The Geometry of Polynomials in Algorithms

Lecture 3: Real Stable Polynomials

Lecturer: Nathan Klein

1 The Multivariate Case

So far, we have only been able to study univariate distributions: single, integer-valued random variables. While fruitful, this can only take us so far. What about distributions over a set of elements $[n]$? In other words, a distribution $\mu : 2^n \to [0,1]$ or more generally $\mu : \mathbb{Z}^n_{\geq 0} \to [0,1]$? Given variables $z_1, \dots, z_n$ and $\alpha \in \mathbb{Z}^n_{\geq 0}$, write $z^\alpha = \prod_{i=1}^n z_i^{\alpha_i}$. Then, the generating polynomial of μ is:

$$p_\mu(z_1, \dots, z_n) = \sum_{\alpha \in \mathbb{Z}^n_{\geq 0}} \mu(\alpha) z^\alpha$$

For example, suppose we flip two independent fair coins. The generating polynomial of the resulting distribution over $\{0,1\}^2$ is then:

$$p(x,y) = \frac{1}{4}(1+x)(1+y).$$

If we set $x = y$, we get back the polynomial we looked at in the first lecture: the number of heads in two flips of a coin. So, this is just a slight step beyond a distribution with a real rooted generating polynomial. So you might hope that this is real rooted too. But it's not: for example, $p(-1, i) = 0$. Even worse, there are infinitely many roots: $p(-1, y) = 0$ for every $y \in \mathbb{C}$. In fact, *every* polynomial in n variables (WLOG assume all variables have positive degree) has a root x in which at least $n-1$ variables are nonreal. Why?

Let's first prove the following claim. Recall an open set $S \subseteq \mathbb{C}^n$ is one for which for every $x \in S$, there is a ball around x in $\mathbb{C}^n$ contained in S.

Claim 1.1. *Let $p \in \mathbb{C}[z_1, \dots, z_n]$ be a non-zero polynomial with finite degree in each variable. Let $S \subseteq \mathbb{C}^n$ be a nonempty open set. Then p cannot vanish everywhere in S.*

Proof. Suppose p vanishes everywhere in S. Let $a \in S$ and $b \in \mathbb{C}^n$. Consider the univariate polynomial:

$$g(t) = p(a + t(b-a))$$

Now, for all sufficiently small t, $a + t(b-a) \in S$. So, $g(t) = 0$ for infinitely many values t, so since it's bounded degree it must be identically 0. Therefore, $p(b) = g(1) = 0$, which means p is identically 0 too, contradiction. $\square$

Now we can prove our claim:

Claim 1.2. *Let $q \in \mathbb{C}[z_1, \dots, z_n]$ be nonconstant for $n \geq 2$. Then q has a root where at least $n-1$ coordinates are not real (and have positive imaginary part).*

Proof. After renaming variables so that q has positive degree in z_n, write

$$q(z_1, \dots, z_n) = \sum_{j=0}^{d} r_j(z_1, \dots, z_{n-1}) z_n^j,$$

where $d \geq 1$ and r_d is not the zero polynomial. Apply the above claim to r_d on the nonempty open set

$$\{(z_1, \ldots, z_{n-1}) \in \mathbb{C}^{n-1} : \operatorname{Im}(z_j) > 0 \text{ for every } j\}.$$

We can therefore choose $s_1, \ldots, s_{n-1}$ with positive imaginary parts such that $r_d(s_1, \ldots, s_{n-1}) \neq 0$. The resulting polynomial $q(s_1, \ldots, s_{n-1}, z_n)$ still has degree $d \geq 1$, so the fundamental theorem of algebra gives a root $z_n = \lambda$. So:

$$q(s_1, \ldots, s_{n-1}, \lambda) = 0,$$

and the first $n-1$ coordinates all have positive imaginary part, as desired. □

So we cannot hope to rule out all zeros with nonreal coordinates. But notice that for our two-coin polynomial, every zero has at least one real coordinate. You might hope the right generalization is exactly that, but unfortunately:

$$p_\mu = \frac{x}{2} + \frac{y}{2}$$

has the root $(-i, i)$. Such a simple distribution, one would hope, would also behave nicely. What's common between the two distributions? That **no zero has** $\operatorname{Im}(x_i) > 0$ **for all** i (or equivalently no zero's coordinates all have negative imaginary part, but these are the same up to conjugation). This will be, it turns out, the right generalization of real rootedness.

A Quick Recap of Complex Numbers

Write $z = a + ib$, where $a, b \in \mathbb{R}$ and $i^2 = -1$. Its real and imaginary parts are $\operatorname{Re}(z) = a$ and $\operatorname{Im}(z) = b$. Its conjugate is $\overline{z} = a - ib$, and:

$$|z|^2 = z\overline{z} = a^2 + b^2.$$

Conjugation respects addition and multiplication, so $\overline{x} + \overline{y} = \overline{x + y}$ and $\overline{x}\,\overline{y} = \overline{xy}$.

