Cooper Johnston

Linear Algebra

Early Draft
December 2024

Contents

Chapter 0

Sets and Proofs

0.1 Sets

We will begin by exploring the concept of a set through what is sometimes called intuitive or naive set theory. There exist more rigorous approaches, axiomatic set theories, but we will not be looking at these; our intuitive treatment of sets will suffice for the purposes of this course.

Definition 0.1.1. A set is a well-defined collection of objects. By "well-defined" we mean that for any set S, any object is either definitely in S or definitely not in S.

An object that is in a set is called an element of that set. We write $x \in S$ to denote that x is an element of the set S.

The set that does not contain any elements is called the empty set, denoted $\varnothing$.

The number of elements in a set is called the cardinality of that set. We write $|S|$ to denote the cardinality of the set S.

One way to describe a set is by listing its elements. For example, we can define A to be the set containing the numbers 3, 6, 9, and 12, denoted by

$$A = \{3, 6, 9, 12\}.$$

Another way is to give a defining property of its elements. For example, A is the set of the first four positive multiples of three, or more mathematically, A is the set of all elements $3n$ such that $n = 1, 2, 3, 4$, denoted by

$$A = \{3n \mid n = 1, 2, 3, 4\}.$$

The latter notation is often called set-builder notation.

We will denote certain special sets of numbers as follows:

- $\mathbb{N}$ is the set of natural numbers, which we will take to start at 1;

- $\mathbb{Z}$ is the set of integers;
- $\mathbb{Q}$ is the set of rational numbers;
- $\mathbb{Q}^+$ is the set of positive rational numbers;
- $\mathbb{R}$ is the set of real numbers;
- $\mathbb{R}^+$ is the set of positive real numbers; and
- $\mathbb{C}$ is the set of complex numbers.

Example 0.1.2. The set of even numbers is the set of all numbers $2n$ where n is an integer, i.e. $\{2n \mid n \in \mathbb{Z}\}$.

Example 0.1.3. The set of *ratio*nal numbers $\mathbb{Q}$ is the set of all numbers that can be expressed as a *ratio* p/q where p and q are integers and $q \neq 0$, i.e.

$$\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0 \right\}.$$

Definition 0.1.4. Let A and B be two sets. B is called a subset of A, denoted $B \subseteq A$, if every element in B is also an element in A, i.e. for every $b \in B$, we have $b \in A$.

B is called a proper subset of A, denoted $B \subset A$, if $B \subseteq A$ and $B \neq A$.

Definition 0.1.5. Let $A_1, A_2, \ldots, A_n$ be non-empty sets. The set

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \ldots, a_n) \mid a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n\}$$

is called the Cartesian product of $A_1, A_2, \ldots, A_n$.

The Cartesian product of a set with itself can be denoted by

$$\underbrace{A \times A \times \cdots \times A}_{n \text{ times}} = A^n.$$

Example 0.1.6. The set of ordered pairs of real numbers, i.e.

$$\{(x, y) \mid x, y \in \mathbb{R}\},$$

can be written as $\mathbb{R} \times \mathbb{R}$ or $\mathbb{R}^2$.

Example 0.1.7. Let $A = \{1, 2\}$ and $B = \{1, 2, 3\}$. Then,

$$\begin{aligned} A \times B &= \{(a, b) \mid a \in A, b \in B\} \\ &= \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3)\}. \end{aligned}$$

Definition 0.1.8. Let A and B be two sets. The set

$$A \cup B = \{c \mid c \in A \text{ or } c \in B\},$$

i.e. the set containing all elements of A as well as all the elements of B, is called the union of A and B. The set

$$A \cap B = \{c \mid c \in A \text{ and } c \in B\},$$

i.e. the set containing all the elements that are in both A and B at the same time, is called the intersect of A and B.

0.2 Mappings

Definition 0.2.1. Let A and B be two non-empty sets, and let $\mathcal{R} \subseteq A \times B$. The set $\mathcal{R}$ is called a relation between A and B. For an ordered pair $(a, b) \in \mathcal{R}$, we say that $\mathcal{R}$ relates a to b.

Definition 0.2.2. Let A and B be two non-empty sets. A relation f between A and B is called a mapping or a function if for every $a \in A$, there exists exactly one $b \in B$ such that f relates a to b. The set A is called the domain of f, and B is called the codomain of f.

We write $f : A \to B$ to denote that f is a mapping with domain A and codomain B; that is, f is a mapping from A to B.

We write $f(a) = b$ or $a \mapsto b$ to denote that f relates a to b; that is, f maps a to b.

Definition 0.2.3. Let A and B be two sets and let $\odot : B \times A \to A$. The mapping $\odot$ is called a binary operation. If $A = B$, i.e. we have $\odot : A \times A \to A$, we say $\odot$ is a binary operation on A.

We write $x \odot y = z$ to denote that $\odot$ maps (x, y) to z.

