International Statistical Review (1990), **58**, 3, pp. 263–277. Printed in Great Britain

A Survey of Multidimensional Medians

Christopher G. Small

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Summary

In this paper we survey a sequence of papers whose primary aim is the generalization of the concept of the median into higher dimensional settings. While a variety of distinct definitions of the median of a multivariate data set are possible these definitions have the common property of producing the usual definition when applied to univariate data or a univariate distribution. Some common ideas of equivariance, symmetry and breakdown are discussed as well as computational convenience for each definition. The extension of these ideas to directional statistics is also discussed.

Key words: Affine; Breakdown; Directional statistics; Equivariance; Invariance; Median; Multivariate distribution; Symmetry.

1 Background and History

The median of a data set arises chiefly in nonparametric problems as a natural robust estimate for the center of a distribution. As a nonparametric and robust estimate for the center of a distribution, it has a different character from the sample mean as is illustrated by different breakdown properties. Thus it suffices to have a single point contaminating a data set and going off to infinity to send the mean to infinity as well. By contrast, at least 50% of the data must be moved to infinity to force the median to do the same. As Sibson (1984) has mentioned, the geometry of quadratic decompositions makes methods based upon squared distances particularly amenable to generalization from one dimension into higher dimensions. Therefore, multivariate analysis has been dominated by theory based upon the multivariate normal distribution. Nevertheless, the class of nonparametric and distribution-free methods have a more appropriate place in the analysis of higher dimensional data than they have hereto been given.

What is the appropriate analog in two or more dimensions of the univariate median? We shall explore a number of distinct answers to this question as they have arisen in the literature. Even if it is not at all clear what the appropriate analog is, there are certain properties it should definitely have. For example, in dimension one, the center of symmetry of a distribution is its median. One would naturally suppose the same to be true in higher dimensions, making the higher dimensional median a natural estimate for the center of symmetry. Even here, however, there is some ambiguity: the center of symmetry of a univariate distribution can be generalized in several ways. In this paper we shall consider the use of multidimensional medians in the estimation and testing of *four* types of symmetry: points of spherical, elliptical, central and angular symmetry which we shall now describe.

Definition 1.1 The distribution of a random vector $\mathbf{X} \in \mathbf{R}^p$ is *spherically symmetric* about $\boldsymbol{\mu}$ provided $\mathbf{X} - \boldsymbol{\mu}$ and $L(\mathbf{X} - \boldsymbol{\mu})$ are identically distributed for any orthogonal linear transformation $L: \mathbf{R}^p \to \mathbf{R}^p$. The distribution of a random vector $\mathbf{X}$ is said to be *elliptically* *symmetric* if there exists some linear transformation $T:\mathbf{R}^p \rightarrow \mathbf{R}^p$ of full rank for which $\mathbf{Y} = T(\mathbf{X})$ has a spherically symmetric distribution. If we relax the criterion of spherical symmetry by restricting L to be the involution $L(\mathbf{x}) = -\mathbf{x}$ then the condition becomes that of *central symmetry* about $\boldsymbol{\mu}$. The concept of *angular symmetry* was suggested by Liu (1988) and is a further relaxation of the symmetry condition. The random vector $\mathbf{X}$ is said to be angularly symmetric about $\boldsymbol{\mu}$ if the standardized vector $(\mathbf{X} - \boldsymbol{\mu})/\|\mathbf{X} - \boldsymbol{\mu}\|$ is centrally symmetric about the origin.

Clearly, any point $\boldsymbol{\mu}$ of spherical symmetry is a point of elliptical symmetry and every point of elliptical symmetry is a point of central symmetry. In turn, any point of central symmetry is a point of angular symmetry. Note that in dimension $p = 1$ a point of angular symmetry is simply a median. A point $\boldsymbol{\mu}$ is necessarily unique if it is a point of spherical or central symmetry, but need not be unique if it is a point of angular symmetry. However, the exceptions to uniqueness in this final case arise from distributions concentrated on a one-dimensional flat. This follows from the results of Milasevic & Ducharme (1987). As a sample drawn from some symmetric distribution will not possess the symmetry properties of its parent distribution the multidimensional median serves the purpose of estimating the point of underlying symmetry.

Closely related to the concept of a point of symmetry is the idea of the equivariance (or invariance) of a median to estimate that point of symmetry. For example, the univariate median is equivariant under monotone transformations of the line: if $X_1, \ldots, X_n$ is a sample with median $\hat{\mu}(X_1, \ldots, X_n)$ and $h:\mathbf{R} \rightarrow \mathbf{R}$ is a monotone transformation, then

$$\hat{\mu}[h(X_1), \ldots, h(X_n)] = h[\hat{\mu}(X_1, \ldots, X_n)].$$

In higher dimensions we expect an estimator of a point of spherical symmetry to be equivariant under the class of transformations which preserve spherical symmetry, namely the rigid Euclidean motions and the isotropic rescalings of $\mathbf{R}^p$. Together these form the similarity transformations of the space, or the conformal affine group. However, for the estimation of elliptical symmetries the assumption of equivariance under the group of affine transformations (i.e. linear transformation coupled with translations) seems more appropriate. A much richer theory of equivariance and invariance can be established by fixing a median of the distribution or sample and considering the class of transformations of the space which leave the median invariant. In dimension one, this is naturally the class of monotone transformations which, of course, does not require specification of the median. In higher dimensions, however, the class will typically depend on the choice of the median point. The success of such a theory is dependent on two issues. In the special case of univariate distributions and samples, does the class of transformations become the class of monotone transformations? Secondly, does the class of transformations contain special cases of transformations which map the entire Euclidean space to a compact set? If the answer to this second question is yes, then the median has some natural robustness against mass at infinity. Closely related to this is the concept of breakdown.

The breakdown of an estimator is a useful property in understanding its robustness. Data can often have contaminating points, and so it is essential to know what proportion of the data can be contaminated without sending the estimator off to arbitrarily large and absurd values. A definition of breakdown was introduced by Hodges (1967) for the univariate case, and by Hampel (1971) more generally. Let $\mathbf{X} = \mathbf{X}_1, \ldots, \mathbf{X}_n$ be a sample and $\hat{\boldsymbol{\mu}}(\mathbf{X})$ an estimator from that sample. Let $\mathbf{X}'$ be a corrupted sample obtained by replacing m of the original points by arbitrary ones and suppose $\hat{\boldsymbol{\mu}}(\mathbf{X}')$ is the corresponding estimator. The maximum bias caused by such contamination is

$$\text{Bias}\,(m; \hat{\boldsymbol{\mu}}) = \sup_{\mathbf{X}'} \|\hat{\boldsymbol{\mu}}(\mathbf{X}') - \hat{\boldsymbol{\mu}}(\mathbf{X})\|;$$

see Donoho & Huber (1983). Infinite values for the bias indicate that the estimator has 'broken down'. The breakdown point can now be defined as

$$\text{BD}\,(\hat{\boldsymbol{\mu}}; \mathbf{X}) = \inf\left\{\frac{m}{n} : \text{Bias}\,(m, \hat{\boldsymbol{\mu}}) = \infty\right\}.$$

The largest possible value for breakdown point of any sensible estimator of location is $\frac{1}{2}$, as is achieved for the univariate median. As we shall see, not all multidimensional medians achieve this value.

