Statistica Sinica **22** (2012), 1589-1610
doi:http://dx.doi.org/10.5705/ss.2010.224

QUANTILE TOMOGRAPHY: USING QUANTILES WITH MULTIVARIATE DATA

Linglong Kong and Ivan Mizera

University of Alberta

Abstract: The use of quantiles to obtain insights about multivariate data is addressed. It is argued that incisive insights can be obtained by considering directional quantiles, the quantiles of projections. Directional quantile envelopes are proposed as a way to condense this kind of information; it is demonstrated that they are essentially halfspace (Tukey) depth levels sets, coinciding for elliptic distributions (in particular multivariate normal) with density contours. Relevant questions concerning their indexing, the possibility of the reverse retrieval of directional quantile information, invariance with respect to affine transformations, and approximation/asymptotic properties are studied. It is argued that analysis in terms of directional quantiles and their envelopes offers a straightforward probabilistic interpretation and thus conveys a concrete quantitative meaning; the directional definition can be adapted to elaborate frameworks, like estimation of extreme quantiles and directional quantile regression, the regression of depth contours on covariates. The latter facilitates the construction of multivariate growth charts—the question that motivated this development.

Key words and phrases: Data depth, growth charts, quantile regression, quantiles.

1. Introduction

The concept of the quantile function is well rooted in the ordering of $\mathbb{R}$. For $0 < p < 1$, the pth *quantile* (or *percentile*, if indexed by $100p$) of a probability distribution P is

$$Q(p) = \inf\{u\colon F(u) \geq p\},$$

where $F(u) = P((-\infty, u])$ is the cumulative distribution function of P; see Eubank (1986) or Shorack (2000). Essentially, Q could be perceived as a function inverse to F; the more sophisticated definition is necessitated by a demand to treat formally cases when there is none, or more than one q satisfying $F(q) = p$. This is a well-known detail: if an alternative definition via the minimization of the integral

$$\int {}_{p-1}|x - q|_p \, P(dx), \qquad \text{where } {}_{p-1}|x|_p = x(p - I(x < 0)),$$

is adopted, then the set, $\mathcal{Q}(p)$, of all minimizing q may be called, with Shorack (2000), a pth *quantile set* of P; a prescription that returns, for all p, a unique element of this (always nonempty, convex, and closed) set then constitutes a *quantile version.* Hyndman and Fan (1996) review quantile versions used in practice; these are implemented as options of the R function `quantile` by Frohne and Hyndman (2004). While the "inf" version, as defined above (not the default of `quantile`, but its option "`type=1`"), is preferred in theory and was used for all pictures and computations in this paper, practice often favors other choices—like the "midpoint" version yielding the sample median for $p = 1/2$.

The potential of quantiles for blunt quantitative statements is well-known, and was noted already when the reflection of Quetelet was endorsed by Edgeworth (1886, 1893) and Galton (1889). The information, say, that 50 is the 0.9th quantile leads to unambiguous conclusion that about 10% of the results are to be expected beyond, and about 90% below 50. Compared to other statistical uses—for which we refer to Parzen (2004) and the references there—this "descriptive grip" is very palpable and hard to imagine in the multivariate setting.

Yet, a natural and legitimate step in the analysis of multivariate datasets is to apply quantiles to univariate functions of the original data, the most immediate of such functions being projections. In Section 2, we exemplify aspects of such exploration: in particular, when projections in all directions are investigated simultaneously, we observe a need for some kind of a summary, and propose in Section 3 "directional quantile envelopes" to this end. The latter turn out to be (if the "inf" quantile version is adhered to) level sets ("contours") of the halfspace (Tukey) depth—already a well-known concept whose directional interpretation is also hardly surprising. The new name is thus feebly justified by the fact that the exact equality to depth contours does not hold true in general for other quantile versions (a fact of mathematical rather than data-analytic significance).

What we see as a potential contribution of this paper is the observation that the directional interpretation of depth contours not only gives them concrete probabilistic interpretation (discussed in Sections 4 and 5) and thus quantitative meaning, but that it enables to adapt them to more elaborate frameworks such as estimation of extreme quantiles and directional quantile regression. In particular, directional interpretation makes it possible, in a simple way, to regress depth contours on covariates (borrowing strength as is typical in regression) and subsequently the construction of bivariate growth charts—the methodology whose pursuit was the original motivation for this paper. The applications of the directional approach are introduced in Section 7, after the discussion of some relevant properties in Section 6.

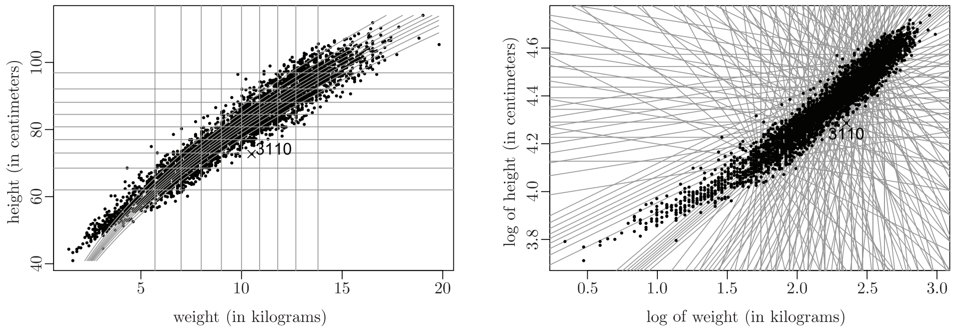

Figure 1. Left panel: multivariate data typically offer insights beyond the marginal view, often through the quantiles of univariate functions of primary variables. Right panel: the plot gets quickly overloaded if multiple directions and indexing probabilities are requested.

2. Quantile Analysis of Projections, Illustrated on an Example

We illustrate the possible objectives of the analysis using quantiles in multivariate setting with an example. The left panel of Figure 1 shows the scatterplot of the weight and height of 4291 Nepali children, aged between 3 and 60 months—the data constituting a part of the Nepal Nutrition Intervention Project-Sarlahi (NNIP-S, principal investigator Keith P. West, Jr., funded by the Agency of International Development). The horizontal and vertical lines show the deciles of height and weight, respectively, of the empirical distributions of the corresponding variables—indicating the simple conclusions that can be reached about the variables. For instance, the points above the upper horizontal line correspond to 10% of the subjects exceeding the others in height; similarly, the points right of the rightmost line correspond to 10% of those exceeding the others in weight.

It would be interesting to know what proportion of the data corresponds to the upper right corner, but this information is not directly available (unless we count the points manually). Regarding the subject labeled by 3, 110, for example, we can only say that its weight is somewhat higher than, but otherwise fairly close to the median; its height is about at the second decile, that is, exceeding about 20% and exceeded by about 80% of its peers. Nevertheless, one could argue that 3, 110 is in certain sense extremal, outstanding from the rest.

A possible way of substantiating this impression quantitatively is to invoke Quetelet's *body mass index* (hereafter *BMI*), defined as the ratio of weight to squared height (in the metric system). The curved lines in the left panel of Figure 1 show the deciles of the empirical distribution of the *BMI*. We can see that in terms of *BMI*, the subject 3, 110 is indeed extreme, belonging to the group of 10% of those with maximal *BMI*.

An expert on nutrition may dispute the relevance of *BMI* for young children, and remind us of possible alternatives—for instance, the Rohrer index (ratio of weight to *cubed* height, hereafter *ROI*). However, we do not think that the problem lies in deciding whether that or another index is to be preferred; the essence of the data may lie well beyond the index-style of description. For example, suppose that we wish to make quantitative statements about the subjects represented by the points in the upper right and lower left rectangles. Since we are not aware of any relevant index related to this objective, we may simply look, in the left panel of Figure 1, at the deciles of some suitable linear combination of weight and height.

Pursuing vague objectives in the nonlinear realm may be hard—there are simply too many choices. A possible solution is to limit the attention only to *linear* functions of the original data; note that the "*BMI* contours" in the left panel of Figure 1 are not that badly approximated by straight lines. We can do even better by taking the *logarithms* of weight and height as primary variables—then we can investigate both *BMI* and *ROI* among their linear combinations, and possibly much more. Therefore, we switch to the logarithmic scale, starting from the right panel of Figure 1.