0.3 Propositional logic

0.4 Proofs

Chapter 1

Vectors

1.1 Fields

Definition 1.1.1. Let F be a set and let $+$ and $\cdot$ be two binary operations on F. The set F, together with the operations $+$ and $\cdot$, is called a field if all of the following axioms are satisfied:

1. $+$ and $\cdot$ are associative, i.e. for all $a, b, c \in F$, we have
$$(a + b) + c = a + (b + c) \quad \text{and} \quad (a \cdot b) \cdot c = a \cdot (b \cdot c).$$

2. $+$ and $\cdot$ are commutative, i.e. for all $a, b \in F$, we have
$$a + b = b + a \quad \text{and} \quad a \cdot b = b \cdot a.$$

3. There exists an element $0_F \in F$, called the additive identity, such that for all $a \in F$, we have
$$a + 0_F = a.$$

4. There exists an element $1_F \in F$, called the multiplicative identity, such that for all $a \in F$, we have
$$a \cdot 1_F = a.$$

5. For every $a \in F$, there exists an element $-a \in F$, called the additive inverse of a, such that
$$a + (-a) = 0_F.$$

6. For every $a \in F$ other than 0_F, there exists an element $a^{-1} \in F$, called the multiplicative inverse of a, such that
$$a \cdot a^{-1} = 1_F.$$

7. $\cdot$ is distributive over $+$, i.e. for all $a, b, c \in F$, we have

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c).$$

The operation $+$ is called addition, and the operation $\cdot$ is called multiplication. For multiplication, we will often use the notation $a \cdot b = ab$.

Note that since the operations $+$ and $\cdot$ are defined as operations on F, when we add or multiply two elements of F, the result must also be an element of F, i.e. for all $a, b \in F$, we must have

$$a + b \in F \quad \text{and} \quad a \cdot b \in F.$$

This is called closure, and we say that a field must be closed under addition and multiplication.

Let us examine some examples of what is and isn't a field.

Example 1.1.2. Show that the set of real numbers $\mathbb{R}$, together with standard addition and multiplication, is a field.

Solution. Much of this proof will involve results that we already know and don't need to show in detail. First, note that since the sum and product of two real numbers is always a real number, $\mathbb{R}$ is closed under standard addition and multiplication. Now we can check the field axioms:

1. Associativity: We already know that standard addition and multiplication are associative.
2. Commutativity: We also know that standard addition and multiplication are commutative.
3. Existence of additive identity: The additive identity is the number 0.
4. Existence of multiplicative identity: The multiplicative identity is the number 1.
5. Existence of additive inverse: For any $x \in \mathbb{R}$, the additive inverse is the number $-x$.
6. Existence of multiplicative inverse: For any $x \in \mathbb{R}$ other than 0, the multiplicative inverse is the number $1/x$.
7. Distributivity: We already know that standard multiplication is distributive over standard addition.

Hence, $\mathbb{R}$ with standard addition and multiplication is a field. □

Example 1.1.3. Show that the set of integers $\mathbb{Z}$, together with standard addition and multiplication, is not a field.

Solution. We need only find one axiom that does not hold. The multiplicative identity is the number 1. Consider the number $2 \in \mathbb{Z}$. There does not exist a number $n \in \mathbb{Z}$ such that $2n = 1$; that is, 2 does not have a multiplicative inverse in $\mathbb{Z}$. Hence, $\mathbb{Z}$ is not a field. □

We will typically denote a field simply by its set, and when working with fields of numbers, we will from now on assume standard addition and multiplication unless otherwise specified. For example, the field "$\mathbb{R}$" is assumed to mean $\mathbb{R}$ together with standard addition and multiplication.

Exercises

Exercise 1.1. Consider the set of complex numbers $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$ where i is the imaginary number, defined such that $i^2 = -1$. We naturally have the operations $+$ and $\cdot$ defined such that for all $(a + bi), (c + di) \in \mathbb{C}$, we have

$$(a + bi) + (c + di) = a + c + bi + di = (a + c) + (b + d)i$$

and

$$\begin{aligned}(a + bi) \cdot (c + di) &= ac + a(di) + (bi)c + (bi)(di) = ac + adi + bci + bdi^2 \\ &= ac + adi + bci - bd = (ac - bd) + (ad + bc)i.\end{aligned}$$

Show that $\mathbb{C}$ with these operations is a field.

1.2 Vector spaces

Definition 1.2.1. Let F be a field and V be a set, and let $\odot : F \times V \to V$ and $\oplus : V \times V \to V$ be two binary operations. V, together with these operations, is called a vector space over F if all of the following axioms are satisfied:

1. $\oplus$ is associative, i.e. for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, we have
$$(\mathbf{u} \oplus \mathbf{v}) \oplus \mathbf{w} = \mathbf{u} \oplus (\mathbf{v} \oplus \mathbf{w}).$$
2. $\oplus$ is commutative, i.e. for all $\mathbf{u}, \mathbf{v} \in V$, we have
$$\mathbf{u} \oplus \mathbf{v} = \mathbf{v} \oplus \mathbf{u}.$$
3. There exists an element $\mathbf{0} \in V$, called the zero vector, such that for all $\mathbf{v} \in V$, we have
$$\mathbf{v} \oplus \mathbf{0} = \mathbf{v}.$$
4. For every $\mathbf{v} \in V$, there exists an element $-\mathbf{v} \in V$, called the additive inverse of $\mathbf{v}$, such that
$$\mathbf{v} \oplus (-\mathbf{v}) = \mathbf{0}.$$

5. For all $a, b \in F$ and $\mathbf{v} \in V$, we have
$$a \odot (b \odot \mathbf{v}) = (ab) \odot \mathbf{v}.$$

6. For every $\mathbf{v} \in V$, we have
$$1_F \odot \mathbf{v} = \mathbf{v}$$
where 1_F is the multiplicative identity of F.

