On Min-Max Majority and Deepest Points

D. Nolan

University of California at Berkeley

Abstract

The asymptotic properties of a multivariate location estimator are obtained in this paper. The estimator examined is based on the notion of half-space depth, where the depth of a point is the minimum probability content of all half spaces containing the point. The location estimator of interest is the deepest point with respect to the empirical measure on half spaces. For angularly symmetric distributions, this estimator is $\sqrt{n}$ consistent. For two dimensions, the exact limit distribution is derived, and the extension of the limit distribution results to higher dimensions is discussed.

Keywords and phrases: robust multivariate location estimation; empirical process; min-max majority; Gaussian process; half space depth.

1 INTRODUCTION

The notion of depth for a point in one dimension dates back to Hotelling (1929) and Chamberlin (1937), where the depth or tailedness of a point t, with respect to a probability measure P, is the minimum of $P[t, \infty)$ and $P(-\infty, t]$. In one dimension, the median is a point with maximum depth. It follows that a natural extension of depth for t in $\mathbb{R}^d$ is the minimum P-content over all closed half spaces with boundary passing through t, and it follows that a point of greatest depth is a multidimensional analog of the median.

Hodges (1955) first extended the notion of depth to $\mathbb{R}^2$ in his bivariate sign test. He developed a test for location based on the smallest portion of the sample contained in any closed half space containing the origin. Tukey (1975) extended this notion of depth to $\mathbb{R}^d$, and Donoho (1982) studied the robustness properties of the deepest point in higher dimensions. Stahel (1981) also studied this location estimator. Huber (1985) noted the projection pursuit flavor of the deepest point – the assignment of depth to a point in $\mathbb{R}^d$ is equivalent to pursuing the one dimensional projection with largest depth. Small (1987) generalized this notion of depth to sets other than half spaces and to regions other than $\mathbb{R}^d$. Related work can be found in Ajne (1968), Randles (1989), Oja and Nyblom (1989), and Liu (1990).

The notion of depth also appears in the econometrics literature in the theory of social choice. There, the deepest points form a "min-max majority set"(Caplin and Nalebuff, 1988). The min-max majority set is a subset of proposals considered least unfavorable to some electorate. In particular, each proposal is represented by a point t in $\mathbb{R}^d$. It is assumed the complete set of proposals forms a convex subset of R^d. Caplin and Nalebuff show that the viability of a proposal is determined by the distribution of voters' most-preferred proposals, i.e. opposition to a specific proposal t is determined by: $1 - \text{depth}(t)$, where t's depth is found using the distribution constructed from the voters' preferences. Proposals of least opposition are those that minimize over t the quantity $1 - \text{depth}(t)$. Thus the name "min-max majority".

When the empirical distribution from a P-sample is used to determine the probability content of a half space, the resulting deepest point can be thought of as a location estimator. Rousseeuw and Leroy (1987, Chapter 7) and Small (1990) provide comparisons of this estimator with other multidimensional location estimators. They assess estimators according to their affine equivariance and breakdown point. The point of greatest depth is affine equivariant (Donoho and Gasko, 1992), and it achieves a breakdown of roughly 1/3 for many configurations of samples. In comparison, affine-equivariant M-estimates of location have breakdown point at best $1/(d+1)$ in $\mathbb{R}^d$ (Marronna, 1976). Estimators based on ellipsoidal peeling (Titterington, 1978) and convex peeling (Barnett, 1976), and Oja's (1983) generalized median have equally low breakdown points. Few affine equivariant location estimators have breakdown properties as good as the deepest point. One that does is a minimum volume ellipsoid estimator (Rousseeuw and Leroy, 1987). Unfortunately, this estimator has the unusually slow rate of convergence of $n^{1/3}$ (Davies, 1992, Kim and Pollard, 1990). Additionally, both the deepest point and the minimum volume ellipsoid estimator are computationally intensive. The search for the best projection requires examination of all $\binom{n}{d}$ hyper-planes determined by d sample points, and to find the smallest ellipsoid to contain half the sample requires examination of all $\binom{n}{[n/2]}$ subsamples of size $[n/2]$. Rousseeuw and Ruts (1996) and Rousseeuw (1993) propose alternative methods for computing the half space depth which are less computationally intensive.

