Approximation of Continuous Functions*

Francis J. Narcowich

October 2014

1 Modulus of Continuity

Recall that every function continuous on a closed interval $-\infty < a \leq x \leq b < \infty$ is uniformly continuous: For every $\epsilon > 0$, there is a $\delta > 0$ such that

$$|f(x) - f(y)| < \epsilon \tag{1.1}$$

as long as $x, y \in [a, b]$ satisfy $|x - y| < \delta$. This differs from the definition of continuity at a single point y in that δ is *independent* of y, and depends only on ε.

Let's turn around the roles of ε and δ. In uniform continuity, we start with ε and look for δ. What we want to do now, is start with δ and, essentially, find ε. With this in mind we make the following definition[1]

Definition 1.1. *The* modulus of continuity *for* $f \in C[0, 1]$ *and* $\delta > 0$ *is defined to be*

$$\omega(f, \delta) = \sup\{|f(x) - f(y)| \colon |x - y| \leq \delta,\ x, y \in [0, 1]\}. \tag{1.2}$$

Example 1.2. *Let* $f(x) = \sqrt{x}$, $0 \leq x \leq 1$. *Show that* $\omega(f, \delta) \leq C\sqrt{\delta}$.

*Updated 10/2/2019

[1] From now on, without loss of generality, we will work with the closed interval $[0, 1]$.

Proof. Let $0 < x < y \le 1$. We note that

$$\begin{aligned}
0 < |\sqrt{y} - \sqrt{x}| &= \frac{y-x}{\sqrt{x}+\sqrt{y}} \\
&= \sqrt{y-x}\Big(\frac{\sqrt{y-x}}{\sqrt{y}+\sqrt{x}}\Big) \\
&= \sqrt{y-x}\frac{\sqrt{1-\sqrt{\frac{x}{y}}}\sqrt{1+\sqrt{\frac{x}{y}}}}{1+\sqrt{\frac{x}{y}}} = \sqrt{y-x}\frac{\sqrt{1-\sqrt{\frac{x}{y}}}}{\sqrt{1+\sqrt{\frac{x}{y}}}} \\
&\le \sqrt{y-x} \le \sqrt{\delta}
\end{aligned}$$

Hence, $\omega(f,\delta) \le \sqrt{\delta}$. To get equality, take $x = 0$ and $y = \delta$. □

Example 1.3. *Suppose that* $f \in C^{(1)}[0,1]$. *Show that* $\omega(f,\delta) \le ||f'||_\infty \delta$.

Proof. Because $f \in C^{(1)}$, we can estimate $f(t) - f(s)$ this way:

$$|f(t) - f(s)| \le \int_s^t |f'(x)|\, dx \le (t-s)\|f'\|_\infty \le \delta \|f'\|_\infty, \tag{1.3}$$

which immediately gives $\omega(f,\delta) \le \delta\|f'\|_\infty$. □

2 Approximation with Linear Splines

One very effective way to approximate a continuous function $f \in C[0,1]$, given a finite set of points in $\{x_j\} \subset [0,1]$ and values $\{f(x_j)\}$ at these points, is to use a "connect-the-dots" approach. The idea is to form a piecewise-linear, continuous function by joining neighboring points $(x_j, f(x_j))$ and $(x_{j+1}, f(x_{j+1})$ with a straight line. This procedure results in a *linear spline*. Linear splines are used for generating plots in many standard programs, such as Matlab or Mathematica.

Defining a space of linear splines starts with sequence of points (or partition of $[0,1]$) $\Delta = \{x_0 = 0 < x_1 < x_2 < \cdots < x_n = 1\}$, which is called a *knot sequence*. Linear splines on $[0,1]$ with knot sequence Δ are the set of all *piecewise* linear functions that are continuous on $[0,1]$ and that (possibly) have corners at the knots, but nowhere else. As described above, we can interpolate continuous functions using linear splines: Let $f \in C[0,1]$ and let $y_j = f(x_j)$. The linear spline $s_f(x)$ is constructed by joining pairs of points (x_j, y_j) and (x_{j+1}, y_{j+1}) with straight lines. The resulting spline is unique and satisfies the interpolation conditions $s_f(x_j) = y_j,\ j = 0, \ldots, n$. The result below gives an estimate of the error made by replacing f by s_f.

