

Linear Algebra

We have learned about linear algebra in a concrete way, such as solving linear equations, matrix operations, etc. In this chapter, we will learn about the abstract concept of linear spaces, which is a generalization of vectors and matrices. First of all, you may ask, why do we need to learn such abstract concepts? check the following example. We have learned about vectors at least in high school, like displacement, velocity, force, etc. actually, when we say vectors, our first impression is that they must have two or three components, which represent a point in a plane or in space. but after we learned about the linear algebra, we learned that something like $(1, 2, 3, 4, 5)$ is also a vector, even though we cannot imagine what its *physical picture* in the real world.

So, basically, the change of the concept of vector to a arbitrary number of components is a generalization or a abstraction to the original concept. so that we can handle more with the new definition. but at the cost of losing some physical meaning. In this lecture, we will repeat this process, to further abstract the concept of vector, so that we can figure out some shared, common, or general properties of these.

let us recall the process of the abstraction from $\mathbb{R}^3$ to $\mathbb{R}^n$, here, I will give a more rigorous definition, as the first step to get you familiar with the algebraic structure.

Recall that if given a structure like (a, b, c, d) , is this a vector? No, definitely not. actually, it is just a tuple, which is an ordered finite sequence or list of elements. btw you can construct it with ordered pairs, anyway. Although we usually by default treat it as a vector, that because we can define a addition and scalar multiplication on it, in a very natural way. but with the structure of a tuple only, we cannot do any operation on it, unless you define some in advance. for example, if you try to run

python

```
1 | (1, 2, 3) + (1, 2, 3)
```

in Python, it will return $(1, 2, 3, 1, 2, 3)$, and if you try to write $(1, 2, 3) + (4, 5, 6)$ on your paper, the reader will think it is $(1 + 4, 2 + 5, 3 + 6)$ by default. that because for the programming languages, the $+$ operator is overloaded to concatenate two tuples. but for the math, especially in the Euclid space, the $+$ operator is defined to add their elements one by one. so you will understand that it is important to declare or at least to know what a symbol exactly means before using it, otherwise you will get confused. so, the tuple is not a vector, but we can define a vector structure on it, and then it becomes a vector.

0.1 APPLICATION OF VECTORS

Then, let's return back to the topic of vectors. for example, we know that force can be decomposed into two orthogonal components, and we use this to analyze mechanics problems. Why can we do that? because the force is a vector. but similarly, a sine signal

$$A \cos(\omega t + \varphi) = (A \cos \varphi) \cos(\omega t) - (A \sin \varphi) \sin(\omega t)$$

can be decomposed into two components $\sin(\omega t)$ and $\cos(\omega t)$, we usually use a constellation to represent this decomposition. Have you ever noticed the similarity between the two cases? therefore, our goal today is to come up with a structure to deal with all the similar cases.

From these two examples above, looking like a list is neither a sufficient condition for being a vector (because addition on lists isn't necessarily vector addition), nor is it a necessary condition (because even if something doesn't look like a list, like a function or a polynomial, it can still behave exactly like a vector). So, Finally, what is a vector? how should we define what a vector is?

Here we will first introduce the concept of field, which are the basic algebraic structures that we will use to define linear spaces.

1 INTRODUCTION TO FIELD

Before we start talking about the vector itself, let's begin from a more basic concept. Our story will starts with a set, the fundamental building block of mathematics. Suppose there is a set S . it is too trivial(boring) just like the tuple we mentioned above. with a set only we can seldom do anything.

So, we want to define a binary operation on the set S , which is a function that maps two elements of the set to the set itself.

DEFINITION 1 (Binary Operation) A binary operation A on a set S is a function

$$A : S \times S \rightarrow S$$

Here is some examples of binary operations:

EXAMPLE 1 (Some Binary Operation Examples)

- $+, -, \times$ defined on $\mathbb{R}$, (is $/$ a binary operation?)
- $\cup, \cap$ defined on sets are also binary operation. (are $\subseteq, \subset$ binary operations?)
- $\wedge, \vee, \oplus$ defined on $\{\top, \perp\}$, (is $\neg$ a binary operation?)

the we can define¹⁾ a **field** is a set S with two binary operations, addition and multiplication, that satisfy certain properties. Note that the name “addition” and “multiplication” are just names, they do not necessarily mean the same as the addition and multiplication of real numbers. The important thing is that these operations satisfy certain properties. then the “addition” and “multiplication” satisfies' the following properties.

