STA4513: Statistical Models of Networks	Fall 2014

Lecture 4 — October 1, 2014

Prof. Daniel M. Roy *Scribe: Mengye Ren*

This lecture closely follows L. Lovasz (2012) and online notes by M. Racz (2010).

1 Overview

In this lecture, we will discuss graph isomorphism and homomorphism, convergence in terms of homomorphism densities, and then define the limit object of a sequence of graphs, which is known as a graphon. We will conclude by defining a metric over the space of graphons which corresponds with the convergence of homomorphism densities.

2 Graphs

We begin by defining simple graphs.

Definition: Simple graphs are graphs with no self loops and no multiple edges. $G = (V(G), E(G))$, where $V(G)$ is the set of vertices and $E(G)$ is the set of edges. $E(G) \subseteq V(G) \times V(G)$.

Definition: Two simple graphs G and G' are isomorphic when there exists a bijection $\varphi : V(G) \to V(G')$ such that $(i, j) \in E(G) \iff \varphi(i, j) \in E(G')$.

Note: Isomorphism are adjacency-preserving and non-adjacency preserving maps.

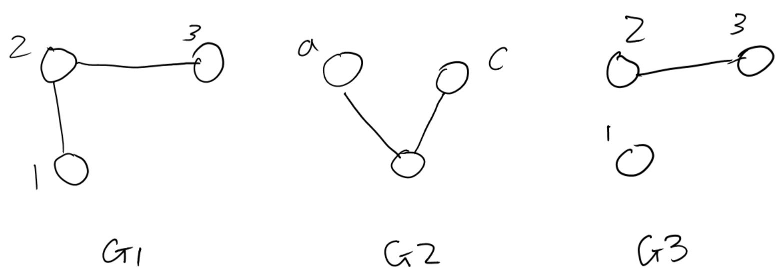

Figure 1: G_1, G_2, and G_3

Example: In the example, G_1 is isomorphic to G_2 but G_1 is not isomorphic to G_3.

2.1 Graph Parameters

Definition: A graph parameter is any function f defined on all simple graphs that is invariant under isomorphism. In other words, $f(G) = f(\varphi(G)) = (f \circ \varphi)(G)$ for all isomorphism φ.

Definition: Let F and G be simple graphs. A homomorphism from F to G is a map $\varphi : V(F) \to V(G)$ such that $(i, j) \in E(F) \implies (\varphi(i), \varphi(j)) \in E(G)$.

Note: We will write $F \to G$ if there exists a homomorphism from F to G.

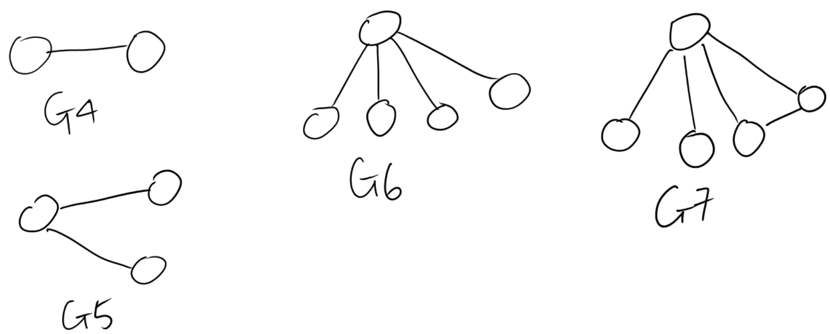

Figure 2: G_4, G_5, G_6, and G_3

Example: In the example, $G_5 \to G_4$ because we can map two distinct vertices in G_5 to a same vertex in G_4. By similar reason, $G_6 \to G_5$. However, there does not exist a homomorphism $G_7 \to G_6$ because of the triangle clique.

Definition: $K_n \equiv$ complete simple graph on $[n]$.

Note: $K_n \to G$ means G contains a K-clique. $G \to K_n$ means G is n-colourable.

Definition: $\hom(F, G) \equiv$ number of homomorphisms from F to G.

