Lecture 6

(6.0) Recall that we have been studying algebraic operations on matrices. - Addition, Scalar multiplication and matrix multiplication

Today we will write "formal algebraic" properties of these operations. Most of these properties are obvious, and are listed here only for their use in the future – to define abstract vector spaces.

Notation. $O_{m\times n}$ denotes the $m\times n$ matrix all of whose entries $= 0$.

(6.1) Algebraic properties of addition and scalar multiplication.

Fix m, n : two positive integers. Assume that A, B and C are three $m\times n$ matrices. Then :

(1) $A + B = B + A$.

(2) $(A+B)+C = A+(B+C)$.

(3) $A + O_{m\times n} = A$.

(4) $P = (-1)A$ is the only matrix such that $A + P = O_{m\times n}$

(5) For every number λ, $\lambda(A+B) = \lambda A + \lambda B$.

(6) For any two numbers λ_1 and λ_2 :

$$(\lambda_1 + \lambda_2)A = \lambda_1 A + \lambda_2 A.$$

$$\lambda_1(\lambda_2 A) = (\lambda_1\lambda_2)A.$$

I am not going to prove these properties since these are clear from the definition of addition & scalar multiplication (These properties (1) – (6) clearly hold if A, B and C were just numbers. When these are matrices, the operations of scalar multiplication and addition are performed entry-by-entry, and each entry is just a number.)

(6.2) Algebraic properties of matrix multiplication

(7) If $\boxed{\begin{array}{l} A \text{ is an } m \times n \text{ matrix} \\ B \text{ is an } n \times p \text{ matrix} \\ C \text{ is a } p \times q \text{ matrix} \end{array}}$, then

$\boxed{(AB)C = A(BC)}$ (Note: both sides are defined, and are $m \times q$ size matrices).

(8) If A_1 and A_2 are $m \times n$ matrices, B is $n \times p$ matrix, then

$$(A_1 + A_2)B = A_1B + A_2B.$$

(9) If A is $m \times n$ matrix, B_1 and B_2 two $n \times p$ matrices then $A(B_1 + B_2) = AB_1 + AB_2$.

(10) A : $m \times n$ matrix
B : $n \times p$ matrix and λ : a number

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

(6.3) Properties listed (1) – (10) in the last two pages look exactly <u>the same</u> as <u>those for ordinary numbers</u>. The point where matrix operations <u>differ</u> in their features from those of numbers is:

$\boxed{AB \neq BA}$ (even when both sides are defined)

for instance $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ two 2×2 matrices.

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$$

$\neq$

$$BA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$$

In mathematics, people refer to this feature as:

"matrix multiplication is <u>NOT commutative</u>".

While property (7): $\boxed{(AB)C = A(BC)}$ is referred to as:

"matrix multiplication is <u>associative</u>."

(6.4) Transpose of a matrix.-

If $A = (a_{ij})$ is an $m \times n$ matrix, then the transpose of A, denoted by A^T, is an $n \times m$ matrix whose $(i,j)^{th}$ entry is $(j,i)^{th}$ entry of A

$$(A^T)_{ij} = a_{ji}$$

Thus A^T is obtained from A by switching the row and column indices. For example -

$$A = \begin{bmatrix} \boxed{1 \quad 2 \quad 3} \\ -2 \quad 4 \quad 0 \end{bmatrix}_{2 \times 3}$$

$$\leadsto \quad A^T = \begin{bmatrix} \boxed{\begin{matrix} 1 \\ 2 \\ 3 \end{matrix}} & \begin{matrix} -2 \\ 4 \\ 0 \end{matrix} \end{bmatrix}_{3 \times 2}$$

(list rows of A as columns of A^T)

$$A = \begin{bmatrix} 1 \\ 2 \\ 7 \\ 4 \end{bmatrix}_{4 \times 1} \quad \leadsto \quad A^T = \begin{bmatrix} 1 & 2 & 7 & 4 \end{bmatrix}_{1 \times 4}$$

(4-dim. vector)

(6.5) The operation of taking transpose allows us to view <u>the dot product of two vectors</u> <u>as matrix multiplication</u>

Recall that last time we introduced <u>dot product</u>:

