Optimal Bayesian inference in preferential attachment networks

Dr. Fengnan Gao

December 4, 2023

Abstract

The preferential attachment (PA) mechanism is a paradigm to model the prevalent rich-get-richer phenomenon in dynamic networks observed in nature and society. In its simplest form, it starts from two nodes connected to each other, and new nodes keep coming in. Each incoming new node establishes only one connection with one of the existing nodes. The probability to choose the node to connect to is proportional to preferences based on a PA function, mapping the natural numbers "degrees" to a positive real number "preference". The PA function is assumed a priori non-decreasing, which means the nodes with high degrees are more likely to inspire new connections, i.e., "the rich get richer". We are interested in recovering the mechanism responsible for the evolution of the network and the mathematical goal is to solve the inference problem of the said PA function. So far only the frequentist side of statistics inference has been investigated. Three solid projects are ideal for PhD students.

The first is to employ the Bayesian method to the estimation of affine preferential attachment networks. Given that the maximum likelihood estimation on the problem leads to great results, this project is to prove a Bernstein-von Mises Theorem type results, with suitable priors defined. The result will be of interest as the classic Bernstein-von Mises Theorem does not apply to models that are highly Markovian and where the state space cannot be formulated with Hilbert spaces.

The second project is to exploit the techniques of nonparametric Bayesian methods to estimate the PA function, where we only assume the sublinearity and monotonicity of the PA function. This will be a bold attempt and challenging. Ideally, we hope to obtain results akin to minimax rates on functions with certain smoothness. New techniques must be invented, and the posterior distribution can only be understood via the proxy of certain branching processes, if we manage to shed light on the problem with some new equivalences between the PA networks and urn models.

The third project deals with finding the root (oldest) node with the Bayesian paradigm. Proper priors will be proposed, and theoretical properties of the posterior contraction, as well as the credible set coverage will be investigated with also a focus on the computationally inexpensive sampling algorithms.

The projects will be conducted with close collaborations from Kolyan Ray (Imperial) and Aad van der Vaart (TU Delft).

References:

1. Crane, H. and Xu, M., 2023. Root and community inference on latent network growth processes using noisy attachment models. Journal of the Royal Statistical Society Series B: Statistical Methodology, p.qkad102.

2. Gao, F. and van der Vaart, A., 2017. On the asymptotic normality of estimating the affine preferential attachment network models with random initial degrees. Stochastic Processes and their Applications, 127(11), pp.3754-3775.

3. Gao, F. and van der Vaart, A., 2021. Statistical Inference in Parametric Preferential Attachment Trees. arXiv preprint arXiv:2111.00832.

4. Gao, F., van der Vaart, A., Castro, R. and van der Hofstad, R., 2017. Consistent estimation in general sublinear preferential attachment trees.