Random Geometric Graphs

A collection of notes from my time working under the supervision of Mathew Penrose on random geometric graph theory and related topics.

Jacob Green
Department of Mathematical Sciences, The University of Bath

October 14, 2024

Abstract

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

Keywords: Random geometric graphs, Poisson Point Processes

Contents

1 Introduction

1.1 What is a random geometric graph?

Definition 1.1 (geometric graph)

Let $\mathcal{X} \subset \mathbb{R}^d$ be a finite collection of points and $r > 0$. The *geometric graph* $G(\mathcal{X}, r)$ is the graph with vertex set $\mathcal{X}$ and edge set $\{\{x, y\} \subset \mathcal{X} \ : \ |x - y| \leq r\}$, where $|\cdot|$ denotes the Euclidean norm.

We can turn this into a random graph by letting the set $\mathcal{X} \subset \mathbb{R}^d$ be random. The resulting structure $G(\mathcal{X}, r)$ is said to be the *random geometric graph* (A.K.A the *Gilbert graph*). The first structure we'll consider is found by uniformly scattering n points in the d-dimensional hypercube $[0,1]^d$.

Example 1.2 (fixed scale geometric graph)

Let $\xi_1, \xi_2, \ldots \overset{\text{i.i.d}}{\sim} \text{Unif}[0,1]^d$ and fix a sequence $(r_n)_{n\geq 1}$ in $\mathbb{R}_{>0}$. Let $\mathcal{X}_n := \{\xi_1, \ldots, \xi_n\}$. Then $G(\mathcal{X}_n, r_n)$ is the *fixed scale geometric graph* at time $n \geq 1$.

We will also consider what happens when the number of vertices is Poisson.

Example 1.3 (Poisson scale geometric graph)

Let $\xi_1, \xi_2, \ldots \overset{\text{i.i.d}}{\sim} \text{Unif}[0,1]^d$ and $N_n \sim \text{Poisson}(n)$ for $n \geq 1$ be independent of $(\xi_1, \xi_2, \ldots)$. Let $\mathcal{P}_n := \{\xi_1, \ldots, \xi_{N_n}\}$ and fix a sequence $(r_n)_{n\geq 1}$ in $\mathbb{R}_{>0}$. Then the random geometric graph $G(\mathcal{P}_n, r_n)$ is the d-dimensional *Poisson scale geometric graph* at time $n \geq 1$.

As one may expect, the vertices of this graph form a Poisson point process in $\mathbb{R}^d$. This was originally an exercise (1.1) in [1].

Proposition 1.4 ($\mathcal{P}_n(\cdot)$ is a Poisson point process)

Let $\mathcal{P}_n$ be the vertex set in the Poisson scale geometric graph. Then $\mathcal{P}_n(\cdot) : A \mapsto |\mathcal{P}_n \cap A|$, where $A \subset [0,1]^d$ is measurable and $|\cdot|$ is the counting measure, is a *Poisson point process* with intensity $\lambda(A)$, $\lambda(\cdot)$ the Lebesgue measure on $\mathbb{R}^n$.

See Appendix A for the definition of the Poisson point process.

Proof. Write $\mathcal{P}_n(A_j) = \sum_{i=1}^{N_n} \mathbf{1}\{\xi_i \in A_j\}$. Then, conditioning on $\{N_n = m\}$, each $\mathcal{P}_n(A_j)$ is a $\text{Binom}(m, \lambda(A_j))$ random variable, making the joint distribution (also conditioned on $\{N_n = m\}$) a $\text{Multinom}(m, \lambda(A_1), \ldots, \lambda(A_k))$. Thus one has

$$\begin{aligned}
\mathbb{P}(\mathcal{P}_n(A_1) = a_1, \ldots, \mathcal{P}_n(A_k) = a_k) &= \underbrace{\left(\frac{e^{-n}n^m}{m!}\right)}_{\text{Poisson}(n;m)} \times \underbrace{\binom{m}{j_1, \ldots, j_k} \prod_{i=1}^{k} (\lambda(A_i))^{j_i}}_{\text{Multinom}(m, \lambda(A_1), \ldots, \lambda(A_k); j_1, \ldots, j_k)} \\
&= \prod_{i=1}^{k} \frac{e^{-n\lambda(A_i)}(n\lambda(A_i))^{j_i}}{(j_i)!}
\end{aligned}$$

which is a product of the distributions $\text{Poisson}(n\lambda(A_i)), 1 \leq i \leq k$, giving us both of our defining properties of a Poisson point process. $\square$

The exercise (1.2) in [1] states the following result, which we'll prove in an analogous way.

