

3. VECTOR SPACES

Currently when you speak of vectors, you usually mean *coordinate vectors* represented either as a row vector in some \mathbb{R}^n or as a column vector in some $\mathbb{R}_{\text{col}}^n$. But vectors exist without reference to coordinate systems. Wherever you are at the moment, look around you and choose some point near you and label it P , then pick a second point and label it Q . Then \overrightarrow{PQ} is a vector. If you want to treat P as the origin then \overrightarrow{PQ} is the position vector of Q . Or you might think of \overrightarrow{PQ} as a movement and any parallel movement, with the same length and direction, equals the vector \overrightarrow{PQ} . Importantly though \overrightarrow{PQ} has no coordinates, or at least doesn't until you make a choice of origin and axes. This is going to be an important aspect of the Linear Algebra I and II courses, namely choosing coordinates sensibly. This will also be an important aspect of the Geometry and Dynamics courses – in Geometry the change between two coordinate systems will need to be an isometry so that the lengths, areas, angles are measured to be the same; in Dynamics an inertial frame would be necessary for Newton's laws to hold and otherwise so-called 'fictitious forces' will arise.

But the vector spaces we will introduce are not just geometrical vectors like these coordinate or coordinateless vectors. A vector space's elements might contain functions, sequences, matrices, equations or, of course, vectors. Importantly, these more abstract vector spaces do have the same algebraic operations in common with the vectors familiar to you: namely, *addition* and *scalar multiplication*.

3.1 What is a vector space?

A real vector space is a non-empty set with operations of addition and scalar multiplication. Formally this means:

Definition 59 A *real vector space* is a non-empty set V together with a binary operation $V \times V \rightarrow V$ given by $(u, v) \mapsto u + v$ (called **addition**) and a map $\mathbb{R} \times V \rightarrow V$ given by $(\lambda, v) \mapsto \lambda v$ (called **scalar multiplication**) that satisfy the *vector space axioms*

- $u + v = v + u$ for all $u, v \in V$ (*addition is commutative*);
- $u + (v + w) = (u + v) + w$ for all $u, v, w \in V$ (*addition is associative*);
- there is $0_V \in V$ such that $v + 0_V = v = 0_V + v$ for all $v \in V$ (*existence of additive identity*);
- for all $v \in V$ there exists $w \in V$ such that $v + w = 0_V = w + v$ (*existence of additive inverses*);

- $\lambda(u + v) = \lambda u + \lambda v$ for all $u, v \in V$, $\lambda \in \mathbb{R}$ (**distributivity** of scalar multiplication over vector addition);
- $(\lambda + \mu)v = \lambda v + \mu v$ for all $v \in V$, $\lambda, \mu \in \mathbb{R}$ (**distributivity** of scalar multiplication over field addition);
- $(\lambda\mu)v = \lambda(\mu v)$ for all $v \in V$, $\lambda, \mu \in \mathbb{R}$ (scalar multiplication interacts well with field multiplication);
- $1v = v$ for all $v \in V$ (identity for scalar multiplication).

\mathbb{R} is referred to as the **field of scalars** or **base field**. Elements of V are called **vectors** and elements of \mathbb{R} are called **scalars**.

Remark 60 There are a lot of axioms on the above list, but the most important in practice are those requiring:

- V has a zero vector 0_V .
- V is closed under addition.
- V is closed under scalar multiplication.

If these three axioms hold, and addition and scalar multiplication are defined naturally, then usually the remaining axioms will follow as a matter of routine checks.

The subsets of \mathbb{R}^3 that are real vector spaces are the origin, lines through the origin, planes through the origin and all of \mathbb{R}^3 . It's perhaps not surprising then that another term for a vector space is a 'linear space'.

Example 61 We write \mathbb{R}^n for the set of n -tuples (v_1, \dots, v_n) with $v_1, \dots, v_n \in \mathbb{R}$. Then \mathbb{R}^n is a real vector space under componentwise addition and scalar multiplication:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \quad (3.1)$$

$$\text{and } \lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n). \quad (3.2)$$

These satisfy the vector space axioms. The zero vector is $(0, 0, \dots, 0)$ and the additive inverse of (v_1, \dots, v_n) is $(-v_1, \dots, -v_n)$.

