1 Graph partitioning using spectral methods

Recall Cheeger’s inequality

$$\frac{d - \lambda_2}{2} \leq h_G \leq \sqrt{2d(d - \lambda_2)}$$  \hspace{1cm} (1)

Here $G$ is a $d$-regular Graph and $h_G = \min_{S,|S| \leq |V|} \frac{E(S,\bar{S})}{|S|}$ denotes the edge-expansion of the graph. Also $\lambda_2 = \max_{x \perp \bar{1}} \frac{x^T A x}{x^T x}$ denotes the Fiedler value of the graph (least non-zero eigenvalue of the adjacency matrix of the graph) and $\bar{1}$ denotes the all ones vector.

Alternate form of Cheeger's inequality which holds for all connected graphs $G$ \[1\]

$$2h_G \geq \lambda_G \geq \alpha_G \geq h_G^2/2$$  \hspace{1cm} (2)

where $G$ is any graph, $h_G$ is the Cheeger constant for the graph defined as $h_G = \min_{S, \text{Vol}(S) \leq \text{Vol}(V)/2} \frac{E(S,\bar{S})}{\text{Vol}(S)}$, $\lambda_G$ denotes Fiedler value of the graph Laplacian (i.e. second largest eigenvalue of the graph laplacian) and $\alpha_G$ represents minimum value among all Cheeger ratios of initial segments of vertices when all vertices are arranged in a line using the eigen-vector associated with $\lambda_2$.

**Proof of** $\frac{d - \lambda_2}{2} \leq h_G$ in (1)

Consider the quadratic form

$$\sum_{ij} A_{ij} (x_i - x_j)^2 = 2d(x^T x) - 2\sum_{ij} x_i A_{ij} x_j.$$  \hspace{1cm} (3)

Recall

$$\lambda_2 = \max_{x \perp \bar{1}} \frac{x^T A x}{x^T x} = \max_{x \perp \bar{1}} \frac{2x^T x - x^T (1/2) \sum_{ij} A_{ij} (x_i - x_j)^2}{x^T x}.$$  \hspace{1cm} (4)

Let $S$ denote the set which achieves the minimum Cheeger ratio i.e. $\frac{E(S,\bar{S})}{\min(|S||\bar{S}|)} = h_G$. Let $p = |S|/n$ and $q = |\bar{S}|/n = 1 - p$. Let

$$X_i = q \text{ if } i \in S$$
$$X_i = -p \text{ if } i \in \bar{S}.$$  \hspace{1cm} (5)

Then $x \cdot \bar{1} = |S|q + |\bar{S}|p = 0 \iff x \perp \bar{1}$. Also, $x^T x = |S|q^2 + |\bar{S}|p^2 = npq^2 + np^2 = npq(p + q) = npq$. Then,

$$d - \lambda_2 = \min_{x \perp \bar{1}} \frac{\sum_{ij} A_{ij} (x_i - x_j)^2}{2x^T x} = \frac{E(S,\bar{S})}{npq} = \frac{nE(S,\bar{S})}{(|S||\bar{S}|)}.$$  \hspace{1cm} (6)

$^1\text{Vol}(S) = \sum_{i \in S} d_i$
Now we know that the sparsity (sp) of the cut \((S, \bar{S})\) is defined as 
\[ sp(S) = \frac{nE(S, \bar{S})}{\min(|S|, |\bar{S}|)} \]  
and using the fact that \(\max(|S|, |\bar{S}|) \geq n/2\), we get \(sp(S) \leq 2 \frac{E(S, \bar{S})}{\min(|S|, |\bar{S}|)} = 2h_G\). Hence we get \(d - \lambda_2 \leq 2h_G\).

In fact if \(\phi\) denotes the sparsest cut (i.e. the cut of minimum sparsity) then it can be shown that 
\[ \phi \geq h_G \geq (1/2)\phi \]  
(7)

Hence, an approximation to the sparsest cut is a 2-approximation to \(h_G\).

**Another explanation:** We have already seen that 
\[ \sum_{i,j} A_{ij}(x_i - x_j)^2 = 2dx^T x - 2x^T Ax \]
\[ \Leftrightarrow \sum_{i,j} (x_i - x_j)^2 = 2nx^T x - 2x^T 1x \]  
(8)

where \(1\) denotes the matrix of all ones, i.e. \(1 \cdot x = 0, \forall x \perp 1\)

Plugging this back in (9), we get,
\[ d - \lambda_2 = \min_{x \perp 1} \frac{\sum_{i,j} A_{ij}(x_i - x_j)^2}{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2} \]  
(9)

However observe that the right side is invariant under the transformation \(x \rightarrow x + c1\). Hence we can choose \(c\) in order to eliminate the constraint \(x \perp 1\).

