Chapter 12

Method of Interlacing Polynomials

Marcus, Spielman, and Srivastava [MSS14] turned the heuristic argument from [BSS14] about expected characteristic polynomial described in Section 10.1 into a powerful probabilistic method. We have already previewed this method in a simple form in Section 11.2 without seeing the details. In this chapter, we will go through the relevant concepts and describe the method in its general form. Then we will see an interesting and relatively simple application to the restricted invertibility problem, in two different ways.

12.1 New Probabilistic Method

In standard probabilistic method, we compute the expectation of a random variable $\mathbb{E}[X]$, and then conclude that there is an outcome in the sample space with value at least or at most $\mathbb{E}[X]$. Consider the minimum eigenvalue problem in Theorem 11.2, in which the quantity of interest is $\lambda_{\min}(\sum_{i \in S} v_i v_i^T)$ for some multi-subset S with $|S| = k$. To prove that there is a multi-subset with large minimum eigenvalue, the standard way is to compute $\mu := \mathbb{E}_{S:|S|=k}\big[\lambda_{\min}\big(\sum_{i \in S} v_i v_i^T\big)\big]$ and then conclude that there is a mult-subset S with $|S| = k$ and $\lambda_{\min}(\sum_{i \in S} v_i v_i^T) \geq \mu$.

Marcus, Spielman, and Srivastava took an unusual route to solve this kind of problems. First, instead of working with the random matrix $A = \sum_{i \in S} v_i v_i^T$ directly, they consider the characteristic polynomial $p_A(x) = \det(xI - A)$ of the random matrix. Note that $\lambda_{\min}(A)$ is simply the minimum root of the characteristic polynomial $\lambda_{\min}(p_A)$. Then, quite surprisingly, instead of computing the expected minimum eigenvalue of a random characteristic polynomial $\mathbb{E}_A[\lambda_{\min}(p_A)]$, they compute the minimum eigenvalue of the expected polynomial $\lambda_{\min}\big(\mathbb{E}_A[p_A]\big)$. The following is an instantiation of their new probabilistic method for the minimum eigenvalue problem, when each vector is chosen independently and uniformly randomly.

Proposition 12.1 (Probabilistic Method for Minimum Eigenvalue)**.** *Suppose $v_1, \ldots, v_m \in \mathbb{R}^n$ are vectors with $\sum_{i=1}^m v_i v_i^T = I_n$. For any $k \geq n$, suppose $r_1, \ldots, r_k$ are independent uniformly random vectors in $\{v_1, \ldots, v_m\}$ and let $A := \sum_{i=1}^k r_i r_i^T$ be a random matrix. Then, with positive probability,*

$$\lambda_{\min}(p_A) \geq \lambda_{\min}\big(\mathbb{E}[p_A]\big).$$

In general, $\mathbb{E}[\lambda_{\min}(p_A)] \neq \lambda_{\min}(\mathbb{E}[p_A])$, and in fact the latter term could be bigger than the former term, and so this proposition is not trivial at all.

Characteristic polynomials have not played an important role in much of spectral graph theory. One disadvantage for instance is that the information about the eigenvectors is lost. Very interestingly, the method by Marcus, Spielman, and Srivastava showed that they often satisfy a number of very nice algebraic identities and are amenable to a set of very elegant analytic tools that do not naturally apply to matrices.

12.2 Interlacing Polynomials

Let $p_1, \ldots, p_m$ be real-rooted polynomials and $q = \sum_{i=1}^m \mu_i p_i$ be a convex combination of $p_1, \ldots, p_m$ where $\sum_{i=1}^m \mu_i = 1$ and $\mu_i \geq 0$ for $1 \leq i \leq m$. Under what conditions can we conclude that say $\max_i \lambda_{\min}(p_i) \geq \lambda_{\min}(q)$? In general, it could be far from true. For example, $p_1 = (x-1)(x-2)$ and $p_2 = (x-3)(x-4)$ are both real-rooted, but their average $\frac{1}{2}(p_1 + p_2)$ is not even real-rooted. Even if assume $p_1 + p_2$ is real-rooted, there is in general no simple relationship between the roots of two polynomials and the roots of their average.