We think of \mathbb{R}^2 as the Cartesian plane, and \mathbb{R}^3 as three-dimensional space. We can also consider $n = 1$: \mathbb{R}^1 is a real vector space, which we think of as the real line. We tend to write it simply as \mathbb{R} .

Notation 62 I will often denote a single coordinate vector (v_1, v_2, \dots, v_n) as \mathbf{v} . I will use this bold notation for coordinate vectors, but vectors, as elements of a vector space, will not be written in bold.

Example 63 The field \mathbb{C} is a real vector space, it is essentially the same as \mathbb{R}^2 as a vector space. (The technical term for 'essentially the same' is 'isomorphic'. More on this later.)

Example 64 For $m, n \geq 1$, the set $\mathcal{M}_{m \times n}(\mathbb{R})$ is a real vector space as previously stated in Remark 20.

Example 65 Let $V = \mathbb{R}^{\mathbb{R}} = \{f: \mathbb{R} \rightarrow \mathbb{R}\}$, the space of all real-valued functions on \mathbb{R} , with addition and scalar multiplication defined pointwise. That is,

$$(f + g)(r) := f(r) + g(r), \quad (\alpha f)(r) := \alpha f(r),$$

for $f, g \in V$ and $\alpha, r \in \mathbb{R}$.

Example 66 Let $V = \{f: \mathbb{R} \rightarrow \mathbb{R}, f \text{ is differentiable}\}$, the space of all differentiable real-valued functions on \mathbb{R} , with addition and scalar multiplication defined pointwise. This is a vector space as, in particular,

$$(f + g)' = f' + g', \quad (\alpha f)' = \alpha f',$$

for $f, g \in V$ and $\alpha \in \mathbb{R}$, so that the sum of two differentiable functions and the scalar multiple of a differentiable function is differentiable. It's these facts that mean V is closed under addition and scalar multiplication.

Example 67 Let $V = \{f: \mathbb{R} \rightarrow \mathbb{R}, f'' = f\}$. This is a vector space mainly as $f'' = f$ is a **linear** differential equation. That is, if f and g are solutions then so is $\alpha f + \beta g$ where α, β are real scalars. The general solution can be written as

$$f(x) = A \cosh x + B \sinh x$$

or as

$$f(x) = Ae^x + Be^{-x}.$$

In expressing the general vector in this way note that in each case we are ‘coordinatizing the space’ and identifying V with \mathbb{R}^2 . But note that the coordinate vector $(1, 0)$ corresponds to different vectors as we are using different choices of coordinates; it corresponds to the vector $\cosh x$ in the first case and to e^x in the second.

Example 68 Let $V = \mathbb{R}^{\mathbb{N}} = \{(x_0, x_1, x_2, \dots) : x_i \in \mathbb{R}\}$. This is the space of all real sequences with addition and scalar multiplication defined componentwise.

Other important sequence spaces are

$$\begin{aligned} l^\infty &= \{(x_n)_{n=0}^\infty : (x_n) \text{ is bounded}\}. \\ c &= \{(x_n)_{n=0}^\infty : (x_n) \text{ converges}\}. \\ c_0 &= \{(x_n)_{n=0}^\infty : (x_n) \text{ converges to } 0\}. \end{aligned}$$

As an exercise, what theorems of analysis (concerning convergence) need to hold for these last three examples all to be vector spaces?

Our main focus in this course will be real vector spaces. However, vector spaces can be defined over any field, as can simultaneous equations be considered over any field. Formally a vector space V over a field \mathbb{F} is a non-zero set V with addition $V \times V \rightarrow V$ and scalar multiplication $\mathbb{F} \times V \rightarrow V$ satisfying the vector space axioms in Definition 59. Common examples of other fields that we will encounter are:

- \mathbb{C} – the field of complex numbers.
- \mathbb{Q} – the field of rational numbers.
- \mathbb{Z}_p – the field of integers modulo a prime number p .

