Intriguing properties of extreme geometric quantiles

Stéphane Girard, Gilles Stupfler

Intriguing properties of extreme geometric quantiles

Stéphane Girard[(1)] & Gilles Stupfler[(2)]

[(1)] Team Mistis, INRIA Rhône-Alpes & LJK, Inovallée, 655, av. de l'Europe,
Montbonnot, 38334 Saint-Ismier cedex, France
[(2)] Aix Marseille Université, CERGAM, EA 4225,
15-19 allée Claude Forbin, 13628 Aix-en-Provence Cedex 1, France

Abstract. A popular way to study the tail of a distribution function is to consider its high or extreme quantiles. While this is a standard procedure for univariate distributions, it is harder for multivariate ones, primarily because there is no universally accepted definition of what a multivariate quantile should be. In this paper, we focus on extreme geometric quantiles. Their asymptotics are established, both in direction and magnitude, under suitable integrability conditions, when the norm of the associated index vector tends to one. In particular, it appears that if a random vector has a finite covariance matrix, then the magnitude of its extreme geometric quantiles grows at a fixed rate which is independent of the asymptotic behaviour of the underlying probability distribution. Moreover, in the special case of elliptically contoured distributions, the respective shapes of the contour plots of extreme geometric quantiles and extreme level sets of the probability density function are orthogonal, in some sense. These phenomena are illustrated on some numerical examples.

AMS Subject Classifications: 62H05, 62G20, 62G32.

Keywords: Extreme quantile, geometric quantile, asymptotic behaviour.

1 Introduction

Let X be a random vector in $\mathbb{R}^d$. Up to now, several definitions of multivariate quantiles of X have been introduced in the statistical literature. We refer to Serfling (2002) for a review of various possibilities for this notion. Here, we focus on the notion of "spatial" or "geometric" quantiles, introduced by Chaudhuri (1996), which generalises the characterisation of a univariate quantile shown in Koenker and Bassett (1978). For a given vector u belonging to the unit open ball B^d of $\mathbb{R}^d$, where $d \geq 2$, the geometric quantile related to u is a solution of the optimisation problem (P) defined by

$$\arg\min_{q\in\mathbb{R}^d} \mathbb{E}(\phi(u, X-q) - \phi(u, X)), \tag{1}$$

with ϕ the loss function

$$\phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, \ (u, t) \mapsto \|t\| + \langle u, t \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the usual scalar product on $\mathbb{R}^d$ and $\|\cdot\|$ is the associated Euclidean norm. Any solution $q(u)$ of the problem (P) is called a u–th quantile. Note that $q(u) \in \mathbb{R}^d$ possesses both a direction and magnitude. It can be seen that geometric quantiles are in fact special cases of M–quantiles introduced by Breckling and Chambers (1988) which were further analysed by Koltchinskii (1997). Besides, such quantiles have various strong properties. First, the quantile related to a vector $u \in B^d$ is unique whenever the distribution of X is not concentrated on a single straight line in $\mathbb{R}^d$ (see Chaudhuri, 1996, or Theorem 2.17 in Kemperman, 1987). Second, although they are not fully affine equivariant, they are invariant under any orthogonal transformation (Chaudhuri, 1996). Third, geometric quantiles characterise the associated distribution. Namely, if two random variables X and Y yield the same quantile function q, then X and Y have the same distribution (Koltchinskii, 1997). Finally, for $u = 0$, the well-known geometric median is obtained, which is the simplest example of a "central" quantile (see Small, 1990). We point out that one may compute an estimation of the geometric median in an efficient way, see Cardot *et al.* (2013).

These properties make geometric quantiles reasonable candidates when trying to define multivariate quantiles, which is why their estimation was studied in several papers. We refer for instance to Chakraborty (2001), for the introduction of a transformation-retransformation procedure to obtain affine equivariant estimates of multivariate quantiles, to Chakraborty (2003) for a generalisation of this notion to a multiresponse linear model and to Dhar *et al.* (2013) for the definition of a multivariate quantile-quantile plot using geometric quantiles. Conditional geometric quantiles can also be defined by substituting a conditional expectation to the expectation in (1). We refer to Cadre and Gannoun (2000) for the estimation of the conditional geometric median and to Cheng and de Gooijer (2007) for the estimation of an arbitrary conditional geometric quantile. The estimation of a conditional median when there is an infinite-dimensional covariate is considered in Chaouch and Laïb (2013).

Our focus in this paper is rather on extreme geometric quantiles, obtained when $\|u\| \to 1$. The theory of univariate extreme quantiles is well established, see for instance the monograph by de Haan and Ferreira (2006). On the contrary, the few works on extreme multivariate quantiles rely on the study of extreme level sets of the probability density function of X when it is absolutely continuous with respect to the Lebesgue measure. We refer for instance to Cai *et al.* (2011) for an application to the estimation of extreme risk regions for financial data or to Einmahl *et al.* (2013) who focus on the case of bivariate distributions with an application to insurance data. In Daouia *et al.* (2013), the estimation of conditional extreme quantiles is addressed, the case of a functional covariate being considered in Gardes and Girard (2012).

In this study, we provide an equivalent of the direction and magnitude of the extreme geometric quantile $q(u)$, $\|u\| \to 1$ under suitable integrability conditions. As a corollary of our results, it appears that the extreme geometric quantiles of any two distributions having the same finite expectation and covariance matrix have equivalent direction and magnitude, regardless of the respective other characteristics of these distributions. Especially, if the distribution of X is absolutely continuous with respect to the Lebesgue measure, then the magnitude of an extreme geometric quantile does not take into account the asymptotic behaviour of the probability density function of X. Besides, the shape of the contour plots of extreme geometric quantiles is uniquely determined by the covariance matrix of the underlying distribution. In the special case of elliptically contoured distributions (see Cambanis *et al.*, 1981), one may see that the respective shapes of the contour plots of extreme geometric quantiles and extreme level sets of the probability density function are even orthogonal, in some sense. Consequently, one should be cautious when using only extreme geometric quantiles as a graphical tool for analysing multidimensional datasets, particularly when trying to detect outliers of a data cloud (Chaouch and Goga, 2010).

The outline of the paper is as follows. The main results are stated in Section 2. Some examples and illustrations on the asymptotic behaviour of extreme geometric quantiles are presented in Section 3. Proofs are deferred to Section 4.

