Weierstrass approximation theorem (A very Striking result)

Q: Suppose $f \in C[a,b]$ (we will consider $[a,b] = [0,1]$: loose no generality at all). Can we "approximate" f by a polynomial $p \in \mathbb{R}[x]$?

Classification/ Ans/ issues

Here "approximate" means uniform metric $(C[a,b], d_{sup})$:

i.e: "Given $\varepsilon > 0$ $\exists$ $p \in \mathbb{R}[x]$ s.t.

$$\|f - p\|_\infty < \varepsilon$$

i.e. $\sup_{x \in [0,1]} |f(x) - p(x)| < \varepsilon$.

$\iff$ Given $f \in C[a,b]$ $\exists$ $\{p_n\} \subseteq \mathbb{R}[x]$ $\ni$ $p_n \xrightarrow{u} f$!

The answer is yes: By 1) Weierstrass (1885). & then also by

2) Bernstein (1911) ← For us.

3) Fejér (1900) ← perhaps more effective: it comes from Fourier series point of view.

4) Stone (1937): More powerful result: replaces $C[0,1]$ by $C(X)$, X compact metric space.

\# Suppose in addition, f is C^∞ - fn. (or C^k fn.). We can appeal to Taylor's polynomial (or even power series) approach. But it is fairly weak approximation.

Notably: i) Taylor approximation is (super) limited to points near a given point, ii) For n-degree poly. approximation, we must know/play with bound of $(n+1)$-th derivative, & iii) finally what worse, $\exists$ $f \in C^\infty(\mathbb{R})$ [namely: $f(x) = e^{-1/x^2}$ if $x \neq 0$ & $f(0) = 0$]

S.t. $f^{(n)}(0) = 0 \quad \forall\ n \geq 0, 1, \dots$

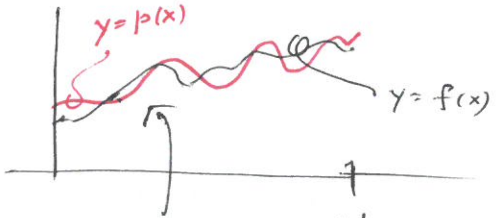
i.e: Taylor's (or power series) approach could be completely misleading !!

— okay —. So:

Thm: (Weierstass approximation thm).

Let $f \in C[0,1]$. Then $\exists\ \{p_n\} \subseteq \mathbb{R}[x] \ni p_n \xrightarrow{\text{unif.}} f$. ($\Leftrightarrow$ if $\varepsilon > 0$ then $\exists\ p \in \mathbb{R}[x] \ni \|f - p\| < \varepsilon$.)

Idea? Introduce "bump" fn/polynomials !!



do it so that the poly p remains inside the "ε-band", i.e: $f(x) - \varepsilon < p(x) < f(x) + \varepsilon \quad \forall\ x \in [0,1]$.

Okay: let's do it (through Bernstein).

Let $n \in \mathbb{N}$. We know

$$\sum_{k=0}^{n} \underbrace{\binom{n}{k} x^k (1-x)^{n-k}}_{:=\ b_k^n} = 1.$$

Binomial formula: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$.

$a \longmapsto x$

$b \longmapsto 1-x$

Def: $\boxed{b_k^n(x) := \binom{n}{k} x^k (1-x)^{n-k}}$, $0 \leq k \leq n$, $n \in \mathbb{N}$.

Called "Bernstein polynomial".

Remark:

1) b_k^n yields the necessary "bump": See through mathematica or Wikipedia picture.

2) $\forall\ n \in \mathbb{N}$ & $0 \leq k \leq n$, b_k^n has a ! maxima at $\boxed{x = \frac{k}{n}}$. [See the pic. again.]

We will use this.

3) $\sum_{k=0}^{n} b_k^n \equiv 1 \quad \forall\ n \in \mathbb{N}$.

4) $\deg b_k^n = n \quad \forall\ 0 \leq k \leq n$.

5) $b_k^n(x) \geq 0 \quad \forall\ x \in [0,1]$.

6) $b^n_k(1-x) = b^n_{n-k}(x) \quad \forall\, x \in [0,1]$. easy

7) $\int_0^1 b^n_k = \frac{1}{n+1}$.

Anyway: (2) [alongwith many other] motivates us to define:

Def: Let $f : [0,1] \to \mathbb{R}$ be a fn. $\forall\, n \in \mathbb{N}$, define the Bernstein polynomial $B_n(f)$ as:

$$(B_n f)(x) = \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) b^n_k(x) \quad \left(= \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}\right).$$

Remark:

1) $B_n : C[0,1] \longrightarrow \mathbb{R}[x]$.
$f \longmapsto B_n f \leftarrow$ a poly. of degree at most n.

2) B_n is linear: $\quad B_n(af+g) = a B_n f + B_n g$
$\forall\, a \in \mathbb{R},\ f, g \in C[0,1]$.