In this paper, we establish asymptotic properties for the sample deepest point. We obtain consistency and $\sqrt{n}$ rates of convergence for the deepest point in $\mathbb{R}^d$. In addition for $d = 2$, we find the asymptotic distribution of the deepest point. The limit distribution is quite curious, as it is an unusual functional of a Gaussian process. In the next sections we state (Section 2) and prove (Section 4) that the deepest point converges in distribution to

$$\arg\sup_t \inf_u [Z(u) \ - \ u'tp_u(u't_0)],$$

where Z is a centered Gaussian process indexed by the unit sphere and p_u is the density of the marginal distribution P_u in the direction u. Each marginal distribution is required to have a unique median at $u't_0$ to establish this result. This requirement reflects the projection pursuit flavor in the notion of depth.

2 RESULTS

To formally define the deepest point, we introduce some notation. We denote by H_{ut} the half space $\{x \in \mathbb{R}^d : \ u'x \geq u't \ , t \in \mathbb{R}^d, \ u \in \mathbb{S}^{d-1}\}$. Also, let P_n represent the empirical measure that places mass $1/n$ on each of $X_1, \ldots, X_n$, a sample from the distribution P on $\mathbb{R}^d$.

1 DEFINITION. A point t_0 is a deepest point of the distribution P provided

$$t_0 \ = \ \arg\sup_{t\in\mathbb{R}^d} \inf_{u\in\mathbb{S}^{d-1}} P(H_{ut}).$$

2 DEFINITION. The deepest point of a sample $X_1, \ldots, X_n$ from P, call it T_n, is any point such that

$$T_n \ = \ \arg\sup_{t\in\mathbb{R}^d} \inf_{u\in\mathbb{S}^d} P_n(H_{ut}).$$

In this paper, we consider only P with t_0 uniquely defined. Nonuniqueness of the supremum in Definition 2 can be handled by any arbitrary, fixed rule. For example, if multiple points maximize $\inf_u P_n H_{ut}$ then choose T_n from those points such that T_n has minimum first coordinate. We assume such a rule is employed in the definition of T_n, and ignore the slight complication that would result in its representation. For the remainder of the paper, assume, for simplicity, that $t_0 = 0$.

1 LEMMA. Suppose for each (admissible) $\epsilon > 0$

$$\inf_u P(H_{u0}) > \sup_{|t|>\epsilon} \inf_u P(H_{ut}).$$

Then $|T_n| \to 0$, in probability.

To strengthen this result to a rate of convergence for T_n, we place more assumptions on P. Let P_u be the one dimensional marginal distribution of $u'X$, and let p_u be the density function of P_u.

2 LEMMA. Assume the following conditions hold:

(i) P_u has a unique median at 0, for all u;

(ii) P_u has a bounded positive density, p_u, at 0, and $p_u(x)$ is continuous in u and x at $x = 0$.

Then $\sqrt{n}T_n = O_p(1)$.

This result can be refined to obtain the asymptotic distribution of T_n for $d = 1, 2$. In one dimension, the conditions of Lemma 2 reduce to the typical requirements for finding the asymptotic distribution of the sample median – that the median is unique and the distribution has a positive, continuous density at the median.

3 LEMMA. Under the assumptions of Lemma 2, for $d = 1, 2$

$$\sqrt{n}T_n \to^{\mathcal{D}} \arg\sup_t \inf_u [Z(u) \;-\; u't\; p_u(0)],$$

where Z is a Gaussian process on the unit sphere with zero means and covariance structure:

$$\mathrm{Cov}[Z(u), Z(v)] \;=\; P(H_{u0}H_{v,0}) \;-\; 1/4.$$

3 DISCUSSION

(i) Lemma 1 says that uniqueness of the deepest point of P guarantees consistency. If $X/|X|$ is invariant under sign changes then the origin is the unique deepest point. In addition, the condition of Lemmas 2 and 3 that the origin is the unique median for each projection P_u is also met (Randles, 1989, Liu, 1990). As another example, Caplin and Nalebuff (1988) show that the uniform distribution over any triangle in $\mathbb{R}^2$ has a unique deepest point at the centroid of the triangle. However, as its depth is 4/9 not 1/2, the projections do not all have unique median at the projected centroid and Lemmas 2 and 3 do not cover this case. In more recent work, Caplin and Nalebuff (1990a,b) produce conditions for uniqueness of the deepest point based on a convexity criterion for the density.