DEFINITION 2 (Field) A field is a set S with two binary operations, $+$ and $\cdot$, such that: $a, b, c \in S$

Associativity of addition $(a + b) + c = a + (b + c)$

Commutativity of addition $a + b = b + a$

Additive identity element²⁾ $a + 0 = a$

Inverse element for addition $\forall a \exists b (a + b = 0)$

1) the term “field” comes from the German word “Körper”, which means “body”, and is different concept from the physical field like electric/magnetic field.

2) note that we use the 0 here is just a notation as like the name $+$ and $\times$

Associativity of multiplication $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

Commutativity of multiplication $a \cdot b = b \cdot a$

Multiplicative identity element $a \cdot 1 = a$

Inverse element for multiplication $a \neq 0 \Rightarrow \exists b(a \cdot b = 1)$

Distributivity of multiplication over addition $a \cdot (b + c) = a \cdot b + a \cdot c$

PROOF. $(0 + a = a)$ trivial. by axiom 2. ■

Actually, we call the structure $(S, +)$ that satisfied the first 4 properties above an **abelian group**. So, if we define the abelian group, then a more simple definition can be given as,

DEFINITION 3 A set with two abelian groups, $(S, +)$ for addition and $(S \setminus \{0\}, \cdot)$, and the multiplication operation is distributive over addition. then we call it a field, and denoted as $(S, +, \cdot)$.

the most common example of a field is the rational numbers, you can easily verify that the rational numbers satisfy all the properties of a field. The real numbers are also a field, and so are the complex numbers.

EXAMPLE 2 (Examples of Fields)

- $\mathbb{Q}$ (rational numbers)
- $\mathbb{R}$ (real numbers)
- $\mathbb{C}$ (complex numbers)

Meanwhile, the integers are *not* a field, because they do not have multiplicative inverses for all non-zero elements (for example $1/2 \notin \mathbb{Z}$), the natural numbers are *not* even a group [Task: search for the definition of groups, and tell why.].

2 VECTOR SPACE

Now We're ready to define what a vector space is.

DEFINITION 4 (Vector Space) A vector space V over $\mathbb{F}$, consists of a set V (whose elements are called vectors) and a field $\mathbb{F}$ (whose elements are called scalars), together with two operations:

Vector addition $+: V \times V \rightarrow V$

Scalar multiplication³⁾ $\cdot: \mathbb{F} \times V \rightarrow V$

These operations must satisfy the following axioms. For vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and scalars $a, b \in \mathbb{F}$:

1. **Associativity:** $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
2. **Commutativity:** $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. **Identity element:** There exists $\mathbf{0} \in V$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
4. **Inverse element:** For every $\mathbf{v} \in V$, there exists $\mathbf{w} \in V$ such that $\mathbf{v} + \mathbf{w} = \mathbf{0}$

3) if it is not confusing, the $\cdot$ could be omitted like $\mathbf{av} := a \cdot \mathbf{v}$

5. **Associativity:** $a(bv) = (ab)v$

6. **Identity:** $1v = v^4)$

7. **Distributivity over vector addition:** $a(u + v) = au + av$

8. **Distributivity over scalar addition:** $(a + b)v = av + bv$

⁵⁾ Note that these eight axioms completely characterize what we mean by a vector space. If a set V with operations satisfies these axioms over some field $\mathbb{F}$, then we call it a vector space over $\mathbb{F}$.

For now, we can eventually answer the question, what is a vector? the answer is simple but abstract: A vector is an element of a vector space. And the vector space is defined by the eight axioms above.

Now let's look at some concrete examples to see how this abstract definition applies to familiar and not-so-familiar cases.

EXAMPLE 3 ($\mathbb{R}^n$)

The most familiar example is

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R}\}$$

over the field $\mathbb{R}$. Here:

- Vector addition: $(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$
- Scalar multiplication: $a \cdot (x_1, \dots, x_n) = (ax_1, \dots, ax_n)$
- Additive identity: $(0, 0, \dots, 0)$
- Additive inverse: $(-x_1, \dots, -x_n)$

You can verify that all eight axioms are satisfied.

