0.1 Matrix multiplication.

0. *Assumed background.*

- 1.1 *Matrices, matrix addition, and scalar multiplication for matrices.*

Abstract. We introduce:—

- matrix multiplication (first for row vectors to column vectors from the left, then in the general situation through presentation in blocks),
- properties of matrix multiplication,
- notions of square matrix and identity matrix,
- notion of positive powers of square matrices,
- presentation of matrix multiplication in terms of blocks.

1. **Definition. (Multiplication of row vector to column vector from the left.)**

   Let A be a row vector with n entries, and B be a column vector with n entries.

   Suppose $A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$, and $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

   though we have a long expression, it's just a 1×1 matrix.

   Then we define the product AB to be the (1×1)-matrix

   $$\begin{bmatrix} a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \end{bmatrix}.$$

   For future convenience we abuse notations to confuse as the number $a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$.

2. **Definition. (Multiplication of matrix to column vector from the left.)**

   Let A be an $(m \times n)$-matrix, and B be a column vector with n entries.

   $$B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

   Suppose $A = \begin{bmatrix} A_1 \\ \hline A_2 \\ \hline \vdots \\ \hline A_m \end{bmatrix}$, in which A_k stands for the k-th row of A for each k.

   We define the product AB to be the column vector with m entries, given by A_k is a row vector

   A number

   $$AB = \begin{bmatrix} A_1 B \\ A_2 B \\ \vdots \\ A_m B \end{bmatrix}.$$

   column vector

   A: $(m \times n)$ matrix

   B: $(n \times 1)$ - matrix

   AB: $(m \times 1)$ matrix

   (For each k, the k-th entry of AB is the number $A_k B$.)

   Remark. Denote the (i, j)-th entry of A by a_{ij}.

   Denote the j-th entry of B by b_j.

   Then the k-th entry of AB is given by the number $a_{k1} b_1 + a_{k2} b_2 + \cdots + a_{kn} b_n$.

   Writing out the entries in the matrices explicitly, we have   only   1 column here

   A_1

   A_m

   $$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_{11}b_1 & + & a_{12}b_2 & + & \cdots & + & a_{1n}b_n \\ a_{21}b_1 & + & a_{22}b_2 & + & \cdots & + & a_{2n}b_n \\ \vdots & & \vdots & & & & \vdots \\ a_{m1}b_1 & + & a_{m2}b_2 & + & \cdots & + & a_{mn}b_n \end{bmatrix}.$$

   A B

   $(m \times 1)$ - matrix

3. **Definition. (Matrix multiplication.)**

   Let A be an $(m \times n)$-matrix, and B be an $(n \times p)$-matrix.

   Suppose $B = \begin{bmatrix} B_1 & | & B_2 & | & \cdots & | & B_p \end{bmatrix}$, in which B_ℓ is the ℓ-th column of B for each ℓ.

   We define the product AB to be the $(m \times p)$-matrix given by

   $$AB = \begin{bmatrix} AB_1 & | & AB_2 & | & \cdots & | & AB_p \end{bmatrix}. \quad \Leftarrow (m \times p) \text{ matrix}$$

1

(For each ℓ, the ℓ-th column of AB is the column vector AB_ℓ with m entries.)

Remark.   Denote the (i,j)-th entry of A by a_{ij}, Denote the (k,ℓ)-th entry of B by $b_{k\ell}$.

Denote the i-th row of A by A_i.

Then the (i,ℓ)-th entry of AB is given by the number $A_i B_\ell = a_{i1}b_{1\ell} + a_{i2}b_{2\ell} + \cdots + a_{in}b_{n\ell}$, and

[handwritten: you're suggested to check this carefully!]

