Linear Algebra W214

Linear Algebra W214

Bruce Bartlett

0.1 Note for the student

You are about to meet linear algebra for the second time. In the first year, we focused on systems of linear equations, matrices, and their determinants. That was good and well, but the time has come for you to return to these topics from a more abstract, mathematical viewpoint.

You should not be scared by *abstraction*. It simply means getting rid of extraneous detail and limiting oneself exclusively to the most important features of a problem. This allows you to understand the problem better. There are less things to worry about! Moreover, if you encounter another problem, which is superficially different but shares the same important features as the original problem, then you could understand it in the same way. This is the power of abstraction.

When we study abstract mathematics, we use the language of *definitions*, *theorems* and *proofs*. Learning to think along these lines (developing abstract mathematical thinking) can be a daunting task at first. But keep trying. One day it will 'click' into place and you will realize it is all much more simple than you had first imagined.

You cannot read mathematics the way you read a novel. You need to have a *pencil and notepad* with you, and you need to actively *engage* with the material. For instance, if you encounter a definition, start by writing down the definition on your notepad. Just the act of writing it out can be therapeutic!

If you encounter a worked example, try to write out the example yourself. Perhaps the example is trying to show that A equals B. Start by asking yourself: Do I understand what 'A' actually means? Do I understand what 'B' actually means? Only then are you ready to consider the question of whether A is equal to B !

Good luck in this new phase of your mathematical training. Enjoy the ride!

Contents

Chapter 1

Abstract vector spaces

1.1 Introduction

1.1.1 Three different sets

We start by playing a game. Recall that in mathematics, a *set* X is just a collection of distinct objects. We call these objects the *elements* of X.

I am going to show you three different sets, and you need to tell me the properties that they all have in common.

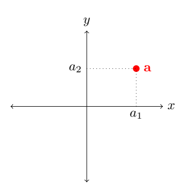
The first set, A, is defined to be the set of all ordered pairs (x, y) where x and y are real numbers.

> Let us pause for a second and translate this definition from English into mathematical symbols. The translation is:
>
> $$A := \{(a_1,\, a_2) : a_1, a_2 \in \mathbb{R}\}. \tag{1.1.1}$$
>
> The := stands for 'is defined to be'. The { and } symbols stand for 'the set of all'. The lone colon : stands for 'where' or 'such that'. The comma in between a and b stands for 'and'. The $\in$ stands for 'an element of'. And $\mathbb{R}$ stands for the set of all real numbers.
>
> Well done — you are learning the language of mathematics!

An element of A is an arbitrary pair of real numbers $\mathbf{a} = (a_1,\, a_2)$. For instance, $(1,\, 2) \in A$ and $(3.891,\, e^{\pi})$ are elements of A. Note also that I am using a boldface $\mathbf{a}$ to refer to an element of A. This is so that we can distinguish $\mathbf{a}$ from its *components* a_1 and a_2, which are just ordinary numbers (not elements of A).

We can visualize an element $\mathbf{a}$ of A as a point in the Cartesian plane whose x-coordinate is a_1 and whose y-coordinate is a_2:

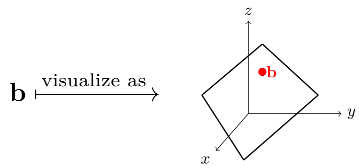
$$\mathbf{a} \xmapsto{\text{visualize as}}$$



The second set, B, is defined to be the set of all ordered real triples $(b_1,\, b_2,\, b_3)$ satisfying $b_1 - b_2 + b_3 = 0$. Translated into mathematical symbols,

$$B := \{(b_1,\, b_2,\, b_3) : b_1, b_2, b_3 \in \mathbb{R} \text{ and } b_1 - b_2 + b_3 = 0\}. \tag{1.1.2}$$

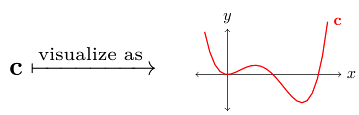
For instance, $(2, 3, 1) \in B$ but $(1, 1, 1) \notin B$. We can visualize an element $\mathbf{b}$ of B as a point in the plane in 3-dimensional space carved out by the equation $x - y + z = 0$:



The third set, C, is the set of all polynomials of degree 4. Translated into mathematical symbols,

$$C := \{\text{polynomials of degree} \leq 4\}. \tag{1.1.3}$$

Recall that the *degree* of a polynomial is the highest power of x which occurs. For instance, $\mathbf{c} = x^4 - 3x^3 + 2x^2$ is a polynomial of degree 4, and so is $\mathbf{p} = 2x^3 + \pi x$. So $\mathbf{c}$ and $\mathbf{p}$ are elements of C. But $\mathbf{r} = 8x^5 - 7$ and $\mathbf{s} = \sin(x)$ are not elements of C. We can visualize an element $\mathbf{c} \in C$ (i.e. a polynomial of degree 4) via its *graph*. For instance, the polynomial $\mathbf{c} = x^4 - 3x^3 + 2x^2 \in C$ can be visualized as:



There you have it. I have defined three different sets: A, B and C, and I have explained how to visualize the elements of each of these sets. On the face of it, the sets are quite different. Elements of A are arbitrary points in $\mathbb{R}^2$. Elements of B are points in $\mathbb{R}^3$ satisfying a certain equation. Elements of C are polynomials.

What features do these sets have in common?

1.1.2 Features the sets have in common

I want to focus on two features that the sets A, B and C have in common.

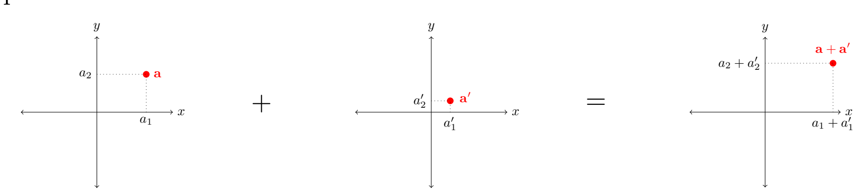
1.1.2.1 Addition

Firstly, in each of these sets, there is a natural *addition operation*. We can add two elements of the set to get a third element.

In set A, we can add two elements $\mathbf{a} = (a_1,\, a_2)$ and $\mathbf{a}' = (a'_1,\, a'_2)$ together by adding their components together, to form a new element $\mathbf{a} + \mathbf{a}' \in A$:

$$\underbrace{(a_1,\, a_2)}_{\mathbf{a}} + \underbrace{(a'_1,\, a'_2)}_{\mathbf{a}'} := \underbrace{(a_1 + a'_1,\, a_2 + a'_2)}_{\mathbf{a}+\mathbf{a}'} \tag{1.1.4}$$

For instance, $(1,\, 3) + (2,\, -1.6) = (3,\, 1.4)$. We can visualize this addition operation as follows:



We can do a similar thing in set B. Suppose we have two elements of B, $\mathbf{b} = (b_1, b_2, b_3)$ and $\mathbf{b}' = (b'_1, b'_2, b'_3)$. Note that, since $\mathbf{b} \in B$, its components satisfy $b_1 - b_2 + b_3 = 0$. Similarly the components of b' satisfy and $b_1 - b_2 + b_3 = 0$. We can add $\mathbf{b}$ and $\mathbf{b}'$ together to get a new element $\mathbf{b} + \mathbf{b}'$ of B, by adding their components together as before:

$$\underbrace{(b_1,\, b_2,\, b_3)}_{\mathbf{b}} + \underbrace{(b'_1,\, b'_2,\, b'_3)}_{\mathbf{b}'} := \underbrace{(b_1 + b'_1,\, b_2 + b'_2,\, b_3 + b'_3)}_{\mathbf{b}+\mathbf{b}'} \tag{1.1.5}$$

We should be careful here. How do we know that the expression on the right hand side is really an element of B? We need to check that it satisfies the equation 'the first component minus the second component plus the third component equals zero'. Let us check that formally:

$$\begin{aligned}(\mathbf{b}+\mathbf{b}')_1 - (\mathbf{b}+\mathbf{b}')_2 + (\mathbf{b}+\mathbf{b}')_3 &= (b_1 + b'_1) - (b_2 + b'_2) + (b_3 + b'_3) \\ &= (b_1 - b_2 + b_3) + (b'_1 - b'_2 + b'_3) \\ &= 0 + 0 \\ &= 0.\end{aligned}$$

We can visualize this addition operation in B in the same way as we did for A.

There is also an addition operation in set C. We can add two polynomials together algebraically by adding their corresponding coefficients:

$$\begin{aligned}&[c_4x^4 + c_3x^3 + c_2x^2 + c_1x^1 + c_0] + [d_4x^4 + d_3x^3 + d_2x^2 + d_1x^1 + d_0] \\ &:= (c_4 + d_4)x^4 + (c_3 + d_3)x^3 + (c_2 + d_2)x^2 + (c_1 + d_1)x^1 + (c_0 + d_0)\end{aligned} \tag{1.1.6}$$

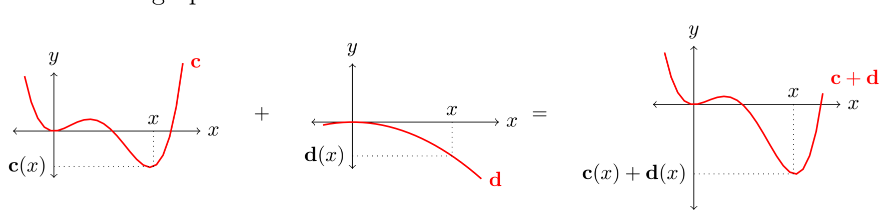
For instance,

$$[2x^4 + x^2 - 3x + 2] + [2x^3 - 7x^2 + x] = 2x^4 + 2x^3 - 6x^2 - 2x + 2.$$

There is another way to think about the addition of polynomials. Each polynomial $\mathbf{c}$ can be thought of as a *function*, in the sense that we can substitute an arbitrary value of x into the polynomial $\mathbf{c}$, and it will output a number $\mathbf{c}(x)$. For instance, if $\mathbf{c}(x) = 3x^2 - 1$, then $\mathbf{c}(2) = 11$. If we think of polynomials as functions in this way, then the addition $\mathbf{c} + \mathbf{d}$ of two polynomials can be thought of as the new function which, when you substitute some number x into it, outputs $\mathbf{c}(x) + \mathbf{d}(x)$. Written mathematically,

$$(\mathbf{c} + \mathbf{d})(x) := \mathbf{c}(x) + \mathbf{d}(x) \tag{1.1.7}$$

Thinking in this way, we can visualize the graph of $\mathbf{c} + \mathbf{d}$ as the graph of $\mathbf{c}$ added to the graph of $\mathbf{d}$:



1.1.2.2 Zero element

In all three sets A, B and C, there is a specific element (the *zero element*) $\mathbf{0}$ which, when you add it to another element, leaves that element unchanged.

In A, the zero element $\mathbf{0}$ is defined by

$$\mathbf{0} := (0, 0) \in A. \tag{1.1.8}$$

When you add this point to another point $(a_1, a_2) \in A$, nothing happens!

$$(0, 0) + (a_1, a_2) = (a_1, a_2).$$

Do not confuse the zero element $\mathbf{0} \in A$ with the real number zero, $0 \in \mathbb{R}$. This is another reason why I am using boldface! (You should use underline to distinguish them.)

In B, the zero element $\mathbf{0}$ is the point $(0, 0, 0) \in B$. When you add this point to another point $(u_1, u_2, u_3) \in B$, nothing happens!

$$(0, 0, 0) + (u_1, u_2, u_3) = (u_1, u_2, u_3).$$

In C, the zero element $\mathbf{0}$ is the *zero polynomial*. If we think algebraically, this is the degree polynomial whose coefficients are all zero:

$$\mathbf{0} = 0x^4 + 0x^3 + 0x^2 + 0x + 0 \tag{1.1.9}$$

If we think of the polynomial as a function, then the zero polynomial $\mathbf{0}$ is the function which returns zero for all values of x, that is $\mathbf{0}(x) = 0$ for all x. Whichever way we think of it, when we add the zero polynomial to another polynomial, nothing happens!

$$\begin{aligned}[0x^4 + 0x^3 + 0x^2 + 0x + 0] + [c_4x^4 + c_3x^3 + c_2x^2 + c_1x + c_0] \\ = [c_4x^4 + c_3x^3 + c_2x^2 + c_1x + c_0]\end{aligned}$$

1.1.2.3 Multiplication by scalars

The last feature all the sets A, B and C have in common is that in each set, you can *multiply* elements of the set by real numbers.

For instance, if $\mathbf{a} = (a_1, a_2)$ is an element of A, then we can multiply it by some arbitrary real number, say 9, to get a new element $9.\mathbf{a}$ of A. We do this multiplication component-wise:

$$9.(a_1, a_2) := (9a_1, 9a_2). \tag{1.1.10}$$

In general, if $k \in \mathbb{R}$ is an arbitrary real number, then we can multiply elements $\mathbf{a} \in A$ by k to get a new element $k.\mathbf{a} \in A$ by multiplying each component of $\mathbf{a}$ by k:

In general, if $k \in \mathbb{R}$ is an arbitrary real number, then we can multiply elements $\mathbf{a} \in A$ by k to get a new element $k.\mathbf{a} \in A$ by multiplying each component of $\mathbf{a}$ by k:

$$\underbrace{k.(a_1, a_2)}_{\text{Multiplying a vector by a scalar}} := (\underbrace{ka_1}_{\text{Multiplying two numbers together}}, \underbrace{ka_2})$$

Just be careful to distinguish scalar multiplication $k.\mathbf{a}$ (written with a .) from ordinary multiplication of real numbers ka_1 (written with no symbol, just using juxtaposition). Later on, because we are lazy, we will stop writing the $\cdot$ explicitly — you have been warned!

Visually, this multiplcation operation *scales* $\mathbf{a}$ by a factor of k. That is why we call it *scalar multiplication*.

There is a similar scalar multiplication in B:

$$k(u_1, u_2, u_3) := (ku_1, ku_2, ku_3) \tag{1.1.11}$$

There is also a scalar multiplication operation in C. We simply multiply each coefficient of a polynomial $\mathbf{c} \in C$ by k:

$$k.[c_4x^4 + c_3x^3 + c_2x^2 + c_1x + c_0] = kc_4x^4 + kc_3x^3 + kc_2x^2 + kc_1x + kc_0 \quad (1.1.12)$$

If we think of the polynomial $\mathbf{c}$ as a function, then this corresponds to *scaling* the graph of the function vertically by a factor of k.

1.1.3 Features that the sets do *not* have

Let us mention a few features that the sets do *not* have, or at least do not have in common.

- Set $A = \mathbb{R}^2$ has a *multiplication operation*. Because we can think of $\mathbb{R}^2$ as the complex plane $\mathbb{C}$, and we know how to multiply complex numbers. There is no clear choice of a multiplication operation on B. The same for C: if you try to multiply two degree 4 polynomials in C together, you will get out a polynomial of degree 8, which does not live in C!
- There is a *'take the derivative'* operation on C,

  $$\mathbf{c} \mapsto \frac{d}{dx}\mathbf{c}$$

  which we will meet again later. Note that taking the derivative decreases the degree of a polynomial by 1, so the result remains in C, and so this is a well defined map from C to C. There is no analogue of this operation in A and B.

  Note that there is no *integration* map from C to C, because integrating a polynomial *increases* the degree by 1, so the result might be a polynomial of degree 5, which does not live in C!

1.1.4 Rules

We have found that each of our three sets A, B and C have an *addition operation* $+$, a *zero vector* $\mathbf{0}$, and a *scalar multiplication* operation $\cdot$. Do these operations satisfy any rules, common to all three sets?

For instance, we can think of the addition operation in A as a function which assigns to each pair of elements $\mathbf{a}$ and $\mathbf{a}'$ in A a new element $\mathbf{a} + \mathbf{a}'$ in A. Does this operation satisfy any rules?

Let us see. Let $\mathbf{a} = (a_1, a_2)$ and $\mathbf{a}' = (a_1', a_2')$ be elements of A. We can add them in two different orders,

$$\mathbf{a} + \mathbf{a}' = (a_1 + a_1', a_2 + a_2')$$

and

$$\mathbf{a}' + \mathbf{a} = (a_1' + a_1, a_2' + a_2).$$

Are these the same? In other words, does the rule

$$\mathbf{a} + \mathbf{a}' = \mathbf{a}' + \mathbf{a} \quad (1.1.13)$$

hold? The answer is *yes*, but why? To check whether two elements of A are equal, we have to check whether each of their components are equal. The first component of $\mathbf{a} + \mathbf{a}'$ is $a_1 + a_1'$. The first component of $\mathbf{a}' + \mathbf{a}$ is $a_1' + a_1$. Is

$a_1 + a_1' = a_1' + a_1$? Yes — because these are just ordinary real numbers (not elements of A anymore), and we know that for ordinary real numbers, you can add them together in either order and get the same result. So the first component of $\mathbf{a}+\mathbf{a}'$ is equal to the first component of $a'+a$. Similarly, we can check that the second component of $\mathbf{a}+\mathbf{a}'$ is equal to the second component of $\mathbf{a}'+\mathbf{a}$. So all the components of $\mathbf{a}+\mathbf{a}'$ are equal to all the components of $\mathbf{a}'+\mathbf{a}$. So, finally, we conclude that $\mathbf{a}+\mathbf{a}' = \mathbf{a}'+\mathbf{a}$.

Does this rule (1.1.13) also hold for the addition operations in B and C? Yes. For instance, let us check that it holds in C. Suppose that $\mathbf{c}$ and $\mathbf{d}$ are polynomials in C. Does the rule

$$\mathbf{c}+\mathbf{d} = \mathbf{d}+\mathbf{c} \tag{1.1.14}$$

hold?

The left and right hand sides of (1.1.14) are elements of C. And elements of C are polynomials. To check if two polynomials are equal, we need to check if they are equal *as functions*, in other words, if you get identical results output from both functions no matter what input value of x you substitute in.

At an arbitrary input value x, the left hand side computes as $(\mathbf{c}+\mathbf{d})(x) = \mathbf{c}(x)+\mathbf{d}(x)$. On the other hand, the right hand side computes as $(\mathbf{d}+\mathbf{c})(x) = \mathbf{d}(x)+\mathbf{c}(x)$. Now, remember that $\mathbf{c}(x)$ and $\mathbf{d}(x)$ are just ordinary numbers (not polynomials). So $\mathbf{c}(x)+\mathbf{d}(x) = \mathbf{d}(x)+\mathbf{c}(x)$, because this is true for ordinary numbers. So for each input value x, $(\mathbf{c}+\mathbf{d})(x) = (\mathbf{d}+\mathbf{c})(x)$. Therefore the polynomials $\mathbf{c}+\mathbf{d}$ and $\mathbf{d}+\mathbf{c}$ are equal, because they output the same values for all numbers x.

There are other rules that also hold in all three sets. For instance, in all three sets, the rule

$$(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z}) \tag{1.1.15}$$

holds for any three elements $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$. Can you find the other common rules?

1.2 Definition of an abstract vector space

Mathematics is about identifying patterns. We have found three different sets, A, B and C, which look very different on the surface but have much in common. In each set, there is an addition operation, a zero vector, and a scalar multiplication operation. Moreover, in each set, these operations satisfy the same rules. Let us now record this pattern by giving it a name and writing down the rules explicitly.

Definition 1.2.1 A **vector space** is a set V equipped with the following data:

D1. An *addition operation*. (That is, for every pair of elements $\mathbf{u}, \mathbf{v} \in V$, a new element $\mathbf{u}+\mathbf{v} \in V$ is defined.)

D2. A *zero vector*. (That is, a special vector $\mathbf{0} \in V$ is marked out.)

D3. A *scalar multiplication operation*. (That is, for each real number k and each element $\mathbf{v} \in V$, a new element $k.\mathbf{v} \in V$ is defined).