7. $\odot$ is distributive over $\oplus$, i.e. for all $a \in F$ and $\mathbf{u}, \mathbf{v} \in V$, we have
$$a \odot (\mathbf{u} \oplus \mathbf{v}) = (a \odot \mathbf{u}) \oplus (a \odot \mathbf{v}).$$

8. For all $a, b \in F$ and $\mathbf{v} \in V$, we have
$$(a + b) \odot \mathbf{v} = (a \odot \mathbf{v}) \oplus (b \odot \mathbf{v}).$$

The elements of F are called scalars and the elements of V are called vectors. The operation $\odot$ is called scalar multiplication and the operation $\oplus$ is called vector addition.

Note that the zero vector is the same as the additive identity under vector addition. Also note that by the way we have defined $\odot$ and $\oplus$, a vector space must be closed under scalar multiplication and vector addition, just as a field must be closed under field addition and multiplication.

Let us now examine some examples of what is and isn't a vector space.

Example 1.2.2. Show that $\mathbb{R}^2$, together with the operations $\oplus$ and $\odot$ defined such that for all $c \in \mathbb{R}$ and $(x, y), (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$, we have
$$(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \quad \text{and} \quad c \odot (x, y) = (cx, cy),$$
is a vector space over $\mathbb{R}$.

Solution. Let $(x, y), (x_1, y_1), (x_2, y_2), (x_3, y_3) \in \mathbb{R}^2$ and let $a, b \in \mathbb{R}$. First we must verify that $\mathbb{R}^2$ is closed under the given operations. Since $(x_1 + x_2), (y_1 + y_2) \in \mathbb{R}$, we see that
$$(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \in \mathbb{R}^2,$$
and similarly, since $cx, cy \in \mathbb{R}$, we also see that
$$c \odot (x, y) = (cx, cy) \in \mathbb{R}^2.$$
Thus, $\mathbb{R}^2$ is closed under $\oplus$ and $\odot$. Now we can check the vector space axioms:

1. Associativity of $\oplus$:

$$\begin{aligned}((x_1, y_1) \oplus (x_2, y_2)) \oplus (x_3, y_3) &= (x_1 + x_2, y_1 + y_2) \oplus (x_3, y_3) \\ &= (x_1 + x_2 + x_3, y_1 + y_2 + y_3) \\ &= (x_1, y_1) \oplus (x_2 + x_3, y_2 + y_3) \\ &= (x_1, y_1) \oplus ((x_2, y_2) \oplus (x_3, y_3)).\end{aligned}$$

2. Commutativity of $\oplus$:

$$(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2, y_1 + y_2) = (x_2 + x_1, y_2 + y_1) = (x_2, y_2) \oplus (x_1, y_1).$$

3. Existence of zero vector: Consider $\mathbf{0} = (0, 0)$. We see

$$(x, y) \oplus \mathbf{0} = (x, y) \oplus (0, 0) = (x + 0, y + 0) = (x, y),$$

so $\mathbf{0}$, as we have defined it, is the zero vector.

4. Existence of additive inverse: Consider $-(x, y) = (-x, -y)$. We see

$$(x, y) \oplus (-(x, y)) = (x, y) \oplus (-x, -y) = (x - x, y - y) = (0, 0) = \mathbf{0},$$

so $-(x, y)$, as we have defined it, is the additive inverse of (x, y).

5. Compatibility of $\odot$ with field multiplication:

$$a \odot (b \odot (x, y)) = a \odot (bx, by) = (abx, aby) = (ab) \odot (x, y).$$

6. Scalar multiplicative identity: Recall that the number 1 is the multiplicative identity of $\mathbb{R}$. We see

$$1 \odot (x, y) = (1x, 1y) = (x, y).$$

7. Distributivity of $\odot$ over $\oplus$:

$$\begin{aligned}a \odot ((x_1, y_1) \oplus (x_2, y_2)) &= a \odot (x_1 + x_2, y_1 + y_2) = (a(x_1 + x_2), a(y_1 + y_2)) \\ &= (ax_1 + ax_2, ay_1 + ay_2) = (ax_1, ay_1) \oplus (ax_2, ay_2) \\ &= (a \odot (x_1, y_1)) \oplus (a \odot (x_2, y_2)).\end{aligned}$$

8. Distributivity of $\odot$ over field addition:

$$\begin{aligned}(a + b) \odot (x, y) &= ((a + b)x, (a + b)y) = (ax + bx, ay + by) \\ &= (ax, ay) \oplus (bx, by) = (a \odot (x, y)) \oplus (b \odot (x, y)).\end{aligned}$$

Hence, $\mathbb{R}^2$ with these operations is a vector space over $\mathbb{R}$. □

Example 1.2.3. Let $P_2(\mathbb{R})$ be the set of all polynomials of degree ≤ 2 over the real numbers, i.e.

$$P_2(\mathbb{R}) = \{p : \mathbb{R} \to \mathbb{R},\ p(x) = a_0 + a_1 x + a_2 x^2 \mid a_0, a_1, a_2 \in \mathbb{R}\}.$$

Show that $P_2(\mathbb{R})$, together with the operations $\oplus$ and $\odot$ defined such that for all $c \in \mathbb{R}$ and $p(x) = a_0 + a_1 x + a_2 x^2, q(x) = b_0 + b_1 x + b_2 x^2 \in P_2(\mathbb{R})$, we have

$$p(x) \oplus q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2$$

and

$$c \odot p(x) = ca_0 + ca_1 x + ca_2 x^2,$$

is a vector space over $\mathbb{R}$.

