Spectral Graph Theory Lecture 22

Expected Characteristic Polynomials

Daniel A. Spielman November 18, 2015

Disclaimer

These notes are not necessarily an accurate representation of what happened in class. The notes written before class say what I think I should say. I sometimes edit the notes after class to make them way what I wish I had said.

There may be small mistakes, so I recommend that you check any mathematically precise statement before using it in your own work.

These notes were last revised on December 8, 2015.

22.1 Overview

Over the next few lectures, we will see two different proofs that infinite families of bipartite Ramanujan graphs exist. Both proofs will use the theory of interlacing polynomials, and will consider the expected characteristic polynomials of random matrices. In today's lecture, we will see a proof that some of these polynomials are real rooted.

At present, we do not know how to use these techniques to prove the existence of infinite families of non-bipartite Ramanujan graphs.

The material in today's lecture comes from [MSS15], but the proof is inspired by the treatment of that work in [HPS15].

22.2 Random sums of graphs

We will build Ramanujan graphs on n vertices of degree d, for every d and even n. We begin by considering a random graph on n vertices of degree d. When n is even, the most natural way to generate such a graph is to choose d perfect matchings uniformly at random, and to then take their sum. I should mention one caveat: some edge could appear in many of the matchings. In this case, we add the weights of the corresponding edges together. So, the weight of an edge is the number of matchings in which it appears.

Let $\boldsymbol{M}$ be the adjacency matrix of some perfect matching on n vertices. We can generate the adjacency matrix of a random perfect matching by choosing a permutation matrix Π uniformly at random, and then forming $\Pi \boldsymbol{M} \Pi^T$. The sum of d independent uniform random perfect machings is then

$$\sum_{i=1}^{d} \Pi_i \boldsymbol{M} \Pi_i^T.$$

In today's lecture, we will consider the *expected characteristic polynomial* of such a graph. For a matrix $\boldsymbol{M}$, we let

$$\chi_x(\boldsymbol{M}) \stackrel{\text{def}}{=} \det(x\boldsymbol{I} - \boldsymbol{M})$$

denote the characteristic polynomial of $\boldsymbol{M}$ in the variable x.

For simplicity, we will consider the expected polynomial of the sum of just two graphs. For generality, we will let them be any graphs, or any symmetric matrices.

Our goal for today is to prove that these expected polynomials are real rooted.

Theorem 22.2.1. *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be symmetric n-by-n matrices and let Π be a uniform random permutation. Then,*

$$\mathbb{E}_{\Pi}\left[\chi_x(\boldsymbol{A} + \Pi \boldsymbol{B} \Pi^T)\right]$$

has only real roots.

So that you will be surprised by this, I remind you that the sum of real rooted polynomials might have no real roots. For example, both $(x-2)^2$ and $(x+2)^2$ have only real roots, but their sum, $2x^2 + 8$, has no real roots.

Theorem 22.2.1 also holds for sums of many matrices. But, for simplicity, we restrict ourselves to considering the sum of two.

22.3 Interlacing

Our first tool for establishing real rootedness of polynomials is interlacing.

If $p(x)$ is a real rooted polynomial of degree n and $q(x)$ is a real rooted polynomial of degree $n-1$, then we say that p and q *interlace* if p has roots $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and q has roots $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n-1}$ that satisfy

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \cdots \geq \lambda_{n-1} \geq \mu_{n-1} \geq \lambda_n.$$

We have seen two important examples of interlacing in this class so far. A real rooted polynomial and its derivative interlace. Similarly, the characteristic polynomial of a symmetric matrix and the characteristic polynomial of a principal submatrix interlace.