This data should satisfy the following rules for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V and for all real numbers k and l:

R1. $\mathbf{v}+\mathbf{w} = \mathbf{w}+\mathbf{v}$

R2. $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$

R3a. $\mathbf{0}+\mathbf{v} = \mathbf{v}$

R3b. $\mathbf{v} + \mathbf{0} = \mathbf{v}$

R4. $k.(\mathbf{v} + \mathbf{w}) = k.\mathbf{v} + k.\mathbf{w}$

R5. $(k + l).\mathbf{v} = k.\mathbf{v} + l.\mathbf{v}$

R6. $k.(l.\mathbf{v}) = (kl).\mathbf{v}$

R7. $1.\mathbf{v} = \mathbf{v}$

R8. $0.\mathbf{v} = \mathbf{0}$

$\Diamond$

We will call the elements of a vector space *vectors*, and we will write them in boldface eg. $\mathbf{v} \in V$. We do this to distinguish vectors from real numbers, which we will call *scalars*, and which don't have a boldface. It is difficult to use boldface in handwriting, so you should write them with an arrow on top, like so: $\vec{v}$.

Also, in this chapter we will write scalar multiplication with a $\cdot$, for instance $k.\mathbf{v}$, but in later chapters we will simply write it as $k\mathbf{v}$ for brevity, so be careful!

To prove that a certain set can be given the structure of a vector space, one therefore needs to do the following:

1. Define a set V.
2. Define the data of an addition operation (D1), a zero vector (D2), and a scalar multiplication operation (D3) on V.
3. Check that this data satisfies the rules (R1) - (R8).

1.3 First example of a vector space

We were led to the definition (Definition 1.2.1) of an abstract vector space by considering the properties of sets A, B and C in Section 1.1. Let us check, for instance, that B indeed satisfies Definition 1.2.1. The others will be left as exercises.

Example 1.3.1 The set B is a vector space. ***1. Define a set B.***
We define

$$B := \{(u_1, u_2, u_3) : u_1, u_2, u_3 \in \mathbb{R} \text{ and } u_1 - u_2 + u_3 = 0\}. \tag{1.3.1}$$

2. Define addition, the zero vector, and scalar multiplication

D1. Addition We define addition as follows. Suppose $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, , v_2, v_3)$ are elements of B. Note that in particular this means $u_1 - u_2 + u_3 = 0$ and $v_1 - v_2 + v_3 = 0$. We define $\mathbf{u} + \mathbf{v}$ by:

$$\mathbf{u} + \mathbf{v} := (u_1 + v_1, u_2 + v_2, u_3 + v_3). \tag{1.3.2}$$

We need to check that this makes sense. We are supposed to have that $\mathbf{u}+\mathbf{v}$ is also an element of B. We can't just write down any definition! To check if $\mathbf{u} + \mathbf{v}$ is an element of B, we need to check if it satisfies equation (1.3.1). Let us check:

$(u_1 + v_1) - (u_2 + v_2) + (u_3 + v_3)$

$$\begin{aligned}
&= (u_1 - u_2 + u_3) + (v_1 - v_2 + v_3) && \text{(this algebra step is true for ordinary numbers)} \\
&= 0 + 0 && \text{(since } \mathbf{u} \text{ and } \mathbf{v} \text{ are in } B) \\
&= 0.
\end{aligned}$$

Therefore, $\mathbf{u} + \mathbf{v}$ is indeed an element of B, so we have written down a well-defined addition operation on B, which takes two arbitrary elements of B and outputs another element of B.

D2. Zero vector We define the zero vector $\mathbf{0} \in B$ as

$$\mathbf{0} := (0, 0, 0). \tag{1.3.3}$$

We need to check that this makes sense. Does $(0, 0, 0)$ really belong to B,in other words, does it satisfy equation (1.3.1)? Yes, since $0-0+0 = 0$. So we have a well-defined zero vector.

D3. Scalar multiplication We define scalar multiplication on B as follows. Let k be a real number and $\mathbf{u} = (u_1, u_2, u_3)$ be an element of B. We define

$$k.\mathbf{u} := (ku_1, ku_2, ku_3). \tag{1.3.4}$$

We need to check that this makes sense. When I multiply a multiply a vector $\mathbf{u}$ in B by a scalar k, the result $k.\mathbf{u}$ is supposed to be an element of B. Does (ku_1, ku_2, ku_3) really belong to B? Let us check that it satisfies the defining equation (1.3.1):

$$\begin{aligned}
&ku_1 - ku_2 + ku_3 \\
&= k(u_1 - u_2 + u_3) && \text{(this algebra step is true for ordinary numbers)} \\
&= k0 && \text{(since } \mathbf{u} \text{ is in } B) \\
&= 0.
\end{aligned}$$

Therefore, $k.\mathbf{u}$ is indeed an element of B, so we have written down a well-defined scalar multiplication operation on B.

3. Check that the data satisfies the rules

We must check that our data D1, D2 and D3 satisfies the rules R1 — R8. So, suppose $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{w} = (w_1, w_2, w_3)$ are in B, and suppose that k and l are real numbers.

R1 We check:

$$\begin{aligned}
&\mathbf{v} + \mathbf{w} \\
&= (v_1 + w_1, v_2 + w_2, v_3 + w_3) && \text{(by defn of addition in } B) \\
&= (w_1 + v_1, w_2 + v_2, w_3 + v_3) && \text{(because } x + y = y + x \text{ is true for real numbers)} \\
&= \mathbf{w} + \mathbf{v}. && \text{(by defn of addition in } B)
\end{aligned} \tag{R1.}$$

R2 We check:

$$\begin{aligned}
&(\mathbf{u} + \mathbf{v}) + \mathbf{w} \\
&= (u_1 + v_1, u_2 + v_2, u_3 + v_3) + \mathbf{w} && \text{(by defn of addition in } B) \\
&= ((u_1 + v_1) + w_1, (u_2 + v_2) + w_2, (u_3 + v_3) + w_3) && \text{(by defn of addition in } B) \\
&= (u_1 + (v_1 + w_1), u_2 + (v_2 + w_2), u_3 + (v_3 + w_3)) && \text{(since } (x + y) + z = x + (y + z) \text{ is true for real nu} \\
&= \mathbf{u} + (v_1 + w_1, v_2 + w_2, v_3 + w_3) && \text{(by defn of addition in } B) \\
&= \mathbf{u} + (\mathbf{v} + \mathbf{w}) && \text{(by defn of addition in } B).
\end{aligned}$$

R3 We check:

$$\begin{aligned}
&\mathbf{0}+\mathbf{v} \\
&= (0,\,0,\,0)+(v_1,\,v_2,\,v_3) && \text{(by defn of the zero vector in } B\text{)} \\
&= (0+v_1,\,0+v_2,\,0+v_3) && \text{(by defn of addition in } B\text{)} \\
&= (v_1,\,v_2,\,v_3) && \text{(because } x+0=x \text{ is true for real numbers)} \\
&= \mathbf{v}.
\end{aligned}$$

By the same reasoning, we can check that $\mathbf{v}+\mathbf{0}=\mathbf{v}$.

R4 We check:

$$\begin{aligned}
&k.(\mathbf{v}+\mathbf{w}) \\
&= k.(v_1+w_1,\,v_2+w_2,\,v_3+w_3) && \text{(by defn of addition in } B \\
&= (k(v_1+w_1,\,k(v_2+w_2),\,k(v_3+w_3)) && \text{(by defn of scalar multipliation in } B\text{)} \\
&= (kv_1+kw_1,\,kv_2+kw_2,\,kv_3+kw_3) && \text{(since } k(x+y)=kx+ky \text{ for real numbers } x,y\text{)} \\
&= (kv_1,\,kv_2,\,kv_3)+(kw_1,\,kw_2,\,kw_3) && \text{(by defn of addition in } B\text{)} \\
&= k.\mathbf{v}+k.\mathbf{w} && \text{(by defn of scalar multiplication in } B\text{)}
\end{aligned}$$

R5 We check:

$$\begin{aligned}
&(k+l).\mathbf{v} \\
&= ((k+l)v_1,\,(k+l)v_2,\,(k+l)v_3) && \text{(by defn of scalar multiplication in } B\text{)} \\
&= (kv_1+lv_1,\,kv_2+lv_2,\,kv_3+lv_3) && \text{(since } (k+l)x=kx+lx \text{ for real numbers)} \\
&= (kv_1,\,kv_2,\,kv_3)+(lv_1,\,lv_2,\,lv_3) && \text{(by defn of addition in } B\text{)} \\
&= k.\mathbf{v}+l.\mathbf{v} && \text{(by defn of scalar multiplication in } B\text{)}
\end{aligned}$$

R6 We check:

$$\begin{aligned}
&k.(l.\mathbf{v}) \\
&= k.(lv_1,\,lv_2,\,lv_3) && \text{(by defn of scalar multiplication in } B\text{)} \\
&= (k(lv_1),\,k(lv_2),\,k(lv_3)) && \text{(by defn of scalar multiplication in } B\text{)} \\
&= ((kl)v_1,\,(kl)v_2,\,(kl)v_3) && \text{(since } k(lx)=(kl)x \text{ for real numbers)} \\
&= (kl).\mathbf{v} && \text{(by defn of scalar multiplication in } B\text{).}
\end{aligned}$$

R7 We check:

$$\begin{aligned}
1.\mathbf{v} &= (1v_1,\,1v_2,\,1v_3) && \text{(by defn of scalar multiplication in } B\text{)} \\
&= (v_1,\,v_2,\,v_3) && \text{(since } 1x=x \text{ for real numbers } x\text{)} \\
&= \mathbf{v}.
\end{aligned}$$

R8 We check:

$$\begin{aligned}
0.\mathbf{v} &= (0v_1,\,0v_2,\,0v_3) && \text{(by defn of scalar multiplication in } B\text{)} \\
&= (0,\,0,\,0) && \text{(since } 0x=0 \text{ for real numbers} \\
&= \mathbf{0} && \text{(by defn of the zero vector in } B\text{).}
\end{aligned}$$

□

Exercises

1. Prove that set A from Section 1.1 together with the addition operation (1.1.4), the zero vector (1.1.8) and the scalar multiplication operation (1.1.10) forms a vector space.

2. Prove that set C from Section 1.1 together with the addition operation (1.1.6), the zero vector (1.1.9) and the scalar multiplication operation (1.1.12) forms a vector space.

3. Define the set C' consisting of all polynomials of degree *exactly* 4 as well as the zero polynomial. Show that if C' is given the addition operation (1.1.6), the zero vector (1.1.9) and the scalar multiplication operation (1.1.12) then C' *does not* form a vector space.

Hint. Give a counterexample!

Solution. Consider the following two polynomials in C':

$$\mathbf{p}(x) = x^4 + x^3,$$
$$\mathbf{q}(x) = -x^4.$$

However, their sum is not in C' since

$$\mathbf{p}(x) + \mathbf{q}(X) = (1-1)x^4 + x^3 = x^3$$

which has degree 3. Hence C' is not closed under addition and so cannot be a vector space.

4. Consider the set

$$X := \{(a_1, a_2) \in \mathbb{R}^2 : a_1 \geq 0, a_2 \geq 0\}$$

equipped with the same addition operation (1.1.4), zero vector (1.1.8) and scalar multiplication operation (1.1.10) as in A. Does X form a vector space? If not, why not?

Solution. X is not a vector space since the additive inverse of an element in X may fail to be in X. For example, consider $(1, 0)$. The additive inverse of $(1, 0)$ would have to be $(-1, 0)$. However, $(-1, 0)$ is certainly *not* in X. Hence X is not a vector space.

Solutions

· Exercises

1.3.3. Solution. Consider the following two polynomials in C':

$$\mathbf{p}(x) = x^4 + x^3,$$
$$\mathbf{q}(x) = -x^4.$$

However, their sum is not in C' since

$$\mathbf{p}(x) + \mathbf{q}(X) = (1-1)x^4 + x^3 = x^3$$

which has degree 3. Hence C' is not closed under addition and so cannot be a vector space.

1.3.4. Solution. X is not a vector space since the additive inverse of an element in X may fail to be in X. For example, consider $(1,0)$. The additive inverse of $(1,0)$ would have to be $(-1,0)$. However, $(-1,0)$ is certainly *not* in X. Hence X is not a vector space.

1.4 More examples and non-examples

Example 1.4.1 A non-example. Define the set V by

$$V := \{\mathbf{a}, \mathbf{b}\}. \tag{1.4.1}$$

Define the addition operation by

$$\begin{aligned} \mathbf{a}+\mathbf{a} &:= \mathbf{a} & \qquad \mathbf{a}+\mathbf{b} &:= \mathbf{a} \\ \mathbf{b}+\mathbf{a} &:= \mathbf{b} & \qquad \mathbf{b}+\mathbf{b} &:= \mathbf{c} \end{aligned}$$

For this to be a well-defined addition operation, we need to check that whenever you add two elements of V together, you get out a well-defined element of V. But $\mathbf{b}+\mathbf{b}=\mathbf{c}$, so adding $\mathbf{b} \in V$ to itself outputs something, namely $\mathbf{c}$, which is not an element of V. So V does not form a vector space since it does not even have a well-defined addition operation. □

Example 1.4.2 Another non-example. Define the set V by

$$V := \{\mathbf{a}, \mathbf{b}\}. \tag{1.4.2}$$

Define the addition operation by

$$\begin{aligned} \mathbf{a}+\mathbf{a} &:= \mathbf{a} & \qquad \mathbf{a}+\mathbf{b} &:= \mathbf{b} \\ \mathbf{b}+\mathbf{a} &:= \mathbf{b} & \qquad \mathbf{b}+\mathbf{b} &:= \mathbf{a} \end{aligned}$$

This is a well-defined addition operation, since whenever you add two elements of V together, you get out a well-defined element of V.

Define the zero vector by

$$\mathbf{0} := \mathbf{a}. \tag{1.4.3}$$

This is well-defined, since $\mathbf{a}$ is indeed an element of V.

Define scalar multiplication by a real number $k \in \mathbb{R}$ by

$$k.\mathbf{a} := \mathbf{a} \text{ and } k.\mathbf{b} := \mathbf{b}. \tag{1.4.4}$$

This is a well-defined scalar multiplication, since it allows us to multiply any element $\mathbf{v} \in V$ by a scalar k and it outputs a well-defined element $k.\mathbf{v} \in V$. □

Checkpoint 1.4.3 Show that these operations satisfy rules R1, R2, R3, R4, R6 and R7, but not R5 and R8.

Solution. R1: We must check whether $\mathbf{v}+\mathbf{w}=\mathbf{w}+\mathbf{v}$ for all $\mathbf{v}, \mathbf{w} \in \{\mathbf{a}, \mathbf{b}\}$. Clearly $\mathbf{a}+\mathbf{a}=\mathbf{a}+\mathbf{a}$ and likewise for $\mathbf{b}$. And finally $\mathbf{b}=\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$.

R2: We must check whether $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \{\mathbf{a}, \mathbf{b}\}$. This requires we check 8 equations in total. For brevity, we shall only present the solution for one of them, the rest are virtually identical. We check whether

$$(\mathbf{a}+\mathbf{b})+\mathbf{b}=\mathbf{a}+(\mathbf{b}+\mathbf{b}$$

To that end, consider:

$$\begin{aligned}\text{LHS} &= (\mathbf{a} + \mathbf{b}) + \mathbf{b} \\ &= \mathbf{b} + \mathbf{b} \\ &= \mathbf{a}.\end{aligned}$$

By a similar method:

$$\begin{aligned}\text{RHS} &= \mathbf{a} + (\mathbf{b} + \mathbf{b}) \\ &= \mathbf{a} + \mathbf{a} \\ &= \mathbf{a}.\end{aligned}$$

R3, R4, R6, and R7 all follow from routine checks.

We shall demonstrate why R5 is not satisfied. For that, we need to find a counterexample. Take $k = 2 = l, \mathbf{v} = \mathbf{b}$. Then

$$\begin{aligned}\text{LHS} &= (2 + 2).\mathbf{b} \\ &= 4.\mathbf{b} \\ &= \mathbf{b}\end{aligned}$$

whereas

$$\begin{aligned}\text{RHS} &= 2.\mathbf{b} + 2.\mathbf{b} \\ &= \mathbf{b} + \mathbf{b} \\ &= \mathbf{a}\end{aligned}$$

Since LHS $\neq$ RHS, R5 cannot be true.

Example 1.4.4 The zero vector space. Define the set Z by

$$Z := \{\mathbf{z}\}. \tag{1.4.5}$$

Note that it contains just a single element, $\mathbf{z}$. Define the addition operation as

$$\mathbf{z} + \mathbf{z} := \mathbf{z} \tag{1.4.6}$$

Define the zero element as

$$\mathbf{0} := \mathbf{z}. \tag{1.4.7}$$

Finally define scalar multiplication by a scalar $k \in \mathbb{R}$ as:

$$k.\mathbf{z} := \mathbf{z}. \tag{1.4.8}$$

□

Checkpoint 1.4.5 Show that this data satisfies the rules R1 to R8.

Example 1.4.6 $\mathbb{R}^n$. Define the set $\mathbb{R}^n$ by

$$\mathbb{R}^n := \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R} \text{ for all } i = 1 \ldots n\}. \tag{1.4.9}$$

Define the addition operation as

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n). \tag{1.4.10}$$

Define the zero element as

$$\mathbf{0} := (0, 0, \ldots, 0). \tag{1.4.11}$$

Define scalar multiplication by

$$k.(x_1, x_2, \ldots, x_n) := (kx_1, kx_2, \ldots, kx_n). \tag{1.4.12}$$

□

Checkpoint 1.4.7 Show that this data satisfies the rules R1 to R8.

Example 1.4.8 $\mathbb{R}^\infty$**.** Define the set $\mathbb{R}^\infty$ by

$$\mathbb{R}^\infty := \{(x_1, x_2, x_3, \ldots,) : x_i \in \mathbb{R} \text{ for all } i = 1, 2, 3, \ldots\} \tag{1.4.13}$$

So an element $\mathbf{x} \in \mathbb{R}^\infty$ is an infinite sequence of real numbers. Define the addition operation componentwise:

$$(x_1, x_2, x_3, \ldots) + (y_1, y_2, y_3, \ldots) := (x_1 + y_1, x_2 + y_2, x_3 + y_3, \ldots). \tag{1.4.14}$$

Define the zero element as

$$\mathbf{0} := (0, 0, 0, \ldots), \tag{1.4.15}$$

the infinite sequence whose components are all zero. Finally, define scalar multiplication componentwise:

$$k.(x_1, x_2, x_3, \ldots) := (kx_1, kx_2, kx_3, \ldots) \tag{1.4.16}$$

Thinking about infinity is an important part of mathematics. Have you watched the movie about me called *The man who knew infinity*?

□

Checkpoint 1.4.9 Show that this data satisfies the rules R1 to R8.

Solution. We shall only check R4 below, the rest are similar.

R4 :Let

$$\begin{aligned} \mathbf{v} &= (v_1, v_2, v_3, \ldots) \\ \mathbf{w} &= (w_1, w_2, w_3, \ldots) \end{aligned}$$

We must check whether $k.(\mathbf{v} + \mathbf{w}) = k.\mathbf{v} + k.\mathbf{w}$.

$$\begin{aligned} \text{LHS} &= k.(\mathbf{v} + \mathbf{w}) \\ &= k.\left[(v_1, v_2, v_3, \ldots) + (w_1, w_2, w_3, \ldots)\right] \\ &= k.\left(v_1 + w_1, v_2 + w_2, v_3 + w_3, \ldots\right) \\ &= (k(v_1 + w_1), k(v_2 + w_2), k(v_3 + w_3), \ldots) \\ &= (kv_1 + kw_1, kv_2 + kw_2, kv_3 + kw_3, \ldots) \\ &= (kv_1, kv_2, kv_3, \ldots) + (kw_1, kw_2, kw_3, \ldots) \\ &= k.(v_1, v_2, v_3, \ldots) + k.(w_1, w_2, w_3, \ldots) \\ &= k.\mathbf{v} + k.{=}w \\ &= \text{LHS} \end{aligned}$$

Example 1.4.10 Functions on a set. Let X be any set. Define the set $\text{Fun}(X)$ of *real-valued functions on* X by

$$\text{Fun}(X) := \{f : X \to \mathbb{R}\}. \tag{1.4.17}$$

Note that the functions can be arbitrary; there is no requirement for them to be continuous, or differentiable. Such a requirement would not make sense, since X could be an arbitrary set. For instance X could be the set $\{a, b, c\}$ —without any further informaiton, it does not make sense to say that a function $f : X \to \mathbb{R}$ is continuous.