In a different direction, in studying the large-sample properties of an estimator $\hat{\boldsymbol{\mu}}$, asymptotic normality such as

$$n^{\frac{1}{2}}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}) \rightarrow N(0, \Sigma)$$

for some Σ is useful. In particular, it gives the rate of convergence as well as the efficiency of the estimator. However, the existence of asymptotic results should not obscure the possibilities for the construction of exact confidence regions based upon some estimators.

Closely related to the concept of a multidimensional median is the concept of a multidimensional quantile among which a median would be defined as the most central quantile. Barnett (1976) has extensively surveyed the available methods for the ordering of multivariate data. Our description will complement and update this work. Eddy (1983, 1985) has developed the approach to quantiles based upon nested sequences of sets. A different approach to quantiles uses the fact that the maximization of the function $-E_F\,|X - \mu|$ can be done by gradients. Of course in dimension one this is simply the derivative. The gradient vector is $P(X > \mu) - P(X < \mu)$ which is a linear transform of the usual quantile. In higher dimensions the gradient vector will typically point inwards to the center of the distribution with a length that is proportional to how exterior the location μ is with respect to the distribution or its empirical analog from the data set.

A starting point for some of the early work in spatial medians was the Twelfth Census of the United States conducted in 1900. Statisticians of the time expressed an interest in studying the flow of population in the United States through the movement over time of a geographical center of population. (There seems to have been no recognition of the importance of the curvature of the Earth in this work! The population was regarded as living on an essentially flat map in which lines of latitude and longitude were orthogonal straight lines.) An early reference to the idea of a spatial median in the statistical literature is to be found in Hayford (1902). A clear distinction is made in that paper between the centroid of a spatial distribution and a median-like estimate of the center of a distribution which is equally as sensitive to probability mass close to the center as it is to probability mass in the extremes. The centroid of the geographical distribution was regarded as an inappropriate measure because of the property that the death of an individual on the periphery of the country such as San Francisco has a greater influence on the centroid than does the death of someone in Indiana. Hayford (1902) suggested the vector of medians of orthogonal coordinates as a solution to this problem but clearly recognized the difficulty that this higher dimensional analog of the median is dependent on the choice of orthogonal coordinates used. Throughout this period, there seems to have been some confusion about the distinction between a centroid and the vector of medians of orthogonal coordinates. See Eells (1930) for the final clarification on this point. Gini & Galvani (1929) introduced the Weber (1909) definition set out in § 2 into the statistical literature. Scates (1933) reexamined the problem of finding the geographical center of the United States, and found it (with the new concept of spatial median) to be '15 miles northwest of Dayton, Ohio'. This was presented as part of a collection of papers devoted to the issue in Vol. 11 of *Metron* which should perhaps have been a watershed for the idea. Other papers in that issue include Griffin (1933) and Galvani (1933). However, these papers did not receive widespread attention. For example, Haldane (1948) rediscovered the same concept of multidimensional median developed by Gini and Galvani, once again within the statistical literature. However, there is no indication in that paper of the mathematical theory that had been developed. Haldane introduced the term 'geometrical median' for the concept to distinguish it from the 'arithmetic median', i.e. the vector of coordinate medians. However, Gower (1974), referring to this paper, terms the same concept the 'mediancentre'. Brown (1983) introduces the term 'spatial median' in the same context. The median due to Oja (1983) has been called the 'generalized median' and the 'affine median'. Donoho & Gasko (1987a) have referred to the median based upon Tukey's depth as the 'multivariate median'. None of these labels is particularly descriptive and so is easily confused with any other. We have decided to adopt the following terminology: all higher dimensional extensions of the univariate median shall be termed 'geometric medians' provided that they are equivariant with respect to rigid Euclidean motions. Definitions such as the vector of medians of coordinates which depend upon the choice of coordinate systems then remain as 'arithmetic medians'. Specific definitions shall be identified either by a simple descriptive label (i.e. halfspace median, convex hull strip median) or, in cases where this becomes too cumbersome, by the originator of the definition (i.e. Oja's median, Fisher's directional median). Thus we shall refer to the median described by Gini and Galvani as the L_1 *median*.

We close this section by reminding the reader that the scope of this survey is restricted to multidimensional medians and not to the much wider topic of robust estimators of location. Our survey does include directional statistics but does not include half-sample methods such as the minimum volume sphere (Rousseeuw, 1984) or minimum volume ellipsoid (Rousseeuw, 1984, 1986). Such estimators imitate the 50% breakdown properties of the univariate median, but are not true multidimensional medians because they do not generalize the usual concept for univariate cases. See Rousseeuw & Leroy (1987) for a discussion of robust regression methods. We shall discuss the sense in which a directional median generalizes a Euclidean median when we return to that topic later.

2 The L_1 Median

Weber (1909) considered the following problem in Location Theory. A company wishes to choose an appropriate location for a warehouse which will service n customers whose planar coordinates are given to be $\mathbf{X}_1, \ldots, \mathbf{X}_n$. We idealize the situation by supposing that the company can locate the warehouse at any coordinates without constraints. We further idealize the problem by supposing that the transportation costs for deliveries from the warehouse to customers are proportional to Euclidean distance. The proposed solution to the location problem suggested by Weber is to locate the warehouse so as to minimize the sum total of transportation costs to all customers. If deliveries are made to all customers equally, then the optimal location is that point $\hat{\boldsymbol{\mu}}$ which minimizes the sum of the distances from $\hat{\boldsymbol{\mu}}$ to X_i.

Definition 2.1. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be n points lying in $\mathbf{R}^p$. We define an L_1 *median* of the set of points $\mathbf{X}_1, \ldots, \mathbf{X}_n$ to any point $\hat{\boldsymbol{\mu}} \in \mathbf{R}^p$ which minimizes.

$$\sum_{i=1}^{n} \|\mathbf{X}_i - \mu\|. \tag{2.1}$$

In the special case where $p = 1$ the L_1 median is well known to reduce to the standard univariate median. At the time of Weber's original work there were few computational tools available for finding such points $\hat{\boldsymbol{\mu}}$ when $p \geqslant 2$. An exception to this is the case where $n = 3$ for which the L_1 median is known to be the Steiner point of the triangle $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3$. The L_1 median of 10 points is shown in Fig. 1.