2 Main results

From now on, we assume that the distribution of X is not concentrated on a single straight line in $\mathbb{R}^d$. We shall rewrite the optimisation problem (P) as

$$\arg\min_{q\in\mathbb{R}^d} \psi(u,q)$$

where

$$\psi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R},\ (u,q) \mapsto \mathbb{E}(\phi(u, X-q) - \phi(u,X)).$$

Chaudhuri (1996) proved that in this context, the solution $q(u)$ of (P) is unique for every $u \in B^d$. The vector $q(u)$ is called the geometric quantile of X associated with u. Define further that $t/\|t\| = 0$ if $t = 0$; if $u \in \mathbb{R}^d$ is such that there is a solution $q(u) \in \mathbb{R}^d$ to problem (P), then the gradient of $q \mapsto \psi(u,q)$ must be zero at $q(u)$, that is

$$u + \mathbb{E}\left(\frac{X-q(u)}{\|X-q(u)\|}\right) = 0. \tag{2}$$

This condition immediately entails that if $u \in \mathbb{R}^d$ is such that problem (P) has a solution $q(u)$, then $\|u\| \le 1$. In fact, we can prove a stronger result:

Proposition 1. *The optimisation problem (P) has a solution if and only if $u \in B^d$.*

Moreover, remarking that for every $u \in B^d$, the function $\varphi(u,\cdot)$ is strictly convex, Chaudhuri (1996) proved the following characterisation of a geometric quantile: for every $u \in B^d$, $q(u)$ is the solution of problem (P) if and only if (2) holds. In particular, this entails that the function $G : \mathbb{R}^d \to B^d$ defined by

$$\forall q \in \mathbb{R}^d,\ G(q) = \mathbb{E}\left(\frac{X-q}{\|X-q\|}\right)$$

is a continuous bijection. Inverting the function G, we obtain that the geometric quantile function $u \mapsto q(u)$ is continuous on B^d.

In most cases however, computing explicitly the function G is a hopeless task, which makes it impossible to obtain a closed-form expression for the geometric quantile function. It is thus of interest to prove general results about the geometric quantile $q(u)$, especially regarding its direction and magnitude. Our first main result focuses on the special case of spherically symmetric distributions.

Proposition 2. *If X has a spherically symmetric distribution then:*

(i) The map $u \mapsto q(u)$ commutes with every linear isometry of $\mathbb{R}^d$. Especially, the norm of a geometric quantile $q(u)$ of a spherically symmetric distribution only depends on the norm of u.

(ii) For all $u \in B^d$, the geometric quantile $q(u)$ and u are collinear:

$$\frac{q(u)}{\|q(u)\|} = \frac{u}{\|u\|}$$

if $u \neq 0$ and $q(0) = 0$ otherwise.

(iii) The function $\|u\| \mapsto \|q(u)\|$ is a continuous increasing function on $[0, 1)$.

(iv) It holds that $\|q(u)\| \to \infty$ as $\|u\| \to 1$.

Although the first and third statement of Proposition 2 cannot be expected to hold true for a random variable which is not spherically symmetric, one may wonder if the second and fourth statement, namely that a geometric quantile shares the direction of its index vector and that the norm of the geometric quantile function tends to infinity on the unit sphere, can be extended to the general case. The next result, which examines the behaviour of the geometric quantile function near the boundary of the open ball B^d, provides an answer to this question.

Theorem 1. *Let S^{d-1} be the unit sphere of $\mathbb{R}^d$.*

(i) It holds that $\|q(v)\| \to \infty$ as $\|v\| \to 1$.

(ii) Moreover, if $v \to u$ with $u \in S^{d-1}$ and $v \in B^d$ then,

$$\frac{q(v)}{\|q(v)\|} \to u.$$

Theorem 1 shows two properties of geometric quantiles: first, the norm of the geometric quantile $q(v)$ associated with v diverges to infinity as $\|v\| \to 1$. In other words, Proposition 2(iv) still holds for any distribution. This is a rather intriguing property of geometric quantiles, since it holds even if the distribution of X has a compact support. Second, if $v \to u \in S^{d-1}$ then the geometric quantile $q(v)$ has asymptotic direction u. Proposition 2(ii) thus remains true asymptotically for any distribution. It is possible to specify the convergences obtained in Theorem 1 under integrability assumptions. Theorem 2 provides a first-order expansion of the direction and of the magnitude of an extreme geometric quantile $q(\alpha u)$ in the direction u, where u is a unit vector and α tends to 1.

Theorem 2. *Let* $u \in S^{d-1}$.

(i) If $\mathbb{E}\|X\| < \infty$ *then*

$$q(\alpha u) - \{\|q(\alpha u)\|u + \mathbb{E}(X - \langle X, u\rangle u)\} \to 0 \quad \textit{as} \quad \alpha \uparrow 1.$$

(ii) If $\mathbb{E}\|X\|^2 < \infty$ *and* Σ *denotes the covariance matrix of* X *then*

$$\|q(\alpha u)\|^2(1-\alpha) \to \frac{1}{2}\left(\operatorname{tr}\Sigma - u'\Sigma u\right) \quad \textit{as} \quad \alpha \uparrow 1.$$

As a consequence of Theorem 2, it appears that the difference between the direction of an extreme geometric quantile and its asymptotic direction u

$$\frac{q(\alpha u)}{\|q(\alpha u)\|} - u = \frac{1}{\|q(\alpha u)\|}\mathbb{E}(X - \langle X, u\rangle u) + \mathrm{o}\left(\frac{1}{\|q(\alpha u)\|}\right) \quad \text{as} \quad \alpha \uparrow 1,$$

essentially depends on the behaviour of X in the orthogonal complement of $\mathbb{R}u$. In particular, since the first-order term on the right-hand side of this equality is the orthogonal projection of $\mathbb{E}(X)$ onto the orthogonal complement of $\mathbb{R}u$, its magnitude is minimum when u and $\mathbb{E}(X)$ are collinear and maximum when they are orthogonal.

Further, we can remark that if X has a finite covariance matrix then the magnitude of an extreme geometric quantile in the direction u essentially depends on the norm of u and on the behaviour of X in the orthogonal complement of $\mathbb{R}u$. Moreover, if the distribution of X is absolutely continuous with respect to the Lebesgue measure, the behaviour of extreme geometric quantiles of X is independent of the asymptotic behaviour of its underlying probability density function. As a consequence, although the geometric quantile function characterises the distribution of X (see Corollary 2.9 in Koltchinskii, 1997), no information can be recovered on the asymptotic behaviour of the distribution of X basing solely on extreme geometric quantiles. This appears surprising compared to the univariate case, when the value of an extreme quantile depends on the tail heaviness of the probability density function of X. Besides, in this case, it holds that

$$\frac{\|q(\beta u)\|}{\|q(\alpha u)\|} = \left(\frac{1-\alpha}{1-\beta}\right)^{1/2}(1 + \mathrm{o}(1))$$

when $\alpha \to 1$ and $\beta \to 1$. In other words, given an arbitrary extreme geometric quantile, one can deduce the asymptotic behaviour of every other extreme geometric quantile sharing its direction, independently of the distribution. This is fundamentally different from the univariate case when deducing the value of an extreme quantile from another one requires the knowledge of the extreme-value index of the distribution, see de Haan and Ferreira (2006), Chapter 4.

