Lecture 1: Real Rooted Polynomials

Lecturer: Shayan Oveis Gharan *Jan 6th*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

1.1 Introduction

Give a univariate (or a multivariate) polynomial p, there are three ways to study p:

By Coefficients: $p(t) = a_0 t^d + a_1 t^{d-1} + \cdots + a_d$,

By Zeros: $p(t) = a_0(t - \lambda_1) \ldots (t - \lambda_d)$,

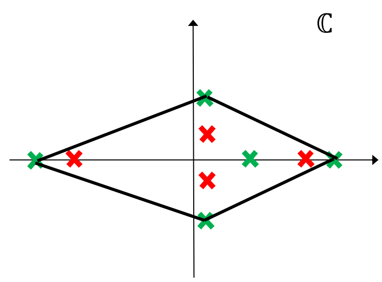
By function values: $p : \mathbb{C} \to \mathbb{C}$

In this note we typically use letter t for the univariate polynomials and z for multivariate polynomials.

In this course we will see interactions of these three views. How can one change when we change the others. More generally, we may also discuss how these views change when we do an operation on p such as differentiation or integration. See the following theorem for a concrete example:

Theorem 1.1 (Gauss-Lucas)**.** *If p is a (nonconstant) polynomial with complex coefficients, all zeros of p' are in the convex hull of the set of zeros of p.*

In the following picture the roots of $p(x) = x^5 - 3x^3 + x^2 - x + 1$ are shown in green and the roots of the derivative in red.



Proof. The idea is to look at the derivative of the log of p. Say, z is a root of p' which is not a root of p. Then,

$$(\log p(z))' = \frac{p'(z)}{p(z)} = 0,$$

The last identity Using the zeros lense of looking at p we can write

$$0 = \frac{p'(z)}{p(z)} = \sum_{i=1}^{d} \frac{1}{z - \lambda_i} = \sum_{i=1}^{d} \frac{\bar{z} - \bar{\lambda}_i}{|z - \lambda_i|^2}$$

where λ_i is the i-th root of p. Therefore,

$$\sum_{i=1}^{d} \frac{1}{|z-\lambda_i|^2}\bar{z} = \sum_{i=1}^{d} \frac{1}{|z-\lambda_i|^2}\bar{\lambda_i}$$

Conjugating both sides and rearranging we get

$$z = \sum_{i=1}^{d} \frac{\frac{1}{|z-\lambda_i|^2}}{\sum_{j=1}^{d} \frac{1}{|z-\lambda_j|^2}}\lambda_i;$$

so, z is a convex combination of the roots of p. □

The logarithmic derivative p'/p goes by many additional names, including the Cauchy Transform, Stieltjes Transform, and barrier function and plays a recurring theme in this course.

Say p is a multivariate polynomial with n variables. The main underlying ingredient is that all (or most) class of polynomials that we see in this course have a nice zero free region in the complex plane ($\mathbb{C}^n$) or the real plane ($\mathbb{R}^n$), in the sense that they do not have any zeros in a large part of the complex or real plane.

The polynomial paradigm says that if a polynomial has a nice zero free region, then we can study interaction of zeros, coeffficints and function values. This was studied in Math for decades.

Connections to CS and Combinatorics Recently, over the last 15 years, there was a new theme of using this machinery: Say we have a complicated mathematical object, say a probability distribution or a combinatorial object. We can encode this object in a complex multivariate polynomial (with exponentially many coefficients) say by writing down the generating polynomial. If the generating polynomial has a nice zero free region, then we can use this machinery to reason about the generating polynomial and prove statements and finally we can prove statements about the underlying object using what we found out for the generating polynomial.

So, here is the thesis that we propose and follow in this course: One way to understand global properties of discrete probability distributions is to encode them as polynomials which has nice zero-free regions and then use the polynomial paradigm to study that.

Given a graph $G = (V, E)$, here are a few fundamental classes of polynomials that we will work with in this course:

Spanning Tree Polynomial: We have a variable x_e for every edge $e \in E$, the spanning tree polynomial is defined as follows:

$$\sum_T \prod_{e \in T} x_e,$$

where the sum is over all trees in G.

Matching Polynomial: We have a variable x_v for every vertex $v \in V$,

$$\sum_M (-1)^{|M|} \prod_{v \text{ saturated in } M} x_v$$

where the sum is over all matchings of G, In the next lecture we will prove nice zero free region properties of the above polynomials. This will imply that the following univariate versions are real rooted

$$\sum_T t^{|T \cap F|} \tag{1.1}$$

$$\sum_M (-1)^{|M|} t^{|M|} \tag{1.2}$$

where in the first sum $F \subseteq E$ is an arbitrary subset of edges of G. In the rest of this lecture we study properties of real rooted polynomials and we use them to study properties of the above polynomials.

