Multiple Sums

P. Sam Johnson

National Institute of Technology Karnataka (NITK)
Surathkal, Mangalore, India

Multiple Sums

The terms of a **(multiple) sum** are governed by two indices j and k:

$$\sum_{1\le j,k\le 3} a_j b_k = a_1 b_1 + a_1 b_2 + \cdots + a_3 b_2 + a_3 b_3.$$

Only one $\sum$ sign is needed, although there is more than one index of summation; $\sum$ denotes a sum over all combinations of indices that apply.

For example,

$$\sum_j \sum_k a_{j,k}[P(j,k)]$$

is an abbreviation for

$$\sum_i \Big(\sum_j a_{j,k}[P(j,k)] \Big)$$

which is the sum, over all integers j, of

$$\sum_k a_{j,k}[P(j,k)]$$

the latter being the sum over all integers k of all terms $a_{j,k}$ for which $P(j,k)$ is true. In such cases we say that the double sum is "summed first on k." A sum that depends on more than one index can be summed first on any one of its indices.

When we talk about a sum of sums, we use two $\sum$'s. In this regard, we have a basic law called **interchanging the order of summation**, which generalizes the associative law.

$$\sum_j \sum_k a_{jk}[P(j,k)] = \sum_{P(j,k)} a_{j,k} = \sum_k \sum_j a_{j,k}[P(j,k)].]$$

The middle term of this law is a sum over two indices. On the left,

$$\sum_j \sum_k$$

stands for summing first on k, then on j. On the right,

$$\sum_k \sum_j$$

stands for summing first on j, then on k. In practice when we want to evaluate a double sum in closed form, it is usually either to sum it first on one index rather than on the other; we get to choose whichever is more convenient.

We illustrate how to manipulate with the double sum using $\sum\sum$'s:

$$
\begin{aligned}
\sum_{1\le j,k\le 3} a_j b_k &= \sum_{j,k} a_j b_k \; [1 \le j, k \le 3] \\
&= \sum_j \sum_k a_j b_k \; [1 \le j \le 3][1 \le k \le 3] && \text{nine terms in no particular order} \\
&= \sum_j \sum_k a_j b_k \; [1 \le j \le 3][1 \le k \le 3] && \text{nine terms are grouped into three} \\
&= \sum_j a_j \; [1 \le j \le 3] \sum_k b_k [1 \le k \le 3] && \text{by distributive law to factor out the } a\text{'s} \\
&= \left(\sum_j a_j [1 \le j \le 3]\right)\left(\sum_k b_k [1 \le k \le 3]\right) && \text{factors out } (b_1 + b_2 + b_3) \text{ for each } j \\
&= \left(\sum_{j=1}^{3} a_j\right)\left(\sum_{k=1}^{3} b_k\right).
\end{aligned}
$$

General distributive law

Hence a **general distributive law**

$$\sum_{\substack{j\in J\\ k\in K}} a_j b_j = \Big(\sum_{j\in J} a_j\Big)\Big(\sum_{k\in K} b_k\Big),$$

valid for all sets of indices J and K.

Here is a useful factorization:

$$[1 \le j \le n][j \le k \le n] = [1 \le j \le k \le n] = [1 \le k \le n][1 \le j \le k].$$

This Iversonian equation allows us to write

$$\sum_{j=1}^{n}\sum_{k=j}^{n} a_{j,k} = \sum_{1\le j\le k\le n} a_{j,k} = \sum_{k=1}^{n}\sum_{j=1}^{n} a_{j,k}. \qquad (1)$$

One of these two sums of sums is usually easier to evaluate than the other. We can use (1) to switch from the hard one to the easy one.

Consider the array

$$\begin{pmatrix} a_1a_1 & a_1a_2 & \cdots & a_na_n \\ a_2a_1 & a_2a_2 & \cdots & a_2a_n \\ \vdots & \vdots & \cdots & \vdots \\ a_na_1 & a_na_2 & \cdots & a_na_n \end{pmatrix}$$

of n^2 products a_ja_k.

