Notes 10: Conductance, Expansion, Normalized Laplacians

1. Conductance and expansion

Last lecture we saw that a graph is connected if and only if the second smallest eigenvalue λ_2 of its Laplacian L_G is strictly larger than the smallest eigenvalue λ_1 (which is zero). Today we will show a robust version of this result: a graph is "well-connected" if and only if λ_2 is much bigger than λ_1.

One way to measure how well a graph $G = (V, E)$ is connected is expansion.

Definition 1.1. Given a graph G with positive edge weights $w \in \mathbb{R}_+^E$, the degree of vertex i is $d(i) := \sum_{j:(i,j)\in E} w_{ij}$ and the total degree of a vertex subset $S \subseteq V$ is $d(S) := \sum_{i\in S} d(i)$.

The conductance of a vertex subset $S \subseteq V$ is

$$\varphi(S) = w(S, \overline{S})/d(S) \ ,$$

where $w(S, \overline{S}) = \sum_{i\in S, j\notin S, (i,j)\in E} w_{ij}$ is the total edge weight across the cut from S to $\overline{S}$.

The expansion of a graph is

$$\varphi(G) = \min_{\substack{S\subseteq V,\ S\neq\emptyset \\ d(S)\leqslant d(V)/2}} \varphi(S) \ .$$

The condition $d(S) \leqslant d(V)/2$ in expansion is equivalent to $d(S) \leqslant d(\overline{S})$.

The total degree of a subset S measures the size of a subset, weighted according to degrees. If the graph is regular (all vertices have the same degree), then $\deg(S)$ is proportional to $|S|$.

The conductance of a subset or the expansion a graph is always between 0 and 1. A graph is disconnected if and only if $\varphi(G) = 0$.

1.1. Complete graph.

What is the expansion of the complete graph, the most well-connected graph? Given a subset $S \subseteq V$ of size k, its conductance is $\dfrac{k(n-k)}{k(n-1)} = \dfrac{n-k}{n-1}$. So for any subset of size at most $n/2$, its conductance is at least $\dfrac{n}{2(n-1)} \approx \dfrac{1}{2}$. Complete graph therefore has expansion roughly 1/2.

1.2. Barbell graph.

The barbell graph on $2n$ vertices consists of two disjoint complete subgraphs, each of size n, that are joined by a single edge. This graph is connected (every vertex has a path to any other vertex), but intuitively not well-connected, since removing the extra edge disconnects the two complete subgraphs.

What is the expansion of this graph? Consider S to be the vertex set of one of the complete subgraphs. Then S has conductance $\dfrac{1}{1+n(n-1)} = O\left(\dfrac{1}{n^2}\right)$. Hence the expansion of the barbell graph is also $O(1/n^2)$.

2. Noramlized Laplacians

We are going to compare graph expansion to Laplacian eigenvalues. We will assume the graph has no isolated vertices (of degree 0).

Recall that $L_G = \sum_{(i,j)\in E} w_{ij}(\mathbb{1}_i - \mathbb{1}_j)(\mathbb{1}_i - \mathbb{1}_j)^\top = D - A$, where D is the diagonal matrix with $D_{ii} = d(i)$ and A is the adjacency matrix. (We will drop subscript G and write $L = L_G$.)

All eigenvalues of L lie in the range $[0, 2\Delta]$, where $\Delta = \max_{i\in V} d(i)$ is the maximum degree:

Proposition 2.1. *$-D \preccurlyeq A \preccurlyeq D$ and $0 \preccurlyeq L \preccurlyeq 2D$. In particular, eigenvalues of A are between $-\Delta$ and Δ, and those of L are between 0 and 2Δ.*

Proof. $D - A = \sum_{(i,j)\in E} w_{ij}(\mathbb{1}_i - \mathbb{1}_j)(\mathbb{1}_i - \mathbb{1}_j)^\top \succcurlyeq 0$.

Similarly $D + A = \sum_{(i,j)\in E} w_{ij}(\mathbb{1}_i + \mathbb{1}_j)(\mathbb{1}_i + \mathbb{1}_j)^\top \succcurlyeq 0$.

Inequalities for L follow from those of $A = D - L$. □

We want to remove the dependence on degree and normalize the Laplacian, so that its eigenvalues are between $[0, 2]$. To this end, we "divide" L by the positive definite matrix D — or rather, multiply by $D^{-1/2}$ on both left and right, so that the resulting matrix is still symmetric.

Definition 2.2. The normalized adjacency matrix is $\mathcal{A} = D^{-1/2}AD^{-1/2}$.

The normalized Laplacian is $\mathcal{L} = D^{-1/2}LD^{-1/2} = D^{-1/2}(D-A)D^{-1/2} = I - \mathcal{A}$.

Claim 2.3. *If n-by-n real symmetric matrix X is positive semidefinite, then so is $Y^\top XY$ for any n-by-m real matrix Y. (simple proof omitted)*

Proposition 2.4. *Eigenvalues of $\mathcal{A}$ are between -1 and 1. Eigenvalues of $\mathcal{L}$ are between 0 and 2.*

Proof. The above proposition showed that $D - A \succcurlyeq 0$. Therefore $I - \mathcal{A} = D^{-1/2}(D-A)D^{-1/2} \succcurlyeq 0$ by the above claim (with $X = D - A, Y = D^{-1/2}$). Equivalently, all eigenvalues of $\mathcal{A}$ are at most 1.