Rather than this technical detail, however, the more important outcome of our exploration is that quantiles of certain functions of variables (in particular, linear combinations) may provide valuable information about multivariate data. Focusing on linear combinations, we realize that it is sufficient to look exclusively at projections; any other linear combination is a multiple of a projection, and the quantile of a multiple is the multiple of the quantile. In other words, we believe that insights about data can be obtained by looking at the directional quantiles.

The right panel of Figure 1 thus shows the plot of the logarithms of weight and height, together with superimposed lines indicating deciles in 20 uniformly spaced directions. While these *directional quantile lines* are an appealing way to present the directional quantile information, we have to admit that the plot becomes quickly overloaded if multiple directions and indexing probabilities are requested. (While our focus here is not exclusively graphical, the task of plotting is probably the most palpable one to epitomize our objectives.) Therefore, we would like to achieve some compression of the directional quantile information; to this end, we propose *directional quantile envelopes.*

3. Directional Quantile Envelopes and Halfspace Depth

Notationally, it is convenient to work with random variables or vectors, and write

$$Q(p) = Q(p, X) = \inf\{u \colon \mathbb{P}[X \leq u] \geq p\},$$

despite that quantiles depend only on the distribution, P, of X. Hereafter, X always stands for a random vector with distribution P; the apparent notational convention is to suppress the dependence on X when no confusion may arise. We call any vector with unit norm in $\mathbb{R}^d$ a *normalized direction*, and denote the set of all such vectors by $\mathbb{S}^{d-1}$. Given a normalized direction $s \in \mathbb{S}^{d-1}$ and $0 < p < 1$, the pth *directional quantile*, in the direction s, is nothing but the pth quantile of the corresponding projection of the distribution of X,

$$Q(p, s) = Q(p, s, X) = Q(p, s^{\mathsf{T}} X).$$

A related notion is the pth *directional quantile hyperplane*, given by the equation $s^{\mathsf{T}} x = Q(p, s)$. For $d = 2$, the hyperplanes amount to lines—which in our figures indicate how directional quantiles divide the data.

The pth directional quantile in the direction s and the $(1-p)$-th directional quantile in the direction $-s$ are not necessarily equal—due to the inf convention employed in their definition. Nonetheless, they often coincide—for instance, it is not possible to distinguish between any p-th and $(1-p)$th directional quantile hyperplanes if for any projection of P, all quantile sets are singletons. A sufficient condition for this is that P has *contiguous support*: there is no intersection of halfspaces with parallel boundaries that has nonempty interior but zero probability P and divides the support of P to two parts. (Note that if the support is not contiguous, it is not connected; however, it may be disconnected and still contiguous.) We believe that contiguous support is a fairly typical virtue of population distributions, and consequently will limit most of our attention to p from $(0, 1/2]$.

Contiguity of the support of X is also one of the things that implies continuity, in s, of the directional quantiles. The following theorem is formulated slightly more generally, to allow for alternative quantile versions and later asymptotic considerations. Using the theorem with $X_n = X$ shows that the directional quantiles depend continuously on s for all empirical, and many population distributions.

Theorem 1. *If the support of X is bounded, then $\mathcal{Q}(p, s)$ is a continuous function of s, for every $p \in (0, 1)$. The same holds true when the support of X is contiguous; moreover, if a sequence of random vectors X_n converges almost surely to X, and $s_n \to s$, then $\mathcal{Q}(p, s_n, X_n)$ converges to $\mathcal{Q}(p, s, X)$ in the Pompeiu-Hausdorff distance, for every $p \in (0, 1)$.*

The terminology of "Pompeiu-Hausdorff" is that of Rockafellar and Wets (1998).

The idea of what constitutes the inner envelope of the directional quantile hyperplanes is quite clear from the left panel of Figure 2. More formally, for

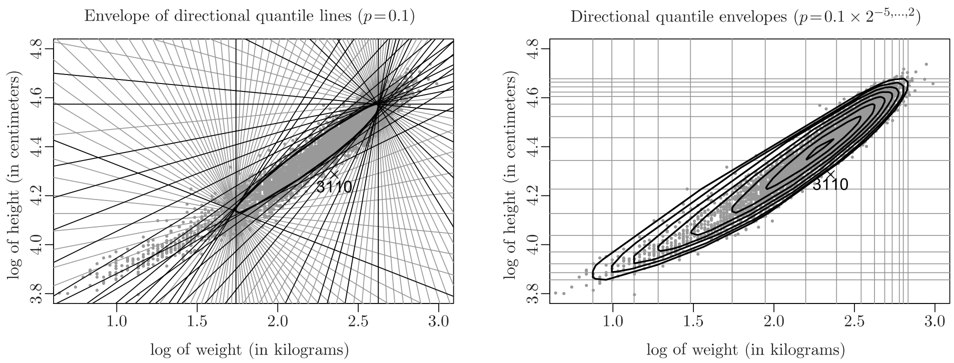

Figure 2. Left panel: for fixed p, we form the inner envelope of the directional quantile lines. Right panel: directional quantile envelopes for $p = 2^i/10$, $i = -5, \ldots, 2$. In the central part, the contours resemble those obtained by fitting normal distribution; in the tail area, they adapt more to the specific shape of the data. Several p can be accommodated simultaneously, and the directional quantile information can be retrieved from the contours in a relatively straightforward retrieval way.

$p \in (0, 1/2]$, the pth *directional quantile envelope* generated by $Q(p, s)$ is defined as the intersection,

$$D(p) = \bigcap_{s \in \mathbb{S}^{d-1}} H(s, Q(p, s)),$$

where $H(s, q) = \{x \colon s^{\mathsf{T}} x \geq q\}$ is the *supporting halfspace* determined by $s \in \mathbb{S}^{d-1}$ and $q \in \mathbb{R}$. In case the intersection is taken only over a subset $A \subseteq \mathbb{S}^{d-1}$ of all possible directions (for instance, in the numerical construction of the envelopes), we write $D_A(p)$; in the spirit of this notation, $D(p) = D_{\mathbb{S}^{d-1}}(p)$. Directional quantile envelopes are convex (being intersections of convex sets) and bounded (for $D_A(p)$, this is true whenever A is not contained in any closed halfspace whose boundary contains the origin). They have a close connection to what is known as (halfspace or Tukey) depth, first considered by Hodges (1955); Tukey (1975) proposed depth contours for plotting bivariate data, in a spirit close to ours. Let P be a distribution in $\mathbb{R}^d$. Recall that the *depth*, $d(x)$, of a point $x \in \mathbb{R}^d$, is defined as $\inf P(H)$, where H runs over all closed halfspaces containing x (or, equivalently, over all closed halfspaces with x lying on their boundary).

Theorem 2. *For every $p \in (0, 1/2]$, the directional quantile envelope is the upper level set of depth: $D(p) = \{x \colon d(x) \geq p\}$.*

Theorem 2 implies that directional quantile envelopes are nonempty for $p \leq 1/(d+1)$, in the two-dimensional case for $p \leq 1/3$, due to a result known as a centerpoint theorem—see Donoho and Gasko (1992) or Mizera (2002).

We remark that Theorem 2 is rigorously true only for the "inf" version of the quantile definition. In practice, some other version may be preferred, for instance, to allow for constructing contours interpolating between various depth level sets. Most of our theorems hold true also for other versions, as can be seen in the Appendix; this fact gives some justification for calling what are essentially "depth contours" by the new name "directional quantile envelopes". All interpolated versions of quantiles yield somewhat smaller envelopes; Rousseeuw and Ruts (1999) point out that this is also the case for the related notion of halfspace trimmed contours of Massé and Theodorescu (1994). These subtle differences vanish in regular situations—for instance, for absolutely continuous distributions with positive densities.

