Nearest-Neighbor Graphs

STEVEN FINCH

April 25, 2008

Consider a set \( \mathcal{P} \) of \( n \) points that are independently and uniformly distributed in the \( d \)-dimensional unit cube. Let \( p \in \mathcal{P} \). There exists almost-surely \( q \in \mathcal{P} \) such that \( q \neq p \) and \(|p - q| < |p - r|\) for all \( r \in \mathcal{P}, r \neq p, r \neq q \). The point \( q \) is called the nearest neighbor of \( p \) and we write \( p \prec q \). Note that \( p \prec q \) does not imply \( q \prec p \).

Draw an edge connecting \( p \) and \( q \) if and only if \( p \prec q \); the resulting graph of \( n \) vertices and \( \leq n \) edges is called the nearest-neighbor graph \( \mathcal{N} \) on \( \mathcal{P} \).

What is the probability, \( \alpha(d) \), given \( p \in \mathcal{P} \), that \( p \prec q \) implies \( q \prec p \)? Such a pair is isolated from the rest of \( \mathcal{N} \), in the sense that the only edge touching \( p \) or \( q \) is the edge that connects \( p \) and \( q \). We have \( [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15] \)

\[
\alpha(1) = \frac{2}{3}, \quad \alpha(2) = \frac{6\pi}{8\pi + 3\sqrt{3}} = 0.6215048968..., \quad \alpha(3) = \frac{16}{27}
\]

and, more generally \( [9] \),

\[
\alpha(d) = \begin{cases} 
\frac{3}{2} + \frac{1}{2} \sum_{k=1}^{\ell} \frac{1 \cdot 3 \cdot \ldots \cdot (2k-1)}{2 \cdot 4 \cdot \ldots \cdot (2k)} \left( \frac{3}{4} \right)^k & \text{if } d = 2\ell + 1, \\
\frac{4}{3} + \frac{\sqrt{3}}{2\pi} \left( 1 + \sum_{k=1}^{\ell-1} \frac{1 \cdot 3 \cdot \ldots \cdot (2k-1)}{3 \cdot 5 \cdot \ldots \cdot (2k+1)} \left( \frac{3}{4} \right)^k \right)^{-1} & \text{if } d = 2\ell.
\end{cases}
\]

Here is a variation of the preceding. Draw an edge connecting \( p \) and \( q \) if and only if \( q \prec p \); the resulting graph of \( n \) vertices and \( \leq n \) edges is called the nearest-neighbor anti-graph \( \mathcal{A} \) on \( \mathcal{P} \). What is the probability, \( \beta(d) \), that \( p \in \mathcal{P} \) is isolated from the rest of \( \mathcal{A} \)? That is, what proportion of points in \( \mathcal{P} \) are not nearest neighbors of any other points? We have \( [16, 17, 18, 19, 20, 21] \)

\[
\beta(1) = \frac{1}{4}, \quad \beta(2) \approx 0.28, \quad \beta(3) \approx 0.30
\]

but the latter two estimates are only simulation-based. To further understand \( \beta(2) \) will occupy us for the remainder of this essay.

\(^{0}\)Copyright © 2008 by Steven R. Finch. All rights reserved.
Define constants $C(0, d) = 1$ and

$$C(k, d) = \int_{\Omega(k, d)} \exp \left[ -\text{Vol} \left( \bigcup_{j=1}^{k} S(x_j) \right) \right] dx_1 dx_2 \ldots dx_k$$

for $k \geq 1$, where $S(x_j)$ is the ball in $\mathbb{R}^d$ of radius $|x_j|$, centered at $x_j$, and

$$\Omega(k, d) = \left\{ (x_1, x_2, \ldots, x_k) \in \mathbb{R}^{dk} : |x_i| \leq |x_i - x_j| \text{ for all } 1 \leq i \neq j \leq k \right\}.$$

It is known that \[19, 22, 23, 24, 25\]

$$\beta(2) = \sum_{k=0}^{6} \frac{(-1)^k}{k!} C(k, 2), \quad \beta(3) = \sum_{k=0}^{12} \frac{(-1)^k}{k!} C(k, 3)$$

and clearly $C(1, d) = 1$, $C(2, 1) = 1/2$. The upper limits of summation are the kissing numbers in $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. A proof that $24$ is the kissing number in $\mathbb{R}^4$ was given only recently \[26, 27\]. Also, $C(6, 2) = 0$ since $\Omega(6, 2)$ is of measure zero.

