Bayesian Models for Graph Data and Invariance in Networks

Peter Orbanz
Columbia University

Collaborators
Daniel M Roy (Cambridge)
Cameron Freer (MIT)

Balazs Szegedy (Toronto)
James R Lloyd and Zoubin Ghahramani (Cambridge)

Random structures

Sequences, graphs, matrices, d-arrays, trees, partitions, ranked lists, discrete measures, countable sets, hypergraphs, ...

General theme

exchangeable random structure

particularly simple distributions

mixing over simple distributions

$$P(X \in \,.\,) = \int_{\mathbf{T}} \mathbf{p}(\,.\,,\theta)\nu(d\theta)$$

- Characterizes "maximal" observation model and parameter space.
- Explains statistical averaging.
- Yields a law of large numbers.

Exchangeable sequences: de Finetti

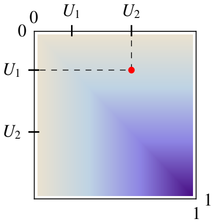
$$X = (X_1, X_2, \dots) \quad \text{exchangeable} \qquad \Leftrightarrow \qquad P(X \in \,.\,) = \int_{\mathbf{M}(\mathcal{X})} \prod_{i=1}^{\infty} \theta(X_i \in \,.\,)\nu(d\theta)$$

EXCHANGEABLE RANDOM GRAPH

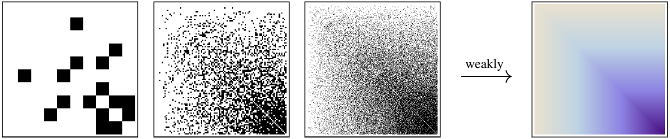
Representation theorem (Aldous, Hoover, Kallenberg)

Any exchangeable graph can be sampled as:

- Sample $W : [0, 1]^2 \longrightarrow [0, 1]$
  (measurable and symmetric)
- Sample $U_1, U_2, \ldots \sim_{\text{iid}} \text{Uniform}[0, 1]$
- Sample edge i,j $\sim \text{Bernoulli}(W(U_i, U_j))$



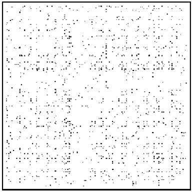
Law of large numbers (Kallenberg '99)

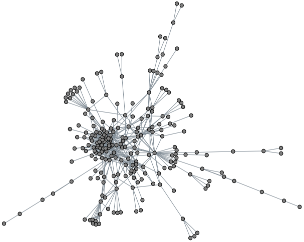


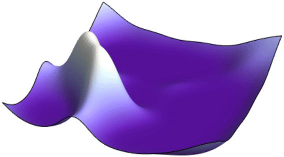
Up to equivalence

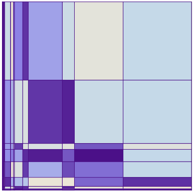
Nonparametric Regression

with Lloyd, Gharamani, Roy (2012)

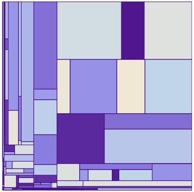




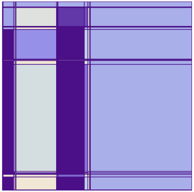




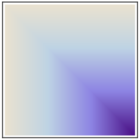
Infinite Relational Model
(Kemp et al, '06)



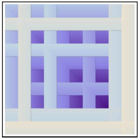
Mondrian process
(Roy & Teh, '08)

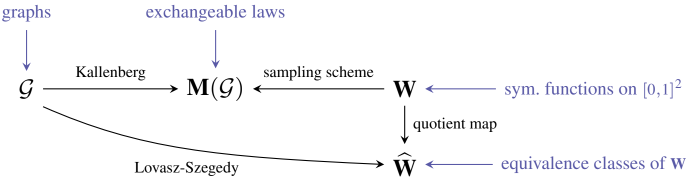


Latent feature relational model
(Miller et al, '09)



$\equiv$





Lifting Theorem (O. & Szegedy, 2011)

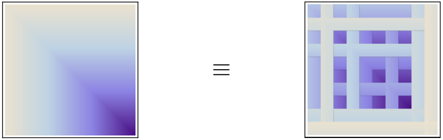
There exists a measurable mapping $\xi : \widehat{\mathbf{W}} \longrightarrow \mathbf{W}$ such that

