The general class of (sparse) random graphs arising from exchangeable point processes

Victor Veitch Daniel M. Roy

Department of Statistical Sciences, University of Toronto

SSC 2016

Big Picture

Motivating Problem

Need families of random graphs for modelling network structures

Example

- network: friendships among n users of a social network
- model: family of random graphs $(G_n)_{n \in \mathbb{N}}$

Sparse Networks

Real world networks are sparsely connected

Random graph models should be *sparse*: $o(n^2)$ edges as number of observed nodes n becomes large.

Problem

No general framework for the statistical analysis of sparsely connected networks.

Sparse Networks

Real world networks are sparsely connected

Random graph models should be *sparse*: $o(n^2)$ edges as number of observed nodes n becomes large.

Problem

No general framework for the statistical analysis of sparsely connected networks.

Sparse Graphs

We have no satisfactory answers to some fundamental questions:

1. how should we parameterize the space of distributions on sparse graphs?
2. what can we learn about a large graph if we observe only a small subgraph? and what do we mean when we say "observe"?

Results

Results

We derive and study a general class of random graphs suitable for modelling network structures.

Special cases

- All dense (graphon) models
- Caron & Fox models
- Sparse graphs with e.g. small world and power law behaviour

Results

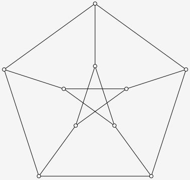
Results

We derive and study a general class of random graphs suitable for modelling network structures.

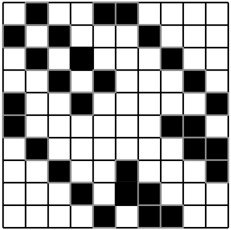
Special cases

- All dense (graphon) models
- Caron & Fox models
- Sparse graphs with e.g. small world and power law behaviour

Graphs, Adjacency, and Pixel Pictures



$$
\begin{matrix}
0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0
\end{matrix}
$$



[Lov12]

Dense Graphs

Graph models

- Basic object: infinite random binary matrix (X_{ij})
- Random graphs: $(G_n)_n$ defined by adjacency matrix as $n \times n$ upper left submatrix

Joint exchangeability of infinite random matrices

$(X_{ij}) \stackrel{d}{=} (X_{\sigma(i)\sigma(j)})$ for all permutations $\sigma \in S_\infty$ of the positive integers

Aldous-Hoover-Kallenberg (translated)

- Primitive of inference: $W : [0,1]^2 \to [0,1]$ (a *graphon*)
- Generative model for (X_{ij}) in terms of W

Dense Graphs

Graph models

- Basic object: infinite random binary matrix (X_{ij})
- Random graphs: $(G_n)_n$ defined by adjacency matrix as $n \times n$ upper left submatrix

Joint exchangeability of infinite random matrices

$(X_{ij}) \stackrel{d}{=} (X_{\sigma(i)\sigma(j)})$ for all permutations $\sigma \in S_\infty$ of the positive integers

Aldous-Hoover-Kallenberg (translated)

- Primitive of inference: $W : [0,1]^2 \to [0,1]$ (a *graphon*)
- Generative model for (X_{ij}) in terms of W

Dense Graphs

Graph models

- Basic object: infinite random binary matrix (X_{ij})
- Random graphs: $(G_n)_n$ defined by adjacency matrix as $n \times n$ upper left submatrix

Joint exchangeability of infinite random matrices

$(X_{ij}) \stackrel{d}{=} (X_{\sigma(i)\sigma(j)})$ for all permutations $\sigma \in S_\infty$ of the positive integers

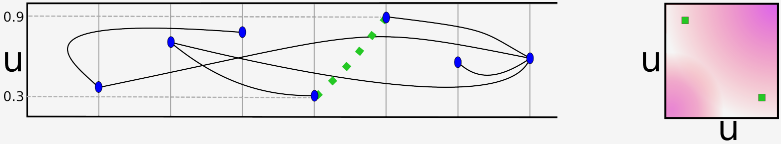
Aldous-Hoover-Kallenberg (translated)

- Primitive of inference: $W : [0,1]^2 \to [0,1]$ (a *graphon*)
- Generative model for (X_{ij}) in terms of W

Graphon Generative Model

Given a graphon $W : [0,1]^2 \rightarrow [0,1]$, sample a random graph by:

1. Assign each vertex i an iid $U[0,1]$ latent random variable
2. Include each edge (i,j) independently with probability $W(U_i, U_j)$



Edge $(4,5)$ is included with probability $W(0.3, 0.9) = W(0.9, 0.3)$.
The graphon W is shown as a heatmap on the right.

Representation Theorem

Recipe for constructing statistical models

- Assume a probabilistic symmetry on some infinite random structure
- Associated representation theorem picks out privileged family of distributions

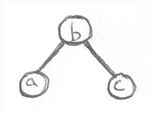
Examples

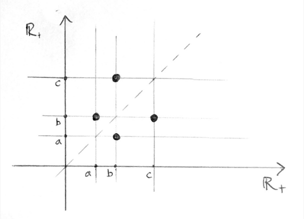
- de Finetti's representation theorem for exchangeable sequences
- Aldous-Hoover-Kallenberg theorem for exchangeable arrays

Caron & Fox 2014: (Sparse) Exchangeable Graphs

Key insights

- adjacency matrix $\rightarrow$ point process on $\mathbb{R}_+^2$
- array joint exchangeability $\rightarrow$ point process joint exchangeability





Graph edges correspond to points on $\mathbb{R}_+^2$

(Sparse) Graph Representation Theorem

Setup

- Random structure: point process on $\mathbb{R}_+^2$
- Finite graph G_s: truncate to $s \times s$ box
- Symmetry: joint exchangeability of point process

