Finding a formula for the sum of squares

Andrew Taylor

February 21 2022

1 Introduction

The goal of this paper is to derive a formula for the sum of squares.

Before working on this, we will explain what a formula is, what a sequence is, and how to add the terms of a sequence.

2 Formulas

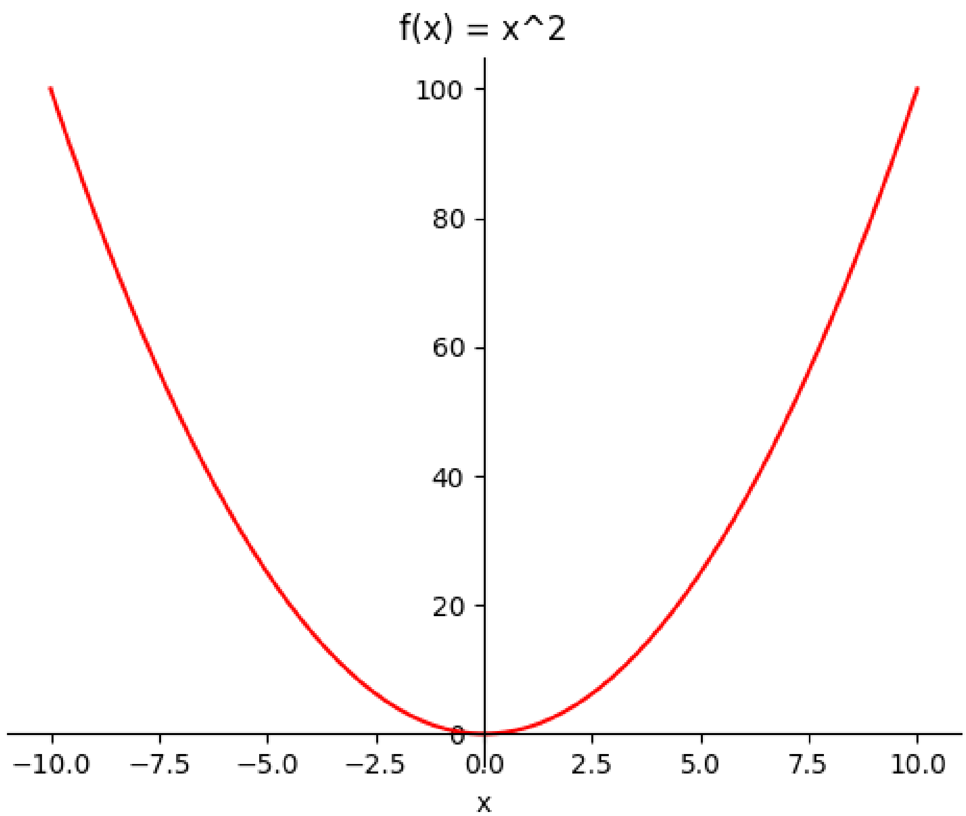
A formula is a function.

A function is a correspondence between two sets X and Y such that each element in X is associated with exactly one element in Y.

Another wording is this:

A function is a correspondence between two sets X and Y such that each element in X corresponds to exactly one element in Y.

On the next page there is a graph of a function.


f(x) = x^2
100
80
60
40
20
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
x

The graph illustrates the function $f(x) = x^2$.

The x and y axes visualize the real numbers. The x axis visualizes the interval $[-10, 10]$ and the y axis visualizes the interval $[0, 100]$.

f has the real numbers as its domain and range. We can completely describe the function by saying $f : \mathbb{R} \to \mathbb{R}$ and $f(x) = x^2$.

The symbol $\mathbb{R}$ denotes the real numbers.

If we travel left or right along the x axis, and stop at any coordinate x_1, we can see that there is one $f(x_1)$ corresponding to x_1.

This is visual evidence that $f(x)$ is a function.

The graph shows that every element in the domain of f is associated with exactly one element in the range of f.

3 Sequences

A sequence is a list of numbers. A sequence can be finite or infinite.

One example of a sequence is the first ten positive squares.

$$1, 4, 9, 16, 25, 36, 49, 64, 81, 100$$

We can write the sequence of squares symbolically as:

$$a_n = n^2$$

The subscript n can be any natural number.

Thus $a_{10} = 10^2 = 100$ is the tenth term in the sequence.

A sequence is a kind of function. For every input we get exactly one output. For every index we get exactly one value.

The symbol $\mathbb{N}$ denotes the natural numbers.

The sequence $a_n = n^2$ is a function $f : \mathbb{N} \to \mathbb{N}$ defined by $f(n) = n^2$.

4 Summation

Let's write out the sum of the first ten squares.

$$1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + 7^2 + 8^2 + 9^2 + 10^2$$

This takes a lot of time and effort. What if we wanted to sum the first one hundred squares? Is there a concise way of writing this?

We can write this concisely using summation.

$$\sum_{k=1}^{100} k^2$$

The summation operator $\sum$ gets its symbol from the Greek letter sigma. The summation operator accepts a sequence, a lower bound, and an upper bound, and sums every term in the sequence from the lower bound to the upper bound, up to and including the upper bound.

$$\sum_{k=1}^{10} k^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + 7^2 + 8^2 + 9^2 + 10^2$$

In the above equation, we see a concise expression on the left hand side, and an expanded expression on the right hand side.