$$A = \begin{bmatrix} \frac{A_1}{} \\ \vdots \\ A_m \end{bmatrix} \qquad B = \begin{bmatrix} B_1 & | & B_2 & | & \cdots & | & B_p \end{bmatrix}$$

$$AB = \begin{bmatrix} A_1B_1 & A_1B_2 & \cdots & A_1B_p \\ A_2B_1 & A_2B_2 & \cdots & A_2B_p \\ \vdots & \vdots & & \vdots \\ A_mB_1 & A_mB_2 & \cdots & A_mB_p \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}b_{j1} & \sum_{j=1}^{n} a_{1j}b_{j2} & \cdots & \sum_{j=1}^{n} a_{1j}b_{jp} \\ \sum_{j=1}^{n} a_{2j}b_{j1} & \sum_{j=1}^{n} a_{2j}b_{j2} & \cdots & \sum_{j=1}^{n} a_{2j}b_{jp} \\ \vdots & \vdots & & \vdots \\ \sum_{j=1}^{n} a_{mj}b_{jp} & \sum_{j=1}^{n} a_{mj}b_{j2} & \cdots & \sum_{j=1}^{n} a_{mj}b_{jp} \end{bmatrix} = \begin{bmatrix} \dfrac{A_1B}{} \\ \dfrac{A_2B}{} \\ \vdots \\ A_mB \end{bmatrix}.$$

4. **Example (1). (Matrix multiplication.)**

(a) Let $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 5 \\ 1 & 6 \\ 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix}$.

Write $A_1 = [\ 1\quad 2\quad 3\quad 4\quad 5\ ]$, $A_2 = [\ 2\quad 3\quad 4\quad 5\quad 6\ ]$, $A_3 = [\ 3\quad 4\quad 5\quad 6\quad 7\ ]$, $A_4 = [\ 4\quad 5\quad 6\quad 7\quad 8\ ]$.

Write $B_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $B_2 = \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix}$.

We have $A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix}$, $B = [\ B_1\ |\ B_2\ ]$.

[handwritten:
$A_1 B_1$
$= 1 \times 0 + 2 \times 1 + 3 \times 2 + 4 \times 3 + 5 \times 4$
  $\quad\quad\quad\quad 2 \quad\quad 6 \quad\quad 12 \quad\quad 20$
$= 40$ *]*

Then

$$AB = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix}[\ B_1\ |\ B_2\ ] = \begin{bmatrix} A_1B_1 & A_1B_2 \\ A_2B_1 & A_2B_2 \\ A_3B_1 & A_3B_2 \\ A_4B_1 & A_4B_2 \end{bmatrix} = \begin{bmatrix} 40 & 115 \\ 50 & 150 \\ 60 & 185 \\ 70 & 220 \end{bmatrix}.$$

(b) Let $A = \begin{bmatrix} 1 & -1 & 1 & 6 & 1 \\ 6 & 4 & 1 & 4 & -2 \\ 2 & 3 & 2 & -1 & 3 \\ 1 & 2 & 3 & 2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & -5 \\ 2 & -4 & 1 \\ -1 & 1 & 2 \\ 4 & 2 & -3 \\ 6 & 3 & 4 \end{bmatrix}$.

We have

$$AB = \begin{bmatrix} 1 & -1 & 1 & 6 & 1 \\ 6 & 4 & 1 & 4 & -2 \\ 2 & 3 & 2 & -1 & 3 \\ 1 & 2 & 3 & 2 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & -5 \\ 2 & -4 & 1 \\ -1 & 1 & 2 \\ 4 & 2 & -3 \\ 6 & 3 & 4 \end{bmatrix} = \cdots = \begin{bmatrix} 28 & 20 & -18 \\ 17 & -13 & -44 \\ 20 & -3 & 12 \\ 10 & -1 & -3 \end{bmatrix}.$$

5. **Theorem (1). (Distributive Laws for addition and multiplication of matrices.)**

(1) *Suppose A is an $(m \times n)$-matrix and B, C are $(n \times p)$-matrices. Then $A(B + C) = (AB) + (AC)$.*

(2) *Suppose A, B are $(m \times n)$-matrices and C is an $(n \times p)$-matrix. Then $(A + B)C = (AC) + (BC)$.*

Remark on notations.   We will dispense with the brackets in '$(AB) + (AC)$', '$(AC) + (BC)$', and simply write '$AB + AC$', '$AC + BC$' respectively.

Proof of Statement (1) of Theorem (1).

Suppose A is an $(m \times n)$-matrix and B, C are $(n \times p)$-matrices.
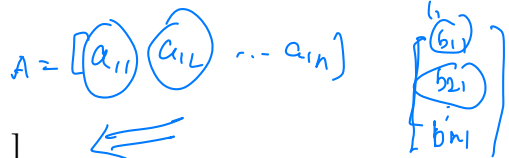
(a) Suppose $m = 1$ and $p = 1$ for the moment. (So A is a row vector and B, C are column vectors.)