Proposition 2.1. *Let $f \in C[0,1]$ and let $\Delta = \{x_0 = 0 < x_1 < \cdots < x_n = 1\}$ be a knot sequence with norm $\|\Delta\| = \max |x_j - x_{j+1}|,\ j = 0, \ldots, n-1$. If s_f is the linear spline that interpolates f at the x_j's, then,*

$$\|f - s_f\|_\infty \leq \omega(f, \|\Delta\|). \tag{2.1}$$

Proof. Consider the interval $I_j = [x_j, x_{j+1}]$. We have on I_j that $s_f(x)$ is a line joining $(x_j, f(x_j))$ and $(x_{j+1}, f(x_{j+1}))$; it has the form

$$s_f(x) = \frac{x_{j+1} - x}{x_{j+1} - x_j} f(x_j) + \frac{x - x_j}{x_{j+1} - x_j} f(x_{j+1})$$

Also, note that we have

$$\frac{x_{j+1} - x}{x_{j+1} - x_j} + \frac{x - x_j}{x_{j+1} - x_j} = 1.$$

Using these equations, we see that $f(x) - s_f(x)$ for any $x \in [x_j, x_{j+1}]$ can be written as

$$\begin{aligned} f(x) - s_f(x) &= f(x)\left(\frac{x_{j+1} - x}{x_{j+1} - x_j} + \frac{x - x_j}{x_{j+1} - x_j}\right) - s_f(x) \\ &= (f(x) - f(x_j))\frac{x_{j+1} - x}{x_{j+1} - x_j} + (f(x) - f(x_{j+1}))\frac{x - x_j}{x_{j+1} - x_j}. \end{aligned}$$

By the definition of the modulus of continuity, $|f(x) - f(y)| \leq \omega(f, \delta)$ for any x, y such that $|x - y| \leq \delta$. If we set $\delta_j = x_{j+1} - x_j$, then we see that on the interval I_j we have

$$\begin{aligned} |f(x) - s_f(x)| &\leq \big|(f(x) - f(x_j))\big|\frac{x_{j+1} - x}{x_{j+1} - x_j} + \big|f(x) - f(x_{j+1})\big|\frac{x - t_j}{x_{j+1} - x_j} \\ &\leq \left(\frac{x_{j+1} - x}{x_{j+1} - x_j} + \frac{x - x_j}{x_{j+1} - x_j}\right)\omega(f, \delta_j) = \omega(f, \delta_j). \end{aligned}$$

Because the modulus of continuity is non decreasing (exercise 5.2(c)) and $\delta_j \leq \|\Delta\|$, we have $\omega(f, \delta_j) \leq \omega(f, \|\Delta\|)$. Consequently, $|f(x) - s_f(x)| \leq \omega(f, \|\Delta\|)$, uniformly in x. Taking the supremum on the right side of this inequality then yields (2.1). $\square$

3 The Weierstrass Approximation Theorem

The completeness of various sets of orthogonal polynomials relies on being able to uniformly approximate continuous functions by polynomials. The Weierstrass Approximation Theorem does exactly that. The proof that we will give here follows the one Sergei Bernstein gave in 1912. The proof is not the "slickest," but it does introduce a number of important things. Here is the statement.

Theorem 3.1 (Weierstrass Approximation Theorem)**.** *Let $f \in C[0,1]$. Then, for every $\varepsilon > 0$ we can find a polynomial p such that $\|f - p\|_{C[0,1]} < \varepsilon$.*

3.1 Bernstein polynomials

Let n be a positive integer. The binomial theorem states that $(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j}$. We define the Bernstein polynomial using the terms in the expansion with $y = 1 - x$:

$$\beta_{j,n}(x) := \binom{n}{j} x^j (1-x)^{n-j}. \tag{3.1}$$

Proposition 3.2. *The Bernstein polynomials $\{\beta_{j,n}\}_{j=0}^{n}$ form a basis for $\mathcal{P}_n$, the space of polynomials of degree n or less.*

