Intriguing properties of extreme geometric quantiles

Stephane 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 vector tends to one. It appears that the behavior of extreme geometric quantiles is totally disconnected from the shape of the associated probability density function. As a consequence, geometric quantiles should not be used as a graphical tool for analyzing multidimensional datasets. We illustrate this phenomenon on some numerical examples.

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

Keywords: Extreme quantile, geometric quantile, asymptotic behavior.

1 Introduction

Up to now, several definitions of multivariate quantiles 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 generalizes the characterization 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 optimization problem (P) defined by

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

with ϕ a loss function defined by

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

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 analyzed 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 characterize 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 generalization 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 extreme geometric quantiles and extreme level sets (of the probability density function) are two fundamentally different notions: the respective shapes of the extreme quantile contour $q(u)$, *i.e.* when $\|u\|$ is constant and close to 1, and of the extreme level sets are not at all similar. In the case of elliptically contoured distributions (Cambanis *et al.* (1981)), these shapes are even orthogonal, in some sense. Moreover, we show that the magnitude of an extreme geometric quantile does not take into account the tail heaviness of the probability density function. As a conclusion, the behavior of high geometric quantiles is totally disconnected from the shape of the associated probability density function. Consequently, geometric quantiles should not be used as a graphical tool for analyzing 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 behavior of extreme geometric quantiles are presented in Section 3. Proofs are deferred to Section 4.

2 Main results

From now on, we assume that X has a probability density function f on $\mathbb{R}^d$, $d \geq 2$. In this case, the optimization problem (P) is equivalent to

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

where

$$\psi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, \ (u,q) \mapsto \int_{\mathbb{R}^d} [\phi(u, x-q) - \phi(u,x)] f(x) dx.$$

From Chaudhuri (1996), since the distribution of X is not concentrated on a straight line in $\mathbb{R}^d$, the solution $q(u)$ of (P) is unique for every $u \in B^d$:

Lemma 1. *For every u belonging to the unit open ball B^d of $\mathbb{R}^d$, the optimization problem*

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

denoted by (P), has a unique solution $q(u)$ in $\mathbb{R}^d$.

The vector $q(u)$ is the geometric quantile of X associated with u. Introducing the convention $t/\|t\| = 0$ if $t = 0$ and remarking that $\varphi(u,\cdot)$ is a strictly convex function, Chaudhuri (1996) proved the following characterization of a geometric quantile: for every $u \in \mathbb{R}^d$, $q(u)$ is the solution of problem (P) if and only if

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

This property makes it possible to prove that the sufficient condition $u \in B^d$ for the existence of $q(u)$ established in Lemma 1 is also necessary:

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

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. Its behavior on the boundary of the unit open ball is somewhat specified in the next result.

Proposition 2. *For every sequence $(u_n) \subset B^d$ such that $\|u_n\| \to 1$, one has $\|q(u_n)\| \to \infty$ as $n \to \infty$. If moreover (u_n) converges to some vector u belonging to the unit sphere S^{d-1} of $\mathbb{R}^d$, then*

$$\frac{q(u_n)}{\|q(u_n)\|} \to u \quad as \quad n \to \infty.$$

Proposition 2 shows two properties of geometric quantiles:

(i) The norm of the geometric quantile $q(u)$ associated with u diverges to infinity as $\|u\| \to 1$. This is a rather disturbing property of geometric quantiles, since it holds even if the density f of X is compactly supported.

(ii) If (u_n) is a sequence of vectors contained in B^d converging to some unit vector u, then the geometric quantile $q(u_n)$ is asymptotically collinear with u_n (and u).

Our first main result examines the behavior of geometric quantiles when the random vector X has an isotropic distribution. In this case, Proposition 2 can be considerably strengthened:

Proposition 3. *If X has an isotropic distribution, i.e. there exists a measurable function $g : \mathbb{R}_+ \to \mathbb{R}_+$ such that for all $x \in \mathbb{R}^d$, $f(x) = g(\|x\|)$, 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 an isotropic 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.

Coming back to the general case, it is possible to specify the convergences obtained in Proposition 2 under integrability assumptions. When X has a finite expectation, Theorem 1 provides a first-order expansion of the direction of an extreme geometric quantile.