When p and q have the same degree, we also say that they interlace if their roots alternate. But, now there are two ways in which their roots can do so, depending on which polynomial has the largest root. If

$$p(x) = \prod_{i=1}^{n}(x - \lambda_i) \quad \text{and} \quad q(x) = \prod_{i=1}^{n}(x - \mu_i),$$

we write $q \to p$ if p and q interlace and for every i the ith root of p is at least as large as the ith root of q. That is, if

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \lambda_n \geq \mu_n.$$

Lemma 22.3.1. *Let p and q be polynomials of degree n and $n-1$ that interlace and have positive leading coefficients. For every $t > 0$, define $p_t(x) = p(x) - tq(x)$. Then, $p_t(x)$ is real rooted and*

$$p(x) \to p_t(x).$$

Proof Sketch. For simplicity, I consider the case in which all of the roots of p and q are distinct. One can prove the general case by dividing out the common repeated roots.

To see that the largest root of p_t is larger than λ_1, note that $q(x)$ is positive for all $x > \mu_1$, and $\lambda_1 > \mu_1$. So, $p_t(\lambda_1) = p(\lambda_1) - tq(\lambda_1) < 0$. As p_t is monic, it is eventually positive and it must have a root larger than λ_1.

We will now show that for every $i \geq 1$, p_t has a root between λ_{i+1} and λ_i. As this gives us $d-1$ more roots, it accounts for all d roots of p_t. For i odd, we know that $q(\lambda_i) > 0$ and $q(\lambda_{i+1}) < 0$. As p is zero at both of these points, $p_t(\lambda_i) > 0$ and $p_t(\lambda_{i+1}) < 0$, which means that p_t has a root between λ_i and λ_{i+1}. The case of even i is similar. □

The converse of this theorem is also true.

Lemma 22.3.2. *Let p and q be polynomials of degree n and $n-1$, and let $p_t(x) = p(x) - tq(x)$. If p_t is real rooted for all $t \in \mathbb{R}$, then p and q interlace.*

Proof Sketch. Recall that the roots of a polynomial are continuous functions of its coefficients, and thus the roots of p_t are continuous functions of t. We will use this fact to obtain a contradiction.

For simplicity,[1] I again just consider the case in which all of the roots of p and q are distinct.

If p and q do not interlace, then p must have two roots that do not have a root of q between them. Let these roots of p be λ_{i+1} and λ_i. Assume, without loss of generality, that both p and q are positive between these roots. We now consider the behavior of p_t for positive t.

As we have assumed that the roots of p and q are distinct, q is positive at these roots, and so p_t is negative at λ_{i+1} and λ_i. If t is very small, then p_t will be close to p in value, and so there must be some small t_0 for which $p_{t_0}(x) > 0$ for some $\lambda_{i+1} < x < \lambda_i$. This means that p_{t_0} must have two roots between λ_{i+1} and λ_i.

As q is positive on the entire closed interval $[\lambda_{i+1}, \lambda_i]$, when t is large p_t will be negative on this entire interval, and thus have no roots inside. As we vary t between t_0 and infinity, the two roots at t_0 must vary continuously and cannot cross λ_{i+1} or λ_i. This means that they must become complex, contradicting our assumption that p_t is always real rooted. □

Together, Lemmas 22.3.1 and 22.3.2 are known as Obreschkoff's Theorem [Obr63].

The following example will be critical.

[1] I thank Sushant Sachdeva for helping me work out this particularly simple proof.

Lemma 22.3.3. *Let $\boldsymbol{A}$ be an n-dimensional symmetric matrix and let $\boldsymbol{v}$ be a vector. Let*

$$p_t(x) = \chi_x(\boldsymbol{A} + t\boldsymbol{v}\boldsymbol{v}^T).$$

Then there is a degree $n-1$ polynomial $q(x)$ so that

$$p_t(x) = \chi_x(\boldsymbol{A}) - tq(x).$$

Proof. Consider the case in which $\boldsymbol{v} = \boldsymbol{\delta}_1$. It suffices to consider this case as determinants, and thus characteristic polynomials, are unchanged by multiplication by rotation matrices.

