Spectral Graph Theory and Cheeger's Inequality

Michael Tang

The goal of this paper is to establish Cheeger's Inequality for d-regular undirected graphs. Along the way, we prove the Courant-Fischer Theorem for the real finite-dimensional case and showcase the power of the probabilistic method. The main proof is based on [4], with some background and examples based on [1], [5], [3]. We assume basic knowledge of linear algebra and probability, but derive graph theory from fundamentals.

1 Graph fundamentals

Definition 1.1. *An undirected graph is a pair of sets $G = (V, E)$, where the elements of V are called vertices and the elements of E are called edges. G is finite if V, E are finite sets. In particular*

$$E \subseteq \{\{u, v\} : u, v \in V, u \neq v\}$$

For convenience, whenever we define a graph G, fix some labeling of the vertices as $v_1, ..., v_n$ where $n = |V|$. Unless otherwise specified, assume all graphs in this paper are undirected and finite, and all vector spaces are real, finite-dimensional, and equipped with the Euclidean dot product.

Definition 1.2. *A pair of vertices $u, v \in V$ are adjacent if $\{u, v\} \in E$. Adjacent vertices are also called neighbors. For a vertex $v \in V$, let the degree $\deg(v)$ be the number of vertices in G adjacent to v. If all vertices of G are adjacent, we say G is complete. If all vertices in a graph G have degree d, we say G is d-regular.*

Definition 1.3 (Adjacency matrix, normalized adjacency matrix)**.** *For a graph G with $n = |V|$, let its adjacency matrix $A \in \mathbb{R}^{n,n}$ be*

$$A_{ij} = \begin{cases} 1 & \textit{if } \{v_i, v_j\} \in E \\ 0 & \textit{otherwise} \end{cases}$$

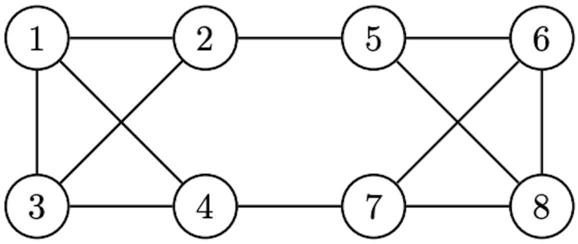
For G d-regular, let its normalized adjacency matrix $M = A/d$.

Notice for undirected graphs A, M are symmetric by construction, and we have $A, M \in \mathbb{R}^{n,n}$. It also follows directly that

Proposition 1.1. *For a d-regular graph G with $n = |V|$, for all i, j, we have that*

$$\sum_{k=1}^{n} A_{ik} = \sum_{k=1}^{n} A_{kj} = d$$
$$\sum_{k=1}^{n} M_{ik} = \sum_{k=1}^{n} M_{kj} = 1$$

Example 1.1. The following graph is 3-regular:



and has adjacency matrix

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

which we see is symmetric and satisfies $\sum_{k=1}^{8} A_{ik} = \sum_{k=1}^{8} A_{kj} = 3$.

Definition 1.4. *The cycle graph $C_n = (V, E)$ has $|V| = n, E = \{\{v_i, v_{i+1}\} : i = 1, ..., n-1\} \cup \{\{1, n\}\}$.*

Definition 1.5. *The complete graph $K_n = (V, E)$ has $|V| = n, E = \{\{u, v\} : u, v \in V, u \neq v\}$.*

Definition 1.6 (Cuts and partitions)**.** *A partition $P = (X_1, ..., X_n)$ of sets X_i is a collection of disjoint subsets $X_i \subseteq X$ with $\cup_i X_i = X$. In other words, each element of X is in exactly one subset X_i. A cut $(S, V - S)$ of a graph $G = (V, E)$ is a partition of V into two subsets. Let the set of crossing edges $E(S, V - S) \subseteq E$ be the set of edges (u, v) with $u \in S, v \in V - S$. (note that we use $-$ in this context to denote set difference) A nontrivial cut $(S, V - S)$ is one where both subsets are non-empty.*

2 Eigenvalues of graphs

Definition 2.1. *For M a self-adjoint operator over an inner product space $\mathcal{V} = \mathbb{R}^n$, and a nonzero $x \in \mathcal{V}$, let Rayleigh's quotient be*

$$R(x) = \frac{\langle x, Mx \rangle}{||x||^2}$$

Theorem 2.1 (Courant-Fischer Theorem)**.** *Let M be a self-adjoint operator over an inner product space $\mathcal{V} = \mathbb{R}^n$ with eigenvalues $\lambda_1 \geq ... \geq \lambda_n$ repeated according to their multiplicities and corresponding eigenvectors $e_1, ..., e_n$. Then*

$$\lambda_1 = \max_{x \neq \mathbf{0}} R(x)$$

and in general

$$\lambda_k = \max_{\substack{x \neq \mathbf{0} \\ x \perp e_1, ..., x \perp e_{k-1}}} R(x)$$