Theorem 1. *Assume that $(u_n) \subset B^d$ is a sequence such that $u_n \to u \in S^{d-1}$ and $(u_n) \subset \mathbb{R}u$. If $\mathbb{E}\|X\| < \infty$ then*

$$q(u_n) - \{\|q(u_n)\|u + \mathbb{E}(X - \langle X, u\rangle u)\} \to 0 \quad as \quad n \to \infty.$$

It appears that the difference between the direction of an extreme geometric quantile and its asymptotic direction u

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

essentially depends on the behavior 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.

Theorem 2 provides an equivalent of the norm of an extreme geometric quantile when X has a finite covariance matrix.

Theorem 2. *Assume that $(u_n) \subset B^d$ is a sequence such that $u_n \to u \in S^{d-1}$ and $(u_n) \subset \mathbb{R}u$. If $\mathbb{E}\|X\|^2 < \infty$ then, letting Σ be the covariance matrix of X, it holds that*

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

As a consequence of Theorem 2, we can remark that if X has a finite covariance matrix then the magnitude of an extreme geometric quantile $q(u_n)$ is essentially a function of the norm of u_n and of the behavior of X in the orthogonal complement of $\mathbb{R}u$. Besides, Theorem 2 entails

$$\frac{\|q(\beta_n u)\|}{\|q(\alpha_n u)\|} = \sqrt{\frac{1-\alpha_n}{1-\beta_n}}(1 + \mathrm{o}(1))$$

when $(\alpha_n), (\beta_n) \subset (0, 1)$ are two arbitrary sequences tending to 1. Therefore, the way the density f behaves far from the origin is not captured by extreme geometric quantiles when X is square integrable. Moreover, in this case, given an arbitrary extreme geometric quantile, one can deduce the asymptotic behavior of every other extreme geometric quantile sharing its direction, independently of the probability density function f. 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 the 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 the unit eigenvector $u_{\max}$ of Σ associated with its largest eigenvalue $\lambda_{\max} > 0$. This is a very counterintuitive property. In particular, if f is the density 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 $u_{\min}$. In such a case, the extreme geometric quantile contour plot and the density level plots have opposite behaviors. For instance, the extreme geometric quantile is furthest from the origin in the direction where the density level is closest to the origin. This phenomenon is directly related to the fact that the behavior 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. For more on specific examples to illustrate this peculiar behavior, we refer to the next section.

3 Examples and illustrations

3.1 About our theoretical results

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 behavior 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 centered 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 minimization of the function ψ or using the equivalent (3). It appears that the equivalent of $q(\alpha u)$ is very close to the true $q(\alpha u)$ for $\alpha \geq 0.99$. Besides, the iso-quantile curves associated with the uniform and Gaussian distributions are very close: high geometric quantiles do not bring any information on the associated probability density function. In particular, as predicted by Proposition 2, 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 centered bivariate Gaussian random pair having covariance matrix $\Sigma = \operatorname{diag}(\sigma_1^2, \sigma_2^2)$ 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. Moreover, the volumes bounded by these respective curves turn out to be very different. Indeed, the volume within the iso-density ellipse is given by $-2\pi\sigma_1\sigma_2 \log(1-\alpha)$ while the volume within the iso-quantile curve is

$$\begin{aligned}\frac{1}{2}\int_0^{2\pi} \|q(\alpha u_\theta)\|^2 d\theta &= \frac{1}{4(1-\alpha)}\int_0^{2\pi} (\sigma_1^2 \sin^2\theta + \sigma_2^2 \cos^2\theta) d\theta (1+\mathrm{o}(1)) \\ &= \frac{\pi(\sigma_1^2+\sigma_2^2)}{4}\frac{1}{1-\alpha}(1+\mathrm{o}(1)).\end{aligned}$$

Therefore, as $\alpha \uparrow 1$, the volume within the iso-quantile curve grows at a polynomial rate, while the volume within the iso-density ellipse increases at a logarithmic rate. As a conclusion, in view of Figures 1 and 2, high geometric quantiles do not appear as a convenient tool for analyzing the tail behavior of multivariate distributions.

