The Formalisation and Applications of the Stone-Weierstrass Theorem

Kexing Ying

Imperial College London

http://imperial.cloud.panopto.eu/Panopto/Pages/Viewer.aspx?id=2d5eb25d-c698-44f3-a1a1-abd700fe1f51

The Stone-Weierstrass Theorem

Let X be a compact metric space and M the set of all continuous function from X to $\mathbb{R}$, then the *Stone-Weierstrass theorem* states that,

> Given an unital subalgebra of M, M_0 (i.e. a subset of M that contains 1 and is closed under addition, multiplication and scalar multiplication) that is closed under the lattice operations and separates points, $\bar{M}_0 = M$; [2]

where we define the lattice operations $\vee, \wedge : (X \to \mathbb{R})^2 \to (X \to \mathbb{R})$ such that, for all $f, g : X \to \mathbb{R}$, where f, g are continuous functions

$$f \vee g = \max\{f, g\}; \quad f \wedge g = \min\{f, g\};$$

$\bar{M}_0$ the closure of M_0 *under uniform convergence to the limit* and we say M_0 *separates points* if and only if for all distinct $x, y \in X$, there exists some $f \in M_0$ such that $f(x) \neq f(y)$.

Application: Weierstrass Theorem

The Weierstrass approximation theorem states

> Let f be a real continuous function on a closed interval $[\alpha, \beta]$. Then for all $\epsilon > 0$, there exists a polynomial P_n such that
>
> $$\sup_{x \in [\alpha, \beta]} |f(x) - P_n(x)| < \epsilon.$$

By section *Lemma: $|x|$ is in the Closure!*, the sequence of polynomial $P_n(x) \to |x|$ uniformly where,

$$P_n(x) := \sum_{k=0}^{n} \left|\alpha \frac{k}{n} - \beta\right| \binom{n}{i} x^i (1-x)^{n-i}$$

Now, as

$$f \vee g = \frac{1}{2}(f + g + |f - g|);$$

$$f \wedge g = \frac{1}{2}(f + g - |f - g|),$$

we see $\bar{\mathcal{P}}$ (the closure of polynomials under uniform convergence) is closed under the lattice operations. Thus, as $\bar{\mathcal{P}}$ form an unital subalgebra of all real to real functions, (as it contains 1, closed under addition, multiplication and scalar multiplication), and as $\bar{\mathcal{P}}$ separates points trivially, $\bar{\bar{\mathcal{P}}} = \bar{\mathcal{P}} = [\alpha, \beta]^{\mathbb{R}}$ for all $\alpha, \beta \in \mathbb{R}, \alpha < \beta$ as required!

Lemma: $|x|$ is in the Closure!

Let $X_1, \cdots X_n$ be a sequence of i.i.d. Bernoulli random variables with parameter x and let $S_n = \sum_{i=1}^{n} X_i$ so $S_n \sim Bin(n, x)$. Then by the *law of the unconscious statistician*, we see,

$$\mathbb{E}\left(\left|\alpha \frac{S_n}{n} - \beta\right|\right) = \sum_{k=0}^{n} \left|\alpha \frac{k}{n} - \beta\right| \Pr(S_n = nk)$$

$$= \sum_{k=0}^{n} \left|\alpha \frac{k}{n} - \beta\right| \binom{n}{i} x^i (1-x)^{n-i} =: P_n(x)$$

for all $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$. Let us now define

$$A_n(\delta) = \left\{\omega : \left|\frac{S_n(\omega)}{n} - x\right| \geq \delta\right\},$$

then, for all $\epsilon > 0$, as $f(x) : [0,1] \to \mathbb{R} := |\alpha x - \beta|$ is continuous on the compact interval $[0,1]$, it is uniformly continuous on the same interval; so there is some $\delta > 0$ such that for all $|x - y| < \delta$, $\frac{\epsilon}{2} > |f(x) - f(y)|$. Now, by LLN, $\Pr(A_n(\delta)) \to 0$, so, choose N such that $|\Pr(A_N(\delta))| < 1/2$, then

$$\frac{\epsilon}{2} > \mathbb{E}\left[\mathbf{1}_{(A_N)^c} \left|f(x) - f\left(\frac{S_N}{N}\right)\right| \mid A_N(\delta)^c\right]$$

$$\geq \left|f(x) - \mathbb{E}\left[f\left(\frac{S_N}{N}\right)\right]\right| \Pr((A_N(\delta))^c).$$

Now as $\mathbb{E}\left[f\left(\frac{S_N}{N}\right)\right] = P_N(x)$, we have $\epsilon/2 > |f(x) - P_N(x)|(1 - \Pr(A_N(\delta))) > |f(x) - P_N(x)|/2$, so $P_n(x) \to f(x) = |\alpha x - \beta|$ uniformly for $x \in [0,1]$ and $P_n(\frac{x+\beta}{\alpha}) \to |x|$ uniformly for all $x \in [\alpha, \beta]$.

