On the Performance of Some Robust Nonparametric Location Measures Relative to a General Notion of Multivariate Symmetry

Yijun Zuo[1]
Arizona State University

Robert Serfling[2]
University of Texas at Dallas

November, 1998 (revised May, 1999)

[1]Department of Mathematics, Arizona State University, Tempe, Arizona 85287-1804, USA.
[2]Programs in Mathematical Sciences, University of Texas at Dallas, Richardson, Texas 75083-0688, USA.

Supported by NSF Grant DMS-9705209.

Abstract

Several robust nonparametric location estimators are examined with respect to several criteria, with emphasis on the criterion that they should agree with the point of symmetry in the case of a symmetric distribution. For this purpose, a broad version of multidimensional symmetry is introduced, namely "halfspace symmetry", generalizing the well-known notions of "central" and "angular" symmetry. Characterizations of these symmetry notions are established, permitting their properties and interrelations to be illuminated. The particular location measures considered consist of several nonparametric notions of multidimensional median: the "L^2" (or "spatial"), "Tukey/Donoho halfspace", "projection", and "Liu simplicial" medians, all of which are robust in the sense of nonzero breakdown point. It is established that the first three of these in general do identify the point of symmetry when it exists, whereas the latter, however, fails to do so in some circumstances. Combining this finding with consideration of other criteria such as affine equivariance, stochastic order preserving, and degree of robustness, we conclude that among these choices, the "halfspace" and "projection" medians, both of which are based on projection pursuit methodology, are the most attractive overall.

AMS 1991 Subject Classification: Primary 62H05 Secondary 62G05.

Key words and phrases: multivariate symmetry; nonparametric location measure; multidimensional median.

1 Introduction

For the problem of nonparametric multivariate location estimation, a variety of interesting notions of multidimensional median have been introduced (see Small (1990) for a partial review). For *symmetric* distributions, it is especially desirable that a location measure agree with the point of symmetry (at least within linear transformation), and in Section 3 we shall formally introduce this "center-locating" criterion.

In this regard, it is preferable to employ as broad a notion of symmetry as possible. Here we consider two well-known versions of multidimensional symmetry, *central* and *angular*, and introduce a new more general notion, *halfspace symmetry.*

Definitions of these notions of symmetry are provided in Section 2, along with characterizations, properties, and interrelations. Using these results, we argue (see Remarks 2.3) that the notion of halfspace symmetry provides

- a natural basis for requiring a "center-locating" property of any reasonable notion of multivariate location measure of "median" type, and
- an important widening of the class of actual distributions which might be usefully approximated by symmetric ones.

In Section 3 several particular robust and nonparametric location measures are studied, namely the following four notions of multidimensional median: the "L^2" (or "spatial"), "Tukey/Donoho halfspace", "projection", and "Liu simplicial" medians, all of which are robust in the sense of nonzero breakdown point. (The terminology "L^2 median" is explained in §3.2.1.) These four medians are examined with respect to the above-mentioned *center-locating* criterion as well as the additional criteria of *affine equivariance*, *stochastic order preserving*, and *degree of robustness.* Regarding the center-locating criterion, it is found (Theorems 3.3 and 3.4) to be satisfied in great generality by all but the simplicial median, the latter failing to meet it in some circumstances.

Another finding (Section 3.3.2) is that two angularly symmetric distributions in $\mathbf{R}^d$ can be stochastically ordered with their points of angular symmetry ordered in the opposite direction. This argues not against the notion of angular symmetry, but rather against requiring "stochastic order preserving" as a property of nonparametric location measures, when notions of symmetry broader than central symmetry are considered.

In Section 4 we review the results of Section 3 and conclude that among these four choices of multivariate median, the halfspace and projection medians are the more attractive overall. We note that these are based on projection pursuit methodology. We also comment briefly on some other multivariate location estimators.

2 Nonparametric Notions of Multivariate Symmetry

Here we define several notions of symmetry apropos to nonparametric multivariate inference and develop their characterizations, properties, and interrelations.

2.1 Definitions and First Properties

For any unit vector $u \in \mathbf{R}^d$ and $t \in \mathbf{R}$, the set of points $H_{u,t} = \{x : u'x \leq t\}$ defines a *closed halfspace* in $\mathbf{R}^d$. Its interior is an *open* halfspace, and the boundary $\{x : u'x = t\}$ defines a *hyperplane*. (All closed and open halfspaces in $\mathbf{R}^d$ may be represented in this fashion.) More generally, a *halfspace* consists of an open halfspace along with a (Borel) subset of its boundary.

We use the notation C^c, C° and $\overline{C}$, respectively, for the *complement, interior*, and *closure* of a set C.

For any random variable $Y \in \mathbf{R}$, we define a *median* of its distribution to be any number c satisfying

$$P(Y \leq c) \geq .5, \quad P(Y \geq c) \geq .5,$$

and we use the notation $\mathrm{Med}(Y)$. (Notions of median for the multidimensional case are introduced in Section 3.)

2.1.1 Central symmetry

A standard notion prevalent in the literature on nonparametric multivariate inference is that a random vector X in $\mathbf{R}^d$ has a distribution *centrally symmetric* about $\theta \in \mathbf{R}^d$ if

$$X - \theta \stackrel{\mathrm{d}}{=} \theta - X, \tag{1}$$

where "$\stackrel{\mathrm{d}}{=}$" denotes "equal in distribution". The following useful equivalent representations of central symmetry are easily established.

Lemma 2.1 *Equivalent to* (1) *are each of*

$$u'(X - \theta) \stackrel{\mathrm{d}}{=} u'(\theta - X), \quad \textit{each unit vector } u \textit{ in } \mathbf{R}^d, \tag{2}$$

and

$$P(X - \theta \in H) = P(X - \theta \in -H), \quad \textit{each closed halfspace } H \subset \mathbf{R}^d. \tag{3}$$

2.1.2 Angular symmetry

A broader notion was introduced by Liu (1988, 1990). A random vector X in $\mathbf{R}^d$ has a distribution *angularly symmetric* about $\theta \in \mathbf{R}^d$ if

$$\frac{(X - \theta)}{\| X - \theta \|} \stackrel{\mathrm{d}}{=} \frac{(\theta - X)}{\| X - \theta \|}, \tag{4}$$

or, equivalently, if $(X - \theta)/ \| X - \theta \|$ has centrally symmetric distribution, where $\| \cdot \|$ stands for Euclidean norm. (Here we adopt the convention that $(X - \theta)/ \| X - \theta \| = 0$ if $\| X - \theta \| = 0$.) It is immediate that, in an obvious terminology,

Lemma 2.2 *C-symmetry about a point θ implies A-symmetry about that point.*

Counterexamples to the converse are easily found.

The following result on *uniqueness* of the point of angular symmetry is established by Liu (1988, 1990). (It also follows from Theorem 2.1 below, which treats uniqueness of the point of symmetry in the case of the more general notion of halfspace symmetry to be introduced.)

Lemma 2.3 *The point (or center) θ of angular symmetry of a random vector $X \in \mathbf{R}^d$, if it exists, is unique unless the distribution of X is concentrated on a line and its probability distribution on that line has more than one median.*

Remarks 2.1 (i) It is easily seen (as noted in Liu (1990), or see Corollary 2.1 below) that if θ is a point of angular symmetry, then any hyperplane passing through θ divides $\mathbf{R}^d$ into two open halfspaces with equal probabilities, which equal 1/2 if the distribution of X is continuous.

The converse is also true: if every hyperplane through a point θ divides $\mathbf{R}^d$ into two open halfspaces with equal probabilities, then θ is a point of angular symmetry. The proof of the latter is not straightforward, however, but follows directly from Theorem 2.1 below.

(ii) It is easily seen (as noted in Randles (1989), or see Corollary 2.1 below) that if θ is a point of angular symmetry, then θ agrees with the median of the conditional distribution of X on any axis through θ.

But here the strict converse fails to hold. Counterexamples can arise when there is nonzero probability mass on the median, permitting complementary open halfspaces on either side of the median to have unequal probability, in which case (by (i) above) angular symmetry fails to hold. Thus, for the *discrete* case, the condition that $\text{Med}(u'X) = u'\theta$ for each unit vector $u \in \mathbf{R}^d$ does not necessarily imply that X is angularly symmetric about θ.

> *Counterexamples.* (1) Let $d = 1$, $P(X = 0) = 1/2$, $P(X = -1) = 1/3$, and $P(X = 1) = 1/6$. Then $\text{Med}(X) = 0$, but by Theorem 2.3 below it is easily checked that X is not angularly symmetric about 0.
>
> (2) Let $d = 2$, $P(X = (0,0)) = 2/5$, $P(X = (1, \pm 1)) = 1/5$, and $P(X = (-1, \pm 1)) = 1/10$. Then it is easy to check that $\text{Med}(u'X) = u'(0,0)$ for any $u \in \mathbf{R}^2$, whereas by Theorem 2.3 below X is not angularly symmetric about $(0,0)$.

For X having a *continuous* distribution, however, the converse does indeed follow by an immediate argument using Theorems 2.4 and 2.6 below. □

2.1.3 Halfspace symmetry

Here we define a new nonparametric notion of multivariate symmetry. A random vector X in $\mathbf{R}^d$ has a distribution *halfspace symmetric* about $\theta \in \mathbf{R}^d$ if

$$P(X \in H) \geq 1/2, \quad \text{each closed halfspace } H \text{ with } \theta \text{ on the boundary.} \tag{5}$$

Since every halfspace containing θ contains a closed halfspace with θ on its boundary, it is equivalent to say simply: "$P(X \in H) \geq 1/2$ for any halfspace H containing θ." Clearly, it also is equivalent to say that any hyperplane passing through θ must divide $\mathbf{R}^d$ into two closed halfspaces each of which has probability at least $1/2$. Thus, in view of Remarks 2.1(i), the following result is immediate.

