Univariate Polynomials

Chris Williams

August 29, 2018

Contents

1 Review

Recall that much of our theory so far has dealt with what Stetter refers to as Intrinsic polynomials.

Definition 1 *Intrinsic univariate polynomials are those for which we assume the coefficients to be exact.*

We will be considering the univariate polynomials,

$$p(x; a) = \sum_{v=0}^{n} \alpha_v x^v = \mathcal{P}_n = \{a^T \mathbf{x} \ : \ a \in \mathbb{C}^{n+1}\},$$

where $\mathbf{x} = (1, x, x^2, \ldots, x^n)^T$.

We will also be looking at the so called empirical polynomials as well.

Definition 2 *The empirical vector* $(\overline{a}, e)$*, with specified value*

$$\overline{a} = (\overline{\alpha}_1, \ldots, \overline{\alpha}_M) \in \mathbb{C}^M \qquad \textit{and tolerance} \qquad e := (\epsilon_1, \ldots, \epsilon_M), \quad \epsilon_j > 0$$

defines a family N_δ*,* $\delta > 0$ *of neighbourhoods*

$$N_\delta(\overline{a}, e) := \{\tilde{a} : \ \|\tilde{a} - \overline{a}\|_e^* \leq \delta\},$$

where $\|\cdot\|_e^*$ *is the weighted dual maximum norm. Specifically*

$$\|a\|_e^* = \max_{i \in \{1,\ldots,n\}} \left\{ \frac{|\alpha_i|}{\epsilon_i} \right\}.$$

2 Univariate Polynomials Root Splitting

We will be considering the univariate polynomials,

$$p(x;a) = \sum_{v=0}^{n} \alpha_v x^v = \mathcal{P}_n = \{a^T \mathbf{x} \ : \ a \in \mathbb{C}^{n+1}\},$$

where $\mathbf{x} = (1, x, x^2, \ldots, x^n)^T$.

We also have that $p \in \mathcal{P}_n$ is called monic if $a_{n+1} = 1$. Now for each $p \in \mathcal{P}^n$ there exist n possibly non distinct numbers $\zeta_v \in \mathbb{C}$ such that

$$p(x) = \alpha_n \prod_{v=1}^{n} (x - \zeta_v).$$

Given this parameterisation of our polynomials it is worthwhile to note that the gcd of two randomly selected polynomials is almost surely one.

Note 1 *Each $p \in \mathcal{P}_n$ has exactly n zeros $\zeta_v \in \mathbb{C}$. The multiplicity of a root ζ is exactly m if and only if*

$$p(\zeta) = p'(\zeta) = \cdots = p^{m-1}(\zeta) = 0, p^m \neq 0.$$

Proof 1 *Note that the derivative of p can be calculated through repeated use of the product rule, thus*

$$\begin{aligned}\frac{d}{dx}p(x) &= \alpha_n \left[\left(\frac{d}{dx}(x - \zeta_1) \right) \prod_{j \neq 1} (x - \zeta_j) + (x - \zeta_1) \cdot \frac{d}{dx} \prod_{j \neq 1} (x - \zeta_j) \right] \\ &= \alpha_n \sum_{i=1}^{n} \prod_{j \neq i} (x - \zeta_j).\end{aligned}$$

The Note follows.

In Chapter 3 we were given the idea of viewing a polynomial $p(x;a) \in \mathcal{P}_n$ as a function of its coefficients $a \in \mathbb{C}^{n+1}$, that is p is a linear function of a.

Definition 3 *For an empirical algebraic problem, with empirical data $(\overline{a}, e)$, the mapping*

$$F : \ N_\delta(\overline{a}, e) \to \mathcal{Z},$$

which assigns to each data value $\tilde{a}$ the exact result $z \in \mathcal{Z}$ of the algebraic problem with this data is called the data $\rightarrow$result mapping.

With this in place we can make sense of the following Proposition.

