The length of the shortest edge of a graph on a sphere

Hiroshi Maehara
College of Education, Ryukyu University, Okinawa, Japan

Abstract

Let S^d denote a unit sphere in the $(d+1)$-dimensional Euclidean space $\mathbf{R}^{d+1}$ $(d \geq 1)$. For a simple graph $G_\mathcal{E}$ with edge set $\mathcal{E}$, take independent random points $x_k, k \in V(G_\mathcal{E})$, on S^d, and let $D_\mathcal{E}$ be the minimum value of the spherical distance between x_i, x_j for $\{i,j\} \in \mathcal{E}$. We prove that $|\mathcal{E}|D_\mathcal{E}^d$ is asymptotically (as $|\mathcal{E}| \to \infty$) distributed according to the exponential distribution with mean $dB(\frac{1}{2}, \frac{d}{2})$, where $B(p,q)$ is the beta function.

1 Introduction

Let S^d denote a unit sphere in the $(d+1)$-dimensional Euclidean space $\mathbf{R}^{d+1}$ $(d \geq 1)$. Consider n random points on S^d distributed independently and uniformly, and let D denote the smallest spherical distance between these n points. What is the distribution of D? To estimate its distribution is useful in statistical applications. If $n = 2$, this problem reduces to the problem of computing the surface area of a spherical cap of a given angular radius on a d-dimensional sphere, and the case $n = 2, d \to \infty$ is considered in [1]. The case $d = 1, n \to \infty$ is partly considered in [3]. Moran [5] considered the bounds for the tail of the distribution of D to study the distribution of the largest sample correlation coefficient between a set of normally distributed variables. Here, we are going to determine the asymptotic distribution of D as $n \to \infty$ for general $d \geq 1$.

Let us consider in more general setting: Let

$$G_\mathcal{E} = (V, \mathcal{E})$$

denote a simple graph with edge-set $\mathcal{E}$ and vertex set V. The number of edges in $\mathcal{E}$ is denoted by N;

$$N = |\mathcal{E}|.$$

Let x_k $(k \in V)$ be random points on S^d distributed independently and uniformly, and let

$$D_\mathcal{E} := \min_{\{i,j\} \in \mathcal{E}} \widehat{x_i x_j},$$

where $\widehat{x_i x_j}$ is the spherical distance between x_i and x_j. Then $D_{\mathcal{E}}$ is a random variable. We are going to determine the asymptotic distribution of $D_{\mathcal{E}}$ as $N \to \infty$.

For a positive constant c, put

$$\alpha = \alpha_N = \left(\frac{c}{N}\right)^{1/d}.$$

For each edge $A = \{i, j\}$ of $G_{\mathcal{E}}$, define a random variable X_A by

$$X_A = \begin{cases} 1 & \text{if } \widehat{x_i x_j} \leq \alpha \\ 0 & \text{if } \widehat{x_i x_j} > \alpha \end{cases}$$

Then "$D_{\mathcal{E}} > \alpha$" is equivalent to "$\sum_{A \in \mathcal{E}} X_A = 0$". Let

$$\beta = \beta_d = dB(\tfrac{1}{2}, \tfrac{d}{2}),$$

where $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$, the beta function. For instance,

$$\beta_1 = \pi,\ \beta_2 = 4,\ \beta_3 = 3\pi/2,\ \beta_4 = 16/3.$$

We prove the following.

Theorem 1 *For each integer $k \geq 0$,*

$$\Pr(\sum_{A \in \mathcal{E}} X_A = k) \to \frac{(c/\beta)^k e^{-c/\beta}}{k!} \quad \text{as } N \to \infty.$$

Since $\Pr(D_{\mathcal{E}} \leq \alpha) = 1 - \Pr(D_{\mathcal{E}} > \alpha) = 1 - \Pr(\sum X_A = 0)$, it follows that $\Pr(D_{\mathcal{E}} \leq \alpha) \to 1 - e^{-c/\beta}$ as $N \to \infty$. Hence,

$$\Pr(ND_{\mathcal{E}}^d \leq c) = \Pr(ND_{\mathcal{E}}^d \leq N\alpha^d) = \Pr(D_{\mathcal{E}} \leq \alpha) \to 1 - e^{-c/\beta}.$$

Hence the next follows.