Lemma 2.4 *A-symmetry about a point θ implies H-symmetry about that point.*

Remarks 2.2 (i) The converse to Lemma 2.4 does not hold. The counterexamples of Remarks 2.1(ii) may be exploited here also, as well as the following.

> *Counterexample.* Let $d = 2$, $P(X = (0.2, 0.1)) = 2/5$, $P(X = (1, \pm 1)) = 1/10$, and $P(X = (-1, \pm 1)) = 1/5$. Then it is easily checked that X is halfspace symmetric about the point $(0.2, 0.1)$ but not angularly symmetric about any point.

(ii) For $d = 1$, distributions are easily constructed which are not C- or A-symmetric about any point, but all distributions are H-symmetric about the median.

(iii) For $d > 1$, however, there are multivariate distributions which are not H-symmetric about any point.

> *Example.* Let $d = 2$, $P(X = (\pm 1, 0)) = 1/3$, and $P(X = (0, 1)) = 1/3$. Then it is easy to see that X is not halfspace symmetric about any point in $\mathbf{R}^2$.

□

Theorem 2.1 *The point (or center) θ of halfspace symmetry of a random vector $X \in \mathbf{R}^d$, if it exists, is unique unless the distribution of X is concentrated on a line and its probability distribution on that line has more than one median.*

PROOF. Suppose that θ_1 and θ_2 are two distinct points of symmetry of the halfspace symmetric distribution of a random vector X. Let H_{θ_1} be an arbitrary closed halfspace with θ_1 on the boundary and containing θ_2, and H_{θ_2} be the closed halfspace with θ_2 on the boundary and containing θ_1 such that its boundary is parallel to that of H_{θ_1}. Then, by the definition of halfspace symmetry, $P(X \in (H_{\theta_1} \cap H_{\theta_2})^\circ) = 0$. It readily follows that $P(X \in L_{\theta_1 \theta_2}) = 1$, where $L_{\theta_1 \theta_2}$ is the line in $\mathbf{R}^d$ passing through θ_1 and θ_2, in which case θ_1 and θ_2 are each a median of the distribution of X on $L_{\theta_1 \theta_2}$. This completes the proof. □

Remarks 2.3 *Relevance of halfspace symmetry.* (i) As shown in Theorem 2.4 below, halfspace symmetry is equivalent to the condition that $\mathrm{Med}(u'X) = u'\theta$ for each unit vector $u \in \mathbf{R}^d$. Such a condition appeals quite naturally as a requirement to be met by any reasonable notion of "multidimensional median". In view of this equivalence, halfspace symmetry provides precisely the broadest possible manifestation of the criterion that any reasonable notion of "multidimensional median" should agree with the point of symmetry in the case of a symmetric distribution. In particular, therefore, this "center-locating" criterion provides a way to compare various notions of multidimensional median that have arisen in the literature (see Section 3).

(ii) Also, although we often use as ideal models distributions which are symmetric, the actual underlying distributions being modeled are typically not symmetric but only can be approximated by "symmetric" ones. Thus we would like the class of nonsymmetric distributions which might be approximated usefully by symmetric ones to be as wide as possible. Halfspace symmetry, as a broader relaxation of the normality assumption on multivariate distributions than that provided by central symmetry or angular symmetry, indeed widens the class of distributions subject to such approximation.

Although, however (as seen in Theorem 2.6 below), halfspace symmetry actually reduces to angular symmetry for all distributions except certain discrete ones with positive probability on the center of halfspace symmetry, true underlying distributions for the phenomena we observe in practice are invariably discrete and, further, it is reasonable to permit the center of an approximating discrete halfspace symmetric distribution to carry some probability mass. Thus halfspace symmetry indeed constitutes an important widening.

(iii) The points in a dataset typically are *not* symmetric in any sense about any point, and thus typical multivariate location estimators do not themselves represent a point of symmetry. We may then ask what is being estimated by these statistics, and a natural answer is given in terms of *functionals*: a statistic defined as some functional evaluated at the empirical cdf estimates that functional evaluated at the underlying population cdf. When the data can be considered as a sample from a distribution *symmetric* in some sense, such sample statistics take on added meaning. It is thus desirable to be able to assume, whenever possible, some sort of symmetric distribution as a model for the data. □

2.2 Characterizations and Interrelations

2.2.1 Characterizations of angular symmetry

Preliminary definitions. A set of N vectors in Euclidean $\mathbf{R}^d$ is said to be *in general position* if every d-element subset is linearly independent, and a set of N hyperplanes through the origin is said to be *in general position* if the corresponding set of normal vectors is in general position. A set of N hyperplanes in general position through the origin partitions $\mathbf{R}^d$ into

$$2\sum_{i=0}^{d-1}\binom{N-1}{i} \tag{6}$$

regions which are proper, nondegenerate polyhedral *convex cones* with the origin as vertices. (See, e.g., Cover and Efron (1967) for discussion and background.) □

The following result is easily understood intuitively, but a completely rigorous proof requires lengthy detail which we defer to the Appendix.

Theorem 2.2 *The following statements are equivalent:*

(i) $X \in \mathbf{R}^d$ *is angularly symmetric about a point* $\theta \in \mathbf{R}^d$;

(ii) $P(X - \theta \in C) = P(X - \theta \in -C)$, *for any circular cone* $C \subset \mathbf{R}^d$ *with the origin as vertex*;

(iii) $P(X - \theta \in C) = P(X - \theta \in -C)$, *for any convex cone* $C \subset \mathbf{R}^d$ *with the origin as vertex;*

(iv) $P(X - \theta \in H) = P(X - \theta \in -H)$, *for any closed halfspace* H *with the origin on the boundary*;

(v) $P(u'(X - \theta) \geq 0) = P(u'(\theta - X) \geq 0)$, *for any unit vector* $u \in \mathbf{R}^d$.

The preceding result yields

Corollary 2.1 *Suppose that the random vector* $X \in \mathbf{R}^d$ *is angularly symmetric about a point* $\theta \in \mathbf{R}^d$. *Then*

(i) $P(X \in H + \theta) = P(X \in -H + \theta) \geq \frac{1}{2}$ *for any closed halfspace* H *with the origin on the boundary. Further, if* X *is continuous, then* $P(X \in H + \theta) = 1/2$*;*

(ii) $\mathrm{Med}(u'X) = u'\theta$ *for any unit vector* $u \in \mathbf{R}^d$.

PROOF. (i) By Theorem 2.2, we immediately have that $P(X \in H + \theta) = P(X \in -H + \theta)$ for any closed halfspace H with the origin on the boundary. Since $P(X \in H + \theta) + P(X \in -H+\theta) \geq 1$, it follows that $P(X \in H+\theta) = P(X \in -H+\theta) \geq 1/2$ and $P(X \in H+\theta) = 1/2$ if X is continuous.

(ii) By Theorem 2.2, we obtain that $P(u'(X - \theta) \geq 0) = P(u'(\theta - X) \geq 0)$ for any unit vector $u \in \mathbf{R}^d$, whence $\mathrm{Med}(u'X) = u'\theta$ for any unit vector $u \in \mathbf{R}^d$. □

In the discrete case, we have a simple characterization of angular symmetry which is helpful in practice for the purpose of checking whether a given multivariate distribution is angularly symmetric.

Theorem 2.3 *Suppose* $X \in \mathbf{R}^d$ *is discrete. Then the following statements are equivalent:*

(i) $X \in \mathbf{R}^d$ *is angularly symmetric about a point* $\theta \in \mathbf{R}^d$;

(ii) $P(X - \theta \in L) = P(X - \theta \in -L)$ *for any ray* L *passing through the origin.*

PROOF. (i) $\Rightarrow$ (ii). Assume (i). By Theorem 2.2, we have that $P(X - \theta \in C) = P(X - \theta \in -C)$ for any circular cone C with the origin as vertex. Consider the limit situation as the cone shrinks to a ray; we thus obtain (ii).

To show that (ii) $\Rightarrow$ (i), note that (ii) implies that $P(X - \theta \in C) = P(X - \theta \in -C)$ for any circular cone C with the origin as vertex, since $P(X - \theta \in C) = \sum_i P(X \in l^i + \theta)$ and $P(X - \theta \in -C) = \sum_i P(X \in -l^i + \theta)$ for some rays $\{l^i\}$ (at most countably many) passing through the origin. Then (i) follows from (ii) of Theorem 2.2. □

2.2.2 Characterizations of halfspace symmetry

In the subsequent discussion we will assume that the point of halfspace symmetry (if any) of a multivariate distribution is unique.

Denote a closed halfspace with θ on its boundary by H_θ, and its reflection about θ by $\widetilde{H}_\theta$. Similarly, denote a ray stemming from θ by L_θ, and its reflection about θ by $\widetilde{L}_\theta$.