Proof. By the Real Spectral Theorem we know $e_1, ..., e_n$ form a basis of $\mathcal{V} = \mathbb{R}^n$. For any nonzero $x \in \mathcal{V}$ (note $||x||^2 = 0$ iff $x = \mathbf{0}$), we can write

$$\begin{aligned} x &= \sum_{i=1}^n \langle x, e_i \rangle e_i \\ \langle x, Mx \rangle &= \langle \sum_{i=1}^n \langle x, e_i \rangle e_i, \sum_{i=1}^n \lambda_i \langle x, e_i \rangle e_i \rangle = \sum_{i=1}^n \langle x, e_i \rangle \langle e_i, \lambda_i e_i \rangle = \sum_{i=1}^n \langle x, e_i \rangle^2 \lambda_i \\ ||x||^2 &= \langle \sum_{i=1}^n \langle x, e_i \rangle e_i, \sum_{i=1}^n \langle x, e_i \rangle e_i \rangle = \sum_{i=1}^n \langle x, e_i \rangle^2 \end{aligned}$$

Thus

$$x = \frac{\langle x, Mx \rangle}{||x||^2} = \frac{\sum_{i=1}^n \langle x, e_i \rangle^2 \lambda_i}{\sum_{i=1}^n \langle x, e_i \rangle^2} \leq \frac{\sum_{i=1}^n \langle x, e_i \rangle^2 \lambda_1}{\sum_{i=1}^n \langle x, e_i \rangle^2} = \lambda_1$$

We also have

$$\frac{\langle e_1, Me_1 \rangle}{||e_1||^2} = \lambda_1$$

so the sup is achieved and thus the max exists. To get the case for general k, we see that

$x \perp e_1, ..., x \perp e_{k-1}$ implies $\langle x, e_i \rangle = 0$ for all $i < k$ so we can similarly write

$$x = \sum_{i=k}^{n} \langle x, e_i \rangle e_i$$
$$\langle x, Mx \rangle = \sum_{i=k}^{n} \langle x, e_i \rangle^2 \lambda_i$$
$$||x||^2 = \sum_{i=k}^{n} \langle x, e_i \rangle^2$$
$$x = \frac{\langle x, Mx \rangle}{||x||^2} = \frac{\sum_{i=k}^{n} \langle x, e_i \rangle^2 \lambda_i}{\sum_{i=k}^{n} \langle x, e_i \rangle^2} \leq \frac{\sum_{i=k}^{n} \langle x, e_i \rangle^2 \lambda_k}{\sum_{i=k}^{n} \langle x, e_i \rangle^2} = \lambda_k$$

and we have

$$\frac{\langle e_k, Me_k \rangle}{||e_k||^2} = \lambda_k$$

so the sup is achieved and thus the max exists. □

Proposition 2.1. *Let $G = (V, E)$ be a d-regular graph and M be its normalized adjacency matrix. Let $n = |V|$, and let the eigenvalues of M be $\lambda_1 \geq ... \geq \lambda_n$ repeated according to their multiplicities and corresponding eigenvectors be $e_1, ..., e_n$. Then $\lambda = 1, e_1 = \mathbf{1}$.*

Proof. Since M has rows that sum to 1 and all entries non-negative, we can fix a λ-eigenvector $v = x_1 f_1 + ... + x_n f_n$ where f_i is the standard basis, and we see that $Mv = \lambda v$ implies $m_{i1} x_1 + ... + m_{in} x_n = \lambda x_i$ for all i, so we can take k such that $x_k \geq x_i$ for all i. $|x_k| > 0$ follows from v nonzero so we can write

$$\begin{aligned}|\lambda| &= \frac{|x_1 f_1 + ... + x_n f_n|}{|x_k|} \\ &\leq a_{k1} \left| \frac{x_1}{x_k} \right| + ... + a_{kn} \left| \frac{x_n}{x_k} \right| \\ &\leq a_{k1} + ... + a_{nk} \\ &= 1\end{aligned}$$

and it is easy to see that taking $v = f_1 + ... + f_n = \mathbf{1}$ gives a 1-eigenvector of M, so for all other eigenvalues λ', $1 \geq \lambda'$ so $1 = \lambda_1$. □

This allows us to write the previous result in a more convenient form for use in proving Cheeger's inequality:

Proposition 2.2. *Let $G = (V, E)$ be d-regular with normalized adjacency matrix M, $|V| = n$, eigenvalues $\lambda_1 \geq ... \geq \lambda_n$ repeated according to their multiplicities*

$$\lambda_2 = \max_{\substack{x \neq \mathbf{0} \\ x \perp \mathbf{1}}} \frac{x^T M x}{x^T x}$$

Proof. $M \in \mathbb{R}^{n,n}$ as an operator is self-adjoint by being symmetric and real (note orthonormality of the standard basis), so we can apply the Courant-Fischer Theorem with the inner product $\langle u, v\rangle = u^T v$ on $k = 2$, and substitute $e_1 = \mathbf{1}$ from 2.1. $\square$

Example 2.1. The Courant-Fischer Theorem allows us to bound λ_2 without any explicit computation of eigenvalues. Consider C_n, and take a vector with a high Rayleigh quotient $R(x)$, such as the one given by

$$x_i = \begin{cases} i - n/4 & \text{if } i \le n/2 \\ 3n/4 - i & \text{if } i > n/2 \end{cases}$$

which we can verify satisfies $\sum_i x_i \neq 0$ so $x \perp \mathbf{1}$, and $x \neq \mathbf{0}$. From this we can compute

$$(Mx)_i = \begin{cases} 1 - n/4 & \text{if } i = 1, n \\ n/4 - 1 & \text{if } i = n/2, n/2+1 \\ x_i & \text{otherwise} \end{cases}$$

since $(Mx)_i = \frac{1}{2}(x_{i-1} + x_{i+1})$ with x_n in place of $i-1$ when $i = 1$ and x_1 in place of $i+1$ when $i = n$. Then, denoting a polynomial expression of n with degree at most k by $P(n^k)$, it is easy to see that

$$\begin{aligned} x^T Mx - x^T x &= \sum_{i=1,n} [(i - n/4)^2 - (i - n/4)(1 - n/4)] + \\ &\quad \sum_{i=n/2,n/2+1} [(3n/4 - i)^2 - (3n/4 - i)(n/4 - 1)] = P(n) \\ x^T x &= \sum_{i=1}^{n/2} (i - n/4)^2 + \sum_{i=n/2+1}^{n} (3n/4 - i)^2 = P(n^3) \end{aligned}$$

so we get

$$\begin{aligned} R(x) &= \frac{x^T Mx}{x^T x} = \frac{x^T x + P(n)}{x^T Mx} = 1 + \frac{P(n)}{P(n^3)} = 1 + \frac{1}{P(n^2)} \\ \lambda_2 &= \max_{\substack{x' \neq \mathbf{0} \\ x' \perp \mathbf{1}}} R(x') \ge R(x) = 1 + \frac{1}{P(n^2)} \end{aligned}$$

giving the asymptotic bound $\lambda_2 \ge 1 + \frac{1}{P(n^2)}$. We will see why this is considered a high $R(x)$ and a strong bound once we prove the Cheeger inequality.

3 Easy direction of Cheeger's inequality

Definition 3.1 (Conductance and edge expansion)**.** *Given a d-regular graph $G = (V, E)$ and a nontrivial cut $(S, V - S)$, let the edge expansion $h(S)$ and conductance $\phi(S)$ of the* *cut be*

$$h(S) = \frac{|E(S, V-S)|}{d \min(|S|, |V-S|)}$$
$$\phi(S) = \frac{|E(S, V-S)|}{\frac{d}{|V|}|S||V-S|}$$

and the conductance $\phi(G)$ *and edge expansion* $h(G)$ *of* G *be*

$$h(G) = \min_S h(G)$$
$$\phi(G) = \min_S \phi(G)$$

with the min taken over all nontrivial cuts $(S, V-S)$ *of* G*. Note in this context we will sometimes use just* S *to denote a cut* $(S, V-S)$*. The edge expansion* h *is also called the Cheeger constant or isoperimetric number, and the conductance* ϕ *is also called the sparsity.*

We can quickly show a relation between these two quantities using their definitions.

Proposition 3.1. *Given a graph* G*, we can write*

$$h(G) \leq \phi(G) \leq 2h(G)$$

Proof. Fix a nontrivial cut $(S, V-S)$ of G, and we see that $|S|, |V-S| \leq |V|$ and thus $\min(|S|, |V-S|) \geq \frac{|S||V-S|}{|V|}$. We also see that at least one of $|S|, |V-S| \leq |V|/2$. WLOG let it be $|S|$, so $|V-S|/|V| \geq 1/2$ and thus $\frac{|S||V-S|}{|V|} \geq |S|/2 = \frac{1}{2}\min(|S|, |V-S|)$. (it is easy to see that an analogous line of reasoning leads to the same result for $|V-S| \leq |V|/2$). This gives us

$$\min(|S|, |V-S|) \geq \frac{|S||V-S|}{|V|} \geq \frac{1}{2}\min(|S|, |V-S|)$$
$$\frac{|E(S, V-S)|}{d \min(|S|, |V-S|)} \leq \frac{|E(S, V-S)|}{\frac{d}{|V|}|S||V-S|} \leq \frac{1}{2}\frac{|E(S, V-S)|}{d \min(|S|, |V-S|)}$$
$$h(S) \leq \phi(S) \leq 2h(S)$$