Solution. Since the coefficients of $p(x) \oplus q(x)$ and $c \odot p(x)$ are in $\mathbb{R}$, we see that $p(x) \oplus q(x)$ and $c \odot p(x)$ are in $P_2(\mathbb{R})$. Thus, $P_2(\mathbb{R})$ is closed under addition and scalar multiplication. Now we can check the vector space axioms:

1. Associativity of $\oplus$: Let $r(x) = c_0 + c_1 x + c_2 x^2 \in P_2(\mathbb{R})$. Then,

$$\begin{aligned}(p(x) \oplus q(x)) \oplus r(x) &= ((a_0 + b_0) + (a_1 + b_1)x + (a_2 + b + 2)x^2) \oplus r(x) \\ &= (a_0 + b_0 + c_0) + (a_1 + b_1 + c_1)x + (a_2 + b_2 + c_2)x^2 \\ &= p(x) \oplus ((b_0 + c_0) + (b_1 + c_1)x + (b_2 + c_2)x^2) \\ &= p(x) \oplus (q(x) \oplus r(x)).\end{aligned}$$

2. Commutativity of $\oplus$:

$$\begin{aligned}p(x) \oplus q(x) &= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 \\ &= (b_0 + a_0) + (b_1 + a_1)x + (b_2 + a_2)x^2 = q(x) \oplus p(x).\end{aligned}$$

3. Existence of zero vector: Consider $o(x) = 0 + 0x + 0x^2$. We see

$$p(x) \oplus o(x) = (a_0 + 0) + (a_1 + 0)x + (a_2 + 0)x^2 = a_0 + a_1 x + a_2 x^2 = p(x),$$

so $o(x)$ is the zero vector.

4. Existence of additive inverse: Consider $-p(x) = -a_0 + (-a_1)x + (-a_2)x^2$. We see

$$\begin{aligned}p(x) \oplus (-p(x)) &= (a_0 + (-a_0)) + (a_1 + (-a_1))x + (a_2 + (-a_2))x^2 \\ &= 0 + 0x + 0x^2 = o(x),\end{aligned}$$

so $-p(x)$, as we have defined it, is the additive inverse of $p(x)$.

5. Compatibility of $\odot$ with field multiplication: Let $d \in F$. Then,

$$\begin{aligned}c \odot (d \odot p(x)) &= c \odot (da_0 + da_1 x + da_2 x^2) = cda_0 + cda_1 x + cda_2 x^2 \\ &= (cd) \odot p(x).\end{aligned}$$

6. Scalar multiplicative identity:

$$1 \odot p(x) = 1a_0 + 1a_1x + 1a_2x^2 = a_0 + a_1x + a_2x^2 = p(x).$$

7. Distributivity of $\odot$ over $\oplus$:

$$\begin{aligned} c \odot (p(x) \oplus q(x)) &= c \odot ((a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2) \\ &= c(a_0 + b_0) + c(a_1 + b_1)x + c(a_2 + b_2)x^2 \\ &= (ca_0 + cb_0) + (ca_1 + cb_1)x + (ca_2 + cb_2)x^2 \\ &= (c \odot p(x)) \oplus (c \odot q(x)). \end{aligned}$$

8. Distributivity of $\odot$ over field addition:

$$\begin{aligned} (c + d) \odot p(x) &= (c + d)a_0 + (c + d)a_1x + (c + d)a_2x^2 \\ &= (ca_0 + da_0) + (ca_1 + da_1)x + (ca_2 + da_2)x^2 \\ &= (c \odot p(x)) \odot (d \odot p(x)). \end{aligned}$$

Hence, $P_2(\mathbb{R})$ with these operations is a vector space over $\mathbb{R}$. □

Example 1.2.4. Show that $\mathbb{R}^3$, together with the operations $\oplus$ and $\odot$ defined such that for all $c \in \mathbb{R}$ and $(x, y, z), (\tilde{x}, \tilde{y}, \tilde{z}) \in \mathbb{R}^3$, we have

$$(x, y, z) \oplus (\tilde{x}, \tilde{y}, \tilde{z}) = (x + \tilde{x}, y + \tilde{y}, 0) \quad \text{and} \quad c \odot (x, y, z) = (cx, cy, cz),$$

is not a vector space over $\mathbb{R}$.

Solution. We need only find one vector space axiom that does not hold. Consider $(1, 1, 1) \in \mathbb{R}^3$. We see

$$(1, 1, 1) \oplus (x, y, z) = (x, y, 0) \neq (1, 1, 1)$$

for any $(x, y, z) \in \mathbb{R}^3$, so there is no additive identity in $\mathbb{R}^3$ under $\oplus$. Hence, $\mathbb{R}^3$ with these operations is not a vector space. □

In these examples and in the definition, we used the symbols $\odot$ and $\oplus$ for scalar multiplication and vector addition, respectively, in order to help distinguish these operations from addition and multiplication on the field. From now on, we will use additive notation

$$\mathbf{u} \oplus \mathbf{v} = \mathbf{u} + \mathbf{v}$$

for vector addition and multiplicative notation

$$c \odot \mathbf{v} = c\mathbf{v}$$

for scalar multiplication, but it is still important to remember the distinction between these operations with vectors and the analogous operations on the field. We will also write expressions with vectors assuming that scalar multiplication takes precedence over vector addition, akin to the standard order of operations.

