

Lecture 2: Cheeger Inequality

In which we generalize the notion of normalized Laplacian to irregular graphs, we extend the basic spectral graph theory results from last lecture to irregular graphs, and we prove the easy direction of Cheeger's inequalities.

1 Irregular Graphs

Let $G = (V, E)$ be an undirected graph, not necessarily regular. We will assume that every vertex has non-zero degree. We would like to define a normalized Laplacian matrix associated to G so that the properties we proved last time are true: that the multiplicity of 0 as an eigenvalue is equal to the number of connected components of G , that the largest eigenvalue is at most 2, and that it is 2 if and only if (a connected component of) the graph is bipartite.

In order to have a matrix such that zero is the smallest eigenvalue, and such that multiplicity of zero is the number of connected component, we want a matrix such that the numerator of the Rayleigh quotient is (a multiple of)

$$\sum_{\{u,v\} \in E} (x_u - x_v)^2$$

and the matrix M such that $\mathbf{x}^T M \mathbf{x}$ is the above expression is the matrix $M = D - A$, where D is the diagonal matrix such that $D_{v,v} = d_v$, the degree of v . The matrix $D - A$ is called the *Laplacian* matrix of G . Note that there is no fixed constant upper bound to the largest eigenvalue of $D - A$; for example, if G is a d -regular bipartite graph, the largest eigenvalue is $2d$, as proved in the last lecture.

Some calculations shows that the right analog of the normalization that we did in the regular case (in which we divided by the degree d) would be to have a matrix whose Rayleigh quotient is

$$\frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{\sum_v d_v x_v^2} = 2 - \frac{\sum_{\{u,v\} \in E} (x_u + x_v)^2}{\sum_v d_v x_v^2} \quad (1)$$

and it's clear that the above expression is at most 2 for every \mathbf{x} , and it is possible to find an \mathbf{x} for which the above expression is 2 if and only if G has a bipartite connected component.

Unfortunately, there is no matrix whose Rayleigh quotient equals (??), because the denominator of a Rayleigh quotient is, by definition, $\sum_v x_v^2$ regardless of the matrix.

One way to work around this problem would be to give a more general form of the variational characterization of eigenvalues, in which we have an arbitrary inner product $\langle \cdot, \cdot \rangle$, and the Rayleigh quotient is defined as $\frac{\langle \mathbf{x}, M\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle}$.

Here we will proceed in a way that is essentially equivalent, but without introducing this additional definitional framework.

The point is that, if we look at the Rayleigh quotient of the vector $D^{1/2}\mathbf{x}$, where $D^{1/2}$ is the diagonal matrix such that $D_{v,v}^{1/2} = \sqrt{d_v}$, then the denominator will indeed be $\mathbf{x}^T D \mathbf{x} = \sum_v d_v x_v^2$, and that we can find a matrix L such that the numerator of the Rayleigh quotient of $D^{1/2}\mathbf{x}$ is $\mathbf{x}^T D^{1/2} L D^{1/2} \mathbf{x} = \sum_{\{u,v\} \in E} (x_u - x_v)^2$, so that the Rayleigh quotient $R_L(D^{1/2}\mathbf{x})$ is indeed the expression in (??).

This matrix L is called the *normalized Laplacian* of G and, by the above observation, it has to be $L = D^{-1/2}(D - A)D^{-1/2} = I - D^{-1/2}AD^{-1/2}$. Note that, in a d -regular graph, we get $L = I - \frac{1}{d}A$, consistent with our definition from the last lecture.

Now the point is that the mapping $\mathbf{x} \rightarrow D^{1/2}\mathbf{x}$ is linear and bijective, so it maps the set of all possible vectors to the set of all possible vectors, and it maps a k -dimensional space to a (possibly different) k -dimensional space.