For each j, denote the $(1, j)$-th entry of A by a_{1j}.

For each k, denote the $(k, 1)$-th entries of B, C respectively by b_{k1}, c_{k1}.

[handwritten: we write A as $A = [a_{11}, a_{12}, \cdots, a_{1n}]$]

The $(k, 1)$-th entry of B, C is given by $b_{k1} + c_{k1}$.

By definition,

(handwritten:) $A = [\overset{\frown}{(a_{11})} \overset{\frown}{(a_{1L})} \cdots a_{1n}]$ $B \begin{bmatrix} \overset{\frown}{(b_1)} \\ \overset{\frown}{(b_{21})} \\ \vdots \\ [b_{n1} \end{bmatrix}$

$$AB = [\ a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}\ ]$$
$$AC = [\ a_{11}c_{11} + a_{12}c_{21} + \cdots + a_{1n}c_{n1}\ ]$$
$$A(B + C) = [\ a_{11}(b_{11} + c_{11}) + a_{12}(b_{21} + c_{21}) + \cdots + a_{1n}(b_{n1} + c_{n1})\ ]$$

Then

$$
\begin{aligned}
A(B + C) &= [\ a_{11}(b_{11} + c_{11}) + a_{12}(b_{21} + c_{21}) + \cdots + a_{1n}(b_{n1} + c_{n1})\ ] \\
&= [\ (a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}) + (a_{11}c_{11} + a_{12}c_{21} + \cdots + a_{1n}c_{n1})\ ] \\
&= [\ a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}\ ] + [\ a_{11}c_{11} + a_{12}c_{21} + \cdots + a_{1n}c_{n1}\ ] \\
&= AB + AC
\end{aligned}
$$

(b) We now still suppose $p = 1$, but leave the value of m un-restricted.   *(handwritten:)* B, C are still Column vectors

For each i, denote the i-th row of A by A_i.

By the calculation above, we have $A_i(B + C) = A_iB + A_iC$.   *(handwritten:)* This is due to (a)

Then

$$
A(B + C) = \begin{bmatrix} A_1 \\ \hline A_2 \\ \hline \vdots \\ \hline A_m \end{bmatrix}(B + C) = \begin{bmatrix} A_1(B + C) \\ A_2(B + C) \\ \vdots \\ A_m(B + C) \end{bmatrix} = \begin{bmatrix} A_1B + A_1C \\ A_2B + A_2C \\ \vdots \\ A_mB + A_mC \end{bmatrix} = \begin{bmatrix} A_1B \\ A_2B \\ \vdots \\ A_mB \end{bmatrix} + \begin{bmatrix} A_1C \\ A_2C \\ \vdots \\ A_mC \end{bmatrix} = AB + AC.
$$

(c) We now leave the values of m, p un-restricted.

For each j, denote the j-th columns of B, C by B_j, C_j respectively.   *(handwritten:)* Try to use $(AB)_{i\ell}$ $A: m \times n$ $B: n \times p$

By the calculation above, we have $A(B_j + C_j) = AB_j + AC_j$.

Then   *(handwritten:)* $= \sum_{j=1}^{n} a_{ij}\, b_{j\ell}$

$$
\begin{aligned}
A(B + C) &= A[\ B_1 + C_1 \mid B_2 + C_2 \mid \cdots \mid B_p + C_p\ ] \\
&= [\ A(B_1 + C_1) \mid A(B_2 + C_2) \mid \cdots \mid A(B_p + C_p)\ ] \\
&= [\ AB_1 \mid AB_2 \mid \cdots \mid AB_p\ ] + [\ AC_1 \mid AC_2 \mid \cdots \mid AC_p\ ] = AB + AC
\end{aligned}
$$

(handwritten:) to move the theorem !

Proof of Statement (2) of Theorem (1).   This is left as an exercise. (Imitate what is done above.)

6. **Theorem (2).**

   Suppose A is an $(m \times n)$-matrix and B is an $(n \times p)$-matrix. Suppose λ is a real number.

