

GENERALIZED SUMS OF DISTANCES

RALPH ALEXANDER

Let K be a compact set in a Euclidean space and let d be a metric on K which is continuous with respect to the usual topology. The generalized energy integral $I(\mu) = \iint d(x, y) d\mu(x) d\mu(y)$ is investigated as μ is allowed to range over the family of signed Borel measures of total mass one concentrated on K . A trick of integral geometry is used to define a class of metrics d , including many standard ones, possessing a number of pleasing properties related to the functional I .

1. Introduction. In his article [4] L. Fejes Tóth discusses problems of the following nature: Suppose K is a compact subset of a Euclidean space and that p_1, \dots, p_n are variable points in K . How can these points be arranged so as to maximize the sum of the $\binom{n}{2}$ distances they determine, and what is the maximum value attained by this sum?

Since these problems are usually very difficult, it is advantageous to look for generalizations which can be attacked. K. B. Stolarsky and the author [1] had considerable success with the following: Let μ vary over all signed Borel measures of total mass one concentrated on K . What can be said about the maximum value of the generalized energy integral $I(\mu) = \iint |x - y| d\mu(x) d\mu(y)$?

If μ is defined by placing mass $1/n$ at each p_i , we see that $n^2 I(\mu) = 2 \sum_{i < j} |p_i - p_j|$. Information concerning the generalized problem has proved most helpful in the study of the original problem, at least for obtaining good estimates. We remark that an explicit solution to the generalized problem is obtained in [1] when K is a finite set.

Using a trick of integral geometry we will define a rather general class of metrics d , including certain of the L^p metrics as well as the great circle metric, which has a number of pleasing properties with respect to the functional $I(\mu) = \iint d(x, y) d\mu(x) d\mu(y)$. Concavity, given by $2I((\mu_1 + \mu_2)/2) \geq I(\mu_1) + I(\mu_2)$, is probably the most useful of these.

Each metric is defined in terms of a measure on hyperplanes, and it turns out that $I(\mu)$ possesses a nice representation as an integral with respect to this measure. This allows an easy derivation of the basic properties of I . As is often the case, these are more apparent in the general situation than in special cases where it is not clear what is important.

We use some standard terminology and notation from measure theory. For example, if μ is a signed measure, $|\mu|$ will denote $\mu^+ + \mu^-$ where $\mu^+ - \mu^-$ is the Jordan decomposition of μ . The book by Halmos [5] contains a thorough discussion of such topics.

2. A definition and a general example. By a Borel set of hyperplanes, or $(m-1)$ -flats, in E^m we mean a collection of hyperplanes whose dual is a Borel set in P^m . However, we are thinking of the hyperplanes as "points". Let η be a measure on the Borel sets of hyperplanes (t denotes a hyperplane), and let K be a fixed compact set in E^m . We make the following basic assumption:

- (1) The η -measure of planes containing points in the convex hull of K is finite
- (2) If p is any point in K , $\eta\{t: p \in t\} = 0$.
- (3) If p, q in K are distinct. $\eta\{t: t \cap \overline{pq} \neq \emptyset\} > 0$.

DEFINITION 2.1. If p, q are points of K , define $d(p, q)$ to be $\eta\{t: t \cap \overline{pq} \neq \emptyset\}$.

PROPOSITION 2.1. *The function d is a metric on K which is continuous with respect to the Euclidean topology.*

Proof. Let p, q, r be points in K . Properties (1), (2), and (3) imply that $d(p, p) = 0$ and that $d(p, q)$ is a positive real number if p and q are distinct. Since any hyperplane which cuts pq must cut at least one of the segments pr, rq , the subadditivity of η immediately implies that $d(p, q) \leq d(p, r) + d(r, q)$.

