Abstract

When a sequence of real numbers is convergent to some finite number, we may approximate the members of the sequence by its limit provided the subscript is large. But we may want a higher accuracy. If we know the speed of convergence, we define a derivative of the sequence at infinity. We also define the second derivative which enables us even better approximations.

Motivation.

When a program has to use a sequence, it is sometimes more handy to use an approximation than a table look-up the entries of which are members of a sequence. Here the meaning of the word approximation is technical because, typically, the approximation has to satisfy a pre-assigned accuracy.

Taylor expansions are polynomial functions students learn about in the beginning course of calculus. But, actually, polynomial approximations are used for fast computations. When we want to approximate a sequence we can hardly use algorithms that calculate best approximations on intervals. Our option would be the use of discrete approximation algorithms.

First of all, we discuss the question of a speed of convergence. If this is calculated, we can see how to use it for a linear approximation. The next step is to investigate the notion of the second derivative. The second derivative at some point may be defined if the function has the first derivative at each point of some nonempty open interval containing that point. This is not the case when we try to define it for a sequence.
Once the question is resolved, we may use a Taylor like expansion to improve our approximation.

We will repeatedly use the following Stolz theorem. Not only is it something like the L’Hospital rule for sequences, it is also very convenient for series.

**Theorem.** (Stolz) Let $a_N$ and $b_N$ be two sequences of real numbers. Assume that both $a_N$ and $b_N$ converge to zero, $b_N$ is strictly decreasing. If

$$\lim_{N \to 0} \frac{a_{N+1} - a_N}{b_{N+1} - b_N} = L$$

then

$$\lim_{N \to 0} \frac{a_N}{b_N} = L.$$  

The proof can be had at Wikipedia:


or just look for Stolz theorem. We will use only the case when $L$ is finite.

**Speed of convergence, example 1**

We all know that $(1 + 1/N)^N$ converges to $e$ as $N$ goes to infinity, but the question is how fast. To be able to compare the speed of convergence of one sequence with another sequence we have to develop a method of such a comparison. To do so, we recall the topic of acceleration of convergence of series in which we compare the remainder of one series to another series.

We compare the sequence $e - (1 + 1/N)^N$ with the sequence $C(N) = 1/N$. We can extend the domain of definition from natural numbers to real numbers by simply using the same formula. We define the limit as $D$, if it exists.

$$D = \lim_{N \to \infty} \frac{e - (1 + 1/N)^N}{1/N} = \lim_{x \to \infty} \frac{e - (1 + x^{-1})^x}{x^{-1}} = \lim_{x \to \infty} \frac{e - \exp(x \ln(1 + x^{-1}))}{x^{-1}}.$$  

We know that both $\lim_{x \to \infty} (e - \exp(x \ln(1 + x^{-1})))$ is zero and $\lim_{x \to \infty} x^{-1}$ is
zero and apply the L’Hospital’s rule. Thus

\[ D = \lim_{x \to \infty} -\exp(x \ln(1 + x^{-1})) \left( \ln(1 + x^{-1}) + \frac{x}{1 + x^{-1}}(-x^{-2}) \right) = \]

\[ e \lim_{x \to \infty} \frac{\ln(1 + x^{-1}) - \frac{x}{1 + x^{-1}}(x^{-2})}{-x^{-2}} = e \lim_{x \to \infty} \frac{\ln(1 + x^{-1}) - \frac{1}{1 + x}}{x^{-2}} = \]

\[ e \lim_{x \to \infty} \frac{\ln(1 + x^{-1}) - (1 + x)^{-1}}{x^{-2}} \]

Since \( \lim_{x \to \infty} (\ln(1 + x^{-1}) - (1 + x)^{-1}) = \ln 1 - 0 = 0 \) and so is the limit of the denominator, we use the L’Hospital rule the second time.

