3. Introduction to Matrix Algebra

Definition 1. Let m, n be positive integers. We define

$$\mathcal{M}_{m\times n} = \{ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} : a_{ij} \in \mathbb{R}\}$$

to be the set of matrices with m rows and n columns with entries in F. We will often use the shorthand notation (a_{ij}) to denote a typical matrix in $\mathcal{M}_{m\times n}$. If $A = (a_{ij}) \in \mathcal{M}_{m\times n}$, then we will sometimes use the notation $\text{ent}_{ij}(A)$ instead of a_{ij}. We declare that if $A, B \in \mathcal{M}_{m\times n}$, then $A = B$ if and only if $a_{ij} = b_{ij}$ for all $i \in \{1, 2, ..., m\}$ and $j \in \{1, 2, ..., n\}$. If $m = n$, then we will simply write $\mathcal{M}_n$ to denote the set of $n \times n$ matricies over the field $\mathbb{R}$.

Notations 2. Let $A = (a_{ij}) \in \mathcal{M}_{m\times n}$. Then we write the columns of A as

$$\mathbf{c}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \mathbf{c}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, ..., \mathbf{c}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \in \mathbb{R}^m.$$

We can now denote the matrix as

$$A = [a_{ij}] = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n].$$

Similarly, we write the rows of A as

$$\begin{aligned} \mathbf{r}_1 &= (a_{11}, a_{12}, ..., a_{1n}) \\ \mathbf{r}_2 &= (a_{21}, a_{22}, ..., a_{2n}) \\ &\vdots \\ \mathbf{r}_m &= (a_{m1}, a_{m2}, ..., a_{mn}) \in \mathbb{R}^n. \end{aligned}$$

We can now denote the matrix as

$$A = (a_{ij}) = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{bmatrix}.$$

Definition 3. We define an addition $\boxplus$ on $\mathcal{M}_{m\times n}$ given by $(a_{ij}) \boxplus (b_{ij}) = (a_{ij} + b_{ij})$. In other words, if two matrices have the same shape, then addition is carried out by adding the corresponding entries. Addition is not defined for matrices of different shape. We define a scalar multiplication $\mathbb{R} \times \mathcal{M}_{m\times n} \to \mathcal{M}_{m\times n}$ given by $\alpha(a_{ij}) = (\alpha a_{ij})$.

Theorem 4. The following properties hold for the set $\mathcal{M}_{m\times n}$ of $m\times n$ matricies with entries in $\mathbb{R}$.
(A1) $A \boxplus (B \boxplus C) = (A \boxplus B) \boxplus C$ for all $A, B, C \in \mathcal{M}_{m\times n}$
(A2) $A \boxplus B = B \boxplus A$ for all $A, B \in \mathcal{M}_{m\times n}$
(A3) The matrix $\mathbf{0}$ is the unique matrix in $\mathcal{M}_{m\times n}$ such that $A \boxplus \mathbf{0} = A$ for all $A \in \mathcal{M}_{m\times n}$
(A4) For each $A \in \mathcal{M}_{m\times n}$, there exists a unique $B \in \mathcal{M}_{m\times n}$ such that $A \boxplus B = \mathbf{0}$
(S1) $1A = A$ for all $A \in \mathcal{M}_{m\times n}$
(S2) $\alpha(A \boxplus B) = \alpha A \boxplus \alpha B$ for all $\alpha \in \mathbb{R}$ and for all $A, B \in \mathcal{M}_{m\times n}$
(S3) $(\alpha + \beta)A = \alpha A \boxplus \beta A$ for all $\alpha, \beta \in \mathbb{R}$ and for all $A \in \mathcal{M}_{m\times n}$
(S4) $\alpha(\beta A) = (\alpha\beta)A$ for all $\alpha, \beta \in \mathbb{R}$ and for all $A \in \mathcal{M}_{m\times n}$

Remark 5.
(1) From now on, we will simply write $A + B$ instead of $A \boxplus B$ for vector addition in $\mathcal{M}_{m\times n}$.
(2) We will denote the unique additive inverse of a vector $A \in \mathcal{M}_{m\times n}$ by $-A$. We write $A - B$ instead of $A + (-B)$.

Theorem 6. The following properties hold in the sset $\mathcal{M}_{m\times n}$ of $m\times n$ matricies with entries in $\mathbb{R}$.
(1) $\alpha\mathbf{0} = \mathbf{0}$ for all $\alpha \in \mathbb{R}$.
(2) $0A = \mathbf{0}$ for all $A \in M_{m\times n}$.
(3) $(-\alpha)A = \alpha(-A) = -\alpha A$ for all $\alpha \in \mathbb{R}$ and for all $A \in M_{m\times n}$.
(4) $\alpha A = \mathbf{0}$ implies $\alpha = 0$ or $A = \mathbf{0}$.

