Spectral Graph Theory Lecture 17

Properties of Expander Graphs

Daniel A. Spielman October 29, 2018

17.1 Overview

We say that a d-regular graph is a good expander if all of its adjacency matrix eigenvalues are small. To quantify this, we set a threshold $\epsilon > 0$, and require that each adjacency matrix eigenvalue, other than d, has absolute value at most ϵd. This is equivalent to requiring all non-zero eigenvalues of the Laplacian to be within ϵd of d.

In this lecture, we will:

1. Show that this condition is equivalent to approximating the complete graph.
2. Prove that this condition implies that the number of edges between sets of vertices in the graph is approximately the same as in a d-regular random graph.
3. Prove Tanner's Theorem: that small sets of vertices have many neighbors.
4. Derive the Alon-Boppana bound, which says that ϵ cannot be asymptotically smaller than $2\sqrt{d-1}/d$. This will tell us that the asymptotically best expanders are the Ramanujan graphs.

Random d-regular graphs are expander graphs. Explicitly constructed expander graphs have proved useful in a large number of algorithms and theorems. We will see some applications of them next week.

17.2 Expanders as Approximations of the Complete Graph

One way of measuring how well two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ approximate each other is to measure the operator norm of their difference: $\boldsymbol{A} - \boldsymbol{B}$. Since I consider the operator norm by default, I will just refer to it as the norm. Recall that the norm of a matrix $\boldsymbol{M}$ is defined to be its largest singular value:

$$\|\boldsymbol{M}\| = \max_{\boldsymbol{x}} \frac{\|\boldsymbol{M}\boldsymbol{x}\|}{\|\boldsymbol{x}\|},$$

where the norms in the fraction are the standard Euclidean vector norms. The norm of a symmetric matrix is just the largest absolute value of one of its eigenvalues. It can be very different for a non symmetric matrix.

For this lecture, we define an ϵ-expander to be a d-regular graph whose adjacency matrix eigenvalues satisfy $|\mu_i| \leq \epsilon d$ for $\mu_i \geq 2$. As the Laplacian matrix eigenvalues are given by $\lambda_i = d - \mu_i$, this is equivalent to $|d - \lambda_i| \leq \epsilon d$ for $i \geq 2$. It is also equivalent to

$$\|\boldsymbol{L}_G - (d/n)\boldsymbol{L}_{K_n}\| \leq \epsilon d.$$

For this lecture, I define a graph G to be an ϵ-approximation of a graph H if

$$(1-\epsilon)H \preccurlyeq G \preccurlyeq (1+\epsilon)H,$$

where I recall that I say $H \preccurlyeq G$ if for all $\boldsymbol{x}$

$$\boldsymbol{x}^T \boldsymbol{L}_H \boldsymbol{x} \leq \boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x}.$$

I warn you that this definition is not symmetric. When I require a symmetric definition, I usually use the condition $(1+\epsilon)^{-1}H \preccurlyeq G$ instead of $(1-\epsilon)H \preccurlyeq G$.

If G is an ϵ-expander, then for all $\boldsymbol{x} \in \mathbb{R}^V$ that are orthogonal to the constant vectors,

$$(1-\epsilon)d\boldsymbol{x}^T\boldsymbol{x} \leq \boldsymbol{x}^T \boldsymbol{L}_G \boldsymbol{x} \leq (1+\epsilon)d\boldsymbol{x}^T\boldsymbol{x}.$$

On the other hand, for the complete graph K_n, we know that all $\boldsymbol{x}$ orthogonal to the constant vectors satisfy

$$\boldsymbol{x}^T \boldsymbol{L}_{K_n} \boldsymbol{x} = n\boldsymbol{x}^T\boldsymbol{x}.$$

Let H be the graph

$$H = \frac{d}{n}K_n,$$

so

$$\boldsymbol{x}^T \boldsymbol{L}_H \boldsymbol{x} = d\boldsymbol{x}^T\boldsymbol{x}.$$

So, G is an ϵ-approximation of H.

