Chapter 5

Graph Clustering Using Ratio Cuts

In this short chapter, we consider the alternative to normalized cut, *called ratio cut*, and show that the methods of Chapters 3 and 4 can be trivially adapted to solve the clustering problem using ratio cuts.

All that needs to be done is to replace the normalized Laplacian L_{sym} by the unormalized Laplacian L, and omit the step of considering Problem $(**_2)$.

In particular, there is no need to multiply the continuous solution Y by $D^{-1/2}$.

The idea of ratio cut is to *replace the volume* $\mathrm{vol}(A_j)$ *of each block* A_j *of the partition by its size* $|A_j|$ (the number of nodes in A_j).

First, we deal with unsigned graphs, the case where the entries in the symmetric weight matrix W are nonnegative.

Definition 5.1. The *ratio cut* $\mathrm{Rcut}(A_1, \ldots, A_K)$ of the partition $(A_1, \ldots, A_K)$ is defined as

$$\mathrm{Rcut}(A_1, \ldots, A_K) = \sum_{i=1}^{K} \frac{\mathrm{cut}(A_j, \overline{A}_j)}{|A_j|}.$$

As in Section 3.3, given a partition of V into K clusters $(A_1, \ldots, A_K)$, if we represent the jth block of this partition by a vector X^j such that

$$X_i^j = \begin{cases} a_j & \text{if } v_i \in A_j \\ 0 & \text{if } v_i \notin A_j, \end{cases}$$

for some $a_j \neq 0$, then

$$\begin{aligned} (X^j)^\top L X^j &= a_j^2 (\mathrm{cut}(A_j, \overline{A_j}) \\ (X^j)^\top X^j &= a_j^2 |A_j|. \end{aligned}$$

Consequently, we have

$$\begin{aligned} \mathrm{Rcut}(A_1, \ldots, A_K) = \sum_{i=1}^{K} \frac{\mathrm{cut}(A_j, \overline{A}_j)}{|A_j|} \\ = \sum_{i=1}^{K} \frac{(X^j)^\top L X^j}{(X^j)^\top X^j}. \end{aligned}$$

On the other hand, the normalized cut is given by

$$\begin{aligned}\mathrm{Ncut}(A_1, \ldots, A_K) &= \sum_{i=1}^{K} \frac{\mathrm{cut}(A_j, \overline{A}_j)}{\mathrm{vol}(A_j)} \\ &= \sum_{i=1}^{K} \frac{(X^j)^\top L X^j}{(X^j)^\top D X^j}.\end{aligned}$$

Therefore, ratio cut is the special case of normalized cut where $D = I$.

If we let

$$\begin{aligned}\mathcal{X} = \Big\{ [X^1 \ldots X^K] \mid X^j = a_j(x_1^j, \ldots, x_N^j),\, x_i^j \in \{1, 0\}, \\ a_j \in \mathbb{R},\ X^j \neq 0 \Big\}\end{aligned}$$

(note that the condition $X^j \neq 0$ implies that $a_j \neq 0$), then the set of matrices representing partitions of V into K blocks is

$$\mathcal{K} = \Big\{X = [X^1 \ \cdots \ X^K] \ \mid \ X \in \mathcal{X}, \ (X^i)^\top X^j = 0, \ 1 \leq i, j \leq K, \ i \neq j\Big\}.$$

Here is our first formulation of K-way clustering of a graph using ratio cuts, called problem PRC1 :

K-way Clustering of a graph using Ratio Cut, Version 1:
Problem PRC1

$$\begin{aligned} &\text{minimize} && \sum_{j=1}^{K} \frac{(X^j)^\top L X^j}{(X^j)^\top X^j} \\ &\text{subject to} && (X^i)^\top X^j = 0, \quad 1 \leq i, j \leq K, \ i \neq j, \\ &&& X \in \mathcal{X}. \end{aligned}$$

The solutions that we are seeking are K-tuples $(\mathbb{P}(X^1), \ldots, \mathbb{P}(X^K))$ of points in $\mathbb{RP}^{N-1}$ determined by their homogeneous coordinates $X^1, \ldots, X^K$.

As in Chapter 3, chasing denominators and introducing a trace, we obtain the following formulation of our minimization problem:

K-way Clustering of a graph using Ratio Cut, Version 2:
Problem PRC2

$$\begin{aligned} &\text{minimize} && \operatorname{tr}(X^\top L X) \\ &\text{subject to} && X^\top X = I, \\ & && X \in \mathcal{X}. \end{aligned}$$

The natural relaxation of problem PRC2 is to drop the condition that $X \in \mathcal{X}$, and we obtain the

Problem $(R{*}_2)$

$$\begin{array}{ll} \text{minimize} & \operatorname{tr}(X^\top L X) \\ \text{subject to} & X^\top X = I. \end{array}$$

This time, since the normalization condition is $X^\top X = I$, we can use the eigenvalues and the eigenvectors of L, and by Proposition A.2, the minimum is achieved by any K unit eigenvectors $(u_1, \ldots, u_K)$ associated with the smallest K eigenvalues

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_K$$

of L.

The matrix $Z = Y = [u_1, \ldots, u_K]$ yields a minimum of our relaxed problem $(R*_2)$.

The rest of the algorithm is as before; we try to find $Q = R\Lambda$ with $R \in \mathbf{O}(K)$, Λ diagonal invertible, and $X \in \mathcal{X}$ such that $\|X - ZQ\|$ is minimum.

In the case of signed graphs, we define the *signed ratio cut* $\mathrm{sRcut}(A_1, \ldots, A_K)$ of the partition $(A_1, \ldots, A_K)$ as

$$\mathrm{sRcut}(A_1, \ldots, A_K) = \sum_{j=1}^{K} \frac{\mathrm{cut}(A_j, \overline{A_j})}{|A_j|} + 2 \sum_{j=1}^{K} \frac{\mathrm{links}^-(A_j, A_j)}{|A_j|}.$$

Since we still have

$$(X^j)^\top \overline{L} X^j = a_j^2(\mathrm{cut}(A_j, \overline{A_j}) + 2\mathrm{links}^-(A_j, A_j)),$$

we obtain

$$\mathrm{sRcut}(A_1, \ldots, A_K) = \sum_{j=1}^{K} \frac{(X^j)^\top \overline{L} X^j}{(X^j)^\top X^j}.$$

Therefore, this is similar to the case of unsigned graphs, with L replaced with $\overline{L}$.

The same algorithm applies, but as in Chapter 4, the signed Laplacian $\overline{L}$ is positive definite iff G is unbalanced.

Modifying the computer program implementing normalized cuts to deal with ratio cuts is trivial (use $\overline{L}$ instead of $\overline{L}_{\mathrm{sym}}$ and don't multiply Y by $\overline{D}^{-1/2}$).

Generally, normalized cut seems to yield "better clusters," but this is not a very satisfactory statement since we haven't defined precisely in which sense a clustering is better than another.

We leave this point as further research.