UChicago REU 2013 Apprentice Program **Summer 2013**

Lecture 10: July 15, 2013

Instructor: Madhur Tulsiani Scribe: David Kim

1 Eigenvalues of the adjacency matrix

All graphs in this lecture will be *undirected*.

Let $G = (V, E)$ be a graph with n vertices. Let A be the adjacency matrix of G. So A is a symmetric real matrix and therefore its eigenvalues are real; let $\mu_1 \geq \mu_2 \geq \dots \mu_n$ be these eigenvalues.

Exercise 1.1 $\sum_{i=1}^n \mu_i = 0$. *(Hint: trace* $\mathrm{Tr}(A)$*)*

Exercise 1.2 $\sum_{i=1}^n \mu_i^2 = 2|E|$. *(Hint:* $\mathrm{Tr}(A^2)$*)*

Exercise 1.3 $(\forall i)(|\mu_i| \leq \deg_{\max})$ *where* $\deg_{\max}$ *is the maximum degree.*

Exercise 1.4 $\mu_1 \geq \deg_{\text{avg}}$ *where* $\deg_{\text{avg}} = 2|E|/n$ *is the average degree. (Hint: Rayleigh quotient)*

Exercise 1.5 $(\forall i)(\mu_1 \geq |\mu_i|)$.

Exercise 1.6 *(a) If G is connected then $\mu_1 > \mu_2$.*
(b) Prove that the converse is false: $\mu_1 > \mu_2$ does not imply connectedness.

Exercise 1.7 *If G is bipartite then the adjacency spectrum is symmetric about the origin:*

$$(\forall i)(\mu_{n-i+1} = -\mu_i).$$

Exercise 1.8 *(a) If G is connected and $\mu_n = -\mu_1$ then G is bipartite.*
(b) Prove that the converse is false: without the connectedness assumption, $\mu_n = -\mu_1$ does not imply that the graph is bipartite.

2 The Laplacian

Let $G = (V, E)$ be a graph with n vertices. Let A be the adjacency matrix of G. Let D be the diagonal matrix with $D_{ii} = \deg(i)$.

Definition 2.1 (The Laplacian) *Given G as above, the Laplacian associated with G is the symmetric $n \times n$ matrix $L := D - A$.*

Next we take a look at the Laplacian quadratic form. Let $x \in \mathbb{R}^n$ be a function on the vertices of $G = (V, E)$. Then

$$\begin{aligned} x^T L x &= x^T (D - A) x \\ &= x^T D x - x^T A x \\ &= \sum_{i=1}^{n} \deg(i) x_i^2 - \sum_{\{i,j\} \in E} 2 x_i x_j \\ &= \sum_{i=1}^{n} \sum_{\{i,j\} \in E} x_i^2 - \sum_{\{i,j\} \in E} 2 x_i x_j \\ &= \sum_{\{i,j\} \in E} (x_i^2 + x_j^2 - 2 x_i x_j) \\ &= \sum_{\{i,j\} \in E} (x_i - x_j)^2 \end{aligned}$$

So the Laplacian is positive semidefinite and therefore its eigenvalues are nonnegative. In fact, the smallest eigenvalue is 0, as any non-zero constant vector is an eigenvector of eigenvalue 0. Thus the eigenvalues of the Laplacian (the "Laplacian eigenvalues") of a graph G can be ordered as follows:

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$$

Exercise 2.2 *A graph is connected if and only if its second smallest Laplacian eigenvalue is positive:* $\lambda_2 > 0$.

Definition 2.3 (The Normalized Laplacian) *The normalized Laplacian of a graph* G *is* $N := D^{-1/2} L D^{-1/2}$, *where* D *and* L *are as above.*

Note that we have

$$\begin{aligned} N &= D^{-1/2} L D^{-1/2} \\ &= D^{-1/2} (D - A) D^{-1/2} \\ &= I - D^{-1/2} A D^{-1/2} \end{aligned}$$

Now consider a d-regular graph G. Then

$$L = D - A = dI - A$$

and if we let $\mu_1 \geq \cdots \geq \mu_n$ be the eigenvalues of A, then $\mu_i = d - \lambda_i$ for $i = 1, \ldots, n$ as $L v_i = \lambda_i v_i \Leftrightarrow A v_i = (d - \lambda_i) v_i$. So we have the following spectrum for A:

$$d = \mu_1 \geq \cdots \geq \mu_n \geq -d$$

The first inequality follows from $\lambda_1 = 0$ and $\mu_1 = d - \lambda_1$ (and alternatively from $\deg_{\max} = \deg_{\text{avg}} = d$, by Exx. 1.3 and 1.4). The last inequality follows from Ex. 1.3.

