

Discussion of ‘Sparse graphs using exchangeable random measures’ by François Caron and Emily B. Fox.

Dr. Benjamin Bloem-Reddy, Columbia University, New York City, U.S.A.
benjamin.reddy@columbia.edu

Research in this area is progressing rapidly, and I would like to draw connections to some concurrent work. See also Janson (2017, Section 5.1) for related ideas.

For every model based on an exchangeable random measure (ERM) (see also Veitch and Roy (2015, 2016), Borgs et al. (2016), Janson (2016)), one can construct a so-called *edge exchangeable* model (Crane and Dempsey, 2015, 2016; Williamson, 2016; Cai et al., 2016; Janson, 2017) that coincides with a finite restriction of the ERM model. The converse is true only in some cases. (See Cai et al. (2016) for an example of when the converse does not hold). I consider the CRM case for concreteness.

Let X_1, X_2, \dots be an exchangeable sequence of edges, and let \mathcal{E}_n be the directed multigraph composed of the first n edges labeled by their order of appearance. If one considers the edge(s) connecting a pair of vertices as having a ‘color’ unique to that pair, Kingman’s paint-box theorem shows that every such graph can be generated by sampling

$$\begin{aligned} \Phi &\sim \mu \\ X_i \mid \Phi &\stackrel{\text{iid}}{\sim} \Phi \quad \text{for } i = 1, \dots, n, \end{aligned}$$

where Φ is a random discrete probability measure sampled from a mixing measure μ (Crane and Dempsey, 2016). Denote such a graph by \mathcal{E}_n^Φ .

Consider when Φ is the normalized CRM product measure defined in Section 5.5.1., $\Phi_\alpha := W_\alpha \times W_\alpha / W_\alpha^{*2}$. Then the graph D_α , with D_α^* edges, and $\mathcal{E}_{D_\alpha^*}^{\Phi_\alpha}$ have the same conditional law,

$$\mathcal{L}(D_\alpha \mid W_\alpha, D_\alpha^*) = \mathcal{L}(\mathcal{E}_{D_\alpha^*}^{\Phi_\alpha} \mid \Phi_\alpha, D_\alpha^*).$$

If μ places non-zero mass only on normalized CRMs with mean measure $\rho(dw)\lambda_\alpha(d\theta)$, then equality in distribution holds unconditionally. However, an ERM model and its counterpart edge exchangeable model intersect only for a particular α . D_α and $\mathcal{E}_{D_\alpha^*}^{\Phi_\alpha}$ grow differently: Fix $\varepsilon > 0$, and let D_α grow to $D_{\alpha+\varepsilon}$, denoting by $D_{\alpha,\varepsilon}^* := D_{\alpha+\varepsilon}^* - D_\alpha^*$ the number of additional edges. Then

$$\mathcal{L}(D_{\alpha+\varepsilon} \mid D_\alpha, W_{\alpha+\varepsilon}, D_{\alpha,\varepsilon}^*) \neq \mathcal{L}(\mathcal{E}_{D_\alpha^*+D_{\alpha,\varepsilon}^*}^{\Phi_\alpha} \mid \mathcal{E}_{D_\alpha^*}^{\Phi_\alpha}, \Phi_\alpha, D_{\alpha,\varepsilon}^*).$$

The inequality reflects a fundamental difference in how the two model classes encode the notion of growth, and offers practitioners guidance when choosing a model appropriate for their data. Edge exchangeable models posit that graphs grow one edge at a time; growth in ERM models is by a random number of edges as α increases. As a consequence, as an edge exchangeable graph grows to arbitrary size, an edge may occur between two vertices that were previously not connected. This is not true for ERM models; exchangeability of the random measure requires that new edges form only between new vertices. Furthermore, ERM models allow for a growing population of edges from which to sample; edge exchangeable models sample from a fixed (possibly infinite) population.

References

- Borgs, Christian, Jennifer T. Chayes, Henry Cohn, and Nina Holden (2016). “Sparse exchangeable graphs and their limits via graphon processes”. In: arXiv: 1601.07134 [math.PR]. URL: <http://arxiv.org/abs/1601.07134>.
- Cai, Diana, Trevor Campbell, and Tamara Broderick (2016). “Edge-exchangeable graphs and sparsity”. In: *Advances in Neural Information Processing Systems 29*. Ed. by D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett. Curran Associates, Inc., pp. 4242–4250.
- Crane, Harry and Walter Dempsey (2015). “A framework for statistical network modeling”. In: arXiv: 1509.08185 [math.ST]. URL: <https://arxiv.org/abs/1509.08185>.
- (2016). “Edge exchangeable models for network data”. In: arXiv: 1603.04571 [math.ST]. URL: <https://arxiv.org/abs/1603.04571>.
- Janson, Svante (2016). “Graphons and cut metric on sigma-finite measure spaces”. In: arXiv: 1608.01833 [math.CO]. URL: <https://arxiv.org/abs/1608.01833>.
- (2017). “On edge exchangeable random graphs”. In: eprint: 1702.06396. URL: <https://arxiv.org/abs/1702.06396>.
- Veitch, Victor and Daniel M. Roy (2015). “The Class of Random Graphs Arising from Exchangeable Random Measures”. In: arXiv: 1512.03099 [math.ST]. URL: <http://arxiv.org/abs/1512.03099>.
- (2016). “Sampling and Estimation for (Sparse) Exchangeable Graphs”. In: arXiv: 1611.00843 [math.ST]. URL: <https://arxiv.org/abs/1611.00843>.
- Williamson, Sinead A. (2016). “Nonparametric Network Models for Link Prediction”. In: *Journal of Machine Learning Research* 17.202, pp. 1–21.