Some geometrical insight into the nature of these medians can be obtained from the following characterization. Let r_{ij} be the jth coordinate of the vector $\mathbf{X}_i - \boldsymbol{\mu}$, and let $r_i^2 = \Sigma r_{ij}^2$, where the sum is over $j = 1, \ldots p$. We perturb μ in the kth coordinate by an amount ε and set the derivative of (2.1) to be zero. From this we obtain

$$\sum_{i=1}^{n} \left[\frac{r_{ik}}{r_i} \right] = 0$$

for all $k = 1, 2, \ldots, p$. If an equation does not have a solution, '$= 0$' is to be interpreted as 'changes sign'. This can occur when the L_1 median coincides with one of the points $\mathbf{X}_i$. A geometrical interpretation is that the centroid of the vectors

$$A(\hat{\boldsymbol{\mu}}; \mathbf{X}_i) = \frac{\mathbf{X}_i - \hat{\boldsymbol{\mu}}}{\|\mathbf{X}_i - \hat{\boldsymbol{\mu}}\|}$$

is the origin in $\mathbf{R}^p$. As mentioned in § 1, the centroid $n^{-1} \sum A(\mu; X_i)$ is the multivariate analog of the quantile in dimension one. It arises as a natural generalization of the sign test to a direction test in higher dimensions. Thus, for example, in dimension $p = 1$ the quantity $A(\mu; X_i)$ reduces to the sign of $X_i - \mu$ and $\sum A(\mu; X_i)$ has a recentered binomial $(n, \frac{1}{2})$ distribution when μ is the true median. This geometrical interpretation makes the robustness properties of the L_1 median easier to interpret. An observation $\mathbf{X}_i$ can be moved out to infinity without changing the median provided it is moved along the directional vector $A(\hat{\boldsymbol{\mu}}; \mathbf{X}_i)$. The class of transformations so obtained is the extension of the class of monotone transformations to higher dimensions that preserves the invariance properties of the L_1 median.

The geometrical interpretation also makes the testing for symmetry relatively straightforward. Under the null hypothesis that $\boldsymbol{\mu}$ is a point of spherical symmetry, the directional vectors $A(\boldsymbol{\mu}; \mathbf{X}_i)$ are uniformly distributed on the unit $(p-1)$-dimensional sphere in $\mathbf{R}^p$ when $\boldsymbol{\mu}$ is the true L_1 median of the distribution. Thus a test for spherical symmetry at $\boldsymbol{\mu}$ can be interpreted as a test for uniformity on the sphere. Mardia (1972) discusses this in detail in the context of directional statistics. The problem of testing for central symmetry at $\boldsymbol{\mu}$ is complicated by the lack of a distribution-free statistic. Under the null hypothesis the directional vectors $A(\boldsymbol{\mu}; \mathbf{X}_i)$ continue to have their first moment centered at the origin $\mathbf{R}^p$. However, their distributions will not be determined by the hypothesis of central symmetry. Brown (1983) has suggested a chi-square statistic based upon the principle that $\sum A(\boldsymbol{\mu}; \mathbf{X}_i)$, where the sum is over $i = 1, \ldots, n$, converges via the central limit theorem

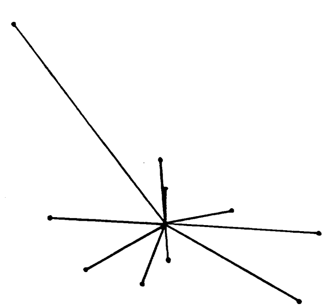

Figure 1. *The L_1 median of* 10 *points.*

to a multivariate normal vector with mean at the origin and with unknown covariance matrix that can be estimated by the sample covariance matrix. This leads to an approximate chi-square test of the hypothesis.

Lopuhaa & Rousseeuw (1987) have investigated the breakdown of the L_1 median and have found it to be $\frac{1}{2}$. A simple geometrical interpretation of this result is that if at least 50% of the data lies on a given point $\boldsymbol{\mu}$ then that point will be the L_1 median of the data. This follows directly from the geometrical interpretation given above.

Brown (1983) has studied the asymptotic properties of the L_1 median. When sampling from a distribution whose true median is $\boldsymbol{\mu}$ the quantity $n^{\frac{1}{2}}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})$ converges to bivariate normal. Surprisingly, Brown notes that when sampling from a multivariate spherically symmetric normal distribution, the asymptotic relative efficiency of the L_1 median increases to one as $p \to \infty$.

We close this section with a few comments about the uniqueness of the L_1 median. By inspection of the equations leading to the median we observe that if all points $\mathbf{X}_i$ of the data set are collinear, i.e have a one-dimensional flat passing through them, then the L_1 median reduces to the usual median on that flat. Thus the median need not be unique. However, this exception aside, the equations have a unique solution. So, the L_1 median is unique for data sets that lie in dimensions 2 or more. See Milasevic & Ducharme (1987) for a detailed discussion of this.

3 Oja's Simplex Median

The Oja median provides an interesting alternative to the L_1 median in that it possesses the additional property of *affine equivariance*. Consider a point $\boldsymbol{\mu}$ in $\mathbf{R}^p$ which is a point of central or elliptical symmetry of a distribution F. Suppose T is an affine transformation from $\mathbf{R}^p$ to $\mathbf{R}^p$ of full rank. We can write $T(\mathbf{X}) = L(\mathbf{X}) + \mathbf{b}$, where L is a linear transformation on $\mathbf{R}^p$. Then it can be seen that $T(\boldsymbol{\mu})$ is a point of central or elliptical symmetry of the distribution FT^{-1} induced on $\mathbf{R}^p$ by T. Suppose $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is a sample from distribution F, and $\hat{\boldsymbol{\mu}}$ is its associated L_1 median. One might naturally wish that $T(\hat{\boldsymbol{\mu}})$ be the L_1 median of the sample $T(\mathbf{X}_1), \ldots, T(\mathbf{X}_n)$ estimating $T(\boldsymbol{\mu})$. However, $T(\hat{\boldsymbol{\mu}})$ is not in general the L_1 median of the transformed sample. The additional restriction that L be an orthogonal transformation needs to be imposed to ensure the result. This failure of affine equivariance is dissatisfying and leads to consideration of affine equivariant spatial medians.

In dimension one, the absolute distance $|X_i - \mu|$ when generalized to $\|\mathbf{X}_i - \boldsymbol{\mu}\|$ in dimension p extends the usual L_1 definition of the median to the L_1 median. However $|X_i - \mu|$ can alternatively be interpreted as the length of a simplex with vertices X_i and μ. Oja (1983) has proposed a simplex definition of spatial median, which we now describe. In a sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$ in $\mathbf{R}^p$ we define $c[\mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_p}; \boldsymbol{\mu}]$ to be the p-dimensional volume of the simplex in $\mathbf{R}^p$ whose vertices are $\mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_p}$ and $\boldsymbol{\mu}$, where $i_1 < i_2 < \ldots < i_p$.

