Weierstrass' approksimasjons teorem

Dette er et meget berømt teorem som ble bevist av den tyske matematikeren WEIERSTRASS i 1907. Weierstrass var en av klippene i tysk matrematikk (for ikke å si i verdensmatematikken) rundt århundreskiftet.

Theorem 1 *Weierstrass' approksimasjons teorem] La $I = [a, b]$ være et lukklet intervall. mengden av polynomer er tett i rommet $C([a, b], \mathbb{R})$ av kontinuerlige reellvaluerte funksjoner på $[a, b]$ utstyrt med sup-norm metrikken.*

Med andre ord, for en hver kontinuerlige reellvaluert funksjon $f(x)$ definert i $[a, b]$ og enhver $\epsilon > 0$, finnes det et polynom $P(x)$ slik at $\rho(f, P) = \sup\{f(x) - P(x) : x \in [a, b]\} < \epsilon$.

I disse notatene (og på forelseningen) gir vi et bevis for dette teoremet som følger en noe annen vei en det Tom gir i sine notater. Det matematiske, analytiske innholdet er det samme, men vi gjennomfører det uten referanser til statistikk.

Det er en del i auditoriet uten bakgrunn i statistikk, og for dem er nok et "myntkastebevis" noe mer krevende enn et ordinært et.

Den statistiske måten er elegant og kanskje oppklarende, men utelukkende for dem med en viss bakgrunn i statistikk. Det nevnes forøvrig i litteraturen at det var gjennom disse statistiske overlegningene at SERGEI NATANOVICH BERNSTEIN fant sitt teorem. Han var til syvende og sist mer av en statistiker enn en (ren-)matematiker.

Vi starter med å gjøre en forenkling:

Lemma 1 *Vi kan anta at $[a, b] = [0, 1]$.*

PROOF. La $p(x)$ være et polynom som avbilder $[0, 1]$ på $[a, b]$ og som er invertibelt, *e.g.*, p kan være strengt voksende. Det enjleste er å bruke $p(x) = (x - a)/(b - a)$ med invers $q(t) = (b - a)t + a$. Hvis $f(x)$ er kontinuerlig i $[a, b]$, så er $f(p(t))$ kontinuerlig i $[0, 1]$. Om $f(p(t))$ kan approksimeres uniformt med polynomet $Q(t)$ over $[0, 1]$, så kan $f(x) = f(p(q(x)))$ approksimeres uniform med $Q(q(x))$ over $[a, b]$. □

Bernstein-polynomer

Byggestenene i *Bernsteins* tilnærmingsmåte er de såkalt *elementære Bernstein-polynomene* . Det finnes ett slik for hvert naturlige tall n og hvert naturlige tall $r \leq n$. Disse polynomene skal vi skrive som $p_{n,r}(x)$, og de er definert ved:

$$p_{n,r}(x) = \binom{n}{r} x^r (1 - x)^{n-r}$$

Det klart at $p_{n,r}(x)$ tar verdiene 0 i begge endepunktene til $[0,1]$, og $p_{n,r}(x)$ er selvsagt positivt over hele $[0,1]$. Deriverer man $p_{n,r}(x)$ finner man at $p_{n,r}(x)' = (r - nx)p_{n-2,r-1}(x)$. Det gir at $p_{n,r}(x)$ har et maksimumspunkt for $x = \frac{r}{n}$ — og ved innsetting finner vi at maksimumsverdien er det fantastiske tallet

$$\frac{(n-r)^{n-r}}{(n-r)!}\frac{n^n}{n!}\frac{r!}{r^r}.$$

Tro det eller ei, men det holder seg i nærheten av 1 iallfall når n er stor, mens $\binom{n}{r}x^r(1-x)^{n-r}$ synker ganske fort med en gang vi beveger oss vekk fra $\frac{r}{n}$. Det betyr at $f(\frac{r}{n})p_{n,r}(x)$ er nær $f(x)$ når x er nær $\frac{r}{n}$, og liten når x lengre unna. Det er disse heuristiske overlegningen som, i en viss forstand, er ledetråden i beviset (om man ikke er statistiker da).