3.2 Estimating an extreme geometric quantile

In this paragraph, we show that basing on a small sample, extreme geometric quantiles may be estimated in an unexpectedly accurate way. To this end, equation (3) is used to suggest the following estimator of an extreme geometric quantile $q(\alpha_n u)$, $\alpha_n \uparrow 1$:

$$\widehat{q}(\alpha_n u) = \left(\frac{\operatorname{tr}\widehat{\Sigma}_n - u'\widehat{\Sigma}_n u}{2}\right)^{1/2} \frac{1}{(1-\alpha_n)^{1/2}}\, u \tag{4}$$

where $\widehat{\Sigma}_n$ is the empirical estimator of the covariance matrix: if $(X_1, \ldots, X_n)$ is a sample of independent copies of the random vector X, then

$$\widehat{\Sigma}_n = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)'(X_i - \overline{X}_n) \text{ with } \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i.$$

The estimator $\widehat{q}(\alpha_n u)$ is illustrated in the case where X is a centered bivariate Gaussian random pair with covariance matrix $\Sigma = \operatorname{diag}(\sigma_1^2, \sigma_2^2)$ where $\sigma_1^2 = 2$ and $\sigma_2^2 = 1$. The sample size is $n = 100$ and $\alpha_n = 0.995$. The estimator is computed on $N = 100$ replications of the sample. For each replication, the mean-squared error between the true iso-quantile curve and the estimated one is evaluated, and the results corresponding to the first, fifth (median) and ninth decile of the error are displayed in Figure 3. It appears that the estimated curve fits the true iso-quantile curve reasonably well, although the sample size is moderate. This may be seen as a consequence of both the fact that the true iso-quantile curve is well approximated by the one computed with the equivalent (3) and that the estimator $\widehat{\Sigma}_n$ is $\sqrt{n}$−consistent. Note that this is very different from the univariate case, when the sample size needed to estimate extreme quantiles has to be much larger.

4 Proofs

Proof of Proposition 1. Let us assume that $u \in \mathbb{R}^d$ is such that problem (P) has a solution $q(u) \in \mathbb{R}^d$. Equation (2) entails

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

Besides, using the Cauchy-Schwarz inequality together with the fact that f is a probability density function yields

$$\|u\|^2 = \left\| \mathbb{E}\left(\frac{X - q(u)}{\|X - q(u)\|} \right) \right\|^2 = \sum_{i=1}^{d} \left(\int_{\mathbb{R}^d} \frac{x_i - q_i(u)}{\|x - q(u)\|} f(x)\, dx \right)^2 \leq \sum_{i=1}^{d} \int_{\mathbb{R}^d} \frac{(x_i - q_i(u))^2}{\|x - q(u)\|^2} f(x)\, dx.$$

Furthermore, equality holds if and only if for all $i \in \{1, \ldots, d\}$, there exists $(\lambda_i, \mu_i) \in \mathbb{R}^2 \setminus \{0\}$ such that

$$\sqrt{f(x)} \left[\lambda_i \frac{x_i - q_i(u)}{\|x - q(u)\|} + \mu_i \right] = 0$$

for almost every $x \in \mathbb{R}^d$, which leads to

$$\lambda_i \frac{x_i - q_i(u)}{\|x - q(u)\|} + \mu_i = 0$$

for every $i \in \{1, \ldots, d\}$ almost everywhere on the support S of f. In particular, since S is not contained in an affine hyperplane of $\mathbb{R}^d$, one must have $\lambda_i, \mu_i \neq 0$. Hence, since $d \geq 2$, this implies that there exists a nonzero linear form θ on $\mathbb{R}^d$ such that $\theta(x - q(u)) = 0$ almost everywhere on S, which is clearly not true since the set $\mathcal{H} = \{x \in \mathbb{R}^d \,|\, \theta(x - q(u)) = 0\}$ is precisely an affine hyperplane of $\mathbb{R}^d$. It follows that

$$\|u\|^2 < \sum_{i=1}^{d} \int_{\mathbb{R}^d} \frac{(x_i - q_i(u))^2}{\|x - q(u)\|^2} f(x)\, dx = \int_{\mathbb{R}^d} f(x)\, dx = 1.$$

Lemma 1 then proves the converse part of the result. ■

Proof of Proposition 2. Let (u_n) be a sequence contained in B^d such that $\|u_n\| \to 1$. Assume that the sequence $(\|q(u_n)\|)$ does not tend to infinity. Up to extracting a subsequence, one can assume that $(\|q(u_n)\|)$ is bounded. Again, up to extraction, one can assume that (u_n) converges to some u_∞ such that $\|u_\infty\| = 1$ and that $(q(u_n))$ converges to some $q_\infty \in \mathbb{R}^d$. For every $q \in \mathbb{R}^d$, the definition of $q(u_n)$ implies that $\psi(u_n, q(u_n)) \leq \psi(u_n, q)$. Letting n tend to infinity and using the continuity of ϕ entails

