Correction to Statistical clustering of temporal networks through a dynamic stochastic block model

Catherine Matias and Vincent Miele

Centre National de la Recherche Scientifique, Sorbonne Université, Université Paris Diderot, Laboratoire de Probabilités, Statistique et Modélisation, Paris, France.
catherine.matias@math.cnrs.fr
Université de Lyon, Université Lyon 1, Centre National de la Recherche Scientifique, Laboratoire de Biométrie et Biologie Évolutive, Villeurbanne, France.
vincent.miele@univ-lyon1.fr

This note provides a complement to the proof of the result on parameter identifiability, namely Proposition 1 in the manuscript, in the binary case. We would like to mention that there exist new results on the identifiability of dynamic stochastic block models (Becker and Holzmann, 2018).

Complement to the proof of Proposition 1. We consider the binary setup. The proof from the manuscript establishes that from the distribution $\mathbb{P}_\theta^Y$, we can obtain the parameters $\{\beta_{ql}^t, 1 \leq t \leq T, 1 \leq q, l \leq Q\}$ up to a permutation $\sigma \in \mathfrak{S}_Q$. Note that with the same argument, we also obtain the stationary distribution α on the nodes groups up to the same permutation. Then it remains to identify the parameter π of the latent transition matrix. The argument given in the manuscript does not apply to the binary case as it relies on the identifiability of the mixtures of the densities ϕ_{ql}^t, which are Bernoulli in the binary case (and thus do not induce identifiable mixtures).

Let us consider the degree of the first node at time t, namely $D_1^t = \sum_{j=2}^N Y_{1j}^t$ whose conditional distribution, given $Z_1^t = q$ is Binomial $\mathcal{B}(N-1, \bar{\beta}_q^t)$ where $\bar{\beta}_q^t = \sum_{l=1}^Q \alpha_l \beta_{ql}^t$. Those parameters $\bar{\beta}_q^t$ are already identified up to a permutation $\sigma \in \mathfrak{S}_Q$. Now, we consider the distribution of the pair (D_1^t, D_1^{t+1}), which is a mixture distribution given by

$$\sum_{1 \leq q, q' \leq Q} \alpha_q \pi(q, q') \mathcal{B}(N-1, \bar{\beta}_q^t) \otimes \mathcal{B}(N-1, \bar{\beta}_{q'}^{t+1}).$$

Teicher (1967) has proved the equivalence between parameter identifiability for the mixtures of a family of distributions and parameter identifiability for the mixtures of finite products from this same family. Thus the above mixture has identifiable parameters as soon as the one dimensional mixtures $\sum_{1 \leq q \leq Q} \alpha_q \mathcal{B}(N-1, \bar{\beta}_q^t)$ and $\sum_{1 \leq q' \leq Q} [\sum_{q=1}^Q \alpha_q \pi(q, q')] \mathcal{B}(N-1, \bar{\beta}_{q'}^{t+1})$ are identifiable. Moreover, Blischke (1964) establishes that these mixtures of Binomial distributions are identifiable as soon as (in our setup) $N - 1 \geq 2Q - 1$ and the parameters $\{\bar{\beta}_q^t, 1 \leq q \leq Q\}$ are pairwise distinct, which is a generic condition. We thus conclude that we can identify the parameters $\{(\alpha_q\pi(q,q'),\bar\beta^t_q,\bar\beta^{t+1}_{q'});1\le q,q'\le Q\}$ up to a permutation on these Q^2 values. Nonetheless, the parameters $\alpha_q,\bar\beta^t_q,\bar\beta^{t+1}_{q'}$ are already identified up to a permutation in $\mathfrak{S}_Q$, so we recover the proportions $\{\alpha_q\pi(q,q');1\le q,q'\le Q\}$ up to a permutation in $\mathfrak{S}_Q$ and then also the latent transition matrix π up to a permutation in $\mathfrak{S}_Q$. □

Acknowledgments The authors would like to thank Ann-Kristin Becker for pointing that an argument was missing in the proof.

References

Becker, A.-K. and H. Holzmann (2018). Nonparametric identification in the dynamic stochastic block model. Technical report.

Blischke, W. (1964). Estimating the parameters of mixtures of binomial distributions. *J. Am. Stat. Assoc. 59*, 510–528.

Teicher, H. (1967). Identifiability of mixtures of product measures. *Ann. Math. Statist 38*, 1300–1302.