ON THE CHEBYSHEV POLYNOMIALS

JOSEPH DICAPUA

Abstract. This paper is a short exposition of several magnificent properties of the Chebyshev polynomials. The author illustrates how the Chebyshev polynomials arise as solutions to two optimization problems. The presentation closely follows *The Chebyshev Polynomials* by Theodore J. Rivlin. The results presented in this paper can be found in Rivlin's book.

Contents

1. Definitions and Properties

One can define the Chebyshev polynomials using de Moivre's formula. For a nonnegative integer n, the Chebyshev polynomial T_n of degree n is defined as follows. Given any $x \in [-1, 1]$ there exists a unique angle $0 \leq \theta \leq \pi$ such that $x = \cos\theta$. Observe that x decreases from 1 to -1 as θ increases from 0 to π. Then T_n is defined pointwise on $[-1, 1]$ by

$$T_n(x) = \cos n\theta.$$

At this point it might not be clear why T_n is a polynomial, but it is not difficult to show that T_n extends uniquely to a real polynomial on all of $\mathbb{R}$. Recall that de Moivre's formula states

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

De Moivre's implies that

$$e^{in\theta} = \cos n\theta + i\sin n\theta = (\cos\theta + i\sin\theta)^n = \sum_{k=0}^{n} \binom{n}{k} i^{n-k} \cos^k\theta \sin^{n-k}\theta,$$

and, from the above, one can derive the trigonometric identity

$$\cos n\theta = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} (-1)^k \cos^{n-2k}\theta \sin^{2k}\theta.$$

Conveniently only even powers of $\sin\theta$ appear in the above expression, so one can replace them with even powers of $\cos\theta$ to obtain

$$\cos n\theta = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \binom{n}{2k} \cos^{n-2k}\theta \left(\sum_{j=0}^{k} (-1)^j \binom{k}{j} \cos^{2j}\theta \right).$$

This is enough to show that T_n is a polynomial in x, but one can simplify the above expression, through careful reindexing, to obtain the equality

$$\cos n\theta = \sum_{j=0}^{\lfloor n/2 \rfloor} (-1)^j \left(\sum_{h=j}^{\lfloor n/2 \rfloor} \binom{n}{2h} \binom{h}{j} \right) \cos^{n-2j}\theta.$$

One can then replace $\cos\theta$ with x, and it becomes clear that T_n has degree n.

$$T_n(x) = \sum_{j=0}^{\lfloor n/2 \rfloor} (-1)^j \left(\sum_{h=j}^{\lfloor n/2 \rfloor} \binom{n}{2h} \binom{h}{j} \right) x^{n-2j}.$$

It will often be easier to work with the original definition for the Chebyshev polynomials. For instance, the trigonometric identity

$$\cos n\theta \cos m\theta = \frac{\cos(n+m)\theta + \cos(n-m)\theta}{2}$$

implies the polynomial identity

$$T_n(x)\, T_m(x) = \frac{T_{n+m}(x) + T_{n-m}(x)}{2} \qquad m \leq n.$$

One can even compose two Chebyshev polynomials to obtain the identity

$$T_{nm}(x) = \cos nm\theta = T_n(\cos m\theta) = T_n(T_m(\cos\theta)) = T_n(T_m(x)).$$

It is not difficult to determine the zeroes and extrema of $T_n(x)$ using the definition $T_n(x) = \cos n\theta$. Consider the n angles

$$\theta_j = \frac{2j-1}{n}\frac{\pi}{2} \qquad 1 \leq j \leq n.$$

These θ_j are distinct, and they all lie between 0 and π. Then define

$$\xi_j = \cos\theta_j = \cos\frac{2j-1}{n}\frac{\pi}{2}.$$

The θ_j lie between 0 and π, so the ξ_j must lie between 1 and -1. The θ_j are all distinct, so the ξ_j are also distinct. T_n is of degree n, and

$$T_n(\xi_j) = 0$$

for each j, so the zeroes of T_n are exactly the n distinct ξ_j.

