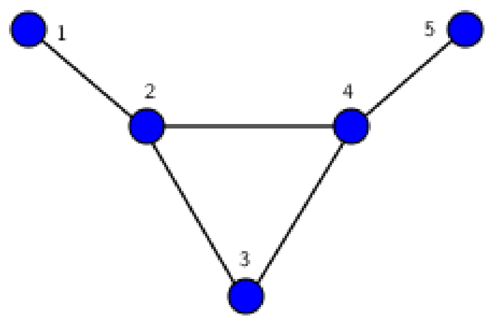
GRAPHONS AND CONVERGENCE

LIAM HODGKINSON

Since its inception, graph theory has become ubiquitous in the study of networks. Not to be confused with the scientific notion of a graph which has more in common with its topological definition, graph theorists define a graph as a pair $G = (V, E)$, where V is a vertex set consisting of $|V|$ vertices, and E is the edge set, itself consisting of pairs of vertices in V, and representing lines (edges) adjoining these vertices. For an undirected graph, the order of this pair is irrelevant. A graph may also be represented using an *adjacency matrix.* For example, the bull graph



can be represented as the adjacency matrix

$$A = \begin{pmatrix} & 1 & & & \\ 1 & & 1 & 1 & \\ & 1 & & 1 & \\ & 1 & 1 & & 1 \\ & & & 1 & \end{pmatrix}.$$

With breakthroughs in modern probability, it became natural to incorporate random elements into the graph theory framework. A *random graph* is a probability distribution over a predetermined set of graphs. They can have a random number of vertices, but for our purposes, we shall assume the number of vertices for a particular random graph is fixed, and so a random graph of size n (n vertices) is simply a probability distribution over $n \times n$ binary symmetric matrices with zero diagonal, or alternatively, over $\{0, 1\}^{\binom{n}{2}}$.

At the same time, a central objective in probability theory is the development of 'limit theorems' demonstrating that under some reasonable transformation (or regularisation/normalisation), a probabilistic object is well approximated by a much simpler object as one of its characteristics tends to some limit. Examples include the law of large numbers, the central limit theorem and Donsker's theorem (also called the functional central limit theorem). It is desirable to develop some technology that combines these two ideas together. Their intersection is the study of *graphons*.

The primary resource for the study of graphons is the marvellous monograph "Large networks and graph limits" by Laszló Lovasz.

Graphons. Let $\mathcal{W}$ be the space of all functions $W: [0,1]^2 \to \mathbb{R}$; such functions are referred to as *kernels*. Let $\mathcal{W}_0$ be the space of all functions $W: [0,1]^2 \to \mathbb{R}$ such that $0 \le W \le 1$. Clearly $\mathcal{W}_0$ is a closed subset of $\mathcal{W}$. The elements of $\mathcal{W}_0$ are called *graphons* and generalise the standard notion of a graph. Indeed, there is a one-to-one correspondence between the subclass of stepfunctions and graphs: for any graph G with $|V(G)| = n$ vertices, let $J_1, \dots, J_n$ be a partition of $[0,1]$, and set W_G as the stepfunction defined for $x, y \in J_i \times J_j$ by $W_G(x,y) = \mathbb{1}\{ij \in E(G)\}$. This stepfunction can be conceived as a functional representation of the adjacency matrix of G, but has the important advantage of remaining on the same space, regardless of n. Many of the ideas from traditional graph theory generalises to graphons. For example, we say that graphons U and W are equivalent if there exists an invertible measure preserving map $\varphi: [0,1] \to [0,1]$ such that $U(\varphi(x), \varphi(y)) =: U^{\varphi}(x,y) = W(x,y)$ almost everywhere (it can be verified that this does indeed define an equivalence relation). This is analogous to isomorphisms for unlabelled graphs. Similarly, we can define the degree function d_W of a graphon W as the analogue of the proportion of external vertices to which a vertex is connected:

$$d_W(x) = \int_0^1 W(x,y)\,dy \sim \frac{\deg x}{n}.$$

Homomorphism Density. For two simple graphs G and H, an adjacency-preserving map φ from $V(G)$ to $V(H)$ is called a *homomorphism* (in other words, if $ij \in E(G)$ then $\varphi(i)\varphi(j) \in E(H)$). Many important questions in graph theory can be phrased in terms of homomorphisms: for example if $G = \triangle$, the existence of a homomorphism $G \to H$ implies the existence of a triangle in H. Additionally, the existence of $K_n \to G$ (for K_n the complete graph of order n) implies G contains a clique with n nodes, while $G \to K_n$ implies that G is n-colorable. The number of homomorphisms is denoted $\hom(G,H)$, given by

$$\hom(G,H) = \sum_{\varphi\colon V(G) \to V(H)} \underbrace{\prod_{uv \in E(G)} \mathbb{1}\{\varphi(u)\varphi(v) \in E(H)\}}_{\varphi \text{ is a homomorphism}},$$

and the homomorphism density $t(G,H)$ is

$$t(G,H) = \frac{\hom(G,H)}{|V(G)|^{|V(F)|}}.$$

The density of triangles $t(\triangle, H)$ is roughly the probability that any three vertices chosen in the graph form a triangular subgraph. The homomorphism density has a surprisingly simple extension to graphons, given for a simple graph F and graphon $W \in \mathcal{W}_0$ by

$$t(F,W)=\int_{[0,1]^{V(F)}}\prod_{ij\in E(F)}W(x_i,x_j)\prod_{i\in V(F)}dx_i.$$

So, for example, the 'density of triangles' in a graphon W is given by

$$t(\triangle,W)=\int_0^1\int_0^1\int_0^1 W(x,y)\,W(y,z)\,W(x,z)\,dx\,dy\,dz,$$

and similarly, the 'density of an n-clique' is

$$t(K_n,W)=\int_{[0,1]^n}\prod_{i<j}^{n}W(x_i,x_j)\,dx_1\cdots dx_n.$$

The Cut Distance. Evidently, graphons intersect functional analysis and graph theory, but we have still not yet touched on one of the main selling points of graphons: the enabling of limit theory. For this, a topology is required, which will hopefully imply that $t(F,U)\to t(F,W)$ whenever $U\to W$. As it turns out, none are more suitable than the cut distance.

The cut norm is defined for $n\times n$ matrices A by

$$\|A\|_\square=\frac{1}{n^2}\max_{S,T\subset\{1,\dots,n\}}\left|\sum_{i\in S,j\in T}A_{ij}\right|,$$

and analogously, the cut norm is defined for kernels $W\in\mathcal{W}$ by

$$\|W\|_\square=\sup_{U,V\subseteq[0,1]}\left|\int_{U\times V}W(x,y)\,dxdy\right|.$$

One interesting feature of the cut metric is that $\int_{[0,1]^2}|W(x,y)|\,dxdy=\|W\|_1\le\sqrt{2n}\,\|W\|_\square$. To incorporate our equivalence class for graphons, we introduce the cut distance $\delta_\square(U,W)$ between graphons U and W defined by

$$\delta_\square(U,W)=\inf_\varphi\|U^\varphi-W\|_\square\le\|U-W\|_\square.$$

The counting lemma states that convergence in the cut distance implies convergence in homomorphism densities, as desired.

Lemma 1 (Counting Lemma)**.** *Let F be a simple graph and let U and W be graphons. Then $|t(F,U)-t(F,W)|\le|E(F)|\,\delta_\square(U,W)$.*

