

# 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  $(x_1, x_2, \dots, x_n)$  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

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

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  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  and  $y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ , 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  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ , 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  $(i, j) \in \{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$  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  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ . 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  $(\mathbb{R}^n, +, s)$ .

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

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

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

1. That we work with  $(\mathbb{R}^n, +, s)$ , 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.  $\times : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ , 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  $(\mathbb{R}^{m \times n}, +, s)$ :

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

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

- 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 product 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

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

Multivariable Calculus:

1. Also assume  $f$  and  $g$  are many times differentiable functions.

2. For a given point  $\bar{x}$ , we collect numbers

$$B_{\alpha j} = \frac{\partial f^\alpha}{\partial x^j}(\bar{x}) = \lim_{h \rightarrow 0} \frac{f^\alpha(\bar{x} + h e_j) - f^\alpha(\bar{x})}{h}.$$

3. Similarly, for a given point  $\bar{y}$ , we collect numbers

$$A_{i\alpha} = \frac{\partial g^i}{\partial y^\alpha}(\bar{y}) = \lim_{h \rightarrow 0} \frac{g^i(\bar{y} + h e_\alpha) - g^i(\bar{y})}{h}.$$

4. Now we consider the composition

$$g \circ f : X \rightarrow Z, \quad g \circ f(x) = g(f(x))$$

and collect numbers

$$C_{ij} = \frac{\partial (g \circ f)^i}{\partial x^j}(\bar{x}) = \lim_{h \rightarrow 0} \frac{(g \circ f)^i(\bar{x} + h e_j) - (g \circ f)^i(\bar{x})}{h}.$$

5. Then  $C$  turns out to equal to  $AB$ , for  $\bar{y} = f(\bar{x})$ .

### 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 not even 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  $(\mathbb{R}^{m \times n}, +, s)$  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.



## Chapter 3

# Linear Independence and Basis

Let us consider elements in  $\mathbb{R}^2$ ,

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad z = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

It is easy to notice that if we are with  $x$  and  $y$  in our hands,  $z$  can be reproduced by

$$z = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \end{pmatrix},$$

a linear combination of  $x$  and  $y$ .

## Linear Combinations and Span

For a given finite number of elements  $a_1, a_2, \dots, a_k \in \mathbb{R}^n$ , the set of all linear combinations of them is denoted by

$$\text{span} \langle a_1, a_2, \dots, a_k \rangle = \{x \in \mathbb{R}^n \mid x \text{ is a linear combination of } a_1, a_2, \dots, a_k\}.$$

### Example

Let

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad \cdots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Then,  $e_1, e_2, \dots, e_n$  span all elements in  $\mathbb{R}^n$ , i.e.,  $\text{span} \langle e_1, e_2, \dots, e_n \rangle = \mathbb{R}^n$ .

## Linear Independence

As far as reproducibility is concerned, we notice that we keep only two elements

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 \\ 2 \end{pmatrix},$$

and every element in  $\mathbb{R}^2$  can be reproduced by linear combination of  $x$  and  $y$ .

This necessitates the notion of linear independence of a few elements in  $\mathbb{R}^n$ .

**Example 1**  $z$  above is linearly dependent with respect to  $x$  and  $y$ .

**Example 2** On the other hand, in  $\mathbb{R}^3$ ,

$$x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \quad z = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

Then,  $x, y, z$  are linearly independent. By no chance,  $z$  is a linear combination of  $x$  and  $y$ .

**Definition 1.** We say  $a_1, a_2, \dots, a_k \in \mathbb{R}^n$  are linearly independent if

$$\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_k a_k = 0 \implies \lambda_1 = \lambda_2 = \dots = \lambda_k = 0.$$

$a_1, a_2, \dots, a_k \in \mathbb{R}^n$  are said to be linearly dependent if they are not linearly independent.

## Basis

**Definition 2.** We say  $a_1, a_2, \dots, a_k$  of  $\mathbb{R}^n$  form a basis of  $\mathbb{R}^n$  if

1.  $a_1, a_2, \dots, a_k$  span  $\mathbb{R}^n$ ,
2.  $a_1, a_2, \dots, a_k$  are linearly independent.