The theory of vector spaces applies equally well for all fields. There can be some differences worth noting though depending on the choice of field.

- A non-zero real vector space is an infinite set. This need not be the case over a finite field like \mathbb{Z}_p .
- When we consider \mathbb{C} as a vector space over \mathbb{R} , then every z can be uniquely written as $x1 + yi$ for two real scalars x and y . But when \mathbb{C} is considered as a vector space over \mathbb{C} , then every z can be uniquely written as $z1$ for a single complex scalar z . (In due course we will appreciate that \mathbb{C} is a 2-dimensional real vector space and a 1-dimensional complex vector space.)

Lemma 69 *Let V be a vector space over \mathbb{F} . Then there is a unique additive identity element 0_V .*

Proof. Suppose that 0 and $0'$ are two elements that have the properties of 0_V . Then

$$\begin{aligned} 0 &= 0 + 0' \quad [\text{as } 0' \text{ is a zero vector}] \\ &= 0' \quad [\text{as } 0 \text{ is a zero vector}] \end{aligned}$$

and so $0 = 0'$, thus showing 0_V to be unique. ■

Remark 70 *Where it will not be ambiguous, we often write 0 for 0_V .*

Lemma 71 *Let V be a vector space over \mathbb{F} . Take $v \in V$. Then there is a unique additive inverse for v . That is, if there are $w_1, w_2 \in V$ with $v + w_1 = 0_V = w_1 + v$ and $v + w_2 = 0_V = w_2 + v$, then $w_1 = w_2$.*

Proof. With the notation introduced above we have

$$\begin{aligned} w_2 &= 0 + w_2 \quad [\text{as } 0 \text{ is a zero vector}] \\ &= (w_1 + v) + w_2 \quad [\text{by a hypothesis}] \\ &= w_1 + (v + w_2) \quad [\text{by associativity}] \\ &= w_1 + 0 \quad [\text{by a hypothesis}] \\ &= w_1. \quad [\text{as } 0 \text{ is a zero vector}] \end{aligned}$$

■

Remark 72 *Using the notation of Lemma 71, we write $-v$ for the unique additive inverse of v .*

Proposition 73 Let V be a vector space over a field \mathbb{F} . Take $v \in V$, $\lambda \in \mathbb{F}$. Then

- (a) $\lambda 0_V = 0_V$;
- (b) $0v = 0_V$;
- (c) $(-\lambda)v = -(\lambda v) = \lambda(-v)$;
- (d) if $\lambda v = 0_V$ then $\lambda = 0$ or $v = 0_V$.
- (e) $-v = (-1)v$.

Proof. (a) We have

$$\begin{aligned}\lambda 0_V &= \lambda(0_V + 0_V) && [\text{definition of additive identity}] \\ &= \lambda 0_V + \lambda 0_V && [\text{distributivity of scalar } \cdot \text{ over vector } +].\end{aligned}$$

Adding $-(\lambda 0_V)$ to both sides, we have

$$0_V = \lambda 0_V.$$

(b) Exercise (hint: in \mathbb{F} we have $0 + 0 = 0$).

(c) We have

$$\begin{aligned}\lambda v + \lambda(-v) &= \lambda(v + (-v)) && [\text{distributivity of scalar } \cdot \text{ over vector } +] \\ &= \lambda 0_V && [\text{definition of additive inverse}] \\ &= 0_V && [\text{by (b)}].\end{aligned}$$

So $\lambda(-v)$ is the additive inverse of λv (by uniqueness), so $\lambda(-v) = -(\lambda v)$.

Similarly, we see that $\lambda v + (-\lambda)v = 0_V$ and so $(-\lambda)v = -(\lambda v)$.

