Exchangeable Array and Random Graph

Dongwoo Kim
ANU

April 12, 2016

1 Exchangeable array for graph

Definition 1 (Jointly exchangeable array)**.** A random 2-array $(X_{ij})_{i,j\in\mathbb{N}}$ is jointly exchangeable if

$$(X_{ij}) \stackrel{d}{=} (X_{\pi(i)\pi(j)}) \quad \text{for } i,j \in \mathbb{N}^2 \tag{1}$$

for any permutation π of $\mathbb{N}$.

Theorem 1 (Aldous, Hoover Theorem [Aldous, 1981, Hoover, 1979])**.** *A random 2-array $(X_{ij})_{i,j\in\mathbb{N}}$ is jointly exchangeable if and only if there exists a random measurable function $f : [0,1]^3 \to \mathbf{X}$ such that*

$$(X_{ij}) \stackrel{d}{=} (f(U_i, U_j, U_{ij})), \tag{2}$$

where $(U_i)_{i\in\mathbb{N}}$ and $(U_{ij})_{ij>i\in\mathbb{N}}$ with $U_{ij} = U_{ij}$ are a sequence and matrix of i.i.d. Uniform[0,1] random variables.

The function f is a symmetric in its first two argument.

For the undirected graph case $\mathcal{X} = \{0,1\}$, the theorem can be simplified further through a random function called *graphon* $W : [0,1]^2 \to [0,1]$, symmetric in its arguments, where

$$f(U_i, U_j, U_{ij}) = \begin{cases} 1 & U_{ij} < W(U_i, U_j) \\ 0 & \text{otherwise} \end{cases}$$

Example 1.1 (Random function prior on function f [Lloyd et al., 2013])**.** Lloyd et al. use a Gaussian process prior to define function f for an undirected graph. They define $W(U_i, U_j) = \phi(\Theta(U_i, U_j))$ where ϕ is a logstic function, and $\Theta(\cdot,\cdot)$ is a continuous function with a Gaussian process prior. Thus, the probability of edge between node i and j is equal to Bernoulli($\phi(\Theta(U_i, U_j))$).

Definition 2 (Sparse Graph)**.** Let the number of nodes in a graph be n. The graph is sparse if the number of edges are $o(n^2)$ or dense if the number of edges are $\Theta(n^2)$.

Remark 1 (Graphon is trivially dense)**.** Every graph represented by graphon W are either empty or dense. The asymptotic proportion of edges is $p = \frac{1}{2}\int W(x,y)dxdy$ and the graph is hence either empty ($p = 0$) or dense (since $O(pn^2) = O(n^2)$).

In [Lloyd et al., 2013], the authors place a Gaussian process prior over graphon $W : [0,1]^2 \to \mathbb{R}$ and transform the output through the logistic function to model the edge probability between nodes.

For the undirected graph case where $X_{ij} = X_{ji}$, one can sample upper triangle of the adjacency matrix, and use the same result for the lower triangle. However, for the directed graph case, by the theorem, both X_{ij} and X_{ji} rely on single parameter U_{ij} which means one should jointly sample $(X_{ij},\ X_{ji})$ together from three parameters U_i, U_j, U_{ij}. Thus, X_{ij} and X_{ji} are not conditionally independent. Also, asymmetric graphon W might be employed for directed random graph. However, [Cai et al., 2015] show that assymetric graphon is inappropriate to impose certain structures on a graph such as the partial ordering and propose a class of priors for directed graphs.

2 Exchangeable array for matrix factorisation

Unlike the graph model, jointly exchangeability for a general matrix where rows and columns represent different entities (i.e. users and items in recommendation system) is

Definition 3 (Separately exchangeable array [Orbanz and Roy, 2015])**.** A random 2-array $(X_{ij})_{i,j\in\mathbb{N}}$ is separately exchangeable if

$$(X_{ij}) \overset{d}{=} (X_{\pi(i)\sigma(j)}) \quad \text{for } i, j \in \mathbb{N}^2 \tag{3}$$

for any permutation π and σ of $\mathbb{N}$.

Corollary 1.1 (Separately exchangeable array)**.** *A random 2-array (X_{ij}) is separately exchangeable if and only if there exists a random measurable function $f : [0,1]^3 \to \mathbb{X}$ such that*

$$(X_{ij}) \overset{d}{=} (f(U_i, U_j, U_{ij})), \tag{4}$$

where $(U_i)_{i\in\mathbb{N}}$, $(U_j)_{j\in\mathbb{N}}$, and $(U_{ij})_{ij\in\mathbb{N}^2}$ are i.i.d. uniform[0,1] random variables.

In the separately exchangeable array case, function f is not a symmetric in its first two arguments.

Example 3.1 (Probabilistic matrix factorisation (PMF) [Salakhutdinov and Mnih, 2008])**.** PMF is one instantiation of Corollary 1.1. Let the generative process of PMF be

$$U_i \sim \mathcal{MN}_d(0, \Sigma_U) \tag{5}$$
$$V_j \sim \mathcal{MN}_d(0, \Sigma_V) \tag{6}$$
$$X_{ij} \sim \mathcal{N}(U_i^\top V_j, \sigma_x) \tag{7}$$

where $\mathcal{MN}_d$ is a d-dimensional zero-mean multivariate normal distribution with covariance Σ. This corresponds to random measurable function $f(U_i, U_j, U_{ij}) = \Phi_1(U_{ij}; \Phi_d(U_i; 0, \Sigma_U)^\top \Phi_d(U_j; 0, \Sigma_U), \sigma_x^2)$, where $\Phi_d(\cdot; \mu, \sigma^2)$ is an inverse-CDF of d-dimensional multivariate function with mean μ and variance σ^2.

3 Sparse exchangeable array

Theorem 2 (Kallenberg's exchangeable theorem [Kallenberg, 1990])**.** *Representation theorems for jointly exchangeable random measures on $\mathbb{R}^2$*

Matrix factorisation for the sparse bipartite graph has been partially discussed in [Caron, 2012].

References

[Aldous, 1981] Aldous, D. J. (1981). Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, 11(4):581–598.

[Cai et al., 2015] Cai, D., Ackerman, N., and Freer, C. (2015). Priors on exchangeable directed graphs. pages 1–17.

[Caron, 2012] Caron, F. (2012). Bayesian nonparametric models for bipartite graphs. *Advances in Neural Information Processing Systems 25 (Proceedings of NIPS)*, pages 1–9.

[Hoover, 1979] Hoover, D. N. (1979). Relations on probability spaces and arrays of random variables. *Preprint, Institute for Advanced Study, Princeton, NJ*, 2.

[Kallenberg, 1990] Kallenberg, O. (1990). Exchangeable random measures in the plane. *Journal of Theoretical Probability*.

[Lloyd et al., 2013] Lloyd, J., Orbanz, P., Ghahramani, Z., and Roy, D. (2013). Random function priors for exchangeable arrays with applications to graphs and relational data. In *Advances in Neural Information Processing Systems 26*, pages 1–9.

[Orbanz and Roy, 2015] Orbanz, P. and Roy, D. M. (2015). Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures. *TPAMI*, pages 1–25.

[Salakhutdinov and Mnih, 2008] Salakhutdinov, R. and Mnih, A. (2008). Bayesian probabilistic matrix factorization using Markov chain Monte Carlo. In *Proceedings of the 25th International Conference on Machine Learning (2008)*, volume 25, pages 880–887.