Euler's formula says $e^{i\theta} = \cos\theta + i\sin\theta$. Thus a non-zero complex number can be written as $re^{i\theta}$, with $r > 0$. Multiplication multiplies the lengths and adds the angles:

$$(re^{i\theta})(se^{i\phi}) = rse^{i(\theta+\phi)}.$$

For example, $e^{\pi i/4} = (1+i)/\sqrt{2}$ and $e^{3\pi i/4} = (-1+i)/\sqrt{2}$. Both have positive imaginary part, but their squares are i and $-i$.

1.1 Real Stability

Definition 1.3 (Stability)**.** *Let $\mathbb{H} = \{z \in \mathbb{C} : \operatorname{Im}(z) > 0\}$ be the open upper half-plane. A polynomial $p \in \mathbb{C}[z_1, \ldots, z_n]$ is stable if:*

$$p(z_1, \ldots, z_n) \neq 0 \qquad \text{whenever } z_1, \ldots, z_n \in \mathbb{H}.$$

If p is stable and all coefficients of p are real, then we say p is real stable.

Fact 1.4. *A non-zero univariate polynomial $p \in \mathbb{R}[t]$ is real rooted if and only if it is real stable.*

Proof. If all roots are real, there are no roots in $\mathbb{H}$. Conversely, suppose p is real stable. If it is not real rooted, it must have a root x with $\operatorname{Im}(x) < 0$. But then:

$$p(\overline{x}) = \sum_{i=0}^{d} a_i \overline{x}^i = \sum_{i=0}^{d} a_i \overline{x}^i = \sum_{i=0}^{d} \overline{a_i x^i} = \overline{p(x)} = 0$$

and $\overline{x}$ has positive imaginary part, contradiction. Notice we crucially used that $a_i = \overline{a_i}$ since the coefficients are real. $\square$

So in one variable, real stability corresponds to real rootedness, as we would want. Notice this is *not* true for complex polynomials. For example, $p(t) = i + t$ is stable but not really rooted, as its root is $-i$.

We give an equivalent, sometimes easier to use, definition of real stability here:

Lemma 1.5 (Alternate Characterization). *A polynomial $p \in \mathbb{R}[z_1, \ldots, z_n]$ is real stable if and only if, for every $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$, the univariate polynomial*

$$q(t) = p(b + ta) = p(b_1 + ta_1, \ldots, b_n + ta_n)$$

is non-zero and real rooted.

Proof. Suppose p is real stable. If $t \in \mathbb{H}$, then $\operatorname{Im}(b_j + ta_j) = a_j \operatorname{Im}(t) > 0$ for every j, so $q(t) \neq 0$. Since q has real coefficients, q is non-zero and real rooted.

For the other direction, suppose $p(z) = 0$ for some $z \in \mathbb{H}^n$. Then $z = b + ia$, where $b_j = \operatorname{Re}(z_j)$ and $a_j = \operatorname{Im}(z_j) > 0$. Then $q(t) = p(b + ta)$ has $q(i) = 0$, contradiction. $\square$

Geometrically, as t varies over $\mathbb{R}$, the point $b + ta$ traces a line through b with positive slope. The lemma says that along every such line, the polynomial must be non-zero and real rooted.

1.2 Examples and Non-Examples

(1) **Linear polynomials.** Suppose $p(z) = c + \sum_{j=1}^{n} a_j z_j$, with real coefficients. When it is not identically zero, this is real stable exactly when all its non-zero a_j have the same sign.

(2) **Independent coins.** For $p_j \in [0, 1]$, the polynomial

$$\prod_{j=1}^{n} ((1 - p_j) + p_j z_j)$$

is real stable.

(3) **Elementary symmetric polynomials.** For $0 \leq k \leq n$, define

$$e_k(z_1, \ldots, z_n) = \sum_{S \subseteq [n] : |S| = k} z^S.$$

where $z^S = \prod_{i \in S} z^i$. You will prove that this is real stable on the homework.

(4) $1 - xy$ is real stable. If $xy = 1$, then neither is 0, and $x = 1/y$. Then $\operatorname{Im}(x)$ has the opposite sign of $\operatorname{Im}(y)$ since $\frac{1}{c} = \frac{\overline{c}}{|c|^2}$ for $c \in \mathbb{C}$.

(5) **Non-examples:** $1 - xyz$ is not real stable – set $x = y = z = e^{2i\pi/3}$.

2 Closure Properties

As with real rooted polynomials, understanding the closure properties of real stable polynomials will help us build up a useful theory. Before we do that, we need the following generalization of the fact that the limit of real rooted polynomials is real rooted:

Theorem 2.1 (Hurwitz's Theorem). *Let $\Omega \subseteq \mathbb{C}^n$ be open and connected. Suppose $p_1, p_2, \ldots$ are bounded degree, have no zeros in Ω, and converge coefficient-wise to a polynomial p. Then either p has no zeros in Ω or p is identically zero.*

For us, the main case is $\Omega = \mathbb{H}^n$, allowing us to show that a limit of real stable polynomials of bounded degree is real stable or zero. Given this, we can derive the following closure properties:

(1) **Products.** If p, q are real stable, then so is pq.

