Stone-Weierstrass Theorem

Ethan Kessinger

Weierstrass Approximation Theorem

Theorem: *Every continuous, non-differentiable, real-valued function on a closed interval* $[a, b] \subseteq \mathbb{R}$ *can be approximated by a sequence of polynomials. i.e.:*

$$\forall f \epsilon C^0([a, b], \mathbb{R})\ \exists (p_n) \subseteq \mathcal{P}([a, b], \mathbb{R}) : \forall \varepsilon > 0\ \exists N \epsilon \mathbb{N} : \forall n \geq N\ \|f - p_n\| < \varepsilon.$$

Where $\mathcal{P}$ *is the set of polynomials and* $\|\cdot\|$ *is the supremum norm for functions. This is equivalent to the set of polynomials being dense in* $C^0([a, b], \mathbb{R})$*.*

Before beginning the proof, it is necessary to intoduce a definition and a few related formulas.

Definition: Let $f \epsilon C^0([0, 1], \mathbb{R})$. The *Bernstein Polynomials* of f are the polynomials which construct the sequence $(p_n) \subseteq \mathcal{P}([a, b], \mathbb{R})$ such that

$$\forall n \epsilon \mathbb{N}\ p_n = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1 - x)^{n-k} \tag{1}$$

We then define the functions $r_k : [a, b] \to \mathbb{R}$:

$$x \mapsto r_k(x) := \binom{n}{k} x^k (1 - x)^{n-k}. \tag{2}$$

From this definition, we can obtain two useful formulas, namely

$$\sum_{k=0}^{n} r_k(x) = 1, \tag{3}$$

and

$$\sum_{k=0}^{n} (k - nx)^2 r_k(x) = nx(1 - x). \tag{4}$$

Taking the sum of r_k as an n^{th} degree binomial expression, we can set $\sum_{k=0}^{n} r_k(x) = (x + y)^n$, where $y = 1 - x$. This naturally implies (3). If we temporarily fix y and instead differentiate the binomial equation twice in terms of x, we obtain

$$n(x+y)^{n-1} = \sum_{k=0}^{n} \binom{n}{k} k x^{k-1} y^{n-k}, \tag{5}$$

and

$$n(n-1)(x+y)^{n-2} = \sum_{k=0}^{n} k(k-1) x^{k-2} y^{n-k}. \tag{6}$$

Setting $y = 1 - x$ again, multiplying (5) by x, and multiplying (6) by x^2 yields

$$\sum_{k=0}^{n} k r_k(x) = nx, \tag{7}$$

and

$$\sum_{k=0}^{n} k(k-1) r_k(x) = n(n-1)x^2. \tag{8}$$

From (7) and (8), we obtain

$$\sum_{k=0}^{n} k^2 r_k(x) = n(n-1)x^2 + \sum_{k=0}^{n} k r_k(x) = n(n-1)x^2 + nx. \tag{9}$$

Using (3), (7), and (9), we construct

$$\sum_{k=0}^{n} (k-nx)^2 r_k(x) = \sum_{k=0}^{n} k^2 r_k(x) - 2nx \sum_{k=0}^{n} k r_k(x) + (nx)^2 \sum_{k=0}^{n} r_k(x) = nx(1-x),$$

which is the expected result in (4).

Proof of Weierstrass Approximation Theorem:

Firstly, instead of considering an arbitrary closed interval $[a, b] \subseteq \mathbb{R}$, we need only consider the closed interval $[0, 1]$, since any closed interval is homeomorphic to $[0, 1]$. Our goal is to show that the sequence of Bernstein polynomials converges uniformly to $f \epsilon C^0([0, 1], \mathbb{R})$. Using the definition of r_k and (3), we can rewrite f and its n^{th} Bernstein polynomial as

$$f(x) = \sum_{k=0}^{n} f(x) r_k(x) \quad \wedge \quad p_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) r_k(x).$$

By doing this, the difference between the two functions can be written as

$$p_n - f = \sum_{k=0}^{n} \left(f\left(\frac{k}{n}\right) - f \right) r_k. \tag{10}$$

Since $[0, 1]$ is compact in $\mathbb{R}$ and f is continuous, f is uniformly continuous on [0,1]. This means, given some $\varepsilon > 0$, we can find a $\delta > 0$ such that

$$\forall x, y \epsilon [0, 1] : |x - y| < \delta \ |f(x) - f(y)| < \frac{\varepsilon}{2}. \tag{11}$$

Using (11), we can group the indices from (10) into two groups, namely, for an arbitrary $x \epsilon [0, 1]$,

$$K_1 = \{k \epsilon \{0, \dots, n\} : |\frac{k}{n} - x| < \delta\} \wedge K_2 = \{0, \dots, n\} \backslash K_1.$$