Proposition 1 *For an m-fold zero ζ of $p(x;a) \in \mathcal{P}_n$, the mapping $F : a \to \zeta(a)$ is Hölder continuous with exponent $\frac{1}{m}$ at the coefficient vector a which generates the multiple zero.*

Stetter notes that this result gives that under a sufficiently small generic perturbation of p, an m-fold zero ζ splits into m simple zeros that differ from ζ by $O(\|\Delta a\|^{\frac{1}{m}})$. To see this explicitly we will look at the example given in Chapter 3.

Example 1 *Consider the intrinsic polynomial with associated empirical polynomial below,*

$$p(x) = (x-\sqrt{2})^3(x+\sqrt{2}) \qquad \overline{p}(x) = x^4 - 2.83088x^3 + 0.00347x^2 + 5.66176x - 4.00694,\ \epsilon_j = 10^{-5}.$$

For the single stable root the magical computer gives us results in $[-1.4142168, -1.4142104]$ *for* $a \in N_1$. *Note anything between* $-1.414217, \ldots, -1.414210$ *is a valid approximate zero of* $(\overline{p}, e)$. *Using a root finder to find the zeros of the 'center' of our empirical polynomial we find it gives* $\{1.41421, 1.4148, 1.41607\}$. *This shows us the splitting of a multiplicity three root to three single multiplicity roots.*

However the equally valid polynomial

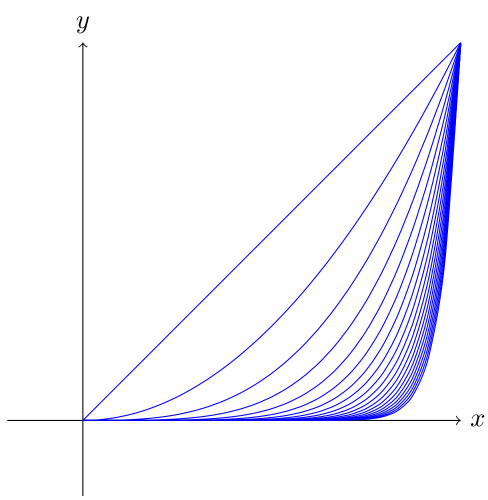
$$\tilde{p}(x) := x^4 - 2.83087x^3 + 0.00348x^2 + 5.66177x - 4.00693 \in N_1(\overline{p}, e)$$

has the zeros (rounded to 5 digits)

$$1.38583 \textit{ and } 1.42963 \pm 0.02578i.$$

3 Orthogonal Polynomials

The set $\mathcal{P} = \mathbb{C}[x]$ of all univariate complex coefficient polynomials is a (countably) infinite dimensional vector space, and the truncations of this space $\mathcal{P}_n$ are of dimension $n+1$. This is somewhat immediate once we recall the natural basis of $\mathcal{P}$ as $\mathbf{p}_0 := \{x^v \mid v \in \mathbb{N}_0\}$. Whilst this basis is very notationally easy to work with it is not the best computationally. Immediately is the fact that all the polynomials in this basis set have a repeated root at zero and roots no where else and are thus subject to the aforementioned splitting we had before.



To find a better way we can equip this space with the inner product

$$\langle p, q\rangle = \int_a^b p(\xi)q(\xi)d\mu(\xi),$$

which now gives us a sense of angles on this vector space. Now a desirable basis of a vector space is one which is in the appropriate sense orthonormal.

