First formalized by Weierstrass in 1885, **approximation theory** concerns the best approximation of arbitrary functions by some class of simpler functions. Weierstrass was originally interested in the approximation of complex-analytic functions by power series, but his explorations lead him to prove that both algebraic and trigonometric polynomials are dense in the space of continuous functions. The structure of an approximation problem involves three central components: a *function class* containing the function to be approximated, a *form* of approximating function, and a *norm* for measuring approximation error. Three types of approximation are commonly sought after [1]:

Definition 1. Let $\mathcal{U}$ be a subspace of some normed linear $\mathcal{F}$, and fix $f \in F$. Then,

I. A function $p^* \in \mathcal{U}$ is a **good** approximation with error $\varepsilon > 0$ if $||f - p^*|| < \varepsilon$.
II. A function $p_B^* \in \mathcal{U}$ is a **best** approximation if $||f - p^*|| \leq ||f - p||$ for any other $p \in \mathcal{U}$.
III. A function $p_N^* \in \mathcal{U}$ is a **near-best** approximation if $||f - p_N^*|| \leq (1+\rho)||f - p_B^*||$ for small $\rho > 0$.

In these notes, we seek to approximate the class $\mathcal{C}[a,b]$ of continuous functions on the interval $[a,b]$ by degree-n polynomials in $\mathcal{P}_n$ using the L^∞-norm to measure fit. This scenario is referred to as **minimax polynomial approximation**, since the best (or *minimax*) approximation solves

$$p_n^* = \arg \min_{p_n \in \mathcal{P}_n} \max_{a \leq x \leq b} |f(x) - p_n(x)| \tag{1}$$

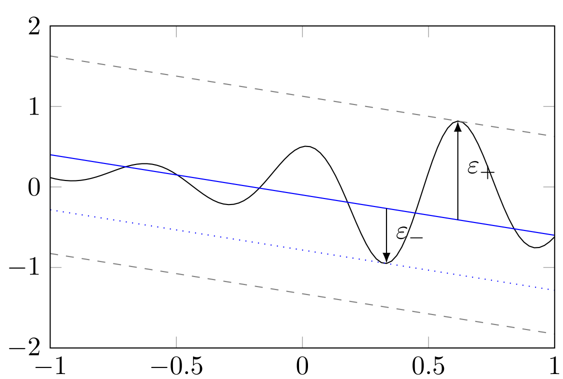
The Alternating Property of Minimax Polynomials

Minimax approximating polynomials of degree n are characterized by the **alternating property**, which requires that the approximation error oscillates between its equal maximum and minimum values a total of $n+2$ times. This result, which we will prove rigorously later, seems outright preposterous at first glance! Chebyshev's original notes (1854) were apparently rather haphazard [2]; we owe our modern understanding of this remarkable fact to Kirchberger (1903), Borel (1905), Haar (1918), and de La Vallée Poussin (1918).

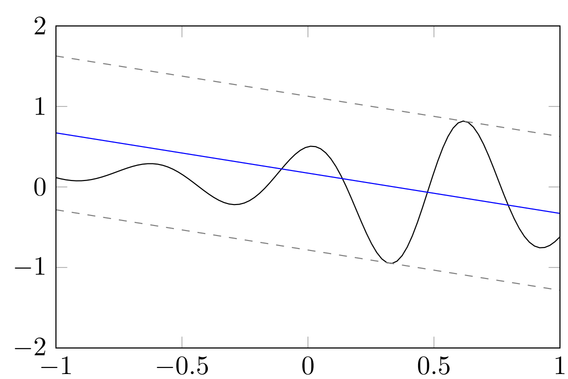
Definition 2. An **alternating set** for $\varphi \in \mathcal{C}[a,b]$ is a sequence $a \leq x_1 < \cdots < x_n \leq b$ of distinct points alternating between $\pm||\varphi||_\infty$, that is, $|\varphi(x_k)| = ||\varphi||_\infty$ and $\varphi(x_{k+1}) = -\varphi(x_k)$.

Definition 3. The polynomial approximation $p_n \in \mathcal{P}_n$ to $f \in \mathcal{C}[a,b]$ is said to satisfy the **alternating property** if there is an alternating set of at least $n+2$ points for the function $f - p_n \in \mathcal{C}[a,b]$.