## Examples

$e_1, e_2, \dots, e_n$  form a basis.

We will prove in later section the following seemingly natural facts:

1. Elements in  $\mathbb{R}^n$  fewer than  $n$  cannot span all of  $\mathbb{R}^n$ .
2. Elements in  $\mathbb{R}^n$  exceeding  $n$  must linearly dependent.
3.  $n$  elements in  $\mathbb{R}^n$  are linearly independent if and only if they do span all of  $\mathbb{R}^n$ .





## Chapter 4

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

We saw that  $(\mathbb{R}^{m \times n}, +, s)$  is so simply defined. The power of  $\mathbb{R}^{m \times n}$  comes from that, (i) there are so many instances in mathematics, science, engineering, and etc., where entities of interests are such simple that we can add and scale them and that (ii) we will learn nice tools for  $\mathbb{R}^{m \times n}$ .

**Example 1** In fact, this is an example of  $\mathbb{C}^{m \times n}$ , with  $\mathbb{C}$  of complex numbers. Wolfgang Pauli, a physicist, introduced *Pauli matrices* in developing theory of quantum mechanics in early 20c.

We consider a multiplication table for a 4 elements set  $\{1, \sigma_x, \sigma_y, \sigma_z\}$ :

Table 4.1: A multiplicative table

	1	$\sigma_x$	$\sigma_y$	$\sigma_z$
1	1	$\sigma_x$	$\sigma_y$	$\sigma_z$
$\sigma_x$	$\sigma_x$	-1	$\sigma_z$	$-\sigma_y$
$\sigma_y$	$\sigma_y$	$-\sigma_z$	-1	$\sigma_x$
$\sigma_z$	$\sigma_z$	$\sigma_y$	$-\sigma_x$	-1

Note that  $\sigma_x \sigma_y \neq \sigma_y \sigma_x$ , and this set cannot be represented by elements in  $\mathbb{R}$ .

One of the way we achieve the table:

We consider

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \right\}$$

for  $1, \sigma_x, \sigma_y, \sigma_z$  from the left.

The set  $\text{span} \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \right\rangle$  contains all the elements that can be generated.

**note:** The role played by  $\mathbb{R}$ , the scalar factor, can be taken by the mathematical object of *field*. But we will stick to  $\mathbb{R}$  in this course.

**Example 2** Colors (Light)

Although there are so many colors such as yellow, green, blue, white, black,  $\dots$ , it turns out that the only three of them, red, green, and blue are independent. All others are linear combination of the three.

We choose  $[0, 1]^3 \subset \mathbb{R}^3$  represents the colors, with

1.  $(1, 0, 0) \simeq$  red,
2.  $(0, 1, 0) \simeq$  green,
3.  $(0, 0, 1) \simeq$  blue,
4.  $(0, 0, 0) \simeq$  black,
5.  $(1, 1, 1) \simeq$  white.

The value ranged in  $[0, 1]$  represents the intensity.

**Example 3** Bacteria with manipulatable genes.

In Bio-Engineering, certain bacteria is called a *vector*. This reflects that a certain kind of bacteria may or may not contain a few genes of our interests, and can be manipulated in the laboratory in the way of addition.

By a reproduction, a bacteria with gene  $A$  and a bacteria with gene  $B$  can result in a bacteria with both genes  $A$  and  $B$ . The number of such genes that can be manipulated in lab. and bacteria can be represented by  $\mathbb{R}^n$ .

**Example 4** Sauce

Sauce can be perfectly represented by  $\mathbb{R}^n$ . They can be added and scaled:

We linear combine the sugar, salt, vinegar, soy sauce, water, what you want, in grams.

So a recipe you have to remember for a certain sauce is an element in  $\mathbb{R}^n$ , and if you want to make sauce a lot, you scale the element.

