Convolution

Let $f(x)$ and g(x) be continuous real-valued functions for $x \in \mathbb{R}$ and assume that f or g is zero outside some bounded set (this assumption can be relaxed a bit). Define the *convolution*

$$(f*g)(x) := \int_{-\infty}^{\infty} f(x-y)g(y)\,dy \tag{1}$$

One preliminary useful observation is

$$f*g = g*f. \tag{2}$$

To prove this make the change of variable $t = x-y$ in the integral (1).

Remark 1 Note that if g is zero outside of the interval $[a,b]$,, then $(f*g)(x) = \int_a^b f(x-y)g(y)\,dy$, so only the values of f on the interval $[x-b, x-a]$ are used. Thus if $x \in [c,d]$, then the convolution only involves the values of f on $[c-b,\, d-a]$.

Remark 2 Similarly, if f is zero outside of the interval $[-\frac{1}{2}, \frac{1}{2}]$ and $x \in [c,\, d]$, then the convolution only involves the values of g on the interval $[c-\frac{1}{2}, d+\frac{1}{2}]$.

SMOOTHNESS OF $f*g$.

Theorem 1 *If $f \in C^1(\mathbb{R})$ then $f*g \in C^1(\mathbb{R})$. Better yet, if $f \in C^k(\mathbb{R})$ and $g \in C^\ell(\mathbb{R})$, then $f*g \in C^{k+\ell}(\mathbb{R})$.*

PROOF This is clearer if we write $h(x) := (f*g)(x)$. Then

$$\frac{h(x)-h(x_0)}{x-x_0} = \int_{-\infty}^{\infty} \frac{f(x-y)-f(x_0-y)}{x-x_0} g(x)\,dx. \tag{3}$$

We will be done if we can show that $[f(x-y)-f(x_0-y)]/(x-x_0)$ converges uniformly to $f'(x_0-y)$. To do this we use the integral form of the mean value theorem:

$$\begin{aligned} f(x-y)-f(x_0-y) &= \int_0^1 \frac{d f(x_0-y+t(x-x_0))}{dt}\,dt \\ &= \left[\int_0^1 f'(x_0-y+t(x-x_0))\,dt\right](x-x_0). \end{aligned}$$

Then

$$\frac{f(x-y)-f(x_0-y)}{x-x_0} - f'(x_0-y) = \int_0^1 [f'(x_0-y+t(x-x_0)) - f'(x_0-y)]\,dt \tag{4}$$

Since f' is assumed continuous and is zero outside of a bounded set, it is uniformly continuous. Thus, given any $\varepsilon > 0$ there is a $\delta > 0$ so that if $|x-x_0| < \delta$ then

$$|f'(z+t(x-x_0)) - f'(z)| < \varepsilon$$

for all values of z. In our case $z = x_0 - y$. Thus the left side of (4) tends to zero uniformly for all choices of x_0 and y. Consequently, $h \in C^1(\mathbb{R})$.

Repeating this we conclude that if $f \in C^k$ then $h \in C^k$. Because of (2) $f^{(k)} * g = g * f^{(k)}$, so we can repeat this reasoning to show that $g * f^{(k)} \in C^\ell$. Thus $f * g \in C^{k+\ell}$. Note that although g might not be zero outside a bounded set, because f is zero outside a bounded set, the integration in $g * f^{(k)}$ is only over a bounded set – in which the derivatives of g are uniformly continuous.

APPROXIMATE IDENTITIES

Let $\varphi_n(t)$ be a sequence of smooth real-valued functions with the properties

$$(a)\ \varphi_n(t) \geq 0, \qquad (b)\ \varphi_n(t) = 0 \ \text{ for } \ |t| \geq 1/n, \qquad (c) \int_{-\infty}^{\infty} \varphi_n(t)\,dt = 1. \tag{5}$$

Note: because of (b), this integral is only over $-1/n \leq t \leq 1/n$.

Assume $f(x)$ is uniformly continuous for all $x \in \mathbb{R}$ and zero outside a bounded set. Define

$$f_n(x) := (f * \varphi_n)(x) = \int_{-\infty}^{\infty} f(x-t)\varphi_n(t)\,dt. \tag{6}$$

Theorem 2 *$f_n(x) \in C^\infty$ converges uniformly to $f(x)$ for all $x \in \mathbb{R}$. Thus, on a compact set any continuous function can be approximated arbitrarily closely in the uniform norm by a smooth function.*

PROOF The smoothness of the approximations f_n is an immediate consequence of Theorem 1.

