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

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

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

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

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

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 $\boldsymbol{u} - \boldsymbol{u}', \boldsymbol{w}' - \boldsymbol{w} \in U_1 \cap U_2 = \{\mathbf{0}\}$, i.e.,

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

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 $\mathbf{0}_U = \mathbf{0}_V$, 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 $\mathbf{0}$ is the only vector in $\mathcal{F}^n$ which is orthogonal to itself, so we have $U \cap U_\perp = \{\mathbf{0}\}$. It follows that $V = U \oplus U_\perp$.

1.2 Isomorphism

Since the underlying structure of a vector space is still a set, the notion of a mapping between two vector spaces is well-defined. However, note that a vector space possesses unique algebraic structures and properties, namely that the linear combinations of any

members of the space are still in the space, so we would like to focus on mapping which preserves such properties.

Definition 1.2.1 ▶ Homomorphism

Let U and V be vector spaces, a **homomorphism** is a mapping $\phi : U \to V$ such that

$$\phi(\boldsymbol{u} + \boldsymbol{v}) = \phi(\boldsymbol{u}) + \phi(\boldsymbol{v})$$

for any $\boldsymbol{u}, \boldsymbol{v} \in U$.

Note that it suffices to only require $\phi(\boldsymbol{u} + \boldsymbol{v}) = \phi(\boldsymbol{u}) + \phi(\boldsymbol{v})$ but not $\phi(c\boldsymbol{u}) = c\phi(\boldsymbol{u})$. Here, note that a vector space is a connected set, so we can easily prove the above claim with some knowledge in mathematical analysis.

Naturally, if a homomorphism is bijective, then it means that the elements in two vector spaces have a one-to-one correspondence. In practice, this means we can treat them as equivalent spaces in some sense.

Definition 1.2.2 ▶ Isomorphism

An **isomorphism** between vector spaces U and V is a homomorphism between them which is bijective.

An interesting fact here is that an isomorphism between any vector spaces is not unique. To see this, let us first consider an arbitrary vector space V. Now, we can always find the trivial isomorphism $\mathrm{id}_V : V \to V$. In fact, any mapping $\phi : \boldsymbol{v} \mapsto c\boldsymbol{v}$ where c is a scalar is clearly an isomorphism from V to V. This means that there are infinitely many isomorphisms from V to itself.

Let U be an arbitrary vector space such that there exists some isomorphism $\psi : V \to U$. We consider the following theorem:

Theorem 1.2.3 ▶ Composition Preserves Isomorphism

Let U, V, W be vector spaces. If $\phi : U \to V$ and $\psi : V \to W$ are isomorphisms, then the composite mapping $\phi \circ \psi : U \to W$ is an isomorphism.

Proof. Since both ϕ and ψ are bijective, it is clear that $\psi \circ \phi$ is bijective. Take any $\boldsymbol{u}, \boldsymbol{v} \in U$, then

$$\psi\big(\phi(\boldsymbol{u} + \boldsymbol{v})\big) = \psi\big(\phi(\boldsymbol{u}) + \phi(\boldsymbol{v})\big) = \psi\big(\phi(\boldsymbol{u})\big) + \psi\big(\phi(\boldsymbol{v})\big)$$

since $\phi(\boldsymbol{u}), \phi(\boldsymbol{v}) \in V$. Therefore, $\psi \circ \phi$ is an isomorphism. □

Using Theorem 1.2.3, we can immediately see that if $\phi : V \to V$ is any isomorphism and $\psi : V \to U$ is an isomorphism, then $\psi \circ \phi$ is an isomorphism between V and U. Therefore, there are infinitely many isomorphisms between V and U.

> *Remark.* If V is isomorphic to U, we write $V \cong U$. Clearly, $\cong$ is an equivalence relation.

Now, observe that for any field $\mathcal{F}$, the set $\mathcal{F}^n$ for any $n \in \mathbb{N}$ is a vector space over $\mathcal{F}$.

Definition 1.2.4 ▸ Finite-Dimensional Vector Space

A vector space V is said to be **finite-dimensional** over a field $\mathcal{F}$ if it it is isomorphic to $\mathcal{F}^n$ for some $n \in \mathbb{N}$. n is called the **dimension** of V.

Obviously, a vector space which is not finite-dimensional is called *infinite-dimensional*. For example, the set of all polynomials is a vector space of infinite dimension.