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

This contradicts Proposition 1, and the proof of the first statement is complete: $\|q(u_n)\| \to \infty$ as $n \to \infty$. To show the second part of Proposition 2, remark that from equation (2),

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

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

$$u_n + \mathbb{E}\left(\left\| \frac{X}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|} \right\|^{-1} \left[\frac{X}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|} \right] \right) = 0. \tag{5}$$

Since the sequence $(q(u_n)/\|q(u_n)\|)$ 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 (5) and applying the dominated convergence theorem, we get $u - u^* = 0$, which completes the proof. ■

Proof of Proposition 3. Equation (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

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

Applying together Lemma 1 and equation (2) completes the proof of the first statement. To prove the second part of Proposition 3, let us consider the following coordinate representation:

$$\mathbb{E}\left(\frac{X}{\|X\|}\right) = \left(\int_{\mathbb{R}^d} \frac{x_1}{\|x\|} g(\|x\|)\, dx,\, \ldots,\, \int_{\mathbb{R}^d} \frac{x_d}{\|x\|} g(\|x\|)\, dx\right).$$

For every $j \in \{1, \ldots, d\}$, the function

$$x_j \mapsto \frac{x_j}{\|x\|} g(\|x\|)$$

is an odd integrable function, so that $\mathbb{E}(X/\|X\|) = 0$ and the case $u = 0$ of the second statement is obtained via equation (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

$$\forall\, j \in \{2, \ldots, d\},\ \int_{\mathbb{R}^d} \frac{x_j}{\|x - v_1 w\|} g(\|x\|) dx = 0, \tag{6}$$

since, for every $j \in \{2, \ldots, d\}$, the function

$$x_j \mapsto \frac{x_j}{\|x - v_1 w\|} g(\|x\|)$$

is an odd integrable function, On the other hand, the dominated convergence theorem entails that the function

$$v_1 \mapsto \int_{\mathbb{R}^d} \frac{x_1 - v_1}{\|x - v_1 w\|} g(\|x\|) dx$$

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

$$u_1 + \int_{\mathbb{R}^d} \frac{x_1 - q_1(u)}{\|x - q_1(u) w\|} g(\|x\|) dx = 0. \tag{7}$$

Consequently, collecting (6) and (7) yields

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

and it only remains to apply equation (2) to finish the proof. ■

Lemma 2 is the first step to prove Theorem 1.

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

$$\|q(u_n)\| \mathbb{E}\left(W_n(X, v)\right) \to -\langle u,\, v\rangle \mathbb{E}\langle X,\, u\rangle \quad \text{as} \quad n \to \infty. \tag{8}$$

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

$$\left\| \frac{x}{\|q(u_n)\|} - \frac{q(u_n)}{\|q(u_n)\|} \right\|^2 = 1 - \frac{2}{\|q(u_n)\|} \left\langle x,\, \frac{q(u_n)}{\|q(u_n)\|} \right\rangle + \frac{\|x\|^2}{\|q(u_n)\|^2}, \tag{9}$$

it follows from a Taylor expansion and Proposition 2 that

$$\|q(u_n)\| W_n(X, v) \to -\langle u,\, v\rangle \langle X,\, u\rangle \quad \text{almost surely as} \quad n \to \infty. \tag{10}$$

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

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

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

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

Consequently, one has

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

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

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

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 (8) is bounded from above by an integrable random variable. One can now recall (10) and apply the dominated convergence theorem to obtain (8). The proof is complete. ■

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

$$\frac{q(u_n)}{\|q(u_n)\|} = \alpha_n u + \sum_{k=1}^{d-1} \beta_{k,n} w_k$$

where $\alpha_n,\, \beta_{1,n},\, \ldots,\, \beta_{d-1,n}$ are real numbers. It straightforwardly follows that

$$\frac{q(u_n)}{\|q(u_n)\|} - u - \frac{1}{\|q(u_n)\|}\left\{\mathbb{E}(X) - \langle \mathbb{E}(X), u\rangle u\right\} = (\alpha_n - 1) + \sum_{k=1}^{d-1} \frac{\|q(u_n)\|\beta_{k,n} - \mathbb{E}\langle X, w_k\rangle}{\|q(u_n)\|} w_k. \tag{11}$$

