Graph limits and flag algebras

lecture notes from IBS-SDU Summer School 2023

Andrzej Grzesik

We start by fixing the general notation. Unless specified otherwise, all graphs that we consider are undirected and simple. If G is a graph, we write $V(G)$ for its vertex set and $E(G)$ for its edge set. The *order* of G, i.e., the number of its vertices, is denoted by $|G|$ and the *size* of G, i.e., the number of its edges, is denoted by $e(G)$. The set of the first k positive integers is denoted by $[k]$.

1 Graph limits

The theory of graph limits was developed in a series of papers by Borgs, Chayes, Lovász, Sós, Szegedy and Vesztergombi. Here we only focus on the case of dense graph convergence and their limits, which is the one relevant to the flag algebra method. The theory of graph limits aims at capturing properties of large graphs. To do so, we first define a notion of convergence of a sequence of graphs and then define an analytic object representing the limit of a convergent sequence. It is possible to talk about convergence of a certain kind of combinatorial objects without having any particular limit representation in mind and the flag algebra method can also be applied in such scenarios.

We define a *density* of a graph H in a graph G, denoted by $d(H,G)$, as the probability that randomly chosen subset of $|H|$ vertices of G induces a graph isomorphic to H, i.e., $d(H,G)$ is equal to the number of induced copies of H in G divided by $\binom{|G|}{|H|}$. We say that a sequence of graphs $(G_n)_{n\in\mathbb{N}}$ of growing orders is *convergent* if for every graph H the sequence of densities $d(H,G_n)$ is convergent.

Examples of convergence sequences are complete graphs K_n, complete bipartite graphs with fixed ratio of part sizes, or Erdős-Renyi random graphs $G(n,p)$ for fixed probability p, which is convergent with probability 1.

Note that an arbitrary sequence of graphs with $o(n^2)$ edges is convergent, so this notion is meaningful only for dense graphs. There are various notions of convergence of bounded degree graphs and results covering the sparse and mixed regimes. More details can be found in the book *Large Networks and Graph Limits* by Lovász.

An equivalent way of defining the convergence is through homomorphisms. The *homomorphism density* of H in G, denoted by $t(H,G)$, is the probability that a random mapping from the vertices of H to the vertices of G is a homomorphism. It is not hard to show that a sequence of graphs $(G_n)_{n\in\mathbb{N}}$ is convergent if and only if the sequence of homomorphism densities $t(H,G_n)$ converges for every graph H.

The limit of a convergent sequence of graphs is represented by an analytic object called a graphon. Formally, a *graphon* is a symmetric measurable function $W : [0,1]^2 \to [0,1]$, where symmetric stands for the property that $W(x,y) = W(y,x)$ for all $x,y \in [0,1]$.

A *W-random graph of order n* is the random graph obtained by sampling n points $x_1,\dots,x_n$ independently and uniformly in the unit interval $[0,1]$ and joining the i-th vertex and the j-th vertex of the graph by an edge with probability $W(x_i,x_j)$. Note that if W is the constant graphon equal to $p \in [0,1]$, then the W-random graph of order n is the Erdős-Rényi random graph $G(n,p)$. The *density* of a graph H in a graphon W, denoted by $d(H,W)$, is the probability that the W-random graph of order $|H|$ is isomorphic to H. The definition of the W-random graph readily yields that

$$d(H,W) = \frac{|H|!}{|\operatorname{Aut}(H)|} \int\limits_{[0,1]^{|H|}} \prod_{v_i v_j \in E(H)} W(x_i,x_j) \prod_{v_i v_j \notin E(H)} (1 - W(x_i,x_j)) \,\mathrm{d}x_1 \cdots \mathrm{d}x_{|H|},$$

where $V(H) = \{v_1, \dots, v_{|H|}\}$ and $\operatorname{Aut}(H)$ is the automorphism group of H.

We can think of the interval $[0,1]$ as the set of vertices, and of the value $W(x,y)$ as the weight of the edge xy. Then the formula above is an infinite analogue of weighted density. Intuitively, one can also think of a graphon as an infinite analogue of the adjacency matrix.