Define the addition operation by

$$(f + g)(x) := f(x) + g(x), x \in X \tag{1.4.18}$$

Make sure you understand what this formula is saying! We start with two functions f and g, and we are defining their sum $f + g$. This is supposed to be another function on X. To define a function on X, I am supposed to write down what value it assigns to each $x \in X$. And that is what the formula says: the value that the function $\mathbf{f} + \mathbf{g}$ assigns to an element $x \in X$ is defined to be the number $f(x)$ plus the number $g(x)$. Remember: f is a function, while $f(x)$ is a number!

Define the zero vector, which we will call z in this example, to be the function which outputs the number 0 for every input value of $x \in X$:

$$z(x) := 0 \text{ for all } x \in X. \tag{1.4.19}$$

Define scalar multiplication by

$$(k.f)(x) := kf(x). \tag{1.4.20}$$

□

Checkpoint 1.4.11 Notation quiz! Say whether the following combination of symbols represents a real number or a function.

1. f
2. $f(x)$
3. $k.f$
4. $(k.f)(x)$

Solution.

1. Function
2. Real Number
3. Function
4. Real Number

Checkpoint 1.4.12 Let $X = \{a, b, c\}$.

1. Write down three different functions f, g, h in $\text{Fun}(X)$.
2. For each of the functions you wrote down in Item 1.4.12.1, calculate (i) $f + g$ and (ii) $3.h$.

Solution.

1.

$$f(a) = 4$$
$$f(b) = 0$$

$$f(c) = 2$$

$$\begin{aligned} g(a) &= 1 \\ g(b) &= 1 \\ g(c) &= 1 \end{aligned}$$

$$\begin{aligned} h(a) &= 0 \\ h(b) &= 3 \\ h(c) &= 0 \end{aligned}$$

2.

$$\begin{aligned} (f+g)(a) &= 5 \\ (f+g)(b) &= 1 \\ (f+g)(c) &= 3 \end{aligned}$$

$$\begin{aligned} (3.h)(a) &= 0 \\ (3.h)(b) &= 9 \\ (3.h)(c) &= 0 \end{aligned}$$

Checkpoint 1.4.13 Show that the data (1.4.18), (1.4.19), (1.4.20) satisfies the rules R1 to R8, so that $\mathrm{Fun}(X)$ is a vector space.

Example 1.4.14 Matrices. The set $\mathrm{Mat}_{n,m}$ of all $n \times m$ matrices is a vector space. See Appendix A for a reminder about matrices. □

Checkpoint 1.4.15 Show that when equipped with the addition operation, zero vector, and scalar multiplication operation as defined in Appendix A, the set $\mathrm{Mat}_{n,m}$ of all $n \times m$ matrices is a vector space.

Example 1.4.16 We will write Col_n for the vector space $\mathrm{Mat}_{n,1}$ of n-dimensional *column vectors*,

$$\mathrm{Col}_n = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} : x_1, \ldots, x_n \in \mathbb{R} \right\}$$

So, Col_n 'is' just $\mathbb{R}^n$, but we make explicit the fact that the components of the vectors are arranged in a column. □

Exercises

1. Define an addition operation on the set $X := \{\mathbf{0}, \mathbf{a}, \mathbf{b}\}$ by the following table:

$+$	$\mathbf{0}$	$\mathbf{a}$	$\mathbf{b}$
$\mathbf{0}$	$\mathbf{0}$	$\mathbf{a}$	$\mathbf{b}$
$\mathbf{a}$	$\mathbf{a}$	$\mathbf{0}$	$\mathbf{a}$
$\mathbf{b}$	$\mathbf{b}$	$\mathbf{a}$	$\mathbf{0}$

This table works as follows. To calculate, for example, $\mathbf{b}+\mathbf{a}$, find the intersection of the row labelled by $\mathbf{b}$ with the column labelled by $\mathbf{a}$. We see that $\mathbf{b}+\mathbf{a} := \mathbf{a}$.

Prove that this addition operation satisfies R1.

2. Prove that the addition operation from Exercise 1.4.1 does not satisfy R2.

3. Define a strange new addition operation $\hat{+}$ on $\mathbb{R}$ by

$$x \hat{+} y := x - y, \qquad x, y \in \mathbb{R}.$$

Does $\hat{+}$ satisfy R2? If it does, prove it. If it does not, give a counterexample.

Solution. No, for example:

$$(1 \hat{+} 2) \hat{+} 3 = (1-2) - 3 = -4.$$

But

$$1 \hat{+} (2 \hat{+} 3) = 1 - (2-3) = 2.$$

4. Construct an operation $\boxplus$ on $\mathbb{R}$ satisfying R1 but not R2.

Hint. Try adjusting the formula from Exercise 1.4.3.

Solution. Define $x \boxplus y = |x-y|$. R1 is satisfied since $x \boxplus y = |x-y| = |y-x| = y \boxplus x$. However, R2 is not satisfied since $(1 \boxplus 2) \boxplus 3 = ||1-2|-3| = 2$ but $1 \boxplus (2 \boxplus 3) = |1 - |2-3|| = 0$

5. Let $\mathbb{R}^+$ be the set of positive real numbers. Define an addition operation $\oplus$, a zero vector z and a scalar multiplication .on $\mathbb{R}^+$ by

$$x \oplus y := xy$$
$$z := 1$$
$$k.x := x^k$$

where $x, y \in \mathbb{R}^+$, and k is a scalar (i.e. an arbitrary real number).

(a) Check that these operations are well-defined. For instance, is $x+y \in \mathbb{R}^+$, as it should be?

(b) Check that these operations satisfy R1 to R8.

We conclude that $\mathbb{R}^+$, equipped with these operations, forms a vector space.

Solution.

(a) Let x, y be two positive reals. Then

$$x \oplus y := xy$$

is certainly also a positive real number. Similarly, for any $k \in R$ and any $x \in \mathbb{R}^+$,

$$k.x := x^k$$

is also positive. To see this, for any fixed k, notice that the graph of the function $f(x) = x^k$restricted to $x \geq 0$ lies above the x-axis.

(b)

6. Consider the operation $\oplus$ on $\mathbb{R}^2$ defined by:

$$(a_1, a_2) \oplus (b_1, b_2) := (a_1 + b_2, a_2 + b_1).$$

(a) Does this operation satisfy R1 ?

(b) Does this operation satisfy R2 ?

Solution.

(a) Yes. This follows from the usual rules for addition in $\mathbb{R}$.

$$(a_1, a_2) \oplus (b_1, b_2) := (a_1 + b_2, a_2 + b_1) = (b_1 + a_1, b_2 + a_2) = (b_1, b_2) \oplus (a_1, a_2).$$

(b) Similar to a, except using the associativity of addition in $\mathbb{R}$.

Solutions

· Exercises

1.4.3. Solution. No, for example:

$$(1 \hat{+} 2) \hat{+} 3 = (1 - 2) - 3 = -4.$$

But

$$1 \hat{+} (2 \hat{+} 3) = 1 - (2 - 3) = 2.$$

1.4.4. Solution. Define $x \boxplus y = |x - y|$. R1 is satisfied since $x \boxplus y = |x - y| = |y - x| = y \boxplus x$. However, R2 is not satisfied since $(1 \boxplus 2) \boxplus 3 = ||1 - 2| - 3| = 2$ but $1 \boxplus (2 \boxplus 3) = |1 - |2 - 3|| = 0$

1.4.5. Solution.

(a) Let x, y be two positive reals. Then

$$x \oplus y := xy$$

is certainly also a positive real number. Similarly, for any $k \in R$ and any $x \in \mathbb{R}^+$,

$$k.x := x^k$$

is also positive. To see this, for any fixed k, notice that the graph of the function $f(x) = x^k$restricted to $x \geq 0$ lies above the x-axis.

(b)

1.4.6. Solution.

(a) Yes. This follows from the usual rules for addition in $\mathbb{R}$.

$$(a_1, a_2) \oplus (b_1, b_2) := (a_1 + b_2, a_2 + b_1) = (b_1 + a_1, b_2 + a_2) = (b_1, b_2) \oplus (a_1, a_2).$$

(b) Similar to a, except using the associativity of addition in $\mathbb{R}$.

1.5 Some results about abstract vector spaces

It is time to use the rules of a vector space to prove some general results.

> We are about to do our first formal proof in the course!

Our first lemma shows that the zero vector $\mathbf{0}$ is the *unique* vector in V which 'behaves like a zero vector'. More precisely:

Lemma 1.5.1 *Suppose V is a vector space with zero vector $\mathbf{0}$. If $\mathbf{0}'$ is a vector in V satisfying*

$$\mathbf{0}' + \mathbf{v} = \mathbf{v} \text{ for all } \mathbf{v} \in V \tag{1.5.1}$$

then $\mathbf{0}' = \mathbf{0}$.

Proof.

$$\begin{aligned} \mathbf{0} &= \mathbf{0}' + \mathbf{0} && \text{using (1.5.1) with } \mathbf{v} = \mathbf{0} \\ &= \mathbf{0}' && (\text{ R3b }) \end{aligned}$$

■

Definition 1.5.2 If V is a vector space, we define the **additive inverse** of a vector $\mathbf{v} \in V$ as

$$-\mathbf{v} := (-1).\mathbf{v}$$

◊

Lemma 1.5.3 *If V is a vector space, then for all $\mathbf{v} \in V$,*

$$-\mathbf{v} + \mathbf{v} = \mathbf{0} \text{ and } \mathbf{v} + (-\mathbf{v}) = \mathbf{0}. \tag{1.5.2}$$

Proof.

$$\begin{aligned} -\mathbf{v} + \mathbf{v} &= (-1).\mathbf{v} + \mathbf{v} && (\text{using defn of } -\mathbf{v}) \\ &= (-1).\mathbf{v} + 1.\mathbf{v} && (\text{R7}) \\ &= (-1+1).\mathbf{v} && (\text{R5}) \\ &= 0.\mathbf{v} \\ &= \mathbf{0} && (\text{R8}) \end{aligned}$$

In addition,

$$\begin{aligned} \mathbf{v} + (-\mathbf{v}) &= -\mathbf{v} + \mathbf{v} && (\text{R1}) \\ &= \mathbf{0} && (\text{by previous proof}) \end{aligned}$$

■

Lemma 1.5.4 *Suppose that two vectors $\mathbf{w}$ and $\mathbf{v}$ in a vector space satisfy $\mathbf{w} + \mathbf{v} = \mathbf{0}$. Then $\mathbf{w} = -\mathbf{v}$.*

Proof.

$$\begin{aligned} \mathbf{w} &= \mathbf{w} + \mathbf{0} && (\text{R3b}) \\ &= \mathbf{w} + (\mathbf{v} + -\mathbf{v}) && \text{by Lemma 1.5.3} \\ &= (\mathbf{w} + \mathbf{v}) + -\mathbf{v} && (\text{R2}) \\ &= \mathbf{0} + -\mathbf{v} && (\text{by assumption}) \\ &= -\mathbf{v} && (\text{R3a}). \end{aligned}$$

■

Let us prove two more lemmas, for practice.

Lemma 1.5.5 *Let V be a vector space and k any scalar. Then*

$$k.\mathbf{0} = \mathbf{0}$$

Proof.

$$\begin{aligned}
k.\mathbf{0} &= k.(0.\mathbf{0}) && (\text{R8 for } \mathbf{v} = \mathbf{0}) \\
&= ((k)(0)).\mathbf{0} && (\text{R6}) \\
&= 0.\mathbf{0} && ((k)(0) = 0 \text{ for any real number } k) \\
&= \mathbf{0} && (\text{R8 for } \mathbf{v} = \mathbf{0})
\end{aligned}$$

■

Lemma 1.5.6 *Suppose that $\mathbf{v}$ is a vector in a vector space V and that k is a scalar. Then*

$$k.\mathbf{v} = \mathbf{0} \Leftrightarrow k = 0 \text{ or } \mathbf{v} = \mathbf{0}.$$

Proof. ***(Proof of $\Leftarrow$).*** Suppose $k = 0$. Then $k.\mathbf{v} = 0.\mathbf{v} = \mathbf{0}$ by R8 of a vector space. On the other hand, suppose $\mathbf{v} = \mathbf{0}$. Then $k.\mathbf{v} = k.\mathbf{0} = \mathbf{0}$ by Exercise 1.5.2.

(Proof of $\Rightarrow$). Suppose $k.\mathbf{v} = \mathbf{0}$. There are two possibilities: either $k = 0$, or $k \neq 0$. If $k = 0$, then we are done. If $k \neq 0$, then $\frac{1}{k}$ exists and we can multiply both sides by it:

$$\begin{aligned}
k.\mathbf{v} &= \mathbf{0} \\
\therefore \frac{1}{k}.(k.\mathbf{v}) &= \frac{1}{k}.\mathbf{0} && (\text{Multiplied both sides by } \tfrac{1}{k}) \\
\therefore \left(\frac{1}{k}k\right).\mathbf{v} &= \mathbf{0} && (\text{On the LHS, we used R6. On the RHS, we used Exercise 1.5.2}) \\
\therefore 1.\mathbf{v} &= \mathbf{0} && (\text{using } \tfrac{1}{k}k = 1) \\
\therefore \mathbf{v} &= \mathbf{0} && (\text{R7})
\end{aligned}$$

Hence in the case $k \neq 0$ we must have $\mathbf{v} = \mathbf{0}$, which is what we wanted to show. ■

Example 1.5.7 Let us practice using the rules of a vector space to perform everyday calculations. For instance, suppose that we are trying to solve for the vector $\mathbf{x}$ appearing in the following equation:

$$\mathbf{v} + 7.\mathbf{x} = \mathbf{w} \tag{1.5.3}$$

We do this using the rules as follows:

$$\begin{aligned}
\mathbf{v} + 7.\mathbf{x} &= \mathbf{w} \\
\therefore -\mathbf{v} + (\mathbf{v} + 7.\mathbf{x}) &= -\mathbf{v} + \mathbf{w} && (\text{Added } -\mathbf{v} \text{ on left to both sides}) \\
\therefore (-\mathbf{v} + \mathbf{v}) + 7.\mathbf{x} &= -\mathbf{v} + \mathbf{w} && (\text{used R2 on LHS}) \\
\therefore \mathbf{0} + 7.\mathbf{x} &= -\mathbf{v} + \mathbf{w} && (\text{used Lemma 1.5.3 on LHS}) \\
\therefore 7.\mathbf{x} &= -\mathbf{v} + \mathbf{w} && (\text{used R3a on LHS}) \\
\therefore \frac{1}{7}.(7.\mathbf{x}) &= \frac{1}{7}.(-\mathbf{v} + \mathbf{w}) && (\text{scalar multiplied both sides by } \tfrac{1}{7}) \\
\therefore (\frac{1}{7}7).\mathbf{x} &= \frac{1}{7}.(-\mathbf{v} + \mathbf{w}) && (\text{used R6 on LHS}) \\
\therefore 1.\mathbf{x} &= \frac{1}{7}.(-\mathbf{v} + \mathbf{w}) && (\text{multiplied } \tfrac{1}{7} \text{ with } 7)
\end{aligned}$$

$$\therefore \mathbf{x} = \frac{1}{7}.(-\mathbf{v} + \mathbf{w}) \qquad \text{(R7)}$$

As the course goes on we will leave out all these steps. But it is important for you to be able to reproduce them all, if asked to do so! □

Exercises

1. Prove that for all vectors $\mathbf{v}$ in a vector space, $-(-\mathbf{v}) = \mathbf{v}$.

Solution. We apply the definition of $-\mathbf{v}$ twice:

$$-(-\mathbf{v}) = (-1).(-\mathbf{v}) = (-1).(-1.(\mathbf{v})).$$

Using Item we get

$$(-1).(-1(\mathbf{v})) = ((-1)(-1)).\mathbf{v} = 1.\mathbf{v}.$$

Finally, a single application of Item allows us to conclude that

$$1.\mathbf{v} = \mathbf{v}$$

2. Let V be a vector space with zero vector $\mathbf{0}$. Prove that for all scalars k, $k.\mathbf{0} = \mathbf{0}$.

Solution. We apply Item to $k.\mathbf{0}$:

$$k.\mathbf{0} = k.(\mathbf{0} + \mathbf{0}).$$

By Item we get

$$k.(\mathbf{0} + \mathbf{0}) = k.\mathbf{0} + k.0.$$

We now know

$$k.\mathbf{0} = k.\mathbf{0} + k.0.$$

Adding the inverse of $k.0$ to both sides we get

$$\mathbf{0} = k.\mathbf{0} + \mathbf{0} = k.0.$$

And we are done.

3. Let V be a vector space. Suppose that a vector $\mathbf{v} \in V$ satisfies

$$5.\mathbf{v} = 2.\mathbf{v}. \tag{1.5.4}$$

Prove that $\mathbf{v} = \mathbf{0}$.

Solution.

$$\begin{aligned}
&5.v = 2.v\\
&\implies 5.v + (-2).v = 2.v + (-2).v\\
&\implies (5-2).v = (2-2).v\\
&\implies 3.v = 0.v\\
&\implies (\frac{1}{3}3).v = (\frac{1}{3}0)v\\
&\implies 1v = 0v\\
&\implies v = 0
\end{aligned}$$

4. Suppose that two vectors $\mathbf{x}$ and $\mathbf{w}$ in a vector space satisfy $2\mathbf{x} + 6\mathbf{w} = \mathbf{0}$. Solve for $\mathbf{x}$, showing explicitly how you use the rules of a vector space, as in Example 1.5.7.

Solution.

$$
\begin{aligned}
& 2\mathbf{x} + 6\mathbf{w} = \mathbf{0} \\
\implies & 2\mathbf{x} + 6\mathbf{w} + (-6\mathbf{w})\mathbf{0} + (-6\mathbf{w}) && \text{(Existence of Inverses)} \\
\implies & 2\mathbf{x} + 0 = -6\mathbf{w} && \text{(R3a)} \\
\implies & 2\mathbf{x} = -6\mathbf{w} && \text{(R3b)} \\
\implies & (\frac{1}{2})(2\mathbf{x}) = ((\frac{1}{2}))(-6\mathbf{w}) \\
\implies & (\frac{1}{2}2)\mathbf{x} = (-\frac{1}{2}6)\mathbf{w}) && \text{(R6)} \\
\implies & 1\mathbf{x} = -3\mathbf{w} \\
\implies & \mathbf{x} = -3\mathbf{w} && \text{(R7)}
\end{aligned}
$$

5. Suppose V is a vector space which is not the zero vector space. Show that V contains infinitely many elements.

Hint 1. Since V is not the zero vector space, there must exist a vector $\mathbf{v} \in V$ such that $\mathbf{v} \neq \mathbf{0}$.

Hint 2. Use the idea of the proof from Exercise 1.5.3.

True or False For each of the following statements, write down whether the statement is true or false, and prove your assertion. (In other words, if you say that it is true, prove that it is true, and if you say that it is false, prove that it is false, by giving an *explicit counterexample*.)

6. If $k.\mathbf{v} = \mathbf{0}$ in a vector space, then it necessarily follows that $k = 0$.

Solution. False. Take $\mathbb{R}^2$ as an example. If $v = (0, 0)$ then $2.(0, 0) = (0, 0)$ but, of course, $2 \neq 0$.