Since $f(x) = f(x)\left(\int_{-\infty}^{\infty} \varphi_n(t)\,dt\right) = \int_{-\infty}^{\infty} f(x)\varphi_n(t)\,dt$,

$$f_n(x) - f(x) = \int_{|t|\leq 1/n} [f(x-t) - f(x)]\varphi_n(t)\,dt. \tag{7}$$

Since f is uniformly continuous, given any $\varepsilon > 0$ there is a $\delta > 0$ such that if $|t| < \delta$ then $|f(x-t) - f(x)| < \varepsilon$ for all x. If $1/n < \delta$, then by (5c)

$$|f_n(x) - f(x)| < \varepsilon \int_{|t|\leq 1/n} \varphi_n(t)\,dt = \varepsilon. \tag{8}$$

Since the right side is independent of x this shows that in the uniform norm $\|f_n - f\|_\infty < \varepsilon$.

Since the operators $T_n(f) := f * \varphi_n \to f$, so in this sense T_n converges to the identity operator I, we sometime call the T_n (or the φ_n) *approximate identities*.

EXAMPLE Assume $f(x)$ is continuous on the interval $[a,b]$. Then $\int_a^b f(x) \sin \lambda x \, dx \to 0$.

PROOF: If $f \in \mathbb{C}^1([a,b])$ this is easy to show by an integration by parts, using $|f'(x)| \leq M$ for some constant M.
If f is only continuous, use Theorem 2 to find a smooth $g(x)$ with $\|f - g\|_\infty < \varepsilon$ on $[a,b]$. Then

$$\left| \int_a^b f(x) \sin \lambda x \, dx \right| \leq \left| \int_a^b [f(x) - g(x)] \sin \lambda x \, dx \right| + \left| \int_a^b g(x) \sin \lambda x \, dx \right|.$$

Since $\|f - g\|_\infty < \varepsilon$, the first term on the right is small. Because g is smooth, the second term goes to zero as $\lambda \to \infty$.

In many applications the condition (5b) is too restrictive.

Theorem 3 *Theorem 1 is valid if you replace* (5b) *with:*

$$\textit{For every } \delta > 0, \quad \lim_{n \to \infty} \int_{|t| > \delta} \varphi_n(t) \, dt = 0. \qquad \text{(5b')}$$

PROOF Replace (7) by

$$\begin{aligned} f_n(x) - f(x) &= \int_{|t| \leq \delta} [f(x-t) - f(x)] \varphi_n(t) \, dt + \int_{|t| > \delta} [f(x-t) - f(x)] \varphi_n(t) \, dt \\ &= J_1 + J_2. \end{aligned}$$

Given $\varepsilon > 0$, pick δ as was done above. Then

$$|J_1| \le \varepsilon \int_{|t|\le\delta} \varphi_n(t)\,dt \le \varepsilon.$$

To estimate J_2, say $|f(x| \le M$ for all x. Then by our assumption on the φ_n,

$$|J_2| \le 2M \int_{|t|>\delta} \varphi_n(t)\,dt \;\to\; 0.$$

This proves that $\|f_n - f\|_\infty \to 0$.

Weierstrass used essentially this argument to prove his Approximation Theorem (see below) with

$$u(x,t) := \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} e^{-\frac{(x-y)^2}{4t}} f(y)\,dy,$$

He was thinking of $t = 1/n \to 0$. Then $\lim_{t\to 0} u(x,t) \to f(x)$. This classical formula was well-known since $u(x,t)$ is the solution of the *heat equation* $u_t = u_{xx}$ for $x \in \mathbb{R}$, $t > 0$ with initial temperature $u(x,0) = f(x)$.

We'll use this idea but with a different integrand to prove

Theorem 4 (WEIERSTRASS APPROXIMATION THEOREM) *Let f be a continuous function. Then on any compact set it can be approximated uniformly by a polynomial.*

PROOF We prove this where f is continuous on a compact set $[a', b']$ in $\mathbb{R}$. The same proof works for a compact set in $\mathbb{R}^n$. As a preliminary step, extend f as a continuous function to a slightly larger interval $[a,b]$ so that this extended function satisfies $f(a) = f(b) = 0$ (for the interval $a \le x \le a'$ use a straight line between the points $(a,0)$ and $(a', f(a'))$, with a similar extension at the right end, $x = b'$). We can now extend f as a continuous function to all of $\mathbb{R}$ by letting $f(x) = 0$ outside of $[a,b]$. By scaling the $x-$axis, we may further assume that $f(x) = 0$ for $|x| \ge 1/2$.

Our approximations are

$$f_n(x) := f * \varphi_n(x) = \int_{-\infty}^{\infty} f(x-t)\varphi_n(t)\,dt. \qquad (9)$$

Because for us (see below) $\varphi_n(x) = 0$ for $|x| \ge 1$ this integral will be only over the interval $|x| \le 1$.

