UNIVERSALITY IN INNER-PRODUCT RANDOM GEOMETRIC GRAPHS

DONALD STRALKUS

1. Introduction

This project concerns latent estimation within the random geometric graph model:

Definition 1.1. *Let μ be a probability measure on $\mathbb{R}^d$ and $X_1, ..., X_n$ be i.i.d samples from μ. Further, let $p : \mathbb{R}^d \times \mathbb{R}^d \to [0, 1]$ be a connection function. Then, a sample from the random geometric graph model $G(n, p(\cdot, \cdot), \mu)$ is an n-vertex graph $G = (V, E)$ where $(i, j) \in E$ with probability $p(X_i, X_j)$.*

See [4] for a survey on random geometric graphs. Throughout, we will assume an inner-product based connection function: $p(X_j, X_k) = \mathbb{1}_{\langle X_j, X_k \rangle \geq t}$ for some threshold t, chosen such that the probability of an edge is p for a parameter $p \in [0, 1]$. Denote this model as $G(n, p, \mu)$. We can equivalently sample a graph in the following way: Draw n samples of μ and construct A (d x n). Then, obtain $A^T A$ and replace each entry with a 1 if it is greater than t and 0 otherwise. Then, remove the 1's on the diagonal. Henceforth, $G \sim G(n, p, \mu)$ will be sampled in this way. The overarching goal of the project is to explore when underlying geometries can be efficiently inferred from observed graphs. At a high-level, the strategy to do so is as follows: For a measure μ, find a measure μ' such that the graphs produced by μ and μ' are similar but inference algorithms are easier to develop on μ'. As a result, it makes sense to study universality properties of the model, which is what this project is exploring.

2. Universality

Conjecture 2.1. *Let $\mathcal{P}$ be the space of probability measures on $\mathbb{R}$ such that all $\mu \in \mathcal{P}$ have mean 0 and unit variance and let $\mathcal{P}^d$ be all such i.i.d product measures of measures in $\mathcal{P}$. Then, $G(n, p, \mu)$ is asymptotically indistinguishable (as $d \to \infty$) for all $\mu \in \mathcal{P}^d$.*

This conjecture is originally motivated by the Marcenko-Pastur law, and has further motivation by the fact that the graphs produced in the setting of the conjecture have the same bulk eigenvalues. All nodes will have approximately same degree so it's reasonable to believe that extremal eigenvalues are universal, which is show by experiment (see appendix). First, we need a lemma describing the threshold's relation to edge probabilities.

Lemma 2.2. *For all $\mu \in \mathcal{P}^d$, $t(p, d) = Q^{-1}(p)\sqrt{d}$ satisfies*

$$\mathbb{P}\left(\langle X, Y \rangle \geq t(p, d)\right) \to p,$$

as $d \to \infty$ where $X, Y \sim \mu$ independently.

Proof. As $d \to \infty$, the CLT gives

$$\begin{aligned}\mathbb{P}(\langle X, Y \rangle \geq t) &= \mathbb{P}\left(\sum_{i=1}^{d} X_i Y_i \geq t\right) \\ &\to \mathbb{P}(N(0, \sqrt{d}) \geq t) \\ &= \mathbb{P}(N(0, 1) \geq t/\sqrt{d})\end{aligned}$$

where X_i, Y_i are standard 1-d Gaussians. Then, if we choose t to be $t(p, d)$, the result follows. ☐

Now, we'll give a proof strategy for the conjecture. Note that the claim that ties it together is not proven (appears at the end of the proof attempt) so this is not complete. It seems plausible but it's not clear to me in what sense the Ito integral is defined and how to work with these objects, so this will require a little more effort.

Proof Attempt of Conjecture 2.1. ☐

Date: April 3rd, 2024.

