 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), Laplace in Eltoft et al. (2006) and Kozubowski et al. (2013), Unimodal Gaussian Mixture in Fraley and Raftery (2002) and Slash in Gómez et al. (2007).

In Table 1, $K_m(x) = \frac{\pi}{2}\frac{I_{-m}(x) - I_m(x)}{\sin(m\pi)}$ denotes the modified Bessel function of the second kind with order m, where $I_m(x) = \displaystyle\sum_{k=0}^{+\infty} \frac{1}{k!\Gamma(k+m+1)}\left(\frac{x}{2}\right)^{2k+m}$ (see Abramowitz et al. (1966)), and $\chi^2_m(x)$ 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

Distribution	Coefficient c_d	Generator $g_d(t)$
Gaussian	$\frac{1}{(2\pi)^{\frac{d}{2}}}$	$\exp(-\frac{t}{2})$
Student, $\nu > 0$	$\frac{\Gamma(\frac{d+\nu}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{(\nu\pi)^{\frac{d}{2}}}$	$\left(1+\frac{t}{\nu}\right)^{-\frac{d+\nu}{2}}$
Logistic	$\frac{\Gamma(\frac{d}{2})}{(2\pi)^{\frac{d}{2}}} \left[\int_0^{+\infty} x^{\frac{d}{2}-1} \cdot \frac{e^{-x}}{(1+e^{-x})^2} dx\right]^{-1}$	$\frac{\exp(-\frac{1}{2}t)}{\left(1+\exp(-\frac{1}{2}t)\right)^2}$
Kotz, $q, r, s > 0$	$\frac{s\Gamma\left(\frac{d}{2}\right)}{\pi^{\frac{d}{2}}\Gamma\left(\frac{2q+d-2}{2s}\right)} r^{\frac{2q+d-2}{2s}}$	$t^{q-1}\exp(-rt^s)$
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, $\lambda > 0$	$\frac{2}{\lambda(2\pi)^{\frac{d}{2}}}$	$\frac{K_{\frac{d}{2}-1}\left(\sqrt{\frac{2}{\lambda}t}\right)}{\left(\sqrt{\frac{\lambda}{2}t}\right)^{\frac{d}{2}-1}}$
Slash	$\frac{2^{\frac{a}{2}-1} a\Gamma\left(\frac{d+a}{2}\right)}{\pi^{\frac{d}{2}}}$	$\frac{\chi^2_{d+a}(t)}{t^{\frac{d+a}{2}}}$

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

do not have this property (except the Kotz distribution with $s = q = 1$, and $r = \frac{1}{2}$, i.e the Gaussian distribution). For consistent models, the non negative random variable ϵ is given in Table 2.

Distribution	ϵ
Gaussian	1
Student, $\nu > 0$	$\frac{\chi_\nu}{\sqrt{\nu}}$
Unimodal Gaussian Mixture	$\sum_{k=1}^{n} \pi_k \delta_{\theta_k}$
Laplace, $\lambda > 0$	$\frac{1}{\sqrt{\mathcal{E}(\frac{1}{\lambda})}}$
Slash	$Beta\,(a, 1)$

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

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 characterizations, using indifferently the notations $X \sim \mathcal{E}_d(\mu, \Sigma, \Phi)$, $X \sim \mathcal{E}_d(\mu, \Sigma, R)$ 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:*

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

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 (R, d)-elliptical random vector with the consistency property is $(R, p)-$elliptical. As a consequence, all marginals are $(R, 1)-$elliptical. The following proposition gives some indications concerning the conditional distributions of elliptical vectors. The proof is already given in Maume-Deschamps et al. (2016).

Proposition 2.6 (Conditional distribution)**.** *Let $X = (X_1, X_2)^T$ be a consistent $(R, d)-$elliptical random vector with 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:*

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

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

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

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

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

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

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

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

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

We have introduced the main definitions and properties of elliptical distributions. With these tools, we can define the notion of elliptical random fields. Indeed, a random field $(X(t))_{t\in T}$ is $R-$elliptical if $\forall n \in \mathbb{N}$, $\forall t_1, ..., t_n \in T$, the vector $(X(t_1), ..., X(t_n))$ is $(R, n)-$elliptical. Obviously, it implies that all the $k-$dimensional subvectors of $(X(t_1), ..., X(t_n))$ are $(R, k)-$elliptical. This assumption corresponds to consistent elliptical distributions properties given in Propositions 2.4, 2.5 and in Kano (1994). We thus focus our study on elliptical distributions with the consistency property.
In the following section, we focus on conditional expectiles of elliptical distributions, applied to our problem of spatial prediction. We consider the following context: $(X(t))_{t\in T}$ is an $R-$elliptical random field defined on some metric space T. We consider N observations at locations $t_1, ..., t_N \in T$, called $(X(t_1), ..., X(t_N))$. Given $X(t_1), ..., X(t_N)$, our aim is to predict, at a site $t \in T$, a functional of the distribution of $X(t)$: the further defined *expectile*. Notice that the vector $(X(t), X(t_1), ..., X(t_N))$ is $(R, N + 1)-$elliptical. Thus, we can denote $X_2 = X(t) \in \mathbb{R}$ and $X_1 = (X(t_1), ..., X(t_N)) \in \mathbb{R}^N$ and restrict ourselves to the study of the conditional distribution of the random variable X_2 given the random vector X_1.

3. Theoretical Expectiles

3.1. **General expression.** Expectiles, introduced by Newey and Powell (1987), may be seen as a generalization of quantiles. In this part, we introduced these quantities with the notion of elicitability (see e.g. Ziegel (2014)). Indeed, let X be a random variable. The $\alpha-$quantile q_α of X is given by the minimization problem:

$$q_\alpha(X) = \underset{q\in\mathbb{R}}{\arg\min}\ \mathbb{E}\left[(\alpha - 1)(X - q)\mathbb{1}_{\{X-q<0\}} + \alpha(X - q)\mathbb{1}_{\{X-q>0\}}\right] \tag{3.1}$$

Equation (3.1) easily leads to the relationship $\mathbb{P}\left(X \le q_\alpha(X)\right) = \alpha$ for continuous distributions. The $\alpha-$expectile $e_\alpha(X)$ is defined as the solution of a similar minimization problem:

$$e_\alpha(X) = \underset{e\in\mathbb{R}}{\arg\min}\ \mathbb{E}\left[(1 - \alpha)(X - e)^2\mathbb{1}_{\{X-e<0\}} + \alpha(X - e)^2\mathbb{1}_{\{X-e\ge0\}}\right] \tag{3.2}$$

The former definition assumes the existence of the first two moments of X, but other definitions can be proposed, involving only the first moment (see e.g. Bellini et al. (2014)). In this part, we give a general expression of conditional expectiles in our elliptical context. Consider the respective cumulative distribution functions Φ_R and Φ_R^*,

$$\begin{cases} \Phi_R(x) & = & \mathbb{P}\left(RU^{(1)} \le x\right), \\ \Phi_{R^*}(x) & = & \mathbb{P}\left(R^*U^{(1)} \le x\right), \end{cases} \tag{3.3}$$

where $U^{(1)}$ is 1 or -1 with probability $\frac{1}{2}$. Φ_R and Φ_{R^*} are respective cumulative distribution functions of the reduced centered $(R, 1)-$elliptical random variable and $(R^*, 1)-$elliptical random variable. With this notation, we can introduce the following definition.

Definition 3.1. *Let $\Psi_R : \mathbb{R}^* \to]-\infty, 0[\cup]1, +\infty[$ be such that:*

$$\Psi_R(x) = \Phi_R(x) + \frac{1}{x}\int_x^{+\infty} y c_1 g_1(y^2)dy\,. \tag{3.4}$$

This function is introduced mainly to simplify future equations and relationships. From Equation (3.4), we can deduce several properties.

Lemma 3.1. *Ψ_R satisfies the following properties:*
- $\Psi_R(-x) = 1 - \Psi_R(x), \forall x \in \mathbb{R}^*$.
- $\Psi_R : \mathbb{R}^* \to]-\infty, 0[\cup]1, +\infty[$ *is bijective and decreasing.*
- $\Psi_R^{-1}(1-\alpha) = -\Psi_R^{-1}(\alpha), \forall \alpha \in \mathbb{R}^*$.

The proof requires very classical and simple calculations, then we do not develop it in details.
We will see that the general expression of expectile is related to the function Ψ_R. We introduce a last lemma before giving this expression.

Lemma 3.2. *Let $X \sim \mathcal{E}_1(\mu, \sigma^2, g_1)$ be an elliptical random variable. We have the following relationship:*

$$\mathbb{E}[X\mathbb{1}_{\{X\geq 0\}}] = \mu\Phi_R\left(\frac{\mu}{\sigma}\right) + \sigma\int_{-\frac{\mu}{\sigma}}^{+\infty} y c_1 g_1(y^2)dy \tag{3.5}$$

Proof. We recall that density of X is given by:

$$f_X(x) = \tfrac{c_1}{\sigma} g_1\left(\tfrac{(x-\mu)^2}{\sigma^2}\right)$$

Then:

$$\mathbb{E}[X\mathbb{1}_{\{X\geq 0\}}] = \int_0^{+\infty} x\frac{c_1}{\sigma}g_1\left(\frac{(x-\mu)^2}{\sigma^2}\right)dx = \mu\int_{-\frac{\mu}{\sigma}}^{+\infty} c_1 g_1\left(y^2\right)dy + \sigma\int_{-\frac{\mu}{\sigma}}^{+\infty} y c_1 g_1\left(y^2\right)dy$$

Since $\Phi_R(x) = \int_{-\infty}^{x} c_1 g_1\left(y^2\right)dy$ and $1 - \Phi_R(x) = \Phi_R(-x)$ (by the symmetry properties of elliptical distributions), we get Equation (3.5). □

We are now able to give a general expression for the $\alpha-$expectile of a univariate elliptical distribution.

