APPROXIMATING CONTINUOUS FUNCTIONS: WEIERSTRASS, BERNSTEIN, AND RUNGE

WILLIE WAI-YEUNG WONG

1. Introduction

This set of notes is meant to describe some aspects of polynomial approximations to continuous functions. It in particular concerns the apparent discord between the *Weierstrass Approximation Theorem* and *Runge's Phenomenon*.

First let us recall the theorem due to Weierstrass:

thm:Weierstrass **Theorem 1** (Weierstrass Approximation, real-variable version)**.** *Let* $f : [0,1] \to \mathbb{R}$ *be continuous. Then there exists a sequence of polynomials* p_n *that approximates* f *uniformly on* $[0,1]$*. That is to say we have*

$$\lim_{n\to\infty} \sup_{x\in[0,1]} |f(x) - p_n(x)| = 0 . \tag{1}$$

In particular, the theorem establishes that the polynomials are *dense* in the set of continuous functions $C([0,1];\mathbb{R})$ in the topology of the uniform norm. Note also by re-scaling, the interval $[0,1]$ in Theorem thm:Weierstrass 1 can be replaced by by any compact interval $[a,b]$.

On the other hand, Runge's example concerns the following function,

$$R(x) = \frac{1}{1+x^2} , \qquad x \in [-5,5] . \tag{2}$$

By Weierstrass's theorem, there exists an approximating sequence of polynomials. Naïvely one may try to construct this sequence by *polynomial interpolation*.

Definition 2. Let $f(x) : [a,b] \to \mathbb{R}$ be continuous. We denote by $I_n[f]$ the nth order equidistant polynomial interpolation of f. That is to say, we denote by $I_n[f]$ the unique nth order polynomial such that

$$I_n[f](a + i(b-a)/n) = f(a + i(b-a)/n) , \qquad \forall i \in \{0, 1, \ldots, n\} .$$

More generally, for each $n \geq 2$, let σ_n be a subset of $[a,b]$ consisting of $n+1$ distinct points, two of which are a and b. We denote by $I_n^\sigma[f]$ the nth order polynomial interpolation of f relative to the points σ_n. That is to say we let $I_n^\sigma[f]$ be the unique nth order polynomial such that

$$I_n^\sigma[f](x) = f(x) , \qquad \forall x \in \sigma_n .$$

Note that by definition, there exists a subsequence $I_{n_k}[R]$ which converges on every rational point in $[-5,5]$ (let $n_k = \prod_{i=1}^{k+1} p_i$ where p_i is the ith smallest prime

Version rev358 of 2013-11-29 12:00:03 -0500 (Fri, 29 Nov 2013).

number). But as it turns out, for the function $R(x)$, we not only have a lack of uniformity in the "convergence" of $I_n[R]$ to R, we have in fact that

$$\limsup_{n\to\infty} \sup_{x\in[-5,5]} |I_n[R](x) - R(x)| = +\infty \ .$$

As we shall see below, the method of polynomial interpolation is not particularly good for constructing the approximation sequence. In turns out that a better approximation can be obtained explicitly using *Bernstein polynomials* (incidentally we can directly prove Theorem 1 using these polynomials).

2. Weierstrass approximation through Bernstein polynomials

We first recall the Bernstein polynomials.

Definition 3. By $\mathfrak{B}_{k,n}(t)$ we denote the kth Bernstein polynomial of order n. They can be explicitly given by

$$\mathfrak{B}_{k,n}(t) = \binom{n}{k} t^k (1-t)^{n-k} \ . \tag{3}$$

We note that they form a partition of unity: $\sum_{k=0}^{n} \mathfrak{B}_{k,n}(t) = [t + (1-t)]^n = 1$.

Definition 4. Let $f : [0,1] \to \mathbb{R}$ be continuous. Denote by $B_n[f]$ the Bernstein-polynomial approximation of f, which we define as

$$B_n[f](x) = \sum_{k=0}^{n} f(k/n)\mathfrak{B}_{k,n}(x) \ .$$

By construction $B_n[f]$ is a polynomial of degree at most n. It is easily checked that since $\mathfrak{B}_{k,n}(0) = \delta_{k,0} = \mathfrak{B}_{n-k,n}(1)$, that $B_n[f](0) = f(0)$ and $B_n[f](1) = f(1)$. We claim that

eq:bernsteinconv

$$\lim_{n\to\infty} \sup_{x\in[0,1]} |B_n[f](x) - f(x)| = 0 \tag{4}$$

which would imply Theorem 1.

Proof of (4). Since $\sum_k \mathfrak{B}_{k,n} = 1$, we have that $B_n[f]$ is a *weighted average*. We will show that as n grows, the weight heavily favours the points where $k/n \approx x$.