Definition: Homomorphism density of F in G:

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

Note: Homomorphism density can be viewed as

$$\mathbb{P}(\text{uniformly random } \varphi \text{ is a homomorphism } F \to G)$$

For a fixed G, $F \to \hom(F, G)$, $F \to t(F, G)$, $G \to \hom(F, G)$, $G \to t(F, G)$ are graph parameters.

3 Convergence Sequence of Graphs

Definition: A sequence of graphs G_n is convergent when, for every simple graph F, $t(F, G_n)$ is convergent. When the limit exists, we will write $t(F) = \lim_{n\to\infty} t(F, G_n)$.

Example:

$$t(K_1) = \lim_{n\to\infty} t(K_1, G_n) = 1$$

$$\frac{1}{2}t(K_2) = \frac{\frac{1}{2}\hom(K_2, G_n)}{|G_n|^{|K_2|}} = \frac{|E(G_n)|}{|V(G_n)|^2}$$

Note: If $|E(G_n)| \in o(|V(G_n)|^2)$ then $t(K_2) = 0 \Rightarrow t(F) = 0$ for all simple graph F such that $K_2 \to F$.

3.1 Why Homomorphism and t?

Definition: $\text{inj}(F, G) \equiv$ number of injective homomorphisms.

Definition: $\text{ind}(F, G) \equiv$ number of embeddings as induced subgraphs.

If $|V(G)| = n$ and $|V(F)| = k \leq n$ then,

$$t_{\text{inj}}(F, G) = \frac{\text{inj}(F, G)}{n(n-1)\dots(n-k+1)}$$

$$t_{\text{ind}}(F, G) = \frac{\text{ind}(F, G)}{n(n-1)\dots(n-k+1)}$$

(Similar to sampling without replacement).

Proposition: As number of vertices grows, the probability of any homomorphism is injective approaches 1. The following bound captures that fact:

$$|t(F, G) - t_{\text{inj}}(F, G)| \leq \frac{1}{|V(G)|}\binom{|V(F)|}{2} \xrightarrow{|V(G)|\to\infty} 0$$

We also have relations between number of injective homomorphisms and number of embedded subgraphs:

$$t_{\text{inj}}(F, G) = \sum_{F'\supseteq F} t_{\text{ind}}(F', G), \quad t_{\text{ind}}(F, G) = \sum_{F'\supseteq F} (-1)^{|E(F')|-|E(F)|} t_{\text{inj}}(F', G)$$

Note F' ranges over those graphs on the same vertex set as F and including at least those edges present in F.

3.2 Limit Object

If $\{G_n\}$ is convergent, then what is the limit? Say $V(G_n) = [n]$. One might guess that $\lim G_n$ is a graph on $\mathbb{N}^+ = \{1, 2, \ldots\}$. But this won't work in the standard model. (Why? It is related to the fact that there is no uniform distribution on $\mathbb{N}$.)

Limit object is a symmetric measurable function $W : [0,1]^2 \to [0,1]$ called a graphon. Here, symmetric means $W(x,y) = W(y,x)$.

For every simple graph F,

$$t(F,W) = \int_{[0,1]} \prod_{(i,j)\in E(F)} W(x_i,x_j)\,\mathrm{d}x_1,\ldots,\mathrm{d}x_k = \mathbb{E}\Big[\prod_{(i,j)\in E(F)} W(X_i,X_j)\Big]$$

where the X_i are i.i.d $U(0,1)$ random variables.

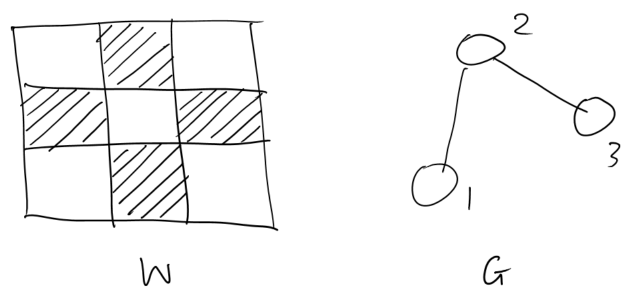

Figure 3: W and G

Example: In this example, $t(F,W) = t(F,G)$ We usually call W of black and white a "pixel picture" or "empirical graphon".