The main insight of Marcus, Spielman and Srivastava is that in several problems of interest, the (characteristic) polynomials satisfy some interlacing properties that would allow us to conclude that $\max_i \lambda_{\min}(p_i) \geq \lambda_{\min}(q)$.

Definition 12.2 (Interlacing Polynomials). *Let p be a degree n polynomial with real roots $\alpha_1 \geq \ldots \geq \alpha_n$ and let q be a degree n or $n-1$ polynomial with real roots $\beta_1 \geq \ldots \geq \beta_n$ (ignoring β_n in the degree $n-1$ case). We say that q interlaces p if their roots alternate and the largest root belongs to p such that*

$$\alpha_1 \geq \beta_1 \geq \alpha_2 \geq \beta_2 \geq \ldots \beta_{n-1} \geq \alpha_n \geq \beta_n.$$

Definition 12.3 (Common Interlacing). *A set of degree n real-rooted polynomials $p_1, \ldots, p_m$ is said to have a common interlacing if there is a polynomial q that interlaces each p_i for $1 \leq i \leq m$.*

Equivalently, $p_1, \ldots, p_m$ have a common interlacing if there are inner-disjoint intervals $I_1 \geq I_2 \geq \ldots \geq I_n$ on the real line such that the k-th largest root of each p_i for $1 \leq i \leq m$ is contained in I_k

An important class of interlacing polynomials are characteristic polynomials of matrices under rank-one updates. The following is also called Cauchy's interlacing theorem, and one can prove it in a similar way as in Cauchy's interlacing Theorem 2.13, using Courant-Fischer Theorem 2.12.

Exercise 12.4 (Cauchy's Interlacing Theorem). *Let $A \in \mathbb{R}^{n\times n}$ be a real symmetric matrix and $v \in \mathbb{R}^n$. Then p_A interlaces p_{A+vv^T}.*

Note that Exercise 12.4 implies that if $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix and $v_1, \ldots, v_m \in \mathbb{R}^n$, then $p_{A+v_1v_1^T}, \ldots, p_{A+v_mv_m^T}$ have a common interlacing.

Common Interlacing and Probabilistic Method

If $p_1, \ldots, p_m$ have a common interlacing, then any convex combination q of $p_1, \ldots, p_m$ is also real-rooted and we can compare the roots of $p_1, \ldots, p_m$ with the roots of q. The proof is a simple application of the intermediate value theorem in the interval I_j for the j-th root for each j.

Theorem 12.5 (Probabilistic Method for Common Interlacing Polynomials). *Suppose $p_1, \ldots, p_m$ are real-rooted polynomials of degree n with positive leading coefficients. Let $\lambda_k(p_j)$ be the k-th largest root of p_j. If $p_1, \ldots, p_m$ have a common interlacing, then for any non-negative numbers $\mu_1, \ldots, \mu_m$ with $\sum_{i=1}^m \mu_i = 1$ and for any $1 \leq k \leq n$,*

$$\min_j \lambda_k(p_j) \leq \lambda_k\big(\mathbb{E}_{j\sim\mu}[p_j]\big) \leq \max_j \lambda_k(p_j)$$

Proof. Let $q = \sum_{i=1}^{m} \mu_i p_i$. Let $u = \max_j \lambda_{\max}(p_j)$ and $l = \min_j \lambda_{\max}(p_j)$. We would like to argue that $\lambda_{\max}(q)$ is contained in $[l, u]$.

First, we argue that $\lambda_{\max}(q) \leq u$. As $p_1, \ldots, p_m$ all have positive leading coefficients, all polynomials are positive in the range (u, ∞). As q is a convex combination of $p_1, \ldots, p_m$, so q is also positive in the range (u, ∞). Therefore, q cannot have a root in the range (u, ∞), and thus $\lambda_{\max}(q) \leq u$.

