From Calculus to Analysis: The Stone-Weierstrass Theorem

Anthony Dokanchi

July 2024

1 Abstract

In an introductory calculus class, students learn about Taylor Series and Taylor Polynomials, which express certain functions as Power Series and closely approximate these functions, respectively. Taylor's Theorem provides an upper bound on the error between functions and their Taylor Polynomials, ensuring that the error decreases as the polynomial degree increases, approaching zero as the degree approaches infinity. However, the theorem only applies to functions with certain higher order derivatives. While most commonly encountered functions like $sin(x)$, $cos(x)$, and e^x are analytic, many real-world functions are not, rendering Taylor Series approximations ineffective. To address this, the Stone-Weierstrass Theorem offers an alternative, stating that any continuous function on a closed interval can be uniformly approximated by a sequence of polynomials. This paper aims to explain and prove the Stone-Weierstrass Theorem, providing necessary (and more) background in analysis principles.

The Stone-Weierstrass Theorem: For every continuous real function f defined on $[a, b]$, there exists a sequence of polynomials $\{P_n\}$ that converges uniformly to f on $[a, b]$.

2 Introduction

Most introductory calculus classes teach the concept of sequence convergence; that is, that there can be an infinitely long sequence of numbers that slowly approach a certain value. The sequence is said to converge to that value if the difference between the value and terms of an arbitrarily large index can become arbitrarily small. They are also taught about the convergence of sequences of functions in the context of Power and Taylor Series, although these concepts are usually not rigorously defined.

Function convergence is a powerful feature of math, as if a sequence of functions is shown to converge to another function, our understanding of the functions in the sequence can be used to inform our understanding of the function they converge to, and vice versa. Thus, the formal definitions and explanations of function convergence can lead to both interesting intuitions and useful discoveries.

This paper will assume a basic understanding of sequence convergence of numbers as would be taught in an introductory calculus course. From there, we will discuss formal definitions of function convergence by defining and examining pointwise convergence and uniform convergence. Finally, we will see the Stone-Weierstrass Theorem and its proof.

3 Pointwise Convergence

Imagine you would like to define what it would mean for a sequence of functions to converge. One reasonable attempt would be the following definition.

Definition 1. Suppose we have a sequence of functions $\{f_n\}$ where each f_i is defined on a set E. The sequence of functions is said to converge **pointwise** to a function f if

$$f(x) = \lim_{n \to \infty} f_n(x)$$

for all $x \in E$.

Pointwise convergence is the most basic type of function convergence, but as such, it has many limitations. For instance, one would hope that if each f_i in the sequence has a certain property (e.g. continuity, differentiability, integrability, etc.), then f might retain these properties. However, this is not the case. Below, two examples are shown of sequences of continuous functions that converge to a discontinuous function.

Example 2. Let

$$f_n(x) = \arctan(nx)$$

and

$$f(x) = \lim_{n \to \infty} f_n(x)$$

for all $x \in \mathbb{R}$. Each f_n is continuous, as it is a composition of arctan and nx, both continuous functions.

We have

$$f(x) = \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \arctan(nx) = \arctan(\lim_{n \to \infty} nx).$$

If $x > 0$, $\lim_{n \to \infty} nx = \infty$, so

$$\arctan(\lim_{n \to \infty} nx) = \arctan(\lim_{n \to \infty} n) = \frac{\pi}{2}.$$

If $x < 0$, $\lim_{n \to \infty} nx = -\infty$, so

$$\arctan(\lim_{n \to \infty} nx) = \arctan(\lim_{n \to -\infty} n) = -\frac{\pi}{2}.$$

