Bull. Aust. Math. Soc. **80** (2009), 353–359
doi:10.1017/S0004972709000707

THE AVERAGE DISTANCE BETWEEN TWO POINTS

BERNHARD BURGSTALLER and FRIEDRICH PILLICHSHAMMER ✉

(Received 19 May 2008)

Abstract

We provide bounds on the average distance between two points uniformly and independently chosen from a compact convex subset of the s-dimensional Euclidean space.

2000 *Mathematics subject classification*: primary 52A22; secondary 60D05.
Keywords and phrases: distance geometry, random convex sets, average distance.

Let X be a compact convex subset of the s-dimensional Euclidean space $\mathbb{R}^s$ and assume that we choose uniformly and independently two points from X. How large is the expected Euclidean distance $\|\cdot\|$ between these two points? In other words, we require the quantity

$$a(X) := \mathbb{E}[\|x-y\|] = \frac{1}{\lambda(X)^2} \int_X \int_X \|x-y\| \, d\lambda(x) \, d\lambda(y),$$

where λ denotes the s-dimensional Lebesgue measure. This problem was stated in [1, 2, 4, 5]. Note that there is a close connection between this problem and that of finding the moments of the length of random chords (see [8, Ch. 4, Section 2] or [9, Ch. 2]).

Trivially $a(X) \le d(X)$, where $d(X) = \max\{\|x-y\| : x, y \in X\}$ is the diameter of X. The following results are well known from the literature.

EXAMPLE 1.

(1) For all compact convex subsets of $\mathbb{R}$ (the intervals) we have $a(X) = d(X)/3$.
(2) If $X \subseteq \mathbb{R}^s$ is a ball with diameter $d(X)$, then

$$a(X) = \frac{s}{2s+1} \beta_s \, d(X),$$

where

$$\beta_s = \begin{cases} \dfrac{2^{3s+1}((s/2)!)^2 s!}{(s+1)(2s)!\pi} & \text{for even } s, \\ \dfrac{2^{s+1}(s!)^3}{(s+1)(((s-1)/2)!)^2(2s)!} & \text{for odd } s. \end{cases}$$

For a proof see [4] or [8]. In particular, if X is a disc in $\mathbb{R}^2$ with diameter $d(X)$, then

$$a(X) = 64d(X)/(45\pi) = 0.45271\ldots\ d(X).$$

(3) If $X \subseteq \mathbb{R}^2$ is a rectangle of sides $a \geq b$, then (see [8])

$$a(X) = \frac{1}{15}\left[\frac{a^3}{b^2} + \frac{b^3}{a^2} + d\left(3 - \frac{a^2}{b^2} - \frac{b^2}{a^2}\right) + \frac{5}{2}\left(\frac{b^2}{a}\log\frac{a+d}{b} + \frac{a^2}{b}\log\frac{b+d}{a}\right)\right],$$

where $d = d(X) = \sqrt{a^2+b^2}$. In particular, if X is a square, then

$$a(X) = (2+\sqrt{2}+5\log(\sqrt{2}+1))\frac{d(X)}{15\sqrt{2}} = 0.36869\ldots d(X).$$

(4) If X is a cube in $\mathbb{R}^s$, then

$$a(X) = \frac{1}{\sqrt{6}}\left(1 - \frac{7}{40s} - \frac{65}{869s^2} + \cdots\right)d(X)$$

and

$$a(X) \leq \frac{1}{\sqrt{6}}\left(\frac{1+2\sqrt{1-3/(5s)}}{3}\right)^{1/2} d(X).$$

For a proof of the asymptotic formula see [5], and for a proof of the upper bound see [2].

(5) If $X \subseteq \mathbb{R}^2$ is an equilateral triangle of side a, then (see [8])

$$a(X) = \frac{3a}{5}\left(\frac{1}{3} + \frac{\log 3}{4}\right).$$

In the following we prove a general bound on $a(X)$ for $X \subseteq \mathbb{R}^s$ with fixed diameter $d(X) = 1$. Furthermore, we present two results which may be useful to give upper and lower bounds on $a(X)$.

