Lecture 2: Zeros and coefficients of polynomials in one variable

Robin Pemantle

University of Pennsylvania

pemantle@math.upenn.edu

Minerva Lectures at Columbia University

02 November, 2016

Real-rooted polynomials

Elementary properties

Let **RR** denote the set of real polynomials all of whose roots are real. Let $\mathbf{RR}^+ \subseteq \mathbf{RR}$ denote the subset of polynomials all of whose roots are in $(-\infty, 0]$.

For probability generating functions the two notions coincide.

A number of properties follow. The most important are central limit behavior and ultra-log-concavity (Newton's inequalities) from which also follow log-concavity, unimodality and proximity of mean and mode.

Theorem (CLT)

Let $\{\mathbf{p_j} : 0 \le \mathbf{j} \le \mathbf{N}\}$ *be a probability sequence with generating polynomial* $\mathbf{f}(\mathbf{x}) = \sum_{\mathbf{j}=0}^{\mathbf{N}} \mathbf{p}(\mathbf{j})\mathbf{x}^{\mathbf{j}}$. *Let*

$$\begin{aligned} \mu &:= \mathbf{f}'(1) = \sum_{\mathbf{j}} \mathbf{j}\mathbf{p}(\mathbf{j}) \\ \sigma^2 &:= \mathbf{f}''(1) + \mu(1-\mu) = \sum_{\mathbf{j}} \mathbf{j}^2\mathbf{p}(\mathbf{j}) - \mu^2 \end{aligned}$$

be the mean and variance respectively. If $\mathbf{f} \in \mathbf{RR}$ *then*

$$\left| \Phi(\mathbf{t}) - \sum_{\mathbf{j} \le \mu + \mathbf{t}\sigma} \mathbf{p}(\mathbf{j}) \right| \le \frac{\mathbf{C}}{\sigma},$$

where Φ *is the normal CDF.*

Proof.

Factor $\mathbf{f}$ as a product of binomials. The constants can be distributed so that each binomial is the generating function $(1-\mathbf{p})+\mathbf{px}$ of a Bernoulli.

Thus the distribution is that of independent Bernoulli trials and the result follows from the Lindeberg-Feller CLT. $\square$

An easy improvement: let LHP be the set of all polynomials whose roots have nonpositive real part.

Polynomials in LHP can be factored into trinomials with nonnegative coefficients. Therefore the same argument yields the same conclusion, $\left|\Phi(\mathbf{t}) - \sum_{\mathbf{j}\le\mu+\mathbf{t}\sigma} \mathbf{p}(\mathbf{j})\right| \le \frac{\mathbf{C}}{\sigma}$ for all $\mathbf{f} \in$ LHP.

A surprising further improvement

Theorem (LPRS2015)

Suppose the zero set of $\mathbf{f}$ *avoids a ball of radius* δ *about 1. Then*

$$\left| \Phi(\mathbf{t}) - \sum_{\mathbf{j} \leq \mu + \mathbf{t}\sigma} \mathbf{p}(\mathbf{j}) \right| \leq \frac{\mathbf{C}_\delta \mathbf{N}}{\sigma^3} .$$

PROOF: Again factor into real trinomials, giving a convolution of signed measures with generating functions

$$\mathbf{f}_\mathbf{z}(\mathbf{x}) := \frac{(\mathbf{x} - \mathbf{z})(\mathbf{x} - \bar{\mathbf{z}})}{(1 - \mathbf{z})(1 - \bar{\mathbf{z}})} .$$

These signed measures satisfy $\log \mathbb{E}_\mathbf{z} \mathbf{e}^{\mathbf{it}} = -\mathbf{it}\mu_\mathbf{z} - \mathbf{t}^2\sigma_\mathbf{z}^2/2 + \mathbf{O}(\mathbf{t}^3)$, uniformly in $\mathbf{z}$, as long as $\mathbf{z}$ stays away from 1. Sum $\mathbf{N}$ of these, recenter to kill μ, and evaluate at $\mathbf{t}/\sigma$ to obtain the result. □

Early literature

From the modern point of view the CLT is obvious.

But in 1967 it was worth a publication in the Annals of Mathematical Statistics [Har67] to show asymptotic normality for Stirling numbers of the second kind by showing that their generating polynomial is real-rooted and then deriving the CLT.

A 1964 article [Dar64] made the reverse connection: from Bernoulli trials to real-rootedness to the other well known consequences of real-rootedness, which we now list.

