The average distance between two points

Bernhard Burgstaller and Friedrich Pillichshammer

Abstract

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

Let X be a compact convex subset of the s-dimensional Euclidean space $\mathbb{R}^s$ and assume 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 ask for the quantity

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

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

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

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

For a proof see [4] or [8]. Especially, 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 $a \ge b$, then we have (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 $d = d(X) = \sqrt{a^2+b^2}$. Especially, if X is a square, then we have

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

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

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

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 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 $\{\delta_x : x \in X\}$ 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 $\{\delta_x : x \in X\}$. Let $\mathcal{F} = \{\frac{1}{n}\sum_{i=1}^n \delta_{x_i} : x_1, \dots, x_n \in X, n \in \mathbb{N}\}$. 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

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

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 1 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, i.e., $\mu_N := N^{-1}\sum_{i=0}^{N-1}\delta_{x_i} \to \lambda'$ with respect to w^*-topology on $\mathcal{M}(X)$. Then by continuity of I we obtain

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

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

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

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

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

Proof. We have

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

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 $d(X) = 1$ we have

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

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

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

The result follows from these bounds. □

Remark 2 Note that it is not true in general that $X \subseteq Y$ implies $a(X) \le a(Y)$. For example, let, for $h > 0$, A_h denote the right triangle with vertices $\{(0,0), (1,0), (1,h)\}$. Then we have

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

On the other hand we have

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

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 $0 < h < h_0$ we have $|a(A_h) - 4/15| < \varepsilon$.

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

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 we have

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

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

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

Lemma 1 *1. Let X and Y be compact sets in $\mathbb{R}^s$ with $X \cap Y = \emptyset$. Then we have*

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

2. Let $X \subseteq Y$ be compact sets in $\mathbb{R}^s$. Then we have

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

Proof. 1. We have

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

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 (**)])

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

Let now $A = X \cup Y$, 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 we have

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

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.

$\square$

Corollary 1 *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 we have*

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

with equality if X is a ball. Especially, 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 1 we obtain $\lambda(K_{\text{in}})a(K_{\text{in}}) \leq \lambda(X)a(X) \leq \lambda(K_{\text{circ}})a(K_{\text{circ}})$ and the result follows from Example 1 (note that the volume of an s-dimensional ball of radius $t > 0$ is given by $\pi^{s/2}t^s/\Gamma(s/2+1)$). □

Remark 3 It follows from a result of Blaschke [3] that for any plane compact convex $X \subseteq \mathbb{R}^2$ we have

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

with equality if X is a disc. In many cases this bound is better than the lower bound from Corollary 1 in the plane case (for example, in Example 2 below). For more information see [8, Chapter 4, Section 2] or [9, Chapter 2, Eq. (2.55)].

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

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

From Remark 3 we even obtain the lower bound $a(X_n) \geq \frac{128}{45\pi}\sqrt{\frac{n}{2\pi}\sin\frac{2\pi}{n}}$ which is slightly better than the lower bound above. Note that $\lim_{n\to\infty}\frac{128}{45\pi}\sqrt{\frac{n}{2\pi}\sin\frac{2\pi}{n}} = \lim_{n\to\infty}\frac{256}{45}\frac{\cos^3\frac{\pi}{n}}{n\sin\frac{2\pi}{n}} = \lim_{n\to\infty}\frac{256}{45}\frac{1}{n\sin\frac{2\pi}{n}} = \frac{128}{45\pi}$.

In some cases the following easy lemma gives better estimates than Corollary 1.

Lemma 2 *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 $a(T(X)) \leq a(X)\|T\|_2$.*

Example 3 Let X be an ellipse $x^2 + y^2/b^2 \leq 1$ in the Euclidean plane with $0 < b \leq 1$. Then $X = T(K)$ where K is the disc with diameter 2 and center in the origin and where $T = \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix}$. It is easy to see that $\|T\|_2 = \max\{1, |b|\} = 1$ and $\|T^{-1}\|_2 = 1/b$. Then from Lemma 2 we obtain

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

whereas from Corollary 1 we would just obtain

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

From Remark 3 we obtain the lower bound $a(X) \geq \sqrt{b}\frac{128}{45\pi}$.

References

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

[2] Anderssen, R. S., Brent, R. P., Daley, D. J. and Moran, P. A. P.: Concerning $\int^1{}_0 \cdots \int^1{}_0 (x^2{}_1 + \cdots + x^2{}_k)^{1/2} dx_1 \cdots dx_k$ and a Taylor series method. SIAM J. Appl. Math. 30 (1976), 22–30.

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

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

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

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

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

[8] Santaló, L. A.: *Integral geometry and geometric probability.* Encyclopedia of Mathematics and its Applications, Vol. 1. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976.

[9] Solomon, H.: *Geometric probability.* No 28 in CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, 1978.

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

Author's Address:
Bernhard Burgstaller,

Friedrich Pillichshammer, Institut für Finanzmathematik, Universität Linz, Altenbergstraße 69, A-4040 Linz, Austria. Email: gerhard.larcher@jku.at, friedrich.pillichshammer@jku.at