Proposition 1.5 (expected degree of k^{th} vertex in fixed scale RGG)

Consider the fixed scale random geometric graph $G(\mathcal{X}_n, r_n)$ with $r_n \to 0$ as $n \to \infty$. Let $D_{k,n}$ be the degree of the first vertex ξ_k. Prove that $\mathbb{E}[D_{k,n}] \sim \theta n r_n^d$

Proof. Write $D_{k,n} = \sum_{i \neq k} \mathbf{1}\{|\xi_k - \xi_i| \leq r\}$, then we have $D_{k,n} \sim \text{Binom}(n-1, p)$ with

$$p = \lambda(B_{r_n}(\xi_k) \cap [0,1]^d)$$

where $B_{r_n}(\xi_k)$ is the hypersphere radius r_n centred at ξ_k and λ the Lebesgue measure on $\mathbb{R}^d$. Now, conditioning on ξ_k, we have

$$\mathbb{E}[D_{k,n}|\xi_k] = (n-1)\lambda(B_{r_n}(\xi_k) \cap [0,1]^d)$$

and hence, by the total law of expectation,

$$\begin{aligned}
\mathbb{E}[D_{k,n}] &= \mathbb{E}[\mathbb{E}[D_{k,n}|\xi_k]] \\
&= (n-1) \int_{[0,1]^d} \lambda(B_{r_n}(\xi_k) \cap [0,1]^d) \mathrm{d}\xi_k \\
&\sim n \int_{[0,1]^d} \lambda(B_{r_n}(\xi_k)) \mathrm{d}\xi_k \sim n(\theta r_n^d)
\end{aligned}$$

Where the first asymptotic equality follows via the dominated convergence theorem (to see this, observe $B_{r_n}(\xi_k) \cap [0,1]^d \to B_{r_n}(\xi_k)$ A.E. and $B_{r_n}(\xi_k)$ also dominating). Note θ is a constant depending on d, and is the coefficient of r_n^d in the formula for the volume of a hypersphere radius r_n. $\square$

1.2 Counting Edges of $G(\mathcal{X}_n, r_n)$ and $G(\mathcal{P}_n, r_n)$

The number of edges $\mathcal{E}_n, \mathcal{E}'_n$ in $G(\mathcal{X}_n, r_n)$ and $G(\mathcal{P}_n, r_n)$ respectively is a Poisson random variable. To prove this, we'll use a so called *dependency graph*. We handle the two RGGs $G(\mathcal{X}_n, r_n)$ and $G(\mathcal{P}_n, r_n)$ seperately.

Counting edges of $G(\mathcal{X}_n, r_n)$

An elementary question to ask is "what is the expected edge count at time n?". Via the previously derived vertex degree asymptotics, we have an easy asymptotic result.

Proposition 1.6 (expected edge count in fixed scale RGG)

Let $\mathcal{E}_n$ and $\mathcal{E}'_n$ be the number of edges in $G(\mathcal{X}_n, r_n)$ and $G(\mathcal{P}_n, r_n)$ respectively. Then,

$$\mathbb{E}[\mathcal{E}_n] \sim \theta n^2 r_n^d / 2$$

Proof of lemma. By the handshaking lemma we have $2\mathbb{E}[\mathcal{E}_n] = \sum_{k=1}^n \mathbb{E}[D_{k,n}] \sim n^2 r_n^d$. $\square$

Now we introduce the *dependency graph*.

Definition 1.7 (dependency graph)

Let $(V, \sim)$ be a finite simple graph w/ edge relation $\sim$ and vertex set V. We say $(V, \sim)$ is a *dependency graph* for the random variables $(W_\alpha)_{\alpha \in V}$ if whenever $A, B \subset V$ are disjoint with no $\alpha \in A$, $\beta \in B$ such that $\alpha \sim \beta$ (i.e. A and B lie in unique connected components) then

$$(W_\alpha)_{\alpha \in A} \perp\!\!\!\perp (W_\beta)_{\beta \in B} \quad \text{i.e. are independent}$$

When we take the random variable associated with each vertex to be 0-1, we obtain the following bound.

