MLE phase transition of Gaussian tensor matching (Negative part of non-symmetric observations)

Jiaxin Hu

May 19, 2022

In previous note 0517, we provide the threshold of ρ^2 for MLE to achieve exact recovery. In this note, we provide the threshold of ρ^2 when it is impossible for MLE to achieve exact recovery. We still consider the non-symmetric case since we still need the relationship between node permutation error and edge permutation error.

The theorem for negative part remains to be a conjecture now. Two places need to be fixed to make it a concrete theorem: (1) choosing a precise constant C_0 that occurs in both positive and negative thresholds; (2) Lemma 3 following Ganassali (2020) should be rigorously extended for tensor case.

1 Preliminary

Non-symmetric correlated Gaussian observations.

Consider two order-m random tensor observations $\mathcal{A}, \mathcal{B}' \in \mathbb{R}^{n^{\otimes m}}$ and use $\boldsymbol{\omega} \in [n]^m$ to index the entries in $\mathcal{A}$ and $\mathcal{B}$. Suppose that for all $\boldsymbol{\omega} \in [n]^m$ and some $\rho \in (0,1)$

$$\begin{pmatrix} \mathcal{A}_{\boldsymbol{\omega}} \\ \mathcal{B}'_{\boldsymbol{\omega}} \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right), \quad \text{and} \quad \begin{pmatrix} \mathcal{A}_{\boldsymbol{\omega}} \\ \mathcal{B}'_{\boldsymbol{\omega}} \end{pmatrix} \text{ is independent with } \begin{pmatrix} \mathcal{A}_{\boldsymbol{\omega}'} \\ \mathcal{B}'_{\boldsymbol{\omega}'} \end{pmatrix} \text{ for all } \boldsymbol{\omega} \neq \boldsymbol{\omega}'. \tag{1}$$

Let π^* be a permutation on $[n]$ with corresponding permutation matrix $\Pi^* \in \{0,1\}^{n \times n}$, and consider the permuted observation $\mathcal{B}$ such that for all $\boldsymbol{\omega} \in [n]^m$

$$\mathcal{B}_{\boldsymbol{\omega}} = \mathcal{B}'_{\pi^* \circ \boldsymbol{\omega}}, \quad \text{or equivalently} \quad \mathcal{B} = \mathcal{B}' \times_1 \Pi^* \times_2 \cdots \times_m \Pi^*.$$

Our goal is to recover π^* (or equivalently Π^*) observing $\mathcal{A}, \mathcal{B}$. **Note that $\mathcal{A}, \mathcal{B}$ are not super-symmetric tensors while the permutation on every mode is the same!**

MLE

By Theorem 1 in note 0402, the MLE of π^*, denoted $\hat{\pi}_{MLE}$, satisfies

$$\hat{\Pi}_{MLE} = \underset{\Pi \in \mathcal{P}_n}{\arg\max} \left\langle \mathcal{A} \times_1 \Pi \times_2 \cdots \times_m \Pi, \mathcal{B} \right\rangle,$$

where $\hat{\Pi}_{MLE}$ is the permutation matrix corresponding to $\hat{\pi}_{MLE}$, and $\mathcal{P}_n$ is the collection for all possible permutation matrices on $[n]$.

2 Theorem

Theorem 1 (Converse part of MLE phase transition with non-symmetric observation, Conjecture)**.** *Consider the observations* $(\mathcal{A}, \mathcal{B})$ *from model* (1) *with true permutation* π^*. *Assume* n *is large enough and*

$$\rho^2 \leq \frac{(C_0 - \varepsilon) \log n}{n^{m-1}},$$

for some positive constant $C_0 > 0$ *and small constant* ε. *Then, the MLE* $\hat{\pi}_{MLE}$ *exactly recovers true permutation* π^* *with probability* $o(1)$.

Remark 1 (Conjecture)**.** The Theorem 1 is a conjecture since (1) the constant C_0 is critical to the proof and should be the same constant C_0 in the positive part in note 0517 and (2) the Lemma 3 should be well-extended for tensor case.

Proof of Theorem 1. Without the loss of generality, assume the true permutation π^* is the identity mapping. With observations $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{n^{\otimes m}}$, consider the loss function

$$\mathcal{L}(\pi, \mathcal{A}, \mathcal{B}) = \langle \mathcal{A} \times_1 \Pi \times_2 \cdots \times_m \Pi, \mathcal{B} \rangle,$$

where $\Pi \in \{0,1\}^{n \times n}$ is the permutation matrix corresponding to π. We define the difference

$$\begin{aligned}\Delta(\pi) :=& \mathcal{L}(\pi, \mathcal{A}, \mathcal{B}) - \mathcal{L}(\pi^*, \mathcal{A}, \mathcal{B}) \\ =& \rho \sum_{\boldsymbol{\omega} \in [n]^m} (\mathcal{A}_{\pi \circ \boldsymbol{\omega}} - \mathcal{A}_{\boldsymbol{\omega}}) \mathcal{A}_{\boldsymbol{\omega}} + \sqrt{1-\rho^2} \sum_{\boldsymbol{\omega} \in [n]^m} (\mathcal{A}_{\pi \circ \boldsymbol{\omega}} - \mathcal{A}_{\boldsymbol{\omega}}) \mathcal{Z}_{\boldsymbol{\omega}},\end{aligned}$$