Application: Trig. Polynomials

Another set of functions that are interesting are the trigonometric polynomials $\mathcal{T}$. As any constant is a trigonometric polynomial, $1 \in \mathcal{T}$ while $\mathcal{T}$ is trivially closed under addition and scalar multiplication. Thus, by considering the identity

$$\sin(nx)\cos(kx) = \frac{1}{2}(\cos(n-k)x - \cos(n+k)x),$$

and similarly the identities of $\sin(nx)\sin(kx)$ and $\cos(nx)\cos(kx)$, $\mathcal{T}$ is closed under multiplication. Thus $\mathcal{T}$ form an unital subalgebra of continuous functions on $[-\pi, \pi]$, and by considering the Fourier series of $|x|$, $\bar{\mathcal{T}}$ is closed under lattice operations and hence, by Stone-Weierstrass, $\bar{\bar{\mathcal{T}}} = \bar{\mathcal{T}} = [-\pi, \pi]^{\mathbb{R}}$.

Proof Outline of Stone-Weierstrass

Lemma 1. For all $f \in M$, f is in $\bar{M}_0$ if and only if for all $x, y \in X$, $\epsilon > 0$, there exists $g \in M_0$ such that $|f(x) - g(x)| < \epsilon$ and $|f(y) - g(y)| < \epsilon$.

Proof. The forward direction is by definition so we consider the reverse. Suppose for all $x, y \in X$, $\epsilon > 0$, there exists $g \in M_0$ such that $|f(x) - g(x)| < \epsilon$ and $|f(y) - g(y)| < \epsilon$ $(*)$. Let us fix x and ϵ and define the mapping

$$S(y) := \{z \in X \mid f(z) - g_y(z) < \epsilon\},$$

where g_y was chosen such that $|f(x) - g_y(x)| < \epsilon$ and $|f(y) - g_y(y)| < \epsilon$ which existence is guaranteed by $(*)$.

Then for all $y \in X$, $y \in S(y)$ so $\bigcup_{y \in X} S(y) = X$. But as X is compact, $\bigcup_{y \in X} S(y)$ admits a finite subcover; so, there exists a finite index set I such that $\bigcup_{i \in I} S(y_i) = X$. Thus, by letting $p_x = \bigvee_{i \in I} g_{y_i}$, we have constructed a function $p_x \in \bar{M}_0$ such that

$$p_x(z) \geq g_{y_i}(z) > f(z) - \epsilon \text{ and } p_x(x) < f(x) + \epsilon$$

for all $z \in X$ and $i \in I$.

Now, by defining a similar mapping, T,

$$T := \{z \in X \mid p_x(z) < f(z) + \epsilon\},$$

we again create a finite subcover of X and thus can create the required function with $\bigwedge_{j \in J} p_{x_j}$ where J is the index set such that $\bigcup_{j \in J} T(x_j) = X$.

Lemma 2. Given S, a subalgebra of $\mathbb{R}^2$, S must be $\{(0,0)\}$, $\{(x, 0) \mid x \in \mathbb{R}\}$, $\{(0.y) \mid y \in \mathbb{R}\}$, $\{(z, z) \mid z \in \mathbb{R}\}$, or $\mathbb{R}^2$ itself.

Lemma 2 was proved by evoking the law of the excluded middle on different propositions.

Given $x, y \in X$, define the boundary points of M_0 to be $\{(f(x), f(y)) \mid f \in M_0\}$. Let M_0, M_1 be closed subalgebras of M under lattice operations, by lemma 1, it is deduced $\bar{M}_0 = \bar{M}_1$ iff. for distinct x, y, $\bar{M}_0$ and $\bar{M}_1$ have the same boundary points.

Now, as boundary points of M_0 form an unital subalgebra of $\mathbb{R}^2$, the boundary points of $\bar{M}_0$ must be $\mathbb{R}^2$ by lemma 2 (the first three options excluded as $1 \in M_0$ and the fourth excluded as M_0 **separate points**) hence, as the boundary points of M is $\mathbb{R}^2$ is $\mathbb{R}^2$, it follows $\bar{M}_0 = M$ as required!