Lemma 2 implies that

$$\|q(u_n)\| \left\langle u_n - \frac{q(u_n)}{\|q(u_n)\|},\, w_k \right\rangle = -\|q(u_n)\|\beta_{k,n} \to -\mathbb{E}\langle X, w_k\rangle \quad \text{as} \quad n \to \infty \tag{12}$$

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

$$\alpha_n^2 + \sum_{k=1}^{d-1} \beta_{k,n}^2 = 1.$$

Proposition 2 shows that $\alpha_n \to 1$ and thus (12) yields:

$$\|q(u_n)\|(1-\alpha_n) = \frac{1}{2}\|q(u_n)\|(1-\alpha_n^2)(1+\mathrm{o}(1)) = \frac{1}{2}\|q(u_n)\| \sum_{k=1}^{d-1} \beta_{k,n}^2 (1+\mathrm{o}(1)) \to 0 \quad \text{as} \quad n \to \infty. \tag{13}$$

Collecting (11), (12) and (13), we obtain

$$\frac{q(u_n)}{\|q(u_n)\|} - u - \frac{1}{\|q(u_n)\|}\left\{\mathbb{E}(X) - \langle \mathbb{E}(X),\, u\rangle u\right\} = \mathrm{o}\left(\frac{1}{\|q(u_n)\|}\right)$$

which completes the proof. ■

Lemma 3 below is a technical tool necessary to show Theorem 2.

Lemma 3. *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 3. 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, equation (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

$$\|q(u_n)\|^2 \mathbb{E}\left(Z_n(X)\right) \to \frac{1}{2}\mathbb{E}\langle X - \langle X, u\rangle u, X\rangle \quad \text{as} \quad n \to \infty. \tag{14}$$

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 (9) 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, Proposition 2 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 Proposition 2 again, we then get

$$\|q(u_n)\|^2 Z_n(X) \to \langle X - \langle X, u\rangle u, X\rangle \quad \text{almost surely as} \quad n \to \infty. \tag{15}$$

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

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

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

$$\varphi(x, r, v)\mathbb{1}_{\{\|x\| < r\}} \leq r^2 \frac{\langle x - \langle x, v\rangle v, x\rangle}{\|x - rv\|^2} \mathbb{1}_{\{\|x\| < r\}} =: \psi(x, r, v)\mathbb{1}_{\{\|x\| < r\}}. \tag{17}$$

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

$$\forall\, r > \|x\|,\ \psi(x,r,v) \le \lim_{r\to+\infty} \psi(x,r,v) = \langle x - \langle x,\, v\rangle v,\, x\rangle \le \|x\|^2. \tag{18}$$

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,

$$\forall\, r > \|x\|,\ \psi(x,r,v) \le \psi\left(x,\ \frac{\|x\|^2}{\langle x,\, v\rangle},\ v\right) = \|x\|^2. \tag{19}$$

Collecting (17), (18) and (19) yields

$$\varphi(x,r,v)\mathbb{1}_{\{\|x\|<r\}} \le \|x\|^2\mathbb{1}_{\{\|x\|<r\}}. \tag{20}$$

Recall now (16) to get $\varphi(x,r,v) \le 2\|x\|^2$ for every $r>0$ and every $v \in S^{d-1}$. Hence,

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

where the right-hand side is an integrable random variable. Use then (15) and the dominated convergence theorem to complete the proof. ■

Proof of Theorem 2. As in the proof of Theorem 1, we start by introducing an orthonormal basis $(u,\, w_1,\, \ldots,\, w_{d-1})$ of $\mathbb{R}^d$ to write

$$u_n = \|u_n\| u \quad \text{and} \quad \frac{q(u_n)}{\|q(u_n)\|} = \alpha_n u + \sum_{k=1}^{d-1} \beta_{k,n} w_k$$

where $\alpha_n,\, \beta_{1,n},\, \ldots,\, \beta_{d-1,n}$ are real numbers. Lemma 2 implies, for all $k \in \{1,\, \ldots,\, d-1\}$,

