

1. Definition. (Matrix Multiplication.)

(a) Let A be a row vector with n entries, given by $A = [a_1 \ a_2 \ \cdots \ a_n]$, and B be a column

vector with n entries, given by $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

We define the product AB to be the (1×1) -matrix $[a_1b_1 + a_2b_2 + \cdots + a_nb_n]$.

For future convenience we abuse notations to confuse as the number $a_1b_1 + a_2b_2 + \cdots + a_nb_n$.

1. Definition. (Matrix Multiplication.)

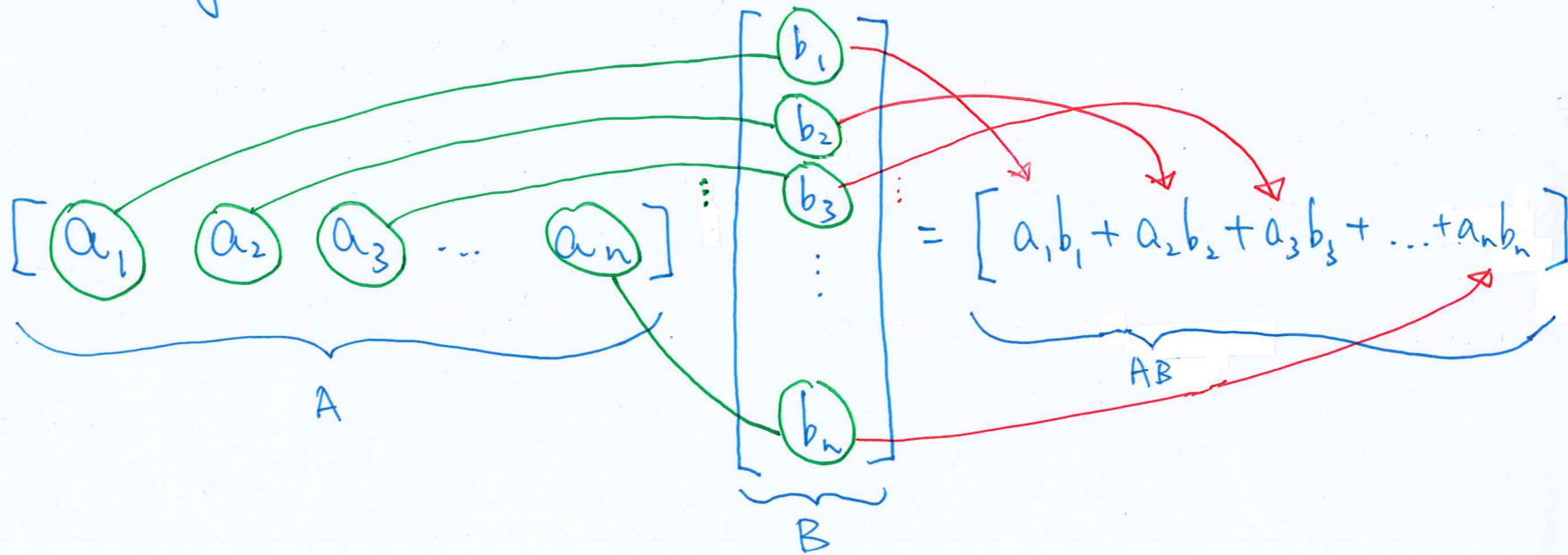
(a) Let A be a row vector with n entries, given by $A = [a_1 \ a_2 \ \cdots \ a_n]$, and B be a column

vector with n entries, given by $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

We define the product AB to be the (1×1) -matrix $[a_1b_1 + a_2b_2 + \cdots + a_nb_n]$.

For future convenience we abuse notations to confuse as the number $a_1b_1 + a_2b_2 + \cdots + a_nb_n$.

Visualization of the above description :



(b) Let A be an $(m \times n)$ -matrix, whose k -th row is denoted by A_k , and B be a column

vector with n entries, given by $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

We define the product AB to be the column vector with m entries, given by $AB =$

$\begin{bmatrix} A_1B \\ A_2B \\ \vdots \\ A_mB \end{bmatrix}$. (The k -th entry of AB is A_kB .)

Remark. Denoting the (i, j) -th entry of A by a_{ij} , the k -th entry of AB is given by

the number $\sum_{j=1}^n a_{kj}b_j = a_{k1}b_1 + a_{k2}b_2 + \cdots + a_{kn}b_n$.