Finally, Theorem 2 provides some information on the shape of an extreme quantile contour. It is readily seen that the global maximum of the function $h_1(u) := \operatorname{tr}\Sigma - u'\Sigma u$ on S^{d-1} is reached at a unit eigenvector $u_{\min}$ of Σ associated with its smallest eigenvalue $\lambda_{\min} > 0$. Thus, the norm of an extreme geometric quantile is asymptotically the largest in the direction where the variance is the smallest. Similarly, the global minimum of h_1 is reached at a unit eigenvector $u_{\max}$ of Σ associated with its largest eigenvalue $\lambda_{\max} > 0$. This phenomenon is directly related to the fact that the behaviour of extreme geometric quantiles in a given direction u only depends on the distribution of X in the orthogonal complement of $\mathbb{R}u$, both in direction and magnitude. In particular, if f is the probability density function associated with an elliptically contoured distribution, the level sets of f coincide with the levels sets of the function $h_2(u) := u'\Sigma u$. The global maximum of h_2 is reached at the eigenvector $u_{\max}$ while the global minimum is reached at $u_{\min}$. The extreme geometric quantile is therefore furthest from the origin in the direction where the density level is closest to the origin. In such a case, the extreme geometric quantile contour plot and the density level plots are in some sense orthogonal, even though they agree when the distribution of X is spherically symmetric. This can be seen as a consequence of the lack of affine-equivariance of geometric quantiles; to tackle this issue, one may apply a transformation-retransformation procedure, see Serfling (2010). Such procedures admit sample analogues, see for instance Chakraborty *et al.* (1998) and Chakraborty (2001), at the possible loss of geometric interpretation, see Serfling (2004).

This behaviour is illustrated on numerical examples in the next section.

3 Examples and illustrations

Our goal in this paragraph is to illustrate our results and especially Theorem 2. To this end, let us note that this result can be rewritten in the following way: for all $u \in \mathbb{R}^d$ such that $\|u\| = 1$,

$$q(\alpha u) = \left(\frac{\operatorname{tr}\Sigma - u'\Sigma u}{2}\right)^{1/2} \frac{1}{(1-\alpha)^{1/2}}\, u\,(1 + \mathrm{o}(1)) \tag{3}$$

as $\alpha \uparrow 1$, which is referred to as the equivalent of an extreme geometric quantile. To make matters easier, we shall focus on the case $d = 2$. In this case, $u \in S^1$ can be represented by an angle and we may write $u = u_\theta = (\cos\theta, \sin\theta)$, $\theta \in [0, 2\pi)$. The iso-quantile curves $\mathcal{C}q_\alpha = \{q(\alpha u_\theta),\ \theta \in [0, 2\pi)\}$ are then considered in order to get a grasp of the behaviour of extreme quantiles in every direction.

We start by considering two different cases for the distribution of X: the first one is the uniform distribution on the square $[-1, 1]^2$ while the second one is the centred bivariate Gaussian distribution with the same covariance matrix $\Sigma = \operatorname{diag}(1/3, 1/3)$. Note that, in the first case, X is compactly supported. In Figure 1, the iso-quantile curves are computed by either a numerical minimisation of the function ψ or using the equivalent (3). It appears that the equivalent of $q(\alpha u)$ is very close to the true value of $q(\alpha u)$ for $\alpha \geq 0.99$. Besides, the iso-quantile curves associated with the uniform and Gaussian distributions are very close: considering only extreme geometric quantiles does not bring information on the tails of the associated probability distribution. In particular, as predicted by Theorem 1, the iso-quantile curves associated with the uniform distribution are not necessarily included in the support of the distribution. The latter remark suggests that in a large number of cases, the iso-quantile curve $\mathcal{C}q_\alpha$ and the iso-density curve of level α, defined as

$$\mathcal{C}f_\alpha = \{x \in \mathbb{R}^d \,|\, f(x) = (1-\alpha)\|f\|_\infty\} \quad \text{where} \quad \|f\|_\infty = \sup_{\mathbb{R}^d} f,$$

for any bounded density function f, have very different shapes. As seen before, this fact is clearly true in the case of a compactly supported distribution. The next example provides an illustration of this property in a case where X is unbounded: we consider the case when X is a centred bivariate Gaussian random pair having covariance matrix $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2)$ where $\sigma_1^2 = 2$ and $\sigma_2^2 = 1$. In this setting, it is straightforward to show that the iso-density curve $\mathcal{C}f_\alpha$ is an ellipse with radii $\sigma_i\sqrt{-2\log(1-\alpha)}$, $i \in \{1, 2\}$. In Figure 2, both the iso-quantile and iso-density curves are represented for a couple of values of α. One may see that the shape of an extremal iso-quantile curve is very different from that of the corresponding iso-density curve. The orientations of both shapes are orthogonal, as already mentioned as a consequence of Theorem 2. As a conclusion, in view of Figures 1 and 2, extreme geometric quantiles should be used with great care to analyse the tail behaviour of multivariate distributions. This is in accordance with the work of Chaouch and Goga (2010) who used extreme geometric quantiles to detect outliers in a two-dimensional dataset extracted from the Pima Indians Diabetes Database [1]. In their study, points outside the iso-quantile curve of order $\alpha = 0.9$ are marked as outliers. The authors note that, since the shape of the data cloud is elliptical, using a transformation-retransformation procedure (Chakraborty, 2001) as a preprocessing step would be more appropriate. Our theoretical results and simulated experiments confirm this empirical remark.

4 Proofs

Proof of Proposition 1. From Chaudhuri (1996), it is known that if $u \in B^d$ then problem (P) has a unique solution $q(u) \in \mathbb{R}^d$. To prove the converse part of this result, use equation (2) to get

$$\left\|\mathbb{E}\left(\frac{X - q(u)}{\|X - q(u)\|}\right)\right\| = \|u\|.$$

Introduce the coordinate representations $X = (X_1, \ldots, X_d)$ and $q(u) = (q_1(u), \ldots, q_d(u))$. Using the Cauchy-Schwarz inequality yields

$$\|u\|^2 = \left\|\mathbb{E}\left(\frac{X - q(u)}{\|X - q(u)\|}\right)\right\|^2 = \sum_{i=1}^d \left[\mathbb{E}\left(\frac{X_i - q_i(u)}{\|X - q(u)\|}\right)\right]^2 \le \sum_{i=1}^d \mathbb{E}\left(\frac{(X_i - q_i(u))^2}{\|X - q(u)\|^2}\right) = 1.$$

