2.1 Matrix Operations.

Notation for matrices

$$A_{m\times n} = [\vec{a}_1 \ \vec{a}_2 \dots \vec{a}_n]$$

$$= \begin{array}{cccc} \vec{a}_1 & \vec{a}_2 & & \vec{a}_n \\ \left[\begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{matrix}\right] \end{array}$$

the numbers in the matrix are called entries.
They are identified by their position within the matrix
for example, a_{21} means the number that is located in row 2 and column 1.

In general, a_{ij} is the number located in row i and column j.

Another notation for A is

$$A = [a_{ij}]$$

The operations of sums and multiplication by scalars are obtained from these operations for vectors.

Def: For $A_{m\times n}$ and $B_{m\times n}$

$$A+B = [\vec{a}_1\ \vec{a}_2 \ldots \vec{a}_n] + [\vec{b}_1\ \vec{b}_2 \ldots \vec{b}_n]$$

$$\stackrel{def}{=} [\vec{a}_1+\vec{b}_1\ \ \vec{a}_2+\vec{b}_2 \ldots \vec{a}_n+\vec{b}_n] =$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \ldots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & & a_{mn}+b_{mn} \end{bmatrix} = [a_{ij}+b_{ij}]$$

Similar

$$cA = c[\vec{a}_1\ \vec{a}_2 \ldots \vec{a}_n] \stackrel{def}{=} [c\vec{a}_1\ c\vec{a}_2 \ldots c\vec{a}_n]$$

$$= \begin{bmatrix} ca_{11} & ca_{12} & \ldots & ca_{1n} \\ ca_{21} & ca_{22} & \ldots & ca_{2n} \\ \vdots & & & \vdots \\ ca_{m1} & ca_{m2} & & ca_{mn} \end{bmatrix} = [ca_{ij}]$$

The usual rules of algebra of real #s. apply to the sums and scalar multiplication of matrices. (See Thm 1).

Matrix Multiplication

Consider the two matrices

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \end{bmatrix}_{2\times 3} \quad \text{and} \quad B = \begin{array}{c} \begin{array}{cc} \vec{b}_1 & \vec{b}_2 \\ \downarrow & \downarrow \end{array} \\ \begin{bmatrix} 2 & 4 \\ 1 & 5 \\ 3 & 1 \end{bmatrix}_{3\times 2} \end{array}$$

We want to define the product of A and B, $\underline{AB}$

Notice that the matrices have <u>different sizes</u>.

The natural definition comes from the product of a matrix and a vector. Consider

$$A\vec{b}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1\begin{bmatrix} 2 \\ 5 \end{bmatrix} + 3\begin{bmatrix} 3 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 9 \\ 3 \end{bmatrix} = \begin{bmatrix} 13 \\ 12 \end{bmatrix}$$

$$A\vec{b}_2 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 5\begin{bmatrix} 2 \\ 5 \end{bmatrix} + 1\begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 8 \end{bmatrix} + \begin{bmatrix} 10 \\ 25 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 17 \\ 34 \end{bmatrix}$$

Therefore, a good candidate for the definition of AB is

$$AB = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 5 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 \end{bmatrix} = \begin{bmatrix} 13 & 17 \\ 12 & 34 \end{bmatrix}$$

Def.- Given $A_{m\times n} = [\vec{a}_1 \dots \vec{a}_n]$ and $B_{n\times p} = [\vec{b}_1 \dots \vec{b}_p]$

The product AB is defined as the matrix $m\times p$ given by

$$AB_{m\times p} = [A\vec{b}_1 \; A\vec{b}_2 \dots A\vec{b}_p]$$

Notice that for the product AB to be defined The number of columns of A should be equal to the number of rows of B.

Therefore, the product AB is not always defined.

In our example, AB is defined but BA is not.

So the familiar property of real numbers

$$ab = ba$$

is not true for matrices, since BA may not even be defined. Also if defined, $BA \neq AB$ in general.

A more efficient way to compute AB which is also used as an alternative definition for the product is illustrated with our previous example:

Notice that

$$A\vec{b}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \end{bmatrix} \begin{array}{c} \vec{b}_1 \\ \downarrow \\ \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \end{array} = 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1\begin{bmatrix} 2 \\ 5 \end{bmatrix} + 3\begin{bmatrix} 3 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 9 \\ 3 \end{bmatrix} = \begin{bmatrix} 2+2+9 \\ 4+5+3 \end{bmatrix} =$$

$$= \begin{bmatrix} 1.2 + 2.1 + 3.3 \\ 2.2 + 5.1 + 1.3 \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \end{bmatrix}$$

Similarly,

$$A\vec{b}_2 = \begin{bmatrix} a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}$$