Hence
\[ d - \lambda_2 = \min_{x \in \mathbb{R}^n} \frac{\sum_{i,j} A_{ij}(x_i - x_j)^2}{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2} \]  
(10)

Now again, recall the sparsest cut
\[ \phi = \min_S \frac{nE(S, S)}{|S||\bar{S}|} \]
\[ = \min_{x \in \{0,1\}^n} \frac{\sum_{i,j} |A_{ij}| |x_i - x_j|}{\frac{1}{n} \sum_{i,j} |x_i - x_j|} \]
\[ = \min_{x \in \{0,1\}^n} \frac{\sum_{i,j} A_{ij}(x_i - x_j)^2}{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2} \]  
(11)

Hence comparing (10) and (11), we get \(d - \lambda_2 \leq \phi\). Also, using (7), we conclude that \(d - \lambda_2 \leq 2h_G\)

**2 Claim:** \(h \leq \sqrt{2d(d - \lambda_2)}\) for \(d\)-regular graphs

**Proof:**

\[ \forall y \in \mathbb{R}^n, \sum_{i,j} A_{ij} |y_i^2 - y_j^2| \leq \sqrt{2dy^Ty - 2dy^T Ay} \sqrt{4dy^Ty} \]
LHS = \sum_{i,j} A_{i,j}^2 |y_i - y_j||y_i + y_j| A_{i,j}^2 \\
\leq \left( \sum_{i,j} A_{i,j} |y_i - y_j|^2 \right)^{\frac{1}{2}} \left( \sum_{i,j} A_{i,j} |y_i + y_j|^2 \right)^{\frac{1}{2}} \\
\leq \left( \sum_{i,j} A_{i,j} |y_i - y_j|^2 \right)^{\frac{1}{2}} \left( 2 \sum_{i,j} 2 A_{i,j} (y_i^2 + y_j^2) \right)^{\frac{1}{2}} \\
\leq \left( \sum_{i,j} A_{i,j} |y_i - y_j|^2 \right)^{\frac{1}{2}} \left( 2d \sum_{i,j} y_i^2 + y_j^2 \right)^{\frac{1}{2}} \\
\leq (2dy^Ty - 2y^TAy)^{\frac{1}{2}} \left( 4dy^Ty \right)^{\frac{1}{2}} \\
\vdash 2: \begin{cases} \text{Let } x \text{ be an eigenvector with eigenvalue } \lambda, \text{ i.e. } xA = \lambda x \text{ s.t. } |\{i : x_i > 0\}| \leq \frac{n}{2} \\
\text{Define } y : y_i = \max\{x_i, 0\}, \text{ then } yA \geq \lambda y \text{ componentwise.} \end{cases} \\
\text{Since } A \text{ is positive, we have} \\
\begin{align*} x_i > 0 & \quad (yA)_i \geq (xA)_i = (\lambda x)_i = (\lambda y)_i \\
 x_i < 0 & \quad (yA)_i \geq (xA)_i \\
\end{align*} \\
\text{For } y \text{ defined before, } \sum_{i,j} A_{i,j} |y_i^2 - y_j^2| \geq 2hy^Ty \\
\text{Arrange the components of } y \text{ in non-increasing order} \\
y(i_1) \geq y(i_2) \geq \cdots \geq y(i_n) \\
\text{With } t \text{ of them strictly greater than } 0, \text{ i.e.} \\
y(i_t) > y(i_{t+1}) = \cdots = y(i_n) = 0 \\
\text{Let } K \text{ be the set such that jump occurs:} \\
K = \{i : y(i_k) > y(i_{k+1})\} \\
\sum_{u,v} A_{u,v} |y(u)^2 - y(v)^2| = 2 \sum_{i=1}^{t} \sum_{j=i+1}^{n} A_{v_i,v_j} (y(v_i)^2 - y(v_j)^2) \\
= 2 \sum_{k \in K} \sum_{i \leq k} \sum_{j > k} A_{v_i,v_j} (y(v_i)^2 - y(v_j)^2) \\
\text{For each } k = 1, \cdots, n, \text{ let} \\
L_k = \{v_i : i \leq k\} \\
L_0 = \phi
Note $\sum_{i\leq k} \sum_{j > k} A_{v_i,v_j} \geq h|L_k| \ (*)$

\[\star \text{RHS} \geq 2 \sum_{k} h|L_k|(y(v_k)^2 - y(v_{k+1})^2)\]
\[= 2h \sum_{k} ((|L_k| - |L_{k'}|)y(v_k)^2\]
\[= 2h \sum_{k} |\{v: y(v) = y(v_k)\}|y(v_k)^2\]
\[= 2d \sum_{v} y(v)^2\]
\[= 2hy^Ty\]

\[h \leq \frac{\sum_{i,j} A_{i,j}(y_i - y_j)^2}{2y^Ty}\]
\[\leq \frac{2dy^Ty - 2y^TAy}{2y^Ty}\]
\[\leq \frac{(2d - \lambda_2)^{\frac{1}{2}}}{2y^Ty}\]

Algorithm (Spectral Graph Partitioning)

1. Compute 2nd eigenvector of $A/L$
2. Perform a sweep cut in some way (i.e. Check the set of best notes derived from the eigenvector)
3. Keep the best cut

Potential Issues

- The actual set returned might not be ‘good’, e.g. close to optimal
- The eigenvector computation may be too expensive.
- Local information may be reliable, but global not. For extremely large graphs, or graphs with local information nice, global properties bad. We may want to cluster locally, pull out a set of nodes that are good near you.

Lots of heuristics are motivated by this: Cut out nearest neighbors, 2nd nearest neighbors, . . .

Can we inherit some of the nice properties of the global spectral? For example the Cheeger’s inequality, sweep cut, . . .

2 ways to be local:

- Bias yourself locally, but still do computation that depend on the size of the graph.
- Have computation that depend on the size of the output, not the size of the graph.

References