Math 171 Notes: Power sums

Reid Huntsinger (rhuntsingerccp@gmail.com)

April 4, 2019

Here is an approach to determining formulas for sums of powers of integers. Sections 2 and 3 below, on integer-valued polynomials and Bernoulli polynomials, concern related issues. Wikipedia has fairly good entries on everything here; look up "Bernoulli polynomial" for example.

1 The sum of the first n squares

Let's start with the formula for the sum of the first n squares. Let's write

$$s(n) = 1^2 + 2^2 + 3^2 + \cdots + n^2$$

and we'll try to determine a "closed-form" formula for s (that is, one that does not involve a number of terms depending on the input).

Since the definition of s does not make sense for n other than a positive integer, we can't use calculus directly. However, we can use the difference quotient with spacing 1, called in this context the *discrete derivative* or *forward difference*. We'll see that enough of the basic ideas work that we can use this to show that s is a certain polynomial of degree 3.

For any function $f : \mathbf{N} \to \mathbf{R}$ (i.e., sequence) define $\triangle f$ to be the function

$$(\triangle f)(n) = f(n+1) - f(n)$$

Then $\triangle f$ has the same domain as f (so it's also a sequence).

Now $(\triangle s)(n) = (n+1)^2$ by definition of s; all but the last term of the sum for $s(n+1)$

gets cancelled by the corresponding term in $s(n)$. That is, for example,

$$\begin{aligned}(\triangle s)(6) &= s(7)-s(6)\\ &= 1^2+2^2+3^2+4^2+5^2+6^2+7^2\\ &\quad -(1^2+2^2+3^2+4^2+5^2+6^2)\\ &= 7^2=49\end{aligned}$$

We can prove a *sum rule* and *constant-multiple rule* for $\triangle$ easily, so $\triangle$ is linear. We've seen how the product rule works for the difference quotient[1], also; we get

$$\triangle(fg)(n)=f(n+1)(\triangle g)(n)+(\triangle f)(n)g(n).$$

This is slightly more complicated than the product rule for derivatives; that's the cost of not taking a limit.

So how does the "power rule" work? It's just the product rule used repeatedly, so it might be a mess, but let's see[2]:

$$\triangle(n)=n+1-n=1$$

OK, no problem there! How about the second power?

$$\triangle(n^2)=(n+1)^2-n^2=2n+1$$

Hmmm, a little messy, but not a problem yet. The third power?

$$\triangle(n^3)=(n+1)^3-n^3=3n^2+3n+1$$

OK, you can see that Pascal's Triangle is going to be involved here. We can write

$$\begin{aligned}\triangle(n^k) &= (n+1)^k-n^k\\ &= \binom{k}{k-1}n^{k-1}+\binom{k}{k-2}n^{k-2}+\cdots+\binom{k}{1}n+1\end{aligned}$$

[1]Notes on "rules for derivatives"

[2]In case the formula which follows looks confusing, here is the explanation for the notation: we are using n to represent the sequence $1,2,3,4,\ldots$ which as a function is the identity function. So $\triangle(n)$ is a function (sequence) $2-1, 3-2, 4-3,\ldots$ which is the constant sequence $1,1,1,1,\ldots$. We should probably write $\triangle(n)(k)=k+1-k=1$ but this might be even more confusing...

which is a polynomial in n of degree $k-1$ with positive integer coefficients.

Now by linearity, we can make a linear combination f of power functions (power sequences?) having $\triangle f = (n+1)^2 = n^2+2n+1$. Looking at the powers up to 3 we've computed, we can write

$$f(n) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$

because

$$\begin{aligned}\triangle(\frac{1}{3}n^3) &= n^2+n+\frac{1}{3}\\ \triangle(\frac{1}{2}n^2) &= n+\frac{1}{2}\\ \triangle(\frac{1}{6}n) &= \frac{1}{6}\end{aligned}$$

So $\triangle f = \triangle s$.

Now we need something like the Mean Value Theorem to let us say that if $\triangle(f) = 0$ then f must be constant. But that's pretty clear here: $f(n+1) - f(n) = 0$ means $f(n+1) = f(n)$. So each term of the sequence $f(n)$ is equal to the next, so they're *all* equal to $f(0)$, which is to say f is constant.