This tells us that $\boldsymbol{L}_G - \boldsymbol{L}_H$ is a matrix of small norm. Observe that

$$(1-\epsilon)\boldsymbol{L}_H \preccurlyeq \boldsymbol{L}_G \preccurlyeq (1+\epsilon)\boldsymbol{L}_H \quad \text{implies} \quad -\epsilon\boldsymbol{L}_H \preccurlyeq \boldsymbol{L}_G - \boldsymbol{L}_H \preccurlyeq \epsilon\boldsymbol{L}_H.$$

As $\boldsymbol{L}_G$ and $\boldsymbol{L}_H$ are symmetric, and all eigenvalues of $\boldsymbol{L}_H$ are 0 or d, we may infer

$$\|\boldsymbol{L}_G - \boldsymbol{L}_H\| \leq \epsilon d. \tag{17.1}$$

17.3 Quasi-Random Properties of Expanders

There are many ways in which expander graphs act like random graphs. Conversely, one can prove that a random d-regular graph is an expander graph with reasonably high probability [Fri08].

We will see that all sets of vertices in an expander graph act like random sets of vertices. To make this precise, imagine creating a random set $S \subset V$ by including each vertex in S independently with probability α. How many edges do we expect to find between vertices in S? Well, for every edge (u, v), the probability that $u \in S$ is α and the probability that $v \in S$ is α, so the probability that both endpoints are in S is α^2. So, we expect an α^2 fraction of the edges to go between vertices in S. We will show that this is true *for all sufficiently large sets* S in an expander.

In fact, we will prove a stronger version of this statement for two sets S and T. Imagine including each vertex in S independently with probability α and each vertex in T with probability β. We allow vertices to belong to both S and T. For how many ordered pairs $(u, v) \in E$ do we expect to have $u \in S$ and $v \in T$? Obviously, it should hold for an $\alpha\beta$ fraction of the pairs.

For a graph $G = (V, E)$, define

$$\vec{E}(S, T) = \{(u, v) : u \in S, v \in T, (u, v) \in E\}.$$

We have put the arrow above the E in the definition, because we are considering ordered pairs of vertices. When S and T are disjoint

$$\left|\vec{E}(S, T)\right|$$

is precisely the number of edges between S and T, while

$$\left|\vec{E}(S, S)\right|$$

counts every edge inside S twice.

The following bound is a slight extension by Beigel, Margulis and Spielman [BMS93] of a bound originally proved by Alon and Chung [AC88].

Theorem 17.3.1. *Let $G = (V, E)$ be a d-regular graph that ϵ-approximates $\frac{d}{n}K_n$. Then, for every $S \subseteq V$ and $T \subseteq V$,*

$$\left|\left|\vec{E}(S, T)\right| - \alpha\beta dn\right| \leq \epsilon dn\sqrt{(\alpha - \alpha^2)(\beta - \beta^2)},$$

where $|S| = \alpha n$ and $|T| = \beta n$.

Observe that when α and β are greater than ϵ, the term on the right is less than $\alpha\beta dn$.

In class, we will just prove this in the case that S and T are disjoint.

Proof. The first step towards the proof is to observe

$$\boldsymbol{\chi}_S^T \boldsymbol{L}_G \boldsymbol{\chi}_T = d\,|S \cap T| - \left|\vec{E}(S, T)\right|.$$

Let $H = \frac{d}{n}K_n$. As G is a good approximation of H, let's compute

$$\boldsymbol{\chi}_S^T \boldsymbol{L}_H \boldsymbol{\chi}_T = \boldsymbol{\chi}_S^T \left(dI - \frac{d}{n}J\right)\boldsymbol{\chi}_T = d\,|S \cap T| - \frac{d}{n}\,|S|\,|T| = d\,|S \cap T| - \alpha\beta dn.$$

So,

$$\left|\left|\vec{E}(S, T)\right| - \alpha\beta dn\right| = \left|\boldsymbol{\chi}_S^T \boldsymbol{L}_G \boldsymbol{\chi}_T - \boldsymbol{\chi}_S^T \boldsymbol{L}_H \boldsymbol{\chi}_T\right|.$$