\[ e \lim_{x \to \infty} \frac{-x^{-2} + (1 + x)^{-2}}{-x^{-3}} = e \lim_{x \to \infty} \frac{-1}{x(1+x)^2} + \frac{1}{(1+x)^2} = \]

\[ e \lim_{x \to \infty} \frac{-1}{x(1+x)^2} + \frac{x}{x(1+x)^2} = e \lim_{x \to \infty} \frac{-1}{x(1+x)^2} = e \]

It is important that our guess was right, this limit is finite and the sequence \( e - (1 + 1/N)^N \) converges to zero as fast as \( 1/N \). Besides this fact, it is also reasonable to use the expression \( e - D/N = e - e/(2N) \) as an approximation to \( 1 + 1/N)^N \) for large \( N \).

**Series, example 2**

Now we use the series \( \sum_{i=0}^{\infty} 1/i! = e \). The function with respect to which we want to calculate the derivative is \( C(N) = 1/(N + 1)! \), thus using the Stolz theorem we get

\[ D = \lim_{N \to \infty} e - \frac{\sum_{i=1}^{N} 1/i!}{1/(N + 1)!} = \lim_{N \to \infty} -\frac{\sum_{i=1}^{N+1} 1/i! + \sum_{i=1}^{N} 1/i!}{(N + 2)!} - \frac{1}{(N + 1)!} = \]

\[ \lim_{N \to \infty} -\frac{1}{(N + 2)!} = \lim_{N \to \infty} \frac{1}{(N + 1)!} = \lim_{N \to \infty} \frac{N + 2}{N + 1} = 1 \]

To compare the speed of convergence of \( \sum_{i=1}^{N} 1/i! \), and \( (1 + 1/N)^N \), it will
suffice to form the ratio
\[ \frac{1/(N + 1)!}{1/N} = \frac{N}{(N + 1)!}. \]

Since this ratio goes to zero as \( N \) tends to infinity, the partial sums \( \sum_{i=1}^{N} 1/i! \) converge faster than \( (1 + 1/N)^N \). It is also more accurate to use the approximation \( e - 1/(N + 1)! \) to calculate partial sums for large \( N \).

**Second derivative**

We can imagine that \( D \) in our calculation may be interpreted as a derivative. To improve the accuracy of our approximations we would want to use the second derivative. Unfortunately, in the case of a series expansion, the derivative is defined at one point only while the definition of the second derivative requires the definition of the derivative on a whole neighborhood of the point at which we want to calculate it. We will show this is not an insurmountable obstacle.

We recall that \( C(N) \) is a comparison sequence if it is positive, decreasing, and converges to zero. A comparison sequence could be any sequence the members of which have the required properties but in practical examples we use those that are easy to calculate. That is why we used \( 1/N \) or \( 1/(N + 1)! \) in our examples.

The first attempt to redefine the second derivative when we look at it from the angle of approximation, is this:

**Definition.** Let \( F(N) \) be a sequence convergent to \( L \), let \( C(N) \) be a comparison sequence for the definition of a derivative \( D \), and let \( C'(N) \) be a comparison sequence for the definition of the second derivative. If the first derivative \( D = \lim_{N \to \infty} (L - F(N))/C(N) \) is nonzero and finite, we define the second derivative as
\[ S = 2 \lim_{N \to \infty} \frac{F(N) - (L + DC(N))}{C'(N)} \]
if this limit exists and is finite.
**Theorem.** Let \( f(x) \) be twice differentiable at a point \( x_0 = 0 \). Let \( f'(x) \) be defined, nonzero, and finite on a neighborhood of \( x_0 = 0 \). Let \( C(N) = 1/N, C'(N) = 1/N^2, L = \lim_{N \to \infty} F(N), \) and \( D = \lim_{N \to \infty} (L - F(N))/C(N) \). Then

\[
2 \lim_{N \to \infty} \frac{f(N) - (L + DC(N))}{C'(N)} = f''(0).
\]

