On a Geometric Notion of Quantiles for Multivariate Data

Probal CHAUDHURI

An extension of the concept of quantiles in multidimensions that uses the geometry of multivariate data clouds has been considered. The approach is based on blending as well as generalization of the key ideas used in the construction of spatial median and regression quantiles, both of which have been extensively studied in the literature. These geometric quantiles are potentially useful in constructing trimmed multivariate means as well as many other L estimates of multivariate location, and they lead to a directional notion of central and extreme points in a multidimensional setup. Such quantiles can be defined as meaningful and natural objects even in infinite-dimensional Hilbert and Banach spaces, and they yield an effective generalization of quantile regression in multiresponse linear model problems. Desirable equivariance properties are shown to hold for these multivariate quantiles, and issues related to their computation for data in finite-dimensional spaces are discussed. $n^{1/2}$ consistency and asymptotic normality of sample geometric quantiles estimating the corresponding population quantiles are established after deriving a Bahadur-type linear expansion. The sampling variation of geometric quantiles is carefully investigated, and estimates for dispersion matrices, which may be used in developing confidence ellipsoids, are constructed. In course of this development of sampling distributions and related statistical properties, we observe several interesting facts, some of which are quite counterintuitive. In particular, many of the intriguing properties of spatial medians documented in the literature appear to be inherited by geometric quantiles.

KEY WORDS: Bahadur representation; Geometric quantiles; L estimation in multidimension; Multiresponse quantile regression; $n^{1/2}$-consistent estimate; Spatial median; Trimmed multivariate mean.

1. INTRODUCTION

Quantiles of univariate data are frequently used in the construction of popular descriptive statistics like the median, the interquartile range, and various measures of skewness and kurtosis based on percentiles. They are also potentially useful in robust estimation of location (e.g., in the construction of L estimates). Regression quantiles (see Efron 1991 and Koenker and Bassett 1978), which are nothing but generalizations of quantiles in a regression setup with a univariate response, have been used in robust estimation of parameters in linear models (see Chaudhuri 1992b and Koenker and Portnoy 1987). Lack of objective basis for ordering multivariate observations is a major problem in extending the notion of quantiles in multidimensions. In a classic paper, Barnett (1976) surveyed several possible techniques for ordering multivariate observations (see also Plackett 1976 and Reiss 1989). In the last decade, Eddy (1982, 1983, 1985) proposed an approach for defining quantiles for multivariate data using certain nested sequence of sets, and Brown and Hettmansperger (1987, 1989) introduced a notion of bivariate quantiles based on Oja's criterion function that arises in the definition of Oja's simplex median (see Oja 1983). Very recently, Abdous and Theodorescu (1992) and Kim (1992) have made some attempts to define quantiles for random vectors, and Einmahl and Mason (1992) have extensively studied certain stochastic processes, which may be viewed as generalizations of the univariate quantile process (see also Pyke 1975, 1984, 1985).

Probal Chaudhuri is Member of Scientific Staff, Division of Theoretical Statistics and Mathematics, Indian Statistical Institute, Calcutta 700035, India. The research presented here began when the author was working at the Department of Statistics, University of Wisconsin-Madison while on leave from Indian Statistical Institute during the 1992–1993 academic year. The research was partially supported by a grant from Indian Statistical Institute. The author thanks Gib Bassett, Roger Koenker, Hannu Oja, and Steve Stigler for many stimulating comments and two anonymous referees, an anonymous associate editor, and Myles Hollander for several helpful suggestions.

All of these attempts are valuable contributions toward multidimensional generalization of univariate quantiles. But something that seems to have received either very little or almost no attention in the existing literature is a comprehensive development of the statistical properties of sample multivariate quantiles that are relevant while using them to analyze data and in making statistical inference about population quantiles. Often the authors (e.g., Abdous and Theodorescu 1992; Barnett 1975; Brown and Hettmansperger 1987, 1989; Eddy 1982, 1983, 1985; Kim 1992) concentrated on introducing certain descriptive statistics that generalize the concept of univariate quantiles or order statistics in the multivariate setup, and they did not spend much effort on exploring the sampling distributions and other properties of multivariate quantiles viewed as estimates of their population analogues. The main emphasis of Einmahl and Mason (1992) and Pyke (1975, 1984, 1985) is on constructing certain stochastic processes and studying their limiting behavior. The processes considered by Einmahl and Mason (1992) are actually real valued in nature, as they are defined through certain real-valued set functions. The proposal of Brown and Hettmansperger (1987, 1989), as well as one of the suggestions of Kim (1992), are based on determinants of matrices formed by random vectors. As a result, their approaches are limited to finite-dimensional spaces and do not have any natural generalization for infinite-dimensional spaces. On the other hand, Abdous and Theodorescu (1992) and Babu and Rao (1988) have explored certain quantiles of random vectors that are defined through the coordinate variables. Consequently, these vectors lack some desirable geometric properties (e.g., they are not rotationally equivariant). It will be appropriate to note here that Evans (1982) considered quantiles of a bivariate normal distribution, and

Journal of the American Statistical Association
June 1996, Vol. 91, No. 434, Theory and Methods

his quantiles are boundary surfaces of certain confidence ellipsoids centered at the mean of the distribution.

The purpose of this article is to investigate a notion of quantiles based on the geometric configuration of multivariate data clouds. These geometric quantiles arise as natural generalizations of the multivariate spatial median considered by Brown (1983), Gower (1974), Haldane (1948), and many others, who used the names "spatial median," "mediancenter," and "geometrical median" to describe it (see Small 1990). Recently Dudley and Koltchinskii (1992) and Koltchinskii (1993) have considered an equivalent notion of quantiles in finite-dimensional Euclidean spaces. In Section 2 we define geometric quantiles in multidimensional spaces and explore their basic properties. We indicate that these multivariate quantiles can be used to construct L estimates for multivariate location and to obtain a natural extension of regression quantiles in multiresponse linear model problems. Also, we gradually expose that this geometric notion of quantiles extends to infinite-dimensional Hilbert and Banach spaces in a very natural and interesting way. In Section 3 we discuss large-sample behavior of multivariate quantiles. We establish a Bahadur-type linear representation and use it to derive asymptotic distributions of sample quantiles. We report several intriguing facts observed in course of the development of these large-sample results. Also, we investigate statistical variability of sample quantiles and the estimation of their dispersion matrices. We show that many of the surprising properties of multivariate spatial median, which were observed by Bose and Chaudhuri (1993) and Chaudhuri (1992a), are inherited by these geometric quantiles. In Section 4 we present some concluding remarks. We provide all technical proofs in the Appendix.

2. GEOMETRIC QUANTILES: DEFINITION AND BASIC PROPERTIES

It is a well-known fact that given any α such that $0 < \alpha < 1$ and $u = 2\alpha - 1$, the sum $\sum_{i=1}^{n}\{|X_i - Q| + u(X_i - Q)\}$ is minimized when $Q =$ the sample αth quantile based on the real-valued observations X_i's (see, e.g., Ferguson 1967). Koenker and Bassett (1978) used the loss function $\Phi(u,t) = |t| + ut$ as a substitute for the squared error loss to estimate the αth regression quantile in a linear regression setup. The case $\alpha = 1/2$ (or, equivalently, $u = 0$) corresponds to sample median, and in this case the definition of the function $\Phi(0,\cdot)$ can be easily extended for a vector-valued second argument. For $\mathbf{t} \in \mathbb{R}^d$, if we define $\Phi(0,\mathbf{t}) = |\mathbf{t}|$, then we get the loss function used for defining the multivariate spatial median (see, e.g., Small 1990). If we have data points $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ in $\mathbb{R}^d$, where $d \geq 2$, then a natural question is how to extend the definition of the function $\Phi(\cdot,\cdot)$ so that it will lead to a multivariate generalization of quantiles. It was observed by Brown and Hettmansperger (1987, 1989) that geometrically it is quite meaningful for a multivariate quantile to have a direction in addition to a magnitude (see also Hettmansperger, Nyblom, and Oja 1994). Observe that the factor $u = 2\alpha - 1$ that appears in the second term in the definition of $\Phi(u,t)$ is a linear transformation of $\alpha \in (0,1)$ that maps the open unit interval (0, 1) onto the open interval $(-1,1)$ in a one-to-one way, and the α's corresponding to extreme quantiles are mapped to values close to $+1$ or -1, whereas those corresponding to central quantiles are mapped to values close to zero. This leads to the idea of indexing d-dimensional multivariate quantiles by elements of the open unit ball $B^{(d)} = \{\mathbf{u}|\mathbf{u} \in \mathbb{R}^d, |\mathbf{u}| < 1\}$. For any $\mathbf{u} \in B^{(d)}$ and $\mathbf{t} \in \mathbb{R}^d$, let us define $\Phi(\mathbf{u},\mathbf{t}) = |\mathbf{t}| + \langle \mathbf{u},\mathbf{t} \rangle$, where $\langle \cdot,\cdot \rangle$ denotes the usual Euclidean inner product. Then the geometric quantile $\hat{\mathbf{Q}}_n(\mathbf{u})$ corresponding to $\mathbf{u}$ and based on d-dimensional data points $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ is defined as

$$\hat{\mathbf{Q}}_n(\mathbf{u}) = \arg\min_{\mathbf{Q} \in \mathbb{R}^d} \sum_{i=1}^{n} \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q}).$$

Observe at this point that a $\mathbf{u}$ for which $|\mathbf{u}|$ is close to 1 corresponds to an extreme quantile, whereas a $\mathbf{u}$ for which $|\mathbf{u}|$ is close to zero corresponds to a central quantile. As pointed out by Small (1990), Weber (1909) considered spatial median (which is just $\hat{\mathbf{Q}}_n(0)$) as a solution to a problem in "location theory" in which $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ are the planar coordinates of n customers, who are served by a company that wants to find an optimal location for its warehouse. If the transportation cost happens to be proportional to the distance, and all customers are equally important for the company, one should try to locate the warehouse as close to the spatial median of the $\mathbf{X}_i$'s as possible, so that the average (or equivalently the total) transportation cost is minimized. On the other hand, for a nonzero $\mathbf{u}$, its magnitude $|\mathbf{u}|$ measures the extent of deviation of the quantile $\hat{\mathbf{Q}}_n(\mathbf{u})$ from the center of the data cloud formed by the $\mathbf{X}_i$'s. Because the vector $\mathbf{u}$ has a direction in addition to its magnitude, this immediately leads to a notion of directional outlyingness of a point with respect to the center of a cloud of observations based on the geometry of the cloud. (For other notions of outlyingness based on various concepts of "data depth," see Donoho and Gasko 1992; Liu 1990, 1992; Stahel 1981; and Tukey 1975).

