Pacific
Journal of
Mathematics

STRONG NEGATIVE TYPE IN SPHERES

Russell Lyons

Volume 307 No. 2 August 2020

PACIFIC JOURNAL OF MATHEMATICS
Vol. 307, No. 2, 2020
https://doi.org/10.2140/pjm.2020.307.383

STRONG NEGATIVE TYPE IN SPHERES

RUSSELL LYONS

It is known that spheres have negative type, but only subsets with at most one pair of antipodal points have strict negative type. These are conditions on the (angular) distances within any finite subset of points. We show that subsets with at most one pair of antipodal points have strong negative type, a condition on every probability distribution of points. This implies that the function of expected distances to points determines uniquely the probability measure on such a set. It also implies that the distance covariance test for stochastic independence, introduced by Székely, Rizzo and Bakirov, is consistent against all alternatives in such sets. Similarly, it allows tests of goodness of fit, equality of distributions, and hierarchical clustering with angular distances. We prove this by showing an analogue of the Cramér–Wold theorem.

1. Introduction

We introduce the topic by borrowing from [Lyons 2014].

Let (X, d) be a metric space. One says that (X, d) has *negative type* if for all $n \geq 1$ and all lists of n red points x_i and n blue points x'_i in X, the sum $2\sum_{i,j} d(x_i, x'_j)$ of the distances between the $2n^2$ ordered pairs of points of opposite color is at least the sum $\sum_{i,j}(d(x_i, x_j) + d(x'_i, x'_j))$ of the distances between the $2n^2$ ordered pairs of points of the same color. It is not obvious that euclidean space has this property, but it is well known. By considering repetitions of x_i and taking limits, we arrive at a superficially more general property: For all $n \geq 1$, $x_1, \ldots, x_n \in X$, and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ with $\sum_{i=1}^n \alpha_i = 0$, we have

$$\sum_{i,j \leq n} \alpha_i \alpha_j d(x_i, x_j) \leq 0. \tag{1-1}$$

We say that (X, d) has *strict negative type* if, for every n and all n-tuples of distinct points $x_1, \ldots, x_n$, equality holds in (1-1) only when $\alpha_i = 0$ for all i. Again, euclidean

Research partially supported by NSF grant DMS-1612363.
MSC2010: primary 44A12, 45Q05, 51K99, 51M10; secondary 62G20, 62H15, 62H20, 62H30.
Keywords: Cramér–Wold, hemispheres, expected distances, distance covariance, equality of distributions, goodness of fit, hierarchical clustering.

spaces have strict negative type. A simple example of a metric space of nonstrict negative type is ℓ^1 on a 2-point space, i.e., $\mathbb{R}^2$ with the ℓ^1-metric.

We say that a (Borel) probability measure μ on X has *finite first moment* if $\int d(o,x)\,\mathrm{d}\mu(x) < \infty$ for some (hence all) $o \in X$; we write $P_1(X,d)$ for the set of such probability measures. Suppose that $\mu_1, \mu_2 \in P_1(X,d)$. By approximating μ_i by probability measures of finite support, we obtain a yet more general property, namely, that when X has negative type,

$$\int d(x_1,x_2)\,\mathrm{d}(\mu_1-\mu_2)^2(x_1,x_2) \le 0. \tag{1-2}$$

We say that (X,d) has *strong negative type* if it has negative type and equality holds in (1-2) only when $\mu_1=\mu_2$. See [Lyons 2018] for an example of a (countable) metric space of strict but not strong negative type. The notion of strong negative type was first defined by Zinger, Kakosyan and Klebanov [1992]. Lyons [2013] used it to show that a metric space X has strong negative type if and only if the theory of distance covariance holds in X just as in euclidean spaces, as introduced by Székely, Rizzo and Bakirov [2007]. Similarly, it allows tests of goodness of fit, equality of distributions, and hierarchical clustering with angular distances: see the review in [Székely and Rizzo 2017]. Lyons [2013] noted that if (X,d) has negative type, then (X,d^r) has strong negative type when $0<r<1$.

