Taylor Expansion

Asaf Pe'er[1]

October 16, 2015

1. Introduction

Taylor expansion is among the most fundamental tools of numerical analysis. It is extremely important and useful tool in very many branches of physics, math and engineering. Thus, personally, I believe that it is an absolutely essential tool to master. In fact, this is exactly how your computer works - from your pocket calculator to supercomputers, when you ask it to do some calculation. Fortunately, Taylor expansion is also simple to understand, and it (used to be) taught already in high school. I bring here the most general derivation of it, as is done in calculus courses and I strongly urge the students to practice it.

1.1. Preliminary: L'Hôpital's rule

Near the end of the proof of Taylor expansion formula, I will have to use the following theorem, known as **"L'Hôpital's rule"**. Similar to Taylor expansion, this is a very strong, yet simple rule that should be familiar to you from your calculus course, if not from high school. I will only bring here the rule, without proving it; hopefully, you have proved it in your calculus course, otherwise, you are welcome to consult any standard calculus textbook. L'Hôpital's rule has several versions, and I will bring two - the basic form, and a second form which we will use.

The basic form: Consider 2 functions, $f(x)$ and $g(x)$, both have derivatives around an arbitrary point a, maybe except at the point a itself [1]. Denote by $f'(x)$ and $g'(x)$ their (first) derivative with respect to x. We assume that the functions fulfill the following:
(1) $g'(x) \neq 0$ for any $x \neq a$.
(2) The functions fulfill the equality

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0. \tag{1}$$

[1]Physics Dep., University College Cork

[1]if you don't fully understand the meaning of this sentence, think of the function $f(x) = 1/x$ around $x = 0$.

Then, if the limit

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{2}$$

exist, then also the limit $\lim_{x\to a} \frac{f(x)}{g(x)}$ exist, and

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}. \tag{3}$$

L'Hôpital's rule is a very useful tool when evaluating limits of indetermined functions.

Example. Try to evaluate:

$$\lim_{x\to 0} \frac{\sin(x)}{x} \tag{4}$$

Direct evaluation leads to 0/0, which is not defined. However, using L'Hôpital's rule (Equation 3), one immediately finds

$$\lim_{x\to 0} \frac{\sin(x)}{x} = \lim_{x\to 0} \frac{\cos(x)}{1} = 1 \tag{5}$$

which is the correct answer. (**Try it !** use you pocket calculator to evaluate $\sin(x)/x$ for $x = 0.1, 0.01, 0.001, ...$)

A slightly different form of L'Hôpital's rule: Let $f(x)$ and $g(x)$ two functions which are defined and have derivatives in the region $[a, b]$, and that fulfill:
(1) $f(a) = g(a) = 0$;
(2) $g'(a) \neq 0$.

Then

$$\lim_{x\to a,\ x>a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}. \tag{6}$$

I did not bring here the full proof of L'Hôpital's rule, which can be found in any standard calculus textbook.

2. Taylor expansion of a polynom

Let us now put L'Hôpital's rule aside (we will use it later), and turn our attention to polynoms. Consider a general polynom of the form

$$P(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3 + ... + a_n \cdot x^n \tag{7}$$

where the coefficients a_0, a_1, ... , a_n are constants (some numbers).

The first derivative of the polynom is

$$P'(x) \equiv \frac{dP(x)}{dx} = a_1 + 2 \cdot a_2 \cdot x + 3 \cdot a_3 \cdot x^2 + ... + n \cdot a_n \cdot x^{n-1} \tag{8}$$

Its second derivative is

$$P''(x) \equiv \frac{d^2P(x)}{dx^2} = 2 \cdot a_2 + 3 \cdot 2 \cdot a_3 \cdot x + ... + n \cdot (n-1) \cdot a_n \cdot x^{n-2} \tag{9}$$

Its third derivative is

$$\frac{d^3P(x)}{dx^3} = 3 \cdot 2 \cdot a_3 + 4 \cdot 3 \cdot 2 \cdot a_4 \cdot x + ... + n \cdot (n-1) \cdot (n-2) \cdot a_n \cdot x^{n-3} \tag{10}$$

and so on; its n-derivative is

$$\frac{d^nP(x)}{dx^n} = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 1 \cdot a_n \tag{11}$$

If we now set $x = 0$ in all the above equations, we can write the constants a_0, a_1, ... , a_n as

$$\begin{array}{rcl} a_0 & = & P(x=0) \equiv P(0); \\ a_1 & = & P'(0); \\ a_2 & = & \frac{1}{2}P''(0); \\ a_3 & = & \frac{1}{3 \cdot 2 \cdot 1}P'''(0); \\ & ... & \\ a_n & = & \frac{1}{n!}\frac{d^nP(0)}{dx^n}. \end{array} \tag{12}$$

We can use this in the polynomial Equation (Equation 7), to write the polynom $P(x)$ for every x, using its n-derivatives at $x = 0$:

$$P(x) = P(0) + P'(0) \cdot x + \frac{1}{2}P''(0) \cdot x^2 + ... + \frac{1}{n!}\frac{d^nP(0)}{dx^n} \cdot x^n \tag{13}$$

Equation 13 is known as **"MacLauren formula"** of a polynom.

