# Chapter 1

# $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$

## The set $\mathbb{R}^n$

- $\mathbb{R}$ is the set of all real numbers.
- Let $n$ be a positive integer.
- We can write $(x_1, x_2, \cdots, x_n)$ of real numbers, which is an ordered $n$-tuple of real numbers.
- $\mathbb{R}^n$ is the set of all ordered $n$-tuples of real numbers.
- An element $x \in \mathbb{R}^n$ can be written in

$$\text{row} \quad x = (x_1, x_2, x_3, \cdots, x_n) \quad \text{or in column} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{or any other form.}$$

  As long as the order of listed $n$ real numbers are seen, there would not be a problem.

# Addition and Scalar multiplication

### Addition

1. If $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ and $y = (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n$, we can add the two to obtain an element in $\mathbb{R}^n$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \\ \vdots \\ x_n + y_n \end{pmatrix},$$

   which is denoted by $x + y \in \mathbb{R}^n$.

### Scalar multiplication

1. In this course, the scalar is a synonym of real number.
2. If $\lambda$ is a real number and $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, then we can scale $x$ to obtain $\lambda x$:

$$\lambda \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \\ \vdots \\ \lambda x_n \end{pmatrix} \in \mathbb{R}^n.$$

# The set $\mathbb{R}^{m\times n}$

- Let $\mathbb{R}$ be again the set of all real numbers.
- Let $m$ and $n$ be two positive integers.
- The set of all $(A_{ij})_{i=1,j=1}^{i=m,j=n}$ of real numbers indexed by $(i,j) \in \{1,2,\cdots,m\} \times \{1,2,\cdots,n\}$ is denoted by $\mathbb{R}^{m\times n}$.
- The notation for an element $A \in \mathbb{R}^{m\times n}$ in this time is more specific. We distinguish two indices $i = 1,2,\cdots,m$ and $j = 1,2,\cdots,n$. We list $mn$ real numbers in a box so that $i$ is a row index, and $j$ is a column index:

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2n} \\ A_{31} & A_{32} & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{m1} & A_{m2} & A_{m3} & \cdots & A_{mn} \end{pmatrix}.$$

- We call an element of $\mathbb{R}^{m\times n}$ an $(m \times n)$ matrix, which reads as "$m$ by $n$ matrix".
- Why ..?

# Addition and Scalar multiplication

### Addition

1. If $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$, i.e., if the matrix $A$ and $B$ are in same shape, we can add the two to obtain an element in $\mathbb{R}^{m\times n}$:

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2n} \\ A_{31} & A_{32} & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{m1} & A_{m2} & A_{m3} & \cdots & A_{mn} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} & B_{13} & \cdots & B_{1n} \\ B_{21} & B_{22} & B_{23} & \cdots & B_{2n} \\ B_{31} & B_{32} & B_{33} & \cdots & B_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ B_{m1} & B_{m2} & B_{m3} & \cdots & B_{mn} \end{pmatrix}$$
$$= \begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} & A_{13}+B_{13} & \cdots & A_{1n}+B_{1n} \\ A_{21}+B_{21} & A_{22}+B_{22} & A_{23}+B_{23} & \cdots & A_{2n}+B_{2n} \\ A_{31}+B_{31} & A_{32}+B_{32} & A_{33}+B_{33} & \cdots & A_{3n}+B_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{m1}+B_{m1} & A_{m2}+B_{m2} & A_{m3}+B_{m3} & \cdots & A_{mn}+B_{mn} \end{pmatrix},$$

which is denoted by $A + B \in \mathbb{R}^{m\times n}$.