A similar strategy is used to determine the relative extrema of T_n. Define

$$\eta_k = \cos\frac{k\pi}{n} \qquad 0 \leq k \leq n.$$

The η_k are distinct and lie between 1 and -1 because the values $\frac{k\pi}{n}$ are all distinct and lie between 0 and π. From earlier work,

$$T_n(\eta_k) = (-1)^k.$$

This implies that the numbers $\eta_1, \eta_2, \ldots, \eta_{n-1}$ are exactly the zeroes of T_n'. This is so because $|T_n(x)| \leq 1$ for $-1 \leq x \leq 1$, implying that $\eta_1, \eta_2, \ldots, \eta_{n-1}$ must be relative maxima for T_n on the interval $(-1, 1)$.

The derivatives of Chebyshev polynomials are the Chebyshev polynomials of the second kind, and they satisfy some nice identities as well. From our previous discussion,

$$\cos n\theta = T_n(\cos\theta),$$

so finding the derivative of $T_n(x)$ with respect to x is equivalent to finding the derivative of $\cos n\theta$ with respect to $\cos\theta$. Applying the chain rule gives

$$\frac{d}{d\cos\theta}\cos n\theta \frac{d}{d\theta}\cos\theta = \frac{d}{d\theta}\cos n\theta = -n\sin n\theta,$$

and this implies

$$\frac{d}{d\cos\theta}\cos n\theta = \frac{n\sin n\theta}{\sin\theta}.$$

The polynomials of the form

$$U_{n-1}(x) = \frac{1}{n}T_n'(x) = \frac{\sin n\theta}{\sin\theta} \qquad n \geq 1$$

are the Chebyshev polynomials of the second kind. The rightmost equality holds only for $0 \leq \theta \leq \pi$ and $x = \cos\theta$. These polynomials satisfy some exciting identities involving the T_n.

The trigonometric identity

$$\sin(n+1)\theta - \sin(n-1)\theta = 2\sin\theta\cos n\theta$$

implies the polynomial identity

$$U_n - U_{n-2} = 2T_n.$$

Similarly the trigonometric identity

$$\sin(n+1)\theta - \cos\theta\sin n\theta = \sin\theta\cos n\theta$$

implies the polynomial identity

$$U_n - xU_{n-1} = T_n.$$

One can even obtain a recursive formula for the Chebyshev polynomials using trigonometric identities. As before, the trigonometric identity

$$\cos n\theta + \cos(n-2)\theta = 2\cos\theta\cos(n-1)\theta$$

implies the polynomial identity

$$T_n = 2xT_{n-1} - T_{n-2} \qquad n \geq 2.$$

This recursive formula can be used to deduce the following polynomial generating function for $T_n(x)$:

$$F(y,x) = \frac{1 - xy}{1 - (2xy - y^2)} = \sum_{n=0}^{\infty} T_n(x)\, y^n.$$

The Chebyshev polynomials are also defined by their extremal properties. Recall that for a real valued function $f : [-1, 1] \to \mathbb{R}$ the supremum norm of f is defined to be

$$||f|| = \max |f(x)|$$

where the max is taken over all $x \in [-1, 1]$. Define $\mathscr{P}_n$ to be the real vector space of real polynomials of degree at most n equipped with the supremum norm.

Theorem 1.1. *Let u be a monic polynomial of degree n such that $|u|$ achieves its maximum value on $[-1, 1]$ at $n + 1$ or more points. Suppose that $p \in \mathscr{P}_n$ and that p is monic. If $p \neq u$, then $||p|| > ||u||$.*

Proof. Let $\gamma_0 \leq \gamma_1 \leq \ldots \leq \gamma_{n-1} \leq \gamma_n$ be the $n + 1$ maximal points of u, so $|u(\gamma_j)| = ||u||$ for $0 \leq j \leq n$. Necessarily $\gamma_0 = -1$ and $\gamma_n = 1$, and one can show that the signs of the values $u(\gamma_j)$ must alternate. If $u(\gamma_j) = u(\gamma_{j+1})$ for some j, $0 \leq j \leq n - 1$, then u must have a critical point in (γ_j, γ_{j+1}). This is impossible because u' has degree $n - 1$, so sgn $u(\gamma_j)$ must equal $(-1)^j$sgn $u(\gamma_0)$ for $0 \leq j \leq n$.

Now assume there exists a monic $p \in \mathscr{P}_n$, $p \neq u$, with $||p|| \leq ||u||$. Define $q = u - p$, so $q \in \mathscr{P}_{n-1}$ because u and p are both monic. The following lemma is the meat and potatoes of the proof.