(ii) Notice the periodicity of the limit process. If $u'v = -1$ then $Z(u) = -Z(v)$. In one dimension: $\mathbb{S}^0 = \{+1, -1\}$; the Gaussian process reduces to $[Z(1),\ Z(-1)]$, where $Z(1)$ has a $\mathcal{N}(0, 1/4)$ distribution and $Z(-1) = -Z(1)$. The "marginal" densities are $p_1(t) = p(t)$ and $p_{-1}(t) = p(-t)$. Therefore, $\min\{Z(1) - tp(0),\ -Z(1) + tp(0)\} \leq 0$, all t. The maximum occurs at 0 for $t = Z(1)/p(0)$, the well-known limit distribution of the sample median.

(iii) If we further restrict the distribution to be spherically symmetric then the one-dimensional marginals are all the same, say p_0. In this case, the asymptotic distribution of $\sqrt{n}T_n$ is

$$\arg\sup_t \inf_u [Z(u) \; - \; u't \; p_0(0)],$$

and the covariance of Z becomes:

$$Cov[Z(u), Z(v)] \;\; = \;\; 1/4 \; - \; \arccos(u'v)/2\pi \; .$$

Note the covariance no longer depends on P and so the limit process is distribution free.

(iv) A maximum depth regression estimator was recently proposed by Rousseeuw and Hubert, and He and Portnoy (1998) present the asymptotic distribution theory for the estimator. They find similar results to those presented here, the limit distribution is determined by the min–max of a function of a Gaussian process. They also show in a simulation study that this limit distribution is close to a normal.

(v) The limit distribution of $\sqrt{n}T_n$ depends in a complex way on P: the distribution enters the covariance structure of Z as well as the deterministic function $u'tp_u(0)$. However, under the assumption of spherical symmetry, confidence regions can feasibly be obtained via simulation, provided a data dependent estimate of $p_0(0)$ is employed. Alternatively, the trimmed convex set of points with sample depth at least α, for some $0 < \alpha < 1/2$, is a confidence region of sorts, but it should be noted that the amount of trimming varies according to the shape of P. See Nolan (1992) for the asymptotic properties of this random convex set of deepest points.

(vi) The proofs are found in the next section, but first here is an explanation of the unusual form of T_n's limit distribution. Rewrite P_n as the sum of a deterministic expectation and a stochastic contribution:

$$P_n H_{ut} \;\; = \;\; PH_{ut} \; + \; \frac{1}{\sqrt{n}} \nu_n H_{ut} \; .$$

Here ν_n is $\sqrt{n}(P_n - P)$. Follow the example of the one-dimensional sample median, and proceed with a one-term Taylor expansion of PH_{ut}, for t in a neighborhood of 0.

$$P_n H_{ut} \;\; = \;\; \frac{1}{2} \; - \; u't \; p_u(0) \; + \; o(u't) \; + \; \frac{1}{\sqrt{n}} \nu_n H_{ut}. \tag{1}$$

The assumption that each projection has a unique median at 0 allows 1/2 to replace PH_{u0} in the above equation. The empirical process ν_n is indexed by half spaces in $\mathbb{R}^d$. Although standard empirical process methodology provides the limit behavior of the process ν_n, it is not adequate for determining the limit behavior of T_n. Rather than simply a maximum of a Gaussian process, this statistic is formed by a max–min of a centered Gaussian process plus the function $-u'tp_u(0)$. Therefore careful application of a continuous mapping theorem is required to obtain T_n's limit distribution. This is a key element to the proof, which is found in the next section. It is also the element that has hindered us from extending the proof to dimensions greater than 2. The problem encountered is outlined at the end of the proof of Lemma 3 in the next section.