**Proof.** We substitute \( h = 1/N \) and see that

\[
D = \lim_{N \to \infty} \frac{F(N) - L}{1/N} = \lim_{h \to 0} \frac{f(h) - f(0)}{h} = f'(0)
\]

Since \( f(x) \) is continuous at 0, we have \( \lim_{h \to 0} (f(h) - f(0)) = 0 \) and obviously \( \lim_{h \to 0} f'(0)h = 0 \). Thus the L’Hospital rule for type 0/0 may be applied to the following

\[
2 \lim_{N \to \infty} \frac{F(N) - (L + DC(N))}{C'(N)} = 2 \lim_{h \to 0} \frac{f(h) - (f(0) + f'(0)h)}{h^2}
\]

yielding

\[
2 \lim_{h \to 0} \frac{f'(h) - f'(0)}{2h} = f''(0).
\]

**Note.** We have only shown that our definition of the second derivative of a sequence is nothing but a generalization of the usual definition of the second derivative of a function. The proof also shows why a constant 2 was put in front of the symbol of a limit.

Another way to define the second derivative is the use of interpolation. When we know the limit \( L \) and derivative \( D \), we can use the linear interpolation \( F(N) \approx L + D \times C(N) \). But we want to use the quadratic interpolation \( ax^2 + bx + c \) in which \( c = L, b = D \) and \( a \) is to be determined so that \( L + D \times C(N) + \)
\(a \times C'(N) = F(N)\). We solve the equation for \(a\) as

\[
a = \frac{F(N) - (L + D \times C(N))}{C'(N)}
\]

But the second derivative of a quadratic polynomial \(ax^2 + bx + c\) is \(2a\) and we are ready to define the second derivative as

\[
2 \lim_{N \to \infty} \frac{F(N) - (L + D \times C(N))}{C'(N)}.
\]

This is obviously the same expression as the one in our previous definition of the second derivative.

We may try to calculate the second derivative in example 1 as:

\[
2 \times \lim_{x \to \infty} \frac{e - (1 + 1/x)^x - e/(2x)}{x^{-2}} = -11e/12
\]

so the approximate formula is \((1 + 1/N)^N \approx e - e/(2N) + 11e/(12N^2)\). This is only an exercise, the question of speed of exact calculations of \(\exp(N \times \ln(1 + 1/N))\) is not that essential.

Now we calculate the second derivative in example 2:

\[
2 \times \lim_{N \to \infty} \frac{e - \sum_{i=1}^{N} 1/i! - 1/(N+1)!}{1/(N(N+1)!)} =
\]

We use the Stolz theorem and obtain

\[
2 \times \lim_{N \to \infty} \frac{-\frac{1}{(N+1)!} - \frac{1}{(N+2)!} + \frac{1}{N+1}}{\frac{1}{N(N+1)!}} =
\]

\[
2 \times \lim_{N \to \infty} \frac{-\frac{1}{(N+2)!} + \frac{N+2}{N(N+2)!}}{\frac{1}{(N+1)!}} =
\]

\[
2 \times \lim_{N \to \infty} \frac{-\frac{1}{(N+2)!} \left( \frac{1}{(N+1)!} - \frac{N+2}{N} \right)}{\frac{1}{(N+1)!}} = 2
\]
We can see that $C'(N) = 1/(N(N+1)!)$ and that it is not a square of $C(N) = 1/(N+1)!$. It follows that, when we use the second derivative, the approximating formula will be

$$\sum_{i=1}^{N} \frac{1}{i!} \approx e - \frac{1}{(N+1)!} + \frac{1}{N(N+1)!}.$$ 