Lemma 1.8 (Poisson approximation lemma for Bernoulli sums)

Let $(\xi_i)_{i \in I}$ be a finite collection of Bernoulli random variables with dependency graph $(I, \sim)$. Set $p_i := \mathbb{P}(\xi_i = 1)$, $p_{ij} := \mathbb{P}(\xi_i = 1; \xi_j = 1)$ and suppose $\lambda := \sum_{i \in I} p_i$ is finite. Then, letting $W := \sum_{i \in I} \xi_i$, we have (LHS is just total variation distance)

$$\sum_{k \geq 0} \Big| \mathbb{P}[W = k] - \mathbb{P}(\text{Po}(\lambda) = k) \Big| \leq \min\{6, 2\lambda^{-1}\} \left(\sum_{i \in I} \sum_{\mathcal{N}(i) \setminus \{i\}} p_{ij} + \sum_{i \in I} \sum_{j \in \mathcal{N}(i)} p_i p_j \right)$$

Proof. Omitted, consult [3]. ◻

Theorem 1.9 (total variation between edge and Poisson distribution)

Let $\lambda_n := \mathbb{E}[\mathcal{E}_n]$. Then

$$\sum_{k\geq 0} |\mathbb{P}(\mathcal{E}_n = k) - \mathbb{P}(\mathrm{Po}(\lambda_n) = k)| = O(nr^d)$$

The idea of this proof is to write $\mathcal{E}_n = \sum_{1\leq i<j\leq n} \mathbf{1}\{|\xi_i - \xi_j| \leq r_n\}$, i.e. the number of edges as the sum of the indicators over all possible edges indicating whether this edge exists. This has a rather obvious dependency graph and hence we can apply the Poisson approximation lemma (1.7) with some work on the asymptotics.

Proof. Let $V\{\{i,j\} : 1 \leq i < j \leq\}$ and define $\sim$ by $\alpha \sim \beta$ if $\alpha \cap \beta \neq \emptyset$ and $\alpha \neq \beta$. Define the random variables $W_\alpha = \mathbf{1}\{|\xi_i - \xi_j| \leq r_n\}$ for $\alpha \in V$ and let $\lambda_n := \sum_{\alpha\in V} p_\alpha$.

Claim: $G(V, \sim)$ is a dependency graph for $(W_\alpha)_{\alpha\in V}$.

Proof of claim. Follows immediately from the independence of the ξ_i. ◻

Let $p_\alpha := \mathbb{P}(W_\alpha = 1)$ and $p_{\alpha\beta} := \mathbb{P}(W_\alpha = 1; W_\beta = 1)$, as in the setup of lemma 1.7.

Claim: $p_\alpha \sim \theta r_n^d$

Proof of claim. Fix $\alpha = \{i,j\} \in V$. Then, letting $\lambda(\cdot)$ be the Lebesgue measure on $\mathbb{R}^d$,

$$\begin{aligned} p_\alpha = \mathbb{P}(|\xi_i - \xi_j| \leq r_n) &= \int_{v\in[0,1]^d} \mathbb{P}(|\xi_i - \xi_j| \leq r_n | \xi_i = v)\mathrm{d}v \\ &= \int_{v\in[0,1]^d} \lambda(B_{r_n}(v) \cap [0,1]^d)\mathrm{d}v \\ &\sim \int_{v\in[0,1]^d} \lambda(B_{r_n}(v))\mathrm{d}v \sim \theta r_n^d \end{aligned}$$

where the first asymptotic equality follows from the dominated convergence theorem. ◻

Claim: $p_{\alpha\beta} \sim (\theta r_n^d)^2$

Proof of claim. An analogous argument works. ◻

Now we have all our ingredients, lets cook. Observe, from the previously computed asymptotics,

$$\lambda_n \sim n^2\theta r_n^d/2 \quad \text{and} \quad \sum_{\alpha\in V} p_\alpha^2 \sim \theta\lambda_n r_n^d$$

and hence by counting $|\mathcal{N}(\alpha)| = 2(n-2)$, obtain

$$\sum_{\alpha\in V}\sum_{\beta\sim\alpha} (p_{\alpha\beta} + p_\alpha p_\beta) \sim \binom{n}{2} \times 2(n-2) \times 2(\theta r_n^d)^2 = O(n\lambda_n r_n^d)$$

which by the Poisson approximation lemma gives

$$\sum_{k\geq 0} |\mathbb{P}(\mathcal{E}_n = k) - e^{-\lambda_n}\lambda_n^k/k!| = O(nr_n^d)$$

as claimed. ◻

Corollary 1.10 (edge distribution of the fixed scale random geometric graph)

