On the sum $\sum_{k=1}^{n} k^p$

Marián Gloser

May 6, 2021

Abstract

When studying integer sums it is often important to have a formula for the sum of first n terms. Sum of particular interest in this paper will be $\sum_{k=1}^{n} k^p$. Formula for the first n terms exists and is named after Johann Faulhaber. It is therefore called Faulhaber's formula (sometimes also Bernoulli's formula). In this paper we shall derive a general formula for the above mentioned sum and also an alternate way to approach the problem which will allow us to calculate the formula for a given n and p.

Introduction

We approach the problem rather intuitively. In the first section we observe how such a formula can be derived for a few special cases. In later section we will solve the general problem in a similar manner as the special cases but with much more caution and consequently we will also prove the result. As mentioned in the Abstract, the purpose of this paper is to derive a general formula which, however, will differ from Faulhaber's formula (although, obviously, they are equivalent). An alternative approach but a similar one is chosen in section 3. Little knowledge of number theory and elementary linear algebra is presumed and special requirements are dealt with in the Appendices.

1 Special cases

1.1 Case p=1

Firstly, we will derive a formula for the sum $\sum_{k=1}^{n} k$. We are going to do it in two different ways.

Notice that the sum for $p = 1$ is

$$\sum_{k=1}^{n} k = 1 + 2 + 3 + \cdots + n. \tag{1}$$

We can see that this sum is equal to $\sum_{k=1}^{n}(n-k+1) = n + (n-1) + \cdots + 1$ (which is essentialy the same sum but in inverse order). Hence (since the sum is finite there are no problems with convergence):

$$\sum_{k=1}^{n} k + \sum_{k=1}^{n}(n-k+1) = 2\sum_{k=1}^{n} k = \sum_{k=1}^{n}(n+1) = n(n+1).$$

And the formula for this case ends up being very easy.

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

We shall now use a different approach, which will resemble the method we will use for the general case. Suppose (1) can be written as a qudratic polynomial in variable n. Then it has the form $p(n) = an^2 + bn + c$.[1] In order for the polynomial to work, it has to give a correct value for every n. Thus after plugging respectively $n = 1, 2, 3$ in the polynomial, we get a system of 3 linear equations, which can be easily solved.

$$\begin{aligned} a + b + c &= 1 \\ 4a + 2b + c &= 3 \\ 9a + 3b + c &= 6 \end{aligned}$$

The solution to this system is $a = \frac{1}{2}, b = \frac{1}{2}, c = 0$. Thus we get the desired result

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

1.2 Case p=2

The first approach we used in the case $p = 1$ is no longer possible in this case. We will thus focus on using the second method. This time we want to find a formula for the sum

$$\sum_{k=1}^{n} k^2 = 1 + 2^2 + 3^2 + \cdots + n^2. \tag{2}$$

[1]Normally polynomials have real arguments. Here, however, n is an integer.

Suppose that (2) can be written as a cubic polynomial in variable n. Then it has the form $p(n) = an^3+bn^2+cn+d$. Plugging now respectively $n = 1, 2, 3, 4$ in the polynomial we get a system of 4 linear equations.

$$\begin{aligned} a + b + c + d &= 1 \\ 8a + 4b + 2c + d &= 5 \\ 27a + 9b + 3c + d &= 14 \\ 64a + 16b + 4c + d &= 30 \end{aligned}$$

We get a solution $a = \frac{1}{3}, b = \frac{1}{2}, c = \frac{1}{6}, d = 0$. And therefore we get that the sum is

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

2 General case

2.1 Existence of a polynomial representing $\sum_{k=1}^{n} k^p$

In the previous section we assumed existence of a polynomial that would represent the sum for a given p. Of course for any given p we could proceed that way, but subsequently we would have to prove the formula we would get. Hence in this section we shall prove the existence of a polynomial representing $\sum_{k=1}^{n} k^p$.

Theorem 1. *For any given natural number p there exists a polynomial $p(n)$, such that $\sum_{k=1}^{n} k^p = p(n)$.*

Proof. We will prove that for every n such sum can be represented by a polynomial. We proceed by induction. Case $n = 1$ is trivial as in this case $p(n) = 1$. Suppose that the theorem holds for some $n = m$. Then

$$\sum_{k=1}^{m+1} k^p = (m+1)^p + \sum_{k=1}^{m} k^p = \sum_{i=0}^{p} \binom{p}{i} m^j + \sum_{k=1}^{m} k^p = p(n) + q(n) = r(n).$$

Where $p(n), q(n), r(n)$ are polynomials, hence we get the desired result. $\square$

2.2 Degree of the polynomial

Once we know that the polynomial exists, we can also derive the degree of such polynomial. Firstly, we define a sequence

$$s_{i,n} = s_{i-1,n+1} - s_{i-1,n}$$
$$s_{0,n} = \sum_{k=0}^{n} a_k.$$

Lemma. *If $s_{i-1,n} = p_{m+1}(n)$ then $s_{i,n} = q_m(n)$.*[2]