Proof. The result stated as above requires a few more complex arguments, so I shall prove a more straightforward result. First note that if $a_1,a_2,b_1,b_2\in[0,1]$, then

$$|a_1a_2-b_1b_2|\le|a_1a_2-a_2b_1|+|a_2b_1-b_1b_2|\le|a_1-b_1|+|a_2-b_2|,$$

and so more generally, for $a_i,b_i\in[0,1]$ for $i=1,\dots,n$,

$$\left|\prod_{i=1}^n a_i-\prod_{i=1}^n b_i\right|\le\sum_{i=1}^n|a_i-b_i|.$$

By definition of the homomorphism density,

$$\begin{aligned}|t(F,U)-t(F,W)| &\leq \int_{[0,1]^{V(F)}}\left|\prod_{ij\in E(F)}U(x_i,x_j)-\prod_{ij\in E(F)}W(x_i,x_j)\right|\prod_{i\in V(F)}dx_i \\ &\leq \sum_{ij\in E(F)}\int_{[0,1]^2}|U(x_i,x_j)-W(x_i,x_j)|\,dx_idx_j \\ &= |E(F)|\,\|U-W\|_1 \leq \sqrt{2n}\,|E(F)|\,\|U-W\|_\square .\end{aligned}$$

□

In essence, the counting lemma states that if two graphons are globally close, then they are also locally close (as expected). The converse is perhaps not so obvious, but still true, as implied by the inverse counting lemma.

Lemma 2 (INVERSE COUNTING LEMMA)**.** *If k is a positive integer such that for every simple graph F on k nodes, $|t(F,U)-t(F,W)|\leq 2^{-k^2}$, then*

$$\delta_\square(U,W)\leq \frac{50}{\sqrt{\log k}}.$$

Examples of Convergence. The $G(n,p)$ Erdos-Renyi random graph is a probability distribution on graphs of size n wherein each potential edge in the graph is included independently with probability p.

Proposition 3. *Let W_n be the graphon for the $G(n,p)$ Erdos-Renyi random graph. Then $\|W_n-p\|_\square \xrightarrow{a.s.} 0$ as $n\to\infty$.*

Proof. Recalling Bernstein's inequality, we have that

$$\begin{aligned}\Pr\left(\left\|W_t^{(n)}-p\right\|_\square>\epsilon\right) &= \Pr\left(\max_{u,v\in\{0,1\}^n}\left|\frac{1}{n^2}\sum_{i,j=1}^n u_iv_j\left[W_t^{(n)}(x_i,x_j)-p\right]\right|\right) \\ &\leq \sum_{u,v\in\{0,1\}^n}\Pr\left(\left|\frac{1}{n^2}\sum_{i,j=1}^n u_iv_j\left[W_t^{(n)}(x_i,x_j)-p\right]\right|>\epsilon\right) \\ &\leq \exp\left(-2n^2\epsilon^2+(n+1)\log 2\right),\end{aligned}$$

which is summable in n, completing the proof. □

Thus, the convergence of graphons can be demonstrated via concentration inequalities for Bernoulli random variables. As an example of this, note that $t(\triangle,p)=p^3$ and so the density of triangles in the Erdos-Renyi random graph converges almost surely to p^3, which is what we might expect, given that the probability that any three vertices are fully connected is p^3 also (regardless of the total number of vertices).

For a more complicated example, let $G_t^{(n)}$ be a Markov chain on the space of simple graphs with n vertices whose edges evolve independently in such a way that, for any vertices i,j, edge

ij is deleted at rate q_t and is added at rate

$$f\left(\frac{\deg i + \deg j}{2n}\right)$$

for some non-negative function $f \in \mathcal{C}^2$. Let $W_t^{(n)}$ denote the graphon of $G_t^{(n)}$ and suppose that $\left\|W_0^{(n)} - W_0\right\|_\square \to 0$ as $n \to \infty$ for some graphon W_0. Define the sequence of graphons W_t by the recursion

$$\frac{d}{dt}W_t\left(x,y\right) = \left[1 - W_t\left(x,y\right)\right] \cdot f\left(\frac{d_{W_t}(x) + d_{W_t}(y)}{2}\right) - q_t W_t\left(x,y\right).$$

Theorem 4 (H.)**.** *For any* $T > 0$, $\sup_{t\in[0,T]}\left\|W_t^{(n)} - W_t\right\|_\square \xrightarrow{a.s.} 0.$