Furthermore, equality holds if and only if for all $i \in \{1, \ldots, d\}$, there exists $\mu_i \in \mathbb{R}$ such that

$$\frac{X_i - q_i(u)}{\|X - q(u)\|} = \mu_i$$

almost surely. In particular, if $w = (\mu_1, \ldots, \mu_d)$, this entails $X \in D = q(u) + \mathbb{R}w$ almost surely, which cannot hold since the distribution of X is not concentrated in a single straight line in $\mathbb{R}^d$. It follows that necessarily $\|u\|^2 < 1$, which is the result. ∎

[1] ftp://ftp.ics.uci.edu/pub/machine-learning-databases/pima-indians-diabetes

Proof of Proposition 2. Note that (2) implies that, for every linear isometry h of $\mathbb{R}^d$ and every $u \in B^d$,

$$h(u) + \mathbb{E}\left(\frac{h(X) - h \circ q(u)}{\|X - q(u)\|}\right) = 0.$$

Since h is a linear isometry, the random vectors X and $h(X)$ have the same distribution and the equality $\|X - q(u)\| = \|h(X) - h \circ q(u)\|$ holds almost surely. It follows that

$$h(u) + \mathbb{E}\left(\frac{X - h \circ q(u)}{\|X - h \circ q(u)\|}\right) = 0.$$

Since $h(u) \in B^d$, it follows that $h \circ q(u) = q \circ h(u)$, which completes the proof of the first statement.

To prove the second part of Proposition 2, start by noting that since X and $-X$ have the same distribution, it holds that $\mathbb{E}\,(X/\|X\|) = 0$. The case $u = 0$ of the second statement is then obtained via (2). If $u \neq 0$, up to using the first part of the result with a suitable linear isometry, we shall assume without loss of generality that $u = (u_1, 0, \ldots, 0)$ for some constant $u_1 \in (0, 1)$. It is then enough to prove that there exists some constant $q_1(u) > 0$ such that $q(u) = (q_1(u), 0, \ldots, 0)$. To this end, let us remark that, on the one hand, if $v_1 \in \mathbb{R}$ and $v = v_1 w \in \mathbb{R}^d$ where $w = (1, 0, \ldots, 0)$ then

$$\forall j \in \{2, \ldots, d\},\ \mathbb{E}\left(\frac{X_j}{\|X - v_1 w\|}\right) = 0, \tag{4}$$

since, for every $j \in \{2, \ldots, d\}$, the random variables X_j and $-X_j$ have the same distribution. On the other hand, the dominated convergence theorem entails that the function

$$v_1 \mapsto \mathbb{E}\left(\frac{X_1 - v_1}{\|X - v_1 w\|}\right)$$

is continuous, converges to 1 at $-\infty$, is equal to 0 at 0 and converges to -1 at $+\infty$. Thus, the intermediate value theorem yields that there exists some constant $q_1(u) > 0$ such that

$$u_1 + \mathbb{E}\left(\frac{X_1 - q_1(u)}{\|X - q_1(u) w\|}\right) = 0. \tag{5}$$

Consequently, collecting (4) and (5) yields

$$u + \mathbb{E}\left(\frac{X - q_1(u) w}{\|X - q_1(u) w\|}\right) = 0$$

and it only remains to apply (2) to finish the proof of the second statement.

To show the third statement, use the first result to obtain that the function $g : \|u\| \mapsto \|q(u)\|$ is indeed well-defined; since the geometric quantile function is continuous, so is g. Assume that g is not increasing: namely, there exist u_1, $u_2 \in B^d$ such that $\|u_1\| < \|u_2\|$ and $\|q(u_1)\| \geq \|q(u_2)\|$. Since $\|q(0)\| = 0$, it is a consequence of the intermediate value theorem that one may find u, $v \in B^d$ such that $\|u\| < \|v\|$ and $\|q(u)\| = \|q(v)\|$. Let h be an isometry such that $h(u/\|u\|) = h(v/\|v\|)$; then

$$\|q(h(u))\| = \|q(u)\| = \|q(v)\| = \|q(h(v))\| \ \text{ and } \ \frac{q(h(u))}{\|q(h(u))\|} = \frac{h(u)}{\|h(u)\|} = \frac{h(v)}{\|h(v)\|} = \frac{q(h(v))}{\|q(h(v))\|}.$$

In other words, $q(h(u))$ and $q(h(v))$ have the same direction and magnitude, so that they are necessarily equal, which entails that $h(u) = h(v)$ because the geometric quantile function is one-to-one. This is a contradiction because $\|h(u)\| = \|u\| < \|v\| = \|h(v)\|$, and the third statement is proven.

Finally, assume that $\|q(u)\|$ does not tend to infinity as $\|u\| \to 1$; since g is increasing, it tends to a finite positive limit r. In other words, $\|q(u)\| < r$ for every $u \in B^d$, which is a contradiction since the geometric quantile function maps B^d onto $\mathbb{R}^d$, and the proof is complete. ■

Proof of Theorem 1. If the first statement were false, then one could find a sequence (v_n) contained in B^d such that $\|v_n\| \to 1$ and such that $(\|q(v_n)\|)$ does not tend to infinity. Up to extracting a subsequence, one can assume that $(\|q(v_n)\|)$ is bounded. Again, up to extraction, one can assume that (v_n) converges to some $v_\infty \in S^{d-1}$ and that $(q(v_n))$ converges to some $q_\infty \in \mathbb{R}^d$. Moreover, it is straightforward to show that for every $u_1,\, u_2,\, q_1,\, q_2 \in \mathbb{R}^d$

$$|\psi(u_1, q_1) - \psi(u_2, q_2)| \leq \{1 + \|u_2\|\} \|q_2 - q_1\| + \|q_1\| \|u_2 - u_1\|$$

so that the function ψ is continuous on $\mathbb{R}^d \times \mathbb{R}^d$. Recall then that the definition of $q(v_n)$ implies that for every $q \in \mathbb{R}^d$, $\psi(v_n, q(v_n)) \leq \psi(v_n, q)$ and let n tend to infinity to obtain

$$q_\infty = \arg\min_{q \in \mathbb{R}^d} \psi(v_\infty, q).$$

Because $v \in S^{d-1}$, this contradicts Proposition 1, and the proof of the first statement is complete: $\|q(v)\| \to \infty$ as $\|v\| \to 1$. To show the second part of Theorem 1, pick a sequence (v_n) of elements of B^d converging to u and remark that from (2),

$$v_n + \mathbb{E}\left(\frac{X - q(v_n)}{\|X - q(v_n)\|}\right) = 0$$

for every integer n. Hence, for n large enough, the following equality holds:

$$v_n + \mathbb{E}\left(\left\|\frac{X}{\|q(v_n)\|} - \frac{q(v_n)}{\|q(v_n)\|}\right\|^{-1} \left[\frac{X}{\|q(v_n)\|} - \frac{q(v_n)}{\|q(v_n)\|}\right]\right) = 0. \tag{6}$$

Since the sequence $(q(v_n)/\|q(v_n)\|)$ is bounded it is enough to show that its only accumulation point is u. Let then u^* be an accumulation point of this sequence. By letting $n \to \infty$ in (6) and applying the dominated convergence theorem, we get $u - u^* = 0$, which completes the proof. ■

Lemma 1 is the first step to prove the first part of Theorem 2.