(d) Suppose that $\lambda v = 0_V$, and that $\lambda \neq 0$. Then λ^{-1} exists in \mathbb{F} , and

$$\lambda^{-1}(\lambda v) = \lambda^{-1}0_V = 0_V \quad [\text{by (a)}].$$

So

$$(\lambda^{-1}\lambda)v = 0_V \quad [\text{scalar } \cdot \text{ interacts well with field } \cdot],$$

showing

$$v = 1v = 0_V \quad [\text{identity for scalar multiplication}].$$

(e) Note that

$$\begin{aligned}v + (-1)v &= 1v + (-1)v && [\text{by a vector space axiom}] \\ &= (1 + (-1))v && [\text{distributivity}] \\ &= 0v && [\text{definition of additive inverse in field}] \\ &= 0_V. && [\text{by (a)}]\end{aligned}$$

Hence by the uniqueness of the additive inverse $(-1)v = -v$. ■

3.2 Subspaces

Whenever we have a mathematical object with some structure, we want to consider subsets that also have that same structure.

Definition 74 Let V be a vector space over \mathbb{F} . A **subspace** of V is a non-empty subset of V that is closed under addition and scalar multiplication, that is, a subset $U \subseteq V$ such that

- (i) $U \neq \emptyset$ (U is non-empty); (this usually involves showing $0_V \in U$).
- (ii) $u_1 + u_2 \in U$ for all $u_1, u_2 \in U$ (U is closed under addition);
- (iii) $\lambda u \in U$ for all $u \in U, \lambda \in \mathbb{F}$ (U is closed under scalar multiplication).

Note that the operations of addition and scalar multiplication referred to are those of V , not some separate, different operations of U .

Definition 75 The sets $\{0_V\}$ and V are always subspaces of V . The subspace $\{0_V\}$ is sometimes called the **zero subspace** or the **trivial subspace**. Subspaces other than V are called **proper subspaces**.

Proposition 76 (Subspace test) Let V be a vector space over \mathbb{F} , let U be a subset of V . Then U is a subspace if and only if

- (i) $0_V \in U$; and
- (ii) $\lambda u_1 + u_2 \in U$ for all $u_1, u_2 \in U$ and $\lambda \in \mathbb{F}$.

Proof. (\Rightarrow) Assume that U is a subspace of V .

$0_V \in U$: Since U is a subspace, it is non-empty, so there exists $u \in U$. Since U is closed under scalar multiplication, $0u = 0_V \in U$.

$\lambda u_1 + u_2 \in U$ for all $u_1, u_2 \in U$ and all $\lambda \in \mathbb{F}$: Take $u_1, u_2 \in U$, and $\lambda \in \mathbb{F}$. Then $\lambda u_1 \in U$ because U is closed under scalar multiplication, so $\lambda u_1 + u_2 \in U$ because U is also closed under addition.

(\Leftarrow) Assume that $0_V \in U$ and that $\lambda u_1 + u_2 \in U$ for all $u_1, u_2 \in U$ and $\lambda \in \mathbb{F}$.

U is non-empty: we note $0_V \in U$.

U is closed under addition: for $u_1, u_2 \in U$ have $u_1 + u_2 = 1u_1 + u_2 \in U$.

U is closed under scalar multiplication: for $u \in U$ and $\lambda \in \mathbb{F}$, have $\lambda u = \lambda u + 0_V \in U$.

So U is a subspace of V . ■

Notation 77 If U is a subspace of the vector space V , then we write $U \leqslant V$.

Proposition 78 Let V be a vector space over \mathbb{F} , and let $U \leqslant V$. Then

- (a) U is a vector space over \mathbb{F} . In fact, the only subsets of V that are vector spaces over \mathbb{F} are the subspaces;
- (b) if $W \leqslant U$ then $W \leqslant V$ (“a subspace of a subspace is a subspace”).

Proof. (a) We need to check the vector space axioms, but first we need to check that we have legitimate operations. Since U is closed under addition, the operation $+$ restricted to U gives

a map $U \times U \rightarrow U$. Likewise since U is closed under scalar multiplication, that operation restricted to U gives a map $\mathbb{F} \times U \rightarrow U$.

