Contents

1

Vector Spaces

1.1 Fields, Scalars and Vectors

In elementary mathematics, we often refer to a vector as an ordered tuple of numbers with a direction and a magnitude. However, there is a much more abstract aspect to the notion of vectors. In fact, let us first generalise the notion of *scalars*, which are taken as complex constants in an elementary level.

In general, we have the following algebraic structure:

Definition 1.1.1 ▸ Field

A **field** is a set $\mathcal{F}$ with two binary operations $\mathcal{F}^2 \to \mathcal{F}$, namely addition and multiplication, such that

1. $u + v = v + u$ for all $u, v \in \mathcal{F}$;
2. $(u + v) + w = u + (v + w)$ for all $u, v, w \in \mathcal{F}$;
3. $uv = vu$ for all $u, v \in \mathcal{F}$;
4. $(uv)w = u(vw)$ for all $u, v, w \in \mathcal{F}$;
5. $u(v + w) = uv + uw$ for all $u, v, w \in \mathcal{F}$;
6. there exists $0 \in \mathcal{F}$ such that $u + 0 = u$ for all $u \in \mathcal{F}$;
7. there exists $1 \in \mathcal{F}$ such that $1u = u$ for all $u \in \mathcal{F}$;
8. for every $u \in \mathcal{F}$, there exists some $v \in \mathcal{F}$ such that $u + v = 0$;
9. for every $u \in \mathcal{F}$, there exists some $v \in \mathcal{F}$ such that $uv = 1$.

One may check that both $\mathbb{R}$ and $\mathbb{C}$ are fields. It turns out that we can also generalise the concept of vectors as any objects which possess properties similar to that of Euclidean vectors, i.e., we can view a vector as a mathematical quantity which can be added up and multiplied by another quantity called a scalar with some axioms which they follow. Rigorously, we define the notion of a *vector space*.

Definition 1.1.2 ▸ Vector Space

A **vector space** is a set V over a field $\mathcal{F}$ with two binary operations, namely

- addition $+\colon V^2 \to V$, and
- scalar multiplication $(\quad)(\quad)\colon \mathcal{F} \times V \to V$,

such that

1. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ for all $\boldsymbol{u}, \boldsymbol{v} \in V$;
2. $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$ for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$;
3. $ab\boldsymbol{v} = a(b\boldsymbol{v})$ for all $a, b \in \mathcal{F}$ and $\boldsymbol{v} \in V$;
4. there exists an **additive identity** or **zero vector** $\mathbf{0} \in V$ such that $\boldsymbol{v} + \mathbf{0} = \boldsymbol{v}$ for all $\boldsymbol{v} \in V$;
5. every $\boldsymbol{v} \in V$ has an **additive inverse** $\boldsymbol{w} \in V$ with $\boldsymbol{v} + \boldsymbol{w} = 0$;
6. there exists a **multiplicative identity** $1 \in \mathcal{F}$ such that $1\boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v} \in V$;
7. $a(\boldsymbol{u} + \boldsymbol{v}) = a\boldsymbol{u} + a\boldsymbol{v}$ and $(a + b)\boldsymbol{u} = a\boldsymbol{u} + b\boldsymbol{u}$ for all $a, b \in \mathcal{F}$ and $\boldsymbol{u}, \boldsymbol{v} \in V$.

Notice that here, the definitions of addition in scalar multiplication in a vector space imply that any vector space must be **closed** under these two operations. Notice also that the operations "addition" and "scalar multiplication" are not necessary the addition and scalar multiplication which we are used to in $\mathbb{R}^n$, but abstract mappings which satisfy the given axioms.

We shall prove a few basic properties regarding vector spaces.

Theorem 1.1.3 ▸ Uniqueness of Additive Identity

Let V be a vector space with $\mathbf{0} \in V$ as an additive identity, then $\mathbf{0}$ is unique.

Proof. Suppose on contrary that there exists $\boldsymbol{u} \in V$ such that $\boldsymbol{v} + \boldsymbol{u} = \boldsymbol{v}$ for all $\boldsymbol{v} \in V$. Since $\mathbf{0} \in V$, we have

$$\mathbf{0} + \boldsymbol{u} = \mathbf{0}.$$

However, $\mathbf{0}$ is the additive identity, so

$$\boldsymbol{u} = \boldsymbol{u} + \mathbf{0} = \mathbf{0} + \boldsymbol{u} = \mathbf{0},$$

i.e. $\mathbf{0}$ is unique. □

Similarly, we can also prove the uniqueness of additive inverse.

Theorem 1.1.4 ▸ Uniqueness of Additive Inverse

Let V be a vector space, then every $\boldsymbol{v} \in V$ has a unique additive inverse.

Proof. Suppose on contrary that there exist $\boldsymbol{u}, \boldsymbol{w} \in V$ both being additive inverse of $\boldsymbol{v}$, then $\boldsymbol{u} + \boldsymbol{v} = \mathbf{0}$ and $\boldsymbol{w} + \boldsymbol{v} = \mathbf{0}$. Therefore,

$$\boldsymbol{u} = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{u} = (\boldsymbol{w} + \boldsymbol{v}) + \boldsymbol{u} = \boldsymbol{w} + (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{w},$$

i.e., $\boldsymbol{v}$ has a unique additive inverse. □

Theorem 1.1.4 justifies the notation $-\boldsymbol{u}$ to denote the additive inverse of $\boldsymbol{u}$. However, so far we have not ascertained the fact that $-\boldsymbol{u} = (-1)\boldsymbol{u}$ (note that the former means the inverse of $\boldsymbol{u}$ while the latter means $\boldsymbol{u}$ multiplied by the scalar -1)! While seemingly innocent, this result is not as easily proven as it looks.

