Section 2.1: The Algebra of the Matrix

Linear algebra often deals with solving systems of linear equations

eg
$$\begin{aligned} -3x_1 - 3x_2 + x_3 &= -8 \\ 5x_1 \qquad\quad - 2x_3 &= 20 \\ x_1 - 4x_2 + 3x_3 &= -9 \end{aligned}$$

Consider the first equation, $-3x_1 - 3x_2 + x_3 = -8$

We can think of the coefficients as the components of a vector $\underline{r}_1 = \begin{bmatrix} -3 \\ -3 \\ 1 \end{bmatrix}$ and the unknowns as the components of a vector $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Thus the equation can be rewritten as $\underline{r}_1 \cdot \underline{x} = -8$.

Likewise if $\underline{r}_2 = \begin{bmatrix} 5 \\ 0 \\ -2 \end{bmatrix}$ and $\underline{r}_3 = \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}$ then the other equations become $\underline{r}_2 \cdot \underline{x} = 20$ and $\underline{r}_3 \cdot \underline{x} = -9$.

We can write this more compactly as a <u>matrix</u>.

<u>Def'n</u>: A <u>matrix</u> is a rectangular array of numbers enclosed in square (or round) brackets. It is said to be an $m \times n$ matrix if it has m rows and n columns. Here, $m \times n$ is the <u>size</u> of the matrix.

eg $A = \begin{bmatrix} -3 & -3 & 1 \\ 5 & 0 & -2 \\ 1 & -4 & 3 \end{bmatrix}$ is a 3×3 matrix

$$B = \begin{bmatrix} 1 & 2 \\ -1 & 4 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} \quad \text{is a } 4\times 2 \text{ matrix}$$

$$C = \begin{bmatrix} 1 & 6 & -2 & 0 & -2 \end{bmatrix} \quad \text{is a } 1\times 5 \text{ matrix}$$

Here, C is called a <u>row matrix</u>.

$$D = \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix} \quad \text{is a } 3\times 1 \text{ Matrix}$$

Here, D is a <u>column matrix</u> but it is also a 3-dimensional vector.

Hence an $m\times n$ matrix can be viewed as being comprised of n vectors of dimension m, so we can write

$$A = \begin{bmatrix} \underline{a}_1 & \underline{a}_2 & \underline{a}_3 & \cdots & \underline{a}_n \\ \downarrow & \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

<u>Def'n</u>: The <u>transpose</u> of an $m\times n$ matrix A is the $n\times m$ matrix whose rows are the columns of A, and whose columns are the rows of A, in the same order. It is denoted A^T.

eg $A = \begin{bmatrix} 2 & 5 & -1 \\ 0 & -3 & 6 \end{bmatrix}$ then $A^T = \begin{bmatrix} 2 & 0 \\ 5 & -3 \\ -1 & 6 \end{bmatrix}$

Thus row matrices and column matrices are transposes of each other, and so row matrices can be viewed as the transposes of vectors.

eg $\begin{bmatrix} -4 \\ 3 \\ 1 \end{bmatrix}^T = \begin{bmatrix} -4 & 3 & 1 \end{bmatrix}$

Hence an $m \times n$ matrix can also be written as

$$A = \begin{bmatrix} \underline{a}_1^T \rightarrow \\ \underline{a}_2^T \rightarrow \\ \vdots \\ \underline{a}_m^T \rightarrow \end{bmatrix}$$

eg Given the matrix $A = \begin{bmatrix} 2 & -1 & 5 \\ 6 & 0 & 0 \end{bmatrix}$,

the columns of A are the vectors $\begin{bmatrix} 2 \\ 6 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \end{bmatrix}$

while the rows of A are the vector transposes $\begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}^T$

and $\begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}^T$.

Each of the numbers in a matrix is called an <u>entry</u>

We refer to an entry by its row and column position.

eg For a 4×3 matrix, the entries of the first row are the $(1,1)$-entry, the $(1,2)$-entry, the $(1,3)$-entry and the $(1,4)$-entry. The entries of the second column are the $(1,2)$-entry, the $(2,2)$-entry and the $(3,2)$-entry.