Then, we know that

$$\chi_x(\boldsymbol{A} + t\boldsymbol{\delta}_1\boldsymbol{\delta}_1^T) = \det(x\boldsymbol{I} - \boldsymbol{A} - t\boldsymbol{\delta}_1\boldsymbol{\delta}_1^T).$$

Now, the matrix $t\boldsymbol{\delta}_1\boldsymbol{\delta}_1^T$ is zeros everywhere except for the element t in the upper left entry. So,

$$\det(x\boldsymbol{I} - \boldsymbol{A} - t\boldsymbol{\delta}_1\boldsymbol{\delta}_1^T) = \det(x\boldsymbol{I} - \boldsymbol{A}) - t\det(x\boldsymbol{I}^{(1)} - \boldsymbol{A}^{(1)}) = \chi_x(\boldsymbol{A}) - t\chi_x(\boldsymbol{A}^{(1)}),$$

where $\boldsymbol{A}^{(1)}$ is the submatrix of $\boldsymbol{A}$ obtained by removing its first row and column. □

We know that $\chi_x(\boldsymbol{A} + t\boldsymbol{v}\boldsymbol{v}^T)$ is real rooted for all t, and we can easily show using the Courant Fischer Theorem that for $t > 0$ it interlaces $\chi_x(\boldsymbol{A})$ from above. Lemmas 22.3.1 and 22.3.2 tell us that these facts imply each other.

We need one other fact about interlacing polynomials.

Lemma 22.3.4. *Let $p_0(x)$ and $p_1(x)$ be two degree n monic polynomials for which there is a third polynomial $r(x)$ that has the same degree as p_0 and p_1 and so that*

$$p_0(x) \rightarrow r(x) \quad \textit{and} \quad p_1(x) \rightarrow r(x).$$

Then for all $0 \leq s \leq 1$,

$$p_s(x) \stackrel{\text{def}}{=} sp_1(x) + (1-s)p_0(x)$$

is a real rooted polynomial.

Sketch. Assume for simplicity that all the roots of r are distinct. Let $\mu_1 > \mu_2 > \cdots > \mu_n$ be the roots of r. Our assumptions imply that both p_0 and p_1 are positive at μ_i for odd i and negative for even i. So, the same is true of their sum p_s. This tells us that p_s must have at least $n-1$ real roots.

We can also show that p_s has a root that is less than μ_n. One way to do it is to recall that the complex roots of a polynomial with real coefficients come in conjugate pairs. So, p_s can not have only one complex root. □

22.4 Sums of polynomials

Our goal is to show that

$$\sum_{\Pi \in S_n} \chi_x(\boldsymbol{A} + \Pi \boldsymbol{B} \Pi^T)$$

is a real rooted polynomial for all symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, where S_n is the set of n-by-n permuation matrices. We will do this by proving it for smaller sets of permutation matrices. To begin, we know it for $S = \{\boldsymbol{I}\}$. We will build up larger sets by swapping coordinates.

This will actually result in a distribution on permuations, so we consider $\sigma : S_n \to \mathbb{R}_{\geq 0}$ and consider sums of the form

$$\sum_{\Pi} \sigma(\Pi) \chi_x(\boldsymbol{A} + \Pi \boldsymbol{B} \Pi^T).$$

For coordinates i and j, let $\Gamma_{i,j}$ be the permutation matrix that just swaps i and j. We call such a permutation a swap. We need the following important fact about the action of swaps on matrices.

Lemma 22.4.1. *Let $\boldsymbol{A}$ be a symmetric matrix. Then, for all i and j, there are vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ so that*

$$\Gamma_{i,j} \boldsymbol{A} \Gamma_{i,j} = \boldsymbol{A} - \boldsymbol{u}\boldsymbol{u}^T + \boldsymbol{v}\boldsymbol{v}^T.$$

Proof. Without loss of generality, let $i = 1$ and $j = 2$. We prove that $\boldsymbol{A} - \Gamma_{i,j} \boldsymbol{A} \Gamma_{i,j}$ has rank 2 and trace 0.