Let us now write $x = x_0 + h$. This is useful in cases where, e.g., we know the value of the polynom at a certain point x_0 (not necessarily 0), and we are interested in knowing its value at a different point, x. Thus, we treat x_0 as fixed, while h is can vary. Use this in Equation 7, we write

$$P(x) = P(x_0 + h) = a_0 + a_1 \cdot (x_0 + h) + a_2 \cdot (x_0 + h)^2 + ... + a_n \cdot (x_0 + h)^n \tag{14}$$

The right hand side of Equation 14 contains n binoms, which all look like $(x_0 + h)^k$, with $k = 0, 1, ..., n$. We can expand all the binoms, and order what we get in powers of h. The result is that we can write our polynom $q(h) \equiv P(x_0 + h)$ in the form

$$q(h) = b_0 + b_1 \cdot h + b_2 \cdot h^2 + ... + b_n \cdot h^n. \tag{15}$$

While we can calculate of course the values of the constants b_0, b_1, b_2 ... b_n from the values of the original numbers a_0, a_1, ... a_n, in fact we don't need to do it; this is because we know from the above analysis, that we can write these numbers as

$$b_k = \frac{1}{k!}\frac{d^k q(h=0)}{dh^k} \tag{16}$$

and since $q(h) \equiv P(x_0 + h)$ and x_0 is assumed fixed, we can write

$$b_k = \frac{1}{k!}\frac{d^k q(h=0)}{dh^k} = \frac{1}{k!}\frac{d^k P(x=x_0)}{dx^k} \tag{17}$$

We can therefore evaluate the polynom at $x = x_0 + h$ as

$$P(x_0+h) = P(x_0) + P'(x=x_0)\cdot h + \frac{1}{2}P''(x_0)\cdot h^2 + ... + \frac{1}{n!}\cdot\frac{d^n P(x=x_0)}{dx^n}\cdot h^n. \tag{18}$$

Alternatively, using $h = x - x_0$ we can write Equation 18 in the form

$$P(x) = P(x_0) + P'(x_0)\cdot(x-x_0) + \frac{1}{2}P''(x_0)\cdot(x-x_0)^2 + ... + \frac{1}{n!}\frac{d^n P(x=x_0)}{dx^n}\cdot(x-x_0)^n \tag{19}$$

Equations 18, 19 are called **"Taylor expansion of the polynom $P(x)$ around the point x_0"**. Its importance lie on the fact that if we know the value of the polynom - and its derivatives - at some point x_0, we can find its value at any other point x by simple summation. The closer the point x is to x_0, the more quicker we get convergence - we can get good estimate of the value of the polynom even if we don't do all the derivatives, provided that $x - x_0$ is small enough, since the contribution of each term gets smaller and smaller.

3. Taylor expansion of an arbitrary function

We are now going to use the above analysis to estimate the value of any arbitrary function $f(x)$, at arbitrary point x, by knowing the values of this function and its derivatives at another (arbitrary) point, x_0. We require that at the point x_0, this function has n derivatives (namely, the function $f(x)$ is "well behaved" near the point x_0).

We begin by writing a polynom:

$$P_n(x) = f(x_0) + \frac{1}{1!}\cdot f'(x_0)\cdot(x-x_0) + ... + \frac{1}{n!}\frac{d^n f(x_0)}{dx^n}\cdot(x-x_0)^n \tag{20}$$

This polynom looks similar to the polynom we derived above, in Equation 19. If we compare term by term the coefficients in Equations 20 to that of Equation 19 we can write:

$$\frac{d^k P(x_0)}{dx^k} = \frac{d^k f(x_0)}{dx^k}, \tag{21}$$

for all values of $k = 0, 1, ..., n$.

