Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures

Yuhua Zhang

September 20, 2019

Definitions
Adjacency matrix: A matrix $X \in \{0,1\}^{n\times n}$, where n is the number of nodes, and $X_{ij} = X_{ji} = 1$ if there's an edge between node i and node j.
The entries in X can be generated from an underlying probability matrix P, with $X_{ij} \sim P_{ij}$.
Exchangeable: Suppose (X_i) is an infinite sequence of random variables in a sample space $\mathbb{X}$. We call (X_i) **exchangeable** if its joint distribution satisfies

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

for every permutation π of $\mathbb{N} := \{1, 2, ...\}$ and every collection of sets A_1, A_2,...
Ergodic distribution/measure: a special family of distributions on X_∞. Each element $\theta \in T$ (T is the parameter space, $T := M(X)$) determines an ergodic distribution. Denote the set of ergodic distribution as

$$\{p_\theta : \theta \in T\} \subset M(X_\infty)$$

The distribution of any exchangeable random structure X_∞ can then be represented as a mixture of these ergodic distributions,

$$\mathbb{P}(X_\infty \in .) = \int_T p_\theta(.)\nu(d\theta)$$

Exchangeable partition: An exchangeable partition is a random partition X_∞ of $\mathbb{N}$ which is invariant under permutations of $\mathbb{N}$. (Probability of a partition depends only on the relative sizes of the block, not on which elements are in which block)
Random matrix and random graph: Denote the random 2-array X_∞ as:

$$X_\infty = (X_{ij}) = \begin{bmatrix} X_{11} & X_{12} & \dots \\ X_{21} & X_{22} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

A random matrix is a random 2-array; a random graph is a random matrix with $X = \{0, 1\}$.
Jointly & Separately exchangeable: A random 2-array (X_{ij}) is called **jointly exchangeable** if

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

for every permutation π of $\mathbb{N}$, and **separately exchangeable** if

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

for every pair of permutations π, π' of $\mathbb{N}$.
(For undirected graph, the jointly exchangeable can be used to describe the adjacency matrix.)
Exchangeable graph: For a graph on a countably infinite vertex set, say $\mathbb{N}$. A random graph G is given by a random subset of $\mathbb{N} \times \mathbb{N}$. A symmetry property of a random graph is the invariance of its distribution to a permutation of its vertex set. In this case, G is said to be an **exchangeable graph**.
That is, if $X = (X_{ij}) \sim P$, then for any permutation π, $(X_{\pi(i)\pi(j)} \sim P)$.
Graphon: A symmetric measurable function from $[0, 1]^2$ to $[0, 1]$ is called a graphon.

Another way to understand graphon:

1) Each vertex j is assigned an independent random value $u_j \sim U[0, 1]$

2) Edge (i, j) is independently included in the graph with probability $f(u_i, u_j)$

Example: Erdos-Renyi Model

Each edge is included independently with probability p. It can be generalized as:

1) Divide the unit square into $k \times k$ blocks;

2) Let f equals to p_{lm} on the l, mth block;

The statistical models of exchangeable simple graphs are parametrized by graphons. The problem of estimating the distribution of an exchangeable graph can be formulated as a regression problem on the unknown function w ($X_{ij} \sim Bern(w(u_i, u_j))$).

Identifiability

Two distinct graphs may parameterize the same random graph. In this case, the two graphons are called weakly isomorphic, and is not identifiable.
Graph limit: A sequence $(g_n)_{n\in\mathbb{N}}$ of graphs converges if $\delta(w_{g_n}, w) \to 0$ for some measurable function $w : [0, 1]^2 \to [0, 1]$. The function w is called the limit of (g_n), and often referred to as a **graph limit**.
Sparse & Dense Graph: Let (g_n) be a sequence of graphs $g_n = (v_n, e_n)$, where g_n has n vertices. We say that the sequence is **sparse** if, as n increases, $|e_n|$ is of size $O(n)$ ($O(.)$: upper-bound). It is called dense if $|e_n| = \Omega(n^2)$ ($\Omega(.)$: lower-bound).
(Graph limit is inherently a theory of dense graph)

Theorems
2-1 (de Finetti) Let $(X_1, X_2, ...)$ be an infinite sequence of random variables with values in a space X. the sequence $X_1, X_2, ...$ is exchangeable if and only if there is a random probability measure Θ on X such that the X_i are conditionally *i.i.d.* given Θ and

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

where ν is the distribution of Θ.
The random measure Θ is called the **directing random measure** of X. Its distribution ν is called the **mixing measure** or **de Finetti measure**.

2-2 If the sequence (X_i) is exchangeable, the empirical distributions

$$\hat{S_n}(.) := \frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}(.)$$

converge to Θ, in the sense that

$$\hat{S_n}(A) \to \Theta(A)$$

as $n \to \infty$ holds with probability for every set A.

2-9 Let X_1, X_2,... be an infinite, exchangeable sequence of random variables with values in a space X. Then there exists a random function F (inverse CDF) from [0,1] to X such that, if U_1, U_2,... is an *i.i.d.* sequence of uniform random variables,

$$(X_1, X_2, ...) = (F(U_1), F(U_2), ...)$$

3-2 (Aldous-Hoover). A random array (X_{ij}) is jointly exchangeable if and only if it can be represented as follows: There is a random function $F : [0,1]^3 \to X$ such that

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

where $(U_i)_{i\in\mathbb{N}}$ and $(U_{\{i,j\}})_{i,j\in\mathbb{N}}$ are, respectively, a sequence and an array of i.i.d. Uniform[0,1] random variables, which are independent of F.
3-6 Let G be a random simple graph with vertex set $\mathbb{N}$ and let X be its adjacency matrix. Then G is an exchangeable graph if nd only if there is a random function W from $[0,1]^2$ to $[0,1]$ such that

$$(X_{ij}) = (\mathbf{1}\{U_{i,j} < W(U_i, U_j)\})$$

where U_i and $U_{i,j}$ are independent i.i.d. uniform variables that are independent of W. (W is a graphon)