The Stone-Weierstrass Theorem: A constructive proof

Erick Ore Matos

September 11, 2024

Contents

1 Introduction

Polynomials are an easy-to-manage algebraic expression. They have useful properties we can exploit in many settings: they are continuous, can be added and multiplied easily, and their computation is straightforward. The Stone-Weierstrass theorem increases its usefulness, as it states that we can approximate any continuous function defined in a compact space using a sequence of polynomials.

We will develop a constructive proof of the theorem in this document. In this sense, some definitions and concepts of basic topology and metric spaces are required. To familiarize yourself with these concepts, I recommend reviewing Rudin et al. (1964) and Johnsonbaugh and Pfaffenberger (2012). The following definitions are particularly relevant: metric spaces, completeness, compactness, density, convergence of sequences of functions, and Taylor's theorem. All of these concepts are used in the proof.

The document is divided into four additional sections. Section 2 proves a relevant result about series, which will be helpful in Section 3.3. Section 3 defines a metric space using the set of continuous functions, the functions we are trying to approximate. Also, we define an unital subalgebra as a subset of this metric space and characterize its uniform closure, the functions we use to approximate. Section 4 presents and proves the main theorem. In Section 5, we will link the main theorem to the specific Weierstrass theorem, a particular case applied to the continuous functions defined in the compact set $[0, 1]$ and using polynomials as the subalgebra.

The proof of the main theorem is a natural consequence of the properties of functions defined over a compact set and the characteristics of the closure of any unital subalgebra with the separating points property. We will show that this subalgebra is dense in the space of continuous functions. Therefore, we can use a sequence in this subalgebra to approximate any continuous function. It turns out that polynomials are an unital subalgebra with the separating points property, therefore, we can use them to approximate any real-valued continuous function defined in a compact spaces.

2 A relevant first result on series

Theorem 2.1. *For any $0 \leq x \leq 1$, there exists coefficients a_i such that $\sqrt{1-x} = \sum_{i=0}^{\infty} a_i x^n$*

Proof. I use Taylor's theorem to prove that the series exists. The general expression for the derivative is as follows:

$$f^{(n)}(x) = \frac{d^n(\sqrt{1-x})}{dx^n} = (-1)^n \frac{(1 \times 3 \times 5 \times ... \times (2(n-1)-1))}{2^n} (1-x)^{1/2-n}$$

If we evaluate the absolute value of this expression at any x_0, we will notice that it is bounded above:

$$|f^{(n)}(x_0)| < \frac{2 \times 4... \times 2(n-1)}{2^n} |1-x_0|^{1/2-n} = \frac{(n-1)!}{2} |1-x_0|^{1/2-n}$$

The Taylor expansion for the function around $x = 0$ is:

$$P_{n-1}(x) = \sum_{i=0}^{n-1} \frac{f^{(i)}(0)}{i!} x^i$$

By Taylor's theorem, a value x_0 exists such that $0 < x_0 < x$, and the difference between this expansion and the true value can be expressed as:

$$|f(x) - P_{n-1}(x)| = \frac{|f^{(n)}(x_0)|}{n!} |x^n| < |1 - x_0|^{1/2} \frac{1}{2n} |\frac{x}{1 - x_0}|^n < \frac{1}{2n}$$

The last inequality is due to $0 < x_0 < x \leq 1$. The series $P_{n-1}(x)$ converges to $f(x)$.

Therefore, we can approximate the function $f(x) = \sqrt{1-x}$ by a Taylor series. The convergence rate is stable and does not depend on a specific point, hence it also converges uniformly. This result will help us understand why some functions will be part of the closure of a set in Section 3.3. Note that the derivation holds even when $x = 1$. $\square$

3 Defining a metric space for continuous function

3.1 A metric spaces using the set of real-valued continuous functions

In this section, I start defining continuity in a topological sense. We will return to this definition in Section 4, as we will define open sets in the image of functions, and from this definition, we will get open sets as preimages. We use the collection of those preimages to obtain a cover of a domain.

Definition 3.1. *Let X and Y be sets and τ_X and τ_Y a topology for each. We say that a function $f : X \to Y$ is continuous at the point x if the preimage of an open set containing $f(x)$ in τ_Y is an open set in τ_X.*

To discuss approximation, we need a way to measure the distance between two real functions. One way to measure this distance is to use the maximum distance between the functions evaluated at each point. When dealing with real numbers, a generalization of the maximum difference is the least upper bound (or supremum) of the differences, which exists if the functions are bounded. This requirement is fulfilled since continuous real-valued functions on compact sets are bounded.