   Then $\lambda(AB) = (\lambda A)B = A(\lambda B)$.

   Remark on notations.   We will dispense with the brackets in '$\lambda(AB)$', '$(\lambda A)B$', and simply write 'λAB'.

   Proof of Theorem (2).   Exercise.   *(handwritten:)* When you carry out multiplication, be careful with sizes of matrices:

7. **Definition. (Square matrix.)**   *(handwritten:)* $m \times n$, $n \times p$, $p \times q$, $q \times r$ ..

   *A matrix with the same number of rows and columns is called a **square matrix**.*

8. **Theorem (3). ('Existence and uniqueness' of 'multiplicative identity' for matrix multiplication.)**

   There is a unique $(n \times n)$-square matrix M such that for any $(n \times n)$-square matrix A, the equalities '$MA = A$', '$AM = A$' hold.

   Proof of Theorem (3).   *(handwritten:)* The number 1 is the multiplicative identity for multiplication of real numbers:

   (handwritten, right:) for addition we want to find Z such that for all A $\boxed{Z + A = A}$

   Define $M = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$.

   Define $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$. ('δ_{ij}' is called the Kronecker symbol.)

   Then the (i, j)-th entry of M is δ_{ij} for each i, j.   *(handwritten:)* $1 \cdot a = a \cdot 1 = a$

   [We intend to verify two things: *(handwritten:)* for any real number a.

3

(1) *The equalities 'MA = A', 'AM = A' hold for any $(n \times n)$-matrix A.*

← *it is important that* $MA = AM = A$ *holds for any* A. *For example,* *if* $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

(2) *If some $(n\times n)$-matrix P possesses the property that both equalities 'PA = A', 'AP = A' hold for any $(n\times n)$-matrix A', then P = M.*

We proceed with (1), (2) separately.]

(1) Let A be an $(n \times n)$-matrix with its (i, j)-th entry given by a_{ij}.

    ∗ By the definition of matrix multiplication, the (i, j)-th entry of MA is given by

$$\delta_{i1}a_{1j} + \delta_{i2}a_{2j} + \delta_{i3}a_{3j} + \cdots + \delta_{in}a_{nj} = \delta_{ii}a_{ij} = a_{ij}.$$

we can choose $M = \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}$, *then* $MA = A = AM$ *where* c *can be any real number, hence it is not unique.*

    Hence $MA = A$.

    ∗ By the definition of matrix multiplication, the (i, j)-th entry of MA is given by

$$a_{i1}\delta_{1j} + a_{i2}\delta_{2j} + a_{i3}\delta_{3j} + \cdots + a_{in}\delta_{nj} = a_{ij}\delta_{jj} = a_{ij}.$$

    Hence $AM = A$.

(2) Let P be an $(n \times n)$-matrix. Suppose $AP = A = PA$ for any $(n \times n)$-matrix A.

Then in particular, (because M is an $(n \times n)$-matrix), we have $MP = M = PM$.

Since P is an $(n \times n)$-matrix, we have $PM = P$ from the calculations above.

Hence $M = PM = P$.

same trick as we remarked for Thm (3) in the previous note.

9. **Definition. (Identity matrix.)**

For each positive integer n, the $(n \times n)$-square matrix whose (k, k)-th entry is 1 for each k and whose every other entry is 0, given explicitly by

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & 0 & & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

is called the $(n \times n)$-identity matrix.

It is denoted by I_n.

10. **Theorem (4).**

Suppose A is an $(m \times n)$-matrix Then the equalities '$I_m A = A$', '$AI_n = A$' hold.

Notice the position of A in the product.

Proof of Theorem (4). Exercise. (Imitate (the relevant portion of) the argument for Theorem (4).)

11. **Theorem (5). (Associativity of matrix multiplication.)**

Suppose A is an $(m \times n)$-matrix, B is an $(n \times p)$-matrix, and C is a $(p \times q)$-matrix. Then $A(BC) = (AB)C$.

Remark. We will give a proof for Theorem (5) later. (The argument is not hard. It is in the same spirit as the argument for Theorem (1), but the work in keeping track of symbols is more 'involved'.)