Proof. $p(x)q(x) = 0$ implies $p(x)$ or $q(x)$ is 0, but since they are stable, whenever $x \in \mathbb{H}^n$, $p(x)q(x) \neq 0$. □

(2) **Identifying variables.** If $p(z_1, z_2, z_3, \ldots, z_n)$ is real stable, then so is $p(z_1, z_1, z_3, \ldots, z_n)$. This is also called diagonalization. (In particular, $p(t, \ldots, t)$ is non-zero and real rooted.)

Proof. Any root of $p(z_1, z_1, z_3, \ldots, z_n)$ is also a root of $p(z_1, z_2, z_3, \ldots, z_n)$. □

(3) **Real translations and external fields.** For $a \in \mathbb{R}^n_{>0}$ and $b \in \mathbb{R}^n$, the polynomial $p(a_1 z_1 + b_1, \ldots, a_n z_n + b_n)$ is real stable. The case $b = 0$ is called an external field.

Proof. Given a zero x of $p(a_1 z_1 + b_1, \ldots, a_n z_n + b_n)$ in $\mathbb{H}^n$, $ax + b$ is a zero of p and is still in $\mathbb{H}^n$ and $a > 0$ and $b \in \mathbb{R}$. □

(4) **Specialization.** For $c \in \mathbb{R}$, setting $z_1 = c$ gives a real stable polynomial or zero.

Proof. Consider for $m \to \infty$:

$$p_m(z_2, \ldots, z_n) = p(c + i/m, z_2, \ldots, z_n).$$

Each p_i is stable, since we plugged in a value with positive imaginary part for z_1. Their coefficients may be complex, but they converge coefficient-wise, with bounded degree, to $p(c, z_2, \ldots, z_n)$, so by Hurwitz's theorem we get the claim. □

(5) **Inversion.** Let d_1 be the degree of p in z_1. Then:

$$z_1^{d_1} p(-1/z_1, z_2, \ldots, z_n)$$

is real stable. We leave this as an exercise.

(6) **Differentiation.** The polynomial $\partial_{z_i} p = \frac{\partial p}{\partial z_i}$ is real stable (or zero) for any $i \in [n]$.

Proof. For simplicity let's differentiate with respect to z_n; it's the same for general z_i. If p does not depend on z_n, its derivative is zero. Otherwise, let $d \geq 1$ be its degree in z_n, and to use our favorite trick:

$$p(z_1, z_2, \ldots, z_n) = \sum_{j=0}^{d} r_j(z_1, \ldots, z_{n-1}) z_n^j.$$

We want to fix $z_1, \ldots, z_{n-1}$ and differentiate the resulting univariate polynomial. But there is a small issue: could fixing $z = z_1, \ldots, z_{n-1}$ make all the terms involving z_n disappear? To rule this out, we first show that $r_d(z) \neq 0$ whenever $z \in \mathbb{H}^{n-1}$. This is because the stable polynomials $(im)^{-d} p(z, im)$ converge coefficient-wise to $r_d(z)$ as $m \to \infty$, as you will only get the terms with degree d in z_n. Since r_d is not the zero polynomial, Hurwitz's theorem says it is stable. Thus fixing the other variables in $\mathbb{H}$ does not decrease the degree in z_n.

Now fix such a $z \in \mathbb{H}^{n-1}$, and let $f(t) = p(z, t)$. The coefficients of f may be complex, but its roots $\lambda_1, \ldots, \lambda_d$ all satisfy $\mathrm{Im}(\lambda_j) \leq 0$ since p is stable.[1] Write f as a product of its linear factors, apply the product rule,[2] and divide by $f(t)$. This gives:

$$\frac{f'(t)}{f(t)} = \sum_{j=1}^{d} \frac{1}{t - \lambda_j}.$$

Plug in $t \in \mathbb{H}$. Each $t - \lambda_j$ has positive imaginary part, so each reciprocal has strictly negative imaginary part. Therefore the sum is non-zero. Since $f(t) \neq 0$, $f'(t) \neq 0$. But differentiating f means differentiating p in its last coordinate, with the other coordinates fixed:

$$f'(t) = \partial_{z_n} p(z, t).$$

We chose $z \in \mathbb{H}^{n-1}$ and $t \in \mathbb{H}$ arbitrarily, so this shows that $\partial_{z_n} p$ is non-zero at every point of $\mathbb{H}^n$. □

Next time we will characterize all stability preservers and use them to give an alternate proof of Heilmann-Lieb.

[1] Note you cannot apply the fact that roots come in conjugate pairs here, since that's only true for univariate polynomials with real coefficients.

[2] Recall $(f_1 \ldots f_n)' = f_1' f_2 \ldots f_n + f_1 f_2' f_3 \ldots f_n + \cdots + f_1 f_2 \ldots f_{n-1} f_n'$.