First, we shall justify that $0\boldsymbol{u} = \mathbf{0}$ for all $\boldsymbol{u} \in V$. Notice that

$$0\boldsymbol{u} = (0+0)\boldsymbol{u} = 0\boldsymbol{u} + 0\boldsymbol{u}.$$

Adding $-(0\boldsymbol{u})$ to both sides of the equation yields $0\boldsymbol{u} = \mathbf{0}$ as desired. From this result we see that

$$(-1)\boldsymbol{u} + \boldsymbol{u} = (-1+1)\boldsymbol{u} = 0\boldsymbol{u} = \mathbf{0}.$$

By uniqueness of additive inverse, we must have $(-1)\boldsymbol{u} = -\boldsymbol{u}$.

Note that by using a similar technique we can prove that $a\mathbf{0} = \mathbf{0}$ for all $a \in \mathcal{F}$, and so $\mathbf{0} = -\mathbf{0}$ as a consequence.

Additionally, note that subtraction is defined as $\mathbf{u} - \mathbf{v} = \boldsymbol{u} + (-1)\boldsymbol{v}$, so the above result allows us to write $\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{u} + (-\boldsymbol{v})$.

1.1.1 Subspaces

Note that a vector space is extended based on a set of vectors, so we can define *subspaces* similarly to the notion of subsets.

Definition 1.1.5 ▸ Subspace

Let V be a vector space. $U \subseteq V$ is called a **subspace** if U is a vector space under addition and scalar multiplication in V.

It is easy to see that the intersection of any number of subspaces of a vector space V is still a subspace of V, but the union might not be so. In particular, we would like to consider a special construct known as *direct sum*.

Definition 1.1.6 ▸ Direct Sum

Let V be a vector space and $U_1, U_2 \subseteq V$ such that $U_1 \cap U_2 = \{\mathbf{0}\}$, then their **direct sum** is defined as

$$U_1 \oplus U_2 := \{\boldsymbol{u}_1 + \boldsymbol{u}_2 : \boldsymbol{u}_1 \in U_1, \boldsymbol{u}_2 \in U_2\}.$$

More generally, we can let U_1 and U_2 be any subsets of V and define $U_1 + U_2$ in the same manner, which is known as the *sum* of U_1 and U_2.

It can be easily proven that for any vector space V, the direct sum of any two subspaces of V is still a subspace of V. A nice property of direct sum can be proven as follows:

Proposition 1.1.7 ▸ Unique Decomposition with Direct Sums

Let $V = U_1 \oplus U_2$, then every $\boldsymbol{v} \in V$ can be uniquely expressed as $\boldsymbol{u} + \boldsymbol{w}$ for some $\boldsymbol{u} \in U_1$ and $\boldsymbol{w} \in U_2$.

Proof. The existence of $\boldsymbol{u}$ and $\boldsymbol{w}$ is trivial by Definition 1.1.6. Suppose there exist $\boldsymbol{u}' \in U_1$ and $\boldsymbol{w}' \in U_2$ such that $\boldsymbol{u} + \boldsymbol{w} = \boldsymbol{u}' + \boldsymbol{w}'$, then we have $\boldsymbol{u} - \boldsymbol{u}' = \boldsymbol{w}' - \boldsymbol{w}$. Note that $\boldsymbol{u} - \boldsymbol{u}' \in U_1$ and $\boldsymbol{w}' - \boldsymbol{w} \in U_2$, so we have $\boldsymbol{u} - \boldsymbol{u}', \boldsymbol{w}' - \boldsymbol{w} \in U_1 \cap U_2 = \{\boldsymbol{0}\}$, i.e.,

$$\boldsymbol{u} - \boldsymbol{u}' = \boldsymbol{w}' - \boldsymbol{w} = \boldsymbol{0}.$$

Therefore, $\boldsymbol{u} = \boldsymbol{u}'$ and $\boldsymbol{w} = \boldsymbol{w}'$, i.e., $\boldsymbol{u}$ and $\boldsymbol{w}$ are unique. □

In some sense, a direct sum of V can be viewed as a "partition" of V into two subsets with a minimal overlap. Note that unlike partition in its real definition, the subspaces U_1 and U_2 here cannot be disjoint sets as both of them have to contain the zero vector in V. More generally, for any subspace $U \subseteq V$, we have $\boldsymbol{0}_U = \boldsymbol{0}_V$, the proof of which should be trivial enough as an exercise to the reader.

In particular, we would like to consider $\mathcal{F}^n$ for a general field $\mathcal{F}$. We can define the dot product operation over $\mathcal{F}^n$ in the same way as $\mathbb{R}^n$. Take any subspace $U \subseteq \mathcal{F}^n$ and define the set

$$U_\perp \coloneqq \{\boldsymbol{u} \in \mathcal{F}^n : \boldsymbol{u} \cdot \boldsymbol{v} = 0 \quad \text{for all } \boldsymbol{v} \in U\},$$

then $\mathcal{F}^n = U \oplus U_\perp$.

To justify this, we first take any $\boldsymbol{v} \in \mathcal{F}^n$. Using some calculus, we can show that there exists

$$\boldsymbol{u}_0 = \operatorname*{argmin}_{\boldsymbol{u} \in U} |\boldsymbol{u} \cdot \boldsymbol{v}|.$$

Let $\boldsymbol{w} = \boldsymbol{v} - \boldsymbol{u}_0$, then clearly $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{u}_0$ where $\boldsymbol{u}_0 \in U$ and $\boldsymbol{w} \in U_\perp$. This implies that $V = U + U_\perp$. Note that $\boldsymbol{0}$ is the only vector in $\mathcal{F}^n$ which is orthogonal to itself, so we have $U \cap U_\perp = \{\boldsymbol{0}\}$. It follows that $V = U \oplus U_\perp$.