Notice that the preceding definition of multidimensional quantiles extends in a natural way when the observations lie in a Hilbert space, which may very well be infinite dimensional in nature. Any Hilbert space is equipped with an inner product, and the open unit ball around the origin is a well-defined concept there. Hence the definition of $\Phi(\cdot,\cdot)$ extends naturally for data in a Hilbert space, where the quantiles continue to be indexed by vectors having norms smaller than 1. Kemperman (1987) introduced and extensively studied a notion of median in Banach spaces. Observe that the second term in the definition of $\Phi(\mathbf{u},\mathbf{t})$ can be viewed as a real-valued linear functional with norm $= |\mathbf{u}| < 1$. In a Banach space, it is natural to replace the second term, which is currently defined as an inner product, by a real-valued linear functional with norm (i.e., functional norm) smaller than 1. In other words, geometric quantiles in a Banach space will be indexed by the elements of the open unit ball around the origin in the dual Banach space of real-valued linear functionals. This yields a generalization of Kemperman's (1987) idea of median into a notion of quantiles in

Banach spaces. Recall at this point a well-known result in elementary functional analysis stating that the dual of a Hilbert space is isometrically isomorphic to the space itself. This is why geometric quantiles in a Hilbert space will be indexed by the elements of the open unit ball around the origin in the space itself. It will be appropriate to point out here that many multivariate versions of median proposed in the literature (e.g., Liu 1990; Oja 1983) are limited to only finite-dimensional spaces because of the very nature of their construction.

2.1 Existence, Uniqueness and Computation

Consider a set of observations $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ in $\mathbb{R}^d$. We begin by addressing the issue of the existence of a minimizer (with respect to $\mathbf{Q}$) of $\sum_{i=1}^n \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q})$. Note that for any fixed $\mathbf{u} \in B^{(d)}$, the function $\Phi(\mathbf{u}, \mathbf{t})$ explodes to infinity as $|\mathbf{t}|$ tends to infinity. Hence $\sum_{i=1}^n \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q})$ must tend to infinity if $|\mathbf{Q}|$ goes to infinity. In other words, the value of the sum will be arbitrarily large for a $\mathbf{Q}$ for which $|\mathbf{Q}|$ is sufficiently large, and one must look for a minimizer within a closed and bounded ball around the origin in $\mathbb{R}^d$. In view of the continuity of $\Phi(\mathbf{u}, \mathbf{t})$ as a function of $\mathbf{t}$, which implies the continuity of $\sum_{i=1}^n \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q})$ as a function of $\mathbf{Q}$, there must be a minimizer $\hat{\mathbf{Q}}_n(\mathbf{u})$ located at a finite distance from the origin in $\mathbb{R}^d$. Next comes the question of uniqueness. Because $\mathbb{R}^d$ equipped with Euclidean norm is a strictly convex Banach space for $d \geq 2$, and $\langle \mathbf{u}, \mathbf{t} \rangle$ is a linear function in $\mathbf{t}$ for every fixed $\mathbf{u} \in B^{(d)}$, it follows from theorem 2.17 of Kemperman (1987, p. 220) that unless all of the data points $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ fall on a straight line in $\mathbb{R}^d$, $\sum_{i=1}^n \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q})$ must be a strictly convex function of $\mathbf{Q}$. This guarantees the uniqueness of the minimizer $\hat{\mathbf{Q}}_n(\mathbf{u})$ in $\mathbb{R}^d$ for any $d \geq 2$ provided that the data points do not lie on a single straight line. Summarizing all these, we now have the following.

Fact 2.1.1. For a set of observations $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ in $\mathbb{R}^d$, the geometric quantile $\hat{\mathbf{Q}}_n(\mathbf{u})$ exists for any given $\mathbf{u} \in B^{(d)}$. Further, for $d \geq 2$, it will be unique if the $\mathbf{X}_i$'s are not all carried by a straight line in $\mathbb{R}^d$.

As a matter of fact, a natural generalization of some of the results of Kemperman (1987) guarantees the uniqueness of geometric quantiles in any strictly convex Banach space unless the observations all lie on one straight line in that space. It is easy to extend some of the arguments used by Valadier (1984) to establish the existence of geometric quantiles for observations in any reflexive Banach space for which the dual Banach space is isometrically isomorphic to the original space (e.g., $\mathbb{R}^d$ or any Hilbert space). But León and Massé (1992) pointed out that a spatial median (or L_1 median as they called it) may not exist in some nonreflexive Banach spaces.

Efficient algorithms to compute spatial median minimizing $\sum_{i=1}^n |\mathbf{X}_i - \mathbf{Q}|$, when the $\mathbf{X}_i$'s are in $\mathbb{R}^d$, have been extensively studied by Bedall and Zimmermann (1979) and Gower (1974). We next state a theorem that gives an important characterization of a geometric quantile in terms of the data points from which it is computed.

Theorem 2.1.2. Consider data points $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ in $\mathbb{R}^d$ and $\hat{\mathbf{Q}}_n(\mathbf{u})$ computed from these observations. If $\hat{\mathbf{Q}}_n(\mathbf{u}) \neq \mathbf{X}_i$ for all $1 \leq i \leq n$, then we will have

$$\sum_{i=1}^n |\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})|^{-1}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})\} + n\mathbf{u} = 0.$$

On the other hand, if $\hat{\mathbf{Q}}_n(\mathbf{u}) = \mathbf{X}_i$ for some $1 \leq i \leq n$, then we will have

$$\left| \sum_{i:1 \leq i \leq n; \mathbf{X}_i \neq \hat{\mathbf{Q}}_n(\mathbf{u})} [|\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})|^{-1}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})\} + \mathbf{u}] \right| \leq (1 + |\mathbf{u}|)[\#\{i : \mathbf{X}_i = \hat{\mathbf{Q}}_n(\mathbf{u})\}].$$

This crucial theorem implies that the algorithms of Bedall and Zimmermann (1979) and Gower (1974) can be modified to yield algorithms for computing geometric quantiles from multivariate observations. Specifically, one can use iterative methods like the "first-order method" (see Gower 1974) or a "Newton–Raphson-type method" (see Bedall and Zimmermann 1979), with the latter usually being much faster than the former. We now describe the main steps involved in the computation of $\hat{\mathbf{Q}}_n(\mathbf{u})$, when $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ are n distinct points in $\mathbb{R}^d$ ($d \geq 2$) not lying on a single straight line (e.g., they may be iid observations with a common absolutely continuous distribution on $\mathbb{R}^d$) and a "Newton–Raphson-type iteration" is used. From now on, all vectors in this article are assumed to be column vectors unless specified otherwise, and the superscript T is used to indicate the transpose of vectors and matrices.

Step 1. For each $1 \leq i \leq n$, one checks whether or not the degeneracy condition

$$\left| \sum_{j:1 \leq j \leq n; j \neq i} \{|\mathbf{X}_j - \mathbf{X}_i|^{-1}(\mathbf{X}_j - \mathbf{X}_i)\} + (n-1)\mathbf{u} \right| \leq (1 + |\mathbf{u}|)$$

is satisfied. If the condition is satisfied for some $1 \leq i \leq n$, then one sets $\hat{\mathbf{Q}}_n(\mathbf{u}) = \mathbf{X}_i$. Otherwise, one moves to the next step and tries to solve the equation

$$\sum_{i=1}^n |\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})|^{-1}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})\} + n\mathbf{u} = 0.$$

Step 2. One needs an initial approximation $\hat{\mathbf{Q}}_n^{(1)}(\mathbf{u})$ of $\hat{\mathbf{Q}}_n(\mathbf{u})$ to start the iteration, and this can be taken to be the vector of medians of real-valued components of the $\mathbf{X}_i$'s or some other suitable point in $\mathbb{R}^d$. Let $\hat{\mathbf{Q}}_n^{(1)}(\mathbf{u}), \ldots, \hat{\mathbf{Q}}_n^{(m)}(\mathbf{u})$ be the successive approximations of $\hat{\mathbf{Q}}_n(\mathbf{u})$ obtained in consecutive iterations. Then $\hat{\mathbf{Q}}_n^{(m+1)}(\mathbf{u})$ is computed as follows. Let

$$\sum_{i=1}^n |\mathbf{X}_i - \hat{\mathbf{Q}}_n^{(m)}(\mathbf{u})|^{-1}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n^{(m)}(\mathbf{u})\} + n\mathbf{u} = \Delta$$

and define

$$\sum_{i=1}^{n} |\mathbf{X}_i - \hat{\mathbf{Q}}_n^{(m)}(\mathbf{u})|^{-1}[\mathbf{I}_d - |\mathbf{X}_i - \mathbf{Q}_n^{(m)}(\mathbf{u})|^{-2} \times \{\mathbf{X}_i - \hat{\mathbf{Q}}_n^{(m)}(\mathbf{u})\}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n^{(m)}(\mathbf{u})\}^T] = \Phi,$$

where $\mathbf{I}_d$ is the $d \times d$ identity matrix. Then unless the $\mathbf{X}_i$'s all lie on a straight line in $\mathbb{R}^d$, Φ will be a positive definite matrix, and one sets $\hat{\mathbf{Q}}_n^{(m+1)}(\mathbf{u}) = \hat{\mathbf{Q}}_n^{(m)}(\mathbf{u}) + \Phi^{-1}\Delta$. Iteration is continued until two successive approximations of $\hat{\mathbf{Q}}_n(\mathbf{u})$ happen to be sufficiently close.

It is easy to prepare a simple computer program to implement the algorithm just described. A FORTRAN program was tried on several bivariate and trivariate data sets that were simulated on a VAX 8650 (with VMS operating system) using IMSL routines. The distributions tried included multivariate versions of Gaussian, Laplace, and Cauchy distributions. In all cases, the algorithm converged after 5–10 iterations, and the total time for running the program was only a few seconds in each case. Because of the unsmooth nature of the function $\sum_{i=1}^{n} \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q})$ when $\mathbf{Q}$ is close to some of the data points, Brown (1985) suggested including an "escape hatch" in the "Newton–Raphson iteration" for computing spatial median. Although this is definitely a wise idea, for the sake of simplicity we did not build any such "escape hatch" into our program, and this did not have any serious effect on the program's performance. We close this section by pointing out that each of the multivariate versions of median proposed by Liu (1990), Tukey (1975), and Oja (1983) is fairly difficult to compute when $d \geq 3$, and this computational difficulty increases at a substantial rate as d increases. The iterative algorithm presented here for computing spatial median or any geometric quantile is easy to use even for high-dimensional data, as the only effect of dimension that one can feel while running the algorithm is during the inversion of the $d \times d$ matrix in Step 2.