Henze \[24, 25\] showed that

$$C(2, d) = \frac{2^{d+1} \pi^{d-1}}{\Gamma(d-1)} \int_0^\infty \int_0^\pi \int_0^\xi \xi^{d-1} \eta^{d-1} \sin(\theta)^{d-2} F_d(\xi, \eta) \, d\theta \, d\eta \, d\xi$$

where

$$\theta_0 = \arccos \left( \frac{\eta}{2\xi} \right),$$

$$F_d(\xi, \eta) = \exp \left[ -f_d(\xi, \gamma) - f_d(\eta, \delta) \right],$$

$$\gamma = \frac{\xi(\xi - \eta \cos(\theta))}{\sqrt{\xi^2 + \eta^2 - 2\xi \eta \cos(\theta)}}, \quad \delta = \frac{\eta(\eta - \xi \cos(\theta))}{\sqrt{\xi^2 + \eta^2 - 2\xi \eta \cos(\theta)}},$$

$$f_d(x, y) = \frac{\pi^{d/2} x^d}{2 \Gamma(d/2 + 1)} \left[ 1 + I \left( \frac{y^2}{x^2}, \frac{1}{2}, \frac{d + 1}{2} \right) \right]$$

and $I$ is the regularized beta function

$$I(z, a, b) = \frac{\Gamma(a + b)}{\Gamma(a) \Gamma(b)} \int_0^z w^{a-1} (1 - w)^{b-1} \, dw.$$

(In \[24\], the definitions of $\gamma$ and $\delta$ were mistakenly reversed; also, the expression within square brackets for $f_d(x, y)$ was unclear.) We obtain

$$C(2, 2) = 0.63317 \ldots = 2(0.316585 \ldots), \quad C(2, 3) = 0.70888 \ldots.$$
Tao & Wu [19] independently showed that

\[
C(2, 2) = \pi \int_{\pi/2}^{3\pi/2} d\theta \int_0^{2\tau} \frac{\tau}{(g(\tau, \theta) + \tau^2h(\tau, \theta))^2} d\tau
\]

where

\[
\begin{align*}
\varphi &= \arcsin \left( \frac{\tau \sin(\theta)}{\sqrt{1 + \tau^2 - 2\tau \cos(\theta)}} \right), \\
\psi &= \arcsin \left( \frac{\sin(\theta)}{\sqrt{1 + \tau^2 - 2\tau \cos(\theta)}} \right).
\end{align*}
\]

(In [19], the absolute value signs in the definitions of \(\varphi\) and \(\psi\) were mistakenly omitted.) Even more elaborate integral formulas apply for \(C(3, 2), C(4, 2), C(5, 2)\). Given the discrepancy between our estimate of \(C(2, 2)\) and their estimate (see the Table), it seems doubtful that their approximation \(\beta(2) = 0.284051\ldots\) is entirely correct.

Table 1 Old and New Calculations of Constants

<table>
<thead>
<tr>
<th>(k)</th>
<th>Tao &amp; Wu estimate of (C(k, 2)/k!)</th>
<th>Current estimate of (C(k, 2)/k!)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.3163335...</td>
<td>0.316585...</td>
</tr>
<tr>
<td>3</td>
<td>0.0329390...</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0.0006575...</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0.0000010...</td>
<td>?</td>
</tr>
</tbody>
</table>

A discrete version of the latter problem appears in [28, 29, 30, 31]. Let all the vertices of the lattice \(\mathbb{Z}^d\) be initially occupied by particles which can annihilate one-by-one their \(2d\) nearest neighbors. More precisely, for each unit-length edge \(\{u, v\}\) of the lattice, there is a Uniform \([0, 1]\) random variable \(T_{\{u, v\}}\) representing the time of an attack along the edge. If vertices \(u, v\) are both occupied immediately prior to time \(T_{\{u, v\}}\), then at time \(T_{\{u, v\}}\) either vertex \(u\) or vertex \(v\) (each with probability \(1/2\)) becomes vacant (that is, one particle annihilates the other). If \(u, v\) are not both occupied at time \(T_{\{u, v\}}\), then there is no change. Once a vertex becomes vacant, it remains vacant permanently. The variables \(T_{\{u, v\}}\), considered over all unit-length edges \(\{u, v\}\), are independent. By time 1, no two surviving particles can be adjacent. When \(d = 1\), the probability that a given vertex remains occupied is \(1/e = 0.3678794411\ldots\). When \(d = 2\), this probability is known to be in the interval \((0.227, 0.306)\) and is approximately 0.25 via simulation. Greater accuracy is desired.
0.1. Appeal for Help. Any assistance in completing Table 1, using the formulation in [19], would be deeply appreciated! We note similarity between this problem and others in [32, 33].

References