Definition 3.2. *Given a set $\mathcal{F}(X)$ comprised by all continuous real functions defined in a compact set X. We define a function $d : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ as:*

$$d(f, g) = \sup_{x \in \kappa} |f(x) - g(x)|$$

$(\mathcal{F}(X), d)$ *is a metric space.*

Finally, the following theorem provides a criterion for the convergence of sequences in this metric space. This criterion allows us to characterize limits in the metric space and utilize concepts such as closures and density.

Theorem 3.1. *Every sequence of functions in* $(\mathcal{F}(X), d)$ *converges if and only if uniformly converges.*

Proof. The following needs to be proven:

$$\forall \epsilon > 0, \exists N, n \geq N : \sup_{x \in X} |f_n(x) - f(x)| \iff \forall \epsilon', \exists N', n \geq N' : |f_n(x) - f(x)| < \epsilon'$$

The first part of the proof follows from the fact that the supremum is an upper bound and, therefore, greater than or equal to each element:

$$|f_n(x) - f(x)| \leq \sup_{x \in X} |f_n(x) - f(x)| < \epsilon$$

The second part requires setting an ϵ a little higher than ϵ' to make the inequality strict:

$$|f_n(x) - f(x)| < \epsilon'$$
$$\sup_{x \in X} |f_n(x) - f(x)| \leq \epsilon' < \epsilon$$

□

3.2 Unital subalgebra with the separating points property

This section characterizes a subset of $\mathcal{F}(X)$. We need to define two binary operations in this subset and ensure they are closed in both operations. Also, we will require a scalar multiplication operation with elements of another field. This subset includes all the functions we will use to approximate the set of all continuous functions.

Definition 3.3. *Let* $\mathcal{A}$ *be a family of functions defined in* X*. We say* $\mathcal{A}$ *is a subalgebra if:*

- $\mathcal{A}$ *is a vector space with two binary operations: an addition "*$+$*" and multiplication operation "*$\times$*", and a scalar operation "*$\cdot$*" over an exterior field* $\mathbb{R}$*.*
- *For* $f, g \in \mathcal{A}$ *and* $c \in \mathbb{R}$*:* $f + g \in \mathcal{A}, f \times g \in \mathcal{A}, c \cdot f \in \mathcal{A}$

We are interested in using a subalgebra defined in our metric space $(\mathcal{F}(X), d)$. There is no restriction in defining the addition and multiplication inside the subalgebra, but the scalar multiplication should be defined over the real numbers $\mathbb{R}$. This feature is required because we want to approximate real functions.

In addition to the definition of the subalgebra, we have mentioned two features of the subset: it being unital and having the separating points property. An unital subalgebra includes an identity multiplicative element. We can understand the separating points property as the requirement that given two different points in X, we can find a function that, when evaluated at these two points, the given values are different.

Definition 3.4. *Let $\mathcal{A}$ be a family of functions on a set X. We say $\mathcal{A}$ separates points in X if to every pair of distinct points $x_1, x_2 \in X$, there corresponds a function $f \in \mathcal{A}$ that $f(x_1) \neq f(x_2)$.*

One consequence of the elements in this unital subalgebra is the existence of a function in $\mathcal{A}$ such that, for any two points in the domain X and any two values in the range, there exists a function attaining those values at the specified points. This property is a crucial feature of the subalgebra, as it allows us to define functions for any two arbitrary points in the domain.

Theorem 3.2. *Let $\mathcal{A}$ be a subalgebra on X with the separating points property. Suppose $x_1, x_2 \in X$ and c_1, c_2 are real constants. Then $\mathcal{A}$ contains a function f such that:*

$$f(x_1) = c_1, f(x_2) = c_2$$

Proof. Given a function $g \in \mathcal{F}(X)$, such that $g(x_1) \neq g(x_2)$, we can define the following functions:

$$f_1(x) = \frac{g(x) - g(x_2)}{g(x_1) - g(x_2)} \in \mathcal{A}$$
$$f_2(x) = \frac{g(x) - g(x_1)}{g(x_2) - g(x_1)} \in \mathcal{A}$$