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

(b) Let A be an $(m \times n)$ -matrix, whose k -th row is denoted by A_k , and B be a column

vector with n entries, given by $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$.

$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$

We define the product AB to be the column vector with m entries, given by $AB =$

$\begin{bmatrix} A_1 B \\ A_2 B \\ \vdots \\ A_m B \end{bmatrix}$. (The k -th entry of AB is $A_k B$.)

$A_k B = [a_{k1} \ a_{k2} \ \dots \ a_{kn}] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = [a_{k1}b_1 + a_{k2}b_2 + \dots + a_{kn}b_n]$

Remark. Denoting the (i, j) -th entry of A by a_{ij} , the k -th entry of AB is given by

the number $\sum_{j=1}^n a_{kj}b_j = a_{k1}b_1 + a_{k2}b_2 + \dots + a_{kn}b_n$.

Now think of this (1×1) -matrix as just a number.

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

(c) Let A be an $(m \times n)$ -matrix, and B be an $(n \times p)$ -matrix, whose ℓ -th column is denoted by B_ℓ .

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

$$AB = [AB_1 \mid AB_2 \mid \cdots \mid AB_p].$$

(The ℓ -th column of AB is AB_ℓ .)

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

the (i, ℓ) -th entry of AB is given by the number $\sum_{j=1}^n a_{ij}b_{j\ell} = a_{i1}b_{1\ell} + a_{i2}b_{2\ell} + \cdots + a_{in}b_{n\ell}$.

Denote the i -th row of A by A_i , we have $A_i B_\ell = \sum_{j=1}^n a_{ij}b_{j\ell}$ for each j , and

$$AB = \begin{bmatrix} A_1 B_1 & A_1 B_2 & \cdots & A_1 B_p \\ A_2 B_1 & A_2 B_2 & \cdots & A_2 B_p \\ \vdots & \vdots & & \vdots \\ A_m B_1 & A_m B_2 & \cdots & A_m B_p \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n a_{1j}b_{j1} & \sum_{j=1}^n a_{1j}b_{j2} & \cdots & \sum_{j=1}^n a_{1j}b_{jp} \\ \sum_{j=1}^n a_{2j}b_{j1} & \sum_{j=1}^n a_{2j}b_{j2} & \cdots & \sum_{j=1}^n a_{2j}b_{jp} \\ \vdots & \vdots & & \vdots \\ \sum_{j=1}^n a_{mj}b_{j1} & \sum_{j=1}^n a_{mj}b_{j2} & \cdots & \sum_{j=1}^n a_{mj}b_{jp} \end{bmatrix} = \begin{bmatrix} \frac{A_1 B}{A_2 B} \\ \vdots \\ \frac{A_m B}{A_m B} \end{bmatrix}$$

$$B = [B_1 | B_2 | \dots | B_p] \text{ with } B_\ell = \begin{bmatrix} b_{1\ell} \\ b_{2\ell} \\ \vdots \\ b_{n\ell} \end{bmatrix}$$

(c) Let A be an $(m \times n)$ -matrix, and B be an $(n \times p)$ -matrix, whose ℓ -th column is denoted by B_ℓ .
The number of entries in each row of A matches the number of entries in each column of B .

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix} \text{ with } A_k = [a_{k1} \ a_{k2} \ \dots \ a_{kn}]$$

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

$$AB = [AB_1 | AB_2 | \dots | AB_p].$$

$$AB_\ell = \begin{bmatrix} A_1 B_\ell \\ A_2 B_\ell \\ \vdots \\ A_m B_\ell \end{bmatrix} = \begin{bmatrix} a_{11}b_{1\ell} + a_{12}b_{2\ell} + \dots + a_{1n}b_{n\ell} \\ a_{21}b_{1\ell} + a_{22}b_{2\ell} + \dots + a_{2n}b_{n\ell} \\ \vdots \\ a_{m1}b_{1\ell} + a_{m2}b_{2\ell} + \dots + a_{mn}b_{n\ell} \end{bmatrix}$$

(The ℓ -th column of AB is AB_ℓ .)

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

the (i, ℓ) -th entry of AB is given by the number $\sum_{j=1}^n a_{ij}b_{j\ell} = a_{i1}b_{1\ell} + a_{i2}b_{2\ell} + \dots + a_{in}b_{n\ell}$.

