MATH888: High-dimensional probability and statistics Fall 2021

Lecture 42 — December 17, 2021

Sebastien Roch, UW-Madison *Scribe: Shuyao Li, Sebastien Roch*

1 Overview

In a previous lecture, we stated the following theorem:

Theorem 1. *In SSBM*$(n, q_{in} = a\frac{\log n}{n}, q_{out} = b\frac{\log n}{n})$*, where* $a > b$*, exact recovery*

- *is possible if* $(\sqrt{a} - \sqrt{b})^2 > 2$
- *is not possible if* $(\sqrt{a} - \sqrt{b})^2 < 2$

In this lecture, we sketch the proof of the first half of the theorem. The lecture is based on [Abb18].

Remark. This result is somewhat different from what we have done this semester. Here we derive a tight result by a tailored argument for a specific. One must also be more careful with some of the constants appearing in the proof.

2 A Key Lemma

Lemma 2 (Lemma 8 in [AFWZ20])**.** *In the setting above,*

$$\mathbb{P}\left\{\mathrm{Bin}(\frac{n}{2}, q_{in}) - \mathrm{Bin}(\frac{n}{2}, q_{out}) \leq \epsilon \log n\right\} \leq n^{-\frac{(\sqrt{a}-\sqrt{b})^2}{2} + \epsilon \log \sqrt{\frac{a}{b}}},$$

where the binomial random variables are independent.

Remark. Note the absence of constants in the exponent on the right-hand side. In fact, the inequality can be strengthened to equality $\mathbb{P}\{\cdots\} = n^{-(\cdots)+o(1)}$, which plays a role in proving the second half of Theorem 1 (impossibility result).

Sketch of Proof. Apply Markov inequality to $\exp\left(\lambda \cdot \left(\mathrm{Bin}(\frac{n}{2}, q_{in}) - \mathrm{Bin}(\frac{n}{2}, q_{out})\right)\right)$ and take $\lambda = -\log\sqrt{\frac{a}{b}}$. Use the explicit moment-generating function for the binomial. □

3 A Technical Proposition

In this section we state a key technical proposition.

Let A' be the adjacent matrix with self-loops (random zeros/ones on diagonal following $Ber(q_{in})$). Let $A := A' - n\frac{q_{in}+q_{out}}{2}\mathbf{1}_n\mathbf{1}_n^T$ and $\bar{A} := \mathbb{E}A = n\frac{q_{in}-q_{out}}{2}\bar{\boldsymbol{\phi}}\bar{\boldsymbol{\phi}}^T$, where $\bar{\boldsymbol{\phi}} = \frac{1}{\sqrt{n}} = \begin{bmatrix} \mathbf{1}_{\frac{n}{2}} \\ -\mathbf{1}_{\frac{n}{2}} \end{bmatrix}$.

We already proved that $\|\boldsymbol{\phi} - \bar{\boldsymbol{\phi}}\| \lesssim \frac{1}{\sqrt{\log n}}$ on the event $\mathcal{E} := \{\|A - \bar{A}\|_2 \lesssim \sqrt{\log n}\}$ with $\mathbb{P}\mathcal{E} \geq 1 - n^{-8}$. For exact recovery, a first attempt would be to show that $\boldsymbol{\phi}$ and $\bar{\boldsymbol{\phi}}$ are close enough in each component, i.e., that with high probability,

$$\left|\boldsymbol{\phi}_i - \bar{\boldsymbol{\phi}}_i\right| \overset{?}{<} \left|\bar{\boldsymbol{\phi}}_i\right|, \quad \forall i \in [n] \tag{1}$$

$$\Longleftrightarrow \|\boldsymbol{\phi} - \bar{\boldsymbol{\phi}}\|_\infty \overset{?}{<} 1/\sqrt{n} \tag{2}$$

This would imply that rounding the components of $\boldsymbol{\phi}$ to their signs would produce with high probability the same signs as $\bar{\boldsymbol{\phi}}$, which solves exact recovery.