If we let $\lambda_1 \leq \dots \leq \lambda_n$ be the eigenvalues of $L = I - D^{-1/2}AD^{-1/2}$, counting repetitions, the variational characterization gives us

$$\lambda_1 = \min_{\mathbf{x} \neq \mathbf{0}} R_L(\mathbf{x}) = \min_{\mathbf{x} \neq \mathbf{0}} R_L(D^{1/2}\mathbf{x}) = \min_{\mathbf{x} \neq \mathbf{0}} \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{\sum_v d_v x_v^2}$$

and

$$\begin{aligned} \lambda_k &= \min_{X \text{ } k\text{-dimensional}} \max_{\mathbf{x} \in X - \{\mathbf{0}\}} R_L(\mathbf{x}) \\ &= \min_{X \text{ } k\text{-dimensional}} \max_{\mathbf{x} \in X - \{\mathbf{0}\}} R_L(D^{1/2}\mathbf{x}) \\ &= \min_{X \text{ } k\text{-dimensional}} \max_{\mathbf{x} \in X - \{\mathbf{0}\}} \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{\sum_v d_v x_v^2} \end{aligned}$$

from which we have that $\lambda_1 = 0$ and that the multiplicity of zero is equal to the number of connected components.

We also have

$$\lambda_n = \max_{\mathbf{x} \neq \mathbf{0}} R_L(\mathbf{x}) = \max_{\mathbf{x} \neq \mathbf{0}} R_L(D^{1/2}\mathbf{x})$$

$$= 2 - \min_{\mathbf{x} \neq \mathbf{0}} \frac{\sum_{\{u,v\} \in E} (x_u + x_v)^2}{\sum_v d_v x_v^2}$$

from which we see that $\lambda_n \leq 2$ and that $\lambda_n = 2$ if and only if one of the connected components of G is bipartite.

2 Edge Expansion, Fiedler's Algorithm and Cheeger's Inequalities

We will now return, for simplicity, to the regular case.

If $G = (V, E)$ is an undirected d -regular graph, and $S \subseteq V$ is a set of vertices, we call the quantity

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

the *edge expansion* of S . The quantity $\phi(S)$ is the average fraction of neighbors outside of S for a random element of S , and it compares the actual number of edges crossing the cut $(S, V - S)$ with the trivial upper bound $d|S|$.

We define the expansion of a cut $(S, V - S)$ as

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

The edge expansion of the graph G is defined as

$$\phi(G) := \min_S \phi(S, V - S) = \min_{S: 1 \leq |S| \leq \frac{|V|}{2}} \phi(S)$$

(Note: it is common in the literature to use the notation $\phi(S)$ to refer to the quantity that we call $\phi(S, V - S)$.)

Finding cuts of small expansion is a problem with several applications. It is an open question if there is a polynomial-time approximation with a constant-factor approximation ratio; a positive answer would refute the “small-set expansion conjecture” which is closely related to the unique games conjecture.

The following algorithm was proposed by Fiedler, and it works well in practice when \mathbf{x} is the eigenvector of λ_2 .

- Input: graph $G = (V, E)$, vector $\mathbf{x} \in \mathbb{R}^V$
 - Sort the vertices according the values x_v , and let v_1, \dots, v_n be the sorted order

- Find a k that minimizes $\phi(\{v_1, \dots, v_k\}, \{v_{k+1}, \dots, v_n\})$, and output such a cut

Note that Fiedler's algorithm can be implemented in time $O(|E| + |V| \log |V|)$, because it takes time $O(|V| \log |V|)$ to sort the vertices, and the cut of minimal expansion that respects the sorted order can be found in time $O(E)$. (To see that this is the case, consider that, in order to find such a cut, we just need to compute the numbers $e_k := E(\{v_1, \dots, v_k\}, \{v_{k+1}, \dots, v_n\})$ for each $k = 1, \dots, n-1$. We see that e_1 is equal to the degree of v_1 , and that, given e_{k-1} , the value of e_k can be computed by just adding to e_{k-1} the number of neighbors of v_k in $\{v_{k+1}, \dots, v_n\}$, and subtracting the number of neighbors of v_k in $\{v_1, \dots, v_k\}$, on operation that can be done in time $O(d_{v_k})$. Thus the total running time is of the order of $\sum_v d_v$, that is, $O(|E|)$.)