Our goal it to find a simple formula for

$$S_{\triangledown} = \sum_{1\le j\le k\le n} a_ja_k$$

the sum of all elements **on or above** the main diagonal of this array.

We can rename (j, k) and (k, j), we get

$$S_{◹} = \sum_{1\leq j\leq k\leq n} a_j a_k = \sum_{1\leq k\leq j\leq n} a_k a_j = \sum_{1\leq k\leq j\leq n} a_j a_k = S_{◺}.$$

Since

$[1 \leq j \leq k \leq n] + [1 \leq k \leq j \leq n] = [1 \leq j, k \leq n] + [1 \leq j = k \leq n]$,

we have $2S_{◹} = S_{◹} + S_{◺} = \sum_{1\leq j,k\leq n} a_j a_k + \sum_{1\leq j=k\leq n} a_j a_k$. By the general distributive law,

$$\sum_{1\leq j,k\leq n} a_j a_k = \Big(\sum_{j=1}^{n} a_j\Big)\Big(\sum_{k=1}^{n} a_k\Big) = \Big(\sum_{k=1}^{n} a_k\Big)^2.$$

Therefore we have

$$S_{◹} = \sum_{1\leq j\leq k\leq n} a_j a_k = \frac{1}{2}\Big(\big(\sum_{k=1}^{n} a_k\big)^2 + \sum_{k=1}^{n} a_k^2\Big)$$

an expression for the upper triangular sum in terms of simple single sums.

Exercises

1. *Using general distributive law, prove that*

$$\Big(\sum_{k=1}^{n} a_k\Big)\Big(\sum_{k=1}^{n} b_k\Big) = n\sum_{k=1}^{n} a_k b_k - \sum_{1\le j<k\le n} (a_k - a_j)(b_k - b_j). \quad (2)$$

The identity (2) yields Chebyshev's monotonic inequalities *as a special case :*

   (a) $\Big(\sum_{k=1}^{n} a_k\Big)\Big(\sum_{k=1}^{n} b_k\Big) \le n\sum_{k=1}^{n} a_k b_k$ *if* $a_1 \le a_2 \le \cdots \le a_n$ *and* $b_1 \le b_2 \le \cdots \le b_n$.

   (b) $\Big(\sum_{k=1}^{n} a_k\Big)\Big(\sum_{k=1}^{n} b_k\Big) \ge n\sum_{k=1}^{n} a_k b_k$ *if* $a_1 \le a_2 \le \cdots \le a_n$ *and* $b_1 \ge b_2 \ge \cdots \ge b_n$.

2. *Prove that* $S_n = \sum_{1\le j<k\le n} \frac{1}{k-j} = nH_n - n$, *by using a method on multiple sums.*

One problem with several solutions

We look at a single example from several different angles. We are going to try to find a closed form for the sum of the first n squares, which we denote it by $\square_n$.

$$\square_n = \sum_{k=0}^{n} k^2, \quad \text{for } n \geq 0.$$

We will see that there are at least 7 different ways to solve this problem.

Method 1 : Guess the answer, prove it by induction.

We may **conjecture** a closed form and we merely have to prove that it is correct.

We may come up with the formula

$$\square_n = \frac{n(n+\frac{1}{2})(n+1)}{3}, \quad \text{for } n \geq 0.$$

which works for small values of n.

Method 2 : Perturb the sum (perturbation method)

We extract the first and last terms of $\Box_{n+1}$ in order to get an equation for $\Box_n$:

$$\Box_n + (n+1)^2 = \sum_{0 \le k \le n} (k+1)^2$$

Hence we get

$$\Box_n + (n+1)^2 = \Box_n + 2 \sum_{0 \le k \le n} k + (n+1).$$

We cannot find $\Box_n$, but we can find $\displaystyle\sum_{0 \le k \le n} k$.

The above derivation reveals a way to sum the first n integers in closed form

$$\sum_{0 \le k \le n} k = \frac{n(n+1)}{2}.$$

Method 2 : Perturb the sum (perturbation method)

Could it be that if we start with the sum of the integers cubed (denoted by C_n)?