As

$$\|\boldsymbol{L}_G - \boldsymbol{L}_H\| \leq \epsilon d,$$

$$\begin{aligned}
\boldsymbol{\chi}_S^T \boldsymbol{L}_H \boldsymbol{\chi}_T - \boldsymbol{\chi}_S^T \boldsymbol{L}_G \boldsymbol{\chi}_T &= \boldsymbol{\chi}_S^T (\boldsymbol{L}_H - \boldsymbol{L}_G) \boldsymbol{\chi}_T \\
&\leq \|\boldsymbol{\chi}_S\| \, \|(\boldsymbol{L}_H - \boldsymbol{L}_G)\boldsymbol{\chi}_T\| \\
&\leq \|\boldsymbol{\chi}_S\| \, \|\boldsymbol{L}_H - \boldsymbol{L}_G\| \, \|\boldsymbol{\chi}_T\| \\
&\leq \epsilon d \, \|\boldsymbol{\chi}_S\| \, \|\boldsymbol{\chi}_T\| \\
&= \epsilon d n \sqrt{\alpha\beta}.
\end{aligned}$$

This is almost as good as the bound we are trying to prove. To prove the claimed bound, recall that $\boldsymbol{L}_H \boldsymbol{x} = \boldsymbol{L}_H(\boldsymbol{x} + c\boldsymbol{1})$ for all c. So, let $\boldsymbol{x}_S$ and $\boldsymbol{x}_T$ be the result of orthogonalizing $\boldsymbol{\chi}_S$ and $\boldsymbol{\chi}_T$ with respect to the constant vectors. By Claim 2.4.2 (from Lecture 2), $\|\boldsymbol{x}_S\| = n(\alpha - \alpha^2)$. So, we obtain the improved bound

$$\boldsymbol{x}_S^T (\boldsymbol{L}_H - \boldsymbol{L}_G) \boldsymbol{x}_T = \boldsymbol{\chi}_S^T (\boldsymbol{L}_H - \boldsymbol{L}_G) \boldsymbol{\chi}_T,$$

while

$$\|\boldsymbol{x}_S\| \, \|\boldsymbol{x}_T\| = n\sqrt{(\alpha - \alpha^2)(\beta - \beta^2)}.$$

So, we may conclude

$$\left| \left| \vec{E}(S,T) \right| - \alpha\beta d n \right| \leq \epsilon d n \sqrt{(\alpha - \alpha^2)(\beta - \beta^2)}.$$

□

We remark that when S and T are disjoint, the same proof goes through even if G is irregular and weighted if we replace $\vec{E}(S,T)$ with

$$w(S,T) = \sum_{(u,v) \in E, u \in S, v \in T} w(u,v).$$

We only need the fact that G ϵ-approximates $\frac{d}{n} K_n$. See [BSS12] for details.

17.4 Vertex Expansion

The reason for the name *expander graph* is that small sets of vertices in expander graphs have unusually large numbers of neighbors. For $S \subset V$, let $N(S)$ denote the set of vertices that are neighbors of vertices in S. The following theorem, called Tanner's Theorem, provides a lower bound on the size of $N(S)$.

Theorem 17.4.1 ([Tan84]). *Let $G = (V, E)$ be a d-regular graph on n vertices that ϵ-approximates $\frac{d}{n} K_n$. Then, for all $S \subseteq V$,*

$$|N(S)| \geq \frac{|S|}{\epsilon^2 (1 - \alpha) + \alpha},$$

where $|S| = \alpha n$.

Note that when α is much less than ϵ^2, the term on the right is approximately $|S|/\epsilon^2$, which can be much larger than $|S|$. We will derive Tanner's theorem from Theorem 17.3.1.

Proof. Let $R = N(S)$ and let $T = V - R$. Then, there are no edges between S and T. Let $|T| = \beta n$ and $|R| = \gamma n$, so $\gamma = 1 - \beta$. By Theorem 17.3.1, it must be the case that

$$\alpha\beta dn \leq \epsilon dn\sqrt{(\alpha - \alpha^2)(\beta - \beta^2)}.$$

The lower bound on γ now follows by re-arranging terms. Dividing through by dn and squaring both sides gives

$$\begin{aligned}
\alpha^2\beta^2 &\leq \epsilon^2(\alpha - \alpha^2)(\beta - \beta^2) && \iff \\
\alpha\beta &\leq \epsilon^2(1-\alpha)(1-\beta) && \iff \\
\frac{\beta}{1-\beta} &\leq \frac{\epsilon^2(1-\alpha)}{\alpha} && \iff \\
\frac{1-\gamma}{\gamma} &\leq \frac{\epsilon^2(1-\alpha)}{\alpha} && \iff \\
\frac{1}{\gamma} &\leq \frac{\epsilon^2(1-\alpha)+\alpha}{\alpha} && \iff \\
\gamma &\geq \frac{\alpha}{\epsilon^2(1-\alpha)+\alpha}.
\end{aligned}$$