Next, we argue that $\lambda_{\max}(q) \geq l$. If $l = u$, then $p_j(u) = 0$ for all $1 \leq j \leq m$, and thus $q(u) = 0$ and hence $\lambda_{\max}(q) = u \geq l$. Henceforth we assume $l < u$. On one hand, note that $q(u) > 0$ as each $p_j(u) \geq 0$ and there exists i with $p_i(u) > 0$ (e.g. the one with $\lambda_{\max}(p_i) = l < u$). On the other hand, since $p_1, \ldots, p_m$ have a common interlacing, $\lambda_2(p_j) \leq l$ for each j, and since $p_1, \ldots, p_m$ all have positive leading coefficients, each polynomial p_j is non-positive in the range $[\lambda_2(p_j), \lambda_1(p_j)]$ with $\lambda_2(p_j) \leq l$ and $\lambda_1(p_j) \geq l$ for all $1 \leq j \leq m$. Therefore, $p_j(l) \leq 0$ for all $1 \leq j \leq m$, and thus $q(l) \leq 0$. Since $q(u) > 0$ and $q(l) \leq 0$, by the intermediate value theorem, there exists $r \in [l, u)$ such that $q(r) = 0$, and therefore $\lambda_{\max}(q) \geq l$.

A similar argument works for any $1 \leq k \leq n$ and is left to the reader (see Lemma 2.11 of [MSS21]). It may be more convenient for the argument to first reduce to the case when $p_1, \ldots, p_m$ have no common roots. □

So, if we could show that a set of polynomials have a common interlacing, then we can apply Theorem 12.5 to show that there exists a polynomial with large/small k-th largest root by showing that *some* weighted average polynomial has large/small k-th largest root.

Common Interlacing and Real-Rootedness

We are thus interested in some general techniques to prove that a set of polynomials have a common interlacing. Note that Theorem 12.5 proves that if $p_1, \ldots, p_m$ are real-rooted and have a common interlacing, then any convex combination of $p_1, \ldots, p_m$ is also real-rooted. It turns out that the converse is also true. This gives us a characterization when a set of real-rooted polynomials have a common interlacing. We use the following simple fact in the proof.

Exercise 12.6 (Common Interlacing is a Pairwise Property)**.** *A set of polynomials $p_1, \ldots, p_m$ have a common interlacing if and only if each pair of polynomials p_i, p_j have a common interlacing for all $1 \leq i \neq j \leq n$.*

We also use the following well-known result from elementary complex analysis without proof.

Theorem 12.7 (Continuity of Roots)**.** *The roots of a polynomial are continuous functions of its coefficients.*

Theorem 12.8 (Common Interlacing and Real-Rootedness)**.** *If $p_1, \ldots, p_m$ are degree n polynomials and all of their convex combinations $\sum_{i=1}^{m} \mu_i p_i$ are real-rooted, then $p_1, \ldots, p_m$ have a common interlacing.*

Proof. By Exercise 12.6, we only need to prove the lemma for two polynomials. We assume without loss of generality that p_1 and p_2 have no common roots, as otherwise we can just divide both polynomials by their common factors, prove that the resulting polynomials have a common interlacing, and conclude that the original polynomials also have a common interlacing.

Let $q_\mu = (1-\mu)\cdot p_1 + \mu \cdot p_2$ for $\mu \in [0,1]$. If we keep track of the roots of q_μ from $\mu = 0$ and $\mu = 1$ as a continuous function of μ, then each root of q_μ is a continuous curve on the complex plane as μ varies from 0 to 1 by Theorem 12.7. Since each q_μ is real-rooted by assumption, the curve of each root j is an interval J_j on the real line, with one endpoint being a root of p_1 and the other endpoint being a root of p_2.