First we compute a few things about $\mathfrak{B}_{k,n}(x)$. Observe that

$$\begin{aligned}
\sum_{k=0}^{n} \frac{k}{n}\mathfrak{B}_{k,n}(x) &= \sum_{k=0}^{n} \frac{k}{n}\binom{n}{k} x^k (1-x)^{n-k} \\
&= \sum_{k=1}^{n} \frac{k}{n}\frac{n}{k}\binom{n-1}{k-1} x\cdot x^{k-1}(1-x)^{(n-1)-(k-1)} \\
&= x\cdot \sum_{j=0}^{n-1}\binom{n-1}{j} x^j (1-x)^{n-1-j} \\
&= x\cdot\sum_{j=0}^{n-1}\mathfrak{B}_{j,n-1}(x) = x
\end{aligned} \tag{5}$$

So in particular $\sum_k (k/n - x)\mathfrak{B}_{k,n}(x) = 0$. Next we see that

eq:variance

$$\begin{aligned}
\sum_{k=0}^{n} \left(\frac{k}{n} - x\right)^2 \mathfrak{B}_{k,n}(x) &= -x^2 + \sum_{k=1}^{n} \frac{k}{n}\binom{n-1}{k-1} x^k (1-x)^{n-k} \\
&= -x^2 + x \cdot \frac{n-1}{n} \sum_{j=0}^{n-1} \left(\frac{j}{n-1} + \frac{1}{n-1}\right)\binom{n-1}{j} x^j (1-x)^{(n-1)-j} \\
&= -x^2 + \frac{n-1}{n} x \cdot \left(\sum_{\ell=0}^{n-2} x \mathfrak{B}_{\ell,n-2}(x) + \frac{1}{n-1}\sum_{j=0}^{n-1} \mathfrak{B}_{j,n-1}(x)\right) \\
&= -x^2 + \frac{n-1}{n} x(x + 1/(n-1)) = \frac{1}{n} x(1-x)
\end{aligned} \tag{6}$$

Now we are in a position to conclude the proof. Fix an $\epsilon > 0$. Using that $f(x)$ is continuous on a compact interval $[0,1]$, we have that $f(x)$ is uniformly continuous, and hence there exists a δ such that $\left|f(x) - f(y)\right| < \epsilon$ whenever $\left|x - y\right| < \delta$. Uniform continuity also implies that $f(x)$ is bounded: $|f(x)| < M$. We consider

$$\begin{aligned}
|B_n[f](x) - f(x)| &\leq \sum_{k=0}^{n} |f(k/n) - f(x)| \mathfrak{B}_{n,k}(x) \\
&= \left(\sum_{|k/n-x|<\delta} + \sum_{|k/n-x|\geq\delta}\right) |f(k/n) - f(x)| \mathfrak{B}_{n,k}(x) .
\end{aligned} \tag{7}$$

For the first sum, by uniform continuity we have that $|f(k/n) - f(x)| < \epsilon$, and since $\sum \mathfrak{B}_{k,n} = 1$ we have that

$$\sum_{|k/n-x|<\delta} |f(k/n) - f(x)| \mathfrak{B}_{n,k}(x) < \epsilon . \tag{8}$$

For the second sum, we estimate

eq:secondsum

$$\sum_{|k/n-x|\geq\delta} |f(k/n) - f(x)| \mathfrak{B}_{n,k}(x) \leq \sum_{|k/n-x|\geq\delta} \frac{2M}{\delta^2} |k/n - x|^2 \mathfrak{B}_{n,k}(x) \leq \frac{2M}{\delta^2 n} \tag{9}$$

by (6) [eq:variance]. So by choosing $n > \frac{2M}{\delta^2 \epsilon}$, we have that

$$|B_n[f](x) - f(x)| \leq 2\epsilon$$

independent of x, and so we obtained the uniform convergence. □

An after-note: the above proof can be interpreted using the probabilistic point of view as saying something about the binomial distribution. Observe that if we have a biased coin that lands with probability p heads and probability $(1-p)$ tails, the probability of receiving k heads after n tosses is exactly $\mathfrak{B}_{k,n}(p)$. So by the *Law of Large Numbers*, we expect that as $n \nearrow \infty$, for a fixed x, the functions $\mathfrak{B}_{k,n}(x)$ will be all mostly close to zero except for the few points where $k/n \approx x$. Hence in the summation $B_n[f](x)$ we expect mostly only contributions from $f(k/n)$ where $k/n \approx x$. By continuity this means that $B_n[f](x) \approx x$. The argument given above makes this intuition rigorous.