Proof. Suppose that $s_{i-1,n} = p_{m+1}(n)$. Then by definition

$$s_{i,n} = s_{i-1,n+1} - s_{i-1,n} = p_{m+1}(n+1) - p_{m+1}(n) =$$
$$\sum_{k=0}^{m+1} a_k(n+1)^k - \sum_{k=0}^{m+1} a_k n^k = \sum_{k=0}^{m+1} a_k((n+1)^k - n^k) =$$
$$\sum_{k=0}^{m+1} a_k(\sum_{j=0}^{k} \binom{k}{j} n^j - n^k) = \sum_{k=0}^{m+1} a_k \sum_{j=0}^{k-1} \binom{k}{j} n^j = q_m(n).$$

□

From this follows

Theorem 2. *If $s_{i-1,n}$ can be represented by a polynomial and $s_{i,n} = p_m(n)$ then $s_{i-1,n} = q_{m+1}(n)$.*

Now notice that $s_{1,n}$ is just the $(n+1)$-th term of the sum $\sum_{k=0}^{n+1} a_k$. If $\sum_{k=0}^{n+1} a_k = \sum_{k=1}^{n+1} k^p$ then the $(n+1)$-th term is $(n+1)^p$, which is p-th degree polynomial. Moreover since $\sum_{k=1}^{n} k^p$ can be represented by a polynomial (see theorem 1), by applying theorem 2 we get

Corollary. *The sum $\sum_{k=1}^{n} k^p$ can be represented as a $(p+1)$-th degree polynomial.*

We shall use this fact in the next section.

[2]Here $p_m(n), q_m(n)$... mean m-th degree polynomials.

3 Solution by Vandermonde matrix

In this section we are going to show a connection between the sum $\sum_{k=1}^{n} k^p$ and the Vandermonde matrix. We start by defining the Vandermonde matrix as a matrix of the form:

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix}$$

where $x_1, x_2, ...x_n$ are real numbers. Suppose now that $\sum_{k=1}^{n} k^p$ can be represented by the polynomial $p(n) = \sum_{k=0}^{p+1} a_k n^k$ (in section 2 we have proved that such a polynomial exists and has degree p+1). We are now going to plug in the polynomial values n=1,2,...p+2 respectively. We get a system of equations:

$$\sum_{k=0}^{p+1} a_k = 1,$$

$$\sum_{k=0}^{p+1} a_k 2^k = 1 + 2^p,$$

$$\vdots$$

$$\sum_{k=0}^{p+1} a_k (p+2)^k = \sum_{k=1}^{p+2} k^p.$$

This system of equations can be represented by a matrix equation.

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2^2 & \dots & 2^{p+1} \\ 1 & 3 & 3^2 & \dots & 3^{p+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & p+2 & (p+2)^2 & \dots & (p+2)^{p+1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{p+1} \end{bmatrix} = \begin{bmatrix} 1 \\ 1+2^p \\ 1+2^p+3^p \\ \vdots \\ \sum_{k=1}^{p+2} k^p \end{bmatrix}$$

Notice that the matrix on the left-hand side is Vandermonde matrix for $x_i = i$ and $n = p + 2$. Since this Vandermonde matrix has non-zero determinant (as is proven in Appendix A), the solution of the equation is unique and hence by solving the equation we would get the coefficients of the polynomial representing $\sum_{k=1}^{n} k^p$.

4 Solution by Lagrange polynomial

In the previous section we were concerned with transforming the problem into one involving Vandermonde matrix. However, we could have proceeded in a different, although quite a similar way. In this section we are going to use polynomial interpolation, more specifically the Lagrange polynomial. Details of this technique are dealt with in Appendix B.

Suppose we are given a set of x values $x = x_1, x_2, ..., x_n$ and a corresponding set of y values $y = y_1, y_2, ..., y_n$ and we want to find a polynomial of the lowest possible degree $p(x)$ such that $p(x_1) = y_1, p(x_2) = y_2, ..., p(x_n) = y_n$. In that case the polynomial we seek is called Lagrange polynomial and the following theorem holds.

Theorem 3. *Given a set of points $(x_1, y_1), (x_2, y_2), (x_n, y_n)$, where each x_i, y_i is a real number and no two x_i are the same, the polynomial of the lowest possible degree (Lagrange polynomial), such that $p(x_1) = y_1, p(x_2) = y_2, ..., p(x_n) = y_n$, has the form*

$$p(x) = \sum_{i=1}^{n} y_i \prod_{\substack{1 \le m \le n \\ m \ne i}} \frac{x - x_m}{x_i - x_m}. \tag{3}$$

Proof is given in Appendix B. We can now use this theorem to find the polynomial representing the sum $\sum_{k=1}^{n} k^p$. Let us now restate the problem in a way more suitable for using theorem 3.