Theorem 2.4 *If any one of the following statements holds for a unique point θ in $\mathbf{R}^d$, then so do the other statements:*

(i) $X \in \mathbf{R}^d$ *is halfspace symmetric about θ;*

(ii) $\mathrm{Med}(u'X) = u'\theta$ *for any unit vector $u \in \mathbf{R}^d$;*

(iii) $P(u'(X-\theta) \geq 0) \geq \frac{1}{2}$ *for any unit vector $u \in \mathbf{R}^d$;*

(iv) $P(X \in (H_\theta^1 \cap H_\theta^2)^\circ) + P(X \in \pi^1 \cap \pi^2) \geq P(X \in (\widetilde{H}_\theta^1 \cap \widetilde{H}_\theta^2)^\circ)$ *for any closed halfspaces $H_\theta^i \subset \mathbf{R}^d$ with θ on the boundaries π^i, $i = 1, 2$;*

(v) $P(X \in (H_\theta^1 \cap H_\theta^2)) + P(X \in \pi^1 \cap \pi^2) \geq P(X \in (\widetilde{H}_\theta^1 \cap \widetilde{H}_\theta^2))$ *for any closed halfspaces $H_\theta^i \subset \mathbf{R}^d$ with θ on the boundaries π^i, $i = 1, 2$.*

Proof. (a) (i) $\Leftrightarrow$ (iii). X is halfspace symmetric about θ iff $P(X \in H + \theta) \geq 1/2$ for any closed halfspace $H \subset \mathbf{R}^d$ with the origin on its boundary. But $X \in (H + \theta)$ for some closed halfspace $H \subset \mathbf{R}^d$ with the origin on its boundary iff $u'(X - \theta) \geq 0$ for a corresponding unit vector $u \in \mathbf{R}^d$ which is normal to H. Thus X is halfspace symmetric about θ iff $P(u'(X-\theta) \geq 0) \geq 1/2$ for any unit vector $u \in \mathbf{R}^d$. Thus (i) $\Leftrightarrow$ (iii).

(b) (ii) $\Leftrightarrow$ (iii). Trivial.

(c) (i) $\Leftrightarrow$ (iv). Suppose that X is halfspace symmetric about $\theta \in \mathbf{R}^d$. Let H_θ^i be a closed halfspace with θ on its boundary π^i, $i = 1, 2$. It is not difficult to see that there are closed halfspaces H_θ^3 and H_θ^4 such that

$$(H_\theta^1 \cap \pi^2) \cup (\widetilde{H}_\theta^2 \cap \pi^1) \subset H_\theta^i \qquad \text{for } i = 3, 4,$$

and for any $\epsilon > 0$

$$\begin{aligned} P(X \in H_\theta^4) &\leq P(X \in (H_\theta^1 \cap H_\theta^2)^\circ) + P(X \in H_\theta^1 \cap \widetilde{H}_\theta^2) + \frac{\epsilon}{2}, \\ P(X \in (H_\theta^3)^\circ) &\geq P(X \in (\widetilde{H}_\theta^1 \cap \widetilde{H}_\theta^2)^\circ) + P(X \in H_\theta^1 \cap \widetilde{H}_\theta^2) \\ &\quad -P(X \in \pi^1 \cap \pi^2) - \frac{\epsilon}{2}. \end{aligned}$$

Since X is halfspace symmetric about $\theta \in \mathbf{R}^d$, we have $P(X \in (H_\theta^3)^\circ) \leq \frac{1}{2} \leq P(X \in H_\theta^4)$. Thus

$$P(X \in (H_\theta^1 \cap H_\theta^2)^\circ) + P(X \in \pi^1 \cap \pi^2) \geq P(X \in (\widetilde{H}_\theta^1 \cap \widetilde{H}_\theta^2)^\circ) - \epsilon,$$

for any $\epsilon > 0$. Since ϵ is arbitrary, it now follows that (i) $\Rightarrow$ (iv).

To show that (iv) $\Rightarrow$ (i), take $H_\theta^1 = H_\theta^2$. Then

$$P(X \in (H_\theta^1)^\circ) + P(X \in \partial H_\theta^1) \geq P(X \in (\widetilde{H}_\theta^1)^\circ).$$

Thus, $P(X \in H_\theta^1) \geq \frac{1}{2}$ for any closed halfspace H_θ^1 with $\theta \in \mathbf{R}^d$ on its boundary, proving that X is halfspace symmetric about θ.

(d) (i) $\Leftrightarrow$ (v). Similar to the proof that (i) $\Leftrightarrow$ (iv). □

When X is discrete, the following necessary conditions can be used to check halfspace symmetry of the underlying distribution.

Theorem 2.5 *Suppose that $X \in \mathbf{R}^d$ is discrete and halfspace symmetric about a unique point $\theta \in \mathbf{R}^d$. Then*

(i) $P(X \in H_\theta^\circ) + P(X = \theta) \geq P\left(X \in (\widetilde{H}_\theta)^\circ\right)$ *for any closed halfspace $H_\theta \subset \mathbf{R}^d$ with θ on the boundary;*

(ii) $P(X \in L_\theta) + P(X = \theta) \geq P(X \in \widetilde{L}_\theta)$ *for any ray $L_\theta \subset \mathbf{R}^d$ stemming from θ.*

Proof. The proof for $d = 1$ is trivial. Now take $d \geq 2$.

(i) Let $H_\theta \subset \mathbf{R}^d$ be a closed halfspace with $\theta \in \mathbf{R}^d$ on its boundary (hyperplane π). Since X is discrete, for any small $\epsilon > 0$, it is not difficult to see that there is a closed halfspace $H_\theta^1 \subset \mathbf{R}^d$ with a boundary hyperplane π_1 such that $P(X \in \pi \cap \pi_1) = P(X = \theta)$ and

$$P(X \in H_\theta^1) \leq P(X \in H_\theta^\circ) + P(X = \theta) + P(X \in (H_\theta^1 \cap \pi - \{\theta\})) + \epsilon,$$

and a closed halfspace $H_\theta^2 \subset \mathbf{R}^d$ with a boundary hyperplane π_2 such that $\pi \cap \pi_2 = \pi \cap \pi_1$ and

$$P(X \in (H_\theta^2)^\circ) \geq P(X \in (\widetilde{H}_\theta)^\circ) + P(X \in (H_\theta^1 \cap \pi - \{\theta\})) - \epsilon.$$

Since X is halfspace symmetric about θ, $P(X \in H_\theta^1) \geq \frac{1}{2}$ and $P(X \in (H_\theta^2)^\circ) \leq \frac{1}{2}$, it follows that

$$P(X \in H_\theta^\circ) + P(X = \theta) \geq P(X \in (\widetilde{H}_\theta)^\circ) - 2\epsilon.$$

Since the above inequality holds true for any sufficiently small ϵ, we have

$$P(X \in H_\theta^\circ) + P(X = \theta) \geq P(X \in (\widetilde{H}_\theta)^\circ),$$

for any closed halfspace $H_\theta \subset \mathbf{R}^d$ with θ on the boundary.

(ii) Let $L_\theta \subset \mathbf{R}^d$ be a ray stemming from θ with $\widetilde{L}_\theta$ as its reflection about θ. Since X is discrete, it is not difficult to see that there is a closed halfspace $H_\theta \subset \mathbf{R}^d$ with a boundary (hyperplane π) containing the ray L_θ such that $P(X \in \pi) = P(X \in L_\theta)$. For any $\epsilon > 0$, it is not difficult to see that there is a closed halfspace $H_\theta^1 \subset \mathbf{R}^d$ containing L_θ, and with θ on its boundary, such that

$$P(X \in H_\theta^1) \leq P(X \in L_\theta) + P(X \in H_\theta^\circ) + \epsilon,$$

and a closed halfspace $H_\theta^2 \subset \mathbf{R}^d$ containing $\widetilde{L}_\theta$, and with θ on its boundary, such that

$$P(X \in (H_\theta^2)^\circ) \geq P(X \in \widetilde{L}_\theta) - P(X = \theta) + P(X \in H_\theta^\circ) - \epsilon.$$

Since X is halfspace symmetric about θ, $P(X \in H_\theta^1) \geq 1/2$ and $P(X \in (H_\theta^2)^\circ) \leq 1/2$, and thus

$$P(X \in L_\theta) + P(X = \theta) \geq P(X \in \widetilde{L}_\theta) - 2\epsilon.$$

Hence $P(X \in L_\theta) + P(X = \theta) \geq P(X \in \widetilde{L}_\theta)$, for any ray $L_\theta \subset \mathbf{R}^d$ stemming from θ. $\square$

2.2.3 Interrelations

In Lemma 2.4 and Remarks 2.2(i), it was seen that angular symmetry implies halfspace symmetry, whereas the converse does not hold. A natural question is then: Under what conditions does halfspace symmetry reduce to angular symmetry? The answer is given in the following result.

Theorem 2.6 *Suppose a random vector X is halfspace symmetric about a unique point $\theta \in \mathbf{R}^d$, and either*
(1) *X is continuous,*
or
(2) *X is discrete and $P(X = \theta) = 0$.*
Then X is angularly symmetric about θ.

Proof. (a) Assume (1). By the definition of halfspace symmetry and the continuity of X, we have $P(X \in H + \theta) = P(X \in -H + \theta) = .5$ for any closed halfspace H with the origin on the boundary. Angular symmetry of X now follows from Theorem 2.2.