In general, $A_{m\times n} B_{n\times p} = (AB)_{m\times p}$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \\ a_{m1} & a_{m2} & & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} (AB)_{11} & \cdots & (AB)_{1p} \\ \vdots & \vdots & \vdots \\ & (AB)_{ij} & \\ (AB)_{m1} & \cdots & (AB)_{mp} \end{pmatrix}$$

Show that $A_{m\times n}$, $B_{n\times p}$

$$A(B\vec{x}) \overset{(*)}{=} (AB)\vec{x}.$$

Proof.-

$$A(B\vec{x}) = A(x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_p\vec{b}_p) =$$

$$= x_1 A\vec{b}_1 + x_2 A\vec{b}_2 + \cdots + x_p A\vec{b}_p =$$

$$= [A\vec{b}_1 \ A\vec{b}_2 \ \cdots \ A\vec{b}_p]\vec{x} = (AB)\vec{x}.$$

Show that for $A_{m\times n}$, $B_{n\times p}$, and $C_{p\times q}$

$$\boxed{(AB)C = A(BC).}$$

Proof.- $C = [\vec{c}_1 \ \ldots \ \vec{c}_q]$

First, $BC \overset{def}{=} [B\vec{c}_1 \ B\vec{c}_2 \ldots \ B\vec{c}_q]$

then $\underline{A(BC)} \overset{def}{=} [A(B\vec{c}_1) \ A(B\vec{c}_2) \ \ldots A(B\vec{c}_q)] \overset{\text{from } (*)}{=}$

$$= [(AB)\vec{c}_1 \ \ (AB)\vec{c}_2 \cdots \ (AB)\vec{c}_q] =$$

$$\overset{def}{=} \underline{(AB)C}$$

Exercise #10 $\quad AB = AC \not\Rightarrow B = C.$

<u>Def.</u>- $A_{n\times n}$ $\quad A^k = \overbrace{A.A\ldots A}^{k}$

$A^0 =$ Identity matrix I.

Transpose of a matrix

Consider $A_{2\times 3}$ again

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 1 \end{bmatrix}$$

a new matrix called $A^T_{3\times 2}$ (A transpose) can be obtained from A interchanging the rows and columns of A. It means

$$A^T = \begin{bmatrix} 1 & 2 \\ 2 & 5 \\ 3 & 1 \end{bmatrix}$$

In general,

<u>Definition</u>. Given $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$

The transpose of A, A^T is a matrix $n\times m$ given by

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & & a_{mn} \end{bmatrix}$$

Thm 3 A, B with appropriate sizes

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

b) $(A+B)^T = A^T + B^T$

c) For any $r \in \mathbb{R}$: $(rA)^T = rA^T$

d) $(AB)^T = B^T A^T$.

All immediate for the definition with the exception of (d).

Exercise #33

Consider the two matrices $A_{m\times r}$ and $B_{r\times n}$

Show that $(AB)^T = B^T A^T$

Proof:- First thing to notice is that $(AB)_{m\times n}$ and $(B^T A^T)_{n\times m}$

$$AB = \begin{array}{r} \\ \\ i\text{th row} \\ \\ \\ \end{array} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{ir} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mr} \end{bmatrix} \begin{array}{c} \begin{array}{ccccc} & & & j\text{th column} & \end{array} \\ \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & & & & & \\ b_{r1} & b_{r2} & \cdots & b_{rj} & \cdots & b_{rn} \end{bmatrix} \end{array}$$

Also,

if $A = [a_{ij}]_{m\times n}$

then $A^T = [a_{ji}]_{n\times m}$

then,

$$AB = [(AB)_{ij}] \overset{(*)}{=} [a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj}]$$

then,

$$(AB)^T = [(AB)^T_{ij}] = [(AB)_{ji}] \overset{\text{from } (*)}{=} [a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jr}b_{ri}]$$

On the other hand,

$$A^T = [\,\hat{a}_{ij}\,] \overset{(1)}{=} [a_{ji}]_{r\times m}\ ,\quad B^T = [\,\hat{b}_{ij}\,] \overset{(2)}{=} [\,b_{ji}\,]_{n\times r}$$

then,

$$B^T A^T = [\,(B^T A^T)_{ij}\,] \overset{\text{from } (*)}{=} [\,\hat{b}_{i1}\,\hat{a}_{1j} + \hat{b}_{i2}\,\hat{a}_{2j} + \cdots + \hat{b}_{ir}\,\hat{a}_{rj}\,]$$

$$\overset{\substack{\text{from } (1)\\ \text{and } (2)}}{=\!=} [\,b_{1i}\,a_{j1} + b_{2i}\,a_{j2} + \cdots + b_{ri}\,a_{jr}\,]$$

$$\overset{\text{from } (*)}{=\!=} [\,(AB)_{ji}\,] = [\,(AB)^T_{ij}\,] = (AB)^T.$$