Define

$$a_\mu(x) := \int d(x,x')\,\mathrm{d}\mu(x')$$

for $x \in X$ and $\mu \in P_1(X,d)$. Lyons [2013] remarked that if (X,d) has negative type, then the map $\alpha\colon \mu \mapsto a_\mu$ is injective on $\mu \in P_1(X)$ if and only if X has strong negative type. (There are also metric spaces not of negative type for which α is injective.)

A list of metric spaces of negative type appears as Theorem 3.6 of [Meckes 2013]. All euclidean spaces have strong negative type; see [Lyons 2013] for a discussion of various proofs.

That real and complex hyperbolic spaces $\mathbb{H}^n$ have negative type was shown in [Gangolli 1967, Section 4], and was made explicit in [Faraut and Harzallah 1974, Corollary 7.4]; that they have strict negative type was shown by Hjorth, Kokkendorff and Markvorsen [Hjorth et al. 2002]. Lyons [2014] showed that real hyperbolic spaces have strong negative type. The remaining constant-curvature, simply connected spaces are spheres.

Let S^n denote the unit-radius sphere centered at the origin of $\mathbb{R}^{n+1}$. Although spheres have negative type (in their intrinsic metric), not even circles have strict negative type. For example, in S^1, take two red points $\{(1,0),(-1,0)\}$ and two blue points $\{(0,1),(0,-1)\}$. Nevertheless, antipodal symmetry is the only obstruction to strict negative type: the main result, Theorem 9.1, of [Hjorth et al. 1998] is that a subset of a sphere has strict negative type if and only if that subset contains at most one pair of antipodal points. We strengthen this to strong negative type:

Theorem 1.1. *If $B \subset S^n$ contains at most one pair of antipodal points, then B has strong negative type.*

We begin by proving a special case:

Theorem 1.2. *If $H \subset S^n$ is an open hemisphere, then H has strong negative type.*

We may parametrize open hemispheres as

$$H_t := \{x \in S^n : t \cdot x > 0\}$$

for $t \in S^n$. A crucial ingredient in the proof of Theorem 1.2 is an analogue of the Cramér–Wold theorem:

Theorem 1.3. *Let H be an open hemisphere in an n-dimensional sphere, S^n. For a finite signed measure μ on H and $t \in S^n$, define $b_\mu(t) := \mu(H \cap H_t)$. The map $\mu \mapsto b_\mu$ is injective. Moreover, if D is a dense subset of S^n, then $\mu \mapsto b_\mu \restriction D$ is injective.*

Let $R\colon x \mapsto -x$ be the reflection in the origin. If K is a Borel subset of S^n such that K and its image under R partition S^n and such that the interior of K is a hemisphere, then call K a *partitioning hemisphere*. Given a probability measure μ on S^n, let μ_R denote the maximal measure that is invariant under R and such that $\mu_R \le \mu$. Note that if μ is a probability measure on S^n that is invariant under R, then $\mu(K) = 1/2$ for every partitioning hemisphere, K. Therefore, for every probability measure μ on S^n with $\mu_R \neq 0$, there is a probability measure $\nu \neq \mu$ such that $\mu(K) = \nu(K)$ for every partitioning hemisphere, K. Moreover, if $\mu_R \neq 0$ but $\mu_R(A) = 0$ for every $(n-1)$-dimensional great sphere A in S^n, then there is another probability measure $\nu \neq \mu$ such that $\mu(H) = \nu(H)$ for every open hemisphere, H. We extend Theorem 1.3 to show the converse, which we use to prove Theorem 1.1:

Theorem 1.4. *Let μ be a probability measure on S^n such that $\mu_R = 0$. If ν is a probability on S^n such that $\mu(K) = \nu(K)$ for every partitioning hemisphere, K, then $\nu = \mu$. Similarly, if ν is a probability on S^n such that $\mu(H_t) = \nu(H_t)$ for every t belonging to a dense subset D of S^n, then $\nu = \mu$.*

The last assertion of Theorem 1.4 is essentially known: see, e.g., Lemmas 2.3 and 2.4 of [Rubin 1999].

