International Journal of Mathematical Archive-2(11), 2011, Page: 2184-2187
Available online through www.ijma.info ISSN 2229 – 5046

APPLICATIONS OF CLUSTERING OF NODES IN DIRECTED WEIGHTED TREE

G. N. Purohit [1] and *Ritu Rani Meherwal[2]

Dept of Mathematics and Statistics, Banasthali University, Rajasthan-304022, India

*E-mail: [1]gn_purohitjaipur@yahoo.co.in, [2]ritu.meherwal@gmail.com**

(Received on: 22-10-11; Accepted on: 07-11-11)

ABSTRACT

This paper presents a concise tutorial on spectral clustering for trees which includes weighted tree and directed weighted tree. We show how to transform directed tree into corresponding undirected tree, therefore allowing a unified treatment to all cases.

Key Words: *Eigenvectors, tree clustering, spectral methods, affinity matrix.*

1. INTRODUCTION:

Spectral graph theory, a subtopic of Discrete Mathematics, has applications in many areas of knowledge such as Computer Science, Chemistry, Network design, Coding theory etc. One of the main applications of Spectral graph theory is in clustering of vertices. However clustering may be of any thing like data, nodes etc. Spectral clustering refers to a class of techniques which are based on the eigen structure of a similarity matrix of partition points into disjoint clusters. Points in the same cluster have high similarity and points in different clusters have low similarity. We can form clusters mainly by using K-means clustering and kernel K-means clustering. Many researchers are working for revealing the secret and power of spectral clustering, using eigenvectors of an affinity matrix induced from a graph to develop a natural grouping of the vertices.

Dhillon[2007] considered an equivalence between the objective functions used in the seemingly different methods, spectral clustering and kernel K-means. In particular, general weighted kernel K-means objective is mathematically equivalent to a weighted graph clustering objective. He exploitrd this equivalence to develop a fast high-quality multilevel algorithm that directly optimizes various weighted graph clustering objectives. Xiang and Gong [2007] discussed two critical issues in spectral clustering (1) how to determine the number of clusters automatically and (2) how to perform effective clustering given noisy and sparse data. Luxburg and Belkin[2008] discussed the consistency of the popular family of spectral clustering algorithms for clustering the data with the help of eigenvectors of graph Laplacian matrices and also develop new methods to increase sample size. Pivovaro [2011] introduced a new fast algorithm for clustering and classification of large collections of text documents. This new algorithm employs and bipartite graph that realizes the word document matrix of the collection.

In this paper result based on the node clustering of an undirected weighted tree proved. Further the same result is extended for a directed weighted tree.

2. UNDIRECTED WEIGHTED TREE CLUSTERING:

For the completeness of paper we introduce some concepts and terminologies. Here, T is the weighted tree and W is its affinity matrix. We can also denote T by $T(W)$. Following notations are used in this paper:

$R^{n\times k}$: Matrix of real numbers of order $n\times k$.
$B_+^{n\times k}$: Binary matrix of order $n\times k$.
$[1,k]$: Set of first k natural numbers i.e. $\{1,2,3....k\}$ and
Γ_V^K: Clustering of V nodes into K clusters.

Definition: For a graph $G=(V,E)$ with $|V|=n$, a partition of V into K subsets, $V_1,V_2,\dots V_K$ such that $V_i\cap V_j=\phi$ for $i\neq j$ and $\bigcup_i V_i=V$, is a K-way partition of the nodes of the graph G.

The algorithm for K-way partitioning of a graph is given as:

***Corresponding author: *Ritu Rani Meherwal[2],*E-mail: [2]ritu.meherwal@gmail.com**

Step-1: Construct K-clusters of the vertices of the vertex set V arbitrarily.

Step-2: Find the centroid for each cluster.

Step-3: Determine the most similar vertices to the centroids and construct new clusters.

Step-4: Repeat steps 2 and 3 until there is no change in centroids obtained in previous iteration and each node lies exactly in one cluster.