We would like to argue that these intervals are pairwise inner-disjoint (i.e. they are disjoint except possibly at the endpoints). Suppose to the contrary that this is not the case, that one endpoint of an interval is contained in the interior of some other interval. This implies that some root r of a polynomial, say p_1, is a root of q_μ for some $0 < \mu < 1$, but then

$$0 = q_\mu(r) = (1-\mu)\cdot p_1(r) + \mu \cdot p_2(r) = \mu \cdot p_2(r) \quad \implies \quad p_2(r) = 0,$$

contradicting that p_1 and p_2 have no common roots. Therefore, these intervals are pairwise inner-disjoint. This implies that the intervals can be arranged so that $J_1 \geq J_2 \geq \ldots \geq J_n$, and thus p_1 and p_2 have a common interlacing. □

By Theorem 12.8, to prove a set of polynomials have a common interlacing (in order to apply the probabilistic method), it is equivalent to proving that all convex combinations of any two polynomials are real-rooted. In the next chapter, we will study methods to prove that a polynomial is real-rooted.

Differential Operators and Common Interlacing

In Section 10.1 and in Section 11.2, the expected characteristic polynomials are of the form $(1-s\partial)p$ for some scalar s. With the results in the previous subsections, we can show that this differential operator preserves real-rootedness and also common interlacing.

Problem 12.9 (Differential Operators and Common Interlacing). *Prove that if p is real-rooted, then $(1-s\partial)p$ is also real-rooted. Furthermore, prove that if $p_1, \ldots, p_m$ have positive leading coefficients and a common interlacing, then $(1-s\partial)p_1, \ldots, (1-s\partial)p_m$ also have a common interlacing.*

With Problem 12.9, it should be straightforward to solve Problem 11.11, and thus completing the proof of Theorem 11.2 using a simple version of the method of interlacing polynomials.

12.3 Interlacing Family

Recall that in Proposition 12.1, the goal is to prove that there is a positive probability that a random matrix $A = \sum_{i=1}^k r_i r_i^T$ satisfies $\lambda_{\min}(p_A) \geq \lambda_{\min}(\mathbb{E}[p_A])$, where there are m^k possibilities of A. To prove this statement by directly applying the probabilistic method in Theorem 12.5, we need to prove that these m^k different characteristic polynomials have a common interlacing. A moment of thought reveals that this is clearly not true in general.

The idea of Marcus, Spielman, and Srivastava is to build a tree structure among these polynomials and show that the children of each internal node have a common interlacing. This idea is similar to the method of conditional expectations used in derandomization.

Definition 12.10 (Interlacing Family). *An interlacing family consists of a finite rooted tree T and a labeling of the nodes $v \in T$ by monic real-rooted polynomials $p_v(x) \in \mathbb{R}[x]$, with two properties:*

1. *Every polynomial $p_v(x)$ corresponding to a non-leaf node v is a convex combination of the polynomials corresponding to the children of v.*
2. *For all nodes $v_1, v_2 \in T$ with a common parent, all convex combinations of $p_{v_1}(x)$ and $p_{v_2}(x)$ are real-rooted.*

We say that a set of polynomials is an interlacing family if they are the labels of the leaves of such a tree.

Note that, by Theorem 12.8 and Exercise 12.6, the second condition implies that all the children have a common interlacing, and it follows from Theorem 12.5 that all convex combinations of all children are real-rooted.

The above definition may look a bit abstract, but in applications the root polynomial will usually simply be the average polynomial of all the leaves, while the internal nodes will usually simply be the average polynomial of the leaves of the corresponding subtrees. Let us see a concrete example that is useful for the minimum eigenvalue problem in Theorem 11.2.

Example 12.11 (Interlacing Family of Multi-Subset of k Vectors). *Let $v_1, \ldots, v_m \in \mathbb{R}^n$. For any $s_1, \ldots, s_k \in [m]$, define*

$$p_{s_1,\ldots,s_k}(x) := \det\left(xI_n - \sum_{i=1}^{k} v_{s_i} v_{s_i}^T\right).$$

The tree T is a complete m-ary tree, with depth k, and thus m^k leaves. Each leaf of the tree is labeled by a sequence $s_1, \ldots, s_k$, representing a path from the root to the leaf, where s_i represents the s_i-th child of the internal node in the $(i-1)$-th level, with the root being in the 0-th level. The polynomials in the internal nodes are defined inductively as

$$p_{s_1,\ldots,s_t}(x) = \frac{1}{m}\sum_{j=1}^{m} p_{s_1,\ldots,s_t,j}(x) = \frac{1}{m^{k-t}} \sum_{s_{t+1},\ldots,s_k} p_{s_1,\ldots,s_k}(x)$$

for any $t < k$ and the root polynomial is

$$p_{\emptyset}(x) = \frac{1}{m^k} \sum_{s_1,\ldots,s_k \in [m]^k} p_{s_1,\ldots,s_k}(x).$$

We will prove in the next chapter that these polynomials $\mathcal{P} := \left\{p_{s_1,\ldots,s_k}(x)\right\}_{s_1,\ldots,s_k \in [m]^k}$ form an interlacing family.