If in (9) the functions φ_n are polynomials, then the approximations $f * \varphi_n$ are also polynomials. However, polynomials will never satisfy the restrictions required of the φ_n. Our $\varphi_n(x)$, defined below, will be polynomials for $|x| \le 1$ and zero for $|x| \ge 1$. As we observed in the Remark 2 after equation (2), since our $f(x) = 0$ for $|x| \ge 1/2$, if $x \in [c,d]$, then the convolution $f_n(x) = f * \varphi_n$ will only use the values of $\varphi_n(x)$ for $x \in c - \frac{1}{2}, d + \frac{1}{2}]$. In particular, if $x \in [-\frac{1}{2}, \frac{1}{2}]$, then the convolution $f_n(x) = f * \varphi_n$ will only use the values of $\varphi_n(x)$ for $x \in [-1, 1]$ – which is exactly where φ_n is a polynomial. Note that if x is in a larger interval, the f_n will converge to f – but the f_n will not be polynomials.

Define the functions $\varphi_n(x)$ as

$$\varphi_n(x) = \begin{cases} \frac{(1-x^2)^n}{c_n} & \text{if } |x| \le 1 \\ 0 & \text{if } |x| > 1 \end{cases}, \qquad (10)$$

where

$$c_n = \int_{-1}^{1} (1-x^2)^n \, dx \tag{11}$$

was chosen so that φ_n satisfies the condition (5c). We will verify the modified property (5b') of Theorem 3 by showing that for any $\delta > 0$ in the region $|x| > \delta$ the functions $\varphi_n(x)$ converge uniformly to zero.

To show this we estimate the constants c_n in equation (11). After the change of variable $t = x^2$

$$c_n = 2\int_0^1 (1-x^2)^n \, dx = \int_0^1 (1-t)^n \frac{dt}{\sqrt{t}}. \tag{12}$$

Since for $n \ge 2$ the second derivative of $(1-t)^n$ is positive for all $0 \le t \le 1$, it is convex and thus lies above its tangent line at $t = 0$. Thus $(1-t)^n \ge 1 - nt$ for $0 \le t \le 1$. Consequently, if $0 \le nt \le 1/2$ we find $1 - nt \ge 1/2$ so $(1-t)^n \ge 1/2$. This estimate on the interval $0 \le t \le \frac{1}{2n}$ therefore gives the inequality

$$c_n \ge \int_0^{\frac{1}{2n}} (1-t)^n \frac{dt}{\sqrt{t}} \ge \frac{1}{2}\int_0^{\frac{1}{2n}} \frac{dt}{\sqrt{t}} = \frac{1}{\sqrt{2n}}. \tag{13}$$

Consequently, if $1 \ge |x| > \delta$, then from the definition (10)

$$0 \le \varphi_n(x) = \frac{(1-x^2)^n}{c_n} \le \sqrt{2n}(1-x^2)^n \le \sqrt{2n}(1-\delta^2)^n.$$

This has the form $\sqrt{2n}\, b^n$ where $0 < b < 1$. Thus

$$\text{For every } \delta > 0, \quad \lim_{n\to\infty} \int_{|x|>\delta} \varphi_n(x)\, dx = 0.$$

Thus we have verified the assumptions of Theorem 3, so our approximations $f_n(x)$, which are polynomials for $|x| \leq 1/2$, converge uniformly to $f(x)$.

EXAMPLE In the function space $L_2([0,1])$ the norm comes from an inner product

$$\langle f, g\rangle := \int_0^1 f(x)\overline{g(x)}\,dx \quad \text{so} \quad \|f\|_2 = \sqrt{\langle f, f\rangle}.$$

We say f and g are *orthogonal* if $\langle f, g\rangle = 0$. Assume that the continuous function f is orthogonal to 1, x, x^2, ..., so

$$\int_0^1 f(x)x^k\,dx = 0, \quad k = 0,1,2,\ldots.$$

We claim the only possibility is that $f(x) \equiv 0$ for all $x \in [0,1]$. In brief, this is because f is orthogonal to all polynomials p, but by the Weierstrass approximation theorem, polynomials are dense in $L_2([0,1])$ so f is essentially orthogonal to itself. Thus $f \equiv 0$. In greater detail, find a polynomial p so that $\|f - p\|_\infty < \varepsilon$ in $[0,1]$. Then $\langle f, p\rangle = 0$ so by the Schwarz inequality

$$\|f\|_2^2 = \langle f, f - p\rangle + \langle f, p\rangle \leq \|f\|_2\|f - p\|_2 \leq \varepsilon\|f\|_2.$$

Then $\|f\|_2 \leq \varepsilon$ for any $\varepsilon > 0$. This gives a contradiction.