Denote the i -th row of A by A_i , we have $A_i B_\ell = \sum_{j=1}^n a_{ij}b_{j\ell}$ for each j , and

$$AB = \begin{bmatrix} A_1 B_1 & A_1 B_2 & \dots & A_1 B_p \\ A_2 B_1 & A_2 B_2 & \dots & A_2 B_p \\ \vdots & \vdots & & \vdots \\ A_m B_1 & A_m B_2 & \dots & A_m B_p \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n a_{1j}b_{j1} & \sum_{j=1}^n a_{1j}b_{j2} & \dots & \sum_{j=1}^n a_{1j}b_{jp} \\ \sum_{j=1}^n a_{2j}b_{j1} & \sum_{j=1}^n a_{2j}b_{j2} & \dots & \sum_{j=1}^n a_{2j}b_{jp} \\ \vdots & \vdots & & \vdots \\ \sum_{j=1}^n a_{mj}b_{j1} & \sum_{j=1}^n a_{mj}b_{j2} & \dots & \sum_{j=1}^n a_{mj}b_{jp} \end{bmatrix} = \begin{bmatrix} A_1 B \\ A_2 B \\ \vdots \\ A_m B \end{bmatrix}$$

$\uparrow \quad \quad \uparrow \quad \quad \dots \quad \uparrow$
 $AB_1 \quad AB_2 \quad \dots \quad AB_p$

2. Examples.

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

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

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

$$\text{We have } A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix}, B = [B_1 \mid B_2]. \text{ Then } AB = \begin{bmatrix} A_1B_1 & A_1B_2 \\ A_2B_1 & A_2B_2 \\ A_3B_1 & A_3B_2 \\ A_4B_1 & A_4B_2 \end{bmatrix} = \begin{bmatrix} 40 & 115 \\ 50 & 150 \\ 60 & 185 \\ 70 & 220 \end{bmatrix}.$$

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

$$\text{We have } AB = \dots = \begin{bmatrix} 28 & 20 & -18 \\ 17 & -13 & -44 \\ 20 & -3 & 12 \\ 10 & -1 & -3 \end{bmatrix}.$$

3. Definition. (Identity matrix.)

For each positive integer n , the $(n \times n)$ -matrix whose (k, k) -th entry is 1 for each k and whose every other entry is 0 is called the identity matrix, and is denoted by I_n .

Remark. $I_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$

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

$$\text{We have } AB = \dots = \begin{bmatrix} 28 & 20 & -18 \\ 17 & -13 & -44 \\ 20 & -3 & 12 \\ 10 & -1 & -3 \end{bmatrix}.$$

3. Definition. (Identity matrix.)

For each positive integer n , the $(n \times n)$ -matrix whose (k, k) -th entry is 1 for each k and whose every other entry is 0 is called the identity matrix, and is denoted by I_n .

Remark. $I_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ \vdots & \vdots & & \dots & & \vdots \\ 0 & 0 & 0 & & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$

$$I_1 = [1]$$

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$I_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

⋮

4. **Theorem (1).** (Basic properties of matrix multiplication.)

The statements below hold:

(a) Suppose A is an $(m \times n)$ -matrix with real entries.

Then $I_m A = A = A I_n$.

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

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

(c) Suppose A is an $(m \times n)$ -matrix and B, C are $(n \times p)$ -matrices.

Then $A(B + C) = (AB) + (AC)$.

(d) Suppose A, B are $(m \times n)$ -matrices and C is an $(n \times p)$ -matrix.

Then $(A + B)C = (AC) + (BC)$.

Remark on terminologies.

(a) Statement (a) is how the ‘Law of Existence of Multiplicative Identity’ for matrix multiplication. The identity matrices of various sizes are the ‘multiplicative identities’ concerned.

(b) Statements (c), (d) are collectively known as the ‘Distributive Laws’ for matrix addition and matrix multiplication.

4. Theorem (1). (Basic properties of matrix multiplication.)

The statements below hold:

Illustration: $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(a) Suppose A is an $(m \times n)$ -matrix with real entries.

Then $I_m A = A = A I_n$.

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

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

(c) Suppose A is an $(m \times n)$ -matrix and B, C are $(n \times p)$ -matrices.

Then $A(B + C) = (AB) + (AC)$.

(d) Suppose A, B are $(m \times n)$ -matrices and C is an $(n \times p)$ -matrix.

Then $(A + B)C = (AC) + (BC)$.

Remark on terminologies.