By linearity, we can get the same statement as with the derivative:

Basic fact *If $\triangle f = \triangle g$ then f and g must differ by a constant.*

Now we almost have our formula: $s(n) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n + c$ and we just need to determine the constant c. To do so, we can evaluate at any n. You might think $n = 0$ would be cheating, so let's take $n = 1$. That gives $s(1) = \frac{1}{3} + \frac{1}{2} + \frac{1}{6} + c$ but $s(1) = 1$ from the definition so $c = 0$.

So that's our formula:

$$\sum_{k=0}^{n} k^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$

You can factor this as $\frac{1}{6}n(2n^2+3n+1) = \frac{1}{6}n(n+1)(2n+1)$ and this is how it's usually written.

Exercise Use this method to find a polynomial expression for $s_3(n) = \sum_{k=1}^{n} k^3$.

Exercise Use the product rule for $\triangle$ to find $\triangle(n^2)$ and $\triangle(n^3)$. Do you get the same functions as we did above? Show that you do, or explain why you don't.

2 Integer-valued polynomials

The polynomial $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$ is interesting in that it is integer-valued when n is an integer, yet does not have integer coefficients. The factored form $\frac{1}{6}n(n+1)(2n+1)$ shows why: either n or $n+1$ must be even, and either $n, n+1$ or $2n+1$ must be a multiple of 3. So the numerator is always a multiple of 6.

There are other integer-valued polynomials we are familiar with. The binomial coefficients are integers, and written like

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$$

we can think of this as a polynomial in n of degree k. To be specific, we let $\binom{x}{k}$ denote the *polynomial* whose factored form is $\frac{1}{k!}x(x-1)\cdots(x-k+1)$.

It turns out that *any* integer-valued polynomial is actually an integer linear combination of these polynomials.

You can see this with the help of the forward-difference operator (extended to deal with functions $\mathbf{R} \to \mathbf{R}$). By definition of $\triangle$,

$$\left[\triangle \binom{\bullet}{k}\right](n) = \binom{n+1}{k} - \binom{n}{k}$$

Now Pascal's Triangle has row $n+1$ made up by adding two consecutive entries in row n. If you work out the indexing, you can see that this says

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$$

So we get

$$\left[\triangle \binom{\bullet}{k}\right](n) = \binom{n}{k-1}$$

Here is how we can use this. If f is an integer-valued polynomial of degree k, then $\triangle f$ is an integer-valued polynomial of degree $k-1$. If we can express $\triangle f$ as an integer linear combination of the binomial coefficient polynomials of degree up to $k-1$, then *the same* linear combination of binomial coefficients of degree one higher differs from f by a constant. The constant must be an integer, so we're done.

Now any degree zero integer-valued polynomial is an integer linear combination of $\binom{n}{0} = 1$. Using the result above (or just algebra), we get that degree 1 integer-valued polynomials are integer-linear combinations of $\binom{n}{0} = 1$ and $\binom{n}{1} = n$. Using the result again, we see that degree 2 integer-valued polynomials are integer linear combinations of the binomial coefficient polynomials, and so on.

Remark This is not necessarily the most efficient way to express a given integer-valued polynomial as a linear combination of binomial coefficient polynomials. An alternative approach is to use the fact that $\binom{x}{k}$ is 0 for $x = 0, 1, 2, \ldots, k-1$ together with the "identity principle" for polynomial functions.

[Exercise based on this]

3 Bernoulli polynomials

The Bernoulli numbers and Bernoulli polynomials were invented for the very problem we're discussing. The *Bernoulli polynomials* are certain polynomials $B_k(x)$ for $k = 0, 1, 2, 3, \ldots$. One way to describe them is to say that they are what we need to make the power rule for $\triangle$ "nice". That is, $B_0 = 1$ and for $k > 0$, B_k is a degree k polynomial with

$$(\triangle B_k)(n) = kn^{k-1} \tag{1}$$

This doesn't quite determine $B_k(x)$; we can specify it completely if we say what the constant term $B_k(0)$ is. There is one for each value of k; these are called the *Bernoulli numbers*, and arise in many areas of mathematics (we'll see them in Math 172 when we discuss power series expansions for the tangent function).

For purposes of the power sums, the constant terms won't matter; any polynomials satisfying (1) will do.