Proposition 3.3. *Let X a $(R, 1)-$elliptical random variable, with parameters μ and σ^2. The $\alpha-$expectile e_α of X is given by:*

$$e_\alpha(X) = \begin{cases} \mu + \sigma\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right) & ,\alpha \neq \frac{1}{2} \\ \mu & ,\alpha = \frac{1}{2} \end{cases} \tag{3.6}$$

Proof. We have to solve the minimization problem:

$$e_\alpha(X) = \arg\min_{x\in\mathbb{R}} \mathbb{E}[\mathcal{S}_\alpha(X-x)]$$

By deriving the loss function, we get:

$$(1-\alpha)\mathbb{E}[X-e_\alpha] + (2\alpha-1)\mathbb{E}\left[(X-e_\alpha)\mathbb{1}_{\{X-e_\alpha\geq 0\}}\right] = 0$$

Using Lemma 3.2, we have the equality:

$$\mathbb{E}[(X-e_\alpha)\mathbb{1}_{\{X-\mu\geq 0\}}] = (\mu-e_\alpha)\Phi_R\left(\frac{\mu-e_\alpha}{\sigma}\right) + \sigma\int_{-\frac{\mu-e_\alpha}{\sigma}}^{+\infty} y c_1 g_1(y^2)dy$$

We add this term in the previous equation:

$$(1-\alpha)(\mu-e_\alpha) + (2\alpha-1)\left[(\mu-e_\alpha)\Phi_R\left(\frac{\mu-e_\alpha}{\sigma}\right) + \sigma\int_{-\frac{\mu-e_\alpha}{\sigma}}^{+\infty} y c_1 g_1(y^2)dy\right] = 0$$

A level of $\alpha = \frac{1}{2}$ leads to the obvious solution $e_\alpha = \mu$. Then, we consider now $\alpha \neq \frac{1}{2}$ (then $e_\alpha \neq \mu$), the equation may be written as follows:

$$\Phi_R\left(\frac{e_\alpha - \mu}{\sigma}\right) + \frac{\sigma}{e_\alpha - \mu}\int_{\frac{e_\alpha-\mu}{\sigma}}^{+\infty} yc_1 g_1(y^2)dy = \frac{\alpha}{2\alpha - 1}$$

Hence $\Psi_R\left(\frac{e_\alpha-\mu}{\sigma}\right) = \frac{\alpha}{2\alpha-1}$, and $e_\alpha = \mu + \sigma\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$ □

Let us focus now on the conditional expectiles of an elliptical vector. Be $X \in \mathbb{R}^{N+1}$ a consistent $(R, N+1)-$elliptical vector, $X_1 \in \mathbb{R}^N$ and $X_2 \in \mathbb{R}$ subvectors of X. Since $X_2|X_1$ is $(R^*, 1)-$elliptical, where R^* is given in Proposition 2.6, it is not difficult to get a general expression for $e_\alpha(X_2|X_1)$.

Proposition 3.4. *Let $X = (X_1, X_2)$ be a $(R, N+1)-$elliptical random vector, with parameters μ and Σ. The $\alpha-$expectile e_α of $X_2|(X_1 = x_1)$ is given by:*

$$e_\alpha(X_2|X_1 = x_1) = \begin{cases} \mu_{2|1} + \sigma_{2|1}\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right) & ,\alpha \neq \frac{1}{2} \\ \mu_{2|1} & ,\alpha = \frac{1}{2} \end{cases} \tag{3.7}$$

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

Proof. Obvious with Proposition 3.3 and Proposition 2.6. □

Our formulas involve the terms $\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$ and $\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$, which are respectively the solutions of $\Phi_R(x) + \frac{1}{x}\int_x^{+\infty} yc_1g_1(y^2)dy = \frac{\alpha}{2\alpha-1}$ and $\Phi_{R^*}(x) + \frac{1}{x}\int_x^{+\infty} yc_1^*g_1^*(y^2)dy = \frac{\alpha}{2\alpha-1}$. Then, in the next subsection, we propose some algorithms to compute these terms.

3.2. **Algorithms.** We propose some algorithms to compute the standardized conditional expectile $\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$, and thus the conditional expectile $e_\alpha(X_2|X_1 = x_1)$ of Equation (3.7). Obviously, the same algorithms will apply for $\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$, using R^* instead of R.
We consider two kinds of algorithms: a MM algorithm (see Hunter and Lange (2004)) and a fixed-point algorithm. We also study the speed of convergence for these two algorithms.

Let us introduce the following definitions and properties that may be found in Frontini and Sormani (2003).

Consider a sequence $\left(e^{(k)}\right)_{k\in\mathbb{N}}$ and a target value $e^* \in \mathbb{R}$. The sequence $\left(e^{(k)}\right)_{k\in\mathbb{N}}$ is said to converge to $e^* \in \mathbb{R}$ with order $p \in [1, \infty)$ if for all k, $\left|e^{(k)} - e^*\right| \leq \epsilon_k$ where ϵ_k is a positive sequence satisfying

$$\exists c > 0 : \lim_{k\to+\infty} \frac{\epsilon_{k+1}}{\epsilon_k^p} = c\,. \tag{3.8}$$

The coefficient c is called the *asymptotic factor.* In particular, one says that the convergence is linear if $p = 1$ and $c < 1$, or quadratic if $p = 2$.

Now, consider a fixed-point algorithm $e^{(k+1)} = f\left(e^{(k)}\right)$, where $e^{(k+1)}, e^{(k)} \in \mathbb{R}$, and $f : \mathbb{R} \to \mathbb{R}$ is a continuous and differentiable function, $k \in \mathbb{N}$. The convergence of the algorithm is insured by $|f'(e)| < 1, \forall e \in \mathbb{R}$. If furthermore f is p-times differentiable with $\forall m < p, f^{(m)}(e^*) = 0$ and $f^{(p)}(e^*) \neq 0$, then the convergence is of order p with an asymptotic factor $c = \dfrac{f^{(p)}(e^*)}{p!}$ (see Frontini and Sormani, 2003).

3.2.1. *MM algorithm.* We now present an approach based on a MM algorithm, in order to compute $\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$ and the associated expectile of Equation (3.7). A similar approach has been proposed in Yang et al. (2015), and some results thus explicitly refer to this paper.

In the following, we denote by $\mathcal{S}_\alpha$ the function $\mathcal{S}_\alpha(x) = (1-\alpha)x^2\mathbb{1}_{\{x<0\}} + \alpha x^2\mathbb{1}_{\{x\geq 0\}}$.

Lemma 3.5 (Yang et al. (2015))**.** *The function $\mathcal{S}'_\alpha$ is Lipschitz:*

$$|\mathcal{S}'_\alpha(a) - \mathcal{S}'_\alpha(b)| \leq 2\max\{1-\alpha, \alpha\}|a - b|, \forall a, b \in \mathbb{R} \tag{3.9}$$

Hence the following quadratic upper bound:

$$\mathcal{S}_\alpha(a) \leq \mathcal{S}_\alpha(b) + \mathcal{S}'_\alpha(b)(a-b) + \max\{1-\alpha, \alpha\}(a-b)^2, \forall a, b \in \mathbb{R} \tag{3.10}$$

Using this result from Yang et al. (2015), we propose a new MM algorithm, detailed hereafter.

Proposition 3.6 (MM algorithm)**.** *The following sequence* $\left(e^{(k)}\right)_{k\in\mathbb{N}}$ *converges to* $\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$*:*

$$\left\{\begin{array}{lcl} e^{(0)} & = & 0 \\ e^{(k+1)} & = & e^{(k)} - \frac{e^{(k)}}{\max\{1-\alpha,\alpha\}}\left[\alpha-(2\alpha-1)\Phi_R\left(e^{(k)}\right)\right] + \\ & & +\dfrac{2\alpha-1}{\max\{1-\alpha,\alpha\}}\displaystyle\int_{e^{(k)}}^{+\infty} y c_1 g_1(y^2)dy \end{array}\right. \tag{3.11}$$

Furthermore, the convergence is linear with an asymptotic factor

$$c = 1 - \frac{1}{\max\{1-\alpha,\alpha\}}\left[\alpha-(2\alpha-1)\Phi_R\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)\right] \tag{3.12}$$

Proof. Let X be the $(R,1)-$elliptical random variable with parameters 0 and 1. At the k^{th} iteration, we call $R^{(k)} = X - e^{(k)}$. It follows for any $e \in \mathbb{R}$:

$$X - e = R^{(k)} - (e - e^{(k)})$$

Thanks to Lemma 3.5, we have the following upper bound:

$$\mathbb{E}\left[\mathcal{S}_\alpha\left(R^{(k)} - (e-e^{(k)})\right)\right] \leq \mathbb{E}\left[\mathcal{S}_\alpha(R^{(k)})\right] - (e-e^{(k)})\mathbb{E}\left[\mathcal{S}'_\alpha(R^{(k)})\right] + \max\{1-\alpha,\alpha\}\left(e-e^{(k)}\right)^2$$

Let us denote $Q\left(e,e^{(k)}\right) = \mathbb{E}\left[\mathcal{S}_\alpha(R^{(k)})\right] - (e-e^{(k)})\mathbb{E}\left[\mathcal{S}'_\alpha(R^{(k)})\right] + \max\{1-\alpha,\alpha\}\left(e-e^{(k)}\right)^2$. When trying to minimize $\mathbb{E}\left[\mathcal{S}_\alpha\left(R^{(k)} - (e-e^{(k)})\right)\right]$, the principle of MM Algorithm is to minimize its upper bound, and to choose $e^{(k+1)} = \arg\min_{e\in\mathbb{R}} Q\left(e,e^{(k)}\right)$. We easily get:

$$e^{(k+1)} = e^{(k)} + \tfrac{1}{2\max\{1-\alpha,\alpha\}}\mathbb{E}\left[\mathcal{S}'_\alpha\left(X-e^{(k)}\right)\right]$$

It remains to calculate $\mathbb{E}\left[\mathcal{S}'_\alpha\left(X-e^{(k)}\right)\right]$.

$$\mathbb{E}\left[\mathcal{S}'_\alpha\left(X-e^{(k)}\right)\right] = 2(1-\alpha)\mathbb{E}\left[X-e^{(k)}\right] + 2(2\alpha-1)\mathbb{E}\left[\left(X-e^{(k)}\right)\mathbb{1}_{\{X-e^{(k)}\geq 0\}}\right]$$

Lemma 3.2 leads to Equation (3.11). Now, let us prove the convergence of the algorithm. It is sufficient to have $|\varphi'(x)| < 1, \forall x$.

$$\varphi'(x) = 1 - \frac{\alpha-(2\alpha-1)\Phi_R(x)}{\max\{1-\alpha,\alpha\}}.$$

Clearly, since $0 < \Phi_R(x) < 1, \forall x \in \mathbb{R}$, $0 < \varphi'(x) < 1$. Hence the convergence. Furthermore, since $\varphi'(x) \neq 0, \forall x \in \mathbb{R}$, the convergence is linear in the sense of Equation (3.8), with asymptotic factor $\varphi'\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)$, hence Equation (3.12). ☐

As expected, $\lim_{\alpha\to 1} c = 1$. Indeed, when α is close to 1, $e_\alpha(X)$ is huge, and its calculation requires more iterations. On the other hand, a level of $\alpha = \frac{1}{2}$ gives $c = 0$, since the sequence in the algorithm is constant, $e^{(k)} = 0$ for all k.