1. We saw many examples of *entities* that can be added and scaled, and sometimes we employ multiplication.
2. They are well-represented by  $\mathbb{R}^{m \times n}$  or  $\mathbb{R}^n$ .
3. In all of examples, we see that there are two different roles:

Representation    vs.    Essential Entities

4. The theory has been developed in Linear Algebra so that the two roles are explicitly detached from each other. The latter, of essential entities, is defined in mathematics as *vector space* and *algebra*.
5. In our course, we use the word *vector* to denote the essential entity, while we simply use the word element of  $\mathbb{R}^n$ , or the matrix in  $\mathbb{R}^{m \times n}$  for the representational use.
6. The connection from the set of essential entities to for instance  $\mathbb{R}^n$ , is done by fixing the basis. Let us be clear on this below.

- The definition of vector space  $X$  over the scalar  $\mathbb{R}$  can be found in math textbooks and webpages. For completeness, we also contained the mathematical definition in the last page of this chapter.
- If you are not familiar with mathematical definitions, consider a vector space  $X$  as a set of symbols that can be added and scaled.
- $k$ -dimensional vector space for  $k$  a positive integer.

**Definition 3.** A vector space  $X$  that admits a basis consists of  $k$  elements is called a  $k$ -dimensional vector space.

**Example** The space of colors admits a basis of three elements. Hence the space is 3-dimensional vector space.

- Connection from the *essential entities* to *representing elements* is done once we fix what are the basis elements in  $X$ .

**Example** The space of sauces consists of  $a = \text{sugar}$ ,  $b = \text{salt}$ ,  $c = \text{vinegar}$ , and  $d = \text{water}$ .

Once we fix  $a$ ,  $b$ ,  $c$ , and  $d$  to be the basis, we will understand an element

$$\begin{pmatrix} 5 \\ 5 \\ 3 \\ 100 \end{pmatrix} \in \mathbb{R}^4$$

to represent a sauce consisting of 5-sugar, 5-salt, 3-vinegar, and 100 water, in grams.

We may want to take the diluted ingredients  $a + 100d$ ,  $b + 100d$ ,  $c + 100d$ , and  $d$  to be the basis. Then

$$\begin{pmatrix} 5 \\ 5 \\ 3 \\ 100 \end{pmatrix} \in \mathbb{R}^4$$

to represent a sauce consisting of 5-sugar, 5-salt, 3-vinegar, and 1400 water, in grams.

- You can see the detached roles of *representation* and *essential entities*.
- And how the basis choice connects the two.

A vector space  $X$  over the scalar  $\mathbb{R}$  is mathematically defined as follows.

1.  $X$  is a nonempty set.
2.  $+: X \times X \rightarrow X$  is well-defined with following properties.
  - (a) For any  $x, y, z \in X$ ,  $x + (y + z) = (x + y) + z$ .
  - (b) For any  $x, y \in X$ ,  $x + y = y + x$ .
  - (c)  $\exists 0 \in X$  such that for any  $x \in X$   $x + 0 = x$ .
  - (d) For any  $x \in X$ ,  $\exists(-x) \in X$  such that  $x + (-x) = 0$ .
3.  $s: \mathbb{R} \times X \rightarrow X$  is well-defined with following properties.
  - (a) For any  $\lambda_1, \lambda_2 \in \mathbb{R}$  and  $x \in X$ ,  $\lambda_1(\lambda_2 x) = (\lambda_1 \lambda_2)x$
  - (b) For any  $x \in X$ ,  $1x = x$ .
4.  $+$  and  $s$  work in compatible ways satisfying the following.
  - (a) For any  $\lambda \in \mathbb{R}$  and  $x, y \in X$ ,  $\lambda(x + y) = \lambda x + \lambda y$ .
  - (b) For any  $\lambda_1, \lambda_2 \in \mathbb{R}$  and  $x \in X$ ,  $(\lambda_1 + \lambda_2)x = \lambda_1 x + \lambda_2 x$ .