Similarly $D+A \succcurlyeq 0$. Therefore $I+\mathcal{A} = D^{-1/2}(D+A)D^{-1/2} \succcurlyeq 0$ by the above claim. Equivalently, all eigenvalues of $\mathcal{A}$ are at least -1.

Eigenvalue bounds for $\mathcal{L}$ follows from eigenvalue bounds for $\mathcal{A} = I - \mathcal{L}$. □

In fact 0 is always an eigenvalue of $\mathcal{L}$, with eigenvector $v_1 = D^{1/2}\mathbb{1}$, because

$$\mathcal{L}v_1 = D^{-1/2}LD^{-1/2}D^{1/2}\mathbb{1} = D^{-1/2}L\mathbb{1} = 0.$$

One can show that L and $\mathcal{L}$ have the same zero eigenspace, via the invertible map $D^{-1/2}$.

3. Cheeger–Alon–Milman inequality

Let $0 = \lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n \leqslant 2$ be the eigenvalues of the normalized Laplacian matrix $\mathcal{L}$ of a graph G.

In the previous lecture, we showed that $\lambda_2 = 0\ (= \lambda_1)$ if and only if the graph is connected.

We now quantify well-connectedness of a graph via λ_2 (the gap between the two smallest eigenvalues).

Theorem 3.1 (Cheeger–Alon–Milman)**.**

$$\frac{\lambda_2}{2} \leqslant \varphi(G) \leqslant \sqrt{2\lambda_2}\,.$$

We first prove the easy direction (left inequality).

By Courant–Fischer, taking $v_1 = D^{1/2}\mathbb{1}$ to be an eigenvector of $\mathcal{L}$ with eigenvalue 0,

$$\lambda_2 = \min_{x \perp v_1} \frac{x^\top \mathcal{L}x}{x^\top x} = \min_{x \perp v_1} \frac{x^\top D^{-1/2}LD^{-1/2}x}{x^\top x} = \min_{D^{1/2}y \perp v_1} \frac{y^\top Ly}{y^\top Dy}\,,$$

where $y = D^{-1/2}x$.

For every vertex subset S, we will construct a vector y satisfying the orthogonality constraint whose Rayleigh quotient is controlled by the conductance of S:

Lemma 3.2. *Every nonempty subset $S \subseteq V$ corresponds to a vector y such that $D^{1/2}y \perp v_1$ and*

$$\frac{y^\top Ly}{y^\top Dy} = \varphi(S)\frac{d(V)}{d(\overline{S})}\,.$$

If $d(S) \leqslant d(V)/2$ (equivalently $d(\overline{S}) \geqslant d(V)/2$), then the quotient $\dfrac{y^\top Ly}{y^\top Dy} \leqslant 2\varphi(S)$. Every nonempty subset $S \subseteq V$ with at most half of the total degree therefore gives us an upperbound $\lambda_2 \leqslant 2\varphi(S)$. Minimizing over all such S yields $\lambda_2 \leqslant 2\varphi(G)$.

It remains to prove the lemma.

The condition $D^{1/2}y \perp v_1$ means $0 = (D^{1/2}y)^\top D^{1/2}\mathbb{1} = y^\top D\mathbb{1} = \sum_{i\in V} d(i)y_i$.

Also, the denominator in the quotient is $y^\top Dy = \sum_{i\in V} d(i)y_i^2$.

In summary,

$$\lambda_2 = \min_{\sum_{i\in V} d(i)y_i = 0} \frac{\sum_{(i,j)\in E} w_{ij}(y_i - y_j)^2}{\sum_{i\in V} d(i)y_i^2}.$$

How to construct y from S? A natural choice is $y = \mathbb{1}_S$, the indicator function for S, i.e. $y_i = 1$ if $i \in S$ and $y_i = 0$ if $i \notin S$.

Then the numerator $\sum_{(i,j)\in E} w_{ij}(y_i - y_j)^2 = w(S, \overline{S})$ and denominator $\sum_{i\in V} d(i)y_i^2 = d(S)$, so the quotient gives us exactly $\varphi(S)$.

But this y fails to satisfy the orthogonality constraint, because $0 \neq \sum_{i\in V} d(i)y_i = d(S)$.

Instead we pick real numbers a and b and assign $y_i = a$ if $i \in S$ and $y_i = b$ if $i \notin S$. We want

$$0 = \sum_{i\in V} d(i)y_i = d(S)a + d(\overline{S})b \; .$$

Solving gives

$$a = \frac{1}{d(S)} \qquad \text{and} \qquad b = \frac{-1}{d(\overline{S})} \; .$$

For this y,

$$\begin{aligned} \frac{\sum_{i\sim j} w_{ij}(y_i - y_j)^2}{\sum_{i\in V} d(i)y_i^2} &= \frac{w(S,\overline{S})\left(\frac{1}{d(S)} + \frac{1}{d(\overline{S})}\right)^2}{d(S)\frac{1}{d(S)^2} + d(\overline{S})\frac{1}{d(\overline{S})^2}} = \frac{w(S,\overline{S})\left(\frac{1}{d(S)} + \frac{1}{d(\overline{S})}\right)^2}{\frac{1}{d(S)} + \frac{1}{d(\overline{S})}} \\ &= w(S,\overline{S})\left(\frac{1}{d(S)} + \frac{1}{d(\overline{S})}\right) = w(S,\overline{S})\frac{d(S) + d(\overline{S})}{d(S)d(\overline{S})} = \frac{w(S,\overline{S})d(V)}{d(S)d(\overline{S})} \; . \end{aligned}$$

We will prove the hard direction (right inequality) of Cheeger–Alon–Milman in the next lecture.