Proof. The dimension of $\mathcal{P}_n$ is $n+1$. Since there are $n+1$ Bernstein polynomials, we need only show that they span $\mathcal{P}_n$. We will show that 1, x are in the span of the Bernstein polynomials, and leave $x^2, \ldots, x^n$ as an exercise. To get 1, set $y = 1 - x$ in the binomial expansion we get $(x + 1 - x)^n = \sum_{j=0}^{n} \binom{n}{j} x^j (1-x)^{n-j}$, so that $1 = \sum_{j=0}^{n} \beta_{j,n}(x)$. For x, we take the partial derivative of $(x+y)^n$ with respect to x to get $n(x+y)^{n-1} = \sum_{j=1}^{n} j \binom{j}{n} x^{j-1} y^{n-j}$. Multiplying this by x, setting $y = 1-x$ and dividing by n, we obtain $x = \sum_{j=1}^{n} \frac{j}{n} \beta_{j,n}(x)$. The others are obtained similarly. $\square$

We will need several identities involving the Bernstein polynomials, which we now list. The last two identities start the sum at $j = 0$, rather than $j = 1$.

$$\left.\begin{aligned} 1 &= \sum_{j=0}^{n} \beta_{j,n}(x) \\ x &= \sum_{j=0}^{n} \frac{j}{n} \beta_{j,n}(x) \\ \frac{1}{n}x + (1 - \frac{1}{n})x^2 &= \sum_{j=0}^{n} \frac{j^2}{n^2} \beta_{j,n}(x). \end{aligned}\right\} \tag{3.2}$$

3.2 Proof of the Weierstrass Approximation Theorem

All of the Bernstein polynomials are positive, except at 0 and 1, where they are 0. When n is very large, the Bernstein polynomial $\beta_{j,n}$ is highly peaked near its maximum at $x = j/n$. That is, a small distance away from j/n, the polynomial $\beta_{j,n}$ is itself quite small. With this in mind, if $f \in C[0, 1]$, define the polynomial

$$f_n(x) := \sum_{j=0}^{n} f(j/n)\beta_{j,n}(x) \in \mathcal{P}_n.$$

The idea here is that near each point j/n the main contribution to the sum of Bernstein polynomials making up f_n should come from the term $f(j/n)\beta_{j,n}(x)$, so f_n should be a good approximation to f.

Proof. (Weierstrass Approximation Theorem) Choose n large; let $\delta > 0$. Using the first identity in (3.2), we have $f(x) = f(x) \cdot 1 = \sum_{j=0}^{n} f(x)\beta_{j,n}(x)$. Thus the difference between f and f_n is

$$E_n(x) := f(x) - f_n(x) = \sum_{j=0}^{n} (f(x) - f(j/n))\beta_{j,n}(x).$$

We want to show that, for sufficiently large n, $\|E_n\|_{C[0,1]} < \varepsilon$. Fix x. We are now going to break the sum into two parts. The first will be all those j for which $|x - j/n| \le \delta$. This is F_n below. The second, G_n consists of all remaining j's – namely, all j such that $|x - j/n| > \delta$. Carrying this out breaks E_n into the sum $E_n = F_n + G_n$, where

$$F_n(x) = \sum_{|x-j/n| \le \delta} (f(x) - f(j/n))\beta_{j,n}(x) \tag{3.3}$$

$$G_n(x) = \sum_{|x-j/n| > \delta} (f(x) - f(j/n))\beta_{j,n}(x). \tag{3.4}$$

Using the triangle inequality on the sum in F_n yields

$$|F_n(x)| \le \sum_{|x-j/n|\le\delta} |f(x) - f(j/n)|\beta_{j,n}(x).$$

Because $|x-j/n| \le \delta$, we also have $|f(x)-f(j/n)| \le \omega(f,\delta)$. Consequently,

$$|F_n(x)| \le \Bigg(\sum_{|x-j/n|\le\delta} \beta_{j,n}(x)\Bigg)\omega(f,\delta) \le \Bigg(\sum_{j=0}^{n} \beta_{j,n}(x)\Bigg)\omega(f,\delta).$$

Using the first identity in (3.2) in the inequality above yields this:

$$|F_n(x)| \le \omega(f,\delta). \tag{3.5}$$

Estimating G_n requires more care. We will treat the case in which $x - j/n > \delta$; the other case is similar. The idea is to think of δ as a unit of measure – like inches or centimeters. Then there will be a smallest k such that $k\delta < x - j/n \le (k+1)\delta$. Write $f(x) - f(j/n)$ this way:

$$\begin{aligned} f(x) - f(\tfrac{j}{n}) &= \big[f(x) - f(\tfrac{j}{n} + k\delta)\big] + \big[f(\tfrac{j}{n} + k\delta) - f(\tfrac{j}{n} + (k-1)\delta))\big] + \\ &\qquad \cdots + \big[f(\tfrac{j}{n} + \delta) - f(j/n)\big] \\ &= \big[f(x) - f(\tfrac{j}{n} + k\delta)\big] + \sum_{m=0}^{k-1} \big[f(\tfrac{j}{n} + (k-m)\delta) - f(\tfrac{j}{n} + (k-m-1)\delta)\big] \end{aligned}$$