We want to find a polynomial $p(n)$, such that $p(n) = \sum_{k=1}^{n} k^p$. By theorem 1 such a polynomial exists and by theorem 2 it has degree p+1. Moreover, it has to hold for each natural n. So we know it has to pass through points $(x_i = i, y_i = \sum_{k=1}^{i} k^p$. We can see that i has to go from 1 to $p+2$ so that the polynomial we get has degree $p+1$. Knowing this we can now plug $x_i = i$ and $y_i = \sum_{k=1}^{i} k^p$ into equation (3).

$$p(n) = \sum_{i=1}^{p+2} \sum_{k=1}^{i} k^p \prod_{\substack{1 \le m \le p+2 \\ m \ne i}} \frac{n - m}{i - m}.$$

Appendices

A Determinant of Vandermonde matrix

In this appendix we will prove the following theorem:

Theorem. *The determinant of Vandermonde matrix is equal to the following expression:*

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} = \prod_{1 \le i < m \le n} (x_m - x_i).$$

Proof. Let us call the determinant $W(x_1, ..., x_n)$, such that

$$W(x_1, ..., x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix}.$$

$W(x_1, ..., x_n)$ is a polynomial of degree $n-1$ in x_n (since x_n^{n-1} appears in the matrix). This polynomial vanishes if x_n takes any of the values $x_1, ..., x_{n-1}$, since then the determinant has two identical columns. Hence the polynomial is divisible by the product $(x_n - x_1)...(x - x_{n-1})$, so that

$$W(x_1, ..., x_n) = a(x_1, ..., x_{n-1}) \prod_{k=1}^{n-1} (x_n - x_k).$$

Here $a(x_1, ..., x_{n-1})$ is the leading coefficient of the polynomial $W(x_1, ..., x_n)$. Expanding the Vandermonde determinant with respect to the last column we see that this coefficient is just $W(x_1, ..., x_{n-1})$. It follows that

$$W(x_1, ..., x_n) = W(x_1, ..., x_{n-1}) \prod_{k=1}^{n-1} (x_n - x_k).$$

Similarly,

$$W(x_1, ..., x_{n-1}) = W(x_1, ..., x_{n-2}) \prod_{k=1}^{n-2} (x_{n-1} - x_k),$$

$$\vdots$$

$$W(x_1, x_2) = W(x_1)(x_2 - x_1),$$

and obviously $W(x_1) = 1$. Multiplying these equalities together we get the result

$$W(x_1, ..., x_n) = \prod_{1 \leq i < m \leq n} (x_m - x_i).$$

$\square$

In particular, if $x_1, ..., x_n$ are all distinct (which is the case in section 3), then the determinant is non-zero.

B Lagrange polynomial

Let us recall and then prove theorem 3:

Theorem ??. *Given a set of points $(x_1, y_1), (x_2, y_2), (x_n, y_n)$, where each x_i, y_i is a real number and no two x_i are the same, the polynomial of the lowest possible degree (Lagrange polynomial), such that $p(x_1) = y_1, p(x_2) = y_2, ..., p(x_n) = y_n$, has the form*

$$p(x) = \sum_{i=1}^{n} y_i \prod_{\substack{1 \leq m \leq n \\ m \neq i}} \frac{x - x_m}{x_i - x_m}.$$

Proof. When substituting $x = x_j$ into the formula, the product becomes 0 for every i except $i = j$ since then m will eventually take the value j and thus the nominator becomes zero and that also means that the whole product is zero. The only non-zero term is when $i = j$. In that case the product becomes 1 and we get $p(x_j) = y_j$. Hence the polynomial interpolates the whole set of given points. Notice also that the degree of the polynomial is at most $n-1$ since each product contains exactly $n-1$ factors, each containing x.

We will now prove that there is one and only one polynomial $p(x)$ of degree $\leq n - 1$ such that $p(x_1) = y_1, p(x_2) = y_2, ..., p(x_n) = y_n$ and therefore that the Lagrange polynomial is of the lowest possible degree.

Suppose that there are at least 2 such polynomials, say $p(x), q(x)$. Then the polynomial $p(x) - q(x) = 0$ for each $x = x_i$. Moreover the degree of the polynomial is $\leq n - 1$. However, the polynomial has n roots, namely $x_1, ..., x_n$, so the degree of the polynomial should be at least n, which is a contradiction. Hence there is only one such polynomial. $\square$

C Faulhaber's formula

In this appendix we show how the Faulhaber's formula looks like in its most common form.

$$\sum_{k=1}^{n} k^p = \frac{1}{p+1} \sum_{j=0}^{p} (-1)^j \binom{p+1}{j} B_j n^{p+1-j}.$$

Where B_j is the j-th Bernoulli number with the convention of $B_1 = -\frac{1}{2}$.

References

[1] Shilov, E. G. *Linear Algebra.* Dover Publications, 1977. 387 pp. ISBN 978-0486635187

[2] Pinter, C. C. *A Book of Abstract Algebra.* Dover Publications, 2010. 384 pp. ISBN 978-0-486-47417-5

[3] Orosi, G. *A Simple Derivation Of Faulhaber's Formula.* ISSN 1607-2510