Definition 4 *For a specified inner product for* $\mathcal{P}$, *a sequence of polynomials* $p_v \in \mathcal{P}_v$ *which satisfies* $\mathcal{P}_n = span(p_0, \ldots, p_n)$ *and*

$$\langle p_v, p_u\rangle = 0 \qquad v \neq u,$$

is a system of orthogonal polynomials. If $\langle p_v, p_v\rangle = 1$ *then this is a orthonormal system.*

Proposition 2 *The representation of some $p \in \mathcal{P}_n$ can be represented in the basis $\mathbf{p}$ is given by,*

$$p(x) = \sum_{v=0}^{n} \frac{\langle p, p_v \rangle}{\langle p_v, p_v \rangle} p_v(x).$$

Furthermore if $p \in \mathcal{P}_N$ for $N > n$, then the projection $p^* \in \mathcal{P}_n$ can be calculated with q satisfying $\langle p_n^*, q - p_n^* \rangle = 0$ as

$$P_n^*(x) = \sum_{v=0}^{n} \frac{\langle q, p_v \rangle}{\langle p_v, p_v \rangle} p_v(x).$$

Now there are some additional nice properties if the measure μ in the inner product in non-singular, that is it has a Lebesgue density. Under this assumption we have:

(1) The polynomials p_v for our orthogonal basis obey a three term recurrence relation. That is a reccurence of the form

$$p_{v+1} = (x - \beta_v) p_v(x) - \gamma_v p_{v-1}(x) \qquad v \geq 1.$$

(2) p_n has n real disjoint zeros in (a, b)

(3) Any polynomial orthogonal to all p_v for $v \in \{0, \ldots, n\}$ has at least $n+1$ zeros in (a, b) with a sign change.

4 Chebyshev Polynomials

The Chebyshev polynomials are those which are constructed through the inner product

$$\langle p, q \rangle := \int_{-1}^{1} \frac{p(\xi) q(\xi)}{\sqrt{1 - \xi^2}} d\xi.$$

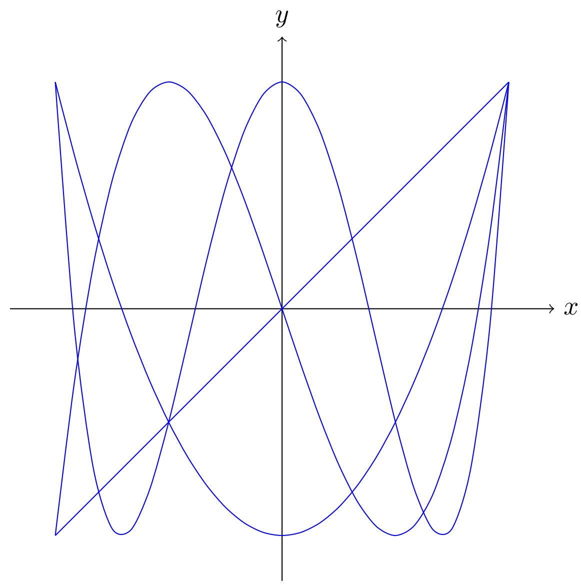
The polynomial system which is orthogonal with respect to this inner product, $\{T_v\}_{v \in \mathbb{N}}$, are the Chebyshev polynomials. They obey the following trigonometric identity

$$T_v(\cos\phi) = \cos(v\phi),$$

which can be used to show the orthogonality condition through a calculation:

$$\begin{aligned} \int_{-1}^{1} \frac{T_v(x) T_u(x)}{\sqrt{1 - x^2}} dx &= \int_0^{\pi} \sin(y) \frac{T_v(\cos(y)) T_u(\cos(y))}{\sqrt{1 - \cos(y)^2}} dy \\ &= \int_0^{\pi} \cos(vy) \cos(uy) dy = 0. \qquad (v \neq u) \end{aligned}$$

Here is a plot of the first few



Some nice properties of these polynomials are:

(1) The obey the recurrence relation

$$T_{v+1}(x) = 2xT_v(x) - T_{v-1}(x). \qquad (v \geq 1)$$

(2) The zeros in $(-1, 1)$ are:

$$\xi_v = \cos\left(\frac{2v_1}{2n}\pi\right). \qquad (v \geq 1)$$

In particular all the zeros are simple zeros.

(3) The extreme points are:

$$\eta_v = \cos\left(\frac{v}{n}\pi\right). \qquad (v \geq 0)$$

Furthermore the extrema values are of magnitude one.

