# Homework 2

Guoxiang Grayson Tong

October 13, 2022

**Problem 1:** Consider a graph: $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, the normalized Cut (NCut) clustering objective for $K = 2$ can be written as:

$$\min_{\mathcal{A}\in\mathcal{V}} \Big(R^{\text{NCut}}(\mathcal{A}, \bar{\mathcal{A}})\Big).$$

Define a vector $\boldsymbol{\alpha} \in \mathbb{R}^{|\mathcal{V}|}$ as:

$$\boldsymbol{\alpha}(u \in \mathcal{V}) = \begin{cases} \sqrt{\dfrac{vol(\bar{\mathcal{A}})}{vol(\mathcal{A})}}\,, & u \in \mathcal{A} \\ -\sqrt{\dfrac{vol(\mathcal{A})}{vol(\bar{\mathcal{A}})}}\,, & u \in \bar{\mathcal{A}} \end{cases} \quad , \tag{1}$$

where $vol(\mathcal{A}) = \sum_{u\in\mathcal{A}} d(u)$.

Note that $\mathbf{L}$ is the laplacian matrix of graph $\mathcal{G}$, and by its property, we have

$$\begin{aligned}
\boldsymbol{\alpha}^T\mathbf{L}\boldsymbol{\alpha} &= \sum_{(u,v),u\in\mathcal{A},v\in\bar{\mathcal{A}}} (\boldsymbol{\alpha}[u] - \boldsymbol{\alpha}[v])^2 \\
&= \sum_{(u,v),u\in\mathcal{A},v\in\bar{\mathcal{A}}} \Big(\frac{vol(\bar{\mathcal{A}}) + vol(\mathcal{A})}{vol(\mathcal{A})} + \frac{vol(\mathcal{A}) + vol(\bar{\mathcal{A}})}{vol(\bar{\mathcal{A}})}\Big) \\
&= \Big(vol(\bar{\mathcal{A}}) + vol(\mathcal{A})\Big)\Big(\frac{\sum_{(u,v),u\in\mathcal{A},v\in\bar{\mathcal{A}}}}{vol(\mathcal{A})} + \frac{\sum_{(u,v),u\in\mathcal{A},v\in\bar{\mathcal{A}}}}{vol(\bar{\mathcal{A}})}\Big) \\
&= \Big(vol(\bar{\mathcal{A}}) + vol(\mathcal{A})\Big)\Big(\frac{\text{cut}(\mathcal{A}, \bar{\mathcal{A}})}{vol(\mathcal{A})} + \frac{\text{cut}(\mathcal{A}, \bar{\mathcal{A}})}{vol(\bar{\mathcal{A}})}\Big) \\
&= \Big(vol(\bar{\mathcal{A}}) + vol(\mathcal{A})\Big)R^{\text{NCut}}(\mathcal{A}, \bar{\mathcal{A}}) \\
&= 2|\mathcal{E}|R^{\text{NCut}}(\mathcal{A}, \bar{\mathcal{A}})\ ,
\end{aligned}$$

since we have

$$\begin{aligned}
\text{cut}(\mathcal{A}, \bar{\mathcal{A}}) &= \frac{1}{2}\Big(\sum_{(u,v),u\in\mathcal{A},v\in\bar{\mathcal{A}}} 1 + \sum_{(u,v),u\in\hat{\mathcal{A}},v\in\mathcal{A}} 1\Big) \\
&= \sum_{(u,v),u\in\mathcal{A},v\in\bar{\mathcal{A}}} 1 \\
&= \sum_{(u,v),u\in\bar{\mathcal{A}},v\in\mathcal{A}} 1\ .,
\end{aligned}$$

and

$$vol(\bar{\mathcal{A}}) + vol(\mathcal{A}) = \sum_{v_i \in \mathcal{A}} d(v_i) + \sum_{v_i \in \bar{\mathcal{A}}} d(v_i) = 2|\mathcal{E}|.$$

As a result, we have the following equivalence

$$\min_{\mathcal{A} \in \mathcal{V}} \Big( R^{\text{NCut}}(\mathcal{A}, \bar{\mathcal{A}}) \Big) \;\Leftrightarrow\; \underset{\boldsymbol{\alpha} \text{ of the form (1)}}{\operatorname{argmin}} \;\; \boldsymbol{\alpha}^T \mathbf{L} \boldsymbol{\alpha}.$$

Now consider the degree matrix

$$\mathbf{D} = diag\Big( d(v_1), d(v_2), \cdots, d(v_{|\mathcal{V}|}) \Big) \ ,$$

such that

$$\begin{aligned} \sum_{i=1}^{|\mathcal{V}|} (\mathbf{D}\boldsymbol{\alpha})_i &= \sum_{i=1}^{|A|} d(v_i)\alpha(v_i \in \mathcal{A}) + \sum_{j=1}^{|\bar{A}|} d(v_j)\alpha(v_i \in \bar{\mathcal{A}}) \\ &= vol(A)\alpha(v_i \in \mathcal{A}) + vol(\bar{A})\alpha(v_i \in \bar{\mathcal{A}}) \\ &= \sqrt{vol(\mathcal{A})vol(\bar{\mathcal{A}})} - \sqrt{vol(\mathcal{A})vol(\bar{\mathcal{A}})} \\ &= 0 \ . \end{aligned}$$

Thus, one constraint of $\boldsymbol{\alpha}$ would be $\sum_{i=1}^{|\mathcal{V}|} (\mathbf{D}\boldsymbol{\alpha})_i = 0$, which is also equivalent to $(\mathbf{D}\boldsymbol{\alpha}) \perp \mathbf{1}$.

Further more, we can have another constraint of $\boldsymbol{\alpha}$ as

$$\begin{aligned} \|\mathbf{D}(\boldsymbol{\alpha} \odot \boldsymbol{\alpha})\|_1 &= \sum_{i=1}^{|A|} d(v_i)\alpha^2(v_i \in \mathcal{A}) + \sum_{j=1}^{|\bar{A}|} d(v_j)\alpha^2(v_i \in \bar{\mathcal{A}}) \\ &= vol(\mathcal{A})\frac{vol(\bar{\mathcal{A}})}{vol(\mathcal{A})} + vol(\bar{\mathcal{A}})\frac{vol(\mathcal{A})}{vol(\bar{\mathcal{A}})} \\ &= 2|\mathcal{E}| \ , \end{aligned}$$

where $\odot$ denotes the Hadamard product.

As above, the equivalent normalized spectral clustering problem of Ncut clustering is then defined as:

$$\begin{cases} \underset{\boldsymbol{\alpha} \text{ of the form (1)}}{\operatorname{argmin}} \;\; \boldsymbol{\alpha}^T \mathbf{L} \boldsymbol{\alpha} \\ s.t. \ (\mathbf{D}\boldsymbol{\alpha}) \perp \mathbf{1} \\ \qquad \|\mathbf{D}(\boldsymbol{\alpha} \odot \boldsymbol{\alpha})\|_1 = 2|\mathcal{E}| \end{cases} \tag{2}$$