Tensor Matching

Jiaxin Hu

December 27, 2021

Problem Setup Let $\mathcal{A}, \mathcal{B}' \in (\mathbb{R}^d)^{\otimes m}$ denote two random Gaussian tensors, and $\mathcal{A}(\omega), \mathcal{B}'(\omega) \in \mathbb{R}$ denote tensor entry indexed by $\omega \in [d]^m$. Consider the bivariate model

$$(\mathcal{A}(\omega), \mathcal{B}'(\omega)) \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right), \quad \text{and} \quad (\mathcal{A}(\omega), \mathcal{B}'(\omega)) \perp (\mathcal{A}(\omega'), \mathcal{B}'(\omega')), \text{ for all } \omega \neq \omega',$$

where the correlation $\rho \in (0, 1]$ and $\perp$ denote the statistical independence. Suppose we observe the tensor pair $\mathcal{A}$ and $\mathcal{B} \stackrel{\text{def}}{=} \mathcal{B}' \circ \pi$, where $\pi : [d] \mapsto [d]$ denotes a permutation on $[d]$, and by definition $\mathcal{B}(i_1, \ldots, i_m) = \mathcal{B}'(\pi(i_1), \ldots, \pi(i_m))$ for all $(i_1, \ldots, i_m) \in [d]^m$.

Here are two questions:

1. How to provably recover *pi* from the input tensors $\mathcal{A}$ and $\mathcal{B}$? (this note)
2. What if we change the bivariate *normal* to bivariate *Bernoulli* distribution?

Intuition The intuition of tensor matching follows the idea in Section 2 Ding et al. (2021). Without loss of generality, let π be the identity permutation and $m = 3$.

Let δ_x denote the Dirac measure at x. Consider two statistics

$$\mu_i = \frac{1}{d^2} \sum_{(j,l) \in [d]^2} \delta_{\mathcal{A}_{ijl}}, \quad \text{and} \quad \nu_k = \frac{1}{d^2} \sum_{(j,l) \in [d]^2} \delta_{\mathcal{B}_{kjl}}.$$

Note that the statistics only include the summation of point masses. Then, the permutation for the index (j, l) does not change the value of the statistics, and the correlation between these two statistics only depends on whether (i, k) is a true pair or a fake pair.

Specifically, (1) If (i, k) is a true pair, $(\mathcal{A}_{ijl}, \mathcal{B}_{kjl}) \sim \mathcal{N}(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix})$ independently for all $(j, l) \in [d]^2$, and thus μ_i, ν_k are correlated. (2) If (i, k) is a fake pair, then $\mathcal{A}_{ijl}$ and $\mathcal{B}_{kjl}$ are independent with each other. Thus μ_i and ν_k are uncorrelated.

Therefore, we can apply the test in Ding et al. (2021) to check correlation between μ_i and ν_k, and thereof to recover the true permutation.

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.