Suppose K^* is the convex hull of K and $B(x, \delta)$ is the set of points in K^* within Euclidean distance δ of x . Let $\{p_i\}$ be a convergent sequence in K with limit p_0 . Observe that $d(p_0, p_i) \leq \eta\{t: t \cap B(p_0, |p_0 - p_i|) \neq \emptyset\} < \infty$. Also, since $\bigcap_i B(p_0, |p_0 - p_i|) = \{p_0\}$, property (2) together with "continuity from above" of η imply that $\lim_{i \rightarrow \infty} d(p_0, p_i) = 0$. The continuity of d as a function on $K \times K$ follows easily as a consequence of the triangle inequality.

Now we give a reasonably general construction of a measure on the hyperplanes of E^m . Let $l(u)$ be the directed line through the origin containing the unit vector u , and let m_u be a Borel measure on $l(u)$; ρ will be a finite Borel measure on the surface of the unit ball. If T is a family of hyperplanes, let $T(u)$ denote those members of T which are orthogonal to u . Put $h(u) = m_u(T(u) \cap l(u))$. Finally, we define $\eta(T) = \int h(u) d\rho(u)$. Whether or not η satisfies our three basic conditions is a separate problem depending on the choice of K .

Next we give four special cases of this construction to serve as examples and for future reference.

(1) Let K be an arbitrary compact subset of E^m . If for each u , m_u is Lebesgue measure and ρ is σ , the usual surface measure of the ball, then $d(p, q) = C \int |\langle p - q, u \rangle| d\sigma(u)$ for a suitable constant C . Hence d is essentially the Euclidean metric.

(2) If in example (1) we replace σ by m atoms of weight one concentrated on $v_1 = (1, 0, \dots, 0), \dots, v_m = (0, \dots, 0, 1)$, we see that $d(p, q) = \sum_i |\langle p - q, v_i \rangle|$. Hence d is the L^1 metric on E^m .

It is known that for $1 \leq p \leq 2$, the L^p metric on E^m may be represented as $\int |\langle p - q, u \rangle| d\rho(u)$ for a suitable measure ρ on the surface of the unit ball. See Bolker's interesting article [3] for further discussion.

For our next two examples we let K be the spherical surface of radius r centered at the origin.

(3) Let m_u consist of a single atom of weight one concentrated at the origin and let $\rho = \sigma$. Here it is seen that d is a constant multiple of the great circle metric on K .

(4) Let m_u consist of a single atom of weight one concentrated at $r_0 u$ where $0 < r_0 < r$ and let $\rho = \sigma$. We do not know if these metrics have received attention in the literature. Various interesting metrics on the sphere K , including the usual Euclidean metric, may be expressed as weighted averages of these metrics as r_0 varies from 0 to r . Also, it is clear that as r_0 tends to zero, the great circle metric is obtained.

Even our general construction is far from being all-inclusive. Indeed if η' is any centrally symmetric Borel measure on the surface of the $(m + 1)$ -ball, the principle of duality gives a corresponding Borel measure on the hyperplanes of m -space.

3. Basic properties of the functional $I(\mu)$. Let A_t and B_t be the two open half-spaces determined by the hyperplane t . We define $v(t)$ to be $\mu(A_t)\mu(B_t)$. Our next result gives the fundamental integral expression for I .

PROPOSITION 3.1. *We have $I(\mu) = 2 \int v(t) d\eta(t)$.*

Proof. We begin by noting that $|v(t)| \leq \frac{1}{4} [\mu(K)]^2$ for any t . Hence v is an integrable function.

Let χ_{xy} be the characteristic function of the open segment from x to y (excluding x and y). Since $\eta(t: x \in t \text{ or } y \in t) = 0$, $d(x, y) = \int \chi_{xy}(t \cap xy) d\eta(t)$.

