Binomial expansion

Jake

September 2024

1 Introduction

In this paper a proof of the binomial theorem will be found

2 $(x+1)^n$ as a polynomial

For the purposes of this section a polynomial will be defined as an expression that can be written in the form $\sum_{r=0}^{n} a_r x^r$ where n is the order of the polynomial and $a_0, a_1, \cdots, a_{n-1}, a_n$ is a sequence of constants with $a_n \neq 0$. By expanding $(x+1)^n$ for small n it can be seen that it seems to always give a polynomial of order n. Since this result is intuitive the proof will be by induction.

$$\text{Assume } (x+1)^k = \sum_{r=0}^{k} a_r x^r$$

$$\text{Where } a_0, a_1, \cdots, a_{n-1}, a_n \text{ is a sequence of constants}$$

$$\text{Then } (x+1)^{k+1} = (x+1)(x+1)^k$$

$$= (x+1)\sum_{r=0}^{k} a_r x^r$$

$$= x\sum_{r=0}^{k} a_r x^r + \sum_{r=0}^{k} a_r x^r$$

$$= \sum_{r=0}^{k} a_r x^{r+1} + \sum_{r=0}^{k} a_r x^r$$

$$
\begin{aligned}
&= \sum_{r=1}^{k+1} a_{r-1}x^r + \sum_{r=0}^{k} a_r x^r \\
&= \left[\sum_{r=1}^{k} a_{r-1}x^r + \sum_{r=1}^{k} a_r x^r\right] + a_0 + a_k x^{k+1} \\
&= \left[\sum_{r=1}^{k} [a_{r-1} + a_r]x^r\right] + a_0 + a_k x^{k+1}
\end{aligned}
$$

To Simplify further we will define a new sequence of constant, $b_0, b_2, \cdots, b_{k+1}$, and we will construct this sequence from the previous one using the following three rules:

- $b_0 = a_0$
- $b_i = a_{i-1} + a_i \quad \text{For} \quad 0 < i < k+1$
- $b_{k+1} + a_k$

$$
\begin{aligned}
(x+1)^{k+1} &= \left[\sum_{r=1}^{k} [a_{r-1} + a_r]x^r\right] + a_0 + a_k x^{k+1} \\
&= \left[\sum_{r=1}^{k} b_r x^r\right] + b_0 + b_{k+1}x^{k+1} \\
&= \sum_{r=0}^{k+1} b_r x^r
\end{aligned}
$$

It is clear that $(x+1)^1$ is a polynomial of order one so the assumption is true for $k = 1$ with $a_0 = a_1 = 1$. Since the assumption has been shown to be true for $k+1$ if it is true for k and is true for one, it must then be true for all natural numbers i.e.

$$
(x+1)^n = \sum_{r=0}^{n} a_r x^r \quad \text{For} \quad n \in \mathbb{N}
$$

Recurrence relation

Our interest now lies with the constant coefficients that come for the sequence $a_0, a_1, \cdots, a_{n-1}, a_n$. It is worth taking time to consider what we did when defining the second sequence in the above proof, by doing this we gave the rules of a recurrence relation of sorts and by checking $(x+1)^1$ we found its initial conditions. The reason I say 'of sorts' is because a typical recurrence relation give the next term in a sequence as a function of previous terms but the one we have here gives a whole new sequence from the terms of the previous one. For simplicity I will call this a two dimensional recurrence relation, each term in a standard or 'one dimensional' recurrence relation can be written as a function of one variable, its position in the sequence, but term is a 2D relation must be given as a function of two variable, position in the sequence and which sequence it is in. In the above proof only a single value was used in the subscript of the terms in the sequence which is why when creating the next iteration we had to change from a to b to indicate the other variable changing. A better description would be to call it $f(n, r)$ where n is the sequence and r is the position in the sequence, so from above $a_r = f(k, r)$ and $b_r = f(k+1, r)$. With this new definition we can rewrite the rules and initial conditions from before.

- $f(k+1, 0) = f(k, 0)$
- $f(k+1, i) = f(k, i-1) + f(k, i) \quad \text{For} \quad 0 < i < k+1$
- $f(k+1, k+1) = f(k, k)$
- $f(1, 0) = f(1, 1) = 1$

Finally, by using the initial conditions to simplify the first and third rules as well as changing the variables to be the ones I want we get,

- $f(n, r) = f(n-1, r-1) + f(n-1, r) \quad \text{For} \quad 0 < r < n$
- $f(n, n) = f(n, 0) = 1$

I don't know how to solve this recurrence relation but it was important to describe it properly like this because we can now find the coefficients a different way and use that solution to solve this relation, and knowing the solution to this will help later on. It is also useful because we see this same recurrence relation in pascals triangle which will be shown later.