We will prove the following result

Theorem 1 (Cheeger's Inequalities) *Let G be an undirected regular graph and $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be the eigenvalues of the normalized Laplacian, with repetitions, then*

$$\frac{\lambda_2}{2} \leq \phi(G) \leq \sqrt{2\lambda_2}$$

Furthermore, if $(S, V-S)$ is the cut found by Fiedler's algorithm given the eigenvector of λ_2 , then

$$\phi(S, V-S) \leq \sqrt{2\lambda_2}$$

Note that, from the *furthermore* part of the Theorem, it follows that, if $(S, V-S)$ is the cut found by Fiedler's algorithm given an eigenvector of λ_2 , we have

$$\phi(S, V-S) \leq 2\sqrt{\phi(G)}$$

which is a worst-case guarantee of the quality of the solution found by the algorithm.

3 Proof that $\frac{\lambda_2}{2} \leq \phi(G)$

Let S be a set of vertices such that $\phi(S, V-S) = \phi(G)$. Recall that for every set S , we have that expansion of S is the same as the Rayleigh quotient of the indicator vector $\mathbf{1}_S$. (The indicator vector of a set S is the 0/1 vector $\mathbf{1}_S$ whose v -th coordinate is 1 if and only if $v \in S$.) So we have

$$R_L(\mathbf{1}_S) \leq \phi(G)$$

$$R_L(\mathbf{1}_{V-S}) \leq \phi(G)$$

also recall that, from the variational characterization of eigenvalues, we have

$$\lambda_2 = \min_{X \text{ 2-dimensional}} \max_{\mathbf{x} \in X - \{\mathbf{0}\}} R_L(\mathbf{x})$$

We will prove the inequality $\lambda_2 \leq 2\phi(G)$ by showing that all the vectors in the 2-dimensional space X of linear combinations of the orthogonal vectors $\mathbf{1}_S, \mathbf{1}_{V-S}$ have Rayleigh quotient at most $2\phi(G)$. This is a consequence of the following useful fact.

Lemma 2 *Let \mathbf{x} and \mathbf{y} be two orthogonal vectors, and let M be a positive semidefinite matrix. Then*

$$R_M(\mathbf{x} + \mathbf{y}) \leq 2 \cdot \max\{R_M(\mathbf{x}), R_M(\mathbf{y})\}$$

PROOF: Let $0 \leq \lambda_1 \leq \dots \leq \lambda_n$ be the eigenvalues of M and $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a corresponding basis of eigenvectors. Let us write $\mathbf{x} = \sum_i a_i \mathbf{v}_i$ and $\mathbf{y} = \sum_i b_i \mathbf{v}_i$.

The Rayleigh quotient of $\mathbf{x} + \mathbf{y}$ is

$$\begin{aligned} \frac{\sum_i \lambda_i (a_i + b_i)^2}{\|\mathbf{x} + \mathbf{y}\|^2} &\leq \frac{\sum_i \lambda_i 2(a_i^2 + b_i^2)}{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2} \\ &= \frac{2R_M(\mathbf{x}) \cdot \|\mathbf{x}\|^2 + 2R_M(\mathbf{y}) \cdot \|\mathbf{y}\|^2}{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2} \leq 2 \max\{R_M(\mathbf{x}), R_M(\mathbf{y})\} \end{aligned}$$

In the first inequality, we used orthogonality of \mathbf{x} and \mathbf{y} to derive $\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ and we used the Cauchy-Schwarz inequality $(a + b)^2 \leq 2a^2 + 2b^2$. \square