Since for fixed $n = |V|$, the functions $\min(x, n-x)$ and $x(n-x)$ taken on $x \in (0, n)$ monotonically decrease together away from a shared maximum at $x = n/2$, the same cut must minimize both $h(G)$ and $\phi(G)$. Letting that cut be S gives $h(G) = h(S), \phi(G) = \phi(S)$ and thus $h(G) \leq \phi(G) \leq 2h(G)$. □

The Cheeger inequality gives us a relation between $h(G)$ and λ_2. We start by proving the easy direction, the lower bound for $h(G)$.

Proposition 3.2. *For a d-regular graph $G = (V, E)$ with $n = |V|$ and eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$ repeated according to their multiplicities we have*

$$1 - \lambda_2 \leq \phi(G)$$

Proof. We can compactly represent a subset, and thus a nontrivial cut S, by a bit vector $x \in \{0,1\}^n - \{\mathbf{0}, \mathbf{1}\}$ where $x_i = 1$ if $v_i \in S$ and $x_i = 0$ otherwise. Note that the removal of $\{\mathbf{0}, \mathbf{1}\}$ corresponds to nontriviality of cuts. Then $\{x_i, x_j\} \in E(S, V - S)$ iff $|x_i - x_j| = 1$, so we can write

$$\begin{aligned} |E(S, V - S)| &= \sum_{\{i,j\} \in E} |x_i - x_j| \\ &= \frac{1}{2} \sum_{i,j} A_{i,j} |x_i - x_j| \\ &= \frac{d}{2} \sum_{i,j} M_{i,j} |x_i - x_j| \end{aligned}$$

where the constant $1/2$ comes from every edge $\{u, v\} \in E$ being counted twice in the sum. Since $|S||V - S| = E'(S, V - S)$ for the complete graph K_n with adjacency matrix A' with diagonal entries 0 and all other entries 1, we get

$$\begin{aligned} |S||V - S| &= \frac{1}{2} \sum_{i,j} A'_{i,j} |x_i - x_j| \\ &= \frac{1}{2} \sum_{i,j} |x_i - x_j| \end{aligned}$$

Note that these expressions involving $|x_i - x_j|$ are actually closely related to our earlier work with λ_2 with inner products, since

$$\begin{aligned} \sum_{i,j} M_{ij} (x_i - x_j)^2 &= \sum_{i,j} M_{ij} (x_i^2 - 2x_i x_j + x_j^2) \\ &= \sum_i x_i^2 (\sum_j M_{ij} + \sum_j M_{ji}) - 2 \sum_{i,j} M_{ij} x_i x_j \\ &= 2x^T x - 2x^T M x \end{aligned}$$

and

$$\begin{aligned} \sum_{i,j} (x_i - x_j)^2 &= \sum_{i,j} (x_i^2 - 2x_i x_j + x_j^2) \\ &= 2n \sum_i x_i^2 - 2 \sum_{i,j} x_i x_j \\ &= 2n x^T x - 2(\sum_i x_i)^2 \end{aligned}$$

Note that when $x \perp \mathbf{1}$, we have $x^T\mathbf{1} = \sum_i x_i = 0$, and thus

$$\sum_{i,j}(x_i - x_j)^2 = 2nx^Tx$$
$$\frac{1}{n}\sum_{i,j}(x_i - x_j)^2 = 2x^Tx$$

In particular, with $x \neq \mathbf{0}$ and thus $x^Tx \neq 0$ we can write

$$\begin{aligned}
2x^TMx &= 2x^Tx - \sum_{i,j} M_{ij}(x_i - x_j)^2 \\
\frac{x^TMx}{x^Tx} &= 1 - \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{2x^Tx} \\
\max_{\substack{x\neq\mathbf{0}\\ x\perp\mathbf{1}}} \frac{x^TMx}{x^Tx} &= 1 - \min_{\substack{x\neq\mathbf{0}\\ x\perp\mathbf{1}}} \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{2x^Tx} \\
1 - \lambda_2 &= \min_{\substack{x\neq\mathbf{0}\\ x\perp\mathbf{1}}} \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{2x^Tx} \\
&= \min_{\substack{x\neq\mathbf{0}\\ x\perp\mathbf{1}}} \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i - x_j)^2}
\end{aligned}$$