(a) Statement (a) is how the 'Law of Existence of Multiplicative Identity' for matrix multiplication. The identity matrices of various sizes are the 'multiplicative identities' concerned.

(b) Statements (c), (d) are collectively known as the 'Distributive Laws' for matrix addition and matrix multiplication.

5. **Lemma (2).** (Special case of associativity of matrix multiplication.)

Suppose A is a row vector with n entries, given by

$$A = [a_1 \ a_2 \ \cdots \ a_n],$$

B is an $(n \times p)$ -matrix whose (j, k) -th entry is denoted by b_{jk} , and C is a column vectors with p entries, given by

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}.$$

Then

$$(AB)C = A(BC) = \sum_{j=1}^n \sum_{k=1}^p a_j b_{jk} c_k.$$

5. Lemma (2). (Special case of associativity of matrix multiplication.)

Suppose A is a row vector with n entries, given by

$$A = [a_1 \ a_2 \ \cdots \ a_n],$$

B is an $(n \times p)$ -matrix whose (j, k) -th entry is denoted by b_{jk} , and C is a column vector with p entries, given by

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}.$$

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \left. \vphantom{\begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix}} \right\} \begin{array}{l} n \text{ rows} \\ p \text{ columns} \end{array}$$

Then

$$(AB)C \stackrel{\circ}{=} A(BC) = \sum_{j=1}^n \sum_{k=1}^p a_j b_{jk} c_k.$$

Interpretation of this equality: whether you multiply A to B first, or multiply B to C first, you will end up with the same product as long as the order A, B, C is this multiplications is maintained.

What is it?
It is the sum of all possible np expressions of the form $a_j b_{jk} c_k$, each expression appearing exactly once.

Proof.

- Denote the k -th column of B by $B_{\text{col-}k}$. We have

$$\begin{aligned} AB &= A \left[B_{\text{col-}1} \mid B_{\text{col-}2} \mid \cdots \mid B_{\text{col-}p} \right] = \left[AB_{\text{col-}1} \mid AB_{\text{col-}2} \mid \cdots \mid AB_{\text{col-}p} \right] \\ &= \left[\sum_{j=1}^n a_j b_{j1} \quad \sum_{j=1}^n a_j b_{j2} \quad \cdots \quad \sum_{j=1}^n a_j b_{jp} \right] \end{aligned}$$

Then

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

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

- Denote the j -th row of B by $B_{\text{row-}j}$. We have

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

Then

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

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

Hence $(AB)C = A(BC)$.

6. Illustration of the idea in the argument for Lemma (2).

$$\text{Let } A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{bmatrix} C = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}.$$

- We have $B = \left[B_{\text{col-1}} \mid B_{\text{col-2}} \mid B_{\text{col-3}} \mid B_{\text{col-4}} \right]$, in which $B_{\text{col-}k} = \begin{bmatrix} b_{1k} \\ b_{2k} \\ b_{3k} \end{bmatrix}$ for each $k = 1, 2, 3, 4$.

Then, for each $k = 1, 2, 3, 4$, we have $AB_{\text{col-}k} = a_1b_{1k} + a_2b_{2k} + a_3b_{3k}$. This is the k -th entry in the (1×4) -row matrix AB . So

$$\begin{aligned} (AB)C &= (a_1b_{11} + a_2b_{21} + a_3b_{31})c_1 \\ &\quad + (a_1b_{12} + a_2b_{22} + a_3b_{32})c_2 \\ &\quad + (a_1b_{13} + a_2b_{23} + a_3b_{33})c_3 \\ &\quad + (a_1b_{14} + a_2b_{24} + a_3b_{34})c_4 \\ &= a_1b_{11}c_1 + a_1b_{12}c_2 + a_1b_{13}c_3 + a_1b_{14}c_4 \\ &\quad + a_2b_{21}c_1 + a_2b_{22}c_2 + a_2b_{23}c_3 + a_2b_{24}c_4 \\ &\quad + a_3b_{31}c_1 + a_3b_{32}c_2 + a_3b_{33}c_3 + a_3b_{34}c_4 \end{aligned}$$

which is the sum of all the $a_jb_{jk}c_k$'s, each copy exactly once.

