Higher Order Group Prediction

ABSTRACT

Aim is to predict higher order groups given a previous group interaction history.

General Terms

Hyperedge, Hypergraph

Keywords

Hyperedge Prediction, Hypergraphs, Convex Optimization

1. PROBLEM STATEMENT

Our input is the history of previous collaborations, $\mathbf{Y} = \{\mathbf{y_1}, \mathbf{y_2}, ..., \mathbf{y_{(T-1)}}\}$. Here, each $\mathbf{y_t} = \{v_1, v_2, ..., v_k\}$ represents a particular hyperedge (group or collaboration) that occurs at time instance t with the given vertices each represented by v_i (total number of vertices being n). Our aim is to predict the possible new hyperedge $\mathbf{y_{new}}$ that occurs at $t = T$.

2. APPROACH 1

We treat this problem as an optimization task of finding the vector $\mathbf{y_{new}}$ which is nearest to all the given previous collaboration vectors $\mathbf{y_i}$, $i \in \{1, 2, ..., (T-1)\}$ and is also sparse. We define the following regularization framework:

$$\arg\min_{\mathbf{y_{new}}} \mathbf{Dist}(\mathbf{Y}, \mathbf{y_{new}}) + \lambda \|\mathbf{y_{new}}\|_1 \tag{1}$$

where $\mathbf{Dist}(\mathbf{Y}, \mathbf{y_{new}})$ represents the distance of different kinds like euclidean distance $\|\mathbf{Y} - \mathbf{1}\mathbf{y_{new}^T}\|_2$, hamming distance etc. and $\mathbf{1}$ is a vector of all ones. The second term is lasso constraint for sparsity. The parameter λ is used for tuning the extent of regularization.

3. APPROACH 2

We treat this problem as an optimization task of finding the stack of vectors $\mathbf{Y_{new}} = [\mathbf{y_{new}^1}, ..., \mathbf{y_{new}^P}]^\mathbf{T}$ which is nearest to all the given previous collaboration vectors $\mathbf{y_i}$, $i \in \{1, 2, ..., (T-1)\}$ and is also sparse. We define the following regularization framework:

$$\arg\min_{\mathbf{Y_{new}}} \sum_{p=1}^{P} \sum_{t=1}^{T-1} \mathbf{Dist}(\mathbf{y_t}, \mathbf{y_{new}^P}) + \lambda \sum_{p=1}^{P} \|\mathbf{y_{new}^P}\|_1 \tag{2}$$

where $\mathbf{Dist}(\mathbf{y_t}, \mathbf{y_{new}^P})$ represents the distance of different kinds like euclidean distance $\|\mathbf{y_t} - \mathbf{y_{new}^P}\|_2$, hamming distance etc. The second term is lasso constraint for sparsity. The parameter λ is used for tuning the extent of regularization. If we add diversity in the set of group obtained by regularizing the covariance matrix then we obtain:

$$\arg\min_{\mathbf{Y_{new}}} \sum_{p=1}^{P} \sum_{t=1}^{T-1} \mathbf{Dist}(\mathbf{y_t}, \mathbf{y_{new}^P}) + \lambda_1 \sum_{p=1}^{P} \|\mathbf{y_{new}^P}\|_1 + \lambda_2 \|\mathbf{Y_{new}} \mathbf{Y_{new}^T}\|_2 \tag{3}$$

We do not see any easy way to deal with the third (covariance) term while taking the gradient. Therefore, another method is to iteratively run the optimization for each of the $\mathbf{y_{new}^P}$ and modify the minimized function by adding constraint of diversity from the previous predictions. Following is the method:

$$\arg\min_{\mathbf{y_{new}^P}} \sum_{t=1}^{T-1} \mathbf{Dist}(\mathbf{y_t}, \mathbf{y_{new}^P}) + \lambda_1 \|\mathbf{y_{new}^P}\|_1 + \lambda_2 \mathbf{exp}\left\{-\sum_{j}^{\mathbf{p}-1} \mathbf{Dist}(\mathbf{y_{new}^P}, \mathbf{y_{new}^j})\right\} \tag{4}$$

for the $\mathbf{p}^{th}$ prediction. For $\mathbf{p} = 1$, the third term is zero. Third term decreases as the distance between the new prediction and the previous prediction increases. This convex term therefore, should take care of diversity needed. (**Is there a better way for diversity modeling or is there a straightforward way of dealing with the gradient for Covariance term in (3) ?**)

4. APPROACH 3

Approach 1 is static in the sense that it does not take into account the time dimension. Moreover, **Approach 1** assumes an oversimplified and over demanding in the sense that it wants to find out new groups which are similar to the whole stack of groups observed in past. Rather, a more realistic approach should consider vectors (groups or hyperedges) that are similar to some subset of previously observed groups. Moreover, it is safe to assume that these subsets are the various communities or clusters of related hyperedges observed in past. We therefore, represent the previous collaboration vectors $\mathbf{y_{ic}}$, $i \in \{1, 2, ..., (T-1)\}$ and $c \in \{1, 2, ..., N_c\}$ (N_c being the total number of communities observed in past). We define the following regularization framework:

$$\arg\min_{\mathbf{Y_{new}}} \sum_{p=1}^{P} \sum_{t=1}^{T-1} \gamma^{-t} \mathbf{Dist}(\mathbf{y_{tc}}, \mathbf{y^{P}_{new}}) + \lambda \sum_{p=1}^{P} \|\mathbf{y^{P}_{new}}\|_1 \quad (5)$$

where $\mathbf{Dist}(\mathbf{y_{tc}}, \mathbf{y^{P}_{new}})$ represents the distance of different kinds like euclidean distance $\|\mathbf{y_{tc}} - \mathbf{1}(\mathbf{y^{P}_{new}})^{\mathbf{T}}\|_2$, hamming distance etc, p is the index of the new vectors predicted for each community, and $\mathbf{1}$ is a vector of all ones. The second term is lasso constraint for sparsity. The parameter λ is used for tuning the extent of regularization and parameter γ penalizes more if the predicted hyperedge is not similar to hyperedges in recent past. Community c for $\mathbf{y_{tc}}$ is decided using any clustering methods like KMeans, or columns of a Dictionary. Adding diversity results in :

$$\arg\min_{\mathbf{Y_{new}}} \sum_{p=1}^{P} \sum_{t=1}^{T-1} \gamma^{-t} \mathbf{Dist}(\mathbf{y_{tc}}, \mathbf{y^{P}_{new}}) + \lambda_1 \sum_{p=1}^{P} \|\mathbf{y^{P}_{new}}\|_1 + \lambda_2 \|\mathbf{Y_{new}} \mathbf{Y^{T}_{new}}\|_2 \quad (6)$$

5. ISSUES

- Is this a reasonable approach to take ? Or there is some trivial issue with it ?,
- We can measure distance from the cluster centers rather than the $\mathbf{y_{tc}}$ directly. More generally we can take distance from a dictionary representing $\mathbf{Y}$.

6. ADDITIONAL AUTHORS

7. REFERENCES