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

leading to

$$\|q(u_n)\|^2 \beta_{k,n}^2 \to [\mathbb{E}\langle X,\, w_k\rangle]^2 \quad \text{as} \quad n \to \infty \tag{21}$$

for all $k \in \{1,\, \ldots,\, d-1\}$. Besides, recall that $q(u_n)/\|q(u_n)\| \in S^{d-1}$ entails

$$\alpha_n^2 + \sum_{k=1}^{d-1} \beta_{k,n}^2 = 1.$$

Hence, Lemma 3 yields

$$\|q(u_n)\|^2 \left[\alpha_n \|u_n\| - 1\right] \to -\frac{1}{2}\mathbb{E}\langle X - \langle X,\, u\rangle u,\, X\rangle \quad \text{as} \quad n\to\infty. \tag{22}$$

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

$$\langle X - \langle X,\, u\rangle u,\, X\rangle = \sum_{k=1}^{d-1} \langle X,\, w_k\rangle^2. \tag{23}$$

Collecting (21), (22) and (23) leads to

$$\|q(u_n)\|^2 \left[1 - \alpha_n\|u_n\| - \frac{1}{2}\sum_{k=1}^{d-1} \beta_{k,n}^2\right] \to \frac{1}{2}\sum_{k=1}^{d-1} \operatorname{Var}\langle X,\, w_k\rangle \quad \text{as} \quad n\to\infty.$$

Therefore,

$$\|q(u_n)\|^2 \left[1 - \alpha_n \|u_n\| - \frac{1}{2}\left(1 - \alpha_n^2\right)\right] \to \frac{1}{2}\sum_{k=1}^{d-1} \mathrm{Var}\langle X,\, w_k\rangle \quad \text{as} \quad n \to \infty, \tag{24}$$

and easy calculations show that

$$1 - \alpha_n \|u_n\| - \frac{1}{2}\left(1 - \alpha_n^2\right) = \frac{1}{2}\left[(1 - \|u_n\|)(1 + \|u_n\|) + (\|u_n\| - \alpha_n)^2\right]. \tag{25}$$

Finally, in view of Lemma 2,

$$\|q(u_n)\| \left\langle u_n - \frac{q(u_n)}{\|q(u_n)\|},\, u \right\rangle \to 0 \quad \text{as} \quad n \to \infty$$

which is equivalent to

$$\|q(u_n)\|^2 \left(\|u_n\| - \alpha_n\right)^2 \to 0 \quad \text{as} \quad n \to \infty. \tag{26}$$

Recalling that $\|u_n\| \to 1$ and collecting (24), (25) and (26) yield

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

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. (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*, to appear.

Dhar, S.S., Chakraborty, B., Chaudhuri, P. (2013) 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.