2.2 Properties and Applications

It is obvious that our geometric quantiles are location equivariant in the sense that if $\mathbf{Y}_i = \mathbf{X}_i + \mathbf{a}$ for all $1 \leq i \leq n$, where $\mathbf{a} \in \mathbb{R}^d$ is a fixed vector, the geometric quantile $\hat{\mathbf{Q}}_n^{+}(\mathbf{u})$ corresponding to $\mathbf{u} \in B^{(d)}$ and based on $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$ will satisfy $\hat{\mathbf{Q}}_n^{+}(\mathbf{u}) = \hat{\mathbf{Q}}_n(\mathbf{u}) + \mathbf{a}$, where $\hat{\mathbf{Q}}_n(\mathbf{u})$ is the geometric quantile based on $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ as before. One of Brown's (1983) main motivations in considering "spatial median" is its rotational equivariance, which is very desirable in the analysis of spatial data, where variables possess isometry. Small (1990) pointed out that a starting point for some of the early work on "spatial median" was the twelfth census of the United States in 1900, when statisticians were interested in investigating the flow of population in the United States by observing the movement of a "geographical center" of the population over time. It was clearly recognized (see, e.g., Hayford 1902) that a median-like estimate of the center of a geographical distribution is preferable to the centroid (i.e., the usual multivariate average), as the centroid may be highly sensitive to the influence of probability masses at the extremes. In fact, one can argue (see Small 1990) that a death or a birth in the periphery of the country should not have more influence on the center of the population than a similar event occurring at the central part of the country. Hayford (1902) proposed the vector of medians of orthogonal coordinates as the "geographical center" but explicitly noted the difficulty arising from the fact that such a multivariate median depends on the choice of the orthogonal coordinates and suffers from lack of equivariance under orthogonal transformations. Eventually, Scates (1933) used "spatial median" to locate the "geographical center" of the U.S. population and found it to be located at a place "15 miles northwest of Dayton, Ohio." Like the "spatial median," our geometric quantiles also happen to be rotationally equivariant. In fact, we have the following in view of the way geometric quantiles are defined.

Fact 2.2.1. As before let $\hat{\mathbf{Q}}_n(\mathbf{u})$ be the geometric quantile corresponding to $\mathbf{u} \in B^{(d)}$ and based on $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ in $\mathbb{R}^d$. Let $\mathbf{A}$ be a $d \times d$ orthogonal matrix and let $\mathbf{a}$ be a fixed d-dimensional vector. Set $\mathbf{v} = \mathbf{A}\mathbf{u}$ so that $|\mathbf{v}| = |\mathbf{u}|$ in view of the orthogonality of $\mathbf{A}$. Suppose that $\hat{\mathbf{Q}}_n^{+}(\mathbf{v})$ is the geometric quantile corresponding to $\mathbf{v} \in B^{(d)}$ and based on $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$, where $\mathbf{Y}_i = \mathbf{A}\mathbf{X}_i + \mathbf{a}$ for all $1 \leq i \leq n$. Then $\hat{\mathbf{Q}}_n^{+}(\mathbf{v}) = \hat{\mathbf{Q}}_n^{+}(\mathbf{A}\mathbf{u}) = \mathbf{A}\hat{\mathbf{Q}}_n(\mathbf{u}) + \mathbf{a}$.

Note that the preceding fact has been stated for data observed in finite-dimensional Euclidean spaces. But it can be easily generalized in an arbitrary Hilbert space, where one can have a concept of equivariance under location transformations and norm-preserving linear transformations (i.e., under any kind of rigid motion of points in the space). More generally, our geometric quantiles will be equivariant under any invertible and distance-preserving affine transformation on a Banach space.

It is quite easy to see that geometric quantiles are equivariant under any homogeneous scale transformation of the coordinates of the multivariate observations, as indicated next.

Fact 2.2.2. If $c > 0$ is a fixed scalar, and $\hat{\mathbf{Q}}_n(\mathbf{u})$ is the geometric quantile corresponding to $\mathbf{u} \in B^{(d)}$ and based on $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ in $\mathbb{R}^d$, then $c\hat{\mathbf{Q}}_n(\mathbf{u})$ will be the geometric quantile based on $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$, where $\mathbf{Y}_i = c\mathbf{X}_i$ for all $1 \leq i \leq n$.

In connection with Fact 2.2.2, it may be appropriate to note that in some situations, one may need to standardize the coordinate variables appropriately before computing the spatial median or any other geometric quantile for a multivariate data set (e.g., when the units of measurements for different coordinate variables happen to be different). In an attempt to make Haldane's (1948) "geometrical median" affine equivariant, Rao (1988) recommended standardizing the observations using the square root of the variance–covariance matrix computed from the data. Such a standardization amounts to the replacement of the Euclidean distance by Mahalanobis's statistical distance (see Mahalanobis 1936). Alternatively, one may use "data-driven coordinate systems" (see Chaudhuri and Sengupta 1993) constructed from appropriately centered observations.

Chaudhuri and Sengupta (1993) introduced and used such invariant coordinate systems to construct affine-invariant sign tests in multidimensions, and recently Chakraborty and Chaudhuri (1994, 1995) used them to construct an affine-equivariant version of multivariate median.

We conclude this section by indicating two potential applications of geometric quantiles.

a. Geometric quantiles can be used to extend the concept of quantile regression from univariate response problems (see Efron 1991 and Koenker and Bassett 1978) to multiresponse linear model situations in the following way. Let $(\mathbf{Y}_1, \mathbf{Z}_1), (\mathbf{Y}_2, \mathbf{Z}_2), \ldots, (\mathbf{Y}_n, \mathbf{Z}_n)$ be a set of observations satisfying the multivariate linear regression model $\mathbf{Y}_i = \mathbf{\Gamma Z}_i + \mathbf{E}_i$ for all $1 \le i \le n$. Here $\mathbf{Y}_i$ is a d-dimensional response vector, $\mathbf{Z}_i$ is a p-dimensional regressor vector, $\mathbf{\Gamma}$ is a $d \times p$ matrix of parameters, and $\mathbf{E}_i$ is a d-dimensional vector of unobservable random errors. Then for $\mathbf{u} \in B^{(d)}$, one can try to estimate $\mathbf{\Gamma}$ by solving the minimization problem $\min_{\mathbf{\Gamma}} \sum_{i=1}^n \Phi(\mathbf{u}, \mathbf{Y}_i - \mathbf{\Gamma Z}_i)$. If $\hat{\mathbf{\Gamma}}_n(\mathbf{u})$ denotes a solution for this minimization problem, then we can identify it as a multiresponse regression quantile estimate of $\mathbf{\Gamma}$. As an extension of spatial median, Bai, Chen, Miao, and Rao (1990) have considered certain minimum Euclidean norm estimates of parameters in multiresponse linear model problems.

b. It is possible to extend the concept of trimmed mean or any L estimate (see, e.g., Serfling 1980) of univariate location to a multivariate setup using geometric quantiles in a natural way. One just needs to form suitable weighted averages of $\hat{\mathbf{Q}}_n(\mathbf{u})$'s as $\mathbf{u}$ varies over an appropriate subset of $B^{(d)}$, keeping in mind that for a $\mathbf{u}$ with $|\mathbf{u}|$ close to zero, we get a central quantile, whereas for a $\mathbf{u}$ with $|\mathbf{u}|$ close to 1, we get an extreme quantile. Specifically, if μ is an appropriately chosen probability measure on $B^{(d)}$ supported on a subset S of $B^{(d)}$, then an L estimate of multivariate location will have the form $\int_S \hat{\mathbf{Q}}_n(\mathbf{u})\mu(d\mathbf{u})$. If S happens to be the sphere with center at the origin and radius $= r$ (i.e., $S = \{\mathbf{u} | \mathbf{u} \in \mathbb{R}^d, |\mathbf{u}| \le r\}$), where r is a constant such that $0 < r < 1$, and μ is taken to be the uniform probability measure on S, $\int_S \hat{\mathbf{Q}}_n(\mathbf{u})\mu(d\mathbf{u})$ is a version of the trimmed multivariate mean. Some recent attempts to construct and study various versions of the trimmed mean for multivariate location using different ideas include those of Donoho and Gasko (1992), Gordaliza (1991), and Nolan (1992). L estimates of parameters in linear models using regression quantiles have been studied extensively by Koenker and Portnoy (1987) (see also Bickel 1973, Chaudhuri 1992b, Ruppert and Carroll 1980, and Welsh 1987a,b), who considered univariate response. $\hat{\mathbf{\Gamma}}_n(\mathbf{u})$'s with $\mathbf{u}$ varying in $B^{(d)}$ defined previously can be utilized to carry the L estimation technique into the domain of multiresponse linear model problems.

3. LARGE-SAMPLE STATISTICAL PROPERTIES

We begin by defining geometric quantiles for a multivariate probability distribution. Let $\mathbf{X}$ be a random vector with a probability distribution on $\mathbb{R}^d$. For $\mathbf{u} \in B^{(d)}$, the quantile $\mathbf{Q}(\mathbf{u})$ of the distribution of $\mathbf{X}$ is defined by

$$\mathbf{Q}(\mathbf{u}) = \arg\min_{\mathbf{Q} \in \mathbb{R}^d} E\{\Phi(\mathbf{u}, \mathbf{X} - \mathbf{Q}) - \Phi(\mathbf{u}, \mathbf{X})\}.$$

Note that $\Phi(\mathbf{u}, \mathbf{X} - \mathbf{Q}) - \Phi(\mathbf{u}, \mathbf{X})$ will always have a finite expectation even though the expectation of $\mathbf{X}$ may not always be finite. When $\mathbf{X}$ has a finite expectation, $\mathbf{Q}(\mathbf{u})$ becomes a minimizer of $E\{\Phi(\mathbf{u}, \mathbf{X} - \mathbf{Q})\}$. Further, in view of our observations in Section 2.1, the existence and the uniqueness of $\mathbf{Q}(\mathbf{u})$ is guaranteed for any $\mathbf{u} \in B^{(d)}$ and $d \ge 2$, provided that the distribution of $\mathbf{X}$ is not supported on a single straight line. Hence uniqueness holds whenever $\mathbf{X}$ has an absolutely continuous distribution on $\mathbb{R}^d$ with $d \ge 2$, and in fact in this case $\mathbf{Q}(\mathbf{u})$ will be the unique solution in $\mathbf{Q}$ of the equation $E\{|\mathbf{X}_i - \mathbf{Q}|^{-1}(\mathbf{X}_i - \mathbf{Q})\} + \mathbf{u} = 0$. For a univariate probability distribution (i.e., when $d = 1$), it is obvious that geometric quantiles of the distribution coincide with usual univariate quantiles indexed by the elements of the open interval $B^{(1)} = (-1, 1)$.

