Stat206: Random Graphs and Complex Networks

Lecture 22: Induced Subgraphs; Heuristics for max-degree, Connectivity

Lecturer: David Aldous *Scribe: Raul Etkin*

22.1 Random Geometric Graph - Induced Subgraphs

22.1.1 Model

- density $f(x)$ on $\mathcal{R}^d$.
- $X_0, X_1, \dots, X_{n-1}$ independent, $f(\cdot)$.
- Real $r > 0$.
- Put edge (i,j) if $\|X_i - X_j\| \leq r$.
- Random Graph $\mathcal{G}_n$.

Study asymptotics as $n \to \infty$ and $r_n \downarrow 0$.

Figure 1: Scaling by $1/r$ of the random geometric graph model. There is an edge $i \leftrightarrow j$ in the scaled graph if $\|Y_i - Y_j\| \leq 1$.

Y_i is approximately a Poisson process of rate Qnr_n^d where $Q = f(X_0)$. We have that degree$(X_0|Q) \approx$ Poisson$(Qnr_n^d v_d)$ where v_d is the volume of the unit ball in d dimensions.

22.1.2 Induced Subgraphs

Fix finite graph G, with $K = |G|$, i.e. the number of vertices of G. Consider "number of copies of G in $\mathcal{G}_n$" $= N_n(G) = \#$ subsets Y of $\{0, 1, \dots, n-1\}$ such that $\mathcal{G}_n$ restricted to Y is isomorphic to G.

Define the indicator function:

$$h_g(Y_1, \dots, Y_K) = \begin{cases} 1 & \text{if graph on vertices at points } (Y_1, \dots, Y_K), \text{ edge iff } \|Y_i - Y_j\| \leq 1, \text{ is isomorphic to } G \\ 0 & \text{otherwise} \end{cases}$$

Then we can write:

$$\begin{aligned} E[N_n(G)] &= \frac{1}{K!} E\left[\sum_{j_1=0}^{n-1} \text{ j's distinct}^{\cdots} \sum_{j_K=0}^{n-1} h_G\left(\frac{X_{j_1}}{r}, \dots, \frac{X_{j_K}}{r}\right)\right] \\ &= \frac{n}{K!} \int_{\mathcal{R}^d} f(x)dx E\left[\sum_{j_2} \text{ j's distinct}^{\cdots} \sum_{j_K} h_G\left(0, \frac{X_{j_2} - x}{r}, \dots, \frac{X_{j_K - x}}{r}\right)\right] \end{aligned}$$

By doing a "change of variables", the density of $\frac{X_j - x}{r_n}$ at any point is $Y \xrightarrow{\sim} f(x) r_n^d$ as $n \to \infty$.

The $E(\cdot)$ term, for fixed $(j_1, \ldots, j_K)$ is approximately given by

$$\int \cdots \int h_G(0, Y_2, \ldots, Y_{K-1}) dY_2 \ldots dY_{K-1} (f(x) r_n^d)^{K-1}$$

Also the number of choices $(j_2, \ldots, j_K) \sim n^{K-1}$.

Then we have:

Proposition 3.1:

$$E[N_n(G)] \sim \frac{1}{K!} n^K r_n^{d(K-1)} \left(\int_{\mathcal{R}^d} f^K(x) dx \right) \left(\int \cdots \int h_G(0, Y_1, \ldots, Y_{K-1}) dY_1 \ldots dY_{K-1} \right) = (*)$$

In sparse regime when $nr_n^d \to$ constant, we have that $E[N_n(G)]$ scales as n.

If $E[N_n(G)] \to \lambda$ (super-sparse regime) then $N_n(G) \xrightarrow{d}$ Poisson(mean=$(*)$).

If $E[N_n(G)] \uparrow \infty$, $\frac{N_n(G) - E[N_n(G)]}{\sqrt{E[N_n(G)]}} \xrightarrow{d}$ Normal$(0, \sigma^2)$, where σ^2 is given by a complicated formula.

This is different from E-R where scaling power depends on $K = |G|$.