It may not be easy to establish that a set of polynomials forms an interlacing family, and in some applications the theory of real stable polynomials is needed to prove so, which we will study in the next chapter.

But once we have established that a family is an interlacing family, we can then easily relate the roots of the root-polynomial to the roots of the polynomials in the leaves. The following theorem follows from a simple induction using Theorem 12.5.

Theorem 12.12 (Probabilistic Method for Interlacing Family). *Let $\mathcal{P}$ be an interlacing family of degree n polynomials with root labeled by $p_{\emptyset}(x)$ and leaves by $\{p_l(x)\}_{l \in L}$ where L is the set of leaves. Then, for any $1 \leq j \leq n$, there exist leaves $a \in L$ and $b \in L$ such that*

$$\lambda_j(p_a) \leq \lambda_j(p_{\emptyset}) \leq \lambda_j(p_b).$$

Proof. The proof is by a simple induction on the depth of the internal node. By Theorem 12.8, the second condition in Definition 12.10 implies that every pair of children of the root node have a common interlacing. By Exercise 12.6, it follows that all the children of the root node have a common interlacing. Then, Theorem 12.5 proves that there is a child a_1 of the root node with $\lambda_j(p_{a_1}) \leq \lambda_j(p_\emptyset)$ and there is a child b_1 of the root node with $\lambda_j(p_{b_1}) \geq \lambda_j(p_\emptyset)$. By induction, there is a leaf node a in the subtree of a_1 with $\lambda_j(p_a) \leq \lambda_j(p_{a_1}) \leq \lambda_j(p_\emptyset)$, and there is a leaf node b in the subtree of b_1 with $\lambda_j(p_b) \geq \lambda_j(p_{b_1}) \geq \lambda_j(p_\emptyset)$. □

12.4 Restricted Invertibility

In this section, we see an interesting application of the techniques developed so far to the restricted invertibility problem. This is not the first application of the method of interlacing family, but it is the simplest as it only involves univariate polynomials, and so we present it first to separate the ideas of the interlacing family method from the theory of real-stable (multivariate) polynomials.

The restricted invertibility problem is a well-studied problem in functional analysis, which says that a matrix of high stable rank has a large column submatrix with large smallest singular value. We consider an equivalent formulation that is very close to the minimum eigenvalue problem in Theorem 11.2.

Definition 12.13 (Restricted Invertibility Problem)**.** *Given $v_1, \ldots, v_m \in \mathbb{R}^n$ and an integer $k < n$, find a subset $S \subseteq [m]$ with $|S| = k$ to maximize $\lambda_k\big(\sum_{i \in S} v_i v_i^T\big)$, where $\lambda_k(A)$ denotes the k-th largest eigenvalue of matrix A.*

To illustrate the method of interlacing family, we only consider the special "isotropy" case when $\sum_{i=1}^m v_i v_i^T = I_n$. We remark that, unlike the minimum eigenvalue problem, it is no longer true that the general case can be reduced to this special case, because of $k < n$. Marcus, Spielman, and Srivastava [MSS21] used the method of interlacing family to derive a sharp result in the isotropy case.