6. Illustration of the idea in the argument for Lemma (2).

$$\text{Let } A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{bmatrix} \quad C = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}.$$

- We have $B = [B_{\text{col-1}} \mid B_{\text{col-2}} \mid B_{\text{col-3}} \mid B_{\text{col-4}}]$, in which $B_{\text{col-}k} = \begin{bmatrix} b_{1k} \\ b_{2k} \\ b_{3k} \end{bmatrix}$ for each $k = 1, 2, 3, 4$.

Then, for each $k = 1, 2, 3, 4$, we have $AB_{\text{col-}k} = a_1b_{1k} + a_2b_{2k} + a_3b_{3k}$. This is the k -th entry in the (1×4) -row matrix AB . So

$$\begin{aligned} (AB)C &= (a_1b_{11} + a_2b_{21} + a_3b_{31})c_1 && \leftarrow \text{This is } (AB_{\text{col-1}})c_1 \\ &+ (a_1b_{12} + a_2b_{22} + a_3b_{32})c_2 && \leftarrow \text{This is } (AB_{\text{col-2}})c_2 \\ &+ (a_1b_{13} + a_2b_{23} + a_3b_{33})c_3 && \leftarrow \text{This is } (AB_{\text{col-3}})c_3 \\ &+ (a_1b_{14} + a_2b_{24} + a_3b_{34})c_4 && \leftarrow \text{This is } (AB_{\text{col-4}})c_4 \\ &= a_1b_{11}c_1 + a_1b_{12}c_2 + a_1b_{13}c_3 + a_1b_{14}c_4 \\ &\quad + a_2b_{21}c_1 + a_2b_{22}c_2 + a_2b_{23}c_3 + a_2b_{24}c_4 \\ &\quad + a_3b_{31}c_1 + a_3b_{32}c_2 + a_3b_{33}c_3 + a_3b_{34}c_4 \end{aligned}$$

Handwritten notes: The matrix AB is shown as $[AB_{\text{col-1}} \mid AB_{\text{col-2}} \mid AB_{\text{col-3}} \mid AB_{\text{col-4}}]$ with each column circled in red. Red arrows point from each circled column to the corresponding term in the expansion of $(AB)C$. Blue arrows point from the text "This is" to the corresponding circled column.

which is the sum of all the $a_jb_{jk}c_k$'s, each copy exactly once.

- We have $B = \begin{bmatrix} B_{\text{row-1}} \\ B_{\text{row-2}} \\ B_{\text{row-3}} \end{bmatrix}$, in which $B_{\text{row-}j} = [b_{j1} \ b_{j2} \ b_{j3} \ b_{j4}]$ for each $j = 1, 2, 3$.

Then, for each $j = 1, 2, 3$, we have $B_{\text{row-}j}C = b_{j1}c_1 + b_{j2}c_2 + b_{j3}c_3 + b_{j4}c_4$. This is the j -th entry in the (3×1) -column matrix BC . So

$$\begin{aligned}
 A(BC) &= a_1(b_{11}c_1 + b_{12}c_2 + b_{13}c_3 + b_{14}c_4) \\
 &\quad + a_2(b_{21}c_1 + b_{22}c_2 + b_{23}c_3 + b_{24}c_4) \\
 &\quad + a_3(b_{31}c_1 + b_{32}c_2 + b_{33}c_3 + b_{34}c_4) \\
 &= a_1b_{11}c_1 + a_1b_{12}c_2 + a_1b_{13}c_3 + a_1b_{14}c_4 \\
 &\quad + a_2b_{21}c_1 + a_2b_{22}c_2 + a_2b_{23}c_3 + a_2b_{24}c_4 \\
 &\quad + a_3b_{31}c_1 + a_3b_{32}c_2 + a_3b_{33}c_3 + a_3b_{34}c_4
 \end{aligned}$$

which is the sum of all the $a_j b_{jk} c_k$'s, each copy exactly once.

Hence $(AB)C = A(BC)$ in this particular case indeed.

- We have $B = \begin{bmatrix} B_{\text{row-1}} \\ B_{\text{row-2}} \\ B_{\text{row-3}} \end{bmatrix}$, in which $B_{\text{row-}j} = [b_{j1} \ b_{j2} \ b_{j3} \ b_{j4}]$ for each $j = 1, 2, 3$.