With this grouping of indices, it becomes possible to arrange (10) so that

$$\forall x \epsilon [0, 1] \ |p_n(x) - f(x)| \leq \sum_{k \epsilon K_1} |f\left(\frac{k}{n}\right) - f(x)| r_k(x) + \sum_{k \epsilon K_2} |f\left(\frac{k}{n}\right) - f(x)| r_k(x).$$

Applying (3) and (11), we obtain

$$|p_n(x) - f(x)| < \frac{\varepsilon}{2} + \sum_{k \epsilon K_2} |f\left(\frac{k}{n}\right) - f(x)| r_k(x). \tag{12}$$

All that is left is to show the second sum is also less than $\frac{\varepsilon}{2}$. We begin by noting that, by (4),

$$nx(1 - x) = \sum_{k=0}^{n} (k - nx)^2 r_k(x) \geq \sum_{k \epsilon K_2} (k - nx)^2 r_k(x).$$

Since $k \epsilon K_2$ implies $|\frac{k}{n} - x| \geq \delta$, we obtain

$$nx(1 - x) \geq \sum_{k \epsilon K_2} (k - nx)^2 r_k(x) \geq \sum_{k \epsilon K_2} (n\delta)^2 r_k(x). \tag{13}$$

With this, and the fact that max$(x(1 - x)) = \frac{1}{4}$, we can rewrite (13) as

$$\sum_{k \epsilon K_2} r_k(x) \leq \frac{nx(1 - x)}{(n\delta)^2} \leq \frac{1}{4n\delta^2}. \tag{14}$$

Since $|f\left(\frac{k}{n}\right) - f(x)| \leq 2M, M = \|f\|$, we can combine (12) and (14) to obtain

$$|p_n(x) - f(x)| < \frac{\varepsilon}{2} + \frac{M}{2n\delta^2} \leq \varepsilon$$

for sufficiently large n. This implies $(p_n) \to f$ uniformly on [0,1], completing the proof. □

Stone-Weierstrass Theorem

Before introducing the theorem itself, it is necessary to include several definitions which are used in the proof. Let X be an arbitrary topological space.

Definition: $\mathcal{A} \subseteq C^0(X;\mathbb{R})$ is a *function algrebra* iff it is closed under addition, scalar multiplication, and function multiplication. i.e.:

$$\forall f_1, f_2, g_1, g_2 \epsilon \mathcal{A} \wedge \alpha_1, \alpha_2, \beta_1, \beta_2 \epsilon \mathbb{R}\ (\alpha_1 f_1 + \beta_1 g_1)(\alpha_2 f_2 + \beta_2 g_2) \epsilon \mathcal{A}.$$

Definition: A function algrebra $\mathcal{A} \subseteq C^0(X;\mathbb{R})$ *vanishes* at a point $p \epsilon X$ iff $\forall f \epsilon \mathcal{A}\ f(p) = 0$.

Definition: A function algrebra $\mathcal{A} \subseteq C^0(X;\mathbb{R})$ *separates points* iff every pair of distinct points have distinct values for some function in the algebra. i.e.:

$$\forall p_1, p_2 \epsilon X : p_1 \neq p_2\ \exists f \epsilon \mathcal{A} : f(p_1) \neq f(p_2).$$

Along with these definitions, two lemmas will be required for the proof.

Lemma I: (2-Point Interpolation) *If a function algebra $\mathcal{A} \subseteq C^0(X;\mathbb{R})$ vanishes nowhere and separates points then there exists an $f \epsilon \mathcal{A}$ with specified values at any pair of distinct points. i.e.:*

$$\forall c_1, c_2 \epsilon \mathbb{R}\ \exists f \epsilon \mathcal{A} \wedge \exists p_1, p_2 \epsilon X : p_1 \neq p_2 \wedge f(p_1) = c_1 \wedge f(p_2) = c_2.$$

Proof of Lemma I:

Let $c_1, c_2 \epsilon \mathbb{R}$ and $p_1, p_2 \epsilon X$. Since $\mathcal{A}$ vanishes nowhere, there exist $g_1, g_2 \epsilon \mathcal{A}$ such that $g_1(p_1) \neq 0 \neq g_2(p_2)$. This means that there exists a $g \epsilon \mathcal{A}$ such that $g = g_1^2 + g_2^2$ and $g(p_1) \neq 0 \neq g(p_2)$ by contruction.

