LECTURE 5: WEIERSTRASS APPROXIMATION

1. The Weierstrass Approximation Theorem

The following result is really useful in Applied Mathematics. It says that any continuous function can be approximated by polynomials:

Theorem: [Weierstraß Approximation Theorem]

If f is continuous on $[a, b]$, then there is a sequence of polynomials P_n such that $P_n \to f$ uniformly.

In practice f is a complicated function, but the P_n are much easier.

Note: This result is **false** on all of $\mathbb{R}$ or on open intervals (a, b)

Proof: STEP 1: Some simplifications:

- Assume $[a, b] = [0, 1]$ (just for sake of notation)
- WLOG, assume $f(0) = f(1) = 0$ because otherwise apply the result to

$$g(x) = f(x) - f(0) - x\,(f(1) - f(0))$$

  Then $g(0) = g(1) = 0$ and if g can be approximated uniformly, then so can f

Date: Monday, July 11, 2022.

- Finally, define f to be zero outside $[0,1]$, so f is uniformly continuous on $\mathbb{R}$

Main Idea: We want to define P_n via an integral:

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

Where Q_n is a helper polynomial defined below.

Note: P_n is indeed a polynomial. If you use the $u-$sub $u = x+t$ then

$$P_n(x) = \int_{-1+x}^{1+x} f(u)Q_n(u-x)du = \int_0^1 f(u)Q_n(u-x)du$$

(The last step is because f is 0 outside $[0,1]$ and $0 \leq x \leq 1$)

And this is a polynomial in x because Q_n is a polynomial. For example, if $Q_n(u-x) = u-x$ then this is $\left(\int_0^1 f(u)udu\right) - x\left(\int_0^1 f(u)du\right)$

STEP 2: Our Helper Function:

$$\text{Let } \; Q_n(x) = c_n\left(1-x^2\right)^n \text{ where } c_n = \frac{1}{\int_{-1}^{1}\left(1-x^2\right)^n dx}$$

This choice of c_n makes $\int_{-1}^{1} Q_n = 1$

Claim: $c_n < \sqrt{n}$

Note: $(1-x^2)^n \geq 1 - nx^2$ on $[0,1]$ This follows from Calculus I by taking the difference and taking the derivative. Therefore

$$\int_{-1}^{1} \left(1-x^2\right)^n dx \overset{\text{Even}}{=} 2\int_{0}^{1} \left(1-x^2\right)^n dx \geq 2\int_{0}^{\frac{1}{\sqrt{n}}} \left(1-x^2\right)^n dx \overset{\text{Note}}{\geq} 2\int_{0}^{\frac{1}{\sqrt{n}}} 1-nx^2 dx$$
$$=2\left(\frac{2}{3\sqrt{n}}\right) > \frac{1}{\sqrt{n}}$$

Since that integral is $\frac{1}{c_n}$ by definition, we get $c_n < \sqrt{n}$ ✓

STEP 3: Main Proof:

Let $\epsilon > 0$ be given. Since f is uniformly continuous, there is $\delta > 0$ such that if $|x-y| < \delta$ then

$$|f(x) - f(y)| < \frac{\epsilon}{2}$$

Let $M = \sup_x |f(x)|$, then

$$|P_n(x) - f(x)| = \left| \int_{-1}^{1} f(x+t)Q_n(t)dt - f(x) \overbrace{\int_{-1}^{1} Q_n(t)dt}^{1} \right|$$
$$\leq \int_{-1}^{1} |f(x+t) - f(x)| \, Q_n(t)dt$$
$$= \left(\int_{-1}^{-\delta} + \int_{-\delta}^{\delta} + \int_{\delta}^{1} \right) |f(x+t) - f(x)| \, Q_n(t)dt$$

Study of the last term:

Notice that from $Q_n = c_n \left(1-x^2\right)^n$ and $c_n \leq \sqrt{n}$, we get that on $[\delta, 1]$, $Q_n \leq \sqrt{n}\left(1-\delta^2\right)^n$ and hence

$$\begin{aligned}\int_{\delta}^{1} |f(x+t) - f(x)| \, Q_n(t) dt \leq & \int_{\delta}^{1} \left(\underbrace{|f(x+t)|}_{\leq M} + \underbrace{|f(x)|}_{\leq M} \right) Q_n(t) \\ \leq & 2M \int_{\delta}^{1} \sqrt{n} \left(1 - \delta^2\right)^n dt \\ \leq & 2M\sqrt{n} \left(1 - \delta^2\right)^n \underbrace{(1-\delta)}_{\leq 1} \\ \leq & 2M\sqrt{n} \left(1 - \delta^2\right)^n\end{aligned}$$

The same estimate holds for the first term.

Study of the middle term:

Notice that if $t \in [-\delta, \delta]$, then $|(x+t) - x| = |t| \leq \delta$ and so by uniform continuity we have $|f(x+t) - f(x)| < \frac{\epsilon}{2}$ so

$$\int_{-\delta}^{\delta} \underbrace{|f(x+t) - f(x)|}_{< \frac{\epsilon}{2}} Q_n(t) dt \leq \frac{\epsilon}{2} \underbrace{\int_{-\delta}^{\delta} Q_n}_{\leq 1} \leq \frac{\epsilon}{2}$$

Therefore, putting everything together, we get

$$\begin{aligned}|P_n(x) - f(x)| \leq \left(2M\sqrt{n} \left(1 - \delta^2\right)^n\right) + \left(\frac{\epsilon}{2}\right) + \left(2M\sqrt{n} \left(1 - \delta^2\right)^n\right) \\ = 4M\sqrt{n} \left(1 - \delta^2\right)^2 + \frac{\epsilon}{2}\end{aligned}$$

Since $\lim_{n \to \infty} \sqrt{n} \left(1 - \delta^2\right)^n = 0$ (because $1 - \delta^2 < 1$) there is N so that if $n > N$ then the first term is $< \frac{\epsilon}{2}$, which makes $|P_n(x) - f(x)| < \epsilon$ $\square$

Note: There is also a way of constructing P_n explicitly, without integrals, by sampling f at as many points as possible. See the proof of Theorem 18 in Chapter 4 of Pugh's book if you're interested. That proof uses Bernstein polynomials.