Representation theorem*

Distribution characterized by a *graphex*: a symmetric, integrable function $W : \mathbb{R}_+^2 \to [0, 1]$

(Sparse) Graph Representation Theorem

Setup

- Random structure: point process on $\mathbb{R}_+^2$
- Finite graph G_s: truncate to $s \times s$ box
- Symmetry: joint exchangeability of point process

Representation theorem*

Distribution characterized by a *graphex*: a symmetric, integrable function $W : \mathbb{R}_+^2 \to [0, 1]$

Graphex Model

Generative Model

Given W an infinite random graph is sampled by:

1. Sample a (latent) unit rate Poisson process Π on $\theta \times \vartheta$.
2. For each pair of points (θ_i, ϑ_i), $(\theta_j, \vartheta_j) \in \Pi$ include edge (θ_i, θ_j) with probability $W(\vartheta_i, \vartheta_j)$.
3. Include θ_i as a vertex whenever θ_i participates in at least one edge.

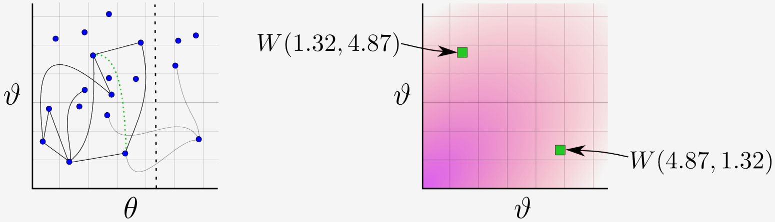


Figure : Graphex Model. **Graphex W is magenta heatmap.**

Sampling Distribution Results

Given graphex W we know:

1. the expected number of vertices and edges as a function of the size s
2. the asymptotic degree distribution
3. the asymptotic connectivity structure for certain families of graphexes.

Punchline

These models include a wide range of interesting graphs.

Estimation

Problem

How can we estimate a graphex?

Setup

- Observation: G_s (and s)
- Want Estimator: $\hat{W}_{G_s} : \mathbb{R}_+^2 \to [0, 1]$

Estimation

Problem

How can we estimate a graphex?

Setup

- Observation: G_s (and s)
- Want Estimator: $\hat{W}_{G_s} : \mathbb{R}_+^2 \to [0,1]$

Dense Graph Estimation

Empirical Graphon

Let $\widetilde{W}_n : [0,1]^2 \to \{0,1\}$ be the step function corresponding to the (arbitrarily permuted) adjacency matrix. $\widetilde{W}_n$ is a general non-parametric estimator for graphons.

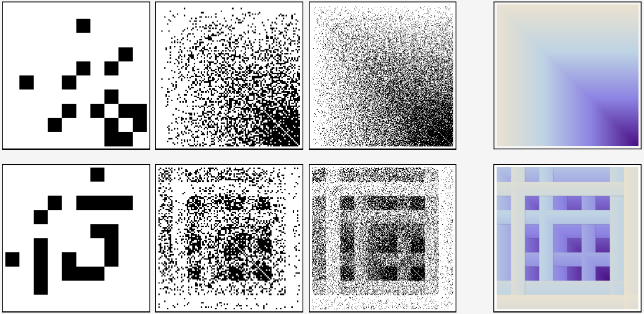


Figure : Empirical Graphon (from Orbanz Roy 2015)

Sparse Graph Estimation

Empirical Graphex

Let $\widetilde{W}^{G_s}$ be the empirical graphex of G_s, and define the empirical graphex $\widehat{W}^{G_s} : [0, \frac{v_s}{s}] \to \{0, 1\}$ by

$$\widehat{W}^{G_s}(x, y) = \widetilde{W}^{G_s}(\frac{v_s}{s}x, \frac{v_s}{s}y)$$

Theorem

Let $\widehat{G}_r^{(s)}$ be generated by $\widehat{W}^{G_s}(x, y)$ and let $\eta_r^{G_s}$ be the (random, G_s measurable) law of η^{G_s}. Then, almost surely,

$$\lim_{s \to \infty} \|\eta_r^{G_s} - \mathscr{L}(G_r)\|_{\mathrm{TV}} \to 0,$$

where $\mathscr{L}(G_r)$ is the law of a size r graph generated by W and $\|\cdot\|_{\mathrm{TV}}$ is the total variation distance.

Sparse Graph Estimation

Empirical Graphex

Let $\widetilde{W}^{G_s}$ be the empirical graphex of G_s, and define the empirical graphex $\widehat{W}^{G_s} : [0, \frac{v_s}{s}] \to \{0,1\}$ by

$$\widehat{W}^{G_s}(x,y) = \widetilde{W}^{G_s}(\frac{v_s}{s}x, \frac{v_s}{s}y)$$

Theorem

Let $\widehat{G}_r^{(s)}$ be generated by $\widehat{W}^{G_s}(x,y)$ and let $\eta_r^{G_s}$ be the (random, G_s measurable) law of η^{G_s}. Then, almost surely,

$$\lim_{s\to\infty} \|\eta_r^{G_s} - \mathscr{L}(G_r)\|_{\mathrm{TV}} \to 0,$$

where $\mathscr{L}(G_r)$ is the law of a size r graph generated by W and $\|\cdot\|_{\mathrm{TV}}$ is the total variation distance.

Sparse Graph Estimation

Estimation is also possible even when the size s is unknown
(but it's a little bit too long to state here)

Summary

Summary

- Representation theorem for sparse random graphs
- Extends dense (graphon) theory to sparse graphs
- Formulas for sampling distribution properties in terms of graphex
- General non-parametric estimator

Arxiv

1512.03099