3 nth derivatives

We want to find nice expressions for the rth derivatives of $(x+1)^n$ and x^n where $r \leq n$, so we will consider the function $f(x) = (x+c)^n$ where c is some constant. By experimenting we see,

$$\begin{aligned}
f(x) &= (x+c)^n \\
f'(x) &= n(x+c)^{n-1} \\
f''(x) &= n(n-1)(x+c)^{n-2} \\
&\vdots \\
f^{(r)}(x) &= \frac{n!}{(n-r)!}(x+c)^{n-r}
\end{aligned}$$

This is obvious from intuition so I will prove proof by induction.

$$\begin{aligned}
\text{Assume } f^{(k)}(x) &= \frac{n!}{(n-k)!}(x+c)^{n-k} \\
\text{Then } f^{(k+1)}(x) &= \frac{d}{dx}(f^{(k)}(x)) \\
&= \frac{d}{dx}\left(\frac{n!}{(n-k)!}(x+c)^{n-k}\right) \\
&= (n-k)\frac{n!}{(n-k)!}(x+c)^{n-k-1} \\
&= \frac{n!}{(n-(k+1))!}(x+c)^{n-(k+1)}
\end{aligned}$$

$$\text{Hence, by induction, } ((x+c)^n)^{(r)} = \frac{n!}{(n-r)!}(x+c)^{n-r} \quad \text{For} \quad r \leq n$$

By considering $c=0$ and $c=1$, this formula becomes what we were looking for.

4 Finding the coefficients

Here we will find the values of the coefficients in the expansion of $(x+1)^n$ by considering the pth derivative of the function where $p \leq n$. We will do this by expressing the function as a polynomial and splitting the polynomial into three part, one for terms of an order less than p, one for the term of order p and a third for the terms with an order greater than p. Then we will show that upon taking the pth derivative the first section is reduced to zero, the second becomes constant and the third has a factor of x in each term. The by considering the pth derivative at zero we can isolate the constant and find the coefficient.

$$\text{Let } f(x) = (x+1)^n$$

$$f(x) = \sum_{r=0}^{n} a_r x^r$$

$$f^{(p)}(x) = \left(\sum_{r=0}^{n} a_r x^r \right)^{(p)}$$

$$= \left[\sum_{r=0}^{p-1} a_r (x^r)^{(p)} \right] + a_p (x^p)^{(p)} + \left[\sum_{r=p+1}^{n} a_r (x^r)^{(p)} \right]$$

$$= \left[\sum_{r=0}^{p-1} a_r \left((x^r)^{(r)} \right)^{(p-r)} \right] + (a_p p!) + \left[\sum_{r=p+1}^{n} a_r \frac{r!}{(r-p)!} x^{r-p} \right]$$

$$= \left[\sum_{r=0}^{p-1} a_r \, (r!)^{(p-r)} \right] + (a_p p!) + x \left[\sum_{r=p+1}^{n} a_r \frac{r!}{(r-p)!} x^{r-(p+1)} \right]$$

$$= a_p p! + x \left[\sum_{r=p+1}^{n} a_r \frac{r!}{(r-p)!} x^{r-(p+1)} \right]$$

$$\implies f^{(p)}(0) = a_p p!$$

$$a_p = \frac{f^{(p)}(0)}{p!}$$

$$= \frac{\frac{n!}{(n-r)!}(0+1)^{n-r}}{p!}$$

$$= \frac{n!}{p!(n-p)!}$$

$$\therefore (x+1)^n = \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} x^r$$

5 generalising

Our previous result implies that,

$$(g(x)+1)^n = \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} g^r(x)$$

for some function $g(x)$. From this it is reasonable to then wonder if,

$$(g(x)+h(x))^n = \sum_{r=0}^{n} v_r g^r(x)$$

Where $h(x)$ is another function and $v_0, v_1, \cdots, v_{n-1}, v_n$ is a sequence of functions of x. Since this assumption is reasonable, I will prove it by induction.

$$\text{Assume } (g(x)+h(x))^k = \sum_{r=0}^{k} v_r g^r(x)$$

$$\text{Where } v_0, v_1, \cdots, v_k \text{ are functions of } x$$

$$\text{Then } (g(x)+h(x))^{k+1} = (g(x)+h(x)) \left[(g(x)+h(x))^k\right]$$

$$= (g(x)+h(x)) \left[\sum_{r=0}^{k} v_r g^r(x)\right]$$

$$= g(x) \left[\sum_{r=0}^{k} v_r g^r(x)\right] + h(x) \left[\sum_{r=0}^{k} v_r g^r(x)\right]$$

$$= \left[\sum_{r=0}^{k} v_r g^{r+1}(x)\right] + \left[\sum_{r=0}^{k} h(x) v_r g^r(x)\right]$$

$$= \left[\sum_{r=1}^{k+1} v_{r-1} g^r(x)\right] + \left[\sum_{r=1}^{k} h(x) v_r g^r(x)\right] + h(x) v_0$$

$$= \left[\sum_{r=1}^{k} v_{r-1} g^r(x)\right] + v_k g^{k+1}(x) + \left[\sum_{r=1}^{k} h(x) v_r g^r(x)\right] + h(x) v_0$$

$$= \left[\sum_{r=1}^{k} [v_{r-1} + h(x)v_r]g^r(x)\right] + h(x)v_0 + v_k g^{k+1}(x)$$

