

The Weierstrass Approximation Theorem states that a function in $C[a, b]$ can be uniformly approximated by a polynomial. One way of expressing this fact is that given $f \in C[a, b]$ and $\epsilon > 0$, there exists $p \in P[a, b]$ such that $|f(x) - p(x)| \leq \epsilon$ for every $x \in [a, b]$. Using the vocabulary of norms, this is equivalent to

$$\|f - p\|_\infty \leq \epsilon.$$

The same idea can also be expressed in terms of the closure of $P[a, b]$ in $C[a, b]$. Recall that given a set A in a metric space M , $x \in \bar{A}$ if and only if for every $\epsilon > 0$, $B_\epsilon(x) \cap A \neq \emptyset$. Hence the Weierstrass Approximation Theorem asserts that $C[a, b] \subseteq \overline{P[a, b]}$. But of course $\overline{P[a, b]} \subseteq C[a, b]$. Hence we have arrived at a concise statement of the theorem.

Theorem 1. (Weierstrass Approximation Theorem) $\overline{P[a, b]} = C[a, b]$, where closure is taken with respect to the uniform norm.

You are already familiar with the idea of writing certain functions as power series. For example,

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^{n+1} x^{2n+1}}{n!}.$$

This series converges pointwise on all of \mathbb{R} (verify this with the ratio test) and therefore uniformly on any fixed interval $[-R, R]$. (Recall Theorem 10.10) Hence, given any $\epsilon > 0$, we can find an N such that

$$\left| \sin(x) - \sum_{n=0}^N \frac{(-1)^{n+1} x^{2n+1}}{n!} \right| \leq \epsilon$$

for every $x \in [-\pi, \pi]$. So \sin can be approximated uniformly by polynomials on $[-\pi, \pi]$. But functions that can be written as power series are special; in particular they are infinitely differentiable – this is a consequence of Theorem 10.10.

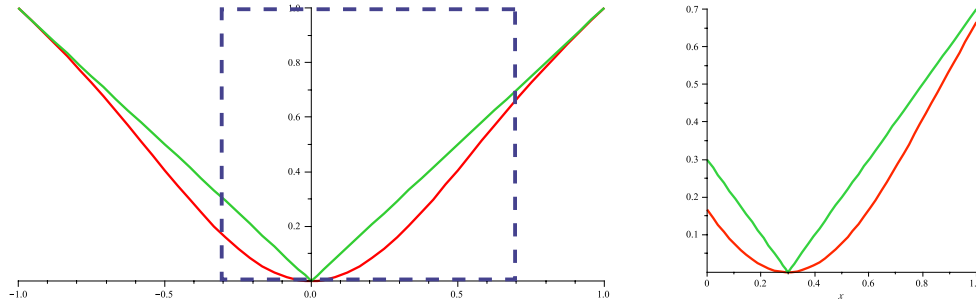
The remarkable part about the Weierstrass Approximation Theorem is that every continuous function, even the non-differentiable ones, can be uniformly approximated by polynomials. Interestingly, the proof of this fact can be reduced to showing that just one non-smooth function, the absolute value function abs , can be uniformly approximated by polynomials.

Proposition 2. $\text{abs} \in \overline{P[-1, 1]}$.

Supposing for the moment we have proved this result, let's see how this results in a fairly easy proof of Theorem 1. Some of this work is made easy because the metric space and the vector space structures of a normed vector space are compatible.

Proposition 3. Let X be a normed linear space and let W be a subspace of X . Then \overline{W} is a subspace of X .

Proof. Let $x, y \in \overline{W}$. Let (x_n) and (y_n) be sequences in W converging to x and y . Then $\|(x + y) - (x_n + y_n)\| \leq \|x - x_n\| + \|y - y_n\|$ and therefore $(x_n + y_n) \rightarrow (x + y)$. Hence $x + y \in \overline{W}$. Similarly, $\alpha x_n \rightarrow \alpha x$ and hence $\alpha x \in \overline{W}$. So \overline{W} is a subspace. \square

Approximating abs_a on $[0, 1]$.