Theorem 2 *For any $t > 0$,*

$$\lim_{|\mathcal{E}| \to \infty} \Pr(|\mathcal{E}| D_{\mathcal{E}}^d \leq t) = 1 - e^{-t/\beta}.$$

□

Thus the asymptotic distribution of $|\mathcal{E}|D_{\mathcal{E}}^d$ is irrelevant to the graph structure of $G_{\mathcal{E}}$. By taking the complete graph K_n as $G_{\mathcal{E}}$, we have the next.

Corollary 1 *Let D denote the smallest spherical distance between n independent random points on S^d. Then*

$$\Pr(\binom{n}{2} D^d \leq t) \to 1 - e^{-t/\beta} \quad \text{as } n \to \infty.$$

□

If we take the complete bipartite graph $K_{p,q}$ as $G_{\mathcal{E}}$, then we have the next.

Corollary 2 *Choose p 'red' points and q 'blue' points, all distributed independently and uniformly on S^d, and let $D_{p,q}$ denote the minimum value of the spherical distance from a red point to a blue point. Then*

$$\Pr(pqD_{p,q}^d < t) \to 1 - e^{-t/\beta} \quad as\ pq \to \infty.$$

□

2 A few lemmas

Put

$$\varphi(\alpha) = \frac{area(C(\alpha))}{area(S^d)},$$

where $C(\alpha)$ denotes a spherical cap of angular radius α, and $aera(\cdot)$ denotes the d-dimensional surface area. For a collection of edges $A_1, A_2, \ldots A_s$ of $G_{\mathcal{E}}$, let

$$G(A_1, A_2, \ldots, A_s)$$

denote the subgraph of $G_{\mathcal{E}}$ with vertex set $A_1 \cup A_2 \cup \ldots \cup A_s$ and the edges $A_1, A_2, \ldots, A_s$.

Lemma 1 *If the graph $G(A_1, A_2, \ldots, A_s)$ contains no cycle, then*

$$\Pr(X_{A_1}X_{A_2}\ldots X_{A_s} = 1) = \varphi(\alpha)^s.$$

Proof. Since $G(A_1, \ldots, A_s)$ contains no cycle, there is an *end*-edge (an edge that is incident to a vertex of degree 1). We may suppose that $A_s = \{j, k\}$ is an end-edge, and that k is a vertex of degree 1. Then, by the symmetry of the sphere, we may assume that x_j is a fixed point on S^d, and hence X_{A_s} is independent of other X_{A_i}s. Now, $X_{A_s} = 1$ if and only if x_k lies in the spherical cap with center x_j and angular radius α. Hence $\Pr(X_{A_s} = 1) = \varphi(\alpha)$. Thus

$$\Pr(X_{A_1}X_{A_2}\ldots X_{A_s} = 1) = \Pr(X_{A_1}\ldots X_{A_{s-1}} = 1)\varphi(\alpha).$$

Since $G(A_1, \ldots, A_{s-1})$ has no cycle, the lemma now follows by induction. □

Lemma 2 *Let $F_{r,\mathcal{E}}$ denote the number of r-tuples $(A_1, A_2, \ldots, A_r) \in \mathcal{E}^r$ such that $G(A_1, \ldots, A_r)$ forms a forest. Then $F_{r,\mathcal{E}} = N^r + O(N^{r-1})$.*

Proof. For any k edges $A_1, A_2, \ldots A_k$ $(k < r)$, the number of edges in the subgraph of $G_{\mathcal{E}}$ induced by $\{A_1, A_2, \ldots A_k\}$ is at most $\binom{2k}{2}$. Hence

$$F_{r,\mathcal{E}} \geq N(N - \binom{2}{2})(N - \binom{4}{2})(N - \binom{6}{2})\cdots(N - \binom{2(r-1)}{2}) = N^r + O(N^{r-1}).$$

On the other hand, clearly $F_{r,\mathcal{E}} \leq N^r$. Hence we have the lemma. □

The next result is proved in [4].

