Exchangeable Random Graphs, Graph Limits, and Graphons

Jiasen Yang

March 31, 2016

Department of Statistics
Purdue University

Outline

Graph Data

Exchangeable Random Graphs

Graph Limits and Graphons

Graphon Estimation

Graph Data

Graph Data

A graph $G = (V, E)$, where

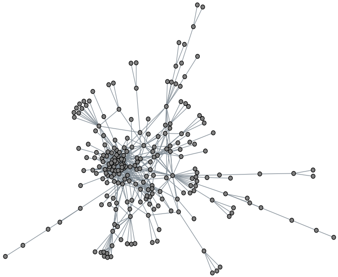
- $V = \{1, 2, \cdots, n\}$ is the set of vertices;
- $E \subseteq V \times V$ is the set of edges,

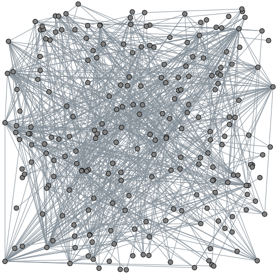
can be represented by an adjacency matrix A, where

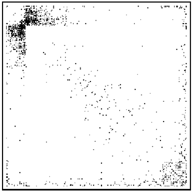
$$A_{ij} = \begin{cases} 1, & \text{if } (i, j) \in E; \\ 0, & \text{otherwise.} \end{cases}$$

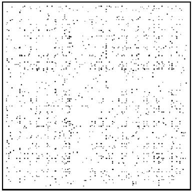
Assume G is simple and undirected $\Rightarrow A$ is binary, symmetric, and $A_{ii} = 0, \forall i$.

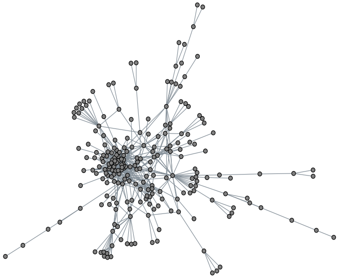
Graph Data (Cont'd)

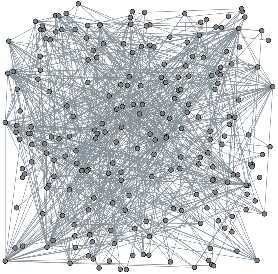




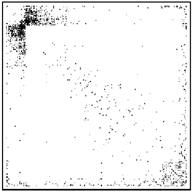


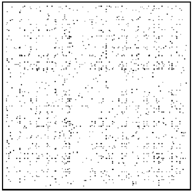






They are the same network!





Exchangeable Random Graphs

Definition

An exchangeable sequence is an infinite sequence $X_1, X_2, \cdots$ of random variables whose joint distribution satisfies

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2, \cdots) = \mathbb{P}(X_{\pi(1)} \in A_1, X_{\pi(2)} \in A_2, \cdots)$$

for every permutation π of $\mathbb{N} := \{1, 2, \cdots\}$ and collection $A_1, A_2, \cdots$ of (measurable) sets.

In words, order of observations does not matter.

i.i.d. $\Rightarrow$ exchangeable; exchangeable $\not\Rightarrow$ *i.i.d.*

P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Exchangeablility (Cont'd)

Example (Polyá's Urn scheme)
Consider an urn with b black balls and w white balls. Draw a ball at random and note its color. Replace the ball together with a balls of the same color. Repeat the procedure *ad infinitum*. Let $X_i = \mathbb{I}\{\text{the } i\text{-th draw yields a black ball}\}$.

The sequence $X_1, X_2, \cdots$ is exchangeable but not *i.i.d.*/Markov.

$$
\begin{aligned}
\text{E.g., } P(X_1 = 1, X_2 = 1, X_3 = 0, X_4 = 1) \\
&= \frac{b}{b+w}\frac{b+a}{b+w+a}\frac{w+a}{b+w+2a}\frac{b+2a}{b+w+3a} \\
&= \frac{b}{b+w}\frac{w+a}{b+w+2a}\frac{b+a}{b+w+a}\frac{b+2a}{b+w+3a} \\
&= P(X_1 = 1, X_2 = 0, X_3 = 1, X_4 = 1).
\end{aligned}
$$

Steffen Lauritzen. http://www.stats.ox.ac.uk/~steffen/teaching/grad/definetti.pdf

Theorem (de Finetti, 1931)
Let $X_1, X_2, \cdots$ be an infinite sequence of random variables with values in a space $\mathcal{X}$. The sequence $X_1, X_2, \cdots$ is exchangeable if and only if there is a random probability measure Θ on $\mathcal{X}$—i.e., a random variable with values in the set $\mathcal{M}(\mathcal{X})$ of probability distributions on $\mathcal{X}$—such that the X_i are conditionally i.i.d. given Θ and

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2, \cdots) = \int_{\mathcal{M}(\mathcal{X})} \prod_{i=1}^{\infty} \theta(A_i)\nu(d\theta)$$

where ν is the distribution of Θ.