4 PROOFS

PROOF OF LEMMA 1: Consistency of T_n is a consequence of the Glivenko-Cantelli lemma for half spaces in $\mathbb{R}^d$ (Gaenssler and Stute, 1979). It says that the empirical measure of a half space converges to its expected measure uniformly over $\mathcal{H}$, the collection of all half spaces in $\mathbb{R}^d$.

Write N_ϵ for $\{t : |t| > \epsilon\}$ and α_o for $\inf_u PH_{u0}$. Let

$$\alpha_\epsilon = \frac{1}{2}(\alpha_o + \sup_{N_\epsilon}\inf_u PH_{ut}).$$

Then

$$\begin{aligned}\mathbb{P}\{|T_n| < \epsilon\} &\geq \mathbb{P}\{\inf_u P_n H_{u0} > \sup_{N_\epsilon}\inf_u P_n H_{ut}\}\\ &\geq \mathbb{P}\{\inf_u P_n H_{u0} > \alpha_\epsilon > \sup_{N_\epsilon}\inf_u P_n H_{ut}\}\\ &\geq \mathbb{P}\{\sup_{\mathcal{H}}|(P_n - P)H_{ut}| < \alpha_o - \alpha_\epsilon\}\end{aligned}$$

The final lower bound converges to 1 by the Glivenko-Cantelli lemma for $\mathcal{H}$.

PROOF OF LEMMA 2: Here we sharpen the consistency result in Lemma 1 to a $\sqrt{n}$ rate of convergence. The argument uses a functional limit theorem for the empirical process and the Taylor-series expansion in (1).

Write q_u for $p_u(0)$. Express $P_n H_{uT_n}$ as in (1),

$$P_n H_{uT_n} = \frac{1}{2} - q_u u'T_n + o(u'T_n) + \frac{1}{\sqrt{n}}\nu_n H_{uT_n}.$$

Empirical process methodology provides the limit theory for the stochastic process ν_n, a random element of $\mathcal{L}^\infty(\mathcal{H})$ the space of real functions on $\mathcal{H}$ that is equipped with the supremum norm. The collection of half spaces is a Vapnik-Cervonenkis class. It easily meets the regularity conditions (Theorem VII.21, Pollard 1984) for convergence of ν_n to a Gaussian process with uniformly continuous sample paths in $\mathcal{L}^\infty(\mathcal{H})$ (uniform continuity is with respect to the $\mathcal{L}^2(P)$ semi-norm on $\mathcal{H}$). In addition, ν_n is stochastically equicontinuous (Lemma VII.15, Pollard 1984). Stochastic equicontinuity of ν_n, along with condition (ii), says that if the random sequence $\{(u, T_n)\}$ converges to $(u, 0)$ in probability then

$$\nu_n(H_{u,T_n} - H_{u0}) \to 0$$

in probability, uniformly in u.

Weak convergence of ν_n and the continuous mapping theorem imply $\sup_{u,t}|\nu_n H_{ut}|$ is $O_p(1)$. Thus

$$\begin{aligned}\inf_u P_n(H_{uT_n}) &= \frac{1}{2} + \inf_u[-q_u u'T_n + o_p(u'T_n) + \frac{1}{\sqrt{n}}\nu_n(H_{uT_n})]\\ &\leq \frac{1}{2} + \inf_u[-q_u u'T_n + o_p(u'T_n)] + \frac{1}{\sqrt{n}}\sup_{u,t}|\nu_n(H_{ut})|\\ &= \frac{1}{2} + \inf_u[-q_u u'T_n + o_p(u'T_n)] + O_p(\frac{1}{\sqrt{n}})\end{aligned}$$

Compare $\inf_u P_n(H_{uT_n})$ with $\inf_u P_n(H_{u0})$.

$$\inf_u P_n(H_{uT_n}) \geq \inf_u P_n H_{u0} = \frac{1}{2} + \inf_u \frac{1}{\sqrt{n}} \nu_n H_{u0}.$$

The process on the right-hand side is $O_p(1/\sqrt{n})$. It follows from the upper bound on $\inf_u P_n H_{uT_n}$ that $\sqrt{n}T_n$ must be $O_p(1)$ provided $0 < \inf_u q_u \leq \sup_u q_u < \infty$. This provision is guaranteed by assumption (ii).