3.1 Bahadur Representation and Asymptotic Distribution

Clearly, if the observations $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ are independent and identically distributed copies of $\mathbf{X}$, then $\hat{\mathbf{Q}}_n(\mathbf{u})$ will act as an estimate of $\mathbf{Q}(\mathbf{u})$ based on those observations. Chaudhuri (1992a) derived a Bahadur-type representation for a class of multivariate location estimates that includes spatial median as a special case (see also Niemiro 1992). Our next theorem establishes a Bahadur-type representation of geometric quantiles, and we use this result to derive the asymptotic distribution and related properties of $\hat{\mathbf{Q}}_n(\mathbf{u})$ to get useful insights into its behavior as an estimate of $\mathbf{Q}(\mathbf{u})$. But before stating the theorem, we introduce some notations. For any $\mathbf{Q} \in \mathbb{R}^d$, define the $d \times d$ symmetric matrix

$$\mathbf{D}_1(\mathbf{Q}) = E[|\mathbf{X} - \mathbf{Q}|^{-1}\{\mathbf{I}_d - |\mathbf{X} - \mathbf{Q}|^{-2}(\mathbf{X} - \mathbf{Q})(\mathbf{X} - \mathbf{Q})^T\}],$$

which will be positive definite unless the distribution of $\mathbf{X}$ is completely supported on a straight line in $\mathbb{R}^d$. Note that the expectation defining $\mathbf{D}_1(\mathbf{Q})$ will exist finitely for $d \ge 2$ whenever $\mathbf{X}$ has a density bounded on compact subsets of $\mathbb{R}^d$. This is a consequence of the fact that for any fixed $\mathbf{y} \in \mathbb{R}^d$ and a density f bounded on compact subsets of $\mathbb{R}^d$, the integral $\int_{\mathbb{R}^d} |\mathbf{x} - \mathbf{y}|^{-1} f(\mathbf{x})\, d\mathbf{x}$ is finite. This fact can be verified by using d-dimensional polar transformation for which the Jacobian determinant involves the $(d-1)$th power of the length of the radius vector (see proposition 3.1 in Bose and Chaudhuri 1993, p. 546, and remark 4 in Chaudhuri 1992a, p. 904). Also, for $\mathbf{Q}_1, \mathbf{Q}_2 \in \mathbb{R}^d$, and $\mathbf{u}, \mathbf{v} \in B^{(d)}$, let us write $\mathbf{D}_2(\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{u}, \mathbf{v})$ to denote the $d \times d$ matrix

$$E[\{|\mathbf{X} - \mathbf{Q}_1|^{-1}(\mathbf{X} - \mathbf{Q}) + \mathbf{u}\}\{|\mathbf{X} - \mathbf{Q}_2|^{-1}(\mathbf{X} - \mathbf{Q}_2) + \mathbf{v}\}^T].$$

Theorem 3.1.1. Assume that $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n, \ldots$ is a sequence of independent and identically distributed d-dimensional random vectors with a common density, which is bounded on every bounded subset of $\mathbb{R}^d$. Then, for any fixed $\mathbf{u} \in B^{(d)}$, we have the following Bahadur-type expansion:

$$\hat{\mathbf{Q}}_n(\mathbf{u}) - \mathbf{Q}(\mathbf{u}) = n^{-1}[\mathbf{D}_1\{\mathbf{Q}(\mathbf{u})\}]^{-1} \times \sum_{i=1}^{n} [|\mathbf{X}_i - \mathbf{Q}(\mathbf{u})|^{-1}\{\mathbf{X}_i - \mathbf{Q}(\mathbf{u})\} + \mathbf{u}] + \mathbf{R}_n(\mathbf{u}),$$

where as n tends to infinity, $\mathbf{R}_n(\mathbf{u})$ is almost surely $O(\log n/n)$ if $d \geq 3$, and when $d = 2, \mathbf{R}_n(\mathbf{u})$ is almost surely $o(n^{-\beta})$ for any fixed β such that $0 < \beta < 1$.

Observe that the condition assumed on the common density of the $\mathbf{X}_i$'s in the statement of Theorem 3.1.1 is much weaker than the condition needed to establish the Bahadur expansion of a univariate quantile, (see Bahadur 1966, Ghosh 1971, Kiefer 1967, and Serfling 1980). Also, the convergence rates for the remainder term $\mathbf{R}_n(\mathbf{u})$ in the theorem is much faster than that for the remainder term in the Bahadur representation of a univariate quantile. Recently Koltchinskii (1993) obtained stronger results related to the asymptotic behavior of $\mathbf{R}_n(\mathbf{u})$ (see also Niemiro 1992, who discussed a Bahadur expansion for spatial median). All these demonstrate that geometric quantiles inherit some of the intriguing asymptotic properties of the spatial median (see remarks 4 and 5 in Chaudhuri 1992a, pp. 904–905).

We now state a theorem concerning the joint asymptotic distribution of several geometric quantiles.

Theorem 3.1.2. Suppose that the condition assumed in Theorem 3.1.1 holds. Let $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ be points in the open unit ball $B^{(d)}$, where k is a fixed positive integer. Then the joint asymptotic distribution of centered and normalized geometric quantiles

$$n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{u}_1) - \mathbf{Q}(\mathbf{u}_1)\}, n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{u}_2) - \mathbf{Q}(\mathbf{u}_2)\}, \ldots, n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{u}_k) - \mathbf{Q}(\mathbf{u}_k)\}$$

will be Gaussian with mean zero. Further, the asymptotic covariance matrix between $n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{u}_r) - \mathbf{Q}(\mathbf{u}_r)\}$ and $n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{u}_s) - \mathbf{Q}(\mathbf{u}_s)\}$, where $1 \leq r, s \leq k$ (note that r and s may or may not be distinct), will be given by

$$[\mathbf{D}_1\{\mathbf{Q}(\mathbf{u}_r)\}]^{-1}[\mathbf{D}_2\{\mathbf{Q}(\mathbf{u}_r), \mathbf{Q}(\mathbf{u}_s), \mathbf{u}_r, \mathbf{u}_s\}][\mathbf{D}_1\{\mathbf{Q}(\mathbf{u}_s)\}]^{-1}.$$

Clearly, the preceding theorem guarantees that the sample geometric quantiles are consistent estimates of corresponding population quantiles, they converge at $n^{-1/2}$ rate, and are asymptotically normally distributed. In fact, this theorem can be used to obtain the limiting distribution (which again will be normal) of multivariate L estimates (see Sec. 2.2) that are defined as weighted averages (i.e., convex combinations) of finitely many geometric quantiles (i.e., if the set S appearing in $\int_S \hat{\mathbf{Q}}_n(\mathbf{u})\mu(d\mathbf{u})$ is a finite set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_s\}$ so that the integral becomes a finite sum of the form $\sum_{i=1}^{s} \hat{\mathbf{Q}}_n(\mathbf{u}_i)\mu(\{\mathbf{u}_i\})$). The multivariate stochastic process $\hat{\mathbf{Q}}_n(\mathbf{u})$ indexed by the vector parameter $\mathbf{u} \in B^{(d)}$ can be viewed as a generalization of the univariate quantile process. In view of Theorem 3.1.2, one can hope that the centered and normalized process $n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{u}) - \mathbf{Q}(\mathbf{u})\}$ with $\mathbf{u}$ varying in $B^{(d)}$ will converge weakly to a Gaussian process (which too will be parameterized by the elements of $B^{(d)}$) under appropriate regularity conditions. Although we will not dig deeper into technical matters (e.g., the oscillation of the sample path of the process and the tightness issues) related to the weak convergence of such a stochastic process, Theorem 3.1.2 can be helpful in identifying the nature of the limiting Gaussian process by utilizing the variance–covariance structure explicitly worked out there.

3.2 Sampling Variation and Related Issues

Theorem 3.1.2 can be used to construct large-sample confidence ellipsoids for $\mathbf{Q}(\mathbf{u})$, provided that we can construct a reasonable estimate of the limiting dispersion matrix of $n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{u}) - \mathbf{Q}(\mathbf{u})\}$ from the data. Estimation of the asymptotic variance of a univariate quantile has been studied extensively in the literature, and it has drawn attention from several leading statisticians. Efron (1982) observed that the standard "delete one jackknife" leads to an inconsistent estimate of the asymptotic variance of univariate median. Later, Shao and Wu (1989) established that "delete k jackknife" yields a consistent estimate of this variance if k is allowed to grow to infinity as the sample size increases. But practical implementation of "delete k jackknife" will require prohibitively complex and expensive computation in the case of large data sets. On the other hand, it is well known that the "standard bootstrap," which resamples from the usual empirical distribution, produces a consistent estimate of the large-sample variance of a univariate quantile (see Babu 1986; Efron 1982; Ghosh, Parr, Singh, and Babu 1984; and Shao 1990) under suitable regularity conditions. But Hall and Martin (1988, 1991) showed that such a bootstrap variance estimate converges at an extremely slow rate, namely $n^{-1/4}$. Hall, DiCiccio, and Romano (1989) pointed out that it is possible to improve this convergence rate substantially by resampling from appropriate kernel density estimates rather than using the "standard bootstrap" based on the unsmoothed empirical distribution. Although the convergence rate always remains slower than $n^{-1/2}$, these authors demonstrated that it can be brought arbitrarily close to $n^{-1/2}$. But to actually achieve such an improvement, one may need to use higher-order kernels, which may lead to negative estimates of density and unnatural variance estimates. Surprisingly, these technical complexities disappear as soon as we start dealing with multivariate observations.

We next exhibit a very simple estimate of the limiting covariance matrix between a pair of centered and normalized geometric quantiles with excellent asymptotic properties. To construct this estimate, we do not use any of the computationally intensive resampling techniques like the bootstrap (smoothed or unsmoothed) or "delete k jackknife." Bose and Chaudhuri (1993) observed a similar phenomenon while constructing the estimate for large-sample dispersion of multivariate spatial median.