Further properties

Proposition (properties of real-rooted polynomials)

1. *(Newton, 1707) A nonnegative coefficient sequence of a polynomial with real roots is log concave. In fact the sequence is* **ultra-logconcave**, *meaning that* $\{a_k/\binom{n}{k}\}$ *is log-concave.*
2. *(Edrai, 1953) A polynomial with nonnegative real coefficients has real roots if and only if its sequence of coefficients* $(a_0, \ldots, a_n)$ *is a* **Pólya frequency sequence**, *meaning that all the minors of the matrix* (a_{i-j}) *have nonnegative determinant.*
3. *Such a sequence is unimodal and its mean is within 1 of its mode.*

Newton's inequalities

What do Newton's inequalities say?

Because the sequence $\{\binom{\mathbf{n}}{\mathbf{k}} : 0 \leq \mathbf{k} \leq \mathbf{n}\}$ is it self log-concave, this says that the sequence $\{\mathbf{a_k}\}$ is "ultra-log-concave", a stronger property than log-concavity.

For the distribution $\mathrm{Bin}\,(\mathbf{n}, \mathbf{p})$ Newton's inequalities hold with equality because the sequence

$$\frac{\mathrm{Bin}\,(\mathbf{n}, \mathbf{p})(\mathbf{k})}{\binom{\mathbf{n}}{\mathbf{k}}} = (1 - \mathbf{p})^{\mathbf{n}} \left(\frac{\mathbf{p}}{1 - \mathbf{p}} \right)^{\mathbf{k}}$$

is log-linear.

Proof of Newton's inequalities

A polynomial of degree $\mathbf{n}$ can be reduced to a quadratic by differentiating $\mathbf{k}$ times with respect to $\mathbf{x}$, reversing the coefficient sequence, then differentiating $\mathbf{n}-\mathbf{k}-2$ more times.

$$a_n x^n \; + \cdots + \; a_{k+2}x^{k+2} + \; a_{k+1}x^{k+1} + \; a_k x^k \; + \cdots + \; a_0 x^0$$

$$\left(\frac{d}{dx}\right)^{k}$$

$$\cdots \; + \; \frac{(k+2)!}{2} a_{k+2}\, x^2 \; + \; (k+1)! a_{k+1}\, x \; + \; k!\, a_k$$

$$n!\left(\frac{d}{dx^{-1}}\right)^{n-k-2}$$

$$\frac{a_{k+2}}{2\binom{n}{k+2}}\, x^2 \; + \; \frac{a_{k+1}}{\binom{n}{k+1}} a_{k+1}\, x \; + \; \frac{a_k}{\binom{n}{k}} a_k$$

$\mathbf{B}^2 \geq 4\mathbf{AC}$:

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

Examples from combinatorics

Derangements

Example (generalized derangements)

Let $\mathbf{D_r(n)}$ denote the subset of permutations in $\mathcal{S}_\mathbf{n}$ all of whose cycles have length at least $\mathbf{r}$. For example, $\mathbf{D}_2(\mathbf{n})$ is the number of derangements. Let $\mathbf{m}(\sigma)$ denote the number of cycles of a permutation σ and let $\mathbf{f_{n,r}(x)} := \sum_{\sigma \in \mathbf{D_r(n)}} \mathbf{x}^{\mathbf{m}(\sigma)}$ be the generating function of cycles in $\mathbf{D_r(n)}$ counted by number of cycles.

Example: $\mathbf{f}_{6,2}(\mathbf{x}) = 15\mathbf{x}^3 + 130\mathbf{x}^2 + 120\mathbf{x}$

Type:	(123456)	(1234)(56)	(123)(456)	(12)(34)(56)
Count:	120	90	40	15

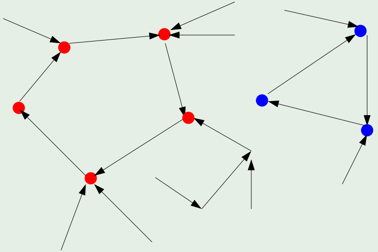
Using a bijection, an induction, and some facts from an early edition of EC I, Brenti [Bre95] showed that $\mathbf{f_{n,r}} \in \mathbf{RR}$.

Function digraphs

Example (functions from $[\mathbf{n}]$ to $[\mathbf{n}]$)

Given a function $\mathbf{f} : [\mathbf{n}] \to [\mathbf{n}]$, the cycle number $\kappa(\mathbf{f})$ denotes the number of cycles in the digraph defined by $\mathbf{f}$ (alternatively the number of irreducible closed classes in the degenerate Markov chain defined by $\mathbf{f}$).