Then, for each $j = 1, 2, 3$, we have $B_{\text{row-}j}C = b_{j1}c_1 + b_{j2}c_2 + b_{j3}c_3 + b_{j4}c_4$. This is the j -th entry in the (3×1) -column matrix BC . So

$$\begin{aligned}
 A(BC) &= a_1(b_{11}c_1 + b_{12}c_2 + b_{13}c_3 + b_{14}c_4) \quad \leftarrow \text{This is } a_1 (B_{\text{row-1}}C) \\
 &+ a_2(b_{21}c_1 + b_{22}c_2 + b_{23}c_3 + b_{24}c_4) \quad \leftarrow \text{This is } a_2 (B_{\text{row-2}}C) \\
 &+ a_3(b_{31}c_1 + b_{32}c_2 + b_{33}c_3 + b_{34}c_4) \quad \leftarrow \text{This is } a_3 (B_{\text{row-3}}C) \\
 &= a_1b_{11}c_1 + a_1b_{12}c_2 + a_1b_{13}c_3 + a_1b_{14}c_4 \\
 &+ a_2b_{21}c_1 + a_2b_{22}c_2 + a_2b_{23}c_3 + a_2b_{24}c_4 \\
 &+ a_3b_{31}c_1 + a_3b_{32}c_2 + a_3b_{33}c_3 + a_3b_{34}c_4
 \end{aligned}$$

which is the sum of all the $a_j b_{jk} c_k$'s, each copy exactly once.

Hence $(AB)C = A(BC)$ in this particular case indeed.

7. Theorem (3). (Associativity of matrix multiplication.)

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

Proof.

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

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

• 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

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

7. **Theorem (3).** (Associativity of matrix multiplication.)

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

Proof.

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

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

• 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 -th row is $A_i B$. The ℓ -th column is C_ℓ . The (i, ℓ) -th entry is $(A_i B)C_\ell$.

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

Then

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

By Lemma (2), the equality $(A_i B)C_\ell = A_i(BC_\ell)$ holds for each i and for each ℓ .

Then $(AB)C = A(BC)$.

- 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 -th row is A_i .

The ℓ -th column is BC_ℓ .

The (i, ℓ) -th entry is $A_i(BC_\ell)$.

By Lemma (2), the equality $(A_i B)C_\ell = A_i(BC_\ell)$ holds for each i and for each ℓ .

Then $(AB)C = A(BC)$.

8. Matrix multiplication for ‘block matrices’, introduced through examples.

(a) Let A be an $(m \times n)$ -matrix, and $B = [B_1 | B_2 | \cdots | B_s]$, in which B_1, B_2, \cdots, B_s be matrices all with n rows and with p_1, p_2, \cdots, p_s columns respectively.

$$\text{Then } AB = A [B_1 | B_2 | \cdots | B_s] = [AB_1 | AB_2 | \cdots | AB_s].$$

Remark. When $s = p$ and B_1, B_2, \cdots, B_p are the individual columns of B , the above formula reduces to the definition for the product AB .

(b) Let $A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_s \end{bmatrix}$, in which A_1, A_2, \cdots, A_s be matrices all with n columns and with

m_1, m_2, \cdots, m_s rows respectively, and B be an $(n \times p)$ -matrix.

$$\text{Then } AB = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_s \end{bmatrix} B = \begin{bmatrix} A_1 B \\ A_2 B \\ \vdots \\ A_s B \end{bmatrix}.$$

Remark. When $s = m$ and A_1, A_2, \cdots, A_m are the individual rows of A , the above formula reduces to the definition for the product AB .

(c) Let $A = [A_1 | A_2 | \cdots | A_s]$, in which A_1, A_2, \cdots, A_s are matrices with m rows and with n_1, n_2, \cdots, n_s columns respectively

Let $B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_s \end{bmatrix}$, in which B_1, B_2, \cdots, B_s are matrices with n_1, n_2, \cdots, n_s rows respec-

tively and with p columns.

(So for each k , the number of columns of A_k and the number of rows of B_k are the same, and note that $A_k B_k$ is an $(m \times p)$ -matrix. Also note that AB is an $(m \times p)$ -matrix)

Then $AB = A_1 B_1 + A_2 B_2 + \cdots + A_s B_s$.