Definition 3.1. An Oja simplex median of the data set $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is a point $\hat{\boldsymbol{\mu}}$ which minimizes

$$\sum_{i_1 < \ldots < i_p} c[\mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_p}; \boldsymbol{\mu}]$$

where the sum is taken over all subsets of integers of the form $1 \leq i_1 < \ldots . < i_p \leq n$.

It can be checked that the Oja median does not have the uniqueness properties of the L_1 median but has the advantage of affine equivariance.

The Oja medium has a geometrical interpretation via gradient vectors that is analogous to that of the L_1 median. For each $(p-1)$-dimensional simplex with vertices $\mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_p}$ chosen as in Definition 3.1, construct a vector $A_{i_1 \ldots i_p}(\boldsymbol{\mu})$ at the origin whose length is proportional to the $(p-1)$-volume of the simplex, and pointing in the same direction as the ray from $\boldsymbol{\mu}$ which passes perpendicularly through the $(p-1)$-flat generated by the simplex. If the vector sum of all these

$$\binom{n}{p}$$

vectors is zero, then $\boldsymbol{\mu}$ is an Oja median for the data. In turn, the gradient vector

$$\binom{n}{p}^{-1} \sum_{i_1 < \ldots < i_p} A_{i_1 \ldots i_p}(\boldsymbol{\mu})$$

is a multidimensional generalization of the quantile appropriate for this median. See Brown & Hettmansperger (1987, 1989) for the development of this idea and in particular for the construction of affine invariant analogs of rank tests and R-estimates in one- and two-sample multivariate settings.

The asymptotic properties of Oja's simplex median have been developed in Oja & Niinimaa (1985). In particular they show the multivariate normality of $n^{\frac{1}{2}}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu})$. In the case where the sample is drawn from a multivariate normal distribution they show that the asymptotic efficiency of this median is uniformly dominates that of the L_1 median. If the multivariate normal distribution is spherically symmetric the asymptotic efficiencies of the two medians are equal. For other cases of multivariate normality, the asymptotic efficiency of Oja's median strictly dominates that of the L_1 median. Despite the promise of this result, there are some disturbing questions as to whether Oja's simplex median is naturally robust. While the influence function is bounded (compare Brown & Hettmansperger (1989)), Oja, Niinimaa & Tableman (1990) have found that this median has 0% breakdown. This is in sharp contrast with the 50% breakdown of the L_1 median.

4 Hotelling's Beach and the Halfspace Median

Hotelling (1929) introduced a different interpretation of the univariate median which generalizes to what we shall call the halfspace median due to Tukey (1975) and Donoho (1982) in higher dimensions. While this halfspace median is based upon the depth function of Tukey (1975), earlier game theoretic arguments by Hotelling provide an introduction to this median, and illustrate how it can be generalized to other settings, indeed, any setting in which a measure of distance between points is defined.

We begin be considering Hotelling's original problem, which, like the L_1 median, is motivated by considerations of locating a site of maximum financial benefit. Consider a beach on which two competitive ice cream vendors wish to locate stands so as to sell as many ice creams as possible. We idealize the beach as a straight line along which customers are to be distributed according to some probability distribution F. We make the assumptions that the customers are all equally interested in ice creams and that in choosing a stand to buy from they will naturally choose the one which is closest. If the two stands are equally close, then one will be selected by the customer at random. Finally we shall suppose that the locations of the two stands must not coincide. How should the two vendors place their stands if vendor I is to locate a stand first? A minimax strategy suggests itself. Suppose vendor I places a stand at x and vendor II places a stand at y. If $x < y$ then vendor I will obtain the customers in $(-\infty, \frac{1}{2}(x+y))$ and vendor II will obtain the customers in $(\frac{1}{2}(x+y), \infty)$ and they share the customer(s) at $\frac{1}{2}(x+y)$. Thus vendor I has a profit proportional to

$$\frac{1}{2}\left[F\left(\frac{x+y}{2}\right)+F\left(\frac{x+y^-}{2}\right)\right].$$

A similar formula holds when $x > y$. The minimum payoff to vendor I is proportional to

$$\min\,[F(x),\, 1-F(x^-)]$$

which is maximized when x is the median of the distribution F. The symmetrized quantile $\min\,[F(x), 1-F(x^-)]$ is a measure of the depth of the point x within the distribution F. If we replace F by the empirical distribution on a sample $X_1, \ldots, X_n$ then the minimax argument above leads us to the median of the empirical distribution, or sample median.

Published in the econometric literature, this game theoretic problem was introduced to model the strategies of competitive companies. This game theoretic interpretation of the median has not attracted much attention in probability and statistics until recently. Steele & Zidek (1980) showed the relationship between the ε-saddlepoint strategy of the second player and the use of shrinkage estimators of location. The importance of symmetry assumptions in such game theoretic arguments was pointed out by Pittenger (1980) who generalized the results of Steele & Zidek (1980) to spherically symmetric distributions in higher dimensions.

The virtue of this argument is that its extension into higher dimensions is straightforward, as, indeed, is its extension into contexts of directional statistics. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be a sample in $\mathbf{R}^p$. Let $\mathscr{H}$ be the class of all closed half spaces in $\mathbf{R}^p$. We generalize the empirical distribution function to $\mathscr{H}$ by setting

$$\hat{F}(H) = \left\{\sum_{i=1}^{n} 1[\mathbf{X}_i \in H]\right\}\Big/ n$$

for each $H \in \mathscr{H}$. Following Tukey, we define the *depth* $D(\boldsymbol{\mu})$ of a point $\boldsymbol{\mu} \in \mathbf{R}^p$ within the data set to be the infimum of $\hat{F}(H)$, where the infimum is taken over all closed half spaces H for which $\boldsymbol{\mu} \in H$.

Definition 4.1. The *halfspace median* of a data set is defined to be the set of points $\boldsymbol{\mu}$ of maximal depth.

The breakdown properties of the halfspace median have been extensively studied by Donoho (1982), Donoho & Huber (1983) and Donoho & Gasko (1987a). In particular the breakdown point of the halfspace median is at least $1/(p+1)$ and as high as $\frac{1}{3}$ in the limit for large samples from a centrally symmetric distribution. As this median is affine equivariant, this compares well with the results for Oja's simplex median. Donoho & Gasko (1987b) have also demonstrated that under assumptions of symmetry, the halfspace median is asymptotically normal. There is an obvious distributional analog to this sample median, which is constructed by replacing the empirical distribution function $\hat{F}$ by a general multivariate distribution F.

