Week 1

1.1 The Real Vector Space $\mathbb{R}^n$

$$n \in \mathbb{N}, \quad \mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \middle| x_i \in \mathbb{R} \right\}$$

We call an element of $\mathbb{R}^n$ a **vector**, typically denoted by a symbol of the form $\vec{v}$.

The vector whose entries are all zero is called the **zero vector**. We denote it by $\vec{0}$.

1.1.1 Important Properties of $\mathbb{R}^n$

- **Addition Law**

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

- **Scalar Multiplication** For $\lambda \in \mathbb{R}$,

$$\lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix}$$

The vectors:

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

form a **basis** of $\mathbb{R}^n$, in the sense that every vector of $\mathbb{R}^n$ may be written uniquely as a **linear combination** of them:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + v_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

1.1.2 Linear Transformations

A map $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$ is called a **linear transformation** (over $\mathbb{R}$) if:

- For all $\vec{v} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$, we have

$$\mathcal{L}(\lambda\vec{v}) = \lambda\mathcal{L}(\vec{v}),$$

  and

- For all $\vec{v}, \vec{w} \in \mathbb{R}^n$, we have:

$$\mathcal{L}(\vec{v} + \vec{w}) = \mathcal{L}(\vec{v}) + \mathcal{L}(\vec{w}).$$

In particular, a linear transformation $\mathcal{L} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ must map the zero vector in $\mathbb{R}^n$ to the zero vector in $\mathbb{R}^m$, since:

$$\mathcal{L}(\vec{0}_{\mathbb{R}^n}) = \mathcal{L}(0 \cdot \vec{0}_{\mathbb{R}^n}) = 0\mathcal{L}(\vec{0}_{\mathbb{R}^n}) = \vec{0}_{\mathbb{R}^m}.$$

Examples:

- Rotation of $\mathbb{R}^2$ about the origin by a fixed angle.
- Reflection of $\mathbb{R}^3$ over the xy-plane.

Given a linear transformation $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$, since every $\vec{v} \in \mathbb{R}^n$ may be written as $v_1\vec{e}_1 + \cdots v_n\vec{e_n}$, we have:

$$\mathcal{L}(\vec{v}) = v_1\mathcal{L}(\vec{e}_1) + \cdots v_n\mathcal{L}(\vec{e}_n).$$

In other words, $\mathcal{L}$ is uniquely determined by where it sends the basis vectors $\vec{e}_1, \ldots, \vec{e}_n$. All information about $\mathcal{L}$ is captured by the following $m \times n$ **matrix**:

$$L = \begin{pmatrix} | & | & & | \\ \mathcal{L}(\vec{e}_1) & \mathcal{L}(\vec{e}_2) & \cdots & \mathcal{L}(\vec{e}_n) \\ | & | & & | \end{pmatrix}$$

It is an array of real numbers with m rows and n columns, where the i-th column is the vector $\mathcal{L}(\vec{e}_i)$ in $\mathbb{R}^m$.

Given a $m \times n$ matrix:

$$A = (A_{ij}) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}$$

We define the multiplication of A with $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ as follows:

$$A\vec{v} = \begin{pmatrix} \sum_{k=1}^n A_{1k}v_k \\ \sum_{k=1}^n A_{2k}v_k \\ \vdots \\ \sum_{k=1}^n A_{mk}v_k \end{pmatrix}$$

Fact: If L is the $m \times n$ matrix which corresponds to a linear transformation $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$, then for all $\vec{v} \in \mathbb{R}^n$, we have:

$$\mathcal{L}(\vec{v}) = L\vec{v}.$$

Exercise. Given an $m \times n$ matrix A, the map $\mathcal{A} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, defined by:

$$\mathcal{A}(\vec{v}) = A\vec{v}, \quad \vec{v} \in \mathbb{R}^n,$$

is a linear transformation.

Exercise. Given two $m \times n$ matrices A and B, $A\vec{v} = B\vec{v}$ for all $\vec{v} \in \mathbb{R}^n$ if and only if $A = B$, i.e.:

$$A_{ij} = B_{ij}, \quad 1 \le i \le m, 1 \le j \le n.$$

Corollary. Each linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ corresponds to the multiplication by a unique $m \times n$ matrix, and vice versa.

Example. Consider the map $\mathcal{L} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by:

$$\mathcal{L}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x + 2y \\ -y \end{pmatrix}$$

Exercise. $\mathcal{L}$ is a linear transformation.

The matrix corresponding to $\mathcal{L}$ is the 2×2 matrix:

$$L = \left(\mathcal{L}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) \quad \mathcal{L}\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)\right) = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$$

We see that indeed:

$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ -y \end{pmatrix} = \mathcal{L}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)$$

1.1.3 Algebraic Operations on Matrices

- Given any $m \times n$ matrix $A = (A_{ij})$ and scalar $\lambda \in \mathbb{R}$, we define the $m \times n$ matrix $\lambda A = ((\lambda A)_{ij})$ as follows:

$$(\lambda A)_{ij} = \lambda \cdot A_{ij}, \quad 1 \le i \le m, 1 \le j \le n.$$

- Given two $m \times n$ matrices $A = (A_{ij})$, $B = (B_{ij})$, we define their sum $A + B = ((A + B)_{ij})$ as follows:

$$(A + B)_{ij} = A_{ij} + B_{ij}, \quad 1 \le i \le m, 1 \le j \le n.$$

- Given an $m \times n$ matrix C and an $n \times l$ matrix D, the product CD is the $m \times l$ matrix $CD = ((CD)_{ij})$ defined by:

$$(CD)_{ij} = \sum_{k=1}^{n} C_{ik} D_{kj}, \quad 1 \le i \le m, 1 \le j \le l.$$

If we view D as an array of l column vectors in $\mathbb{R}^n$:

$$D = \begin{pmatrix} | & & | \\ \vec{d}_1 & \cdots & \vec{d}_l \\ | & & | \end{pmatrix},$$

then

$$CD = \begin{pmatrix} | & & | \\ C\vec{d}_1 & \cdots & C\vec{d}_l \\ | & & | \end{pmatrix}.$$

Note: Given an $m \times n$ matrix C and an $n' \times l$ matrix D, the product CD is defined **IF AND ONLY IF** $n = n'$.

Example. Let $A = \begin{pmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{pmatrix}$ and $B = \begin{pmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{pmatrix}$, then:

$$AB = \begin{pmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{pmatrix}$$

$$= \begin{pmatrix} 3(-2) - 2(4) & 3(1) - 2(1) & 3(3) - 2(6) \\ 2(-2) + 4(4) & 2(1) + 4(1) & 2(3) + 4(6) \\ 1(-2) - 3(4) & 1(1) - 3(1) & 1(3) - 3(6) \end{pmatrix} = \begin{pmatrix} -14 & 1 & -3 \\ 12 & 6 & 30 \\ -14 & -2 & -15 \end{pmatrix}$$

On the other hand:

$$BA = \begin{pmatrix} -1 & -1 \\ 20 & -22 \end{pmatrix} \neq AB$$

Exercise. If C is an $m \times n$ matrix corresponding to a linear map $\mathcal{C} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, and D is a $n \times l$ matrix corresponding to a linear map $\mathcal{D} : \mathbb{R}^l \longrightarrow \mathbb{R}^n$, then the product CD is the $m \times l$ matrix corresponding to the composition of linear maps:

$$\mathcal{C} \circ \mathcal{D} : \mathbb{R}^l \longrightarrow \mathbb{R}^m.$$