Lemma 1.2. *Suppose that $q(\gamma_i)$ and $q(\gamma_{i+h})$ are nonzero and that $q(\gamma_{i+j}) = 0$ for $0 < j < h$. Then q has at least h zeroes in $[\gamma_i, \gamma_{i+h}]$ counted with multiplicity.*

Proof. The hypothesis gives us $h-1$ zeroes for q directly. The parity of the number of zeroes counted with multiplicity is the same as the parity of h. This is because sgn $q(\gamma_{i+h})$ = sgn $u(\gamma_{i+h}) = (-1)^h$ sgn $u(\gamma_i) = (-1)^h$ sgn $q(\gamma_i)$. Then the number of zeroes is at least h. □

Let $S = \{x \in [-1, 1] \mid x = \gamma_j$ for some $0 \leq j \leq n$ and $q(x) \neq 0\}$. Note that S has at least two elements because $q \in \mathscr{P}_{n-1}$. Pick integers m and M so that $\gamma_m = \min S$ and $\gamma_M = \max S$. Then q has at least $m + (M - m) + (n - M) = n$ zeroes in $[-1, 1]$. This is a contradiction because $q \in \mathscr{P}_{n-1}$, so the assumption that such a p exists must be false. □

Corollary 1.3. *Let $p \in \mathscr{P}_n$ such that $||p|| = 1$ and p has at least $n+1$ extrema on $[-1, 1]$. Then $p = \pm 1$ or $p = \pm T_n$.*

2. A Result on Linear Functionals on $\mathscr{P}_n$

Throughout this section X is a compact subset of $\mathbb{R}^m$ and V is a k-dimensional subspace of $C(X)$, the space of real valued continuous functions on X. If $v \in V$ the extremal points of v are the set of points $x \in X$ such that $|v(x)| = ||v||$. Recall that for a bounded linear functional F on a normed linear space V, the supremum norm is defined to be

$$||F|| = \sup_{v \in V} \frac{|Fv|}{||v||}$$

where $|Fv|$ denotes the usual absolute value of the complex number Fv. All linear functionals on V will necessarily be bounded linear functionals because V is finite dimensional. A nonzero $v \in V$ is extremal for F if $Fv = ||F||$ and $||v|| = 1$.

Let C_n be the convex subset of $\mathscr{P}_n$ defined by

$$C_n = \{p \in \mathscr{P}_n \mid \max_{0 \le j \le n} |p(\eta_j)| \le 1\}$$

where $\eta_0, \dots, \eta_n$ are as usual the extremal points of T_n. Define $\tilde{T}_n$ to be the unique monic scalar multiple of T_n. The goal of this section is to prove the following result on linear functionals.

Theorem 2.1. *Let F be a linear functional on $\mathscr{P}_n$. Suppose that F is such that v has n distinct roots implies $Fv \neq 0$. Suppose further that neither 1 nor -1 is extremal for F. Then for $p \in C_n$, $||Fp|| \le ||F\tilde{T}_n||$ with equality holding iff $p = \pm\tilde{T}_n(x)$.*

Further knowledge of linear functionals and approximations will be used in the proof.

Definition 2.2. A canonical representation of a real linear functional F on V is defined as follows. If there exists a set of r points $y_1, \dots, y_r$ where $y_i \in X$ and $r \le \dim V$ and there exist corresponding real numbers $\alpha_1, \dots, \alpha_r$ such that

$$||F|| = \sum_{i=1}^{r} |\alpha_i| \text{ and } Fv = \sum_{i=1}^{r} \alpha_i v(y_i) \text{ for all } v \in V,$$

then F is said to have a canonical representation.

The following result on best approximations is used in order to prove that every real linear functional on V has a canonical representation.

Theorem 2.3. *Let w be an element of $C(X)$, and let $w_1, \dots, w_k$ be a basis for V. Define $\overline{w_i}(x) = w(x)w_i(x)$, so $\overline{w_i} \in C(X)$. Let S be the subset of $\mathbb{R}^k$ defined by*

$$S = \{(\overline{w_1}(y), \dots, \overline{w_k}(y)) \text{ such that } y \in X \text{ and } |w(y)| = ||w||\}.$$

Then $||w+v|| \ge ||w||$ for every $v \in V$ iff the origin of $\mathbb{R}^k$ is contained in the convex hull of some r points of S where $r \le k+1$.

