1.2 Vector Spaces

The notion of the vector presented in the previous section is here re-cast in a more formal and abstract way, using some basic concepts of Linear Algebra and Topology. This might seem at first to be unnecessarily complicating matters, but this approach turns out to be helpful in unifying and bringing clarity to much of the theory which follows.

Some background theory which complements this material is given in Appendix A to this Chapter, §1.A.

1.2.1 The Vector Space

The vectors introduced in the previous section obey certain rules, those listed in §1.1.3. It turns out that many other mathematical objects obey the same list of rules. For that reason, the mathematical structure defined by these rules is given a special name, the **linear space** or **vector space**.

First, a **set** is any well-defined list, collection, or class of objects, which could be finite or infinite. An example of a set might be

$$B = \{x \mid x \le 3\} \tag{1.2.1}$$

which reads "*B* is the set of objects *x* such that *x* satisfies the property $x \le 3$". Members of a set are referred to as **elements**.

Consider now the **field**[1] of real numbers *R*. The elements of *R* are referred to as **scalars**. Let *V* be a non-empty set of elements $\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots$ with rules of **addition** and **scalar multiplication**, that is there is a **sum** $\mathbf{a} + \mathbf{b} \in V$ for any $\mathbf{a}, \mathbf{b} \in V$ and a **product** $\alpha\mathbf{a} \in V$ for any $\mathbf{a} \in V$, $\alpha \in R$. Then *V* is called a **(real)**[2] **vector space** over *R* if the following eight axioms hold:

1. *associative law for addition*: for any $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$, one has $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$
2. *zero element*: there exists an element $\mathbf{o} \in V$, called the zero element, such that $\mathbf{a} + \mathbf{o} = \mathbf{o} + \mathbf{a} = \mathbf{a}$ for every $\mathbf{a} \in V$
3. *negative* (or *inverse*): for each $\mathbf{a} \in V$ there exists an element $-\mathbf{a} \in V$, called the negative of **a**, such that $\mathbf{a} + (-\mathbf{a}) = (-\mathbf{a}) + \mathbf{a} = \mathbf{o}$
4. *commutative law for addition*: for any $\mathbf{a}, \mathbf{b} \in V$, one has $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
5. *distributive law, over addition of elements of V*: for any $\mathbf{a}, \mathbf{b} \in V$ and scalar $\alpha \in R$, $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$
6. *distributive law, over addition of scalars*: for any $\mathbf{a} \in V$ and scalars $\alpha, \beta \in R$, $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$

---

[1] A **field** is another mathematical structure (see Appendix A to this Chapter, §1.A). For example, the set of complex numbers is a field. In what follows, the only field which will be used is the familiar set of real numbers with the usual operations of addition and multiplication.

[2] "real", since the associated field is the reals. The word *real* will usually be omitted in what follows for brevity.

7. *associative law for multiplication*: for any $\mathbf{a} \in V$ and scalars $\alpha, \beta \in R$, $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$
8. *unit multiplication*: for the unit scalar $1 \in R$, $1\mathbf{a} = \mathbf{a}$ for any $\mathbf{a} \in V$.

The set of vectors as objects with "magnitude and direction" discussed in the previous section satisfy these rules and therefore form a vector space over *R*. However, despite the name "vector" space, other objects, which are *not* the familiar geometric vectors, can also form a vector space over *R*, as will be seen in a later section.

1.2.2 Inner Product Space

Just as the vector of the previous section is an element of a vector space, next is introduced the notion that the vector dot product is one example of the more general **inner product**.

First, a **function** (or **mapping**) is an assignment which assigns to *each* element of a set *A* a *unique* element of a set B, and is denoted by

$$f : A \to B \tag{1.2.2}$$

An **ordered pair** (a,b) consists of two elements *a* and *b* in which one of them is designated the first element and the other is designated the second element The **product set** (or **Cartesian product**) $A \times B$ consists of all ordered pairs (a,b) where $a \in A$ and $b \in B$:

$$A \times B = \{(a,b) \mid a \in A, b \in B\} \tag{1.2.3}$$

Now let *V* be a real vector space. An **inner product** (or **scalar product**) on *V* is a mapping that associates to each ordered pair of elements **x**, **y**, a scalar, denoted by $\langle x, y \rangle$,

$$\langle \cdot,\cdot \rangle : V \times V \to R \tag{1.2.4}$$

that satisfies the following properties, for $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\alpha \in R$:

1. *additivity*: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
2. *homogeneity*: $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$
3. *symmetry*: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
4. *positive definiteness*: $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ when $\mathbf{x} \neq \mathbf{o}$

From these properties, it follows that, if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for all $\mathbf{y} \in V$, then $\mathbf{x} = 0$

A vector space with an associated inner product is called an **inner product space**.

Two elements of an inner product space are said to be **orthogonal** if

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0 \tag{1.2.5}$$

and a set of elements of *V*, $\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots\}$, are said to form an **orthogonal set** if every element in the set is orthogonal to every other element:

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0, \quad \langle \mathbf{x}, \mathbf{z} \rangle = 0, \quad \langle \mathbf{y}, \mathbf{z} \rangle = 0, \quad \text{etc.} \tag{1.2.6}$$

The above properties are those listed in §1.1.4, and so the set of vectors with the associated dot product forms an inner product space.

Euclidean Vector Space

The set of real triplets (x_1, x_2, x_3) under the usual rules of addition and multiplication forms a vector space R^3. With the inner product defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3$$

one has the inner product space known as (three dimensional) **Euclidean vector space**, and denoted by *E*. This inner product allows one to take distances (and angles) between elements of *E* through the norm (length) and metric (distance) concepts discussed next.