Notice that also that in the construction of $B_n[f]$, the only points at which $B_n[f]$ is guaranteed to be equal to f are the two endpoints 0 and 1.

Notice also that in the proof above, the rate of convergence can be quantitatively estimated if we have Lipschitz control on f. That is, suppose we know that $\left|f(x)-f(y)\right| \leq N\left|x-y\right|$ for all $x,y \in [0,1]$, then we have that $\delta \geq \epsilon/N$, and also that eq:secondsum (9) can be replaced by

eq:secondsumprime

$$\sum_{|k/n-x|\geq\delta} |f(k/n)-f(x)|\,\mathfrak{B}_{n,k}(x) \leq \sum \frac{N^3}{\epsilon^2}\,|k/n-x|^2\,\mathfrak{B}_{n,k}(x) \leq \frac{N^3}{\epsilon^2 n}\ . \tag{9'}$$

So we have a rate of convergence of at least $\epsilon \approx n^{-1/3}$.

3. Divergence of interpolation polynomials

In this section we will discuss a bit the case of divergence, namely that of the Runge example. A good part of the material is taken from Epperson[1]. The exact behaviour of the Runge example requires a dose of complex analysis.

Suppose now that $f(x)$ extends to $f(z)$, a function that is complex analytic in a neighborhood of the subset $[a,b]\times\{0\}$ of the complex plane. For the explicit function $R(x)$ we take $R(z) = (1+z^2)^{-1}$ which is holomorphic away from $\pm i$.

Let $\sigma_n = \{z_0,\ldots,z_n\}$ be a collection of distinct points in the region where $f(z)$ is holomorphic. We let $I_n^\sigma[f]$ be the polynomial interpolation of f based on the control points z_j. Then we have that $f(z) - I_n^\sigma[f](z)$ is holomorphic whenever f is, and vanishes at $z_0,\ldots,z_n$. This means that we can write

$$g_n^\sigma(z) \stackrel{\text{def}}{=} \frac{f(z)-I_n^\sigma[f](z)}{w_n^\sigma(z)}$$

where

$$w_n^\sigma(z) = \prod_{i=0}^{n}(z-z_i) \tag{10}$$

such that $g_n^\sigma(z)$ is holomorphic in the domain of holomorphicity of $f(z)$. Now, let γ be a closed contour, and let z be a point belonging to the domain bounded by γ. Writing $G_n^\sigma(\zeta) = g_n^\sigma(\zeta)/(\zeta - z)$ and applying to it the residue theorem, we have that

$$\frac{1}{2\pi i}\oint_\gamma G_n^\sigma(\zeta)\,\mathrm{d}z = g_n^\sigma(z) + \sum_k \mathrm{Res}(G_n^\sigma, a_k)$$

where $\{a_k\}$ enumerates the poles of $f(z)$ contained inside the domain bounded by γ. We can rewrite as

$$\frac{1}{2\pi i}\oint_\gamma G_n^\sigma(\zeta)\,\mathrm{d}z = g_n^\sigma(z) + \sum_k \frac{1}{w_n^\sigma(a_k)\cdot(a_k - z)}\mathrm{Res}(f,a_k)\ .$$

So finally, after rearranging the terms we get

eqn:errorestimate

$$f(z) - I_n^\sigma[f](z) = \frac{1}{2\pi i}\oint_\gamma \frac{w_n^\sigma(z)}{w_n^\sigma(\zeta)}\frac{f(\zeta)}{\zeta - z}\,\mathrm{d}z - \sum_k \frac{w_n^\sigma(z)}{w_n^\sigma(a_k)}\frac{\mathrm{Res}(f,a_k)}{a_k - z}\ . \tag{11}$$

So we see that the key to understanding the error behaviour as $n\to\infty$ is through understanding the function $w_n^\sigma(z)$. Obviously, the limiting behaviour of $w_n^\sigma(z)$ as $n\to\infty$ would depend on the sequence of sets σ_n.

---

[1] James Epperson, "On the Runge example", *Amer. Math. Monthly* **94** (1987), 329–341

Denote by δ_{σ_n} the counting measure attached to σ_n. Let us assume that $\frac{1}{n+1}\delta_{\sigma_n} \to \rho$ in measure. Then

$$\frac{1}{n+1}\log|w_n^\sigma(z)| = \frac{1}{n+1}\int \log|z-\zeta| \, \mathrm{d}\delta_{\sigma_n}(\zeta)$$

and hence for each fixed point z outside the support of ρ we have that

$$\lim_{n\to\infty}\frac{1}{n+1}\log|w_n^\sigma(z)| = \int \log|z-\zeta| \, \mathrm{d}\rho(\zeta) .$$

Hence we can interpret the function

eqn:winfty

$$\bar{w}_\infty^\sigma(z) \stackrel{\text{def}}{=} \lim_{n\to\infty}|w_n^\sigma(z)|^{\frac{1}{n+1}} = \exp\left(\int \log|z-\zeta| \, \mathrm{d}\rho(\zeta)\right) . \tag{12}$$

By differentiating under the integral sign, which we can do as long as we stay away from the support of the measure ρ, we have that $\bar{w}_\infty^\sigma(z)$ is real-valued and smooth.

