Random Geometric Graphs

Jacob Green
Department of Mathematical Sciences, The University of Bath

December 7, 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

Preface

Contents

1 Introduction

Consider a (finite) collection of points in space, $X \subset \mathbb{R}^d$. We wish to model the situation where points $x, y \in X$ that are "close" are connected. In the graph theoretic sense this would be done by connecting x and y with an edge. Given we are in Euclidean space (but noting this logic can be extendend to arbitrary metric spaces) we can define "closeness" with the usual metric on $\mathbb{R}^d$, the Euclidean norm $\|\cdot\|$. Hence we connect $x \sim y$ iff $\|x - y\| \leq r$ for some $r > 0$. The resulting graph, the so called *geometric graph*, is denoted $G(X, r)$.

How do we make this random? In the classical Erdős-Renyi random graph our randomness comes from the connection criterion on vertices. Here our notion of connectivity is deterministic, so we instead allow the points (vertices) $X \subset \mathbb{R}^d$ to be random. Some classic (and the examples we'll study in this handout) examples are $X = (X_1, \ldots, X_n) \sim \mathrm{Unif}[0,1]^d$ and $X \sim \mathrm{PPP}_{[0,1]^d}(\lambda)$, that is X follows a Poisson point process intensity λ in the unit hypercube $[0,1]^d$.

A classic use case for this model would be the spread of infection in a plant population. If we know roughly the density of plants in a given habitat and how close two plants must be for a given disease to spread between them, we may model the spread of this disease with a random geometric graph (with Poisson vertices). Then we may answer questions like "What is the maximum number of plants that could be affected by releasing this disease into the population?" by looking at its connected components, and other similar questions by looking at graph theoretic properties of the random geometric graph.

In this handout we are concerned with sequences of random geometric graphs on the hypercube $[0,1]^d$. We will work to prove several asymptotic properties of these graphs in the limiting density ($n \to \infty$ or $\lambda \to \infty$ in our previous examples). Let us define our sequences of interest.

Definition 1.1 (fixed/poisson vertex random geometric graphs)

Let $n \in \mathbb{N}$ and let $\xi_1, \xi_2, \ldots$ be independent points in $\mathbb{R}^d$. Let $N_n \sim \mathrm{Po}(n)$ independently of $\xi_1, \xi_2, \ldots$ and consider the two point sets

$$\mathcal{X}_n = \{\xi_1, \ldots, \xi_n\} \qquad \text{and} \qquad \mathcal{P}_n = \{\xi_1, \ldots, \xi_{N_n}\}$$

Let $(r_n)_{n \geq 1}$ be a sequence of positive reals. Then we define, respectively, for $n \geq 1$ the *fixed vertex random geometric graphs* and *Poisson vertex random geometric graphs* as

$$G(\mathcal{X}_n, r_n) \qquad \text{and} \qquad G(\mathcal{P}_n, r_n)$$

We will study properties as $n \to \infty$ in the various cases on the size of r_n as $n \to \infty$. The three types of limit in particular we care about are the so called *sparse limit* where $nr_d^n \to 0$, the *thermodynamic limit* $nr_n^d = \Theta(1)$ and the *dense limit* $nr_n^d \to \infty$.

It is not immediately obvious that our Poisson vertex random geometric graph $G(\mathcal{P}_n, r_n)$ corresponds to a Poisson point process in $[0,1]^d$, as suggested earlier. Let us prove this fact.

Proposition 1.2 (fixed vertex RGG is a Poisson point process)

The set $\mathcal{P}_n$ defines a Poisson point process in $[0,1]^d$. That is, for Borel $A, A_1, \ldots, A_k \in [0,1]^d$

(I) $\mathcal{P}_n(A) \sim \mathrm{Po}(n\lambda(A))$

(II) $\mathcal{P}_n(A_1), \ldots, \mathcal{P}_n(A_k)$ are independent random variables for $A_1, \ldots, A_k$ disjoint

where $\mathcal{P}_n(X)$ denotes the number of points in $\mathcal{P}_n \cap X$ and $\lambda(\cdot)$ is the Lebesgue measure in $\mathbb{R}^d$.