Let $\mathcal{F}_n$ be a subset of $\{1, 2, \ldots, n\}$ such that $\#(\mathcal{F}_n) = f_n$. Consider $\mathbf{u}, \mathbf{v} \in B^{(d)}$ (here $\mathbf{u}$ and $\mathbf{v}$ may or may not be distinct). Define $\hat{\mathbf{Q}}_n^*(\mathbf{u})$ and $\hat{\mathbf{Q}}_n^*(\mathbf{v})$ as the geometric quantiles corresponding to $\mathbf{u}$ and $\mathbf{v}$, based on the $\mathbf{X}_i$'s for which $i \in \mathcal{F}_n$. In other words,

$$\hat{\mathbf{Q}}_n^*(\mathbf{u}) = \arg\min_{\mathbf{Q}\in\mathbb{R}^d} \sum_{i\in\mathcal{F}_n} \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q})$$

and

$$\hat{\mathbf{Q}}_n^*(\mathbf{v}) = \arg\min_{\mathbf{Q}\in\mathbb{R}^d} \sum_{i\in\mathcal{F}_n} \Phi(\mathbf{v}, \mathbf{X}_i - \mathbf{Q}).$$

Next, we set

$$\mathbf{A}_n^{(i)}(\mathbf{u},\mathbf{v}) = [|\mathbf{X}_i - \hat{\mathbf{Q}}_n^*(\mathbf{u})|^{-1}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n^*(\mathbf{u})\} + \mathbf{u}] \times [|\mathbf{X}_i - \hat{\mathbf{Q}}_n^*(\mathbf{u})|^{-1}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n^*(\mathbf{u})\} + \mathbf{v}]^T$$

and

$$\mathbf{B}_n^{(i)}(\mathbf{u}) = |\mathbf{X}_i - \hat{\mathbf{Q}}_n^*(\mathbf{u})|^{-1}[\mathbf{I}_d - |\mathbf{X}_i - \hat{\mathbf{Q}}_n^*(\mathbf{u})|^{-2} \times \{\mathbf{X}_i - \hat{\mathbf{Q}}_n^*(\mathbf{u})\}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n^*(\mathbf{u})\}^T].$$

Then we have the following theorem, which describes the asymptotic behavior of $\mathbf{F}_n(\mathbf{u},\mathbf{v}) = (n - f_n)^{-1}\sum_{i\in\mathcal{F}_n^c} \mathbf{A}_n^{(i)}(\mathbf{u},\mathbf{v})$ and $\mathbf{G}_n(\mathbf{u}) = (n - f_n)^{-1}\sum_{i\in\mathcal{F}_n^c} \mathbf{B}_n^{(i)}(\mathbf{u})$, where $\mathcal{F}_n^c$ is the set theoretic complement of $\mathcal{F}_n$ in $\{1, 2, \ldots, n\}$.

Theorem 3.2.1. Suppose that both $n^{-1}f_n$ and $1 - n^{-1}f_n$ remain bounded away from zero as n tends to infinity and that the condition assumed in Theorem 3.1.1 holds. Then for $d \geq 2$, we have $\mathbf{F}_n(\mathbf{u},\mathbf{v}) - \mathbf{D}_2\{\mathbf{Q}(\mathbf{u}), \mathbf{Q}(\mathbf{v}), \mathbf{u}, \mathbf{v}\} = O_P(n^{-1/2})$ as n tends to infinity. Also, for $d \geq 3$, we have $\mathbf{G}_n(\mathbf{u}) - \mathbf{D}_1\{\mathbf{Q}(\mathbf{u})\} = O_P(n^{-1/2})$ as n tends to infinity. But when $d = 2$, we have only $\mathbf{G}_n(\mathbf{u}) - \mathbf{D}_1\{\mathbf{Q}(\mathbf{u})\} = o_p(n^{-\beta})$ as n tends to infinity for any fixed constant β such that $0 < \beta < 1/2$.

In view of the positive definiteness of $\mathbf{D}_1\{\mathbf{Q}(\mathbf{u})\}$ for any $\mathbf{u} \in B^{(d)}$, the foregoing theorem guarantees that $\{\mathbf{G}_n(\mathbf{u})\}^{-1}\mathbf{F}_n(\mathbf{u},\mathbf{v})\{\mathbf{G}_n(\mathbf{v})\}^{-1}$ will be a consistent estimate of the limiting covariance matrix between $n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{u}) - \mathbf{Q}(\mathbf{u})\}$ and $n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{v}) - \mathbf{Q}(\mathbf{v})\}$. Further, when $d \geq 3$, this estimate will converge at $n^{-1/2}$ rate, whereas when $d = 2$, it will converge at a rate arbitrarily close to $n^{-1/2}$. By taking $\mathbf{u} = \mathbf{v}$, we get an estimate for the limiting dispersion matrix of $n^{1/2}\{\hat{\mathbf{Q}}_n(\mathbf{u}) - \mathbf{Q}(\mathbf{u})\}$. It can be used to get an estimate for the large-sample "generalized variance" (see Wilks 1932) of geometric quantiles. Also, $\{\mathbf{G}_n(\mathbf{u})\}^{-1}\mathbf{F}_n(\mathbf{u},\mathbf{u})\{\mathbf{G}_n(\mathbf{u})\}^{-1}$ and $\hat{\mathbf{Q}}_n(\mathbf{u})$ can be utilized together to construct confidence ellipsoids for $\mathbf{Q}(\mathbf{u})$, and Theorems 3.1.2 and 3.2.1 will ensure the asymptotic accuracy of such confidence sets.

We conclude this section by noting that estimates $\mathbf{F}_n(\mathbf{u},\mathbf{v})$ and $\mathbf{G}_n(\mathbf{u})$ both depend on the choice of the set $\mathcal{F}_n$ in view of their construction, and as a result these estimates are not invariant under a permutation of the labels of the data points. One way of symmetrizing such an asymmetric function of the data is to form a simple average of various estimates corresponding to different possible choices of $\mathcal{F}_n$. As a matter of fact, it is not difficult to see from the arguments used in the proofs given in the Appendix that such an averaging will not affect the asymptotic properties of the original estimates very much (see also remark (c), section 4, in Bose and Chaudhuri 1993, pp. 548–549). It will be appropriate to note here that Bai et al. (1990) proposed some estimate for the asymptotic dispersion of least Euclidean distances estimates of parameters in multiresponse linear models. But their estimate is known to be weakly consistent only, and it is not clear at what rate this estimate converges.

4. SOME CONCLUDING REMARKS

1. As we have already noted in several places, geometric quantiles can be defined as meaningful and natural objects for probability distributions (including empirical distributions associated with data) supported on very general Banach spaces. In this connection, one interesting fact is that for a d-dimensional random vector $\mathbf{X}$, the vector of marginal quantiles corresponding to different real-valued components of $\mathbf{X}$ (see, e.g., Abdous and Theodorescu 1992 and Babu and Rao 1988) is also a version of a geometric quantile. When $\mathbb{R}^d$ is metrized using the l_1-norm defined as $|\mathbf{x}|_1 = |x_1| + |x_2| + \cdots + |x_d|$ for $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$, we get a Banach space that is geometrically very different from the Hilbert space $\mathbb{R}^d$ metrized with the standard Euclidean metric. The dual space of $\mathbb{R}^d$ metrized using the l_1-norm can be identified with $\mathbb{R}^d$ having the l_∞-norm defined as $|\mathbf{x}|_\infty = \max_{1\leq i\leq d}|x_i|$. So the open unit ball around the origin in that dual space is the d-dimensional hypercube $\{\mathbf{u} | \mathbf{u} \in \mathbb{R}^d, |\mathbf{u}|_\infty < 1\}$. This clearly demonstrates how one can view the vector of marginal quantiles of a d-dimensional random vector as a geometric quantile in $\mathbb{R}^d$ equipped with the l_1-norm.

2. We mentioned at the beginning of Section 1 that univariate quantiles are quite useful in constructing descriptive statistics such as interquartile range and various measures of skewness and kurtosis. One can use the d-dimensional Lebesgue measure of the set $\{\hat{\mathbf{Q}}_n(\mathbf{u}) | |\mathbf{u}| \leq .5\}$ as a multivariate analog of interquartile range based on geometric quantiles. In general, for fixed $r \in (0,1)$, consider the set $\{\hat{\mathbf{Q}}_n(\mathbf{u}) | |\mathbf{u}| \leq r\}$, which can be viewed as a quantile ball of radius r, and let $\lambda(r)$ denote the d-dimensional Lebesgue measure of this set. Then $\lambda(r)$ can be used as a measure of dispersion. Also, for suitable $r, s \in (0,1)$ such that $r < s$, the ratio $\lambda(r)/\lambda(s)$ can be used a measure of kurtosis in multidimension. Note that this generalizes quantile-based measures of kurtosis used in a univariate setup. It is easy to see that if a probability distribution is spherically symmetric around a fixed point, then the quantile ball of radius r associated with that distribution will be a sphere with the same center of symmetry for any $r \in (0,1)$, and this is true even in general Banach spaces, where spherical symmetry of a probability measure is equivalent to its invariance under distance-preserving affine transformations. Therefore, one can use these quantile balls computed from the data to detect the possible presence (or evidence of deviation from) spherical symmetry. Further, like some well-known quantile-based measures of skewness for univariate data, the quantity $\sup_{|\mathbf{u}|=r}|\hat{\mathbf{Q}}_n(\mathbf{u}) + \hat{\mathbf{Q}}_n(-\mathbf{u}) - 2\hat{\mathbf{Q}}_n(0)|$

$\{\lambda(r)\}^{-1/d}$ can be used as a measure of multivariate skewness for some appropriate choice of $r \in (0,1)$. Note that in the univariate case with $r = .5$, the foregoing becomes the standard quartile-based measure of skewness. The power $-1/d$ of the volume $\lambda(r)$ of the quantile ball of radius r, which has been suggested by a referee, makes the skewness measure invariant under any homogeneous scale transformation (i.e., scalar multiplication) of the multivariate data points.

3. In the univariate case, quantiles can be obtained by inverting the cumulative distribution function. We have already observed at the beginning of Section 3 that for a random vector $\mathbf{X}$ having an absolutely continuous distribution in $\mathbb{R}^d$, the equation $E|\mathbf{X} - \mathbf{Q}(\mathbf{u})|^{-1}\{\mathbf{X} - \mathbf{Q}(\mathbf{u})\} = -\mathbf{u}$ holds. In other words, $\mathbf{Q}(\mathbf{u})$ can be obtained by inverting the function (from $\mathbb{R}^d$ into $\mathbb{R}^d$) that maps $\mathbf{y} \in \mathbb{R}^d$ into $E|\mathbf{X} - \mathbf{y}|^{-1}(\mathbf{X} - \mathbf{y}) = -\mathbf{u}$ (see also Dudley and Koltchinskii 1992 and Koltchinskii 1993). Recall from the statement of Theorem 2.1.2 that when $\hat{\mathbf{Q}}_n(\mathbf{u}) \neq \mathbf{X}_i$ for all $1 \leq i \leq n$, we have the equation $n^{-1}\sum_{i=1}^n |\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})|^{-1}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})\} = -\mathbf{u}$, which is the empirical or sample version of the previous equation.

