Different idea for hypergraph matching

Jiaxin Hu

February 4, 2022

This note aims to compare the current ideas of tensor matching from me and Chanwoo.

	Jiaxin, Feb 3	Chanwoo, Jan 3	Jiaxin, Feb 3, after meeting
Setup	Two correlated asymmetric graphs $A, B \in \{0,1\}^{n\times m}$ with two row and column latent permutations $\pi_1^* : [n] \mapsto [n]$ and $\pi_2^* : [m] \mapsto [m]$. **Want to find two permutations** π_1^*, π_2^*.	Two correlated order-3 symmetric hypergraphs $\mathcal{A}, \mathcal{B} \in \{0,1\}^{n\otimes 3}$ with a signle permutation $\pi^* : [n] \mapsto [n]$. **Want to find the permutation** π^*.	Two correlated order-3 symmetric hypergraphs $\mathcal{A}, \mathcal{B} \in \{0,1\}^{n\otimes 3}$ with a signle permutation $\pi^* : [n] \mapsto [n]$. **Want to find the permutation** π^*.
Derivation of distance statistics Z	Derive the row and column distance statistics respectively. Take row statistics as an example. 1. Define connected sets $N_A(i) = \{j \in [m] : A_{ij} = 1\},$ $N_B(k) = \{j \in [m] : B_{kj} = 1\},$ with $a_i = \|N_A(i)\|, b_k = \|N_B(k)\|$. 2. Define "degree" of vertex $j \in [m]$ $a_j^{(i)} = \frac{1}{\sqrt{(n-1)q(1-q)}} \sum_{l\neq i}(A_{lj} - q),$ $b_j^{(k)} = \frac{1}{\sqrt{(n-1)q(1-q)}} \sum_{l\neq j}(B_{lj} - q).$ 3. Obtain empirical distributions $\bar{\mu}_i = \frac{1}{a_i} \sum_{j\in N_A(i)} \delta_{a_j^{(i)}} - \overline{B}(n-1, q),$ $\bar{\nu}_k = \frac{1}{b_k} \sum_{j\in N_B(i)} \delta_{b_k^{(i)}} - \overline{B}(n-1, q).$ 4. Obtain row distance Z_{ik} with $\bar{\mu}_i, \bar{\nu}_k$ and tuning parameter L as (Ding et al., 2021). 5. Repeat previous steps for column distance.	1. Define open neighbourhood $\mathcal{N}_{\mathcal{A}}(i) = \{(i_2, i_3) \in [n]^2, \mathcal{A}_{i,i_2,i_3} = 1\},$ and closed neighbourhood $\mathcal{N}_{\mathcal{A}}[i] = \mathcal{N}_{\mathcal{A}}(i) \cup \{(i, i_3) : i_3 \in [n]\},$ with $a_i = \|N_A(i)\|$. Similar for $\mathcal{N}_{\mathcal{B}}(k), \mathcal{N}_{\mathcal{B}}[k]$ and b_k. 2. Define "outdegree" of vertex $j \in [n]$ $a_j^{(i)} = \frac{1}{\sqrt{n^2 - a_i - nq(1-q)}} \sum_{\omega\notin\mathcal{N}_{\mathcal{A}}[i]} (\mathcal{A}_{\omega,j} - q).$ Similar for $b_j^{(k)}$. 3. Define dual neighbourhood $\mathcal{N}^*_{\mathcal{A}}(i) = \{j : \exists\omega \in \mathcal{N}_{\mathcal{A}}(i), \mathcal{A}_{\omega,j} = 1\},$ and $a_i^* = \|\mathcal{N}^*_{\mathcal{A}}(i)\|$. Similar for $\mathcal{N}^*_{\mathcal{B}}(k)$ and b_k^*. 4. Obtain empirical distributions $\bar{\mu}_i = \frac{1}{a_i^*} \sum_{j\in\mathcal{N}^*_{\mathcal{A}}(i)} \delta_{a_j^{(i)}} - \overline{B}(n^2 - a_i - n, q),$ and similar for $\bar{\nu}_k$. 5. Obtain distance Z_{ik} with $\bar{\mu}_i$ and $\bar{\nu}_k$..	1. Define open neighbourhood $\mathcal{N}_{\mathcal{A}}(i)$ as Chanwoo. Define the connected set $\mathcal{C}_{\mathcal{A}}(i) = \{i \in [n] : i \in \omega, \omega \in \mathcal{N}_{\mathcal{A}}(i)\} \cup \{i\},$ and $c_a(i) = \|\mathcal{C}_{\mathcal{A}}(i)\|$. Then $a_i = P^2_{c_a(i)-1}$ (See caption). Similar for $\mathcal{C}_{\mathcal{B}}(k), c_b(k)$. 2. Define "innerdegree" set of j and i $\mathcal{D}_j(i) = \{(i_1, i_2) : i_1, i_2 \in \mathcal{C}_{\mathcal{A}}(i)/\{j\}\} \cup$ $\{(i_1, i_2) : \text{at least one of } i_1, i_2 = j\},$ and note that $\|\mathcal{D}_j(i)\| = P^2_{c_a(i)-1} = a_i + 2n - 1$. Define "outdegree" of vertex $j \in [n]$ $a_j^{(i)} = \frac{1}{\sqrt{n^2 - \|\mathcal{D}_j(i)\|}} \sum_{\omega\notin\mathcal{D}_j(i)} (\mathcal{A}_{\omega,j} - q),$ similar for $b_j^{(k)}$. 3. Obtain empirical distribution $\bar{\mu}_i = \frac{1}{c_a(i) - 1} \sum_{j\in\mathcal{C}_{\mathcal{A}}(i)/\{j\}} \delta_{a_j^{(i)}} - \overline{B}(n^2 - a_i - 2n + 1, q),$ and similar for $\bar{\nu}_k$. 4. Obtain distance Z_{ik} with $\bar{\mu}_i$ and $\bar{\nu}_k$.

Table 1: Note that P_b^a is the number of permutation of a elements out of b elements. Particularly, $P_b^2 = b(b-1)$.

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.