Denote by $\mathcal{M}(X)$ the space of all regular Borel probability measures on X. It is well known that $\mathcal{M}(X)$ equipped with the w^*-topology becomes a compact convex space. For $x \in X$, let $\delta_x \in \mathcal{M}(X)$ be the point measure concentrated on x. It is easy to show that the set $\{\delta_x \mid x \in X\}$ is the set of all extreme points of $\mathcal{M}(X)$ and hence from the Krein–Milman theorem we find that $\mathcal{M}(X)$ is the w^*-closure of the convex hull of $\{\delta_x \mid x \in X\}$. Let $\mathcal{F} = \{(1/n)\sum_{i=1}^{n} \delta_{x_i} \mid x_1, \ldots, x_n \in X, n \in \mathbb{N}\}$. Then one can show that $\mathcal{F}$ is the set of all convex combinations with rational coefficients of extreme points of $\mathcal{M}(X)$. Now, since $\mathbb{Q}$ is dense in $\mathbb{R}$, we deduce from the above considerations that $\mathcal{F}$ is dense in $\mathcal{M}(X)$.

For any $\mu \in \mathcal{M}(X)$, we define

$$I(\mu) := \int_X \int_X \|x - y\| \, d\mu(x) \, d\mu(y).$$

It is known that the mapping $I : \mathcal{M}(X) \to \mathbb{R}$ is continuous with respect to the w^*-topology on $\mathcal{M}(X)$ (see [10, Lemma 1]). Note that $a(X) = I(\lambda')$ where λ' is the normalized Lebesgue measure on X.

REMARK 2. Let X be a compact subset of $\mathbb{R}^s$ and let $(x_n)_{n\geq 0}$ be a sequence which is uniformly distributed in X with respect to the normalized Lebesgue measure λ' on X, that is, $\mu_N := N^{-1}\sum_{i=0}^{N-1} \delta_{x_i} \to \lambda'$ with respect to w^*-topology on $\mathcal{M}(X)$. Then by continuity of I we obtain

$$\frac{1}{N^2}\sum_{i,j=0}^{N-1} \|x_i - x_j\| = I(\mu_N) \to I(\lambda') = a(X) \quad \text{as } N \to \infty.$$

THEOREM 3. *Let X be a compact subset of $\mathbb{R}^s$ with diameter $d(X) = 1$. Then*

$$a(X) \leq \sqrt{\frac{2s}{s+1}} \frac{2^{s-2}\Gamma(s/2)^2}{\Gamma(s-1/2)\sqrt{\pi}},$$

where Γ denotes the gamma function. For $s = 2$ this bound can be improved to

$$a(X) \leq \frac{229}{800} + \frac{44}{75}\sqrt{2-\sqrt{3}} + \frac{19}{480}\sqrt{5} = 0.678442\ldots.$$

PROOF. We have

$$a(X) = I(\lambda') \leq \sup_{\mu \in \mathcal{M}(X)} I(\mu).$$

Since $I : \mathcal{M}(X) \to \mathbb{R}$ is continuous with respect to the w^*-topology on $\mathcal{M}(X)$ and $\mathcal{F}$ is dense in $\mathcal{M}(X)$ we obtain

$$\sup_{\mu \in \mathcal{M}(X)} I(\mu) = \sup_{n \in \mathbb{N}, x_1, \ldots, x_n \in X} \frac{1}{n^2}\sum_{i,j=1}^{n} \|x_i - x_j\|.$$

It was shown by Nickolas and Yost [6] that, for all $x_1, \ldots, x_n \in X \subseteq \mathbb{R}^s$ with $d(X) = 1$,

$$\frac{1}{n^2}\sum_{i,j=1}^{n} \|x_i - x_j\| \leq \sqrt{\frac{2s}{s+1}} \frac{2^{s-2}\Gamma(s/2)^2}{\Gamma(s-1/2)\sqrt{\pi}}.$$

For $s = 2$ it was shown by Pillichshammer [7] that, for all $x_1, \ldots, x_n \in \mathbb{R}^2$ with $\|x_i - x_j\| \le 1$,

$$\frac{1}{n^2}\sum_{i,j=1}^{n} \|x_i - x_j\| \le \frac{229}{800} + \frac{44}{75}\sqrt{2-\sqrt{3}} + \frac{19}{480}\sqrt{5} = 0.678442\ldots.$$