Now we have the following theorem,

Theorem: 2.1 *Weighted tree node clustering is equivalent to the trace maximization problem of a symmetric matrix.*

Proof. Let $W \in R^{N\times N}$ be the symmetric affinity matrix induced from a weighted tree. By K-way partitioning of $T(W)$ we can find set of K-clusters $V_1, V_2, \dots V_K$ of V denoted by Γ_V^K. Now we define general weighted association objective function of Γ_V^K i.e.

$$GWAssoc(\Gamma_V^K) = \frac{1}{K}\sum_{k=1}^{K} \frac{links(V_k, V_k)}{w(V_k)}, \tag{2.1}$$

where

$$links(A, B) = \sum_{i\in A, j\in B} W(i, j)$$

A and B are the clusters of the vertices of V and $links(A, B)$ denotes the sum of the edge weight between nodes in A and B and $w(V_K) = \sum_{i\in V_K} w_i$, w_i is the sum of the weights of the edges incident on a node of the tree.

Now we use the general weighted association objective function of Γ_V^K. For this we have to maximize the inter cluster relationship i.e.

$$\max\left(GWAssoc\left(\Gamma_V^K\right)\right) = \frac{1}{K}\sum_{k=1}^{K} \frac{links(V_k, V_k)}{w(V_k)} \tag{2.2}$$

Now we use $N \times K$ partition matrix X to represent Γ_V^K. Let $X = \{x_1, x_2 \dots x_K\}$, where x_k is a column vector for V_K. Also $x_i = \{0,0,0\dots,1,0,0..0\}$, 1 is at the ith position if $v_i \in V_i$.

We can rewrite $links(V_k, V_k)$ and $w(V_k)$ as

$$links(V_k, V_k) = x_k^T W x_k$$

$$w(V_k) = x_k^T \phi x_k,$$

where $\phi \in R_+^{N\times N}$ denotes a diagonal matrix with $\phi_{i,i}$ associated with weight of v_i. Then equation(2.2) can be written as

$$max(GWAsso(\Gamma_V^K)) = \frac{1}{K}\sum_{K}^{k=1} \frac{x_k^T W x_k}{x_k^T \phi x_k} \tag{2.3}$$

The details of reduction of equation(2.3) in the form of equation(2.4) are given in appendix-I.

$$max(GWAsso(\Gamma_V^K)) = max(J_u) = \frac{1}{K} tr\left(\frac{X^T W X}{X^T \phi X}\right)$$

$$= \frac{1}{K} tr(\tilde{X}^T \phi^{-1/2} W \phi^{-1/2} \tilde{X})$$

(2.4)

where, $\tilde{X} \in R_+^{N\times K} = \frac{X}{(XX^T)^{1/2}}$ denotes its orthonormal version.

By relaxing the strict non-negativity constraints, that is allowing $\widetilde{K}$ to contain negative values while preserving its orthonormality, according to the Ky Fan theorem [4], the global optimum of J_u can be obtained by assigning

$$\hat{X} = [u_1, u_2, u_3 \dots, u_k]Q, \tag{2.5}$$

where $u_1, u_2, u_3 \ldots, u_k \in C^{N\times K}$ denotes the first K eigen vectors of $\phi^{-1/2}W\phi^{-1/2}$ and $\hat{X} \in C^{N\times K}$ denotes a relaxed version of $\tilde{X}$ and $Q \in R^{K\times K}$ denotes an arbitrary orthonormal matrix. Hence, equation(2.5) present a solution for equation (2.4).

3. DIRECTED WEIGHTED TREE CLUSTERING:

In the previous section, we proved a result for undirected weighted tree clustering. Now, we extend it for clustering of nodes in a directed tree. The basic terminologies and notations have already been discussed in the previous section. Now we prove the following theorem

Theorem 3.1 *Directed weighted tree node clustering is equivalent to the trace maximization problem of a symmetric matrix.*

