Cliques in dense inhomogeneous random graphs

Martin Doležal

Institute of Mathematics of the Czech Academy of Sciences

joint work with Jan Hladký *and* András Máthé

5.11.2015

Motivation

$\mathbb{G}(n, p)$ .... Erdős–Rényi random graph (here $n \in \mathbb{N}$ and $p \in (0, 1)$):

- vertex set $\{1, \ldots, n\}$
- each edge is included with probability p

$\omega(\mathbb{G}(n, p))$ ..... the order of the largest clique in $\mathbb{G}(n, p)$

Theorem (Grimmet and McDiarmid 1975; Matula 1976)

It holds a.a.s. that

$$\omega(\mathbb{G}(n, p)) = (1 + o(1)) \cdot \frac{2}{\log(1/p)} \cdot \log n .$$

Graphons

A graphon (introduced by Lovász and Szegedy in 2006) is a symmetric measurable function $W\colon [0,1]^2 \to [0,1]$.

FACTS:

- Each finite graph can be represented as a graphon.
- The space of graphons, with the so called *cut metric*, is a metric compactification of the space of finite graphs.

More general random graphs

$\mathbb{G}(n, W)$ .... random graph introduced by Lovász and Szegedy:

- we sample n random independent points $x_1, \ldots, x_n \in [0, 1]$
- the vertex set is $\{1, \ldots, n\}$
- i and j are connected by an edge with probability $W(x_i, x_j)$

NOTE: If $W = p$ a.e. then $\mathbb{G}(n, W) = \mathbb{G}(n, p)$.

$\omega(\mathbb{G}(n, W))$ .... the order of the largest clique in $\mathbb{G}(n, W)$

Example

The clique number (normalized by $\log n$) is not continuous with respect to the cut metric.

Let $f : \mathbb{N} \to \mathbb{N}$ be such that $f(n) \to \infty$ but $\frac{f(n)}{\log n} \to 0$.

G_n ..... n-vertex $f(n)$-partite Turán graph

- On one hand, we have $\frac{\omega(G_n)}{\log n} = \frac{f(n)}{\log n} \to 0$.
- On the other hand, the graph G_n has $\binom{n}{2} - f(n)\binom{\frac{n}{f(n)}}{2} \approx \frac{n^2}{2}$ edges, and so the sequence $\{G_n\}_{n=1}^{\infty}$ converges to the constant graphon $W \equiv 1$.

<u>CONCLUSION</u>: When examining the clique number $\omega(\mathbb{G}(n, W))$, we cannot straightforwardly use the cut metric.

Our main result

Theorem

Suppose that $\operatorname{ess\,inf} W > 0$. *Then a.a.s.*

$$\omega(\mathbb{G}(n, W)) = (1 + o(1)) \cdot \kappa(W) \cdot \log n \ ,$$

where

$$\kappa(W) = \sup \left\{ \frac{2\|h\|_1^2}{\int_x \int_y h(x)h(y) \log\left(1/W(x,y)\right)} : \ h \in L^1([0,1]), \ \ h \geq 0 \right\} \ .$$

In the supremum above, it suffices to range only over characteristic functions, i.e.

$$\kappa(W) = \sup \left\{ \frac{2\lambda(A)^2}{\int_{A\times A} \log\left(1/W(x,y)\right)} : \ A \subset [0,1], \ \ \lambda(A) > 0 \right\} \ .$$

Our main result

Recall:

Suppose that $\operatorname{ess\,inf} W > 0$. Then a.a.s.

$$\omega(\mathbb{G}(n, W)) = (1 + o(1)) \cdot \kappa(W) \cdot \log n \,, \tag{1}$$

where

$$\kappa(W) = \sup \left\{ \frac{2\lambda(A)^2}{\int_{A\times A} \log\left(1/W(x,y)\right)} : A \subset [0,1],\ \lambda(A) > 0 \right\} .$$

If $\kappa(W) = +\infty$ then (1) reads as $\omega(\mathbb{G}(n, W)) \gg \log n$.

<u>NOTE</u>: If $W = p$ a.e. then $\kappa(W) = \frac{2}{\log(1/p)}$.

Heuristics of the proof

Let $[0,1] = I_1 \sqcup I_2$ and $p_{11}, p_{12}, p_{22} \in (0,1)$.

Consider the the 2-step graphon $W\colon [0,1]^2 \to [0,1]$ defined by

$$W(x,y) = \begin{cases} p_{11} & \text{if } x,y \in I_1, \\ p_{22} & \text{if } x,y \in I_2, \\ p_{12} & \text{otherwise.} \end{cases}$$

<u>OUR TASK</u>: For every $c > 0$, decide whether there typically exists a clique of order $c \log n$ in $\mathbb{G}(n, W)$.

So let us fix $c > 0$.