Proof. Each ξ_i is in $A \subset [0,1]^d$ with probability $\lambda(A)$. Thus, conditioned on $N_n = m$, $\mathcal{P}_n(A)$ is a $\mathrm{Binom}(m, \lambda(A))$. Further, considering disjoint $A_1, \ldots, A_k \subset [0,1]^d$ and conditioning on $N_n = m$, we see that the random vector $(\mathcal{P}_n(A_1), \ldots, \mathcal{P}_n(A_k))$ is $\mathrm{Multinom}(k, (\lambda(A_1), \ldots, \lambda(A_k)))$. Suppose now that

$A_1, \ldots, A_k$ partition $[0,1]^d$, then, letting $m = \sum_{i=1}^k j_k$,

$$\mathbb{P}[\mathcal{P}_n(A_1) = j_1, \ldots, \mathcal{P}_n(A_k) = j_k] = \underbrace{\left(\frac{e^{-n} n^m}{m!}\right)}_{\mathrm{Po}(m)} \times \underbrace{\binom{m}{j_1, \ldots, j_k} \prod_{i=1}^k \lambda(A_i)^{j_i}}_{\mathrm{Multinom}(k, (\lambda(A_1), \ldots, \lambda(A_k)))}$$
$$= \prod_{i=1}^k \underbrace{\frac{e^{-n\lambda(A_i)}(n\lambda(A_i))^{j_i}}{j_i!}}_{\mathrm{Po}(n\lambda(A_i))}$$

are independent $\mathrm{Po}(n\lambda(A_i))$, which demonstrates the result. □

Before moving on to edge counts, let us prove one elementary property of random geometric graphs to get a feel for these objects.

Proposition 1.3 (expected vertex degrees)

Let $D_{i,n}$ denote the degree of vertex ξ_i in $G(\mathcal{X}_n, r_n)$ and $D'_{i,n}$ the degree of vertex i in $G(\mathcal{X}_n, r_n)$ suppose $r_n \to 0$. Then

$$\mathbb{E}[D_{i,n}], \mathbb{E}[D'_{i,n}] \sim \theta_d n r_n^d$$

as $n \to \infty$, where θ_d is the volume of the unit ball in $\mathbb{R}^d$.

Note that due to points near the boundary we do not get equality. If we were to condition on the event $B_{\xi_i}(r_n) \subset [0,1]^d$ for $n \geq 1$ we would have equality.

Proof. The idea is to use the law of total expectation and the dominated convergence theorem. Observe, conditioning on $\xi_i = x \in \mathbb{R}^d$, that $D_{i,n} \sim \mathrm{Binom}(n-1, \lambda(B_x(r_n) \cap [0,1]^d))$ where $B_x(r_n)$ is the d-ball centered at x with radius r_n under the Euclidean norm. Observe,

$$\begin{aligned} r_n^{-d}\mathbb{E}[D_{i,n}] &= \int_{x \in [0,1]^d} r_n^{-d}\mathbb{E}[D_{i,n} | \xi_i = x] \mathrm{d}x \\ &= (n-1) \int_{x \in [0,1]^d} r_n^{-d} \lambda(B_x(r_n) \cap [0,1]^d) \mathrm{d}x \\ &\sim n\theta_d \end{aligned}$$

We apply the dominated convergence theorem in the asymptotic equality, noting that

$$r_n^{-d}\lambda(B_x(r_n) \cap [0,1]^d) \to \theta_d$$

almost everywhere in $[0,1]^d$ and is clearly dominated by 1. The result follows immediately.

□

Remark 1.4 (applying the dominated convergence theorem in 1.3)

We multiply through by r_n^{-d} or our integrand $\lambda(B_x(r_n) \cap [0,1]^d)$ will tend pointwise to 0 and we will get $n^{-1}\mathbb{E}[D_{i,n}] \to 0$ which is true (assuming $r_n \to 0$) but uninteresting.

2 Edge Counts

Throughout this section we will denote the number of edges of $G(\mathcal{X}_n, r_n)$ by $\mathcal{E}_n$ and the number edges of $G(\mathcal{P}_n, r_n)$ by $\mathcal{E}'_n$. Our goal is to find the limiting distribution of these two quantities.

Before doing so, we'll first investigate the expected number of edges in such a graph. We apply a classical lemma from graph theory to find the expected number of edges in $G(\mathcal{X}_n, r_n)$. Then for the expected number of edges in $G(\mathcal{P}_n, r_n)$ we'll use properties of the conditional expectation.

Proposition 2.0 (expected edge counts)

In the notation outlined above,
$$\mathbb{E}[\mathcal{E}_n], \mathbb{E}[\mathcal{E}'_n] \sim n^2 r_d^n \theta_d / 2$$

Proof. By the handshaking lemma, linearity of expectations and proposition 1.3, one has

$$2\mathbb{E}[\mathcal{E}_n] = \sum_{i=1}^{n} \mathbb{E}[D_{1,n}] \sim n^2 \theta_d r_n^d$$

and the result follows by dividing through by 2. Now by the total law of expectations one has

$$\mathbb{E}[\mathcal{E}'_n] = \mathbb{E}[\mathbb{E}[\mathcal{E}'_n | N_n = m]] =$$

□

2.1 Edge Distribution

2.2 Normal Approximation

3 Maximal Degree