4. In a univariate setup, the concepts of ranks and quantiles are closely related. Recently, Jan and Randles (1994) and Mottonen and Oja (1995) considered some notions of multivariate signs and ranks that have some natural relationships with our geometric quantiles. Note that the d-dimensional vector $\sum_{i:\mathbf{X}_i \neq \mathbf{y}} |\mathbf{X}_i - \mathbf{y}|^{-1}(\mathbf{X}_i - \mathbf{y})$ can be viewed as a descriptive statistic that determines the geometric position of the point $\mathbf{y} \in \mathbb{R}^d$ with respect to the data cloud formed by the observations $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$, and this leads to a concept of vector-valued ranks in multidimensions. On the other hand, it should be pointed out here that the direction vector associated with the notion of ranks arising from Oja's median (Oja 1983) used by Brown and Hettmansperger (1987, 1989) and Hettmansperger et al. (1994) is not (unlike the vector $\mathbf{u}$ associated with geometric quantiles) normalized to lie inside the open unit ball $B^{(d)}$, and it is not bounded in general. This is noteworthy, as a normalized direction vector is more useful than an unbounded direction vector in judging the closeness to (or deviation from) the center of a multivariate data cloud.

APPENDIX: PROOFS

Proof of Theorem 2.1.2

Clearly, for any $\mathbf{x}, \mathbf{h} \in \mathbb{R}^d$ such that $\mathbf{x} \neq 0$, we have

$$\lim_{t \to 0+} t^{-1}\{\Phi(\mathbf{u}, \mathbf{x} + t\mathbf{h} - \Phi(\mathbf{u}, \mathbf{x})\} = \langle |\mathbf{x}|^{-1}\mathbf{x} + \mathbf{u}, \mathbf{h}\rangle,$$

and for any $\mathbf{h} \in \mathbb{R}^d$, we have

$$\lim_{t \to 0+} t^{-1}\{\Phi(\mathbf{u}, t\mathbf{h}) - \Phi(\mathbf{u}, 0)\} = |\mathbf{h}| + \langle \mathbf{u}, \mathbf{h}\rangle.$$

Now, using the convexity of $\Phi(\mathbf{u}, \mathbf{x})$ as a function of $\mathbf{x}, \mathbf{Q} = \hat{\mathbf{Q}}_n(\mathbf{u})$ minimizes $\sum_{i=1}^n \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q})$ if and only if

$$\lim_{t \to 0+} t^{-1}\left[\sum_{i=1}^n \Phi\{\mathbf{u}, \mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u}) + t\mathbf{h}\} - \sum_{i=1}^n \Phi\{\mathbf{u}, \mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})\}\right] \geq 0$$

for all $\mathbf{h} \in \mathbb{R}^d$. In other words, we must have

$$\sum_{i:1 \leq i \leq n; \mathbf{X}_i \neq \hat{\mathbf{Q}}_n(\mathbf{u})} \{|\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})|^{-1}\langle \mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u}), \mathbf{h}\rangle + \langle \mathbf{u}, \mathbf{h}\rangle\} + \sum_{i:1 \leq i \leq n; \mathbf{X}_i = \hat{\mathbf{Q}}_n(\mathbf{u})} \{|\mathbf{h}| + \langle \mathbf{u}, \mathbf{h}\rangle\} \geq 0$$

for all $\mathbf{h} \in \mathbb{R}^d$. Note that we can replace $\mathbf{h}$ by $-\mathbf{h}$ in the preceding inequality to obtain a second version of it. The proof of the theorem is now complete after observing that $||\mathbf{h}| \pm \langle \mathbf{u}, \mathbf{h}\rangle| \leq (1 + |\mathbf{u}|)|\mathbf{h}|$.

Proof of Theorem 3.1.1

We assume that the reader is familiar with the arguments used in developing the main technical results of Chaudhuri (1992a). We split the proof into several paragraphs to clearly expose the key ideas and observations.

First, note that $\Phi(\mathbf{u}, \mathbf{t})$ tends to infinity as $|\mathbf{t}|$ tends to infinity for any fixed $\mathbf{u}$ such that $|\mathbf{u}| < 1$. Then, arguing along the same line as in the proof of lemma 5.2 of Chaudhuri (1992a, pp. 906–907), there exists a constant $K_1 > 0$ such that almost surely $\sum_{i=1}^n \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q}) > \sum_{i=1}^n \Phi(\mathbf{u}, \mathbf{X}_i)$ for all n sufficiently large if $|\mathbf{Q} - \mathbf{Q}(\mathbf{u})| > K_1$. In other words, because $\hat{\mathbf{Q}}_n(\mathbf{u})$ minimizes $\sum_{i=1}^n \Phi(\mathbf{u}, \mathbf{X}_i - \mathbf{Q})$, we must have $|\hat{\mathbf{Q}}_n(\mathbf{u}) - \mathbf{Q}(\mathbf{u})| \leq K_1$ almost surely for all n sufficiently large.

Next consider fact 5.5 of Chaudhuri (1992a, p. 909) in the special case $m = 1$, which corresponds to the case of spatial median. Then our Theorem 2.1.2, which implies that the d-dimensional vector $\sum_{i=1}^n |\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})|^{-1}\{\mathbf{X}_i - \hat{\mathbf{Q}}_n(\mathbf{u})\} + n\mathbf{u}$ remains bounded in magnitude (adopt the convention that $|\mathbf{x}|^{-1}\mathbf{x} = 0 \in \mathbb{R}^d$ if $\mathbf{x} = 0 \in \mathbb{R}^d$ as in Chaudhuri 1992a, p. 900), can be viewed as an extension of this fact for arbitrary geometric quantiles. Consequently, an easy generalization of proposition 5.6 of Chaudhuri (1992a, pp. 910–911) implies the existence of a constant $K_2 > 0$ such that almost surely $|\hat{\mathbf{Q}}_n(\mathbf{u}) - \mathbf{Q}(\mathbf{u})| \leq K_2 n^{-1/2}(\log n)^{1/2}$ for all n sufficiently large. Recall here that $\mathbf{Q}(\mathbf{u})$ satisfies $E[|\mathbf{X}_i - \mathbf{Q}(\mathbf{u})|^{-1}\{\mathbf{X}_i - \mathbf{Q}(\mathbf{u})\} + \mathbf{u}] = 0$, and lemmas 5.3 and 5.4 of Chaudhuri (1992a, pp. 907–909) can be suitably modified to imply that the magnitude of the d-dimensional vector $\sum_{i=1}^n |\mathbf{X}_i - \mathbf{Q}|^{-1}(\mathbf{X}_i - \mathbf{Q}) + n\mathbf{u}$ will explode to infinity almost surely as n tends to infinity, unless $\mathbf{Q}$ lies inside a ball in $\mathbb{R}^d$ with center at $\mathbf{Q}(\mathbf{u})$ and radius of the order $O(n^{-1/2}[\log n]^{1/2})$.

Let $\mathcal{G}_n$ be the subset of $\mathbb{R}^d$ defined as

$$\mathcal{G}_n = \{\mathbf{W} | \mathbf{W} = (w_1, w_2, \ldots, w_d) \in \mathbb{R}^d, n^4 w_i = \text{an integer}, \text{ and } |w_i| \leq K_2 n^{-1/2}(\log n)^{1/2} \text{ for all } 1 \leq i \leq n\}.$$

For $\mathbf{W} \in \mathcal{G}_n$, define

$$\Lambda_n\{\mathbf{Q}(\mathbf{u}), \mathbf{W} + \mathbf{Q}(\mathbf{u})\} = n^{-1}\sum_{i=1}^n [|\mathbf{X}_i - \mathbf{Q}(\mathbf{u})|^{-1}\{\mathbf{X}_i - \mathbf{Q}(\mathbf{u})\} - |\mathbf{X}_i - \mathbf{Q}(\mathbf{u}) - \mathbf{W}|^{-1} \times \{\mathbf{X}_i - \mathbf{Q}(\mathbf{u}) - \mathbf{W}\} + E(|\mathbf{X}_i - \mathbf{Q}(\mathbf{u}) - \mathbf{W}|^{-1} \times \{\mathbf{X}_i - \mathbf{Q}(\mathbf{u}) - \mathbf{W}\} + \mathbf{u})].$$

Then each term in the sum defining $\Lambda_n\{\mathbf{Q}(\mathbf{u}), \mathbf{W} + \mathbf{Q}(\mathbf{u})\}$ has mean zero for every $\mathbf{W} \in \mathcal{G}_n$. Further, if $\Theta_n\{\mathbf{Q}(\mathbf{u}), \mathbf{W} + \mathbf{Q}(\mathbf{u})\}$ denotes the variance–covariance matrix of a term, then we have the following:

a. For $d \geq 3$, $\max_{\mathbf{W}\in\mathcal{G}_n}|\Theta_n\{\mathbf{Q}(\mathbf{u}), \mathbf{W} + \mathbf{Q}(\mathbf{u})\}| = O(n^{-1}\log n)$ as n tends to infinity.
b. For $d = 2$, $\max_{\mathbf{W}\in\mathcal{G}_n}|\Theta_n\{\mathbf{Q}(\mathbf{u}), \mathbf{W} + \mathbf{Q}(\mathbf{u})\}| = o(n^{-\beta})$ as n tends to infinity for any fixed β such that $0 < \beta < 1$.

Observe that both a and b follow from a minor modification of the results stated in lemma 5.7 of Chaudhuri (1992a, p. 911). Then, arguing along the same line as in the proof of lemma 5.9 of Chaudhuri (1992a, p. 912), we can conclude the following:

a. When $d \geq 3, \max_{\mathbf{W}\in\mathcal{G}_n}|\Lambda_n\{\mathbf{Q}(\mathbf{u}), \mathbf{W} + \mathbf{Q}(\mathbf{u})\}| = O(n^{-1}\log n)$ almost surely as n tends to infinity.
b. When $d = 2, \max_{\mathbf{W}\in\mathcal{G}_n}|\Lambda_n\{\mathbf{Q}(\mathbf{u}), \mathbf{W} + \mathbf{Q}(\mathbf{u})\}| = o(n^{-\beta})$ almost surely as n tends to infinity for any fixed β such that $0 < \beta < 1$.