3.2.2. *Fixed-point algorithm.* It can be shown that $e_\alpha(X)$ satisfies a fixed-point relationship. In this paragraph, we detail this relationship and the corresponding fixed-point algorithm.

Proposition 3.7 (Fixed-point algorithm)**.** *The following sequence* $\left(e^{(k)}\right)_{k\in\mathbb{N}}$ *converges to* $\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$*:*

$$\left\{\begin{array}{rc} e^{(0)} = & 0 \\ e^{(k+1)} = & \dfrac{(1-2\alpha)\int_{e^{(k)}}^{+\infty} y c_1 g_1(y^2)dy}{(2\alpha-1)\Phi_R\left(e^{(k)}\right)-\alpha} \end{array}\right. \tag{3.13}$$

Furthermore, the convergence is quadratic (in the sense of Equation (3.8) *with* $p = 2$*), with an asymptotic factor*

$$c = \frac{(2\alpha-1)c_1 g_1\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)^2\right)}{2\left[(2\alpha-1)\Phi_R\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)-\alpha\right]} \tag{3.14}$$

Proof. Let X the $(R,1)-$elliptical random variable with parameters 0 and 1. Since $\mu = 0$ and $\sigma = 1$, $\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right) = e_\alpha(X)$ and thus:

$$\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right) = \underset{e\in\mathbb{R}}{\arg\min}\ \mathbb{E}\left[\mathcal{S}_\alpha(X-e)\right]$$

Then, $\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$ is such that $\mathbb{E}\left[\mathcal{S}'_\alpha\left(X-\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)\right] = 0$, i.e:

$$2(1-\alpha)\mathbb{E}\left[\left(X-\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)\right] + 2(2\alpha-1)\mathbb{E}\left[\left(X-\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)\mathbb{1}_{\{X-\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\geq 0\}}\right] = 0$$

With the previous equation, and the formula of Lemma 3.2, we get the following fixed-point relation

$$\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right) = \frac{(1-2\alpha)\int\limits_{\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)}^{+\infty} yc_1g_1(y^2)dy}{(2\alpha-1)\Phi_R\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)-\alpha}.$$

As in the MM algorithm given above, we have a recursive algorithm $e^{(k+1)} = \varphi(e^{(k)})$. Then, the convergence is insured by $\left|\varphi'\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)\right| < 1$. The quadratic convergence is insured if $\varphi'\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right) = 0$ and $\varphi''\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right) \neq 0$. We have

$$\varphi'(x) = (2\alpha-1)c_1g_1(x^2)\frac{x(2\alpha-1)\Phi_R(x) - \alpha x + (2\alpha-1)\int\limits_x^{+\infty} yc_1g_1(y^2)dy}{\left[(2\alpha-1)\Phi_R(x)-\alpha\right]^2}.$$

Using Equation (3.1), we deduce $\varphi\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right) = 0$. With this relationship, we not only show that the algorithm converges, but this convergence is at least quadratic. We straightforward calculations, and get:

$$\varphi''\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right) = \frac{(2\alpha-1)c_1g_1\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)^2\right)}{\left[(2\alpha-1)\Phi_R\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)-\alpha\right]}.$$

If $\alpha \neq \frac{1}{2}$, $\varphi''\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right) \neq 0$ and the convergence is quadratic. Furthermore, the asymptotic factor c is easily deduced, with the relationship $c = \frac{\varphi''\left(\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)\right)}{2}$, hence Equation (3.14). $\square$

Considering the convergence rate, this fixed-point algorithm is more efficient than the MM algorithm. In the following, unless specified, we thus only consider this fixed-point algorithm.

3.3. **Examples.** In this subsection, we give some examples of theoretical conditional expectiles calculated by our algorithms for Gaussian, Student and Gaussian Mixture distributions. The difficulty is to calculate the terms $\Phi_{R*}\left(e^{(k)}\right)$ and $\int\limits_{e^{(k)}}^{+\infty} yc_1^*g_1^*(y^2)dy$. In the general case, we have no guarantee of obtaining closed-form formulas for these quantities. Nevertheless, the calculation is possible in the three examples mentioned above. As the fixed-point algorithm (3.13) is asymptotically faster than the MM algorithm (3.11), we just give results for the fixed-point algorithm, but the MM algorithm works as well.
Recall that in all cases, we have the relationship

$$e_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right).$$

Then, for all the examples, we give an algorithm to calculate $\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$.

3.3.1. *Gaussian example.* The gaussian case, as usual, is the most simple case. Since $R^* \overset{d}{=} R$, calculations are very simple, and lead to the following sequence which converges to $\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$:

$$\left\{\begin{array}{rcl} e^{(0)} = & & 0 \\ e^{(k+1)} = & & \frac{(1-2\alpha)\varphi\left(e^{(k)}\right)}{(2\alpha-1)\Phi\left(e^{(k)}\right)-\alpha} \end{array}\right. \tag{3.15}$$

3.3.2. *Student example.* In the Student case, the conditional radius R^* is not equal to R, but calculations are still possible. Indeed, using Equations (2.8), Algorithm (3.13) becomes

$$e^{(k+1)} = \frac{(1-2\alpha)\int\limits_{e^{(k)}}^{+\infty} y c_{2|1} g_{N+1}\left(q_1 + (x_2-\mu_{2|1})^T\Sigma^{-1}(x_2-\mu_{2|1})\right)dy}{(2\alpha-1)\Phi_\nu\left(e^{(k)}\right)-\alpha},$$

where $c_{2|1}$ and g_n are given in Proposition 2.6 and Table 1 . Hence $\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$ is obtained by:

$$\left\{\begin{array}{rc} e^{(0)} = & 0 \\ e^{(k+1)} = & \dfrac{(2\alpha-1)\frac{\Gamma\left(\frac{\nu+N+1}{2}\right)}{\Gamma\left(\frac{\nu+N}{2}\right)\sqrt{\pi}}\frac{\sqrt{\nu+q_1}}{1-\nu-N}\left(1+\frac{1}{\nu}\frac{\left(e^{(k)}\right)^2}{\left(1+\frac{1}{\nu}q_1\right)}\right)^{\frac{1-\nu-N}{2}}}{(2\alpha-1)\Phi_{\nu+N}\left(e^{(k)}\sqrt{\frac{\nu+N}{\nu+q_1}}\right)-\alpha} \end{array}\right. \tag{3.16}$$

3.3.3. *Gaussian Mixture example.* We do the same kind of calculations as in the Gaussian and Student cases, and get the following algorithm.

$$\left\{\begin{array}{rc} e^{(0)} = & 0 \\ e^{(k+1)} = & \dfrac{(1-2\alpha)\sum\limits_{i=1}^n \pi_i\theta_i^{N-1}\exp\left(-\frac{\theta_i^2}{2}q_1\right)\varphi\left(\theta_i e^{(k)}\right)}{(2\alpha-1)\sum\limits_{i=1}^n \pi_i\theta_i^N\exp\left(-\frac{\theta_i^2}{2}q_1\right)\Phi\left(\theta_i e^{(k)}\right)-\alpha\sum\limits_{i=1}^n \pi_i\theta_i^N\exp\left(-\frac{\theta_i^2}{2}q_1\right)} \end{array}\right. \tag{3.17}$$

Some numerical applications and illustrations are given in Section 6.

In some other cases, it is difficult to express $\Phi_{R*}\left(e^{(k)}\right)$ or $\int\limits_{e^{(k)}}^{+\infty} y c_1^* g_1^*(y^2)dy$. Obviously, one can approximate numerically these values, but this approximation may lead to a poor prediction of the expectile, especially for extreme levels of α. This is why we propose two prediction methods. The first one, introduced in the next section, is called expectile regression.

4. Expectile Regression

Expectile Regression, introduced by Newey and Powell (1987), by analogy to Quantile Regression (see Koenker and Bassett (1978)), is an usual way to estimate conditional expectiles. If $X_1 \in \mathbb{R}^N$ and $X_2 \in \mathbb{R}$, the $\alpha-$expectile of $X_2|(X_1 = x_1)$ is approximated by

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

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

$$(\beta^*, \beta_0^*) = \underset{\beta\in\mathbb{R}^N,\beta_0\in\mathbb{R}}{\arg\min}\ \mathbb{E}\left[\mathcal{S}_\alpha(X_2 - \beta^T X_1 - \beta_0)\right]. \tag{4.2}$$

and where the scoring function $\mathcal{S}_\alpha$ (see Ziegel (2014)) is

$$\mathcal{S}_\alpha(x) = (1-\alpha)x^2\mathbb{1}_{\{x<0\}} + \alpha x^2\mathbb{1}_{\{x>0\}}. \tag{4.3}$$

Obviously, if $\alpha = \frac{1}{2}$, the scoring function $\mathcal{S}_{\frac{1}{2}}(x)$ is the least square loss function, and the minimization 4.2 gives the kriging vector $\beta^* = \left(\mathbb{E}[X_1X_1^T] - \mathbb{E}[X_1]\mathbb{E}[X_1]^T\right)^{-1}(\mathbb{E}[X_1X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2])$, and $\beta_0^* = \mathbb{E}[X_2] - \beta^{*T}\mathbb{E}[X_1]$. If $\alpha \neq \frac{1}{2}$, the problem is more difficult to solve, and we usually need simulations and stochastic algorithms. Fortunately, in our elliptical context, we do not need it to express β^* and β_0^*. The following lemma will be useful to calculate β^*.