PROOF OF LEMMA 3: Lemma 2 establishes that the estimator is stochastically bounded. So, given $\epsilon > 0$, there is an M_ϵ for which $\sqrt{n}|T_n| < M_\epsilon$ with probability at least $1-\epsilon$, eventually.

$$\begin{aligned} \sqrt{n}T_n &= \sqrt{n} \arg \sup_{\{t:|t|\leq M_\epsilon/\sqrt{n}\}} \inf_u [\frac{1}{2} - q_u u't + o(u't) + \frac{1}{\sqrt{n}}\nu_n H_{u0} + o_p(\frac{1}{\sqrt{n}})] \\ &\approx \arg \sup_{\{t:|t|\leq M_\epsilon\}} \inf_u [-q_u u't + \nu_n H_{u0}] \end{aligned}$$

Stochastic equicontinuity of ν_n implies the o_p term is uniform in u and so can be ignored. The process ν_n converges in distribution to a Gaussian process Z on $\mathcal{H}$ with zero means and covariance structure as stated in the theorem. The behavior of $\sqrt{n}T_n$ is determined by Z on those half spaces that pass through the origin. There is no harm in the simplification that treats Z as a process indexed by u on the unit sphere, rather than by H_{u0} in $\mathcal{H}$. The process in square brackets above converges in distribution to $Z(u) - q_u u't$. If the functional $\arg\sup_t \inf_u [x(u) - q_u u't]$ is continuous at those functions x on which the process Z concentrates then an application of the continuous mapping theorem would give:

$$\sqrt{n}T_n \rightarrow^{\mathcal{D}} \arg\sup_t \inf_u [Z(u) - q_u u't].$$

The functional

$$\arg\sup_t \inf_u [x(u) - q_u u't]$$

is continuous at x, provided

$$\inf_u [x(u) - q_u u't]$$

has a unique maximum. Abbreviate the function $\inf_u [x(u) - q_u u't]$ to $m_x(t)$.

As pointed out in Section 3, in one dimension $m_x(t)$ is uniquely maximized at $m_x(0)$. Thus we need only address the two dimensional case. (In fact, we proceed working in arbitrary dimension, until forced to restrict $d = 2$.)

The argument is based on three properties of the minimizers of $m_x(t)$. The term "minimizer" is coined for those unit vectors v such that

$$m_x(t) = x(v) - q_v v't \tag{2}$$

Of particular interest is the collection of minimizers, $\mathcal{M}_x(t^*)$, of $m_x(t^*)$, where t^* is such that $m_x(t^*) = \sup_t m_x(t)$. The first property to be established is the following:

(A) PROPERTY: If t^* maximizes m_x then there does not exist a unit vector $\tilde{u}$ such that

$$\inf_{u \in \mathcal{M}_x(t^*)} \tilde{u}'u \ > \ 0.$$

In words, the minimizers in $\mathcal{M}_x(t^*)$ are not all contained in some open spherical cap; an open spherical cap is a subset of $\mathbb{S}^{d-1}$ of the form $\mathbb{S}^{d-1} \cap H$, where H is an open half space with boundary passing through the origin.

PROOF of PROPERTY (A). Suppose the contrary, that there exists a unit vector $\tilde{u}$ where $\tilde{u}'u > \epsilon$ for all u in $\mathcal{M}_x(t^*)$. This unit vector can be used to construct $\tilde{t}$ with

$$\inf_u \ x(u) - q_u u'\tilde{t} \ > \ \inf_u x(u) - q_u u' t^*,$$

which contradicts the maximality of t^*.