We now show that $P[0, 1]$ is dense in $C[0, 1]$ by a series of steps. First, we show that any translate of the absolute value function is in $\overline{P[0, 1]}$. We define

$$\text{abs}_a(x) = |x - a|.$$

Now if $a \geq 1$ or $a \leq 0$, then abs_a restricted to $[0, 1]$ is linear and therefore certainly in $\overline{P[0, 1]}$. So we need only consider the cases where $0 < a < 1$. By translating the domain, approximating abs_a on $[0, 1]$ by a polynomial is equivalent to approximating the standard absolute value function by a polynomial on the translated domain $[-a, 1 - a]$. Since $0 < a < 1$, we know that $[-a, 1 - a] \subseteq [-1, 1]$, so Proposition 2 can be applied to produce the approximating polynomial.

Lemma 4. For any $a \in \mathbb{R}$, $\text{abs}_a \in \overline{P[0, 1]}$.

Proof. If $a \leq 0$ or $a \geq 1$, abs_a is linear on $[0, 1]$ and hence in $P[0, 1]$.

Suppose $0 < a < 1$ and let $\epsilon > 0$. Let p be a polynomial such that

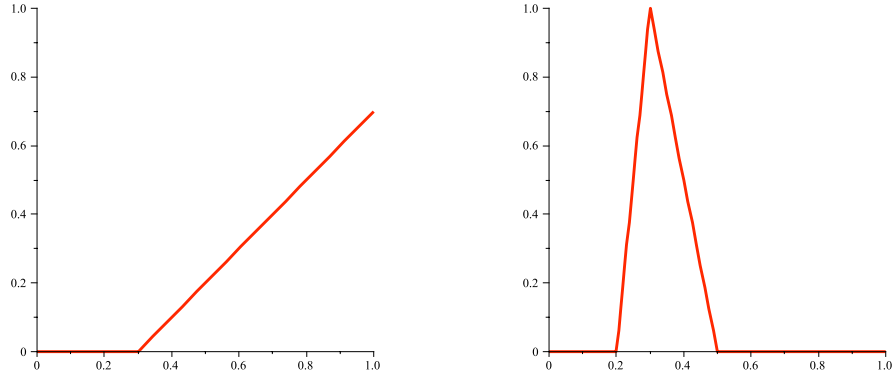
$$|p(x) - \text{abs}(x)| < \epsilon$$

for every $x \in [-1, 1]$. Define $q(x) = p(x - a)$, so q is a polynomial. Then

$$\begin{aligned} \sup_{x \in [0, 1]} |q(x) - \text{abs}_a(x)| &= \sup_{x \in [0, 1]} |p(x - a) - |x - a|| \\ &= \sup_{x \in [-a, 1 - a]} |p(x) - |x|| \\ &\leq \sup_{x \in [-1, 1]} |p(x) - \text{abs}(x)| \leq \epsilon. \end{aligned}$$

Hence $\|q - \text{abs}_a\|_{C[0, 1]} \leq \epsilon$. Since q is a polynomial and $\epsilon > 0$ is arbitrary, $\text{abs}_a \in \overline{P[0, 1]}$. \square

Combining Lemma 4 and Proposition 3, we know that any linear combination of the functions abs_a belongs to $\overline{P[0, 1]}$. A function $f \in C[0, 1]$ is called piecewise linear if there is a partition $0 = x_0 < x_1 < \dots < x_n = 1$ such that the restriction of f to each interval $[x_{k-1}, x_k]$ is linear; we denote by $\text{PL}[0, 1]$ the collection of all such functions. Clearly any linear combination of functions of the form abs_a belongs to $\text{PL}[0, 1]$. In fact, by taking linear combinations of the functions abs_a we obtain nearly all of $\text{PL}[0, 1]$; the only missing functions are the constants.

Functions R_k and H_k

Proposition 5. Let $f \in \text{PL}[0, 1]$, and let $0 = x_0 < x_1 < \dots < x_n = 1$ be a partition such that f is linear on each interval $I_k = [x_{k-1}, x_k]$. Then f is a linear combination of the functions 1 and $\{\text{abs}_{x_k} : 1 \leq k \leq n\}$.

Proof. For each k , define

$$R_k(x) = \frac{1}{2} (\text{abs}_{x_k}(x) + (x - x_k)).$$

Notice that $R_k(x) = 0$ if $x \leq x_k$, and R_k is linear with slope 1 for $x \geq x_k$. Each R_k is a linear combination of abs_{x_k} , 1 , and abs_{x_n} , so it is enough to show that f is a linear combination of the functions R_k .

