Proof of the associative law for matrices

camillejr.github.io/science-docs

Motivation

I was reviewing the MIT course 18.06 in Linear Algebra by Professor Gilbert Strang and the associative law for matrices first appeared on the blackboard in more or less this way:

$$(AB)C = A(BC) \tag{1}$$

Professor Strang said then:

"It's not that easy to prove that this is correct. You have to go into the gory details of matrix multiplication, do it both ways and see that you come out the same."

So I said: *let's do it!*

Proof

Suppose now we have matrices A size $(m \times n)$, B size $(n \times p)$ and C size $(p \times r)$. We are going to prove that the relation from the eq.(1) holds for these matrices.

We start off with the general equation for the $i^{th}j^{th}$ element of the matrix that is a result of multiplication of two matrices A and B, it can be found in a linear algebra textbook to be:

$$\gamma_{ij}^{AB} = \sum_{k=1}^{n} (a_{ik}b_{kj}) \tag{2}$$

And the general equation for the $i^{th}j^{th}$ element of the matrix that results in multiplication of matrices B and C first is:

$$\gamma_{ij}^{BC} = \sum_{k=1}^{p} (b_{ik}c_{kj}) \tag{3}$$

We write the general equation for the $i^{th}j^{th}$ element of the matrix that results in complete multiplication

$(AB)C$:

$$\gamma_{ij}^{(AB)C} = \sum_{l=1}^{p} \left(\left(\sum_{k=1}^{n} (a_{ik}b_{kl}) \right) c_{lj} \right) \tag{4}$$

And we write the general equation for the $i^{th}j^{th}$ element of the matrix that results in complete multiplication $A(BC)$:

$$\gamma_{ij}^{A(BC)} = \sum_{h=1}^{n} \left(a_{ih} \left(\sum_{k=1}^{p} (b_{hk}c_{kj}) \right) \right) \tag{5}$$

Now we are going to take eq.(4) and eq.(5) and write the sums element by element.

We start with eq.(4):

$$\gamma_{ij}^{(AB)C} = \sum_{l=1}^{p} \left(\left(\sum_{k=1}^{n} (a_{ik}b_{kl}) \right) c_{lj} \right) =$$

$$\sum_{l=1}^{p} \left(a_{i1}b_{1l}c_{lj} + \cdots + a_{in}b_{nl}c_{lj} \right) =$$

$$= a_{i1}b_{11}c_{1j} + \cdots + a_{in}b_{n1}c_{1j} + \tag{6}$$

$$+ a_{i1}b_{12}c_{2j} + \cdots + a_{in}b_{n2}c_{2j} +$$

$$\vdots$$

$$+ a_{i1}b_{1p}c_{pj} + \cdots + a_{in}b_{np}c_{pj}$$

For eq.(5) we have:

$$\gamma_{ij}^{A(BC)} = \sum_{h=1}^{n} \left(a_{ih} \left(\sum_{k=1}^{p} (b_{hk}c_{kj}) \right) \right) =$$

$$\sum_{h=1}^{n} \left(a_{ih}b_{h1}c_{1j} + \cdots + a_{ih}b_{hp}c_{pj} \right) =$$

$$= a_{i1}b_{11}c_{1j} + \cdots + a_{i1}b_{1p}c_{pj} + \tag{7}$$

$$+a_{i2}b_{21}c_{1j} + \cdots + a_{i2}b_{2p}c_{pj}+$$

$$\vdots$$

$$+a_{in}b_{n1}c_{1j} + \cdots + a_{in}b_{np}c_{pj}$$

In the eq.(6) we have the final sum written in p "rows" and in each row there is n elements of the sum.

In the eq.(7) we have the final sum written in n "rows" and in each row there is p elements of the sum.

Notice now, that the sum of all the first elements from each row of eq.(6) is the same as the sum of all the first's row elements in the eq.(7) (marked in orange in the equations).

The same holds for the second elements from each row of eq.(6) with the second's row elements in eq.(7) and so on.

If we were to write all the elements in an array (that would be size $(p \times n)$ for eq.(6) and $(n \times p)$ for eq.(7)), it would be equivalent to saying that taking the sum of elements from any column q of the case $(AB)C$ is the same as the sum of elements from row q of the case $A(BC)$.

This concludes the proof.[1]

References

[1] *Lec 2 | MIT 18.06 Linear Algebra*, Spring 2005
https://youtu.be/QVKj3LADCnA

---

[1]Especially that I'm running out of letters to index things.