Notice that $f_1(x_1) = 1$, $f_1(x_2) = 0$, $f_2(x_1) = 0$, $f_2(x_2) = 1$. We can define the function:

$$f(x) = c_1 \times f_1(x) + c_2 \times f_2(x) \in \mathcal{A}$$

We can easily verify that $f(x_1) = c_1$ and $f(x_2) = c_2$. □

3.3 The closure of an unital subalgebra with the separating points property

We have characterized a particular subset in $(\mathcal{F}, d)$: the unital subalgebra with the separating points property denoted as $\mathcal{A}$. We can define a closure of this set, denoted as

$\mathcal{B}$, by taking any limits of the elements in the subalgebra in the metric space of function $(\mathcal{F}(X), d)$. We will derive an essential feature of this set: the maximum and the minimum of two elements of the subalgebra belong to the closure.

Lemma 3.3. *Let $(\mathcal{F}(X), d)$ be the set of all continuous functions defined in a compact space X. Let $\mathcal{A}$ be an unital subalgebra with the separating points property defined in that metric space, and let $\mathcal{B}$ be its closure. Then:*

- *$f \in \mathcal{B}$ and $f(x) \geq 0$ for every $x \in X$ then $\sqrt{f} \in \mathcal{B}$*
- *$f \in \mathcal{B}$ then $|f| \in \mathcal{B}$*
- *$f, g \in \mathcal{B}$, given:*

$$\max(f, g) = \begin{cases} f(x) \text{ if } f(x) \geq g(x) \\ g(x) \text{ if } f(x) < g(x) \end{cases} \quad \min(f, g) = \begin{cases} f(x) \text{ if } f(x) \leq g(x) \\ g(x) \text{ if } f(x) > g(x) \end{cases}$$

Then $max(f, g), min(f, g) \in \mathcal{B}$.

Proof. We can easily verify that for any two functions in the closure $f, g \in \mathcal{B}$, the sum and the product also belong to the closure, which is a consequence of Definition 3.4. We normalize the function f to have a range between $[0, 1]$. This is feasible as we require f to be non-negative and to be defined in a compact space, so it is just a rescale operation. We define $g = 1 - f$, $g \in \mathcal{A}$ since f and the multiplicative identity 1 are both elements of $\mathcal{A}$. By Theorem 2.1, there exists coefficients a_i such that:

$$(1 - g(x))^{1/2} = \sum_{i=1}^{\infty} a_i g(x)^n$$

As we know, $g \in \mathcal{B}$, then $g^n \in \mathcal{B}$, since any product of the elements of the subalgebra is in $\mathcal{B}$. To prove that it converges in the metric space $(\mathcal{F}(X), d)$ we require to converge uniformly. Then we need that for all ϵ should exists a N such that for all $n \geq N$ the following holds:

$$\sup |(1 - g(x))^{1/2} - \sum_{i=1}^{n} a_i g(x)^n| < \epsilon$$

In our proof of Theorem 2.1, we noticed that the convergence rate does not depend on the specific point at which it is calculated. Then, for any $g(x)$, the sequence we have defined converges uniformly and $\sqrt{f(x)} = \sqrt{1 + g(x)} \in \mathcal{B}$.

The next two items of the lemma are straightforward to prove. We define a function $g \in \mathcal{B}$ such that $g = \sqrt{f^2} = |f|$. The product of two functions belongs to the closure. Also, the root of a function also belongs to the closure. Then $|f| \in \mathcal{B}$.

Finally, we can write the $\max$ and $\min$ operators as:

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

As f and g belong to the closure, their sum and multiplication by a scalar also belong to it. We have proven that the absolute value of a function in the closure is also an element of this set.. Therefore, the minimum and maximum belong to the closure. □

4 The Stone-Weierstrass theorem

Theorem 4.1. *Let $\mathcal{A}$ be an unital subalgebra in $(\mathcal{F}(X), d)$ with the separating points property, then $\mathcal{A}$ is dense in $\mathcal{F}(X)$. Equivalently, the closure of $\mathcal{A}$ in $(\mathcal{F}(X), d)$, denoted as $\mathcal{B}$, is $\mathcal{F}(X)$.*

The proof is divided into Lemmas 4.2 and 4.3. Most of the details were developed previously, so the reader will find this proof to be a compilation of previous definitions and theorems.