La nå $f(x)$ være en kontinuerlig funskjon på intetvallet $[0,1]$. La videre n være et naturlige tall. Vi definerer *Bernstein-polynomet av grad n tilordnet f* som følgende lineærkombinasjon av elementære Bernstein-polynomer:

$$B_n(f;x) = \sum_{r=0}^{n} f(\frac{r}{n})\binom{n}{r}x^r(1-x)^{n-r}. \qquad (✸)$$

Vi skal bevise følgende:

Theorem 2 *La $f(x)$ være en kontinuerlig funskjon på intervallet $[0,1]$. La videre $\epsilon > 0$ være gitt. Da finnes et naturlig tall N slik at bare $n \geq N$, så er*

$$|B_n(f;x) - f(x)| < \epsilon$$

for alle x i $[0,1]$.

Det første vi skal gjøre er å etablere tre spesialtilfeller av teoremet som viser seg å være nyttig. Først prøver vi oss ganske bekjedent med den enkleste funksjonen, nemlig den som er som konstant lik 1: Vi har

Lemma 2

$$\sum_{r=0}^{n}\binom{n}{r}x^r(1-x)^{n-r} = 1. \qquad (✸)$$

PROOF. Dette følger direkte av binomialformelen siden $(x + 1 - x)^n = 1$. □

Vi legger listen en smule høyere, og prøver oss med funksjonen x:

Lemma 3

$$B_n(x;x) = \sum_{r=0}^{n} \frac{r}{n}\binom{n}{r} x^r(1-x)^{n-r} = x$$

Proof. Denne er noe verre, men fortsatt svært overkommelig, så vi klarerer det uten å rive på første forsøk:

Vi bruker at $\frac{r}{n}\binom{n}{r} = \binom{n-1}{r-1}$. Det gir oss

$$\begin{aligned}\sum_{r=0}^{n} \frac{r}{n}\binom{n}{r} x^r(1-x)^{n-r} &= x\sum_{r=1}^{n}\binom{n-1}{r-1}x^{r-1}(1-x)^{n-1-r-1} \\ &= x\sum_{s=0}^{n-1}\binom{n-1}{s}x^s(1-x)^{n-1-s} = x\end{aligned}$$

der vi har skiftet summasjonsvariabel med $s = r-1$, og brukt lemma 2. Vi har også brukt at det første leddet — altså det tilsvarende $r = 0$ — i den første summen forsvinner, siden r er en faktor i det. □

Det tredje lemmaet vi trenger, handler om funksjonen x^2:

Lemma 4

$$B_n(x^2;x) = \ast = \sum_{r=0}^{n}(\frac{r}{n})^2\binom{n}{r}x^r(1-x)^{n-r} = x^2 + \frac{1}{n}x(1-x).$$

Vi bemerker at i motsetning til de to tidligere tilfellene, er ikke Bernstein-polynomet $B_n(x^2;x)$ for funksjonen x^2 lik funksjonen selv. Men siden $x(1-x)$ er begrenset av $1/4$, ser vi at $B_n(x^2;x)$ konvergerer uniformt mot x^2.

Proof. Ideen er først å gjøre samme manipulasjon som i forrige lemma: Det gir

$$\ast = \sum_{r=0}^{n}(\frac{r}{n})^2\binom{n}{r}x^r(1-x)^{n-r} = x\sum_{s=0}^{n-1}(\frac{s+1}{n})\binom{n-1}{s}x^s(1-x)^{n-1-s}.$$

Vi splitter summen i to, bruker at $\frac{s+1}{n-1} = \frac{n}{n-1}\frac{s}{n} + \frac{1}{n-1}$. Det gir

$$\begin{aligned}\ast &= x\frac{n-1}{n}\sum_{s=0}^{n-1}\frac{s}{n-1}\binom{n-1}{s}x^s(1-x)^{n-1-s} + \frac{x}{n}\sum_{s=0}^{n-1}\binom{n-1}{s}x^s(1-x)^{n-1-s} \\ &= x^2 - \frac{1}{n}x^2 + \frac{x}{n} = x^2 + \frac{1}{n}x(1-x).\end{aligned}$$

etter de to foregående lemmaene. □

Selve approksimasjonen

Så til selve verket. Vi skal vise at

$$\left|\sum_{r=0}^{n} f(\frac{r}{n})\binom{n}{r}x^r(1-x)^{n-r} - f(x)\right|$$

er liten når n er stor, og det uniformt i x. Presist betyr at hver gang ϵ er gitt, finnes en N slik at bare $n > N$ så er

$$\left|\sum\sum_{r=0}^{n} f(\frac{r}{n})\binom{n}{r}x^r(1-x)^{n-r} - f(x)\right| < \epsilon \qquad (\star)$$

for alle $x \in [0,1]$.