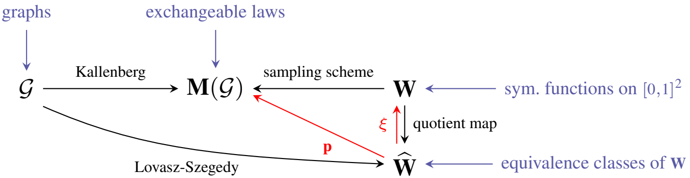
$$\xi(\hat{w}) \in [\hat{w}]_{\equiv} \qquad \text{for all } \hat{w} \in \widehat{\mathbf{W}} .$$

This result is *not* constructive.

LIFTINGS OF GRAPH LIMITS

with B. Szegedy





Lifting Theorem (O. & Szegedy, 2011)

There exists a measurable mapping $\xi : \widehat{\mathbf{W}} \longrightarrow \mathbf{W}$ such that

$$\xi(\hat{w}) \in [\hat{w}]_{\equiv} \qquad \text{for all } \hat{w} \in \widehat{\mathbf{W}} \; .$$

This result is *not* constructive.

WHY ARE GRAPHONS SO COMPLICATED?

	limit object	randomness
de Finetti	unique	two layers
Aldous-Hoover	not unique	three layers

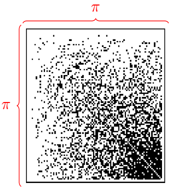
Permutations as constraints

$$P(X \in \,.\,) = \int \mathbf{p}(\,.\,,\theta)\nu(d\theta)$$

Informally: more constraints $\rightarrow$ stronger representation result

Exchangeability

Structure	index set	invariance under
sequences	$\mathbb{N}$	*all* permutations of $\mathbb{N}$
graphs	$\mathbb{N}^2$	product permutations $\pi \otimes \pi$



Exchangeable graphs are dense (or empty)

$$p = \int_{\triangle} W(x,y)dxdy \qquad \hat{p}_n = \frac{\#\text{ edges observed in } G_n}{\#\text{ edges in complete graph}} = p \cdot \Theta(n^2) = \Theta(n^2)$$

This is a consequence of looking at the graph "globally".

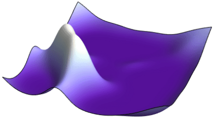
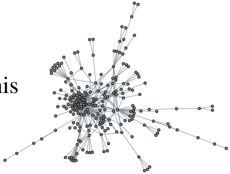
More generally

Exchangeable random structures have a density property.

Exchangeable binary sequence: cond. i.i.d. Bernoulli(P) $\rightarrow$ limiting ratio of 1s is P

Can we model network structure (not just edge density)?

Under an exchangeable model, we will never

sample from  and see something like this 

Probabilistic symmetries

random structure with with symmetry property → $P(X \in \,.\,)$

particularly simple distributions → $\mathbf{p}(\,.\,,\theta)$

mixing over simple distributions → $\nu(d\theta)$

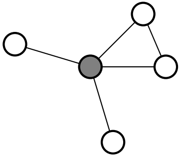
$$P(X \in \,.\,) = \int_{\mathbf{T}} \mathbf{p}(\,.\,,\theta)\nu(d\theta)$$

In principle, this also works for other symmetries than exchangeability.

Problem

Can we find a type of invariant random graphs that:

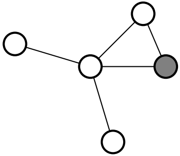
1. Can generate network structures?
2. Have useful statistical properties?



Involution invariance

A random graph with a marked location (vertex) v is **involution invariant** if the distribution of the neighborhood of v is invariant under moving the location along the graph.

This is a property of a random **rooted graph** (G, v).



Applicability

- Has a limit theory, but limit object contains much less information than dense graph limit.
- Much too weak for use in statistics.
- From graph theoretic perspective: Various interesting properties do not carry over to the limit.

Intuition: Too few constraints (1 shift per shift-length).

SUMMARY

Under exchangeability assumption

- Aldous-Hoover theorem explains statistical averaging, convergence of empirical measures and parameter space.
- Resulting graphs are dense, ie models are misspecified.

Sparse graphs

- Various models, but no statistical framework.
- To obtain a de Finetti-like result, we would need an invariance weaker than exchangeability, but stronger than shift invariance.
- No useful example of such an invariance is known.