For the normalized Laplacian, we have in the d-regular case:

$$N = I - \frac{1}{d}A$$

If we let $\nu_1 \geq \cdots \geq \nu_n$ be the eigenvalues of N, then $\nu_i = 1 - \frac{1}{d}\mu_i = \frac{\lambda_i}{d}$ for $i = 1, \ldots, n$. (λ_i and μ_i are the eigenvalues of L and A, respectively)

3 Conductance of a Graph

A *cut* in the graph $G = (V, E)$ is a partition of the vertices into $(S, \overline{S})$ where $S \subseteq V$ and $\overline{S} = V \setminus S$. We write $E(S, \overline{S})$ to denote the set of edges connecting a vertex in S to a vertex in $\overline{S}$ (the edges "leaving S").

The *volume* of $S \subseteq V$ is $\text{vol}(S) = \sum_{i \in S} \deg(i)$. The volume counts the pairs (v, e) where $v \in S, e \in E$ and v, e are incident. Note that if G is d-regular then $\text{vol}(S) = d|S|$.

Definition 3.1 (Edge Expansion of a subset) *The* edge expansion *of a subset $S \subseteq V$ is*

$$\varphi(S) := \frac{|E(S, \overline{S})|}{\text{vol}(S)}.$$

This is the proportion of the indicences (v, e) where e leaves S among all incidences counted by the volume.

Definition 3.2 (Conductance of a graph) *The conductance of a graph G is*

$$\Phi_G := \min_{0 < \text{vol}(S) \leq |E|} \varphi(S)$$

The conductance is also called the *edge expansion* of G.

Assume G is d-regular. In this case the minimum in the definition of the conductance is over all subsets $S \subseteq V$ such that $0 < |S| \leq n/2$.

If we represent $S \subseteq V$ by its incidence vector $x \in \{0, 1\}^n$ then

$$\varphi(S) = \frac{|E(S, \overline{S})|}{d|S|} = \frac{\frac{1}{2}\sum_{i,j} A_{i,j}|x_i - x_j|}{d\sum_i x_i^2} = \frac{\sum_{\{i,j\} \in E}(x_i - x_j)^2}{d\sum x_i^2}$$

and Φ_G is obtained by $x \in \{0, 1\}^n, x \neq 0, \mathbb{1}$ ($|S| \leq n/2$) which minimizes the above expression. But this is exactly the expression of the Rayleigh quotient, $R_N(x)$, associated with the normalized Laplacian of G. Thus, the eigenvalues of the Laplacian, optimized over $x \in \mathbb{R}^n, x \neq 0$, can be seen as a continuous *relaxation* of the conductance.

Cheeger's Inequality bounds the conductance of a graph in terms of the second smallest eigenvalue of the normalized Laplacian.

4 Cheeger's Inequality

Let λ_i denote the i-th smallest eigenvalue of the *normalized Laplacian* of the graph G:

$$0 = \lambda_1 \leq \cdots \leq \lambda_n \leq 2$$

Theorem 4.1 (Cheeger's Inequalities)

$$\frac{\lambda_2}{2} \leq \Phi_G \leq \sqrt{2\lambda_2}$$

4.1 Proof of the First Inequality

We first prove the easy direction: $\frac{\lambda_2}{2} \leq \Phi_G$ We give the proof for the case when G is d-regular.

Proof: Let $(S, \overline{S})$ be a cut such that $|S| \leq \frac{n}{2}$.
As before, N denotes the normalized Laplacian and $R_N(x)$ its Rayleigh quotient.

Claim 4.2 $\exists x \in \mathbb{R}^n, x \neq 0, x \perp \mathbb{1}$, *such that* $R_N(x) \leq 2 \cdot \varphi(S)$.