For $0 < k < n$, define

$$H_k = \frac{1}{2} \left(\frac{R_{k-1} - R_k}{x_k - x_{k-1}} + \frac{R_{k+1} - R_k}{x_{k+1} - x_k} \right);$$

for $k = 0$ and $k = n$ define

$$H_0 = \frac{1}{2} \left(\frac{R_1 - R_0}{x_1 - x_0} \right) \quad \text{and} \quad H_n = \frac{1}{2} \left(\frac{R_{n-1} - R_n}{x_n - x_{n-1}} \right).$$

It is easy to verify that

$$H_k(x_j) = \begin{cases} 1 & k = j \\ 0 & k \neq j. \end{cases}$$

Hence

$$\sum_{k=0}^n f(x_k) H_k$$

is a piecewise linear function that agrees with f at each point x_k . We conclude that

$$f = \sum_{k=0}^n f(x_k) H_k.$$

Since each H_k is a linear combination of the functions R_j , we are done. \square

Along with a couple of facts proved in class (and in Carothers) we can now prove the Weierstrass Approximation Theorem, at least for the domain $[0, 1]$.

Proposition 6. $C[0, 1] = \overline{P[0, 1]}$.

Proof. From Lemma 4 we know that each function abs_a belongs to $\overline{P[0, 1]}$. Hence from Proposition 3 and Proposition 5 we know that $\text{PL}[0, 1] \subseteq \overline{P[0, 1]}$. Hence $\overline{\text{PL}[0, 1]} \subseteq \overline{P[0, 1]}$. From the proof of Carothers 11.2 it follows that $C[0, 1] = \overline{\text{PL}[0, 1]}$. Hence $\overline{P[0, 1]} = C[0, 1]$. \square

Exercise 1: Use Proposition 6 to prove the Weierstrass Approximation Theorem for an arbitrary interval $[a, b]$. Hint: Given $f \in C[a, b]$, define $g(x) = f(a + x(b - a))$. Approximate g in $C[0, 1]$ by $p \in P[0, 1]$, and define $q(x) = p((x - a)/(b - a))$.

It remains to prove Proposition 2, which we do now.

Proof. For $0 \leq x \leq 1$, define $P_0(x) = 0$ and for $k \geq 0$ define

$$P_{k+1}(x) = P_k(x) + \frac{x - P_k^2}{2}.$$

We claim that $0 \leq P_k(x) \leq \sqrt{x}$ for every $k \geq 0$ and that $P_{k+1} \geq P_k$ for every k . This is certainly true for $k = 0$. Suppose $0 \leq P_k(x) \leq \sqrt{x}$. Then

$$P_{k+1}(x) = P_k(x) + \frac{x - P_k^2}{2} \geq P_k(x)$$

so $P_{k+1}(x) \geq 0$. But also, since $0 \leq P_k(x) \leq \sqrt{x} \leq 1$, we have

$$P_{k+1}(x) = P_k(x) + \frac{x - P_k^2}{2} = P_k(x) + \frac{1}{2}(\sqrt{x} + P_k(x))(\sqrt{x} - P_k(x)) \leq P_k(x) + (\sqrt{x} - P_k(x)) = \sqrt{x}.$$

Hence $P_{k+1}(x) \leq \sqrt{x}$. We have therefore shown inductively that $0 \leq P_k(x) \leq \sqrt{x}$ for every $k \geq 0$. As seen above, this also implies that $P_{k+1} \geq P_k(x)$.

It follows that for any fixed $x \in [0, 1]$, $P_k(x)$ is monotone increasing and converges to a limit $P(x)$. But then $P(x)$ satisfies

$$P(x) = P(x) + \frac{x - P(x)^2}{2}$$

and hence

$$P(x)^2 = x.$$

Since $P(x) \geq 0$, we conclude that $P(x) = \sqrt{x}$ and P_k converges pointwise to the square root function. Since the convergence is monotone and the limit function is continuous, Dini's theorem implies that the convergence is actually uniform.

Now let $\epsilon > 0$. Pick k so that $|P_k(x) - \sqrt{x}| < \epsilon$ for all $x \in [0, 1]$. Define $q(y) = P_k(y^2)$ for $y \in [-1, 1]$, so q is a polynomial. Then for any $y \in [-1, 1]$,

$$|q(y) - \text{abs } y| = |P_k(y^2) - \sqrt{y^2}| < \epsilon$$

since $y^2 \in [0, 1]$. Since $\epsilon > 0$ is arbitrary, we conclude that $\text{abs} \in P[0, 1]$.