4. Indexing, Illustrated on the Multivariate Normal Distribution

As can be seen on the right panel of Figure 2, suppressing the underlying directional quantile lines (still shown in the left panel of Figure 2) allows for accommodating several p simultaneously. In the central part, the contours have elliptical shape, resembling the density contours of the multivariate normal distribution. Indeed, Theorem 4 below implies that directional quantile envelopes coincide with the density contours for any elliptic distribution—in particular, for the multivariate normal. In such a context, an intriguing question of practical importance is that of indexing: which particular contours of the fitted normal distribution would correspond to which p?

Our definition of directional quantile envelopes leads to what we call "indexing by the tangent mass". It is illustrated by the grey shading in the right panel of Figure 3: given any point of the contour, the halfplane passing through the point and tangent to the contour contains exactly p of the mass of the fitted multivariate normal distribution. This extrapolates the univariate fact that pth and $(1-p)$th quantiles mark the boundaries of the halfspaces containing exactly p of the distribution mass. For the standard normal distribution, the contour corresponding to p is that matching the univariate quantiles indexed by p and $1-p$ when projected on the coordinate axes, and can be found by transforming to the standard form and the subsequent inverse transformation.

Thinking in terms of elliptic confidence sets suggests another alternative, "indexing by the enclosed mass", which generalizes the univariate fact that the pth and $(1-p)$th quantiles together leave $2p$ of the distribution mass outside their convex hull. In such a case, the contours corresponding to deciles would be those enclosing 0.8, 0.6, 0.4, 0.2 of the mass of the fitted normal distribution, together with the contour consisting of the single point located at the mode. This type of indexing can be seen in the right panel of Figure 3; it shows now that the subject represented by the point 3,110 lies in the outstanding 20% of the sample;

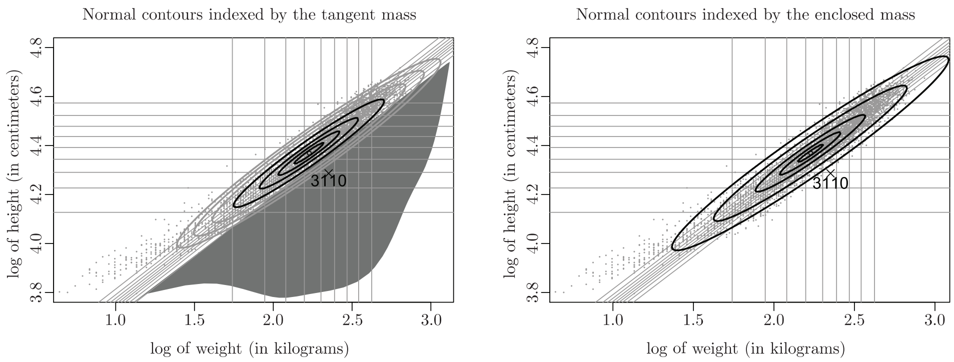

Figure 3. Left panel: if indexed by the tangent mass, the contours of the fitted normal distribution theoretically match projected quantiles. The half-plane tangent to the contour and passing through the point contains exactly p of the mass of the fitted multivariate normal distribution. Right panel: if indexed by the mass they enclose, the contours of the fitted normal distribution do not interact well with directional and marginal quantiles.

however, this exceptionality is somewhat "generic"—expressed not only through the company of similar subjects with large weight given the height, but also by the company of those with small weight given the height, and of those with small height and weight altogether.

Contrary to that, the 10% extremality of 3,110 suggested by the left panel can be interpreted as substantial: it is carried by the company of subjects with similar nature, those with large weight given the height. Note that the boundary of the greyed halfspace is almost identical with the line indicating the (0.9)th quantile of the *BMI*; hence the picture shows that in this case, the extremality of 3,110 may be interpreted in terms of the *BMI*. Also, unlike the indexing by the enclosed mass, indexing by the tangent mass interacts well with marginal and directional quantiles.

Methods based on fitting normal distribution are still somewhat central to multivariate statistics; we find it thus encouraging that for the normal distribution, our proposed contours coincide with the contours of its density, the objects that Evans (1982) defines, in one of the first papers on the subject, to be "bivariate quantiles". However, the approach based on fitting normal distribution would have all virtues of an ideal, if the hypothesized distribution would be "closely followed" by the data—as for our example occurs in the central part, as can be seen in the right panel of Figure 2, but not that much on the fringe of the data cloud. The elliptic contours would be considerably off there, but we can see that the directional quantile envelopes adapt to the specific shape of the data, and thus behave in a rather nonparametric way.

5. Recovery of Directional Quantile Information

Obviously, directional quantile envelopes suppress some information contained in directional quantiles; a question of importance is how far it is possible to get this information back. Let e be a point lying on the boundary, ∂E, of a bounded convex set $E \subset \mathbb{R}^d$. A *tangent* of E at e is any hyperplane (line) containing e that has empty intersection with the interior of E. Such a line determines the corresponding *tangent halfspace*, the halfspace that has the tangent as its boundary and its interior does not contain any point of E. The *maximal mass at a hyperplane* is defined as $\Delta(P) = \sup\{\mathbb{P}[s^\tau X = c] \colon s \in \mathbb{S}^{d-1},\ c \in \mathbb{R}\}$. The following theorem provides a practical guideline for recovering the directional quantile information and is thus essential in interpreting directional quantile envelopes.

Theorem 3. *Let P be a distribution in $\mathbb{R}^d$, and let $p \in (0, 1/2]$. If H is a tangent halfspace of $D(p)$, then $p \le P(H) \le 2p+\Delta(P)$. Moreover, $p \le P(H) \le p+\Delta(P)$ if ∂H is the unique tangent of $D(p)$ at some point from $H \cap \partial D(p)$; in particular, $P(H) = p$ if $\Delta(P) = 0$.*

If $A \in \mathbb{S}^{d-1}$ is a finite set of directions and H is a tangent halfspace of $D_A(p)$, then $P(H) \le 2p + \Delta(P)$ and $p \le P(H) \le p + \Delta(P)$ if ∂H is the unique tangent of $D_A(p)$ at some *point from $H \cap \partial D(p)$. In particular, $P(H) = p$ if $\Delta(P) = 0$.*

The left panel of Figure 4 shows the situation in which the tangent to the directional quantile envelope is unique. For a population distribution with $\Delta(P) = 0$, one can uniquely identify the directional quantile in the direction perpendicular to the tangent. The visual determination of the uniqueness of the tangent may be slightly in the eye of beholder; if this is undesirable, then one may switch to a strictly finite-sample viewpoint in which the directional quantile envelopes of empirical probability distributions are polygons and the uniquely identifiable directional quantile lines are those that contain a boundary segment of the polygon. If the tangent is not unique, the situation shown in the right panel of Figure 4, then the exact identification of the directional quantile line is not possible; nevertheless, the inequality $P(H) \le 2p$ given by Theorem 3 allows at least for its approximate localization (especially when the plotted envelopes are so chosen that p follows a geometric progression with multiplier 1/2, as in the right panel of Figure 2; note that such choice gives approximately equispaced contours for normal distribution in the tail area).

A boundary point of a convex set that admits more than one tangent is called *rough*. It is known—see Theorem 2.2.4 of Schneider (1993)—that such points are quite exceptional and, in particular, for any closed convex set in $\mathbb{R}^2$, the set of rough points is at most countable. Convex, closed subsets of $\mathbb{R}^d$ having no rough points are called *smooth*, consistent with the natural geometric perception of

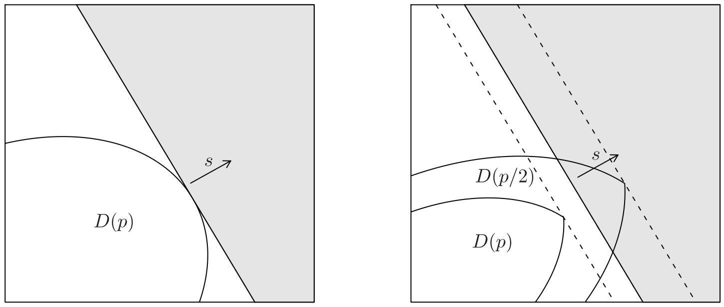

Figure 4. Left panel: if the tangent line to the pth directional quantile envelope is unique, then the tangential halfspace is the pth directional quantile halfspace, in the given direction. Right panel: if the tangent line is nonunique, then this directional quantile halfspace lies between pth and $(p/2)$th directional quantile envelope.

the boundary in this case. If $D(p)$ is smooth, then the collection of its tangent halfspaces is in one-one correspondence with the collection of pth directional quantile halfspaces, with the same boundaries, but in opposite directions.