P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Exchangeable Random Graphs

Adj. matrix of undirected graph = binary & symmetric array

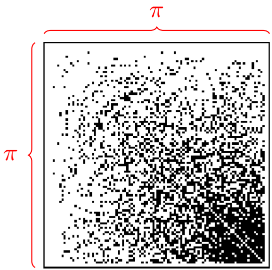
$\Rightarrow$ de Finetti exchangeability too strong!

$\Rightarrow$ Need generalization of exchangeability to arrays.

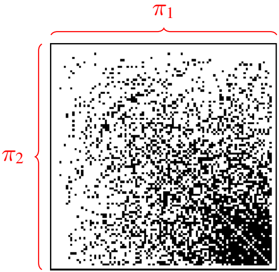
Adj. matrix of undirected graph = binary & symmetric array

$\Rightarrow$ de Finetti exchangeability too strong!

$\Rightarrow$ Need generalization of exchangeability to arrays.



$\sigma = \pi \otimes \pi$



$\sigma = \pi_1 \otimes \pi_2$

Definition

A random array $(X_{ij})_{i,j\in\mathbb{N}}$ is called separately exchangeable if

$$(X_{ij}) \stackrel{d}{=} (X_{\pi(i)\pi'(j)})$$

holds for every pair of permutations π, π' of $\mathbb{N}$.

Definition

A random array $(X_{ij})_{i,j\in\mathbb{N}}$ is called jointly exchangeable if

$$(X_{ij}) \stackrel{d}{=} (X_{\pi(i)\pi(j)})$$

holds for every permutation π of $\mathbb{N}$.

P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Theorem (Aldous, 1981-Hoover, 1979)

A random array $(X_{ij})_{i,j\in\mathbb{N}}$ is

- *jointly exchangeable iff. $\exists$ a random measurable function $F : [0,1]^3 \to \mathcal{X}$ s.t.*

$$(X_{ij}) \stackrel{d}{=} (F(U_i, U_j, U_{\{\{i,j\}\}}))$$

- *separately exchangeable iff. $\exists$ a random measurable function $F : [0,1]^3 \to \mathcal{X}$ s.t.*

$$(X_{ij}) \stackrel{d}{=} (F(U_i, U'_j, U_{ij}))$$

where $U_i, U_{\{\{i,j\}\}}, U'_j, U_{ij} \stackrel{i.i.d.}{\sim}$ Uniform$[0,1]$.

P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Definition

An undirected *random* graph $G = (\mathbb{N}, E)$ with an infinite (countable) vertex set $\mathbb{N}$ and a *random* edge set E is called an exchangeable random graph if its *random* adjacency matrix $A = (A_{ij})$ is a jointly exchangeable array.

Thus, G is exchangeable if its distribution is invariant under relabeling of the vertices.

P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Adjacency matrix $A = (A_{ij})$ symmetric & binary:

$$A_{ij} \stackrel{d}{=} F(U_i, U_j, U_{\{i,j\}}) \stackrel{d}{=} \mathbb{I}\{U_{\{i,j\}} < W(U_i, U_j)\}$$

P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Exchangeable Random Graphs

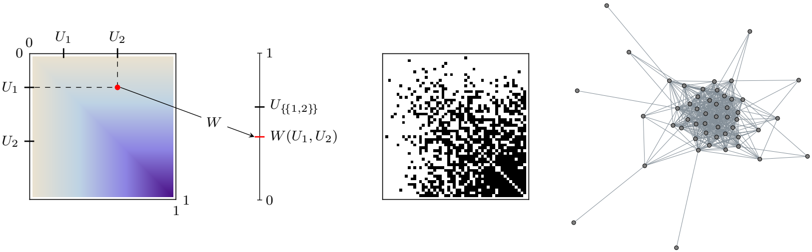
Adjacency matrix $A = (A_{ij})$ symmetric & binary:

$$A_{ij} \stackrel{d}{=} F(U_i, U_j, U_{\{i,j\}}) \stackrel{d}{=} \mathbb{I}\{U_{\{i,j\}} < W(U_i, U_j)\}$$

Sample $W : [0,1]^2 \to [0,1]$ measurable & symmetric (graphon);

Sample $U_1, U_2, \cdots \stackrel{i.i.d.}{\sim}$ Uniform$[0,1]$;

Sample $A_{ij} \sim$ Bernoulli$(W(U_i, U_j))$ for $i < j$.