## Summary

1. For an  $m$ -dimensional vector space  $(X, +, s, \{b_1, b_2, \dots, b_m\})$  with  $m$  basis vectors  $b_1, b_2, \dots, b_m$ , we consider the representation of its element for example

$$\begin{pmatrix} 3 \\ -5 \\ 0 \\ \vdots \\ 0.8 \end{pmatrix} \in \mathbb{R}^{m \times 1}$$

to represent  $3b_1 - 5b_2 + 0b_3 + \dots + 0.8b_m$ .

2. Be careful when  $X$  itself is  $\mathbb{R}^m$ , which is the very possible example. One way we will follow in this course is that, for the representational use, we make use of

$$\mathbb{R}^{m \times 1} \quad \text{of one column matrices.}$$

and try best to distinguish the essential entity  $X$  and its representation.

## Chapter 5

# Transformation of Solid shapes

We will spend long times studying transformations of geometrical shapes.

### Examples

Cube, Airplane, Right Hand.

#### 1. Rotation

- How do we make an airplane heading to a certain direction ?

## 2. Dilation

- Dilation in all direction.
- Dilation differently for semi-axes.
- Dilation differently for a few non-orthogonal directions.

## 3. Reflection with respect to a hyperplane.

- Right hand vs. Left hand.
- Composition of many reflections.



#### 4. Projection

- Same projection twice ?
- Composition of many projections.

#### 5. Shearing

For those solid shapes and transformations, we do the following.

For solid shapes:

- We let the set

$$X = \{\text{Locations}\} \quad \text{with respect to a fixed origin.}$$

- Location with respect to an origin can be scaled by a real number.

We let locations be added using the parallelogram law.

You can check  $X$  of locations to a fixed origin is a vector space.

More precisely,  $X$  is equipped with the basis locations  $e_1, e_2, e_3$  for the case of 3-dimensions. In general, up to  $e_n$  for  $n$ -dimensional locations.

- A solid shape is then a subset occupying a region of locations.

$$G \subset X$$

A transformation of a shape can be then a mapping from Location space to itself.

- A transformation is then a mapping

$$T : X \rightarrow X.$$

- We let the set of all linear transformations from Location space  $X$  to itself:

$$L_{X,X} = \{T : X \rightarrow X \mid T \text{ is linear}\}.$$

- Of course, in general we can think of

$$\begin{aligned} X, & \quad \text{the space of } n\text{-dimensional locations to some fixed origin,} \\ Y, & \quad \text{the space of } m\text{-dimensional locations to some fixed origin,} \\ L_{X,Y}, & \quad \text{the space of linear mappings } T : X \rightarrow Y. \end{aligned}$$

**Representation:**

- We will use

$$\mathbb{R}^{n \times 1}, \quad \mathbb{R}^{n \times n} \quad \text{for } X \text{ and } L_{X,X}.$$

- Indeed,  $n \times n$  matrix  $A$  and  $n \times 1$  column  $x$  can be multiplied to produce another  $n \times 1$  column  $Ax$ , representing a new location.
- By techniques in  $\mathbb{R}^{n \times n}$  (or in general  $\mathbb{R}^{m \times n}$ ), you will learn the surprising fact

that every linear transformation is a few compositions of rotation, dialation, reflection, shearing, and projections. No others.

The composition of linear transformations will be stacking matrix multiplications

$$T = A_k A_{k-1} A_{k-2} \cdots A_2 A_1$$

of  $n \times n$  matrices.

**Home Work**

Suppose  $A \in \mathbb{R}^{k \times \ell}$ ,  $B \in \mathbb{R}^{\ell \times m}$ , and  $C \in \mathbb{R}^{m \times n}$  so that we can multiply  $A$  and  $B$ , and multiply  $B$  and  $C$ . Try to check that

$$(AB)C = A(BC).$$

A map  $T : X \rightarrow Y$  is linear if

for a linear combination  $\lambda_1 x_1 + \lambda_2 x_2 + \cdots \lambda_k x_k$

$T(\lambda_1 x_1 + \lambda_2 x_2 + \cdots \lambda_k x_k)$  must equal to  $\lambda_1 T(x_1) + \lambda_2 T(x_2) + \cdots \lambda_k T(x_k)$ .