Additionally, since $\mathcal{A}$ separates points, we can find an $h \epsilon \mathcal{A}$ such that $h(p_1) \neq h(p_2)$. We then construct the matrix

$$H = \begin{bmatrix} g(p_1) & g(p_1)h(p_1) \\ g(p_2) & g(p_2)h(p_2) \end{bmatrix}.$$

Since $g(p_1), g(p_2) \neq 0$ and $h(p_1) \neq h(p_2)$,

$$\det H = g(p_1)g(p_2)h(p_2) - g(p_1)g(p_2)h(p_1) = g(p_1)g(p_2)(h(p_2) - h(p_1)) \neq 0.$$

This implies that the columns of H are linearly independent and there is a solution $(f_1, f_2) \epsilon \mathcal{A} \times \mathcal{A}$ such that

$$f_1 g + f_2 g h = f \epsilon \mathcal{A} \wedge f(p_1) = c_1 \wedge f(p_2) = c_2,$$

completing the proof. □

Lemma II.i: *The closure of a function algebra* $\mathcal{A} \subseteq C^0(X;\mathbb{R})$ *is a function algebra.*

Proof of Lemma II.i:

The conclusion is easily obtained via the definition of a function algebra.

Lemma II.ii: *If* $\mathcal{A} \subseteq C^0(X;\mathbb{R})$ *is a function algebra and* $f \epsilon \overline{\mathcal{A}}$ *then* $|f| \epsilon \overline{\mathcal{A}}$*.*

Proof of Lemma II.ii:

Let $f \epsilon \overline{\mathcal{A}}$ and $\varepsilon > 0$. Using the Weierstrass Approximation Theorem, we can find a polynomial $p \epsilon \mathcal{P}([-\|f\|, \|f\|]; \mathbb{R})$ such that

$$\sup\{|p(\omega) - |\omega|| : |\omega| \leq \|f\|\} < \frac{\varepsilon}{2} \tag{15}$$

since $|\cdot|$ is continuous on the interval $[-\|f\|, \|f\|]$. Additionally, since $|p(0) - 0| < \frac{\varepsilon}{2}$, the constant term is less than $\frac{\varepsilon}{2}$. In order to elimiate the constant term of p, we define $q \epsilon \mathcal{P}([-\|f\|, \|f\|]; \mathbb{R})$ such that

$$\forall \omega \epsilon [-\|f\|, \|f\|] \ \omega \mapsto q(\omega) := p(\omega) - p(0). \tag{16}$$

Using (15) and (16), we obtain

$$\sup\{|q(\omega) - |\omega|| : |\omega| \leq \|f\|\} < \varepsilon. \tag{17}$$

We can write $q(\omega)$ as an n^{th} degree polynomial so that

$$q(\omega) = a_1\omega + \cdots + a_n\omega^n,$$

for some $n \epsilon \mathbb{N}$. Let $g \epsilon C^0(X;\mathbb{R})$ such that

$$g = a_1 f + \cdots + a_n f^n. \tag{18}$$

Since $\overline{\mathcal{A}}$ is a function algebra by Lemma II.i and $f \epsilon \overline{\mathcal{A}}$, g is also in $\overline{\mathcal{A}}$. Using (17) and (18), we obtain

$$\forall x \epsilon X \ |g(x) - |f(x)|| = |q(f(x)) - |f(x)|| < \varepsilon.$$

This allows us to construct a sequence of functions in $\overline{\mathcal{A}}$ converging uniformly to $|f|$, completing the proof. □

Lemma II: (Inclusion of Pointwise Maximum and Minimum) *If* $f, g \epsilon \overline{\mathcal{A}} \subseteq C^0(X;\mathbb{R})$ *then* $\max(f,g), \min(f,g) \epsilon \overline{\mathcal{A}}$ *where*

$$\max(f,g) = \frac{f+g}{2} + \frac{|f-g|}{2} \quad \wedge \quad \min(f,g) = \frac{f+g}{2} - \frac{|f-g|}{2}.$$

Proof of Lemma II:

The conclusion is easily obtained via Lemma II.ii and the definition of a function algebra.