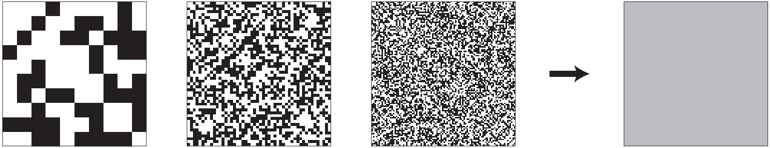


P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Graph Limits and Graphons

Graph Limits

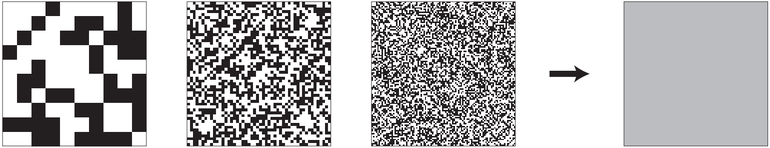
Rado graph: $A_{ij} = A_{ji} \overset{i.i.d.}{\sim} \text{Bernoulli}(1/2)$, $i \neq j \in \mathbb{N}$.



D. Glasscock, *What is a graphon?*, Notices of the AMS 62 (2015), no. 1.

Graph Limits

Rado graph: $A_{ij} = A_{ji} \overset{i.i.d.}{\sim} \text{Bernoulli}(1/2),\ i \neq j \in \mathbb{N}$.

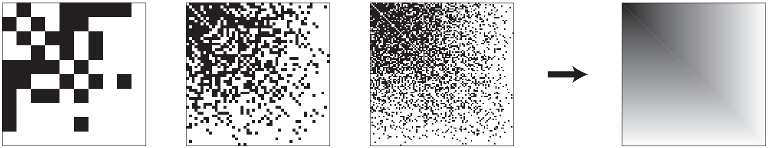


Limit graphon (*graph* function*on*):

$$f : [0,1]^2 \to [0,1],\ (x, y) \mapsto \frac{1}{2}.$$

D. Glasscock, *What is a graphon?*, Notices of the AMS 62 (2015), no. 1.

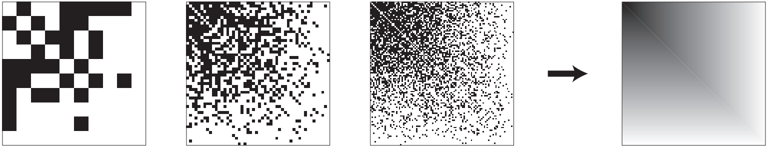
Growing uniform attachment: Let $G_1 = \bullet$. For $n \geq 2$, construct G_n from G_{n-1} by adding one new vertex, then, drawing an edge between each pair of non-adjacent vertices with probability $\frac{1}{n}$.



D. Glasscock, *What is a graphon?*, Notices of the AMS 62 (2015), no. 1.

Graph Limits (Cont'd)

Growing uniform attachment: Let $G_1 = \bullet$. For $n \geq 2$, construct G_n from G_{n-1} by adding one new vertex, then, drawing an edge between each pair of non-adjacent vertices with probability $\frac{1}{n}$.



Limit graphon:

$$f : [0,1]^2 \to [0,1],\ (x,y) \mapsto 1 - \max(x,y).$$

Definition

A labeled graphon is a symmetric, Lebesgue-measurable *a.e.* function $W : [0,1]^2 \to [0,1]$. A labeled graphon determines the equivalence class of graphons

$$[W] = \{W^{\phi} : (x,y) \mapsto W(\phi(x), \phi(y))\},$$

where ϕ is an invertible, measure-preserving transformation of $[0,1]$. Such equivalence classes are called unlabeled graphons.

D. Glasscock, *What is a graphon?*, Notices of the AMS 62 (2015), no. 1.

Definition

The cut distance between two labeled graphons W and U is

$$\delta_\square(W,U) = \inf_{\substack{\phi,\psi \text{ m.p.t.} \\ \text{of } [0,1]}} \sup_{\substack{S,T\subseteq[0,1] \\ \text{measurable}}} \left| \int_{S\times T} W^\phi(x,y) - U^\psi(x,y) \right|.$$

Theorem

Every graphon is the $\delta_\square$-limit of a sequence of finite graphs.