Although the assumption of smoothness may sound optimistically mild, the examples in Rousseeuw and Ruts (1999) show that distributions with depth contours having a few rough points are not that uncommon. However, it may be argued that all these examples have somewhat contrived flavor, especially when the support of the distribution is some regular geometric figure; it is not unlikely that typical population distributions have smooth depth contours—but we were not able to find a suitable formal condition reinforcing this belief, beyond the somewhat restricted realm of elliptically-contoured distributions. Recall that a distribution is called *elliptic* if it can be transformed by an affine transformation to a circularly symmetric, rotationally-invariant distribution. The following theorem, in particular, confirms the fact mentioned earlier: normal contours allow for the retrieval of *all* directional quantile lines.

Theorem 4. *The directional quantile envelopes of any elliptic distribution are smooth.*

Even if the tangent line at a boundary point of a directional quantile envelope is nonunique, it does not necessarily mean that the information about certain directional quantiles is lost. Although the directional quantile is not retrievable from the envelope *directly*, in a straightforward manner, it may be possible to reconstruct it from the totality of all envelopes. Formally this means that the collection of directional quantile envelopes determines the distribution uniquely. Surprisingly, this plausible property has not yet been rigorously proved in full generality, positive answers have been established only for partial cases:

depth functions uniquely characterize empirical (Struyf and Rousseeuw (1999)), and more generally atomic (Koshevoy (2002)) distributions, and also absolutely continuous distributions with compact support (Koshevoy (2001)). A small step in this direction is the following result of Kong and Zuo (2010) concerning distributions with smooth depth contours.

Theorem 5. *If the directional quantile envelopes $D(p)$ of a probability distribution P in $\mathbb{R}^d$ with contiguous support have smooth boundaries for every $p \in (0, 1/2)$, then there is no other probability distribution with the same directional quantile envelopes.*

6. Invariance, Approximation, and Estimation

Our considerations have been more probabilistic than statistical to this point; to advance the latter, the most straightforward way is to invoke the principle called "naïve statistics" by Hájek and Vorlíčková (1977), "analogy" by Goldberger (1968) and Manski (1988), and "the plug-in principle" by Efron and Tibshirani (1993): apply the general definition to empirical distributions.

From the general point of view, we are interested in the *population quantile information*, directional quantiles of some *population distribution* from which our data come, in some sampling manner,To facilitate theoretical analysis of typical cases, it is often reasonable to posit some assumptions on this distribution; while a membership in a parametric family, or ellipticity may be considered too stringent, continuity assumptions are often acceptable. Our general strategy is to estimate the result of the evaluation of a functional on the population distribution via the application of the same functional to the empirical distribution supported by the data. Specifically, we estimate, for fixed p, the directional quantiles $Q(p, s)$ by $\hat{Q}(p, s)$, and then use these estimates to generate the estimated directional quantile envelope.

Recall that an operator (we use this word to indicate that unlike a "function", an "operator" can be set-valued) assigning a point or a set T in $\mathbb{R}^d$ to a collection of datapoints $x_i \in \mathbb{R}^d$, is called *affine equivariant*, if its value is $BT + b$ when evaluated from the datapoints $Bx_i + b$, for any nonsingular matrix B and any $b \in \mathbb{R}^d$. (If T is a set, then the transformations are performed elementwise.) By Theorem 2, the estimated and population directional quantile envelopes are the level sets of depth applied to the empirical and population distributions, respectively; the properties of depth then imply their affine equivariance. When directional quantiles are estimated by some other means, for instance as a response of a quantile regression, the affine equivariance of the resulting envelopes may be not that clear. Nevertheless, the affine equivariance still holds under mild assumptions on the directional quantile estimators. Recall that an operator that assigns a point, or set of points, T, in $\mathbb{R}$, to a random variable X is called *translation equivariant*, if its value for $X+b$ coincides with $T+b$, and *scale equivariant*, if its value for cX coincides with cT. It is a direct consequence of the definition that the directional quantile operator $Q(p,\cdot)$ is translation and scale equivariant for any $0<p<1$ (and for the "inf" version; for every other version, the equivariance has to be checked individually—usually a straightforward task).

Theorem 6. *Suppose that directional quantile estimators $\hat{Q}(p,s)$ are translation and scale equivariant for all $s \in \mathbb{S}^{d-1}$ and fixed p. Then the directional quantile envelope generated by these estimators is affine equivariant.*

The next question we investigate is how quantile directional envelopes behave when not determined exactly, but only approximately. The numerical motivation for this stems from the fact that in practice we are not able to take all directions to construct the directional quantile envelope—what is constructed is rather an approximate envelope $D_A(p)$, and we are interested in the quality of this approximation. Obviously, $D(p) \subseteq D_A(p)$; moreover, we believe that a decent collection, A, of directions that reasonably fill $\mathbb{S}^{d-1}$ makes the approximation satisfactory. While the experimental evidence does not contradict this belief—for the left panel of Figure 2 we used only 100 uniformly spaced directions, for the right panel of Figure 2 we took 1009, and hardly any difference can be seen for $p=0.1$—some theoretical support would be desirable too. The statistical motivation for the investigation of the approximation effects comes from the fact that our directional quantiles are typically not the "true", but "estimated" ones. If we believe that this estimation is consistent—we can show, in some customary probabilistic framework, that estimates become more and more precise, say, with growing sample size—then we are again interested whether this consistent behavior of individual directional quantiles translates into something analogous for their envelopes.

Theorem 7. *Suppose $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$ is a sequence of closed sets with union dense in a closed set $A \subseteq \mathbb{S}^{d-1}$ that is not contained in any closed halfspace whose boundary contains the origin. If, for every sequence $s_n \in A_n$ that converges to $s \in A$, $q_n(s_n)$ converges to $q(s)$, then $\bigcap_{s\in A_n} H(s,q_n(s))$ converges to $\bigcap_{s\in A} H(s,q(s))$ in the Pompeiu-Hausdorff distance—provided either the limit set is the closure of its interior, or is a singleton and the sets in the sequence are nonempty.*

We illustrate the use of this theorem on two examples. In the first, we take $q_n(s)=q(s)=Q(p,s)$; the successive approximations $D_{A_1}(p) \supseteq D_{A_2}(p) \supseteq \ldots$ then approach $D_A(p)$ in the Pompeiu-Hausdorff distance. Typically, A_n are finite, while $A=\mathbb{S}^{d-1}$; the only requirements is that the directional quantiles $Q(p,s)$ depend on s in a continuous way—for instance, P satisfies the assumptions of

Theorem 1. The second example leads to a proof of consistency, similar to those given by He and Wang (1997), of $\hat{D}_n(p)$ to $D(p)$, when the $\hat{D}_n(p)$ arise via applying the definition of directional quantile envelopes to empirical distributions that converge weakly almost surely to the sampled population distribution P; the required assumptions are those of Theorem 1, continuous or bounded support of P, and the nondegeneracy of the limit $D(p)$ (in general we cannot guarantee that the $\hat{D}_n(p)$ are nonempty). The Skorokhod representation then yields random variables X_n converging almost surely to random variables X, such that the laws of X_n and X are the corresponding empirical distributions and P, respectively; Theorem 1 then implies the convergence assumption required by Theorem 7.