It is easy to see that at least two alternating points must exist for any minimax approximation. Otherwise, there is some "wiggle room" left over that can be used to construct a better approximation simply by shifting the original, as in Figure 1 below.



(a) Original Approximation



(b) Shifted Approximation

Figure 1: Unless two alternating points exist, we can always find a better approximation.

Lemma 1 (following [1]). *Suppose $p^* \in \mathcal{P}_n$ is the best approximation to $f \in \mathcal{C}[a,b]$ out of $\mathcal{P}_n$. Then f has two alternating points, that is, there are at least two distinct points $x_1, x_2 \in [a,b]$ such that*

$$f(x_1) - p(x_1) = p(x_2) - f(x_2) = ||f - p||_\infty$$

(1) The function $|f - p^*|$ is continuous with compact support, so attains its maximum value $||f - p^*||_\infty$ at some point $x_1 \in [a,b]$. Define the upper and lower approximation errors

$$\varepsilon_+ = \max_{a \le x \le b} f(x) - p^*(x) \qquad \varepsilon_- = \min_{a \le x \le b} f(x) - p^*(x)$$

(2) Assume towards contradiction that $\varepsilon_+ \neq \varepsilon_-$. We may take the upper margin to be larger,

$$f(x_1) - p^*(x_1) = ||f - p||_\infty = \varepsilon_+ > -\varepsilon_-$$

(3) In particular, $\varepsilon_+ + \varepsilon_- > 0$ and so $q = p^* + (\varepsilon_+ + \varepsilon_-)/2 \in \mathcal{P}_n$ is distinct from p^*. It is clear that

$$||f - q||_\infty \le \left(\frac{\varepsilon_+ + \varepsilon_-}{2} \right) < \varepsilon_+ = ||f - p^*||_\infty$$

(4) Therefore, q is a better approximation to f than p^*, a contradiction! □

From the lemma, it is not too difficult to prove that the best approximating constant to a function $f \in \mathcal{C}[a,b]$ is the average of its maximum and minimum values. Similar reasoning applies for the best degree-n approximating polynomial, provided that we have a large enough alternating set. Before proving the equioscillation theorem, we characterize best approximations by a related property of the extreme set[1].

Theorem 1 (Kolmogorov Criterion). *Let $\mathcal{U}$ be a finite-dimensional subspace of $\mathcal{C}[K]$, with K compact. Then a function $p^* \in \mathcal{U}$ is a best approximation to $f \in \mathcal{C}[K]$ if and only if no other approximation $q \in \mathcal{U}$ has the same sign as $f - p^*$ on its extreme set $\mathcal{Z}_{f-p^*} = \{x \in K \mid |f(x) - p^*(x)| = ||f - p||_\infty\}$. That is,*

$$\min_{x \in \mathcal{Z}} [f(x) - p^*(x)] q(x) \le 0 \quad \forall\, q \in \mathcal{U}$$

Proof. Note that $p^* \in \mathcal{U}$ is a best approximation to $f \in \mathcal{C}[K]$ if and only if $0 \in \mathcal{U}$ is a best approximation to $\varphi = f - p^*$. Therefore, it suffices prove 0 is a best approximation to $\varphi \in \mathcal{C}[K]$ if and only if no other $q \in U$ has the same sign as φ on its extreme set.

(1) $\underline{\Longleftarrow}$ Suppose $0 \in \mathcal{U}$ is not a best approximation to $\varphi \in \mathcal{C}[K]$.
   (a) Then there is a better approximation $q \in \mathcal{U}$ with $||\varphi - q||_\infty < ||\varphi||_\infty$.
   (b) For any $x \in \mathcal{Z}_\varphi$ in the extreme set, unless $\operatorname{sign} \varphi(x) = \operatorname{sign} q(x)$, we must have

$$|\varphi(x) - q(x)| \ge |\varphi(x)| = ||\varphi||_\infty$$

   (c) This would imply $||\varphi - q||_\infty > ||\varphi||_\infty$, a contradiction! Thus $\operatorname{sign} q$ agrees with $\operatorname{sign} \varphi$ on $\mathcal{Z}_\varphi$.

For the converse, we demonstrate that any function violating the Kolmogorov criterion can be modified to yield a better approximation. If q has the same sign as φ on its extreme set and we subtract a sufficiently small multiple of q from φ, the difference will be of strictly smaller magnitude than φ near the extreme set. Provided that the multiple we subtract is small enough, the difference will not exceed the extreme value outside the extreme set, giving a strictly better approximation.