$\rightarrow$ Define $x \in \mathbb{R}^n$ as follows:

$$x_i = \begin{cases} \dfrac{1}{|S|} & \text{if } i \in S \\ -\dfrac{1}{|\overline{S}|} & \text{if } i \in \overline{S} \end{cases}$$

Then $x \perp \mathbb{1}$ since

$$\mathbb{1}^T x = \sum_{i=1}^{n} x_i = |S| \cdot \frac{1}{|S|} + |\overline{S}| \cdot \left(-\frac{1}{|\overline{S}|}\right) = 0$$

Also, the Rayleigh quotient evaluates to

$$R_N(x) = \frac{x^T N x}{x^T x} = \frac{1}{d} \cdot \frac{\sum_{\{i,j\} \in E} (x_i - x_j)^2}{x^T x} = \frac{1}{d} \cdot \frac{|E(S, \overline{S})| \cdot \left(1/|S| + 1/|\overline{S}|\right)}{|S| \cdot (1/|S|^2) + |\overline{S}| \cdot (1/|\overline{S}|^2)}$$

as only the edges going across the cut $(S, \overline{S})$ contribute to the sum in the numerator. Thus we have

$$R_N(x) = \frac{|E(S, \overline{S})|}{d|S|} \cdot \frac{|S| + |\overline{S}|}{|\overline{S}|} \leq 2 \cdot \varphi(S)$$

since $|S| \leq n/2$. □
The claim also holds for S and the corresponding x for which $\varphi(S) = \Phi_G$. Thus,

$$\lambda_2 = \min_{v \perp \mathbb{1}} \frac{v^T N v}{v^T v} \leq \frac{x^T N x}{x^T x} \leq 2 \cdot \varphi(S) = 2 \cdot \Phi_G$$

■

4.2 Proof of the Second Inequality

For the other direction of the inequality, we will prove a slightly weaker bound: $\Phi_G \leq 2\sqrt{\lambda_2}$. Again, we give the proof for the case when G is d-regular.

Proof: First we prove the following main lemma:

Lemma 4.3 (Spectral Partitioning) *Given $x \in \mathbb{R}^n$ such that $x \neq 0, x \perp \mathbb{1}$, we can find $|S| \leq n/2$, such that $\varphi(S) \leq 2\sqrt{R_N(x)}$.*

Our proof is algorithmic; it is a linear time algorithm known as *spectral partitioning.*

Observation 1. Multiplying x by a constant $c \neq 0$ does not change the Rayleigh quotient, i.e., $R_N(x) = R_N(cx)$.

Observation 2. Adding $c\mathbb{1}$ to x decreases the Rayleigh quotient.
(proof) Let $y = x + c\mathbb{1}$. Then $R_N(y) \leq R_N(x)$, since $y^T N y = x^T N x$ (the numerator stays same), while $y^T y = (x + c\mathbb{1})^T (x + c\mathbb{1}) = x^T x + ||c\mathbb{1}||^2$, since $x \perp \mathbb{1}$ (the denominator increases). □

Now shift and scale x to get $z \in [-1, 1]^n$ such that $z_{\lceil \frac{n}{2} \rceil} = 0$, i.e., the median is at 0 (think of $z \in \mathbb{R}^n$ as points, $z_1 \leq \cdots \leq z_n$, on the real line). We can then obtain a cut $(S, \overline{S})$ by cutting the line, say at $t \in \mathbb{R}$, and taking $S = \{1, \ldots, i\}$ for all $z_i \leq t$. Note that in this way, there are at most n cuts possible.

Claim 4.4 *One of these cuts is good enough for us.*

This is a linear-time algorithm, since the algorithm calculates only these cuts, and given the value of the cut for $(S, \overline{S})$ where $S = \{1, \ldots, i\}$, to find the value of the next cut, $S = \{1, \ldots, i+1\}$, one only needs to look at the edges incident to vertex $i+1$; hence each edge will be looked at only twice.

So given $z \in [-1, 1]^n$, we use a probabilistic argument to show that one of these cuts is good enough:

- choose $t \in [0, 1]$ uniformly at random
- let $S = \{i : z_i^2 \geq t\}$.

We use $\mathbb{P}[A]$ to denote the probability of event A and $\mathbb{E}[X]$ to denote the expected value of the random variable X.
Now $\mathbb{P}[i \in S] = z_i^2$. We will show that

$$\frac{\mathbb{E}[|E(S, \overline{S})|]}{d \cdot \mathbb{E}[|S|]} \leq \sqrt{2R_N(z)}$$

in other words, $\exists t \in [0, 1]$ s.t. $S = \{i : z_i^2 \geq t\}$ gives $\varphi(S) \leq \sqrt{2R_N(z)}$.