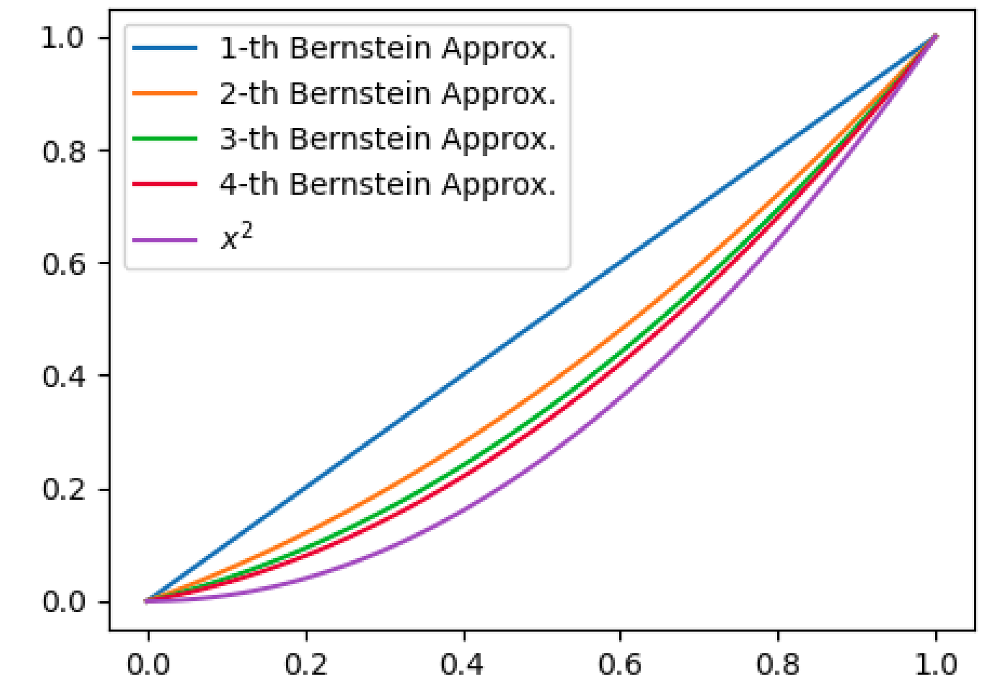
The theorem was also formalised using the *Lean*: http://github.com/JasonKYi/stone-weierstrass/tree/master/src

Further Discussion & Interpolation

The Stone-Weierstrass theorem proved the density of polynomial functions, but does not provide a construction. We can achieve this using the Bernstein polynomials. Consider the approximation of x^2 for $x \in [0,1]$ using Bernstein polynomials,

$$B_n(x) = \frac{1}{n}x + \frac{n-1}{n}x^2.$$

By using calculus, we find $\sup_{x \in [0,1]} |x^2 - B_n(x)| = (4n)^{-1}$; much worse than just using x^2 as its own approximation.



Consider instead the Lagrange interpolation [3]. Given $(x_0, y_0), \cdots, (x_n, y_n)$ be n distinct points, we want polynomial $L_n(x) = \sum_{i=0}^{n} a_i x^i$, $L_n(x_i) = y_i$ $\implies V[a_i] = [y_i]$ where V is the Vandermonde matrix such that $V_{i,j} = x_{i-1}^{j-1}$. As $V_{i,j}$ is invertible by the Vandermonde determinant, we found an unique coefficient vector interpolating these points.

Let $f : [a, b] \to \mathbb{R}$ be a continuous, we can approximate f using the polynomial that interpolates the set of equidistant points from a to b. With Python, we find that for $n = 3$ for x^2 on $[0,1]$, this method results in a polynomial approximation of $L_3(x) = x^2$.

However, $\|f - P_n\|_\infty$ does not necessarily tends to 0 exemplified by *Runge's function*. To mitigate this, one can instead use *Spline interpolation* however this results in an approximant that is a piecewise polynomial instead of just a polynomial.

References

[1] Stone, M.H. (1948) The Generalized Weierstrass Approximation Theorem. Mathematics Magazine 21, no. 5 : 237-54.

[2] Gaddy, P. The Stone-Weierstrass Theorem and its Applications to L^2 Spaces.

[3] Humpherys, J. (2020) Foundations of Applied Mathematics Volume 2: Algorithms, Approximation, Optimization Society for Industrial and Applied Mathematics