The result follows from these bounds. □

REMARK 4. Note that it is not true in general that $X \subseteq Y$ implies $a(X) \le a(Y)$. For example, for $h > 0$, let A_h denote the right triangle with vertices $\{(0, 0), (1, 0), (1, h)\}$. Then

$$\begin{aligned} a(A_h) &= \frac{4}{h^2}\int_0^1\int_0^{hx_1}\int_0^1\int_0^{hx_2}\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}\,dy_2\,dx_2\,dy_1\,dx_1 \\ &\ge 4\int_0^1\int_0^1\frac{1}{h^2}\int_0^{hx_1}\int_0^{hx_2}|x_1-x_2|\,dy_2\,dy_1\,dx_2\,dx_1 \\ &= 4\int_0^1\int_0^1|x_1-x_2|x_1x_2\,dx_2\,dx_1 = \frac{4}{15}. \end{aligned}$$

On the other hand,

$$a(A_h) \le 4\int_0^1\int_0^1 x_1x_2\sqrt{(x_1-x_2)^2+h^2}\,dx_2\,dx_1$$

and hence $\lim_{h\to 0^+} a(A_h) = 4/15$. Thus for any $\varepsilon > 0$ there is a $h_0 > 0$ such that, for all $0 < h < h_0$, $|a(A_h) - 4/15| < \varepsilon$.

For $l > 0$, let B_l be the rectangle with vertices $\{(0, 0), (1, 0), (1, -l), (0, -l)\}$. Then from Example 1 we obtain $\lim_{l\to 0^+} a(B_l) = 1/3$. Thus for any $\varepsilon > 0$ there is a $l_0 > 0$ such that, for all $0 < l < l_0$, $|a(B_l) - 1/3| < \varepsilon$.

Now let $\varepsilon, \delta > 0$. Choose $0 < h < \min\{1, h_0\}$, and $0 < l < \min\{1, l_0\}$ small enough such that $\lambda(B_l) < \delta\lambda(A_h)$ and let $C_{h,l} := A_h \cup B_l$. Then

$$\begin{aligned} a(C_{h,l}) &= \frac{\lambda(A_h)^2}{(\lambda(A_h)+\lambda(B_l))^2}a(A_h) + \frac{\lambda(B_l)^2}{(\lambda(A_h)+\lambda(B_l))^2}a(B_l) \\ &\quad + \frac{2}{(\lambda(A_h)+\lambda(B_l))^2}\int_{A_h}\int_{B_l}\|x-y\|\,d\lambda(x)\,d\lambda(y) \\ &< a(A_h) + \left(\frac{\delta}{1+\delta}\right)^2 a(B_l) + \frac{3\delta}{1+\delta} < \frac{4}{15} + \varepsilon + \delta^2\left(\frac{1}{3}+\varepsilon\right) + 3\delta. \end{aligned}$$

Hence if we choose $1/60 > \varepsilon > 0$ and $\delta > 0$ small enough we can obtain $a(C_{h,l}) < 3/10$. Of course $B_l \subseteq C_{h,l}$, but

$$a(B_l) \ge \tfrac{1}{3} - \varepsilon \ge \tfrac{19}{60} > \tfrac{3}{10} > a(C_{h,l}).$$

LEMMA 5.

(1) *Let X and Y be compact sets in $\mathbb{R}^s$ with $\lambda(X \cap Y) = 0$. Then*

$$\lambda(X \cup Y)a(X \cup Y) \geq \lambda(X)a(X) + \lambda(Y)a(Y).$$

(2) *Let $X \subseteq Y$ be compact sets in $\mathbb{R}^s$. Then*

$$\lambda(X)a(X) \leq \lambda(Y)a(Y).$$

PROOF. (1) We have

$$\begin{aligned} a(X \cup Y) = {} & \frac{\lambda(X)^2}{(\lambda(X) + \lambda(Y))^2} a(X) + \frac{\lambda(Y)^2}{(\lambda(X) + \lambda(Y))^2} a(Y) \\ & + 2 \frac{\lambda(X)\lambda(Y)}{(\lambda(X) + \lambda(Y))^2} \frac{1}{\lambda(X)\lambda(Y)} \int_X \int_Y \|x - y\| \, d\lambda(x) \, d\lambda(y). \end{aligned}$$

