The Class of Random Graphs Arising from Exchangeable Random Measures Victor Veitch and Daniel M. Roy

`https://arxiv.org/abs/1512.03099`

Sampling and Estimation for (Sparse) Exchangeable Graphs Victor Veitch and Daniel M. Roy

`https://arxiv.org/abs/1611.00843`

And maybe also this survey:

Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures (with Peter Orbanz) IEEE Trans. Pattern Anal. Mach. Intelligence (PAMI), 2014.

`https://arxiv.org/abs/1312.7857`

The talk will be informal but based on slides I've produced for previous talks. Here's a recent abstract:

Statistical models of sparse networks from symmetries

The statistical analysis of network data rests on a foundation of random graph models, yet existing models are inadequate because they cannot model sparse graphs or fail to meet basic statistical criteria. We introduce a new class of random graphs, defined by the exchangeability of their vertices, and indexed by the expected number of edges, rather than the number of vertices. A straightforward adaptation of a result by Kallenberg yields a representation theorem: every such random graph is characterized by three (potentially random) components: a nonnegative real I in $\mathbb{R}_+$, a measurable function $S : \mathbb{R}_+ \to \mathbb{R}_+$, and a symmetric measurable function $W : \mathbb{R}_+^2$ to $[0, 1]$, where both S and W satisfy weak integrability conditions. We call the triple (I, S, W) a graphex, in analogy to graphons, which characterize the (dense) exchangeable graphs indexed by the cardinality of their vertex sets. I will present some results about the structure and consistent estimation of these random graphs, and what role they can play in the analysis of real-world networks.

I'm coming from a statistical modeling perspective. What is an observation? What is the model?

We need a framework for the statistical analysis of networks:

- family of distribution on large graphs (populations)
- distribution over observed (sub)graphs (samples)

These are often conflated.

Real world networks are sparsely connected.

- Random graph models should be sparse, $o(n^2)$ deges among n vertices.

Problem: There is no suitable general framework.

- Graphons cannot get you this. The problem is exchangeability.
- Do we need to abandon exchangeability to get sparsity? No. (Caron and Fox 2014, Kallenberg 1990)

We introduce a general class of random graphs suitable for modelling sparsely connected network structures. The math structures we exploit are exchangeable random measures.

Special cases:

- dense exchangeable graphs (graphon model)
- Caron and Fox's nonparametric Bayesian models of sparse graphs
- Sparse graphs with mall world, power law behaviors.

Preferential attachment yields these.

- Exchangeability can be related to a sampling process.
- Sampling process is perverse for large (sparse) graphs.

Example:

- Initialize: sum 1, total 2.
- Next draw: Bernoulli sum/total. Add y to sum, 1 to total.

The probability of a path depends only on where it ends; all that matters is the total counts of 0's and 1's.

There's another description of this process, identical at the level of what numbers pop out.

- Let theta be uniform from $[0, 1]$.
- Return bernoulli(theta).

Note this doesn't include mutation.

A sequence of random variables Y is exchangeable when

$$(Y_1, \ldots, Y_n) \overset{d}{=} (Y_{\pi(1)}, \ldots, Y_{\pi(n)})$$

for all $n \in \mathbb{N}$, permutation π of $[n]$.

Theorem 0.1 (De Finetti)**.** *TFAE:*

- *Y is exchangeable.*
- *Y is conditionally iid given some θ*
- *Exists f such that*

$$Y \overset{d}{(} f(\theta, U_1), f(\theta, U_2), \ldots)$$

for iid uniform $\theta, U_1, U_2, \ldots$.

Let's say n is large; look at a prefix $k \ll n$. That sequence is exchangeable. Insofar as n grows much larger than k, it looks like it is the draw from a structure like this.

Given a sample, I need to think it's a prefix of an infinite exchangeable sequence.

Represent a graph as an adjacency matrix, or as a function on $[0, 1]^2$, "graphical graphon".

Joint exchangeability on an array X on $\mathbb{N}^2$ means

$$(X_{ij}) \overset{d}{=} (X_{\sigma(i)\sigma(j)}).$$

Fix $n \in \mathbb{N}$, graphon $W : [0, 1]^2 \to [0, 1]$ symmetric, measurable. Generative model for $G(n, W)$:

1. Let $\xi_1, \xi_2, \ldots$ be iid uniform $[0, 1]$.
2. Connect vertices i and j for $i < j \le n$, independently with probability $W(\xi_i, \xi_j)$.

Write $G(W) = G(\infty, W)$.

Theorem 0.2 (Aldous-Hoover). *Let X be infinite symmetric adjacency matrix. TFAE:*

1. *X is jointly exchangeable.*
2. *X is W-random graph for some (potentially random) graphon W.*

Making the graph larger and larger, viewed from the sampling processes perspective, it's possible that the law of the sampling process will converge. The sequence of graph converges to a measurable function.