4 First Part of the Analysis of Fiedler's Algorithm

The vector $\mathbf{1} = (1, \dots, 1)$ is an eigenvector for 0, which is the smallest eigenvalue of the normalized Laplacian of G , and so, from the variational characterization of eigenvalues, we have that

$$\lambda_2 = \min_{\mathbf{x} \perp \mathbf{1}} \frac{\sum_{\{u,v\} \in E} (x_u - x_v)^2}{d \sum_v x_v^2}$$

and that any eigenvector \mathbf{x} of λ_2 is a minimizer of the above expression. We will prove that $\phi(G) \leq \sqrt{2\lambda_2}$ and that the *Furthermore* part of Theorem ?? is true by showing the following stronger result:

Lemma 3 *Let \mathbf{x} be a vector orthogonal to $\mathbf{1}$ and let $(S, V - S)$ be the cut found by Fiedler's algorithm given \mathbf{x} . Then*

$$\phi(S, V - S) \leq \sqrt{2R_L(\mathbf{x})}$$

This stronger form is useful, because often one runs Fiedler's algorithm on an approximate eigenvector, and Lemma ?? shows that one gets a guarantee on the quality of the resulting cut that does not require \mathbf{x} to be an eigenvector, as long as its Rayleigh quotient is small.

We divide the proof of Lemma ?? into two parts: we analyze the performance of the algorithm given a vector \mathbf{x} that, instead of being orthogonal to $\mathbf{1}$, has the property of having non-negative entries and at most $|V|/2$ non-zero entries, and we show that analyzing the performance of the algorithm on vectors of the former type reduces to analyzing the performance on vectors of the latter type.

Lemma 4 *Let $\mathbf{y} \in \mathbb{R}_{\geq 0}^V$ be a vector with non-negative entries. Then there is a $0 < t \leq \max_v \{y_v\}$ such that*

$$\phi(\{v : y_v \geq t\}) \leq \sqrt{2R_L(\mathbf{y})}$$

Lemma 5 *Let $\mathbf{x} \in \mathbb{R}^V$ be orthogonal to $\mathbf{1}$. Then there is a vector $\mathbf{y} \in \mathbb{R}_{\geq 0}^V$ with at most $|V|/2$ non-zero entries such that*

$$R_L(\mathbf{y}) \leq R_L(\mathbf{x})$$

Furthermore, for every $0 < t \leq \max_v \{y_v\}$, the cut $(\{v : y_v \geq t\}, \{v : y_v < t\})$ is one of the cuts considered by Fiedler's algorithm on input \mathbf{x} .

Let us quickly see how to prove Lemma ?? given Lemma ?? and Lemma ??. Let \mathbf{x} be a vector orthogonal to $\mathbf{1}$, and let $(S_F, V - S_F)$ be the cut found by Fiedler's algorithm given \mathbf{x} . Let \mathbf{y} be the non-negative vector with at most $|V|/2$ positive entries and such that $R_L(\mathbf{y}) \leq R_L(\mathbf{x})$ as promised by Lemma ??. Let $0 < t \leq \max_v \{y_v\}$ be a threshold such that

$$\phi(\{v : y_v \geq t\}) \leq \sqrt{2R_L(\mathbf{y})} \leq \sqrt{2R_L(\mathbf{x})}$$

as promised by Lemma ??. The set $S_t := \phi(\{v : y_v \geq t\})$ contains at most $|V|/2$ vertices, and the cut $(S_t, V - S_t)$ is one of the cuts considered by Fiedler's algorithm on input \mathbf{x} , and so

$$\phi(S_F, V - S_F) \leq \phi(S_t, V - S_t) = \phi(S_t) \leq \sqrt{2R_L(\mathbf{x})}$$

We will prove Lemma ?? next time. We conclude this lecture with a proof of Lemma ??.