We also have the following fact:

Proposition 1.5. *If μ is a probability measure on S^n such that $\mu(K) = 1/2$ for every partitioning hemisphere, K, then μ is R-invariant. Similarly, if μ is a probability measure on S^n such that $\mu(H_t) = \mu(H_{-t})$ for a dense set of t, then μ is R-invariant.*

The last assertion is again essentially known: see [Schneider 1970, Korollar 3.2]. The work of [Rubin 1999] and [Schneider 1970], as well as other authors who study related questions, uses spherical harmonics. This is a powerful tool that leads to more general results, although those extensions do not seem relevant to negative type. We give elementary proofs that rely only on the Cramér–Wold theorem for euclidean spaces:

Theorem 1.6. *If μ is a complex Borel measure on $\mathbb{R}^n$ such that $\mu(H) = 0$ for every open halfspace in $\mathbb{R}^n$, then $\mu = 0$.*

Proof. For $t \in S^{n-1}$, define μ_t on $\mathbb{R}$ by

$$\mu_t(-\infty, a) := \mu\{x \in \mathbb{R}^n : \ t \cdot x < a\} \quad (a \in \mathbb{R}).$$

Then $\mu_t = 0$, whence its Fourier transform $\widehat{\mu}_t$ satisfies $\widehat{\mu}_t(b) = 0$ for all $b \in \mathbb{R}$. Because $\widehat{\mu}_t(b) = \widehat{\mu}(bt)$, it follows that the Fourier transform of μ also vanishes, whence so does μ. $\square$

Even this theorem can be proved without Fourier analysis — see [Walther 1997] or [Lyons and Zumbrun 2018].

2. Proofs

Proof of Theorem 1.3. By the bounded convergence theorem, $b_\mu \restriction D$ determines $b_\mu(t)$ for all t such that $\mu(\partial(H \cap H_t)) = 0$, and therefore $b_\mu \restriction D$ determines all of b_μ by continuity from below: for every $t \in S^n$, there are $s_k \in S^n$ such that $\mu(\partial(H \cap H_{s_k})) = 0$ and $H \cap H_{s_k}$ increase to $H \cap H_t$.

We may take H to be the upper open hemisphere, $\{(t_1, t_2, \ldots, t_{n+1}) \in S^n : t_{n+1} > 0\}$. Define $\phi \colon H \to \mathbb{R}^{n+1}$ by

$$\phi(t_1, \ldots, t_{n+1}) := (t_1/t_{n+1}, \ldots, t_n/t_{n+1}, 1).$$

Then ϕ is a homeomorphism from H to the affine hyperplane

$$H' := \{(t_1, t_2, \ldots, t_{n+1}) \in \mathbb{R}^{n+1} : \ t_{n+1} = 1\},$$

namely, $\phi(t)$ is the intersection of H' with the line through the origin and t. Furthermore, ϕ maps $H \cap H_t$ to an open halfspace in H' and every open halfspace in H' is the image under ϕ of some $H \cap H_t$. Therefore, b_μ determines the measures of all open halfspaces with respect to the pushforward $\phi_*\mu$ on H'. The classical theorem of Cramér and Wold applied to H' shows that this determines $\phi_*\mu$, which in turn determines μ. $\square$

Proof of Theorem 1.2. Write σ for the volume measure on S^n normalized to have mass π. Then for all $x_1, x_2 \in S^n$, we have

$$d(x_1, x_2) = \int |\mathbf{1}_{H_t}(x_1) - \mathbf{1}_{H_t}(x_2)|^2 \, \mathrm{d}\sigma(t).$$

This well-known fact is easy to see: By rotation-invariance of σ, the right-hand side depends only on $d(x_1, x_2)$. By considering three points on a great circle, we find that the dependence is linear. Finally, by taking antipodal points, we verify that the constant of linearity is 1.