7. If $k.\mathbf{v} = \mathbf{0}$ in a vector space, then it necessarily follows that $\mathbf{v} = \mathbf{0}$.

8. The empty set can be equipped with data D1 , D2 , D3 satisfying the rules of a vector space.

Solution. False. In order for the empty set to be a vector space, it must have a zero vector. That is, we must be able to find some $v \in$ the empty set satisfying the axioms for the zero vector. However, since the empty set has no elements in it, by definition, we cannot ever hope to find such a v. Hence the empty set can never be a vector space.

9. Rule R3b of a vector space follows automatically from the other rules.

Solution. True. Combining R1 and R3a gives R3b .

10. Rule R7 of a vector space follows automatically from the other rules.

Solution. False. Let V be a non-zero vector space (such as $\mathbb{R}^2$). Redefine scalar multiplication as follows

$$k.v := 0 \text{ for all scalars k and all vectors v.}$$

Then V will satisfy all the rules of a vector space except R7. Thus it is not the case that R7 follows from the other rules.

Solutions

- **Exercises**

1.5.1. Solution. We apply the definition of $-\mathbf{v}$ twice:

$$-(-\mathbf{v}) = (-1).(-\mathbf{v}) = (-1).(-1.(\mathbf{v})).$$

Using Item we get

$$(-1).(-1(\mathbf{v})) = ((-1)(-1)).\mathbf{v} = 1.\mathbf{v}.$$

Finally, a single application of Item allows us to conclude that

$$1.\mathbf{v} = \mathbf{v}$$

1.5.2. Solution. We apply Item to $k.\mathbf{0}$:

$$k.\mathbf{0} = k.(\mathbf{0} + \mathbf{0}).$$

By Item we get

$$k.(\mathbf{0} + \mathbf{0}) = k.\mathbf{0} + k.0.$$

We now know

$$k.\mathbf{0} = k.\mathbf{0} + k.0.$$

Adding the inverse of $k.0$ to both sides we get

$$\mathbf{0} = k.\mathbf{0} + \mathbf{0} = k.0.$$

And we are done.

1.5.3. Solution.

$$\begin{aligned}
&5.v = 2.v \\
&\implies 5.v + (-2).v = 2.v + (-2).v \\
&\implies (5-2).v = (2-2).v \\
&\implies 3.v = 0.v \\
&\implies (\frac{1}{3}3).v = (\frac{1}{3}0)v \\
&\implies 1v = 0v \\
&\implies v = 0
\end{aligned}$$

1.5.4. Solution.

$$\begin{aligned}
&2\mathbf{x} + 6\mathbf{w} = \mathbf{0} \\
&\implies 2\mathbf{x} + 6\mathbf{w} + (-6\mathbf{w})\mathbf{0} + (-6\mathbf{w}) && \text{(Existence of Inverses)} \\
&\implies 2\mathbf{x} + 0 = -6\mathbf{w} && \text{(R3a)} \\
&\implies 2\mathbf{x} = -6\mathbf{w} && \text{(R3b)} \\
&\implies (\frac{1}{2})(2\mathbf{x}) = ((\frac{1}{2}))(-6\mathbf{w}) \\
&\implies (\frac{1}{2}2)\mathbf{x} = (-\frac{1}{2}6)\mathbf{w}) && \text{(R6)} \\
&\implies 1\mathbf{x} = -3\mathbf{w} \\
&\implies \mathbf{x} = -3\mathbf{w} && \text{(R7)}
\end{aligned}$$

True or False 1.5.6. Solution. False. Take $\mathbb{R}^2$ as an example. If $v = (0,0)$ then $2.(0,0) = (0,0)$ but, of course, $2 \neq 0$.

1.5.8. Solution. False. In order for the empty set to be a vector space, it must have a zero vector. That is, we must be able to find some $v \in$ the empty set satisfying the axioms for the zero vector. However, since the empty set has no elements in it, by definition, we cannot ever hope to find such a v. Hence the empty set can never be a vector space.

1.5.9. Solution. True. Combining R1 and R3a gives R3b .

1.5.10. Solution. False. Let V be a non-zero vector space (such as $\mathbb{R}^2$). Redefine scalar multiplication as follows

$$k.v := 0 \text{ for all scalars k and all vectors v.}$$

Then V will satisfy all the rules of a vector space except R7. Thus it is not the case that R7 follows from the other rules.

1.6 Subspaces

In this section we will introduce the notion of a *subspace* of a vector space. This notion will allow us to quickly establish many more examples of vector spaces.

1.6.1 Definition of a subspace

Definition 1.6.1 A subset $U \subseteq V$ of a vector space V is called a **subspace** of V if:

- For all $\mathbf{u}, \mathbf{u}' \in U$, $\mathbf{u} + \mathbf{u}' \in U$
- $\mathbf{0} \in U$
- For all scalars k and all vectors $\mathbf{u} \in U$, $k.\mathbf{u} \in U$

$\Diamond$

Lemma 1.6.2 *If U is a subspace of a vector space V, then U is also a vector space, when we equip it with the same addition operation, zero vector and scalar multiplication as in V.*

Proof. Since U is a subspace, we know that it actually makes sense to "equip it with the same addition operation, zero vector and scalar multiplication as in V". (If U was *not* a subspace, then we might have for instance $\mathbf{u}, \mathbf{u}' \in U$ but $\mathbf{u} + \mathbf{u}' \notin U$, so the addition operation would not make sense.)

So we simply need to check the rules R1 to R8. Since these rules hold for all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V, they certainly hold for all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in U. So R1 to R8 are satisfied. ■

1.6.2 Examples of subspaces

Example 1.6.3 Line in $\mathbb{R}^2$. A line L through the origin in $\mathbb{R}^2$ is a subspace of $\mathbb{R}^2$:

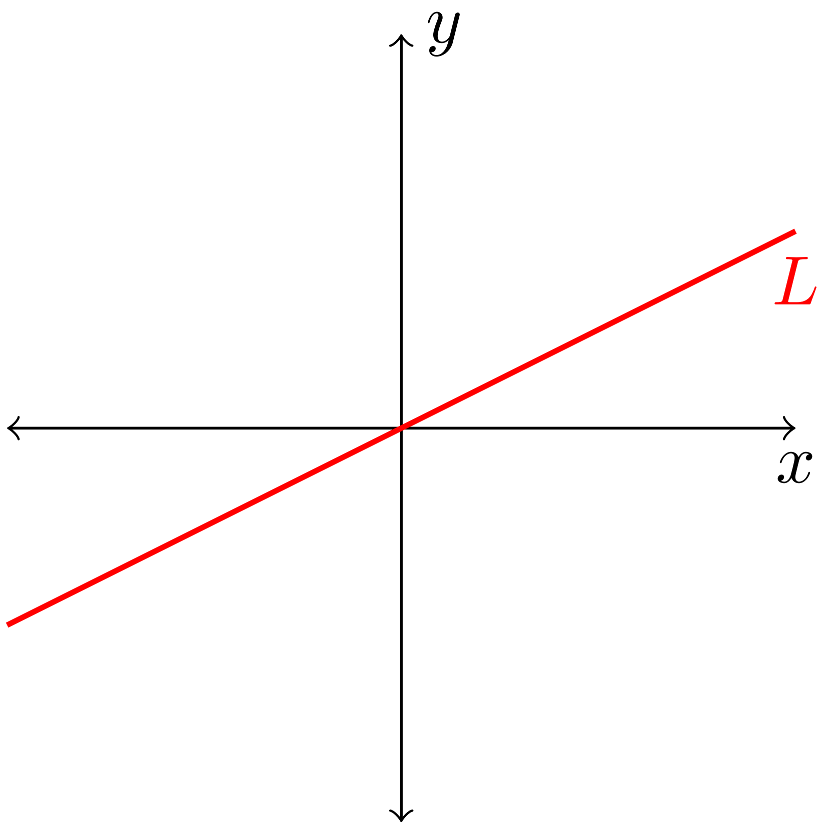


Figure 1.6.4: A line through the origin in $\mathbb{R}^2$.

Indeed, recall that L can be specified by a homogenous linear equation of the form:

$$L = \{(x, y) \in \mathbb{R}^2 : ax + by = 0\} \tag{1.6.1}$$

for some constants a and b. So, if $\mathbf{v} = (x, y)$ and $\mathbf{v}' = (x', y')$ lie on L, then their sum $\mathbf{v}+\mathbf{v}' = (x+x', y+y')$ also lies on L, because its components satisfy the defining equation (1.6.1):

$$\begin{aligned}
&a(x + x') + b(y + y') \\
&= (ax + by) + (ax' + by') \\
&= 0 + 0 && \text{(since } ax + by = 0 \text{ and } ax' + by' = 0\text{)} \\
&= 0.
\end{aligned}$$

This also makes sense geometrically: if you look at the picture, then you will see that adding two vectors $\mathbf{v}, \mathbf{v}'$ on L by the head-to-tail method will produce another vector on L. □

Checkpoint 1.6.5 Complete the proof that L is a subspace of $\mathbb{R}^2$ by checking that the zero vector is in L, and that multiplying a vector in L by a scalar outputs a vector still in L.

Example 1.6.6 Lines and planes in $\mathbb{R}^3$. Similarly a line L or a plane P through the origin in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$:

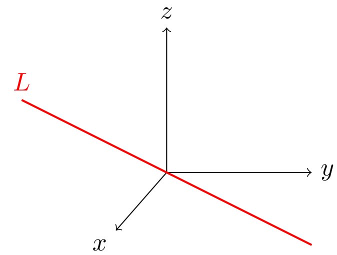


Figure 1.6.7: A line through the origin in $\mathbb{R}^3$.

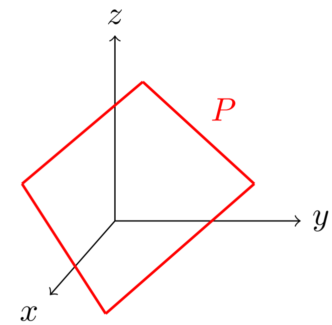


Figure 1.6.8: A plane through the origin in $\mathbb{R}^3$.

□

Example 1.6.9 Zero vector space. If V is a vector space, the the set $\{\mathbf{0}\} \subseteq V$ containing just the zero vector $\mathbf{0}$ is a subspace of V. □

Checkpoint 1.6.10 Check this.

Example 1.6.11 Non-example: Line not through origin. Be careful though — not *every* line $L \subset \mathbb{R}^2$ is a subspace of $\mathbb{R}^2$. If L does not go through the origin, then $\mathbf{0} \notin L$, so L is not a subspace.

Another reason that L is not a subspace is that it is not closed under addition: when we add two nonzero vectors $\mathbf{v}$ and $\mathbf{v}'$ on L, we end up with a vector $\mathbf{v} + \mathbf{v}'$ which does not lie on L:

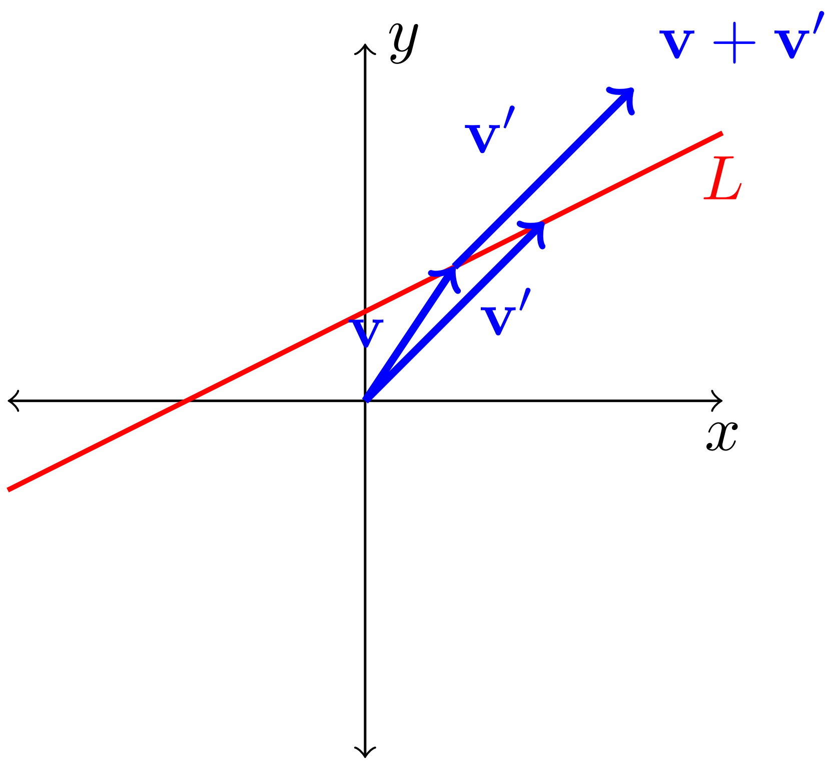


Figure 1.6.12: A line which does not pass through the origin is not closed under addition.

□

Example 1.6.13 Hyperplanes orthogonal to a fixed vector. This example generalizes Example 1.6.6 to higher dimensions. Let $\mathbf{v} \in \mathbb{R}^n$ be a fixed nonzero vector. The *hyperplane orthogonal to* $\mathbf{v}$ is the set W of all vectors orthogonal to $\mathbf{v}$, that is,

$$W := \{\mathbf{w} \in \mathbb{R}^n : \mathbf{v} \cdot \mathbf{w} = 0\}. \tag{1.6.2}$$

You will prove in Exercise Checkpoint 1.6.14 that W is a subspace of $\mathbb{R}^n$.

For example, consider the vector $\mathbf{v} = (1, 2, 3) \in \mathbb{R}^3$. Then the hyperplane orthogonal to $\mathbf{v}$ is

$$W = \{\mathbf{w} \in \mathbb{R}^3 : \mathbf{v} \cdot \mathbf{w} = 0\}. \tag{1.6.3}$$

If we write $\mathbf{w} = (w_1, w_2, w_3)$ then $\mathbf{v} \cdot \mathbf{w} = 0$ translates into the equation

$$w_1 + 2w_2 + 3w_3 = 0. \tag{1.6.4}$$

So, W can be regarded as the set of vectors in $\mathbb{R}^3$ whose components satisfy (1.6.4). □

Checkpoint 1.6.14 $\mathbf{v} \in \mathbb{R}^n$

$$W := \{\mathbf{w} \in \mathbb{R}^n : \mathbf{v} \cdot \mathbf{w} = 0\}.$$

$\mathbb{R}^n$

Solution.

Closed under addition.. Suppose $\mathbf{w}, \mathbf{w}' \in W$.
That is, $\mathbf{v} \cdot \mathbf{w} = 0$ and $\mathbf{v} \cdot \mathbf{w}' = 0$.
We must show that $\mathbf{w} + \mathbf{w}' \in W$.
That is, we must show that $\mathbf{v} \cdot (\mathbf{w} + \mathbf{w}') = 0$.
Indeed,

$$\begin{aligned} \mathbf{v} \cdot (\mathbf{w} + \mathbf{w}) &= \mathbf{v} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}' \\ &= 0 + 0 \\ &= 0. \end{aligned}$$

Contains the zero vector.. Since $\mathbf{v} \cdot \mathbf{0} = 0$, we conclude that $\mathbf{0} \in W$.

Closed under scalar multiplication.. Suppose $\mathbf{w} \in W$ and k is a scalar.
That is, $\mathbf{v} \cdot \mathbf{w} = 0$.
We must show that $k.\mathbf{w} \in W$.
That is, we must show that $\mathbf{v} \cdot (k.\mathbf{w}) = 0$.
Indeed,

$$\begin{aligned} \mathbf{v} \cdot (k.\mathbf{w}) &= k.(\mathbf{v} \cdot \mathbf{w}) \\ &= (k)(0) \\ &= 0. \end{aligned}$$

Example 1.6.15 Continuous functions as a subspace. The set

$$\mathrm{Cont}(I) := \{\mathbf{f} : I \to \mathbb{R},\ \mathbf{f} \text{ continuous}\}$$

of all continuous functions on an interval I is a subspace of the set $\mathrm{Fun}(I)$ of *all* functions on I. Let us check that it satisfies the definition. You know from earlier courses that:

- If $\mathbf{f}$ and $\mathbf{g}$ are continuous functions on I, then $\mathbf{f} + \mathbf{g}$ is also a continuous function.

- The zero function 0 defined by $\mathbf{0}(x) = 0$ for all $x \in I$ is a continuous function.
- If $\mathbf{f}$ is a continuous function, and k is a scalar, then $k.\mathbf{f}$ is also continuous.

Hence, by Lemma 1.6.2, $\mathrm{Cont}(I)$ is a vector space in its own right. □

Example 1.6.16 Differentiable functions as a subspace. Similarly, the set

$$\mathrm{Diff}(I) := \{\mathbf{f} : I \to \mathbb{R},\ \mathbf{f} \text{ differentiable}\}$$

of differentiable functions on an open interval I is a subspace of $\mathrm{Fun}(I)$. □

Checkpoint 1.6.17 Check this. Also, is $\mathrm{Diff}(I)$ a subspace of $\mathrm{Cont}(I)$?

Example 1.6.18 Vector spaces of polynomials. A *polynomial* is a function $\mathbf{p} : \mathbb{R} \to \mathbb{R}$ of the form

$$\mathbf{p}(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0. \tag{1.6.5}$$

for some fixed real coefficients $a_0, \ldots, a_n$. Two polynomials $\mathbf{p}$ and $\mathbf{q}$ are *equal* if they are equal *as functions*, that is if $\mathbf{p}(x) = \mathbf{q}(x)$ for all $x \in \mathbb{R}$. The *degree* of a polynomial is the highest power of x which occurs in its formula.

For example, $2x^3 - x + 7$ is a polynomial of degree 3, while $x^5 - 2$ is a polynomial of degree 5. We write the set of *all* polynomials as Poly and the set of all polynomials having degree less than or equal to n as Poly_n. □

Checkpoint 1.6.19 Check that Poly and Poly_n are subspaces of $\mathrm{Cont}(\mathbb{R})$.

Example 1.6.20 Polynomials in many variables. A *monomial* in two variables x and y is an expression of the form $x^m y^n$ for some nonnegative integers m, n. The *total degree* of the monomial is $m + n$.For example,

$$x^3y^2, xy^7, \quad \text{and } x^3$$

are monomials in the variables x and y of total degree

$$5, 8, \quad \text{and } 3$$

respectively. A *polynomial in two variables* is a linear combination of monomials. The *degree* of the polynomial is the highest total degree of the monomials occuring in the linear combination. For instance,

$$p = 5x^3y^2 - 3xy^7, \quad q = xy^2 - x^3 + 3y^2 \tag{1.6.6}$$

are polynomials in two variables with total degree 8 and 3 respectively.

We write Poly^2 for the set of *all* polynomials in two variables, and Poly^2_n for the set of all two-variable polynomials of total degree less than or equal to n. For instance, consier the polynomial

$$p = 5x^3y^2 - 3xy^7$$

The total degree of p is 8. So:

- $p \in \mathrm{Poly}^2$,
- $p \in \mathrm{Poly}^2_8$
- $p \in \mathrm{Poly}^2_{12}$

- $p \ni \mathrm{Poly}^2_7$

We can regard a polynomial p in two variables as a special kind of function

$$p : \mathbb{R}^2 \to \mathbb{R}.$$

In this way, we can regard Poly_2 (and similarly Poly^2_n) as a subset of the vector space $\mathrm{Fun}(\mathbb{R}^2)$ of *all* functions on $\mathbb{R}^2$. Indeed, they are *subspaces*, as the reader will check.

Two polynomials p and q in variables x and y are defined to be *equal* if and only if all their corresponding coefficients are equal. This is equivalent to the statement that $p(x, y) = q(x, y)$ for all $(x, y) \in \mathbb{R}^2$.