Shortly, we are going to claim that the polynom $P_n(x)$ is a good approximation to the value of $f(x)$ at the point x. However, in order to do it properly, we need first to approximate the error that we introduce. This is done by looking at the residual:

$$R_n(x) = f(x) - P_n(x) \tag{22}$$

which can be evaluated at any point x, not necessarily x_0. If we calculate the residual at x_0, we can use Equation 21 to deduce that

$$\frac{d^k R_n(x_0)}{dx^k} = 0, \tag{23}$$

for $k = 0, 1, ..., n$. Namely, the first n derivatives of the residual, evaluated at x_0 are 0.

Let us now look at the $k-$th derivative of the function $(x - x_0)^n$ $(k = 0, 1, ..., n)$. It is

$$\frac{d^k[(x - x_0)^n]}{dx^k} = n \cdot (n - 1) \cdot ... \cdot (n - k + 1) \cdot (x - x_0)^{n-k} \tag{24}$$

If we evaluate this function at $x = x_0$, we get:

$$\left.\frac{d^k[(x - x_0)^n]}{dx^k}\right|_{x=x_0} = 0 \tag{25}$$

for $k = 0, 1, ..., n - 1$, and

$$\left.\frac{d^n[(x - x_0)^n]}{dx^n}\right|_{x=x_0} = n! \tag{26}$$

(for $k = n$).

Now is the point that we use L'Hôpital's rule to evaluate the residual $R_n(x)$ in Equation 22:

$$\lim_{x\to x_0} \frac{R_n(x)}{(x - x_0)^n} = \lim_{x\to x_0} \frac{\frac{d^n R_n(x)}{dx^n}}{n!} = 0 \tag{27}$$

Let us now denote

$$\alpha \equiv \alpha(x) \equiv \frac{R_n(x)}{(x - x_0)^n}. \tag{28}$$

From Equation 22 and 20 we can write

$$f(x) = P_n(x)+R_n(x) = f(x_0)+\frac{1}{1!}\cdot f'(x_0)\cdot(x-x_0)+...+\frac{1}{n!}\frac{d^n f(x_0)}{dx^n}\cdot(x-x_0)^n+\alpha(x)\cdot(x-x_0)^n \tag{29}$$

where we know that $\alpha(x)$ fulfills: $\lim_{x\to x_0} \alpha(x) = 0$. Equation 29 is known as **Taylor expansion of the function $f(x)$ around the point x_0, up to the n-th derivative**.

The function $R_n(x)$ is called **the n-th residual of** $f(x)$, and we saw that it is of the order $\mathcal{O}((x-x_0)^n)$.

A Taylor expansion of a function is **unique** (again, I am not bringing you the full proof).

An alternative way of writing a Taylor expansion is by writing $\Delta x \equiv x - x_0$. Putting in Equation 29 gives

$$f(x_0+\Delta x) = f(x_0) + \frac{1}{1!}\cdot f'(x_0)\cdot(\Delta x) + ... + \frac{1}{n!}\frac{d^n f(x_0)}{dx^n}\cdot(\Delta x)^n + \mathcal{O}((\Delta x)^n) \qquad (30)$$

4. Examples

Let us give a few examples, to demonstrate the importance and usefulness of a Taylor expansion.

4.1. The exponential function

Calculate the function e^x (e is the natural exponent).

Answer. If we write $f(x) = e^x$, we know that all of its derivatives are similar, $f'(x) = f''(x) = ... = \frac{d^n f(x)}{dx^n} = e^x$. Furthermore, we know that $f(0) = e^0 = 1$.

We can now use Taylor expansion (Equation 29) to evaluate e^x to whatever accuracy we are interested in. For example, if we take the 5th order, we get:

$$f(x) = e^x = 1 + \frac{1}{1!}\cdot x + \frac{1}{2!}\cdot x^2 + \frac{1}{3!}\cdot x^3 + \frac{1}{4!}\cdot x^4 + \frac{1}{5!}\cdot x^5 + ... \qquad (31)$$

and we know that the error we introduce is not larger than $x^5/5!$. If we set $x = 1$, we get

$$e = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + ... = \frac{163}{60} = 2.717... \qquad (32)$$

and the error is no larger than $1/120 = 0.0083$. As I stated, this is exactly how your computer calculates e^x when you ask it to do so.

4.2. Trigonometric function

Let us evaluate $\sin(x)$.

Answer. We write $f(x) = \sin(x)$. We know: $f'(x) = \cos(x)$; $f''(x) = -\sin(x)$; $f'''(x) = -\cos(x)$; $f^{(4)}(x) = \sin(x)$; $f^{(5)}(x) = \cos(x)$; and so on.

Around $x = 0$, we get: $f(0) = 0$; $f'(0) = 1$; $f''(0) = 0$; $f'''(0) = -1$; $f^{(4)}(0) = 0$; $f^{(5)}(0) = 1$; and so on.

