MA431 Spectral Graph Theory: Lecture 8

Ahmad Abdi Neil Olver*

Abstract

We prove Cheeger's inequality, and then discuss expander graphs.

18 Proof of Cheeger's inequality

Recall the statement of Cheeger's inequality.

Theorem 18.1 (Cheeger's inequality)**.** *With ν_2 denoting the second eigenvalue of $\mathcal{L}$, we have*

$$\tfrac{1}{2}\nu_2 \leq \phi(G) \leq \sqrt{2\nu_2}.$$

We will prove that under the assumption that G is d-regular, meaning that $\mathcal{L} = \frac{1}{d}L$. The proof in the general case follows similar lines, though some additional technical awkwardnesses do need to be overcome.

18.1 The easy direction

By the CFT (and using $\mathcal{L} = \frac{1}{d}L$), we have

$$\nu_2 = \min_{x \neq 0, x \perp \mathbf{1}} \frac{x^\top L x}{d x^\top x} =: R(x).$$

Observe that $R(\mathbf{1}_S) = |\delta(S)| / \operatorname{vol}(S)$, and so $\phi(S) = \min\{R(\mathbf{1}_S), R(\mathbf{1}_{V\setminus S})\}$. However, $\mathbf{1}_S$ and $\mathbf{1}_{V\setminus S}$ are not generally orthogonal to $\mathbf{1}$, so we cannot deduce that $\nu_2 \leq \phi(S)$ from this. Instead, we choose an appropriate convex combination of $\mathbf{1}_S$ and $\mathbf{1}_{V\setminus S}$ that is orthogonal to $\mathbf{1}$ (we took a different approach in the lecture).

So let $z = \frac{1}{|S|}\mathbf{1}_S - \frac{1}{|V\setminus S|}\mathbf{1}_{V\setminus S}$, so that $z \perp \mathbf{1}$. Then $\nu_2 \leq R(z)$. We then just compute:

$$\begin{aligned}
R(z) &= \frac{\sum_{vw\in E}(z_v - z_w)^2}{d\sum_{v\in V} z_v^2} \\
&= \frac{|\delta(S)| \cdot \left(\frac{1}{|S|} + \frac{1}{|V\setminus S|}\right)^2}{d\left(\frac{|S|}{|S|^2} + \frac{|V\setminus S|}{|V\setminus S|^2}\right)} \\
&= \frac{|\delta(S)|}{d} \cdot \left(\frac{1}{|S|} + \frac{1}{|V\setminus S|}\right) \\
&\leq 2\frac{|\delta(S)|}{d} \cdot \max\{1/|S|, 1/|V\setminus S|\} \\
&= 2\phi(S).
\end{aligned}$$

Thus $\nu_2 \leq 2\phi(G)$ as required.

18.2 The hard direction

We now want to show that $\phi(G) \leq \sqrt{2\nu_2}$. Let x be an eigenvector of $\mathcal{L}$ corresponding to ν_2. Also assume that $|\{i : x_i > 0\}| \leq |V|/2$; otherwise simply replace x with $-x$. Note that x does have some positive component, since $x \perp \mathbf{1}$ and $x \neq 0$.

Now define $y \in \mathbb{R}^V$ by setting $y_v = \max\{x_v, 0\}$ for all $v \in V$. We can assume that $\max_v y_v = 1$, simply by scaling.

Lemma 18.2. $R(y) \leq R(x) = \nu_2$.

Proof. Let $Q = \{v : y_v > 0\}$. Then for $v \in Q$,

$$(\mathcal{L}y)_v = y_v - \tfrac{1}{d} \sum_{w:vw \in E} y_w \leq x_v - \tfrac{1}{d} \sum_{w:vw \in E} x_v = (\mathcal{L}x)_v = \nu_2 x_v = \nu_2 y_v.$$

Thus

$$y^\top \mathcal{L} y = \sum_{v \in Q} y_v (\mathcal{L}y)_v \leq \sum_{v \in Q} \nu_2 y_v^2 = \nu_2 \sum_{v \in V} y_v^2,$$

as required. □

For any $\tau \in (0,1)$, let

$$S_\tau = \{v : y_v^2 \geq \tau\}.$$

Note that $S_\tau \neq \emptyset$ and $|S_\tau| \leq |V|/2$ for all $\tau \in (0,1)$. We will now choose τ uniformly at random from $[0,1]$ and see that "on average" this choice works.