As with Oja's simplex median, the halfspace median is generally not a unique point. An argument is given in Small (1987) to show that the set of points $\boldsymbol{\mu}$ of maximal depth is a closed, bounded and convex set.

There appears to be little hope of obtaining general conditions under which the sample halfspace median is unique. Note that if F is centrally symmetric about $\boldsymbol{\mu}$ then the depth is uniquely maximized at the value $D(\boldsymbol{\mu}) = \frac{1}{2}$, and therefore the median is uniquely located at $\boldsymbol{\mu}$. For the general class of absolutely continuous distributions on $\mathbf{R}^p$ the condition of central symmetry can be replaced by the conditions of Theorem 3.3 of Small (1987) to obtain uniqueness. These conditions, while quite general, are too technical to discuss in detail here. However, it can be shown as a consequence of Proposition 4 of the same paper, that strict positivity of the density function is sufficient to prove uniqueness of the halfspace median for $\mathbf{R}^2$.

Another virtue of applying the halfspace median and Tukey's depth statistic to dimension 2 is *the existence of an exact test for angular symmetry*. This is a useful bonus. As was mentioned for the L_1 median, an exact test based on the corresponding quantile function for that median was available for testing spherical symmetry, but not for testing for the more general property of angular symmetry. In fact, an exact test is available for testing the hypothesis that $D(\boldsymbol{\mu}_0) = \frac{1}{2}$. See Hodges (1955) and Ajne (1968). Let $\hat{D}(\boldsymbol{\mu}_0)$ be the empirical depth at $\boldsymbol{\mu}_0$ based upon a planar sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$. The standardization of the data set about $\boldsymbol{\mu}_0$ by replacing each $\mathbf{X}_i$ by $\boldsymbol{\mu}_0 + (\mathbf{X}_i - \boldsymbol{\mu}_0)/\|\mathbf{X}_i - \boldsymbol{\mu}_0\|$ leaves the empirical depth function invariant. The depth function for the standardized data is then immediately seen to be Ajne's test for central symmetry for directional data. The calculation of the distribution of this test statistic was done by Johansen under the assumption of directional uniformity. See Ajne (1968). However, the argument used only the assumption of central symmetry of the directional distribution. Corresponding to an exact test for angular symmetry at a point $\boldsymbol{\mu}_0 \in \mathbf{R}^2$ we can construct exact confidence regions. A $100(1-\alpha)\%$ confidence region for the point $\boldsymbol{\mu}$ of angular symmetry is

$$\{\mathbf{x} \in \mathbf{R}^2 : \hat{D}(\mathbf{x}) > d_\alpha\}$$

where $P[\hat{D}(\boldsymbol{\mu}) > d_\alpha] = 1 - \alpha$. The basic principles of this argument above were first worked out by Hodges (1955). Recent work by Oja & Nyblom (1989) extends this theory to a family of bivariate sign tests. Their paper develops an asymptotic theory of such tests and compares the asymptotic relative efficiencies.

We close this section by noting that the halfspace median is equivariant under a much richer class of transformations than the affine ones. This is developed extensively in Small (1987). The key idea is to relax the affine class to the class of transformations $T: \mathbf{R}^p \to \mathbf{R}^p$ which map closed halfspaces containing the median $\boldsymbol{\mu}$ into convex sets and for which the preimage set of any halfspace which does not contain the median $\boldsymbol{\mu}$ is also a convex set. See Theorem 4·3 of Small (1987).

5 The Simplicial Depth Median

An affine equivariant median and an affine equivariant ranking can be obtained from the simplicial depth function of Liu (1988, 1990). To motivate it, we go back to the usual sample median in one dimension, and observe that it can be characterized by its lying in the greatest number of intervals constructed from the data points. In this sense, it can be viewed as being deep inside the data cloud. To generalize to higher dimensions, it suffices to replace intervals by p dimensional simplexes in $\mathbf{R}^p$. The empirical simplicial depth function is defined to be

$$\mathrm{SDF}(\boldsymbol{\mu}) = \binom{n}{p+1}^{-1} \sum_{1 \leqslant i_1 < \ldots < i_{p+1} \leqslant n} 1[\boldsymbol{\mu} \in S(\mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_{p+1}})]$$

where $S(\mathbf{x}_1, \ldots, \mathbf{x}_{p+1})$ is the simplex with vertices $\mathbf{x}_1, \ldots, \mathbf{x}_{p+1}$.

Definition 5.1. A *simplicial depth median* is a point $\hat{\boldsymbol{\mu}}$ which maximizes the function $\mathrm{SDF}(\boldsymbol{\mu})$.

The distribution version of this is just

$$\text{SDF}(\boldsymbol{\mu}) = P_F[\boldsymbol{\mu} \in S(\mathbf{X}_1, \ldots, \mathbf{X}_{p+1})].$$

It is evident that these depth functions are affine equivariant.

The results of Cover & Efron (1967) can be used to determine the value of the simplicial depth function at a point of angular symmetry. Suppose $\boldsymbol{\mu}$ is a point of angular symmetry of a distribution F on $\mathbf{R}^p$ and suppose $\mathbf{X}_1, \ldots, \mathbf{X}_{p+1}$ is a sample from F. Then

$$P_F[\boldsymbol{\mu} \in S(\mathbf{X}_1, \ldots, \mathbf{X}_{p+1})] = 2^{-p}.$$

Thus the expected value of $\text{SDF}(\boldsymbol{\mu})$ is also 2^{-p}. Conditions for the consistency of the simplicial depth median are given in Liu (1988). For an absolutely continuous angularly symmetric distribution on $\mathbf{R}^p$ the simplicial depth function decreases monotonically from 2^{-p} in all directions away from $\boldsymbol{\mu}$.

6 Convex Hull Stripping and Related Methods

The idea of finding the center of a data set by successively stripping away outlying values is a well established idea which has become quite computationally attractive with modern computing power. Green (1981) has surveyed the class of stripping methods for spatial data sets. So we shall only describe them here in terms of their relationship to the class of spatial medians that we study. In dimension one, the median of a data set can be regarded as the innermost order statistic, and can be found as follows. If $X_{(1)}$, $X_{(2)}, \ldots, X_{(n)}$ are the order statistics of a sample of size n then the median is what remains if the lowest and highest order statistics are successively deleted ('stripped' or 'peeled') from the data set. If n is odd, then eventually a single order statistic will remain, namely the median. To generalize this concept to higher dimensions, we note that the interval $[X_{(1)}, X_{(n)}]$ defined by the lowest and highest order statistics is the convex hull of the data set, for which these two order statistics are the vertices. Such a definition is immediately applicable in higher dimensions. From a set of points $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ we delete those points which are vertices of the convex hull of the n points. This procedure is repeated until a set of points remain which are all vertices of their convex hull. It may well be that this set consists of a single point which is then regarded as the convex hull strip median. If this set is the convex hull of more than one point, then the centroid of the convex set may conveniently be taken as the convex hull strip median. This is analogous to averaging the two middle order statistics for a univariate sample of an even number of observations. A convex hull strip of 10 points is shown in Fig. 2.