Because of this result, we may write '$(AB)C$', '$A(BC)$' simply as ABC, unless we want to emphasize that associativity of matrix multiplication is used.

In the light of this result, the definition below, for the notion of positive integral powers of square matrices, makes perfect sense.

A of size $(n \times n)$

12. **Definition. (Positive integral powers of square matrices.)**

Let n be a positive integer. Suppose A is a square matrix.

*The n-**th power of** A is defined to be the square matrix $\underbrace{AAA \cdots AAA}_{n \text{ copies of } A}$.*

It is denoted by A^n.

13. **Example (2). (Positive integral powers of square matrices.)**

(a) Let $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$. (Note that A itself is neither I_4 nor $-I_4$.)

strictly upper triangular matrix

We have $A^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $A^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$, $A^4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = I_4$.

Question: can you give another matrix X ($\neq I_4$, $-I_4$) such that $X^4 = I_4$?

(b) Let $B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. (Note that B itself is not the zero matrix.)

for real numbers, $a^2 = 1$ ⇒ $a = \pm 1$
$a^4 = 1$ ⇒ $a = \pm 1$.

if x is a number, $x^t = 0$ ⇒ $x = 0$.

We have $B^2 = \begin{bmatrix} 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, $B^3 = \begin{bmatrix} 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 24 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, $B^4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 24 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, $B^5 = O_{5\times 5}$.

But for matrices, this doesn't hold.

Remark. The various parts of Example (2) have told us that the statements (†), (‡) are false:—

(†) *Suppose that A is a $(n \times n)$-square matrix with real entries. Further that suppose there is some positive p such that $A^p = I_n$. Then $A = I_n$ or $A = -I_n$.* ✗

(‡) *Suppose that B is a square matrix. Further suppose that B is not the zero matrix. Then, for each positive integer p, B^p is not the zero matrix.* ✗

The content of part (a) of Example (2) is referred to as a counter-example against the statement (†).

The content of part (b) of Example (2) is referred to as a counter-example against the statement (‡).

14. **Proof of Theorem (5).**

Suppose A is an $(m \times n)$-matrix, B is an $(n \times p)$-matrix, and C is a $(p \times q)$-matrix.

(a) Suppose $m = 1$ and $p = 1$ for the moment. (So A is a row vector and C is a column vector.)

Suppose $A = [\; a_1 \quad a_2 \quad \cdots \quad a_n \;]$, and $C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$.

$\dfrac{(A\,B) \cdot C}{(1 \times p) \quad (p \times 1)}$ ⇒ *We should get an 1×1 - matrix.*

Denote the (j,k)-th entry by b_{jk}. Denote the j-th row of B by $B_{\text{row-}j}$. Denote the k-th column of B by $B_{\text{col-}k}$.

(So $B = \begin{bmatrix} B_{\text{col-1}} \mid B_{\text{col-2}} \mid \cdots \mid B_{\text{col-}p} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} B_{\text{row-1}} \\ B_{\text{row-2}} \\ \vdots \\ B_{\text{row-}n} \end{bmatrix}$.)

We claim that each of $(AB)C$, $A(BC)$ is the sum of all the $a_j b_{jk} c_k$'s, each copy exactly once.

- We have

$$AB = A\begin{bmatrix} B_{\text{col-1}} \mid B_{\text{col-2}} \mid \cdots \mid B_{\text{col-}p} \end{bmatrix} = \begin{bmatrix} AB_{\text{col-1}} \mid AB_{\text{col-2}} \mid \cdots \mid AB_{\text{col-}p} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^{n} a_j b_{j1} & \sum_{j=1}^{n} a_j b_{j2} & \cdots & \sum_{j=1}^{n} a_j b_{jp} \end{bmatrix}$$

$1 \times p$ matrix ≈ row vector

$AB_{\text{col-1}} : [\, a_1, a_2, \cdots a_n \,] \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix}$

$= a_1 b_{11} + a_2 b_{21} + \cdots + a_n b_{n1}$

Then

$$(AB)C = \left(\sum_{j=1}^{n} a_j b_{j1}\right) c_1 + \left(\sum_{j=1}^{n} a_j b_{j2}\right) c_2 + \cdots + \left(\sum_{j=1}^{n} a_j b_{jp}\right) c_p = \sum_{k=1}^{p} c_k \left(\sum_{j=1}^{n} a_j b_{jk}\right) = \sum_{k=1}^{p}\sum_{j=1}^{n} a_j b_{jk} c_k.$$

So $(AB)C$ is the sum of all the $a_j b_{jk} c_k$'s, each copy exactly once.