Definition 7. Let m, p, n be positive integers. We define a multiplication $\cdot : \mathcal{M}_{m\times p} \times \mathcal{M}_{p\times n} \to \mathcal{M}_{m\times n}$ given by

$$[a_{ij}]\,[b_{ij}] = [c_{ij}] \quad \text{where} \quad c_{ij} = \Sigma_{k=1}^{p} a_{ik} b_{kj} \quad \text{and} \quad \begin{array}{c} 1 \le i \le m \\ 1 \le j \le n \end{array}.$$

In other words,

$$\mathrm{ent}_{ij}(AB) = \sum_{k=1}^{p} \mathrm{ent}_{ik}(A)\,\mathrm{ent}_{kj}(B)$$

Observation 8. There are several ways to view matrix multiplication.

(1) Let

$$A = \begin{bmatrix} \mathbf{r}_1(A) \\ \mathbf{r}_2(A) \\ \vdots \\ \mathbf{r}_m(A) \end{bmatrix} \in \mathcal{M}_{m \times p}$$

and

$$B = [\mathbf{c}_1(B)\ \mathbf{c}_2(B)\ \cdots\ \mathbf{c}_n(B)] \in \mathcal{M}_{p \times n}.$$

Then $\mathbf{r}_i(A), \mathbf{c}_j(B) \in \mathbb{R}^p$ and

$$AB = \begin{bmatrix} \mathbf{r}_1(A) \cdot \mathbf{c}_1(B) & \mathbf{r}_1(A) \cdot \mathbf{c}_2(B) & \cdots & \mathbf{r}_1(A) \cdot \mathbf{c}_n(B) \\ \mathbf{r}_2(A) \cdot \mathbf{c}_1(B) & \mathbf{r}_2(A) \cdot \mathbf{c}_2(B) & \cdots & \mathbf{r}_2(A) \cdot \mathbf{c}_n(B) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_m(A) \cdot \mathbf{c}_1(B) & \mathbf{r}_m(A) \cdot \mathbf{c}_2(B) & \cdots & \mathbf{r}_m(A) \cdot \mathbf{c}_n(B) \end{bmatrix}.$$

Therefore, and we have a description of each entry of the product:

$$\mathrm{ent}_{ij}(AB) = \mathbf{r}_i(A) \cdot \mathbf{c}_j(B)$$

(2) Using the notation above, have

$$AB = [A\mathbf{c}_1(B)\ A\mathbf{c}_2(B)\ \cdots\ A\mathbf{c}_n(B)]\,.$$

We have a description of the columns of the product. That is,

$$\mathbf{c}_j(AB) = A\mathbf{c}_j(B).$$

We can show that the columns of AB are linear combinations of the columns of A. In other words, $\mathbf{c}_j(AB) \in \mathrm{Span}(\mathbf{c}_1(A), \mathbf{c}_2(A), ..., \mathbf{c}_p(A))$. We will see this below.

(3) Finally, we have a description of the rows of the product. That is,

$$\mathbf{r}_i(AB) = a_{i1}\mathbf{r}_1(B) + a_{i2}\mathbf{r}_2(B) + ... + a_{ip}\mathbf{r}_p(B)$$

In other words, $\mathbf{r}_i(AB) \in \mathrm{Span}(\mathbf{r}_1(B), \mathbf{r}_2(B), ..., \mathbf{r}_p(B))$.

Theorem 9. Matrix multiplication is is an associative operation in the sense that $A \in \mathcal{M}_{m\times p}$, $B \in \mathcal{M}_{p\times q}$, $C \in \mathcal{M}_{q\times n}$ implies
(1) $AB \in \mathcal{M}_{m\times q}$ so that $(AB)C \in \mathcal{M}_{m\times n}$,
(2) $BC \in \mathcal{M}_{p\times n}$ so that $A(BC) \in \mathcal{M}_{m\times n}$, and
(3) $(AB)C = A(BC)$.