Theorem 1.2.5. *Let V be a vector space over a field F. For all $\mathbf{v} \in V$ and $a \in F$, we have*

1. $0_F\mathbf{v} = \mathbf{0}$;
2. $a\mathbf{0} = \mathbf{0}$;
3. *If $a\mathbf{v} = \mathbf{0}$, then $a = 0_F$ or $\mathbf{v} = \mathbf{0}$; and*
4. $(-1_F)\mathbf{v} = -\mathbf{v}$.

Proof.

1. Since 0_F is the additive identity on the field, we know $c+0_F = c$. In particular, $0_F + 0_F = 0_F$. Thus,

$$\begin{aligned}
0_F\mathbf{v} &= (0_F + 0_F)\mathbf{v} && \\
0_F\mathbf{v} &= 0_F\mathbf{v} + 0_F\mathbf{v} && \text{distributive property} \\
0_F\mathbf{v} + (-(0_F\mathbf{v})) &= 0_F\mathbf{v} + 0_F\mathbf{v} + (-(0_F\mathbf{v})) && \\
\mathbf{0} &= 0_F\mathbf{v} + 0_F\mathbf{v} + (-(0_F\mathbf{v})) && \text{vector additive inverse} \\
\mathbf{0} &= 0_F\mathbf{v} + (0_F\mathbf{v} + (-(0_F\mathbf{v}))) && \text{associative property} \\
\mathbf{0} &= 0_F\mathbf{v} + \mathbf{0} && \text{vector additive inverse} \\
\mathbf{0} &= 0_F\mathbf{v}. && \text{vector additive identity}
\end{aligned}$$

2. Since the zero vector $\mathbf{0}$ is the additive identity under vector addition, we know $\mathbf{v} + \mathbf{0} = \mathbf{v}$. In particular, $\mathbf{0} + \mathbf{0} = \mathbf{0}$. Thus,

$$\begin{aligned}
a\mathbf{0} &= a(\mathbf{0} + \mathbf{0}) && \\
a\mathbf{0} &= a\mathbf{0} + a\mathbf{0} && \text{distributive property} \\
a\mathbf{0} + (-(a\mathbf{0})) &= a\mathbf{0} + a\mathbf{0} + (-(a\mathbf{0})) && \\
\mathbf{0} &= a\mathbf{0} + a\mathbf{0} + (-(a\mathbf{0})) && \text{vector additive inverse} \\
\mathbf{0} &= a\mathbf{0} + (a\mathbf{0} + (-(a\mathbf{0}))) && \text{associative property} \\
\mathbf{0} &= a\mathbf{0} + \mathbf{0} && \text{vector additive inverse} \\
\mathbf{0} &= a\mathbf{0}. && \text{vector additive identity}
\end{aligned}$$

3. Suppose $a\mathbf{v} = \mathbf{0}$. We have two possible cases for a. If $a = 0_F$, then we are already finished with the proof. Otherwise, if $a \neq 0_F$, we can multiply both sides with the multiplicative inverse (on the field) of a:

$$\begin{aligned}
a\mathbf{v} &= \mathbf{0} && \\
a^{-1}(a\mathbf{v}) &= a^{-1}\mathbf{0} && \\
(a^{-1}a)\mathbf{v} &= a^{-1}\mathbf{0} && \text{vector space axiom 5} \\
1_F\mathbf{v} &= a^{-1}\mathbf{0} && \text{field multiplicative inverse} \\
\mathbf{v} &= a^{-1}\mathbf{0} && \text{scalar multiplicative identity} \\
\mathbf{v} &= \mathbf{0}. && \text{part 2 of this theorem}
\end{aligned}$$

4. We will start by using the zero vector:

$$\begin{aligned}
-\mathbf{v} &= -\mathbf{v} + \mathbf{0} \\
&= -\mathbf{v} + 0_F\mathbf{v} && \text{part 1 of this theorem} \\
&= -\mathbf{v} + (1_F + (-1_F))\mathbf{v} && \text{field additive inverse} \\
&= -\mathbf{v} + (1_F\mathbf{v} + (-1_F)\mathbf{v}) && \text{distributive property} \\
&= -\mathbf{v} + (\mathbf{v} + (-1_F)\mathbf{v}) && \text{scalar multiplicative identity} \\
&= (-\mathbf{v} + \mathbf{v}) + (-1_F\mathbf{v}) && \text{associative property} \\
&= \mathbf{0} + (-1_F)\mathbf{v} && \text{vector additive inverse} \\
&= (-1_F)\mathbf{v}. && \text{vector additive identity}
\end{aligned}$$ ■

Note that in the above proof, the distributive property used in parts 1 and 2 comes from vector space axiom 7, while the distributive property used in part 4 comes from axiom 8.