EXAMPLE 4 (Polynomials)

Let $\mathcal{P}_n(\mathbb{R})$ be the set of all polynomials of degree at most n with real coefficients:

$$\mathcal{P}_n(\mathbb{R}) = \{a_0 + a_1x + a_2x^2 + \dots + a_nx^n \mid a_i \in \mathbb{R}\}$$

Here:

- Vector addition: $(a_0 + a_1x + \dots) + (b_0 + b_1x + \dots) = (a_0 + b_0) + (a_1 + b_1)x + \dots$
- Scalar multiplication: $c \cdot (a_0 + a_1x + \dots) = (ca_0) + (ca_1)x + \dots$
- Additive identity: $0 = 0 + 0x + 0x^2 + \dots$
- Additive inverse: $-(a_0 + a_1x + \dots) = (-a_0) + (-a_1)x + \dots$

Notice how polynomials, which don't "look like" vectors in the geometric sense, still satisfy all the vector space axioms!

4) where 1 is the multiplicative identity in $\mathbb{F}$

5) In other words, if we reuse the definition, the $(V, +_V)$ is an abelian group.

EXAMPLE 5 (Functions)

Let $\mathbb{F}^S$ be the set of all functions mapping a nonempty set S to $\mathbb{F}$. We define $\forall f, g \in \mathbb{F}^S$

- Addition: $(f + g)(x) = f(x) + g(x)$
- Scalar multiplication: $(c \cdot f)(x) = c \cdot f(x)$
- additive identity: const function $0(x) = 0$ for all $x \in S$
- additive inverse: $(-f)(x) = -f(x)$

then $\mathbb{F}^S$ is a vector space over $\mathbb{F}$.

This also forms a vector space over $\mathbb{F}$. Functions can be thought of as “vectors” in a very abstract sense, where the “components” are the function values at each point in the domain. and we can restrict this function vector space to a subset of functions, while still satisfying the vector space axioms.

EXAMPLE 6 (Continuous Functions)

Let $C(I)$ be the set of all continuous functions defined on an interval $I \rightarrow \mathbb{F}$. The structure is defined basically the same as in Example 5, but with the additional property that the functions are continuous over the interval I .

$$C(I) = \left\{ f \in \mathbb{F}^I \mid \left(\forall x_0 \in I \right) \lim_{x \rightarrow x_0} f(x) = f(x_0) \right\}$$

by the continuous functions’ properties, we know that the sum of two continuous functions is also continuous, and the scalar multiplication of a continuous function is also continuous. So this set also forms a vector space over $\mathbb{F}$.

Furthermore, let $C^1(I)$ be the set of all continuously differentiable functions⁶⁾ $I \rightarrow \mathbb{F}$. The structure is defined similarly, but with the additional property that the functions have *continuous first derivatives*. similarly, we can define $C^n(I)$ as the set of all functions with continuous derivatives up to order n .

They are all vector spaces over $\mathbb{F}$. the proof is left to the reader as exercises.

This is the example we started with!

EXAMPLE 7 (Solutions to Linear Differential Equations)

The set of all solutions to the differential equation

$$\{ y \in \mathbb{F}^I \mid y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0 \}$$

is a vector space over $\mathbb{F}$. where $a_i \in \mathbb{F}$ are known coefficients. The addition and scalar multiplication are defined the same as in Example 5.

6) which means both the function and its derivative are continuous

Some useful properties of this vector space can be derived from the definition (axioms) directly. and these properties should be proved before we use them. they are obviously true, but quite tricky to prove.

PROPOSITION 1 (Unique Additive Identity) A vector space has a *unique* additive identity.

PROOF. Suppose there are two additive identities $\mathbf{0}_1$ and $\mathbf{0}_2$ in a vector space V . Then: $\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_1$ (by the definition of additive identity) But also, $\mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_2$ (by the same definition) Therefore, $\mathbf{0}_1 = \mathbf{0}_2$ (by the commutativity of addition). ■

PROPOSITION 2 (Unique Additive Inverse) Every element in a vector space has a unique additive inverse.