Considerable discussion of convex hull stripping as a way of partially ordering a multivariate data set can be found in Barnett (1976). In the discussion that follows that

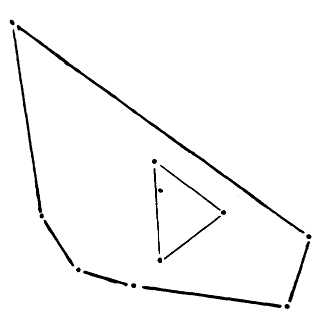

Figure 2. *Convex hull strip of* 10 *points.*

paper, Seheult, Diggle & Evans (1976) proposed the centroid of the innermost convex hull as a multivariate median. While this median resembles the halfspace median based upon Tukey's depth, they are quite distinct and indeed on small data sets, the convex hull strip median can be easily constructed by the most basic of techniques. The median has the additional desirable property of affine equivariance which follows from the affine equivariance of the convex hull construction itself.

In view of the computational convenience of the convex hull strip median, one naturally asks whether its sampling properties can be obtained. Here the situation is not as promising. All the multidimensional medians that we have considered up to now have both sample and distribution versions. The sample definition is usually obtained from the distribution version by replacing the distribution by the empirical distribution. However, this is not straightforward for the convex hull strip median as there is no obvious analog of the ith convex hull peel. Furthermore, the distribution theory for convex hull of samples is quite incomplete. Asymptotic results for simple distributions (e.g. uniform on a convex set) are known. See Renyi & Sulanke (1963, 1964). However, in general only integral representations are available even for the simplest properties of the convex hulls such as the expected number of vertices. This problem is compounded with successive strips or peels of the outermost convex hulls of the data. The influence curves and breakdown properties of the median have also proved difficult to obtain. Donoho & Gasko (1987a) have found that the breakdown point for a sample of size n in $\mathbf{R}^p$ is bounded above by $(n+p+1)/[(n+2)(p+1)]$. However, this result is optimistic: Donoho (1982) reports that the breakdown of a Gaussian sample of size n goes to zero as $n \to \infty$. Results are complicated by the fact that the convex hull strip median is not a continuous function of the data represented in $\mathbf{R}^{p \times n}$.

Related peeling methods are available. The convex hull of a set of points can be generalized to the set within a given class of smallest volume that contains the sample. Thus convex hull peeling can be replaced by minimal volume ellipsoid peeling, which peels 3, 4 or 5 points with probability one from absolutely continuous distributions in dimension two. Thus the peeling is finer than that using convex hulls. Daniels (1952) used minimal covering circle methods to investigate dispersion and obtained some distribution theory.

Green (1981) has suggested a measure of depth that leads to a multidimensional median. Like the measure of depth based upon stripping, it has no obvious distribution analog. If a graph is constructed from the data set by joining adjacent points in the Delaunay triangulation then the Delaunay depth of a sample point is the number of edges of the shortest path joining the point to a vertex of the convex hull of the data set. (A subset of $p+1$ points of a sample in $\mathbf{R}^p$ is said to form a Delaunay simplex if there exists a sphere of dimension $p-1$ passing through the $p+1$ points and containing no other point of the sample in its interior. We understand by a 0-dimensional sphere a pair of distinct points so that a Delaunay interval in $\mathbf{R}$ is simply an interval formed by adjacent order statistics. The corresponding graph on the data set is constructed by joining any two points which belong to a common Delaunay simplex.) The depth function that this produces and its associated median have not been extensively studied.

7 Medians for Directional Data

In this section we shall discuss the research directions of extending medians to directional data. By directional data we refer to situations where the random variables and distributions are defined on spheres rather than Euclidean spaces. For example, data on the location of epicenters of earthquakes are genuine spherical data because on this scale the Earth is roughly a sphere. In other problems, such as observations of wind directions only the direction is relevant. The natural state space for data on directions is a circle or sphere. In general we consider data and distributions on a sphere $\mathbf{S}^p \subset \mathbf{R}^{p+1}$ defined as $\{\mathbf{x}: \|\mathbf{x}\| = 1\}$.

Because such data lie naturally on a sphere it is usually inappropriate to construct directional medians by transforming $\mathbf{S}^p$ to $\mathbf{R}^p$ via a stereographic projection. Instead we prefer to generalize the basic ideas in use in Euclidean space to the directional setting. For example, the analog of a straight line and Euclidean distance are an arc of a great circle and arc length respectively on the sphere. The analog of a halfspace in $\mathbf{R}^p$ becomes a hemisphere on $\mathbf{S}^p$. Thus the analog of simplex on $\mathbf{S}^p$ is a region formed by $p + 1$ points that is bounded by the

$$\binom{p+1}{2}$$

arcs of great circles joining all pairs. This is the intersection of all closed hemispheres containing the $p + 1$ points. There is even an analog of the group of linear transformations on the sphere. Define $U(\mathbf{x}) = [T(\mathbf{x})]/\|T(\mathbf{x})\|$, where $T: \mathbf{R}^{p+1} \to \mathbf{R}^{p+1}$ is a linear transformation. With these directional analogs, a number of multidimensional medians can be adapted from Euclidean settings to spheres.

There is some question as to whether directional medians are statistically useful on a sphere. Because spheres are compact it can be argued that all continuous estimators are naturally robust. On the other hand it is easy to construct samples in which there are a few discordant values which would ideally be discounted in an estimation problem. We will not pursue the issue further but refer the reader to Ko & Guttorp (1988).

Mardia (1972) and Fisher (1985) have extended the L_1 Euclidean median to the unit circle and $\mathbf{S}^p$ respectively. For any points $\mathbf{x}, \mathbf{y} \in \mathbf{S}^p$ let $\langle \mathbf{x}, \mathbf{y} \rangle$ be the inner product of vectors $\mathbf{x}, \mathbf{y}$. Then the length of the geodesic arc joining $\mathbf{x}$ and $\mathbf{y}$ is

$$\rho(\mathbf{x}, \mathbf{y}) = \cos^{-1}(\langle \mathbf{x}, \mathbf{y} \rangle).$$

The total dispersion of a sample $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ on $\mathbf{S}^p$ about a point $\boldsymbol{\mu} \in \mathbf{S}^p$ is

$$D(\boldsymbol{\mu}) = n^{-1} \sum_{i=1}^{n} \rho(\mathbf{X}_i, \boldsymbol{\mu})$$

and the point $\hat{\boldsymbol{\mu}}$ which minimizes this is the Mardia–Fisher spherical median. The gradient vector field associated with the minimization of this function is the vector field of quantiles on $\mathbf{S}^p$ and a cross section of the tangent bundle. Fisher (1985) has established the asymptotic normality of this median. The asymptotic efficiency of this median relative to the directional mean was found to dominate the asymptotic efficiency of the L_1 median relative to the centroid for the corresponding Euclidean case. Note that on the unit circle $\mathbf{S}^1$ the diameter passing through the median and its antipodal point divides the data set into two semicircles with half of the data in each. The analogy with the univariate Euclidean median is clear. The asymptotic efficiency of the Mardia–Fisher median for circular data has also been studied by Ducharme & Milasevic (1987b).