To Simplify further we will define a new sequence of functions of x, $u_0, u_2, \cdots, u_{k+1}$, and we will construct this sequence from the previous one using the following rules:

- $u_0 = h(x)v_0$
- $u_i = v_{i-1} + h(x)v_i \quad \text{For} \quad 0 < i < k+1$
- $u_{k+1} = v_k$

With the rule for this new sequence defined we can continue the simplification

$$\begin{aligned}
(g(x) + h(x))^{k+1} &= \left[\sum_{r=1}^{k} [v_{r-1} + h(x)v_r]g^r(x)\right] + h(x)v_0 + v_k g^{k+1}(x) \\
&= \left[\sum_{r=1}^{k} u_r g^r(x)\right] + u_0 + u_{k+1} g^{k+1}(x) \\
&= \sum_{r=0}^{k+1} u_r g^r(x)
\end{aligned}$$

But note that $(g(x) + h(x))^1 = h(x) + g(x)$ hence our initial assumption is true for $k = 1$ with $v_0 = h(x)$ and $v_1 = 1$, but since it is true for one and it has been shown that if it is true for k it must be true for $k+1$ we can conclude by induction that our assumption holds for all natural numbers, i.e.

$$(g(x) + h(x))^n = \sum_{r=0}^{n} v_r g^r(x) \quad \text{For} \quad n \in \mathbb{N}$$

Another recurrence relation

This is reminiscent of the corresponding proof for $(x+1)^n$ being a polynomial and just as before we can use the rules to set up a recurrence relation and try solving it to find the values of the sequence of functions. By following the same process as before we arrive at,

- $f(n,r) = f(n-1,r-1) + h(x)f(n-1,r) \quad \text{For} \quad 0 < r < n$
- $f(n,0) = h^n(x)$
- $f(n,n) = 1$

We can use the solution to the recurrence relation from earlier as a test solution but we know that it won't work because the presents of the $h(x)$ in the new relation. To account for this we can add a factor of $F(n,r)$ which we can then look for, So we try $\frac{n!}{r!(n-r)!}F(n,r)$

$$\frac{n!}{r!(n-r)!}F(n,r) = \frac{(n-1)!}{(r-1)!(n-r)!}F(n-1,r-1) + h(x)\frac{(n-1)!}{r!(n-1-r)!}F(n-1,r)$$

$$\frac{n!}{(n-r)!}F(n,r) = \frac{(r)(n-1)!}{(n-r)!}F(n-1,r-1) + H(x)\frac{(n-1)!}{(n-1-r)!}F(n-1,r)$$
$$n!F(n,r) = (r)(n-1)!F(n-1,r-1) + h(x)(n-r)(n-1)!F(n-1,r)$$
$$nF(n,r) = (r)F(n-1,r-1) + h(x)(n-r)F(n-1,r)$$

we know that $F(n,r)$ will have $h(x)$ in it so we can try $F(n,r) = h^{M(n,r)}(x)$

$$nh^{M(n,r)}(x) = rh^{M(n-1,r-1)}(x) + nh^{M(n-1,r)+1}(x) - rh^{M(n-1,r)+1}(x)$$
$$n(h^{M(n,r)}(x) - h^{M(n-1,r)+1}(x)) = r(h^{M(n-1,r-1)}(x) - h^{M(n-1,r)+1}(x))$$

So if we can find a function of n and r such that $M(n,r) = M(n-1,r)+1$ and $M(n-1,r-1) = M(n-1,r)+1$ then it will work. By thinking about it $M(n,r) = n-r$ works and satisfies the other requirements hence, $f(n,r) = \frac{n!}{r!(n-r)!}h^{n-r}(x)$. Finally we conclude that,

$$(g(x)+h(x))^n = \sum_{r=0}^{n} \frac{n!}{r!(n-r)!}h^{n-r}(x)g^r(x) \quad \text{For} \quad n \in \mathbb{N}$$

6 Conclusion

I will be the first to admit that the final part of the proof is both unsatisfying, I spent quite a bit of time playing with these 2d recurrence relations and found some interesting stuff, I even found a more satisfying proof of the solution to the final recurrence relation that didn't require guessing the answer but the proof is messy and would fit better in to a paper of recurrence relations in general, a paper that I now have a great desire to start work on. I am held back however, by my poor knowledge of these relations and would need to take some time to learn some more as well as work on it myself, the proof given here will do for now but there is more to come.

I am aware that there is a connection between recurrence relations and differential equations which leads me to wonder what the corresponding concept from differential equations could be for recurrence relations of functions of more than one variable. Despite my limited knowledge, I believe it is the case that multi variable functions are what give rise to partial differential equations so this is one possible candidate for it, if this is true it may be some time before I can understand 2d recurrence relations.