(b) Assume (2) and suppose that X is not angularly symmetric about the point $\theta \in \mathbf{R}^d$. By Theorem 2.3, there is a line L through θ with rays L_θ^+ and L_θ^- on either side of θ, such that $P(X \in L_\theta^+) \neq P(X \in L_\theta^-)$. Assume, without loss of generality, that $P(X \in L_\theta^+) > P(X \in L_\theta^-) + \epsilon$ for some $\epsilon > 0$. Since X is discrete, it is clear that there is a closed halfspace H_θ containing L such that $P(X \in \partial H_\theta) = P(X \in L)$. Also, there exists $H_\theta^+ \subset \mathbf{R}^d$, a closed halfspace which just includes L_θ^+ and excludes L_θ^-, such that $P(X \in \partial H_\theta^+) = 0$ and

$$P(X \in H_\theta^+) > P(X \in (H_\theta)^\circ) + P(X \in L_\theta^+) - \frac{\epsilon}{2}.$$

Similarly, there exists $H_\theta^- \subset \mathbf{R}^d$, a closed halfspace which just includes L_θ^- and excludes L_θ^+, such that $P(X \in \partial H_\theta^-) = 0$ and

$$P(X \in H_\theta^-) < P(X \in (H_\theta)^\circ) + P(X \in L_\theta^-) + \frac{\epsilon}{2}.$$

By the definition of halfspace symmetry and the properties of H_θ^+ and H_θ^-,

$$P(X \in H_\theta^+) = P(X \in H_\theta^-) = \frac{1}{2}.$$

Hence

$$P(X \in (H_\theta)^\circ) + P(X \in L_\theta^+) - \frac{\epsilon}{2} < \frac{1}{2} < P(X \in (H_\theta)^\circ) + P(X \in L_\theta^-) + \frac{\epsilon}{2},$$

which contradicts the assumption about $P(X \in L_\theta^+)$ and $P(X \in L_\theta^-)$. Thus $P(X \in L_\theta^+) = P(X \in L_\theta^-)$ for any line L through θ with rays L_θ^+ and L_θ^- on either side of θ. Angular symmetry of X about θ now follows from Theorem 2.3. □

3 Performance Characteristics of Selected Location Measures

Besides the multivariate *mean*, a number of versions of *multidimensional median* have become relatively popular as multivariate location measures. Here we focus on four notions of multivariate median: the "L^2" (or "spatial"), "Tukey/Donoho halfspace", "projection", and "Liu simplicial" medians. In answer to the question

> What properties should ideally be satisfied by any nonparametric multivariate location measure?

we consider the well-known properties *affine equivariance* and *stochastic order preserving* popularized by Bickel and Lehmann (1975) in the one-dimensional case and Oja (1983) in the higher-dimensional case, and also we introduce a *center-locating* property. By the latter is meant that a location measure, when evaluated on any *symmetric* distribution, should agree with the point of symmetry. These three properties are explicitly formulated in Section 3.1. The above four multidimensional medians are defined in Section 3.2 and their performance investigated in Section 3.3. Comparisons and conclusions are deferred to Section 4.

3.1 Three criteria

Let $\mathcal{F}$ be a class of probability distributions in $\mathbf{R}^d$ and $\psi : \mathcal{F} \to \mathbf{R}^d$ an $\mathbf{R}^d$-valued functional defined on $\mathcal{F}$. For a random vector $X \in \mathbf{R}^d$, denote its distribution function by F_X and for convenience write $\psi(X)$ for $\psi(F_X)$. For $x = (x_1, \ldots, x_d)'$ and $y = (y_1, \ldots, y_d)'$ in $\mathbf{R}^d$, the notation $x \leq y$ will denote the coordinatewise ordering, $x_i \leq y_i$ for $i = 1, \ldots, d$.

For random variables and distributions in $\mathbf{R}$, we denote by $X \overset{\text{st}}{\leq} Y$ that X is *stochastically smaller than* Y, i.e., that $F_X(x) \geq F_Y(x)$ for every x. More generally, for X and Y having distributions in $\mathbf{R}^d$, we define X to be *stochastically smaller* than Y if $\int g\, dF_X \leq \int f\, dF_Y$ for every real bounded coordinatewise increasing g on $\mathbf{R}^d$.

For a functional $\psi : \mathcal{F} \to \mathbf{R}^d$ to be a multivariate nonparametric *location measure*, Bickel and Lehmann (1975) and Oja (1983) require two conditions:

C1 ψ is "affine equivariant": $\psi(AX + b) = A\psi(X) + b$, for every nonsingular $d \times d$ matrix A and vector $b \in \mathbf{R}^d$;

C2 ψ is "stochastic order preserving": If $X \overset{\text{st}}{\leq} Y$, then $\psi(X) \leq \psi(Y)$.

Roughly speaking, C1 requires that the location measure ψ should not depend on the underlying coordinate system, nor on the scales of the underlying measurements, and C2 requires that ψ should take on larger values for random vectors which typically are "larger". Both of C1 and C2 are satisfied, for example, by the *mean functional*

$$\psi_0(F) = \int x dF(x). \quad F \in \mathcal{F}.$$

In the case of a distribution which is *symmetric* under a given notion of symmetry, it is reasonable to adopt the point of symmetry as its "location" (at least within linear transformation) and thus to require of any location measure ψ that it be able to identify this location. Specifically, we introduce

C3 ψ is "center locating": For any $F \in \mathcal{F}$ which is symmetric about a unique point $\theta \in \mathbf{R}^d$ under a given notion of symmetry, $\psi(F) = \theta$ should hold (at least within linear transformation).

In the sequel, our primary purpose is to confine attention to location measures which satisfy C1 (at least nearly so) and investigate which ones satisfy C3. Regarding C2, we shall see in Section 3.2 that it is too stringent and thus will give it lesser attention.

The following easily proved result shows that C3 is easily satisfied for typical kinds of location measure in the case of *centrally symmetric* distributions. (For $d = 1$ a related result is given by Corollary 1.3.19 of Randles and Wolfe (1979), to the effect that odd translation statistics have distributions symmetric about the point of symmetry of the distribution of the observations.)

Theorem 3.1 *Suppose that $T : \mathcal{F} \to \mathbf{R}^d$ is an odd and translation equivariant functional, that is, $T(F_{-X}) = -T(F_X)$ and $T(F_{X+b}) = T(F_X) + b$ for any X and vector b in $\mathbf{R}^d$. Let X be centrally symmetric about θ in $\mathbf{R}^d$. Then $T(F_X) = \theta$.*

The foregoing result yields

Corollary 3.1 *If ψ satisfies C1, then for C-symmetric F it also satisfies C3.*

Remarks 3.1 (i) The mean functional ψ_0 satisfies C1 and thus for C-symmetric F satisfies C3 as well.

(ii) For more general notions of symmetry, however, property C1 fails to imply C3. It is not difficult to construct distributions in $\mathbf{R}^d$ which are angularly symmetric about a point θ but have means taking values $\theta' \neq \theta$. Thus the mean functional ψ_0, while it satisfies C1 in general, cannot satisfy C3 for arbitrary angularly symmetric distributions.

(iii) Consequently, when notions of multivariate symmetry more general than central symmetry are considered, the mean functional fails to adequately represent a notion of multivariate "center". For this purpose *median* functionals become important, and several varieties will be considered below. □

3.2 Selected nonparametric multivariate location measures

Here we formally define the L^2 (or "spatial"), Tukey/Donoho halfspace, projection, and Liu simplicial depth medians.

These may be viewed from a common perspective by considering them as points of maximal depth of certain corresponding depth functions. For a given distribution F on $\mathbf{R}^d$, any function $D(\cdot, F)$ on $\mathbf{R}^d$ which provides an F-based center-outward ordering of points in $\mathbf{R}^d$ may be regarded (potentially, at least) as a *depth function*. Points of *maximal depth* thus serve as nonparametric notions of multivariate location. For many depth functions, in one dimension the point of maximal depth is the usual one-dimensional median. These depth functions thus also provide a source of various notions of *multidimensional* median.

Thus we shall indicate below the relevant depth functions associated with the above medians. For discussion of these and other depth functions, and of depth functions in general, see Zuo and Serfling (1998).

3.2.1 The L^2 median

A classical notion of multidimensional median is the L^2 or spatial median. (See Small (1990) for general discussion.) The L^2-median of a distribution F in $\mathbf{R}^d$ is defined by

$$\theta \in \arg \inf_{x \in \mathbf{R}^d} E\|X - x\|,$$

where X has distribution F and $\|\cdot\|$ denotes the usual Euclidean norm. This may be viewed as the point of maximal depth for the L^2 *depth function*

$$L^2\mathrm{D}(x; F) = (1 + E\|X - x\|)^{-1}.$$

A sample analogue estimator is given by the L^2-median of the usual empirical distribution $\widehat{F}_n(x) = n^{-1}\sum_{i=1}^{n} \mathbf{I}\{X_i \leq x\}$, $x \in \mathbf{R}^d$, based on a sample $X_1, \ldots, X_n$ from F.

Remarks 3.2 *A remark on terminology.* In $\mathbf{R}$, the median may be characterized as a value c that minimizes $E|X - c|$. One approach toward formulation of a median for distributions in a Banach space with norm $\|\cdot\|$ is to generalize this characterization, i.e., to minimize $E\|X - c\|$ (see Kemperman (1987)). This may be interpreted as minimizing the L^1-norm of the distance $\|X - c\|$, and on this basis the resulting median is sometimes called the L^1 median (e.g., Small (1990)). This is insufficiently specific, however, as the resulting median depends on the choice of norm $\|X - c\|$. In particular, for $\|X - c\|$ denoting the L^2 norm in $\mathbf{R}^d$, we obtain the above "spatial" median. On the other hand, for $\|X - c\|$ denoting the L^1 norm in $\mathbf{R}^d$, we obtain the "coordinatewise" median, i.e., the vector of coordinatewise medians. It is more informative, therefore, to name the resulting median the "L^p median" when $\|X - c\|$ is taken to be the L^p norm. Thus "L^p" indicates the choice of distance to be minimized, and "median" indicates that the minimizing is with respect to the L_1 norm. See Dodge and Rousson (1997) for further discussion. □

3.2.2 The Tukey/Donoho halfspace median

An early notion of depth function for the purpose of data analysis was the "halfspace depth" proposed by Tukey (1975), who suggested its role in defining multivariate analogues of univariate rank and order statistics via depth-induced "contours". Further development was stimulated by Donoho (1982), who studied its properties and focused on the associated notion of halfspace median. The *halfspace depth* of a point x in $\mathbf{R}^d$ with respect to a probability distribution F on $\mathbf{R}^d$ is defined to be the minimum probability mass carried by a closed halfspace containing x, i.e.,

$$HD(x; F) = \inf\{P_F(H) : H \text{ a closed halfspace}, \ x \in H\}, \ x \in \mathbf{R}^d.$$

A point θ maximizing $HD(\cdot, F)$ is called an H-median of F. A sample analogue estimator is given by the H-median of $\widehat{F}_n$.