1.2 Real-rooted Polynomials

We start by recalling some properties of real-rooted polynomials.

In the following simple lemma we show that imaginary roots of univariate polynomials come in conjugate pairs.

Lemma 1.2. *For any $p \in \mathbb{R}[t]$, if $p(a + \imath b) = 0$, then $p(a - \imath b) = 0$.*

Proof. Say $p(t) = \sum_{i=0}^{d} a_i t^{d-i}$. Then,

$$\begin{aligned} p(a + \imath b) = 0 &= \overline{p(a + \imath b)} \\ &= \sum_{i=0}^{d} a_i \overline{(a + \imath b)^i} \\ &= \sum_{i=0}^{d} a_i \overline{(a + \imath b)}^i \\ &= \sum_{i=0}^{d} a_i (a - \imath b)^i = p(a - \imath b). \end{aligned}$$

□

1.3 Connection to Sum of Independent Bernoullis

There is a close connection between the real-rooted polynomials and the sum of Bernoulli random variables. Given a probability distribution μ, over $[d] = \{0, 1, \ldots, d\}$ where $\mathbb{P}_\mu [i] = a_i$. For $Z \sim \mu$,

$$p_\mu(t) = \mathbb{E}\left[t^Z\right] = \sum_{i=0}^{d} a_i t^i.$$

Lemma 1.3. *$p_\mu(t)$ is real-rooted if and only if μ can be written as the sum of independent Bernoulli random variables.*

Proof. Suppose $p_\mu(t)$ is real-rooted. Since the coefficients of $p_\mu(t)$ are non-negative, the roots are non-positive. Therefore, we can write

$$p_\mu(t) = a_0 \prod_{i=1}^{d} (t + \lambda_i),$$

where $\lambda_1, \dots, \lambda_d \geq 0$. Since $p(1) = 1$,

$$\frac{1}{a_0} = \prod_{i=1}^{d} (1 + \lambda_i),$$

Now, let $q_i = \frac{1}{1+\lambda_i}$; we show that μ is the distribution of d Bernoulli random variables where the i-th one has success probability q_i. Equivalently, let $\lambda_i = \frac{1-q_i}{q_i}$, we have

$$\begin{aligned}
p_\mu(t) &= \frac{1}{\prod_{i=1}^{d}(1+\lambda_i)} \prod_{i=1}^{d} \left(t + \frac{1-q_i}{q_i}\right) \\
&= \frac{1}{\prod_{i=1}^{d} \frac{1}{q_i}} \sum_{S \subseteq [d]} t^{|S|} \prod_{i \notin S} \frac{1-q_i}{q_i} \\
&= \sum_{S \subseteq [d]} t^{|S|} \prod_{i \in S} q_i \prod_{i \notin S} (1 - q_i) \\
&= \sum_{k=0}^{d} t^k \sum_{S:|S|=k} \prod_{i \in S} q_i \prod_{i \notin S} (1 - q_i). \qquad (1.3)
\end{aligned}$$

Observe that the probability that exactly k of the Bernoullis occur is exactly equal to a_k.

Conversely, if μ is the sum of d Bernoullis with success probabilities $q_1, \dots, q_d$, we can write

$$p_\mu(t) = \sum_{k=0}^{d} t^k \sum_{S:|S|=k} \prod_{i \in S} q_i \prod_{i \notin S} (1 - q_i)$$

Following (1.3), we conclude that p_μ is real rooted. □

Consequently, we can use Chernoff types of bound to show that a_i's which are far from the expectation $\sum_{k=0}^{d} k \cdot a_k$ are very small.

Corollary 1.4. *Let μ be a uniformly random spanning tree of $G = (V, E)$ and $F \subseteq E$ fixed. Then the random variable $Z = |T \cap F|$ is a summation of independent Bernoulli random variables.*

Note that the appearance of different edges are certainly not independent!

Proof. The idea is to use the real rooted polynomial of (1.1). In particular, the polynomial

$$\mathbb{E}\left[x^Z\right] = \frac{1}{\#T} \sum_{T} x^{|T \cap F|}$$

is a real rooted polynomial; so $\mathbb{E}\left[x^Z\right] = \mathbb{E}\left[t^{Y_1 + \cdots + Y_{|F|}}\right]$ where $Y_1, \dots, Y_{|F|}$ are independent Bernoullis. □

1.4 Log Concavity

Next we show that these sequence are highly concentrated by means of ultra log concavity.