- We have

$$BC = \begin{bmatrix} \underline{B_{\text{row-1}}} \\ \underline{B_{\text{row-2}}} \\ \vdots \\ B_{\text{row-}n} \end{bmatrix} C = \begin{bmatrix} \underline{B_{\text{row-1}}C} \\ \underline{B_{\text{row-2}}C} \\ \vdots \\ B_{\text{row-}n}C \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{p} b_{1k}c_k \\ \sum_{k=1}^{p} b_{2k}c_k \\ \vdots \\ \sum_{k=1}^{p} b_{nk}c_k \end{bmatrix}$$

(handwritten: $\in (n\times 1)$ -matrix i.e, a column vector.)

Then

$$A(BC) = a_1\left(\sum_{k=1}^{p} b_{1k}c_k\right) + a_2\left(\sum_{k=1}^{p} b_{2k}c_k\right) + \cdots + a_n\left(\sum_{k=1}^{p} b_{nk}c_k\right) = \sum_{j=1}^{n} a_j\left(\sum_{k=1}^{p} b_{jk}\right) = \sum_{j=1}^{n}\sum_{k=1}^{p} a_j b_{jk}c_k.$$

So $A(BC)$ is also the sum of all the $a_j b_{jk} c_k$'s, each copy exactly once.

(b) We now leave the values of m, p un-restricted.

For each i, denote the i-th row of A by A_i. (So $A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}$.)   *(handwritten: Each A_i is a row vector)*

For each ℓ, denote the ℓ-th column of C by C_ℓ. (So $C = [\ C_1\ |\ C_2\ |\ \cdots\ |\ C_q\ ]$.)

(handwritten: $A_i B : (1\times p)$ -matrix for each $1\le i\le m$.)

- We have $AB = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} B = \begin{bmatrix} A_1 B \\ A_2 B \\ \vdots \\ A_m B \end{bmatrix}$.

Then $(AB)C = \begin{bmatrix} A_1 B \\ A_2 B \\ \vdots \\ A_m B \end{bmatrix}[\ C_1\ |\ C_2\ |\ \cdots\ |\ C_q\ ] = \begin{bmatrix} (A_1 B)C_1 & (A_1 B)C_2 & \cdots & (A_1 B)C_q \\ (A_2 B)C_1 & (A_2 B)C_2 & \cdots & (A_2 B)C_q \\ \vdots & \vdots & & \vdots \\ (A_m B)C_1 & (A_m B)C_2 & \cdots & (A_m B)C_q \end{bmatrix}$.

The (i, ℓ)-th entry of $(AB)C$ is $(A_i B)C_\ell$ for each i, ℓ.

- We have $BC = B[\ C_1\ |\ C_2\ |\ \cdots\ |\ C_q\ ] = [\ BC_1\ |\ BC_2\ |\ \cdots\ |\ BC_q\ ]$.

Then $A(BC) = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}[\ BC_1\ |\ BC_2\ |\ \cdots\ |\ BC_q\ ] = \begin{bmatrix} A_1(BC_1) & A_1(BC_2) & \cdots & A_1(BC_q) \\ A_2(BC_1) & A_2(BC_2) & \cdots & A_2(BC_q) \\ \vdots & \vdots & & \vdots \\ A_m(BC_1) & A_m(BC_2) & \cdots & A_m(BC_q) \end{bmatrix}$.

The (i, ℓ)-th entry of $A(BC)$ is $A_i(BC_\ell)$ for each i, ℓ.

By the calculations above, the equality $(A_i B)C_\ell = A_i(BC_\ell)$ holds for each i and for each ℓ.

Then $(AB)C = A(BC)$ by the definition of matrix equality.

(handwritten: Suggestion ; Using the formula : $(AB)_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj}$ [A: $m\times n$ matrix, B: $n\times p$ matrix]

to prove the associativity law ; $A(BC) = (AB)\,C$.)