Proof. We verify part (3) with the computation

$$
\begin{aligned}
& \operatorname{ent}_{ij}[(AB)C)] \\
= \ & \sum_{k=1}^{q} \operatorname{ent}_{ik}(AB) \operatorname{ent}_{kj}(C) \quad (\text{def of the mult. } \mathcal{M}_{m\times q} \times \mathcal{M}_{q\times n} \to \mathcal{M}_{m\times n}) \\
= \ & \sum_{k=1}^{q} \left(\sum_{l=1}^{p} \operatorname{ent}_{il}(A) \operatorname{ent}_{lk}(B) \right) \operatorname{ent}_{kj}(C) \quad (\text{def of the mult. } \mathcal{M}_{m\times p} \times \mathcal{M}_{p\times q} \to \mathcal{M}_{m\times q}) \\
= \ & \sum_{k=1}^{q} \left(\sum_{l=1}^{p} \operatorname{ent}_{il}(A) \operatorname{ent}_{lk}(B) \operatorname{ent}_{kj}(C) \right) \quad (\text{Distribution}) \\
= \ & \sum_{l=1}^{p} \left(\sum_{k=1}^{q} \operatorname{ent}_{il}(A) \operatorname{ent}_{lk}(B) \operatorname{ent}_{kj}(C) \right) \quad (\text{Interchance Finite Sums}) \\
= \ & \sum_{l=1}^{p} \left(\operatorname{ent}_{il}(A) \sum_{k=1}^{q} \operatorname{ent}_{lk}(B) \operatorname{ent}_{kj}(C) \right) \quad (\text{Since } \operatorname{ent}_{il}(A) \text{ constant in } \sum_{k=1}^{q}) \\
= \ & \sum_{l=1}^{p} \operatorname{ent}_{il}(A) \operatorname{ent}_{lj}(BC) \quad (\text{def of the mult. } \mathcal{M}_{p\times q} \times \mathcal{M}_{q\times n} \to \mathcal{M}_{p\times n}) \\
= \ & \operatorname{ent}_{ij}(A(BC)) \quad (\text{def of the mult. } \mathcal{M}_{m\times p} \times \mathcal{M}_{p\times n} \to \mathcal{M}_{m\times n})
\end{aligned}
$$

Definition 10. We define $I_n = (\delta_{ij}) \in \mathcal{M}_n$ where

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

For example,

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Theorem 11. Let $A, A' \in \mathcal{M}_{m\times p}$, $B, B' \in \mathcal{M}_{p\times n}$. Then
(1) $I_m A = A = A I_n$
(2) $(A + A')B = AB + A'B$
(3) $A(B + B') = AB + AB'$
(4) $(cA)B = cAB = A(cB)$ for all $c \in \mathbb{R}$.

Proof. We prove (2) and leave the remaining parts as an exercise.

$$
\begin{aligned}
& \operatorname{ent}_{ij}((A+A')B) \\
= \;& \sum_{k=1}^{p} \operatorname{ent}_{ik}(A+A') \operatorname{ent}_{kj}(B) \\
= \;& \sum_{k=1}^{p} (\operatorname{ent}_{ik}(A) + \operatorname{ent}_{ik}(A')) \operatorname{ent}_{kj}(B) \\
= \;& \sum_{k=1}^{p} (\operatorname{ent}_{ik}(A) \operatorname{ent}_{kj}(B) + \operatorname{ent}_{ik}(A') \operatorname{ent}_{kj}(B)) \\
= \;& \sum_{k=1}^{p} \operatorname{ent}_{ik}(A) \operatorname{ent}_{kj}(B) + \sum_{k=1}^{p} \operatorname{ent}_{ik}(A') \operatorname{ent}_{kj}(B) \\
= \;& \operatorname{ent}_{ij}(AB) + \operatorname{ent}_{ij}(A'B) \\
= \;& \operatorname{ent}_{ij}(AB + A'B).
\end{aligned}
$$

Systems of Linear Equations

Definition 12. A system of linear equations is a collection of m hyperplanes in $\mathbb{R}^n$

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\
&\;\vdots \\
a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n &= b_m.
\end{aligned} \tag{$\mathcal{S}$}
$$

In matrix notation, we have

$$
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.
$$

In a more compact notation, we write

$$
A\mathbf{x} = \mathbf{b}.
$$

It is easy to check that

$$
x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + ... + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}
$$

and so

$$\mathbf{b} \in \mathrm{Span}(\mathbf{c}_1(A), A\mathbf{c}_2(A), ..., \mathbf{c}_n(A)).$$

Incidently, if we set $\mathbf{x} = \mathbf{c}_j(B)$ we verify the statement in Observation 8(2) that $\mathbf{c}_j(AB)$ is a linear combination of the vectors $\mathbf{c}_1(A), \mathbf{c}_2(A), ..., \mathbf{c}_p(A)$.

Question 13. Given a linear system $A\mathbf{x} = \mathbf{b}$:
(1) Does there exists a solution vector $\mathbf{x}$?
(2) If a solution exists is it unique?

Definition 14. If the linear system $A\mathbf{x} = \mathbf{b}$ has a solution, then it is called a consistent linear system.

Exercises Section 2.1.

1. Prove Theorem 11(3 and 4).

2. Exercise 2.1.5 of the text.

3. Exercise 2.1.6 of the text.

4. Exercise 2.1.7 of the text.

5. Exercise 2.1.8 of the text.

6. Exercise 2.1.11 of the text.

7. Exercise 2.1.12(a,b,c,e) of the text.

8. Exercise 2.1.14 of the text.