The General Binomial Theorem: Part 1

Introduction

One of most basic algebraic formulas which a student encounters in high school curriculum is the following

$$(a+b)^2 = a^2 + 2ab + b^2$$

and its variant for $(a-b)^2$. And after many exercises and problems later one encounters another formula of similar nature namely

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

and one wonders if there are similar formulas for higher powers of $(a+b)$.

This optimism is amply rewarded and further down the mathematical curriculum one encounters a general formula for positive integral powers of $(a+b)$ and the result is important enough to deserve a name: *The Binomial Theorem*. The Binomial Theorem states that if n is a positive integer then

$$(a+b)^n = a^n + \binom{n}{1}a^{n-1}b + \cdots + \binom{n}{r}a^{n-r}b^r + \cdots + \binom{n}{n-1}ab^{n-1} + b^n \quad (1)$$

and the right hand side above is sometimes written compactly as

$$\sum_{r=0}^{n} \binom{n}{r}a^{n-r}b^r$$

The symbol $\binom{n}{r}$ is called a *binomial coefficient* and is defined for all numbers n and non-negative integers r by

$$\binom{n}{0} = 1, \quad \binom{n}{r} = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!} \text{ for } r \geq 1 \quad (2)$$

The binomial theorem has an easy proof based on induction and it is purely an algebraic result.

Our topic of discussion today is an extension of the above result for the case when n is not necessarily a positive integer. In that case the result is known by the name *The General Binomial Theorem* or *Binomial Theorem for General Index* and it transcends the powers of algebra and belongs more properly to the field of mathematical analysis. We turn to this powerful result next.

The General Binomial Theorem

The general binomial theorem does not try to deal with the expression $(a+b)^n$ for all values of n, a, b rather it is a conditional result and is presented in the following form:

If x, n are real numbers with $|x| < 1$ then

$$(1 + x)^n = 1 + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots = \sum_{k=0}^{\infty} \binom{n}{k}x^k \tag{3}$$

Thus the formula for $(1 + x)^n$ is no longer a finite expression but rather an infinite series and we have a condition that $|x| < 1$. The series involved is called the binomial series and it is absolutely convergent when $|x| < 1$ and divergent when $|x| > 1$. We will talk later about its behavior when $|x| = 1$. The proof of the above formula is difficult and it belongs to *the infamous category of theorems whose proofs lie beyond the scope of the book/syllabus.*

In order to prove the general binomial theorem we need two results from differential calculus:

- $\lim\limits_{x \to a} \dfrac{x^n - a^n}{x - a} = na^{n-1}$ for real n, a and $a > 0$. This result is already established as one of the standard limits in an earlier post.

- **Taylor's Theorem with Cauchy's Form of Remainder**: This we discuss next along with its proof.

Taylor's Theorem with Cauchy's Form of Remainder

We have already encountered a version of Taylor's theorem with Peano's form of remainder in an earlier post (see another proof available on MSE). Here we need a stronger version of Taylor's theorem and undoubtedly it needs stronger hypotheses to remain valid. We state the theorem below:

Taylor's Theorem: *Let n be a positive integer. If $f(x)$ is a function such that $f^{(n-1)}(x)$ is continuous in $[a, a + h]$ and $f^{(n)}(x)$ exists in $(a, a + h)$ then*

$$f(a + h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \cdots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(a) + R_n \tag{4}$$

where

$$R_n = \frac{(1 - \theta)^{n-p}h^n f^{(n)}(a + \theta h)}{p(n-1)!} \tag{5}$$

for some number $\theta \in (0, 1)$ and any chosen integer p with $1 \le p \le n$.