where the second equality follows from the fact that $\mathcal{B} = \rho \mathcal{A} + \sqrt{1-\rho^2} \mathcal{Z}$, where $\mathcal{Z}_{\boldsymbol{\omega}} \sim N(0,1)$ for all $\boldsymbol{\omega} \in [n]^m$ independently and $\mathcal{Z}$ is independent with $\mathcal{A}$. By the derivation in note 0517, we have

$$\mathbb{P}(\Delta(\pi) \geq 0) = \mathbb{E}[\mathbb{E}[\mathbb{1}\{\Delta(\pi) \geq 0\} | \mathcal{A}]] = \mathbb{E}\left[\mathbb{P}\left(N(0,1) \geq \frac{\rho \|\pi \circ \mathcal{A} - \mathcal{A}\|_F}{2\sqrt{1-\rho^2}} \middle| \mathcal{A}\right)\right] \tag{2}$$

To show the failure of MLE, we need to show

$$\begin{aligned}\mathbb{P}(\hat{\pi}_{MLE} = \pi^*) \leq o(1) \quad &\Leftrightarrow \quad \mathbb{P}(\text{for all } \pi \neq \pi^*, \Delta(\pi) < 0) \leq o(1) \\ &\Leftrightarrow \quad \mathbb{P}(|\{\pi \in \mathcal{P}_n : \pi \neq \pi^*, \Delta(\pi) \geq 0\}| \geq 1) \geq 1 - o(1),\end{aligned}$$

where $\mathcal{P}_n$ is the collection for all possible permutation matrices on $[n]$.

Let $D_\pi = \{i \in [n] : \pi(i) \neq i\}$ denote the set of unfixed points of π and $D_\pi^{m,E} = \{\boldsymbol{\omega} \in [n]^m : \pi \circ \boldsymbol{\omega} \neq \boldsymbol{\omega}\}$ denote the set of unfixed order-m edges of π. Define the variable

$$X = \sum_{\pi \in \mathcal{P}_n : |D_\pi| = 2} \mathbb{1}\{\Delta(\pi) \geq 0\},$$

which counts the number of permutations that have node permutation error 2 and have larger likelihoods than true permutations. Hence, for the failure of MLE, it suffices to show that

$$\mathbb{P}(X \geq 1) \geq 1 - o(1).$$

For simplicity, let $d = |D_\pi^{m,E}| = n^m - (n-2)^m$ denote the number of unfixed edged for all π such that $|D_\pi| = 2$. Consider the event

$$E(\mathcal{A}) := \{\text{for all } \pi \neq \pi^*, \quad 2d^E(1-\epsilon_n) \leq \|\pi \circ \mathcal{A} - \mathcal{A}\|_F^2 \leq 2d^E(1+\epsilon_n)\},$$

where $\epsilon_n = \frac{C}{2}\sqrt{\log n / n^{m-1}}$. By Proposition 2, we have $\mathbb{P}(E^c(\mathcal{A})) = o(1)$. Let $\tilde{X} = X\mathbb{1}\{E(\mathcal{A})\}$ and notice that $X \geq \tilde{X}$. Then, it suffices to show

$$\mathbb{P}(\tilde{X} \geq 1) \geq 1 - o(1). \tag{3}$$

By the Paley-Zygmund inequality Lemma 1 and conclusions for the first and second moments of $\tilde{X}$ in Lemmas 2 and 3, we have

$$\mathbb{P}(\tilde{X} \geq 1) = \mathbb{P}(\tilde{X} \geq n^{-\epsilon_0}\mathbb{E}[\tilde{X}]) \geq (1 - n^{-\epsilon_0})(1 - o(1)) = (1 - o(1)),$$

where ϵ_0 is some small constant smaller than ε.

□

Lemma 1 (Paley-Zygmund inequality)**.** *Let Z be a positive random variable with finite variance. Then, for all $c \in [0,1]$, we have*

$$\mathbb{P}(Z \geq c\mathbb{E}[Z]) \geq (1-c^2)\frac{(\mathbb{E}[Z])^2}{\mathbb{E}[Z^2]}.$$

Lemma 2 (First moment of $\tilde{X}$)**.** *Consider the variable $\tilde{X}$ defined in (3). We have*

$$\mathbb{E}[\tilde{X}] \geq (1 - o(1))\frac{n(n-1)\sqrt{1-\rho}}{\sqrt{2\pi}\rho\sqrt{d^E(1+\epsilon_n)}} \exp\left(-\frac{\rho^2 d^E(1+\epsilon_n)}{4(1-\rho^2)}\right)$$

Suppose $\rho^2 \leq \frac{(C_0-\varepsilon)\log n}{n^{m-1}}$, for some positive constants $C_0 > 0$ and ε. Then, we have

$$\mathbb{E}[\tilde{X}] \geq Cn^{\varepsilon_0},$$

for some positive constant C and $\varepsilon_0 < \varepsilon$.