PROOF: (Of Lemma ??) First we observe that, for every constant c ,

$$R_L(\mathbf{x} + c\mathbf{1}) \leq R_L(\mathbf{x})$$

because the numerator of $R_L(\mathbf{x} + c\mathbf{1})$ and the numerator of $R_L(\mathbf{x})$ are the same, and the denominator of $R_L(\mathbf{x} + c\mathbf{1})$ is $\|\mathbf{x} + c\mathbf{1}\|^2 = \|\mathbf{x}\|^2 + \|c\mathbf{1}\|^2 \geq \|\mathbf{x}^2\|$.

Let m be the median value of the entries of \mathbf{x} , and call $\mathbf{x}' := \mathbf{x} - m\mathbf{1}$. Then we have $R_L(\mathbf{x}') \leq R_L(\mathbf{x})$, and the median of the entries of \mathbf{x}' is zero, meaning that \mathbf{x}' has at most $|V|/2$ positive entries and at most $|V|/2$ negative entries. We will refer to the vertices v such that $x'_v > 0$ as the *positive* vertices, and the vertices v such that $x'_v < 0$ as the *negative* vertices.

We write

$$\mathbf{x}' = \mathbf{x}^+ - \mathbf{x}^-$$

where $x_v^+ = x'_v$ if v is positive and $x_v^+ = 0$ otherwise; similarly, $x_v^- = -x'_v$ if v is negative, and $x_v^- = 0$ otherwise. Note that \mathbf{x}^+ and \mathbf{x}^- are orthogonal, non-negative, and each of them has at most $|V|/2$ nonzero entries. Note also that, for every t , the cut defined by the set $\{v : x_v^+ \geq t\}$ is one of the cuts considered by Fiedler's algorithm on input \mathbf{x} , because it is the cut

$$(\{v : x_v < t + m\}, \{v : x_v \geq t + m\})$$

Similarly, for every t , the cut defined by the set $\{v : x_v^- \geq t\}$ is one of the cuts considered by Fiedler's algorithm on input \mathbf{x} , because it is the cut

$$(\{v : x_v \leq m - t\}, \{v : x_v > m - t\})$$

It remains to show that at least one of \mathbf{x}^+ or \mathbf{x}^- has Rayleigh quotient smaller than or equal to the Rayleigh quotient of \mathbf{x}' (and, hence, of \mathbf{x}). We claim that

$$\begin{aligned} R_L(\mathbf{x}') &= \frac{\sum_{\{u,v\}} (x_u - x_v)^2}{\|\mathbf{x}'\|^2} = \frac{\sum_{\{u,v\}} ((x_u^+ - x_v^+) - (x_u^- - x_v^-))^2}{\|\mathbf{x}^+\|^2 + \|\mathbf{x}^-\|^2} \\ &\geq \frac{\sum_{\{u,v\}} (x_u^+ - x_v^+)^2 + (x_u^- - x_v^-)^2}{\|\mathbf{x}^+\|^2 + \|\mathbf{x}^-\|^2} \\ &= \frac{R_L(\mathbf{x}^+) \cdot \|\mathbf{x}^+\|^2 + R_L(\mathbf{x}^-) \cdot \|\mathbf{x}^-\|^2}{\|\mathbf{x}^+\|^2 + \|\mathbf{x}^-\|^2} \geq \min\{R_L(\mathbf{x}^+), R_L(\mathbf{x}^-)\} \end{aligned}$$

The only step that we need to justify is that for every edge $\{u, v\}$ we have

$$((x_u^+ - x_v^+) - (x_u^- - x_v^-))^2 \geq (x_u^+ - x_v^+)^2 + (x_u^- - x_v^-)^2$$

If $\{u, v\}$ is an edge between two non-positive vertices, or between two non-negative vertices, then the left-hand side and the right-hand side are clearly equal. If it is an edge between a positive vertex u and a negative vertex v , then the left-hand side is equal to $(x_u^+ + x_v^-)^2$, and the right-hand side is equal to $(x_u^+)^2 + (x_v^-)^2$. \square