Thus if the (i,j)-entry of a matrix A is a_{ij} then we can write $A = [a_{ij}]$.

eg $A = \begin{bmatrix} 2 & 5 & -1 \\ 0 & -3 & 6 \end{bmatrix}$ $\quad a_{11} = 2$

$a_{13} = -1$

$a_{21} = 0$

eg If $A = [a_{ij}]$ and $B = [b_{ij}]$ for which $B = A^T$ then $b_{ij} = a_{ji}$.

<u>Def'n</u>: Two matrices A and B are <u>equal</u> if they have the same number of rows, the same number of columns, and all corresponding entries to be equal.

eg If $A = \begin{bmatrix} 4 & 2a \\ a+b & -1 \end{bmatrix}$ and $B = \begin{bmatrix} x^2 & 6 \\ -3 & y+4 \end{bmatrix}$

find all values of a, b, x, y for which $A = B$.

Since A and B are both 2×2 matrices we set

$$4 = x^2 \longrightarrow \boxed{x = \pm 2}$$
$$2a = 6 \longrightarrow \boxed{a = 3}$$
$$a+b = -3 \longrightarrow 3+b = -3 \rightarrow \boxed{b = -6}$$
$$-1 = y+4 \longrightarrow \boxed{y = -5}$$

Scalar multiplication for a matrix states that if $A = [a_{ij}]$ then

$$kA = [ka_{ij}] \quad \text{for any constant } k.$$

eg If $A = \begin{bmatrix} 5 & -2 \\ 3 & 0 \end{bmatrix}$ then $(-1)A = \begin{bmatrix} -5 & 2 \\ -3 & 0 \end{bmatrix}$.

Here, $(-1)A$ is the negative of A and can be denoted by $-A$.

Likewise we can define the matrix addition of $A = \begin{bmatrix} \underline{a_1} & \underline{a_2} & \cdots & \underline{a_n} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$

and $B = \begin{bmatrix} \underline{b_1} & \underline{b_2} & \cdots & \underline{b_n} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$ to be the matrix

$$A+B = \begin{bmatrix} \underline{a_1+b_1} & \underline{a_2+b_2} & \cdots & \underline{a_n+b_n} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

eg If $A = \begin{bmatrix} 2 & 1 & 3 \\ -1 & 2 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & 6 \end{bmatrix}$

then $A + B = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 2 & 6 \end{bmatrix}$

Likewise we can define matrix subtraction in the obvious way.

The <u>zero matrix</u> is the matrix $\underline{0}$ all of whose entries are zero.

<u>Theorem</u>: <u>Properties of Scalar Multiplication and Matrix Addition</u>

Let A, B, C be matrices and k, ℓ be scalars then

1. closure under addition: $A+B$ is a matrix
2. commutativity of addition: $A+B = B+A$
3. associativity of addition: $(A+B)+C = A+(B+C)$
4. zero: there is a zero matrix such that $A + \underline{0} = \underline{0} + A = A$
5. negative: for every matrix A, there is a negative $-A$ such that $A + (-A) = (-A) + A = \underline{0}$
6. closure under scalar multiplication: kA is a matrix
7. associativity of scalar multiplication: $k(\ell A) = (k\ell)A$
8. unity: $1A = A$
9. distributivity: $k(A+B) = kA + kB$ and $(k+\ell)A = kA + \ell A$
10. if $kA = \underline{0}$ then $k = 0$ or $A = \underline{0}$

eg Let $A = \begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$.

Solve the equation $A + 2X = 6B$ for the matrix X.

We can write

$$2X = 6B - A$$

$$X = 3B - \tfrac{1}{2}A$$

$$= 3\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 0 \\ -3 & 6 \end{bmatrix} - \begin{bmatrix} 3/2 & 1 \\ -1/2 & 1/2 \end{bmatrix}$$

$$\boxed{= \begin{bmatrix} 3/2 & -1 \\ -5/2 & 11/2 \end{bmatrix}}$$

To understand matrix multiplication, first suppose we have a row matrix A and a column matrix B. Then $A = [\,\underline{a}^T \longrightarrow\,]$ and $B = \begin{bmatrix} \underline{b} \\ \downarrow \end{bmatrix}$. Then we define $AB = \underline{a}^T\underline{b} = [\,\underline{a}\cdot\underline{b}\,]$.

eg $$\begin{bmatrix} 2 & 0 & -3 & -1 \end{bmatrix}\begin{bmatrix} 5 \\ -4 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 10+0-3+1 \end{bmatrix} = \begin{bmatrix} 8 \end{bmatrix}$$

This is a 1×1 matrix.