[1] At this point, it would be nice to reason inductively about minimax polynomials, perhaps applying Lemma 1 to polynomials of successively lower degree, stopping at zero. Unfortunately, such a satisfying proof eludes me.

(2) $\underline{\Longrightarrow}$ Suppose $0 \in \mathcal{U}$ is a best approximation to $\varphi \in \mathcal{C}[K]$.

(a) Assume towards contradiction that Kolmogorov's criterion fails. Then for some $\varepsilon > 0$ and $q \in \mathcal{U}$,

$$\varphi(x)q(x) = ||\varphi||_\infty q(x) > 2\varepsilon \qquad \forall\, x \in \mathcal{Z}_\varphi$$

(b) Because φ is continuous, there is an open set $G \subset K$ containing $\mathcal{Z}$ on which

$$\varphi(x)q(x) > \varepsilon \qquad \forall\, x \in G$$

(c) Claim: The function $\varphi - \lambda q \in \mathcal{U}$ is a better approximation to f than zero is, for some $\lambda > 0$.

(i) For any $x \in G$ in a neighborhood of the extreme set, we have

$$\begin{aligned}
|\varphi(x) - \lambda q(x)|^2 &= (\varphi(x) - \lambda q(x))^2 && (2)\\
&= \varphi(x)^2 - 2\lambda\varphi(x)q(x) + \lambda^2 q(x)^2 && (3)\\
&< ||\varphi||_\infty^2 - \lambda\varepsilon + \lambda^2 ||q||_\infty^2 && (4)\\
&= ||\varphi||_\infty^2 - \lambda(\varepsilon - \lambda||q||_\infty^2) && (5)
\end{aligned}$$

thus if $0 < \lambda < \varepsilon/||q||_\infty^2$, we have $||\varphi - \lambda q||_\infty < ||\varphi||_\infty$.

(ii) The complement $F \equiv K \setminus G$ contains no extreme points $\mathcal{Z}$, so $|\varphi(x)| < ||\varphi||_\infty$ on F. Because F is closed, there exists $\delta > 0$ such that

$$|\varphi(x)| < ||\varphi(x)||_\infty - 2\delta \qquad \forall\, x \in K \setminus G$$

(iii) Therefore, if $\lambda < \delta/||q||_\infty$, then for any $x \in K \setminus G$,

$$\begin{aligned}
|\varphi(x) - \lambda q(x)| &\le |\varphi(x)| + \lambda|q(x)|\\
&< ||\varphi||_\infty - 2\delta + \lambda||q||_\infty\\
&< ||\varphi||_\infty - \delta
\end{aligned}$$

(iv) Together, we find that if $\lambda < \min\left\{\frac{\varepsilon}{||q||_\infty^2}, \frac{\delta}{||q||_\infty}\right\}$, then $||\varphi - \lambda q||_\infty < ||\varphi||_\infty$.

(d) This is a contradiction! Thus every best approximation satisfies Kolmogorov's criterion. □

Corollary 1 (Kolmogorov Criterion, Restated)**.** *Let $\mathcal{U}$ be a finite-dimensional subspace of $\mathcal{C}[K]$, with K compact. Then a function $p^* \in \mathcal{U}$ is a best approximation to $f \in \mathcal{C}[K]$ if and only if*

$$\max_{x \in \mathcal{Z}} [f(x) - p^*(x)]q(x) \ge 0 \quad \forall\, q \in \mathcal{U}$$

We now establish a series of results which assert that for p_n to be a best approximation, it is both *necessary* and *sufficient* that the alternating property holds, that only one polynomial has this property, and that there is only one best approximation [1].

Theorem 2 (Chebyshev Equioscillation)**.** *A polynomial $p^* \in \mathcal{P}_n$ is a best approximation to $f \in \mathcal{C}[a,b]$ if and only if there is an alternating set for $f - p$ consisting of at least $n+2$ points.*

(1) $\underline{\Longrightarrow}$ Suppose $\varphi \equiv f - p^*$ takes the value $||\varphi||_\infty$ on only $m < n+2$ points.

(a) We will apply Kolmogorov's criterion to show that there is a better approximation of degree m.