Thus

$$I(\mu) = \int \int \int \chi_{xy}(t \cap \overline{xy}) d\eta(t) d\mu(x) d\mu(y).$$

We now may apply Fubini's theorem to move the η -integral to the outside. We consider the integral

$$\int \int \chi_{xy}(t \cap \overline{xy}) d\mu(x) d\mu(y).$$

Since $\chi_{xy}(t \cap \overline{xy}) = 0$ unless $x \in A_t$, $y \in B_t$ or $x \in B_t$, $y \in A_t$, we can evaluate the integral directly as

$$2(\mu \times \mu)\{A_t \times B_t\} = 2\mu(A_t)\mu(B_t) = 2v(t).$$

The result follows.

LEMMA 3.1. *Let $H = \{t: |\mu| (t \cap K) \neq 0\}$. Then $\eta(H) = 0$.*

Proof. Let χ_x be the characteristic function of the point x so that $|\mu|(t \cap K) = \int \chi_x(t \cap x) d|\mu|(x)$. We may again apply Fubini's theorem to assert

$$\int \int \chi_x(t \cap x) d|\mu|(x) d\eta(t) = \int \int \chi_x(t \cap x) d\eta(t) d|\mu|(x).$$

The inner integral on the right is zero by property (2) of the measure η , and the result follows.

Because of Lemma 3.1 we may assume that the integral of Theorem 3.1 is taken only over those hyperplanes which intersect K in a set of μ -measure zero. This lemma also implies that $\mu(A_t) + \mu(B_t) = 1$ for almost all t .

PROPOSITION 3.2. *For any μ we have $I(\mu) \leq \frac{1}{2} \eta(T)$ where T is the set of hyperplanes having nonempty intersection with the convex hull K^* .*

Proof. It is clear that $v(t) = 0$, if $t \cap K^* \neq 0$ since μ is supported by K . For almost all t $\mu(A_t) + \mu(B_t) = 1$; whence $v(t) \leq \frac{1}{4}$, and $2 \int v(t) d\eta(t) \leq \frac{1}{2} \eta(T)$ for any μ . This completes the proof.

For the special case of the Euclidean metric, it was established in [1] that $I(\mu)$ was bounded above using indirect arguments involving metric embeddings in the sequence space l^2 .

Next, let μ_1 and μ_2 be two Borel measures of the type we are considering which have common support K . For almost all t , $\mu_1(A_t) + \mu_1(B_t) = \mu_2(A_t) + \mu_2(B_t) = 1$. Hence

$$|\mu_1(A_t) - \mu_2(A_t)| = |\mu_1(B_t) - \mu_2(B_t)|;$$

call this number $C(t)$. We define a variance by

$$\text{Var}(\mu_1, \mu_2) = \int C(t)^2 d\eta(t).$$

PROPOSITION 3.3. *Let μ_1 and μ_2 be signed Borel measures of total mass 1 supported by K . Then*

$$2I((\mu_1 + \mu_2)/2) - [I(\mu_1) + I(\mu_2)] = \text{Var}(\mu_1, \mu_2).$$

Proof. Let $\mu_1(A_t) = k_1$, $\mu_2(A_t) = k_2$. For almost all t $\mu_1(B_t) = 1 - k_1$, $\mu_2(B_t) = 1 - k_2$. Applying Proposition 3.1, the left side of our equation may be written

$$4 \int \left[\frac{k_1 + k_2}{2} \left(1 - \frac{k_1 + k_2}{2} \right) - \frac{1}{2} k_1(1 - k_1) - \frac{1}{2} k_2(1 - k_2) \right] d\eta(t).$$

The integrand immediately simplifies to $C(t)^2/4$.

It is natural to call I strictly concave if $\text{Var}(\mu_1, \mu_2) > 0$, whenever $\mu_1 \neq \mu_2$. It can be shown that I is strictly concave when d is the Euclidean metric. If d is the great circle metric and μ_1, μ_2 are centrally symmetric measures on a sphere, we see that $\text{Var}(\mu_1, \mu_2) = 0$, and hence I is not strictly concave in this situation.