Note that the number of columns of A must be the same as the number of rows of B.

Now suppose that we have a matrix $A = \begin{bmatrix} \underline{a}_1^T \rightarrow \\ \underline{a}_2^T \rightarrow \\ \vdots \\ \underline{a}_m^T \rightarrow \end{bmatrix}$

Now we can define $A\underline{b} = \begin{bmatrix} \underline{a}_1^T \underline{b} \\ \underline{a}_2^T \underline{b} \\ \vdots \\ \underline{a}_m^T \underline{b} \end{bmatrix} = \begin{bmatrix} \underline{a}_1 \cdot \underline{b} \\ \underline{a}_2 \cdot \underline{b} \\ \vdots \\ \underline{a}_m \cdot \underline{b} \end{bmatrix}$

eg $\begin{bmatrix} 2 & -5 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 - 5 + 3 \\ 0 + 0 - 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$

Thus the system of linear equations from the start of Section 2.1 can be written as a matrix equation $A\underline{x} = \underline{b}$ where

$$A = \begin{bmatrix} -3 & -3 & 1 \\ 5 & 0 & -2 \\ 1 & -4 & 3 \end{bmatrix} \quad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \underline{b} = \begin{bmatrix} -8 \\ 20 \\ -9 \end{bmatrix}$$

Consider the product of a row matrix with a standard basis vector.

eg $$\begin{bmatrix} 5 & 3 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 3 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix}$$

In general, the product of a row matrix A with $\underline{e}_i$ will be the 1×1 matrix whose entry is the ith entry of A.

Now we replace the row matrix with any matrix A.

eg $$\begin{bmatrix} -1 & 4 \\ 9 & 8 \end{bmatrix} \underline{e}_1 = \begin{bmatrix} -1 & 4 \\ 9 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 9 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 4 \\ 9 & 8 \end{bmatrix} \underline{e}_2 = \begin{bmatrix} -1 & 4 \\ 9 & 8 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

Thus the product of A with $\underline{e}_i$ is the ith column matrix of A.

To multiply two matrices A and B, we write

$$AB = A \begin{bmatrix} \underline{b}_1 & \underline{b}_2 & \cdots & \underline{b}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} A\underline{b}_1 & A\underline{b}_2 & \cdots & A\underline{b}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

eg $\begin{bmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 & 6 & 0 \\ 0 & 2 & 3 & 4 \\ -1 & 0 & 5 & 8 \end{bmatrix}$

$$= \begin{bmatrix} 4 & 1 & 25 & 12 \\ -4 & 2 & 23 & 36 \end{bmatrix}$$

In general, the (i,j) entry of AB is the dot product of ~~column~~ row i of A with column j of B.

Thus the number of columns of A must be the same as the number of rows of B.

If A is an $m \times n$ matrix and B is an $n \times p$ matrix then AB will be an $m \times p$ matrix.

eg $\begin{bmatrix} 2 & 1 & 6 & 0 \\ 0 & 2 & 3 & 4 \\ -1 & 0 & 5 & 8 \end{bmatrix} \begin{bmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{bmatrix}$ cannot be computed

We can define the power of a matrix A^n but only for a natural number n and only for a square matrix A which has an equal number of rows and columns.

eg $\begin{bmatrix} 3 & -3 \\ 0 & -4 \end{bmatrix}^2 = \begin{bmatrix} 3 & -3 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} 3 & -3 \\ 0 & -4 \end{bmatrix} = \begin{bmatrix} 9 & 3 \\ 0 & 16 \end{bmatrix}$

Even if AB and BA can both be computed, in general $AB \neq BA$. Thus matrix multiplication is <u>not</u> commutative.