A related but distinct directional median also based upon the L_1 median has been proposed by Ducharme & Milasevic (1987a). A data set $\mathbf{X}_1, \ldots, \mathbf{X}_n$ on $\mathbf{S}^p$ can be regarded as lying in $\mathbf{R}^{p+1}$ with L_1 median $\hat{\boldsymbol{\mu}}$. The normalized vector $\tilde{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}/\|\hat{\boldsymbol{\mu}}\|$ lies again on the sphere and is another generalization of the L_1 median called the normalized L_1 median. Some asymptotics are developed for this median in Ducharme & Milasevic (1987a).

The basic idea of generalizing Oja's simplex median to the directional setting would seem to be sound. So far this avenue does not seem to have been pursued. As mentioned above, the simplex generated by $p+1$ points is simply the intersection of all closed hemispheres containing these points. The p-dimensional volume generalizes to the content of the spherical complex on $\mathbf{S}^p$. An avenue which has been considered is the generalization of the halfspace median to the hemisphere median. See Small (1987). We recall that the saddlepoint strategy that gives rise to the halfspace median can immediately be generalized to any metric space. The directional analog is constructed by replacing the halfspaces of the Euclidean setting with the class of closed hemispheres. Thus the hemisphere depth of a point $\boldsymbol{\mu} \in \mathbf{S}^p$ for a distribution F is

$$\sup\{\mathrm{F(H)} : \boldsymbol{\mu} \in H\}$$

where H is a hemisphere of $\mathbf{S}^p$. Sample versions with $\hat{F}$ are available. A hemisphere median is then a point of maximal depth. Small (1987) also demonstrated that this median is equivariant under a rich class of transformations that includes the analogs of linear transformations mentioned above. Liu's simplicial depth median also generalizes. The simplicial depth of a point $\boldsymbol{\mu} \in \mathbf{S}^p$ is the proportion of simplexes built out of the

$$\binom{n}{p+1}$$

subsets of $p+1$ sample points which contain $\boldsymbol{\mu}$. The spherical simplicial depth function and its associated median are discussed by Liu & Singh (1991) who provide a different motivation for the hemisphere median.

8 Conclusions

A comparison of the various medians presented in this paper shows that no one definition is exclusively the natural generalization of the univariate median into higher dimensions. The L_1-median has a variety of good properties but is naturally applicable to spherically symmetric models. Among affine equivariant medians, the halfspace median has good breakdown properties but is computationally cumbersome. A variety of techniques based upon peeling methods have no obvious distribution analogs but are more straightforward to compute. In this paper I have tried to highlight the central issues about multidimensional medians in order that these strengths and weaknesses can be assessed. Much, especially among affine equivariant medians, remains to be done to compare the properties of the medians considered.

Acknowledgements

This paper has been written with the advice and suggestions of many people. I would like to thank Bruce Brown, David Donoho, Soren Johansen, Mara Tableman and David Tyler for their help. Special thanks go to Regina Liu for help with the references and some of the topics of this survey. This paper was written with the support of an NSERC Operating Grant.

References

Ajne, B. (1968). A simple test for uniformity of a circular distribution. *Biometrika* **55,** 343–354.
Barnett, V. (1976). The ordering of multivariate data (with discussion). *J.R. Statist. Soc.* A **139,** 318–354.
Bedall, F.K. & Zimmermann, H. (1979). The mediancentre. *Appl. Statist.* **23,** 325–328.
Brown, B.M. (1983). Statistical uses of the spatial median. *J.R. Statist. Soc.* B **45,** 25–30.
Brown, B.M. & Hettmansperger, T.P. (1987). Affine invariant rank methods in the bivariate location model. *J.R. Statist. Soc.* B **49,** 301–310.