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

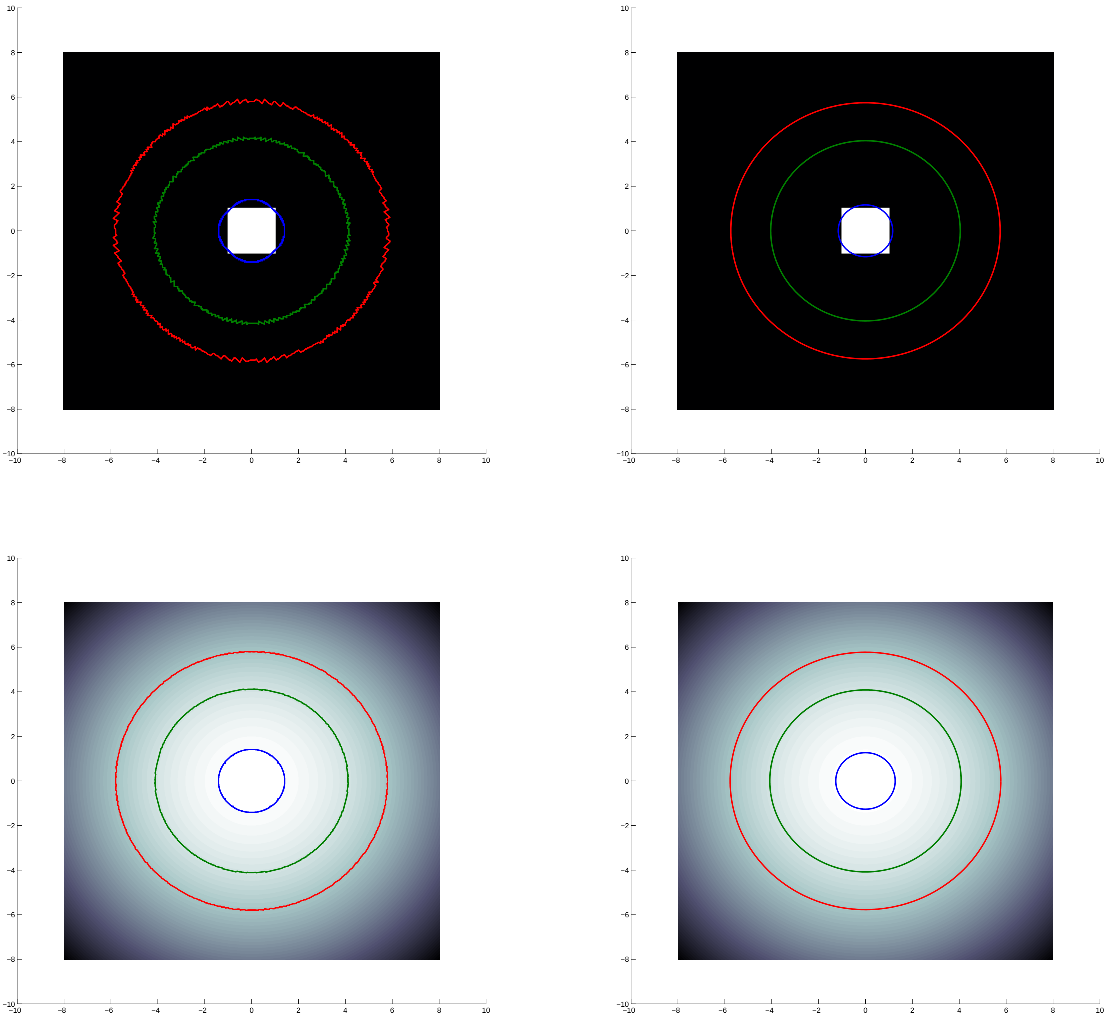

Figure 1: Comparison between a numerical method and the use of the equivalent (3) for the computation of extreme geometric quantiles. Top: uniform distribution, bottom: Gaussian distribution. Left: numerical procedure, right: equivalent (3). Full line: iso-quantile curve of level α; blue: $\alpha = 0.9$,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

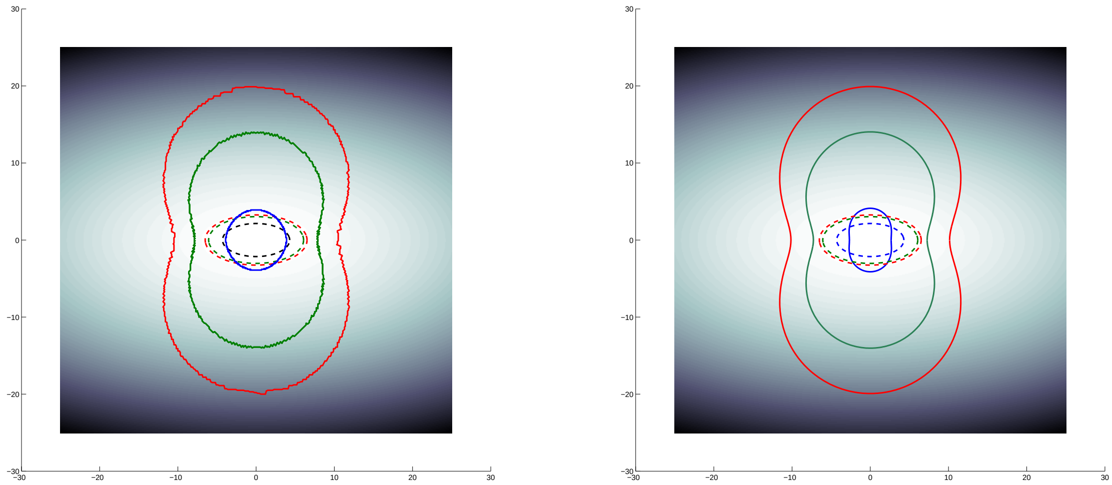

Figure 2: Comparison between the iso-quantile and the iso-density curves. Iso-quantile curves are computed using either a numerical procedure (left) or the equivalent (3) (right). Full line: iso-quantile curve of level α, dashed line: iso-density curve of level α; blue: $\alpha = 0.9$, green: $\alpha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