Lemma 4.1. *Let* (X_1, X_2) *a consistent* $(R, 2)-$*elliptical random vector, with parameters* $\mu = (\mu_1, \mu_2)$ *and* $\Sigma = \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$. *We have the following relationship:*

$$\mathbb{E}[X_1X_2\mathbb{1}_{\{X_2\geq 0\}}] = \left(\mu_1\mu_2 - \rho\frac{\sigma_1}{\sigma_2}\mu_2^2\right)\Phi_R\left(\frac{\mu_2}{\sigma_2}\right) + (\mu_1\sigma_2 + \rho\sigma_1\mu_2)\int\limits_{-\frac{\mu_2}{\sigma_2}}^{+\infty} y c_1 g_1(y^2)dy + \rho\sigma_1\sigma_2\int\limits_{-\frac{\mu_2}{\sigma_2}}^{+\infty} y^2 c_1 g_1(y^2)dy \tag{4.4}$$

Proof. We write:

$$\mathbb{E}[X_1X_2\mathbb{1}_{\{X_2\geq 0\}}] = \int\limits_0^{+\infty} x_2 f_{X_2}(x_2)\left(\int\limits_{-\infty}^{+\infty} x_1\frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)}dx_1\right)dx_2$$

We have $\int\limits_{-\infty}^{+\infty} x_1 \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)} dx_1 = \mathbb{E}[X_1|X_2]$. Using Equation (2.6), we get:

$$\mathbb{E}[X_1 X_2 \mathbb{1}_{\{X_2 \geq 0\}}] = \int\limits_0^{+\infty} x_2 f_{X_2}(x_2) \left(\mu_1 + \rho \tfrac{\sigma_1}{\sigma_2}(x_2 - \mu_2)\right) dx_2$$

We recall that density of X_2 is $f_{X_2}(x_2) = \frac{c_1}{\sigma_2} g_1 \left(\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)$. Equation (4.4) is obtained by using Lemma 3.2. □

We are now able to give the expression of β^*, and the result is exactly the same we have calculated in Maume-Deschamps et al. (2016), i.e β^* corresponds to the vector of simple Kriging weights. For the sake of simplicity, 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 $\left(X_{1j}, X_2 - \beta^{*T} X_1 - \beta_0^*\right)$ 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^*)**.** *$\forall \alpha \in [0,1]$, the optimal β^* is given by:*

$$\beta^* = \Sigma_{11}^{-1} \Sigma_{12} \tag{4.5}$$

Proof. We recall the minimization problem which verifies (β^*, β_0^*):

$$\underset{\beta \in \mathbb{R}^N, \beta_0 \in \mathbb{R}}{\arg\min} \; (1-\alpha)\mathbb{E}\left[(X_2 - \beta^T X_1 - \beta_0)^2\right] - (1-2\alpha)\mathbb{E}\left[(X_2 - \beta^T X_1 - \beta_0)^2 \mathbb{1}_{\{X_2 - \beta^T X_1 - \beta_0 \geq 0\}}\right]$$

A quick gradient approach gives the following equation system

$$\begin{cases} (1-\alpha)\mathbb{E}\left[X_1(X_2 - \beta^{*T} X_1 - \beta_0^*)\right] + (1-2\alpha)\mathbb{E}\left[X_1(X_2 - \beta^{*T} X_1 - \beta_0^*)\mathbb{1}_{\{X_2 - \beta^{*T} X_1 - \beta_0^* \geq 0\}}\right] & = & 0 \\ (1-\alpha)\mathbb{E}\left[X_2 - \beta^{*T} X_1 - \beta_0^*\right] + (1-2\alpha)\mathbb{E}\left[(X_2 - \beta^{*T} X_1 - \beta_0^*)\mathbb{1}_{\{X_2 - \beta^{*T} X_1 - \beta_0^* \geq 0\}}\right] & = & 0 \end{cases}$$

We know that $X_2 - \beta^{*T} X_1 - \beta_0^*$ is $(R,1)-$elliptical with parameters $\overline{\mu} = \mu_2 - \beta^T \mu_1 - \beta_0$ and $\overline{\Sigma} = \Sigma_{22} - 2\beta^T \Sigma_{12} + \beta^T \Sigma_{11} \beta$. If we denote $\overline{\sigma} = \sqrt{\overline{\Sigma}}$, and use Lemmas 3.2 and 4.1, we get the following system, $\forall j \in \{1, ..., N\}$

$$\begin{cases} (1-\alpha)\overline{\mu} + (1-2\alpha)\overline{\mu}\Phi_R\left(\frac{\overline{\mu}}{\overline{\sigma}}\right) + (1-2\alpha)\overline{\sigma} \int\limits_{\frac{\overline{\mu}}{\overline{\sigma}}}^{+\infty} y c_1 g_1(y^2) dy & = & 0 \\ (1-2\alpha)\left(\mu_{1j}\overline{\mu} - \rho_j \frac{\sigma_{1j}}{\overline{\sigma}} \overline{\mu}^2\right)\Phi_R\left(\frac{\overline{\mu}}{\overline{\sigma}}\right) + (1-2\alpha)\left(\mu_{1j}\overline{\sigma} - \rho_j \sigma_{1j}\overline{\sigma}\right) \int\limits_{\frac{\overline{\mu}}{\overline{\sigma}}}^{+\infty} y c_1 g_1(y^2) dy & & \\ \qquad + (1-2\alpha)\rho_j \frac{\sigma_{1j}}{\overline{\sigma}} \int\limits_0^{+\infty} x_2^2 f_{X_2}(x_2) dx_2 & = & 0 \end{cases}$$

Adding the first equation and the second one, it remains

$$-(1-2\alpha)\rho_j \frac{\sigma_{1j}}{\overline{\sigma}} \overline{\mu}^2 \Phi_R\left(\frac{\overline{\mu}}{\overline{\sigma}}\right) - (1-2\alpha)\rho_j \sigma_{1j}\overline{\sigma} \int\limits_{\frac{\overline{\mu}}{\overline{\sigma}}}^{+\infty} y c_1 g_1(y^2) dy + (1-2\alpha)\rho_j \frac{\sigma_{1j}}{\overline{\sigma}} \int\limits_0^{+\infty} x_2^2 f_{X_2}(x_2) dx_2 + (1-\alpha)\rho_j \sigma_{1j}\overline{\sigma} = 0$$

Obviously, $\rho_j = 0$. We have seen in Maume-Deschamps et al. (2016)(Proposition 4.2), that $\rho_j = 0, \forall j$ was equivalent to $\beta^* = \Sigma_{11}^{-1}\Sigma_{12}$. □

In the following remark, we emphasis the role played by the affine constant β_0^* in the affine model of Equation (4.1). Indeed, considering a linear model without this constant would lead to highly undesirable properties.

Remark 1 (Linear models pitfalls)**.** *Consider a linear expectile regression given by:*

$$\hat{e}_\alpha^{linear}(X_2|X_1 = x_1) = \beta^{*T} x_1 \,,$$

where β^* is the solution of the following minimization problem $\beta^* = \underset{\beta \in \mathbb{R}^N}{\arg\min}\, \mathbb{E}\left[\mathcal{S}_\alpha(X_2 - \beta^T X_1)\right]$, and where the scoring function $\mathcal{S}_\alpha$ is defined as previously by Equation (4.3). This corresponds to the case where $\beta_0^* = 0$. Then

(i) $\hat{e}_\alpha^{linear}(X_2|X_1 = x_1)$ *does not depend on* α*: the linear expectile regression is obviously not suited for other levels than* $\alpha = 1/2$.

(ii) *If* $\Sigma_{12} = 0_{\mathbb{R}^N}$*, which means that* X_2 *and* X_1 *are not correlated, then* $\hat{e}_\alpha^{linear}(X_2|X_1 = x_1) = 0$*, while in the uncorrelated case, we would expect to predict the conditional expectile by the unconditional one* $e_\alpha(X_2)$.

(iii) *Consider an elliptical random field* $(X(t))_{t\in T}$ *and let* $X_1 = (X(t_1), ..., X(t_n)) \in \mathbb{R}^N$, $X_2 = X(t) \in \mathbb{R}$*, so that* Σ_{12} *depends on* t*. Assume that there exists* $t_0 \in \overline{T}$ *such that* $\lim_{t\to t_0} ||\Sigma_{12}|| = 0$*, for a given norm* $||.||$*, then* $\lim_{t\to t_0} \hat{e}_\alpha^{linear}(X_2|X_1 = x_1) = 0$*, while, as before, we would expect* $\lim_{t\to t_0} \hat{e}_\alpha^{linear}(X(t))$.

The result of Remark 1 means that a linear model $\hat{e}_\alpha(X_2|X_1 = x_1) = \beta^{*T} x_1$ will always return the same value, $\forall \alpha \in [0, 1]$. It is a reason why we need the term β_0^*, if we hope approximate $e_\alpha(X_2|X_1 = x_1)$. To calculate it, we give the following lemma.

Lemma 4.3. *$X_2 - \beta^{*T} X_1$ is $(R, 1)-$elliptical with parameters $\mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1$ and $\Sigma_{2|1} = \sigma_2^2 - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.*

Proof. $X_2 - \beta^{*T} X_1$, as 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, $\mathbb{E}[X_2 - \beta^{*T} X_1] = \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1$. The second parameter equals $(-\beta^*, 1)^T \Sigma(-\beta^*, 1)$, thus:

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

or, in matrix form:

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

Thanks to Equation (4.5), we have $\beta^* = \Sigma_{11}^{-1}\Sigma_{12}$. Then the following equation holds:

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

□

With the help of Lemma 4.3, we can give the Expectile Regression Predictor, and its distribution in the following theorem.

Theorem 4.4 (Expectile Regression Predictor)**.**

$$\begin{cases} \beta^* = & \Sigma_{11}^{-1}\Sigma_{12} \\ \beta_0^* = & \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1 + \sigma_{2|1}\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right) \end{cases} \tag{4.6}$$

The Expectile Regression Predictor with level $\alpha \in [0, 1]$ *is given by:*

$$\hat{e}_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\Psi_R^{-1}\left(\frac{\alpha}{2\alpha - 1}\right) \tag{4.7}$$

Furthermore,

$$\hat{e}_\alpha(X_2|X_1) \sim \mathcal{E}_1\left(\mu_2 + \sigma_{2|1}\Psi_R^{-1}\left(\frac{\alpha}{2\alpha - 1}\right), \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}, g_1\right) \tag{4.8}$$

Proof. β^* is given in Equation (4.5). β_0^* satisfies:

$$\beta_0^* = \underset{\beta_0\in\mathbb{R}}{\arg\min}\ \mathbb{E}\left[\mathcal{S}_\alpha(X_2 - \beta^{*T} X_1 - \beta_0)\right]$$

Thus, β_0^* is the $\alpha-$expectile 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 expectile formula of Equation (3.6), we get:

$$\beta_0^* = \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1 + \sigma_{2|1}\Psi_R^{-1}\left(\frac{\alpha}{2\alpha - 1}\right)$$

We can now express our Expectile Regression Predictor of X_2 given $X_1 = x_1$:

$$\hat{e}_\alpha(X_2|X_1 = x_1) = \beta^* x_1 + \beta_0^* = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1) + \sigma_{2|1}\Psi_R^{-1}\left(\frac{\alpha}{2\alpha - 1}\right)$$

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

Since $\hat{e}_\alpha(X_2|X_1) = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) + \sigma_{2|1}\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$, with $X_1 \sim \mathcal{E}_N(\mu_1, \Sigma_{11}, g_N)$. We get, using Proposition 2.4,

$$\hat{e}_\alpha(X_2|X_1) \sim \mathcal{E}_1\left(\mu_2 + \sigma_{2|1}\Psi_R^{-1}\left(\frac{\alpha}{2\alpha-1}\right), (\Sigma_{21}\Sigma_{11}^{-1})\Sigma_{11}(\Sigma_{11}^{-1}\Sigma_{12}), g_1\right)$$

Hence the result. □

We have given in Theorem 4.4 our first expectile predictor. As an illustration, we propose to calculate this predictor in several cases.