Lemma 1. *Let $(u_n) \subset B^d$ be a sequence such that $u_n \to u \in S^{d-1}$. If $\mathbb{E}\|X\| < \infty$ then, for all $v \in \mathbb{R}^d$,*

$$\|q(u_n)\| \left\langle u_n - \frac{q(u_n)}{\|q(u_n)\|},\, v \right\rangle \to -\mathbb{E}\langle X - \langle X,\, u\rangle u,\, v\rangle \quad \textit{as} \quad n \to \infty.$$

Proof of Lemma 1. Let $v \in \mathbb{R}^d$ and $W_n(\cdot, v) : \mathbb{R}^d \to \mathbb{R}$ be the function defined by

$$W_n(x, v) = \left[\left\|\frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|}\right\|^{-1} - 1\right] \left\langle \frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|},\, v \right\rangle.$$

For n large enough, (2) entails

$$\left\langle u_n - \frac{q(u_n)}{\|q(u_n)\|},\, v\right\rangle + \mathbb{E}\left(W_n(X,v)\right) + \frac{1}{\|q(u_n)\|}\mathbb{E}\langle X,\, v\rangle = 0.$$

It is therefore enough to show that

$$\|q(u_n)\|\mathbb{E}\left(W_n(X,v)\right) \to -\langle u,\, v\rangle\mathbb{E}\langle X,\, u\rangle \quad \text{as} \quad n \to \infty. \tag{7}$$

Since, for every $x \in \mathbb{R}^d$,

$$\left\|\frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|}\right\|^2 = 1 - \frac{2}{\|q(u_n)\|}\left\langle x,\, \frac{q(u_n)}{\|q(u_n)\|}\right\rangle + \frac{\|x\|^2}{\|q(u_n)\|^2}, \tag{8}$$

it follows from a Taylor expansion and Theorem 1 that

$$\|q(u_n)\|W_n(X,v) \to -\langle u,\, v\rangle\langle X,\, u\rangle \quad \text{almost surely as} \quad n \to \infty. \tag{9}$$

Besides,

$$\begin{aligned}
&\left|\left\|\frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|}\right\|^{-1} - 1\right| \\
= \ &\left\|\frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|}\right\|^{-1}\left[1 + \left\|\frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|}\right\|\right]^{-1}\left|\frac{2}{\|q(u_n)\|}\left\langle x,\, \frac{q(u_n)}{\|q(u_n)\|}\right\rangle - \frac{\|x\|^2}{\|q(u_n)\|^2}\right|,
\end{aligned}$$

and the Cauchy-Schwarz inequality yields

$$\left\|\frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|}\right\|^{-1}\left\langle \frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|},\, v\right\rangle \leq \|v\|.$$

Thus, using the triangular inequality and the Cauchy-Schwarz inequality, it follows that

$$|W_n(x,v)| \leq \left[1 + \left\|\frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|}\right\|\right]^{-1}\frac{\|x\|}{\|q(u_n)\|}\left[2 + \frac{\|x\|}{\|q(u_n)\|}\right]\|v\|.$$

Consequently, one has

$$\|q(u_n)\|\,|W_n(x,v)|\,\mathbb{1}_{\{\|x\|\leq\|q(u_n)\|\}} \leq 3\|v\|\|x\|\mathbb{1}_{\{\|x\|\leq\|q(u_n)\|\}}.$$

Furthermore, the reverse triangle inequality entails, for $x \in \mathbb{R}^d$ such that $\|x\| > \|q(u_n)\|$

$$\left[1 + \left\|\frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|}\right\|\right]^{-1} \leq \frac{\|q(u_n)\|}{\|x\|},$$

and therefore,

$$\|q(u_n)\|\,|W_n(x,v)|\,\mathbb{1}_{\{\|x\|>\|q(u_n)\|\}} \leq 3\|v\|\|x\|\mathbb{1}_{\{\|x\|>\|q(u_n)\|\}}.$$

Finally,

$$\|q(u_n)\|\,|W_n(X,v)| \leq 3\|v\|\|X\|$$

so that the integrand in (7) is bounded from above by an integrable random variable. One can now recall (9) and apply the dominated convergence theorem to obtain (7). The proof is complete. ■

Lemma 2 below is a technical tool necessary to show the second result in Theorem 2.

Lemma 2. *Let $(u_n) \subset B^d$ be a sequence such that $u_n \to u \in S^{d-1}$. If $\mathbb{E}\|X\|^2 < \infty$ then*

$$\|q(u_n)\|^2 \left\langle u_n - \frac{q(u_n)}{\|q(u_n)\|}, \frac{q(u_n)}{\|q(u_n)\|} \right\rangle \to -\frac{1}{2}\mathbb{E}\langle X - \langle X, u\rangle u, X\rangle \quad \textit{as} \quad n \to \infty.$$

Proof of Lemma 2. Let $Z_n : \mathbb{R}^d \to \mathbb{R}$ be the function defined by

$$Z_n(x) = 1 + \|x - q(u_n)\|^{-1} \left\langle x - q(u_n), \frac{q(u_n)}{\|q(u_n)\|} \right\rangle.$$

For n large enough, (2) yields

$$\left\langle u_n - \frac{q(u_n)}{\|q(u_n)\|}, \frac{q(u_n)}{\|q(u_n)\|} \right\rangle + \mathbb{E}\left(Z_n(X)\right) = 0$$

and it thus remains to prove that

$$\|q(u_n)\|^2 \mathbb{E}\left(Z_n(X)\right) \to \frac{1}{2}\mathbb{E}\langle X - \langle X, u\rangle u, X\rangle \quad \text{as} \quad n \to \infty. \tag{10}$$

To this end, rewrite Z_n as

$$Z_n(x) = 1 - \left\| \frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|} \right\|^{-1} \left[1 - \frac{1}{\|q(u_n)\|} \left\langle x, \frac{q(u_n)}{\|q(u_n)\|} \right\rangle\right].$$

Recall from (8) that for every $x \neq q(u_n)$,

$$\left\| \frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|} \right\|^{-1} = \left[1 - \frac{2}{\|q(u_n)\|} \left\langle x, \frac{q(u_n)}{\|q(u_n)\|} \right\rangle + \frac{\|x\|^2}{\|q(u_n)\|^2}\right]^{-1/2}.$$