2. The Stone-Weierstrass Theorem

There is a nice generalization of this to more general function spaces.

Definition: A set $\mathcal{A}$ of functions is called an **algebra** if:

(1) $f, g \in \mathcal{A} \Rightarrow f + g \in \mathcal{A}$

(2) $f \in \mathcal{A}$ and c is a constant $\Rightarrow cf \in \mathcal{A}$

(3) $f, g \in \mathcal{A} \Rightarrow fg \in \mathcal{A}$

(like a vector space, but also closed under multiplication)

Definition: $\mathcal{A}$ **separates points** if for every $x \neq y$ there is $f \in \mathcal{A}$ such that $f(x) \neq f(y)$

That is, there are so many functions in $\mathcal{A}$ that we can distinguish x and y with functions. The negation of this is that there are $x \neq y$ such that for all f, we have $f(x) = f(y)$, it wouldn't be possible to distinguish x and y using functions

Definition: $\mathcal{A}$ **vanishes at no point** if for every x there is $f \in \mathcal{A}$ such that $f(x) \neq 0$

The negation is there is x such that for all f we have $f(x) = 0$, so all the functions would be 0 at that point.

Theorem: [Stone-Weierstraß Theorem] If K is a compact set and $\mathcal{A}$ is an algebra of real continuous functions on K that separates points and vanishes nowhere, then $\mathcal{A}$ is dense in $C(K)$ (continuous functions on K with sup norm)

So given $f \in C(K)$ there is a sequence $f_n \in \mathcal{A}$ such that $f_n \to f$ uniformly.

Example: $\mathcal{A}$ = polynomials in $C[a,b]$, which is the Weierstraß Approximation Theorem above

Example: $\mathcal{A}$ = trigonometric polynomials in $C[0, 2\pi]$, that is polynomials of the form

$$\sum_{k=0}^{n} a_k \cos(kx) + \sum_{k=0}^{n} a_k \sin(kx)$$

Which is the perfect transition to the next chapter, which is about power series and Fourier series ☺

Note: The same thing also works for complex functions provided you also require that if $f \in \mathcal{A}$ then $\overline{f} \in \mathcal{A}$ where $\overline{f}(x) = \overline{f(x)}$ for all x.

3. Power series

Recap: A **power series** is a series of the form

$$\sum_{n=0}^{\infty} c_n x^n$$

It's a special type of series of functions, the functions being power functions

Recall: [Radius of convergence]

$$\text{If } R =: \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}}, \text{ then}$$

(1) If $x \in (-R, R)$, the series converges and defines a function

$$f(x) = \sum_{n=0}^{\infty} c_n x^n$$

In that case, we say that f has a **power series expansion** around $x = 0$ or that f is **analytic**

(2) If $x \notin [-R, R]$, the series diverges.

What we want to show now is that more is true strictly inside $(-R, R)$

Theorem: If $r < R$, the power series converges **uniformly** on $[-r, r]$

Proof:[1] Yet another application of the Weierstraß M-test. Let s be such that $r < s < R$.

[1] The proof is taken from Theorem 11 in Chapter 4 of Pugh's book

Notice $\limsup_{n\to\infty} \sqrt[n]{|c_n|} = \frac{1}{R} < \frac{1}{s}$ so by definition of lim sup, we have for all large n, $\sqrt[n]{|c_n|} < \frac{1}{s}$.

Hence, if $|x| \leq r$ then

$$|c_n x^n| \leq \left(\frac{r}{s}\right)^n =: M_n$$

But since $\sum_n \left(\frac{r}{s}\right)^n$ converges (Geometric series), by the Weierstraß M test, the series $\sum c_n x^n$ converges uniformly when $x \in [-r, r]$ □

Note: It follows from uniform convergence of a series that f is continuous on $(-R, R)$

Theorem: [Integration and Differentiation] If $f(x) = \sum_{n=0}^{\infty} c_n x^n$ and $|x| < R$, then

$$\int_0^x f(t)dt = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1} = \sum_{n=1}^{\infty} \frac{c_{n-1}}{n} x^n$$
$$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1}$$

Proof:[2] Term-by-term integration is valid on $[-r, r]$ for every $r < R$ by uniform convergence. We just need to check that the radius of convergence is the same, but

$$\limsup_{n\to\infty} \sqrt[n]{\left|\frac{c_{n-1}}{n}\right|} = \limsup_{n\to\infty} \left(|c_{n-1}|^{\frac{1}{n-1}}\right)^{\frac{n-1}{n}} \left(\frac{1}{n}\right)^{\frac{1}{n}} = \left(\frac{1}{R}\right)^1 (1) = \frac{1}{R}$$

So indeed the radius of convergence is the same.

[2] The proof is taken from Theorem 12 in Chapter 4 of Pugh's book

For the derivative, a similar calculation shows that the radius of convergence is the same. Since the derivative series converges uniformly on every $[-r, r]$, the original series can be differentiated term by term $\square$

The result above shows that power series $f(x) = \sum c_n x^n$ is infinitely differentiable, since the series can be differentiated as many times as we want inside $(-R, R)$