Exercises

Exercise 1.2. Let F be a field and let $P_n(F)$ be the set of all polynomials of degree $\leq n$ over F, i.e.

$$P_n(F) = \left\{ p : F \to F,\ p(z) = \sum_{k=0}^{n} a_k z^k \mid a_0, a_1, \ldots, a_n \in F \right\}.$$

Show that $P_n(F)$, together with addition and scalar multiplication defined such that for all $c \in F$ and $p(z) = \sum_{k=0}^{n} a_k z^k, q(z) = \sum_{k=0}^{n} b_k z^k \in P_n(F)$, we have

$$p(z) + q(z) = \sum_{k=0}^{n} (a_k + b_k) z^k \quad \text{and} \quad cp(z) = \sum_{k=0}^{n} c a_k z^k,$$

is a vector space over F.

Exercise 1.3. Show that the zero vector of a vector space is unique, i.e. show that if $\mathbf{0}$ and $\tilde{\mathbf{0}}$ are two zero vectors in the same vector space, then $\mathbf{0} = \tilde{\mathbf{0}}$.

1.3 Subspaces

Definition 1.3.1. Let V be a vector space over a field F and let W be a non-empty subset of V. The set W, together with vector addition and scalar multiplication as defined on V, is called a subspace of V if W with these operations is a vector space over F.

The subspaces $W = \{\mathbf{0}_V\}$ and $W = V$ are called the trivial subspaces of V.

Theorem 1.3.2. *Let V be a vector space over a field F and let W be a non-empty subset of V. Then, W is a subspace of V if and only if for all $\mathbf{u}, \mathbf{v} \in W$ and $a \in F$, we have the closure conditions*

$$\mathbf{u} + \mathbf{v} \in W \quad \textit{and} \quad a\mathbf{u} \in W.$$

Proof.

$\Rightarrow$ Suppose W is a subspace of V. Then, W is a vector space, so W must be closed under vector addition and scalar multiplication.

$\Leftarrow$ Suppose that the closure conditions hold. Since the operations for vector addition and scalar multiplication are inherited from the vector space V, we know vector space axioms 1 and 2 (commutativity and associativity), 5 and 6 (properties of scalar multiplication), and 7 and 8 (distributivity) are satisfied. Therefore, we need only show:

3. Existence of zero vector: Since $a\mathbf{u} \in W$ for all $a \in F$, choose $a = 0_F$. Then, by part 1 of Theorem 1.2.5, we have $0_F\mathbf{u} = \mathbf{0} \in W$.

4. Existence of additive inverse: Now choose $a = -1_F$. Then, by part 4 of Theorem 1.2.5, we have $(-1_F)\mathbf{u} = -\mathbf{u} \in W$.

Hence, W is a vector space. ■

Lemma 1.3.3. *If W is a subspace of a vector space V, then $\mathbf{0}_V \in W$ and $\mathbf{0}_W = \mathbf{0}_V$.*

Proof. Exercise 1.4

Theorem 1.3.4. *Let U be a vector space. If V and W are subspaces of U, then $V \cap W$ is also a subspace of U.*

Proof. Suppose V and W are subspaces of U. Then, by Lemma 1.3.3, $\mathbf{0}_U \in V$ and $\mathbf{0}_U \in W$, so $\mathbf{0}_U \in V \cap W$. This means that $V \cap W \neq \emptyset$, and clearly $V \cap W \subseteq U$. We will now use Theorem 1.3.2.

Let $\mathbf{u}_1, \mathbf{u}_2 \in V \cap W$. Then $\mathbf{u}_1, \mathbf{u}_2 \in V$ and $\mathbf{u}_1, \mathbf{u}_2 \in W$, so since V and W are vector spaces, $\mathbf{u}_1 + \mathbf{u}_2 \in V$ and $\mathbf{u}_1 + \mathbf{u}_2 \in W$. Thus,

$$\mathbf{u}_1 + \mathbf{u}_2 \in V \cap W.$$

Let $\mathbf{u} \in V \cap W$ and let a be a scalar. Then, $a\mathbf{u} \in V$ and $a\mathbf{u} \in W$, so

$$a\mathbf{u} \in V \cap W.$$

Hence, $V \cap W$ is a subspace of U. ■

Corollary 1.3.5. *Let V be a vector space and let $W_1, W_2, \ldots, W_n$, $n \geq 2$ be subspaces of V. Then, $\cap_{k=1}^{n} W_k$ is also a subspace of V.*

Proof. Exercise 1.5

Exercises

Exercise 1.4. Prove Lemma 1.3.3.

Exercise 1.5. Prove Corollary 1.3.5.

1.4 Linear combinations

Definition 1.4.1. Let V be a vector space over a field F and let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$ and $a_1, a_2, \dots, a_n \in F$. A vector of the form

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \sum_{k=1}^{n} a_k\mathbf{v}_k$$

is called a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.

Definition 1.4.2. Let V be a vector space over a field F and let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \subseteq V$. The set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, i.e.

$$\{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \mid a_1, a_2, \dots, a_n \in F\},$$

is called the span of S, denoted $\operatorname{span}(S)$.

If $\operatorname{span}(S) = V$, we say that the set S spans V.

Theorem 1.4.3. *Let V be a vector space and let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \subseteq V$. Then,*

1. $\operatorname{span}(S)$ *is a subspace of V; and*
2. *If W is a subspace of V and $S \subseteq W$, then* $\operatorname{span}(S) \subseteq W$. *In other words,* $\operatorname{span}(S)$ *is the smallest subspace of V that contains all the vectors in S.*

Proof.