With this choice of τ, we have

$$\mathbb{E}[\mathrm{vol}(S_\tau)] = d \sum_{v \in V} \Pr(y_v^2 \geq \tau) = d \sum_{v \in V} y_v^2,$$

from linearity of expectation. We also have

$$
\begin{aligned}
\mathbb{E}[|\delta(S_\tau)|] &= \sum_{vw \in E} \Pr(vw \in \delta(S_\tau)) \\
&= \sum_{vw \in E} \Pr\big((y_v^2 < \tau \wedge y_w^2 \geq \tau) \vee (y_v^2 < \tau \wedge y_w^2 \geq \tau)\big) \\
&= \sum_{vw \in E} |y_v^2 - y_w^2| \\
&= \sum_{vw \in E} |y_v - y_w|(y_v + y_w) \\
&\leq \sqrt{\sum_{vw \in E} (y_v - y_w)^2} \cdot \sqrt{\sum_{vw \in E} (y_v + y_w)^2} \\
&\leq \sqrt{R(y) \cdot d \sum_{v \in V} y_v^2} \cdot \sqrt{2 \sum_{vw \in E} (y_v^2 + y_w^2)} \\
&= \sqrt{2R(y)} \cdot d \sum_{v \in V} y_v^2.
\end{aligned}
$$

(We have used Cauchy-Schwartz, and the simple inequality $2y_v y_w \leq y_v^2 + y_w^2$). Since $R(y) \leq \nu_2$, we can now deduce that

$$\mathbb{E}[|\delta(S_\tau)|] \leq \sqrt{2\nu_2}\mathbb{E}[\mathrm{vol}(S_\tau)].$$

Thus there must exist a choice t for which $|\delta(S_t)| \leq \sqrt{2\nu_2}\,\mathrm{vol}(S_t)$, and we're done.

Note that the proof gives an algorithm to find a cut S for which $\phi(S) \leq \sqrt{2\nu_2}$. Simply try all distinct choices of S_τ (of which there are less than n) and choose one of minimum conductance. This will generally not return a cut of minimum conductance, but chaining both the easy and hard directions of Cheeger together we do have that this algorithm returns a cut S for which $\phi(S) \leq 2\sqrt{\phi(G)}$.

19 Expander graphs

We will now really restrict our attention to regular graphs (not just for streamlining a proof). In what follows, unless otherwise stated, $G = (V, E)$ will always refer to a connected, d-regular graph on n vertices, with Laplacian L and adjacency matrix A. The spectrum of L is $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$. The spectrum of A is then $d > d - \lambda_2 \geq \cdots \geq d - \lambda_n$ (since $L = dI - A$) . Note that $\lambda_n \leq 2d$, since the spectrum of A is contained in $[-d, d]$.

We will say that G has *edge expansion* α if

$$\frac{|\delta(S)|}{|S|} \geq \alpha \qquad \text{for all } \emptyset \subsetneq S \subseteq V, |S| \leq n/2.$$

In other words, if $\phi(G) \geq \alpha/d$. Since the easy direction of Cheeger tells us that $\lambda_2/(2d) \leq \phi(G)$, we can deduce that the edge expansion of G is at least $\lambda_2/2$. This motivates the following definition.

Definition 19.1. A d-regular graph G with Laplacian spectrum $\lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$ is called a (spectral) *one-sided β-expander* for some $\beta \in (0, 1]$ if $\lambda_2 \geq \beta d$. If in addition, $\lambda_n \leq 2d - \beta d$, then G is called a *two-sided β-expander*.

So a one-sided β-expander has edge expansion $\beta d/2$. What's the motivation for the two-sided definition?

Firstly, simple random walk on a two-sided expander will mix fast. Lazy random walk (or a continuous time random walk) will mix fast on a one-sided expander; this we saw already. But if G is bipartite, or close to bipartite, simple random walk may not mix at all, or very slowly. It's the largest eigenvalue of A in *absolute value* aside from the d eigenvalue that matters.

Secondly, two-sided expanders can be thought of as good sparse approximations to the complete graph, in the following sense. Recall that K_n denotes the complete graph on n vertices.