In the same way, we can talk about 3-variable polynomials, and so on, eg.

$$r = 5x^3y^2z + 3xy - 4xz^3$$

is a polynomial of total degree 6 in the variables x, y, z. We can regard a polynomial in k variables as a special kind of function

$$r : \mathbb{R}^k \to \mathbb{R}.$$

We write Poly^k for the vector space of all k-variable polynomials. We can regard Poly^k as a subspace of the vector space $\mathrm{Fun}(\mathbb{R}^k)$ of *all* functions on $\mathbb{R}^k$. Similary, we write Poly^k_n for the vector space of all k-variable polynomials of total degree less than or equal to n, and we regard it too as a subspace of $\mathrm{Fun}(\mathbb{R}^n)$. □

When we say simply 'p is a polynomial' we will mean that p is a polynomial in a single variable, i.e. $p \in \mathrm{Poly}$. Note that $\mathrm{Poly} = \mathrm{Poly}^1$ and that $\mathrm{Poly}_n = \mathrm{Poly}^1_n$.

Example 1.6.21 Trigonometric polynomials. A *trigonometric polynomial* is a function $\mathbf{T} : \mathbb{R} \to \mathbb{R}$ of the form

$$\mathbf{T}(x) = a_0 + \sum_{k=1}^{n} a_k \cos(kx) + \sum_{k=1}^{n} b_k \sin(kx). \tag{1.6.7}$$

The *degree* of a trigonometric polynomial is the highest multiple of x which occurs inside one of the sines or cosines in its formula. For instance,

$$3 - \cos(x) + 6\sin(3x)$$

is a trigonometric polynomial of degree 3. We write the set of *all* trigonometric polynomials as Trig and the set of all trigonometric polynomials having degree less than or equal to n as Trig_n. □

Checkpoint 1.6.22 Show that Trig and Trig_n are subspaces of $\mathrm{Cont}(\mathbb{R})$.

Checkpoint 1.6.23 Consider the function $\mathbf{f}(x) = \sin^3(x)$. Show that $\mathbf{f} \in \mathrm{Trig}_3$ by writing it in the form (1.6.7). Hint: use the identities

$$\sin(A)\sin(B) = \frac{1}{2}(\cos(A - B) - \cos(A + B))$$
$$\sin(A)\cos(B) = \frac{1}{2}(\sin(A - B) + \sin(A + B))$$

$$\cos(A)\cos(B) = \frac{1}{2}(\cos(A-B) + \cos(A+B))$$

which follow easily from the addition formulae

$$\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B$$
$$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B.$$

1.6.3 Solutions to homogenous linear differential equations

A homogenous nth order linear ordinary differential equation on an interval I is a differential equation of the form

$$a_n(x)y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + \cdots + a_1(x)y'(x) + a_0(x)y(x) = 0, \quad x \in I \tag{1.6.8}$$

where $y^{(k)}$ means the nth derivative of y. A *solution* to the differential equation is just some function $y(x)$ defined on the interval I which satisfies (1.6.8).

Example 1.6.24 An example of a 2nd order homogenous linear differential equation. For instance,

$$x^2y'' - 3xy' + 5y = 0, \quad x \in (0, \infty) \tag{1.6.9}$$

is a homogenous 2nd order linear differential equation on the interval $(0, \infty)$, and

$$y_1(x) = x^2 \sin(\log x) \tag{1.6.10}$$

is a solution to (1.6.9). Similarly,

$$y_2(x) = x^2 \cos(\log x) \tag{1.6.11}$$

is also a solution to (1.6.9).

We can use SageMath to check that these are indeed solutions to (1.6.9). Click the `Evaluate (Sage)` button --- it should output 'True', indicating that y_1 is indeed a solution to the differential equation.

```
def solves_de(y):
 return bool(x^2 *diff(y,x,2) -3*x*diff(y,x) + 5*y == 0)

y1 = x^2*sin(log(x))

solves_de(y1)
```

Edit the code above to check whether y_2 is a solution of the differential equation (1.6.9).

We can also plot the graphs of y_1 and y_2. Again, click on `Evaluate (Sage)`.

```
y1 = x^2*sin(log(x))
y2 = x^2*cos(log(x))

plot([y1, y2], (x, 0, 1), legend_label=['y1', 'y2'])
```

Play with the code above, and plot some different functions.

□

Checkpoint 1.6.25 Check by hand that (1.6.10) and (1.6.11) are indeed solutions of the differential equation (1.6.9).

Suppose we are given an nth order homogenous linear differential equation of the form (1.6.8) on some interval $I \subseteq \mathbb{R}$. Write V for the set of *all* solutions to the differential equation. That is,

$$V := \{y : a_n(x)y^{(n)}(x) + \cdots + a_1(x)y'(x) + a_0(x)y(x) = 0\} \tag{1.6.12}$$

We can regard V as a subset of the set of *all* functions on the interval I:

$$V \subseteq \mathrm{Fun}(I)$$

Checkpoint 1.6.26 Show that V is a subspace of $\mathrm{Fun}(I)$.

So, by Lemma 1.6.2, we conclude that *the set of solutions to a homogenous linear differential equation is a vector space.*

Example 1.6.27 Continuation of Example 1.6.24. Consider the differential equation from Example 1.6.24. We saw that

$$y_1 = x^2 \sin(\log x), \quad y_2 = x^2 \cos(\log x)$$

are solutions. So, any linear combination of y_1 and y_2 is *also* a solution. For instance,

$$y = 2y_1 + 5y_2$$

is also a solution. Let us check this in SageMath.

```
def solves_de(y):
  return bool(x^2 *diff(y,x,2) -3*x*diff(y,x) + 5*y == 0)

y1 = x^2*sin(log(x))
y2 = x^2*sin(cos(x))

solves_de(2*y1 + 5*y2)
```

□

Example 1.6.28 A non-example: Solutions to a nonlinear ODE. We saw in the previous example that linear ordinary differential equations (ODEs) are well-behaved - a linear combination of solutions is still a solution. This need not occur in the nonlinear case. For example, consider the nonlinear ODE

$$y' = y^2. \tag{1.6.13}$$

The general solution is given by

$$y_c = \frac{1}{c - x}$$

where c is a constant. For instance,

$$y_1 = \frac{1}{1-x}, \quad y_2 = \frac{1}{2-x}$$

are solutions.

Use the SageMath script below to check whether the linear combination $y_1 + y_2$ is also a solution.

```
y = function('y')(x)

def solves_de(f):
```

```
 return bool(diff(f,x) - f^2 == 0)

y1 = 1/(1-x)
y2 = 1/(2-x)

solves_de(y1+y2)
```

The answer is `False`! So linear combinations of solutions to the nonlinear differential equation (1.6.13) are no longer solutions, in general. □

Example 1.6.29 Finding the general solution to a differential equation in SageMath. Let us use SageMath to find the general solution of the following ordinary differential equation

$$y'' + 2y' + y = 0. \tag{1.6.14}$$

We can do this as follows. Note that we need to be a bit more careful now, first defining our variable x and then declaring that y is a function of x.

```
var('x')
y = function('y')(x)

diff_eqn = diff(y,x,2) +2*diff(y,x,1) + 5*y == 0
  desolve(diff_eqn,y)

desolve(diff_eqn, y)
```

SageMath reports that the general solution is given in terms of two unspecified constants `_K1` and `_K2` as `(_K2*cos(2*x) + _K1*sin(2*x))*e^(-x)`.

If we set `_K1` equal to 1 and `_K2` equal to 0 in the general solution, we will get a particular solution y_1 of the differential equation.

```
var('x,␣_K1,␣_K2')
y = function('y')(x)

diff_eqn = diff(y,x,2) +2*diff(y,x,1) + 5*y == 0

my_soln = desolve(diff_eqn,y)

y1 = my_soln.substitute(_K1==1, _K2==0)

y1
```

SageMath is telling us that $y_1 = e^{-x}\sin(2x)$ is a particular solution.

Edit the code to set `_K2` equal to 0 and `_K1` equal to 1 in the general solution to get a different particular solution y_2. What is y_2?

□

1.6.4 Exercises

1. Show that the set

$$V := \{(a, -a, b, -b) : a, b \in \mathbb{R}\}$$

is a subspace of $\mathbb{R}^4$.

2. Show that the set

$$V := \{\text{polynomials of the form } \mathbf{p}(x) = ax^3 + bx^2 - cx + a,\ a, b, c \in \mathbb{R}\}$$

is a subspace of Poly_3.

3. Let $b \in \mathbb{R}$. Prove that

$$V := \{(x_1,\, x_2,\, x_3) \in \mathbb{R}^3 : 2x_1 - 3x_2 + 5x_3 = b\}$$

is a subspace of $\mathbb{R}^3$ if and only if $b = 0$.

4. Consider the set

$$V := \{\mathbf{f} \in \text{Diff}((-1, 1)) : f'(0) = 2\}$$

Is V a subspace of $\text{Diff}((-1, 1))$? If you think it is, *prove* that it is. If you think it is not, *prove* that it is not!

5. Consider the set

$$V := \{(x_1,\, x_2,\, x_3,\, \ldots) \in \mathbb{R}^\infty : \lim_{n\to\infty} x_n = 0\}$$

Is V a subspace of $\mathbb{R}^\infty$? If you think it is, *prove* that it is. If you think it is not, *prove* that it is not!

6. Is $\mathbb{R}^+ := \{\mathbf{x} \in \mathbb{R} : \mathbf{x} \geq 0\}$ a subspace of $\mathbb{R}$? If you think it is, *prove* that it is. If you think it is not, *prove* that it is not!

7. Give an example of a nonempty subset U of $\mathbb{R}^2$ which is closed under addition and under taking additive inverses (i.e. if $\mathbf{u}$ is in U then $-\mathbf{u}$ is in V), but U is not a subspace of $\mathbb{R}^2$.

8. Give an example of a nonempty subset V of $\mathbb{R}^2$ which is closed under scalar multiplication, but V is not a subspace of $\mathbb{R}^2$.

The next 4 exercises will help acquaint the reader with the concept of the sum of two subspaces. First, we'll need a definition.

Definition 1.6.30 Let V be a vector space. Suppose U and W are two subspaces of V. The sum $U + W$ of U and W is defined by

$$U + V = \{\mathbf{u} + \mathbf{w} \in V \ : \ \mathbf{u} \in U, \mathbf{w} \in W\}$$

$\diamondsuit$

In the exercises below, V, U, W will be as above.

9. Show that $U + V$ is a subspace of V.

10. Show that $U + V$ is, in fact, the smallest subspace of V containing both U and V.

11. If $W \subset U$ what is $U + W$?

12. Can you think of two subspaces of $\mathbb{R}^2$ whose sum is $\mathbb{R}^2$? Similarly, can you think of two subspaces of $\mathbb{R}^2$ whose sum is not all of $\mathbb{R}^2$?

1.6.5 Solutions

Chapter 2

Finite-dimensional vector spaces

In this course we concentrate on *finite-dimensional* vector spaces, which we will define in this chapter.

> *Warning*: From now on, I will use shorthand and write scalar multiplication $k.\mathbf{v}$ simply as $k\mathbf{v}$!

2.1 Linear combinations and span

We start with some basic definitions.

Definition 2.1.1 A **linear combination** of a finite collection $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of vectors in a vector space V is a vector of the form

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_b\mathbf{v}_n \tag{2.1.1}$$

where $a_1, a_2, \ldots, a_n$ are scalars. If all the scalars a_i are 0, we say that it is the **trivial** linear combination. ♢

Example 2.1.2 First example of a linear combination. In $\mathbb{R}^3$, $(6, 2, -14)$ is a linear combination of $(-3, 1, 2)$ and $(-2, 0, 3)$ because

$$(6, 2, -14) = 2(-3, 1, 2) - 6(-2, 0, 3).$$

□

Example 2.1.3 Checking if a vector is a linear combination of other vectors. In $\mathbb{R}^4$, is $\mathbf{v} = (2, -1, 3, 0)$ a linear combination of

$$\mathbf{v}_1 = (1, 3, 2, 0), \mathbf{v}_2 = (5, 1, 2, 4), \text{ and } \mathbf{v}_3 = (-1, 0, 2, 1)?$$

To check this, we need to check if the equation

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3, \tag{2.1.2}$$

which is an equation in the unknowns a_1, a_2, a_3, has any solutions. Let us write out (2.1.2) explicitly:

$$(2, -1, 3, 0) = a_1(1, 3, 2, 0) + a_2(5, 1, 2, 4) + a_3(-1, 0, 2, 1) \tag{2.1.3}$$

$$\therefore (2, -1, 3, 0) = (a_1 + 5a_2 - a_3, 3a_1 + a_2, 2a_1 + 2a_2 + 2a_3, 4a_2 + a_3) \tag{2.1.4}$$

(2.1.4) is an equation between two vectors in $\mathbb{R}^4$. Two vectors in $\mathbb{R}^4$ are equal if and only if their corresponding coefficients are equal. So, (2.1.2) is equivalent to the following system of simultaneous linear equations:

$$a_1 + 5a_2 - a_3 = -2 \tag{2.1.5}$$
$$3a_1 + a_2 = -1 \tag{2.1.6}$$
$$2a_1 + 2a_2 + 2a_3 = 3 \tag{2.1.7}$$
$$4a_2 + a_3 = 0 \tag{2.1.8}$$

In other words, our question becomes: do equations (2.1.5)–(2.1.8) have a solution?

This is the kind of problem you already know how to solve by hand, from first year. We can also use SageMath to do it for us. We simply tell it what our unknown variables are, and then ask it to solve the equation. Press `Evaluate (Sage)` to see the result.

```
var('a1, a2, a3')

solve([a1 + 5*a2 - a3 == 2,
      3*a1 + a2 == -1,
      2*a1 + 2*a2 + 2*a3 ==3,
      4*a2 + a3 == 0],
      [a1, a2, a3])
```

SageMath returns an empty list `[]`. In other words, there are no solutions to equations (2.1.5)–(2.1.8). Therefore $\mathbf{v}$ cannot be expressed as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. □

Example 2.1.4 Checking if a polynomial is a linear combination of other polynomials. In Poly_2, can $p = x^2 - 1$ be expressed as a linear combination of

$$p_1 = 1 + x^2,\ p_2 = x - 3,\ p_3 = x^2 + x + 1,\ p_4 = x^2 + x - 1?$$

To check this, we need to check if the equation

$$p = a_1p_1 + a_2p_2 + a_3p_3 + a_4p_4, \tag{2.1.9}$$

which is an equation in the unknowns a_1, a_2, a_3, a_4, has any solutions. Let us write out (2.1.9) explicitly, grouping together powers of x:

$$\begin{aligned} p &= a_1p_1 + a_2p_2 + a_3p_3 + a_4p_4 \\ \therefore x^2 - 1 &= a_1(1 + x^2) + a_2(x - 3) + a_3(x^2 + x + 1) + a_4(x^2 + x - 1) \\ \therefore -1 + x^2 &= (a_1 - 3a_2 + a_3 - a_4) + (a_2 + a_3 + a_4)x + (a_1 + a_3 + a_4)x^2 \end{aligned}$$

Now, two polynomials are equal if and only if all their coefficients are equal. So, (2.1.9) is equivalent to the following system of simultaneous linear equations:

$$a_1 - 3a_2 + a_3 - a_4 = -1 \tag{2.1.10}$$
$$a_2 + a_3 + a_4 = 0 \tag{2.1.11}$$
$$a_1 + a_3 + a_4 = 1 \tag{2.1.12}$$

In other words, our question becomes: do equations (2.1.10)–(2.1.12) have a solution? We ask SageMath.

```
var('a1, a2, a3, a4')
```

```
solve([a1 - 3*a2 + a3 - a4 == -1,
       a2 + a3 + a4 == 0,
       a1 + a3 - a4 == 1],
       [a1, a2, a3, a4])
```

```
[[a1 == 2*r1 + 2/3, a2 == (2/3), a3 == -r1 + 1/3, a4 == r1]]
```

Here, `r1` and `r2` are to be interpreted as free paramters. I'm going to call them s and t instead, because that's what we usually call our free parameters! So, equations (2.1.10)–(2.1.12) have *infinitely* many solutions, parameterized by two free parameters s and t. In particular, there exists *at least one* solution. For instance, if we take $s = 2$ and $t = 1$ (a totally arbitrary choice!), we get the following solution:

$$a_1 = \frac{8}{3},\ a_2 = \frac{2}{3}, a_3 = -\frac{5}{3}, a_4 = 1 \tag{2.1.13}$$

$$\text{i.e. } p = \frac{8}{3}p_1 + \frac{2}{3}p_2 - \frac{5}{3}p_3 + p_4 \tag{2.1.14}$$

You should expand out the right hand side of (2.1.14) by hand and check that it indeed is equal to p.

We conclude that p can indeed be expressed as a linear combination of p_1, p_2, p_3 and p_4. □

Example 2.1.5 Define the functions $\mathbf{f}, \mathbf{f}_1, \mathbf{f}_2 \in \text{Diff}$ by

$$\mathbf{f}(x) = \cos^3 x, \mathbf{f}_1(x) = \cos(x), \mathbf{f}_2(x) = \cos(3x).$$

Then $\mathbf{f}$ is a linear combination of $\mathbf{f}_1$ and $\mathbf{f}_2$ because of the identity $\cos(3x) = \frac{1}{4}(3\cos x + \cos(3x))$. See Example 1.6.21. In other words,

$$\mathbf{f} = \frac{3}{4}\mathbf{f}_1 + \frac{1}{4}\mathbf{f}_2.$$

This example shows that $\mathbf{f}$ is also a trigonometric polynomial, even though its original formula $\mathbf{f}(x) = \cos(3x)$ was not in the form (1.6.7). □

Definition 2.1.6 We say that a list of vectors $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V **spans** V if every vector $\mathbf{v} \in V$ is a linear combination of the vectors from $\mathcal{B}$. ♢

Example 2.1.7 $\mathbb{R}^2$ is spanned by

$$\mathbf{e}_1 := (1,\, 0),\ \ \mathbf{e}_2 := (0, 1)$$

because every vector $\mathbf{v} = (a_1,\, a_2)$ can be written as the linear combination

$$\mathbf{v} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2.$$

□

Example 2.1.8 Checking if a list of vectors space the vector space. Is $\mathbb{R}^2$ spanned by the following list of vectors?

$$\mathbf{f}_1 := (-1, 2),\ \mathbf{f}_2 := (1, 1),\ \mathbf{f}_3 := (2, -1)$$

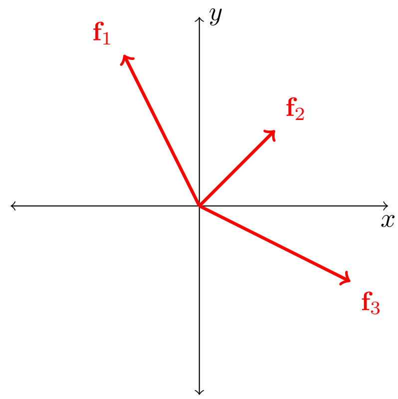


Figure 2.1.9: A list of vectors which spans $\mathbb{R}^2$.

Solution. To check this, we need check if every vector $\mathbf{v} \in V$ can be written as a linear combination of $\mathbf{f}_1, \mathbf{f}_2$ and $\mathbf{f}_3$.