Now, assume that ρ has empty interior: that is to say that for every point $y \in \operatorname{supp}(\rho)$ and every open set $V \ni y$, $V \setminus \operatorname{supp}(\rho) \neq \emptyset$. Furthermore assume that the function $\bar{w}_\infty^\sigma$ defined by (12) extends continuously to the support of ρ. Then since $|w_n^\sigma(z)|^{\frac{1}{n+1}}$ are continuous functions, we conclude that the point-wise convergence established in (12) implies the convergence also on the support of ρ.

Now we can discuss the convergence and divergence of the polynomial interpolation. Henceforth z is a fixed point and γ is a closed curve not intersecting the support of ρ (which guarantees that $\bar{w}_\infty^\sigma|_\gamma > 0$), *with the property that*

(1) z is in the domain bounded by γ;
(2) $\bar{w}_\infty^\sigma(z) < \inf_{\zeta\in\gamma}\bar{w}_\infty^\sigma(\zeta)$.

Then passing to the limit we have that

$$\lim_{n\to\infty}\left|\frac{w_n^\sigma(z)}{w_n^\sigma(\zeta)}\right| \le \lim_{n\to\infty}\left(\frac{\bar{w}_\infty^\sigma(z)}{\inf_{\zeta\in\gamma}\bar{w}_\infty^\sigma(\zeta)}\right)^{n+1} = 0 .$$

Hence (11) implies that

eqn:simplifiederror

$$\lim_{n\to\infty} f(z) - I_n^\sigma[f](z) = -\lim_{n\to\infty}\sum_k \frac{w_n^\sigma(z)}{w_n^\sigma(a_k)}\frac{\operatorname{Res}(f,a_k)}{a_k - z} . \tag{13}$$

Now, if z is a value such that we can choose γ to satisfy the above properties, and such that for every pole a_k enclosed by γ the value $\bar{w}_\infty^\sigma(z) < \bar{w}_\infty^\sigma(a_k)$, then using the exact same argument as above, we can conclude that $I_n^\sigma[f](z) \to f(z)$. On the other hand, if for every γ admissible we must include a_k such that $\bar{w}_\infty^\sigma(z) > \bar{w}_\infty^\sigma(a_k)$, then generically the right hand side must blow-up.

For the Runge example, the measure ρ coming from the limit of the equidistant interpolation is the Lebesgue measure on the interval $[-5,5]$. Evaluating the integral explicitly we have that at the two end-points $\bar{w}_\infty^\sigma(\pm 5) = \exp(10\ln 10 - 10)$, while at the midpoint $\bar{w}_\infty^\sigma(0) = \exp(10\ln 5 - 10)$. For the singularities of $R(z)$ at $\pm i$, we compute

$$2\int_0^5 \ln\sqrt{1+x^2}\,\mathrm{d}x = 5\ln 26 - 10 + 2\arctan 5 \approx 9.03728\ldots .$$

Observing that

$$23 \approx 10\ln 10 > 5\ln 26 + 2\arctan 5 \approx 19 > 10\ln 5 \approx 16$$

we see that in the segment $x \in [-5,5]$ we have both convergent and divergent points.

In fact, we see from the above analysis the following stronger form of Runge's phenomenon:

Theorem 5. *Let $I = [a,b]$ be an interval. Let σ_n a sequence of finite subsets of $n+1$ distinct points, two of which are a,b. Assume that $\frac{1}{n+1}\delta_{\sigma_n}$ converges in measure to ρ, a measure on $[a,b]$ that is absolutely continuous with respect to the Lebesgue measure[2]. If we assume further that the associated*

eqn:infsup

$$\inf_I \bar{w}^\sigma_\infty < \sup_I \bar{w}^\sigma_\infty \,, \tag{14}$$

then we can find an appropriate real analytic function f defined on I such that $I^\sigma_n[f]$ converges to f pointwise on some non-empty subset of I, while diverges on some non-empty subset of I.