### Scalar multiplication

1. If $\lambda$ is a real number and $A \in \mathbb{R}^{m\times n}$, then we can scale $A$ to obtain $\lambda A$:

$$\lambda \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2n} \\ A_{31} & A_{32} & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{m1} & A_{m2} & A_{m3} & \cdots & A_{mn} \end{pmatrix} = \begin{pmatrix} \lambda A_{11} & \lambda A_{12} & \lambda A_{13} & \cdots & \lambda A_{1n} \\ \lambda A_{21} & \lambda A_{22} & \lambda A_{23} & \cdots & \lambda A_{2n} \\ \lambda A_{31} & \lambda A_{32} & \lambda A_{33} & \cdots & \lambda A_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda A_{m1} & \lambda A_{m2} & \lambda A_{m3} & \cdots & \lambda A_{mn} \end{pmatrix}.$$

In summary,

1. We have the set $\mathbb{R}^n$, equipped with the addition and the scalar multiplication, that we denote by $\left(\mathbb{R}^n, +, s\right)$.

$$
\begin{aligned}
+ &: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \\
s &: \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n.
\end{aligned}
$$

2. We havethe set $\mathbb{R}^{m \times n}$, equipped with the addition and the scalar multiplication, that we denote by $\left(\mathbb{R}^{m \times n}, +, s\right)$.

$$
\begin{aligned}
+ &: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \\
s &: \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n.
\end{aligned}
$$

1. That we work with $\big(\mathbb{R}^n, +, s)$, i.e., that $\mathbb{R}^n$ are equipped with the addition and the scalar multiplication may, more importantly, mean that we do not perform other operations.

   **These are illegal expressions:**

   (a) $\mathbb{R}^n + \mathbb{R}$, $n \geq 2$:

   (b) $\mathbb{R}^n + \mathbb{R}^m$, $n \neq m$.

   (c) product in general. $\times : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, though in some dimensions we can define meaningful product.

   (d) comparison in general.

2. Likewise, for the case of $\mathbb{R}^{m\times n}$, we do not perform other operations, unless otherwise defined later of this course.

- We have so far
$$(\mathbb{R}^n, +, s) \quad \text{and} \quad (\mathbb{R}^{m\times n}, +, s),$$
respectively a set equipped with the addition and scalar multiplication.

- Expression that makes use of the addition and scalar multiplication, and use only of them such as
$$\lambda_1 x + \lambda_2 y + \lambda_3 z + \lambda_4 w, \quad x, y, z, w \in \mathbb{R}^n$$
comes as an important expression to us.

- The expression is called a "Linear Combination" of $x, y, z, w$; Linear, meaning that the scalar multiplication $x \mapsto \lambda x$ is to make something proportional, and Combination, meaning here we are adding things.

The final comment on $(\mathbb{R}^n, +, s)$ and $(\mathbb{R}^{m\times n}, +, x)$:

As stated earlier, in some dimensions we may additionally define product, for example in $\mathbb{R}^2$ we can define

$$(x_1, x_2) \times (y_1, y_2) = (v_1, v_2), \quad \text{where} \quad v_1 + iv_2 = (x_1 + ix_2)(y_1 + iy_2).$$

However, all of such operations are treated in this course to be exceptional and come as addendum, emphasizing that there are only $(i)$ the addition and $(ii)$ the scalar multiplication we are allowed to operate so far.

# Chapter 2

# Matrix Matrix multiplication

- Emphasizing that there are no other general product operations, now we define the unique product operation between matrices.
- For given matrix $A \in \mathbb{R}^{\ell \times m}$ and matrix $B \in \mathbb{R}^{m \times n}$, we define the matrix $C \in \mathbb{R}^{\ell \times n}$ to be the product denoted by $AB$ so that

$$C_{ij} = \sum_{\alpha=1}^{m} A_{i\alpha} B_{\alpha j} \quad \text{for } i = 1, 2, \cdots, \ell \text{ and } j = 1, 2, \cdots, n.$$

- Importantly, the product is defined only for the case where the second component of shape of $A$ and the first component of shape of $B$ are the same, and the product is not defined for all remaining cases.
- Why .. ?

The defined product can be regarded in a few different ways.