So, let $\mathbf{v} = (v_1, v_2)$ be a fixed, but arbitrary, vector in $\mathbb{R}^2$. We need to check if the following equation has a solution for a_1, a_2, a_3:

$$\mathbf{v} = a_1\mathbf{f}_1 + a_2\mathbf{f}_2 + a_3\mathbf{f}_3 \tag{2.1.15}$$

Let us write this equation out explicitly:

$$\begin{aligned} \mathbf{v} &= a_1\mathbf{f}_1 + a_2\mathbf{f}_2 + a_3\mathbf{f}_3 && (2.1.16)\\ \therefore (v_1, v_2) &= a_1(-1,2) + a_2(1,1) + a_3(2,-1) && (2.1.17)\\ \therefore (v_1, v_2) &= (-a_1 + a_2 + 2a_3, 2a_1 + a_2 - a_3) && (2.1.18) \end{aligned}$$

(2.1.18) is an equation between two vectors in $\mathbb{R}^2$. Two vectors in $\mathbb{R}^2$ are equal if and only if their corresponding coefficients are equal. So, (2.1.18) is equivalent to the following system of simultaneous linear equations:

$$\begin{aligned} -a_1 + a_2 + 2a_3 &= v_1 && (2.1.19)\\ 2a_1 + a_2 - a_3 &= v_2 && (2.1.20) \end{aligned}$$

In other words, the original question

> Is $\mathbb{R}^2$ spanned by $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$?

is equivalent to the question

> Can we always solve (2.1.19)–(2.1.20) for a_1, a_2, a_3, no matter what the fixed constants $\mathbf{v}_1, \mathbf{v}_2$ are?

You already know how to solve simultaneous linear equations such as (2.1.19)–(2.1.20) by hand:

$$\begin{aligned} -a_1 + a_2 + 2a_3 &= v_1 && (2.1.21)\\ 2a_1 + a_2 - a_3 &= v_2 && (2.1.22)\\ & && (2.1.23)\\ \therefore -a_1 + a_2 + 2a_3 &= v_1 && (2.1.24) \end{aligned}$$

$$3a_2 + 3a_3 = 2v_1 + v_2 \qquad R2 \to R2 + 2R1 \tag{2.1.25}$$

$$\tag{2.1.26}$$

$$\text{Let } a_3 = t \tag{2.1.27}$$

$$\therefore a_2 = \frac{1}{3}(2v_1 + v_2) - t \tag{2.1.28}$$

$$\therefore a_1 = -\frac{1}{3}(-v_1 + v_2) + t \tag{2.1.29}$$

In other words, no matter what v_1, v_2 are, there are always infinitely many solutions (they are parameterized a free parameter t) to (2.1.19)–(2.1.20), and hence to our original equation (2.1.15). That is, we can express *any* $\mathbf{v} \in \mathbb{R}^2$ as a linear combination of the vectors $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$... and in fact there are *infinitely* many ways to do it, parameterized by a free parameter t!

For instance, suppose we try to write $\mathbf{v} = (2, 3)$ as a linear combination of $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$. If we take our general solution ((2.1.27)–(2.1.29)), and set $t = 0$, then we get

$$a_1 = \frac{1}{3},\ a_2 = \frac{7}{3},\ a_3 = 0$$
$$\text{i.e. } \mathbf{v} = \frac{1}{3}\mathbf{f}_1 + \frac{7}{3}\mathbf{f}_2$$

Or we could take, say, $t = 1$. Then our solution will be

$$a_1 = \frac{4}{3},\ a_2 = \frac{4}{3},\ a_3 = 1$$
$$\text{i.e. } \mathbf{v} = \frac{4}{3}\mathbf{f}_1 + \frac{4}{3}\mathbf{f}_2 + \mathbf{f}_3$$

There are infinitely many solutions. But the important point is that *there is always a solution to (2.1.15)*, no matter what $\mathbf{v}$ is. Therefore, the vectors $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ indeed span $\mathbb{R}^2$.

Finally, let us solve this problem using SageMath. Working by hand, we arrive at the simultanous linear equations (2.1.19)–(2.1.20), and then put it into a Sage cell:

```
var('a1, a2, a3, v1, v2')

solve([-a1 + a2 + 2*a3 == v1,
       2*a1 + a2 - a3  == v2],
       [a1, a2, a3])
```

Note that I needed to tell Sage that `v1` and `v2` are variables, and that I am asking it to solve for `a1`, `a2` and `a3`. On my computer, Sage outputs:

```
[[a1 == r1 - 1/3*v1 + 1/3*v2, a2 == -r1 + 2/3*v1 + 1/3*v2, a3 == r1]]
```

Here, `r1` is to be interpreted as our free parameter t. So Sage is giving us the same solution as we found by hand, (2.1.27)–(2.1.29). □

Example 2.1.10 $\mathbb{R}^n$ is spanned by

$$\mathbf{e}_1 := (1, 0, \ldots, 0),\ \ \mathbf{e}_2 := (0, 1, \ldots, 0),\ \ \ldots,\ \ \mathbf{e}_n := (0, 0, \ldots, 0, 1) \tag{2.1.30}$$

because every vector $\mathbf{v} = (a_1, a_2, \ldots, a_n)$ can be written as the linear combination

$$\mathbf{v} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n. \tag{2.1.31}$$

□

Checkpoint 2.1.11 Check equation (2.1.31).

Solution.

$$\mathbf{v} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n = a_1(1,\ldots,0) + \cdots a_n(0,\ldots,1) = (a_1,\ldots,0) + \cdots + (0,\ldots,a_n) = (a_1,\ldots,a_n)$$

The next theorem provides a very convenient method for checking whether a given set W spans V if one already knows that U spans V.

Theorem 2.1.12 *Let V be a vector space. Suppose $U = \{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ spans V. Furthermore, suppose that $W = \{\mathbf{w}_1,\ldots,\mathbf{w}_n\}$ has the property that for any $\mathbf{u}_i \in U$, $\mathbf{u}_i$ is a linear combination of elements in W. That is, there exist scalars $a_{i,1},\ldots,a_{i,m}$ such that*

$$\mathbf{u}_i = \sum_{k=1}^{m} a_{i,1}\mathbf{w}_k.$$

Then W also spans V.

Proof. Let $\mathbf{v} \in V$ be arbitrary. Since U spans V, there exist constants c_i such that

$$\mathbf{v} = \sum_{i=1}^{n} c_i\mathbf{u}_i.$$

By assumption,

$$\mathbf{u}_i = \sum_{k=1}^{m} a_{i,1}\mathbf{w}_k.$$

Substituting the second equation into the first, we obtain

$$\mathbf{v} = \sum_{i=1}^{n} c_i \sum_{k=1}^{m} a_{i,1}\mathbf{w}_k.$$

Using the usual rules for addition and multiplication in a vector space, we find that v is a linear combination of the w_i's. Since $v \in V$ was arbitrary, we conclude that W must also span V. ■

Exercises

1. Recall from 1st year that a function $f : \mathbb{R} \to \mathbb{R}$ is even if $f(-x) = f(x)$ and odd if $f(-x) = -f(x)$. Show that every vector in the vector space $\mathrm{Fun}(\mathbb{R})$ is a linear combination of an even function and an odd function.

Solution. The solution is relatively straightforward. Define the following two functions:

$$f_{\text{even}}(x) = \frac{1}{2}\left(f(x) + f(-x)\right), \quad f_{\text{odd}}(x) = \frac{1}{2}\left(f(x) - f(-x)\right)$$

It is easy to see that, as the names suggest, f_{even} is an even function and $f_{\text{odd}}(x)$ is an odd function. We simply sum f_{even} and f_{odd}:

$$f_{\text{even}}(x) + f_{\text{odd}}(x) = \frac{1}{2}\left(f(x) + f(-x)\right) + \frac{1}{2}\left(f(x) - f(-x)\right) = f(x).$$

2. Suppose $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4$ spans V. Prove that $\mathbf{v}_1 - \mathbf{v}_2, \mathbf{v}_2 - \mathbf{v}_3, \mathbf{v}_3 - \mathbf{v}_4, \mathbf{v}_4$ also spans V.

Solution. If we are given that $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4$ spans V then to show that any other collection of vectors in V spans V it suffices to show that each of $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4$ can be written has a linear combination of the new collection by Theorem 2.1.12. With this observation in hand, the exercise has an easy solution.

$$\begin{aligned}
\mathbf{v}_1 &= (\mathbf{v}_1 - \mathbf{v}_2) + (\mathbf{v}_2 - \mathbf{v}_3) + (\mathbf{v}_3 - \mathbf{v}_4) + \mathbf{v}_4 \\
\mathbf{v}_2 &= (\mathbf{v}_2 - \mathbf{v}_3) + (\mathbf{v}_3 - \mathbf{v}_4) + \mathbf{v}_4 \\
\mathbf{v}_3 &= (\mathbf{v}_3 - \mathbf{v}_4) + \mathbf{v}_4 \\
\mathbf{v}_4 &= \mathbf{v}_4
\end{aligned}$$

3. Consider the following polynomials in Poly_2:

$$\mathbf{r}_1(x) := 3x^2 - 2,\ \mathbf{r}_2(x) := x^2 + x,\ \mathbf{r}_3(x) := x + 1,\ \mathbf{r}_4(x) := x - 1$$

(a) Can the polynomial $\mathbf{p}$ with $\mathbf{p}(x) = x^2 + 1$ be written as a linear combination of $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, $\mathbf{r}_4$?

(b) If so, in *how many ways* can this be done?

Solution.

(a) We must set up the appropriate system of linear equations:

$$\begin{aligned}
& a\mathbf{r}_1(x) + b\mathbf{r}_2(x) + c\mathbf{r}_3(x) + d\mathbf{r}_4(x) &&= \mathbf{p}(x) \\
\implies & a(3x^2 - 2) + b(x^2 + x) + c(x + 1) + d(x - 1) &&= x^2 + 1
\end{aligned}$$

After grouping like powers of x we obtain

$$x^2(3a + b) + x(b + c + d) + (-2a + c - d) = x^2 + 1.$$

We equate coefficients on both sides of the equation to obtain the following system of linear equations:

$$\begin{aligned}
3a + b + 0c + 0d &= 1, \\
0a + 1b + 1c + 1d &= 0, \\
-2a + 0b + 1c + -1d &= 1.
\end{aligned}$$

Using your preferred method for solving a system of linear equations (such as Gauss reduction), we obtain a solution set of the form:

$$\begin{aligned}
& d \text{ is free}, \\
& a = 2 + 2d, \\
& b = -5 - 6d, \\
& c = 5 + 5d.
\end{aligned}$$

And so $\mathbf{p}(x)$ is indeed a linear combination of $\mathbf{r}_1(x), \mathbf{r}_2(x), \mathbf{r}_3(x), \mathbf{r}_4(x)$.

(b) Since d is free in the above solution set, we can write $\mathbf{p}(x)$ as a linear combination of $\mathbf{r}_1(x), \mathbf{r}_2(x), \mathbf{r}_3(x), \mathbf{r}_4(x)$ in an uncountably infinite number of ways (one for each real number!).

4. Suppose that the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\mathbf{e}_4$ span a vector space V. Show that the vectors $\mathbf{f}_1 := \mathbf{e}_2 - \mathbf{e}_1$, $\mathbf{f}_2 := \mathbf{e}_3 - \mathbf{e}_2$, $\mathbf{f}_3 := \mathbf{e}_4 - \mathbf{e}_3$, $\mathbf{f}_4 := \mathbf{e}_4$ also span V.

Solution. You could choose to show this directly or we could use a clever approach based on **2** and Theorem 2.1.12. From **2**, we know that $\mathbf{e}_1 - \mathbf{e}_2, \mathbf{e}_2 - \mathbf{e}v_3, \mathbf{e}_3 - \mathbf{e}_4, \mathbf{e}_4$ must span V. But if these vectors span V, then non-zero multiples of the vectors also span V. Thus $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4$ must span V.

5. Show that the polynomials

$$\mathbf{q}_0(x) := 1,\ \mathbf{q}_1(x) := x,\ \mathbf{q}_2(x) := 2x^2 - 1,\ \mathbf{q}_3(x) := 4x^3 - 3x$$

span Poly_3.

Solution. Once again we base our strategy on Theorem 2.1.12. Pick a spanning set for Poly_3. We'll use $1, x, x^2, x^3$, since it's the simplest. $1, x$ are certainly spanned by $\mathbf{q}_0x), \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ since $1 = \mathbf{q}_0(x)$ and $x = \mathbf{q}_1(x)$. It can easily be seen that

$$x^2 = \frac{1}{2}\mathbf{q}_2(x) + \frac{1}{2}\mathbf{q}_0(x)x^3 = \frac{1}{4}\mathbf{q}_3(x) + \frac{3}{4}\mathbf{q}_1(x),$$

completing the proof.

6. Let $\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a list of vectors in a vector space V. Suppose that $\mathcal{S}$ spans V. Suppose that w is another vector in V. Prove that the list of vectors $\mathcal{S}' = \{\mathbf{w}, \mathbf{v}_1, \ldots, \mathbf{v}_n\}$ also spans V.

7. Let $\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a list of vectors in a vector space V. Suppose that $\mathcal{S}$ spans V. Suppose that one of the vectors in the list, say $\mathbf{v}_r$, can be expressed as a linear combination of the preceding vectors:

$$\mathbf{v}_r = a_1\mathbf{v}_1 + \cdots + a_{r-1}\mathbf{v}_{r-1} \tag{2.1.32}$$

Suppose that we remove $\mathbf{v}_r$ from $\mathcal{S}$, to arrive at a new list

$$\mathcal{T} = \{\mathbf{v}_1, \ldots, \hat{\mathbf{v}}_r, \ldots, \mathbf{s}_n\}$$

Prove that $\mathcal{T}$ also spans V.

Solution. We must show that every vector $\mathbf{v} \in V$ can be written as a linear combination of the vectors from $\mathcal{T}$. So let $\mathbf{v} \in V$. Since $\mathcal{S}$ spans V, we know we can write $\mathbf{v}$ as a linear combination of the vectors from $\mathcal{S}$:

$$\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_r\mathbf{v}_r + \cdots + b_n\mathbf{v}_n \tag{2.1.33}$$

Substituting (2.1.32) into (2.1.36) gives

$$\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_r(a_1\mathbf{v}_1 + \cdots + a_{r-1}\mathbf{v}_{r-1}) + b_{r+1}\mathbf{v}_{r+1} + \cdots + b_n\mathbf{v}_n \tag{2.1.34}$$

$$= (b_1 + b_r a_1)\mathbf{v}_1 + \cdots + (b_r + b_r a_{r-1})\mathbf{v}_{r-1} + b_{r+1}\mathbf{v}_{r+1} + \cdots + b_n\mathbf{v}_n \tag{2.1.35}$$

Equation (2.1.38) shows that we can express $\mathbf{v}$ as a linear combination of the vectors from $\mathcal{T}$. Hence $\mathcal{T}$ spans V.

Solutions

· Exercises

2.1.1. Solution. The solution is relatively straightforward. Define the following two functions:

$$f_{\text{even}}(x) = \frac{1}{2}\left(f(x) + f(-x)\right), \quad f_{\text{odd}}(x) = \frac{1}{2}\left(f(x) - f(-x)\right)$$

It is easy to see that, as the names suggest, f_{even} is an even function and $f_{\text{odd}}(x)$ is an odd function. We simply sum f_{even} and f_{odd}:

$$f_{\text{even}}(x) + f_{\text{odd}}(x) = \frac{1}{2}\left(f(x) + f(-x)\right) + \frac{1}{2}\left(f(x) - f(-x)\right) = f(x).$$

2.1.2. Solution. If we are given that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ spans V then to show that any other collection of vectors in V spans V it suffices to show that each of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ can be written has a linear combination of the new collection by Theorem 2.1.12. With this observation in hand, the exercise has an easy solution.

$$\begin{aligned}
\mathbf{v}_1 &= (\mathbf{v}_1 - \mathbf{v}_2) + (\mathbf{v}_2 - \mathbf{v}_3) + (\mathbf{v}_3 - \mathbf{v}_4) + \mathbf{v}_4 \\
\mathbf{v}_2 &= (\mathbf{v}_2 - \mathbf{v}_3) + (\mathbf{v}_3 - \mathbf{v}_4) + \mathbf{v}_4 \\
\mathbf{v}_3 &= (\mathbf{v}_3 - \mathbf{v}_4) + \mathbf{v}_4 \\
\mathbf{v}_4 &= \mathbf{v}_4
\end{aligned}$$

2.1.3. Solution.

(a) We must set up the appropriate system of linear equations:

$$\begin{aligned}
& a\mathbf{r}_1(x) + b\mathbf{r}_2(x) + c\mathbf{r}_3(x) + d\mathbf{r}_4(x) && = \mathbf{p}(x) \\
\implies & a(3x^2 - 2) + b(x^2 + x) + c(x + 1) + d(x - 1) && = x^2 + 1
\end{aligned}$$

After grouping like powers of x we obtain

$$x^2(3a + b) + x(b + c + d) + (-2a + c - d) = x^2 + 1.$$

We equate coefficients on both sides of the equation to obtain the following system of linear equations:

$$\begin{aligned}
3a + b + 0c + 0d &= 1, \\
0a + 1b + 1c + 1d &= 0, \\
-2a + 0b + 1c + -1d &= 1.
\end{aligned}$$

Using your preferred method for solving a system of linear equations (such as Gauss reduction), we obtain a solution set of the form:

$$\begin{aligned}
& d \text{ is free}, \\
& a = 2 + 2d, \\
& b = -5 - 6d, \\
& c = 5 + 5d.
\end{aligned}$$

And so $\mathbf{p}(x)$ is indeed a linear combination of $\mathbf{r}_1(x), \mathbf{r}_2(x), \mathbf{r}_3(x), \mathbf{r}_4(x)$.

(b) Since d is free in the above solution set, we can write $\mathbf{p}(x)$ as a linear combination of $\mathbf{r}_1(x), \mathbf{r}_2(x), \mathbf{r}_3(x), \mathbf{r}_4(x)$ in an uncountably infinite number of ways (one for each real number!).

2.1.4. Solution. You could choose to show this directly or we could use a clever approach based on **2** and Theorem 2.1.12. From **2**, we know that $\mathbf{e}_1 - \mathbf{e}_2, \mathbf{e}_2 - \mathbf{e}v_3, \mathbf{e}_3 - \mathbf{e}_4, \mathbf{e}_4$ must span V. But if these vectors span V, then non-zero multiples of the vectors also span V. Thus $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4$ must span V.

2.1.5. Solution. Once again we base our strategy on Theorem 2.1.12. Pick a spanning set for Poly_3. We'll use $1, x, x^2, x^3$, since it's the simplest. $1, x$ are certainly spanned by $\mathbf{q}_0 x), \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ since $1 = \mathbf{q}_0(x)$ and $x = \mathbf{q}_1(x)$. It can easily be seen that

$$x^2 = \frac{1}{2}\mathbf{q}_2(x) + \frac{1}{2}\mathbf{q}_0(x)x^3 = \frac{1}{4}\mathbf{q}_3(x) + \frac{3}{4}\mathbf{q}_1(x),$$

completing the proof.

2.1.7. Solution. We must show that every vector $\mathbf{v} \in V$ can be written as a linear combination of the vectors from $\mathcal{T}$. So let $\mathbf{v} \in V$. Since $\mathcal{S}$ spans V, we know we can write $\mathbf{v}$ as a linear combination of the vectors from $\mathcal{S}$:

$$\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_r\mathbf{v}_r + \cdots + b_n\mathbf{v}_n \tag{2.1.36}$$

Substituting (2.1.32) into (2.1.36) gives

$$\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_r(a_1\mathbf{v}_1 + \cdots + a_{r-1}\mathbf{v}_{r-1}) + b_{r+1}\mathbf{v}_{r+1} + \cdots + b_n\mathbf{v}_n \tag{2.1.37}$$

$$= (b_1 + b_r a_1)\mathbf{v}_1 + \cdots + (b_r + b_r a_{r-1})\mathbf{v}_{r-1} + b_{r+1}\mathbf{v}_{r+1} + \cdots + b_n\mathbf{v}_n \tag{2.1.38}$$

Equation (2.1.38) shows that we can express $\mathbf{v}$ as a linear combination of the vectors from $\mathcal{T}$. Hence $\mathcal{T}$ spans V.