PROOF. Suppose V is a vector space. Let $\mathbf{v} \in V$. Suppose $\mathbf{w}$ and $\mathbf{w}'$ are additive inverses of $\mathbf{v}$. Then

$$\mathbf{w} = \mathbf{w} + \mathbf{0} = \mathbf{w} + (\mathbf{v} + \mathbf{w}') = (\mathbf{w} + \mathbf{v}) + \mathbf{w}' = \mathbf{0} + \mathbf{w}' = \mathbf{w}'$$

By the uniqueness of the additive identity, the notation $-\mathbf{v}$ is well-defined as the unique additive inverse of $\mathbf{v}$. and we can define the subtraction operation as $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w})$.

PROPOSITION 3 $0\mathbf{v} = \mathbf{0}$ for every $\mathbf{v} \in V$

PROOF. Let $\mathbf{v} \in V$. By the definition of scalar multiplication, we have: $0\mathbf{v} = (0 + 0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}$ Suppose $-0\mathbf{v}$ is the additive inverse of $0\mathbf{v}$, so that $0\mathbf{v} + (-0\mathbf{v}) = \mathbf{0}$. Then we have:

$$\mathbf{0} = 0\mathbf{v} + (-0\mathbf{v}) = 0\mathbf{v} + 0\mathbf{v} + (-0\mathbf{v}) = 0\mathbf{v}$$

PROPOSITION 4 $a\mathbf{0} = \mathbf{0}$ for every $a \in \mathbb{F}$.

PROOF. Let $a \in \mathbb{F}$ and $\mathbf{0} \in V$ be the additive identity. Then: $a\mathbf{0} = a(\mathbf{0} + \mathbf{0}) = a\mathbf{0} + a\mathbf{0}$ By the definition of additive identity, we have: $a\mathbf{0} + (-a\mathbf{0}) = \mathbf{0}$ Therefore,

$$\mathbf{0} = a\mathbf{0} + (-a\mathbf{0}) = a\mathbf{0} + a\mathbf{0} + (-a\mathbf{0}) = a\mathbf{0}$$

PROPOSITION 5 $(-1)\mathbf{v} = -\mathbf{v}$ for every $\mathbf{v} \in V$.

PROOF.

$$\mathbf{v} + (-1)\mathbf{v} = 1\mathbf{v} + (-1)\mathbf{v} = (1 + (-1))\mathbf{v} = 0\mathbf{v} = \mathbf{0}$$

This equation says that $(-1)\mathbf{v}$, when added to $\mathbf{v}$, gives $\mathbf{0}$. Thus $(-1)\mathbf{v}$ is the additive inverse of $\mathbf{v}$, as desired. ■

3 SUBSPACES

Now that we understand vector spaces, let's talk about subspaces. A subspace is essentially a "vector space within a vector space."

DEFINITION 5 (Subspace) A subset $U \subseteq V$ is called a **subspace** of V if U is also a vector space with the same scalar field, additive identity, addition, and scalar multiplication as on V .

PROPOSITION 6 A subset $U \subseteq V$ over the same $\mathbb{F}$ is a **subspace** of V if and only if:

1. $\mathbf{0} \in U$
2. U is closed under vector addition: $\forall \mathbf{u}, \mathbf{v} \in U, \mathbf{u} + \mathbf{v} \in U$
3. U is closed under scalar multiplication: $\forall \mathbf{v} \in U, \forall a \in \mathbb{F}, a\mathbf{v} \in U$

PROOF. If U is a subspace of V , then U satisfies the three conditions above by the definition of vector space. Conversely, suppose U satisfies the three conditions above. The first condition ensures that the additive identity of V is in U . The second condition ensures that addition makes sense on U . The third condition ensures that scalar multiplication makes sense on U . If $\mathbf{u} \in U$, then $-\mathbf{u}$ (which equals $(-1)\mathbf{u}$ by Proposition 5) is also in U by the third condition above. Hence every element of U has an additive inverse in U . The other parts of the definition of a vector space, such as associativity and commutativity, are automatically satisfied for U because they hold on the larger space V . Thus U is a vector space and hence is a subspace of V . ■

Note that if these three conditions are satisfied, then W automatically inherits all the vector space axioms from V , so (W, F) is itself a vector space. meanwhile, if W is a subset of V but does not satisfy these conditions, it won't be a subspace.