The proof is based on Rolle's Theorem. Let us put $b = a + h$ and define a function F by

$$F(x) = f(b) - f(x) - (b - x)f'(x) - \cdots - \frac{(b - x)^{n-1}}{(n-1)!}f^{(n-1)}(x)$$

and another function G by

$$G(x) = F(x) - \left(\frac{b - x}{b - a}\right)^p F(a)$$

where p is some integer between 1 and n. Clearly $G(x)$ is continuous on $[a, b]$ and differentiable on (a, b) and we have $G(a) = G(b) = 0$. Therefore by Rolle's Theorem there is a value $c \in (a, b)$ such that $G'(c) = 0$. Moreover since $c \in (a, b)$ we can write $c = a + \theta(b - a)$ for some $\theta \in (0, 1)$ and thus $c = a + \theta h$. The derivative $G'(x)$ is given by

$$G'(x) = \frac{p(b-x)^{p-1}}{(b-a)^p} F(a) - \frac{(b-x)^{n-1}}{(n-1)!} f^{(n)}(x)$$

and $G'(c) = 0$ implies that

$$\frac{p}{(b-a)^p} F(a) = \frac{(b-c)^{n-p}}{(n-1)!} f^{(n)}(c)$$

or

$$F(a) = \frac{(1-\theta)^{n-p} h^n f^{(n)}(a+\theta h)}{p(n-1)!} = R_n$$

Noting the value of $F(a)$ we see that the proof of the theorem is complete.

The term R_n is said to be the remainder after n terms in the Taylor's series expansion for $f(a+h)$. When $p = n$ we obtain the Lagrange's form of Remainder namely

$$R_n = \frac{h^n}{n!} f^{(n)}(a+\theta h) \qquad \text{(Lagrange)}$$

and if $p = 1$ then we get Cauchy's form of Remainder

$$R_n = \frac{(1-\theta)^{n-1} h^n}{(n-1)!} f^{(n)}(a+\theta h) \qquad \text{(Cauchy)}$$

Note that the statement of Taylor's theorem and the proof above assumes that $h > 0$ but it is easily seen that it holds even if $h < 0$. A particular instance of this theorem is used to find infinite series expansion of certain functions. If we put $a = 0$ and $h = x$ in the Taylor's theorem we obtain the following result which goes by the name of *Maclaurin's series*:

$$f(x) = f(0) = xf'(0) + \frac{x^2}{2!} f''(0) + \cdots + \frac{x^{n-1}}{(n-1)!} f^{(n-1)}(0) + R_n \qquad (6)$$

where

$$R_n = \frac{(1-\theta)^{n-p} x^n f^{(n)}(\theta x)}{p(n-1)!}$$

for some $\theta \in (0,1)$ and p an integer between 1 and n.

As can be seen the value of R_n depends on θ as well x and p. Also note that the value of θ itself is based on x. Normally for p we use one of the two choices mentioned above (i.e. use Lagrange's or Cauchy's form of remainder). Let's then write $R_n(x)$ for R_n assuming that a reasonable choice of p has been made. If for certain values of x we can ensure that $R_n(x) \to 0$ as $n \to \infty$ then from equation (6) we obtain an infinite series for $f(x)$ as

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!} f''(0) + \cdots + \frac{x^n}{n!} f^{(n)}(0) + \cdots \qquad (7)$$

and this is valid for all those values of x for which $R_n(x) \to 0$ as $n \to \infty$. In practice we usually show that for values of x in a certain range and any $\theta \in (0,1)$ the expression

$R_n(x) \to 0$ as $n \to \infty$ and thus obtain the above Maclaurin's series for $f(x)$.

The easiest examples of such series are obtained for $f(x) = \sin x, \cos x, e^x$ as follows:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots \tag{8a}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots \tag{8b}$$

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots \tag{8c}$$

and in each case it is easily proven that $R_n(x) \to 0$ as $n \to \infty$ for all real values of x. Therefore the above series expansions are valid for all real values of x.