Lemma 19.2. *If G is a two-sided $(1 - \epsilon)$-expander (for some $0 < \epsilon < 1$) then*

$$(1 - \epsilon)L(K_n) \preceq \tfrac{n}{d} L \preceq (1 + \epsilon)L(K_n).$$

In other words, for any $x \in \mathbb{R}^V$,

$$(1 - \epsilon)x^\top L(K_n)x^\top \leq \tfrac{n}{d} x^\top Lx \leq (1 + \epsilon)x^\top L(K_n)x.$$

This lemma tells us that for G a two-sided $(1 - \epsilon)$-expander if we take the weighted graph (G, w), where $w_e = n/d$ for all $e \in E$, then the corresponding weighted Laplacian $L_w = \frac{n}{d} L$ is "close" to the Laplacian of the complete graph. Take, for example, $x = \mathbf{1}_S$ for some set S; then we can deduce from this that

$$(1 - \epsilon)|\delta_{K_n}(S)| \leq \tfrac{n}{d} |\delta_G(S)| \leq (1 + \epsilon)|\delta_{K_n}(S)|.$$

That is, the wight across any cut is roughly the same in (G, w) as in K_n. There is a whole theory on finding sparse graphs that "approximate" some given graph; expanders are sparse graphs that approximate the complete graph.

Proof. Note that all eigenvalues of $L(K_n)$ aside from the 0-eigenvalue have value n, by observing that $L(K_n) = nI - J$, where J is the all-ones matrix. We clearly have $\mathbf{1}^\top L(K_n)\mathbf{1} = 0 = \mathbf{1}^\top L\mathbf{1}$. So consider any $x \perp \mathbf{1}$. Then $x^\top L(K_n)x = nx^\top x$, whereas

$$d(1 - \epsilon)x^\top x \leq x^\top Lx \leq d(1 + \epsilon)x^\top x$$

by the CFT. This shows the claim. □

Trivially, the complete graph K_n is an $(n - 1)$-regular two-sided $(1 - \frac{1}{n-1})$-expander. So in that sense, expanders certainly exist. But what's much more interesting, and much less obvious, is the existence of *sparse* expanders. Fix the degree d; for n very large, do β-expanders exist for any fixed positive $\beta > 0$? This is already not obvious. Happily, these amazing objects do exist. They have many useful properties (and are excellent candidates to reach for when looking for counterexamples or extremal examples to many problems). We won't say much (some further facts will appear as exercises). Here is one basic fact: expanders have the following quasirandom property.

Lemma 19.3. *If G is a d-regular two-sided $(1-\epsilon)$-expander on n vertices, then for any two disjoint sets $S, T \subseteq V$,*

$$\left| E(S,T) - d\frac{|S|\cdot|T|}{n} \right| \leq \epsilon d\sqrt{|S|\cdot|T|}.$$

Here, $E(S,T)$ denotes the number of edges with one endpoint in S and the other in T.

We leave the proof as an exercise. Note that $d|S|\cdot|T|/n$ is the expected number of edges between S and T in a random graph, where every edge is independently present with probability d/n (so that the average degree is d).

20 Ramanujan graphs

The next natural question is whether for a fixed degree d, there exists a $\beta > 0$ such that (one or two sided) β-expanders of arbitrarily large size exist, and if so, how to construct them. We're going to skip ahead a bit and ask a stronger question. What is the *best* expansion we can hope for? I.e., what is the largest value of β we can hope for (as a function of d)?

The following theorem gives a partial answer; it shows an upper bound on β.

Theorem 20.1 (Alon-Bopanna bound)**.** *For G d-regular, and λ_2 the second eigenvalue of its Laplacian,*

$$\lambda_2 \leq d - 2\sqrt{d-1} + o(1),$$

where the $o(1$ term hides terms that go to zero as the size of G goes to infinity.

So in other words, the largest possible β for which we can have arbitrarily large one- or two-sided β-expanders is $1 - \frac{2\sqrt{d-1}}{d}$. We won't prove this theorem; a weaker bound of $d - \sqrt{d} + o(1)$ is much easier to obtain however, and is left as an exercise. Instead, we will see where the $2\sqrt{d-1}$ comes from.