From above we know $x \perp \mathbf{1}$ is equivalent to $\sum_i x_i = 0$, and we see that this expression is invariant under scaling or shifting x by a constant, as $x' = x + k\mathbf{1}$ satisfies $|x'_i - x'_j| = |x_i - x_j|$ for all i, j, and any scalar multiple can be factored out in both the numerator and the denominator. Thus, we can take the min over all $x \neq k\mathbf{1}, k \in \mathbb{Z}$ (which gives $\sum_{i,j}|x_i - x_j| = 0$) where for any x with $\sum_i x_i \neq 0$ we can instead evaluate the same expression over $x' = x - \frac{1}{n}\sum_i x'_i$ satisfying $\sum_i x'_i = 0$ and thus $x' \perp \mathbf{1}$. Thus we can write

$$1 - \lambda_2 = \min_{x\neq k\mathbf{1}, k\in\mathbb{Z}} \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i - x_j)^2}$$

and since $\mathbb{R}^n - \{k\mathbf{1}, k \in \mathbb{Z}\} \supseteq \{0,1\}^n - \{\mathbf{0}, \mathbf{1}\}$, we get

$$\begin{aligned}
1 - \lambda_2 &\leq \min_{x\in\{0,1\}^n - \{\mathbf{0},\mathbf{1}\}} \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i - x_j)^2} \\
&= \min_S \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i - x_j)^2} \\
&= \min_S \frac{\frac{1}{2}|E(S, V - S)|}{\frac{d}{2n}|S||V - S|} \\
&= \phi(G)
\end{aligned}$$

where the min is taken over all nontrivial cuts $(S, V - S)$ of G. $\square$

This gives us

Proposition 3.3 (Easy direction of Cheeger's inequality)**.** *For a d-regular graph $G = (V, E)$ with eigenvalues $\lambda_1 \geq ... \geq \lambda_n$ repeated according to their multiplicities we have*

$$\frac{1-\lambda_2}{2} \leq h(G)$$

Proof. By 3.2 and 3.1 we get $1 - \lambda_2 \leq \phi(G) \leq 2h(G)$. □

4 Difficult direction of Cheeger's inequality

To prove the other direction, we use the probabilistic method.

Proposition 4.1. *Let $G = (V, E)$ be d-regular with normalized adjacency matrix M and eigenvalues $\lambda_1 \geq ... \geq \lambda_n$ repeated according to their multiplicities. Also let $n = |V|, x \in \mathbb{R}^n$ such that $x \neq \mathbf{0}, x \perp \mathbf{1}$. Then there exists a nontrivial cut $(S, V - S)$ of G satisfying*

$$h(S) \leq \sqrt{2\delta}$$

where

$$\delta = \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i - x_j)^2}$$

Proof. Fix some arbitrary labeling of vertices by $v_1, ..., v_n$. Intuitively, d-regularity gives us some freedom with the choice of cuts, and it turns out it suffices to consider the $n-1$ cuts of the form $(S, V - S)$ where $S = \{v_1, ..., v_i\}, i < n$. In particular, we will determine S by the value of a real, continuous random variable $t \in [x_1, x_n]$ with probability density function $f(t) = 2|t|$. This gives

$$\mathbb{P}[a \leq t \leq b] = \int_a^b 2|t|dt$$

for $x_1 \leq a \leq b \leq x_n$. We see that for $a, b \geq 0$

$$\mathbb{P}[t \in [a, b]] = \int_a^b 2t = b^2 - a^2 = |a^2 - b^2|$$

for $a, b \leq 0$

$$\mathbb{P}[t \in [a, b]] = -\int_a^b 2t = -(b^2 - a^2) = |a^2 - b^2|$$

and for $a \leq 0, b \geq 0$

$$\mathbb{P}[t \in [a, b]] = -\int_a^0 2t + \int_0^b 2t = a^2 + b^2$$

so we conclude

$$\mathbb{P}[t \in [a,b]] = \begin{cases} |a^2 - b^2| & \text{if } a, b \text{ have the same sign} \\ a^2 + b^2 & \text{otherwise} \end{cases}$$

and we let $S = \{v_i : x_i \leq t\}$.

We can also make a few simplifying assumptions. Since we only look at sums and products of x_i for all i, WLOG reorder M and x such that $x_1 \leq ... \leq x_n$. Furthermore, in the proof for 3.2 we showed that δ is invariant under scaling or shifting x, so WLOG we can shift (i.e. by adding $-x_{\lfloor n/2 \rfloor}$) so that the entry $x_{\lfloor n/2 \rfloor} = 0$ and then scale (i.e. noting that $x \neq \mathbf{0}$ gives at least one of $x_1, x_n \neq 0$ so $x_1^2 + x_n^2 \neq 0$, by multiplying by $\frac{1}{\sqrt{x_1^2 + x_n^2}}$) so that $x_1^2 + x_n^2 = 1$.