So what is the connection with the power-sum problem? The sum

$$s_m(n) = 1^m + 2^m + 3^m + \cdots + n^m$$

has $(\triangle s_m)(n) = (n+1)^m$. We can expand this using Pascal's triangle, getting

$$(\triangle s_m)(n) = \sum_{k=0}^{m} \binom{m}{k} n^k$$

so by linearity, the function

$$f(n) \quad = \quad \sum_{k=0}^{m} \binom{m}{k} \frac{1}{k+1} B_{k+1}(n)$$

has $\triangle f = \triangle s_m$. So $s_m = f + c$ for some constant c.

It's actually even easier than that. It turns out that if we define a new function $g(n) = f(n+1)$ by shifting values one place left[3], then

$$(\triangle g)(n) = g(n+1) - g(n) = f(n+2) - f(n+1) = (\triangle f)(n+1).$$

We say $\triangle$ is "shift-invariant". Therefore, since $(\triangle B_k)(n) = kn^{k-1}$ then the shifted version of B_k should have forward difference $k(n+1)^{k-1}$.

So we can say that $B_{m+1}(x+1)$ has forward difference $(m+1)(x+1)^m$. Therefore,

$$s_m(n) = \frac{1}{m+1} B_{m+1}(n+1) + c$$

We can evaluate the constant c by taking $n = 0$ to get $c = -\frac{1}{m+1} B_{m+1}(1)$.

In other words, for *any* choice of polynomial satisfying (1), we have

$$s_m(n) = \frac{1}{m+1} \left[B_{m+1}(n+1) - B_{m+1}(1) \right]$$

It would be nice to have an explicit formula for these polynomials. Here is where a choice of constant term comes in to make things easier.

First note one consequence of (1) is that $B_k(1) = B_k(0)$ for $k > 1$. This looks like it should depend on the constant term, but in fact it doesn't: it just says the coefficients other than the constant term must add to 0 (think about what $B_k(1)$ means).

Proof. $B_k(1) - B_k(0) = (\triangle B_k)(0) = k0^{k-1} = 0$ if $k > 1$. If $k = 1$ we have $B_1(x+1) - B_1(x) = 1$ so $B_1(x) = x + c$ for some constant c.

The shift-invariance gets us a formula:

$$B_k(x+1) = k \sum_{j=0}^{k-1} \binom{k-1}{j} \frac{1}{j+1} B_{j+1}(x) + c_k$$

again for some constant c_k.

We can simplify the binomial coefficient expression in the above: $k\binom{k-1}{j}\frac{1}{j+1} = \binom{k}{j+1}$. So our formula is

$$B_k(x+1) = \sum_{j=0}^{k-1} \binom{k}{j+1} B_{j+1}(x) + c_k \tag{2}$$

Now we can explain how the constant term is chosen. We choose the constant terms

[3] The *value* which is in position $n+1$ in the f sequence is in position n of the new g sequence.

$B_1(0), B_2(0), \ldots$ successively so that relation (2) holds with $c_k = 1$, for $k > 1$. (We have $c_1 = 1$ by definition of B_1).

To see that this is possible, write (2) as

$$B_k(x+1) = \sum_{j=1}^{k} \binom{k}{j} B_j(x) + 1$$

(changing the index of summation). Now define the zero-degree Bernoulli polynomial $B_0(x)$ to be 1. Then (2) becomes

$$B_k(x+1) = \sum_{j=0}^{k} \binom{k}{j} B_j(x). \tag{3}$$

Writing this out for $k = 2, 3, 4, \ldots$ gives equations which determine $B_{k-1}(x)$ given the lower-degree Bernoulli polynomials, and the condition that $\triangle B_k = kx^{k-1}$. Just subtract the term with $j = k$ from both sides to get

$$B_k(x+1) - B_k(x) = \sum_{j=0}^{k-1} \binom{k}{j} B_j(x)$$

and now the left-hand side is kx^{k-1}, while the right-hand side is a linear combination of $B_{k-1}(x)$ and lower degree Bernoulli polynomials. Solve for $B_{k-1}(x)$:

$$\binom{k}{k-1} B_{k-1}(x) = kx^{k-1} - \sum_{j=0}^{k-2} \binom{k}{j} B_j(x).$$

This gives a *recursive formula* for the Bernoulli polynomials, namely

$$B_{k-1}(x) = x^{k-1} - \sum_{j=0}^{k-2} k^{-1} \binom{k}{j} B_j(x).$$

So they are uniquely determined by the condition that $\triangle B_k = kx^{k-1}$ and our choice of constants.