For each $\mu \in \mathcal{P}^d$, consider the following equivalent model: Associate to each a node a random walk $X_i(t)$ with d time steps, where each step is taken with respect to μ. Then, let $I = \{1/d, 2/d, ...1\}$ between two nodes X_i and X_j, connect an edge if

$$\sum_{t \in I} (X_i(t) - X_i(t - 1/d)) (X_j(t) - X_j(t - 1/d)) \geq t(p, d).$$

Now, if we replace $t(p, d)$ with $Q^{-1}(p)\sqrt{d}$ and normalize the $X_j(t) - X_j(t - 1/d)$ term, we see that this term becomes a Brownian motion for all $\mu \in \mathcal{P}^d$ (Donsker's Invariance Principle). Then, in the $d \to \infty$ limit, $G(n, p, \mu)$ equivalent to this model: For each node draw a μ-white noise $X_i(\cdot)$ on $[0, 1]$. For two nodes X_i, X_j, form X_j into its corresponding Brownian motion B_j. Then, connect an edge between X_i and X_j if $\int X_i(t) dB_j$ (the integral of the remaining white noise with respect to the Brownian motion) is greater than $Q^{-1}(p)$.

First, notice that the choice of turning X_j into the integrator is arbitrary since in our derivation above, we could have equivalently normalized X_i. Denote this model as the $\mu-$white noise model. We'll now argue that $G(n, p, \mu)$ is equivalent to the Gaussian white noise model for all μ. The universality statement we need to prove is that for a Brownian motion B_t and a $\mu-$white noise $F(t)$ on [0,1], both drawn independently,

$$\int_0^1 F(t) dB_t$$

is the same random variable for all μ.

We'll now consider the setting of anisotropic Gaussians. For a vector $\alpha \in \mathbb{R}^{2^d}$ with $\alpha_i > 0$ for all i, denote $\mathcal{P}^\alpha$ to be the class of probability measures on $\mathbb{R}^d$ such that $\mu \in \mathcal{P}^\alpha$ is a product measure with independent entries, where every entry in the same dyadic interval I has mean 0 and variance α_I. Note that the dimension d is implicitly defined through α.

Proposition 2.3. *Let $X_i, X_j \sim N(0, \Sigma_\alpha)$ independently. If $t_{p,\alpha}$ satisfies $\mathbb{P}(\langle X_i, X_j\rangle \geq t_{p,\alpha}) = p$, then*

$$t_{p,\alpha} = \frac{Q^{-1}(p)}{\mathbb{E}_{x\sim N(0,D_\alpha^2)}\left(\frac{1}{||x||}\right)}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-x^2/2}\,dx$.

Proof. Let $f(x)$ denote the density of $N(0, D_\alpha)$. The law of total probability gives that

$$\mathbb{P}(\langle X_i, X_j\rangle \geq t_{p,\alpha}) = \int_{\mathbb{R}^n}\int_{T_x} f(y)\,dy\,f(x)\,dx \tag{1}$$

where $T_x = \{y \in \mathbb{R}^d : \langle y, x\rangle \geq t_{p,\alpha}\}$. Letting $T = \int_{T_x} f(y)\,dy$, we have that

$$T = \frac{1}{(2\pi)^{n/2}\sqrt{\det D_\alpha}}\int_{T_x}\exp\left(-\frac{1}{2}y^T D_\alpha^{-1}y\right)\,dy.$$

Substituting $y = \sqrt{D_\alpha}z$ gives

$$\begin{aligned}
T &= \frac{1}{(2\pi)^{n/2}\sqrt{\det D_\alpha}}\int_{T_x^*}\exp\left(-\frac{1}{2}(\sqrt{D_\alpha}z)^T D_\alpha^{-1}\sqrt{D_\alpha}z\right)\det\sqrt{D_\alpha}\,dz\\
&= \frac{1}{(2\pi)^{n/2}}\int_{T_x^*}\exp\left(-\frac{1}{2}z^T z\right)\,dz,
\end{aligned}$$

where $T_x^* = \{z \in \mathbb{R}^d : \langle z, \tilde{x}\rangle \geq t_{p,\alpha}\}$ with $\tilde{x} = \sqrt{D_\alpha}x$. Choose any orthogonal matrix B such that $B\tilde{x} = ||\tilde{x}||e_n$. Substituting $z = B^{-1}u$, the domain of integration becomes $\mathbb{R}^{d-1}\times[c,\infty)$ where $c = \frac{t_{p,\alpha}}{||\tilde{x}||}$ and we obtain

$$\begin{aligned}
T &= \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^{d-1}\times[c,\infty)}\exp\left(-\frac{1}{2}u^T(B^{-1})^T B^{-1}u\right)\,du\\
&= \frac{1}{\sqrt{2\pi}}\int_c^\infty\exp\left(-\frac{1}{2}u^2\right)\,du\\
&= Q(c),
\end{aligned}$$

where $Q(\cdot)$ is the Q-function of a standard one-dimensional Gaussian. Substituting this back into (1), we obtain

$$\mathbb{E}_{x\sim N(0,D_\alpha)}\left(Q\left(\frac{t_{p,\alpha}}{||\sqrt{D_\alpha}x||}\right)\right) = p.$$