Shown: a function with cycle number 2



Let $\mathbf{g_n}(\mathbf{x}) := \sum_{\mathbf{f}} \mathbf{x}^{\kappa(\mathbf{f})}$ be the generating polynomial counting functions by their cycle number. Brenti [Bre89] shows that $\mathbf{g}$ is real-rooted.

Family of determinants

Example (hyperbolicity of the determinant)

Let $\mathbf{A}$ be a nonnegative definite Hermitian matrix and let $\mathbf{B}$ be any Hermitian matrix. Set

$$\mathbf{f}(\mathbf{z}) := \det(\mathbf{z}\mathbf{A} + \mathbf{B}) .$$

In the lingo of [Går51], this is the hyperbolic property of the determinant. The multivariate version of this plays a big role in the theory of strong Rayleigh distributions.

PROOF: Multiplying on the right by $\mathbf{A}^{-1}$ it suffices to prove the result for $\mathbf{A} = \mathbf{I}$. The eigenvalues of $\mathbf{z}\mathbf{I} + \mathbf{B}$ are those of $\mathbf{B}$ shifted by $\mathbf{z}$, so the determinant of $\mathbf{B}$ has $\mathbf{n}$ real roots iff the determinant of $\mathbf{z}\mathbf{I} + \mathbf{B}$ does. $\square$

Rooted spanning forests

Example (rooted spanning forests)

Let $\mathbf{G}$ be a finite graph. If $\mathbf{F}$ is a forest on $\mathbf{G}$ (an acyclic subgraph) let $|\mathbf{F}|$ denote the number of components and let $\gamma(\mathbf{F})$ denote the product of the component sizes (the number of ways of choosing a root for each component).

Let $\mathbf{f_G}(\mathbf{x}) := \sum_{\mathbf{F}} \gamma(\mathbf{F})\mathbf{x}^{|\mathbf{F}|}$ count the rooted forests by number of components. Then $\mathbf{f_G} \in \mathbf{RR}$.

Proof: Let $\mathbf{A}$ be the matrix with $\mathbf{A_{ij}} = -1$ when there is an edge from $\mathbf{i}$ to $\mathbf{j}$, $\mathbf{A_i i} = \deg(\mathbf{i})$ and $\mathbf{A_{ij}} = 0$ otherwise. Then $\mathbf{f_G}(\mathbf{z}) = \det(\mathbf{A} + \mathbf{zI})$ [Sta97, attributed to Kelmans] and the result follows from nonnegative-definiteness of $\mathbf{A}$.

Matchings

A weighted matching of a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ is a subset of $\mathbf{E}$ consisting of disjoint edges, together with a weight function $\mathbf{w} : \mathbf{E} \to \mathbb{R}^+$.

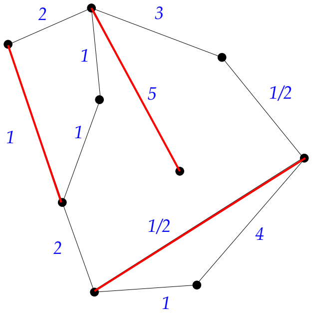


Figure: The red matching has weight $5/2$ and defect $n - 2|M| = 3$

Matchings

Example (matchings)

Let $\mathcal{M}$ denote the set of matchings of G, weighted multiplicatively by a weight function w.

The matching defect polynomial f is defined by

$$f(x) = \sum_{M \in \mathcal{M}} (-1)^{|M|} w(M) x^{|E|-2|M|} .$$

The Heilmann-Lieb Theorem states that the matching polynomial is real-rooted.

The proof uses an interlacing argument, to be examined shortly.

Closure properties

Establishing the *RR* property

How does one establish the *RR* property? Three general methods.

1. Interlacing
2. Elementary closure properties
3. Pólya-Schur theory

Multivariate analogues of these methods are crucial to the Borcea-Brändén-Liggett theory of strong Rayleigh distributions.

Interlacing

For a recursively defined sequence of polynomials $\{P_n\}$, one can sometimes inductively that $P_n \in RR$ and the $n-1$ roots of P_{n-1} interlace the n roots of P_n.