- We will write and analyze matrices for rotations, reflections, dilations, shearings, and projections one by one from next lecture.
- This is to study orthogonal matrices, diagonal matrices, symmetric matrices, triangular matrices, and so on, which we will introduce.

We have  $x \in \mathbb{R}^{n \times 1}$  and  $A \in \mathbb{R}^{n \times n}$ , by which, the new location

$$y = Ax \in \mathbb{R}^{n \times 1}$$

We proceed with  $n = 3$ .

Today, we introduce a few kinds of matrices. We will need only one more technique to deal with axes that are not coordinate axes. The point is that other than ones we introduce today + the technique, you do not need to know any further kinds of matrices.

## Identity Matrix

1. We consider the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$$

2. It is a special element in  $\mathbb{R}^{3 \times 3}$ , and denoted by  $I$ .

3. Note that for any  $x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ ,

$$y = Ax = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

4. Be careful that  $I$  is not the one below !

$$I \neq \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

- Diagonal entries are numbers in diagonal of the square.
- Off diagonal entries are entries not in diagonal.
- All of diagonal entries of  $I$  are 1
- All of off diagonal entries of  $I$  are 0.

## Diagonal Matrices

1. Let  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ . We consider a matrix

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

2. First, let us consider the special cases  $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0$

(a) Example 1 :  $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$   $Ax =$

(b) Example 2 :  $x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$   $Ax =$

(c) Example 3 :  $x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$   $Ax =$

(d) Example 4 :  $x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$   $Ax =$

3.  $\lambda_i = 0$ : One component is shrinken to 0.
4.  $\lambda_i < 0$ : Dilate with sign flip.
5. More matrices for the dilations below will be introduced later.
- (a) Principal axes are not coordinate axes.
- (b) Dilation with directions not orthogonal to each other.

## Elementary Rotation Matrices

1. Let  $\theta \in \mathbb{R}$ . We consider a matrix

$$\begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

- (a) Example 1 :  $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$   $Ax =$
- (b) Example 2 :  $x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$   $Ax =$
- (c) Example 3 :  $x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$   $Ax =$
- (d) Example 4 :  $x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$   $Ax =$

2. We consider a matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

(a) Example 1 :  $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$   $Ax =$

(b) Example 2 :  $x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$   $Ax =$

(c) Example 3 :  $x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$   $Ax =$

(d) Example 4 :  $x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$   $Ax =$



3. We consider a matrix

$$\begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

(a) Example 1 :  $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$      $Ax =$

(b) Example 2 :  $x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$      $Ax =$

(c) Example 3 :  $x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$      $Ax =$

(d) Example 4 :  $x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$      $Ax =$

We will introduce more rotation matrices later.

## Elementary Orthogonal Projection Matrices

1. We consider a matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

(a) Example 1 :  $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$      $Ax =$

(b) Example 2 :  $x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$      $Ax =$

(c) Example 3 :  $x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$      $Ax =$

(d) Example 4 :  $x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$      $Ax =$

2. We will introduce more orthogonal projection matrices onto the projecting plane that is not aligned with coordinate.
3. There are non-orthogonal projections, but we will not use them much. We will introduce them later.

## Elementary Reflection Matrices

1. We consider a matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

(a) Example 1 :  $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$      $Ax =$

(b) Example 2 :  $x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$      $Ax =$

(c) Example 3 :  $x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$      $Ax =$

(d) Example 4 :  $x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$      $Ax =$

2. We will introduce more reflection matrices whose symmetry hyperplane is not aligned with coordinate.

## Elementary Shearing Matrices

1. We consider a matrix

$$\begin{pmatrix} \lambda_1 & \alpha & \beta \\ 0 & \lambda_2 & \gamma \\ 0 & 0 & \lambda_3 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

2. This kind of matrices are called upper triangular matrices.

(a) Example 1 :  $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$   $Ax =$

(b) Example 2 :  $x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$   $Ax =$

(c) Example 3 :  $x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$   $Ax =$

(d) Example 4 :  $x = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$   $Ax =$

1. As soon we are able to freely go back and forth from one representation to another representation w.r.t choice of basis vectors, we will know in full generality about matrices of above kinds.
2. Again, you recall that you do not need to know more kinds other than those.