□

If instead of $N(S)$ we consider $N(S) - S$, then T and S are disjoint, so the same proof goes through for weighted, irregular graphs that ϵ-approximate $\frac{d}{n}K_n$.

17.5 How well can a graph approximate the complete graph?

Consider applying Tanner's Theorem with $S = \{v\}$ for some vertex v. As v has exactly d neighbors, we find

$$\epsilon^2(1 - 1/n) + 1/n \geq 1/d,$$

from which we see that ϵ must be at least $1/\sqrt{d + d^2/n}$, which is essentially $1/\sqrt{d}$. But, how small can it be?

The Ramanujan graphs, constructed by Margulis [Mar88] and Lubotzky, Phillips and Sarnak [LPS88] achieve

$$\epsilon \leq \frac{2\sqrt{d-1}}{d}.$$

We will see that if we keep d fixed while we let n grow, ϵ cannot exceed this bound in the limit. We will prove an upper bound on ϵ by constructing a suitable test function.

As a first step, choose two vertices v and u in V whose neighborhoods to do not overlap. Consider the vector $\boldsymbol{x}$ defined by

$$\boldsymbol{x}(i) = \begin{cases} 1 & \text{if } i = u, \\ 1/\sqrt{d} & \text{if } i \in N(u), \\ -1 & \text{if } i = v, \\ -1/\sqrt{d} & \text{if } i \in N(v), \\ 0 & \text{otherwise.} \end{cases}$$

Now, compute the Rayleigh quotient with respect to $\boldsymbol{x}$. The numerator is the sum over all edges of the squares of differences across the edges. This gives $(1 - 1/\sqrt{d})^2$ for the edges attached to u and v, and $1/d$ for the edges attached to $N(u)$ and $N(v)$ but not to u or v, for a total of

$$2d(1 - 1/\sqrt{d})^2 + 2d(d-1)/d = 2\left(d - 2\sqrt{d} + 1 + (d-1)\right) = 2\left(2d - 2\sqrt{d}\right).$$

On the other hand, the denominator is 4, so we find

$$\frac{\boldsymbol{x}^T L \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = d - \sqrt{d}.$$

If we use instead the vector

$$\boldsymbol{y}(i) = \begin{cases} 1 & \text{if } i = u, \\ -1/\sqrt{d} & \text{if } i \in N(u), \\ -1 & \text{if } i = v, \\ 1/\sqrt{d} & \text{if } i \in N(v), \\ 0 & \text{otherwise,} \end{cases}$$

we find

$$\frac{\boldsymbol{y}^T L \boldsymbol{y}}{\boldsymbol{y}^T \boldsymbol{y}} = d + \sqrt{d}.$$

This is not so impressive, as it merely tells us that $\epsilon \geq 1/\sqrt{d}$, which we already knew. But, we can improve this argument by pushing it further. We do this by modifying it in two ways. First, we extend $\boldsymbol{x}$ to neighborhoods of neighborhoods of u and v. Second, instead of basing the construction at vertices u and v, we base it at two edges. This way, each vertex has $d-1$ edges to those that are farther away from the centers of the construction.

The following theorem is attributed to A. Nilli [Nil91], but we suspect it was written by N. Alon.