2.2 Linear independence

Definition 2.2.1 A list of vectors $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V is called **linearly independent** if the equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots k_n\mathbf{v}_n = \mathbf{0} \tag{2.2.1}$$

has only the trivial solution $k_1 = k_2 = \cdots = k_n = 0$. Otherwise (if (2.2.1) has a solution with at least one scalar $k_i \neq 0$) the list $\mathcal{B}$ is called **linearly dependent**. ♢

Remark 2.2.2 Zero vector implies linear dependence. Suppose one of the vectors $\mathbf{v}_i$ in the list $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is the zero vector $\mathbf{0}$. Then $\mathcal{B}$ is linearly dependent, since the equation (2.2.1) has the nontrivial solution

$$0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \cdots + 0\mathbf{v}_n = \mathbf{0},$$

in other words,

$$k_1 = 0, \ldots, k_{i-1} = 0, k_i = 1, k_{i+1} = 0, \ldots, k_n = 0.$$

So: *a linearly independent list ofvectors never contains the zero vector!*

Example 2.2.3 The list of vectors $\mathbf{f}_1 = (-1, 2)$, $\mathbf{f}_2 = (1, 1)$ from Example 2.1.8 is linearly independent, because the equation

$$k_1(-1, 2) + k_2(1, 1) = (0, 0)$$

is equivalent to the system of equations

$$-k_1 + k_2 = 0, \quad 2k_1 + k_2 = 0 \tag{2.2.2}$$

which has only the trivial solution $k_1 = 0$ and $k_2 = 0$. □

Checkpoint 2.2.4 Check that (2.2.2) has only the trivial solution.

Solution.

$$(2k_1 + k_2) - (-k_1 + k_2) = 0 = 3k_2 \implies k_2 = 0 \implies k_1 = 0 \text{ too.}$$

Example 2.2.5 The list of vectors $\mathbf{f}_1 = (-1, 2)$, $\mathbf{f}_2 = (1, 1)$, $\mathbf{f}_3 = (2, -1)$ from Example 2.1.8 is linearly dependent, because the equation

$$k_1(-1, 2) + k_2(1, 1) + k_3(2, -1) = (0, 0) \tag{2.2.3}$$

is equivalent to the system of equations

$$-k_1 + k_2 + 2k_3 = 0, \quad 2k_1 + k_2 - k_3 = 0 \tag{2.2.4}$$

which has a one-dimensional vector space of solutions parameterized by t,

$$k_1 = t, \; k_2 = -t, \; k_3 = t, t \in \mathbb{R}. \tag{2.2.5}$$

For instance, for $t = 2$, we have

$$2(-1, 2) - 2(1, 1) + 2(2, -1) = (0, 0)$$

so that (2.2.3) has nontrivial solutions. □

Checkpoint 2.2.6 Check that (2.2.4) has the solution set (2.2.5).

Solution. We have a system of consistent homogenous linear equations so we know there exists at least one solution, namely the trivial solution. Since we have 3 unknowns but only 2 equations, we do not have a unique solution. Let k_1 be free, i.e. $k_1 = t, t \in \mathbb{R}$. Then

$$(-k_1+k_2+2k_3)-(2k_1+k_2-k_3) = 0 = -3k_1+3k_3 \implies -k_1+k_3 = 0 \implies k_3 = t. -t+k_2+2t = 0 \implies k_2 = -t$$

Example 2.2.7 The list of polynomials

$$\mathbf{q}_0(x) := 1, \; \mathbf{q}_1(x) := x, \; \mathbf{q}_2(x) := 2x^2 - 1, \; \mathbf{q}_3(x) := 4x^3 - 3x$$

from Example 2.1.5 is linearly independent in Poly_3. This is because the equation

$$k_0\mathbf{q}_0 + k_1\mathbf{q}_1 + k_2\mathbf{q}_2 + k_3\mathbf{q}_3 = \mathbf{0}$$

becomes the following equation between polynomials:

$$4k_3x^3 + 2k_2x^2 + (-3k_3 + k_1)x + (-k_2 + k_0) = 0$$

This is equivalent to the following system of equations,

$$4k_3 = 0, \quad 2k_2 = 0, \quad -3k_3 + k_1 = 0, \quad k_0 - k_2 = 0$$

which has only the trivial solution $k_0 = k_1 = k_2 = k_3 = 0$. □

Here are two more ways to think about linearly dependent lists of vectors.

Proposition 2.2.8 Equivalent Formulations of Linear Dependence. *Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a list of vectors in a vector space V. The following statements are equivalent:*

1. *The list of vectors $\mathcal{B}$ is linearly dependent.*
2. *(Linear Combination of Other Vectors)One of the vectors in the list $\mathcal{B}$ is a linear combination of the other vectors in $\mathcal{B}$.*

3. *(Linear Combination of Preceding Vectors)Either* $\mathbf{v}_1 = \mathbf{0}$*, or for some* $r \in \{2, 3, \ldots, n\}$*,* $\mathbf{v}_r$ *is a linear combination of* $\mathbf{v}_1$*,* $\mathbf{v}_2$*, …,* $\mathbf{v}_{r-1}$*.*

Proof. We will show that (1) $\Leftrightarrow$ (2), (1) $\Rightarrow$ (3) and (3) $\Rightarrow$ (2), and conclude that each statement implies the others.

(1) $\Rightarrow$ (2). Suppose that $\mathcal{B}$ is linearly dependent. This means that there are scalars k_1, k_2, …, k_n, not all zero, such that

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n = \mathbf{0}. \tag{2.2.6}$$

Let k_s be one of the nonzero coefficients. Then, by taking the other vectors to the other side of the equation, and multuplying by $\frac{1}{k_s}$ we can solve for $\mathbf{v}_s$ in terms of the other vectors:

$$\mathbf{v}_s = -\frac{k_1}{k_s}\mathbf{v}_1 - \ldots - \frac{k_n}{k_s}\mathbf{v}_n \qquad\qquad \text{(No } \mathbf{v}_i \text{ terms on RHS)}$$

Therefore, (2) is true.

(2) $\Rightarrow$ (1). Suppose that one of the vectors in the list, say $\mathbf{v}_s$, is a linear combination of the others vectors. That is,

$$\mathbf{v}_s = k_1\mathbf{v}_1 + \ldots + k_n\mathbf{v}_n \qquad\qquad \text{(No term on RHS.)}$$

Rearranging this equation gives:

$$k_1\mathbf{v}_1 + \ldots + (-1)\mathbf{v}_s + \ldots + k_n\mathbf{v}_n = \mathbf{0}. \tag{2.2.7}$$

Not all the coefficients on the LHS of (2.2.7) are zero, since the coefficient of $\mathbf{v}_s$ is equal to -1. Therefore, $\mathcal{B}$ is linearly dependent.

(1) $\Rightarrow$ (3). Suppose that the list $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent. This means that there are scalars k_1, k_2, …, k_n, not all zero, such that

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n = \mathbf{0}. \tag{2.2.8}$$

Let $r \in \{1, 2, \ldots, n\}$ be the largest index such that $k_r \neq 0$. (We are told that not all the k_i are zero, so this makes sense.) If $r = 1$, then (2.2.8) is simply the equation

$$k_1\mathbf{v}_1 = \mathbf{0}, \text{where } k_1 \neq 0.$$

Therefore $\mathbf{v}_1 = \mathbf{0}$ by Lemma 1.5.6, and we are done. On the other hand, suppose $r \neq 1$. Then (2.2.8) becomes the equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r = \mathbf{0}, \text{where } k_r \neq 0.$$

By dividing by k_r, we can now solve for $\mathbf{v}_r$ in terms of the preceding vectors $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_{r-1}$:

$$\therefore \mathbf{v}_r = -\frac{k_1}{k_r}\mathbf{v}_1 - \frac{k_2}{k_r}\mathbf{v}_2 - \cdots - \frac{k_{r-1}}{k_r}\mathbf{v}_{r-1}$$

Therefore, (3) is true.

(3) $\Rightarrow$ (2) Suppose that (3) is true. In other words, either:

- $\mathbf{v}_1 = \mathbf{0}$. Therefore, $\mathcal{B}$ is linearly dependent, by Remark 2.2.2. In other words, (1) is true. Therefore, since we have already proved that (1) $\Rightarrow$ (2), we conclude that (2) is true.

- For some $r \in \{2, \ldots, n\}$, $\mathbf{v}_r$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_{r-1}$. In this case, clearly $\mathbf{v}_r$ is a linear combination of the other vectors in $\mathcal{B}$, so (2) is true.

In both cases, (2) is true. So, (3) $\Rightarrow$ (2). ■

Example 2.2.9 We saw in Example 2.2.5, using the definition of linear dependence, that the list of vectors $\mathbf{f}_1 = (-1, 2)$, $\mathbf{f}_2 = (1, 1)$, $\mathbf{f}_3 = (2, -1)$ in $\mathbb{R}^3$ is linearly dependent. Give two alternative proofs of this, using Proposition 2.2.8.

Solution 1. We check Item 2 of Proposition 2.2.8. That is, we check if one of the vectors in the list is a linear combination of the other vectors. Indeed, we observe by inspection that

$$\mathbf{f}_2 = \mathbf{f}_1 + \mathbf{f}_3. \tag{2.2.9}$$

Hence, $\mathcal{B}$ is linearly dependent.

Solution 2. We check Item 3 of Proposition 2.2.8. That is, we check:

- Is $\mathbf{f}_1 = \mathbf{0}$? No.
- Is $\mathbf{f}_2$ is a scalar multiple of $\mathbf{f}_1$? No.
- Is $\mathbf{f}_3$ is a linear combination of $\mathbf{f}_1$ and $\mathbf{f}_2$? Yes, since

  $$\mathbf{f}_3 = -\mathbf{f}_1 + \mathbf{f}_2.$$

  Hence, $\mathcal{B}$ is linearly dependent.

□

Proposition 2.2.10 Bumping Off Proposition. *Suppose $\mathcal{L} = \{\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_m\}$ is a linearly independent list of vectors in a vector space V, and that $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\}$ spans V. Then $m \leq n$.*

Proof. Start with the original spanning list of vectors

$$\mathcal{S} = \{\ \mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\} \tag{2.2.10}$$

and consider the 'bloated' list

$$\mathcal{S}' = \{\mathbf{l}_1, \mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_n\} \tag{2.2.11}$$

Now, since $\mathcal{S}$ spans V, we know that $\mathbf{l}_1$ can be written as a linear combination of the vectors $\mathbf{s}_1, \dots, \mathbf{s}_n$. Therefore, by Item 2 of Proposition 2.2.8, we know that $\mathcal{S}'$ is linearly dependent. Thus, by Item 3 of Proposition 2.2.8, either:

- $\mathbf{l}_1 = \mathbf{0}$. This cannot be true, since then $\mathcal{L}$ would be linearly dependent by Remark 2.2.2, contradicting our initial assumption.
- or one of the $\mathbf{s}$-vectors, say $\mathbf{s}_r$, can be expressed as a linear combination of the preceding vectors. We can then remove $\mathbf{s}_r$ from the list $\mathcal{S}'$ ('bump it off'), and the resulting list

  $$\mathcal{S}_1 := \{\mathbf{l}_1, \mathbf{s}_1, \mathbf{s}_2, \cdots, \hat{\mathbf{s}_r}, \cdots, \mathbf{s}_n\} \qquad (\mathbf{s}_r \text{ omitted}) \tag{2.2.12}$$

  will still span V, by Exercise 2.1.7.

We can go on in this way, each time transferring another one of the $\mathbf{l}$-vectors into the list, and removing another one of the $\mathbf{s}$-vectors, and still have a list which spans V:

$$\begin{aligned} \mathcal{L} &= \{\mathbf{l}_1, \dots, \mathbf{l}_m\} & \qquad \mathcal{S} &= \{\mathbf{s}_1, \dots, \mathbf{s}_n\} \\ \mathcal{L}_1 &= \{\mathbf{l}_2, \dots, \mathbf{l}_m\} & \qquad \mathcal{S}_1 &= \{\mathbf{l}_1, \underbrace{\mathbf{s}_1, \dots, \mathbf{s}_n}_{n-1}\} \end{aligned}$$

$$\mathcal{L}_2 = \{\mathbf{l}_3, \ldots, \mathbf{l}_m\} \qquad\qquad \mathcal{S}_2 = \{\mathbf{l}_2, \mathbf{l}_1, \underbrace{\mathbf{s}_1, \ldots, \mathbf{s}_n}_{n-2}\}$$

$$\vdots \qquad\qquad\qquad \vdots$$

Now, suppose that $m > n$. When we reach the nth stage of this process, we will have $\mathcal{S}_n = \{\mathbf{l}_n, \ldots, \mathbf{l}_1\}$, and it will span V. Therefore, in particular, $\mathbf{l}_{n+1}$ (which we know exists, since $m > n$) will be a linear combination of $\mathbf{l}_1, \ldots, \mathbf{l}_n$. But then, by Item 2 of Proposition 2.2.8, we conclude that $\mathcal{L}$ is linearly dependent. But we were told in the beginning that $\mathcal{L}$ is linearly *in*dependent, so we have a contradiction. Hence, our assumption that $m > n$ must be false. Therefore, we must have $m \leq n$. ∎

Exercises

1. Show that the list of vectors (2, 3, 1), (1, −1, 2), (7, 3, c) is linearly dependent in $\mathbb{R}^3$ if and only if $c = 8$.

Solution. We set up a linear equation and find the necessary conditions on c. Suppose some linear combination of the vectors equals 0:

$$k_1(2,\, 3,\, 1) + k_2(1,\, -1,\, 2) + k_3(7,\, 3,\, c) = \mathbf{0} = (0, 0, 0)$$

This vector equation gives rise to a system of 3 linear equations:

$$2k_1 + k_2 + 7k_3 = 0, 3k_1 - k_2 + 3k_3 = 0, k_1 + 2k_2 + ck_3 = 0.$$

The corresponding matrix equation is

$$\begin{bmatrix} 2 & 1 & 7 \\ 3 & -1 & 3 \\ 1 & 2 & c \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This matrix is non-invertible if and only if its determinant is 0. Furthermore, the matrix being non-invertible will mean we can find a non-trivial solution to the intial equation. We compute the determinant:s

$$\det \begin{bmatrix} 2 & 1 & 7 \\ 3 & -1 & 3 \\ 1 & 2 & c \end{bmatrix} = -5c + 40$$

which is 0 if and only if $c = 8$.

2. The list of vectors in $\mathrm{Mat}_{2,2}$ given by

$$\mathbf{v}_1 = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix}, \mathbf{v}_5 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is linearly dependent (you will prove this in Exercise 2.3.6.4, but for the sake of this question you may assume it to be true). Go through the same steps as in Example 2.2.9 to find the first vector in the list which is either the zero vector or a linear combination of the preceding vectors.

Solution. Firstly note that $\mathbf{v}_1$ is non-zero, so we consider $\mathbf{v}_2$. $\mathbf{v}_2$ cannnot be a scalar multiple of $\mathbf{v}_1$ by considering the matrix entry in position (1,2). We now consider $\mathbf{v}_3$. Suppose

$$a \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} + b \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$$

This gives rise to a system of four linear equations. In particular, we have the equation for the matrix entry in position (1,2):

$$2a + 0b = 0$$

And hence $a = 0$. But clearly

$$b\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$$

for any choice of b. Hence $\mathbf{v}_3$ is not a scalar multiple of $\mathbf{v}_1$ and $\mathbf{v}_2$. We consider $\mathbf{v}_4$ next. Suppose

$$a\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} + b\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} + c\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix}$$

The equation for the entry in position (1,2) is simply

$$2a = 3$$

and so $a = \frac{3}{2}$. The corresponding equation for the entry in position (1,1) is thus

$$\frac{3}{2} + b + c = 0.$$

Using this result, we consider the equation for the entry in position (2,2) and compute:

$$\frac{3}{2} + b + 3c = -1 \implies \frac{3}{2} + b + c + 2c = -1 \implies 2c = -1 \implies c = -\frac{1}{2}$$

and so $b = -1$. SHOW THAT THIS IS INCONSISTENT WITH (2,1).

3. Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a linearly independent list of vectors in a vector space V. Suppose that $\mathbf{v}$ is a vector in V which cannot be written as a linear combination of the vectors in $\mathcal{B}$. Show that the list $\mathcal{B}' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}\}$ is still linearly independent. (Hint: Use the Linear Combination of Preceding Vectors Proposition.)

4. Consider the vector space of functions on the closed unit interval, $\mathrm{Fun}([0,1])$. Show that for any $n \in \mathbb{N}$, we can find n linear independent vectors in $\mathrm{Fun}([0,1])$.

5. (Bonus) Try adapt the argument in the question above to show that for any $n \in \mathbb{N}$, we can find n linear independent vectors in $\mathrm{Cont}([0,1])$, the vector space of all continuous real valued functions on $[0,1]$.

Solutions

· Exercises

2.2.1. Solution. We set up a linear equation and find the necessary conditions on c. Suppose some linear combination of the vectors equals 0:

$$k_1(2,\, 3,\, 1) + k_2(1,\, -1,\, 2) + k_3(7,\, 3,\, c) = \mathbf{0} = (0, 0, 0)$$

This vector equation gives rise to a system of 3 linear equations:

$$2k_1 + k_2 + 7k_3 = 0, 3k_1 - k_2 + 3k_3 = 0, k_1 + 2k_2 + ck_3 = 0.$$

The corresponding matrix equation is

$$\begin{bmatrix} 2 & 1 & 7 \\ 3 & -1 & 3 \\ 1 & 2 & c \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This matrix is non-invertible if and only if its determinant is 0. Furthermore, the matrix being non-invertible will mean we can find a non-trivial solution to the intial equation. We compute the determinant:s

$$\det \begin{bmatrix} 2 & 1 & 7 \\ 3 & -1 & 3 \\ 1 & 2 & c \end{bmatrix} = -5c + 40$$

which is 0 if and only if $c = 8$.

2.2.2. Solution. Firstly note that $\mathbf{v}_1$ is non-zero, so we consider $\mathbf{v}_2$. $\mathbf{v}_2$ cannnot be a scalar multiple of $\mathbf{v}_1$ by considering the matrix entry in position (1,2). We now consider $\mathbf{v}_3$. Suppose

$$a \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} + b \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$$

This gives rise to a system of four linear equations. In particular, we have the equation for the matrix entry in position (1,2):

$$2a + 0b = 0$$

And hence $a = 0$. But clearly

$$b \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$$

for any choice of b. Hence $\mathbf{v}_3$ is not a scalar multiple of $\mathbf{v}_1$ and $\mathbf{v}_2$. We consider $\mathbf{v}_4$ next. Suppose

$$a \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} + b \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} + c \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix}$$

The equation for the entry in position (1,2) is simply

$$2a = 3$$

and so $a = \frac{3}{2}$. The corresponding equation for the entry in position (1,1) is thus

$$\frac{3}{2} + b + c = 0.$$

Using this result, we consider the equation for the entry in position (2,2) and compute:

$$\frac{3}{2} + b + 3c = -1 \implies \frac{3}{2} + b + c + 2c = -1 \implies 2c = -1 \implies c = -\frac{1}{2}$$

and so $b = -1$. SHOW THAT THIS IS INCONSISTENT WITH (2,1).

2.3 Basis and dimension

In this section we introduce the notions of:

- a *basis* of a vector space, and
- the *dimension* of a vector space.

Then we compute the dimensions of the vector spaces we have introduced up to now. We end by explaining the *sifting algorithm* which allows us to prove some useful results concerning bases and dimension.

Definition 2.3.1 A list of vectors $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ in a vector space V is called a **basis** for V if it is linearly independent and spans V. $\Diamond$

Theorem 2.3.2 Invariance of dimension. *If $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n\}$ are bases of a vector space V, then $m = n$.*

Proof. This is a consequence of Proposition 2.2.10 (the Bumping Off Proposition). Since the **b**-vectors are linearly independent and the **c**-vectors span V, we have $m \leq n$. On the other hand, since the **c**-vectors are linearly independent and the **b**-vectors span V, we have $n \leq m$. Hence $m = n$. ■

Definition 2.3.3 A vector space V is **finite-dimensional** if there exists a basis $\mathcal{B}$ for V. In that case, the **dimension** of V is the number of vectors in the basis $\mathcal{B}$. A vector space is **infinte-dimensional** if it is not finite-dimensional. $\Diamond$

Note that the concept of 'dimension of a vector space' is only well-defined because of Theorem 2.3.2.