The question whether I is strictly concave in the case of the metric defined by example (4) is interesting. Suppose K is a circle of radius r and K' is the concentric circle of radius r_0 . A line tangent to K' will divide K into two arcs of length l_1 and l_2 . A technical but straightforward argument shows that I is strictly concave if and only if l_1/l_2 is irrational. For spheres of higher dimension it seems "intuitive" that I will always be strictly concave. However, a detailed proof could be involved.

The final result of this section is obtained by combining our results with the Menger-Schoenberg theory of metric embedding. The essential features of the proof are to be found in [1] where the special case of the Euclidean metric is treated. Although we regard this as a very interesting and useful theorem, we will omit the proof.

PROPOSITION 3.4. *Let the metric d on the compact set K arise from a measure on hyperplanes. Then the metric space $(K, d^{\frac{1}{2}})$ can be isometrically embedded on (the surface of) a Hilbert sphere. Furthermore, the number $\sup_{\mu} I(\mu)$ is equal to $2\rho_0^2$ where ρ_0 is the least radius for which such an embedding is possible.*

We remark that Proposition 3.2 assures us that ρ_0 is finite. In [1] it was proved by explicit construction that the metric space $([0, 1], d^{\frac{1}{2}})$, where d is the Euclidean metric on the line, can be isometrically embedded on a Hilbert sphere of radius $\frac{1}{2}$. Since $d^{\frac{1}{2}}(0, 1) = 1$, this shows that $\rho_0 = \frac{1}{2}$. Now, however, we observe that the idea used in the proof of Proposition 3.2, taking points as hyperplanes and μ to be Lebesgue measure, shows that I achieves a unique maximal value when μ consists of atoms of weight $\frac{1}{2}$ concentrated at zero and one. In this case $I(\mu) = \frac{1}{2}$, and Proposition 3.4 allows us to deduce that $\rho_0 = \frac{1}{2}$.

4. Brief discussion of applications to extremal problems.

PROPOSITION 4.1. *Let K be a finite set in a Euclidean space, and suppose that the metric d arises from a measure on hyperplanes. Then the number $\sup_{\mu} I(\mu)$ may be explicitly computed.*

Proof. By Proposition 3.4 the finite metric space $(K, d^{\frac{1}{2}})$ can be isometrically embedded on a Euclidean sphere. Clearly ρ_0 is the radius of the sphere of minimal dimension. Standard methods may be employed for computing the radius of the circumsphere of the simplex.

If I is strictly concave, it is easy to show that $(K, d^{\frac{1}{2}})$ embeds as a nondegenerate simplex in E^m where $m + 1$ is the number of points in K . In this case the computation of ρ_0 is straightforward.

PROPOSITION 4.2. *Let K be a Euclidean sphere and d arise from a measure on hyperplanes. Furthermore suppose that d is invariant under orthogonal transformations. Then $I(\mu) \leq I(\sigma)$ where σ is the normalized surface measure of K .*

Proof. Suppose that I is strictly concave. If $\mu \neq \sigma$, there will be an orthogonal transformation τ such that $\tau\mu \neq \mu$. Since $I(\tau\mu) = I(\mu)$, $I(\mu) < I(\frac{1}{2}(\mu + \tau\mu))$. Thus $I(\mu)$ is uniquely maximal when $\mu = \sigma$.

If I is not strictly concave, the functional I_{λ} associated with the metric d^{λ} will be, if $0 < \lambda < 1$. Letting $\lambda \rightarrow 1$ yields the result. The ideas of Schoenberg [9] may be employed to prove the strict concavity of I_{λ} .

We observe that if d is the great circle metric on K , then $I(\mu) = \frac{1}{2}\pi r$ for any centrally symmetric μ . The papers of Sperling [10] and Nielson [8] treat this case when μ consists of n atoms of weight $1/n$.