Theorem: *If X is a compact, hausdorff topological space and $\mathcal{A} \subseteq C^0(X,\mathbb{R})$ is a function algebra that vanishes nowhere and separates points, then $\mathcal{A}$ is dense in $C^0(X,\mathbb{R})$.*

Proof of Stone-Weierstrass Thoerem:

Let $f \epsilon C^0(X;\mathbb{R})$ and $\varepsilon > 0$. The goal is to construct a function $g \epsilon \overline{\mathcal{A}}$ such that

$$\forall x \epsilon X \ f(x) - \varepsilon < g(x) < f(x) + \varepsilon.$$

Let $p, q \epsilon X$ be distinct points. Using Lemma I, we can find $h_{pq} \epsilon \mathcal{A}$ such that

$$h_{pq}(p) = f(p) \quad \wedge \quad h_{pq}(q) = f(q). \tag{19}$$

If we let q vary, then we can find an open neighborhood $U_q \subseteq X$ of each q so that

$$x \epsilon U_q \Rightarrow f(x) - \varepsilon < h_{pq}(x), \tag{20}$$

since h_{pq} is a continuous function. The compactness of X implies that there is an open subcovering composed of finitely many neighborhoods $U_{q_1}, \ldots, U_{q_n}$. If we define $g_p \epsilon C^0(X;\mathbb{R})$ such that

$$\forall x \epsilon X \ g_p(x) = \max(h_{pq_1}(x), \ldots, h_{pq_n}(x)), \tag{21}$$

then $g_p \epsilon \overline{\mathcal{A}}$ by Lemma II, and, by using (19), (20), and (21),

$$\forall x \epsilon X \ f(x) - \varepsilon < g_p(x) \quad \wedge \quad g_p(p) = f(p). \tag{22}$$

If we let p vary, then, by continuity, we can find an open neighborhood $V_p \subseteq X$ of each p so that

$$x \epsilon V_p \Rightarrow g_p(x) < f(x) + \varepsilon. \tag{23}$$

Again, since X is compact, we can find a finite subcovering composed of neighborhoods $V_{p_1}, \ldots, V_{p_n}$. By defining $g \epsilon C^0(X;\mathbb{R})$ so that

$$\forall x \epsilon X \; g(x) = \min(g_{p_1}(x), \dots, g_{p_n}(x)), \tag{24}$$

so $g \epsilon \overline{\mathcal{A}}$. We can apply (22) and (23) to show

$$\forall x \epsilon X \; f(x) - \varepsilon < g(x) < f(x) + \varepsilon,$$

so we can construct a sequence of functions in $\overline{\mathcal{A}}$ converging uniformly to f, completing the proof. □

Remarks on Stone-Weierstrass Theorem:

1.) If X was not hausdorff, then any function algebra $\mathcal{A} \subseteq C^0(X; \mathbb{R})$ would not be able to separate points, so the property is implicitly required for the theorem.

2.) As an example of a noncompact space in which Stone-Weierstrass fails in general, set $X = \mathbb{R}$ with the topology of uniform convergence and define the function algebra $\mathcal{A} \subseteq C^0(\mathbb{R}; \mathbb{R})$ so that $\mathcal{A} = C^0_b(\mathbb{R}; \mathbb{R})$. $\mathcal{A}$ vanishes nowhere and separates points, but the distance between any function in $\mathcal{A}$ and an undbounded function in $C^0(\mathbb{R}; \mathbb{R})$ is infinite under the suprememum norm.

Corollary of Stone-Weierstrass Theorem: *Any 2π-periodic continuous function on $\mathbb{R}$ can be approximated by trigonometric polynomials of the form*

$$T(x) = c_0 + \sum_{k=1}^{n} a_k \text{cos} kx + \sum_{k=1}^{n} b_k \text{sin} kx.$$

Proof of Corollary:

We begin by noticing that we can parameterize the unit circle $S^1 \subseteq \mathbb{R}^2$ by the interal $[0, 2\pi) \subseteq \mathbb{R}$ so that $x \mapsto (\text{cos} x, \text{sin} x)$, in which case, all 2π-periodic continuous functions are also continuous on S^1, which is compact. Using the Sone-Weierstrass theorem, it is sufficient to show that the set of trigonometric polynomials $\mathcal{T} \subseteq C^0(S^1)$ is an algebra that separates points and vanishes nowhere. Using the formulas

$$\begin{aligned}
\text{cos} kx \, \text{sin} jx &= \frac{1}{2}(\text{sin}(k+j)x + \text{sin}(k-j)x) \\
\text{cos} kx \, \text{cos} jx &= \frac{1}{2}(\text{cos}(k+j)x + \text{cos}(k-j)x) \\
\text{sin} kx \, \text{sin} jx &= \frac{1}{2}(\text{cos}(k-j)x - \text{cos}(k+j)x),
\end{aligned}$$

it is simple to show $\mathcal{T}$ is closed under function product, and closure under addition and scalar multiplication are trivial. It is similarly trivial to show $\mathcal{T}$ vanishes nowhere. To show that $\mathcal{T}$ separates points, we need only note that if $\text{sin} a = \text{sin} b$ then $\text{cos} a \neq \text{cos} b$, and vice versa, completing the proof. □