4.1. **Examples.** In this subsection, we apply the Expectile Regression Predictor $\hat{e}_\alpha(X_2|X_1)$ given in Theorem 4.4 on several examples, in order to compare it with the theoretical conditional expectiles $e_\alpha(X_2|X_1)$. We recall the formulas:

$$\left\{\begin{array}{ll} e_\alpha(X_2|X_1 = x_1) = & \mu_{2|1} + \sigma_{2|1}\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right) \\ \hat{e}_\alpha(X_2|X_1 = x_1) = & \mu_{2|1} + \sigma_{2|1}\Psi_{R}^{-1}\left(\frac{\alpha}{2\alpha-1}\right) \end{array}\right.$$

Obviously, if $R \neq R^*$, we do not have the same algorithms to compute $e_\alpha(X_2|X_1)$ and $\hat{e}_\alpha(X_2|X_1)$. We have seen in Section 3 that we were able to compute $e_\alpha(X_2|X_1)$ in the Gaussian, Student and Gaussian Mixture cases. In the Laplace and Slash cases, we cannot compute $e_\alpha(X_2|X_1)$, but we are able to compute $\hat{e}_\alpha(X_2|X_1)$. In the following, we thus apply the Expectile Regression Predictor for these last two cases.

4.1.1. *Laplace example.* The Laplace case is interesting, because we have no explicit algorithm to calculate $\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$, since Φ_{R^*} is unknown in this case. However, it is possible to calculate $\Psi_{R}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$. Indeed, the following algorithm converges to this value.

$$\left\{\begin{array}{lcl} e^{(0)} = & & 0 \\ e^{(k+1)} = & & \dfrac{(1-2\alpha)\exp\left(-\sqrt{\frac{2}{\lambda}}e^{(k)}\right)\left(\sqrt{\frac{\lambda}{2}}+e^{(k)}\right)}{(2\alpha-2)-(1-2\alpha)\exp\left(-\sqrt{\frac{2}{\lambda}}e^{(k)}\right)}, \text{ if } e^{(k)} > 0 \\ e^{(k+1)} = & & \dfrac{(1-2\alpha)\exp\left(\sqrt{\frac{2}{\lambda}}e^{(k)}\right)\left(\sqrt{\frac{\lambda}{2}}-e^{(k)}\right)}{(2\alpha-1)\exp\left(\sqrt{\frac{2}{\lambda}}e^{(k)}\right)-2\alpha}, \text{ if } e^{(k)} \leq 0 \end{array}\right. \tag{4.9}$$

4.1.2. *Slash example.* Consider the Slash distribution as in Table 2. As in the Laplace case, we have no explicit algorithm for $\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$, but we can approximate $\Psi_{R}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$:

$$\left\{\begin{array}{lcl} e^{(0)} = & & 0 \\ e^{(k+1)} = & & \dfrac{(1-2\alpha)2^{\frac{a}{2}-1}\frac{a}{a-1}\frac{\Gamma\left(\frac{1+a}{2}\right)}{\sqrt{\pi}}\left(\frac{\chi^2_{1+a}\left(\left(e^{(k)}\right)^2\right)}{\left(e^{(k)}\right)^{a-1}} + \frac{2^{\frac{1-a}{2}}}{\Gamma\left(\frac{1+a}{2}\right)}\exp\left(-\frac{\left(e^{(k)}\right)^2}{2}\right)\right)}{(2\alpha-1)\left[\Phi\left(e^{(k)}\right) - sgn(e^{(k)})\frac{2^{\frac{a}{2}-1}\Gamma\left(\frac{1+a}{2}\right)}{\sqrt{\pi}}\frac{\chi^2_{1+a}\left(\left(e^{(k)}\right)^2\right)}{\left|e^{(k)}\right|^a}\right]-\alpha}, \text{ if } e^{(k)} \neq 0 \\ e^{(k+1)} = & & (2\alpha-1)\frac{a}{a-1}\sqrt{\frac{2}{\pi}}, \text{ if } e^{(k)} = 0 \end{array}\right. \tag{4.10}$$

Numerical applications of this example are proposed in Section 6.

We have seen that, in general, our Expectile Regression Predictor is not equal to the theoretical conditional expectiles. The difference between these quantities is an error term, which may be huge, especially for high levels of α (see Section 6). This is why, in the following section, we propose another predictor for extreme levels of expectiles.

5. Extremal Expectiles

In the previous part, we have proposed some algorithms to calculate the Expectile Regression Predictor. Let us recall:

$$\left\{\begin{array}{ll} e_\alpha(X_2|X_1 = x_1) = & \mu_{2|1} + \sigma_{2|1}\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right) \\ \hat{e}_\alpha(X_2|X_1 = x_1) = & \mu_{2|1} + \sigma_{2|1}\Psi_{R}^{-1}\left(\frac{\alpha}{2\alpha-1}\right) \end{array}\right.$$

Notice that the difference lies in the use of the radius R^* or R. Whereas the distribution of R^* is in general hard to obtain, the one of R may be known or estimated. Our aim is thus to etablish a relation between $\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$ and $\Psi_{R}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$ for extremal values of α, i.e for $\alpha \to 0$ or $\alpha \to 1$. We have done a similar study in Maume-Deschamps et al. (2016), with the same kind of assumption(but for Ψ_R):

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

$$\lim_{x \to +\infty} \frac{\overline{\Psi}_{R^*}(x)}{\overline{\Psi}_R(x^\gamma)} = \ell \,, \tag{5.1}$$

where $\overline{\Psi} = 1 - \Psi$.

We recall $\Psi_R(x) = \Phi_R(x) + \dfrac{1}{x} \displaystyle\int_x^{+\infty} y c_1 g_1(y^2) dy$, and $\Psi_{R^*}(x) = \Phi_{R^*}(x) + \dfrac{1}{x} \displaystyle\int_x^{+\infty} y c_1^* g_{N+1}(q_1 + y^2) dy$, with $c_1^* = \frac{c_{N+1}}{c_N g_N(q_1)}$ (see Equation 2.8). Then, the coefficients γ and ℓ satisfy:

$$\lim_{x \to +\infty} \frac{x^{\gamma - 1} \displaystyle\int_x^{+\infty} y c_1^* g_{N+1}(q_1 + y^2) dy}{\gamma \displaystyle\int_{x^\gamma}^{+\infty} y c_1 g_1(y^2) dy} = \ell \tag{5.2}$$

Proposition 5.1. *Gaussian, Student, Unimodal Gaussian Mixture and Slash distributions satisfy Assumptions 1 and 2 with coefficients γ and ℓ given in Table 3.*

The proof is detailed in Appendix section 8.

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

TABLE 3. Coefficients γ and ℓ for classical consistent elliptical distributions, where $q_1 = (x_1 - \mu_1)^T \Sigma_{11}^{-1} (x_1 - \mu_1)$.

Thanks to Equation 5.2, we have, under Assumption 1, the following equivalence:

$$\overline{\Psi}_{R^*}(x) \underset{x \to +\infty}{\sim} \ell \overline{\Psi}_R(x^\gamma) \tag{5.3}$$

Our aim is now to get an equivalence relationship between functions $\Psi_{R^*}^{-1}$ and $\Psi_{R^*}^{-1}$. For that purpose, we refer to the paper of Djur**c**ić and Torga**s**ev (2001), which gives some conditions to obtain the equivalence of inverse functions if these functions are equivalent.

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

Clearly, our two equivalent functions $\overline{\Psi}_{R^*}(x)$ and $\ell \overline{\Psi}_R(x^\gamma)$ are not $\varphi-$functions for several reasons: $\lim\limits_{x \to 0} \overline{\Psi}_{R^*}(x) = \lim\limits_{x \to 0} \ell \overline{\Psi}_R(x) = -\infty$, $\lim\limits_{x \to +\infty} \overline{\Psi}_{R^*}(x) = \lim\limits_{x \to +\infty} l \overline{\Psi}_R(x^\gamma) = 0^-$. We have to transform these functions in order to find an equivalence. Let us define first some more definitions and properties. The following is the definition of a general class of functions K_c, which contains in particular Regularly Varying functions. The results will thus be more general than those derived from Karamata's theorem.

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

$$\lim_{\substack{x \to +\infty \\ \lambda \to 1}} \frac{f(\lambda x)}{f(x)} = 1 \tag{5.4}$$

In order to inverse the equivalence 5.3, we do the following assumption.

Assumption 2. *Let Ω and Ω_* be*

$$\begin{cases} \Omega(x) &= -\frac{1}{\ell \overline{\Psi}_R(x^\gamma)} \\ \Omega_*(x) &= -\frac{1}{\overline{\Psi}_{R^*}(x)} \end{cases} \tag{5.5}$$

Then Ω^{-1} or Ω_*^{-1} belongs to the class K_c.

Using Definitions 5.1, 5.4, and Assumptions 1, 2, we are able to inverse equivalence 5.3, hence the following proposition.

Proposition 5.2. *Under Assumptions 1 and 2, we have*

$$\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right) \underset{\alpha\to 1}{\sim} \left[\Psi_R^{-1}\left(1-\frac{\alpha-1}{(2\alpha-1)\ell}\right)\right]^{\frac{1}{\gamma}} \tag{5.6}$$

Remark that $\Psi_{R^*}^{-1}\left(1-\frac{\alpha-1}{(2\alpha-1)\ell}\right)$ corresponds to an expectile of level $\frac{(2\alpha-1)\ell+1-\alpha}{(2\alpha-1)\ell+2(1-\alpha)}$. Obviously, if $\ell=1$, this level is α.
We now define two predictors $\hat{\hat{e}}_{\alpha\uparrow}$ and $\hat{\hat{e}}_{\alpha\downarrow}$. Using our several analytical results, we will prove later that they are equivalent to the theoretical expectile for extreme values of α.