It follows from this equality, Theorem 1 and a Taylor expansion that

$$Z_n(x) = \frac{1}{2\|q(u_n)\|^2} \left\langle x - \left\langle x, \frac{q(u_n)}{\|q(u_n)\|} \right\rangle \frac{q(u_n)}{\|q(u_n)\|}, x \right\rangle (1 + \mathrm{o}(1))$$

for all $x \in \mathbb{R}^d$. Using Theorem 1 again, we then get

$$\|q(u_n)\|^2 Z_n(X) \to \langle X - \langle X, u\rangle u, X\rangle \quad \text{almost surely as} \quad n \to \infty. \tag{11}$$

To conclude the proof, let $\varphi : \mathbb{R}^d \times \mathbb{R}_+ \times S^{d-1} \to \mathbb{R}$ be the nonnegative function defined by

$$\varphi(x, r, v) = r^2 \left[1 + \frac{\langle x - rv, v\rangle}{\|x - rv\|}\right].$$

Note that $\|q(u_n)\|^2 Z_n(x) = \varphi(x, \|q(u_n)\|, q(u_n)/\|q(u_n)\|)$. Besides, the Cauchy-Schwarz inequality yields

$$\varphi(x, r, v)\mathbb{1}_{\{\|x\| \geq r\}} \leq 2r^2 \mathbb{1}_{\{\|x\| \geq r\}} \leq 2\|x\|^2 \mathbb{1}_{\{\|x\| \geq r\}}. \tag{12}$$

Furthermore, φ can be rewritten as

$$\varphi(x, r, v) = r^2 \left[\frac{\langle x - \langle x, v\rangle v, x\rangle}{\|x - rv\| \left[\|x - rv\| - \langle x - rv, v\rangle\right]}\right].$$

Let us now remark that, if $\|x\| < r$, then, by the Cauchy-Schwarz inequality,

$$\langle x - rv, v\rangle = \langle x, v\rangle - r < 0$$

which makes it clear that

$$\varphi(x,r,v)\mathbb{1}_{\{\|x\|<r\}} \leq r^2 \frac{\langle x - \langle x,\, v\rangle v,\, x\rangle}{\|x - rv\|^2} \mathbb{1}_{\{\|x\|<r\}} =: \psi(x,r,v)\mathbb{1}_{\{\|x\|<r\}}. \tag{13}$$

Since $\|x - rv\|^2 = \|x\|^2 - 2r\langle x,\, v\rangle + r^2$, the function $\psi(x,\cdot,v)$ is derivable on $(\|x\|,\, +\infty)$ and some easy computations yield

$$\frac{\partial \psi}{\partial r}(x,r,v) = 2r\left[\|x\|^2 - r\langle x,\, v\rangle\right] \frac{\langle x - \langle x,\, v\rangle v,\, x\rangle^2}{\|x - rv\|^4}.$$

If $\langle x,\, v\rangle \leq 0$ then $\psi(x,\cdot,v)$ is increasing on $(\|x\|,\, +\infty)$ and thus

$$\forall\, r > \|x\|,\ \psi(x,r,v) \leq \lim_{r\to+\infty} \psi(x,r,v) = \langle x - \langle x,\, v\rangle v,\, x\rangle \leq \|x\|^2. \tag{14}$$

Otherwise, if $\langle x,\, v\rangle > 0$ then $\psi(x,\cdot,v)$ reaches its global maximum over $[\|x\|,\, +\infty)$ at $\|x\|^2/\langle x,\, v\rangle$ and therefore,

$$\forall\, r > \|x\|,\ \psi(x,r,v) \leq \psi\left(x,\, \frac{\|x\|^2}{\langle x,\, v\rangle},\, v\right) = \|x\|^2. \tag{15}$$

Collecting (13), (14) and (15) yields

$$\varphi(x,r,v)\mathbb{1}_{\{\|x\|<r\}} \leq \|x\|^2 \mathbb{1}_{\{\|x\|<r\}}. \tag{16}$$

Recall now (12) to get $\varphi(x,r,v) \leq 2\|x\|^2$ for every $r > 0$ and every $v \in S^{d-1}$. Hence,

$$\|q(u_n)\|^2 Z_n(X) = \varphi(X, \|q(u_n)\|, q(u_n)/\|q(u_n)\|) \leq 2\|X\|^2$$

where the right-hand side is an integrable random variable. Use then (11) and the dominated convergence theorem to complete the proof. ■

Proof of Theorem 2. To show the first statement, let $(u,\, w_1,\, \ldots,\, w_{d-1})$ be an orthonormal basis of $\mathbb{R}^d$ and consider the following expansion :

$$\frac{q(\alpha u)}{\|q(\alpha u)\|} = b(\alpha)u + \sum_{k=1}^{d-1} \beta_k(\alpha) w_k \tag{17}$$

where $b(\alpha)$, $\beta_1(\alpha)$, $\ldots$, $\beta_{d-1}(\alpha)$ are real numbers. It straightforwardly follows that

$$\frac{q(\alpha u)}{\|q(\alpha u)\|} - u - \frac{1}{\|q(\alpha u)\|}\left\{\mathbb{E}(X) - \langle \mathbb{E}(X), u\rangle u\right\} = (b(\alpha)-1)u + \sum_{k=1}^{d-1} \frac{\|q(\alpha u)\|\beta_k(\alpha) - \mathbb{E}\langle X, w_k\rangle}{\|q(\alpha u)\|} w_k. \tag{18}$$

Lemma 1 implies that

$$\|q(\alpha u)\| \left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, w_k \right\rangle = -\|q(\alpha u)\|\beta_k(\alpha) \to -\mathbb{E}\langle X, w_k\rangle \quad \text{as} \quad \alpha \uparrow 1 \tag{19}$$

for all $k \in \{1,\, \ldots,\, d-1\}$. Besides, let us note that $q(\alpha u)/\|q(\alpha u)\| \in S^{d-1}$ entails

$$b^2(\alpha) + \sum_{k=1}^{d-1} \beta_k^2(\alpha) = 1. \tag{20}$$

Theorem 1 shows that $b(\alpha) \to 1$ as $\alpha \uparrow 1$ and thus (19) yields:

$$\|q(\alpha u)\|(1-b(\alpha)) = \frac{1}{2}\|q(\alpha u)\|(1-b^2(\alpha))(1+\mathrm{o}(1)) = \frac{1}{2}\|q(\alpha u)\|\sum_{k=1}^{d-1}\beta_k^2(\alpha)(1+\mathrm{o}(1)) \to 0 \quad \text{as} \quad \alpha \uparrow 1. \tag{21}$$

Collecting (18), (19) and (21), we obtain

$$\frac{q(\alpha u)}{\|q(\alpha u)\|} - u - \frac{1}{\|q(\alpha u)\|}\left\{\mathbb{E}(X) - \langle \mathbb{E}(X),\, u\rangle u\right\} = \mathrm{o}\left(\frac{1}{\|q(\alpha u)\|}\right) \quad \text{as} \quad \alpha \uparrow 1$$

which is the first result.