Definition 1.5 (Log Concavity)**.** *A sequence $\{a_0, a_1, \ldots, a_d\}$ of nonnegative numbers is said to be log-concave if for all $0 < k < d$,*

$$a_{k-1} \cdot a_{k+1} \leq a_k^2.$$

We say the sequence is ultra log concave if for all $0 < k < d$,

$$\frac{a_{k-1}}{\binom{d}{k-1}} \cdot \frac{a_{k+1}}{\binom{d}{k+1}} \leq \left(\frac{a_k}{\binom{d}{k}}\right)^2.$$

Note that any log concave sequence of nonnegative numbers is unimodal, i.e., there is a number k such that

$$\cdots \leq a_{k-2} \leq a_{k-1} \leq a_k \geq a_{k+1} \geq \ldots.$$

Next, we want to show that the coefficients of a real-rooted degree d polynomial are ultra log concave. This is also known as the Newton inequalities.

Using the above argument this implies that a sum of independent Bernoulli random variables is a log-concave probability distribution, that is if a_i is the probability that exactly i of them occur, then $a_{i-1} \cdot a_{i+1} \leq a_i^2$ for all i. Consequently, any such distribution is unimodal.

1.4.1 Closure Properties of Real-rooted Polynomials

In this part we show that the coefficients of a real rooted polynomial with nonnegative coefficients are (ultra) log concave. The proof uses closure properties of real-rootedness. We start by describing these properties and then we prove the claim.

Given a real-rooted polynomial, usually it is a non-trivial task to verify real-rootedness without actually computing the roots. In the next section we will see how the extension of real-rootedness to multivariate polynomials can help us with this task.

One way to verify real-rootedness of a polynomial $p(t)$ is to start from a polynomial $q(t)$ that is real-rooted and then use a real-rooted preserving operator to derive $p(t)$ from $q(t)$. Now, let us study some basic operations that preserve real-rootedness.

i) If $p(t)$ is real-rooted then so is $p(c \cdot t)$ for any nonzero $c \in \mathbb{R}$.

ii) If $p(t)$ is a degree d real-rooted polynomial then so is $t^d p(1/t)$.

iii) If $p(t)$ is real-rooted then so is $p'(t)$. This follows from the by the Gauss-Lucas theorem.

Lemma 1.6 (Newton Inequalities)**.** *For any real-rooted polynomial $p(t) = \sum_{i=0}^{d} a_i t^i$, if $a_0, \ldots, a_d \geq 0$, then it is ultra log concave, i.e., for any $1 \leq i \leq d-1$,*

$$\left(\frac{a_i}{\binom{n}{i}}\right)^2 \geq \frac{a_{i-1}}{\binom{n}{i-1}} \cdot \frac{a_{i+1}}{\binom{n}{i+1}}$$

Proof. We apply the closure properties a number of times. First, by the closure of the derivative

$$p_1(t) = \frac{\mathrm{d}^{i-1}}{\mathrm{d}t^{i-1}} p(t)$$

is real-rooted. This shaves off all of the coefficients $a_0, \ldots, a_{i-2}$. By (ii),

$$p_2(t) = t^{d-i+1} p_1(1/t)$$

is real rooted. This reverse the coefficients. By (iii)

$$p_3(t) = \frac{\mathrm{d}^{d-i-1}}{\mathrm{d}t^{d-i-1}} p_2(t)$$

is real-rooted. This shaves off all of the coefficients $a_{i+2}, \ldots, a_d$. So, $p_3(t)$ is a degree 2 real-rooted polynomial,

$$p_3(t) = \frac{d!}{2} \left(\frac{a_{i-1}}{\binom{d}{i-1}} t^2 + \frac{2a_i}{\binom{d}{i}} t + \frac{a_{i+1}}{\binom{d}{i+1}} \right).$$

The above polynomial is real-rooted if and only if its discriminant is non-negative. This implies the lemma. □

As an immediate application, recall the claim the matching polynomial, see (1.2), is real rooted. This implies that for any graph $G = (V, E)$, the sequence $m_0, m_1, \ldots, m_n$, where m_i is the number of matchings of size i in G, is an ultra log-concave sequence.