Then we can apply linearity of expectation, although we have to be careful to not over-count in the case where $|S| = |V - S|$. Here, we choose to only count the contribution by members of $|S|$ in that case, so let X_i be the event that v_i in the strictly smaller subset or $|S| = |V - S|, v_i \in S$. Thus

$$\mathbb{E}[\min(|S|, |V - S)] = \sum_i \mathbb{E}[1_{X_i}] = \sum_i \mathbb{P}[X_i]$$

When n is odd, the special case $|S| = |V - S|$ is impossible. Note that for $i < n/2, v_i \in V - S$ implies $|V - S| > n/2$ and thus $V - S$ is the larger subset, so X_i is equivalent to $v_i \in S$ with $|S| < n/2$, which is equivalent to $t \in [x_i, 0)$. It is easy to see that the analogous statement holds, that for $i > n/2$, X_i is equivalent to $t \in (0, x_i]$. We see

$$\mathbb{P}[t \in [x_i, 0)] = \mathbb{P}[t \in (0, x_i]] = \mathbb{P}[t \in [0, x_i]] = x_i^2$$

since the distribution is continuous, so

$$\mathbb{E}[\min(|S|, |V - S)] = \sum_i x_i^2$$

When n is even, for $i \leq n/2$, X_i occurs when either $|V - S| > n/2$ and $v_i \in S$ from above, or $|S| = |V - S| = n/2$ and $v_i \in S$. Thus, X_i is equivalent to $v_i \in S$ with $|S| \leq n/2$, which is equivalent to $t \in [x_i, x_{n/2+1})$. It is easy to see that the analogous statement holds, that for $i > n/2$, X_i is equivalent to $t \in (x_{n/2+1}, x_i]$. We see

$$\begin{aligned} \mathbb{E}[\min(|S|, |V - S)] &= \sum_{i=1}^{n/2} \mathbb{P}[t \in [x_i, x_{n/2+1})] + \sum_{i=n/2+1}^{n} \mathbb{P}[t \in (x_{n/2+1}, x_i]] \\ &= \sum_{i=1}^{n/2} x_i^2 + x_{n/2+1}^2 + \sum_{i=n/2+1}^{n} x_i^2 - x_{n/2+1}^2 \\ &= \sum_i x_i^2 \end{aligned}$$

Next we see the why this distribution is especially useful, as it allows for an intuitive view of crossing edges: an edge is a crossing edge if t falls between the components of x that correspond to the vertices. In other words, $\{v_i, v_j\} \in E(S, V-S)$ with WLOG $i < j$ iff $t \in [x_i, x_j]$. This probability is either $|x_i^2 - x_j^2|$ or $x_i^2 + x_j^2$, which we can bound by

$$\mathbb{P}[\{v_i, v_j\} \in E(S, V-S)] \leq |x_i - x_j|(|x_i| + |x_j|)$$

since

$$\begin{aligned} |x_i^2 - x_j^2| &= |x_i - x_j||x_i + x_j| \leq |x_i - x_j|(|x_i| + |x_j|) \\ x_i^2 + x_j^2 &\leq (|x_i| + |x_j|)|x_i| + (|x_i| + |x_j|)|x_j| = |x_i - x_j|(|x_i| + |x_j|) \end{aligned}$$

where the first line follows from the Triangle Inequality and the second from x_i, x_j having different signs, which yields $|x_i - x_j| = |x_i| + |x_j| \geq |x_i|, |x_j|$.

We can substitute this into a larger expression for expectation using linearity of expectation and indicator variables to get

$$\begin{aligned} \mathbb{E}[\frac{1}{d}E(S, V-S)] &= \frac{1}{d}\mathbb{E}[\frac{1}{2}\sum_{i,j} A_{ij}\mathbf{1}_{\{v_i,v_j\}\in E(S,V-S)}] \\ &= \frac{1}{2}\sum_{i,j} M_{ij}\mathbb{E}[\mathbf{1}_{\{v_i,v_j\}\in E(S,V-S)}] \\ &= \frac{1}{2}\sum_{i,j} M_{ij}\mathbb{P}[\{v_i, v_j\} \in E(S, V-S)] \\ &\leq \frac{1}{2}\sum_{i,j} M_{ij}|x_i - x_j|(|x_i| + |x_j|) \\ &\leq \frac{1}{2}\sqrt{\sum_{i,j} M_{ij}(x_i - x_j)^2}\sqrt{\sum_{i,j} M_{ij}(|x_i| + |x_j|)^2} \end{aligned}$$

where at the last step we apply Cauchy-Schwarz using the Euclidean dot product on $a = \sum_{i,j}\sqrt{M_{ij}}(x_i - x_j), b = \sum_{i,j}\sqrt{M_{ij}}(|x_i| + |x_j|)$, getting $|\langle a, b\rangle| \leq ||a||||b||$.