We now show how to construct such a $\tilde{t}$. Let

$$\delta \ = \ \inf_{\{u : u'\tilde{u} \leq \frac{1}{2}\epsilon\}} x(u) - q_u u' t^* - m_x(t^*).$$

By assumption δ is positive. Also let $\dot{p} = \sup_u q_u$ and $\bar{p} = \inf_u q_u$; both are finite and positive by assumption (ii). Take $\tilde{t}$ to be $t^* - \frac{1}{2}\delta \tilde{u}/\dot{p}$. Then for $u \in \{u : u'\tilde{u} \leq \frac{1}{2}\epsilon\}$,

$$x(u) \ - \ q_u u'\tilde{t} \ > \ m_x(t^*) \ + \ \frac{1}{2}\delta$$

and for $u \in \{u'\tilde{u} > \frac{1}{2}\epsilon\}$

$$x(u) - q_u u'\tilde{t} \ > \ m_x(t^*) \ + \ \frac{1}{4}\delta\epsilon\bar{p}/\dot{p}.$$

This contradicts the fact that t^* maximizes m_x. Property (A) is established.

Next suppose t^*, s^* both maximize m_x and $t^* \neq s^*$. The convex combination $(s^* + t^*)/2$ also maximizes m_x, because for $v \in \mathbb{S}^{d-1}$,

$$\begin{aligned} x(v) - q_v v'\frac{1}{2}(s^* + t^*) \quad &\geq \quad x(v) - q_v \max(v's^*, v't^*) \\ &\geq \quad \inf_u [x(u) - q_u u't^*]. \end{aligned}$$

The first inequality is strict unless $v's^* = v't^*$; the second inequality is strict unless $v \in \mathcal{M}_x(t^*)$ and $v \in \mathcal{M}_x(s^*)$. These two statements imply

(B) PROPERTY: If t^* and s^* maximize m_x then $(s^* + t^*)/2$ maximizes m_x and

$$\mathcal{M}_x(\frac{1}{2}(s^* + t^*)) = \mathcal{M}_x(t^*) \cap \mathcal{M}_x(s^*).$$

Consequently, $v't^* = v's^*$ for $v \in \mathcal{M}_x((s^* + t^*)/2)$. Distinctness of t^* and s^* and Property (A) imply:

(C) PROPERTY: The dimension of the space spanned by $\mathcal{M}_x((s^*+t^*)/2)$ is at least 1 and at most $d-1$.

In light of these findings we are now ready to show that m_Z has a unique maximum, with probability one. In order for both properties (A) and (C) to hold, it must be that $\mathcal{M}_Z((s^*+t^*)/2) = \{v, -v\}$ for some v in $\mathbb{S}^1$. Write r^* for $(s^*+t^*)/2$. From the perfect negative correlation of $Z(v)$ and $Z(-v)$ we arrive at:

$$\begin{aligned} Z(-v) &= \inf_u [Z(u) - q_u u' r^*] \; - \; q_{-v} v' r^* \\ Z(-v) &= -\inf_u [Z(u) - q_u u' r^*] \; - \; q_v v' r^*. \end{aligned}$$

Together, these two equalities imply

$$\inf_u [Z(u) - q_u u' r^*] \; = \; 0,$$

because $q_{-v} = q_v$. This constraint on $m_Z(t^*)$ is very restrictive, we now show that it implies uniqueness of t^*, almost surely.

To show that m_Z has a unique maximum it suffices to show that for any arbitrary pair of points s_0, t_0 there exist neighborhoods $N(t_0)$, $N(s_0)$ such that

$$\mathbb{P}\{\sup_{t \in N(t_0)} m_Z(t) \; = \; \sup_{s \in N(s_0)} m_Z(s) \; = \; \sup_{r \in R} m_Z(r)\} = 0, \tag{3}$$

where R is a compact subset of $\mathbb{R}^2$. So, consider some pair s_0, t_0. For this pair of points, there exists a unit vector u_0 such that $u_0'(s_0 - t_0) = |s_0 - t_0|$. Take $N(t_0)$ and $N(s_0)$ to be balls of radius $\frac{1}{4}|s_0 - t_0|$ about t_0, s_0, respectively. For any $t \in N(t_0)$

$$Z(u) - q_u u' t \; \leq \; Z(u) - q_u u' t_0 + \frac{1}{4}|s_0 - t_0| q_u \; .$$