For any regular Borel probability measures μ and ν on a subset A of the Euclidean space $\mathbb{R}^s$ we have (see [10, Equation (∗∗)])

$$2 \int_A \int_A \|x - y\| \, d\mu(x) \, d\nu(y) \geq I(\mu) + I(\nu).$$

Now let $A = X \cup Y$, let μ be the probability measure on A which is the normalized Lebesgue measure on X and which is zero on Y and let ν be the probability measure on A which is the normalized Lebesgue measure on Y and which is zero on X. Then

$$\begin{aligned} \frac{2}{\lambda(X)\lambda(Y)} \int_X \int_Y \|x - y\| \, d\lambda(x) \, d\lambda(y) &= 2 \int_A \int_A \int_Y \|x - y\| \, d\mu(x) \, d\nu(y) \\ &\geq \int_{X \cup Y} \int_{X \cup Y} \|x - y\| \, d\mu(x) \, d\mu(y) + \int_{X \cup Y} \int_{X \cup Y} \|x - y\| \, d\nu(x) \, d\nu(y) \\ &= a(X) + a(Y). \end{aligned}$$

Hence

$$\begin{aligned} (\lambda(X) + \lambda(Y))^2 a(X \cup Y) &\geq \lambda(X)^2 a(X) + \lambda(Y)^2 a(Y) + \lambda(X)\lambda(Y)(a(X) + a(Y)) \\ &= (\lambda(X)a(X) + \lambda(Y)a(Y))(\lambda(X) + \lambda(Y)). \end{aligned}$$

(2) This assertion follows from the first one. □

COROLLARY 6. *Let $X \subseteq \mathbb{R}^s$ be compact and convex and let $r = r(X)$ be the in-radius and $R = R(X)$ be the circumradius of X. Then*

$$\frac{\pi^{s/2}}{\Gamma(s/2 + 1)} \frac{2s}{2s + 1} \beta_s r^{s+1} \leq \lambda(X)a(X) \leq \frac{\pi^{s/2}}{\Gamma(s/2 + 1)} \frac{2s}{2s + 1} \beta_s R^{s+1}$$

with equality if X is a ball. In particular, for $s = 2$ we have

$$\frac{128}{45} r^3 \leq \lambda(X)a(X) \leq R^3 \frac{128}{45}$$

with equality if X is a disc.

PROOF. Let K_{in} be the in-ball and let K_{circ} be the circumscribed ball of X. From Lemma 5 we obtain $\lambda(K_{\text{in}})a(K_{\text{in}}) \le \lambda(X)a(X) \le \lambda(K_{\text{circ}})a(K_{\text{circ}})$ and the result follows from Example 1 (note that the volume of an s-dimensional ball of radius $t > 0$ is given by $\pi^{s/2}t^s/\Gamma(s/2+1)$). $\square$

REMARK 7. It follows from a result of Blaschke [3] that, for any plane compact convex $X \subseteq \mathbb{R}^2$,

$$a(X) \ge \frac{128}{45\pi}\sqrt{\frac{\lambda(X)}{\pi}}$$

with equality if X is a disc. In many cases this bound yields better results than the lower bound from Corollary 6 in the plane case (see Examples 8 and 10 below). For more information see [8, Ch. 4, Section 2] or [9, Ch. 2, Equation (2.55)].