The first few Bernoulli numbers are $b_0 = 1$, $b_1 = -\frac{1}{2}$, $b_2 = \frac{1}{6}$,

Note that we now have a formula for the power functions expressed in terms of Bernoulli polynomials:

$$kx^{k-1} = \sum_{j=0}^{k-1} \binom{k}{j} B_j(x)$$

===========

Another important property of the Bernoulli polynomials (considered as polynomial functions $\mathbf{R} \to \mathbf{R}$) is that they "act" like the power functions under the *derivative*:

$$B'_k = kB_{k-1} \tag{4}$$

The easiest way to establish this is to show that the polynomials $C_k = \frac{1}{k+1}B'_{k+1}$ satisfy the defining conditions (1) and (3).

This follows because $\triangle$ commutes with the derivative
so $\triangle(B'_{k+1}) = (\triangle B_{k+1})' = \big((k+1)x^k\big)' = (k+1)kx^{k-1}$ so $\triangle C_k = kx^{k-1}$
and differentiating (2) for index $k+1$ gives

$$B'_{k+1}(x+1) \;=\; \sum_{j=0}^{k} \binom{k+1}{j+1} B'_{j+1}(x)$$

which, after dividing both sides by $k+1$, says

$$C'_k(x+1) = \sum_{j=0}^{k} \frac{1}{k+1}\binom{k+1}{j+1}(j+1)C'_j(x)$$

and using the same binomial identity in reverse as we did to get (2), we get (3) for C_k.

==== ==

This says, in particular, that $B'_k(0) = kB_{k-1}(0)$. So the $k-1$st Bernoulli number, times k, appears also in the Bernoulli polynomial B_k but as the coefficient of x. In this way you can determine that *all* the coefficients of a Bernoulli polynomial are "made from" Bernoulli numbers.

Let's see if we can determine the way the Bernoulli polynomials are built up from the Bernoulli numbers. Let's write $b_k = B_k(0)$ for the kth Bernoulli number. Then we saw that B'_k has coefficient of x equal to kb_{k-1}. Its constant term is b_k by definition. How about the higher terms? Well, formula (4) can be used again, to get $B''_k = kB'_{k-1} = k(k-1)B_{k-2}$.

Now

$$\begin{aligned} B_k''(0) &= \text{coefficient of } x^2 \text{ times } 2 \\ &= k(k-1)b_{k-2} \end{aligned}$$

So the coefficient of x^2 in B_k is $k(k-1)/2 = \binom{k}{2}$ times b_{k-2}. Repeating this, we get

$$\begin{aligned} B_k'''(0) &= \text{coefficient of } x^3 \text{ times } 3! \\ &= k(k-1)(k-2)B_{k-3}(0) \\ &= k(k-1)(k-2)b_{k-3} \end{aligned}$$

so the coefficient of x^3 in B_k is $k(k-1)(k-2)/3! = \binom{k}{3}b_{k-3}$. In general, we can write

$$B_k(x) = \sum_{m=0}^{k} \binom{k}{m} b_{k-m} x^m$$

OK, so knowing the Bernoulli numbers is the same as knowing the Bernoulli polynomials. Is there a simple formula for the Bernoulli numbers?

Not that anyone knows! There are many *equations* that relate them, perhaps the simplest being

$$\sum_{j=1}^{k} \binom{k}{j} b_{k-j} = 0$$

which is just (3) evaluated at $x = 0$. This can be used recursively to compute b_n given values for $b_{n-1}, b_{n-2}, \ldots, b_1, b_0$.

The Bernoulli numbers arise in remarkably many places. For example, Euler evaluated the infinite series

$$\frac{1}{1^p} + \frac{1}{2^p} + \frac{1}{s^p} + \cdots$$

called a p-series in calculus textbooks, for p an even integer, to get

$$\sum_{k=1}^{\infty} \frac{1}{n^{2k}} = (-1)^{k-1} 2^{2k-1} \frac{b_{2k}}{(2k)!} \pi^{2k}$$

Interestingly, very little is known about the value of the sum of the p-series for *odd* integer values of p. The most dramatic result was R. Apery's 1979 proof that the sum is irrational for $p = 3$.

The p-series converges for $p > 1$ and defines the famous function known as the Riemann zeta function. Perhaps the most famous open problem in mathematics, the Riemann hypothesis, concerns properties of this function (extended to have domain the complex numbers other than 1).

Further information can be found in Omar Hijab's *Introduction to Calculus and Classical Analysis*, and Wikipedia's page on the "Basel problem".