Proof of Lemma 2. Note that the number of permutations π such that $|D_\pi| = 2$ is equal to $\binom{n}{2} = \frac{n(n-1)}{2}$. Let π be an arbitrary permutation satisfying $|D_\pi| = 2$. Then, we have

$$\begin{aligned}
\mathbb{E}[\tilde{X}] &= \frac{n(n-1)}{2}\mathbb{P}(\Delta(\pi) \geq 0, E(\mathcal{A})) \\
&= \frac{n(n-1)}{2}\mathbb{E}\left[\mathbb{P}\left(N(0,1) \geq \frac{\rho\|\pi \circ \mathcal{A} - \mathcal{A}\|_F}{2\sqrt{1-\rho^2}}\middle|\mathcal{A}\right)\mathbb{1}\{E(\mathcal{A})\}\right] \\
&\geq \frac{n(n-1)}{2}\mathbb{E}\left[(1-o(1))\frac{2\sqrt{1-\rho}}{\sqrt{2\pi}\rho\|\pi \circ \mathcal{A} - \mathcal{A}\|_F}\exp\left(-\frac{\rho^2\|\pi \circ \mathcal{A} - \mathcal{A}\|_F^2}{8(1-\rho^2)}\right)\mathbb{1}\{E(\mathcal{A})\}\right] \\
&\geq (1-o(1))\frac{n(n-1)\sqrt{1-\rho}}{\sqrt{2\pi}\rho\sqrt{d^E(1+\epsilon_n)}}\exp\left(-\frac{\rho^2 d^E(1+\epsilon_n)}{4(1-\rho^2)}\right),
\end{aligned}$$

where the second equation follows the derivation in (2), and the last two inequalities follow from the fact that $\mathbb{P}(N(0,1) \geq t) \geq \frac{t}{\sqrt{2\pi}(t^2+1)} \exp(-t^2/2)$, the event $E(\mathcal{A})$ such that $\|\pi \circ \mathcal{A} - \mathcal{A}\|_F^2 \leq 2d(1+\epsilon_n)$, and $d^E \geq 2n^{m-1}$ by Proposition 1.

Suppose $\rho^2 \leq \frac{(C_0-\varepsilon)\log n}{n^{m-1}}$. With $d^E \geq 2n^{m-1}$, we further have

$$\mathbb{E}[\tilde{X}] \geq C' \frac{n^2}{\sqrt{\log n}} \exp\left(-\frac{d^E(C_0-\varepsilon)\log n}{4n^{m-1}}\right) \geq Cn^{\epsilon_0},$$

where C' is some positive constant and the second inequality holds by choosing C_0 such that $n^2 \exp\left(-\frac{d^E C_0 \log n}{4n^{m-1}}\right) = c$ for some constant c. □

Lemma 3 (Second moment of $\tilde{X}$). *Consider the variable $\tilde{X}$ defined in (3). Then, we have*

$$\frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]} \geq 1 - o(1).$$

Proof of Lemma 3. Follow the proof of Lemma 3.2 in Ganassali (2020). □

Proposition 1 (Relationship between unfixed points and unfixed edges). *Suppose we have a permutation π on $[n]$. Let $D_\pi = \{i \in [n] : \pi(i) \neq i\}$ denote the set of unfixed points of π and $D_\pi^{m,E} = \{\boldsymbol{\omega} \in [n]^m : \pi \circ \boldsymbol{\omega} \neq \boldsymbol{\omega}\}$ denote the set of unfixed order-m edges. Then, we have*

$$n^{m-1}|D_\pi| \leq |D_\pi^{m,E}| \leq mn^{m-1}|D_\pi|.$$

Proposition 2 (Edge disagreement with permutation π). *Suppose we have an order-m observation $\mathcal{A} \in \mathbb{R}^{n^{\otimes m}}$ with i.i.d. standard Gaussian entries. Let $D_\pi^{m,E} = \{\boldsymbol{\omega} \in [n]^m : \pi \circ \boldsymbol{\omega} \neq \boldsymbol{\omega}\}$ denote the set of unfixed order-m edges. We have the expectation*

$$\mathbb{E}\left[\|\pi \circ \mathcal{A} - \mathcal{A}\|_F^2\right] = 2|D_\pi^{m,E}|,$$

and there exists a positive constant C such that

$$\left|\|\pi \circ \mathcal{A} - \mathcal{A}\|_F^2 - 2|D_\pi^{m,E}|\right| \leq C|D_\pi^{m,E}|\sqrt{\frac{\log n}{n^{m-1}}},$$

with high probability.

References

Ganassali, L. (2020). Sharp threshold for alignment of graph databases with gaussian weights. *arXiv preprint arXiv:2010.16295.*