Use these results in Taylor expansion formula (Equation 29), we get:

$$\begin{aligned} \sin(x) &= f(x) = 0 + \tfrac{1}{1!} \cdot 1 \cdot x + \tfrac{1}{2!} \cdot 0 \cdot x^2 + \tfrac{1}{3!} \cdot (-1) \cdot x^3 + \tfrac{1}{4!} \cdot 0 \cdot x^4 + \tfrac{1}{5!} \cdot 1 \cdot x^5 + \mathcal{O}(x^5) \\ &= x - \tfrac{x^3}{3!} + \tfrac{x^5}{5!} + \mathcal{O}(x^5) \end{aligned} \tag{33}$$

Clearly, as x approaches 0, the approximation gets better and better. For very small angles, often one can take only the first term in the Taylor expansion and approximate $\sin(\theta) \approx \theta$, as the error is $\mathcal{O}(\theta^3)$. **Please go ahead and check this formula for various values of x, by using your pocket calculator !!**

Exercise. Evaluate $\cos(x)$ and $\tan(x)$ using Taylor expansion around $x = 0$.

4.3. square root

Let us evaluuate $\sqrt{1+x}$, where $x \ll 1$.

Answer. We have: $f(x) = \sqrt{x}$. Thus $f'(x) = \frac{1}{2\sqrt{x}}$, $f''(x) = -\frac{1}{4x^{3/2}}$, $f'''(x) = \frac{3}{8x^{5/2}}$ etc.

Easiest: around $x = 1$, we have $f(1) = 1$; $f'(1) = \frac{1}{2}$; $f''(1) = -\frac{1}{4}$; $f'''(1) = \frac{3}{8}$; etc.

We thus have:

$$\begin{aligned} \sqrt{1+x} &= 1 + \tfrac{1}{1!} \cdot \tfrac{1}{2} \cdot x + \tfrac{1}{2!} \cdot \left(-\tfrac{1}{4}\right) \cdot x^2 + \tfrac{1}{3!} \cdot \left(\tfrac{3}{8}\right) \cdot x^3 + ... \\ &= 1 + \tfrac{x}{2} - \tfrac{1}{8}x^2 + \tfrac{1}{16}x^3 + ... \end{aligned} \tag{34}$$

(Note that this result is correct for $|x| < 1$, otherwise the sum does not converge).

Tables of Taylor expansion of basic functions can be found in any standard mathematical handbook, as well as on the web.

5. Proof of L'Hopital's rule.

OK. I am kind enough to bring you the mathematical proof of L'Hopital's rule.

Let us assume that all the condition of the theorem are fullfilled in the upper vicinity of the point a: namely, assume the existence of $h > 0$ such that $g'(x) \neq 0$ for all $a < x \leq a + h$.

Without loss of generality, we can assume that $g(a) = f(a) = 0$ (we work with nicely behaved functions; otherwise, we can change their values at a). Both $g(x)$ and $f(x)$ are continuous in the range $[a, a+h]$. Furthermore, the first derivative of $g(x)$ exists inside this range.

This implies that there is no point x, such that $a < x \leq a+h$ in which $g(x) = 0$. The reason is, that if such a point would have existed, then there would exist a second point, c, such that $a < c < x$ in which $g'(c) = 0$, which contradicts the first condition (namely, $g'(x) \neq 0$ for any $x \neq a$). (This is known as Rolle's theorem, and can be understood very easily using a plot).

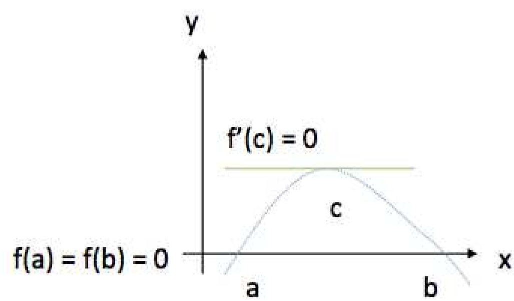

Fig. 1.— Demonstration of Rolle's theorem.

Thus, we are forced to conclude that $g(x) \neq 0$ for any x in the range $a < x <\leq a+h$. This means, that for any such x we choose, there is a point c, $a < c < x$ such that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(c)}{g'(c)} \tag{35}$$

(this is known as Cauchy's mean value theorem).

When x approaches a, so does c, and so we have

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{c \to a} \frac{f'(c)}{g'(c)}. \tag{36}$$

The same is true if we were to consider the lower vicinity of a.