We also can manipulate this recurrence relationship to evaluate these polynomials with the following algorithm. If we take β to be the coefficient vector for a polynomial $p \in \mathcal{P}_n$ expressed with the Chebyshev basis, i.e. $p(x) = \sum_{v=0}^{n} \beta_v T_v(x)$ then the recursion

$$\begin{aligned} \beta_{n-1} &\leftarrow \beta_{n-1} + 2\xi\beta_n \\ \beta_v &\leftarrow \beta_v + 2\xi\beta_{v+1} - \beta_{v+2} \qquad (v \in (n-2):1) \\ \beta_0 &\leftarrow \beta_0 + \xi\beta_1 - \beta_2 \end{aligned}$$

yields $p(\xi) = \beta_0$. Now there is a useful bound

$$\|p\|_{\max} = \max_{x\in[-1,1]} |p(x)| = \max_{x\in[-1,1]} \left|\sum_{v=0}^{n} \beta_v T_v\right| \leq \max_{x\in[-1,1]} \left|\sum_{v=0}^{n} \beta_v\right|.$$

The coefficients of p represented in the Chebyshev basis are usually quite small so this bound is far more useful than using the basis of powers of x.

5 Backwards Error

We begin by giving the definition of an equivalent data set for some empirical value $\tilde{z}$.

Definition 5 *For an empirical algebraic problem with the data $\to$ result function $F : \mathcal{A} \to \mathcal{Z}$ and for a given result $\tilde{z} \in \mathcal{Z}$, the equivalent data set for z is defined as*

$$\begin{aligned}\mathcal{M}(\tilde{z}) :&= \{\tilde{a} \in \mathcal{A} : F(\tilde{a}) = \tilde{z}\} \\ &= \left\{ a \in \mathcal{A} : \sum_{j \in \tilde{J}} (\tilde{\alpha}_j - \overline{\alpha}_j)\tilde{z}^j + \overline{p}(\tilde{z}) = 0 \right\}.\end{aligned}$$

With this in place we may think of the above set as the possible solutions for the result of some empirical data. In particular we are interested in the verification task of seeing if any of the possible solutions lie in our neighbourhood $N_\delta(\overline{a}, e)$. More generally we can find the minimum δ such that the above set and our neighbourhood overlap.

Definition 6 *The above situation can be described as:*

$$\min_{a \in \mathcal{M}(z)} \|a - \overline{a}\|_e^*, \tag{1}$$

which is also referred to as the backward error of the approximate rewult $\tilde{z}$ for the empirical data problem with data $(\overline{a}, e)$.

Proposition 3 *The above quantity can be explicitly calculated for the zeros of polynomials through:*

$$\delta(\tilde{z}) = \frac{|\overline{p}(\tilde{z})|}{\|(\tilde{z}^j)\|_e} = \frac{|\overline{p}(\tilde{z})|}{\sum_{j \in \tilde{J}} \epsilon_j |\tilde{z}|^j}.$$

Back to our previous example.

Example 2 *Recall the polynomial $\overline{p}(x) = x^4 - 2.83088x^3 + 0.00347x^2 + 5.66176x - 4.00694$ with $\epsilon_i = 10^{-5}$. Now we wish to determine whether or not the value $\tilde{z} = 1.43244$ is a valid root of this problem. Using the aove formula*

$$\delta(\tilde{z}) = \frac{1.5 \cdot 10^{-5}}{(1 + \tilde{z} + \tilde{z}^2 + \tilde{z}^3) \cdot 10^{-5}} \approx 0.2,$$

so this seems like a reasonable solution to our empirical polynomial. If we now used the same formula to test if this was a multiple root ne finds:

$$\delta(\tilde{z}) = \frac{|\overline{p}'(\tilde{z})|}{(1 + 2\tilde{z} + 3\tilde{z}^2) \cdot 10^{-5}} \approx 26,$$

which indicates this might just be a simple root result to our problem.