The same inequality holds if we replace t by s and t_0 by s_0. Therefore, if $t^* \in N(t_0)$ maximizes m_Z and $s^* \in N(s_0)$ maximizes m_Z then

$$\begin{aligned} Z(-u_0) + q_{-u_0} u_0' t_0 &\geq -\frac{1}{4}|s_0 - t_0| q_{-u_0} \; + \; Z(-u_0) + q_{-u_0} u_0' t^* \\ &\geq -\frac{1}{4}|s_0 - t_0| q_{-u_0} \\ \text{and } Z(u_0) - q_{u_0} u_0' s_0 &\geq -\frac{1}{4}|s_0 - t)0| q_{u_0}. \end{aligned}$$

Add the appropriate sides of these two inequalities, recalling $Z(u) = -Z(-u)$ and $q_u = q_{-u}$, to arrive at the contradiction $-|s_0 - t_0| \; \geq \; -\frac{1}{2}|s_0 - t_0|$. Therefore (3) has been established. This completes the proof.

REMARK. Properties A, B, and C do not depend on dimension. The first time that the dimension enters the proof is in the reliance on the fact that these properties force $(t^*+s^*)/2$ to have minimizers v and $-v$, which implies the global supremum is 0. In higher dimensions, unfortunately, these properties do not place the same constraint on the value of the global supremum. Instead, in three dimensions, Property (C) implies that if there are multiple maxima, the dimension of the space spanned by $\mathcal{M}_z((s^*+t^*)/2)$ is either 1 or 2. The first possibility, that the dimension is 1, is covered by the two–dimensional problem. The second possibility implies that there are at least 3 minimizers of $(t^*+s^*)/2$, which fall on a great circle orthogonal to t^*-s^*. Additionally, these minimizers are oriented in such a way that there is no empty arc of length π or greater. An arc is "empty" if it contains no minimizers.

As in the two-dimensional case, consider the event in (3), that both $N(s_0)$ and $N(t_0)$ contain points that achieve the global supremum of m_Z. Write G^* for the great circle orthogonal to t^*-s^* that contains $\mathcal{M}_Z((s^*+t^*)/2)$, and write v^* for the unit vector in the direction t^*-s^*.

By properties (A) and (C), there must exist $v_1, v_2, v_3 \in G^*$ where

$$Z(v_i) - v_i' t^* q_{v_i} \;=\; m_Z(t^*) \qquad i = 1, 2, 3,$$

and where the v_i oriented such that there is no empty arc of length π or more. The v_1, v_2, v_3 can be used to find weights w_1, w_2, $w_3 > 0$ such that $\sum w_i = 1$ and

$$\sum_i w_i v_i' t^* \;=\; \sum_i w_i v_i' s^* \;=\; 0.$$

Then $m_Z(t^*)$ can be expressed as a weighted sum of the $Z(v_i)$:

$$m_Z(t^*) \;=\; \sum_i w_i Z(v_i)/[q_{v_i} \sum_j w_j/q_{v_j}],$$

and a new Gaussian process $Y(\alpha)$, $-1 \le \alpha \le 1$ can be defined,

$$Y(\alpha) \;=\; \sum_i \tilde{w}_i(\alpha) Z(v_i(\alpha)),$$

where $v_i(\alpha) = \alpha v^* + \sqrt{1-\alpha^2}\, v_i$ and

$$\tilde{w}_i(\alpha) \;=\; [w_i/q_{v_i(\alpha)}] / \sum_j [w_j/q_{v_j(\alpha)} \;+\; w_j/q_{v_j(-\alpha)}]\;.$$

Note that for $0 < \alpha \le 1$,

$$\frac{1}{2}Y(\alpha) + \frac{1}{2}Y(-\alpha) > Y(0) + \alpha |s_0 - t_0| \dot{p}/8.$$

We conjecture that the Gaussian process meets this condition with probability 0. To support the conjecture we note that for a v^* that does not depend on the data, the process

$$\frac{1}{2}Y(\alpha) + \frac{1}{2}Y(-\alpha) - Y(0),$$

(which is 0 at $\alpha = 0$) is centered and has nondegenerate increments. Therefore, (Dudley, 1972)

$$\inf_\alpha \frac{1}{2}Y(\alpha) + \frac{1}{2}Y(-\alpha) - Y(0)$$

has a density on $\mathbb{R}^+$. This implies that it has a unique minimum at 0 with probability 0. Therefore condition (3) is met. Unfortunately, v^* does depend on the data, and the argument must be strengthened to hold for any v.