At this point, let $\mathbf{Q}_n^{\#}(\mathbf{u})$ be a point in $\mathbb{R}^d$ such that $\mathbf{Q}_n^{\#}(\mathbf{u}) - \mathbf{Q}(\mathbf{u}) \in \mathcal{G}_n$, and $\mathbf{Q}_n^{\#}(\mathbf{u})$ is closest to $\hat{\mathbf{Q}}_n(\mathbf{u})$ in the Euclidean distance. If there are several possible choices for such a $\mathbf{Q}_n^{\#}(\mathbf{u})$, then we can choose any one of them. Then we can write

$$\begin{aligned} n^{-1}\sum_{i=1}^{n} &[|\mathbf{X}_i - \mathbf{Q}(\mathbf{u})|^{-1}\{\mathbf{X}_i - \mathbf{Q}(\mathbf{u})\} + \mathbf{u}] \\ &= \Lambda_n\{\mathbf{Q}(\mathbf{u}), \mathbf{Q}_n^{\#}(\mathbf{u})\} + n^{-1} \\ &\quad \times \sum_{i=1}^{n} [|\mathbf{X}_i - \mathbf{Q}_n^{\#}(\mathbf{u})|^{-1}\{\mathbf{X}_i - \mathbf{Q}_n^{\#}(\mathbf{u})\} + \mathbf{u}] \\ &\quad - E[|\mathbf{X}_1 - \mathbf{Q}_n^{\#}(\mathbf{u})|^{-1}\{\mathbf{X}_1 - \mathbf{Q}_n^{\#}(\mathbf{u})\} + \mathbf{u}]. \end{aligned}$$

It is quite easy to verify (see the inequality (6) in the proof of proposition 5.6 in Chaudhuri 1992a, p. 910) that

$$n^{-1}\sum_{i=1}^{n} [|\mathbf{X}_i - \mathbf{Q}_n^{\#}(\mathbf{u})|^{-1}\{\mathbf{X}_i - \mathbf{Q}_n^{\#}(\mathbf{u})\} + \mathbf{u}] = O(n^{-1}\log n)$$

almost surely as n tends to infinity. On the other hand, it is straightforward to check (cf. fact 5.8 in Chaudhuri 1992, p. 912) the following:

a. For $d \geq 3$, we will have

$$\begin{aligned} \max_{\mathbf{W}\in\mathcal{G}_n} |E(&|\mathbf{X}_1 - \mathbf{Q}(\mathbf{u}) - \mathbf{W}|^{-1}\{\mathbf{X}_1 - \mathbf{Q}(\mathbf{u}) - \mathbf{W}\} + \mathbf{u}) \\ &+ [\mathbf{D}_1\{\mathbf{Q}(\mathbf{u})\}]\mathbf{W}| = O(n^{-1}\log n) \quad \text{as} \quad n \to \infty. \end{aligned}$$

b. For $d = 2$ and a fixed β such that $0 < \beta < 1$, we will have

$$\begin{aligned} \max_{\mathbf{W}\in\mathcal{G}_n} |E(&|\mathbf{X}_1 - \mathbf{Q}(\mathbf{u}) - \mathbf{W}|^{-1}\{\mathbf{X}_1 - \mathbf{Q}(\mathbf{u}) - \mathbf{W}\} + \mathbf{u}) \\ &+ [\mathbf{D}_1\{\mathbf{Q}(\mathbf{u})\}]\mathbf{W}| = o(n^{-\beta}) \quad \text{as} \quad n \to \infty. \end{aligned}$$

The proof of the theorem is now complete, using the positive definiteness of the matrix $\mathbf{D}_1\{\mathbf{Q}(\mathbf{u})\}$ and the fact that $\mathbf{Q}_n^{\#}(\mathbf{u}) - \hat{\mathbf{Q}}_n(\mathbf{u}) = O(n^{-4})$ in view of the definitions of $\mathcal{G}_n$ and $\mathbf{Q}_n^{\#}(\mathbf{u})$.

Proof of Theorem 3.1.2

The proof follows by applying the Cramér–Wold (1936) device (see, e.g., Serfling 1980) and the central limit theorem to the linear term in the Bahadur expansion derived in Theorem 3.1.1.

Proof of Theorem 3.2.1

Because $1 - n^{-1}f_n$ remains bounded away from zero as n tends to infinity, and $\mathbf{A}_n^{(i)}(\mathbf{u}, \mathbf{v})$ is a bounded random matrix, it is obvious that

$$\mathbf{F}_n(\mathbf{u}, \mathbf{v}) - \mathbf{D}_2\{\hat{\mathbf{Q}}_n^*(\mathbf{u}), \hat{\mathbf{Q}}_n^*(\mathbf{v}), \mathbf{u}, \mathbf{v}\} = O_P(n^{-1/2})$$

as n tends to infinity. Note that here we are using the fact that the $\mathbf{X}_i$'s with $i \in \mathcal{F}_n$ and the $\mathbf{X}_i$'s with $i \in \mathcal{F}_n^c$ form two independent subsamples. Also, when $d \geq 2$, some of the arguments used in the proof of the main theorem of Bose and Chaudhuri (1993) can be appropriately modified to establish the existence of a nonnegative random variable Z_n such that $Z_n = O_P(1)$ as n tends to infinity and

$$\begin{aligned} |\mathbf{D}_2\{\hat{\mathbf{Q}}_n^*(\mathbf{u}), \hat{\mathbf{Q}}_n^*(\mathbf{v}), \mathbf{u}, \mathbf{v}\} - \mathbf{D}_2\{\mathbf{Q}(\mathbf{u}), \mathbf{Q}(\mathbf{v}), \mathbf{u}, \mathbf{v}\}| \\ \leq Z_n\{\max(|\hat{\mathbf{Q}}_n^*(\mathbf{u}) - \mathbf{Q}(\mathbf{u})|, |\hat{\mathbf{Q}}_n^*(\mathbf{v}) - \mathbf{Q}(\mathbf{v})|\}. \end{aligned}$$

Because $n^{-1}f_n$ also remains bounded away from zero as n tends to infinity, Theorems 3.1.1 and 3.1.2 now guarantee that

$$\max(|\hat{\mathbf{Q}}_n^*(\mathbf{u}) - \mathbf{Q}(\mathbf{u})|, |\hat{\mathbf{Q}}_n^*(\mathbf{v}) - \mathbf{Q}(\mathbf{v})|) = O_P(n^{-1/2})$$

as n tends to infinity.

Next, observe that for $d \geq 3$, $\mathbf{G}_n(\mathbf{u}) - \mathbf{D}_1\{\hat{\mathbf{Q}}_n^*(\mathbf{u})\}$ is an average of conditionally iid terms, each of which has zero conditional mean and finite conditional second moment given the $\mathbf{X}_i$'s for which $i \in \mathcal{F}_n$ (see the results in sec. 3 in Bose and Chaudhuri 1993). Therefore, we must have

$$\mathbf{G}_n(\mathbf{u}) - \mathbf{D}_1\{\hat{\mathbf{Q}}_n^*(\mathbf{u})\} = O_P(n^{-1/2})$$

as n tends to infinity. But when $d = 2$, the conditional second moment may not be finite, and in that case a result of Bose and Chandra (1993) used in the proof of the main theorem of Bose and Chaudhuri (1993) (see the case $d = 2$ there) guarantee that

$$\mathbf{G}_n(\mathbf{u}) - \mathbf{D}_1\{\hat{\mathbf{Q}}_n^*(\mathbf{u})\} = o_P(n^{-\beta})$$

as n tends to infinity, where β is any constant such that $0 < \beta < 1/2$. Finally, a straightforward application of some of the crucial observations made in course of the development of the proof of the main theorem of Bose and Chaudhuri (1993) yield the following:

a. If $d \geq 3$, then there exists a nonnegative random variable V_n such that $V_n = O_P(1)$ as n tends to infinity, and

$$|\mathbf{D}_1\{\hat{\mathbf{Q}}_n^*(\mathbf{u})\} - \mathbf{D}_1\{\mathbf{Q}(\mathbf{u})\}| \leq V_n|\hat{\mathbf{Q}}_n^*(\mathbf{u}) - \mathbf{Q}(\mathbf{u})|.$$

b. If $d = 2$ and β is any constant such that $0 < \beta < 1$, then we will have

$$|\mathbf{D}_1\{\hat{\mathbf{Q}}_n^*(\mathbf{u})\} - \mathbf{D}_1\{\hat{\mathbf{Q}}_n^*(\mathbf{u})\}||\hat{\mathbf{Q}}_n^*(\mathbf{u}) - \mathbf{Q}(\mathbf{u})|^{-\beta} = O_P(1).$$

The proof of the theorem is now complete, using the $n^{1/2}$ consistency of $\hat{\mathbf{Q}}_n^*(\mathbf{u})$ as an estimate of $\mathbf{Q}(\mathbf{u})$.

[Received August 1993. Revised October 1995.]

REFERENCES

Abdous, B., and Theodorescu, R. (1992), "A Note on the Spatial Quantile of a Random Vector," *Statistics and Probability Letters* 13, 333–336.

Babu, G. J. (1986), "A Note on Bootstrapping the Variance of Sample Quantile," *Annals of the Institute of Statistical Mathematics*, 38, 439–443.

Babu, G. J., and Rao, C. R. (1988), "Joint Asymptotic Distribution of Marginal Quantile Functions in Samples From a Multivariate Population," *Journal of Multivariate Analysis*, 27, 15–23.

Bahadur, R. R. (1966), "A Note on Quantiles in Large Samples," *The Annals of Mathematical Statistics*, 37, 577–580.

Bai, Z. D., Chen, N. R., Miao, B. Q., and Rao, C. R. (1990), "Asymptotic Theory of Least Distances Estimate in Multivariate Linear Models," *Statistics*, 21, 503–519.

Barnett, V. (1976), "The Ordering of Multivariate Data" (with comments), *Journal of the Royal Statistical Society*, Ser. A, 139, 318–354.

Bedall, F. K., and Zimmermann, H. (1979), "Algorithm AS 143, the Mediancenter," *Applied Statistics*, 28, 325–328.

Bickel, P. J. (1973), "On Some Analogues of Linear Combinations of Order Statistics in the Linear Model," *The Annals of Statistics*, 1, 597–616.

Bose, A., and Chandra, T. K. (1993), "Cesaro Uniform Integrability," *Sankhyā*, Ser. A, 55, 12–28.