Unfortunately the above inequality does not hold all the way down to the exact recovery threshold, which makes the problem more difficult. However, note that it is not necessary to have (1) in order to obtain the correct sign by rounding $\boldsymbol{\phi}$: one can have a large gap for $\left|\boldsymbol{\phi}_i - \bar{\boldsymbol{\phi}}_i\right|$ which still produces the good sign as long as this gap in "on the right side", i.e., $\boldsymbol{\phi}_i$ can be much larger than $\bar{\boldsymbol{\phi}}_i$ if $\bar{\boldsymbol{\phi}}_i$ is positive and $\boldsymbol{\phi}_i$ can be much smaller than $\bar{\boldsymbol{\phi}}_i$ if $\bar{\boldsymbol{\phi}}_i$ is negative.

Proposition 3 ([Abb18, Theorem 4.10])**.** *For large enough n,*

$$\mathbb{P}\left(\min_{s=\pm 1} \|s\boldsymbol{\phi} - A\bar{\boldsymbol{\phi}}/\bar{\lambda}\|_\infty \leq \frac{c}{\sqrt{n}\log\log n}\right) \geq 1 - Cn^{-2}$$

Remark. A notable difference from Davis-Kahan-based approach is the $A\bar{\boldsymbol{\phi}}/\bar{\lambda}$ on the LHS instead of $\bar{\boldsymbol{\phi}}$. Recall that $\boldsymbol{\phi}$ is the leading eigenvector of A. So, by the power method, $A\bar{\boldsymbol{\phi}}$ (normalized) is closer to $\boldsymbol{\phi}$ while having properties similar to $\bar{\boldsymbol{\phi}}$ in analysis (specifically its signs). To analyze it, we will see later that entries of $A\bar{\boldsymbol{\phi}}$ behave like $\text{Bin}(\frac{n}{2}, q_{in}) - \text{Bin}(\frac{n}{2}, q_{out})$.

We will not prove this theorem here. We mention one key idea that comes up in the analysis (the leave-one-out trick). First, we use triangle inequality to split into two terms

$$\|\boldsymbol{\phi} - A\bar{\boldsymbol{\phi}}/\bar{\lambda}\|_\infty \leq \|\boldsymbol{\phi} - A\boldsymbol{\phi}/\bar{\lambda}\|_\infty + \|A(\boldsymbol{\phi} - \bar{\boldsymbol{\phi}})/\bar{\lambda}\|_\infty = \|\boldsymbol{\phi} - \lambda\boldsymbol{\phi}/\bar{\lambda}\|_\infty + \|A(\boldsymbol{\phi} - \bar{\boldsymbol{\phi}})/\bar{\lambda}\|_\infty$$

The first term is easy to bound with Weyl's inequality, while the second term is trickier. One difficulty in estimating $\|A(\boldsymbol{\phi} - \bar{\boldsymbol{\phi}})\|_\infty$ is that A and $(\boldsymbol{\phi} - \bar{\boldsymbol{\phi}})$ are dependent since $\boldsymbol{\phi}$ is an eigenvector of A. Thus, to bound the m-th component of $A(\boldsymbol{\phi} - \bar{\boldsymbol{\phi}})$, namely $A_m(\boldsymbol{\phi} - \bar{\boldsymbol{\phi}})$, where A_m is the m-row of A, we cannot use directly a concentration result that applies to expressions of the kind $A_m w$ where w is an independent test vector. To decouple the dependencies, we use a leave-one-out technique.

Define n auxiliary matrices $\left\{A^{(m)}\right\}_{m=1}^n \subseteq \mathbb{R}^{n \times n}$ as follows: for any $m \in [n]$, let

$$\left(A^{(m)}\right)_{ij} = A_{ij}\delta_{\{i \neq m, j \neq m\}}, \quad \forall i, j \in [n]$$

where δ_A is the indicator function on the event A. Therefore, $A^{(m)}$ is obtained from A by zeroing out the m-th row and column. Let $\boldsymbol{\phi}^{(m)}$ be the leading eigenvector of $A^{(m)}$. We can write

$$(A(\boldsymbol{\phi} - \bar{\boldsymbol{\phi}}))_m = A_m(\boldsymbol{\phi} - \bar{\boldsymbol{\phi}}) = A_m\left(\boldsymbol{\phi} - \boldsymbol{\phi}^{(m)}\right) + A_m\left(\boldsymbol{\phi}^{(m)} - \bar{\boldsymbol{\phi}}\right)$$