Lemma 3 *For any integer $p \geq 3$, the number of p-cycles in $G_{\mathcal{E}}$ does not exceed* $\left(\frac{2^{p/2}}{2p}\right) |\mathcal{E}|^{p/2}$. □

Finally, we need a lemma concerning Poisson convergence. Let $Y_1, Y_2, \dots Y_N$ be random variables which may take two values 0 and 1 only, and let $Y = Y_1 + Y_2 + \dots + Y_N$. Then, for a positive integer r, the *r-th binomial moment* U_r of Y is defined by

$$U_r = \sum E(Y_{i_1} Y_{i_2} \dots Y_{i_r}),$$

where the summation is over all $1 \leq i_1 < i_2 < \dots < i_r \leq N$, that is, over all r-subsets of $\{1, 2, \dots, N\}$. Then the following lemma holds. For the proof, see Palmer [6] pp. 139-141.

Lemma 4 *Suppose that for each $r = 1, 2, 3, \dots$,*

$$\lim_{N \to \infty} U_r = \frac{\mu^r}{r!}, \quad 0 < \mu < \infty.$$

Then for each integer $k \geq 0$,

$$\lim_{N \to \infty} \Pr(Y = k) = \frac{e^{-\mu} \mu^k}{k!}.$$

□

3 Proof of Theorem 1

First, we show that

$$N\varphi(\alpha) = \frac{c}{\beta} + O(N^{-2/d}) \text{ (as } N \to \infty).$$

Let v_d denote the volume of the d-dimensional unit ball. Then, since

$$(\sin \alpha)^d v_d < area(C(\alpha)) < \alpha^d v_d,$$

and since $\alpha = (c/N)^{1/d}$, $(\sin \alpha)^d = (\alpha + O(\alpha^3))^d = \alpha^d + O(\alpha^{d+2})$ as $N \to \infty$, we have

$$area(C(\alpha)) = \alpha^d v_d (1 + O(\alpha^2)) = (c/N) v_d (1 + O(N^{-2/d})) \text{ as } N \to \infty.$$

Since

$$v_d = \frac{2\pi^{d/2}}{d\Gamma(\frac{d}{2})} \text{ and } area(S^d) = \frac{2\pi^{(d+1)/2}}{\Gamma(\frac{d+1}{2})}$$

(see, e.g., Kendall [2] p.35), we have

$$\begin{aligned} N\varphi(\alpha) &= c \times \frac{2\pi^{d/2}}{d\Gamma(\frac{d}{2})} \times \frac{\Gamma(\frac{d+1}{2})}{2\pi^{(d+1)/2}}(1 + O(N^{-2/d})) \\ &= c \times \frac{\Gamma(\frac{d+1}{2})}{d\sqrt{\pi}\Gamma(\frac{d}{2})}(1 + O(N^{-2/d})) = \frac{c}{dB(\frac{1}{2}, \frac{d}{2})}(1 + O(N^{-2/d})) \\ &= \frac{c}{\beta} + O(N^{-2/d}). \end{aligned}$$

Proof of Theorem 1.