1. The product can be regarded as to build new $n$ linear combinations of columns of $A$, as designated by numbers in $B$:

2. The product can be regarded as to build new $\ell$ linear combinations of rows of $B$, as designated by numbers in $A$:

3. The product can be regarded as to build new $\ell n$ real numbers, out of rows of $A$ and columns of $B$.

# The produce is useful in many places.

### Perfect for the chain rule in multivariable calculus

Suppose that $X = \mathbb{R}^\ell$, $Y = \mathbb{R}^m$, and $Z = \mathbb{R}^n$, and

$$f : X \to Y, \quad g : Y \to Z.$$

Multivariable Calculus:

1. Also assume $f$ and $g$ are many times diffentiable functions.
2. For a given point $\bar{x}$, we collect numbers
$$B_{\alpha j} = \frac{\partial f^\alpha}{\partial x^j}(\bar{x}) = \lim_{h \to 0} \frac{f^\alpha(\bar{x} + he_j) - f^\alpha(\bar{x})}{h}.$$
3. Similarly, for a given poit $\bar{y}$, we collect numbers
$$A_{i\alpha} = \frac{\partial g^i}{\partial y^\alpha}(\bar{y}) = \lim_{h \to 0} \frac{g^i(\bar{y} + he_\alpha) - g^i(\bar{y})}{h}.$$
4. Now we consider the composition
$$g \circ f : X \to Z, \quad g \circ f(x) = g(f(x))$$
and collect numbers
$$C_{ij} = \frac{\partial (g \circ f)^i}{\partial x^j}(\bar{x}) = \lim_{h \to 0} \frac{(g \circ f)^i(\bar{x} + he_j) - (g \circ f)^i(\bar{x})}{h}.$$
5. Then $C$ turns out to equal to $AB$.

### Writing a system of linear equations

If we are given

$$\begin{cases} 3x - 7y + 4z = -2, \\ 9x - 2y - 6z = 0, \\ -5x + 3y - 11z = -8 \end{cases}$$

It is nice that we can write also

$$\begin{pmatrix} 3 & -7 & 4 \\ 9 & -2 & -6 \\ -5 & 3 & -11 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ -8 \end{pmatrix}$$

Using $\ell = 3$, $m = 3$, and $n = 1$ in the left-hand-side.

## The question of why..?

1. Why does it have to be in that way?
2. Why not we define for $A, B \in \mathbb{R}^{m\times n}$ a product with the same shape in $\mathbb{R}^{m\times n}$ ?

$$\begin{pmatrix} A_{11}B_{11} & A_{12}B_{12} & A_{13}B_{13} & \cdots & A_{1n}B_{1n} \\ A_{21}B_{21} & A_{22}B_{22} & A_{23}B_{23} & \cdots & A_{2n}B_{2n} \\ A_{31}B_{31} & A_{32}B_{32} & A_{33}B_{33} & \cdots & A_{3n}B_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{m1}B_{m1} & A_{m2}B_{m2} & A_{m3}B_{m3} & \cdots & A_{mn}B_{mn} \end{pmatrix}$$

   (a) Actually you can. In most of computer language, this entry-wise product in the same shape is provided by the broadcasting.

3. All why-questions come to the following one,

   Would other people be interested in the new definition ?

## A few important observations

1. Important: You don't mess up with the order of multiplication here.
2. Unlike multiplication we know for two real numbers, for the Matrix-Matrix multiplication for $A \in \mathbb{R}^{\ell \times m}$ and $B \in \mathbb{R}^{m \times n}$, the product $BA$ is even not defined in general. This is simply because in general $n \neq \ell$.
3. In the special case of that $\ell = m = n$, $BA$ is defined. However even in such a case
$$AB \neq BA \quad \text{in general.}$$

## Summary

1. We have $\left(\mathbb{R}^{m \times n}, +, s\right)$ for $m$ and $n$ positive integers.
2. Linear combinations of members of $\mathbb{R}^{m \times n}$ are expressions using $+$ and $s$ such as
$$\lambda_1 A_1 + \lambda_2 A_2 + \cdots \lambda_j A_j \in \mathbb{R}^{m \times n}.$$
3. We have a definition of multiplication for $A \in \mathbb{R}^{\ell \times m}$ and matrix $B \in \mathbb{R}^{m \times n}$ resulting in $AB \in \mathbb{R}^{\ell \times n}$.

Knowledge so far will be applied to mathematics, science, and engineering, and will be extremely important and powerful.