1. Now, we learn the technique to deal with axes oblique to coordinate axes.
2. To be concrete, let us be given an objective to come up with dilation but with respect to three axes, specified by three location points

$$\begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

with dilation factors  $\lambda_1 = \frac{1}{2}$ ,  $\lambda_2 = 2$ , and  $\lambda_3 = 1$ .

3. We will use the three as our basis and seek a *new representations* of locations.
4. Typically, as seen in the example, we would know old representations of your preferred temporal basis.

## I know $B$ . The Old representation of your new basis

Let us list those old representations of your temporal basis in columns of a  $n \times n$  matrix  $B$ :

$$B = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} | & | & | \\ b_1 & b_2 & b_3 \\ | & | & | \end{pmatrix}.$$

All you need to remember is the following equality:

If your new representation of a certain location  $x$  is  $\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix}$ ,

$$\begin{pmatrix} | & | & | \\ b_1 & b_2 & b_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

is nothing but its old representation. To remember:

$$B_{ij'} a_{j'} = a_i.$$

Because

$$\begin{pmatrix} | & | & | \\ b_1 & b_2 & b_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix},$$

(recall that  $B$  is the Old representations of you new basis)

if we want to know what are  $\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix}$ , the temporal representation, we get it by having

$$B^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

---

We will re-visit the inverse matrix later to study it properly, but here we just admit that  $B^{-1}$  of our matrix  $B$  exists with the property we want.

**Definition 4.** For a given  $A \in \mathbb{R}^{3 \times 3}$ , if there exists a matrix  $L \in \mathbb{R}^{3 \times 3}$  such that

$$AL = LA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which must be then unique,  $L$  is said to be the inverse matrix of  $A$ , and denoted by  $A^{-1}$ .

---

If we multiply  $B^{-1}$  both sides,

- (RHS) =  $B^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix},$
- (LHS) =  $B^{-1} B \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} = \text{(RHS)}.$

Now, we are able to achieve our objective of non-elementary dialation:

- Get the new representation, do the elementary dilation in new representation, and don't forget to return the result, not in your representation, but in old representation.

This is to conduct:

$$\begin{aligned}
 x = \begin{pmatrix} a \\ b \\ c \end{pmatrix} &\mapsto \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\
 &\mapsto \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\
 &\mapsto \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = y
 \end{aligned}$$

Because I know that the inverse matrix

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

I can compute one by one.



## Example 2: non-elementary reflection

- We want to compose a matrix that does the reflection with respect to a hyperplane we want.
- That I know what the hyperplane is: I know representations of normal vector and tangent vectors of the hyperplane.
- What we have to do:
  1. Get the new representation of  $x$  in temporal basis of normal and tangents.
  2. Flip the sign of normal. (This is an elementary reflection).
  3. Return back to the old representation.

- For example, let the hyperplane is normal to  $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ .

1. I know one choice of basis:

$$b_1 = \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix}, \quad b_2 = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}, \quad b_3 = \begin{pmatrix} -\frac{\sqrt{6}}{6} \\ -\frac{\sqrt{6}}{6} \\ \frac{2\sqrt{6}}{6} \end{pmatrix}.$$

2. The matrix  $B$  of old representation of new temporal basis is

$$B = \begin{pmatrix} \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & 0 & \frac{2\sqrt{6}}{6} \end{pmatrix}$$

3. We then conduct

$$\begin{aligned} x = \begin{pmatrix} a \\ b \\ c \end{pmatrix} &\mapsto \begin{pmatrix} \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & 0 & \frac{2\sqrt{6}}{6} \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\ &\mapsto \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & 0 & \frac{2\sqrt{6}}{6} \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\ &\mapsto \begin{pmatrix} \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & 0 & \frac{2\sqrt{6}}{6} \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & 0 & \frac{2\sqrt{6}}{6} \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = y \end{aligned}$$