If $x = 0$, then

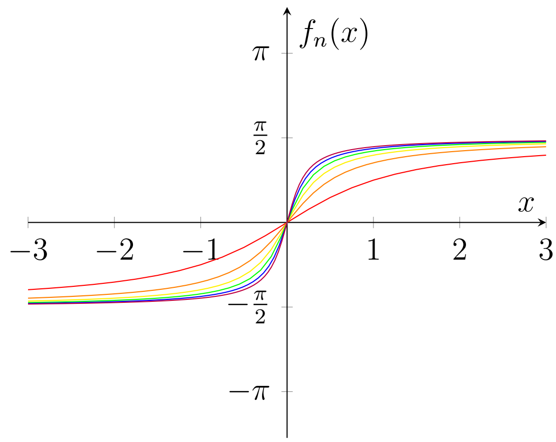
$$\arctan(\lim_{n\to\infty} nx) = \arctan(\lim_{n\to\infty} 0n) = \arctan(0) = 0.$$

Thus, we obtain

$$f(x) = \begin{cases} -\frac{\pi}{2}, & x < 0 \\ 0, & x = 0 \\ \frac{\pi}{2}, & x > 0 \end{cases}$$

which is discontinuous.

$f_n(x) = \arctan(nx)$



Example 3. Let

$$f_n(x) = x^n$$

and

$$f(x) = \lim_{n\to\infty} f_n(x).$$

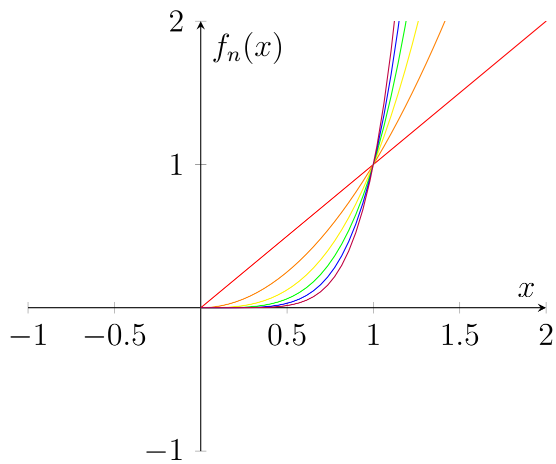
Each f_n is continuous and differentiable; for each choice of n, f_n is a polynomial. However, f is not continuous or differentiable. For all x on $[0, 1)$, $\lim_{n\to\infty} f_n(x) = 0$. However, at $f_n(1) = 1$ for all n. Thus, we have

$$f(x) = \begin{cases} 0, & 0 \le x < 1 \\ 1, & x = 1 \end{cases}$$

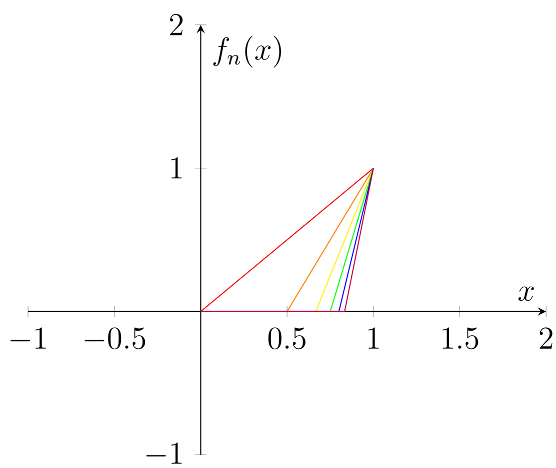
which is discontinuous.

The same line of reasoning applies to many other sequences of continuous functions $\{f_n\}$ such that for any $0 \leq x < 1$, $\lim_{n\to\infty} f_n(x) = 0$, and $f_n(1) = c \neq 0$ for all n (for example, $f_n(x) = \max(0, n(x-1)+1)$ or $f_n(x) = n^{x-1}$).

