Chapter 14 : Matrix Multiplication.

Matrix multiplication shows up as a key tool in a number of very concrete applications (not necessarily related to linear systems), such as probability theory, economics, geometry, quantum theory, for solving linear systems, for vector spaces, etc.

* Matrix multiplication is a generalization of the usual multiplication of numbers.

Example: $W = \begin{pmatrix} 12 & 10 & 14 \\ 8 & 8 & 10 \end{pmatrix}$, $H = \begin{pmatrix} 10 & 0 \\ 10 & 10 \\ 5 & 10 \end{pmatrix}$

$$C = WH = \begin{pmatrix} 12 & 10 & 14 \\ 8 & 8 & 10 \end{pmatrix} \begin{pmatrix} 10 & 0 \\ 10 & 10 \\ 5 & 10 \end{pmatrix} =$$

$$\begin{pmatrix} 12\times10 + 10\times10 + 14\times5 & 12\times0 + 10\times10 + 14\times10 \\ 8\times10 + 8\times10 + 10\times5 & 8\times0 + 8\times10 + 10\times10 \end{pmatrix}$$

$$= \begin{pmatrix} 290 & 240 \\ 210 & 180 \end{pmatrix}$$

Remarks:

(1) we organized things so that the rows of the answer correspond to the rows of the first matrix; the columns of the answer correspond to the columns of the second matrix; and the variable over which we summed corresponds to both the columns of the first and the rows of the second matrix.

(2) we calculated the entry in row i and column j of the product by taking the dot product of row i of the 1st matrix and column j of the 2nd matrix.

Definition: If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then their product AB is the $m \times p$ matrix whose (i,j) entry is the dot product of the ith row of A with jth column of B.

So, if $A = [a_{ij}]$ and $B = [b_{ij}]$, then $AB = C = [c_{ij}]$, where

$$\underbrace{c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}} = [a_{i1} \;\; a_{i2} \; \cdots \; a_{in}] \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix}.$$

In particular, "AB" only makes sense if the number of columns of A equals the number of rows of B.

Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ 4x + 5y + 6z \end{pmatrix} = x\begin{pmatrix} 1 \\ 4 \end{pmatrix} + y\begin{pmatrix} 2 \\ 5 \end{pmatrix} + z\begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

(A linear system or a linear combination).

Example:

$$(1 \;\; 2 \;\; 3)\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = (a+2c+3e \quad b+2d+3f)$$

$$= 1\,(a \;\; b) + 2\,(c \;\; d) + 3\,(e \;\; f)$$

linear combination of rows!

Properties:

(1) Sometimes we can calculate AB, but BA is "not defined"!

example:

$$A=\begin{pmatrix}1&2\\3&4\end{pmatrix}_{2\times 2},\quad B=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}_{2\times 3} \Rightarrow AB=\begin{pmatrix}9&12&15\\19&26&33\end{pmatrix}_{2\times 3},$$

but "BA" is not defined, because the number of columns of B is not the same as the number of rows of A.

(2) In general, $AB \neq BA$. (the product is not commutative).

example:

$$A=\begin{pmatrix}1&2\\3&4\end{pmatrix},\quad B=\begin{pmatrix}0&1\\1&0\end{pmatrix}$$

$$\left.\begin{aligned} AB &= \begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}=\begin{pmatrix}2&1\\4&3\end{pmatrix},\\ BA &= \begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1&2\\3&4\end{pmatrix}=\begin{pmatrix}3&4\\1&2\end{pmatrix}\end{aligned}\right\} AB \neq BA.$$

example: $A=\begin{pmatrix}1\\2\\3\end{pmatrix}$, $B=\begin{pmatrix}4\\5\\6\end{pmatrix}$, then both AB and BA are not defined. However,

- $A^T B = (1\ \ 2\ \ 3)\begin{pmatrix}4\\5\\6\end{pmatrix} = 32$, which is just the dot product of the two vectors A & B.
- $B^T A = (4\ \ 5\ \ 6)\begin{pmatrix}1\\2\\3\end{pmatrix} = 32$, which is again the dot product of two vectors (we know that dot product of vectors is in fact commutative).

· Note: A 1×1 matrix $[a]$ is identified with its single entry a.

· $$AB^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}_{3\times 1} \begin{pmatrix} 4 & 5 & 6 \end{pmatrix}_{1\times 3} = \begin{pmatrix} 4 & 5 & 6 \\ 8 & 10 & 12 \\ 12 & 15 & 18 \end{pmatrix}$$ (tensor product!)

(3) It can happen that $AB=0$, but neither A nor B is zero.

example: $C = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$, then

$$CD = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

$(CD = 0)$

4) It can happen that $AC = BC$, but $C \neq 0$ and $A \neq B$. In other words, we cannot cancel out C, even if $C \neq 0$.

example: $A = \begin{pmatrix} 4 & -1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 7 & -1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$

$$AC = \begin{pmatrix} 4 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 8 \end{pmatrix},$$

$$AB = \begin{pmatrix} 1 & 0 \\ 7 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 8 \end{pmatrix},$$

that is $AC = BC$, but $B \neq C$!

* properties of transpose operation on matrices:

(1) $(A+B)^T = A^T + B^T$,

(2) $(kA)^T = kA^T$, $k \in \mathbb{R}$,

(3) $(A^T)^T = A$.

* "Identity matrix" : square matrix whose diagonal elements are 1s,

for instance:

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and it exists for every "$\underline{n}$".