We can write this difference in the form

$$\begin{bmatrix} a_{11} - a_{22} & a_{12} - a_{21} & a_{13} - a_{23} & a_{14} - a_{24} & \dots \\ a_{21} - a_{12} & a_{22} - a_{11} & a_{23} - a_{13} & a_{24} - a_{14} & \dots \\ a_{31} - a_{32} & a_{32} - a_{31} & 0 & \dots & \\ a_{41} - a_{42} & a_{42} - a_{41} & 0 & \dots & \\ \dots & & & & \end{bmatrix} = \begin{bmatrix} \alpha & \beta & \boldsymbol{y}^T \\ -\beta & -\alpha & -\boldsymbol{y}^T \\ \boldsymbol{y} & -\boldsymbol{y} & 0_{n-2} \end{bmatrix}$$

for some numbers α, β and some column vector $\boldsymbol{y}$ of length $n - 2$. If $\alpha \neq \beta$ then the sum of the first two rows is equal to $(c, -c, 0, \dots, 0)$ for some $c \neq 0$, and every other row is a scalar multiple of this. On the other hand, if $\alpha = \beta$ then the first two rows are linearly dependent, and all of the other rows are multiples of $(1, -1, 0, \dots, 0)$. □

Lemma 22.4.2. *Let σ be such that for all symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$,*

$$p_x(\boldsymbol{A}, \boldsymbol{B}) \stackrel{\text{def}}{=} \sum_{\Pi \in S} \sigma(\Pi) \chi_x(\boldsymbol{A} + \Pi \boldsymbol{B} \Pi^T)$$

is real rooted. Then, for every $0 < s < 1$ and pair of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, for every symmetric $\boldsymbol{A}$ and $\boldsymbol{B}$ the polynomial

$$(1 - s) p_x(\boldsymbol{A}, \boldsymbol{B}) + s p_x(\boldsymbol{A} - \boldsymbol{u}\boldsymbol{u}^T + \boldsymbol{v}\boldsymbol{v}^T, \boldsymbol{B})$$

is real rooted.

Proof. Define

$$r_t(x) = p_x(\boldsymbol{A} + t\boldsymbol{v}\boldsymbol{v}^T, \boldsymbol{B}).$$

By assumption, $r_t(x)$ is real rooted for every $t \in \mathbb{R}$. By Lemma 22.3.3, we can write

$$r_t(x) = r_0(x) - tq(x),$$

where $q(x)$ has degree $n-1$ and both r_0 and q have positive leading coefficients. So, by Lemma 22.3.2 $q(x)$ interlaces $r_0(x) = p_x(\boldsymbol{A}, \boldsymbol{B})$. Lemma 22.3.1 then tells us that

$$p_x(\boldsymbol{A}, \boldsymbol{B}) \to p_x(\boldsymbol{A} + \boldsymbol{v}\boldsymbol{v}^T, \boldsymbol{B}).$$

The same argument tells us that

$$p_x(\boldsymbol{A} - \boldsymbol{u}\boldsymbol{u}^T + \boldsymbol{v}\boldsymbol{v}^T, \boldsymbol{B}) \to p_x(\boldsymbol{A} + \boldsymbol{v}\boldsymbol{v}^T, \boldsymbol{B}).$$

This tells us that $p_x(\boldsymbol{A}, \boldsymbol{B})$ and $p_x(\boldsymbol{A} - \boldsymbol{u}\boldsymbol{u}^T + \boldsymbol{v}\boldsymbol{v}^T, \boldsymbol{B})$ both interlace $r_1(x)$ from below. We finish by applying Lemma 22.3.4 to conclude that every convex combination of these polynomials is real rooted. □