Recall from our proof of 3.2 that when $x \perp \mathbf{1}, x \neq \mathbf{0}$, we have $\sum_{i,j}(x_i - x_j)^2 = 2n\sum_i x_i^2$. Rewriting our definition of δ, we get

$$\begin{aligned} \delta &= \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i - x_j)^2} \\ \sum_{i,j} M_{ij}(x_i - x_j)^2 &= \delta\frac{1}{n}\sum_{i,j}(x_i - x_j)^2 \\ &= 2\delta\sum_i x_i^2 \end{aligned}$$

Also note $2x_i^2 + 2x_j^2 - (|x_i| + |x_j|)^2 = (|x_i| - |x_j|)^2 \geq 0$ implies $2x_i^2 + 2x_j^2 \geq (|x_i| + |x_j|)^2$, so we get

$$\begin{aligned}\sum_{i,j} M_{ij}[|x_i| + |x_j|]^2 &\leq \sum_{i,j} M_{ij}(2x_i^2 + 2x_j^2)\\ &= 2\sum_i(\sum_j M_{ij}x_j^2 + \sum_j M_{ji}x_j^2)\\ &= 4\sum_i x_i^2\end{aligned}$$

by properties of M. We thus find that

$$\begin{aligned}\mathbb{E}[\frac{1}{d}E(S, V-S)] &\leq \frac{1}{2}\sqrt{\sum_{i,j} M_{ij}(x_i - x_j)^2}\sqrt{\sum_{i,j} M_{ij}[|x_i| + |x_j|]^2}\\ &\leq \frac{1}{2}\sqrt{2\delta\sum_i x_i^2}\sqrt{4\sum_i x_i^2}\\ &= \sqrt{2\delta}\sum_i x_i^2\end{aligned}$$

Combining this with our previous result for $\mathbb{E}[\min(|S|, |V-S)]$, we get

$$\begin{aligned}\frac{\mathbb{E}[\frac{1}{d}E(S, V-S)]}{\mathbb{E}[\min(|S|, |V-S)]} &\leq \frac{\sqrt{2\delta}\sum_i x_i^2}{\sum_i x_i^2} = \sqrt{2\delta}\\ \mathbb{E}[\frac{1}{d}E(S, V-S) - \sqrt{2\delta}\min(|S|, |V-S)] &\leq 0\end{aligned}$$

from linearity of expectation. Thus there must exist a nontrivial cut $(S, V-S)$ satisfying

$$\begin{aligned}\frac{1}{d}E(S, V-S) - \sqrt{2\delta}\min(|S|, |V-S) &\leq 0\\ h(S) \leq \frac{E(S, V-S)}{d\min(|S|, |V-S)} &\leq \sqrt{2\delta}\end{aligned}$$

□

Theorem 4.1 (Cheeger's inequality)**.** *For a d-regular graph $G = (V, E)$ with eigenvalues $\lambda_1 \geq ... \geq \lambda_n$ repeated according to their multiplicities we have*

$$\frac{1-\lambda_2}{2} \leq h(G) \leq \sqrt{2(1-\lambda_2)}$$

Proof. The left-hand inequality comes from 3.3. To get the right-hand inequality, we can apply 4.1 to an eigenvector $x = e_2$ of λ_2, which satisfies $x \perp \mathbf{1}, x \neq \mathbf{0}$. This yields

$h(G) \leq h(S) \leq \sqrt{2\delta}$ and in particular for $x = e_2$ we have

$$\begin{aligned}
\delta &= \frac{\sum_{i,j} M_{ij}(x_i - x_j)^2}{\frac{1}{n}\sum_{i,j}(x_i - x_j)^2} \\
&= \frac{2x^T x - 2x^T M x}{2x^T x} \\
&= \frac{2x^T x(1-\lambda_2)}{2x^T x} \\
&= 1 - \lambda_2
\end{aligned}$$

where the second step comes from equivalent expressions for the numerator and denominator expressions derived in the proof of 3.2. This gives the desired

$$h(G) \leq \sqrt{2(1-\lambda_2)}$$

□

Example 4.1. Cheeger's inequality allows us to approximate $h(G)$ for arbitrary G. Take the complete graph K_4 with

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

and from the respective normalized adjacency matrix $M = A/4$ we can compute the eigenvalue $\lambda_1 = 1$ with eigenvectors $(1,1,1,1)$ and the eigenvalue $\lambda_2 = \lambda_3 = \lambda_4 = -1/4$ with eigenvectors $(-1,0,0,1), (-1,0,1,0), (-1,1,0,0)$. K_4 is 3-regular, so substituting for λ_2, Cheeger's inequality gives us the approximation $1 \leq h(G) \leq 2$.