Definition 5.3 (Extremal Expectiles Predictors)**.** *Define*

$$\begin{cases} \hat{\hat{e}}_{\alpha\uparrow}(X_2|X_1=x_1) & =\mu_{2|1}+\sigma_{2|1}\left[\Psi_R^{-1}\left(1-\frac{\alpha-1}{(2\alpha-1)\ell}\right)\right]^{\frac{1}{\gamma}} \\ \hat{\hat{e}}_{\alpha\downarrow}(X_2|X_1=x_1) & =\mu_{2|1}-\sigma_{2|1}\left[\Psi_R^{-1}\left(1-\frac{\alpha}{(2\alpha-1)\ell}\right)\right]^{\frac{1}{\gamma}} \end{cases} \tag{5.7}$$

Using Equation (5.6), we are able to prove the asymptotic equivalences between our predictors $\hat{\hat{e}}_{\alpha\uparrow}$, $\hat{\hat{e}}_{\alpha\downarrow}$ and the theoretical expectiles, respectively for $\alpha\to 1$ and $\alpha\to 0$.

Theorem 5.3 (Equivalence with theoretical expectiles)**.** *Under Assumptions 1 and 2,*

$$\begin{cases} \hat{\hat{e}}_{\alpha\uparrow}(X_2|X_1=x_1) & \underset{\alpha\to 1}{\sim} e_\alpha(X_2|X_1=x_1) \\ \hat{\hat{e}}_{\alpha\downarrow}(X_2|X_1=x_1) & \underset{\alpha\to 0}{\sim} e_\alpha(X_2|X_1=x_1) \end{cases} \tag{5.8}$$

Proof. We recall the relationship $e_\alpha(X_2|X_1=x_1)=\mu_{2|1}+\sigma_{2|1}\Psi_{R^*}^{-1}\left(\frac{\alpha}{2\alpha-1}\right)$. With Equation (5.6), the first equivalence is immediate. Concerning the second one, i.e when $\alpha\to 0$, we use the symmetry properties of elliptical distributions: we know that $e_\alpha(X_2|X_1=x_1)=\mu_{2|1}-\sigma_{2|1}\Psi_{R^*}^{-1}\left(\frac{1-\alpha}{1-2\alpha}\right)$. Or if $\alpha'=1-\alpha$, we have $e_\alpha(X_2|X_1=x_1)=\mu_{2|1}-\sigma_{2|1}\Psi_{R^*}^{-1}\left(\frac{\alpha'}{2\alpha'-1}\right)$, with $\alpha'\to 1$. Using Equation (5.6), we get

$$e_\alpha(X_2|X_1=x_1) \underset{\alpha'\to 1}{\sim} \mu_{2|1}-\sigma_{2|1}\left[\Psi_R^{-1}\left(1-\frac{\alpha'-1}{(2\alpha'-1)\ell}\right)\right]^{\frac{1}{\gamma}}.$$

Replace α' by $1-\alpha$ to the second equivalence. □

From Theorem 5.3, we can deduce that $\hat{\hat{e}}_{\alpha\uparrow}$, $\hat{\hat{e}}_{\alpha\downarrow}$ are approximations of e_α, respectively when α is close to 1 or 0. We propose, in Appendix section 8, to check if Assumptions 1 and 2 are fulfilled, and the calculation of ℓ and γ for each example. In Section 6, we propose some numerical applications and graphical illustrations. Let us now consider some numerical illustrations on the examples we have given, in order to compare our extremal predictors with the Expectile Regression Predictor and the theoretical expectiles.

6. Numerical study

In order to give a visual overview of the predictors, we have plotted in Figure 1, the conditional expectiles 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 predict the expectile 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 would be applicable without these assumptions. Parameters for the Gaussian Mixture example are $\theta_1=1$, $\theta_2=2$ and $p=0.9$.

Of course, for the Gaussian process, the curves coincide. For the other examples, Expectile Regression Predictors seem very far from the theoretical curves, especially in the Slash case. On the other hand, Extremal predictors perform significantly better; they look closer to the target conditional expectiles here. We propose to use the following $RMSE$ in order to quantify the difference between expectile regression and theoretical expectile.

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

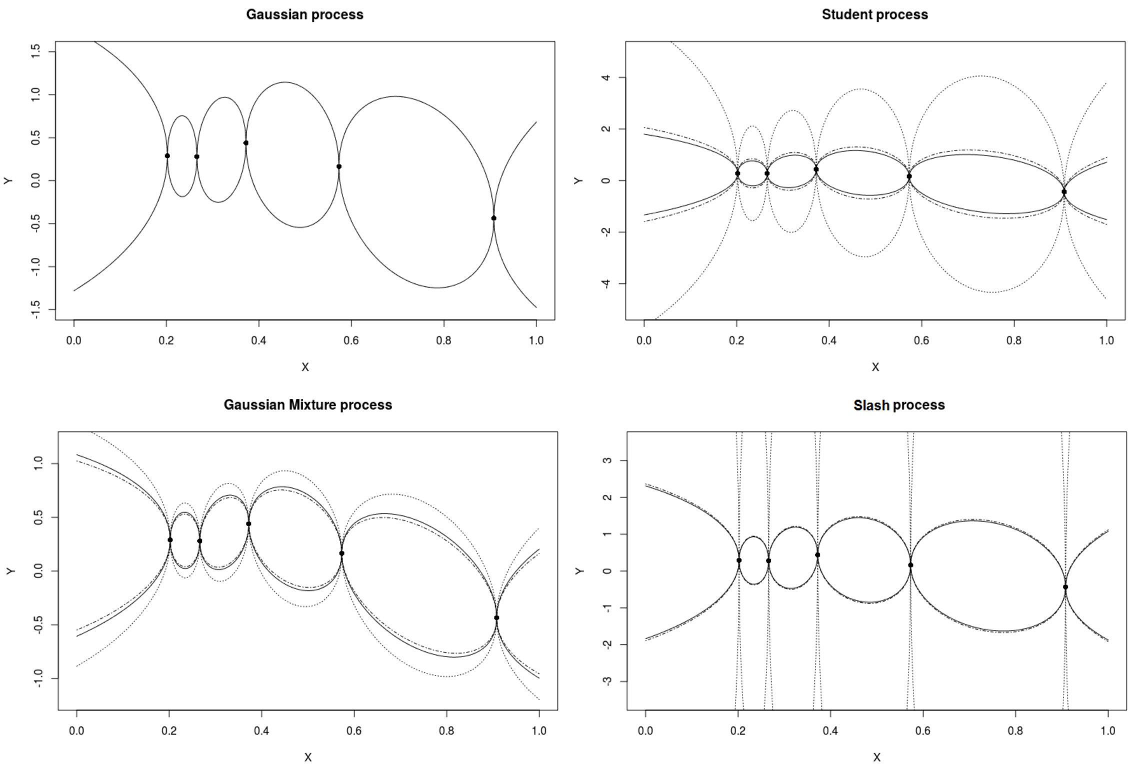

Figure 1. Expectile Regression Predictor dotted, theoretical expectiles in solid lines, and Extremal Expectile Predictors dashed, for levels $\alpha = 0.9995$ and 0.0005

We also consider the $RMSE$ for the Extremal Predictor:

$$RMSE(\hat{\hat{e}}_\alpha) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\Big(e_\alpha(X_2^{(i)}|X_1 = x_1) - \hat{\hat{e}}_\alpha(X_2^{(i)}|X_1 = x_1)\Big)^2}. \tag{6.2}$$

The RMSE measures the average error in the prediction of the conditional expectiles. 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.

α	Gaussian		Student		Unimodal GM		Slash	
	$r(\hat{e}_\alpha)$	$r(\hat{\hat{e}}_\alpha)$	$r(\hat{e}_\alpha)$	$r(\hat{\hat{e}}_\alpha)$	$r(\hat{e}_\alpha)$	$r(\hat{\hat{e}}_\alpha)$	$r(\hat{e}_\alpha)$	$r(\hat{\hat{e}}_\alpha)$
0.5	0	0	0	0	0	0	0	0
0.6	0	0	0.031	0.275	0.002	0.021	0.042	0.294
0.7	0	0	0.066	0.293	0.005	0.041	0.091	0.284
0.8	0	0	0.112	0.290	0.008	0.062	0.160	0.255
0.9	0	0	0.197	0.272	0.014	0.083	0.302	0.205
0.95	0	0	0.300	0.252	0.021	0.091	0.513	0.162
0.995	0	0	0.913	0.190	0.069	0.078	2.368	0.066
0.9995	0	0	2.384	0.142	0.153	0.032	8.840	0.031
0.999995	0	0	13.432	0.079	0.152	$8.089 \cdot 10^{-06}$	97.567	0.015

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

As expected, the error of $\hat{e}_\alpha$ is increasing with α, and may be huge, like in the Slash case. On the other hand, the error of $\hat{\hat{e}}_\alpha$ tends to 0 when α tends to 1. This trend is highlighted by the $ee-$plots (theoretical expectile vs expectile prediction plots) in Figure 2.