Theorem 12.14 (Restricted Invertibility in the Isotropy Case)**.** *Suppose $v_1, \ldots, v_m \in \mathbb{R}^n$ are vectors with $\sum_{i=1}^m v_i v_i^T = I_n$. Then, for every integer $k \leq n$, there exists a subset $S \subset [m]$ with $|S| = k$ and*

$$\lambda_k\left(\sum_{i \in S} v_i v_i^T\right) \geq \left(1 - \sqrt{\frac{k}{n}}\right)^2 \frac{n}{m}.$$

Although this result is sharp for a large regime of k, we do not know whether it is tight when $k \approx n$. The following question is closely related to Question 11.15

Question 12.15 (Restricted Invertibility when $k \approx d$)**.** *When $m = O(n)$ and $k = n - 1$, the lower bound in Theorem 12.14 is $\Omega(1/n^2)$. Is this tight? To my knowledge, the best lower bound that we can hope for in this regime is $\Omega(1/n)$.*

Ravichandran [Rav18] presented a different way to use the interlacing family method to derive the results in [MSS21], with an additional application of proving a quantitative Gauss-Lucas theorem which we may mention later. We will present both approaches, as this will allow us to see two different interlacing families for the problem.

Interlacing Family of Multi-Subset of Vectors

The proof in [MSS21] uses the interlacing family that we described in Example 12.11. We have not proved that it is indeed an interlacing family yet, but we assume it is in this subsection. Then, to apply the probabilistic method in Theorem 12.12, we just need to compute the polynomial in the root of the tree and bound its k-th eigenvalue. The calculations for the expected characteristic polynomial in Section 10.1 and in Exercise 11.8 can be used to compute the root polynomial from the leaves up.

Exercise 12.16 (Root Polynomial in Example 12.11)**.** *When $\sum_{i=1}^m v_i v_i^T = I_n$, the root polynomial $p_\emptyset$ in Example 12.11 is*

$$p_\emptyset(x) = \left(1 - \frac{1}{m}\partial_x\right)^k x^n$$

Note that the k-th largest root of $p_\emptyset(x)$ is simply the smallest root of the polynomial $x^{-(n-k)}p_\emptyset(x) = x^{-(n-k)}\big(1 - \frac{1}{m}\partial_x\big)^k x^n$. Marcus, Spielman, and Srivastava observed that it is a slight transformation of an associated Laguerre polynomial and a known result by Krasikov implies that

$$\lambda_k(p_\emptyset) \geq \left(1 - \sqrt{\frac{k}{n}}\right)^2 \frac{n}{m}.$$

Therefore, we can conclude from Theorem 12.12 that there is a leaf with the k-th largest eigenvalue at least $\lambda_k(p_\emptyset)$, proving Theorem 12.14.

It is quite amazing that the method of interlacing family reduces the restricted invertibility problem to a pure mathematical problem of bounding the smallest root of a well-known polynomial. So, the heuristic argument in Section 10.1 can indeed be made precise, using the method of interlacing family, at least for the restricted invertibility problem.

Question 12.17 (Polynomial Proof for Spectral Sparsification)**.** *Can you prove the spectral sparsification result in Theorem 10.1 by turning the heuristic argument in Section 10.1 into a precise proof (possibly using the method of interlacing family of polynomials)?*

Interlacing Family of Principle Submatrices

Ravichandran's approach is based on the family of characteristic polynomials of principal submatrices of a matrix. Let $A \in \mathbb{R}^{n\times n}$ be a real symmetric matrix. For $1 \leq i \leq n$, let $A_{\{i\}} \in \mathbb{R}^{(n-1)\times(n-1)}$ be the principal submatrix of A obtained by deleting the i-th row and i-th column of A. Note that the characteristic polynomials $p_{A_{\{1\}}}, \ldots, p_{A_{\{n\}}}$ of $A_{\{1\}}, \ldots, A_{\{n\}}$ have a common interlacing by Cauchy's interlacing Theorem 2.13. So, by Theorem 12.5,

$$\max_i \left\{\lambda_{\max}\big(A_{\{i\}}\big)\right\} \geq \lambda_{\max}\left(\sum_{i=1}^m p_{A_{\{i\}}}\right) \geq \min_i \left\{\lambda_{\max}\big(A_{\{i\}}\big)\right\}.$$

Ravichandran noted that there is a very nice formula for $\sum_{i=1}^m p_{A_{\{i\}}}$ and observed that it can be used to define an interlacing family of the characteristic polynomials of principal submatrices.