We now turn to the proof of the second statement. Recall (17) and use Lemma 1 to obtain, for all $k \in \{1,\, \dots,\, d-1\}$,

$$\|q(\alpha u)\|\left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, w_k\right\rangle \to -\mathbb{E}\langle X,\, w_k\rangle \quad \text{as} \quad \alpha \uparrow 1$$

leading to

$$\|q(\alpha u)\|^2\beta_k^2(\alpha) \to [\mathbb{E}\langle X,\, w_k\rangle]^2 \quad \text{as} \quad \alpha \uparrow 1 \tag{22}$$

for all $k \in \{1,\, \dots,\, d-1\}$. Recall (20) and use Lemma 2 to get

$$\|q(\alpha u)\|^2\left[\alpha b(\alpha) - 1\right] \to -\frac{1}{2}\mathbb{E}\langle X - \langle X,\, u\rangle u,\, X\rangle \quad \text{as} \quad \alpha \uparrow 1. \tag{23}$$

Since $(u,\, w_1,\, \dots,\, w_{d-1})$ is an orthonormal basis of $\mathbb{R}^d$, one has the identity

$$\langle X - \langle X,\, u\rangle u,\, X\rangle = \sum_{k=1}^{d-1}\langle X,\, w_k\rangle^2. \tag{24}$$

Collecting (22), (23) and (24) leads to

$$\|q(\alpha u)\|^2\left[1 - \alpha b(\alpha) - \frac{1}{2}\sum_{k=1}^{d-1}\beta_k^2(\alpha)\right] \to \frac{1}{2}\sum_{k=1}^{d-1}\mathrm{Var}\langle X,\, w_k\rangle \quad \text{as} \quad \alpha \uparrow 1.$$

Therefore,

$$\|q(\alpha u)\|^2\left[1 - \alpha b(\alpha) - \frac{1}{2}\left(1 - b^2(\alpha)\right)\right] \to \frac{1}{2}\sum_{k=1}^{d-1}\mathrm{Var}\langle X,\, w_k\rangle \quad \text{as} \quad \alpha \uparrow 1, \tag{25}$$

and easy calculations show that

$$1 - \alpha b(\alpha) - \frac{1}{2}\left(1 - b^2(\alpha)\right) = \frac{1}{2}\left[(1-\alpha)(1+\alpha) + (\alpha - b(\alpha))^2\right]. \tag{26}$$

Finally, in view of Lemma 1,

$$\|q(\alpha u)\|\left\langle \alpha u - \frac{q(\alpha u)}{\|q(\alpha u)\|},\, u\right\rangle \to 0 \quad \text{as} \quad \alpha \uparrow 1$$

which is equivalent to

$$\|q(\alpha u)\|^2\left(\alpha - b(\alpha)\right)^2 \to 0 \quad \text{as} \quad \alpha \uparrow 1. \tag{27}$$

Collecting (25), (26) and (27) yield

$$\|q(\alpha u)\|^2(1-\alpha) \to \frac{1}{2}\sum_{k=1}^{d-1}\mathrm{Var}\langle X,\, w_k\rangle \quad \text{as} \quad \alpha \uparrow 1.$$

Remarking that, for every orthonormal basis $(e_1, \ldots, e_d)$ of $\mathbb{R}^d$,

$$\sum_{k=1}^{d} \mathrm{Var}\langle X,\, e_k\rangle = \sum_{k=1}^{d} e_k' \Sigma e_k = \mathrm{tr}\, \Sigma$$

completes the proof of Theorem 2. ∎

References

Breckling, J., Chambers, R. (1988) M-quantiles, *Biometrika* **75**(4): 761–771.

Cadre, B., Gannoun, A. (2000) Asymptotic normality of consistent estimate of the conditional L_1-median, *Annales de l'Institut de Statistique de l'Université de Paris* **44**: 13–35.

Cai, J.-J., Einmahl, J.H.J., de Haan, L. (2011) Estimation of extreme risk regions under multivariate regular variation, *The Annals of Statistics* **39**(3): 1803–1826.

Cambanis, S., Huang, S., Simons, G. (1981) On the theory of elliptically contoured distributions, *Journal of Multivariate Analysis* **11**(3): 368–385.

Cardot, H., Cénac, P., Zitt, P.-A. (2013) Efficient and fast estimation of the geometric median in Hilbert spaces with an averaged stochastic gradient algorithm, *Bernoulli* **19**(1): 18–43.

Chakraborty, B., Chaudhuri, P., Oja, H. (1998) Operating transformation retransformation on spatial median and angle test, *Statistica Sinica* **8**: 767–784.

Chakraborty, B. (2001) On affine equivariant multivariate quantiles, *Annals of the Institute of Statistical Mathematics* **53**(2): 380–403.

Chakraborty, B. (2003) On multivariate quantile regression, *Journal of Statistical Planning and Inference* **110**(2): 109–132.

Chaouch, M., Goga, C. (2010) Design-based estimation for geometric quantiles with application to outlier detection, *Computational Statistics and Data Analysis* **54**: 2214–2229.

Chaouch, M., Laïb, N. (2013) Nonparametric multivariate L_1-median regression estimation with functional covariates, *Electronic Journal of Statistics* **7**: 1553–1586.

Chaudhuri, P. (1996) On a geometric notion of quantiles for multivariate data, *Journal of the American Statistical Association* **91**(434): 862–872.

Cheng, Y., De Gooijer, J.G. (2007) On the u–th geometric conditional quantile, *Journal of Statistical Planning and Inference* **137**: 1914–1930.

Daouia, A., Gardes, L., Girard, S. (2013) On kernel smoothing for extremal quantile regression, *Bernoulli* **19**: 2557–2589.

Dhar, S.S., Chakraborty, B., Chaudhuri, P. (2014) Comparison of multivariate distributions using quantile-quantile plots and related tests, *Bernoulli*, to appear.

Einmahl, J.H.J., de Haan, L., Krajina, A. (2013) Estimating extreme bivariate quantile regions, *Extremes* **16**(2): 121–145.

Gardes, L., Girard, S. (2012) Functional kernel estimators of large conditional quantiles, *Electronic Journal of Statistics* **6**: 1715–1744.

de Haan, L., Ferreira, A. (2006) Extreme value theory: an introduction, Springer, New York.

Kemperman, J.H.B. (1987) The median of a finite measure on a Banach space, in *Statistical data analysis based on the* L^1*–norm and related methods*, Ed. Y. Dodge, Amsterdam: North Holland, pp. 217–230.

Koenker, R., Bassett, G. (1978) Regression quantiles, *Econometrica* **46**: 33–50.

Koltchinskii, V.I. (2002) M-estimation, convexity and quantiles, *Annals of Statistics* **25**(2): 435–477.