Bose, A., and Chaudhuri, P. (1993), "On the Dispersion of Multivariate Median," *Annals of the Institute of Statistical Mathematics*, 45, 541–550.

Brown, B. M. (1983), "Statistical Use of the Spatial Median," *Journal of the Royal Statistical Society*, Ser. B, 45, 25–30.

——— (1985), "Spatial Median," in *Encyclopedia of Statistical Science*, Vol. 8, New York: John Wiley, pp. 574–575.

Brown, B. M., and Hettmansperger, T. P. (1987), "Affine Invariant Rank Methods in the Bivariate Location Model," *Journal of the Royal Statistical Society*, Ser. B, 49, 301–310.

——— (1989), "An Affine Invariant Bivariate Versions of the Sign Test," *Journal of the Royal Statistical Society*, Ser. B, 51, 117–125.

Chakraborty, B., and Chaudhuri, P. (1994), "On a Transformation and Retransformation Technique for Constructing Affine Equivariant Multivariate Median," Technical Report No. 22/94, Division of Theoretical Statistics and Mathematics, Indian Statistical Institute, Calcutta.

——— (1995), "On an Adaptive Transformation and Retransformation Estimate of Multivariate Location," Technical Report No. 11/95, Division of Theoretical Statistics and Mathematics, Indian Statistical Institute, Calcutta.

Chaudhuri, P. (1992a), "Multivariate Location Estimation Using Extension of R-Estimates Through U-Statistics Type Approach," *The Annals of Statistics*, 20, 897–916.

——— (1992b), "Generalized Regression Quantiles: Forming a Useful Toolkit for Robust Linear Regression," in L_1 *Statistical Analysis and Related Methods*, ed. Y. Dodge, Amsterdam: North-Holland, pp. 169–185.

Chaudhuri, P., and Sengupta, D. (1993), "Sign Tests in Multidimension: Inference Based on the Geometry of the Data Cloud," *Journal of the American Statistical Association*, 88, 1363–1370.

Cramér, H., and Wold, H. (1936), "Some Theorems on Distribution Functions," *Journal of the London Mathematical Society*, 11, 290–295.

Donoho, D. L., and Gasko, M. (1992), "Breakdown Properties of Location Estimates Based on Halfspace Depth and Projected Outlyingness," *The Annals of Statistics*, 20, 1803–1827.

Dudley, R. M., and Koltchinskii, V. (1992), "The Spatial Quantiles," unpublished manuscript.

Eddy, W. F. (1982), "Convex Hull Peeling," in *COMPSTAT 1982 for IASC*, Vienna: Pysica-Verlag, pp. 42–47.

——— (1983), "Set Valued Ordering of Bivariate Data," in *Stochastic Geometry, Geometric Statistics, and Stereology*, eds. R. V. Ambartsumian and W. Weil, Leipzig: Tuebner, pp. 79–90.

——— (1985), "Ordering of Multivariate Data," in *Computer Science and Statistics: The Interface*, ed. L. Billard, Amsterdam: North–Holland, pp. 25–30.

Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, Philadelphia: SIAM.

——— (1991), "Regression Percentiles Using Asymmetric Squared Error Loss," *Statistica Sinica*, 1, 93–125.

Einmahl, J. H. J., and Mason, D. (1992), "Generalized Quantile Processes," *The Annals of Statistics*, 20, 1062–1078.

Evans, M. (1982), "Confidence Bands for Bivariate Quantiles," *Communications in Statistics, Part A—Theory and Methods*, 11, 1465–1474.

Ferguson, T. (1967), *Mathematical Statistics: A Decision Theoretic Approach*, New York: Academic Press.

Ghosh, J. K. (1971), "A New Proof of the Bahadur Representation of Quantiles and an Application," *The Annals of Mathematical Statistics*, 42, 1957–1961.

Ghosh, M., Parr, W. C., Singh, K., and Babu, G. J. (1984), "A Note on Bootstrapping the Sample Median," *The Annals of Statistics*, 12, 1130–1135.

Gordaliza, A. (1991), "On the Breakdown Point of Multivariate Location Estimators Based on Trimming Procedures," *Statistics & Probability Letters*, 11, 387–394.

Gower, J. C. (1974), "Algorithm AS 78: The Mediancenter," *Applied Statistics*, 23, 466–470.

Haldane, J. B. S. (1948), "Note on the Median of a Multivariate Distribution," *Biometrika*, 35, 414–415.

Hall, P., DiCiccio, T. J., and Romano, J. P. (1989), "On Smoothing and the Bootstrap," *The Annals of Statistics*, 17, 692–704.

Hall, P., and Martin, M. A. (1988), "Exact Convergence Rate of the Bootstrap Quantile Variance Estimator," *Probability Theory & Related Fields*, 80, 261–268.

——— (1991), "On the Error Incurred in Using the Bootstrap Variance Estimate When Constructing Confidence Intervals for Quantiles," *Journal of Multivariate Analysis*, 38, 70–81.

Hayford, J. F. (1902), "What is the Center of an Area, or the Center of a Population?," *Journal of the American Statistical Association*, 8, 47–58.

Hettmansperger, T. P., Nyblom, J., and Oja, H. (1994), "Affine Invariant Multivariate One Sample Sign Tests," *Journal of the Royal Statistical Society*, Ser. B, 56, 221–234.

Jan, S. L., and Randles, R. H. (1994), "A Multivariate Signed Sum Test for the One Sample Location Problem," *Journal of Nonparametric Statistics*, 4, 49–63.

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.

Kiefer, J. (1967), "On Bahadur's Representation of Sample Quantiles," *The Annals of Mathematical Statistics*, 38, 1323–1342.

Kim, S. J. (1992), "A Metrically Trimmed Mean as a Robust Estimator of Location," *The Annals of Statistics*, 20, 1534–1547.

Koenker, R., and Bassett, G. (1978), "Regression Quantiles," *Econometrica*, 46, 33–50.

Koenker, R., and Portnoy, S. (1987), "L Estimation for Linear Models," *Journal of the American Statistical Association*, 82, 851–857.

Koltchinskii, V. (1993), "Bahadur–Kiefer Approximation for Spatial Quantiles," unpublished manuscript.

León, C. A., and Massé, J. C. (1992), "A Counterexample on the Existence of the L_1 Median," *Statistics and Probability Letters*, 13, 117–120.

Liu, R. Y. (1990), "On a Notion of Data Depth Based on Random Simplices," *The Annals of Statistics*, 18, 405–414.

——— (1992), "Data Depth and Multivariate Rank Tests," in L_1 *Statistical Analysis and Related Methods*, ed. Y. Dodge, Amsterdam: North-Holland, pp. 279–294.

Mahalanobis, P. C. (1936), "On the Generalized Distance in Statistics," in *Proceedings of the National Institute of Science, India*, Vol. 12, pp. 49–55.

Mottonen, J., and Oja, H. (1995), "Multivariate Spatial Sign and Rank Methods," *Journal of Nonparametric Statistics*, 5, 201–213.

Niemiro, W. (1992), "Asymptotics for M-Estimators Defined by Convex Minimization," *The Annals of Statistics*, 20, 1514–1533.

Nolan, D. (1992), "Asymptotics for Multivariate Trimming," *Stochastic Processes and Their Applications*, 42, 157–169.

Oja, H. (1983), "Descriptive Statistics for Multivariate Distributions," *Statistics and Probability Letters*, 1, 327–332.

Plackett, R. L. (1976), Comment on "The Ordering of Multivariate Data," by V. Barnett, *Journal of the Royal Statistical Society*, Ser. A, 139, 344–346.

Pyke, R. (1975), "Multidimensional Empirical Processes: Some Comments," in *Stochastic Processes and Related Topics, Proceedings of the Summer Research Institute on Statistical Inference for Stochastic Processes*, Vol. 2, New York: Academic Press, pp. 45–58.

——— (1984), Comment on "Some Limit Theorems for Empirical Processes," by E. Giné and J. Zinn, *The Annals of Probability*, 12, 996–997.

——— (1985), "Opportunities for Set Indexed Empirical and Quantile Processes in Inference," in *Bulletin of the International Statistical Institute: Proceedings of the 45th Session, Invited Papers*, Vol. 51, Voozburg, The Netherlands: International Statistical Institute, pp. 25.2.1–25.2.12.

Rao, C. R. (1988), "Methodology Based on the L_1 Norm in Statistical Inference," *Sankhyā*, Ser. A, 50, 289–313.

Reiss, R. D. (1989), *Approximate Distributions of Order Statistics With Applications to Nonparametric Statistics*, New York: Springer.

Ruppert, D., and Carroll, R. J. (1980), "Trimmed Least Squares Estimation in the Linear Model," *Journal of the American Statistical Association*, 75, 828–838.

Scates, D. E. (1933), "Locating the Median of the Population in the United States," *Metron*, 11, 49–66.

Serfling, R. J. (1980), *Approximation Theorems of Mathematical Statistics*, New York: John Wiley.

Shao, J. (1990), "Bootstrap Estimation of the Asymptotic Variances of Statistical Functions," *Annals of the Institute of Statistical Mathematics*, 42, 737–752.

Shao, J., and Wu, C. F. J. (1989), "A General Theory for Jackknife Variance Estimation," *The Annals of Statistics*, 17, 1176–1197.

Small, C. G. (1990), "A Survey of Multidimensional Medians," *International Statistical Review*, 58, 263–277.

Stahel, W. A. (1981), "Robuste Schatzungen: Infinitesimale Optimalitat Und Schatzungen Von Kovarianmatrizen," unpublished Ph.D. dissertation, ETH, Zurich.

Tukey, J. W. (1975), "Mathematics and Picturing Data," in *Proceedings of the International Congress of Mathematicians, Vancouver 1974*, Vol. 2, ed. R. D. James, Montreal: Canadian Mathematical Congress, p. 523–531.

Valadier, M. (1984), "La Multiapplication Medianes Conditionelles," *Zeitschrift Wahrscheinlichkeitsrechnung*, 67, 279–282.

Weber, A. (1909), *Uber Den Standord Der Industrien, Tubingen.* English translation by C. J. Freidrich (1929), *Alfred Weber's Theory of Location of Industries*, Chicago: Chicago University Press.

Welsh, A. H. (1987a), "The Trimmed Mean in the Linear Model" (with comments), *The Annals of Statistics*, 15, 20–45.

——— (1987b), "One-Step L Estimation in Linear Model," *The Annals of Statistics*, 15, 626–641.

Wilks, S. S. (1932), "Certain Generalizations in the Analysis of Variance," *Biometrika*, 24, 471–494.