3) Let $\underbrace{f \geq g}_{\text{i.e. } f(x) \geq g(x)\ \forall x}$ in $C[0,1]$. Then $B_n(f) \geq B_n(g)$. $\leftarrow$ B_n is monotonic

[Indeed, enough to prove: $B_n(f) \geq 0$ if $f(x) \geq 0\ \forall x$.
Straightaway follows from (5) & $f\left(\frac{k}{n}\right) \geq 0$].

4) $|B_n f| \leq B_n g$ if $|f| \leq g$. $\leftarrow$ We need this.
[$|f| \leq g \iff -g \leq f \leq g$. Next: apply (3)].

5) $B_n 1 = 1$ [by (3)].

6) Let $f(x) = x\ \forall x$. Then $B_n f = f$ (i.e: $B_n x = x$).

$\therefore\ B_n f = \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k}$.

$$= \frac{1}{n}\, x \sum_{k=0}^{n} k \binom{n}{k} x^k (1-x)^{n-k} = x$$

– why? [Hint: Use $\frac{d}{da}(a+b)^n = n(a+b)^{n-1}$]

$\Rightarrow n(a+b)^{n-1} = \sum_{k=0}^{n} k\binom{n}{k} a^{k-1} \cdot b^{n-k}$

7) Use again, diff., & get:

$$B_n x^2 = x^2 + \frac{x-x^2}{n}.$$

You can go on like this. VERY INTERESTING.

Proof of Weierstrass approx. thm: [We need $\{B1, Bx, Bx^2\}$, & some basic properties (as remarked earlier).]

Let $f \in C[0,1]$, $\varepsilon > 0$. $\because f$ is unif. cont. $\exists\ \delta > 0$ s.t.

$$|f(x) - f(y)| < \varepsilon/2 \quad \forall\ x, y \in [0,1],\ |x-y| < \delta.$$

Set $M := \sup_{x \in [0,1]} |f(x)|$. Pick & fix $a \in [0,1]$.

Then $\forall x \in [0,1]$ $|f(x) - f(a)| \leq \frac{\varepsilon}{2} + \frac{2M}{\delta^2}(x-a)^2$ ——— (*)

Trivial. If $|x-a| < \delta$, then $|f(x) - f(a)| < \frac{\varepsilon}{2}$.

If $|x-a| \geq \delta$, then $|f(x) - f(a)| \leq 2M \leq 2M\frac{(x-a)^2}{\delta^2}$

Curious $= \frac{2M}{\delta^2}(x-a)^2 \leq \frac{\varepsilon}{2} + \frac{2M}{\delta^2}(x-a)^2$

Then $\forall x \in [0,1]$,

$\because B_n$ is linear.

$$|(B_nf)(x) - f(a)| \overset{1.}{=} \Big|\Big(B_n\big(f - \underbrace{f(a)}_{\text{Const ant fn}}\big)\Big)(x)\Big|$$

$$\overset{(*)\ \&\ (4)}{\leq} B_n\left(\frac{\varepsilon}{2} + \frac{2M}{\delta^2}(x-a)^2\right).$$

$$\overset{\text{Linearity of } B_n}{=} \frac{\varepsilon}{2} + \frac{2M}{\delta^2}\underbrace{B_n(x-a)^2}$$

$= B_n(x^2 - 2ax + a^2)$ ~~$= B_n\left(x^2 + \frac{x-x^2}{n}\right)$~~

$= \left(x^2 + \frac{x-x^2}{n}\right) - 2ax + a^2$

$= (x-a)^2 + \frac{x-x^2}{n}$.

~~$= \frac{\varepsilon}{2} + \frac{2M}{\delta^2}$~~

$$= \frac{\varepsilon}{2} + \frac{2M}{\delta^2}(x-a)^2 + \frac{2M}{\delta^2}\left(\frac{x-x^2}{n}\right).\ \forall x \in [0,1].$$

In particular:

$\overset{x=a}{\Longrightarrow} |(B_nf)(a) - f(a)| \leq \frac{\varepsilon}{2} + \frac{2M}{\delta^2}\left(\frac{a-a^2}{n}\right) \leq \frac{\varepsilon}{2} + \frac{M}{2\delta^2 n}$.

[$\because \max\{a - a^2 : 0 \leq a \leq 1\} = \frac{1}{4}$]

$$\Rightarrow \left|(B_n f)(a) - f(a)\right| \leqslant \frac{\varepsilon}{2} + \frac{M}{2\delta^2 n}. \quad \forall a \in [0,1].$$

~~Choose~~

sup of LHS.

$$\Rightarrow \quad \|B_n f - f\| \leqslant \frac{\varepsilon}{2} + \frac{M}{2\delta^2 n}.$$

Choose $\underbrace{N \geqslant \frac{M}{\delta^2 \varepsilon}}$. Then $\forall n \geqslant N$,

$$\Rightarrow \frac{M}{2\delta^2 N} < \frac{\varepsilon}{2}.$$

$$\|B_n f - f\| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \qquad \square$$

— x —

Thank you :)