Matrices

If m and n are positive integers, an m by n (written $m \times n$) matrix with real number entries is a rectangular array of real numbers having m rows and n columns. The set of all $m \times n$ matrices with real entries will be denoted by $M_{m,n}(\mathbb{R})$. A generic element of $M_{m,n}(\mathbb{R})$ is written as:

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

$a_{ij} \in \mathbb{R}$ for all $i = 1, \cdots, m$ and $j = 1, \cdots, n$. Here a_{ij} denotes the i,j-th entry of A. a_{ij} is the entry of A which appears in both the i-th row and j-th column of A.

<u>Examples</u>:

1) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ is a 2×3 matrix

its 1,1-entry is 1
1,2-entry is 2
1,3-entry is 3
2,1-entry is 4
2,2-entry is 5
2,3-entry is 6

2) $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is a 3×1 matrix

its 1,1-entry is 1
its 2,1-entry is 2
its 3,1-entry is 3

For convenience, we will frequently use $A = [a_{ij}]_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ to denote an $m \times n$ matrix. When the size of A is understood, we'll just write $A = [a_{ij}]$.

Note that a $1 \times n$ matrix looks very much like a vector in $\mathbb{R}^n$. The only difference being the commas between the components in the vector. It is useful, on occasion, to overlook this minor difference and view a $1 \times n$ matrix as a vector in $\mathbb{R}^n$ or view a vector in $\mathbb{R}^n$ as a $1 \times n$ matrix.

Although it is less obvious, an $m \times n$ matrix can be viewed as a vector in $\mathbb{R}^{mn}$. If $A = [a_{ij}] \in M_{m,n}(\mathbb{R})$, we can associate A to the vector

$$[\underbrace{a_{11}, \cdots, a_{1n}}_{\substack{\text{1-st row} \\ \text{of } A}}, \underbrace{a_{21}, \cdots, a_{2n}}_{\substack{\text{2-nd row} \\ \text{of } A}}, \cdots, \underbrace{a_{m1}, \cdots, a_{mn}}_{\substack{m\text{-th row} \\ \text{of } A}}].$$

In words, A is identified with a vector by juxtaposing rows. When $m > 1$, this identification of $M_{m,n}(\mathbb{R})$

with $\mathbb{R}^{mn}$ is far less frequently encountered than the case when $m=1$ - but it does suggest the following definitions.

<u>Definitions</u>: Suppose $r \in \mathbb{R}$ and $A, B \in M_{m,n}(\mathbb{R})$ where $A = [a_{ij}]$, $B = [b_{ij}]$.

1) We say $A = B$ precisely when $a_{ij} = b_{ij}$ for all $i = 1, \cdots, m$ and $j = 1, \cdots, n$.

2) The sum of A and B, denoted by $A+B$, is the $m \times n$ matrix $S = [s_{ij}]$, where $s_{ij} = a_{ij} + b_{ij}$ for all $i = 1, \cdots, m$ and $j = 1, \cdots, n$.

3) Scalar multiplication, denoted by rA, is defined to be the $m \times n$ matrix $P = [p_{ij}]$ where $p_{ij} = ra_{ij}$ for all $i = 1, \cdots, m$ and $j = 1, \cdots, n$.

Note: Matrix equality and matrix addition require two matrices of the same size.

Examples:

1) $$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 1+7 & 2+8 & 3+9 \\ 4+10 & 5+11 & 6+12 \end{bmatrix}$$

$$= \begin{bmatrix} 8 & 10 & 12 \\ 14 & 16 & 18 \end{bmatrix}$$

2) $2\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} (2)(1) & (2)(2) & (2)(3) \\ (2)(4) & (2)(5) & (2)(6) \end{bmatrix}$

$$= \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{bmatrix}$$

The following Theorem can be viewed as a consequence of Theorem 1 – using our identification of $M_{m,n}(\mathbb{R})$ with $\mathbb{R}^{mn}$.

<u>Theorem 2</u>: Matrix addition and scalar multiplication satisfy the following ten properties:

1) If $A, B \in M_{m,n}(\mathbb{R})$, then $A+B \in M_{m,n}(\mathbb{R})$.

2) If $A, B \in M_{m,n}(\mathbb{R})$, then $A+B = B+A$

3) If $A, B, C \in M_{m,n}(\mathbb{R})$, then

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

4) If $O_{m,n}$ denotes the $m \times n$ matrix, all of whose entries equal 0, and $A \in M_{m,n}(\mathbb{R})$, then $A + O_{m,n} = A$.

5) If $A \in M_{m,n}(\mathbb{R})$, there is a $-A \in M_{m,n}(\mathbb{R})$ such that $A + (-A) = O_{m,n}$.

6) If $r \in \mathbb{R}$ and $A \in M_{m,n}(\mathbb{R})$, then $rA \in M_{m,n}(\mathbb{R})$.