Therefore, if μ_1 and μ_2 are probabilities on S^n, we may write

$$\int d(x_1, x_2)\, \mathrm{d}(\mu_1 - \mu_2)^2(x_1, x_2) \\ = \iint |\mathbf{1}_{H_t}(x_1) - \mathbf{1}_{H_t}(x_2)|^2\, \mathrm{d}(\mu_1 - \mu_2)^2(x_1, x_2)\, \mathrm{d}\sigma(t).$$

Expanding the square in the integrand and using the facts that

$$\int \mathbf{1}_{H_t}(x)\, \mathrm{d}\nu^2(x, y) = \nu(H_t)\nu(S^n)$$

and

$$\int \mathbf{1}_{H_t}(x)\mathbf{1}_{H_t}(y)\, \mathrm{d}\nu^2(x, y) = \nu(H_t)^2$$

for any finite signed measure, ν, we obtain that

$$\int d(x_1, x_2)\, \mathrm{d}(\mu_1 - \mu_2)^2(x_1, x_2) = -2 \int (\mu_1(H_t) - \mu_2(H_t))^2\, \mathrm{d}\sigma(t).$$

It is evident from this that (S^n, d) has negative type. In order to prove (H, d) has strong negative type, it suffices to show that if μ_1 and μ_2 are concentrated on H and satisfy $\mu_1(H_t) = \mu_2(H_t)$ for σ-a.e. t, then $\mu_1 = \mu_2$. But this is immediate from Theorem 1.3. □

Given any signed measure θ, define the antisymmetric measure $\bar{\theta} := \theta - R_* \theta$, where $R_* \theta$ is the pushforward of θ by R. For positive θ with $\theta_R = 0$, we have $\theta = \bar{\theta}^+$, the positive part of $\bar{\theta}$. For positive θ without assuming that $\theta_R = 0$, we have

$$2\,\theta(S^n) \ge |\bar{\theta}|(S^n), \text{ with equality if and only if } \theta_R = 0. \tag{2-1}$$

Lemma 2.1. *Let μ and ν be probability measures on S^n. If $\mu(K) = \nu(K)$ for every partitioning hemisphere, K, then $\bar{\mu} = \bar{\nu}$. Similarly, if $\mu(H_t) = \nu(H_t)$ for every $t \in D$, where D is a dense subset of S^n, then $\bar{\mu} = \bar{\nu}$.*

Proof. We claim that there is an $(n-1)$-dimensional great sphere A in S^n with $\mu(A) = \nu(A) = 0$. To see this, we build A inductively by dimension. First, because only countably many points have positive mass, there is a pair A_0 of antipodal points with $\mu(A_0) = \nu(A_0) = 0$. Second, all uncountably many 1-dimensional great spheres in S^n that contain A_0 have pairwise intersections exactly A_0, whence there is a 1-dimensional great sphere $A_1 \supset A_0$ with $\mu(A_1) = \nu(A_1) = 0$. We may continue this procedure recursively, finding a k-dimensional great sphere $A_k \supset A_{k-1}$ for $1 \le k \le n-1$ with $\mu(A_k) = \nu(A_k) = 0$. Finally, take $A := A_{n-1}$.

Let H be one of the two open hemispheres comprising $S^n \setminus A$. Note that $\mu(H) = \nu(H)$ under either assumption (in the second case, we use a continuity argument like that at the start of the proof of Theorem 1.3).

Let K be a partitioning hemisphere. Because $\mu(A) = 0$ and $\mu(S^n) = 1$, we have

$$\begin{aligned}\bar{\mu}(H \cap K) &= \mu(H \cap K) + \mu(H \cap RK) - \mu(H \cap RK) - \mu(RH \cap RK) \\ &= \mu(H) - \mu(RK) = \mu(H) + \mu(K) - 1.\end{aligned}$$

A similar equation holds for ν. Hence, the assumption that $\mu(K) = \nu(K)$ for every partitioning hemisphere, K, yields

$$\bar{\mu}(H \cap K) = \bar{\nu}(H \cap K)$$

for every such K.