EXAMPLE 8 (Lines through the origin)

In $\mathbb{R}^2$, any line passing through the origin forms a subspace. For instance:

$$U = \{(x, y) \in \mathbb{R}^2 \mid y = 2x\} = \{(t, 2t) \mid t \in \mathbb{R}\}$$

You can verify:

- $(0, 0) \in U$
- If $(t_1, 2t_1), (t_2, 2t_2) \in U$, then $(t_1, 2t_1) + (t_2, 2t_2) = (t_1 + t_2, 2(t_1 + t_2)) \in U$
- If $(t, 2t) \in U$ and $a \in \mathbb{R}$, then $a \cdot (t, 2t) = (at, 2at) \in U$

Similarly, any plane or line through the origin in $\mathbb{R}^3$ is a subspace. but note that those that do not pass through the origin are not subspaces.

EXAMPLE 9 (Even/Odd functions)

In the vector space of $\mathbb{F}^S$, the set of even/odd functions forms a subspace.

- It contains the $0(x) = 0$ function.
- If $f(x)$ and $g(x)$ are even/odd, then $(f + g)(x) = f(x) + g(x)$ is also even/odd.
- If $f(x)$ is even/odd and $c \in \mathbb{R}$, then $(cf)(x) = cf(x)$ is also even/odd.

EXAMPLE 10 (Continuous Functions)

The set $C(I)$ of all continuous functions on an interval I forms a subspace of the vector space of all functions $\mathbb{F}^I$.

- constant value function $0(x) = 0$ is in $C(I)$
- If $f, g \in C(I)$, then $(f + g)(x) = f(x) + g(x)$ is also continuous, so $f + g \in C(I)$
- If $f \in C(I)$ and $c \in \mathbb{R}$, then $(cf)(x) = cf(x)$ is also continuous, so $cf \in C(I)$

EXAMPLE 11 (Solutions to Homogeneous Linear Differential Equations)

The set of all solutions to a linear homogeneous differential equation in Example 7 forms a subspace of the vector space of functions.

- The zero function is a solution (the trivial solution).
- If y_1 and y_2 are solutions, then $(y_1 + y_2)(x) = y_1(x) + y_2(x)$ is also a solution.
- If y is a solution and $c \in \mathbb{R}$, then $(cy)(x) = cy(x)$ is also a solution.

4 SUM OF SUBSPACES

Now let's talk about the sum of two subspaces. Given two subspaces U and W of a vector space V , their sum, denoted $U + W$, is defined as:

DEFINITION 6 (Sum of Subspaces) The sum of two subspaces U and W of a vector space V is the set:

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U, \mathbf{w} \in W\}$$

but note that $U \cup W$ is not the same as $U + W$. The sum $U + W$ is a vector space itself, and it contains all possible sums of vectors from U and W whereas $U \cup W$ just combines the elements of both subspaces so it will not must be a vector space.

EXAMPLE 12

Suppose U is the subspace of all vectors of the form $(x, 0)$ in $\mathbb{R}^2$, and W is the subspace of all vectors of the form $(0, y)$. Then:

$$U + W = \{(x, y) \mid x \in \mathbb{R}, y \in \mathbb{R}\} = \mathbb{R}^2$$

forms the $\mathbb{R}^2$ plane while $U \cup W = \{(x, 0) \mid x \in \mathbb{R}\} \cup \{(0, y) \mid y \in \mathbb{R}\}$ is just the union of x -axis and y -axis, which is not a vector space.

PROPOSITION 7 (Sum of Subspaces is the smallest containing subspace) The sum $V_1 + \dots + V_m$ of subspaces $V_1, \dots, V_m$ of V is the smallest subspace of V containing each of $V_1, \dots, V_m$.

PROOF. The reader can verify that $V_1 + \dots + V_m$ contains the additive identity $\mathbf{0}$ and is closed under addition and scalar multiplication. Thus it is a subspace of V .

The subspaces $V_1, \dots, V_m$ are all contained in $V_1 + \dots + V_m$ (to see this, consider sums $\mathbf{v}_1 + \dots + \mathbf{v}_m$ where all except one of the $\mathbf{v}_k$'s are $\mathbf{0}$). Conversely, every subspace of V containing $V_1, \dots, V_m$ contains $V_1 + \dots + V_m$ (because subspaces must contain all finite sums of their elements). Thus $V_1 + \dots + V_m$ is the smallest subspace of V containing $V_1, \dots, V_m$. ■