Proof. For a proof of theorem 2.3 the reader is referred to [1]. □

Theorem 2.4. *Every real linear functional on V has a canonical representation.*

Proof. The case in which $k = 1$ is trivial, so assume $k > 1$. The set of $v \in V$ such that $||v|| \leq 1$ is compact, so there must be an extremal element v_0. Let Z denote the kernel of F. By the rank nullity theorem, dim $Z = k - 1$. For $v \in Z$,

$$||F||||v_0|| = |Fv_0| = |F(v + v_0)| \leq ||F||||v + v_0||$$

implying $||v + v_0|| \geq ||v_0||$. By theorem 2.3 there exist r extremal points of v_0, $r \leq$ dim $Z = k$, and positive real scalars $\lambda_1, \ldots, \lambda_r$ such that

$$\sum_{i=1}^{r} \lambda_i = 1 \quad \text{ and } \quad \sum_{i=1}^{r} \lambda_i v_0(y_i)v(y_i) = 0$$

for every $v \in Z$.

Now suppose v is an arbitrary element of V. Note that the element $u = (Fv)v_0 - (Fv_0)v$ is in Z. Then

$$(Fv) \sum_{i=1}^{r} \lambda_i v_0(y_i)^2 = (Fv_0) \sum_{i=1}^{r} \lambda_i v_0(y_i)v(y_i)$$

implying

$$(Fv) \sum_{i=1}^{r} \lambda_i ||v_0||^2 = ||F||||v_0|| \sum_{i=1}^{r} \lambda_i v_0(y_i)v(y_i).$$

Then

$$Fv = ||F|| \sum_{i=1}^{r} [\lambda_i \text{ sgn } v_0(y_i)]v(y_i),$$

and letting the α_i from 2.2 equal

$$\alpha_i = \frac{\lambda_i \text{ sgn } v_0(y_i)}{\sum_{i=1}^{r} \lambda_i} ||F||$$

completes the proof of the theorem. □

The next intermediate result gives some more information about the nature of Chebyshev polynomials.

Theorem 2.5. *Let F be a real linear functional on $\mathscr{P}_n$. If F has some canonical representation with $r = n + 1$ then F has a unique extremal. This unique extremal is one of ± 1 or $\pm T_n$.*

Proof. Let v_0 be an extremal for F. Then

$$\sum_{j=1}^{n+1} |\alpha_j| = ||F|| = Fv_0 = \sum_{j=1}^{n+1} \alpha_j v_0(y_j)$$

implying $v_0(y_j) = \text{sgn } \alpha_j$ for $1 \leq j \leq n+1$. Therefore v_0 has $n+1$ extremal points on $[-1, 1]$. The theorem follows from 1.3 because F is a functional on $\mathscr{P}_n$. □

At last it is possible to prove theorem 2.1

Proof. The first step is to show that F has a canonical representation with $r = n+1$. F must have some canonical representation

$$Fp = \sum_{j=1}^{r} \alpha_j p(y_j).$$

If $r \leq n$ one can construct a polynomial $p_0 \in \mathscr{P}_n$ such that $p_0(y_j) = 0$, $1 \leq j \leq r \leq n$, and p_0 has n distinct zeroes. Therefore $r = n+1$, and $|F\tilde{T}_n|$ must equal $||F||$ by theorem 2.5. It is also true that $y_j = \eta_{j-1}$ for $1 \leq j \leq n+1$. For $p \in \mathscr{P}_n$,

$$|Fp| \leq \sum_{j=1}^{r} |\alpha_j||p(y_j)| \leq \sum_{j=1}^{r} |\alpha_j| = ||F|| = |FT_n(x)|$$

with equality holding iff $p(y_j) = \text{sgn } \alpha_j$ for $1 \leq j \leq n+1$. In light of corollary 1.3 equality is only possible if $p = \pm\tilde{T}_n$. □

Acknowledgments. I thank my mentor Preston Wake for guiding my studies this summer. I thank Professor May organizing and directing the REU. I thank Professor Sally for being an excellent teacher over the past year.

References

[1] Theodore J. Rivlin. The Chebyshev Polynomials. John Wiley & Sons, 1974.