Remark 1.7. *Newton was trying to characterize the set of real-rooted polynomials via inequalities satisfied by their coefficients. It turns out the ultra log-concavity does not characterize real-rootedness (i.e., it is necessary but not sufficient). However, it does characterize another natural class of polynomials, namely: a nonnegative sequence $a_0, \ldots, a_d$ is ULC iff there are convex compact sets A and B in $\mathbb{R}^d$ such that*

$$\sum_{i=0}^{d} a_i t^i = Vol(tA + B)$$

where Vol is the Lebesgue measure in $\mathbb{R}^d$ and $tA + B$ is the Minkowski sum. This was shown by Shephard in 1960 (see [She60]).

1.5 Approximation Roots

Given a polynomial

$$p(t) = t^d + a_1 t^{d-1} + \cdots + a_d = (t - \lambda_1) \ldots (t - \lambda_d).$$

Viéte in 1580s showed that

$$\begin{aligned} a_1 &= -\sum_{i=1}^{d} \lambda_i, \\ a_2 &= \sum_{i<j} \lambda_i \lambda_j, \\ &\vdots \\ a_d &= (-1)^d \prod_{i=1}^{d} \lambda_i \end{aligned}$$

These are called elementary symmetric polynomials in $\lambda_1, \ldots, \lambda_d$ (up to absolute values). In particular,

$$e_k(\lambda_1, \ldots, \lambda_d) = \sum_{S \in \binom{d}{k}} \prod_{i \in S} \lambda_i.$$

Let m_k be the moment functions of the roots, i.e.,

$$m_k(\lambda_1, \ldots, \lambda_k) = \sum_{i=1}^{d} \lambda_i^k.$$

Newton and Girard in 1600s observed that one can calculate the moment functions from elementary symmetric polynomials, and thus from the coefficients of p.

$$m_k(\lambda_1, \ldots, \lambda_d) = (-1)^{k-1} k e_k(\lambda_1, \ldots, \lambda_k) + \sum_{i=1}^{k-1} (-1)^{k-1-i} e_{k-i}(\lambda_1, \ldots, \lambda_d) m_i(\lambda_1, \ldots, \lambda_d).$$

For example, $m_1 = e_1$ and

$$m_2 = e_1 m_1 - 2e_2,$$

and so on. Also, observe that to compute the k-th moment function we just need to know $e_1, \ldots, e_k$ or $a_1, \ldots, a_k$. The following corollary is immediate:

Corollary 1.8. *Given the first k coefficient $a_0, \ldots, a_k$ of a real rooted polynomial $p = \sum_i a_i t^{d-i}$. For any $\epsilon > 0$, if $k \geq \log(d)/\epsilon$ then we can approximate the largest root of p (in absolute value) within $1 \pm \epsilon$ multiplicative factor.*

Proof. Wlog assume k is even. Given $a_0, \ldots, a_k$ we can compute $m_k = \sum_{i=1}^{d} \lambda_i^k$. Say λ_1 is the largest root of p in absolute value. It follows that

$$|\lambda_1|^k \leq m_k \leq d|\lambda_1|^k$$

Therefore,

$$|\lambda_1| \leq m_k \leq d^{1/k}|\lambda_1|$$

But since $k \geq \log(d)/\epsilon$, $d^{1/k} \leq e^{\epsilon}$ as desired. □

1.6 Testing Real Rootedness

In general, we can compute the roots of a polynomial up to a very good accuracy. But, how can we test if a given polynomial is real rooted?

Theorem 1.9 (Hermite-Sylvester)**.** *A polynomial $p(t) = \prod_{i=1}^{d}(t - \lambda_i)$ is real rooted iff the corresponding Hankel matrix $H \in \mathbb{R}^{d \times d}$ is PSD. H is a symmetric matrix where for any i, j,*

$$H_{i,j} = m_{i+j-2}.$$

Since given p we can compute the moments in polynomial time, this reduces testing real-rootedness of p to PSD-ness of H. To test the latter we compute the characteristic polynomial of H, $\det(tI - H)$. So, it is enough to test if $\det(tI - H)$ has a negative root. We use the following fact:

Fact 1.10. *A real-rooted polynomial p has all non-positive roots iff all coefficients of p are positive. Similarly, it has all non-negative roots iff its coefficients alternate in sign.*

This gives a strongly polynomial time algorithm to test if a given polynomial p is real rooted.

References

[She60] G. Shephard. "Inequalities between mixed volumes of convex sets". In: *Mathematika* 7.02 (1960), pp. 125–138 (cit. on p. 1-6).