Proof of The General Binomial Theorem

Now it is time to apply Taylor's theorem on function $f(x) = (1 + x)^n$ where n, x are real numbers and $x > -1$ so that $1 + x > 0$. Clearly the limit formula

$$\lim_{x \to a} \frac{x^n - a^n}{x - a} = na^{n-1} \tag{9}$$

mentioned above implies that the derivative of $g(x) = x^n$ is $g'(x) = nx^{n-1}$ for real n and $x > 0$. Therefore $f'(x) = n(1 + x)^{n-1}$ and it is easily seen that $f^{(m)}(x)$ exists for all positive integers m and $x > -1$. Moreover we have

$$f^{(m)}(0) = n(n-1)(n-2) \cdots (n-m+1) = m! \binom{n}{m}$$

and by Taylor's theorem we have

$$f(x) = f(0) + xf'(0) + \cdots + \frac{x^{m-1}}{(m-1)!} f^{(m-1)}(0) + R_m(x)$$

or

$$(1 + x)^n = 1 + \binom{n}{1} x + \cdots + \binom{n}{m-1} x^{m-1} + R_m(x) \tag{10}$$

where we use the Cauchy's form of remainder

$$R_m(x) = \frac{(1 - \theta)^{m-1} x^m f^{(m)}(\theta x)}{(m-1)!}$$

$$= \frac{n(n-1) \cdots (n-m+1)}{(m-1)!} \cdot \frac{(1-\theta)^{m-1} x^m}{(1+\theta x)^{m-n}}$$

If $|x| < 1$ then it is easily checked that the expression $\dfrac{1 - \theta}{1 + \theta x}$ lies in $(0, 1)$ for all $\theta \in (0, 1)$. Further $(1 + \theta x)^{n-1}$ is less than $(1 + |x|)^{n-1}$ if $n > 1$ and it is less than $(1 - |x|)^{n-1}$ if $n < 1$. Hence

$$|R_m(x)| < |n|(1 \pm |x|)^{n-1} \left| \binom{n-1}{m-1} \right| |x|^m = r_m \tag{11}$$

Clearly we can see that

$$\lim_{m \to \infty} \frac{r_{m+1}}{r_m} = |x| < 1$$

therefore $r_m \to 0$ as $m \to \infty$. It follows from equation (11) that $R_m(x) \to 0$ as $m \to \infty$ for all real values of n and x with $|x| < 1$. Therefore taking limits as $m \to \infty$ in equation (10) we get the general binomial theorem

$$(1 + x)^n = 1 + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots = \sum_{m=0}^{\infty} \binom{n}{m} x^m$$

for all real n, x with $|x| < 1$.

Note that our proof of the binomial theorem is based on the derivative formula $(x^n)' = nx^{n-1}$ which in turn is based on the limit formula (9) and therefore one should not use binomial theorem in proving the derivative formula for x^n. However many calculus textbooks perform a kind of intellectual fraud by presenting the following proof of derivative formula for x^n:

$$\begin{aligned}
(x^n)' &= \lim_{h \to 0} \frac{(x + h)^n - x^n}{h} \\
&= \lim_{h \to 0} \frac{x^n(1 + (h/x))^n - x^n}{h} \\
&= \lim_{h \to 0} \frac{x^n\left(1 + \dfrac{nh}{x} + \dfrac{n(n-1)h^2}{2x^2} + \cdots\right) - x^n}{h} \\
&= \lim_{h \to 0} nx^{n-1} + \frac{n(n-1)}{2}hx^{n-2} + \cdots \\
&= nx^{n-1}
\end{aligned}$$

The use of general binomial theorem is justified because as $h \to 0$ we can ensure that $|h/x| < 1$. The real problem lies in the circular nature of the proof above because the binomial theorem itself is proved using this derivative formula. Note that if n is a positive integer then it is possible to use the binomial theorem to express $(x + h)^n$ as a finite sum and then obtain the same result.

We can salvage the above calculation of derivative of x^n if we can somehow establish the binomial theorem for general index without the use of derivatives. Surprisingly it is possible to prove the general binomial theorem without using theorems of differential calculus and we will have a look at it in the next post.

By Paramanand Singh

Wednesday, July 20, 2016

Labels: Mathematical Analysis