MLE for Gaussian Matching

Jiaxin Hu

March 30, 2022

1 MLE for Gaussian Tensor matching

Recall that we have two random tensors $\mathcal{A}, \mathcal{B}' \in \mathbb{R}^{n^{\otimes m}}$, where $\mathcal{A}(\omega)$ and $\mathcal{B}'(\omega)$ denote the tensor entry indexed by $\omega = (i_1, \ldots, i_m) \in [n]^m$. Suppose $\mathcal{A}$ and $\mathcal{B}'$ are super-symmetric; i.e., $\mathcal{A}(\omega) = \mathcal{A}(f(\omega)), \mathcal{B}(\omega) = \mathcal{B}'(f(\omega))$ for any function f permutes the indices in ω for all $\omega \in [n]^m$. Consider the bivariate generative model for the entries $\{\omega : 1 \leq i_1 \leq \cdots \leq i_m \leq n\}$

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

where the correlation $\rho \in (0,1)$ and $\perp$ denote the statistical independence.

Suppose we observe the tensor pair $\mathcal{A}$ and $\mathcal{B} \overset{\text{def}}{=} \mathcal{B}' \circ \pi^*$, where $\pi^* : [n] \mapsto [n]$ denotes a permutation on $[d]$, and $\mathcal{B}(i_1, \ldots, i_m) = \mathcal{B}'(\pi(i_1), \ldots, \pi(i_m))$ for all $(i_1, \ldots, i_m) \in [n]^m$. Equivalently, let $\Pi^* \in \{0,1\}^{n \times n}$ denote the permutation matrix, where

$$\Pi^*_{ij} = \begin{cases} 1 & \pi^*(i) = j \\ 0 & \text{otherwise} \end{cases}.$$

We have

$$\mathcal{B}' = \mathcal{B} \times_1 \Pi^* \times_2 \cdots \times_m \Pi^*.$$

Let $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ and $\Sigma^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$. The likelihood function of $\pi : [n] \mapsto [n]$ (or corresponding $\Pi \in \{0,1\}^{n \times n}$) with given observations $\mathcal{A}, \mathcal{B}$ is

$$\begin{aligned} \mathcal{L}(\pi|\mathcal{A}, \mathcal{B}) &= \frac{1}{[\det(2\pi\Sigma)]^{n^m}} \exp\left(-\frac{1}{2} \sum_{\omega \in [n]^m} (\mathcal{A}(\omega), \mathcal{B}(\pi \circ \omega))\Sigma^{-1}(\mathcal{A}(\omega), \mathcal{B}(\pi \circ \omega))^T\right) \\ &= \frac{1}{[\det(2\pi\Sigma)]^{n^m}} \exp\left(-\frac{1}{2(1-\rho^2)} \sum_{\omega \in [n]^m} \mathcal{A}(\omega)^2 + \mathcal{B}(\pi \circ \omega)^2 - 2\rho\mathcal{A}(\omega)\mathcal{B}(\pi \circ \omega)\right). \end{aligned}$$

Note that $\sum_{\omega \in [n]^m} \mathcal{B}(\pi \circ \omega)^2 = \sum_{\omega \in [n]^m} \mathcal{B}(\omega)^2$. Then, the only term dependent to π in $\mathcal{L}$ is $\mathcal{A}(\omega)\mathcal{B}(\pi \circ \omega)$. Thus, with fixed $\rho \in (0,1)$, we have the MLE that satisfies

$$\hat{\pi} = \underset{\pi:[n]\mapsto[n]}{\arg\max}\, \mathcal{L}(\pi|\mathcal{A}, \mathcal{B}) = \underset{\pi:[n]\mapsto[n]}{\arg\max} \sum_{\omega \in [n]^m} \rho\mathcal{A}(\omega)\mathcal{B}(\pi \circ \omega) = \underset{\Pi \in \{0,1\}^{n \times n}}{\arg\max} \langle \mathcal{A}, \mathcal{B} \times_1 \Pi \times_2 \cdots \times_m \Pi \rangle.$$

Solving $\hat{\pi}$ can be viewed as a special case of *multiway assignment problem* which is a higher-order generalization of *quadratic assignment problem*.

References