1.3 Basis

Recall that a *linear combination* of vectors is in the form of

$$\sum a_i \boldsymbol{v}_i = a_1 \boldsymbol{v}_1 + a_2 \boldsymbol{v}_2 + \cdots.$$

In case where no confusion is caused, this can be abbreviated as $a_i \boldsymbol{v}_i$. Recall also that a *span* of a set of vectors is defined as

$$\operatorname{span}(V) \coloneqq \{a_i \boldsymbol{v}_i : \boldsymbol{v}_i \in V\},$$

where a_i's are scalars. A span of a subset of a vector space V is clearly a subspace of V. We also know that a set of vectors S is said to be *linearly independent* if and only if $a_i \boldsymbol{v}_i = \boldsymbol{0}$ implies that $a_i = 0$ for all $i = 1, 2, \cdots$. A *basis* of a vector space V is a linearly independent set S such that $\operatorname{span}(S) = V$. One may check that if S is a basis for V, then any $\boldsymbol{v} \in V$ can be **uniquely** expressed as a linear combination of the members of S, but S itself is not unique. In particular, the coefficients in this linear combination is known as the *components* of $\boldsymbol{v}$ relative to S.

We will see that the basis is closely related to the dimension of vector spaces. First, let us consider the trivial basis for $\mathcal{F}^n$.

Definition 1.3.1 ▶ Canonical Basis

The **canonical basis** for $\mathcal{F}^n$ is defined as $\{\boldsymbol{e}_i : i = 1, 2, \cdots, n\}$, where each $\boldsymbol{e}_i$ is a column vector with $1_\mathcal{F}$ in its i-th row and $0_\mathcal{F}$ in the other rows.

It is easy to see that the number of vectors in any basis of a finite-dimensional vector space V is uniquely equal to its dimension.

Proposition 1.3.2 ▶ Dimension as Cardinality of Basis

Let V be a finite-dimensional vector space with dimension n and basis S, then $n = |S|$.

Proof. Note that $V \cong \mathcal{F}^n$. Let $\boldsymbol{v} \in V$ be an arbitrary vector, then there is some $\boldsymbol{u} \in \mathcal{F}^n$ and a bijection $\phi : \mathcal{F}^n \to V$ such that

$$\begin{aligned}\boldsymbol{v} &= \phi(\boldsymbol{u}) \\ &= \phi\left(\sum_{i=1}^{n} a_i \boldsymbol{e}_i\right) \\ &= \sum_{i=1}^{n} a_i \phi(\boldsymbol{e}_i),\end{aligned}$$

where $a_i \in \mathcal{F}$ for $i = 1, 2, \cdots, n$. This means that V is spanned by at most n vectors and so its basis is finite. Suppose on contrary that $|S| = m < n$, then for any $\boldsymbol{w} \in \mathcal{F}^n$, there is some $\boldsymbol{r} \in V$ such that

$$\begin{aligned}\boldsymbol{w} &= \phi^{-1}(\boldsymbol{r}) \\ &= \phi^{-1}\left(\sum_{i=1}^{m} b_i \boldsymbol{s}_i\right) \\ &= \sum_{i=1}^{m} b_i \phi^{-1}(\boldsymbol{s}_i),\end{aligned}$$

where $b_i \in \mathcal{F}$ for $i = 1, 2, \cdots, m$. This means that $\mathcal{F}^n$ is spanned by m vectors, which is a contradiction. Therefore, $|S| = n$. □

An immediate corollary from Proposition 1.3.2 is that a finite-dimensional vector space always has a unique dimension, as otherwise it will have two bases with different cardinalities.

Note that since every vector in a vector space V can be uniquely expressed as a linear combination of a basis S for V, this really means that we can view the notion of a basis equivalently as a bijection between $\mathcal{F}^n$ and V, i.e., for any $(a_1, a_2, \cdots, a_n) \in \mathcal{F}^n$, we can map the tuple to a vector in V whose components are exactly $a_1, a_2, \cdots, a_n$.

Now, let us denote a basis for V by the mapping z. Notice that for any $\boldsymbol{c}_1, \boldsymbol{c}_2 \in \mathcal{F}^n$, we have

$$z(\boldsymbol{c}_1 + \boldsymbol{c}_2) = z(\boldsymbol{c}_1) + z(\boldsymbol{c}_2),$$

so a basis is nothing more but an isomorphism!