1. By assumption, $\operatorname{span}(S) \neq \emptyset$, and since V is closed under addition and scalar multiplication, $\operatorname{span}(S) \subseteq V$. Now we will show that $\operatorname{span}(S)$ is closed under addition and scalar multiplication. Let $\mathbf{u}, \mathbf{v} \in \operatorname{span}(S)$. Then,

$$\mathbf{u} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \quad \text{and} \quad \mathbf{v} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n$$

for some scalars $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$. Thus,

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n + b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_n\mathbf{v}_n \\ &= a_1\mathbf{v}_1 + b_1\mathbf{v}_1 + a_2\mathbf{v}_2 + b_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n + b_n\mathbf{v}_n \\ &= (a_1 + b_1)\mathbf{v}_1 + (a_2 + b_2)\mathbf{v}_2 + \cdots + (a_n + b_n)\mathbf{v}_n \in \operatorname{span}(S).\end{aligned}$$

Similarly, letting c be an arbitrary scalar, we see

$$c\mathbf{u} = c(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n) = ca_1\mathbf{v}_1 + ca_2\mathbf{v}_2 + \cdots + ca_n\mathbf{v}_n \in \text{span}(S).$$

Hence, by Theorem 1.3.2, $\text{span}(S)$ is a subspace of V.

2. Let W be a subspace of V such that $S \subseteq W$. Then, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in W$. Since W is a vector space, W is closed under addition and scalar multiplication, so any linear combination of vectors in W is also in W. In particular, any linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is in W. Hence, $\text{span}(S) \subseteq W$. ■

Definition 1.4.4. Let V be a vector space over a field F and let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subseteq V$ and $a_1, a_2, \ldots, a_n \in F$. The set S is called linearly independent if

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$$

only if $a_1 = a_2 = \cdots = a_n = 0_F$. Otherwise, if there exists such a linear combination where at least one of the coefficients $a_1, a_2, \ldots, a_n$ is non-zero, we say S is linearly dependent.

Theorem 1.4.5. *Let V be a vector space and let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subseteq V$ where $n \geq 2$. The set S is linearly dependent if and only if at least one of the vectors in S can be expressed as a linear combination of the others.*

Proof.

⇒ Suppose S is linearly dependent. Then, there exist scalars $a_1, a_2, \ldots, a_n$, not all zero, such that $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{c}_n = \mathbf{0}$. Say $a_k \neq 0$ for some $k \in \{1, 2, \ldots, n\}$. Then,

$$\begin{aligned}
\mathbf{0} &= a_1\mathbf{v}_1 + \cdots + a_{k-1}\mathbf{v}_{k-1} + a_k\mathbf{v}_k + a_{k+1}\mathbf{v}_{k+1} + \cdots + a_n\mathbf{v}_n \\
-a_k\mathbf{v}_k &= a_1\mathbf{v}_1 + \cdots + a_{k-1}\mathbf{v}_{k-1} + a_{k+1}\mathbf{v}_{k+1} + \cdots + a_n\mathbf{v}_n \\
\mathbf{v}_k &= -a_k^{-1}a_1\mathbf{v}_1 + \cdots + (-a_k^{-1})a_{k-1}\mathbf{v}_{k-1} + (-a_k^{-1})a_{k+1}\mathbf{v}_{k+1} + \\
&\quad \cdots + (-a_k^{-1})a_n\mathbf{v}_n.
\end{aligned}$$

⇐ Suppose that for some $\mathbf{v}_k \in S$, there exist scalars $b_1, \ldots, b_{k-1}, b_{k+1}, \ldots, b_n$ such that $\mathbf{v}_k = b_1\mathbf{v}_1 + \cdots + b_{k-1}\mathbf{v}_{k-1} + b_{k+1}\mathbf{v}_{k+1} + \cdots + b_n\mathbf{v}_n$. Then,

$$\begin{aligned}
\mathbf{0} &= b_1\mathbf{v}_1 + \cdots + b_{k-1}\mathbf{v}_{k-1} + (-\mathbf{v}_k) + b_{k+1}\mathbf{v}_{k+1} + \cdots + b_n\mathbf{v}_n \\
\mathbf{0} &= b_1\mathbf{v}_1 + \cdots + b_{k-1}\mathbf{v}_{k-1} + (-1)\mathbf{v}_k + b_{k+1}\mathbf{v}_{k+1} + \cdots + b_n\mathbf{v}_n,
\end{aligned}$$

where we see there is at least one coefficient, -1, that is definitely non-zero. Hence, S is linearly dependent. ■

1.5 Basis and dimension

Definition 1.5.1. A set $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ of vectors in a vector space V is called a basis of V if S is linearly independent and $\text{span}(S) = V$. The vectors in S are called basis vectors.

Theorem 1.5.2. *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis of a vector space V. Then, every $\mathbf{v} \in V$ can be written as a unique linear combination of the vectors in S.*

Theorem 1.5.3. *Let V be a vector space. If there exists a basis of V with n vectors, then every subset of V with more than n vectors is linearly dependent.*

Corollary 1.5.4. *Let V be a vector space. If there exists a basis of V with n vectors, then every basis of V has n vectors.*

Definition 1.5.5. Let V be a vector space that has a basis with n vectors. The number n is called the dimension of V, denoted $\dim(V)$. If $V = \{\mathbf{0}\}$, then $\dim(V) = 0$.