To obtain some idea of the magnitude of the approximation error, we proceed as follows (for simplicity, we limit our scope to the two-dimensional setting). Let $d \in \partial D$. The directions of all tangents of a convex set D at d generate a convex cone, $T_D(d)$. Let $c_D(d)$ be the maximal cosine between its two directions, the cosine of the maximal angle between two extremal normalized directions in $T_D(d)$,

$$c(d) = \sup\left\{\frac{s^{\top} t}{\|s\|\|t\|} : s, t \in T_D(d)\right\} = \sup\left\{s^{\top} t \colon s, t \in T_D(d) \cap \mathbb{S}^{d-1}\right\}.$$

In fact, this cosine is the same as the maximal cosine of the directions in the normal cone $N_D(d)$; see Rockafellar and Wets (1998), Chapter 6. We can see that $c_D(d) \le 1$, the equality holding if and only if $T_D(d)$ consists of single direction—when D has a unique tangent at d. Let $\kappa_D = \sup_{d \in \partial D} \sqrt{2/(1 + c(d))}$, the reciprocal of the cosine of the *half* of the maximal angle between directions in the tangent cone. Apparently, $\kappa_D \ge 1$, the equality holding true for smooth D. On the other hand, κ_D can be $+\infty$ for degenerate D, the sets with empty interior.

Theorem 8. *Let $A \subseteq \mathbb{S}^1$ be a set of directions, and let $\hat{q}(s)$ and $q(s)$ be two functions on A. If $\hat{D} = \bigcap_{s \in A} H(s, \hat{q}(s))$ and $D = \bigcap_{s \in A} H(s, q(s))$ are nondegenerate, then $\kappa_{\hat{D}}$ and κ_D are finite and*

$$d\big(\hat{D}, D\big) \le \max\{\kappa_{\hat{D}}, \kappa_D\} \sup_{s \in A} |\hat{q}(s) - q(s)|,$$

where d denotes the Pompeiu-Hausdorff distance.

7. Directional Quantile Envelopes Beyond Simple Location Setting

So far, our methodology was demonstrated in the simple location setting, the situations in which there are no covariates and the estimation is performed via the application of the quantile operators to empirical distributions. In this section, we show how the directional definition can be used in more sophisticated constructions.

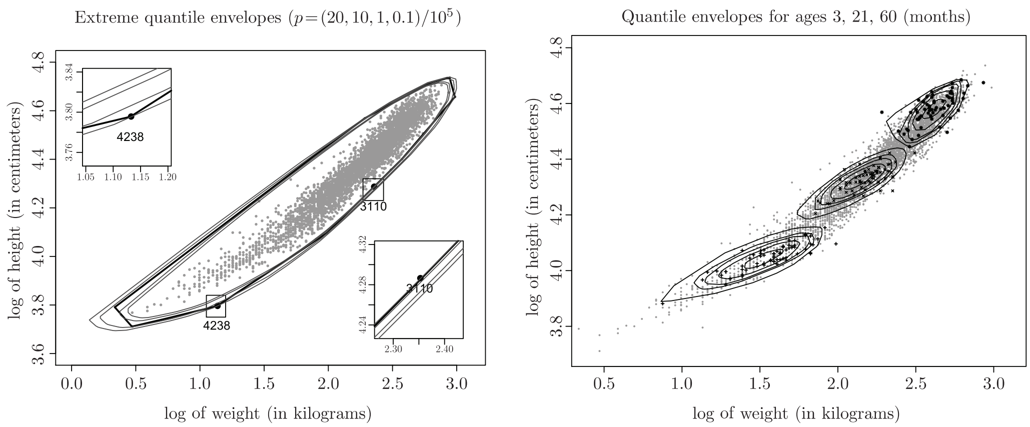

Figure 5. Left panel: directional extremal quantiles, derived from the corresponding univariate analogs, and the convex hull, the empirical extremal quantile. Right panel: imagine an animation in which the directional quantile envelopes slowly ascend upward along the data cloud, demonstrating the dependence on the increasing covariate, age.

The first application is to extreme quantiles. It is apparent that this type of analysis calls for other than empirical estimators of population quantiles. If , say, 100 observations are available, then their maximum, the pth empirical quantile for any $p > 0.99$, may not be found satisfactory for estimating a threshold with exceedance probability less than, say, 0.001. This is well known, and we do not propose any new take on the subject, nor side with any of the approaches that can be found in Beirlant et al. (2004), Reiss and Thomas (2007), Resnick (2007), and the references given there. Our only message here is that once the problem is satisfactorily handled in the univariate case, the directional philosophy allows for an immediate extension to the multivariate setting. The result can be seen in the left panel of Figure 5. The estimated extreme quantiles, for $p = 10^{-6}$, 10^{-5}, 10^{-4}, and 2×10^{-4} are confronted with the convex hull of the data, the empirical estimate for any $p \leq (2.33)10^{-4}$. The plot seems to provide some information about the extent of extremality of the points labeled by 3,110 and 4,238; a closer inspection reveals that 3,110 lies on the (2×10^{-3})th directional quantile envelope, with 4,238 on the (10^{-6})th one. The worth of this information crucially depends on the properties of the estimates of extreme quantiles in the univariate case; in this case we chose those we found in the R package `evir`, McNeil and Stephens (2007), due to their nonparametric flavor and the ready availability of their implementation.

Our second application is to bivariate growth charts. While the relevant proposal of Wei (2008), roughly characterized as quantile regression in polar coordinates, is conceptually capable of delivering interpretable contours, its practical

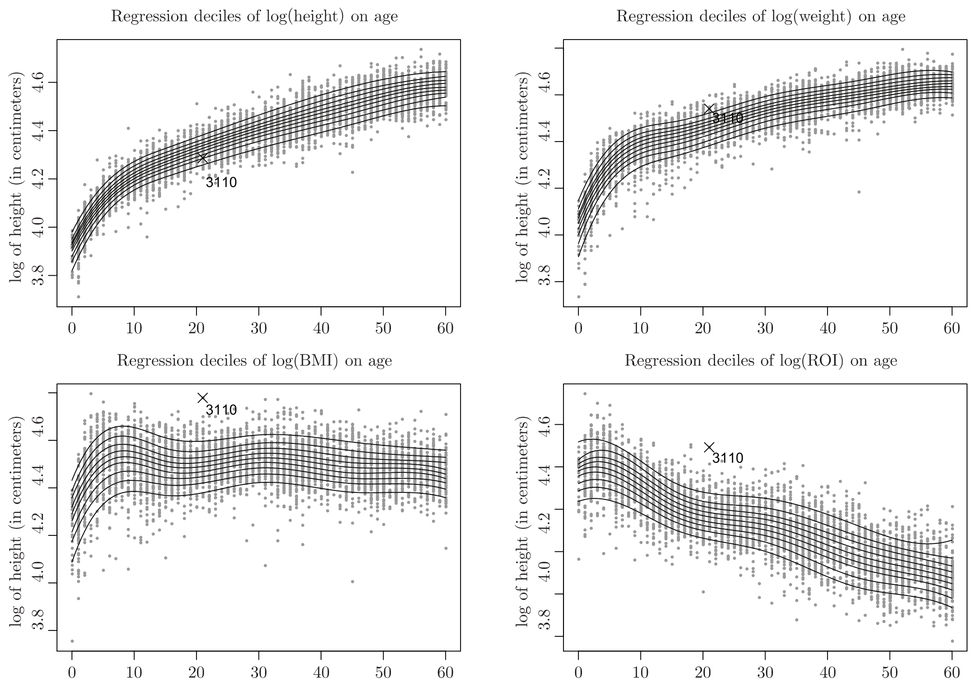

Figure 6. The “growth charts”, quantiles of various linear combinations of the primary variables, regressed on the covariate, age.

application is plagued by its considerable dependence on the underlying nonparametric regression methodology, implemented in the R package `cobs` by Ng and Maechler (2006), and by the necessity to select the origin of the coordinate system—a somewhat ad hoc task that nevertheless seriously influences the result. These and other shortcomings (in particular, the tendency of the estimated contours to intersect themselves) led to the pursuit of the approach we propose here; some precursor ideas were outlined already by Salibian-Barrera and Zamar (2006). It should be noted that our way of indexing is different from that of Wei (2008), so her approach may be, after all, viewed as complementary rather than alternative to ours.