7. Conclusion

In this paper, we focused on conditional expectiles prediction, for elliptical random fields with the consistency property. We have shown that theoretical expectiles were relying on a radius R^* whose distribution was,

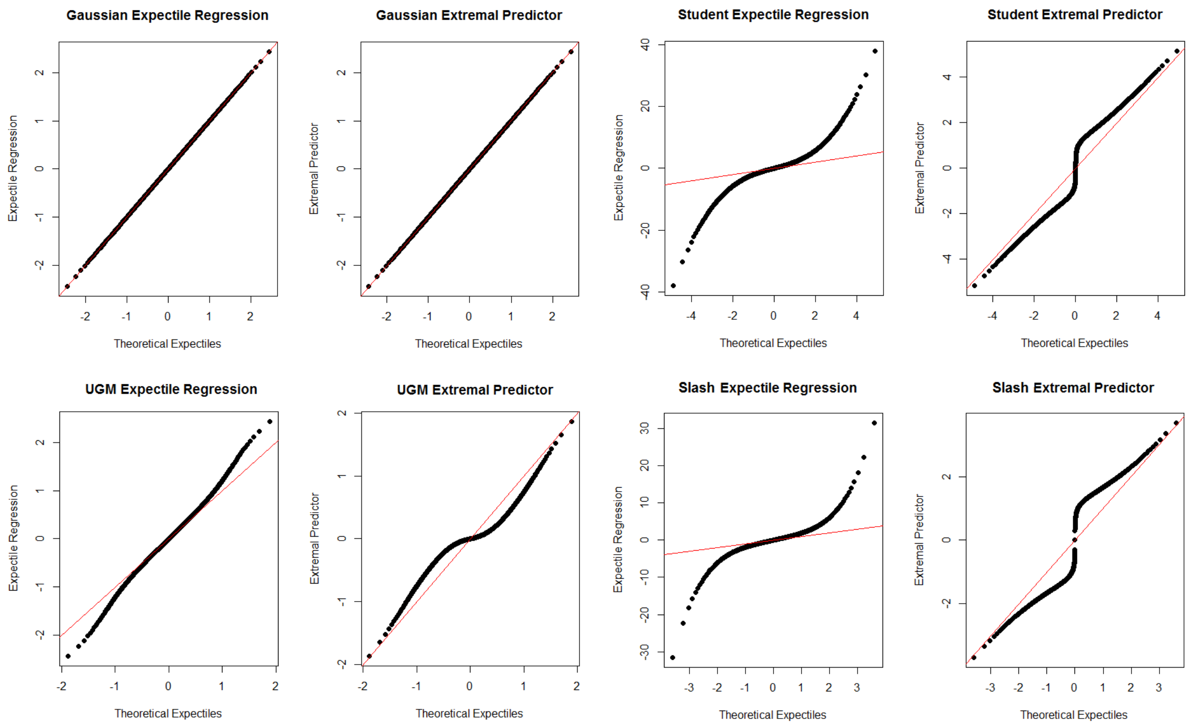

Figure 2. Theoretical expectiles vs expectile predictions for Gaussian, Student, UGM and Slash examples. Expectile regression appears on the left of each pair of panels, extremal predictor on the right.

in the general case, difficult to obtain. We thus have proposed two different methods to predict conditional expectiles.
The first one is to use expectile regression, i.e to express the conditional expectile as an 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 expectile (when $\alpha \to 0$ or $\alpha \to 1$). We have seen, in a first time, that we can obtain some iterative algorithms in our case of consistent elliptical random fields. Furthermore, we have given the distribution of the expectile regression (Theorem 4.4). We have seen that expectile regression is not suited to non Gaussian distributions.
A second predictor is given in order to cope with expectile regression problems for extremal expectile levels. We have shown that the proposed extremal expectile predictor is equivalent to the true conditional expectile for extreme expectile levels. We have also illustrated on several numerical examples the better performance of this predictor for extreme levels.
As a perspective, these prediction methods require the knowledge of the distribution of the covariates vector X_1. We have not explored the prediction procedure when the X_1's distribution is estimated (parametrically e.g.).
Finally, we would like to emphasize that we have given examples in dimension $d = 1$, but all the results may be used in higher dimensions.

8. Appendix

Proof of Proposition 5.1.

Gaussian example. Firstly, we have to calculate the limit in Equation (5.2), with $g_n(t) = \exp\left(-\frac{t}{2}\right), \forall n \in \mathbb{N}$, $c_1 = \frac{1}{\sqrt{2\pi}}$ and $c_1^* = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{q_1}{2}\right)$. Then, we consider

$$\lim_{x \to +\infty} \frac{x^{\gamma-1} \displaystyle\int_x^{+\infty} y \exp\left(-\frac{y^2}{2}\right) dy}{\gamma \displaystyle\int_{x^\gamma}^{+\infty} y \exp\left(-\frac{y^2}{2}\right) dy}.$$

If we take $\gamma = 1$, we directly get the limit ℓ equal to 1.
Now, we have to prove that Ω^{-1} belongs to the class K_c. For that purpose, we consider the limit

$$\lim_{\substack{x\to+\infty\\ \lambda\to 1}} \frac{\Omega^{-1}(\lambda x)}{\Omega^{-1}(x)}.$$

We have previously seen that $\Omega^{-1}(x) = \left[\Psi_R^{-1}\left(1+\frac{1}{\ell x}\right)\right]^{\frac{1}{\gamma}}$. Then we have to consider the limit of $\frac{\Psi_R^{-1}\left(1+\frac{1}{\ell\lambda x}\right)}{\Psi_R^{-1}\left(1+\frac{1}{\ell x}\right)}$. With $\delta = \lambda - 1$ and $y = \frac{1}{x}$, this limit becomes

$$\lim_{(y,\delta)\to(0,0)} \frac{\Psi_R^{-1}\left(1+\frac{y}{\ell(\delta+1)}\right)}{\Psi_R^{-1}\left(1+\frac{y}{\ell}\right)}.$$

In order to calculate this kind of limit, we move to polar coordinates, i.e we take $\delta = r\cos(\theta)$, $y = r\sin(\theta)$, and verify whether the limit when $r \to 0$ exists and is not related to θ. Hence

$$f(\theta) = \lim_{r\to 0} \frac{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)}{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)}$$

The two terms both tend to 0, then we can use the l'Hôpital's rule, and

$$f(\theta) = \lim_{r\to 0} \frac{\frac{\sin(\theta)(r\cos(\theta)+1)-r\sin(\theta)\cos(\theta)}{(r\cos(\theta)+1)^2}}{\sin(\theta)} \frac{\Psi_R'\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)\right)}{\Psi_R'\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)\right)}$$

The ratio on the left clearly tends to 1. Furthermore, we have seen in the proof of Lemma 3.1 that $\Psi_R'(x) = -\frac{1}{x^2}\int_x^{+\infty} yc_1g_1(y^2)dy$. In the Gaussian case, this value is equal to $-\frac{1}{x^2}\varphi(x)$, and $\ell = 1$. Then, we get

$$f(\theta) = \lim_{r\to 0} \frac{\left[\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{r\cos(\theta)+1}\right)\right]^2}{\left[\Psi_R^{-1}\left(1+r\sin(\theta)\right)\right]^2} \frac{\varphi\left(\Psi_R^{-1}\left(1+r\sin(\theta)\right)\right)}{\varphi\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{r\cos(\theta)+1}\right)\right)}$$

The ratio on the left is clearly equal to $f(\theta)^2$. Concerning the term on the right, we apply once again the l'Hôpital's rule. Since $\varphi'(x) = -x\varphi(x)$, it leds to

$$f(\theta) = f(\theta)^2 \lim_{r\to 0} \frac{\left[\Psi_R^{-1}\left(1+r\sin(\theta)\right)\right]^3}{\left[\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{r\cos(\theta)+1}\right)\right]^3} = \frac{1}{f(\theta)} = 1$$

Student example. At a first stage, we calculate the limit in Equation (5.2) with $g_n(t) = \left(1+\frac{t}{\nu}\right)^{-\frac{n+\nu}{2}}, \forall n \in \mathbb{N}$, $c_1 = \frac{\Gamma\left(\frac{1+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{1}{\sqrt{\nu\pi}}$ and $c_1^* = \frac{\Gamma\left(\frac{N+1+\nu}{2}\right)}{\Gamma\left(\frac{N+\nu}{2}\right)}\frac{1}{\sqrt{\nu\pi}}\left(1+\frac{q_1}{\nu}\right)^{\frac{N+\nu}{2}}$. It remains

$$\lim_{x\to+\infty} \frac{x^{\gamma-1}\int_x^{+\infty} yc_1^*\left(1+\frac{q_1+y^2}{\nu}\right)^{-\frac{N+1+\nu}{2}} dy}{\gamma\int_{x^\gamma}^{+\infty} yc_1\left(1+\frac{y^2}{\nu}\right)^{-\frac{1+\nu}{2}} dy} = \ell.$$

The two functions under the integrals are polynomials, then easy to integrate, hence

$$\lim_{x\to+\infty} \frac{x^{\gamma-1}c_1^*\frac{\nu}{\nu+N-1}\left(1+\frac{q_1+x^2}{\nu}\right)^{\frac{1-N-\nu}{2}}}{\gamma c_1\frac{\nu}{\nu-1}\left(1+\frac{x^{2\gamma}}{\nu}\right)^{\frac{1-\nu}{2}}} = \ell.$$

Then, if we take $\gamma = \frac{N+\nu}{\nu}$, and replace c_1 and c_1^* by their values, we get

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

Now, we have to check if Assumption 2 is fulfilled, i.e if the function Ω^{-1} belongs to the class K_c. Let us calculate the limit

$$\lim_{\substack{x\to+\infty\\ \lambda\to 1}} \frac{\Omega^{-1}(\lambda x)}{\Omega^{-1}(x)}.$$

Using the same changes of variables as in the Gaussian case, we consider

$$f(\theta) = \lim_{r\to 0} \frac{\Psi_R'\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)\right)}{\Psi_R'\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)\right)}$$

In the Student case, $\Psi_R(x)' = -\frac{1}{x^2}\frac{\Gamma\left(\frac{1+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{1}{\sqrt{\nu\pi}}\frac{\nu}{\nu-1}\left(1+\frac{x^2}{\nu}\right)^{\frac{1-\nu}{2}}$. Then, it only remains

$$f(\theta) = \lim_{r\to 0} \frac{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)^2}{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)^2}\left(\frac{1+\frac{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)^2}{\nu}}{1+\frac{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)^2}{\nu}}\right)^{\frac{1-\nu}{2}}$$

On the left, the ratio tends to $f(\theta)^2$. The term on the right equals $\lim_{r\to 0}\left(\frac{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)^2}{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)^2}\right)^{\frac{1-\nu}{2}} = f(\theta)^{\nu-1}$. Finally, we get the relationship

$$f(\theta) = f(\theta)^{\nu+1}, \nu > 0$$

Hence $f(\theta) = f(\theta)^{\frac{1}{\gamma}} = 1$.