Corollary 22.4.3. *Let σ be such that for all symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$,*

$$p_x(\boldsymbol{A}, \boldsymbol{B}) \stackrel{\text{def}}{=} \sum_{\Pi \in S} \sigma(\Pi) \chi_x(\boldsymbol{A} + \Pi \boldsymbol{B} \Pi^T)$$

is real rooted. Then, for every $0 < s < 1$ and for every symmetric $\boldsymbol{A}$ and $\boldsymbol{B}$ the polynomial

$$\sum_{\Pi \in S} s\sigma(\Pi)\chi_x(\boldsymbol{A} + \Pi \boldsymbol{B} \Pi^T) + (1-s)\sigma(\Pi)\chi_x(\boldsymbol{A} + \Gamma_{i,j}\Pi \boldsymbol{B} \Pi^T \Gamma_{i,j}^T)$$

is real rooted.

Proof. Recall that

$$\chi_x(\boldsymbol{A} + \Gamma_{i,j}\Pi \boldsymbol{B} \Pi^T \Gamma_{i,j}^T) = \chi_x(\Gamma_{i,j}^T \boldsymbol{A} \Gamma_{i,j} + \Pi \boldsymbol{B} \Pi^T) = \chi_x(\Gamma_{i,j} \boldsymbol{A} \Gamma_{i,j}^T + \Pi \boldsymbol{B} \Pi^T).$$

The corollary now follows from the previous lemma. □

22.5 Random Swaps

We will build a random permutation out of random swaps. A random swap is specified by coordinates i and j and a swap probability s. It is a random matrix is that is equal to the identity with probability $1-s$ and $\Gamma_{i,j}$ with probability s. Let $\boldsymbol{S}$ be a random swap.

In the language of random swaps, we can express Corollary 22.5.1 as follows.

Corollary 22.5.1. *Let* Π *be a random permutation matrix drawn from a distribution so that for all symmetric matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$*,*

$$\mathbb{E}\left[\chi_x(\boldsymbol{A} + \Pi\boldsymbol{B}\Pi^T)\right]$$

is real rooted. Let $\boldsymbol{S}$ *be a random swap. Then,*

$$\mathbb{E}\left[\chi_x(\boldsymbol{A} + \boldsymbol{S}\Pi\boldsymbol{B}\Pi^T\boldsymbol{S}^T)\right]$$

is real rooted for every symmetric $\boldsymbol{A}$ *and* $\boldsymbol{B}$*.*

All that remains is to show that a uniform random permutation can be assembled out of random swaps. The trick to doing this is to choose the random swaps with swap probabilities other than 1/2. If you didn't do this, it would be impossible as there are $n!$ permutations, which is not a power of 2.

Lemma 22.5.2. *For every* n*, there exists a finite sequence of random swaps* $\boldsymbol{S}_1, \ldots, \boldsymbol{S}_k$ *so that*

$$\boldsymbol{S}_1\boldsymbol{S}_2\ldots\boldsymbol{S}_k$$

is a uniform random permutation.

Proof. We prove this by induction. We can generate a random permutation on $1, \ldots, n$ by first choosing which item maps to n, and then generating a random permutation on those that remain. To this end, we first form a sequence that gives a random permtuation on the first $n-1$ elements. We then compose this with a random swap that exchanges elements 1 and n with probability $1 - 1/n$. At this point, the element that maps to n will be uniformly random. We then compose with yet another sequence that gives a random permutation on the first $n-1$ elements. □

References

[HPS15] Chris Hall, Doron Puder, and William F Sawin. Ramanujan coverings of graphs. *arXiv preprint arXiv:1506.02335*, 2015.

[MSS15] Adam W Marcus, Nikhil Srivastava, and Daniel A Spielman. Interlacing families IV: Bipartite Ramanujan graphs of all sizes. *arXiv preprint arXiv:1505.08010*, 2015. appeared in Proceedings of the 56th IEEE Symposium on Foundations of Computer Science.

[Obr63] Nikola Obrechkoff. *Verteilung und berechnung der Nullstellen reeller Polynome.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1963.