**Poisson distribution**

The probability mass function of the Poisson distribution $X$, with parameter $\lambda > 0$, is defined as

$$P(X = i) = e^{-\lambda} \lambda^i / i!$$

for $i = 0, 1, 2, \ldots$, the distribution function is

$$e^{-\lambda} \sum_{i=0}^{N} \lambda^i / i!$$

To calculate the limit

$$D = \lim_{N \to \infty} e^{\lambda} \sum_{i=0}^{N} \frac{\lambda^i}{i!}$$

we use the Stolz theorem, so

$$D = \lim_{N \to \infty} \frac{-\lambda^{N+1}}{(N+1)!} = \lim_{N \to \infty} -\frac{1}{(N+1)!} = \lim_{N \to \infty} \frac{-1}{N+2} = 1$$

The second derivative is

$$D = \lim_{N \to \infty} e^\lambda \frac{\lambda^{N+1}}{(N+1)!} - \sum_{i=0}^{N} \frac{\lambda^i}{i!}$$

$$D = \lim_{N \to \infty} \frac{-\lambda^{N+2}}{(N+2)!} + \frac{\lambda^{N+1}}{(N+1)!} - \frac{\lambda^{N+1}}{(N+1)!} = \lim_{N \to \infty} \frac{-\lambda^{N+2}}{N+2} = \lim_{N \to \infty} \frac{-N \lambda}{N+2} = -\lambda$$

We have thus developed the approximation

$$\sum_{i=0}^{N} \frac{\lambda^i}{i!} \approx e^\lambda - \frac{\lambda^{N+1}}{(N+1)!} + \frac{\lambda^{N+2}}{N(N+1)!}$$

**Riemann zeta**

We define $\zeta(s) = \sum_{i=1}^{\infty} \frac{1}{i^s}$ for $s > 1$ It is well known that the series is
convergent for such $s$ and $\zeta(s)$ is well defined. We consider only natural $s$ for the sake of simplicity. Our aim is the study of the speed of convergence of partial sums:

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} \frac{1}{i} - \zeta(s)}{N^t}$$

We are going to use the Stolz theorem for type $0/0$. We use an unknown parameter $t$, its value will be determined later.

$$\lim_{N \to \infty} \frac{1}{(N+1)^t - N^t} = \lim_{N \to \infty} \frac{1}{N^t - (N + 1)^t} =$$

$$\lim_{N \to \infty} \frac{N^t (N + 1)^{t-s}}{N^t - (N + 1)^t}$$

We get a finite nonzero limit if $2t - s = t - 1$. In this case when $t = s - 1$ the limit is $-(s - 1)^{-1}$. We can now write an approximation of the partial sum as

$$\sum_{i=1}^{N} \frac{1}{i^s} \approx \zeta(s) - \frac{1}{(s - 1) N^{s-1}}$$

Now we write down the second derivative and use the Stolz theorem $0/0$.

$$\lim_{N \to \infty} \frac{\sum_{i=1}^{N} \frac{1}{i^s} - (\zeta(s) - \frac{1}{(s-1) N^{s-1}})}{\frac{1}{N^s}} = \lim_{N \to \infty} \frac{\sum_{i=1}^{N} \frac{1}{i^s} - \zeta(s) + \frac{1}{(s-1) N^{s-1}}}{\frac{1}{N^s}} =$$

$$\lim_{N \to \infty} \frac{\frac{1}{(N+1)^s} + \frac{1}{(s-1)(N+1)^{s-1}} - \frac{1}{(s-1) N^{s-1}}}{\frac{1}{(N+1)^s} - \frac{1}{N^s}} = \lim_{N \to \infty} \frac{(s-1) N^{s-1} - (N+1)^s}{(s-1) N^{s-1} (N+1)^s - (s-1) N^{s-1} (N+1)^{s-1}} =$$

$$\lim_{N \to \infty} \frac{(s-1) N^{s} + (N+1) N^{s-1} - N(N+1)^s}{(s-1) N^{s} (N+1)^s - N(N+1)^s}.$$
if we take \( t = s \).

\[
\sum_{i=1}^{N} \frac{1}{i^s} \approx \zeta(s) - \frac{1}{(s-1)N^{s-1}} + \frac{1}{2N^s}
\]