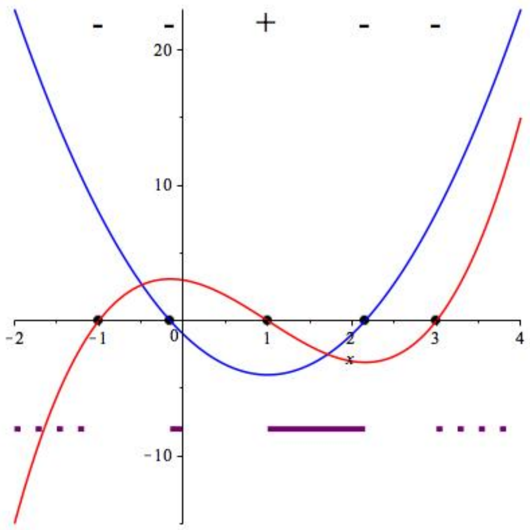
For example, if

$$P_{n+1} = \alpha_n \cdot x \cdot P_n - \beta_n \cdot P_{n-1}$$

for $\alpha_n, \beta_n > 0$, then by checking the signs of P_{n+1} at the zeros of P_n and P_{n-1}, one can identify $n-1$ zeros of P_{n+1} interlacing the zeros of P_n.

Examination of the leading term shows P_{n+1} has a zero greater (and also one less) than any zero of P_n.

Sign changes



Examples

$$\begin{array}{lrcl} \text{Chebyshev:} & C_{n+1} & = & 2xC_n - C_{n-1} \\ \text{Laguerre:} & L_{n+1} & = & (2n-1-x)L_n - nL_{n-1} \\ \text{Hermite:} & H_{n+1} & = & x\,H_n - H_n' \end{array}$$

Chebyshev polynomials are of the correct form, as are Laguerre polynomials (moving the origin to $2n-1$).

Hermite polynomials use H_n' in place of H_{n-1} so one obtains interlacing without assuming it for induction.

The matching defect polynomial requires only a little more care in the induction.

Elementary closure properties of RR

(i) Scaling: if $f \in RR$ then $f(bz) \in RR$.
geometric reweighting

(ii) Translation: if $f \in RR$ then $f(a+z) \in RR$.
binomial killing

(iii) Differentiation: if $f \in RR$ then $f' \in RR$.
size biasing

(iv) Product: if $f, g \in RR$ then $fg \in RR$.
convolution

(v) Inversion: if $f \in RR$ has degree n then $z^n f(1/z) \in RR$.
reversal

In each case, the class RR^+ is also preserved, providing a and b are positive in (i) and (ii) respectively.

Hadamard product

The previous properties follow more or less immediately from the definitions. A useful closure property that is more difficult but still long understood is the Hadamard or term-by-term product.

Proposition (Hadamard product)

If $f(x) := \sum_k a_k x^k$ and $g(x) := \sum_k b_k$ are in RR^+ then $f \odot g(x) := \sum a_k b_k x^k \in RR^+$ as well.

In fact the original proof of E. Maló (1895) shows the slightly stronger result that $f \in RR, g \in RR^+$ implies $f \odot g \in RR$.

We will prove this result using Pólya-Schur theory. Another proof follows from multivariate results in the next lecture.

Pólya-Schur theory

One might ask whether a converse holds: if multiplication by the coefficients of f preserves real-rootedness then must f be real rooted?

No!

In fact there is a century-old complete characterization of sequences (finite or infinite), coefficientwise multiplication by which preserves RR. We will state but not prove it.

Definition (multiplier sequence)

A sequence $\lambda := \{\lambda_0, \lambda_1, \lambda_2, \ldots\}$ is called a **multiplier sequence** *if for every polynomial $g = \sum_{k=0}^{n} a_k x^k \in RR^+$, the polynomial $\sum_{k=0}^{n} \lambda_k a_k x^k$ is also in RR^+.*

Theorem (Pólya-Schur 1914)

Denote by Φ the formal power series $\Phi(z) := \sum_{k=0}^{\infty} \frac{\lambda_k}{k!} z^k$. Then the following are equivalent.

(i) *λ is a multiplier sequence;*

(ii) *Φ is an entire function and is the limit, uniformly on compact sets, of the polynomials with all zeros real and of the same sign;*

(iii) *Φ is entire and either $\Phi(z)$ or $\Phi(-z)$ has a representation*

$$Cz^n e^{\alpha_0 z} \prod_{k=1}^{\infty} (1 + \alpha_k z)$$

where n is a nonnegative integer, C is real, and α_k are real, nonnegative and summable.

Trinomial

Example (trinomial)

Up to renormalization, a three term sequence is $(1, 1, \alpha)$. This is a multiplier sequence if and only if $\alpha \leq 1/2$.

Note that $1 + x + x^2/2 \notin RR$, showing that the converse of the Hadamard product result is false.

Whereas a trinomial $A + Bx + Cx^2$ in RR satisfies $B^2 \leq 4AC$, a sequence of arbitrary length in RR that begins $(A, B, C, \ldots)$ must satisfy only $B^2 \leq 2AC$.