Theorem 17.5.1. *Let G be a d-regular graph containing two edges (u_0, u_1) and (v_0, v_1) that are at distance at least $2k+2$. Then*

$$\lambda_2 \leq d - 2\sqrt{d-1} + \frac{2\sqrt{d-1} - 1}{k+1}.$$

Proof. Define the following neighborhoods.

$$\begin{aligned} U_0 &= \{u_0, u_1\} \\ U_i &= N(U_{i-1}) - \cup_{j<i} U_j, \quad \text{for } 0 < i \leq k, \\ V_0 &= \{v_0, v_1\} \\ V_i &= N(V_{i-1}) - \cup_{j<i} V_j, \quad \text{for } 0 < i \leq k. \end{aligned}$$

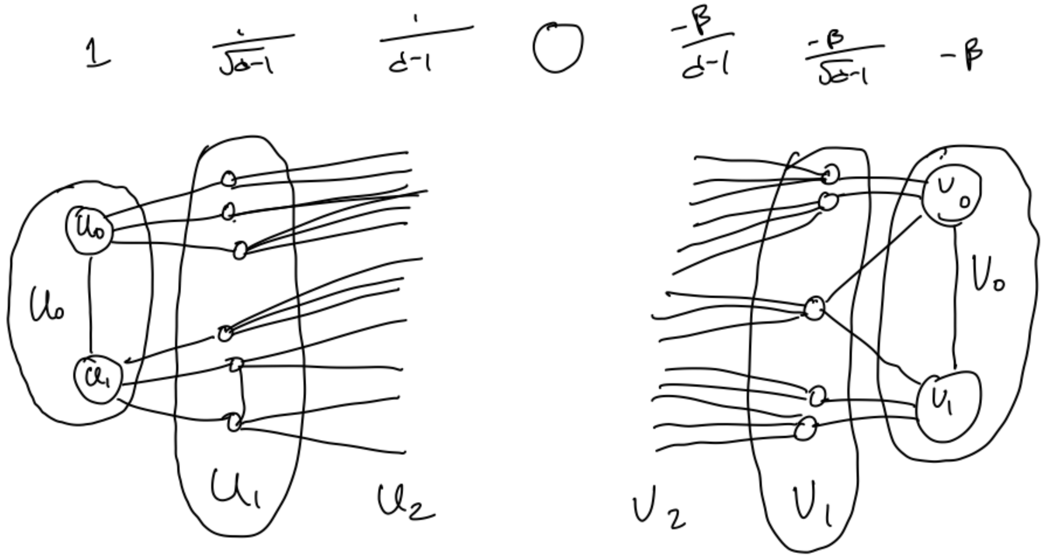

Figure 17.1: The construction of $\boldsymbol{x}$.

That is, U_i consists of exactly those vertices at distance i from U_0. Note that there are no edges between any vertices in any U_i and any V_j.

Our test vector for λ_2 will be given by

$$\boldsymbol{x}(a) = \begin{cases} \frac{1}{(d-1)^{-i/2}} & \text{for } a \in U_i \\ -\frac{\beta}{(d-1)^{-i/2}} & \text{for } a \in V_i \\ \\ 0 & \text{otherwise.} \end{cases}$$

We choose β so that $\boldsymbol{x}$ is orthogonal to $\mathbf{1}$.

We now find that the Rayleigh quotient of $\boldsymbol{x}$ with respect to $\boldsymbol{L}$ is at most

$$\frac{X_0 + \beta^2 Y_0}{X_1 + \beta^2 Y_1},$$

where

$$X_0 = \sum_{i=0}^{k-1} |U_i| \, (d-1) \left(\frac{1 - 1/\sqrt{d-1}}{(d-1)^{-i/2}} \right)^2 + |U_k| \, (d-1)^{-k+1}, \text{ and } X_1 = \sum_{i=0}^{k} |U_i| \, (d-1)^{-i}$$

and

$$Y_0 = \sum_{i=0}^{k-1} |V_i| \, (d-1) \left(\frac{1 - 1/\sqrt{d-1}}{(d-1)^{-i/2}} \right)^2 + |V_k| \, (d-1)^{-k+1}, \text{ and } Y_1 = \sum_{i=0}^{k} |V_i| \, (d-1)^{-i}.$$

By my favorite inequality, it suffices to prove upper bounds on X_0/X_1 and Y_0/Y_1. So, consider

$$\frac{\sum_{i=0}^{k-1} |U_i| (d-1) \left(\frac{1-1/\sqrt{d-1}}{(d-1)^{-i/2}}\right)^2 + |U_k| (d-1)^{-k+1}}{\sum_{i=0}^{k} |U_i| (d-1)^{-i}}.$$

