# A Tutorial on Taylor Series

Corrections to Dr Ian Rudy (http://people.ds.cam.ac.uk/iar1/contact.html) please.

A Taylor Series is a way of approximating some function of interest $f(x)$ with a polynomial. This technique is extremely useful, often because the polynomial is easier to analyse than the original function, particularly when we are interested in small changes in $f$ caused by small changes in $x$.

Rather than aim to approximate the function for all $x$, we concentrate on approximating it near to some specific value of $x$, say $x = a$. It is possible to work out roughly what we need to do with a geometric argument. Imagine first that we know the value of the function at $a$ (ie $f(a)$) but nothing else. In that case, our best guess of the value of the function at some point $a+x$, which is near $a$, would be

$$f(a+x) = f(a) \qquad [1]$$

This is probably not a good approximation: it would work only if the function is flat or nearly flat. Now let's imagine that we also know the value of the gradient of the function at $a$ (ie $f'(a)$). Then we could make a better approximation using:

$$f(a+x) = f(a) + xf'(a) \qquad [2]$$

The new term on the right hand side is the change in the function if we extrapolate the function from $a$ to $a+x$, assuming it to be a straight line. (Remember that the gradient is defined as the change in $f$ divided by the change in $x$, so, rearranging, we find that the change in $f$ is just the gradient $f'(a)$ multiplied by the change in $x$ in going from $a$ to $a+x$.)

[2] is probably a better approximation than [1], but still neglects the possibility that the gradient of the function might change in going from $a$ to $a+x$. This possibility would be suggested by $f''(a)$ being non-zero, and although we cannot now make a simple geometric argument to prove it, it turns out that we can take account of the gradient changing by adding another term to the right hand side:

$$f(a+x) = f(a) + xf'(a) + \frac{x^2 f''(a)}{2} \qquad [3]$$

This is probably a better approximation than [2].

You can see that what is forming here is a polynomial, in $x$, and you would be right if you suspected that we can carry on adding further terms, gradually obtaining a better and better approximation to the value of the function at $a+x$. However, by this point the geometrical logic has ceased to be useful to us, so we will use a different way of working out the later terms in the series. Let's just assume it *is* possible to approximate the function at $x$ using a polynomial:

$$f(a+x) = k_0 + k_1 x + k_2 x^2 + k_3 x^3 + \ldots \qquad [4]$$

Where $k_0, k_1, k_2 \ldots$ are unknown coefficients, which we need to find.

We can work out the value of $k_0$ very quickly, by choosing *x* to be equal to 0. This gives:

$$k_0 = f(a) \qquad [5]$$

We can find the value of $k_1$ by differentiating [4] with respect to *x*, to get:

$$f'(a+x) = k_1 + 2k_2 x + 3k_3 x^2 + \ldots \qquad [6]$$

and then, choosing *x* to be equal to 0, we get:

$$k_1 = f'(a) \qquad [7]$$

In similar fashion, if we differentiate [6] with respect to *x*, and choose *x* to be equal to 0, we get:

$$k_2 = f''(a)/2 \qquad [8]$$

and, repeating the process:

$$k_3 = f'''(a)/3! \qquad [9]$$

You can probably now see the pattern that is forming:

$$k_n = f^{(n)}(a)/n! \qquad [10]$$

Putting [5], [7], [9] and [10] into [4], we get:

$$f(a+x) = f(a) + xf'(a) + \frac{x^2 f''(a)}{2} + \frac{x^3 f'''(a)}{3!} + \ldots$$

$$\ldots + \frac{x^n f^{(n)}(a)}{n!} + \ldots \qquad [11]$$

This not only confirms the polynomial approximation we worked out with our geometrical logic in [2], but extends it to as many terms as we care to take.

Equation [11] is the formula for a Taylor Series. We talk of it being the Taylor Series of $f(x)$ *around* $x = a$. Notice this definition - if you are asked to find a Taylor Series around $x = 3$, the value 3 is that of $a$, not *x*. This might seem confusing at first, but it's just the way the terminology is defined, and you will soon get used to it.

In general, the approximation will get better and better as we take more terms, but if $x$ is very close to 0, then the higher order terms get very small very quickly, and so a decent approximation can be made with just a few terms. In fact, often we need only the first two or three terms on the right hand side in order to get a good approximation, if $x$ is very close to 0.

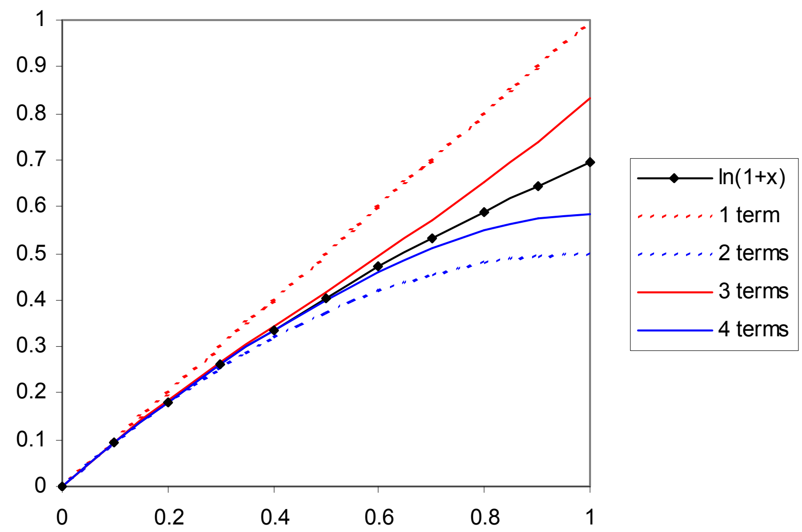
Here is an example of how to construct a Taylor Series. We will find the series for ln(1+$x$) around $x$=0. I'd recommend you create two columns, one listing the function and its derivatives, and the other listing the values of the derivatives at $x = a$:

$$\begin{aligned} f(x) &= \ln(1+x) & f(a) &= 0 \\ f'(x) &= (1+x)^{-1} & f'(a) &= 1 \\ f''(x) &= -(1+x)^{-2} & f''(a) &= -1 \\ f'''(x) &= 2(1+x)^{-3} & f'''(a) &= 2 \\ f^{(4)}(x) &= -3!(1+x)^{-4} & f^{(4)}(a) &= -3! \end{aligned}$$

And so from [11]:

$$f(a+x) = \ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \qquad [12]$$

The graph below shows how this Taylor Series improves as we take more terms in the series:




By far the most common value of $a$ is zero, as in the above example. In such cases, the resulting series is often known as a *Maclaurin* Series, though it is just a special case of a Taylor Series.

It turns out that there is only one Taylor Series for a given function about a given point. For this reason, people often use the phrases "Taylor Series" and "power series" interchangeably, because there is no ambiguity as to which series one is referring. The uniqueness of the series for a given function about a given point means that it does not matter how you find it: the differentiation method above is one method, but sometimes there are quicker alternatives. (We will discuss these in supervisions.)

Finally, we should briefly mention the issue of the error term. As the graph above demonstrates, if one cuts off the series after a certain number of terms (say, $n$ terms), there will be an error. To a first approximation, the error is just the $(n+1)^{th}$ term in the series, if $x$ is very close to 0, because the terms after that will all be very small in comparison. More precisely, it can be shown that the error is exactly equal to the $(n+1)^{th}$ term, if the derivative of $f$ is evaluated at some point between $a$ and $a+x$. We do not actually know which value in this range to use, but we may be able to get an upper limit on the error in this way.