Unimodal Gaussian Mixture example. Let us calculate the limit in Equation (5.2). In the Unimodal GM case, $g_d(t) = \sum_{k=1}^{n} \pi_k \theta_k^d \exp\left(-\frac{\theta_k^2}{2}t\right)$, $c_1 = \frac{1}{\sqrt{2\pi}}$ and $c_1^* = \frac{1}{\sqrt{2\pi}\sum_{k=1}^{n}\pi_k\theta_k^N \exp\left(-\frac{\theta_k^2}{2}q_1\right)}$. The calculation of limit (5.2) gives

$$\lim_{x\to+\infty} \frac{x^{\gamma-1}\sum_{k=1}^{n}\theta_k^{N-1}c_1^*\exp\left(-\frac{\theta_k^2}{2}q_1\right)\exp\left(-\frac{\theta_k^2}{2}x^2\right)}{\gamma\sum_{k=1}^{n}\pi_k\frac{1}{\theta_k}c_1\exp\left(-\frac{\theta_k^2}{2}x^{2\gamma}\right)}.$$

From now, let us consider that $\gamma = 1$. Asymptotically, we only consider the terms $\exp\left(-\frac{\theta_k^2}{2}x^2\right)$ with the highest coefficient $-\frac{\theta_k^2}{2}$, i.e the smallest θ_k. Let k^* such that $\theta_{k^*} = \min\{\theta_1, ..., \theta_n\}$. We have

$$\ell = \lim_{x\to+\infty}\frac{c_1^*\pi_{k^*}\theta_{k^*}^{N-1}\exp\left(-\frac{\theta_{k^*}}{2}q_1\right)\exp\left(-\frac{\theta_{k^*}}{2}x^2\right)}{c_1\pi_{k^*}\theta_{k^*}^{-1}\exp\left(-\frac{\theta_{k^*}}{2}x^2\right)}.$$

After simplifications, and replacing c_1 and c_1^* by their values, we get

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

We have just seen that Assumption 1 is satisfied. We have to check now if Assumption 2 is fulfilled. As usual, we consider the limit

$$f(\theta) = \lim_{r\to 0} \frac{\Psi_R'\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)\right)}{\Psi_R'\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)\right)}$$

In the UGM case, we have the relationship $\Psi_R'(x) = -\frac{1}{x^2}\sum_{k=1}^{n}\frac{\pi_k}{\theta_k}\varphi(\theta_k x)$. Rewriting $\Psi_R'(x)$ in the previous limit, we get

$$f(\theta) = \lim_{r\to 0}\frac{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)^2}{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)^2}\frac{\sum_{k=1}^{n}\frac{\pi_k}{\theta_k}\varphi\left(\theta_k\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)\right)}{\sum_{k=1}^{n}\frac{\pi_k}{\theta_k}\varphi\left(\theta_k\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)\right)}.$$

The term on the left is equal to $f(\theta)^2$. In order to calculate the limit of the second ratio, we consider only the leading terms in the numerator and denominator, i.e the terms with the biggest $-\frac{\theta_k^2}{2}$, or the smallest θ_k. Let k^* be such that $\theta_{k^*} = \min\{\theta_1, ..., \theta_n\}$. It remains

$$f(\theta) = f(\theta)^2\lim_{r\to 0}\frac{\varphi\left(\theta_{k^*}\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)\right)}{\varphi\left(\theta_{k^*}\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)\right)}.$$

We have calculated, in the Gaussian case, this kind of limit, and proved that it was equal to 1.

Slash example. We consider the limit in Equation (5.2), with $g_n(t) = \frac{\chi^2_{n+a}(t)}{t^{\frac{n+a}{2}}}, \forall n \in \mathbb{N}^*$, $c_1^* = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{N+1+a}{2}\right)}{\Gamma\left(\frac{N+a}{2}\right)} \frac{q_1^{\frac{N+a}{2}}}{\chi^2_{N+a}(q_1)}$ and $c_1 = \frac{2^{\frac{a}{2}-1} a \Gamma\left(\frac{1+a}{2}\right)}{\sqrt{\pi}}$. Using integrations by parts, we get

$$\ell = \lim_{x\to+\infty} \frac{x^{\gamma-1} \frac{c_1^*}{N+a-1} \left[\left(q_1+x^2\right)^{\frac{1-N-a}{2}} \chi^2_{N+1+a}\left(q_1+x^2\right) + \frac{2^{\frac{1-N-a}{2}} \exp\left(-\frac{q_1+x^2}{2}\right)}{\Gamma\left(\frac{N+1+a}{2}\right)} \right]}{\gamma \frac{c_1}{a-1} \left[x^{\gamma(1-a)} \chi^2_{1+a}(x^{2\gamma}) + \frac{2^{\frac{1-a}{2}} \exp\left(-\frac{x^2}{2}\right)}{\Gamma\left(\frac{1+a}{2}\right)} \right]}.$$

Asymptotically, it only remains

$$\ell = \lim_{x\to+\infty} \frac{(a-1)x^{\gamma-1} c_1^* x^{1-N-a}}{\gamma c_1 (N+a-1) x^{\gamma(1-a)}}.$$

Using $\gamma = \frac{N}{a}+1$, and replacing c_1^* and c_1 by their values, we directly obtain the value of ℓ given in Table 3. Concerning Assumption 2, we consider

$$f(\theta) = \lim_{r\to 0} \frac{\Psi'_R\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)\right)}{\Psi'_R\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)\right)}.$$

In the Slash case, $\Psi'_R(x) = -\frac{1}{x^2} 2^{\frac{a}{2}-1} \frac{a}{a-1} \frac{\Gamma\left(\frac{1+a}{2}\right)}{\sqrt{\pi}} \left(\frac{\chi^2_{1+a}(x^2)}{|x|^{a-1}} + \frac{2^{\frac{1-a}{2}}}{\Gamma\left(\frac{1+a}{2}\right)} \sqrt{2\pi}\varphi(x) \right)$. Hence the limit

$$f(\theta) = \lim_{r\to 0} \frac{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)^2}{\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)^2} \frac{\frac{\chi^2_{1+a}\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)^2\right)}{\left|\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)\right|^{a-1}} + \frac{2^{\frac{1-a}{2}}}{\Gamma\left(\frac{1+a}{2}\right)} \sqrt{2\pi}\varphi\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)\right)}{\frac{\chi^2_{1+a}\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)^2\right)}{\left|\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)\right|^{a-1}} + \frac{2^{\frac{1-a}{2}}}{\Gamma\left(\frac{1+a}{2}\right)} \sqrt{2\pi}\varphi\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)\right)}$$

The ratio on the left is obvously equal to $f(\theta)^2$. Concerning the term on the right, the ratios $\frac{\chi^2_{1+a}(x)}{x}$ clearly tends to 0 when $x \to \infty$. Then, the limit may be written more easily

$$f(\theta) = f(\theta)^2 \lim_{r\to 0} \frac{\varphi\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell}\right)\right)}{\varphi\left(\Psi_R^{-1}\left(1+\frac{r\sin(\theta)}{\ell(r\cos(\theta)+1)}\right)\right)}$$

We have already calculated this limit in the Gaussian case, and seen the relationship $f(\theta) = \frac{1}{f(\theta)} = 1$, hence Assumption 2.

Proof of Proposition 5.2. We firstly introduce two lemmas.

Lemma 8.1 (Djurcić and Torgasev (2001))**.** *Suppose that f and g are two strictly increasing $\varphi-$functions, and that at least one of the functions f^{-1}, g^{-1} belongs to the class K_c, and $f(x) \underset{x\to\infty}{\sim} g(x)$. Then $f^{-1}(x) \underset{x\to\infty}{\sim} g^{-1}(x)$*

Lemma 8.2. *If $\gamma > 0$, then Ω and Ω_* are $\varphi-$functions. Furthermore, under Assumption 1, we have*

$$\Omega(x) \underset{x\to\infty}{\sim} \Omega_*(x) \tag{8.1}$$

The proof of Lemma 8.2 is straightforward, using properties given in Lemma 3.1. Using these two lemmas, we can now give the proof of Proposition 5.2.

Proof. By quick calculations, we get $\Omega^{-1}(x) = \left[\Psi_R^{-1}\left(1+\frac{1}{\ell x}\right)\right]^{\frac{1}{\gamma}}$ and $\Omega_*^{-1}(x) = \Psi_{R^*}^{-1}\left(1+\frac{1}{x}\right)$. Thanks to Assumption 1, we have $\Omega(x) \underset{x\to\infty}{\sim} \Omega_*(x)$ (Lemma 8.2), with Ω and Ω_* $\varphi-$functions. With Assumption 2, Ω^{-1} or Ω_*^{-1} belongs to the class K_c. Then, we can apply Lemma 8.1. Hence the equivalence $\Omega^{-1}(x) \underset{x\to\infty}{\sim} \Omega_*^{-1}(x)$. In other words,

$$\Psi_{R^*}^{-1}\left(1+\frac{1}{x}\right) \underset{x\to\infty}{\sim} \left[\Psi_R^{-1}\left(1+\frac{1}{\ell x}\right)\right]^{\frac{1}{\gamma}}$$

If we do the change of variable $1+\frac{1}{x} = \frac{\alpha}{2\alpha-1}$, we get the result (5.6). □

References

Abramowitz, M., Stegun, I. A., et al. (1966). Handbook of mathematical functions. *Applied mathematics series*, 55:62.

Bellini, F., Klar, B., Müller, A., and Rosazza Gianin, E. (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, 54:41–48.

Cambanis, S., Huang, S., and Simons, G. (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, (11):368–385.

Djurcić, D. and Torgasev, A. (2001). Strong Asymptotic Equivalence and Inversion of Functions in the Class *Kc*. *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.

Fraley, C. and Raftery, A. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458):611–630.

Frontini, M. and Sormani, E. (2003). Some variant of Newton′s method with third-order convergence. *Applied Mathematics and Computation*, 140:419–426.

Gómez, H. W., Quintana, F. A., and Torres, F. J. (2007). A new family of slash-distributions with elliptical contours. *Statistics & Probability Letters*, 77:717–725.

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. (2004). A Tutorial on MM Algorithms. *The American Statistician*, 58(1):30–37.

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., Podgórskib, and K., 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).

Maume-Deschamps, V., Rullière, D., and Usseglio-Carleve, A. (2016). Spatial Quantile Predictions for Elliptical Random Fields. preprint.

Nadarajah, S. (2003). The Kotz Type Distribution with applications. *Statistics*, 37(4):341–358.

Nadarajah, S. and Kotz, S. (2004). *Multivariate T-Distributions and Their Applications.* Cambridge University Press.

Newey, W. and Powell, J. (1987). Asymmetric Least Squares Estimation and Testing. *Econometrica*, 55:819–847.

Sobotka, F. and Kneib, T. (2012). Geoadditive expectile regression. *Computational Statistics and Data Analysis*, 56:755–767.

Yang, Y., Zhang, T., and Zou, H. (2015). Flexible Expectile Regression in Reproducing Kernel Hilbert Space. preprint.

Ziegel, J. (2014). Coherence and ellicitability. *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: usseglio@math.univ-lyon1.fr