If the limit $\lambda_n \to \lambda \in (0, \infty)$ exists and $nr_n^d \to 0$, then $\mathcal{E}_n \xrightarrow{\mathcal{D}} \text{Poisson}(\lambda)$

Proof. By proposition 1.6, we have $\lambda \sim \lambda_n \sim \theta n^2 r_n^d/2 \Rightarrow n^2 r_n^d \sim 2\lambda/\theta \in (0, \infty)$ which forces $nr_n^d \to 0$ and hence lemma 1.8 applies giving $\mathcal{E}_n \xrightarrow{\mathcal{D}} \text{Poisson}(\lambda)$. □

Counting edges of $G(\mathcal{P}_n, r_n)$

We now spend the rest of the section proving similar results for $G(\mathcal{P}_n, r_n)$. Before proceeding we have the following lemma from [1]. We write $\mathcal{P}_{<\infty}(A)$ for the family of finite subsets of A.

Lemma 1.11 (Mecke formula)

Let $k \in \mathbb{N}$. For any measurable $f : (\mathbb{R}^d)^k \times \mathcal{P}_{<\infty}([0,1]^d) \to \mathbb{R}$, when the expectation exists,

$$\mathbb{E} \sum_{X_1,\ldots,X_k \in \mathcal{P}_n}^{\neq} f(X_1, \ldots, X_k, \mathcal{P}_n \setminus \{X_1, \ldots, X_n\}) = n^k \int \cdots \int \mathbb{E} f(x_1, \ldots, x_k, \mathcal{P}_n) \mathrm{d}x_1 \ldots \mathrm{d}x_k$$

where $\sum^{\neq}$ denotes the sum over the ordered k-tuples of distinct points (in $\mathcal{P}_n$).

Proof. The idea is to condition on the number of points in $\mathcal{P}_n$. Observe,

$$\begin{aligned}
&\mathbb{E} \sum_{X_1,\ldots,X_k \in \mathcal{P}_n}^{\neq} f(X_1, \ldots, X_k, \mathcal{P}_n \setminus \{X_1, \ldots, X_n\}) \\
&= \mathbb{E}\left[\mathbb{E}\left[\sum_{X_1,\ldots,X_k \in \mathcal{P}_n}^{\neq} f(X_1, \ldots, X_k, \mathcal{P}_n \setminus \{X_1, \ldots, X_n\}) \middle| |\mathcal{P}_n| = m\right]\right] \\
&= \sum_{m \geq k} \left(\frac{e^{-n} n^m}{m!}\right) \sum_{X_1,\ldots,X_k \in \{\xi_1,\ldots,\xi_m\}}^{\neq} \mathbb{E} f(X_1, \ldots, X_k, \{X_{k+1}, \ldots, X_m\}) \\
(\dagger) \qquad &= \sum_{m \geq k} \left(\frac{e^{-n} n^m}{m!}\right) \left(\prod_{i=1}^{k} (m - i + 1)\right) \mathbb{E} f(X_1, \ldots, X_k, \{X_{k+1}, \ldots, X_m\}) \\
&= \sum_{m \geq k} \left(\frac{e^{-n} n^m}{m!}\right) \left(\prod_{i=1}^{k} (m - i + 1)\right) \int_{[0,1]^d} \cdots \int_{[0,1]^d} f(x_1, \ldots, x_k, \{x_{k+1}, \ldots, x_m\}) \mathrm{d}x_1 \ldots \mathrm{d}x_m \\
&= n^k \sum_{m \geq k} \left(\frac{e^{-n} n^{m-k}}{(m-k)!}\right) \int_{[0,1]^d} \cdots \int_{[0,1]^d} f(x_1, \ldots, x_k, \{x_{k+1}, \ldots, x_m\}) \mathrm{d}x_1 \ldots \mathrm{d}x_m \\
&= n^k \int_{[0,1]^d} \cdots \int_{[0,1]^d} f(x_1, \ldots, x_k, \{x_{k+1}, \ldots, x_m\}) \mathrm{d}x_1 \ldots \mathrm{d}x_m \\
&= n^k \int_{[0,1]^d} \cdots \int_{[0,1]^d} \mathbb{E} f(x_1, \ldots, x_k, \mathcal{P}_n) \mathrm{d}x_1 \ldots \mathrm{d}x_k
\end{aligned}$$

Where ($\dagger$) follows from the ξ_i being i.i.d, so all expectations are equal and it suffices to count distinct k-tuples in $\{\xi_1, \ldots, \xi_m\}$. [CHECK THIS PROOF W/ PENROSE] □

We can now use this lemma to find the expected number of edges in $G(\mathcal{P}_n, r_n)$.