Denote the r-th binomial moment of $\sum X_A$ by U_r, that is,

$$U_r = \sum E(X_{A_1} X_{A_2} \dots X_{A_r}),$$

where the summation is over all r-collections $\{A_1, A_2, \dots, A_r\}$ of edges of $G_{\mathcal{E}}$. Then, by Lemma 4, it will be sufficient to show that $U_r \to (c/\beta)^r / r!$ as $N \to \infty$. We show that

$$U_r = \frac{(c/\beta)^r}{r!} + O(N^{-2/d}) + O(N^{-1/2}).$$

Let $\mathcal{F}_r$ denote the set of those r-collections $\{A_1, A_2, \dots, A_r\}$ such that $G(A_1, A_2, \dots, A_r)$ are forests. By Lemma 2, we have $|\mathcal{F}_r| = F_{r,\mathcal{E}}/r! = N^r/r! + O(N^{r-1})$. Since

$$E(X_{A_1} \dots X_{A_r}) = \varphi(\alpha)^r$$

for $\{A_1, \dots, A_r\} \in \mathcal{F}_r$ by Lemma 1, the contribution of the elements in $\mathcal{F}_r$ to U_r is

$$\begin{aligned} \frac{N^r}{r!}(1 + O(N^{-1}))\varphi(\alpha)^r &= \frac{(N\varphi(\alpha))^r}{r!}(1 + O(N^{-1})) \\ &= \frac{(c/\beta + O(N^{-2/d}))^r}{r!}(1 + O(N^{-1})) \\ &= \frac{(c/\beta)^r}{r!} + O(N^{-2/d}) + O(N^{-1}). \end{aligned}$$

For $\{A_1, \dots, A_r\}$ not contained in $\mathcal{F}_r$, let $\gamma = \gamma(A_1, \dots, A_r)$ denote the *girth* of $G(A_1, \dots, A_r)$, that is, the length of the minimum cycle in $G(A_1, \dots, A_r)$, and let $\nu = \nu(A_1, \dots, A_r)$ denote the number of edges in a maximal spanning forest of $G(A_1, \dots, A_r)$. Then, clearly $\gamma \le \nu$. For each (p, q, r), $3 \le p \le q < r$, let

$$\mathcal{E}(p, q; r)$$

denote the set of those $\{A_1, \dots, A_r\}$ for which $\gamma = p, \nu = q$. If $\{A_1, \dots, A_r\} \in \mathcal{E}(p, q; r)$, then the number of vertices in $G(A_1, \dots, A_r)$ is at most $2q$. Hence, by Lemma 3,

$$|\mathcal{E}(p, q; r)| \leq \binom{2^{p/2}}{2p} N^{p/2} \cdot N^{q-(p-1)} \cdot \binom{\binom{2q}{2}}{r - q - 1} = O(N^{q-(p/2-1)}).$$

Since

$$E(X_{A_1} \dots X_{A_r}) \leq \varphi(\alpha)^q,$$

for $\{A_1, \dots, A_r\} \in \mathcal{E}(p, q; r)$, the contribution of the elements in $\mathcal{E}(p, q; r)$ to U_r is at most

$$O(N^{q-(p/2-1)})\varphi(\alpha)^q = O(N^{-(p/2-1)}).$$

Therefore

$$U_r = \frac{(c/\beta)^r}{r!} + O(N^{-2/d}) + O(N^{-1/2}).$$

□

Acknowledgment. Thanks to Professor Satoshi Kuriki (The Institute of Statistical Mathematics, Tokyo) who aroused my interest in this subject with many useful comments.

References

[1] P. Frankl and H. Maehara, Some geometric applications of the beta distribution, *Ann. Inst. Statist. Math.* 42(1990) 463–474.

[2] M. G. Kendall, *The Geometry of n Dimensions*, Griffin, London, 1961.

[3] H. Maehara, On the intersection graphs of random arcs on a circle. In *Random Graphs '87*, ed. M. Karonski *et al*, John Wiley, New York 1990, 159–173.

[4] H. Maehara, An inequality on the size of a set in a Cartesian product, preprint.

[5] P. A. P. Moran, Testing the largest of a set of correlation coefficients, *Austral. J. Statist.* 22(1980) 289–297.

[6] E. M. Palmer, *Graphical Evolution.* John Wiley, New York 1985.