7) If $r \in \mathbb{R}$ and $A, B \in M_{m,n}(\mathbb{R})$, then

$r(A+B) = rA + rB$.

8) If $r, s \in \mathbb{R}$ and $A \in M_{m,n}(\mathbb{R})$, then

$(r+s)A = rA + sA$.

9) If $r, s \in \mathbb{R}$ and $A \in M_{m,n}(\mathbb{R})$, then

$r(sA) = (rs)A$.

10) If $A \in M_{m,n}(\mathbb{R})$, then $1A = A$.

The Transpose of a Matrix

Definition: Suppose $A \in M_{m,n}(\mathbb{R})$ and $A = [a_{ij}]$. The transpose of A, denoted by A^t, is the $n \times m$ matrix $A^t = [\alpha_{ij}]$ such that $\alpha_{ij} = a_{ji}$ for all $i = 1, \cdots, n$ and $j = 1, \cdots, m$.

Probably the best way to see how A^t relates to A is to view the i-th row of A^t as a $1 \times m$ matrix.

i-th row of $A^t = [\alpha_{i1}\ \alpha_{i2} \cdots \alpha_{im}]$

Since $\alpha_{ij} = a_{ji}$ for all i, j, we have

i-th row of $A^t = [\alpha_{i1}\ \alpha_{i2} \cdots \alpha_{im}]$

$$= [a_{1i}\ a_{2i} \cdots a_{mi}].$$

Now note that the i-th column of A is

$$i\text{-th column of } A = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}.$$

Comparing the i-th row of A^t with the i-th column of A, we see that the i-th row of A^t is just the i-th column of A — which is "twisted" so as to become a row.

Actually, the term "twisted" used in

the last sentence can be replaced by a more appropriate term - namely "transpose". To see why, observe that the i-th column of A is a $m \times 1$ matrix. As such, more appropriate notation would be

$$\mathbb{X} = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{m1} \end{bmatrix}$$

- where $x_{k1} = a_{ki}$ for $k = 1, \cdots, m$.

Now $\mathbb{X}^t$ would be a $1 \times m$ matrix, so

$$\mathbb{X}^t = [y_{11}\ y_{12} \cdots y_{1m}]$$

where $y_{1j} = x_{j1}$ for $j = 1, \cdots, m$.

Thus,

$$X^t = [x_{11}\ x_{21}\ \cdots\ x_{m1}].$$

Since $x_{k1} = a_{ki}$ for $k = 1, \cdots, m$, we obtain

$$X^t = [a_{1i}\ a_{2i}\ \cdots\ a_{mi}].$$

Finally, since

$$X = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix},$$

we obtain

$$\begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}^t = [a_{1i}\ a_{2i}\ \cdots\ a_{mi}].$$

It follows that the i-th row of A^t

is the transpose of the i-th column of A. An analogous discussion leads to showing that the j-th column of A^t is the transpose of the j-th row of A. Thus, A^t is obtained from A by changing rows into columns and colums into rows.

Examples:

1)
$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

2) $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}^t = \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$

Theorem 5: Suppose $r \in \mathbb{R}$ and $A, B \in M_{m,n}(\mathbb{R})$. Then

1) $(A^t)^t = A$,

2) $(A+B)^t = A^t + B^t$, and

3) $(rA)^t = rA^t$.

Proof of 1): First note that A^t is a $n \times m$ matrix, so $(A^t)^t$ is a $m \times n$ matrix. Thus, both A and $(A^t)^t$ are $m \times n$ matrices.

If $A^t = [b_{ij}]_{\substack{1 \le i \le n \\ 1 \le j \le m}}$, then

$b_{ij} = a_{ji}$ for $i = 1, \cdots, n$ and $j = 1, \cdots, m$.

If $(A^t)^t = [c_{ij}]_{\substack{1 \le i \le m \\ 1 \le j \le n}}$, then

$c_{ij} = b_{ji}$ for $i = 1, \cdots, m$, $j = 1, \cdots, n$.

Now note that

$$\begin{aligned}
i,j\text{-th entry of } (A^t)^t &= c_{ij} \\
&= b_{ji} \\
&= a_{ij} \\
&= i,j\text{-th entry of } A.
\end{aligned}$$

Since the i,j-th entry of $(A^t)^t$ equals the i,j-th entry of A, we conclude that $(A^t)^t = A$.

HW #7: Complete the proof of Theorem 5.

As we have seen, in general a matrix and its transpose can be a different sizes. For the obvious reason, an $n \times n$ matrix is called a square matrix. If A is a square matrix, then both A and A^t are $n \times n$ matrices, for some n. As such, it is possible for a square matrix to satisfy an equation, such as

$$A = A^t$$

or $$A = -A^t.$$

We say that the square matrix A is symmetric if $A = A^t$ and we say A is skew-symmetric if $A = -A^t$.