Truncating to the first three terms necessitates a factor of $1/2$.

Inverse factorials

Example (finite inverse factorial sequence)

Generalizing this, we consider the finite sequence $\{1/k!\}_{k=0}^{n}$.

To see this is a multiplier sequence, observe it is the reverse of $\{1/(n-k)!\}_{k=0}^{n}$. This has exponential generating function $C(1+x)^n$, therefore satisfies (ii).

Taking a limit, this implies that $\{1/k!\}_{k=0}^{\infty}$ is a multiplier sequence. This fact, due to Laguerre, pre-dates the Pólya-Schur theorem.

If one allows also a factor of e^{bx^2}, one obtains the so-called Laguerre-Pólya class, which are exponential generating functions for multiplier sequences maping RR^+ to RR.

Back to Hadamard products

Proof that RR^+ is closed under Hadamard products.

The classical proof of closure under Hadamard products is via the Pólya-Schur Theorem.

Fix $f = \sum_{k=0}^{n} a_k x_k \in RR$.

Applying multiplier sequence $\{1/k!\}$ (either the finite or infinite sequence) shows that $\Phi(x) := \sum_{k=0}^{n} a_k x^k/k!$ has all real roots.

Hence $\{a_k\}$ is a multiplier sequence by criterion (ii) of the Pólya-Schur Theorem. □

Conditioned Bernoullis

Example

Let $\{X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots Y_m\}$ be independent Bernoulli variables with arbitrary means. Let $S := \sum_{i=1}^n X_i$ and $T := \sum_{i=1}^m Y_i$. Fix k and let $P(x) = \sum_{i=0}^n p_j x^j$ be the generating function for the conditional law of S given $S + T = k$, that is, $p_j := \mathbb{P}(S = j \,|\, S + T = k)$. Then $P \in RR^+$.

Let $f, g \in RR^+$ be the generating polynomials for S and T.

<u>Case 1: $k \geq m$.</u>

$$P = f \odot x^k g(1/x)\,.$$

<u>Case 2: $k < m$.</u>

$$x^{m-k} P = x^{m-k} f \odot x^{-m} g(1/x)\,.$$

$\square$

End of Lecture 2

References I

J. Borcea and P. Brändén.
Applications of stable polynomials to mixed determinants.
Duke J. Math., 143:205–223, 2008.

J. Borcea, P. Brändén, and T. Liggett.
Negative dependence and the geometry of polynomials.
J. AMS, 22:521–567, 2009.

P. Brändén, J. Haglund, M. Visontai, and D. Wagner.
Proof of the monotone column permanent conjecture.
In *FPSAC 2009*, volume AK, pages 443–454, Nancy, 2009. Assoc. Discrete Math. Theor. Comput. Sci.

F. Brenti.
Unimodal, log-concave and Pólya frequency sequences in combinatorics.
Memoirs of the AMS, 413:106+viii, 1989.

F. Brenti.
Log-concanve and unimodal sequences in algebra, combinatorics and geometry: an update.
In *Jerusalem Combinatorics '93*, volume 178, pages 71–89, Providence, 1994. AMS.

References II

F. Brenti.
Combinatorics and total positivity.
J. Comb. Theory, ser. A, 71:175–218, 1995.

T. Craven and G. Csordas.
Complex zero decreasing sequences.
Meth. Appl. Anal., 2:420–441, 1995.

J. N. Darroch.
On the distribution of the number of successes in independent trials.
Ann. Math. Stat., 35:1317–1321, 1964.

L. Gårding.
Linear hyperbolic partial differential equations with constant coefficients.
Acta Math., 85:1–62, 1951.

L. H. Harper.
Stirling behavior is asymptotically normal.
Ann. Math. Stat., 38:410–414, 1967.

J. Lebowitz, B. Pittel, D. Ruelle, and D. Speer.
Central limit theorem, Lee-Yang zeros, and graph counting polynomials.
J. Comb. Theory, Ser. A, 141:147–183, 2016.

References III

G. Pólya and J. Schur.
Über zwei Arten von Faktorenfolgen in der Theorie der algebraischen Gleichungen.
J. Reine Angew. Math., 144:89–113, 1914.

R. Stanley.
Log-concanve and unimodal sequences in algebra, combinatorics and geometry.
Annals of the New York Academy of Sciences, 576:500–535, 1989.

R. P. Stanley.
Enumerative Combinatorics. Vol. 1.
Cambridge University Press, Cambridge, 1997.
With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original.

M. Visontai.
Real-rooted polynomials in combinatorics.
PhD thesis, University of Pennsylvania, 2012.