Theorem: For every convergent graph sequence G_n, there is a graphon W such that $t(F,G_n) \to t(F,W)$ (for every simple graph F). Moreover, the function W is unique up to a measure preserving transformation (m.p.t.).

Note: A function T is said to be a measure-preserving transformation if and only if $T(U) \overset{d}{=} U$, where $U \sim U(0,1)$.

Definition: W^T is the graphon given by $W^T(x,y) \equiv W(T(x),T(y))$.

To understand the theorem statement, note that, if W is a limit, then so is W^T, for any measure-preserving transformation T. Two graphons W, W' are said to be weakly isomorphic if there are two measure-preserving transformations T, T' such that $W^T = W'^{T'}$ almost everywhere.

Theorem: Every graphon W arises as the limit of a convergent graph sequence.

$$\{\text{space of graphons}\}/\{\text{weak isomorphism}\}$$
$$\cong$$
$$\{\text{homomorphism densities arising from limits of finite graphs}\}$$
$$\cong$$
$$\{\text{convergent graph sequences}\}/\{\text{same limit}\}$$

4 Distance Measure of Graphons

Definition: The edit distance is defined as follows:

$$d_1(W, W') \equiv \|W - W'\|_1 = \int \left|W(x, y) - W'(x, y)\right| \, \mathrm{d}x\mathrm{d}y$$

Definition: The cut distance is defined as follows:

$$d_\Box(W, W') \equiv \|W - W'\|_\Box = \sup_{S,T \subseteq [0,1]} \left| \int_{S \times T} W(x, y) - W'(x, y) \, \mathrm{d}x\mathrm{d}y \right|$$

Definition: The cut-metric is defined as follows:

$$\delta_\Box(W, W') \equiv \inf_{\varphi \in \text{m.p.t.}} d_\Box(W, W'^{\varphi})$$

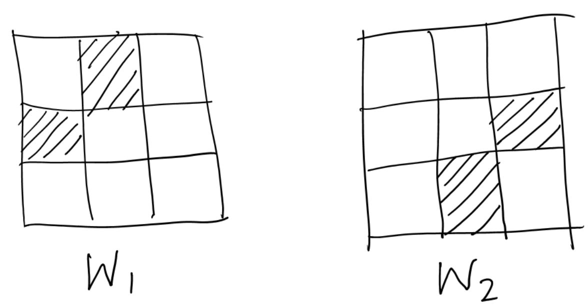

Figure 4: W_1 and W_2

Example: In the figure above, $d_1(W_1, W_2) = 4/9$, $d_\Box(W_1, W_2) = 2/9$, $\delta_\Box(W_1, W_2) = 0$.

Proposition: $\|W\|_\Box \leq \|W\|_1$

Lemma: (weak-regularity lemma): For every graphon W, there is a "step function graphon U" (see figure) with k parts such that

$$\|W - U\|_{\square} < \frac{2}{\sqrt{\log k}}.$$

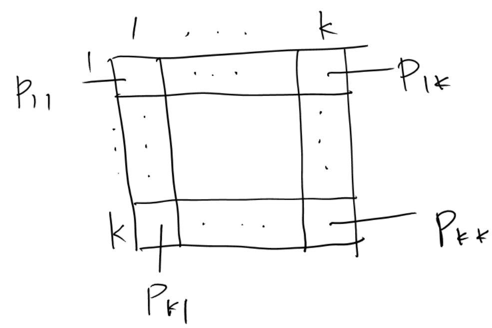

Figure 5: Parititition of a Graphon

Note the partition of $[0, 1]$ needs not to be an equipartition. This lemma provides us with a tool to approximate graphons using step functions.