HW #8: Suppose A is a square matrix.

1) Show that $\frac{1}{2}(A + A^t)$ is a symmetric matrix.

2) Show that $\frac{1}{2}(A - A^t)$ is a skew-symmetric matrix.

3) Show that A can be written as the sum of a symmetric matrix and a skew-symmetric matrix.

Matrix Multiplication

The product AB of the matrices A and B is defined precisely when

$$\begin{array}{c} \text{the number of} \\ \text{columns of } A \end{array} = \begin{array}{c} \text{the number of} \\ \text{rows of } B \end{array}.$$

Thus, AB is defined whenever A is a $m \times n$ matrix and B is a $n \times k$ matrix - where m, n and k are positive integers. Note that in this case, AB is defined - but BA will only be defined when $m = k$.

Suppose now that $A = [a_{ij}]$ is a $m \times n$ matrix and $B = [b_{ij}]$ is a $n \times k$ matrix. Then

$$AB = P = [p_{ij}]$$

where P is a $m \times k$ matrix and

$$p_{ij} = \sum_{s=1}^{n} a_{is} b_{sj} \quad \text{for all } i = 1, \cdots, m$$

and $j = 1, \cdots, k$.

For computational purposes, it is useful to note that each p_{ij} is a dot product. Indeed,

$$p_{ij} = \sum_{s=1}^{n} a_{is} b_{sj}$$

$$= [a_{i1}, a_{i2}, \cdots, a_{in}] \cdot [b_{1j}, b_{2j}, \cdots, b_{nj}]$$

$$= [a_{i1} \; a_{i2} \; \cdots \; a_{in}] \cdot \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix}^{t}$$

- where we are identifying row matrices with vectors. Thus, p_{ij} is just the dot product of the i-th row of A with the j-th column of B.

Examples:

1) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

$$= \begin{bmatrix} [1,2,3] \cdot [1,3,5] & [1,2,3] \cdot [2,4,6] \\ [4,5,6] \cdot [1,3,5] & [4,5,6] \cdot [2,4,6] \end{bmatrix}$$

$$= \begin{bmatrix} 1+6+15 & 2+8+18 \\ 4+15+30 & 8+20+36 \end{bmatrix}$$

$$= \begin{bmatrix} 22 & 28 \\ 49 & 64 \end{bmatrix}$$

2) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

$$= \begin{bmatrix} [1,2]\cdot[1,4] & [1,2]\cdot[2,5] & [1,2]\cdot[3,6] \\ [3,4]\cdot[1,4] & [3,4]\cdot[2,5] & [3,4]\cdot[3,6] \\ [5,6]\cdot[1,4] & [5,6]\cdot[2,5] & [5,6]\cdot[3,6] \end{bmatrix}$$

$$= \begin{bmatrix} 1+8 & 2+10 & 3+12 \\ 3+16 & 6+20 & 9+24 \\ 5+24 & 10+30 & 15+36 \end{bmatrix}$$

$$= \begin{bmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{bmatrix}$$

<u>HW #9</u>: Compute the following products of matrices - if they exist.

1) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$

2) $\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

3) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

4) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

5) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

6) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

7) $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$

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

9) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

Theorem 6: Matrix multiplication has the following properties:

1) If $r \in \mathbb{R}$, $A \in M_{m,n}(\mathbb{R})$ and $B \in M_{n,k}(\mathbb{R})$, then

$$r(AB) = (rA)B = A(rB).$$

2) If $A \in M_{m,n}(\mathbb{R})$ and $B, C \in M_{n,k}(\mathbb{R})$, then

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

3) If $A, B \in M_{m,n}(\mathbb{R})$ and $C \in M_{n,k}(\mathbb{R})$, then

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

4) If $A \in M_{m,n}(\mathbb{R})$ and $B \in M_{n,k}(\mathbb{R})$, then

$$(AB)^t = B^t A^t.$$

5) If $A \in M_{m,n}(\mathbb{R})$, $B \in M_{n,k}(\mathbb{R})$ and $C \in M_{k,\ell}(\mathbb{R})$, then $(AB)C = A(BC)$.

Proof of 4): Initially note that AB is a $m \times k$ matrix - so $(AB)^t$ is $k \times m$. On the other hand, B^t is $k \times n$ and A^t is $n \times m$ - so $B^t A^t$ is a $k \times m$ matrix. Thus, $(AB)^t$ and $B^t A^t$ are both $k \times m$ matrices.