We will get an expression for the integers squared? Let's try it.

$$C_n + (n+1)^3 = \sum_{0 \le k \le n} (k+1)^3.$$

Hence

$$3C_n = (n+1)^3 - 3(n+1)n/2 - (n+1).$$

Thus

$$3C_n = n\left(n + \frac{1}{2}\right)(n+1).$$

Method 3 : Build a repertorie (repertorie method)

Consider the recursion

$$\begin{aligned}\Box_n &= 0\\ \Box_n &= \Box_{n-1} + n^2. \end{aligned} \tag{3}$$

The generalization of (3) is

$$\begin{aligned}R_n &= \alpha\\ R_n &= R_{n-1} + \beta + \gamma n + \delta n^2, \quad \text{for } n \geq 1\end{aligned}$$

which has the solution of the form

$$R_n = A(n)\alpha + B(n)\beta + C(n)\gamma + D(n)\delta, \quad \text{for } n \geq 1.$$

If we consider our recurrence relation $\Box_n = R_n$ if we set $\alpha = \beta = \gamma = 0$ and $\delta = 1$. Thus

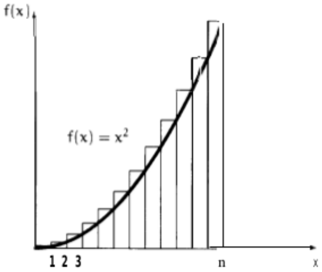
$$\Box_n = \frac{n(n + \frac{1}{2})(n+1)}{3}.$$

Method 4 : Replace sums by integrals

Since $\square_n$ is the sum of the area of rectangles whose sizes are

$$1 \times 1, 1 \times 4, \ldots, 1 \times n^2,$$

it is approximately equal to the area under the curve $f(x) = x^2$ between 0 and n. The area under the curve is $\int_0^n x^2 dx = \frac{n^3}{3}$.



Method 4 : Replace sums by integrals

Let us examine the error in the approximation,

$$E_n = \square_n - \frac{1}{3}n^3.$$

E_n is the sum of areas of the wedge-shaped error terms and E_n satisfies the simpler recurrence

$$\begin{aligned} E_n &= \square_n - \frac{1}{3}n^3 \\ &= E_0 + \sum_{k=1}^{n}(k - \frac{1}{3}) = \sum_{k=1}^{n}(k - \frac{1}{3}). \end{aligned}$$

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

Method 5 : Expand and contract (going from single sum to a double sum)

$$\begin{aligned}
\Box_n &= \sum_{1\le k\le n} k^2 = \sum_{1\le j\le k\le n} k \\
&= \sum_{1\le j\le n}\ \sum_{j\le k\le n} k = \sum_{1\le j\le n} \left(\frac{j+n}{2}\right)(n-j+1) \\
&= \frac{1}{2}\sum_{1\le j\le n} [n(n+1)+j-j^2] \\
&= \frac{1}{2}n^2(n+1)+\frac{1}{4}n(n+1)-\frac{1}{2}\Box_n \\
&= \frac{1}{2}n\left(n+\frac{1}{2}\right)(n+1)-\frac{1}{2}\Box_n.
\end{aligned}$$

Method 6 : Finite Calculus

Infinite calculus is based on the properties of the **derivative operator** D, defined by

$$Df(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

Finite calculus is based on the properties of the **difference operator** Δ, defined by

$$\Delta f(x) = f(x+1) - f(x).$$

The symbols D and Δ are called **operators** because they operate on functions to give new functions ; they are functions of functions that produce functions.

Next we shall discuss finite calculus.

References

1. **Graham, Knuth and Patashnik**, *"Concrete Mathematics – A Foundation for Computer Science"*, Pearson Education.
2. **Marko Petkovsek, Herbert S. Wilf and Doron Zeilberger**, "$A = B$", AK Peters Ltd., Wellesley, Massachusetts.
3. **Herbert S. Wilf**, *"Generatingfunctionology"*, Third Edition, AK Peters Ltd., Wellesley, Massachusetts.