Now every set $H \cap H_t$ is of the form $H \cap K$ for some partitioning hemisphere, K. It follows that

$$\bar{\mu}(H \cap H_t) = \bar{\nu}(H \cap H_t)$$

for every t, whence by Theorem 1.3, it follows that $\bar{\mu} = \bar{\nu}$.

We now prove the second assertion of the lemma. Note that in the preceding proof, we did not use the full strength of the assumption that $\mu(K) = \nu(K)$ for every partitioning hemisphere, K, but only that $\mu(K_t) = \nu(K_t)$ for partitioning hemispheres K_t satisfying $K_t \cap H = H_t \cap H$ for $t \in D$; we may also require that $K_t \cap RH = \overline{H_t} \cap RH$. Let u be such that $H_u = H$, and let s_k be on the geodesic segment from u to t with $s_k \notin \{u, t\}$, $s_k \to t$, and $\mu(\partial H_{s_k}) = \nu(\partial H_{s_k}) = 0$. (Such s_k exist because $\mu(A) = \nu(A) = 0$.) Let $t_{k,j} \in D$ converge to s_k as $j \to \infty$. By the bounded convergence theorem, $\lim_{j\to\infty} \mu(H_{t_{k,j}}) = \mu(H_{s_k})$ and similarly for ν, whence $\mu(H_{s_k}) = \nu(H_{s_k})$. In addition, we have $\lim_{k\to\infty} \mu(H_{s_k}) = \mu(K_t)$ and similarly for ν. Hence, $\mu(K_t) = \nu(K_t)$ for every $t \in D$, whence $\bar{\mu} = \bar{\nu}$. □

Proof of Theorem 1.4. By Lemma 2.1, either assumption implies that $\bar{\mu} = \bar{\nu}$. We may conclude from (2-1) that $2 = 2\,\nu(S^n) \geq |\bar{\nu}|(S^n) = |\bar{\mu}|(S^n) = 2\,\mu(S^n) = 2$, and hence, again from (2-1), that $\nu_R = 0$. Since also $\mu_R = 0$, we obtain the desired conclusion, $\mu = \bar{\mu}^+ = \bar{\nu}^+ = \nu$. □

Proof of Theorem 1.1. If B contains no antipodal points, then every μ concentrated on B has $\mu_R = 0$, whence the proof that B has strong negative type is exactly as for Theorem 1.2, using Theorem 1.4 in place of Theorem 1.3.

If B contains one antipodal pair, $\{x, Rx\}$, then it still suffices to show that for probabilities μ and ν concentrated on B, the assumption $\mu(H_t) = \nu(H_t)$ for a dense set of t implies $\mu = \nu$. By Lemma 2.1, such an assumption yields $\bar{\mu} = \bar{\nu}$. Because $\bar{\mu} = \overline{\mu - \mu_R}$ and $\mu - \mu_R$ is a positive measure with $(\mu - \mu_R)_R = 0$, and similarly for ν, we obtain $\overline{\mu - \mu_R} = \overline{\nu - \nu_R}$ and $\mu - \mu_R = \overline{\mu - \mu_R}^+ = \overline{\nu - \nu_R}^+ = \nu - \nu_R$.

Therefore, $\mu_R(H_t) = \nu_R(H_t)$ for a dense set of t. Because μ_R and ν_R are supported by $\{x, Rx\}$, it follows that $\mu_R = \nu_R$, and so $\mu = \nu$, as desired. $\square$

Proof of Proposition 1.5. For both assertions, we may apply Lemma 2.1 to the pair of measures μ and $R_*\mu$, getting $\bar{\mu} = -\bar{\mu}$, whence $\bar{\mu} = 0$. Thus, $\mu = R_*\mu$, as desired. $\square$

3. Acknowledgement

I thank Marcos Matabuena for asking me about strong negative type for the angular metric on compositional data, i.e., on the probability simplex.