Theorem 1.5.6. *Let V be a vector space and let $\dim(V) = n$. Then,*

1. *If a subset $S \subseteq V$ with n vectors is linearly independent, then S is a basis of V; and*
2. *If a subset $S \subseteq V$ with n vectors spans V, then S is a basis of V.*

Appendix A

Solutions to Exercises

Chapter 1

Solution 1.1. Let $a, b, c, d, e, f \in \mathbb{R}$. First we must verify that $\mathbb{C}$ is indeed closed under the given operations. Since $(a + c), (b + d) \in \mathbb{R}$, we see that

$$(a + bi) + (c + di) = (a + c) + (b + d)i \in \mathbb{C},$$

and similarly, since $(ac - bd), (ad + bc) \in \mathbb{R}$, we also see that

$$(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i \in \mathbb{C}.$$

Thus, $\mathbb{C}$ is closed under $+$ and $\cdot$. Now we can check the field axioms:

1. Associativity:

$$\begin{aligned}((a + bi) + (c + di)) + (e + fi) &= (a + c + (b + d)i) + (e + fi)\\ &= a + c + e + (b + d + f)i\\ &= a + bi + (c + e) + (d + f)i\\ &= (a + bi) + ((c + di) + (e + fi)).\end{aligned}$$

$$\begin{aligned}((a + bi) \cdot (c + di)) \cdot (e + fi) &= (ac - bd + (ad + bc)i) \cdot (e + fi)\\ &= (ac - bd)e - (ad + bc)f + ((ac - bd)f + (ad + bc)e)i\\ &= ace - bde - adf - bcf + (acf - bdf + ade + bce)i\\ &= a(ce - df) - b(cf + de) + (a(cf + de) + b(ce - df))i\\ &= (a + bi) \cdot (ce - df + (cf + de)i)\\ &= (a + bi) \cdot ((c + di) \cdot (e + fi)).\end{aligned}$$

2. Commutativity:

$$(a + bi) + (c + di) = (a + b) + (c + d)i = (b + a) + (d + c)i = (c + di) + (a + bi).$$

$$(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i = (ca - db) + (da + cb)i = (c + di) \cdot (a + bi).$$

3. Existence of additive identity: Consider $0 = 0 + 0i \in \mathbb{C}$. We see

$$(a + bi) + (0 + 0i) = (a + 0) + (b + 0)i = a + bi.$$

4. Existence of multiplicative identity: Consider $1 = 1 + 0i \in \mathbb{C}$. We see

$$(a + bi) \cdot (1 + 0i) = (1a - 0b) + (0a + 1b)i = a + bi.$$

5. Existence of additive inverse: For $a + bi$, consider $-a - bi \in \mathbb{C}$. We see

$$(a + bi) + (-a - bi) = (a - a) + (b - b)i = 0 + 0i = 0.$$

6. Existence of multiplicative inverse: Suppose $a + bi \neq 0$. Consider[1]

$$\frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i \in \mathbb{C}.$$

We see

$$\begin{aligned}(a + bi) \cdot \left(\frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i \right) &= \frac{a^2}{a^2 + b^2} + \frac{b^2}{a^2 + b^2} + \left(\frac{ab}{a^2 + b^2} - \frac{ba}{a^2 + b^2} \right) i \\ &= \frac{a^2 + b^2}{a^2 + b^2} + 0i = 1.\end{aligned}$$

7. Distributivity:

$$\begin{aligned}(a + bi) \cdot ((c + di) + (e + fi)) &= (a + bi) \cdot (c + e + (d + f)i) \\ &= a(c + e) - b(d + f) + (a(d + f) + b(c + e))i \\ &= ac + ae - bd - bf + (ad + af + bc + be)i \\ &= ac - bd + (ad + bc)i + ae - bf + (af + be)i \\ &= ((a + bi) \cdot (c + di)) + ((a + bi) \cdot (e + fi)).\end{aligned}$$

Hence, $\mathbb{C}$ with these operations is a field. □

Solution 1.3. By the definition of the zero vector, for any vector $\mathbf{v}$, we have $\mathbf{v} + \mathbf{0} = \mathbf{v}$ and $\mathbf{v} + \tilde{\mathbf{0}} = \mathbf{v}$. In particular, $\mathbf{0} = \mathbf{0} + \tilde{\mathbf{0}} = \tilde{\mathbf{0}} + \mathbf{0} = \tilde{\mathbf{0}}$. □

Solution 1.4 (Proof of Lemma 1.3.3). Since W is a vector space, there must exist a zero vector $\mathbf{0}_W \in W$, and since $W \subseteq V$, we have $\mathbf{0}_W \in V$. By uniqueness of the zero vector (see Exercise 1.3), $\mathbf{0}_W = \mathbf{0}_V$. □

Solution 1.5 (Proof of Corollary 1.3.5). We can prove this result by induction. The base case (where $n = 2$) is Theorem 1.3.4. For our inductive hypothesis, suppose $U = \cap_{k=1}^{m} W_k$ is a subspace of V. By Theorem 1.3.4, if W_{m+1} is a subspace of V, then $U \cap W_{m+1} = \cap_{k=1}^{m+1} W_k$ is also a subspace of V. Hence, $\cap_{k=1}^{n} W_k$ is a subspace of V for all $n \geq 2$. □

[1] This can be found by setting $(a + bi) \cdot (x + yi) = 1$ and solving for x and y.