Now let $A = [a_{ij}]$, $B = [b_{ij}]$ and set $AB = C = [c_{ij}]$. Then $c_{ij} = \sum_{s=1}^{n} a_{is} b_{sj}$. Letting $C^t = [\gamma_{ij}]$, it follows that

$$\gamma_{ij} = c_{ji} = \sum_{s=1}^{n} a_{js} b_{si}.$$

Writing $B^t = [\beta_{ij}]$ and $A^t = [\alpha_{ij}]$, it follows that $\beta_{ij} = b_{ji}$ and $\alpha_{ij} = a_{ji}$.

We now obtain:

$$\begin{aligned}
\text{the } i,j\text{-th entry of } B^tA^t &= \sum_{s=1}^{n} \beta_{is}\alpha_{sj} \\
&= \sum_{s=1}^{n} b_{si} a_{js} \\
&= \sum_{s=1}^{n} a_{js} b_{si} \\
&= \gamma_{ij} \\
&= i,j\text{-th entry of } C^t \\
&= i,j\text{-th entry of } (AB)^t.
\end{aligned}$$

Thus, $(AB)^t = B^t A^t$.

Proof of 5) : First note that AB is a $m \times k$ matrix - so $(AB)C$ is $m \times \ell$. Similarly, BC is $n \times \ell$ - so $A(BC)$ is also $m \times \ell$.

Now let

$A = [a_{ij}]$ $i = 1, \cdots, m$, $j = 1, \cdots, n$,

$B = [b_{ij}]$ $i = 1, \cdots, n$, $j = 1, \cdots, k$, and

$C = [c_{ij}]$ $i = 1, \cdots, k$, $j = 1, \cdots, \ell$.

Also, set

$AB = D = [d_{ij}]$ $i = 1, \cdots, m$, $j = 1, \cdots, k$

where $d_{ij} = \sum_{s=1}^{n} a_{is} b_{sj}$, and

$BC = E = [e_{ij}]$ $i = 1, \cdots, n$, $j = 1, \cdots, \ell$

where $e_{ij} = \sum_{t=1}^{k} b_{it} c_{tj}$.

Then

$$\text{the } i,j\text{-th entry of } (AB)C = \sum_{t=1}^{k} d_{it} c_{tj}$$
$$= \sum_{t=1}^{k} \left(\sum_{s=1}^{n} a_{is} b_{st} \right) c_{tj}$$
$$= \sum_{t=1}^{k} \left(\sum_{s=1}^{n} a_{is} b_{st} c_{tj} \right).$$

Also,

$$\text{the } i,j\text{-th entry of } A(BC) = \sum_{s=1}^{n} a_{is} e_{sj}$$
$$= \sum_{s=1}^{n} a_{is} \left(\sum_{t=1}^{k} b_{st} c_{tj} \right)$$
$$= \sum_{s=1}^{n} \left(\sum_{t=1}^{k} a_{is} b_{st} c_{tj} \right).$$

Finally, observe that the two sums,

$\sum_{t=1}^{k} \left(\sum_{s=1}^{n} a_{is} b_{st} c_{tj} \right)$ and $\sum_{s=1}^{n} \left(\sum_{t=1}^{k} a_{is} b_{st} c_{tj} \right)$,

are sums of the same nk real numbes – namely, $a_{is} b_{st} c_{tj}$ where $s = 1, \cdots, n$ and $t = 1, \cdots, k$. These sums differ only in the order in which these nk real numbers are added together. Since the sum of any finite set of real number remains the same - no matter the order in which they are added together - we conclude that

the i,j-th entry of $(AB)C$

$=$ the i,j-th entry of $A(BC)$.

This completes the proof of 5).

HW #10: Prove properties 1) and 2) of Theorem 6.

As you saw in HW #9, matrices such as $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ have a special property

when involved in a product of matrices, which is defined. These are two examples of identity matrices.

In general, the $n \times n$ identity matrix, denoted by $I_n \in M_{n,n}(\mathbb{R})$, is defined to be

$$I_n = [\delta_{ij}]$$

where δ_{ij} denotes the Kronecker delta function, i.e.

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}.$$

Thus,

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

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

<u>Proposition</u> 3: Let $A = [a_{ij}] \in M_{m,n}(\mathbb{R})$.

Then

1) $I_m A = A$, and

2) $A I_n = A$.

<u>Proof of 1)</u>: Note that both $I_m A$ and A are $m \times n$ matrices.

Now note that

the i,j-th entry of $I_m A = \sum_{s=1}^{m} \delta_{is} a_{sj}$.

Since $\delta_{is} = \begin{cases} 0 & \text{if } s \neq i \\ 1 & \text{if } s = i \end{cases}$,

$\sum_{s=1}^{m} \delta_{is} a_{sj} = \delta_{ii} a_{ij} = a_{ij}$. It follows

that

the i,j-th entry of $I_m A$

= the i,j-th entry of A.

<u>HW #11</u> : Complete the proof of

Proposition 3.