3.2.3 The projection median

Define the *outlyingness* of a point x to be the worst case "outlyingness" with respect to the one-dimensional median in any one-dimensional projection of x, that is,

$$O(x; F) \equiv \sup_{\|u\|=1} \frac{\mid u'x - \mathrm{Med}(u'X) \mid}{\mathrm{MAD}(u'X)}, \tag{7}$$

where X has distribution F, MAD denotes the univariate median absolute deviation defined for univariate Y as $\mathrm{MAD}(Y) = \mathrm{Med}(|Y - \mathrm{Med}(Y)|)$, and $\| \cdot \|$ is the Euclidean norm. As discussed in Zuo and Serfling (1998), a corresponding *projection depth function* is given by

$$\mathrm{PD}(x; F) = (1 + O(x; F))^{-1}.$$

We then define the P-median as the point of maximal depth, or equivalently

$$\theta \in \arg \inf_{x \in \mathbf{R}^d} O(x; F).$$

A sample analogue estimator is given by putting $\widehat{F}_n$ for F. See Tyler (1994), p. 1032, for a different approach leading to formulation of the sample P-median.

For one-dimensional datasets $X = \{X_1, \ldots, X_n\}$, the quantity

$$O_n(x) \equiv |x - \mathrm{Med}_{1 \le i \le n}\{X_i\}| / \ \mathrm{MAD}_{1 \le i \le n}\{X_i\}$$

has long been used as a robust outlyingness measure of a point $x \in \mathbf{R}$ with respect to the center (median) of the dataset, and corresponding robust location estimators are given by weighting the observations according to their outlyingness. That is, introducing a bounded, continuous, decreasing, and nonnegative weight function $w(\cdot)$ with $xw(x)$ bounded, and defining weights $w_i = w(O_n(X_i))$, a "W-statistic" for robust location estimation is given by

$\sum_i w_i X_i / \sum_i w_i$. A particular case is the popular "biweight" estimator based on the bisquare weight function. See Mosteller and Tukey (1977), pp. 205–208 and 353, for this development.

For multivariate data, Stahel (1981) and Donoho (1982) independently generalized the one-dimensional outlyingness measure to arbitrary dimension d, defining $O_n(x)$ to be the worst case outlyingness of $x \in \mathbf{R}^d$ in any one-dimensional projection of x and the dataset X, and developed analogues of the above W-statistics to produce affine equivariant estimators having high breakdown point regardless of the dimension d. See also Donoho and Gasko (1992) and Tyler (1994).

3.2.4 The Liu simplicial depth median

Liu (1990) introduced another notion of depth. Namely, the *simplicial depth* of a point x in $\mathbf{R}^d$ with respect to a probability distribution F on $\mathbf{R}^d$ is defined to be the probability that x belongs to a random simplex in $\mathbf{R}^d$, i.e.,

$$SD(x; P) = P_F(x \in S[X_1, \ldots, X_{d+1}]), \;\; x \in \mathbf{R}^d,$$

where $X_1, \ldots, X_{d+1}$ is a random sample from F and $S[x_1, \ldots, x_{d+1}]$ denotes the d-dimensional simplex with vertices $x_1, \ldots, x_{d+1}$, that is, the set of all points in $\mathbf{R}^d$ that are convex combinations of $x_1, \ldots, x_{d+1}$. A point θ maximizing SD is called a *simplicial depth median* and denoted "S-median". The usual sample version of simplicial depth function is

$$SD_n(x) = \binom{n}{d+1}^{-1} \sum_{1 \le i_1 < \cdots < i_{d+1} \le n} \mathbf{I}\{x \in S[X_{i_1}, \ldots, X_{i_{d+1}}]\}, \;\; x \in \mathbf{R}^d,$$

which pointwise in x is a U-statistic, and which yields a sample S-median.

3.3 Performance analysis

Here we examine the four medians under consideration with respect to criteria given by properties C1–C3 along with degree of robustness as measured by breakdown point.

3.3.1 Property C1

Property C1 is not quite satisfied by the L^2-median. It is rigid-body equivariant (i.e., location and rotationally equivariant), however, and on this basis continues to receive acceptance in practice. Moreover, modified versions through normalization of observations achieving full affine equivariance through normalization of the observations have been introduced (for discussion, see Rao (1988), Small (1990), Chakraborty and Chaudhuri (1996, 1998, 1999), Chakraborty, Chaudhuri, and Oja (1998), Chaudhuri (1996), and Zuo and Serfling (1998)).

It is easily checked that the other three medians do satisfy C1:

Theorem 3.2 *The H-median, P-median, and S-median all satisfy C1.*

3.3.2 Property C2

When notions of symmetry broader than C-symmetry are considered, property C2 is too stringent a requirement for nonparametric location inference. As shown in the following two examples,

- two A-symmetric distributions in $\mathbf{R}^d$ can be stochastically ordered with their points of angular symmetry ordered in the opposite direction, and
- the S-medians of two stochastically ordered distributions can be ordered in the opposite direction.

Example 3.1 Let $d = 2$ and put $x_1 = (2,2)$, $x_1' = (-4,-4)$, $x_2 = (2,3)$, $x_2' = (-2,-5)$, $y_1 = (6,3)$, $y_1' = (-2,-1)$, $y_2 = (3,6)$, $y_2' = (-2,-4)$, $\theta_1 = (1,1)$, and $\theta_2 = (0,0)$. Define X and Y by $P(X = x_1) = P(X = x_1') = P(Y = y_1) = P(Y = y_1') = \frac{1}{5}$, $P(X = x_2) = P(X = x_2') = P(Y = y_2) = P(Y = y_2') = \frac{3}{10}$, $X(\omega_i) = x_i$, $X(\omega_i') = x_i'$, $Y(\omega_i) = y_i$, and $Y(\omega_i') = y_i'$, for $i = 1,2$. Then it is easy to see that X and Y are angularly symmetric about the unique points θ_1 and θ_2, respectively, and that $X \overset{\text{a.s.}}{\leq} Y$. However, $\theta_1 \not\leq \theta_2$, and thus C2 is not satisfied. □

Example 3.2 Let $d = 2$ and let distribution F place probability 1/40 on the point $(1,0)$, 1/40 on $(-1,0)$, 1/40 on $(0,-1)$, 3/40 on $(0,+1)$, 22/40 on $(-.4,-.4)$, and 12/40 on $(-.1,+.1)$. Then it is immediate that X is H-symmetric about $\theta = (-.4,-.4)$, which is thus the H-median by Theorem 3.4 below. Also, it is straightforward to check that the S-median, however, is the point $(-.1,+.1) \neq \theta$.

Now change the distribution F to the stochastically larger distribution F^* by shifting the probability 12/40 assigned to $(-.1,+.1)$ to the point $(+.1,+.1)$ instead. In this case the S-median changes *not* to a greater point but rather to a *lower* point, in particular the point $\theta = (-.4,-.4)$. (Note that the H-median, however, does *not* change when F is changed to F^*.) □

These examples suggest that any notion of multidimensional median which identifies the point of symmetry (i.e., which satisfies C3) for at least all A-symmetric distributions cannot in general satisfy C2 as well. As will be seen below, this applies to all four of the medians under present consideration.

3.3.3 Property C3

Using Theorem 3.1 for the L^2-median and Corollary 3.1 for the H-, P-, and S-medians, we obtain

Theorem 3.3 *For F C-symmetric, the L^2-median, H-median, P-median, and S-median all satisfy C3.*

We inquire, however, whether these medians also satisfy C3 in the case of A- or H-symmetric distributions. The following two results answer this question. The proofs are deferred to the Appendix.

Theorem 3.4 *For F H-symmetric about a unique point of symmetry, the L^2-median, H-median, and P-median all satisfy C3.*

Theorem 3.5

(i) *For F A-symmetric about a unique point of symmetry, the S-median satisfies C3.*

(ii) *For F H-symmetric about a unique point of symmetry, however, the S-median can fail to satisfy C3.*

3.3.4 Robustness considerations

Finally, we review the (known) large-sample breakdown points of the sample versions of the above four medians. (Finite sample breakdown point comparisons are similar.) We have

- *The L^2-median.* The breakdown point is 1/2 (see Kemperman (1987) and Lopuhaa and Rousseeuw (1991)). This is also satisfied by the affine equivariant modification given by Chakraborty and Chaudhuri (1999).
- *The H-median.* The breakdown point is 1/3 (see Donoho and Gasko (1992)).
- *The P-median.* The breakdown point is 1/2 (see Tyler (1994) and Zuo (1999)).
- *The S-median.* The breakdown point is nonzero but bounded above by $1/(d+2)$ (see Chen (1995)).

4 Concluding Discussion

In view of the findings of §3.3.2, we give priority to properties C1 and C3 (over C2). We also consider robustness as measured by breakdown point.

The H- and P-medians each satisfy both the *affine equivariance* (C1) and the *center-locating* (C3) criteria, and each enjoys a relatively high breakdown point that does not diminish as the dimension d increases. The L^2-median, however, fails to meet C1 fully, although it does meet C3 fully and has a good breakdown point. Finally, although satisfying C1 fully, the S-median fails to meet C3 fully, and its breakdown point diminishes as d increases. On the basis of these considerations, the H- and P-medians appear most attractive overall. Although the latter are based on projection pursuit methodology and thus computationally intensive, new approaches and advances in computation are overcoming this disadvantage.