Now for the axioms.

Commutativity and associativity of addition are inherited from V .

There is an additive identity (by the subspace test).

There are additive inverses: if $u \in U$ then multiplying by $-1 \in \mathbb{F}$ shows that $-u = (-1)u \in U$.

The remaining four properties are all inherited from V . That is, they apply to general vectors of V and vectors in U are vectors in V .

(b) This is immediate from the definition of a subspace. ■

Proposition 79 *Let V be a vector space. Take $U, W \leqslant V$. Then $U + W \leqslant V$ and $U \cap W \leqslant V$, where*

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

Indeed $U + W$ is the smallest subspace of V which contains U and W and $U \cap W$ is the largest subspace of V which is contained in both U and W .

Proof. (a) As $U \leqslant V$ and $W \leqslant V$ then $0_V \in U$ and $0_V \in W$ so that $0_V = 0_V + 0_V \in U + W$.

Say that $v_1, v_2 \in U + W$ and $\lambda \in \mathbb{F}$. By definition there exist $u_1, u_2 \in U$ and $w_1, w_2 \in W$ such that

$$v_1 = u_1 + w_1, \quad v_2 = u_2 + w_2,$$

and then

$$\lambda v_1 + v_2 = \lambda(u_1 + w_1) + u_2 + w_2 = (\lambda u_1 + u_2) + (\lambda w_1 + w_2) \in U + W$$

as $\lambda u_1 + u_2 \in U$ and $\lambda w_1 + w_2 \in W$ because $U \leqslant V$ and $W \leqslant V$.

(b) The statements concerning the intersection are left as exercises. ■

3.3 Further examples

Example 80 Consider a system of **homogeneous** linear equations with real coefficients a_{ij} :

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots && \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

(We say this is homogeneous because all the real numbers on the right are 0.)

Let V be the set of real solutions of this linear system. Then V is a real vector space. This becomes more apparent if we write the equations in matrix form. We see the system corresponds to $A\mathbf{x} = \mathbf{0}$, where $A = (a_{ij}) \in \mathcal{M}_{m \times n}(\mathbb{R})$, $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ is an $n \times 1$ column

vector of variables, and $\mathbf{0}$ is shorthand for $0_{n \times 1}$. Each element of V can be thought of as an $n \times 1$ column vector of real numbers.

To show that V is a vector space, we show that it is a subspace of $\mathbb{R}_{\text{col}}^n$.

Clearly V is non-empty, because $\mathbf{0} \in V$.

For $\mathbf{v}_1, \mathbf{v}_2 \in V$, we have $A\mathbf{v}_1 = \mathbf{0}$ and $A\mathbf{v}_2 = \mathbf{0}$, so $A(\mathbf{v}_1 + \mathbf{v}_2) = A\mathbf{v}_1 + A\mathbf{v}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$, so $\mathbf{v}_1 + \mathbf{v}_2 \in V$. So V is closed under addition.

For $\mathbf{v} \in V$ and $\lambda \in \mathbb{F}$, we have $A(\lambda\mathbf{v}) = \lambda(A\mathbf{v}) = \lambda\mathbf{0} = \mathbf{0}$, so $\lambda\mathbf{v} \in V$. So V is closed under scalar multiplication.

So $V \leq \mathbb{R}_{\text{col}}^n$, and so V is a vector space.

Example 81 The set $\mathbb{R}[x]$ of all real polynomials in a variable x is a real vector space. We will show that it is a subspace of $\mathbb{R}^\mathbb{R}$. Addition and scalar multiplication are defined by

$$\left(\sum a_n x^n \right) + \left(\sum b_n x^n \right) = \sum (a_n + b_n) x^n, \quad \lambda \left(\sum a_n x^n \right) = \sum (\lambda a_n) x^n.$$

As the sums are finite, then the addition and scalar multiple are also finite and hence polynomials. Finally the zero function is a polynomial.