1.2.3 Normed Space

Let *V* be a real vector space. A **norm** on *V* is a real-valued function,

$$\| \ \| : V \to R \tag{1.2.7}$$

that satisfies the following properties, for $\mathbf{x}, \mathbf{y} \in V$, $\alpha \in R$:

1. *positivity*: $\|\mathbf{x}\| \geq 0$
2. *triangle inequality*: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$
3. *homogeneity*: $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$
4. *positive definiteness*: $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{o}$

A vector space with an associated norm is called a **normed vector space**. Many different norms can be defined on a given vector space, each one giving a different normed linear space. A natural norm for the inner product space is

$$\|\mathbf{x}\| \equiv \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \tag{1.2.8}$$

It can be seen that this norm indeed satisfies the defining properties. When the inner product is the vector dot product, the norm defined by 1.2.8 is the familiar vector "length".

One important consequence of the definitions of inner product and norm is the **Schwarz inequality**, which states that

$$\left|\langle \mathbf{x}, \mathbf{y} \rangle\right| \le \|\mathbf{x}\|\|\mathbf{y}\| \tag{1.2.9}$$

One can now define the **angle** between two elements of *V* to be

$$\theta : V \times V \to R, \qquad \theta(\mathbf{x}, \mathbf{y}) \equiv \cos^{-1}\left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}\right) \tag{1.2.10}$$

The quantity inside the curved brackets here is necessarily between -1 and $+1$, by the Schwarz inequality, and hence the angle θ is indeed a real number.

1.2.4 Metric Spaces

Metric spaces are built on the concept of "distance" between objects. This is a generalization of the familiar distance between two points on the real line.

Consider a set *X*. A **metric** is a real valued function,

$$d(\cdot,\cdot) : X \times X \to R \tag{1.2.11}$$

that satisfies the following properties, for $\mathbf{x}, \mathbf{y} \in X$:

1. positive: $d(\mathbf{x}, \mathbf{y}) \ge 0$ and $d(\mathbf{x}, \mathbf{x}) = 0$, for all $\mathbf{x}, \mathbf{y} \in X$
2. strictly positive: if $d(\mathbf{x}, \mathbf{y}) = 0$ then $\mathbf{x} = \mathbf{y}$, for all $\mathbf{x}, \mathbf{y} \in X$
3. symmetry: $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$, for all $\mathbf{x}, \mathbf{y} \in X$
4. triangle inequality: $d(\mathbf{x}, \mathbf{y}) \le d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$, for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$

A set *X* with an associated metric is called a **metric space**. The set *X* can have more than one metric defined on it, with different metrics producing different metric spaces.

Consider now a normed vector space. This space naturally has a metric defined on it:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \tag{1.2.12}$$

and thus the normed vector space *is* a metric space. For the set of vectors with the dot product, this gives the "distance" between two vectors $\mathbf{x}, \mathbf{y}$.

1.2.5 The Affine Space

Consider a set *P*, the elements of which are called **points**. Consider also an associated vector space *V*. *P* is an **affine space** when:

(i) given two points $p, q \in P$, one can define a **difference**, $q - p$ which is a unique element **v** of *V*, i.e. $\mathbf{v} \equiv \mathbf{v}(q, p) = q - p \in V$ (called a **translation vector**),
(ii) given a point $p \in P$ and $\mathbf{v} \in V$, one can define the **sum** $\mathbf{v} + p$ which is a unique point *q* of *P*, i.e. $q = \mathbf{v} + p \in P$,

and for which the following property holds: for $p, q, r \in P$:

$$(q - r) + (r - p) = (q - p)$$

From the above, one has for the affine space that $p - p = \mathbf{o}$ and $q - p = -(p - q)$, for all $p, q \in P$.

One can take the sum of vectors, according to the structure of the vector space, but one cannot take the sum of points, only the difference between two points.

A key point is that there is no notion of **origin** in the affine space. There is no special or significant point in the set *P*, unlike with the vector space, where there is a special zero element, **o**, which has its own axiom (see axiom 2 in §1.2.1 above).

Suppose now that the associated vector space is a Euclidean vector space, i.e. an inner product space. Define the **distance** between two points through the inner product associated with *V*,

$$d(p, q) = \|q - p\| = \sqrt{\langle q - p, q - p \rangle} \tag{1.2.13}$$

It can be shown that this mapping $d : P \times P \to R$ is a metric, i.e. it satisfies the metric properties, and thus *P* is a metric space (although it is not a vector space). In this case, *P* is referred to as **Euclidean point space**, **Euclidean affine space** or, simply, **Euclidean space**.

Whereas in Euclidean vector space there is a zero element, in Euclidean point space there is none – apart from that, the two spaces are the same and, apart from certain special cases, one does not need to distinguish between them.

Note: one can generalise the simple affine space into a vector space by choosing some fixed $o \in P$ to be an origin. In that case, $\mathbf{v} \equiv \mathbf{v}(p, o) = p - o$ is called the **position vector** of *p* relative to *o*. Then one can define the sum of two points through $p + q = o + (\mathbf{v} + \mathbf{w})$, where $\mathbf{v} = p - o$, $\mathbf{w} = q - o$.[3]

[3] One also has to define a scaling, e.g. $\alpha p \equiv 0 + \alpha \mathbf{v}$, where α is in the associated field (of real numbers).