Brown, B.M. & Hettmansperger, T.P. (1989). An affine invariant bivariate version of the sign test. *J.R. Statist. Soc.* B **51,** 117–126.
Cover, T. & Efron, B. (1967). Geometrical probability and random points on a hypersphere. *Ann. Math. Statist.* **38,** 213–220.
Daniels, H.E. (1952). The covering circle of a sample from a circular normal distribution. *Biometrika* **62,** 531–546.
Donoho, D.L. (1982). Breakdown properties of multivariate location estimators. Ph.D. Qualifying Paper, Department of Statistics, Harvard University.
Donoho, D.L. & Gasko, M. (1987a). Multivariate generalizations of the median and trimmed mean, I. Technical Report #133, Department of Statistics, University of California at Berkeley.
Donoho, D.L. & Gasko, M. (1987b). Multivariate generalizations of the median and trimmed mean, II. Technical Report #128 Department of Statistics, University of California at Berkeley.
Donoho, D.L. & Huber, P.J. (1983). The notion of a breakdown point. In *A Festschrift for Erich L. Lehmann,* Ed. P. Bickel, K. Doksum and J.L. Hodges, Jr., pp. 157–184. Belmont, Calif.: Wadsworth.
Ducharme, G.R. & Milasevic, P. (1987a). Spatial median and directional data. *Biometrika* **74,** 212–215.
Ducharme, G.R. & Milasevic, P. (1987b). Some asymptotic properties of the circular median. *Comm. Statist.* A **16,** 659–664.
Eddy, W.F. (1982). Convex hull peeling. *COMPSTAT 1982* for IASC, pp. 42–47. Vienna: Physica-Verlag.
Eddy, W.F. (1983). Set-valued orderings for bivariate data. In *Stochastic Geometry, Geometric Statistics, and Stereology,* Ed. R.V. Ambartsumian and W. Weil, pp. 79–90. Leipzig: Tuebner
Eddy, W.F. (1985). Ordering of multivariate data. In *Computer Science and Statistics: The Interface.* Ed. L. Billard, pp. 25–30.
Eells, W.C. (1930). A mistaken conception of the center of population. *J. Am. Statist. Assoc.* **25,** 33–40.
Fisher, N.I. (1985). Spherical medians *J.R. Statist. Soc.* B **47,** 342–348.
Fisher, N.I., Lewis, T. & Embleton, B.J.J. (1987). *Statistical Analysis of Spherical Data.* Cambridge University Press.
Galambos, J. (1975). Order statistics of samples from multivariate distributions. *J. Am. Statist. Assoc.* **70,** 674–680.
Galvani, L. (1933). Sulla determinazione del centro di gravita e del centro mediano di una popolazione, con applicazioni alla popolazione italiana censita il 1 dicembre 1921. *Metron* **11,** 17–48.
Gini, C. & Galvani L. (1929). Di talune estensioni dei concetti di media ai caratteri qualitativi. *Metron* **8,** 3–209.
Gower, J.C. (1974). The mediancentre. *Appl. Statist.* **23,** 466–470.
Green, P.J. (1976). Discussion of paper by V. Barnett. *J.R. Statist. Soc.* A **139,** 350.
Green, P.J. (1981). Peeling bivariate data. In *Interpreting Multivariate Data,* Ed. V. Barnett, pp. 3–20. New York: Wiley.
Griffin, F.L. (1933). The center of population for various continuous distributions of population over areas of various shapes. *Metron* **11,** 11–16.
Haldane, J.B.S. (1948). Note on the median of a multivariate distribution. *Biometrika* **35,** 414–415.
Hampel, F.R. (1971). A general qualitative definition of robustness. *Ann. Math. Statist.* **42,** 1887–1896.
Hayford, J.F. (1902). What is the center of an area, or the center of a population? *J. Am. Statist. Assoc.* **8,** 47–58.
Hodges, J.L. (1955). A bivariate sign test. *Ann. Math. Statist.* **26,** 523–527.
Hodges, J.L., Jr. (1967). Efficiency in normal samples and tolerance of extreme values for some estimates of location. *Proc. 5th Berkeley Symp.* **1,** 163–178.
Hotelling, H. (1929). Stability in competition. *Econom. J.* **39,** 41–57.
Kimber, A.C. (1985). A note on the detection and accommodation of outliers relative to Fisher's distribution on the sphere. *Appl. Statist.* **34,** 169–172.
Ko, D. & Guttorp, P. (1988). Robustness of estimators for directional data. *Ann. Statist.* **16,** 609–618.
Lenth, R.V. (1981). Robust measures of location for directional data. *Technometrics* **23,** 77–81.
Liu, R.Y. (1988). On a notion of simplicial depth. *Proc. Nat. Acad. Sci. USA.* **85,** 1732–1734.
Liu, R.Y. (1990). On a notion of data depth based upon random simplices. *Ann. Statist.* **18,** 405–414.
Liu, R.Y. & Singh, K. (1991). Ordering directional data: concepts of data depth on circles and spheres. *Ann. Statist.* To appear.
Lopuhaa, H.P. & Rousseeuw, P.J. (1987). Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. Technical Report 87–14, Faculty of Mathematics and Informatics, Delft University of Technology.
Mardia, K.V. (1972). *The Statistics of Directional Data.* New York: Academic Press.
Mardia, K.V. (1975). Statistics of directional data (with discussion). *J.R. Statist. Soc.* B **37,** 349–393.
Milasevic, P. & Ducharme, G.R. (1987). Uniqueness of the spatial median. *Ann. Statist.* **15,** 1332–1333.
Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statist. Prob. Letters.* **1,** 327–332.
Oja, H. & Niinimaa, A. (1985). Asymptotic properties of the generalized median in the case of multivariate normality. *J.R. Statist. Soc.* B **47,** 372–377.
Oja, H., Niinimaa, A. & Tableman, M. (1990). The finite sample breakdown point of the Oja bivariate median and of the corresponding half samples version. Unpublished report.
Oja, H. & Nyblom, J. (1989). On bivariate sign tests. *J. Am. Statist. Assoc.* **84,** 249–259.
Pittenger, A.O. (1980). Success probabilities for second guessers. *J. Appl. Prob.* **17,** 1133–1137.

Renyi, A. & Sulanke, R. (1963). Über die konvexe Hulle von *n* züfallig gewahlten Punkten I. *Z. Wahr. verw. Geb.* **2,** 75–84.
Renyi, A. & Sulanke R. (1964). Über die konvexe Hulle von *n* züfallig gewahlten Punkten II. *Z. Wahr. verw. Geb.* **3,** 138–147.
Rousseeuw, P.J. (1984). Least median of squares regression. *J. Am. Statist. Assoc.* **79,** 871–880.
Rousseeuw, P.J. (1986). Multivariate estimation with high breakdown point. In *Mathematical Statistics and Applications,* Ed. W. Grossmann, G. Pflug, I. Vincze and W. Wertz, pp. 283–297. Dordrecht: Reidel.
Rousseeuw, P.J. & Leroy, A.M. (1987). *Robust Regression and Outlier Detection.* New York: Wiley.
Scates, D.E. (1933). Locating the median of the population in the United States. *Metron* **11,** 49–66.
Seheult, A.H., Diggle, P.J. & Evans, D.A. (1976). Discussion of paper by V. Barnett. *J.R. Statist. Soc.* A **139,** 351–352.
Sibson, R. (1984). Multivariate analysis. *J.R. Statist. Soc.* A **147,** 198–207.
Small, C.G. (1987). Measures of centrality of multivariate and directional distributions. *Can. J. Statist.* **15,** 31–39.
Stahel, W.A. (1981). Robuste Schatzungen: Infinitesimale Optimalität and Schatzungen von Kovarianzmatrizen, Ph.D. Thesis, ETH Zurich, Switzerland.
Steele, J.M. & Zidek, J. (1980). Optimal strategies for second guessers. *J. Am. Statist. Assoc.* **75,** 596–601.
Tukey, J.W. (1975). Mathematics and the picturing of data In *Proc. International Congress of Mathematicians,* Vancouver, 1974, **2,** pp. 523–31. ISBN 0-919558-04-6.
Watson, G.S. (1983). *Statistics on Spheres.* New York: Wiley.
Watson, G.S. (1986). Some estimation theory on the sphere. *Ann. Inst. Statist. Math.* **38,** 263–275.
Weber, A. (1909). Über den Standort der Industrien, Tubingen. English translation by Freidrich, C.J. (1929), *Alfred Weber's Theory of Location of Industries,* University of Chicago Press.
Wehrly, T.E. & Shine, E.P. (1981). Influence curves of estimators for directional data. *Biometrika* **68,** 334–335.

Résumé

L'auteur revoit une série d'articles qui généralisent de plusieurs manières le concept de médiane dans les espaces multidimensionnels. Ces différentes mesures reproduisent la definition usuelle de la médiane dans le cas de loi unidimensionnelle. Il examine en outre les concepts d'équivariance, de symmétrie, de rupture ainsi que la difficulté de calcul de la médiane pour chaque definition. Enfin, il généralise ces idées pour les lois directionnelles.

[*Received April* 1989, *accepted June* 1990]