Although the L^2-median is less competitive according to the above criteria, the modified "transformation–retransformation" data-based version due to Chakraborty and Chaudhuri

(1999) is not only affine equivariant but has breakdown point 1/2. Combined with its favorable computational ease, this makes it a strong contender worth further investigation.

The well-known "coordinatewise" median (i.e., vector of coordinatewise medians) has favorable breakdown point of 1/2, but fails seriously to meet C1. A modified "transformation–retransformation" data-based version due to Chakraborty and Chaudhuri (1996, 1998, 1999), retains, however, the high breakdown point while also achieving affine equivariance. It too is computationally attractive and merits further investigation, although we comment that in one dimension it reduces only approximately to the sample median.

We also comment that the transformation-retransformation estimators do not have a convenient functional representation that illuminates what parameter is being estimated, nor are they representable as points of maximal depth of a well-defined general depth function.

Other versions of multidimensional median might also be considered. For example, the "Oja" median (Oja (1983)) has received considerable attention in the literature. Indeed, it satisfies C1 and C2 (Oja (1983)) and can be shown to satisfy C3 for H-symmetric F. Its breakdown point, however, is zero (Niinimaa, Oja, and Tableman (1990)), so we exclude it from contention.

Finally, we comment that our criteria (mainly C1, C3, and robustness) are not the only valid choices. Others may prefer to give greater emphasis to C2, (although we have exhibited a shortcoming of it), or to give more serious consideration to computational ease, or to include consideration of asymptotic efficiency. We have adopted a particular set of criteria that enables a useful degree of differentation among several leading types of multidimensional median.

5 Acknowledgment

The authors greatly appreciate the contributions of Dr. G. L. Thompson, who provided an improved proof of Theorem 2.2 and other helpful insights and comments. We also acknowledge gratefully the helpful suggestions of three anonymous referees.

Appendix: Selected Proofs

Proof of Theorem 2.2. Assume without loss of generality that $\theta = 0$ and define $Y = X/\|X\|$ for $X \neq 0$ and $Y = 0$ for $X = 0$. Then X is angularly symmetric about the origin iff Y is centrally symmetric about the origin.

(a) (i) $\Leftrightarrow$ (ii). By Lemma 2.1, Y is centrally symmetric iff $P(Y \in H) = P(Y \in -H)$ for any closed halfspace H in $\mathbf{R}^d$. Since Y is distributed on the set consisting of the unit hypersphere $\mathbf{S}^{d-1}$ and 0 in $\mathbf{R}^d$, we have $P(Y \in H) = P(Y \in -H)$ iff $P(Y \in H \cap \mathbf{S}^{d-1}) = P(Y \in -H \cap \mathbf{S}^{d-1})$ for any closed halfspace H in $\mathbf{R}^d$. But $P(Y \in H \cap \mathbf{S}^{d-1}) = P(Y \in -H \cap \mathbf{S}^{d-1})$ for any closed halfspace H in $\mathbf{R}^d$ iff $P(X \in C) = P(X \in -C)$ for any closed circular cone C with the origin as vertex, since $Y \in H \cap \mathbf{S}^{d-1}$ for some closed halfspace H

iff $X \in C$ for a corresponding closed circular cone C with the origin as the vertex. Hence (i) $\Leftrightarrow$ (ii).

(b) (i) $\Leftrightarrow$ (iii). Now Y is centrally symmetric about the origin iff $P(Y \in B) = P(Y \in -B)$, for any Borel set B in $\mathbf{R}^d$. Since Y is distributed on $\{\mathbf{S}^{d-1}, 0\}$, Y is centrally symmetric about the origin iff $P(Y \in B \cap \mathbf{S}^{d-1}) = P(Y \in -B \cap \mathbf{S}^{d-1})$ for any Borel set B in $\mathbf{R}^d$. Since $X \in C$ for some convex cone C with the origin as vertex iff $Y \in B \cap \mathbf{S}^{d-1}$ for a corresponding Borel set B in $\mathbf{R}^d$, Y centrally symmetric about the origin implies that $P(X \in C) = P(X \in -C)$ for any closed convex cone C with the origin as vertex, proving that (i) $\Rightarrow$ (iii).

To prove that (iii) $\Rightarrow$ (i), note that if $P(X \in C) = P(X \in -C)$ for any closed convex cone C with the origin as vertex, then $P(Y \in C \cap \mathbf{S}^{d-1}) = P(Y \in -C \cap \mathbf{S}^{d-1})$. Since any ellipsoid can be approximated by convex hulls, and since $P(A_n) \uparrow P(A)$ if $A_n \uparrow A$, we have that if $P(Y \in C \cap \mathbf{S}^{d-1}) = P(Y \in -C \cap \mathbf{S}^{d-1})$ for any closed convex cone C with the origin as the vertex, then $P(Y \in H \cap \mathbf{S}^{d-1}) = P(Y \in -H \cap \mathbf{S}^{d-1})$ for any closed halfspace $H \in \mathbf{R}^d$, i.e., Y is centrally symmetric about the origin by Lemma 2.1. Hence (iii) $\Rightarrow$ (i).

(c) (i) $\Leftrightarrow$ (iv). Now, from the proof of (i) $\Leftrightarrow$ (ii), we have that X is angularly symmetric about the origin iff $P(Y \in H \cap \mathbf{S}^{d-1}) = P(Y \in -H \cap \mathbf{S}^{d-1})$ for any closed halfspace $H \subset \mathbf{R}^d$. Since $P(X \in H) = P(-X \in H)$ iff $P(Y \in H \cap \mathbf{S}^{d-1}) = P(Y \in -H \cap \mathbf{S}^{d-1})$ for any closed halfspace H with the origin on the boundary, (i) $\Rightarrow$ (iv).

To prove that (iv) $\Rightarrow$ (i), take $d = 2$ for the sake of simplicity. First we show that if $P(X \in H) = P(X \in -H)$ for any closed halfspace H with the origin on the boundary, then

$$P(X \in H_1 \cap H_2) = P(X \in -H_1 \cap -H_2) \tag{A.1}$$

for any continuous or discrete random vector $X \in \mathbf{R}^2$ and any closed halfspaces H_1 and H_2 with the origin on their boundaries.

1°. Suppose that X is continuous. In this case we have that

$$\begin{aligned} & P(X \in H_1 \cap H_2) + P(X \in (H_1 \cap -H_2)) \\ = \ & P(X \in -H_1 \cap -H_2) + P(X \in (H_1 \cap -H_2)) \end{aligned}$$

for any closed halfspaces H_1 and H_2 with the origin on the boundaries. Thus (A.1) holds.

2°. Suppose that X is discrete. Assume (A.1) is violated. Then there exist closed halfspaces H_1 and H_2 with the origin on their boundaries and $\epsilon > 0$ such that

$$P(X \in H_1 \cap H_2) > P(X \in -H_1 \cap -H_2) + \epsilon. \tag{A.2}$$

Now it is not very difficult to see that there are closed halfspaces H^3 and H^4 through the origin such that $P(X \in H^3) = P(X \in H^4)$, and

$$\begin{aligned} P(X \in H^3) & \geq P(X \in H_1 \cap H_2) + P(X \in (H^1 \cap -H^2)^\circ) - \frac{\epsilon}{2}, \\ P(X \in H^4) & \leq P(\in -H_1 \cap -H_2) + P(X \in (H^1 \cap -H^2)^\circ) + \frac{\epsilon}{2}. \end{aligned}$$

The above two inequalities imply that $P(X \in H_1 \cap H_2) < P(X \in -H_1 \cap -H_2) + \epsilon$, contradicting (A.2). Hence (A.1) holds.

Now, by (A.1), we have $P(Y \in (H_1 \cap H_2) \cap \mathbf{S}^1) = P(Y \in (-H_1 \cap -H_2) \cap \mathbf{S}^1)$ for any closed halfspace H_1 and H_2 with the origin on the boundary. Thus $P(Y \in C) = P(Y \in -C)$, for any closed arc C on the unit circle. But the set of all closed arcs on the unit circle forms a *π-system* (e.g., Billingsley (1986), p. 34) and generates all Borel sets on the unit circle. By the unique extension theorem for π-systems (Billingsley (1986), Theorem 3.3), $Y \stackrel{\mathrm{d}}{=} -Y$ on the unit circle. Hence X is angularly symmetric about the origin, proving that (iv) $\Rightarrow$ (i).

(d) (iv) $\Leftrightarrow$ (v). Straightforward. $\square$

Proof of Theorem 3.4 for the L^2-median.

Case 1. *F is A-symmetric.* Considering $E\|X - t\|$ as a function of $t \in \mathbf{R}^d$, by vector differentiation we have

$$\begin{aligned}
\frac{d(E\|X-t\|)}{dt} &= \frac{d(\int \|x-t\|\, dF_X(x))}{dt} = \int \frac{d\|x-t\|}{dt}\, dF_X(x) = \int \frac{x-t}{\|x-t\|}\, dF_X(x) \\
&= E\left[\frac{X-t}{\|X-t\|}\right]. \qquad \text{(A.3)}
\end{aligned}$$

For X angularly symmetric about θ, we have

$$E\left[\frac{X-\theta}{\|X-\theta\|}\right] = E\left[\frac{\theta - X}{\|X-\theta\|}\right] = 0.$$

Then convexity of $\|\cdot\|$ and (A.3) imply that θ is the L^2-median.

Case 2. *Other H-symmetric F.* By Theorem 2.6, it remains to consider only the case of discrete H-symmetric X with nonzero probability p_θ on the unique center θ of H-symmetry.