Since $Q(\cdot)$ is strictly decreasing, the above statement implies

$$\mathbb{E}_{x\sim N(0,D_\alpha)}\left(\frac{t_{p,\alpha}}{||\sqrt{D_\alpha}x||}\right) = Q^{-1}(p).$$

By linearity of expectation it, it suffices to show that

$$\mathbb{E}_{x\sim N(0,D_\alpha)}\left(\frac{1}{||\sqrt{D_\alpha}x||}\right) = \mathbb{E}_{x\sim N(0,D_\alpha^2)}\left(\frac{1}{||x||}\right)$$

to complete the proof. Substituting $x = (\sqrt{D_\alpha})^{-1}y$ into the left hand side, we obtain

$$\begin{aligned}
\mathbb{E}_{x\sim N(0,D_\alpha)}\left(\frac{1}{||\sqrt{D_\alpha}x||}\right) &= \frac{\det(\sqrt{D_\alpha})^{-1}}{(2\pi)^{n/2}\sqrt{\det D_\alpha}}\int_{\mathbb{R}^d}\exp\left(-\frac{1}{2}((\sqrt{D_\alpha})^{-1}y)^T D_\alpha^{-1}(\sqrt{D_\alpha})^{-1}y\right)\frac{1}{||y||}\,dy\\
&= \frac{1}{(2\pi)^{n/2}\sqrt{\det D_\alpha^2}}\int_{\mathbb{R}^d}\exp\left(-\frac{1}{2}y^T D_\alpha^{-2}y\right)\frac{1}{||y||}\,dy\\
&= \mathbb{E}_{x\sim N(0,D_\alpha^2)}\left(\frac{1}{||x||}\right).
\end{aligned}$$

□

It's reasonable to suggest that this derivation could be universal in $\mathcal{P}^\alpha$, but this remains to be proved. Then, we have an analagous conjecture for the anisotropic case:

Conjecture 2.4. *The class $\mathcal{P}^\alpha$ is universal with respect to $G(n, p, \mu)$ as $d \to \infty$.*

3. Appendix

3.1. **Universal Bulk Spectra.** We'll first adapt the results from [3, 5, 1, 2] to derive the universal empirical spectral distribution for $G \sim G(n,p,\mu)$ for $\mu \in \mathcal{P}^d$. Let

$$c_k = \mathbb{E}_{Z\sim N(0,1)}\left(f(z)h_k(z)\right))$$

be the k^{th} Hermite coefficient of $f(x;p,d) = \mathbb{1}(x > t(p,d))$ and $\kappa > 0$ satisfying $\kappa = \frac{n}{d^l} - O(d^{-1/2})$. Further let

$$\gamma_a = \begin{cases} c_l^2 & \text{if } l \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}, \gamma_b = \begin{cases} c_l\sqrt{l!\kappa} & \text{if } l \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases} \gamma_c = \sum_{k=l_c}^{\infty} c_k^2$$

where $l_c = \begin{cases} l+1 & \text{if } l \in \mathbb{N} \\ \lceil l \rceil & \text{otherwise} \end{cases}$. Finally, let $E = \gamma_a/\gamma_b$, $F = \gamma_a/\gamma_b^2, G = \gamma_a/\gamma_b^2\sqrt{\gamma_c} + \gamma_a/\gamma_b$. Then, for all $\mu \in \mathcal{P}^d$, the empirical spectral distribution of G/n where $G \sim G(n,p,\mu)$ is given by

$$f(x) = \begin{cases} 0 & \text{if } D \leq 0 \\ \frac{\sqrt{3}}{2}\left((\sqrt{D(x)} + R(x))^{1/3} + (\sqrt{D(x)} - R(x))^{1/3}\right) & \text{otherwise} \end{cases}$$

where

$$D(x) = Q^3(x) + R^2(x)$$
$$Q(x) = (3\alpha_1(x) - \alpha_2^2(x))/9$$
$$R(x) = (9\alpha_2(x)\alpha_1(x) - 26\alpha_0 - 2\alpha_2^3(x))/54$$
$$\alpha_0 = \frac{F^2}{E(G-E^2)}$$
$$\alpha_1(x) = \frac{(E+Fx)F}{E(G-E^2)}$$
$$\alpha_2(x) = \frac{(G+Ex)F}{E(G-E^2)}.$$

Below is a plot demonstrating the universality but we could not verify the above formula in the code. It was either a flaw in the code or in the equation, but the equation should generally be right up to algebraic error.