What can you learn about facebook by sampling neighborhoods? That's an open problem.

We address sampling with a uniform sampling. It gives relative frequency of subgraphs. Statistically it's a little perverse because we don't sample this way.

Ex. movie ratings: force people to watch movies. Expect underlying ratings to satisfy exchangeability.

Let $W, W_1, W_2, \ldots$ be a sequence of graphons. Write $W_n \to_\square W$ if $G(W_n) \to G(W)$ in distribution. Through random sampling they produce similar subgraphs.

Theorem 0.3 (Kallenberg 99). *Let $\widehat{W}_n$ be empirical graphon of $G(n, W)$. Then $W_n \to_\square W$ a.s.*

Multivariate sampling and the estimation problem for exchangeable arrays.

Corollary 0.4. *Exchangeable graphs are dense or empty a.s.*

Use linearity of expectation and exchangeability,

$$\mathbb{E}\,(\text{edges}) = \binom{n}{2} \mathbb{E}[X_{1,2}].$$

Caron and Fox 2014: go beyond dense structures. Each vertex gets a real-valued label. An edge between 1 and 2 means a point at (θ_1, θ_2). We can think of the points as a set in $\mathbb{R}_+^2$. We've lost the disconnected vertices. A random sample from a sparse graph has many disconnected vertices.

Call this an adjacency measure: each point $(\theta_i, \theta_j) \in \mathbb{R}_+^2$ is an edge ij.

Kallenberg exchangeable graph (Veitch-Roy 15)

Fix a size $s \in \mathbb{R}_+$. Fix a graphex $W : \mathbb{R}_+^2 \to [0, 1]$ symmetric, measurable.

Generative model for $K(s, W)$,

- Let $\Pi = \sum_i \delta_{(\theta_i, \xi_i)}$ be a unit-rate Poisson process on $[0, s] \times \mathbb{R}_+$.
- For every edge $i < j$, $i, j \in \mathbb{N}$, add edge (θ_i, θ_j) with probability $W(\xi_i, \xi_j)$.

Also star structures and isolated edges...

Sample from an infinite strip $[0,s]\times\mathbb{R}$. Decide whether to put an edge $(\theta_i,\xi_i),(\theta_j,\xi_j)$ by flipping $W(\xi_i,\xi_j)$. I forget about the isolated vertices.

It is like the limit of a graphon as it gets more sparse.

How to adjust to do preferential attachment? If you run preferential attachment for a long time and choose a vertex at random it looks locally treelike. You won't see that here; you see isolated edges.

If W is integrable, in expectation you get finitely many edges.

Let X be infinite adjacency measure. TFAE:

1. X is jointly exchangeable.
2. X is Kallenberg exchangeable graph for some (potentially random) graphex W.

At particular times, structure shows up.

Graphons can be represented as graphexes. Define on square of size $c\times c$. Let

$$W(x,y)=\begin{cases}\widetilde{W}(x/c,y/c), & x,y\leq c\\ 0, & \text{else.}\end{cases}$$

Call W a c-dilation of $\widetilde{W}$.

Generative model:

$$\begin{aligned} N_s &\sim Poi(c\cdot s) && (1)\\ \{\theta_i\} &\sim Uni[0,s] && (2)\\ \{\xi_i\} &\sim Uni[0,1] && (3)\\ G_s\{(\theta_i,\theta_j)\}|\xi_i,\xi_j &\sim Bernoulli(\widetilde{W}(\xi_i,\xi_j)). && (4)\end{aligned}$$

How many edges, vertices do we expect to see?

$$\mathbb{E}[e_s]=\frac{1}{2}s^2\left\|W\right\|_1.\mathbb{E}[v_s] \qquad = s\int_{\mathbb{R}_+}(1-e^{-s\mu_W(\lambda)})\,d\lambda. \tag{5}$$

$\mathbb{E}[v_s]$ depends on tail.

Ex. polynomial tail power law. Exponential tail in middle. Compactly supported dense.

For example, $W(x,y)=(x\neq y)(x+1)^{-2}(y+1)^{-2}$ gives

$$\begin{aligned} \mathbb{E}[e_s] &= s^2 && (6)\\ \mathbb{E}[v_s] &= \sqrt{\frac{\pi}{3}}s^{\frac{3}{2}} && (7)\\ P(D_s=k|G_s) &\xrightarrow{p} \frac{\Gamma(-\frac{1}{2}+k)}{2\sqrt{\pi}k!}k^{-\frac{3}{2}} && (8)\end{aligned}$$