Example 82 Let n be a non-negative integer. The set of polynomials $c_n x^n + \dots + c_1 x + c_0$ with $c_0, c_1, \dots, c_n \in \mathbb{R}$ (that is, real polynomials with degree $\leq n$) is a real vector space, and a subspace of $\mathbb{R}[x]$.

Example 83 Let X be a set. Define $\mathbb{R}^X := \{\text{functions } f \text{ with } f: X \rightarrow \mathbb{R}\}$, the set of real-valued functions on X . This is a real vector space with operations of pointwise addition and pointwise multiplication by a real number: for $x \in X$, we define

$$(f + g)(x) = f(x) + g(x) \quad \text{and} \quad (\lambda f)(x) = \lambda f(x).$$

Example 84 We can study the solutions of a homogeneous linear second-order differential equation. These are twice-differentiable real functions y that satisfy an equation of the form

$$y'' + a(x)y' + b(x)y = 0.$$

This equation is **linear** because y and its derivatives occur only to the first power and are not multiplied together. And it is **homogeneous** because of the 0 on the right-hand side. Such equations are important in many applications of mathematics.

The set S of solutions of this homogeneous linear second-order differential equation is a vector space, a subspace of $\mathbb{R}^\mathbb{R}$. Note S is clearly non-empty (the 0 function satisfies the differential equation), and if $w = u + \lambda v$ where $u, v \in S$ and $\lambda \in \mathbb{R}$, then

$$\begin{aligned} w'' + a(x)w' + b(x)w &= (u'' + \lambda v'') + a(x)(u' + \lambda v') + b(x)(u + \lambda v) \\ &= (u'' + a(x)u' + b(x)u) + \lambda(v'' + a(x)v' + b(x)v) \\ &= 0, \end{aligned}$$

so $w \in S$. So, by the Subspace Test, $S \leq \mathbb{R}^\mathbb{R}$.

This generalises to homogeneous linear differential equations of any order.

Example 85 What are the subspaces of \mathbb{R} ?

Let $V = \mathbb{R}$, let U be a non-trivial subspace of V . Then there exists $u \in U$ with $u \neq 0$. Take $x \in \mathbb{R}$. Let $\lambda = \frac{x}{u}$. Then $x = \lambda u \in U$, because U is closed under scalar multiplication. So $U = V$.

So the only subspaces of \mathbb{R} are $\{0\}$ and \mathbb{R} .

Example 86 What are the subspaces of \mathbb{R}^2 ?

Let $V = \mathbb{R}^2$, let U be a non-trivial subspace of V . Then there exists $\mathbf{u} \in U$ with $\mathbf{u} \neq (0, 0)$, say $\mathbf{u} = (a, b)$. We have $\langle \mathbf{u} \rangle = \{\lambda \mathbf{u} : \lambda \in \mathbb{R}\} \subseteq U$. (Such ‘spans’ will more generally be defined in the next chapter.)

Case 1: $\langle \mathbf{u} \rangle = U$.

If $a \neq 0$, then let $m = \frac{b}{a}$. Then $\langle \mathbf{u} \rangle = \{(x, y) \in \mathbb{R}^2 \mid y = mx\}$.

If $a = 0$, then $\langle \mathbf{u} \rangle = \{(0, y) \mid y \in \mathbb{R}\}$.

So if U is, geometrically, a line in \mathbb{R}^2 through the origin, and every such line in \mathbb{R}^2 through the origin corresponds to a subspace.

Case 2: $\langle \mathbf{u} \rangle \neq U$.

Then there is some $\mathbf{v} = (c, d) \in U \setminus \langle \mathbf{u} \rangle$.

Consider the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Applying any sequence of EROs to this matrix gives a matrix whose rows are in U . The matrix must have RRE form $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. [Why?] So U contains the vectors $(1, 0)$ and $(0, 1)$, and hence $U = \mathbb{R}^2$. [Why?]

So the subspaces of \mathbb{R}^2 are $\{0\}$, lines in \mathbb{R}^2 through the origin and \mathbb{R}^2 .