Distance between functions

It is useful to re-cast some of the notions about convergence of functions in terms of a suitable notion of "distance" between functions. The basic idea is to base this on the difference between the values of the functions.

Maxima and minima of continuous functions

Given a continuous function f on an interval $[a, b]$, we want to show that there is a point w in $[a, b]$ with the property that $f(w) \geq f(x)$ for every x in the interval $[a, b]$.

Given a positive integer k, there is a partition $a = x_0 < x_1 < \cdots < x_n = b$ of the interval so that for x, y in the interval $[x_{i-1}, x_i]$, we have $|f(x) - f(y)| < 1/k$. It follows that

$$f(x) \leq f(x_i) + 1/k \text{ for } x \text{ in } [x_{i-1}, x_i]$$

Combining this for all i, we see that

$$f(x) \leq \max\{f(x_0), f(x_1), \ldots, f(x_n)\} + 1/k \text{ for } x \text{ in } [a, b]$$

Since we are taking the maximum of a *finite* set of numbers, this maximum is attained at (at least) one of the x_i's; let us denote this as w_k and let $z_k = f(w_k) + 1/k$.

Since $(w_k)_{k\geq 1}$ is a sequence of points in the interval $[a, b]$, it has a convergent subsequence $(w_{k_p})_{p\geq 1}$. For example, we know that $\limsup(w_k)_{k\geq 1}$ is the limit of a suitable subsequence of $(w_k)_{k\geq 1}$. We also note that the limit $w = \lim(w_{k_p})_{p\geq 1}$ lies in the interval $[a, b]$. Since the function f is continuous at w, it follows that $f(w) = \lim(f(w_{k_p}))_{p\geq 1}$. Now

$$\lim(z_{k_p})_{p\geq 1} = \lim(f(w_{k_p} + 1/k_p)_{p\geq 1} = \lim(f(w_{k_p}))_{p\geq 1} + \lim(1/k_p)_{p\geq 1} = f(w)$$

Given any x in the interval $[a, b]$ we have $f(x) \leq z_{k_p}$ for all $p \geq 1$. It follows that $f(x) \leq f(w)$. Hence, we have the required result.

The approach for the minimum of a continuous function is similar. We can also just observe that the minimum of f is the negative of the maximum of $-f$.

The norm and uniform convergence

We can now define the "magnitude" or the norm $\|f\|$ of a continuous function f on $[a, b]$ as the maximum value of $|f|$ on this interval. (Recall that we have earlier shown that $|f|$ is also a continuous function.) Occasionally, when we want to be explicit about the interval, we also write this as $\|f\|_{[a,b]}$.

The distance between two continuous functions f and g can now be prescribed as $\|f - g\|$. It is the maximum of the differences $|f(x) - g(x)|$ as x varies over $[a, b]$. With this notion of distance, the notions of uniform convergence can be re-written as follows.

A sequence of continuous functions $(f_n)_{n\geq 1}$ converges uniformly to a function f if, given any positive integer k, there is a positive integer n_k so that $\|f_n - f\| < 1/k$ for all $n \geq n_k$.

A sequence of continuous functions $(f_n)_{n\geq 1}$ satisfies the Cauchy criterion for convergence if, given any positive integer k, there is a positive integer n_k so that $\|f_n - f_m\| < 1/k$ for all $n, m \geq n_k$. In this case, (as shown in an earlier section) there is a continuous function f so that $(f_n)_{n\geq 1}$ converges to it.

It is worth noting that above two paragraphs are almost identical to the definitions given for convergence of a sequence of numbers except that distance between numbers has been replaced by the distance between functions.

Approximations

In an earlier section we saw that, given a continuous function f on $[a, b]$ and a positive integer k, we can find a piecewise linear continuous function g with the property that $|f(x) - g(x)| < 1/k$ for all x in the interval $[a, b]$. This can be re-interpreted as saying that $\|f - g\| < 1/k$.

We can say this in a different way. Given any continuous function f on $[a, b]$ which we want to approximate to a given accuracy $1/k$, there is a piecewise linear continuous function g which gives such an approximation. Note that g is likely to be much easier to calculate than f in general, especially on a computer.

Sometimes, we want a "formula" for the approximation. This is the content of the "Weierstrass approximation theorem" which we will merely state and not prove.

> Given any continuous function f on $[a, b]$ which we want to approximate to a given accuracy $1/k$, there is a *polynomial* function g on $[a, b]$ which gives such an approximation.

Though it may appear that such a "formula" is better than a piecewise linear approximation, the latter is often easier to calculate on a computer than a polynomial. However, this polynomial approximation is very useful in theory.

In the case of the interval $[0, 1]$ one can give a very explicit sequence of polynomials that converges to a given continuous function f. These are called the Bernstein polynomials:

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

By translating and scaling the interval $[a, b]$, this result can be extended to all intervals.

It is important to note that the coefficient of x^p in $f_n(x)$ will *not* converge in general. Hence, the Weierstrass approximation theorem should *not* be seen as a way to associate a power series to a continuous function. On the other hand, if the function *is* given by a power series, then the partial sums of the power series do give an approximation in the sense of Weierstrass' theorem.

Another point to note is that $f_n(k/n)$ can be different from $f(k/n)$. Hence, the polynomial $f_n(x)$ does not "interpolate" the given values of f. Even so, it can be seen to be a good approximation for large enough n.

Functionals

A functional is a mathematical object that associates a number to a function. In other words, it is a function on a space of functions. This may seem like a complicated idea but we will see some simple examples.

One way to associate a number to a function is to *evaluate* the function at some point. In other words, given c in the interval $[a, b]$, we can define ev_c on the space of continuous functions on $[a, b]$ as

$$\text{ev}_c(f) = f(c) \text{ for } f \text{ a continuous function on } [a, b]$$

If $(f_n)_{n\geq 1}$ converges to f in norm, then $\|f_n - f\|$ can be made as small as one likes by taking n sufficiently large. By the definition of $\|f_n - f\|$, it follows that $|f_n(c) - f(x)|$ can be made as small as one lines by taking n sufficiently large. In other words, we can see easily that $(f_n(c))_{n\geq 1}$ converges to $f(c)$. In other words, we have shown that if $(f_n)_{n\geq 1}$ converges to f then $(\text{ev}_c(f_n))_{n\geq 1}$ converges to $\text{ev}_c(f)$. This can be re-phrased as saying the ev_c is a *continuous* functional.

A different and equally important continuous functional is the one that associates to f the integral $I(f, [a, b])$ as defined in the previous section. We will examine this more closely in later sections.