EXAMPLE 8. For $n \in \mathbb{N}$, $n \ge 3$, let $X_n \subseteq \mathbb{R}^2$ be the regular n-gon with vertices on the unit circle. Then $\lambda(X_n) = (n/2)\sin(2\pi/n)$, $R = 1$ and $r = \cos(\pi/n)$. Hence we obtain

$$\frac{256}{45}\frac{\cos^3(\pi/n)}{n\sin(2\pi/n)} \le a(X_n) \le \frac{256}{45}\frac{1}{n\sin(2\pi/n)}.$$

From Remark 7 we even obtain the lower bound

$$a(X_n) \ge \frac{128}{45\pi}\sqrt{\frac{n}{2\pi}\sin\frac{2\pi}{n}}$$

which is slightly better than the lower bound above. Note that

$$\lim_{n\to\infty}\frac{128}{45\pi}\sqrt{\frac{n}{2\pi}\sin\frac{2\pi}{n}} = \lim_{n\to\infty}\frac{256}{45}\frac{\cos^3(\pi/n)}{n\sin(2\pi/n)} = \lim_{n\to\infty}\frac{256}{45}\frac{1}{n\sin(2\pi/n)} = \frac{128}{45\pi}.$$

In some cases the following easy lemma gives better estimates than Corollary 6.

LEMMA 9. *Let X be a compact subset of $\mathbb{R}^s$ and let $T : \mathbb{R}^s \to \mathbb{R}^s$ be a linear mapping with norm $\|T\|_2$. Then we have $a(T(X)) \le a(X)\|T\|_2$.*

EXAMPLE 10. Let X be an ellipse $x^2 + y^2/b^2 \le 1$ in the Euclidean plane with $0 < b \le 1$. Then $X = T(K)$ where K is the disc with diameter 2 and center in the origin and where $T = \left(\begin{smallmatrix}1 & 0\\ 0 & b\end{smallmatrix}\right)$. It is easy to see that $\|T\|_2 = \max\{1, |b|\} = 1$ and $\|T^{-1}\|_2 = 1/b$. Then from Lemma 9 we obtain

$$b\frac{128}{45\pi} = ba(K) \le a(X) \le a(K) = \frac{128}{45\pi}$$

whereas from Corollary 6 we would just obtain

$$b^2\frac{128}{45\pi} \le a(X) \le \frac{1}{b}\frac{128}{45\pi}.$$

From Remark 7 we obtain the lower bound $a(X) \ge \sqrt{b}(128/45\pi)$.

References

[1] V. S. Alagar, 'The distribution of the distance between random points', *J. Appl. Probab.* **13** (1976), 558–566.

[2] R. S. Anderssen, R. P. Brent, D. J. Daley and P. A. P. Moran, 'Concerning $\int_0^1 \cdots \int_0^1 (x_1^2 + \cdots + x_k^2)^{1/2}\,dx_1 \cdots dx_k$ and a Taylor series method', *SIAM J. Appl. Math.* **30** (1976), 22–30.

[3] W. Blaschke, 'Eine isoperimetrische Eigenschaft des Kreises', *Math. Z.* **1** (1918), 52–57.

[4] S. R. Dunbar, 'The average distance between points in geometric figures', *College Math. J.* **28** (1997), 187–197.

[5] D. J. Gates, 'Asymptotics of two integrals from optimization theory and geometric probability', *Adv. Appl. Probab.* **17** (1985), 908–910.

[6] P. Nickolas and D. Yost, 'The average distance property for subsets of Euclidean space', *Arch. Math. (Basel)* **50** (1988), 380–384.

[7] F. Pillichshammer, 'A note on the sum of distances under a diameter constraint', *Arch. Math. (Basel)* **77** (2001), 195–199.

[8] L. A. Santaló, *Integral Geometry and Geometric Probability*, Encyclopedia of Mathematics and its Applications, 1 (Addison-Wesley, Reading, MA, 1976).

[9] H. Solomon, *Geometric Probability*, CBMS-NSF Regional Conference Series in Applied Mathematics, 28 (SIAM, Philadelphia, PA, 1978).

[10] R. Wolf, 'On the average distance property and certain energy integrals', *Ark. Mat.* **35** (1997), 387–400.

BERNHARD BURGSTALLER, Doppler Institute for Mathematical Physics,
Trojanova 13, 12000 Prague, Czech Republic
e-mail: bernhardburgstaller@yahoo.de

FRIEDRICH PILLICHSHAMMER, Institut für Finanzmathematik,
Universität Linz, Altenbergstraße 69, A-4040 Linz, Austria
e-mail: friedrich.pillichshammer@jku.at