22.2 Heuristics for Max-degree and Connectivity

22.2.1 Background

Let $M_n = \max(X_1, \ldots, X_n)$, X_i independent, Poisson(λ). As $K \to \infty$, $P(X \geq K) \sim P(X = K) = e^{-\lambda} \frac{\lambda^K}{K!}$. Then,

$$P(M_n < K) = [1 - P(X \geq K)]^n \approx \exp\left(-ne^{-\lambda} \frac{\lambda^K}{K!} \right)$$

For a typical large n there exists $K(n)$ such that $ne^{-\lambda} \frac{\lambda^K}{K!} \gg 1$ for $K = K(n)$, and $ne^{-\lambda} \frac{\lambda^K}{K!} \ll 1$ for $K = K(n) + 1$.

Then we have the concentrated property:

$$P(M_n = K(n)) \approx 1$$

Consider the solution of $ne^{-\lambda} \frac{\lambda^K}{K!} = 1$:

$$\begin{aligned} \log n + K \log \lambda &= K(\log K - 1) \\ K &= \frac{\log n}{\log \log n} \end{aligned}$$

Then, more crudely we have

$$M_n \sim (1 \pm o(1)) \frac{\log n}{\log \log n}$$

as $n \to \infty$ regardless of λ. Note that this result is very asymptotic.

22.2.2 Maximum vertex degree

First consider an Erdös-Renyi random graph $\mathcal{G}(n, \lambda/n)$. We have:

- degree$(v) \approx$ Poisson(λ)
- max-degree $= (1 + o(1)) \frac{\log n}{\log \log n}$

Now, for a Geometric random graph, with $nr_n^d \to$ constant, if $\sup_x(f(x)) < \infty$ we also get max-degree $= (1 + o(1)) \frac{\log n}{\log \log n}$.

22.2.3 Connectivity

The obstacle to connectivity are the isolated (degree=0) vertices.

For an **Erdös-Renyi random graph** $\mathcal{G}(n, p_n)$ we have

- degree$(v) \approx$ Poisson(np_n).
- $P(v \text{ is isolated}) \sim \exp(-np_n)$.

Then,

$$E[\# \text{ of isolated vertices}] \sim n \exp(-np_n) \begin{cases} \to 0 & \text{for } p_n = (1+\epsilon)\frac{\log n}{n} \\ \to \infty & \text{for } p_n = (1-\epsilon)\frac{\log n}{n} \end{cases}$$

With more work one can obtain

$$P(\mathcal{G}(n, p_n) \text{ is connected}) \begin{cases} \to 1 & \text{for } p_n = (1+\epsilon)\frac{\log n}{n} \\ \to 0 & \text{for } p_n = (1-\epsilon)\frac{\log n}{n} \end{cases}$$

For a **random geometric graph**:

Let $f(x) = f_0$ be uniform on a d-dimensional torus.

$$\text{degree}(v) \approx \text{Poisson}(nf_0 v_d r_n^d)$$

Defining $p_n = f_0 v_d r_n^d$ we get $P(v \text{ is isolated})$ similar as before. Then,

$$P(\mathcal{G}_n \text{ is connected}) \begin{cases} \to 1 & \text{for } f_0 v_d r_n^d = (1+\epsilon)\frac{\log n}{n} \\ \to 0 & \text{for } f_0 v_d r_n^d = (1-\epsilon)\frac{\log n}{n} \end{cases}$$

This is true but misleading. We are interested in working in a square, but to avoid edge conditions we have worked with a torus.

Figure 2: Boundary effect when working with a square instead of a torus.

\# of vertices within r_n of boundary $\sim nr_n$.

$$E(\# \text{ non-boundary vertices isolated}) \sim nr_n \exp\left(-\frac{1}{2}np_n\right) = (**)$$

Take $p_n = \frac{\log n}{n}$, so that $\exp(-np_n) = 1/n$. Then,

$$(**) \sim nr_n \frac{1}{n^{1/2}} = n^{1/2} r_n \sim n^{\frac{1}{2}-\frac{1}{d}} \log n$$

which $\to \infty$ for $d \geq 2$, where the last approximation used $r_n \sim n^{-1/d} \log n$.

So at threshold for "no isolated non-boundary vertices" there will be isolated near-boundary vertices.

Min-degree for geometric random graphs involves fussy issues of boundary effects.