Example 4.2 (Tightness of Cheeger's inequality)**.** Recall that from our application of the Courant-Fischer Theorem in 2.1 we know

$$\lambda_2 \geq 1 + \frac{1}{P(n^2)}$$

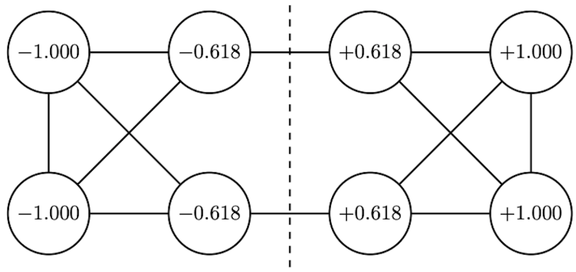
Any non-trivial cut $(S, V-S)$ has $E(S, V-S) = 2$, so the cut that minimizes $h(S)$ will maximize $\min(|S|, |V-S|)$ and thus have $|S| = \lfloor n/2 \rfloor$. This gives $h(G) = \frac{2}{2\lfloor n/2 \rfloor} \geq 2/n$. Note C_n is 2-regular, so Cheeger's inequality gives $\sqrt{2(1-\lambda_2)} \geq h(G) \geq 2/n$ and thus $\lambda_2 \leq 1 - 2/n^2 = 1 + \frac{1}{P(n^2)}$. Thus, the upper bound of Cheeger's inequality is asymptotically tight, and this tightness is achieved in cycles.

Example 4.3 (Sparsest cut problem)**.** Given a graph $G = (V, E)$, the problem of determining the cut $(S, V-S)$ that achieves the minimum $h(S) = h(G)$ has been shown to be NP-hard, meaning there is no known polynomial-time algorithm that solves it. ([2] presents a proof) Note the name of the problem actually alludes to the sparsity ϕ, since h, ϕ are sometimes used interchangeably as we have seen that they are related quantities minimized by the same cuts. However, the proof of the Cheeger inequality from 4.1 suggests a polynomial-time algorithm for finding a relatively small cut, where we guarantee $h \leq \sqrt{2(1-\lambda_2)}$:

1. Compute λ_2 and a corresponding eigenvector x
2. Label the vertices such that $x_1 \geq ... \geq x_n$
3. Try all cuts of the form $(S, V - S) = (\{v_1, ..., v_i\}, \{v_{i+1}, ..., v_n\})$ and return the one with smallest $h(S)$

which takes advantage of known polynomial-time algorithms for computing eigenvalues and eigenvectors of real symmetric matrices, which is the runtime-limiting step.

For instance, applying this algorithm to the graph from 1.1, we get $\lambda_2 = \sqrt{5}/3$ and



with the vertex labels representing the entries of x and the best cut represented by the dotted line (the computation is by [3]). We can verify that this cut has $h(S) = \frac{2}{3(4)} = \frac{1}{6}$ and that this satisfies $0.127 \approx \frac{1-\sqrt{5}/3}{2} \leq \frac{1}{6} \leq \sqrt{2(1-\sqrt{5}/3)} \approx 0.714$.

References

[1] Jiaqi Jiang. An introduction to spectral graph theory. `http://math.uchicago.edu/~may/REU2012/REUPapers/JiangJ.pdf`, May 2012.

[2] David W. Matula and Farhad Shahrokhi. Sparsest cuts and bottlenecks in graphs. *Discrete Applied Mathematics*, 27(1):113–123, 1990.

[3] Thomas Sauerwald and He Sun. Lecture 4 from spectral graph theory. `https://resources.mpi-inf.mpg.de/departments/d1/teaching/ws11/SGT/Lecture4.pdf`, Winter 2011.

[4] Luca Trevisan. Lectures 1-4 from cs359g: Graph partitioning and expanders. `http://theory.stanford.edu/~trevisan/cs359g/#notes`, January 2011.

[5] Anupam Gupta (Scribe) Umesh Vazirani, Satish Rao (Lecturers). Lecture 9 from cs270: Combinatorial algorithms and data structures. `https://people.eecs.berkeley.edu/~satishr/cs270/sp11/rough-notes/Sparse-cuts-spectra.pdf`, Spring 2017.