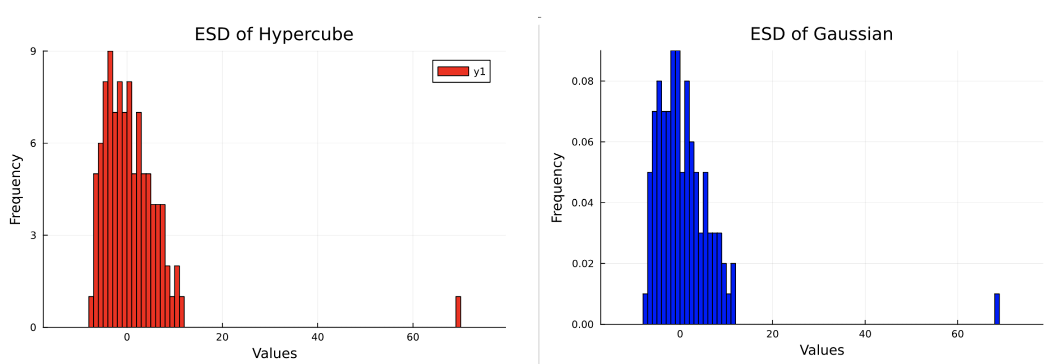

Figure 1. Empirical spectral distributions for d-dimensional hypercube and Gaussian graphs with settings: $l = 1, d = 100, n = round(Int, d^l), p = 1/3$

3.2. **Miscellaneous arguments.** Now, we see how hamming distance between nodes (i.e. whether or not they agree on the parity of each dimension) interacts with edge probabilities.

Lemma 3.1. *For μ standard d-dimensional Gaussian and $l \in [d]$, if $X, Y \sim \mu$ conditioned on being separated by hamming distance l, then as both d and $d - l \to \infty$, the probability of X, Y sharing an edge converges to*

$$Q\left(\frac{t(p,d) - 2/\pi(d-2l)}{(1-4/\pi^2)(\sqrt{d-l}+\sqrt{l})}\right)$$

where $Q(\cdot)$ is the Q-function for a standard Gaussian.

Proof. The probability that X and Y share an edge is

$$\mathbb{P}\left(\sum_{i=1}^{d-l} |X_i Y_i| - \sum_{i=d-l+1}^{d} |X_i Y_i| \geq t(p,d)\right).$$

Then, the CLT holds in this setting and the random variables converges in law to

$$\begin{aligned} &N\left(2(d-l)/\pi, (1-4/\pi^2)\sqrt{d-l}\right) - N\left(2l/\pi, (1-4/\pi^2)\sqrt{l}\right) \\ &\sim N\left(2(d-2l)/\pi, (1-4/\pi^2)(\sqrt{d-l}+\sqrt{l}\right) \end{aligned}$$

Then, normalizing the Gaussian we obtain the result. □

References

[1] Singer A. Cheng, X. The spectrum of random inner-product kernel matrices. 2013.

[2] Tom P. Davis. A general expression for hermite expansions with applications. 2024.

[3] Lue Yue M. Mckenna Benjamin Yau Horng-Tzer Dubova, Sofiia. Universality for the global spectrum of random inner-product kernel matrices in the polynomial regime. 2023.

[4] Quentin Duchemin and Yohann De Castro. Random geometric graph: Some recent developments and perspectives. In Rafał Adamczak, Nathael Gozlan, Karim Lounici, and Mokshay Madiman, editors, *High Dimensional Probability IX*, volume 80 of *Progress in Probability*, pages 265–297. Birkhäuser, Cham, 2023.

[5] Yau Horng-Tzer Lu, M. Yue. An equivalence principle for the spectrum of random inner-product kernel matrices with polynomial scalings. 2023.