Next suppose K is an arbitrary compact set. Except when K is finite, we have not dealt with the problem of whether there generally exists an extremal signed measure μ_0 such that $I(\mu_0) = \sup_{\mu} I(\mu)$. The difficulty, of course, is that signed measures are not weakly

compact. Yet, all evidence indicates that μ_0 will always exist if d arises from a measure on hyperplanes. It may be that this issue can be settled by an existing theorem in the vast literature on potential theory. Since the positive measures of mass one are weakly compact, we record the following proposition. The case where d is the Euclidean metric is due to Björck [2].

PROPOSITION 4.3. *Let K be a compact set and let d arise from a measure on hyperplanes in such a manner that the functional I is strictly concave. Then as μ varies over positive Borel measures of mass one concentrated on K , there will be a unique μ_0 such that $I(\mu_0) = \sup_{\mu} I(\mu)$.*

5. The work of J. B. Kelly. Kelly [6], [7] has done some interesting work which is closely related to the results of this paper. He calls a semimetric d " n -hypermetric" if given any $2n + 1$ points in the space $p_1, \dots, p_n, q_1, \dots, q_{n+1}$, we always have

$$\sum_{i < j} d(p_i, p_j) + \sum_{i < j} d(q_i, q_j) \leq \sum_{i,j} d(p_i, q_j).$$

Thus an ordinary metric is at least 1-hypermetric. We note here that a slight modification of the proof of Proposition 3.3 shows that if d arises from a measure on hyperplanes, then d is n -hypermetric for all n . Kelly's papers contain many important results concerning hypermetrics in general.

We close with an obvious question. Let d be a metric on the Euclidean space E^m which is continuous with respect to the usual topology. When does d arise from a measure on the hyperplanes of E^m ? Two conditions must be met: (1) If p, q, r are collinear in the given order, then $d(p, r) = d(p, q) + d(q, r)$; (2) d must be n -hypermetric for all n . Are they sufficient?

It is not hard to show that any Minkowski metric on the plane arises from a measure on the lines; see [3]. So, we would like to see at least one example of a plane metric satisfying condition (1) but not condition (2).

We wish to thank the referee for a number of comments which improved the clarity of this article. Also, we should point out the results of K. B. Stolarsky [11] which deal with the case where K is a Euclidean sphere. For spheres of higher dimension his estimates of distance sums are by far the best known.

REFERENCES

1. R. Alexander and K. B. Stolarsky, *Some extremal problems of distance geometry*, (to appear in Trans. Amer. Math. Soc.).

2. G. Björck, *Distributions of positive mass which maximize a certain generalized energy integral*, Ark. Mat., **3** (1955), 255–269. MR 17, # 1198.
3. E. D. Bolker, *A class of convex bodies*, Trans. Amer. Math. Soc., **145** (1969), 323–345.
4. L. Fejes Tóth, *On the sum of distances determined by a pointset*, Acta Math. Acad. Sci. Hungar., **7** (1956), 397–401.
5. P. R. Halmos, *Measure Theory*, Van Nostrand, Princeton, 1950.
6. J. B. Kelly, *Metric inequalities and symmetric differences*, Inequalities II (Oved Shisha, ed.), Academic Press, New York, 1970, 193–212.
7. ———, *Hypermetric spaces and metric transforms*, Inequalities III (Oved Shisha, ed.), Academic Press, New York, 1972, 149–158.
8. F. Nielson, *Om summen af afstandene mellem n punkter på en Kugleflade*, Nordisk Mat. Tidsskr., **13** (1965), 45–50.
9. I. J. Schoenberg, *On certain metric spaces arising from Euclidean spaces by change of metric and their imbedding in Hilbert space*, Ann. of Math., **38** (1937), 787–793.
10. G. Sperling, *Lösung einer elementargeometrischen Frage von Fejes Tóth*, Arch. Math., **11** (1960), 69–71.
11. K. B. Stolarsky, *Sums of distances between points on a sphere. II*, Proc. Amer. Math. Soc., **41** (1973), 575–582.

Received December 13, 1973 and in revised form September 17, 1974.

UNIVERSITY OF ILLINOIS