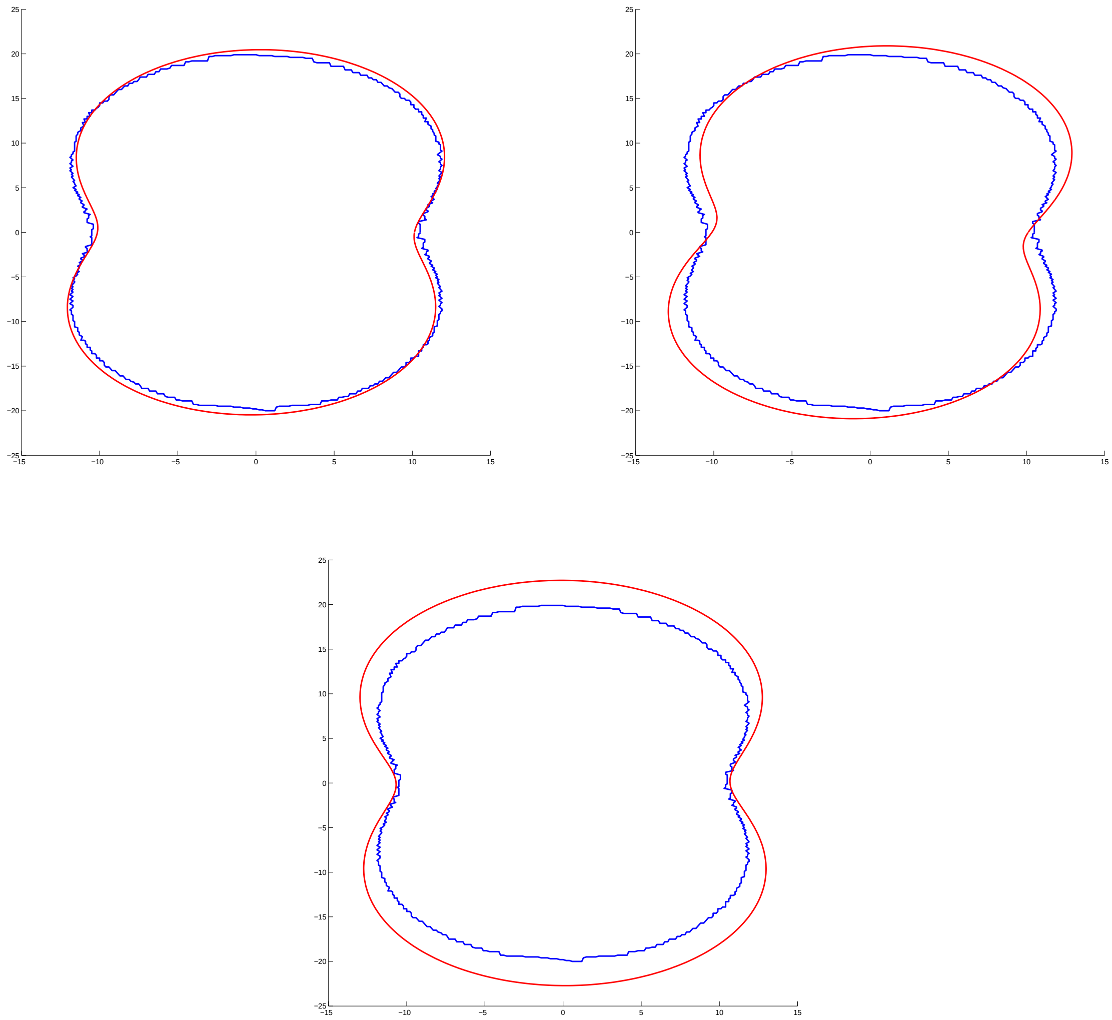

Figure 3: Estimation of the iso-quantile curves. Blue line: true iso-quantile curve of level $\alpha_n = 0.995$, red line: estimated iso-quantile curve of level $\alpha_n = 0.995$. Top left: 10% best result, top right: median result, bottom: 10% worst result.