(b) Let $\mathcal{Z}_\varphi$ be the extreme set of $\varphi = f - p^*$, and define *high* and *low* points of φ as

$$\mathcal{Z}_{\text{high}} = \{x \in [a,b] \mid \varphi(x) = ||\varphi||_\infty\} \qquad \mathcal{Z}_{\text{low}} = \{x \in [a,b] \mid \varphi(x) = -||\varphi||_\infty\}$$

(c) By hypothesis, there exist disjoint open intervals $U_1, \ldots, U_m \subset [a,b]$ covering $\mathcal{Z}_\varphi$ such that

(i) each interval contains either only low points or only high points, and

(ii) adjacent intervals contain extreme points of opposite sign.

(d) Assume that U_1 contains only high points. Construct the offending polynomial $q \in \mathcal{P}_n$ as follows:

(i) Select points z_k between adjacent intervals, such that $U_1 < z_1 < U_2 < \cdots < z_{m-1} < U_m$.

   (ii) Define $q(x) = \prod_{k=1}^{m-1}(z_k - x)$ of degree $m - 1 \leq n$.
  (e) Observe that q has the same sign as φ on the extreme set $\mathcal{Z}_\varphi$, since q alternates sign across adjacent intervals U_k, starting out positive. Noting that $\mathcal{P}_n$ is a subspace of $\mathcal{C}[a, b]$, Kolmogorov's criterion tells us that the polynomial $p^* \in \mathcal{P}_n$ cannot be a best approximation to f.
  (f) This is a contradiction! Thus at least $m \geq n + 2$ alternating points must exist.

(2) $\underline{\Longleftarrow}$ Suppose the approximation $p^* \in \mathcal{P}_n$ to $f \in \mathcal{C}[a, b]$ satisfies the alternating property.
  (a) Then $f - p^*$ achieves its maximum magnitude at $n + 2$ distinct points with alternating sign.
  (b) Certainly, no other function $q \in \mathcal{P}_n$ has the same sign as $f - p^*$ on its extreme set, since a nonzero element of $\mathcal{P}_n$ cannot have $n + 1$ zeros. Thus $p^* \in \mathcal{P}_n$ is a best approximation to f. □

Note that because $f - p^*$ alternates $n + 2$ times, it must have at least $n + 1$ zeros, and thus p^* actually *nterpolates* f at some $n + 1$ points. This theorem also applies more generally for a special class of function spaces called **Haar spaces**, which share with polynomials the property that each function is uniquely determined by its values on $n + 1$ points. For example, a version of the equioscillation theorem derived for trigonometric polynomials is commonly used to characterize minimax approximations for periodic functions. It is worth noting that for algebraic polynomials, the number of alternating points required is $\dim \mathcal{P}_n + 1$.

Uniqueness of Best Polynomial Approximation

Theorem 3 (Carothers, Thm. 4.5)**.** *The polynomial of best approximation to $f \in \mathcal{C}[a, b]$ out of $\mathcal{P}_n$ is unique.*

Proof. Suppose $p, q \in \mathcal{P}_n$ are both best approximations to $f \in \mathcal{C}[a, b]$, satisfying $||f - p||_\infty = ||f - q||_\infty \equiv E$. Then, their average $r \equiv \frac{p+q}{2} \in \mathcal{P}_n$ is also a best approximation, since $f - r$ is the average of $f - p$ and $f - q$. By the Chebyshev equioscillation theorem, then, $f - r$ has an alternating set $x_0, x_1, \ldots, x_{n+1}$ containing $n + 2$ points, and so

$$(f - p)(x_k) + (f - q)(x_k) = \pm 2E \qquad \text{(alternating for } k = 1, \ldots, n) \tag{6}$$

However, since $|f - p|$ and $|f - q|$ are bounded by E, we must have $f - p = f - q$ on the extreme set. In particular, the degree-n polynomial $q - p$ has $n + 2$ zeros, so must vanish. Hence $q = p$. □

References

[1] John C. Mason and David C. Handscomb. *Chebyshev Polynomials*. CRC Press, 2003.

[2] Neal L. Carothers. A Short Course on Approximation Theory. *Bowling Green State University*, 1998.

[3] Alexei Shadrin. Approximation Theory, Lecture 5. University of Cambridge, Mathematical Tripos Part III, 2005. URL http://www.damtp.cam.ac.uk/user/na/PartIIIat/b05.pdf.

[4] Douglas N. Arnold. A Concise Introduction to Numerical Analysis. *Institute for Mathematics and its Applications, Minneapolis*, 2001.