For descriptive simplicity, we consider only the case $d = 2$. By Theorem 4.17 of Kemperman (1987) and rigid-body equivariance of the L^2 median, we may assume w.l.o.g. that θ is the origin and all probability mass is carried by the unit circle. We need to show that

$$E\|X - \theta\| \le E\|X - x\|, \qquad \forall x \in \mathbf{R}^2. \qquad \text{(A.4)}$$

Assume w.l.o.g. that $x = (0, y)$, that $x_1, \ldots, x_n, \ldots$ fall on the left open semi-circle with $x_n \ne (-1, 0)$ for $n = 1, 2, \ldots$, and that $P(X = x_i) = p_i > 0$, $d_i = \|x - x_i\|$ and $d_i \le d_j$ for $i \le j$. Assume that $x_\alpha = (-1, 0)$, $x_\beta = (0, 1)$ and $x_\gamma = (0, -1)$, with $P(X = x_\alpha) = p_\alpha$, $P(X = x_\beta) = p_\beta$ and $P(X = x_\gamma) = p_\gamma$. Denote the point $-x_i$ by x_i' for $i = 1, 2, \ldots$.

Starting at x_γ, travel clockwise along the circle until hitting the first point, say x_i, denote by $\mathrm{arc}(x_\beta, x_i']$ the open-closed arc which is the symmetric part about the origin of the arc just covered, and assume $P(X \in \mathrm{arc}(x_\beta, x_i']) = p_i'$. Starting from x_i now, travel clockwise along the circle until hitting the next point, say x_j, assume that $P(X \in \mathrm{arc}(x_i', x_j']) = p_j'$. Continue in this fashion and stop before hitting x_α. Starting at x_β now, travel counterclockwise along the circle performing similar steps as before, that is, labeling the probability mass on the open-closed arcs which are the symmetric parts about the origin of the arcs covered, and stopping before hitting x_α. Denote by p'_α the probability mass on the open-open arc on the right semi-circle, which contains $(1,0)$ and is the symmetric part about the origin of the open-open arc on the left semi-circle which contains x_α and is never covered above.

Now it is not difficult to show that

$$
\begin{aligned}
&E\|X-x\| - E\|X-\theta\| \\
&\quad \geq \quad (\sqrt{1+y^2}-1)(p_\beta+p_\gamma) + y(p_\alpha+p_\theta-p'_\alpha) + \sum_{i=1}[(d_i-1)p_i + (d'_i-1)p'_i] \\
&\quad = \quad (\sqrt{1+y^2}-1)(p_\beta+p_\gamma) \\
&\qquad\qquad +(d_1-1)(p_\alpha+p_\theta-p'_\alpha+\sum_{i=1}p_i-\sum_{i=1}p'_i) + (d_1+d'_1-2)p'_1 \\
&\qquad\qquad +(y+1-d_1)(p_\alpha+p_\theta-p'_\alpha) + \sum_{i=2}[(d_i-d_1)p_i + (d'_i+d_1-2)p'_i)] \\
&\quad = \quad (\sqrt{1+y^2}-1)(p_\beta+p_\gamma) \\
&\qquad\qquad +(d_1-1)(p_\alpha+p_\theta-p'_\alpha+\sum_{i=1}p_i-\sum_{i=1}p'_i) + (d_1+d'_1-2)p'_1 \\
&\qquad\qquad +(d_2-d_1)(p_\alpha+p_\theta-p'_\alpha+\sum_{i=2}p_i-\sum_{i=2}p'_i) + (d_2+d'_2-2)p'_2 \\
&\qquad\qquad +(y+1-d_2)(p_\alpha+p_\theta-p'_\alpha) + \sum_{i=3}[(d_i-d_2)p_i + (d'_i+d_2-2)p'_i)] \\
&\quad = \quad (\sqrt{1+y^2}-1)(p_\beta+p_\gamma) \\
&\qquad\qquad +(d_1-1)(p_\alpha+p_\theta-p'_\alpha+\sum_{i=1}p_i-\sum_{i=1}p'_i) + (d_1+d'_1-2)p'_1 \\
&\qquad\qquad +(d_2-d_1)(p_\alpha+p_\theta-p'_\alpha+\sum_{i=2}p_i-\sum_{i=2}p'_i) + (d_2+d'_2-2)p'_2 \\
&\qquad\qquad +\cdots \\
&\qquad\qquad +(d_n-d_{n-1})(p_\alpha+p_\theta-p'_\alpha+\sum_{i=n}p_i-\sum_{i=n}p'_i) + (d_n+d'_n-2)p'_n \\
&\qquad\qquad +(y+1-d_n)(p_\alpha+p_\theta-p'_\alpha) + \sum_{i=n+1}[(d_i-d_n)p_i + (d'_i+d_n-2)p'_i)] \\
&\quad = \quad (\sqrt{1+y^2}-1)(p_\beta+p_\gamma) \\
&\qquad\qquad +(d_1-1)(p_\alpha+p_\theta-p'_\alpha+\sum_{i=1}p_i-\sum_{i=1}p'_i) + (d_1+d'_1-2)p'_1
\end{aligned}
$$

$$
\begin{aligned}
&+(d_2-d_1)(p_\alpha+p_\theta-p'_\alpha+\sum_{i=2}p_i-\sum_{i=2}p'_i)+(d_2+d'_2-2)p'_2\\
&+\cdots\\
&+(d_n-d_{n-1})(p_\alpha+p_\theta-p'_\alpha+\sum_{i=n}p_i-\sum_{i=n}p'_i)+(d_n+d'_n-2)p'_n\\
&+\cdots.
\end{aligned}
$$

By Theorem 2.4, we immediately have that

$$
\begin{aligned}
p_\alpha+p_\theta &\geq p'_\alpha,\\
p_\alpha+p_\theta+\sum_{i=m}^{n}p_i &\geq p'_\alpha+\sum_{i=m}^{n}p'_i \qquad \forall\, m\leq n \ \text{ and } \ m\geq 1.
\end{aligned}
$$

In conjunction with the fact that $d_i+d'_i\geq 2,\ \forall\, i\geq 1$, we obtain (A.4). □

Proof of Theorem 3.4 for the H-median. Suppose that F is H-symmetric about a unique point $\theta\in\mathbf{R}^d$. By the definition of H-symmetry, we have $P_F(H_\theta)\geq 1/2$, for any closed halfspace H with $\theta\in\partial H$. It follows that $HD(\theta;F)\geq 1/2$. Now suppose that there is a point $x_0\in\mathbf{R}^d$, $x_0\neq\theta$, such that $HD(x_0;F)>1/2$. Then $P_F(H)>1/2$ for any closed halfspace H with $x_0\in\partial H$, which implies that F is also H-symmetric about x_0, contradicting the assumption that F is H-symmetric about a unique point $\theta\in\mathbf{R}^d$. Therefore, $HD(\theta;F)=\sup_{x\in\mathbf{R}^d}HD(x;F)$, i.e., θ is the H-median of F. □

Proof of Theorem 3.4 for the P-median. Suppose that F is H-symmetric about a unique point $\theta\in\mathbf{R}^d$. Then, for X having distribution F, we have by Theorem 2.4 that $\mathrm{Med}(u'X)=u'\theta$, each unit vector $u\in\mathbf{R}^d$, which readily yields that θ is the P-median of F. □

Proof of Theorem 3.5. (i)

Case 1. F *absolutely continuous.* (Modeled after the proof of Theorem 3 of Liu (1990).) Let F be continuous and angularly symmetric about a unique point in $\theta\in\mathbf{R}^d$. For any $x\in\mathbf{R}^d$, $x\neq\theta$, we have that

$$D(\theta)-D(x)=P(\theta\in S[X_1,\ldots,X_{d+1}])-P(x\in S[X_1,\ldots,X_{d+1}])$$

where $D(\cdot)$ denotes the simplicial depth function with respect to F, $X_1\ldots,X_{d+1}$ is a sample from F, and $S[X_1,\ldots,X_{d+1}]$ is the simplex consisting of vertices $X_1\ldots,X_{d+1}$. Let $\overrightarrow{\theta x}$ be the vector in $\mathbf{R}^d$ starting at θ and ending at x. Then

$$
\begin{aligned}
&D(\theta)-D(x)\\
&\quad= P(\ \overrightarrow{\theta x}\text{ leaves but not enters }S[X_1,\ldots,X_{d+1}])\\
&\qquad -P(\ \overrightarrow{\theta x}\text{ enters but not leaves }S[X_1,\ldots,X_{d+1}])
\end{aligned}
$$

$$= \; P(\overrightarrow{\theta x} \text{ leaves } S[X_1, \ldots, X_{d+1}]) - P(\overrightarrow{\theta x} \text{ enters } S[X_1, \ldots, X_{d+1}])$$

$$= \; (d+1) \int_{\overrightarrow{\theta x} \cap HP(x_1,\ldots,x_d) \neq \emptyset} P(X_{d+1} \in H_\theta(x_1, \ldots, x_d)\, dF(x_1) \ldots dF(x_d)$$

$$-(d+1) \int_{\overrightarrow{\theta x} \cap HP(x_1,\ldots,x_d) \neq \emptyset} P(X_{d+1} \in H_x(x_1, \ldots, x_d)\, dF(x_1) \ldots dF(x_d)$$

where $HP(x_1, \ldots, x_d)$ is the hyperplane determined by $x_1, \ldots, x_d$ and $H_y(x_1, \ldots, x_d)$ is the halfspace containing y and with $x_1, \ldots, x_d$ on the boundary. Angular symmetry about θ now implies that $H_\theta(x_1, \ldots, x_d) \geq .5 \geq H_x(x_1, \ldots, x_d)$, whence $D(\theta) \geq D(x)$, $x \in \mathbf{R}^d$, i.e., θ is the S-median of F.