Proof: Let $W \in R_+^{N\times N}$ be the affinity matrix induced from a directed weighted graph, and ϕ_i and ϕ_o be diagonal weighted matrices associated with indegree and outdegree of vertices in $T(W)$ respectively. We define a diagonal weight matrix of $T(W)$ with

$$\phi_{io} = \sqrt{\phi_i \phi o} \tag{3.1}$$

Now, by K-way partitioning of $T(W)$ we can find set of K-clusters $V_1, V_2, \ldots V_K$ of V denoted by Γ_V^K. We use the general weighted association objective function of Γ_V^K. For this we have to maximize the inter cluster relationship i.e

$$max(GWAssoc(\Gamma_V^K)_1) = \frac{1}{K}\sum_K^{k=1} \frac{links(V_k, V_k)}{w(V_k)}$$

$$= \frac{1}{K}\sum_K^{k=1} \frac{x_k^T W x_k}{x_k^T \phi_{io} x_k}$$

$$= \frac{1}{K} tr\left(\frac{X^T W X}{X^T \phi_{io} X}\right)$$

$$= \frac{1}{K} tr(\tilde{X}^T \phi_{io}^{-1/2} W \phi_{io}^{-1/2} \tilde{X}) \tag{3.2}$$

Similarly for $T(W^T)$, we have

$$max\left(GWAssoc(\Gamma_V^K)_2\right) = \frac{1}{K} tr(\tilde{X}^T \phi_{io}^{-1/2} W^T \phi_{io}^{-1/2} \tilde{X}) \tag{3.3}$$

Adding equation's (3.2)and (3.3), we get

$$max(J) = \frac{1}{K} tr\left(\tilde{X}^T \phi_{io}^{-\frac{1}{2}} (W + W^T) \phi_{io}^{-\frac{1}{2}} \tilde{X}\right), \tag{3.4}$$

which is the trace maximization problem of a symmetric matrix $\phi_{io}^{-1/2}(W + W^T)\phi_{io}^{-1/2}$.

REFERENCES:

[1] Beineke L. W., Wilson R. J. and Cameron P. J., Topics in Algebraic Graph Theory, Cambridge University Press, Cambridge, 2004.

[2] Dhillon I. S., Guan Y., and Kulis B., Kernel K-means, spectral clustering and normalized cuts, Proc. 10th ACM Knowledge Discovery and Data Mining Conference, (2004), 551-556.

[3] Dhillon I. S., Guan Y., and Kulis B., Weighted graph cuts without eigenvectors: A multilevel approach, IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(11) (2007), 1944-1957.

[4] Nakirc I. and Veselirc K., Wielandt and Ky-Fan theorem for matrix pairs, Linear Algebra and its Applications, 369(17) (2003), 77-93.

[5] Zha H., He X., Ding C., Simon H., and Gu M., Spectral relaxation for K-means clustering, Proc. 14th Advances in Neural Information Processing Systems, (2001), 1057-1064.

Appendix-I

$$\sum_{K}^{k=1} \frac{x_k^T W x_k}{x_k^T \phi x_k} = \frac{x_1^T W x_1}{x_1^T \emptyset x_1} + \frac{x_2^T W x_2}{x_2^T \emptyset x_2} + \cdots \frac{x_K^T W x_K}{x_K^T \emptyset x_K}$$

Where,

$$x_k = [x_{k1}, x_{k2}, \dots, x_{kn}], \qquad W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ \dots & \dots & \dots & \dots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix} \qquad \emptyset = \begin{bmatrix} \emptyset_{11} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \emptyset_{nn} \end{bmatrix}$$

$$\sum_{K}^{k=1} \frac{x_k^T W x_k}{x_k^T \phi x_k} = \frac{\left[\sum_{j=1}^{n} \sum_{i=1}^{n} x_{1i}\, w_{ij} x_{1j}\right]}{\left[\sum_{i=1}^{n} x_{1i}\, \emptyset_{ii} x_{1i}\right]} + \frac{\left[\sum_{j=1}^{n} \sum_{i=1}^{n} x_{2i}\, w_{ij} x_{2j}\right]}{\left[\sum_{i=1}^{n} x_{2i}\, \emptyset_{ii} x_{2i}\right]} + \cdots + \frac{\left[\sum_{j=1}^{n} \sum_{i=1}^{n} x_{ni}\, w_{ij} x_{nj}\right]}{\left[\sum_{i=1}^{n} x_{ni}\, \emptyset_{ii} x_{ni}\right]}$$

$$= max\left\{trace\left(\frac{X^T W X}{X^T \emptyset X}\right)\right\},$$

Where,

$$X = [x_1, x_2, \dots, x_k]$$

****