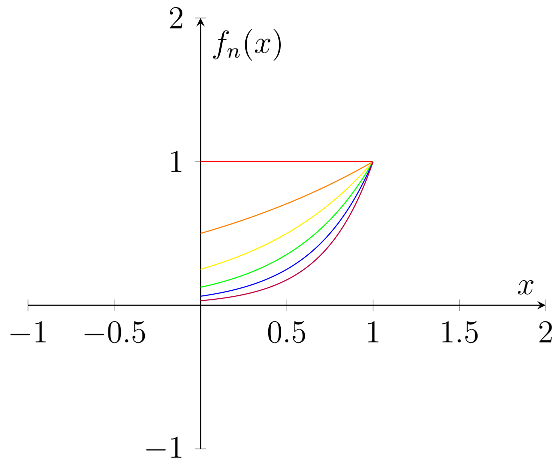
$f_n(x) = x^n$



$f_n(x) = \max((0, n(x-1)+1))$



$f_n(x) = n^{x-1}$



The previous examples demonstrate that a sequence of functions $\{f_n\}$ with certain properties can converge pointwise to a function f that does not have these properties. Above, continuity and differentiability were demonstrated, but it is also true for a number of other properties.

Thus, our problem becomes one of the properties of convergence: is there a stronger statement that can be made about the manner in which a sequence of functions converges that allows us to make stronger statements about what properties are conserved?

4 Uniform Convergence

Definition 4. A sequence of functions $\{f_n\}$ defined on a set E converges **uniformly** to a function f on E if for every $\varepsilon > 0$ there exists an integer N such that for all $n \geq N$, $x \in E$,

$$|f_n(x) - f(x)| < \varepsilon.$$

The definition is very similar to that of pointwise convergence; in fact, any function that converges uniformly also converges pointwise. The difference is that for a function to converge pointwise, for each $x \in E$ there must be a value of N (which may depend on x) such that $|f_n(x) - f(x)| \leq \varepsilon$ individually, whereas for it to converge uniformly, there must be a value of N such that $|f_n(x) - f(x)| \leq \varepsilon$ for all $x \in E$ simultaneously.

Uniform convergence is a much stronger property than pointwise convergence, so when a sequence of functions uniformly converges, we are able to make much more powerful statements about the properties of these functions.

Theorem 5 (Cauchy Criterion for Uniform Convergence)**.** *If (and only if) a sequence of functions $\{f_n\}$ converges on E, then for all $\varepsilon > 0$, there is some integer N such that $m \geq N$, $n \geq N$, $x \in E$ implies*

$$|f_n(x) - f_m(x) \leq \varepsilon.$$

Proof. Say $\{f_n\}$ converges to f. There must be some N such that for any $n > N$, $x \in E$ implies

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{2}.$$

This means that if $m \geq N, n \geq N, x \in E$,

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

To prove the converse, we will take without proof that $\{f_n\}$ converges to f pointwise and will prove that this convergence is uniform. Take any $\varepsilon > 0$, and choose N such that $|f_n(x) - f_m(x)| \leq \varepsilon$. Letting $m \to \infty$, we get

$$\lim_{m\to\infty} |f_n(x) - f_m(x)| = |f_n(x) - \lim_{m\to\infty} f_m(x)| = |f_n(x) - f(x)| \leq \varepsilon$$

proving the theorem.

QED

Theorem 6. *If $\{f_n\}$ is a sequence of functions that converges uniformly to f on a set E, then*

$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t)$$

Proof. We set

$$\lim_{t\to x} f_n(t) = A_n.$$

Our goal is now to prove that

$$\lim_{t\to x} f(t) = \lim_{n\to\infty} A_n.$$

Say we have some $\varepsilon > 0$. Since $\{f_n\}$ converges uniformly, there is a value N such that for all inputs in the domain of f, $n \geq N, m \geq N$ implies

$$|f_n(t) - f_m(t)| \leq \varepsilon.$$

Letting $t \to x$, we get

$$\lim_{t\to x} |A_n(t) - A_m(t)| \leq \varepsilon$$

$$\left|\lim_{t\to x} f_n(t) - \lim_{t\to x} f_m(t)\right| \leq \varepsilon.$$

$$|A_n(t) - A_m(t)| \leq \varepsilon.$$

By the Cauchy Criterion, A converges to some value. Let this value be A.