Theorem 12.18 (Thompson's Theorem [Tho66])**.** *Let $A \in \mathbb{R}^{n\times n}$ be a real symmetric matrix, and let $A_{\{1\}}, \ldots, A_{\{n\}}$ be the $(n-1)\times(n-1)$ principal submatrices of A. Then*

$$\sum_{i=1}^m p_{A_{\{i\}}} = p_A'.$$

Theorem 12.19 (Interlacing Family of Principal Submatrices [Rav18])**.** *Let $A \in \mathbb{R}^{n\times n}$ be a real symmetric matrix. For $S \subseteq [n]$, let A_S be the principal submatrix of A obtained by deleting the rows and columns of S. For any $0 \leq k \leq n$, the set of characteristic polynomials $\{p_{A_S}\}_{|S|=k}$ forms an interlacing family, with the root polynomial being $p_A^{(k)}(x)$ which is the k-th derivative of $p_A(x)$.*

Proof. There are $\binom{n}{k}$ polynomials in this family. We organize them as the leaves of a tree T of depth k, where each node in T at depth i corresponds to a subset $S \subseteq [n]$ of size i and a principle matrix A_S of size $(n-i) \times (n-i)$. The root is at depth 0, and it corresponds to the empty set and the original matrix $A_\emptyset = A$. The i-th node at depth 1 corresponds to the singleton subset $\{i\}$ and the principal submatrix $A_{\{i\}}$. Inductively, given a node of T at depth i which corresponds to a subset $X \subseteq [n]$ of size i, it has $n-i$ children which correspond to the subsets $X \cup \{j\}$ for each $j \in [n] \setminus X$. The tree then has $n \times (n-1) \times \ldots \times (n-k+1) = k! \cdot \binom{n}{k}$ leaves, where each subset of size k is associated with $k!$ leaves of T (one for each permutation).

Next we define the polynomials in the nodes of T. For a leaf node, let S be the corresponding subset of size k, the polynomial is simply p_{A_S} which is the characteristic polynomial of A_S. Inductively, from the leaves to the root, the polynomial of an internal node of T is defined as the sum of the polynomials of its children.

Now we compute the polynomials in the nodes of T. The leaves at depth k are the base cases. For a node at depth $k-1$, it corresponds to a subset $X \subseteq [n]$ of size $k-1$, with the polynomials in its children being $p_{A_{X\cup\{j\}}}$ for $j \in [n] \setminus X$. By Thompson's Theorem 12.18,

$$\sum_{j\in[n]\setminus X} p_{A_{X\cup\{j\}}} = p'_{A_X}.$$

For a node at depth l which corresponds to a subset $Y \subseteq [n]$ of size l, the induction hypothesis is that the polynomials at its children are $p^{(k-l-1)}_{A_{Y\cup\{j\}}}$ for $j \in [n] \setminus Y$. Then, by Thompson's theorem, the polynomial at this node is

$$\sum_{j\in[n]\setminus Y} p^{(k-l-1)}_{A_{Y\cup\{j\}}} = \Big(\sum_{j\in[n]\setminus Y} p_{A_{Y\cup\{j\}}}\Big)^{(k-l-1)} = \Big(p'_{A_Y}\Big)^{(k-l-1)} = p^{(k-l)}_{A_Y},$$

proving the induction step. Therefore, for the root node, the polynomial is $p_A^{(k)}$ as stated.

Finally, we check that these polynomials satisfy the conditions in Definition 12.10. Property (1) is satisfied as the polynomial at a non-leaf node is the sum of the polynomials of its children, which is the same as the average polynomial up to a scalar which does not change the locations of the roots. For property (2), first we consider the case that the non-leaf node is at depth $k-1$, then the polynomials at its children have a common interlacing by Cauchy's interlacing Theorem 2.13, and thus the second property is satisfied by Theorem 12.8. Note that common interlacing is preserved by the differential operator ∂_x, using the same proof as in Problem 12.9. Therefore, for a node at depth l which corresponds to a subset $Y \subseteq [n]$ of size l, the polynomials $p^{(k-l-1)}_{A_{Y\cup\{j\}}}$ for $j \in [n]-Y$ at its children have a common interlacing, because $p_{A_{Y\cup\{j\}}}$ for $j \in [n]-Y$ have a common interlacing by Cauchy's interlacing theorem and applying the differential operator ∂_x at each of these polynomials (multiple times) preserves the common interlacing property. We conclude that the polynomials at the leaves form an interlacing family, with the root polynomial being $p^{(k)}(A)$. □

As a consequence, the method of interlacing family in Theorem 12.12 implies the following bound.