Sketch of proof. By continuity there exists a point z with $\mathfrak{Im}\, z > 0$ such that $\bar{w}^\sigma_\infty(z)$ between the supremum and infimum of $\bar{w}^\sigma_\infty(z)$ on I. Then the function $f(\zeta) = \frac{1}{(\zeta - z)(\zeta - \bar{z})}$ is complex analytic in a neighborhood of I and takes real values on I. □

A natural next question to ask is: when is (eqn:infsup 14) voided? This would require $\mathrm{d}\rho = \rho'(x)\,\mathrm{d}x$ where $\rho'(x)$ is Lebesgue integrable, such that

$$\int_a^b \log|\xi - x|\,\rho'(\xi)\,\mathrm{d}\xi \text{ is constant in } x \in [a,b]\,. \tag{15}$$

As it turns out, one of the answers is given by Chebyshev nodes. When the limiting density $\rho'(x) \propto 1/\sqrt{1-x^2}$ (taking, for simplicity, $-a = b = 1$), a result of classical potential theory gives that

$$w^\sigma_\infty(z) \propto \log \frac{2}{\left|z + \sqrt{z^2-1}\right|}$$

where the complex function $\sqrt{z^2-1}$ is defined (branch selection) so that $\lim_{|z|\to\infty} \frac{\sqrt{z^2-1}}{z} = 1$. With this branch chosen, we have that for $x \in \mathbb{R}$ and $|x| \leq 1$, $\sqrt{x^2-1} = \sqrt{1-x^2}\,i$, and hence the denominator

$$\left|x + \sqrt{x^2-1}\right| = 1$$

is constant along the interval.

For completeness, we give the computation for the Chebyshev nodes. Taking the transformation $x = \cos\theta$, we have that

$$2\int_{-1}^{1} \log|z-x|\,\frac{\mathrm{d}x}{\sqrt{1-x^2}} = \int_{-\pi}^{\pi} \log|z - \cos\theta|\,\mathrm{d}\theta\,.$$

[2]This condition is not strictly necessary, but this helps avoid cases where the integral defining $\bar{w}^\sigma_\infty$ is divergent.

We write $\cos\theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta})$. This prompts us to use the so-called *Joukowski transformation*

$$z = \frac{1}{2}(\zeta + \zeta^{-1})$$

which is defined for every $z \notin [-1,1]$. This is because then we simplify

$$|z - \cos\theta| = \frac{1}{2}\left|\zeta + \zeta^{-1} - e^{i\theta} - e^{-i\theta}\right| = \frac{1}{2}\left|\zeta - e^{i\theta}\right|\left|\zeta^{-1} - e^{-i\theta}\right| .$$

We can solve for ζ in terms of z to get $\zeta = z \pm \sqrt{z^2 - 1}$. We choose the plus sign and the definition of the square root as above, which implies that $|\zeta| \geq 1$. The integral then becomes

$$\int_{-\pi}^{\pi} \log \frac{\left|\zeta - e^{i\theta}\right|\left|\zeta^{-1} - e^{-i\theta}\right|}{2} \, \mathrm{d}\theta .$$

By the properties of the real valued logarithm it suffices to consider the integral

$$\int_{-\pi}^{\pi} \log\left|\zeta - e^{i\theta}\right| \, \mathrm{d}\theta$$

which is physically the potential associated to a uniform distribution of charge on the unit circle. As is well-known that for $|\zeta| \geq 1$ this potential is proportional to $\log|\zeta|$. Putting this all together gives us the expression for $\bar{w}^{\sigma}_{\infty}(z)$, and taking the limit we get the desired conclusion.

4. Some final remarks

The discussion above with regards to the Chebyshev nodes show only that they can be used to get good polynomial interpolation of *real analytic* functions. In practice it turns out that Chebyshev nodes are good for all *absolutely continuous functions*. On the other hand, once we go to the continuous case there are two interestingly opposing theorems.

First, by Weierstrass approximation we know that we have uniform convergence to any continuous function by some sequence of polynomials. As in turns out due to the Chebyshev alternation theorem, we can find one such sequence that intersects the original function in sufficiently many points, showing that this sequence can be constructed by interpolation. However, the nodes here exists *a fortiori*, and are different for each continuous function.

On the other hand, from an argument using the converse of the uniform boundedness principle, one can in fact show that for every set of subsets σ_n there exists a continuous function for which the sequence of interpolating polynomials diverges.

The Wikipedia article on "polynomial interpolation"[3] contains many references in its section on convergence properties.

École Polytechnique Fédérale de Lausanne, Switzerland
E-mail address: `willie.wong@epfl.ch`

[3] `http://en.wikipedia.org/wiki/Polynomial_interpolation`