We say that a graphon W is a *limit* of a convergent sequence $(G_n)_{n\in\mathbb{N}}$ of graphs if

$$d(H,W) = \lim_{n\to\infty} d(H,G_n) \text{ for every graph } H.$$

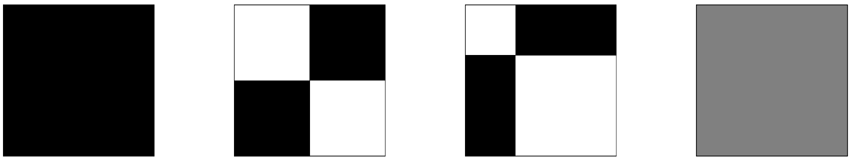

Figure 1: Graphons that are limits of the convergent sequences $(K_n)_{n\in\mathbb{N}}$, $(K_{n,n})_{n\in\mathbb{N}}$, $(K_{n,2n})_{n\in\mathbb{N}}$ and $(G(n,1/2))_{n\in\mathbb{N}}$, respectively. We use white for value 0, black for 1 and shades of grey for the intermediate values. The origin is in the top left corner.

For any graphon W, a rather straightforward application of the Azuma-Hoeffding inequality yields that the sequence of W-random graph of order n is convergent and W is its limit. Hence, every graphon is a limit of a convergent sequence of graphs. Showing that every convergent sequence of graphs has a limit graphon is more involved.

Theorem 1 (Lovász, Szegedy, 2006)**.** *For every convergent sequence of graphs there exists a graphon that is a limit of this sequence.*

Notice that the limit is not unique, because if we consider any measure preserving map $\varphi : [0,1] \to [0,1]$, then the graphon W^φ defined as $W^\varphi(x,y) = W(\varphi(x),\varphi(y))$ has the same density of any subgraph H as a graphon W. For example, the two graphons depicted in Figure 2 are both limits of the same convergent sequence $(G_n)_{n\in\mathbb{N}}$ of graphs where $G_n = K_{n,n}$. We say that graphons U and W are *weakly isomorphic* if $d(H,U) = d(H,W)$ for every graph H, i.e., the graphons U and W are limits of the same sequences of graphs.

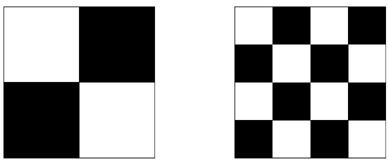

Figure 2: Two weakly isomorphic graphons.

The following result gives a characterization of weakly isomorphic graphons.

Theorem 2 (Borgs, Chayes, Lovász, 2010)**.** *Two graphons U and W are weakly isomorphic if and only if there exist measure preserving maps $\varphi, \psi : [0,1] \to [0,1]$ such that $U^\varphi = W^\psi$ almost everywhere.*

A *kernel* is a symmetric measurable function $W : [0,1]^2 \to \mathbb{R}$. In order to say that two graphons (or kernels) are close to each other, we define the *cut norm* on the space of kernels by

$$\|W\|_\square = \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W(x,y)\,\mathrm{d}x\,\mathrm{d}y \right|,$$

where the supremum is taken over all measurable subsets S and T. It is easy to verify that indeed it is a norm. It also defines a *cut metric* by $d_\square(U,W) = \|U - W\|_\square$ for any graphons U and W.

We still need to consider all possible changes of the graphon by measure preserving maps (somewhat equivalent to permuting vertices of a graph). Therefore, we consider the set $\overline{S}_{[0,1]}$ of all measure preserving maps $[0,1] \to [0,1]$ and the set $S_{[0,1]}$ of all invertible measure preserving maps $[0,1] \to [0,1]$. We define the *cut distance* of two graphons by

$$\delta_\square(U,W) = \inf_{\varphi\in S_{[0,1]}} d_\square(U, W^\varphi).$$

It is easy to verify that

$$\delta_\square(U,W) = \inf_{\varphi\in S_{[0,1]}} d_\square(U^\varphi, W) = \inf_{\varphi\in \overline{S}_{[0,1]}} d_\square(U, W^\varphi) = \inf_{\varphi,\psi\in \overline{S}_{[0,1]}} d_\square(U^\psi, W^\varphi).$$

An important property of the cut distance is the following theorem.

Theorem 3 (Borgs, Chayes, Lovász, Sós, Vesztergombi, 2008)**.** *Two graphons U and W are weakly isomorphic if and only if $\delta_\square(U,W) = 0$.*

Identifying weakly isomorphic graphons we get the set $\widetilde{\mathcal{W}}_0$ of *unlabeled graphons*. The above theorem gives that the cut distance is a metric on $\widetilde{\mathcal{W}}_0$. The crucial fact is that this space is compact.

Theorem 4 (Lovász, Szegedy, 2007)**.** *The space $\widetilde{\mathcal{W}}_0$ with metric $\delta_\square$ is compact.*