Since $\frac{j}{n} + (m+1)\delta - \frac{j}{n} - m\delta = \delta$, for $m = 1, \ldots, k-1$ each term satisfies $|f(\frac{j}{n} + m\delta) - f(\frac{j}{n} + (m+1)\delta)| \le \omega(f,\delta)$. Also, because $|x - \frac{j}{n} - k\delta| \le \delta$, the first term satisfies $|f(x) - f(\frac{j}{n} + k\delta)| \le \omega(f,\delta)$. There are $k+1$ terms in the sum, so $|f(x) - f(\frac{j}{n})| \le (k+1)\omega(f,\delta)$. Since $k\delta < x - j/n$, we have $(k+1)\delta < \delta + (x-j/n)$. As we mentioned earlier, a similar argument will for $j/n > x$ will give $(k+1)\delta < \delta + j/n - x)$. Thus, for both cases, we have that $(k+1)\delta < \delta + |x - j/n|$. Dividing by δ, we see that $k+1 < 1 + |x-j/n|/\delta$; consequently,

$$|f(x) - f(j/n)| \le (k+1)\omega(f,\delta) \le \big(1 + |x - \tfrac{j}{n}|/\delta\big)\omega(f,\delta).$$

What we do next is use a trick that will help us to get an explicit bound on $|G_n(x)|$. The trick is to replace $|x-j/n|/\delta$ by $|x-j/n|^2/\delta^2$. Since $k\delta < |x-j/n|$, we have $\frac{|x-j/n|}{\delta} > 1$, and we thus also have $\frac{|x-j/n|}{\delta} < \frac{|x-j/n|^2}{\delta^2}$. Using this in the previous inequality results in

$$|f(x) - f(j/n)| < \left(1 + \frac{|x-j/n|^2}{\delta^2}\right)\omega(f,\delta). \tag{3.6}$$

Using the triangle inequality on the sum in (3.4) and bounding each term by (3.6), we obtain

$$
\begin{aligned}
|G_n(x)| &< \left(\sum_{j=0}^{n} \left(1 + \frac{|x - j/n|^2}{\delta^2} \right) \beta_{j,n}(x) \right) \omega(f, \delta) \\
&< \left(\sum_{j=0}^{n} \big(1 + \frac{x^2}{\delta^2} - \frac{2xj}{\delta^2 n} + \frac{j^2}{\delta^2 n^2} \big) \beta_{j,n}(x) \right) \omega(f, \delta) \qquad (3.7)
\end{aligned}
$$

Using the three identities in (3.2) allows us to do all of the sums in (3.7). The result, after a little algebra, is $|G_n(x)| < \big(1 + \frac{x-x^2}{\delta^2 n}\big)\omega(f, \delta)$. Since the maximum of $x - x^2$ over $[0, 1]$ is $1/4$, we have that

$$
|G_n(x)| < \big(1 + \frac{1}{4n\delta^2}\big)\omega(f, \delta). \qquad (3.8)
$$

Combining (3.5), (3.8) and $|E_n(x)| = |F_n(x) + G_n(x)| \leq |F_n(x)| + |G_n(x)|$ yields

$$
|E_n(x)| < \big(2 + \frac{1}{4n\delta^2}\big)\omega(f, \delta). \qquad (3.9)
$$

The parameter $\delta > 0$ is free for us to choose. Take it to be $\delta = n^{-1/2}$. With this choice we arrive at

$$
|f(x) - f_n(x)| = |E_n(x)| < \frac{9}{4}\omega(f, n^{-1/2}). \qquad (3.10)
$$

Taking the maximum of $|f(x) - f_n(x)|$ over $x \in [0, 1]$ gives us $\|f - f_n\| < \frac{9}{4}\omega(f, n^{-1/2})$. Choosing n so large that $\frac{9}{4}\omega(f, n^{-1/2}) < \varepsilon$ then completes the proof. □

Previous: orthonormal sets and expansions
Next: pointwise convergence of Fourier series