Serfling, R. (2002) Quantile functions for multivariate analysis: approaches and applications, *Statistica Neerlandica* **56**(2): 214–232.

Serfling, R. (2004) Nonparametric multivariate descriptive measures based on spatial quantiles, *Journal of Statistical Planning and Inference* **56**(2): 214–232.

Serfling, B. (2010). Equivariance and invariance properties of multivariate quantile and related functions, and the role of standardization, *Journal of Nonparametric Statistics* **22**: 915–936.

Small, C.G. (1990) A survey of multidimensional medians, *International Statistical Review* **58**(3): 263–277.

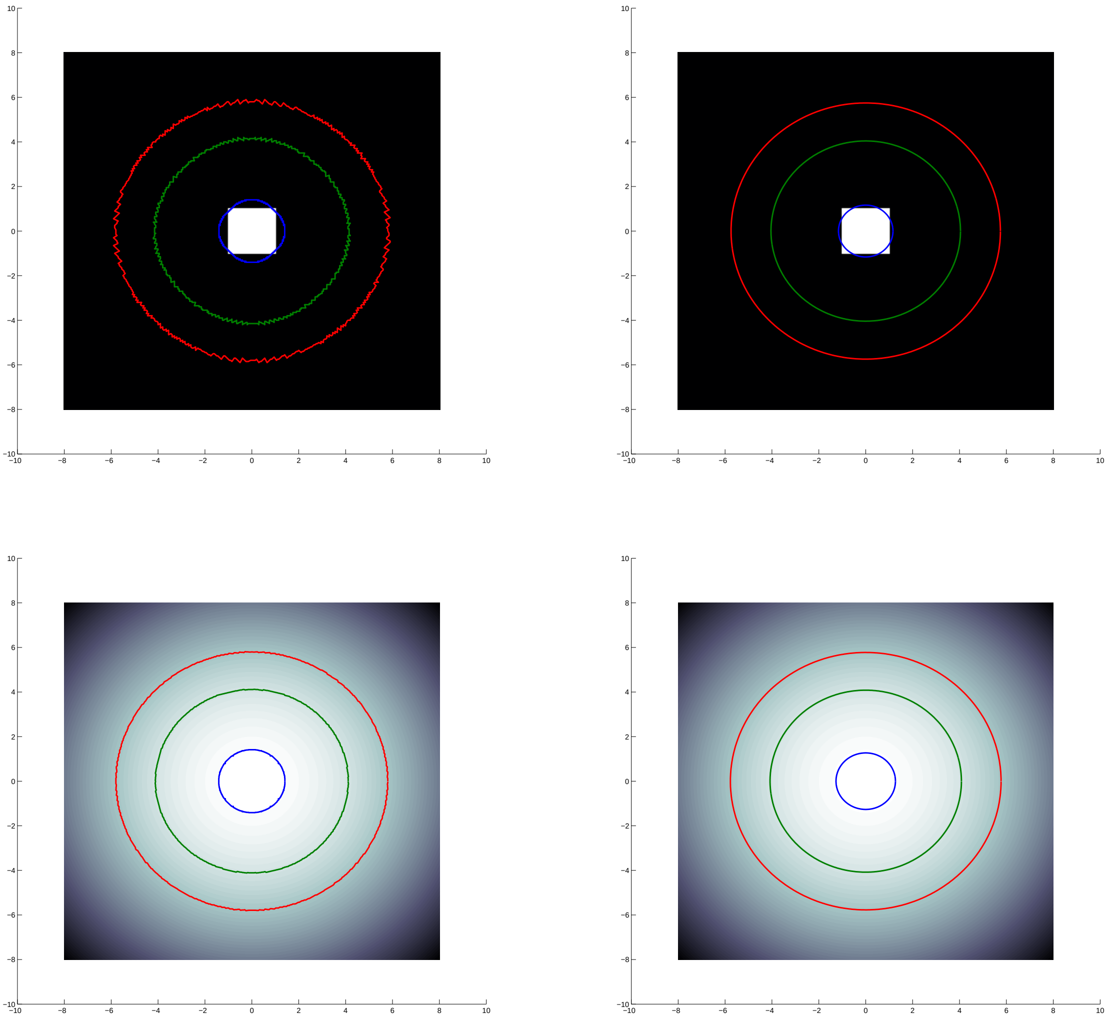

Figure 1: Comparison between a numerical method and the use of the equivalent (3) for the computation of extreme geometric quantiles. Top: uniform distribution, bottom: Gaussian distribution. Left: numerical procedure, right: equivalent (3). Full line: iso-quantile curve of level α; blue: $\alpha = 0.9$, green: $\alpha = 0.99$, red: $\alpha = 0.995$. Density levels are represented with shades of grey.

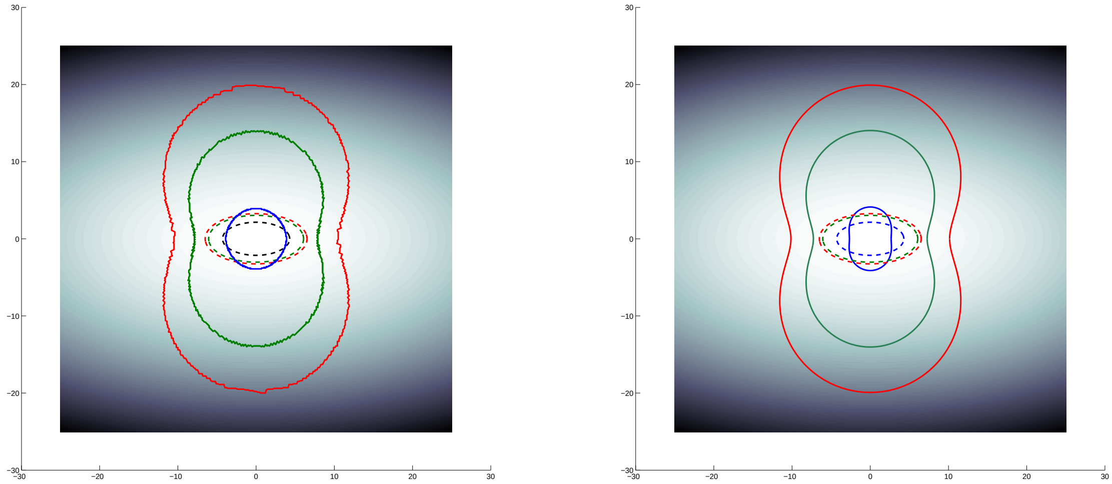

Figure 2: Comparison between the iso-quantile and the iso-density curves. Iso-quantile curves are computed using either a numerical procedure (left) or the equivalent (3) (right). Full line: iso-quantile curve of level α, dashed line: iso-density curve of level α; blue: $\alpha = 0.9$, green: $\alpha = 0.99$, red: $\alpha = 0.995$. Other density levels are represented with shades of grey.