Let $z^{(1)}$ denote the positive part of $z \in \mathbb{R}^n$ and $z^{(2)}$ denote negative part of z, defined by

$$z_i^{(1)} = \max\{z_i, 0\}$$
$$z_i^{(2)} = \min\{z_i, 0\}.$$

We observe that the set $S = \{i : z_i^2 \geq t\}$ for both $z^{(1)}$ and $z^{(2)}$ will be one of the n cuts seen by the algorithm, and we can show that either $R_N(z^{(1)})$ or $R_N(z^{(2)}) \leq 2R_N(z)$. Applying the results to $z^{(1)}$ or $z^{(2)}$ proves the lemma.

Claim 4.5 *At least one of $R_N(z^{(1)})$ and $R_N(z^{(2)})$ must be $\leq 2R_N(z)$.*

Observation 3. $(z^{(1)})^T(z^{(1)}) + (z^{(2)})^T(z^{(2)}) = z^T z$.

Observation 4. $(z^{(1)})^T N(z^{(1)}) \leq z^T N z$, $(z^{(2)})^T N(z^{(2)}) \leq z^T N z$.
(proof) We have

$$(z^{(1)})^T N(z^{(1)}) = \frac{1}{d} \cdot \sum_{\{i,j\} \in E} (z_i^{(1)} - z_j^{(1)})^2$$
$$z^T N z = \frac{1}{d} \cdot \sum_{\{i,j\} \in E} (z_i - z_j)^2$$

If both $z_i, z_j \geq 0$, then $(z_i - z_j)^2 = (z_i^{(1)} - z_j^{(1)})^2$. If both $z_i, z_j < 0$, then $(z_i^{(1)} - z_j^{(1)})^2 = 0$. Otherwise, z_i and z_j have opposite signs, and one of $z_i^{(1)}$ or $z_j^{(1)}$ is 0, hence $(z_i - z_j)^2 \geq (z_i^{(1)} - z_j^{(1)})^2$. An analogous argument gives $(z^{(2)})^T N(z^{(2)}) \leq z^T N z$. □

By Observation 3, one of $(z^{(1)})^T(z^{(1)})$ and $(z^{(2)})^T(z^{(2)})$ must be $\geq \frac{z^T z}{2}$. Therefore, $R_N(z^{(1)})$ or $R_N(z^{(2)})$ must be $\leq 2R_N(z)$.

Claim 4.6 *Given $z \in [-1, 1]^n$, let us choose $t \in [0, 1]$ uniformly at random. Let $S = \{i : z_i^2 \geq t\}$. Then*

$$\frac{\mathbb{E}[|E(S, \overline{S})|]}{d \cdot \mathbb{E}[|S|]} \leq \sqrt{2R_N(z)}$$

Once this has been verified, we conclude that $\mathbb{E}\left[|E(S, \overline{S})| - d\sqrt{2R_N(z)}|S|\right] \leq 0$ and therefore $\exists t \in [0, 1]$ s.t. $\dfrac{|E(S, \overline{S})|}{d|S|} \leq \sqrt{2R_N(z)}$.
(proof) For $i = 1, \ldots, n$, define Y_i s.t.

$$Y_i := \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{o.w.} \end{cases}$$

Then we have

$$\mathbb{E}[|S|] = \sum_{i=1}^{n} \mathbb{E}[Y_i] = \sum_{i=1}^{n} z_i^2$$

and also

$$\begin{aligned}
\mathbb{E}[|E(S,\overline{S})|] &= \sum_{\{i,j\}\in E} |z_i^2 - z_j^2| \\
&= \sum_{\{i,j\}\in E} |z_i - z_j||z_i + z_j| \\
&\le \sqrt{\sum_{\{i,j\}\in E} |z_i - z_j|^2}\sqrt{\sum_{\{i,j\}\in E} |z_i + z_j|^2} \\
&\le \sqrt{\sum_{\{i,j\}\in E} |z_i - z_j|^2}\sqrt{\sum_{\{i,j\}\in E} (2z_i^2 + 2z_j^2)} \\
&= \sqrt{\sum_{\{i,j\}\in E} |z_i - z_j|^2}\sqrt{2d\sum_{i} z_i^2}
\end{aligned}$$

where the first inequality follows from Cauchy-Schwarz, the second from $z_i^2 + z_j^2 \ge 2|z_i z_j|$. Therefore, we have

$$\frac{\mathbb{E}[|E(S,\overline{S})|]}{d\mathbb{E}[|S|]} \le \sqrt{2R(z)}.$$

This completes the proof of □

Now apply Claim 3.6 to $z^{(1)}$ and $z^{(2)}$; by Claim 3.5, one of them will show the existence of a good cut satisfying Claim 3.3 and Claim 3.4. (Note that at least one of $z^{(1)}$ and $z^{(2)}$ is nonzero. Also, having set the median to 0, we do not need to worry about $|S| \le n/2$.) That is, assuming WLOG that $R_N(z^{(1)}) \le 2R_N(z)$, $\exists S = \{i : (z_i^{(1)})^2 \ge t\}$