Det første vi gjør er å bruke ✹ til å skrive om $\star$:

$$\left|\sum_{r=0}^{n}(f(\frac{r}{n}) - f(x))\binom{n}{r}x^r(1-x)^{n-r}\right| < \epsilon$$

I denne summen er x og n fast og r varierer fra 0 til n, og trikset som får det hele til å virke, er å dele opp leddene i denne summen i to grupper, og gjøre hver av de tilsvarende summene små.

Vi vet at f er kontinuerlig — og defor uniformt kontinuerlig siden $[0,1]$ er kompakt. Så om $\frac{r}{n}$ ligger nær x kan vi forvente at $f(\frac{r}{n}) - f(x)$ er liten, så akkurat de leddene i summen har vi kontroll over.

Vi vil være presise, og gir oss en $\epsilon > 0$. Siden f er uniformt kontinuerlig, kan vi finne en δ slik at bare $\left|\frac{r}{n} - x\right| < \delta$, så er $\left|f(\frac{r}{n}) - f(x)\right| < \epsilon/2$. Videre setter vi S_δ være mengden av de r slik at $\left|\frac{r}{n} - x\right| < \delta$. Summerer vi over disse verdien av r, finner vi

$$\begin{aligned} ❄ &= \left|\sum_{r\in S_\delta} f(\frac{r}{n}) - f(x)\binom{n}{r}x^r(1-x)^{n-r}\right| \le \sum_{r\in S_\delta}\left|f(\frac{r}{n}) - f(x)\right|\binom{n}{r}x^r(1-x)^{n-r} \\ &< \epsilon\sum_{r\in S_\delta}\binom{n}{r}x^r(1-x)^{n-r} \le \epsilon\sum_{r=0}^{n}\binom{n}{r}x^r(1-x)^{n-r} = \epsilon/2 \end{aligned}$$

Så hva med summen resten av leddene? Vi skal gjøre bruk at de tre lemmaene våre. Vi finner

$$\left|\sum_{x\notin S_\delta}(\frac{r}{n} - x)^2\binom{n}{r}x^r(1-x)^{n-r}\right| \le \sum_{r=0}^{n}\left|(\frac{r}{n} - x)^2\right|\binom{n}{r}x^r(1-x)^{n-r} = \frac{1}{n}x(1-x),$$

men om $r \notin S_\delta$, så er $(\frac{r}{n} - x)^2 \geq \delta^2$. Det gir

$$\sum_{r \notin S_\delta} \binom{n}{r} x^r (1-x)^{n-r} \leq \frac{1}{\delta^2} \sum_{r \notin S_\delta} (\frac{r}{n} - x)^2 \binom{n}{r} x^r (1-x)^{n-r}$$
$$\leq \frac{1}{\delta^2} \sum_{r=0}^{1} (\frac{r}{n} - x)^2 \binom{n}{r} x^r (1-x)^{n-r} = \frac{1}{\delta^2} \frac{1}{n} x(1-x) \leq \frac{4}{n\delta^2}$$

Nå er f kontinuerlig på en kompakt, defor begrenset. For en stor nok konstant, M vil derfor $|f(x)| \leq M$ for alle $x \in [1, 0]$. Vi finner:

$$❄ = \sum_{r \notin S_\delta} \left| f(\frac{r}{n}) - f(x) \right| \binom{n}{r} x^r (1-x)^{n-r} \leq \frac{8M}{n\delta^2}$$

Nå kan δ være liten og M stor, så $\frac{M}{\delta^2}$ kan være enorm; men poenget er at den *ikke* avhenger av n, så velger vi n gigantisk, blir $\frac{8M}{n\delta^2}$ så lite vi bare vil, f.eks mindre enn $\epsilon/2$.

Setter vi de to rrupperingene sammen, finner vi

$$\left| \sum_{r=0}^{n} f(\frac{r}{n}) \binom{n}{r} x^r (1-x)^{n-r} - f(x) \right| \leq ❄ + ❄ < \epsilon/2 + \epsilon/2 = \epsilon. \qquad (\star\star)$$

Puuh!