Lemma 4.2. *Given a real-valued continuous function $f \in \mathcal{F}(X)$, a point $x \in X$, and $\epsilon > 0$ there exists a function g_x in the closure $\mathcal{B}$ such that, for all $x' \in X$ such that $x' \neq x$:*

- $g_x(x) = f(x)$
- $g_x(x') > f(x') - \epsilon$

Proof. Given two points $x, x_i \in X$, we define a function such that $h_i(x) = f(x)$ and $h_i(x_i) = f(x_i)$. By Theorem 3.2, this function exists because of the separating points property.

Then, we get a ball $B_\epsilon(f(x_i))$ for each of the points in the domain $x_i \in X$. We will have an open set $J_\epsilon(f(x_i))$ in X as a preimage for that ball as f is continuous. In any of these open sets, for any $x_i \in J_\epsilon(f(x_i))$ by continuity it holds the following:

$$-\epsilon < h_i(x_i') - f(x_i) < \epsilon$$

The union of all balls generated by all the points in the domain will be a cover for X. As X is compact, we can select a finite subcover $\mathcal{I}$ such that:

$$X = \bigcup_{i \in \mathcal{I}} J_\epsilon(f(x_i))$$

If we take the maximum of the functions h_i, we have a function $g_x = \max_{i \in \mathcal{I}}(h_i)$ that bounds $f - \epsilon$ above. For a given x and every $x' \in X$ such that $x' \neq x$ it holds the following:

$$g_x(x) = f(x)$$
$$g_x(x') > f(x') - \epsilon$$

We can observe this procedure in Figure 1a and 1d for two different ϵ. We have many functions h_i that approximate perfectly f at the point x. The function g_x approximates perfectly f at x, and in the rest of the domain, it bounds $f - \epsilon$ above. □

Lemma 4.3. *Given a continuous real-valued function $f \in \mathcal{F}(X)$ there exists a function g in the closure $\mathcal{B}$ such that for every $x \in X$:*

$$|g(x) - f(x) < \epsilon|$$

This final lemma implies any function in $\mathcal{F}(X)$ can be approximated by a sequence of functions in our subalgebra $\mathcal{A}$.

Proof. We know that for every point x in the domain of f, we can define a function g_x that bounds f above. This step involves using those functions to construct a function that belongs to the closure such that the distance to f is lower than any ϵ.

For every x_i, we can define a ball of radius ϵ for the function g_{x_i} around the value $f(x_i)$. By continuity stated as in Definition 3.1, for every x_i we will have an open set $V_\epsilon(f(x_i)) \subset X$ such that, for all $x'_i \in V_\epsilon(f(x_i))$:

$$-\epsilon < g_{x_i}(x'_i) - f(x_i) < \epsilon$$

Similarly to the previous step, the collection of all these open sets forms an open cover for X. By compactness of X we will have a finite subcover $\mathcal{I}'$:

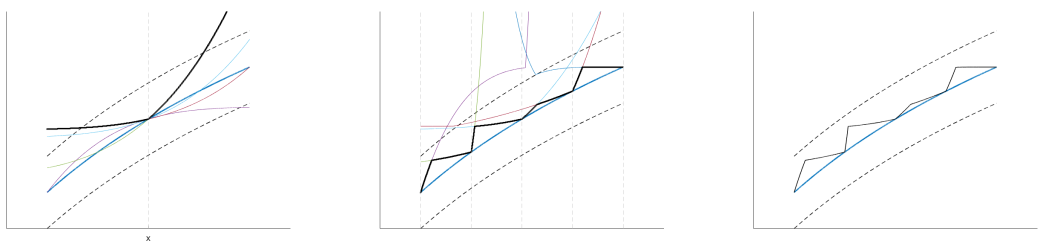
$$X = \bigcup_{i \in \mathcal{I}'} V_\epsilon(f(x_i))$$

If we take the minimum over these functions, we end up with a function $g = \min_{i \in \mathcal{I}'}(g_{x_i})$, which approximates f in an ϵ radius. For every $x \in X$:

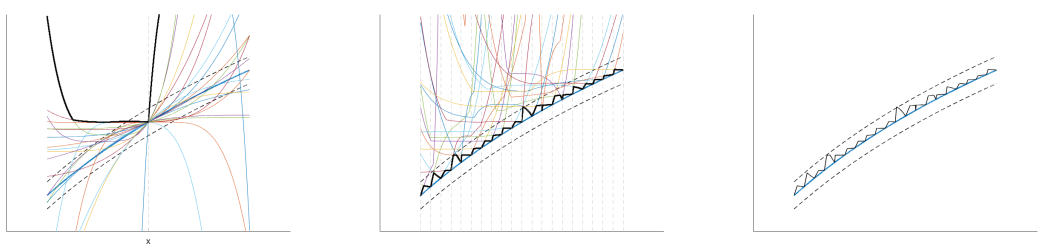
$$\begin{aligned} g(x) &> f(x) - \epsilon && \text{By construction of } g_x \\ g(x) - f(x) &< \epsilon && \text{As } g \text{ is the minimum of } g_x \\ |g(x) - f(x)| &< \epsilon \end{aligned}$$

We can observe this step in Figure 1b and 1e for two different values of ϵ. We already have many functions g_{x_i} that approximate f at the many points x_i since Lemma 4.2. The function g is obtained as the minimum of these functions, which gets closer to f for a given ϵ. $\square$

Figure 1: Approximation using Stone-Weierstrass Theorem



(a) $g_x = \max_{i \in \mathcal{I}}(h_i)$ (b) $g = \min_{i \in \mathcal{I}'}(g_{x_i})$ (c) Approximation using ϵ



(d) $g_x = \max_{i \in \mathcal{I}}(h_i)$ (e) $g = \min_{i \in \mathcal{I}'}(g_{x_i})$ (f) Approximation using $\epsilon' < \epsilon$

5 The Weierstrass theorem

This section will derive the Weierstrass theorem as a corollary of the Theorem 4.1. We will prove that the set $[0, 1]$ is compact and polynomials are an unital subalgebra with the separating points property.

Proposition 5.1. *The set [0,1] is compact.*

Proof. The interval $[0, 1]$ is complete. We prove it by contradiction. Any element of a sequence x_n is in $[0, 1]$. If the limit of the sequence is outside $[0, 1]$, that would imply that there exists an ϵ so small such that an $x_n > 1$, which contradicts the premise (same logic for limits below 0). Therefore, all limits are in $[0, 1]$, and the metric space is complete.

The interval $[0, 1]$ is totally bounded. We can take any ϵ, and build the following intervals:

$$(-\epsilon, \epsilon), (0, 2\epsilon), (\epsilon, 3\epsilon), ..., (n\epsilon, (n + 1)\epsilon)$$

Choose n to be the largest integer such that $n\epsilon < 1$ and $(n+1)\epsilon \geq 1$. We can easily verify that for every ϵ, this collection of open intervals is a cover for $[0, 1]$. Therefore, $[0, 1]$ is totally bounded. As this set is complete and totally bounded, it is compact. □

Proposition 5.2. *Polynomials are an unital subalgebras with the separating points property.*

Proof. Let $P(X)$ be the set of polynomials defined over a set $X \subset \mathbb{R}$. Polynomials are functions $p : \mathbb{R} \to \mathbb{R}$ that have the following specification: $p_a(x) = \sum_{i=0}^{n} a_i x^n$ where $a_i \in \mathbb{R}$.

Polynomials have the inner operations $+$ and $\times$ and are closed in those operations. They also have a scalar multiplication by any real number, which is also closed. They are an unital subalgebra, and $p(x) = 1$ is the multiplicative identity. They have the separating points property as there exist polynomials such that when evaluated at two different points, they give different values. □

Corollary 5.3. *Suppose f is a continuous real-valued function defined on the real interval $[0, 1]$. For every $\epsilon > 0$, there exists a polynomial $p(x)$ such that for all x in $[a, b]$, we have $|f(x) - p(x)| < \epsilon$.*

Proof. Proposition 5.1 shows $[0, 1]$ is compact, and Proposition 5.2 shows polynomials are an unital subalgebra with the separating points property. Therefore, by Theorem 4.1 polynomials are dense in the space of continuous functions defined in $[0, 1]$. This result implies that any continuous function defined in $[0, 1]$ can be approximated using polynomials. □

References

Johnsonbaugh, Richard and William E Pfaffenberger, *Foundations of mathematical analysis*, Courier Corporation, 2012.

Rudin, Walter et al., *Principles of mathematical analysis*, Vol. 3, McGraw-hill New York, 1964.