Figure 6 shows the deciles of several projections of the vector response, consisting of the logarithms of weight and height, regressed on the covariate, age in months. While such “growth charts” facilitate useful insights, the user may wish to confront them from a directional perspective—in a related covariate-dependent context. Such a desire stumbles upon the inevitable fact that our graphical universe is two-dimensional; animations and interactive graphics are certainly possible, but in the traditional setting we can merely opt for the plotting of the directional quantiles for some fixed value(s) of the covariate—as in the right panel of Figure 5 that shows the predicted envelopes for three values of age (selected so that the resulting envelopes do not overplot, rather than pursuing any other objective). The highlighted datapoints represent the subjects of the particular age. If we computed directional quantile envelopes from these points separately, the resulting contours would be rougher, and would vary from one value of age to another; the contours given in the right panel of Figure 5 borrow strength from other ages, constructing quantile envelopes from a number of quantile regressions, like those seen in Figure 6.

Once again, our focus is on how quantile regression blends into directional quantile philosophy; from this perspective, our rendering of nonparametric quantile regression rather avoided than explored potential challenges, and we refer to Koenker (2005) and references there for the fine aspects of the methodology. In view of Theorem 6, our main concern is whether the estimates are translation (regression) and scale equivariant, to yield affine equivariant envelopes—which is true if the fits in Figure 6 and the right panel of Figure 5 are obtained via regression splines, the methodology used by Wei et al. (2005) in their paper on growth charts. We used the automated knot selection furnished by the R package `splines`, R Development Core Team (2007), and fitted quantile regressions by the R package `quantreg`, Koenker (2007). The smoothing parameter was selected by eyeballing the plots included in Figure 6, and then adopting a universal smoothing parameter for all directions in the right panel of Figure 5. We are aware of the possible shortcomings in certain engineering details—for instance that unlike in our situation, one can easily imagine data exhibiting more signal-to-noise in certain directions than in others, the fact that would have to be reflected in varying smoothing parameters; we hope to address these problems in future research, as well as explore alternative possibilities for nonparametric quantile regression in the construction of growth charts.

8. Final Remarks

This paper is a shortened version of the preprint of Kong and Mizera (2008), dropping any mention of impasses like "quantile biplots" (so that they are not confused with the concepts that we really champion); space considerations led us also to omit the discussion of fine aspects of growth charts; finally, we do not discuss any computational details, as those were rendered obsolete by developments—we hope to address computational aspects in a separate publication. We still maintain that our objective was not to propose any "multivariate quantile" generalization of the univariate concept, akin to those reviewed in Serfling (2002).

We believe that directional quantile envelopes—which are, essentially, depth contours—are a possible way to condense directional quantile information, the information carried by the quantiles of projections. In typical circumstances, they allow for relatively faithful and straightforward retrieval of the directional quantile information; the methodology offers straightforward probabilistic interpretations, and the estimated quantile envelopes are affine equivariant under mild equivariance assumptions on the estimators of directional quantiles. Most importantly, the directional interpretation can be adapted to elaborate frameworks requiring more sophisticated quantile estimation methods than evaluating quantiles for empirical distributions, including estimation of extreme quantiles and directional quantile regression.

Acknowledgement

We are indebted to Ying Wei for turning our attention to multivariate growth charts, as well as for many insights in Wei (2008), and to Roger Koenker for valuable discussions. The directional approach to depth contours was pioneered in the unpublished master thesis of Benoît Laine, as reported by Koenker (2005)—in, however, quite significantly more complicated version fitting not directional quantiles, but directional quantile regressions. Another important forerunner was Salibian-Barrera and Zamar (2006). This research was supported by the Natural Sciences and Engineering Research Council of Canada; some of the results originate from the doctoral dissertation of Kong (2009).

Appendix. Proofs

We define the pth *directional quantile set* to be the quantile set of the corresponding projection:

$$\mathcal{Q}(p,s)=\mathcal{Q}(p,s,X)=\mathcal{Q}(p,s^{\mathsf T}X).$$

Proof of Theorem 1. Since quantile sets are bounded intervals, it is sufficient to prove the convergence of their endpoints to $\inf \mathcal{Q}(p,s^{\mathsf T}X)=\inf\{u\colon \mathbb{P}[s^{\mathsf T}X\le u]\ge p\}$ and $\sup \mathcal{Q}(p,s^{\mathsf T}X)=\sup\{u\colon \mathbb{P}[s^{\mathsf T}X\ge u]\le (1-p)\}$.

Suppose the support of X is bounded and let $q=\inf \mathcal{Q}(p,s^{\mathsf T}X)$. We have $\mathbb{P}[s^{\mathsf T}X\le q]\ge p$ and $\mathbb{P}[s^{\mathsf T}X\le q-\varepsilon]<p$. If the support of the distribution of X is bounded, we have $\|X\|\le M$ almost surely; by the Schwarz inequality, $|(s-s_n)^TX|\le M\|s-s_n\|$ and therefore

$$p\le \mathbb{P}[s^{\mathsf T}X\le q]=\mathbb{P}[s_n^{\mathsf T}X\le q-(s-s_n)^{\mathsf T}X]\le \mathbb{P}[s_n^{\mathsf T}X\le q+M\|s-s_n\|],$$

which means that $\inf \mathcal{Q}(p,s_n^TX)\le q+M\|s-s_n\|$. In a similar fashion, we obtain that $\inf \mathcal{Q}(p,s_n^TX)\ge q-M\|s-s_n\|-\varepsilon$, due to $\mathbb{P}[s_n^{\mathsf T}X\le q-M\|s-s_n\|-\varepsilon]\le \mathbb{P}[s^{\mathsf T}X\le q-\varepsilon]<p$. Letting $\varepsilon\to 0$, we obtain $q-M\|s-s_n\|\le \inf \mathcal{Q}(p,s_n^{\mathsf T}X)\le q+M\|s-s_n\|$, and therefore $\inf \mathcal{Q}(p,s_n^{\mathsf T}X)\to \inf \mathcal{Q}(p,s^{\mathsf T}X)$

and thus also $Q(p, s_n, X_n)$ to $Q(p, s, X)$. The convergence of $\sup \mathcal{Q}(p, s_n^\tau X) \to \sup \mathcal{Q}(p, s^\tau X)$ is proved analogously.

If the support of the distribution of X is contiguous, then all directional quantile sets in the limit are singletons. Pompeiu-Hausdorff convergence then follows from the "outer convergence" of quantile sets in the sense of Rockafellar and Wets (1998), see also Mizera and Volauf (2002): any limit point, x, of any sequence $x_n \in \mathcal{Q}(p, s_n, X_n)$ lies in $\mathcal{Q}(p, s, X)$. This can be easily seen in an elementary way, observing that $x_n \in \mathcal{Q}(p, s_n, X_n)$ entails

$$p \le \limsup_{n\to\infty} \mathbb{P}[s_n^\tau X_n \le x_n] \le \mathbb{P}[s^\tau X \le x],$$
$$1 - p \le \limsup_{n\to\infty} \mathbb{P}[s_n^\tau X_n \ge x_n] \le \mathbb{P}[s^\tau X \ge x].$$

Since under the contiguous support assumption the quantiles are unique, this second part of the theorem holds true for every quantile version.

Proof of Theorem 2. If $y \in D(p)$, then $y \in H(p, s)$ for every $s \in \mathbb{S}^{d-1}$ and thus $P(\{x\colon s^\tau x \ge s^\tau y\}) \ge p$ for all $s \in \mathbb{S}^{d-1}$; therefore $d(x) \ge p$. Conversely, if $d(y) \ge p$, then for every $s \in \mathbb{S}^{d-1}$ we have $P(\{x\colon s^\tau x \ge s^\tau y\}) \ge p$. It follows that $s^\tau y \ge Q(p, s)$ and thus $y \in H(p, s)$. Hence $y \in D(p)$. As mentioned, this theorem is true only for the "inf" definition, other quantile versions give smaller envelopes.

Proof of Theorem 3. See Kong and Zuo (2010).