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad \rightsquigarrow \quad \boxed{\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n}$$

Two n-dim. vectors in $\mathbb{R}^n$ (i.e. two $n \times 1$ matrices)

<u>Note</u>: $\underbrace{\vec{u} \cdot \vec{v}}_{\text{dot product}} = \vec{u}^T \vec{v}$

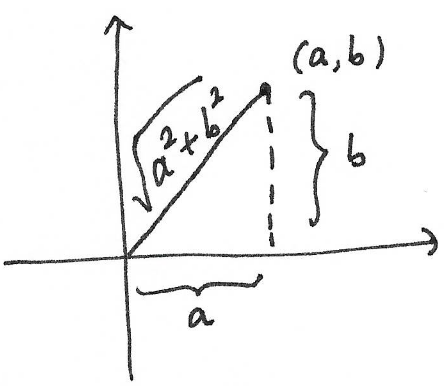
matrix multiplication of an $1 \times n$ matrix with $n \times 1$ matrix, resulting in a 1×1 matrix, i.e. a number

Since $$\vec{u}^T \vec{v} = [u_1 \ u_2 \ \dots \ u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = \vec{u} \cdot \vec{v}.$$

(6.6) <u>Euclidean length or norm of a vector.</u>

Recall, from Calculus, that length of the line segment joining $(0,0)$ to (a,b) is given by

$$\sqrt{a^2 + b^2}$$



This notion generalizes to $\mathbb{R}^n$ as follows:

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \rightsquigarrow \|\vec{u}\| = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2}$$

(<u>norm</u> of $\vec{u}$,
or <u>length</u> of $\vec{u}$)

Note: $\boxed{\|\vec{u}\|^2 = \vec{u} \cdot \vec{u} = \vec{u}^T \vec{u}}$

e.g. $\vec{x} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

- $\vec{x}^T \vec{y} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = 1(2) + 0(1) + (-1)3 = -1.$

- $\vec{x}^T \vec{x} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 1^2 + 0^2 + (-1)^2 = 2$

  So $\|\vec{x}\| = \sqrt{2}$.

- $\vec{y}^T \vec{y} = \begin{bmatrix} 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = 2^2 + 1^2 + 3^2 = 14$

  So $\|\vec{y}\| = \sqrt{14}$.

- $\|\vec{x} - \vec{y}\| = \left\| \begin{bmatrix} -1 \\ -1 \\ -4 \end{bmatrix} \right\| = \sqrt{\begin{bmatrix} -1 & -1 & -4 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ -4 \end{bmatrix}} = \sqrt{(-1)^2 + (-1)^2 + (-4)^2} = \sqrt{18}.$

(6.7) Transpose vs. Matrix addition and multiplication

Let A and B be two $m \times n$ matrices, and C be an $n \times p$ matrix. Then

(1) $(A+B)^T = A^T + B^T$. (both sides are $n \times m$ matrices)

(2) $(AC)^T = C^T A^T$.

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

Note that (1) and (3) follow easily from the definitions.

Let us see a proof of (2).

$$\underset{m \times n}{A} = (a_{ij}) \qquad \underset{n \times p}{C} = (c_{k\ell}) \rightsquigarrow AC \text{ is } m \times p \text{ matrix}$$

$$\rightsquigarrow (AC)^T \text{ is } p \times m \text{ matrix}$$

For $1 \leq i \leq p$; $1 \leq j \leq m$;

$$\boxed{\begin{aligned}(i,j)^{th} \text{ entry of } (AC)^T &= (j,i)^{th} \text{ entry of } AC \\ &= a_{j1}c_{1i} + a_{j2}c_{2i} + \dots + a_{jn}c_{ni}\end{aligned}} \quad -(*)$$

Similarly $(i,j)^{th}$ entry of $C^T A^T$

$$= (C^T)_{i1}(A^T)_{1j} + (C^T)_{i2}(A^T)_{2j} + \dots + (C^T)_{in}(A^T)_{nj}$$

$$= c_{1i}a_{j1} + c_{2i}a_{j2} + \dots + c_{ni}a_{jn} \quad -(**)$$

Comparing $(*)$ and $(**)$ we see that $\boxed{(AC)^T = C^T A^T}$.