Theorem (Lovász-Szegedy)

Let $\mathcal{G}$ be the space of unlabeled graphons (modulo weak isometry). The metric space $(\mathcal{G}, \delta_\square)$ is compact.

D. Glasscock, *What is a graphon?*, Notices of the AMS 62 (2015), no. 1.

Graphon Estimation

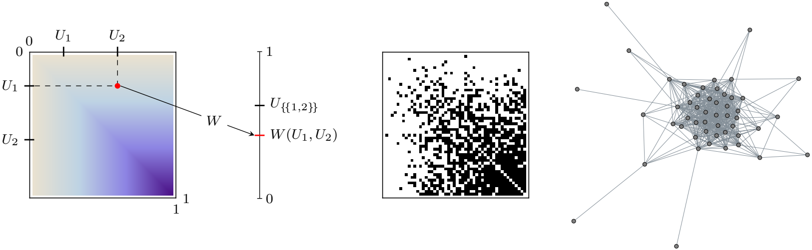
Adjacency matrix $A = (A_{ij})$ symmetric & binary:

$$A_{ij} \overset{d}{=} F(U_i, U_j, U_{\{i,j\}}) \overset{d}{=} \mathbb{I}\{U_{\{i,j\}} < W(U_i, U_j)\}$$

Sample $W : [0,1]^2 \to [0,1]$ measurable & symmetric (graphon);

Sample $U_1, U_2, \cdots \overset{i.i.d.}{\sim}$ Uniform$[0,1]$;

Sample $A_{ij} \sim$ Bernoulli$(W(U_i, U_j))$ for $i < j$.



P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Problem: Estimate graphon W from a set of observed networks.

Nonparametric regression: estimate W from $\{(U_i, U_j), A_{ij}\}_{i,j\in\mathbb{N}}$.

Challenge: design points $\{(U_i, U_j)\}$ are latent.

Proposed estimators:

- Stochastic blockmodel approximation [Airoldi *et.al.*, 2013], [Gao *et.al.*, 2015]
- Histogram estimator (sorting-and-smoothing) [Chan and Airoldi, 2014]
- Gaussian process model [Lloyd *et.al.*, 2012], [Orbanz and Roy, 2015]

E. M. Airoldi, T. B. Costa, and S. H. Chan. (2013) Stochastic Blockmodel Approximation of a Graphon: Theory and Consistent Estimation. *NIPS '13*.
Gao, C., Lu, Y. and Zhou, H. H. (2015). Rate-Optimal Graphon Estimation. *Ann. Statist.* 43(6):2624-2652.
S.H. Chan and E.M. Airoldi (2014). A Consistent Histogram Estimator for Exchangeable Graph Models. *ICML '14*.
J.R. Lloyd, P. Orbanz, Z. Ghahramani, and D.M. Roy (2012). Random Function Priors for Exchangeable Arrays with Applications to Graphs and Relational Data. *NIPS '12*.
P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

- Community detection
- Link prediction:

$$\mathbb{P}((i,j) \in E) = \mathbb{P}(A_{ij} = 1) = W(U_i, U_j)$$

- Network comparison & hypothesis testing

Bickel, P. J. and Chen, A. (2009). A Nonparametric View of Network Models and Newman-Girvan and Other Modularities. *Proc. Natl. Acad. Sci. USA* 106:21068-21073.

Bad News

Real-world graphs are **sparse** (finite # of edges per vertex).

- Power-law degree distribution;
- Small-world phenomena.

P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Real-world graphs are **sparse** (finite # of edges per vertex).

- Power-law degree distribution;
- Small-world phenomena.

Theorem (Misspecification)

If a random graph is exchangeable, it is either dense or empty.

Proof.

For a random graph G_n with n vertices, the expected proportion of present edges is

$$p := \int_{[0,1]^2} W(x,y)\,dx\,dy.$$

If $p = 0$, G_n is empty; if $p > 0$, G_n has $p \cdot \binom{n}{2} = \Theta(n^2)$ edges in expectation. $\square$

P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.

Open Problem

The limit object of a convergent sequence of sparse graphs in the cut metric is always the empty graphon.

Representation theorems for sparse random graphs?

P. Orbanz and D. M. Roy (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 37(2):437-461.
F. Caron and E.B. Fox (2015). Sparse Graphs using Exchangeable Random Measures. *arXiv:1401.1137*, 2015.

Thank you!