Proof of Theorem 4. By rotational invariance, the directional quantile envelopes of any circularly symmetric distribution are circles; since elliptic distributions are those that can be transformed to the circular symmetric ones by an affine transformation, the theorem follows from their affine equivariance (and holds true for any quantile version).

Proof of Theorem 5. See Kong and Zuo (2010).

Proof of Theorem 6. Let B be a nonsingular matrix and b a vector. First, we verify the transformation rule for the supporting halfspace of the directional quantile: for every $s \in \mathbb{S}^{d-1}$ and every $p \in (0, 1)$,

$$H(B^* s/\|B^* s\|, Q(p, s, BX + b)) = BH(s, Q(p, s, X)) + b, \tag{A.1}$$

where $B^* = (B^{-1})^\tau$. If B is orthogonal, then $B^* = B$, and if B is diagonal (more generally, symmetric), then $B^* = B^{-1}$. Indeed, the equation satisfied by x in $BH(s, (Q, p, s, X))$, $s^\tau \left(B^{-1}x\right) \le Q(p, s, X)$, is equivalent to $((B^{-1})^\tau s)^\tau x = (B^* s)^\tau x \le Q(p, s, X)$. The norm of s is one, but not necessarily that of $B^* s$; therefore, we divide both sides by $\|B^* s\|$, to write

$$(1/\|B^* s\|)(B^* s)^\tau x \le (1/\|B^* s\|)Q(p, s, X). \tag{A.2}$$

By the scale equivariance of the quantile operator, and by the relationship $Q(p,s,AX) = Q(p,A^\tau s,X)$ that follows directly from the definition, the right-hand side of (A.2) is $Q(p,s,X/\|B^* s\|) = Q(p,s/\|B^* s\|,X) = Q(p,B^* s/\|B^* s\|, BX)$. Since the transformation $BX + b$ is one-to-one, the transformed intersection of halfspaces is the intersection of transformed halfspaces. Therefore, the transformed directional quantile envelope is, by (A.1),

$$\bigcap_{s\in\mathbb{S}^{d-1}} (BH(p,s,X)+b) = \bigcap_{s\in\mathbb{S}^{d-1}} H(p,B^* s/\|B^* s\|, BX+b).$$

The proof is concluded by observing that $s \mapsto B^* s/\|B^* s\|$, where $B^* = (B^{-1})^\tau$, is a one-to-one transformation of $\mathbb{S}^{d-1}$ onto itself—as can be seen by the direct verification involving its inverse, $t \mapsto B^\tau t/\|B^\tau t\|$. The proof is the same for any quantile version.

Proof of Theorem 7. To prove convergence with respect to Pompeiu-Hausdorff distance, we exploit the following. The sequence $\bigcap_{s\in A_n} H(s,q_n(s))$, together with the limit $\bigcap_{s\in A} H(s,q(s))$, is contained in a bounded set starting from some n, since the sets A_n are approaching a dense set in A, and the latter is not contained in any halfspace whose boundary contains the origin; therefore this property is shared by A_n starting from some n, which means that $\bigcap_{s\in A_n} H(s,\inf_{k\geq n} q_n(s))$ is the desired bounded set. For uniformly bounded sequences, the convergence in Pompeiu-Hausdorff distance follows from the convergence in Painlevé-Kuratowski sense; see Rockafellar and Wets (1998), 4.13. The latter means that a general sequence of sets K_n converges to K if (i) every limit point of any sequence $x_n \in K_n$ lies in K, and (ii) every point from K is a limit of a sequence $x_n \in K_n$; see also Mizera and Volauf (2002).

For sequences of closed sets with "solid" limits (sets that are closures of their interiors), the Painlevé-Kuratowski convergence follows from the "rough" convergence, defined by Lucchetti, Salinetti, and Wets (1994) to require (i) and (ii)' every limit point of every sequence $y_n \in (\operatorname{int} K_n)^c$ is in $(\operatorname{int} K)^c$. The inner convergence requirement of Painlevé-Kuratowski definition is thus replaced by the outer convergence for "closed complements"; see also Lucchetti, Torre, and Wets (1993).

Suppose that $y \in \operatorname{int} K$. Then y belongs to all but finitely many K_n; otherwise, there would be a subsequence n_i such that $y \in (\operatorname{int} K_{n_i})^c$ and, by (ii)', $y \in (\operatorname{int} K)^c$. Hence, every y from the relative interior of K is a limit of an (eventually constant) sequence $y_n \in K_n$. To obtain (ii) for every $x \in K$, consider a sequence y_k of points from (nonempty) $\operatorname{rint} K$ such that $y_n \to y$; the desired sequence x_n is then obtained by a "diagonal selection": for every y_k, there is n_k such that $y_k \in K_i$ for every $i \geq k$; set $x_n = y_k$ for every $n_k \leq n < n_{k+1}$.

Thus, it is sufficient to prove (i) and (ii)'. Suppose that x is a limit point of a sequence $x_n \in \bigcap_{s\in A_n} H(s, q_n(s))$. Then there is a subsequence such that $s_n^{\mathsf{T}} x_n \geq q_n(s_n)$ for every $s_n \in A_n$; every $s \in A$ is a limit of a sequence $s_n \in A_n$, therefore the assumptions of the theorem imply that $s^{\mathsf{T}} x \geq q(s)$; hence $x \in \bigcap_{s\in A} H(s, q(s))$. This proves (i) and the theorem for the singleton case, since then the Painlevé-Kuratowski convergence is implied by (i) once the sets in the sequence are nonempty.

Suppose now that x is a limit point of a sequence $x_n \in (\operatorname{int}\bigcap_{s\in A_n} H(s, q_n(s)))^c$, that is, a limit of some subsequence of x_n. Every such x_n satisfies $s_n^{\mathsf{T}} x_n \leq q_n(s_n)$ for some $s_n \in A_n$. By the compactness of A, there is $s \in A$ that is a limit of a subsequence of s_n; passing to the limit along the appropriate subsequences, we obtain that $s^{\mathsf{T}} x \leq q(s)$, by the assumptions of the theorem. This means that $x \in \left(\operatorname{int}\bigcap_{s\in A} H(s, q(s))\right)^c$.

Proof of Theorem 8. As $\hat{D}$ and D are compact convex sets, we have $d(\hat{D}, D) = d(\partial\hat{D}, \partial D)$. Let $\varepsilon = \sup_{s\in A} |\hat{q}(s) - q(s)|$; we will show that for any $x \in \partial D$, $d(x, \partial\hat{D}) \leq \kappa_D \varepsilon$. Let $\tilde{q}(s) = q(s) - \varepsilon$ and $\tilde{D} = \bigcap_{s\in A} H(s, \tilde{q}(s))$.

For simplicity, we assume that $\tilde{D}$ is also nondegenerate. We have that $\tilde{D} \subseteq \hat{D}$, and also $\tilde{D} \subseteq D$, the latter set being congruent to $\tilde{D}$. If $\kappa_D(x) > 1$, then x is a vertex of D. Since $d(x, \tilde{x}) = \kappa_D(x)\varepsilon$, where $\tilde{x}$ is the corresponding congruent vertex in $\partial\tilde{D}$, it follows that $d(x, \partial\tilde{D}) \leq \kappa_D \varepsilon$. If $\kappa_D(x) = 1$, then by Theorem 24.1 of Rockafellar (1996) there exists a sequence $x_n \neq x$, $x_n \in \partial D$, such that $x_n \to x$ and $s_n \to s$, $\kappa_D(x_n) = 1$, where s_n and s are the directions of the tangent lines passing through x_n and x, respectively. There are two possibilities.