For now, let's focus on the numerator,

$$\begin{aligned}
&\sum_{i=0}^{k-1} |U_i| (d-1) \left(\frac{1-1/\sqrt{d-1}}{(d-1)^{-i/2}}\right)^2 + |U_k| (d-1)(d-1)^{-k} \\
&= \sum_{i=0}^{k-1} \frac{|U_i|}{(d-1)^i}(d-2\sqrt{d-1}) + \frac{|U_k|}{(d-1)^k}(d-1) \\
&= \sum_{i=0}^{k-1} \frac{|U_i|}{(d-1)^i}(d-2\sqrt{d-1}) + \frac{|U_k|}{(d-1)^k}(d-2\sqrt{d-1}) + \frac{|U_k|}{(d-1)^k}(2\sqrt{d-1}-1) \\
&= \sum_{i=0}^{k} \frac{|U_i|}{(d-1)^i}(d-2\sqrt{d-1}) + \frac{|U_k|}{(d-1)^k}(2\sqrt{d-1}-1).
\end{aligned}$$

To upper bound the Rayleigh quotient, we observe that the left-most of these terms contributes

$$\frac{\sum_{i=0}^{k} \frac{|U_i|}{(d-1)^i}(d-2\sqrt{d-1})}{\sum_{i=0}^{k} |U_i| (d-1)^{-i}} = d-2\sqrt{d-1}.$$

To bound the impact of the remaining term,

$$\frac{|U_k|}{(d-1)^k}(2\sqrt{d-1}-1),$$

note that

$$|U_k| \leq (d-1)^{k-i} |U_i|.$$

So, we have

$$\frac{|U_k|}{(d-1)^k} \leq \frac{1}{k+1} \sum_{i=0}^{k} \frac{|U_i|}{(d-1)^i}.$$

Thus, the last term contributes at most

$$\frac{2\sqrt{d}-1}{k+1}$$

to the Rayleigh quotient.

□

17.6 Open Problems

What can we say about λ_n? In a previous iteration of this course, I falsely asserted that the same proof tells us that

$$\lambda_n \geq d + 2\sqrt{d-1} - \frac{2\sqrt{d-1}-1}{k+1}.$$

But, the proof did not work.

Another question is how well a graph of average degree d can approximate the complete graph. That is, let G be a graph with $dn/2$ edges, but let G be irregular. While I doubt that irregularity helps one approximate the complete graph, I do not know how to prove it.

We can generalize this question further. Let $G = (V, E, w)$ be a weighted graph with $dn/2$ edges. Can we prove that G cannot approximate a complete graph any better than the Ramanujan graphs do? I conjecture that for every d and every $\beta > 0$ there is an n_0 so that for every graph of average degree d on $n \geq n_0$ vertices,

$$\frac{\lambda_2}{\lambda_n} \leq \frac{d - 2\sqrt{d-1}}{d + 2\sqrt{d-1}} + \beta.$$

References

[AC88] Noga Alon and Fan Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72:15–19, 1988.

[BMS93] Richard Beigel, Grigorii Margulis, and Daniel A. Spielman. Fault diagnosis in a small constant number of parallel testing rounds. In *SPAA '93: Proceedings of the fifth annual ACM symposium on Parallel algorithms and architectures*, pages 21–29, New York, NY, USA, 1993. ACM.

[BSS12] Joshua Batson, Daniel A Spielman, and Nikhil Srivastava. Twice-Ramanujan sparsifiers. *SIAM Journal on Computing*, 41(6):1704–1721, 2012.

[Fri08] Joel Friedman. *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*. Number 910 in Memoirs of the American Mathematical Society. American Mathematical Society, 2008.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[Mar88] G. A. Margulis. Explicit group theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problems of Information Transmission*, 24(1):39–46, July 1988.

[Nil91] A. Nilli. On the second eigenvalue of a graph. *Discrete Math*, 91:207–210, 1991.

[Tan84] R. Michael Tanner. Explicit concentrators from generalized n-gons. *SIAM Journal Alg. Disc. Meth.*, 5(3):287–293, September 1984.