References

[Faraut and Harzallah 1974] J. Faraut and K. Harzallah, "Distances hilbertiennes invariantes sur un espace homogène", *Ann. Inst. Fourier* (*Grenoble*) **24**:3 (1974), xiv, 171–217. MR Zbl

[Gangolli 1967] R. Gangolli, "Positive definite kernels on homogeneous spaces and certain stochastic processes related to Lévy's Brownian motion of several parameters", *Ann. Inst. H. Poincaré Sect. B* (*N.S.*) **3** (1967), 121–226. MR Zbl

[Hjorth et al. 1998] P. Hjorth, P. Lisoněk, S. Markvorsen, and C. Thomassen, "Finite metric spaces of strictly negative type", *Linear Algebra Appl.* **270** (1998), 255–273. MR Zbl

[Hjorth et al. 2002] P. G. Hjorth, S. L. Kokkendorff, and S. Markvorsen, "Hyperbolic spaces are of strictly negative type", *Proc. Amer. Math. Soc.* **130**:1 (2002), 175–181. MR Zbl

[Lyons 2013] R. Lyons, "Distance covariance in metric spaces", *Ann. Probab.* **41**:5 (2013), 3284–3305. MR Zbl

[Lyons 2014] R. Lyons, "Hyperbolic space has strong negative type", *Illinois J. Math.* **58**:4 (2014), 1009–1013. MR Zbl

[Lyons 2018] R. Lyons, "Errata to "Distance covariance in metric spaces" [MR3127883]", *Ann. Probab.* **46**:4 (2018), 2400–2405. MR Zbl

[Lyons and Zumbrun 2018] R. Lyons and K. Zumbrun, "A calculus proof of the Cramér–Wold theorem", *Proc. Amer. Math. Soc.* **146**:3 (2018), 1331–1334. MR Zbl

[Meckes 2013] M. W. Meckes, "Positive definite metric spaces", *Positivity* **17**:3 (2013), 733–757. MR Zbl

[Rubin 1999] B. Rubin, "Inversion and characterization of the hemispherical transform", *J. Anal. Math.* **77** (1999), 105–128. MR Zbl

[Schneider 1970] R. Schneider, "Über eine Integralgleichung in der Theorie der konvexen Körper", *Math. Nachr.* **44** (1970), 55–75. MR Zbl

[Székely and Rizzo 2017] G. J. Székely and M. L. Rizzo, "The energy of data", *Annual Review of Statistics and Its Application* **4**:1 (2017), 447–479.

[Székely et al. 2007] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, "Measuring and testing dependence by correlation of distances", *Ann. Statist.* **35**:6 (2007), 2769–2794. MR Zbl

[Walther 1997] G. Walther, "On a conjecture concerning a theorem of Cramér and Wold", *J. Multivariate Anal.* **63**:2 (1997), 313–319. Addendum, *J. Multivariate Anal.* **67**(2) (1998), 431. MR Zbl

[Zinger et al. 1992] A. A. Zinger, A. V. Kakosyan, and L. B. Klebanov, "A characterization of distributions by mean values of statistics and certain probabilistic metrics", *J. Soviet Math.* **59**:4 (1992), 914–920. Stability problems for stochastic models. Translated from *Problemy Ustoichivosti Stokhasticheskikh Modelei, Trudy Seminara*, 1989, pp. 47–55. MR Zbl

Received May 16, 2019. Revised April 4, 2020.

RUSSELL LYONS
DEPARTMENT OF MATHEMATICS
INDIANA UNIV, BLOOMINGTON
BLOOMINGTON, IN
UNITED STATES

rdlyons@indiana.edu

PACIFIC JOURNAL OF MATHEMATICS

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

msp.org/pjm

See inside back cover or msp.org/pjm for submission instructions.

The subscription price for 2020 is US $520/year for the electronic version, and $705/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 1945-5844 electronic, 0030-8730 printed) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS

Volume 307 No. 2 August 2020