Illustration.

$$\text{Let } A = \left[\begin{array}{ccc|ccc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \end{array} \right], \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \\ b_{51} & b_{52} \\ b_{61} & b_{62} \end{bmatrix}.$$

$$\text{Let } A_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad A_2 = \begin{bmatrix} a_{14} \\ a_{24} \\ a_{34} \end{bmatrix}, \quad A_3 = \begin{bmatrix} a_{15} & a_{16} \\ a_{25} & a_{26} \\ a_{35} & a_{36} \end{bmatrix}.$$

$$\text{Let } B_1 = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}, B_2 = \begin{bmatrix} b_{41} & b_{42} \end{bmatrix}, B_3 = \begin{bmatrix} b_{51} & b_{52} \\ b_{61} & b_{62} \end{bmatrix}.$$

$$\text{We have } A = \left[A_1 \mid A_2 \mid A_3 \right], B = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}. \text{ Note that}$$

$$A_1 B_1 = \begin{bmatrix} \sum_{k=1}^3 a_{1k} b_{k1} & \sum_{k=1}^3 a_{1k} b_{k2} \\ \sum_{k=1}^3 a_{2k} b_{k1} & \sum_{k=1}^3 a_{2k} b_{k2} \\ \sum_{k=1}^3 a_{3k} b_{k1} & \sum_{k=1}^3 a_{3k} b_{k2} \end{bmatrix}, A_2 B_2 = \begin{bmatrix} a_{14} b_{41} & a_{14} b_{42} \\ a_{24} b_{41} & a_{24} b_{42} \\ a_{34} b_{41} & a_{34} b_{42} \end{bmatrix}, A_3 B_3 = \begin{bmatrix} \sum_{k=5}^6 a_{1k} b_{k1} & \sum_{k=5}^6 a_{1k} b_{k2} \\ \sum_{k=5}^6 a_{2k} b_{k1} & \sum_{k=5}^6 a_{2k} b_{k2} \\ \sum_{k=5}^6 a_{3k} b_{k1} & \sum_{k=5}^6 a_{3k} b_{k2} \end{bmatrix}.$$

Then

$$A_1 B_1 + A_2 B_2 + A_3 B_3 = \begin{bmatrix} \sum_{k=1}^6 a_{1k} b_{k1} & \sum_{k=1}^6 a_{1k} b_{k2} \\ \sum_{k=1}^6 a_{2k} b_{k1} & \sum_{k=1}^6 a_{2k} b_{k2} \\ \sum_{k=1}^6 a_{3k} b_{k1} & \sum_{k=1}^6 a_{3k} b_{k2} \end{bmatrix} = AB$$

indeed.

(d) Let $A_{11}, A_{12}, A_{21}, A_{22}, B_{11}, B_{12}, B_{21}, B_{22}$ be matrices. Suppose that

- the number of rows of A_{11}, A_{12} are the same,
- the number of rows of A_{21}, A_{22} are the same,
- the number of columns of B_{11}, B_{21} are the same,
- the number of columns of B_{12}, B_{22} are the same,
- the number of columns of each of A_{11}, A_{21} is the same as the number of rows of each of B_{11}, B_{12} ,
- the number of columns of each of A_{12}, A_{22} is the same as the number of rows of each of B_{21}, B_{22} .

(So there are integers $m_1, m_2, n_1, n_2, p_1, p_2$ so that A_{11} is an $(m_1 \times n_1)$ -matrix, A_{12} is an $(m_1 \times n_2)$ -matrix, A_{21} is an $(m_2 \times n_1)$ -matrix, A_{22} is an $(m_2 \times n_2)$ -matrix, B_{11} is an $(n_1 \times p_1)$ -matrix, B_{12} is an $(n_1 \times p_2)$ -matrix, B_{21} is an $(n_2 \times p_1)$ -matrix, B_{22} is an $(n_2 \times p_2)$ -matrix.)

$$\text{Define } A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right], \quad B = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array} \right].$$

$$\text{Then } AB = \left[\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array} \right].$$

Illustration.

Let

$$A = \left[\begin{array}{cc|cccc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \end{array} \right], \quad B = \left[\begin{array}{cc|cccc} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} & b_{17} \\ b_{21} & b_{22} & b_{23} & a_{24} & b_{25} & b_{26} & b_{27} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} & b_{37} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} & b_{46} & b_{47} \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} & b_{56} & b_{57} \\ b_{61} & b_{62} & b_{63} & b_{64} & b_{65} & b_{66} & b_{67} \end{array} \right].$$