Since f uniformly converges and A converges, we can choose some n such that

$$|f(t) - f_n(t)| \leq \frac{\varepsilon}{3}$$

for all $t \in E$ and

$$|A_n - A| \leq \frac{\varepsilon}{3}.$$

There is some neighborhood V such that if $t \in V \cap E$ and $t \neq x$,

$$|f_n(t) - A_n| \leq \frac{\varepsilon}{3}.$$

Finally, since $|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|$, we can substitute from the last three equalities to obtain

$$|f(t) - A| \leq \varepsilon.$$

The above equality holds for any $\epsilon > 0$ and $t \in V \cap E$ and $t \neq x$, so it must hold as $t \to x$. Thus, we have

$$\lim_{t \to x} f(t) = A = \lim_{n \to \infty} A_n$$

proving the theorem.

QED

The previous theorem shows that if we have $f_n \to f$ uniformly, then if each f_n is continuous, f must be continuous.

The following are properties that are true of uniformly convergent sequences that will not be proven in this paper.

Theorem 7. *If a sequence of differentiable functions $\{f_n\}$ converges uniformly to a function f, and the sequence of derivatives $\{f'_n\}$ converges uniformly to a fucntion g, then*

$$f'(x) = \lim_{n \to \infty} f'_n(x) = g(x)$$

which is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}x} \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \frac{\mathrm{d}}{\mathrm{d}x} f_n(x) = g(x)$$

Theorem 8. *If a sequence of integrable functions $\{f_n\}$ converges uniformly to a function f, then*

$$\int_a^b f(x) = \lim_{n \to \infty} \int_a^b f_n(x)$$

which is equivalent to

$$\int_a^b \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \int_a^b f_n(x)$$

5 The Stone-Weierstrass Theorem

Lemma 9. *For all real $x \in [0,1]$ and $n \geq 1$, $(1-x^2)^n \geq 1 - nx^2$.*

Proof. Consider the function $f(x) = (1-x^2)^n - 1 + nx^2$. We have $f(0) = 0$ and $f'(x) = 2nx(1-(1-x^2)^{n-1})$.

For $x \in [0,1]$ and $n \geq 1$, $0 \leq 1-x^2 \leq 1$, so $0 \leq 1-(1-x^2)^{n-1} \leq 1$. Also, for nonnegative x and n, $2nx \geq 0$. Thus, $f'(x) = 2nx(1-(1-x^2)^{n-1}) \geq 0$ for all $x \in [0,1]$ and $n \geq 1$.

Thus, for $x \in [0,1]$, $f(x)$ is constant or increasing, so $f(x) \geq f(0)$. Substituting, we get $(1-x^2)^n - 1 + nx^2 \geq 0$, implying $(1-x^2)^n \geq 1 - nx^2$. QED

Theorem 10 (The Stone-Weierstrass Theorem)**.** *For every continuous real function f defined on $[a,b]$, there exists a sequence of polynomials $\{P_n\}$ that converges uniformly to f on $[a,b]$.*

Proof. It suffices to prove the theorem for functions on the interval $[0,1]$. If we wish to apply the theorem for a function f on $[a,b]$, consider

$$g(x) = f((b-a)x + a)$$

which exists on $[0,1]$. Suppose there is a sequence of polynomials P_n that converges uniformly to g. In that case, there also must be a sequence of polynomials that converges uniformly to f, since $f(x) = g(\frac{1}{b-a}x - \frac{a}{b-a})$, and if we have $|P_n(x) - g(x)| \leq \varepsilon$ for $x \in [0,1]$, then

$$|P_n(\frac{1}{b-a}x - \frac{a}{b-a}) - g(\frac{1}{b-a}x - \frac{a}{b-a})| = |P_n(\frac{1}{b-a}x - \frac{a}{b-a}) - f(x)| < \varepsilon$$

for $x \in [a,b]$. Outside the interval $[0,1]$, we can define our function to have the value 0 for all inputs.