eg If $A = \begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$ then

$$AB = \begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -3 & 12 \\ 2 & -8 \end{bmatrix}$$

$$BA = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ -6 & -9 \end{bmatrix}$$

<u>Theorem</u>: <u>Properties of Matrix Multiplication</u>

Let A, B, C be matrices of appropriate size, and k be a scalar. Then

1. associativity: $(AB)C = A(BC)$
2. distributivity over addition: $(A+B)C = AC+BC$
   $A(B+C) = AB+AC$
3. scalar commutativity: $A(kB) = (kA)B = k(AB)$

But without commutativity, some "obvious" results do not hold.

eg $(A+B)^2 \neq A^2 + 2AB + B^2$

Instead, $(A+B)^2 = (A+B)(A+B)$
$= A^2 + AB + BA + B^2$

but we cannot assume that $AB = BA$.

A matrix A does commute with the zero matrix $\underline{0}$,

because $A\underline{0} = \underline{0}A = \underline{0}$

It ~~is~~ also commutes with the identity matrix I, which is the square matrix formed by setting all the diagonal entries — that is, the (i,i) entries — equal to 1, and all the off-diagonal entries equal to 0.

eg The 3-dimensional identity matrix is

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

We also denote the $n \times n$ identity matrix as I_n.

In general, $$I_n = \begin{bmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 & \cdots & \underline{e}_n \\ \downarrow & \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

Thus $AI = IA = A$.

eg $$\begin{bmatrix} 1 & -7 & -2 \\ 4 & 0 & -4 \\ 0 & -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -7 & -2 \\ 4 & 0 & -4 \\ 0 & -2 & 3 \end{bmatrix}$$

A <u>linear combination</u> of matrices $A_1, A_2, A_3, \dots, A_n$ is a matrix of the form

$$k_1 A_1 + k_2 A_2 + k_3 A_3 + \dots + k_n A_n$$

where $k_1, k_2, k_3, \dots, k_n$ are scalars.

Now consider a vector $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$. Then

$$\begin{aligned} A\underline{x} = A\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} &= A(x_1\underline{e}_1 + x_2\underline{e}_2 + \dots + x_n\underline{e}_n) \\ &= Ax_1\underline{e}_1 + Ax_2\underline{e}_2 + \dots + Ax_n\underline{e}_n \\ &= x_1(A\underline{e}_1) + x_2(A\underline{e}_2) + \dots + x_n(A\underline{e}_n) \end{aligned}$$

But $A\underline{e}_i$ is the ith column of A so

$$A\underline{x} = x_1\begin{bmatrix} a_1 \\ \downarrow \end{bmatrix} + x_2\begin{bmatrix} a_2 \\ \downarrow \end{bmatrix} + \dots + x_n\begin{bmatrix} a_n \\ \downarrow \end{bmatrix}$$

which means that $A\underline{x}$ is the linear combination of the columns of A where the coefficients are the components of $\underline{x}$ in the appropriate order.

eg $$\begin{bmatrix} 1 & -3 & 0 \\ -2 & -4 & 8 \end{bmatrix}\begin{bmatrix} 6 \\ -2 \\ 3 \end{bmatrix} = 6\begin{bmatrix} 1 \\ -2 \end{bmatrix} - 2\begin{bmatrix} -3 \\ -4 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 8 \end{bmatrix}$$

Now suppose we have a diagonal matrix $D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{bmatrix}$

Then $$AD = A \begin{bmatrix} d_1\underline{e}_1 & d_2\underline{e}_2 & \cdots & d_n\underline{e}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} Ad_1\underline{e}_1 & Ad_2\underline{e}_2 & \cdots & Ad_n\underline{e}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} d_1(A\underline{e}_1) & d_2(A\underline{e}_2) & \cdots & d_n(A\underline{e}_n) \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} d_1\underline{a}_1 & d_2\underline{a}_2 & \cdots & d_n\underline{a}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

eg $$\begin{bmatrix} 5 & -2 & -1 \\ 3 & 0 & 3 \\ 9 & -7 & 8 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} 10 & 2 & 3 \\ 6 & 0 & -9 \\ 18 & 7 & -24 \end{bmatrix}$$