and thus

$$\begin{aligned}\left|(A(\boldsymbol{\phi}-\bar{\boldsymbol{\phi}}))_m\right| &\le \left|A_m\left(\boldsymbol{\phi}-\boldsymbol{\phi}^{(m)}\right)\right| + \left|A_m\left(\boldsymbol{\phi}^{(m)}-\bar{\boldsymbol{\phi}}\right)\right| \\ &\le \|A_m\|_2\left\|\boldsymbol{\phi}-\boldsymbol{\phi}^{(m)}\right\|_2 + \left|A_m\left(\boldsymbol{\phi}^{(m)}-\bar{\boldsymbol{\phi}}\right)\right|\end{aligned}$$

The first term is very small since A_m and A are very close and so do $\boldsymbol{\phi}^{(m)}$ and $\boldsymbol{\phi}$. Note that the second term is better-behaved now since A_m and $\boldsymbol{\phi}^{(m)}-\bar{\boldsymbol{\phi}}$ are independent. That second term can be analyzed using a tailored application of Markov's inequality (see Lemma 7 in [AFWZ20]).

4 Putting everything together

Now we assume

$$(\sqrt{a}-\sqrt{b})^2 > 2 \tag{3}$$

$$\|A-\bar{A}\|_2 \lesssim \sqrt{\log n} \qquad \text{i.e. on event } \mathcal{E} \tag{4}$$

$$\min_{s=\pm 1}\|s\boldsymbol{\phi} - A\bar{\boldsymbol{\phi}}/\bar{\lambda}\|_\infty \le \frac{c}{\sqrt{n}\log\log n} \tag{5}$$

We sketch the proof of Theorem 1 given Lemma 2 and Proposition 3.

Sketch of Proof of Theorem 1. Assume conditions (3), (4) and (5) hold. Then it suffices to prove that $F_+ \cap F_-$ happens with probability $1-o(1)$, where

$$\begin{aligned}F_+ &= \{\frac{1}{\bar{\lambda}}(A\bar{\boldsymbol{\phi}})_i \ge \frac{2\epsilon}{(a-b)\sqrt{n}} \qquad \text{for all } i = 1\ldots n/2\} \\ F_- &= \{\frac{1}{\bar{\lambda}}(A\bar{\boldsymbol{\phi}})_i \le -\frac{2\epsilon}{(a-b)\sqrt{n}} \qquad \text{for all } i = n/2+1\ldots n\}.\end{aligned}$$

Now we need to show that

$$\mathbb{P}\left(\left|(A\bar{\boldsymbol{\phi}}/\bar{\lambda})_1\right| \ge \frac{2\varepsilon}{(a-b)\sqrt{n}}\right) \le \frac{1}{n^{1+\Omega(1)}},$$

and that is exactly what Lemma 2 implies, since

$$(A\bar{\boldsymbol{\phi}})_i \overset{d}{=} (A'\bar{\boldsymbol{\phi}})_i \overset{d}{=} \frac{1}{\sqrt{n}}\left(\mathrm{Bin}(\frac{n}{2}, q_{in}) - \mathrm{Bin}(\frac{n}{2}, q_{out})\right)$$

and

$$\bar{\lambda} = \frac{n(q_{in}-q_{out})}{2} = \frac{a-b}{2}\log n.$$

Finally, it remains to choose ϵ small enough that

$$n^{-\frac{(\sqrt{a}-\sqrt{b})^2}{2}+\epsilon\log\sqrt{\frac{a}{b}}} = n^{-(1+\Omega(1))},$$

which is equivalent to make

$$-\frac{(\sqrt{a}-\sqrt{b})^2}{2} + \epsilon\log\sqrt{\frac{a}{b}} < -1,$$

which is possible because inequality (3) is strict. □

References

[Abb18] Emmanuel Abbe. Community Detection and Stochastic Block Models. *Foundations and Trends® in Communications and Information Theory*, 14(1-2):1–162, June 2018. Publisher: Now Publishers, Inc.

[AFWZ20] Emmanuel Abbe, Jianqing Fan, Kaizheng Wang, and Yiqiao Zhong. Entrywise eigenvector analysis of random matrices with low expected rank. *The Annals of Statistics*, 48(3):1452–1474, June 2020. Publisher: Institute of Mathematical Statistics.