X_n .... the number of cliques of order $c \log n$ in $\mathbb{G}(n, W)$
$Y_n^{\alpha_1,\alpha_2}$ ..... the number of cliques in $\mathbb{G}(n, W)$ consisting of $\alpha_1 \log n$ vertices represented in I_1 and $\alpha_2 \log n$ vertices represented in I_2

First moment argument

We have

$$\mathbf{E}[X_n] = \sum_{\substack{\alpha_1,\alpha_2 \geq 0 \\ \alpha_1+\alpha_2=c \\ \alpha_1 \log n,\ \alpha_2 \log n \in \mathbb{N}_0}} \mathbf{E}\left[Y_n^{\alpha_1,\alpha_2}\right] .$$

This sum has only a $\Theta(\log n)$-many summands, so we can expect that

- $\mathbf{E}[X_n] \to +\infty \quad \Leftrightarrow \quad \exists \alpha_1, \alpha_2 \colon \mathbf{E}[Y_n^{\alpha_1,\alpha_2}] \to +\infty$,
- $\mathbf{E}[X_n] \to 0 \quad \Leftrightarrow \quad \forall \alpha_1, \alpha_2 \colon \mathbf{E}[Y_n^{\alpha_1,\alpha_2}] \to 0$.

Let us write $\beta_1 = \lambda(I_1)$ and $\beta_2 = \lambda(I_2)$. Then for every α_1, α_2 we have

$$\mathbf{E}[Y_n^{\alpha_1,\alpha_2}] \approx \binom{\beta_1 n}{\alpha_1 \log n} \binom{\beta_2 n}{\alpha_2 \log n} p_{11}^{\binom{\alpha_1 \log n}{2}} p_{22}^{\binom{\alpha_2 \log n}{2}} p_{12}^{\alpha_1 \alpha_2 \log^2 n}$$

$$\approx \exp\left(\log^2 n \left(\alpha_1 + \alpha_2 + \frac{\alpha_1^2}{2} \log p_{11} + \frac{\alpha_2^2}{2} \log p_{22} + \alpha_1 \alpha_2 \log p_{12} \right) \right) .$$

First moment argument

Recall:

$$\mathbf{E}[Y_n^{\alpha_1,\alpha_2}] \approx \binom{\beta_1 n}{\alpha_1 \log n}\binom{\beta_2 n}{\alpha_2 \log n} p_{11}^{\binom{\alpha_1 \log n}{2}} p_{22}^{\binom{\alpha_2 \log n}{2}} p_{12}^{\alpha_1\alpha_2 \log^2 n}$$

$$\approx \exp\left(\log^2 n \left(\alpha_1 + \alpha_2 + \frac{\alpha_1^2}{2}\log p_{11} + \frac{\alpha_2^2}{2}\log p_{22} + \alpha_1\alpha_2 \log p_{12}\right)\right) .$$

NOTE: The right hand side does not depend on the values of β_1, β_2.

It follows that whether $\mathbf{E}[Y_n^{\alpha_1,\alpha_2}] \to 0$ or $\mathbf{E}[Y_n^{\alpha_1,\alpha_2}] \to +\infty$ depends on the sign of

$$\alpha_1 + \alpha_2 + \frac{1}{2}\left(\alpha_1^2 \log p_{11} + \alpha_2^2 \log p_{22} + \alpha_1\alpha_2 \log p_{12} + \alpha_2\alpha_1 \log p_{12}\right) .$$

First moment argument

CONCLUSION: We try to maximize $\alpha_1 + \alpha_2$ (where $\alpha_1, \alpha_2 \geq 0$) with respect to the condition

$$\alpha_1 + \alpha_2 + \frac{1}{2}\left(\alpha_1^2 \log p_{11} + \alpha_2^2 \log p_{22} + \alpha_1\alpha_2 \log p_{12} + \alpha_2\alpha_1 \log p_{12}\right) \geq 0 \,.$$

In case of a general graphon W, we substitute summation by integration, so we try to maximize $\|f\|_1$ (where $f \in L^1([0,1])$, $f \geq 0$) with respect to the condition

$$\Gamma(f, W) := \int_x f(x) + \frac{1}{2}\int_x \int_y f(x)f(y) \log W(x,y) \geq 0 \,.$$

In this way, we get an alternative formula for $\kappa(W)$:

$$\kappa(W) = \sup\left\{\|f\|_1 \colon f \in L^1([0,1]),\ f \geq 0,\ \Gamma(f,W) \geq 0\right\}$$

First and second moment argument

By the previous arguments, we can determine whether $\mathbf{E}[X_n] \to 0$ or $\mathbf{E}[X_n] \to +\infty$.

1. If $\mathbf{E}[X_n] \to 0$ then by Markov's inequality, there are a.a.s. no cliques of order c.

2. If $\mathbf{E}[X_n] \to +\infty$ then we would like to use the second moment argument.

To do this, we need to verify that $\mathbf{E}[X_n]^2 = (1 + o(1))\mathbf{E}[X_n^2]$.

This is straightforeard in case of a constant graphon W (i.e. in case of the Erdős–Rényi random graph) but it does not hold in general!

Second moment argument

SOLUTION: We find $A \subseteq [0,1]$ such that the subgraphon $U := W \restriction_{A \times A}$ of W has the following properties:

- $\kappa(U) \approx \kappa(W)$,
- the second moment argument works for the subgraphon U.

It follows that a.a.s., the random graph $\mathbb{G}(n, U)$ contains a clique of order $c \log n$.

It is easy to conclude that the same is true for $\mathbb{G}(n, W)$.

□

The End