Let $\mathbb{T}_d$ denote the infinite d-regular tree (We have not discussed infinite graphs, and we will not go into the formalism here; but it should be reasonably clear what this object is.) Instead of the adjacency matrix, we have an adjacency *operator*, which acts on the space of square integrable vectors on $\mathbb{V} := V(\mathbb{T}_d)$. That is, $\mathbb{A} = A(\mathbb{T}_d)$ is defined by

$$(\mathbb{A}x)_v = \sum_{w: vw \in E(\mathbb{T}_d)} x_w \qquad \text{for all } x \in \ell_2(\mathbb{V}).$$

Although $\mathbb{T}_d$ is d-regular, the all-ones vector is not an eigenvector, and d is not an eigenvalue. The reason is that $\mathbf{1} \notin \ell_2(\mathbb{V})$. As such, the relevant quantity measuring the spectral expansion of $\mathbb{T}_d$ is the spectral radius of $\mathbb{A}$; we do not need to discard any eigenvalues from consideration. The spectral radius of $\mathbb{A}$ (which we will denote by $\rho(\mathbb{T}_d)$) is $\sup\{|\lambda| : \lambda - A \text{ is not invertible}\}$. It turns out that $\rho(\mathbb{T}_d) = 2\sqrt{d-1}$. This explains the value; the

Alon-Bopanna bound is basically saying that for very large d-regular graphs, the spectral expansion cannot be better than that of the infinite d-regular tree.

We won't quite prove that $\rho(\mathbb{T}_d) = 2\sqrt{d-1}$, since we don't want to deal with the analytic subtleties of infinite dimensions (where, e.g., the spectrum is not fully determined by the set of eigenvalues). Instead we'll go back to finite trees, and show that in the limit of large d-ary trees, we obtain precisely the value $2\sqrt{d-1}$. Recall that $\rho(G)$ denotes the spectral radius of the adjacency matrix of G.

Lemma 20.2. *The following hold.*

1. *If T is a tree with all degrees bounded by d, then $\rho(T) \leq 2\sqrt{d-1}$.*
2. *Let T_k denote the complete d-ary tree of height k. Then $\rho(T_k) \to 2\sqrt{d-1}$ as $k \to \infty$.*

Proof. **(1).** Let $T = (W, F)$, and fix a root $r \in W$. For all $v \neq r$, let $p(v)$ denote the parent of v, and for all $v \in W$, let $C(v)$ denote the set of children of v. Then for any $x \in \mathbb{R}^W$,

$$\begin{aligned} x^\top A x &= 2 \sum_{v \in W} \sum_{w \in C(v)} x_v x_w \\ &\leq \sum_{v \in W} \sum_{w \in C(v)} \left(\frac{x_v^2}{\sqrt{d-1}} + \sqrt{d-1} x_w^2 \right) \\ &\leq \sum_{v \in W} x_v^2 \left(\frac{d-1}{\sqrt{d-1}} + \sqrt{d-1} \right) \\ &= 2\sqrt{d-1} \sum_{v \in W} x_v^2. \end{aligned}$$

The first inequality is just the fact that $\alpha x^2 + \frac{1}{\alpha} y^2 \geq 2xy$ for any $x, y \geq 0$, $\alpha > 0$. The choice of α is so that when all the x_v^2 terms are collected, the contribution coming from when v appears as a child (just once) and the contribution coming from when it appears as a parent (up to $d-1$ times) balances.

(2). Exercise. $\square$

Definition 20.3. A d-regular graph G is called *Ramanujan* if it is a two-sided $(1 - \frac{2\sqrt{d-1}}{d})$-expander.

A bipartite d-regular graph G is called *bipartite Ramanujan* if it is a one-sided $(1 - \frac{2\sqrt{d-1}}{d})$-expander.

Happily, Ramanujan graphs of arbitrarily large size do exist! Lubotzky, Phillips and Sarnak (and independently Margulis) proved that if $d-1$ is prime, then there are arbitrarily large Ramanujan graphs of degree d. The construction of LPS is based on something called the "Ramanujan conjecture", which is what motivated them to give these objects this name. Friedman proved that random d-regular graphs are "almost Ramanujan" with high probability: they are two-sided $(1 - \frac{2\sqrt{d-1}}{d} - o(1))$-expanders, where the $o(1)$ hides terms that go to zero as the size of the graph goes to infinity. It is even conjectured that random d-regular graphs *are* Ramanujan with high probability. Despite this, it's not currently known whether Ramanujan graphs exist for all degrees d. What we will see in this course is a result by Marcus, Spielman and Srivastava: for every degree $d \geq 3$, there are bipartite Ramanujan graphs of arbitrary size.