We take a look at a simple example when \( s = 2 \), \( N = 20 \); then the partial sum is \( \sum_{i=1}^{20} 1/i^2 = 1.596163 \). The value of the zeta function is \( \zeta(2) = 1.644934 \), the linear approximations is \( \zeta(s) - \frac{1}{(s-1)N^{s-1}} = \zeta(2) - \frac{1}{20} = 1.594934 \), and the quadratic approximation is \( \zeta(s) - \frac{1}{(s-1)N^{s-1}} + \frac{1}{2N^s} = \zeta(2) - \frac{1}{20} + \frac{1}{2 \times 20^2} = 1.596184 \). Thus the error of the quadratic approximation is \( |1.596163 - 1.596184| = 0.000021 \).

**Derangement**

A derangement is a permutation of the elements of a set such that no element appears in its original position. A derangement is a permutation with no fixed points. There are several ways to count the number of derangements leading to the formula

\[
N! \sum_{i=1}^{N} \frac{(-1)^i}{i!}
\]

where \( N \) is the number of elements in the set. When we ask for the probability that a permutation is a derangement, we divide by the number \( N! \) of all permutations of \( N \) elements and obtain

\[
P(N) = \sum_{i=1}^{N} \frac{(-1)^i}{i!}
\]

We recall that the power series expansion for \( e^x \) is

\[
e^x = \sum_{i=1}^{\infty} \frac{x^i}{i!}
\]

and we get

\[
\lim_{N \to \infty} P(N) = e^{-1} = 1/e.
\]

The series for \( e^{-1} \) is an alternating series causing a little trouble but it can be quickly handled by considering even number of terms in partial sums. It means
we study the partial sums from one to $2N$, such as

\[
\lim_{N \to \infty} \frac{1}{2N} \sum_{i=1}^{2N} \frac{(-1)^i}{i!} = \lim_{N \to \infty} \frac{1}{2N} \sum_{i=1}^{2N+2} \frac{(-1)^i}{i!} - \sum_{i=1}^{2N} \frac{(-1)^i}{i!}
\]

\[
\lim_{N \to \infty} \frac{-1}{(2N+2)!} + \frac{1}{(2N+1)!} = \lim_{N \to \infty} \frac{-1 + 2N + 2}{(2N+2)!}
\]

\[
\lim_{N \to \infty} \frac{2N+1}{(2N+2)!} = \lim_{N \to \infty} \frac{2N+1}{4N^3 + 10N^2 + 7N + 2} = \frac{1}{2}
\]

We can now use the formula

\[
\sum_{i=1}^{2N} \frac{(-1)^i}{i!} \approx \frac{1}{e} - \frac{1}{2N(2N)!}
\]

or, if $M$ is even, then

\[
\sum_{i=1}^{M} \frac{(-1)^i}{i!} \approx \frac{1}{e} - \frac{1}{M \times M!}
\]

As a numerical example we try $M = 2N = 10$. We get the exact probability of a derangement $P(10) = 0.3678792$ from the first principle. The linear approximation is $1/e - 1/(M \times M!) = 1/e - 1/(10 \times 10!) = 0.3678763$ while $1/e = 0.3678794$ and, clearly, the approximation of $P(10)$ by $1/e$ provides a sufficient accuracy. This happens because the number $1/(10 \times 10!) = 2.8e - 8$ is so small that its subtraction makes no practical difference.

**Hilbert space.**

We will not review the definition of a Hilbert space. We will only recall the definition of the $l_p$ space.

**Definition.** The space of infinite sequences $x = (x_1, x_2, \ldots)$ for which $\sum_{i=1}^{\infty} x_i^2$
is convergent is called an $l_2$ space.

We call a sequence $C(N)$ a comparison sequence if it is positive, decreasing, and converges to zero. In our examples those were $1/N$ or $1/(N + 1)!$, but we chose a comparison sequence in such a way that its members of which are easy to calculate.