The summation operator is called a trinary operator because it accepts three operands: a sequence, a lower bound, and an upper bound.

$$\sum_{k=L}^{M} a_k$$

In the expression above, a_k is the sequence, L is the lower bound of summation, and M is the upper bound of summation. The variable k is the index of summation.

We can also write the operation using a second notation, shown below.

$$\sum(a_k, L, M)$$

Summation is an operation that adds the terms of a sequence or a subsequence. It can add all of the terms of a sequence, or some of the terms of a sequence. It can add a finite number of terms, or an infinite number of terms. When we use summation on an infinite number of terms, we call it a series.

We are looking for a formula that sums the first n squares. We can write the sum of the first n squares concisely in this way:

$$\sum_{k=1}^{n} k^2$$

5 The sum of the first n natural numbers

Lemma 1. *Let $a_k = k$ be the sequence of natural numbers. Then*

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

Proof. Let's write out the sum of the first n natural numbers.

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

By rearranging terms, we can write the equation in two ways.

$$\sum_{k=1}^{n} k = 1 + 2 + 3 + ... + n$$
$$\sum_{k=1}^{n} k = n + (n-1) + (n-2) + ... + 1$$

Now let's add the two equations, so that 1 is paired with n, 2 is paired with (n-1), 3 is paired with (n-2), and so on.

$$2 * \sum_{k=1}^{n} k = (1+n) + (2 + (n-1)) + (3 + (n-2)) + ... + (n+1)$$

This gives us n pairs of numbers that sum to $n + 1$.

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

Now let's divide both sides of the equation by 2.

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

□

6 Telescoping sums

Lemma 2. *Let* a_k *be a sequence of real numbers. Then*

$$\sum_{k=1}^{n}(a_k - a_{k-1}) = a_n - a_0$$

Proof. We can expand the sum, and rearrange the terms, so that every term but the first and the last cancel out.

$$\begin{aligned}\sum_{k=1}^{n}(a_k - a_{k-1}) &= (a_n - a_{n-1}) + (a_{n-1} + a_{n-2}) + ... + (a_1 - a_0) \\ &= a_n + (-a_{n-1} + a_{n-1}) + (-a_{n-2} + a_{n-2}) + ... + (-a_1 + a_1) - a_0 \\ &= a_n - a_0\end{aligned}$$

□

7 The sum of the first n squares

Theorem 1. *Let* a_k *be the sequence* $a_k = k^2$*. Then*

$$\sum_{k=1}^{n} a_k = \frac{n(n+1)(2n+1)}{6}$$

Proof. Let b_k be the sequence $b_k = k^3$.

From lemma 2 we know that

$$\begin{aligned}\sum_{k=1}^{n}(b_k - b_{k-1}) &= b_n - b_0 \\ &= n^3 - 0^3 \\ &= n^3\end{aligned}$$

This gives us equation 1.

$$\sum_{k=1}^{n}(b_k - b_{k-1}) = n^3 \tag{1}$$

We also know that

$$\begin{aligned}\sum_{k=1}^{n}(b_k - b_{k-1}) &= \sum_{k=1}^{n}[k^3 - (k-1)^3] \\ &= \sum_{k=1}^{n}[k^3 - (k^3 - 3k^2 + 3k - 1)] \\ &= \sum_{k=1}^{n}(3k^2 - 3k + 1) \\ &= 3 * \sum_{k=1}^{n} k^2 - 3 * \sum_{k=1}^{n} k + \sum_{k=1}^{n} 1\end{aligned}$$

Substituting our result from lemma 1 into the above equation, we get

$$\begin{aligned}\sum_{k=1}^{n}(b_k - b_{k-1}) &= 3\sum_{k=1}^{n} k^2 - 3\sum_{k=1}^{n} k + \sum_{k=1}^{n} 1 \\ &= 3\sum_{k=1}^{n} k^2 - \frac{3n(n+1)}{2} + n \\ &= 3\sum_{k=1}^{n} k^2 - \frac{3n^2 + 3n}{2} + \frac{2n}{2} \\ &= 3\sum_{k=1}^{n} k^2 - \frac{3n^2 + n}{2}\end{aligned}$$

This gives us equation 2.

$$\sum_{k=1}^{n}(b_k - b_{k-1}) = 3\sum_{k=1}^{n} k^2 - \frac{3n^2 + n}{2} \tag{2}$$

We can now set equations 1 and 2 equal to each other.

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

Adding to both sides of the equation, and simplifying, we get

$$\begin{aligned}3\sum_{k=1}^{n} k^2 &= n^3 + \frac{3n^2 + n}{2} \\ &= \frac{2n^3 + 3n^2 + n}{2}\end{aligned}$$

Dividing both sides of the equation by 3, we get

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

$\square$

8 Conclusion

The key insight is that

$$\sum_{k=1}^{n} a_k$$

is a term in the sequence

$$\sum_{k=1}^{n} (b_k - b_{k-1})$$

We see this relationship in the following equation:

$$\begin{aligned}\sum_{k=1}^{n} (b_k - b_{k-1}) &= \sum_{k=1}^{n} [k^3 - (k-1)^3] \\ &= \sum_{k=1}^{n} (3k^2 - 3k + 1)\end{aligned}$$

We perform substitutions using lemmas 1 and 2 and solve for

$$\sum_{k=1}^{n} k^2$$

This gives us the formula

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$