Proposition 1.12 (expected edge count in Poisson scale RGG)

Let $G(\mathcal{P}_n, r_n)$ be the Poisson scale RGG and $\mathcal{E}_n$ be the number of edges at time n. Then,

$$\mathbb{E}[\mathcal{E}_n] \sim \theta n^2 r_n^d/2$$

Proof. Write $\mathcal{E}_n = \frac{1}{2}\sum^{\neq}_{X_1,X_2\in\mathcal{P}_n} \mathbf{1}\{\|X_1 - X_2\| \le r_n\}$. Then we can use Mecke's formula as follows.

$$\begin{aligned}
\mathbb{E}[\mathcal{E}_n] &= \frac{1}{2}\mathbb{E}\left[\sum^{\neq}_{X_1,X_2\in\mathcal{P}_n} \mathbf{1}\{\|X_1 - X_2\| \le r_n\}\right] \\
\text{(Mecke)} \quad &= \frac{n^2}{2}\int_{[0,1]^d}\int_{[0,1]^d} \mathbb{E}[\mathbf{1}\{\|x_1 - x_2\| \le r_n\}]\mathrm{d}x_1\mathrm{d}x_2 \\
&= \frac{n^2}{2}\int_{[0,1]^d}\int_{[0,1]^d} \mathbf{1}\{\|x_1 - x_2\| \le r_n\}\mathrm{d}x_1\mathrm{d}x_2 \\
&= \frac{n^2}{2}\int_{[0,1]^d} \lambda(B(x_1, r_n)\cap[0,1]^d)\mathrm{d}x_1 \sim \theta n^2 r_n^d/2
\end{aligned}$$

where the final asymptotic equality follows from the dominated convergence theorem and volume of a d-dimensional hypersphere. Note we have already seen the asymptotics for the final integral. ☐

Rest follows on from Exercise (2.2), too small brained to figure it out right now will attempt again tomorrow :<

1.3 A central limit theorem for our edge counts

2 Appendices

2.1 Appendix A: Poisson Point Processes

For a full treatment, consult [2]. Here I will simply give the relevent definitions and results from this text, leaving proofs (unless containing a particularly important idea) to [2].

Definition 2.13 (point process)

Let $(\mathbb{X}, \mathcal{X})$ be a measure space and let $\mathbf{N}(\mathbb{X}) \equiv \mathbf{N}$ be the family of measures that can be written as a countable sum of finite measures on $\mathbb{X}$ with image in $\mathbb{N}_0$. Let $\mathcal{N}(\mathbb{X}) \equiv \mathcal{N}$ be the σ-algebra generated by the collection of subsets $\{\mu \in \mathbf{N} : \mu(A) = k\}$ over $A \in \mathcal{X}, k \in \mathbb{N}_0$. A *point process* on $\mathbb{X}$ is a random element η of the measure space $(\mathbf{N}, \mathcal{N})$

Definition 2.14 (Poisson point process)

Let $\mathbb{X}$ be a space and λ an s-finite measure on $\mathbb{X}$. A *Poisson point process* with intensity measure λ is a point process η on $\mathbb{X}$ with

(i) $\eta(A) \sim \text{Poisson}(\lambda(A))$, that is $\eta(A)$ is Poisson with parameter $\lambda(A)$

(ii) If $A_1, \ldots, A_n \in \mathbb{X}$ are pairwise disjoint then $\eta(A_1), \ldots, \eta(A_n)$ are independent.

References

[1] Michael Krivelevich, Konstantinos Panagiotou, Mathew Penrose, and Colin McDiarmid. *Random Graphs, Geometry and Asymptotic Structure*. London Mathematical Society Student Texts. Cambridge University Press, 2016.

[2] Günter Last and Mathew Penrose. *Lectures on the Poisson Process*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017.

[3] Mathew Penrose. *Random Geometric Graphs*. Oxford University Press, 05 2003.