Write

$$A_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad A_{12} = \begin{bmatrix} a_{13} & a_{14} & a_{15} & a_{16} \\ a_{23} & a_{24} & a_{25} & a_{26} \end{bmatrix},$$
$$A_{21} = \begin{bmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \\ a_{51} & a_{52} \end{bmatrix}, \quad A_{22} = \begin{bmatrix} a_{33} & a_{34} & a_{35} & a_{36} \\ a_{43} & a_{44} & a_{45} & a_{46} \\ a_{53} & a_{54} & a_{55} & a_{56} \end{bmatrix},$$

and

$$B_{11} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \quad B_{12} = \begin{bmatrix} b_{13} & b_{14} & b_{15} & b_{16} & b_{17} \\ b_{23} & a_{24} & b_{25} & b_{26} & b_{27} \end{bmatrix}.$$
$$B_{21} = \begin{bmatrix} b_{31} & b_{32} \\ b_{41} & b_{42} \\ b_{51} & b_{52} \\ b_{61} & b_{62} \end{bmatrix}, \quad B_{22} = \begin{bmatrix} b_{33} & b_{34} & b_{35} & b_{36} & b_{37} \\ b_{43} & b_{44} & b_{45} & b_{46} & b_{47} \\ b_{53} & b_{54} & b_{55} & b_{56} & b_{57} \\ b_{63} & b_{64} & b_{65} & b_{66} & b_{67} \end{bmatrix}.$$

So $A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right]$, $B = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array} \right]$. We have

$$\begin{aligned}
 A_{11}B_{11} + A_{12}B_{21} &= \dots = \left[\begin{array}{c|c} \sum_{k=1}^6 a_{1k}b_{k1} & \sum_{k=1}^6 a_{1k}b_{k2} \\ \hline \sum_{k=1}^6 a_{2k}b_{k1} & \sum_{k=1}^6 a_{2k}b_{k2} \end{array} \right], \\
 A_{11}B_{12} + A_{12}B_{22} &= \dots = \left[\begin{array}{c|c|c|c|c} \sum_{k=1}^6 a_{1k}b_{k3} & \sum_{k=1}^6 a_{1k}b_{k4} & \sum_{k=1}^6 a_{1k}b_{k5} & \sum_{k=1}^6 a_{1k}b_{k6} & \sum_{k=1}^6 a_{1k}b_{k7} \\ \hline \sum_{k=1}^6 a_{2k}b_{k3} & \sum_{k=1}^6 a_{2k}b_{k4} & \sum_{k=1}^6 a_{2k}b_{k5} & \sum_{k=1}^6 a_{2k}b_{k6} & \sum_{k=1}^6 a_{2k}b_{k7} \end{array} \right], \\
 A_{21}B_{11} + A_{22}B_{21} &= \dots = \left[\begin{array}{c|c} \sum_{k=1}^6 a_{3k}b_{k1} & \sum_{k=1}^6 a_{3k}b_{k2} \\ \hline \sum_{k=1}^6 a_{4k}b_{k1} & \sum_{k=1}^6 a_{4k}b_{k2} \\ \hline \sum_{k=1}^6 a_{5k}b_{k1} & \sum_{k=1}^6 a_{5k}b_{k2} \end{array} \right],
 \end{aligned}$$

$$A_{21}B_{12} + A_{22}B_{22} = \cdots = \left[\begin{array}{ccccc} \sum_{k=1}^6 a_{3k}b_{k3} & \sum_{k=1}^6 a_{3k}b_{k4} & \sum_{k=1}^6 a_{3k}b_{k5} & \sum_{k=1}^6 a_{3k}b_{k6} & \sum_{k=1}^6 a_{3k}b_{k7} \\ \sum_{k=1}^6 a_{4k}b_{k3} & \sum_{k=1}^6 a_{4k}b_{k4} & \sum_{k=1}^6 a_{4k}b_{k5} & \sum_{k=1}^6 a_{4k}b_{k6} & \sum_{k=1}^6 a_{4k}b_{k7} \\ \sum_{k=1}^6 a_{5k}b_{k3} & \sum_{k=1}^6 a_{5k}b_{k4} & \sum_{k=1}^6 a_{5k}b_{k5} & \sum_{k=1}^6 a_{5k}b_{k6} & \sum_{k=1}^6 a_{5k}b_{k7} \end{array} \right].$$

So the equality $AB = \left[\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array} \right]$ holds indeed.

The recurrent feature in the manipulation is the fact that the (i, j) -th entry of AB , which is

$$\sum_{k=1}^6 a_{ik}b_{kj},$$

can be re-written as

$$\sum_{k=1}^2 a_{ik}b_{kj} + \sum_{k=3}^6 a_{ik}b_{kj}.$$