If there is N such that $s_n = s$ for any $n > N$, then there must be two points, denoted by y_1 and y_2, in $\partial H(s, q(s)) \cap \partial D$ such that $\kappa_D(y_1) > 1$ and $\kappa_D(y_2) > 1$. That is, y_1 and y_2 are two vertices of D and there is no other vertex between y_1 and y_2 of D. Suppose that $\tilde{y}_1$ and $\tilde{y}_2$ are points congruent to them on $\tilde{D}$; then $\tilde{y}_1$ and $\tilde{y}_2$ are two vertices of $\tilde{D}$ and there is no other vertex between $\tilde{y}_1$ and $\tilde{y}_2$ of $\tilde{D}$ as well. In other words, we have a trapezoid with vertices y_1, y_2, $\tilde{y}_1$ and $\tilde{y}_2$, and x lies on one of the bases. A simple geometric calculation then shows the existence of a point, y, lying on the base constructed by $\tilde{y}_1$ and $\tilde{y}_2$, such that $d(x, y) \leq \max\{\kappa_D(y_1), \kappa_D(y_2)\}\varepsilon$, that is, $d(x, \partial\tilde{D}) \leq \kappa_D \varepsilon$.

Suppose that there is an infinite subsequence of s_n such that $s_n \neq s$. Let $\partial\tilde{D}$ by $\tilde{x}_n$ and $\tilde{x}$ be the congruent counterparts of x_n and x, respectively; let s_n and s be the corresponding directions. Let $y_n = \partial H(s_n, q(s_n)) \cap \partial H(s, q(s))$ and $\tilde{y}_n = \partial\tilde{H}(s_n, q(s_n)) \cap \partial\tilde{H}(s, q(s))$. We have that $y_n \to x$, $\tilde{y}_n \to \tilde{x}$, and $d(y_n, \tilde{y}_n) = \sqrt{2}\varepsilon/\sqrt{1 + s_n^T s}$. As $d(y_n, \tilde{y}_n) \to d(x, \tilde{x})$ and $\sqrt{2}\varepsilon/\sqrt{1 + s_n^T s} \to \varepsilon$, we have $d(x, \tilde{x}) = \varepsilon$, which means $d(x, \partial\tilde{D}) \leq \kappa_D \varepsilon$ again.

Taking into account that $\tilde{D} \subseteq \hat{D}$, we obtain that $d(x, \partial\hat{D}) \leq \kappa_D \varepsilon$ for any $x \in \partial D$. The theorem follows from this and the symmetric inequality, $d(x, \partial D) \leq \kappa_{\hat{D}} \varepsilon$ for any $x \in \partial\hat{D}$, which can established in an analogous way.

References

Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. (2004). *Statistics of Extremes: Theory and Applications*. Wiley, Chichester.

Donoho, D. L. and Gasko, M. (1992). Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *Ann. Statist.* **20**, 1803-1827.

Edgeworth, F. Y. (1886). Problems in probabilities. *London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science,* 5*th series* **22**, 371-384.

Edgeworth, F. Y. (1893). Exercises in the calculation of errors. *London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science,* 5*th series* **36**, 98-111.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to Bootstrap*. Chapman and Hall, New York.

Eubank, R. L. (1986). Quantiles. In *Encyclopedia of Statistical Sciences, Volume 7* (Edited by S. Kotz, N. L. Johnson and C. B. Read), 424-432. Wiley, New York.

Evans, M. (1982). Confidence bands for bivariate quantiles. *Commun. Statist.-Theor. Meth.* **11**, 1465-1474.

Frohne, I. and Hyndman, R. J. (2004). `quantile`. A function in R starting from version 2.0.0, `http://www.r-project.org`.

Galton, F. (1888-1889). Co-relations and their measurement, chiefly from anthropometric data. *Proc. Royal Soc. London* **45**, 135-145.

Goldberger, A. S. (1968). *Topics in Regression Analysis*. Macmillan, New York.

Hájek, J. and Vorlíčková, D. (1977). *Mathematical Statistics*. SPN, Praha. [in Czech].

He, X. and Wang, G. (1997). Convergence of depth contours for multivariate datasets. *Ann. Statist.* **25**, 495-504.

Hodges, Jr, J. L. (1955). A bivariate sign test. *Ann. Math. Statist.* **26**, 523-527.

Hyndman, R. J. and Fan, Y. (1996). Sample quantiles in statistical packages. *Amer. Statist.* **50**, 361-365.

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press, Cambridge.

Koenker, R. (2007). `quantreg`: Quantile Regression. R package version 4.10, `http://www.r-project.org`.

Kong, L. (2009). On multivariate quantile regression: Directional approach and application with growth charts. *PhD thesis*, University of Alberta.

Kong, L. and Mizera, I. (2008). Quantile tomography: using quantiles with multivariate data. *ArXiv preprint* arXiv:0805.0056v1.

Kong, L. and Zuo, Y. (2010). Smooth depth contours characterize underlying distribution. *J. Multivariate Anal.* **101**, 2222-2226.

Koshevoy, G. A. (2001). Projections of lift zonoids, the Oja depth and Tukey depth. Preprint.

Koshevoy, G. A. (2002). The Tukey depth characterizes the atomic measure. *J. Multivariate Anal.* **83**, 360-364.

Lucchetti, R., Salinetti, G. and Wets, R. J.-B. (1994). Uniform convergence of probability measures: topological criteria. *J. Multivariate Anal.* **51**, 252-264.

Lucchetti, R., Torre, A. and Wets, R. J.-B. (1993). Uniform convergence of probability measures: topological criteria. *Canad. Math. Bull.* **36**, 197-208.

Manski, C. F. (1988). *Analog Estimation Methods in Econometrics*. Chapman and Hall, New York.

Massé, J.-C. and Theodorescu, R. (1994). Halfplane trimming for bivariate distributions. *J. Multivariate Anal.* **48**, 188-202.

McNeil, A. and Stephens, A. (2007). `evir`: Extreme Values in R. R package version 1.5, `http://www.maths.lancs.ac.uk/~stephena/`.

Mizera, I. (2002). On depth and deep points: A calculus. *Ann. Statist.* **30**, 1681-1736.

Mizera, I. and Volauf, M. (2002). Continuity of halfspace depth contours and maximum depth estimators: diagnostics of depth-related methods. *J. Multivariate Anal.* **83**, 365-388.

Ng, P. T. and Maechler, M. (2006). `cobs`: COBS - Constrained B-splines. R package version 1.1-3.5, `http://wiki.r-project.org/rwiki/doku.php?id=packages:cran:cobs`.

Parzen, E. (2004). Quantile probability and statistical data modeling. *Statist. Sci.* **19**, 652-662.

R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. `http://www.R-project.org`.

Reiss, R.-D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values*. Birkhäuser Verlag, Basel.

Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.

Rockafellar, R. T. (1996). *Convex Analysis*. Princeton University Press, Princeton.

Rockafellar, R. T. and Wets, R. J.-B. (1998). *Variational analysis*. Springer-Verlag, Berlin.

Rousseeuw, P. J. and Ruts, I. (1999). The depth function of a population distribution. *Metrika* **49**, 213-244.

Salibian-Barrera, M. and Zamar, R. (2006). Discussion of Conditional growth charts by Y. Wei and X. He. *Ann. Statist.* **34**, 2113-2118.

Schneider, R. (1993). *Convex bodies: The Brunn-Minkowski theory*. Cambridge University Press, Cambridge.

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

Shorack, G. R. (2000). *Probability for Statisticians*. Springer-Verlag, New York.

Struyf, A. and Rousseeuw, P. J. (1999). Halfspace depth and regression depth characterize the empirical distribution. *J. Multivariate Anal.* **69**, 135-153.

Tukey, J. W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians*, Vol. **2**, 523-531. Canad. Math. Congress, Quebec.

Wei, Y. (2008). An approach to multivariate covariate-dependent quantile contours with application to bivariate conditional growth charts. *J. Amer. Statist. Assoc.* **103**, 397-409.

Wei, Y., Pere, A., Koenker, R. and He, X. (2005). Quantile regression methods for reference growth charts. *Statist. in Medicine* **25**, 1369-1382.

Department of Mathematical and Statistical Sciences, University of Alberta, CAB 632, Edmonton, Alberta, T6G 2G1 Canada.

E-mail: lkong@ualberta.ca

Department of Mathematical and Statistical Sciences, University of Alberta, CAB 632, Edmonton, Alberta, T6G 2G1 Canada.

E-mail: mizera@stat.ualberta.ca

(Received October 2010; accepted September 2011)