**Definition.** Let $C(N)$ be a comparison sequence. A subset of $l_2$ the elements of which are those sequences $x = (x_1, x_2, \ldots)$ for which

$$\lim_{N \to \infty} \left( S_x - \sum_{i=1}^{N} x_i^2 \right)/C(N) < \infty,$$

where $S_x = \sum_{i=1}^{\infty} x_i^2$, is called a $C$ (convergence) speed subset.

**Note:** It is obvious that $S_x = \sum_{i=1}^{\infty} x_i^2 \geq \sum_{i=1}^{N} x_i^2$ and there is no need to use the absolute value of the limit.

As an example we take the comparison sequence $C(N) = 1/N$. The $1/N$ speed sequences are those, for which

$$\lim_{N \to \infty} N(S_x - \sum_{i=1}^{N} x_i^2) < \infty,$$

where $S_x = \sum_{i=1}^{\infty} x_i^2$.

**Theorem.**

Let $C(N)$ be a comparison function. Then a $C$ speed subset of the $l_2$ space is a linear subset of $l_2$.

**Proof.** The element $x = (0, 0, \ldots)$ is obviously a zero element of $l_2$. If $a$ is a constant and $x = (x_1, x_2, \ldots)$ is any element of the $C$-speed subset, then so is $ax = (ax_1, ax_2, \ldots)$ because

$$\lim_{N \to \infty} (S_{ax} - \sum_{i=1}^{N} a^2 x_i^2)/C(N) = a^2 \lim_{N \to \infty} (S_x - \sum_{i=1}^{N} x_i^2)/C(N) < \infty,$$
where $S_x = \sum_{i=1}^{\infty} x_i^2$ and $S_{ax} = \sum_{i=1}^{\infty} a^2 x_i^2$, thus $S_{ax} = a^2 S_x$.

We skip the proof of all the other properties of the linear subspace because they are as easy as this one.

**Theorem.**

Let $C(N)$ be any comparison function. Then a $C$ speed subset of the $l_2$ space is a dense linear subset of $l_2$.

**Proof.** Let $x = (x_1, x_2, \ldots)$ be any element of the $l_2$ space. Let us define the sequence $y_K = (x_1, x_2, \ldots, x_K, 0, 0, \ldots)$ of elements of $l_2$. Obviously, if $N > K$, we have, by definition, $S_{y_K} = \sum_{i=1}^{K} x_i^2 + \sum_{i=K+1}^{\infty} 0 = \sum_{i=1}^{K} x_i^2$. Thus, for any comparison function $C(N)$, we have for some $K > 0$

$$\lim_{N \to \infty} \frac{(S_{y_K} - \sum_{i=1}^{N} x_i^2)}{C(N)} = 0.$$ 

Let $N_0 = K$. Then for $N > N_0$ we have

$$\lim_{N \to \infty} \frac{(S_{y_K} - \sum_{i=1}^{N} x_i^2)}{C(N)} = \lim_{N \to \infty} \frac{(S_{y_K} - \sum_{i=1}^{K} x_i^2)}{C(N)} = 0$$

because $S_{y_K} - \sum_{i=1}^{K} x_i^2 = 0$ for $N > N_0$.

**Conclusion**

So far, we could use only a numerical evidence both to check if the formulas are correct and also to see if they are accurate enough for practical calculations to replace the actual calculations of sums. Numerical evidence may be sufficient for numerical calculation but it is not a rigorous proof of anything. We may recall what we studied in the topic of Taylor expansion in the first semester of calculus to see that the formula for a remainder is exactly what is missing in our paper. This is one open question.

Another open question is the definition of higher derivatives of a sequence at infinity. For example, if there are infinitely many of these derivatives, we
can form something like an expansion for such a sequence that would obviously characterize the behavior of a sequence at infinity. The question is what is the meaning of such an expansion.

**Acknowledgement**

I am grateful to Vixra for publishing my paper. Just try to imagine what all those official real scientists would say. This is unpredictable.