$$\varphi(S) \le \sqrt{2R_N(z^{(1)})} \le 2\sqrt{R_N(z)} \le 2\sqrt{R_N(x)}$$

To finish, apply the main lemma (spectral partitioning) to the second eigenvector v_2 of the second smallest eigenvalue λ_2 of the normalized Laplacian. Then the output S will satisfy:

$$\Phi_G \le \varphi(S) \le 2\sqrt{R_N(v_2)} = 2\sqrt{\lambda_2}.$$

■

4.3 Proof of the Tight Bound

We are getting the 2 outside the root in $2\sqrt{\lambda_2}$; we will try to get it inside now.

To prove $\Phi_G \le \sqrt{2\lambda_2}$, we prove the following lemma:

Lemma 4.7 *Given $x \perp \mathbb{1}$, $\exists S \subseteq V$ s.t. $|S| \le n/2$ and $\varphi(S) \le \sqrt{2R_N(x)}$.*

Proof: Again, shift and scale x to z s.t. $z_{\lceil n/2 \rceil} = 0$ and $z_1^2 + z_n^2 = 1$. Choose $t \in [z_1, z_n]$, where t is a random variable with density $2|t|$, defined only between z_1 and z_n.

Let $S = \{i : z_i \le t\}$. Then for $a \ge z_1$ and $b \le z_n$,

$$\mathbb{P}\left[t \in [a,b]\right] = \int_a^b 2|r|dr = \operatorname{sgn}(b) \cdot b^2 - \operatorname{sgn}(a) \cdot a^2$$

Claim 4.8 $\dfrac{\mathbb{E}[|E(S,\overline{S})|]}{d\mathbb{E}[\min(|S|,|\overline{S}|)]} \le \sqrt{2R(z)}$

(pf.) For $i = 1, \ldots, n$ let Y_i be such that

$$Y_i = \begin{cases} 1 & \text{if } (i \in S \text{ and } |S| \le n/2) \text{ OR } (i \in \overline{S} \text{ and } |\overline{S}| \le n/2) \\ 0 & \text{o.w.} \end{cases}$$

Then we have

$$\begin{aligned} \mathbb{E}[\min(|S|,|\overline{S}|)] &= \sum_i \mathbb{E}[Y_i] = \sum_i z_i^2 \\ \mathbb{E}[|E(S,\overline{S})|] &= \sum_{\{i,j\} \in E} \mathbb{P}\left[t \in [z_i, z_j)\right] \\ &= \sum_{\{i,j\} \in E} |\operatorname{sgn}(z_j) \cdot z_j^2 - \operatorname{sgn}(z_i) \cdot z_i^2| \\ &\le \sum_{\{i,j\} \in E} |z_j - z_i|(|z_j| + |z_i|) \end{aligned}$$

One can easily check $|\operatorname{sgn}(z_j) \cdot z_j^2 - \operatorname{sgn}(z_i) \cdot z_i^2| \le |z_j - z_i|(|z_j| + |z_i|)$ holds for all cases of the signs. The rest of the proof is same as the previous proof using Cauchy-Schwarz. □

Having established the lemma, using v_2, the eigenvector to the second eigenvalue λ_2 of the normalized Laplacian, the set S of the above analysis immediately satisfies

$$\Phi_G \le \varphi(S) \le \sqrt{2R_N(v_2)} = \sqrt{2\lambda_2}$$

■

5 Some Generalizations of Cheeger's Inequality

Here we found sets S_1, S_2 s.t. $S_2 = \overline{S_1}$, $\varphi(S_1) \le \sqrt{2\lambda_2}, \varphi(S_2) \le \sqrt{2\lambda_2}$. We can also find sets $S_1, \ldots, S_k$ that are disjoint s.t. $\varphi(S_i) \le O(k^2)\sqrt{\lambda_k}$ for each $i = 1, \ldots, k$. More recent results also show that we can find $S_1, \ldots, S_{\frac{k}{2}}$ s.t. $\varphi(S_i) \le O(\sqrt{\lambda_k \log k})$ (cannot do better than $\log k$).

Look up: Dirichlet Boundary Conditions