Theorem 12.20 (Ravichandran's Theorem [Rav18]). *Let $A \in \mathbb{R}^{n \times n}$ be a real symmetric matrix. There exists a subset $S \subset [n]$ with $|S| = k$ and*

$$\lambda_{\max}(A_S) \leq \lambda_{\max}\left(p_A^{(k)}\right).$$

Ravichandran applies Theorem 12.20 to the restricted invertibility problem in the following way. Given $v_1, \ldots, v_m \in \mathbb{R}^n$ with $\sum_{i=1}^m v_i v_i^T = I_n$, let $V \in \mathbb{R}^{n \times m}$ be the matrix with the i-th column be v_i. Consider the $m \times m$ matrix $B = I_m - V^T V$. For a subset $S \subseteq [m]$ with $|S| = k$, check that

$$\lambda_k\left(\sum_{i \in S} v_i v_i^T\right) = 1 - \lambda_{\max}\left(B_{[m] \setminus S}\right).$$

So the restricted invertibility problem is reduced to finding a subset X of size $m - k$ with small maximum eigenvalue $\lambda_{\max}(B_X)$. Using Theorem 12.20,

$$\max_{S:|S|=k} \lambda_k\left(\sum_{i \in S} v_i v_i^T\right) = 1 - \min_{X:|X|=m-k} \lambda_{\max}(B_X) \geq 1 - \lambda_{\max}\left(p_B^{(m-k)}\right) = 1 - \lambda_{\max}\left(\partial_x^{m-k}(x-1)^{m-n} x^n\right),$$

as the matrix B has eigenvalue 1 with multiplicity $m - n$ and eigenvalue 0 with multiplicity n, because $V^T V$ has the same spectrum as $V V^T = I_n$ by Fact 2.28. Therefore, once again, we have reduced the bound in the restricted invertibility problem to a pure mathematical problem about the maximum root of a well-studied polynomial.

Discussions

We end with two concluding remarks. One is that instead of looking up the known results for the roots of the specific polynomial in Exercise 12.16 and $\partial_x^{m-k}(x-1)^{m-n}x^n$ in Ravichandran's approach, we can use the results in Lemma 11.9 and Problem 11.13 from the barrier method to bound the roots of these polynomials. So, combining the method of interlacing family with the barrier method would give self-contained proofs of Theorem 12.14.

Another is that the proofs are constructive in that they give polynomial time algorithms to find such a subset. We leave this to the reader to check.

12.5 References

[BSS14] Joshua D. Batson, Daniel A. Spielman, and Nikhil Srivastava. Twice-ramanujan sparsifiers. *SIAM Rev.*, 56(2):315–334, 2014. 97, 98, 99, 100, 103, 108, 110, 117

[MSS14] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. Ramanujan graphs and the solution of the Kadison-Singer problem. In *Proceedings of the international Congress of Mathematicians*, volume 3, pages 363–386. 2014. 117

[MSS21] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families iii: Sharper restricted invertibility estimates. *Israel Journal of Mathematics*, pages 1–28, 2021. 104, 111, 119, 122, 123

[Rav18] Mohan Ravichandran. Principal submatrices, restricted invertibility, and a quantitative gauss-lucas theorem. *International Mathematics Research Notices*, 2020(15):4809–4832, 2018. 122, 124, 125

[Tho66] R. C. Thompson. Principal submatrices of normal and Hermitian matrices. *Illinois Journal of Mathematics*, 10(2):296 – 308, 1966. 123