5 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. *J.R. Statist. Soc. Ser. A* **138**:318-344.

Caplin, A. and Nalebuff, B. (1988) On 64%-majority rule. *Econometrica* **56**:787-814.

Caplin, A. and Nalebuff, B. (1991a) Aggregation and social choice: A mean voter theorem. *Econometrica* **59**:1–23.

Caplin, A. and Nalebuff, B. (1991b) Aggregation and imperfect competition: On the existence of equilibrium. *Econometrica* **59**:25–59.

Chamberlin, E. (1937) *The Theory of Monopolistic Competition.* Harvard University Press, Harvard.

Davies, L. (1992) The asymptotic of Rousseeuw's minimum volume ellipsoid. *Ann. Statist.* **20**:1828-1843.

Donoho, D.L. (1982), "Breakdown properties of multivariate location estimators". Ph.D. Qualifying paper, Dept of Statistics, Harvard University.

Donoho, D.L. and Gasko, M. (1992) Breakdown properties of location estimates based on half space depth and projection outlyingness. *Ann. Statist.* **20**:1803–1827.

Dudley, R.M. (1972), Sample functions of the Gaussian process. *Ann. Statist.* **1**:66-103.

Gaenssler, P. and Stute, W. (1979) Empirical processes: a survey of results for independent and identically distributed random variables. *Ann. Statist.* **7**:193-243.

He, X. and Portnoy, S. (1998) Asymptotics of the deepest line. Chapter 5 in *Statistical Inference and Related Topics, A Festschrift in Honor of A.K.Md.E. Saleh* to appear.

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

Hotelling, H. (1929) Stability in competition. *Econom. J.* **39**:41-57.

Huber, P.J. (1985), Projection pursuit. *Ann. Statist.* **13**:435-475.

Kim, J. and Pollard, D. (1990) Cube root asymptotics. *Ann. Statist.* **18**:191-219.

Liu, R. (1990) On a notion of data depth based on random simplices. *Ann. Statist.* **18**:405-414

Maronna, R.A. (1976) Robust M-estimates of multivariate location and scatter. *Ann. Statist.* **4**:51-67.

Nolan, D. (1992) Asymptotics for multivariate trimming. *Stoch. Proc. Appl.* **42**:157–169.

Oja, H. (1983) Descriptive statistics for multivariate distributions. it Statist. and Probab. Letters **1**:327-332.

Oja, H. and Nyblom, J. (1989) Bivariate sign tests. *J. Amer. Statist. Assoc.* **84**:249-259.

Pollard, D. (1984), *Convergence of Stochastic Processes.* Springer-Verlag, New York.

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

Rousseeuw, P.J. (1993) A resampling design for computing high-breakdown regression. *Stat. Prob. Letters* **18**:125-128.

Rousseeuw, P.J. and Leroy, A.M. (1987) *Robust Regression and Outlier Detection.* John Wiley and Sons, New York.

Rousseeuw, P.J. and Ruts, I. (1996) Bivariate location depth. *J. Royal Statist. Soc. Ser. C* **45**:516–526.

Small, C.G. (1987) Measures of centrality for multivariate and directional distributions. *Can. J. Statist.* **15**:31-39.

Small, C.G. (1990) A survey of multidimensional medians. *Int. Statist. Rev.* **58**:263–277.

Stahel, W.A. (1981) "Robuste Schatzungen: Infinitesimale Optimalitat and Schatzungen von Kovarianzmatrizen", Ph.D. thesis, ETH, Zurich.

Titterington, D.M. (1978) Estimation of correlation coefficients by ellipsoidal trimming. *Appl. Statist.* bf 27:227-234.

Tukey, J.W. (1975) Mathematics and the picturing of data. In *Proc. Intl. Congress of Mathematicians*, Vancouver **2**:523-31.