* "zero matrix": Any matrix (square or not), whose elements are all zeros:

$$0_{2\times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad 0_{3\times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

* other properties of matrix products: (A, B, C matrices & $k \in \mathbb{R}$)

(1) $(AB)C = A(BC)$,

(2) $A(B+C) = AB + AC$,

(3) $(B+C)A = BA + CA$,

(4) $k(AB) = (kA)B = A(kB)$,

(5) $(AB)^T = B^T A^T$ (Note the order),

(6) $AI = A$ and $IB = B$,

(7) If A is $m\times n$, then $A\,0_{n\times p} = 0_{m\times p}$ and $0_{q\times m} A = 0_{q\times n}$.

proofs:

(1) $A=[a_{ij}]$, $B=[b_{ij}]$, $C=[c_{ij}]$. Suppose A is $m\times n$, B is $n\times p$ and C is $p\times q$. Then, (i,j)th entry of $(AB)C$ is

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

$$= \sum_{k=1}^{n} a_{ik}\left(\sum_{l=1}^{p} b_{kl}c_{lj}\right) = A(BC) \checkmark$$

(2), (3), (4), (5), (6), (7): exercises!

examples: (assume that the multiplications are defined).

(1) Simplify $(A+B)(C+D)$:

$$(A+B)(C+D) = (A+B)C + (A+B)D = AC+BC+AD+BD.$$

(2) Simplify $(A+B)(A-B)$:

$$(A+B)(A-B) = (A+B)A - (A+B)B = A^2+BA-AB-B^2.$$

($AB \neq BA$ in general).

* Cayley-Hamilton theorem:

suppose A is a 2×2 square matrix, then we can define $A^2 = AA$ and $A^3 = AAA$ etc.

So, for instance if $A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$, then

$$A^2 = AA = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 2 & 4 \end{pmatrix}, \text{ and}$$

$$A^3 = AAA = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 5 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 9 & 10 \\ 10 & 4 \end{pmatrix}, \text{ etc.}$$

Here, $tr(A)$ (= sum of diagonal entries) $= 1$, and also $\det(A)$, for any $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\det(A) = ad - bc$, which here equals -4.

Then, the "characteristic Polynomial" of A is

$$\boxed{x^2 - tr(A)x + \det(A),} \text{ for any } 2\times2 \text{ matrix } A.$$

• Let's plug in A for x (and put an identity matrix at the end after $\det(A)$ so that everything becomes a matrix), then

$$\begin{pmatrix} 5 & 2 \\ 2 & 4 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} - 4\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \; !!$$

* **Block multiplication**: (simplify calculations)

Example:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Let's divide A into blocks by setting:

$$B = I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and thinking of A as :

$$A = \begin{pmatrix} B & 0_{2\times 3} \\ 0_{3\times 2} & C \end{pmatrix}, \text{ then,}$$

$$A^2 = AA = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix} = \begin{pmatrix} B^2 & 0 \\ 0 & C^2 \end{pmatrix} \cdots$$

$$\vdots$$

$$A^{100} = \begin{pmatrix} B^{100} & 0 \\ 0 & C^{100} \end{pmatrix}, \text{ where } B^{100} = I_2^{100} = I_2,$$

and $C^2 = 0$, so $C^{100} = (C^2)^{50} = 0$, and $A^{100} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$.

* Back to Linear systems :

Example :

$$\left.\begin{aligned} x + 2y + 3z &= 4, \\ x - y + z &= 2, \\ y - 3z &= 0 \end{aligned}\right\} \Rightarrow A\vec{x} = b, \text{ where}$$

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \\ 0 & 1 & -3 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix},$$

that is

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \\ 0 & 1 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}.$$

or,

$$x\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} + z\begin{pmatrix} 3 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix},$$

in vector form, which can be solved by row reducing the augmented matrix $[A|b]$:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 1 & -1 & 1 & 2 \\ 0 & 1 & -3 & 0 \end{array}\right).$$

In term of block multiplication, we can write

$$A = [C_1, C_2, C_3],$$ where C_i is the ith column of A, and

then,

$$Ax = [C_1, C_2, C_3]\begin{bmatrix} x \\ y \\ z \end{bmatrix} = xC_1 + yC_2 + zC_3.$$

This is true for any size of matrix A:

If $A = [C_1 \; C_2 \; \cdots \; C_n]$, then $Ax = x_1C_1 + x_2C_2 + \cdots + x_nC_n$,

which is just a linear combination of the columns of A.

Facts :

(1) Ax is a linear combination of the columns of A (with coefficients given by the column vector x).

(2) $Ax = b$ is consistent if and only if b is a linear combination of the columns of A.

(3) $Ax = 0$ has a unique solution if and only if the columns of A are LI.

Definition (14.2): Suppose $A = [C_1 C_2 \cdots C_n]$ be an $m \times n$ matrix with columns C_i. Set $\mathrm{Col}(A) = \mathrm{span}\{C_1, C_2, \dots, C_n\}$. This is a subspace of $\mathbb{R}^m$, called the "Column space of A".

(4) $\mathrm{Col}(A) = \mathbb{R}^m$ if & only if the columns of A span $\mathbb{R}^m$, if and only if $Ax = b$ is consistent for any $b \in \mathbb{R}^m$, if & only if there is no zero row in the REF of A.

(5) The columns of A are linearly independent if & only if $Ax=0$ has a unique solution, if and only if there is a leading one in every column of an REF of A, if and only if $\operatorname{rank}(A)$ equals the number of columns of A.

(6) If A is an $n \times n$ matrix, then the columns of A form a basis for $\mathbb{R}^n$ if and only if they are linearly independent, if and only if $\operatorname{rank}(A) = n$, if and only if $\operatorname{Col}(A) = \mathbb{R}^n$, if and only if the columns of A span $\mathbb{R}^n$, if and only if no REF of A has a zero row.