Case 2. *F discrete.* Assume that F on $\mathbf{R}^d$ is discrete and angularly symmetric about a unique point $\theta \in \mathbf{R}^d$. For simplicity, we treat the case $d = 2$. The argument is readily generalized to $d > 2$.

In the above notation, we need to establish that $D(\theta) \geq D(x)$, $x \in \mathbf{R}^d$. For this, it is clear that we need only consider the simplices which contain x but not θ, or contain θ but not x. The definition of the simplicial median and Theorem 2.4 yield that there are only two cases which need be taken into account.

> Case 2A. *The vertices of the triangles containing x lie on two lines intersecting at θ.*

Assume AB and CD intersect at θ, x_{p_i} lies on θB, $x_{p'_i}$ lies on $A\theta$, x_{q_i} lies on θD, and $x_{q'_i}$ lies on $C\theta$, $i \in \{1, 2, \ldots\}$. For convenience, we sometimes use p_i to denote either the probability mass p_i on x_{p_i} or the point x_{p_i} itself, interchangeably. Assume that

$$\sum p_i = p, \;\; \sum p'_i = p', \;\; \sum q_j = q, \;\; \sum q'_j = q'.$$

Then by Theorem 2.4 $p = p'$, $\;q = q'$. Consider an arbitrary point p_i, and assume that $p_i x$ intersects θD at point y_{p_i} and that

$$P(X \in (\theta, \; y_{p_i})) = q_{i1}, \;\; P(X \in [y_{p_i}, \; D)) = q_{i2}.$$

Then it is not difficult to show that

$$P(x \in \Delta(X_1, X_2, X_3) \not\ni \theta)/3! = \sum_i p_i q_{i1} q_{i2} + \sum_j q_j p_{j1} p_{j2},$$

where p_{j1} and p_{j2} are defined similarly to q_{i1} and q_{i2}, and $\Delta(y_1, y_2, y_3)$ denotes the triangle with vertices y_1, y_2, y_3. On the other hand, it is also not difficult to see that

$$P(\theta \in \Delta(X_1, X_2, X_3) \not\ni x)/3! \geq \sum_i p_i q_{i1} q' + \sum_j q_j p_{j1} p'.$$

Thus

$$P(x \in \Delta(X_1, X_2, X_3)) \leq P(\theta \in \Delta(X_1, X_2, X_3)).$$

Case 2B. *The vertices of the triangles containing x lie on three lines intersecting at θ.*

Assume AB, CD and EF intersect at θ, x_{p_i} lies on θB, $x_{p'_i}$ lies on $A\theta$, x_{q_i} lies on θD, $x_{q'_i}$ lies on $C\theta$, x_{r_i} lies on θF, and $x_{r'_i}$ lies on $E\theta$, $i \in \{1, 2, \ldots\}$. Assume that

$$\sum p_i = p \quad \sum p'_i = p', \quad \sum q_j = q,$$

$$\sum q'_j = q', \quad \sum r_k = r, \quad \sum r'_k = r'.$$

By Theorem 2.4 we have that $p = p', \quad q = q', \quad r = r'$. Consider an arbitrary point p_i, connect p_i and x, intersecting θD at y_{p_i}, θF at z_{p_i}. Assume that

$$P(X \in (\theta, y_{p_i})) = q^1_{p_i}, \quad P(X \in [y_{p_i}, D)) = q^2_{p_i},$$

$$P(X \in (\theta, z_{p_i})) = r^1_{p_i}, \quad P(X \in [z_{p_i}, F)) = r^2_{p_i}.$$

Then it is not difficult to show that

$$\begin{aligned} &P(x \in \Delta(X_1, X_2, X_3) \not\ni \theta)/3! \\ &\leq \ \sum p_i q^1_{p_i} r^2_{p_i} + \sum p_i q^2_{p_i} r^1_{p_i} + \sum p'_i q^2_{p'_i} r^1_{p'_i} + \sum p'_i r^2_{p'_i} q^1_{p'_i} \end{aligned}$$

where $r^1_{p'_i}$, $r^2_{p'_i}$, $q^1_{p'_i}$ and $q^2_{p'_i}$ are defined similarly to $q^k_{p_i}$ and $r^k_{p_i}$ for $k = 1, 2$. It is also not very difficult to prove that

$$\begin{aligned} &P(\theta \in \Delta(X_1, X_2, X_3) \not\ni x))/3! \\ &\geq \ \sum p'_i \max_i \{q^1_{p_i}\} r' + \sum p_i q' r^1_{p_i} + \sum p_i \sum p'_i \sum r'_k + \sum q_j \sum q'_j \sum r'_k. \end{aligned}$$

(Note that it is very important that we have avoided the repeated use of the triangles containing θ which have been employed in the proof of Case 2A.) Thus

$$\begin{aligned} &P(\theta \in \Delta(X_1, X_2, X_3) \not\ni x)/3! \\ &\geq \ \sum p_i \max_i \{q^1_{p_i}\} r' + \sum p_i q' r^1_{p_i} + (p^2 + q^2) r \\ &\geq \ \sum p_i q^1_{p_i} r^2_{p_i} + \sum p_i q^2_{p_i} r^1_{p_i} + 2pqr \\ &\geq \ \sum p_i q^1_{p_i} r^2_{p_i} + \sum p_i q^2_{p_i} r^1_{p_i} + \sum p'_i q^2_{p'_i} r^1_{p'_i} + \sum p'_i q^1_{p'_i} r^2_{p'_i}, \end{aligned}$$

and thus $P(\theta \in \Delta(X_1, X_2, X_3)) \geq P(x \in \Delta(X_1, X_2, X_3))$.

Cases 2A and 2B combined now imply that $\theta = \arg\max_{x \in \mathbf{R}^2} P(x \in S[X_1, \ldots, X_{d+1}])$, completing the proof of (i).

(ii) It suffices to exhibit an H-symmetric F for which the point of symmetry is unique but differs from the S-median. Such an F is given in Example 3.2 above. □

References

[1] Bickel, P. J., and Lehmann, E. L. (1975). Descriptive statistics for nonparametric models. II. Location. *Ann. Statist.* **4** 1045–1069.

[2] Chakraborty, B., and Chaudhuri, P. (1996). On a transformation and re-transformation technique for constructing an affine equivariant multivariate median. *Proc. Amer. Math. Soc.* **124** 2539–2547.

[3] Chakraborty, B., and Chaudhuri, P. (1998). On an adaptive transformation–retransformation estimate of multivariate location. *J. R. Statist. Soc. B* **60** 145–157.

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

[5] Chakraborty, B., and Chaudhuri, P. (1999). A note on robustness of multivariate median. *Statist. Probab. Lettr.* **41** to appear.

[6] Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data. *J. Amer. Statist. Assoc.* **91** 862–872

[7] Chen, Z. (1995). Bounds for the breakdown point of the simplicial median. *J. Multiv. Analysis* **55** 1–13.

[8] Cover, T. M., and Efron, B. (1967). Geometrical probability and random points on a hypersphere. *Ann. Math. Statist.* **38** 213–220.

[9] Dodge, Y., and Rousson, V. (1997). Multivariate L_1 mean. Preprint.

[10] Donoho, D. L. (1982). Breakdown properties of multivariate location estimators. Ph.D. Qualifying Paper, Department of Statistics, Harvard University.

[11] Donoho, D. L., and Gasko, M. (1992). Breakdown properties of location estinates based on halfspace depth and projected outlyingness. *Ann. Statist.* **20** 1803–1827.

[12] 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* (Y. Dodge, ed.), pp. 217–230. North-Holland.

[13] Liu, R. Y. (1988). On a notion of simplicial depth. *Proc. Nat. Acad. Sci. USA* **85** 1732–1734.

[14] Liu, R. Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18** 405–414.

[15] Lopuhaa, H. P. and Rousseeuw, P. J. (1991). Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *Ann. Statist.* **19** 229–248.

[16] Mosteller, C. F., and Tukey, J. W. (1977). *Data Analysis and Regression.* Addison-Wesley, Reading, Mass.

[17] Niinimaa, A., Oja, H. and Tableman, M. (1990). The finite-sample breakdown point of the Oja bivariate median. *Statist. Prob. Letters* **10** 325–328.

[18] Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statist. Prob. Letters* **1** 327–333.

[19] Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections. *J. Amer. Statist. Assoc.* **84** 1045–1050.

[20] Randles, R. H., and Wolfe, D. A. (1979). *Introduction to the Theory of Nonparametric Statistics.* Wiley, New York.

[21] Rao, C. R. (1988). Methodology based on the L_1 norm in statistical inference. *Sankhyā Ser. A* **50** 289–313.

[22] Small, C. G. (1990). A survey of multidimensional medians. *Intern. Statist. Inst. Rev.* **58** 263–277.

[23] Stahel, W. A. (1981). Robuste Schätzungen: Infinitesimale Optimalität und Schätzungen von Kovarianzmatrizen. Ph.D. thesis, Eidgenossiche Technische Hochschule, Zürich.

[24] Tukey, J. W. (1975). Mathematics and picturing data. *Proc. Intern. Congr. Math. Vancouver 1974* **2** 523–531.

[25] Tyler, D. E. (1994). Finite sample breakdown points of projection based multivariate location and scatter statistics. *Ann. Statist.* **22** 1024–1044.

[26] Zuo, Y. (1999). Affine equivariant multivariate location estimates with best possible breakdown points. Submitted.

[27] Zuo, Y., and Serfling, R. (1998). General notions of statistical depth function and some related convergence results. Submitted.