Additionally, it suffices to prove the theorem for functions with roots at 0 and 1. If we wish to apply the theorem for a function f where this does not hold, consider $g(x) = f(x) - f(0) + x(f(1) - f(0))$. It is easily seen that $g(0) = g(1) = 0$. Suppose there is a sequence of polynomials P_n that converges uniformly to g. In that case, there also must be a sequence of polynomials that converges uniformly to f, as $f(x) = g(x) + x(f(1) - f(0)) + f(0)$, which is the sum of g and a degree 1 polynomial.

We can define a sequence of functions Q_n such that

$$Q_n(x) = c_n(1 - x^2)^n$$

where c_n is a constant chosen to ensure

$$\int_{-1}^{1} Q_n(x)dx = 1.$$

From this and from Lemma 9, we obtain

$$1 = \int_{-1}^{1} Q_n(x)dx \tag{1}$$
$$= c_n \int_{-1}^{1} (1 - x^2)^n dx \tag{2}$$
$$= 2c_n \int_{0}^{1} (1 - x^2)^n dx \tag{3}$$
$$\geq 2c_n \int_{0}^{1/\sqrt{n}} (1 - x^2)^n dx \tag{4}$$
$$\geq 2c_n \int_{0}^{1/\sqrt{n}} (1 - nx^2) dx \tag{5}$$
$$= \frac{4c_n}{3\sqrt{n}} \tag{6}$$
$$> \frac{c_n}{\sqrt{n}}. \tag{7}$$

Thus, we have $\frac{c_n}{n} < 1$, so $c_n < \sqrt{n}$.
It follows that for $\delta \leq |x| \leq 1$, $Q_n(x) \leq \sqrt{n}(1 - \delta^2)^n$.
Now we define a sequence of functions P_n such that

$$P_n(x) = \int_{-1}^{1} f(x + t)Q_n(t)dt.$$

Our goal is now to show that P_n is a sequence of polynomials that converges to f uniformly.

Since earlier we set $f(x) = 0$ for all x not in $[0, 1]$, we can reduce the bounds of the integral to $[0 - x, 1 - x] = [-x, 1 - x]$. Then, by a change of variable t to $t - x$, we have

$$P_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)dt = \int_0^1 f(t)Q_n(t-x)dt = \int_0^1 f(t)c_n(1-(t-x)^2)^n dt.$$

This last integral (and thus P_n) is a polynomial in x, as $(1-(t-x)^2)^n$ can be expanded using the binomial theorem, and all appearances of t will become constants under the definite integral.

For $\varepsilon > 0$ we take $\delta > 0$ such that $|y-x| < \delta$ implies $|f(y) - f(x)| < \frac{\varepsilon}{2}$. Setting $M = sup|f(x)|$, on $0 \leq x \leq 1$ we have

$$\begin{aligned}
&|P_n(x) - f(x)| \\
&= \left| \int_{-1}^{1} f(x+t)Q_n(t) - f(x) \int_{-1}^{1} Q_n(t) \right| \\
&= \left| \int_{-1}^{1} (f(x+t) - f(x))Q_n(t)dt \right| \\
&\leq \int_{-1}^{1} |(f(x+t) - f(x))Q_n(t)|\, dt \\
&\leq 2M \int_{-1}^{-\delta} Q_n(t)dt + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} Q_n(t)dt + 2M \int_{\delta}^{1} Q_n(t)dt \\
&\leq 4M\sqrt{n}(1-\delta^2)^n + \frac{\varepsilon}{2} \\
&< \varepsilon
\end{aligned}$$

for n of sufficient size. We have that for all large enough n and for all $x \in [0,1]$, $|P_n(x) - f(x)| < \varepsilon$, proving that P_n converges uniformly to f and thus proving the theorem. QED

Acknowledgments

I would like to thank Sawyer Dobson for helpful discussions around the content and format of the paper.

References

[1] Walter Rudin (1953) *Principles of Mathematical Analysis*, McGraw-Hill, Inc.