Guarantee for Ding's algorithm

Jiaxin Hu

April 9, 2022

In previous notes, we find that non-iterative clean-up can not relax the order of the number of seeds. We want to check whether this phenomenon also happens in Ding's work. This note shows that the non-iterative clean-up also does not benefit the seeded algorithm in terms of the order of seeds number.

Ding's Algorithm

For self-consistency, we recall the model and seeded matching without non-iterative clean-up in the Algorithm 3 of Ding et al. (2021). Let $\boldsymbol{A}, \boldsymbol{B} \in \{0,1\}^{n\times n}$ denote two correlated Erdős-Rényi graphs following $\mathcal{G}(n,q,\rho)$; i.e., with permutation π^* on $[n]$

$$\boldsymbol{B}_{ij} \sim \begin{cases} \text{Ber}(\rho) & \text{if } \boldsymbol{A}_{\pi^*(i)\pi^*(j)} = 1 \\ \text{Ber}\left(\frac{q(1-\rho)}{1-q}\right) & \text{if } \boldsymbol{A}_{\pi^*(i)\pi^*(j)} = 0 \end{cases}, \quad (\boldsymbol{A}_{\pi^*(i)\pi^*(j)}, \boldsymbol{B}_{ij}) \perp (\boldsymbol{A}_{\pi^*(i')\pi^*(j')}, \boldsymbol{B}_{i'j'}) \text{ for all } (i,j) \neq (i',j').$$

Suppose we have the seed set with corresponding perfect matching $\pi_0 : S \mapsto T$, where $S, T \subset [n]$ and $\pi_0(i) = \pi^*(i)$ for all $i \in S$. Let $s := |S|$ denote the number seeds corresponding to π_0. Ding et al. (2021) adapts following Algorithm 1 for seeded matching.

Algorithm 1 Seeded matching without non-iterative clean-up in Ding et al. (2021)

Input: Binary matrices $\boldsymbol{A}, \boldsymbol{B} \in \{0,1\}^{n\times n}$, the perfect matching $\pi_0 : S \mapsto T$.

1: For each $i \in S^c$ and $k \in T^c$, define $n_{ik} = \sum_{j\in S} \boldsymbol{A}_{ij}\boldsymbol{B}_{k\pi_0(j)}$.

2: Define a bipartite graph with vertex sets $S^c \times T^c$ and adjacency matrix H given by $H_{ik} = \mathbb{1}\{n_{ik} \geq \kappa\}$ for each $i \in S^c$ and $k \in T^c$, where $\kappa = sq\rho/2$.

3: Find a perfect matching $\tilde{\pi}_1 : S^c \mapsto T^c$ such that

$$\tilde{\pi}_1 = \underset{\pi: S^c \mapsto T^c \text{ is a perfect matching}}{\arg\max} \sum_{i\in S^c} H_{i\pi(i)}. \tag{1}$$

Let π_1 denote a permutation on $[n]$ such that $\pi_1|_S = \pi_0$ and $\pi_1|_{S^c} = \tilde{\pi}_1$.

Output: Estimated permutation π_1.

Guarantee for Algorithm 1 to exactly recover.

Lemma 1 (Guarantee for Algorithm 1). *Assume $sq\rho \gtrsim \log n$ and $\rho \geq 12q$. The output π_1 from Algorithm 1 is equal to true permutation π^* with probability tends to 1.*

Remark 1. Lemma 1 indicates that under the same seed condition, $sq\rho \gtrsim \log n$, Ding's seeded algorithm achieves exact recovery without non-iterative clean-up.

Proof of Lemma 1. Without loss of generality, assume that π^* is the identity mapping; i.e., $\pi^*(i) = i$ for all $i \in [n]$. To show the exact recovery of π_1, we need to show that $\tilde{\pi}_1$ in (1) is equal to $\pi^*|_{S^c}$; i.e., for all perfect matching $\pi : S^c \mapsto T^c$

$$\sum_{i \in S^c} \mathbb{1}\{n_{ii} \geq \kappa\} \geq \sum_{i \in S^c} \mathbb{1}\{n_{i\pi(i)} \geq \kappa\}.$$

Hence, it suffices to show that with probability tends to 1

$$\min_{i \in S^c} n_{ii} \geq \max_{i,k \in S^c, i \neq k} n_{ik}.$$

Note that $n_{ii} \sim \text{Binom}(s, q\rho)$ and $n_{ik} \sim \text{Binom}(s, q^2)$ for all $i, k \in S^c$. Recall $\kappa = sq\rho/2$. By Proposition 1 and union bound, we have

$$\mathbb{P}\left(\min_{i \in S^c} n_{ii} \leq \kappa\right) \leq n \exp\left(-\frac{1}{8} sq\rho\right),$$

and

$$\mathbb{P}\left(\max_{i,k \in S^c, i \neq k} n_{ik} \geq \kappa\right) \leq n^2 2^{-sq\rho/2} \leq n^2 \exp\left(-\frac{1}{4} sq\rho\right).$$

Therefore, by the assumption that $sq\rho \gtrsim \log n$, we have

$$\mathbb{P}\left(\min_{i \in S^c} n_{ii} \geq \max_{i,k \in S^c, i \neq k} n_{ik}\right) \geq 1 - n \exp\left(-\frac{1}{8} sq\rho\right) - n^2 \exp\left(-\frac{1}{4} sq\rho\right) \to 1,$$

as $n \to \infty$.

$\square$

Proposition 1 (Tail bound for Binomial variables)**.** *Let $X \sim$ Binom(n,p). Then,*

$$\mathbb{P}(X \leq (1-t)np) \leq \exp\left(-\frac{t^2}{2} np\right), \text{ for all } t \in [0, 1],$$

and

$$\mathbb{P}(X \geq R) \leq 2^{-R}, \quad \text{for all } R \geq 6np.$$

Proof. See Appendix A in Ding et al. (2021). $\square$

References

Ding, J., Ma, Z., Wu, Y., and Xu, J. (2021). Efficient random graph matching via degree profiles. *Probability Theory and Related Fields*, 179(1):29–115.