The case of the zero vector space $Z = \{\mathbf{0}\}$ is not explicitly handled in Definition 2.3.3 and we treat it as a special case. Namely, we *define* the dimension of the zero vector space Z to be 0. So, by definition, Z is finite-dimensional, and its dimension equals 0.

2.3.1 Dimensions of some familiar vector spaces

Example 2.3.4 Standard basis for $\mathbb{R}^n$. The list of vectors

$$\mathbf{e}_1 := (1, 0, \dots, 0), \quad \mathbf{e}_2 := (0, 1, \dots, 0), \quad \dots, \quad \mathbf{e}_n := (0, 0, \dots, 0, 1)$$

is a basis for $\mathbb{R}^n$. We already saw in Example 2.1.10 that this list spans $\mathbb{R}^n$. We need to check that it is linearly independent. So, suppose that

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

Expanding out the left hand side in components using the definition of the standard basis vectors $\mathbf{e}_i$, this becomes the equation

$$(a_1, 0, 0, \dots, 0) + (0, a_2, 0, \dots, 0) + \cdots + (0, 0, 0, \dots, a_n) = (0, 0, 0, \dots, 0).$$

In other words, we have

$$(a_1, a_2, a_3, \dots, a_n) = (0, 0, 0, \dots, 0)$$

which says precisely that $a_1 = a_2 = a_3 = \cdots = a_n = 0$, which is what we needed to prove. Thus the list of vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is linearly independent, and is hence a basis for $\mathbb{R}^n$. So $\mathbb{R}^n$ has dimension n. □

Example 2.3.5 A basis for $\mathbb{R}^4$. Check whether the following list of vectors

$$\mathbf{v}_1 = (1,0,2,-3),\ \mathbf{v}_2 = (1,3,-1,2),\ \mathbf{v}_3 = (0,1,2,-1), \mathbf{v}_4 = (1,2,3,4) \quad (2.3.1)$$

is a basis for $\mathbb{R}^4$.

Solution. First we check if the list of vectors is linearly independent. Consider the equation

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 + a_4\mathbf{v}_4 = \mathbf{0} \quad (2.3.2)$$

$$\therefore a_1(1,0,2,-3) + a_2(1,3,-1,2) + a_3(0,1,2,-1) + a_4(1,2,3,4) = (0,0,0,0) \quad (2.3.3)$$

$$\therefore (a_1 - a_2 + a_4, 3a_2 + a_3 + 2a_4, 2a_1 - a_2 + 2a_3 + 3a_4, -3a_1 + 2a_2 - a_3 + 4a_4) = (0,0,0,0) \quad (2.3.4)$$

So the list of vectors is linearly indepedent if and only if the following equations have only the trivial solution $a_1 = 0, a_2 = 0, a_3 = 0, a_4 = 0$:

$$a_1 - a_2 + a_4 = 0 \quad (2.3.5)$$

$$3a_1 + a_3 + 2a_4 = 0 \quad (2.3.6)$$

$$2a_1 - a_2 + 2a_3 + 3a_4 = 0 \quad (2.3.7)$$

$$-3a_1 + 2a_2 - a_3 + 4a_4 = 0 \quad (2.3.8)$$

We can compute the solutions to equations (2.3.5)–(2.3.8) by hand, or using SageMath.

```
var('a1, a2, a3, a4')

solve([a1 - a2 + a4 == 0,
      3*a1 + a3 + 2*a4 == 0,
      2*a1 - a2 + 2*a3 + 3*a4 == 0,
      -3*a1 + 2*a2 - a3 + 4*a4 == 0],
      [a1, a2, a3, a4])
```

SageMath outputs:

```
[[a1 == 0, a2 == 0, a3 == 0, a4 == 0]]
```

So indeed, equations (2.3.5)–(2.3.8) have only the trivial solution. Therefore the list of vectors is linearly independent.

Next, we need to check that the list of vectors spans $\mathbb{R}^4$. (There is a shorter way of doing this, using Corollary 2.3.27 below, but for now we prove it from first principles.) So, let $\mathbf{w} = (w_1, w_2, w_3, w_4)$ be an arbitrary vector in $\mathbb{R}^4$. We need to show that there exists at least one way to express $\mathbf{w}$ as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$. In other words, we need to check if there exists at least one solution to the following equation:

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 + a_4\mathbf{v}_4 = \mathbf{w} \quad (2.3.9)$$

$$\therefore a_1(1,0,2,-3) + a_2(1,3,-1,2) + a_3(0,1,2,-1) + a_4(1,2,3,4) = (w_1,w_2,w_3,w_4) \quad (2.3.10)$$

$$\therefore (a_1 - a_2 + a_4, 3a_2 + a_3 + 2a_4, 2a_1 - a_2 + 2a_3 + 3a_4, -3a_1 + 2a_2 - a_3 + 4a_4) = (w_1,w_2,w_3,w_4) \quad (2.3.11)$$

So the list of vectors spans $\mathbb{R}^4$ if and only if the following equations for a_1, a_2, a_3, a_4 always have a solution, no matter what the values of w_1, w_2, w_3, w_4 are:

$$a_1 - a_2 + a_4 = w_1 \tag{2.3.12}$$
$$3a_1 + a_3 + 2a_4 = w_2 \tag{2.3.13}$$
$$2a_1 - a_2 + 2a_3 + 3a_4 = w_3 \tag{2.3.14}$$
$$-3a_1 + 2a_2 - a_3 + 4a_4 = w_4 \tag{2.3.15}$$

We can compute the solutions to equations (2.3.12)–(2.3.15) by hand, or using SageMath:

```
var('a1, a2, a3, a4, w1, w2, w3, w4')

solve([a1 - a2 + a4 == w1,
       3*a1 + a3 + 2*a4 == w2,
       2*a1 - a2 + 2*a3 + 3*a4 == w3,
       -3*a1 + 2*a2 - a3 + 4*a4 == w4],
       [a1, a2, a3, a4])
```

Note that we ask SageMath to solve for `a1, a2, a3, a4`, since `w1, w2, w3, w4` are regarded as constants in the equation... we are not trying to solve for them, they are fixed, but arbitrary! SageMath outputs:

```
[[a1 == 1/9*w1 + 7/18*w2 - 2/9*w3 - 1/18*w4, a2 == -2/3*w1 + 5/12*w2
- 1/6*w3 + 1/12*w4, a3 == -7/9*w1 - 2/9*w2 + 5/9*w3 - 1/9*w4, a4 ==
2/9*w1 + 1/36*w2 + 1/18*w3 + 5/36*w4]]
```

In other words, there does indeed exist a solution, no matter what (w_1, w_2, w_3, w_4) is. For instance, if $(w_1, w_2, w_3, w_4) = (3, 1, 2, 4)$, then the solution is

$$a_1 = \frac{1}{18}, a_2 = -\frac{19}{12}, a_3 = -\frac{17}{9}, a_4 = \frac{49}{36}.$$

In other words,

$$(3, 1, 2, 4) = \frac{1}{18}\mathbf{v}_1 - \frac{19}{12}\mathbf{v}_2 - \frac{17}{9}\mathbf{v}_3 + \frac{49}{36}\mathbf{v}_4.$$

Since there exists a solution to equation (2.3.9) for each vector $\mathbf{w} \in \mathbb{R}^4$, we conclude that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ spans $\mathbb{R}^4$.

Hence $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is a basis for $\mathbb{R}^4$, since it is linearly independent and spans $\mathbb{R}^4$. □

Example 2.3.6 The list of polynomials

$$\mathbf{p}_0(x) := 1,\ \mathbf{p}_1(x) := x,\ \mathbf{p}_2(x) := x^2,\ \ldots,\ \mathbf{p}_n(x) := x^n$$

is a basis for Poly_n, so $\text{Dim}\,\text{Poly}_n = n + 1$. Indeed, this list spans Poly_n by definition, so we just need to check that it is linearly independent. Suppose that

$$a_0\mathbf{p}_0 + a_1\mathbf{p}_1 + a_2\mathbf{p}_2 + \cdots + a_n\mathbf{p}_n = \mathbf{0}.$$

This is an equation between functions, so it holds for all $x \in \mathbb{R}$! In other words, for all $x \in \mathbb{R}$, the following equation holds:

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0 \tag{2.3.16}$$

Think about this carefully. Equation (2.3.16) represents an *infinite* family of equations for the unknowns $a_0, a_1, \ldots, a_n$. There is one equation for each value of $x \in \mathbb{R}$. For example:

$$(x = 1) \qquad a_0 + a_1 + a_2 + \cdots + a_n = 0 \tag{2.3.17}$$
$$(x = -1) \qquad a_0 - a_1 + a_2 + \cdots + (-1)^n a_n = 0 \tag{2.3.18}$$
$$(x = 2) \qquad a_0 + 2a_1 + 4a_3 + \cdots + 2^n a_n = 0 \tag{2.3.19}$$
$$(x = 3) \qquad a_0 + 3a_1 + 9a_3 + \cdots + 3^n a_n = 0 \tag{2.3.20}$$
$$\vdots \tag{2.3.21}$$

Suppose we find values for $a_0, a_1, a_2, \ldots, a_n$ which solve *all* these infinitely many equations (2.3.17)–(2.3.21). We can now change our point of view. Namely, substitute these fixed values for $a_0, a_1, \ldots, a_n$ into Equation (2.3.16) and regard Equation (2.3.16) as an equation for the unknown x (the coefficients $a_0, a_1, \ldots, a_n$ are now *fixed*.) We conclude that *every* $x \in \mathbb{R}$ is a root of this polynomial equation!

But, we know from algebra that a polynomial equation of the form (2.3.16) with nonzero coefficients has *at most* n roots $x_1, x_2, \ldots, x_n$. So, in order for (2.3.16) to hold for *all* real numbers x, the coefficients must be zero, i.e. $a_0 = a_1 = a_2 = \cdots = a_n = 0$, which is what we needed to show. □

Example 2.3.7 Suppose X is a finite set. Then $\mathrm{Fun}(X)$ is finite-dimensional, with dimension $|X|$, with basis given by the functions f_a, $a \in X$, defined by:

$$f_a(x) := \begin{cases} 1 & \text{if } x = a \\ 0 & \text{otherwise} \end{cases} \tag{2.3.22}$$

We will prove this in a series of exercises.

The formula on the right hand side of (2.3.22) occurs so often in mathematics we give it a symbol of its own, δ_{ab} (the 'Kronecker delta'). This symbol stands for the formula: "If $a = b$, return a 1. If $a \neq b$, return a 0". In this language, we can rewrite the definition of the functions f_a as

$$f_a(x) := \delta_{ax}. \tag{2.3.23}$$

□

Checkpoint 2.3.8 Suppose $X = \{a, b, c\}$.

1. Evaluate the function f_b at each $x \in X$.

2. Show that $\{f_a, f_b, f_c\}$ is a basis for $\mathrm{Fun}(X)$.

Checkpoint 2.3.9 Now let X be an arbitrary finite set. Consider the collection of functions

$$\mathcal{B} = \{\mathbf{f}_a : a \in X\}$$

Show that $\mathcal{B}$ is a basis for $\mathrm{Fun}(X)$.

Example 2.3.10 Trig_n is $(2n+1)$-dimensional, with basis

$$\mathbf{T}_0(x) := 1, \ \ \mathbf{T}_1(x) := \cos x, \ \mathbf{T}_2(x) := \sin x, \ \mathbf{T}_3(x) := \cos 2x,$$
$$\mathbf{T}_4(x) := \sin 2x, \ \ldots, \ \mathbf{T}_{2n-1}(x) := \cos nx, \ \mathbf{T}_{2n}(x) := \sin nx.$$

You know that these functions span Trig_n, by definition. They are also linearly independent, though we will not prove this. □

Example 2.3.11 The dimension of $\mathrm{Mat}_{n,m}$ is nm, with basis given by the matrices

$$\mathsf{E}_{ij}, i = 1 \ldots n,\ j = 1 \ldots m$$

which have a 1 in the ith row and jth column and zeroes everywhere else.

> Usually A is a matrix, and A_{ij} is the element of the matrix at position (i, j). But now E_{ij} is a matrix in its own right! Its element at position (k, l) will be written as $(\mathsf{E}_{ij})_{kl}$. I hope you don't find this too confusing. In fact, we can write down an elegant formula for the elements of E_{ij} using the Kronecker delta symbol:
>
> $$(\mathsf{E}_{ij})_{kl} = \delta_{ik}\delta_{jl} \tag{2.3.24}$$

□

Checkpoint 2.3.12 Check that (2.3.24) is indeed the correct formula for the matrix elements of E_{ij}.

Example 2.3.13 The standard basis of $\mathrm{Mat}_{2,2}$ is

$$\mathsf{E}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathsf{E}_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathsf{E}_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \mathsf{E}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

□

Example 2.3.14 The standard basis of Col_n is

$$\mathbf{e}_1 := \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 := \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

□

Example 2.3.15 Dimension of a hyperplane. Let $\mathbf{v} \in \mathbb{R}^n$ be a fixed vector, and consider the hyperplane $W \subset \mathbb{R}^n$ orthogonal to $\mathbf{v}$ as in Example 1.6.13. Then you will prove in Exercise Checkpoint 2.3.17 that $\mathrm{Dim}(W) = n - 1$.

For instance, consider the specific example from Example 1.6.13, namely the plane $W \subset \mathbb{R}^3$ of vectors orthogonal to $\mathbf{v} = (1, 2, 3)$. In other words,

$$W = \{(w_1, w_2, w_3) \in \mathbb{R}^3 : w_1 + 2w_2 + 3w_3 = 0\}. \tag{2.3.25}$$

There is no 'standard' basis for W. But, here is one basis (as good as any other):

$$\mathbf{a} = (1, 0, -\frac{1}{3}), \quad \mathbf{b} = (0, 1, -\frac{2}{3}). \tag{2.3.26}$$

You will show this is indeed a basis for W in Checkpoint 2.3.16 below. I computed these vectors as follows. To obtain $\mathbf{a}$, I simply set $w_1 = 1, w_2 = 0$ and then solved for w_3 using Equation (2.3.25). Similarly, for $\mathbf{b}$, I simply set $w_1 = 0, w_2 = 1$ and then solved for w_3 using (2.3.25).

There is nothing special about my method above for computing a basis for W. Here is a different basis for W, which I arrived at by choosing random values of w_1 and w_2 and then calculating what w_3 must be in order to satisfy Equation (2.3.25):

$$\mathbf{u} = (1, 2, -\frac{5}{3}), \quad \mathbf{v} = (-4, 2, 0). \tag{2.3.27}$$

In any event, we see that $\mathrm{Dim}(W) = 2$. □

Checkpoint 2.3.16 Show that the list of vectors $\{\mathbf{a}, \mathbf{b}\}$ from (2.3.26) in Example 2.3.15 is a basis for W.

Checkpoint 2.3.17 Let $\mathbf{v} \in \mathbb{R}^n$ be a fixed vector, and set

$$W := \{\mathbf{w} \in \mathbb{R}^n : \mathbf{v} \cdot \mathbf{w} = 0\}$$

Prove that $\mathrm{Dim}(W) = n - 1$.

Hint. Find a basis for the solution space of the equation determining W.

2.3.2 Dimension of space of solutions to a homogenous linear differential equation

We will now compute the dimension of the vector space of solutions to a homogenous linear ordinary differential equation. We will need the following theorem from the theory of differential equations, which we won't prove.

Theorem 2.3.18 Existence and uniqueness of solutions to linear ODE's.

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y^{(1)} + a_0(x)y == 0 \tag{2.3.28}$$

$Ia_0(x), \cdots, a_{n-1}(x)I$

$$y(x_0) = c_0 \tag{2.3.29}$$
$$y^{(1)}(x_0) = c_1 \tag{2.3.30}$$
$$\vdots \tag{2.3.31}$$
$$y^{(n-1)}(x_0) = c_{n-1} \tag{2.3.32}$$

$x_0 \in Ic_0, \ldots, c_{n-1}y(x)I$ *(2.3.28)(2.3.29)–(2.3.32)*

Example 2.3.19 Application of existence and uniqueness theorem of solutions to ODE. Consider the ODE

$$x^2y^{''} - 3xy^{'} + 5y = 0 \quad \text{on } (0, \infty) \tag{2.3.33}$$

from Example 1.6.24. To apply Theorem 2.3.18, we first rewrite it in the form

$$y^{''} - \frac{3}{x}y' + \frac{5}{x^2}y = 0 \quad \text{on } (0, \infty). \tag{2.3.34}$$

The coefficient functions $\frac{3}{x}$ and $\frac{5}{x^2}$ are continuous on $(0, \infty)$ so we can apply Theorem 2.3.18. Choose, say, $x_0 = 1$ and two arbitrary numbers c_0, c_1. Then Theorem 2.3.18 says that there exists a unique solution $y(x)$ to the differential equation (2.3.34) satisfying the initial conditions:

$$y(1) = c_0 \tag{2.3.35}$$
$$y'(1) = c_1 \tag{2.3.36}$$

Let us verify this in SageMath. First, we ask SageMath to find the most general solution to the differential equation (2.3.34):

```
x = var('x')
y = function('y')(x)
```

```
ode = diff(y,x,2) - 3/x * diff(y,x,1) + 5/x^2 * y == 0

show(desolve(ode, y))
```

SageMath tells us that the most general solution to the differential equation (2.3.34) is

$$y = K_1 x^2 \sin(\log(x)) + K_2 x^2 \cos(\log(x)). \tag{2.3.37}$$

Let us now apply the initial conditions (2.3.35)–(2.3.36). We can compute $y(1)$ and $y'(1)$ using the formula for y from (2.3.37). So (2.3.35)–(2.3.36) becomes (check this):

$$K_2 = c_0 \tag{2.3.38}$$
$$K_1 + 2K_2 = c_1 \tag{2.3.39}$$

Equations (2.3.38)–(2.3.39) have a unique solution, namely $K_1 = c_1 - 2c_0, K_2 = c_0$. So indeed, for any initial conditions (2.3.35)–(2.3.36), the differential equation (2.3.34) has a unique solution, namely:

$$y = (c_1 - 2c_0)x^2 \sin(\log(x)) + c_0 x^2 \cos(\log(x))$$

For instance, if our initial conditions were

$$y(1) = 1 \tag{2.3.40}$$
$$y'(1) = 0 \tag{2.3.41}$$

then the unique solution is

$$y = -2x^2 \sin(\log(x)) + x^2 \cos(\log(x)). \tag{2.3.42}$$

You can also check this explicitly in SageMath, using the `ics=[1,1,0]` option of `desolve` (the first number is the value of x_0, the second number is the value of $y(x_0)$, and the third number is the value of y'(x_0), etc.):

```
x = var('x')
y = function('y')(x)

ode = diff(y,x,2) - 3/x * diff(y,x,1) + 5/x^2 * y == 0

show(desolve(ode, y, ics=[1,1,0]))
```

SageMath outputs the same solution as in (2.3.42). Similarly, if our initial conditions were

$$y(1) = 1 \tag{2.3.43}$$
$$y'(1) = 0 \tag{2.3.44}$$

then the unique solution is

$$y = x^2 \sin(\log(x)).$$

□

2.3.3 Dimensions of subspaces

We now consider dimensions of subspaces of vector spaces.

Proposition 2.3.20 *Let W be a subspace of a finite-dimensional vector space V. Then W is finite-dimensional, and* $\mathrm{Dim}(W) \leq \mathrm{Dim}(V)$.

Proof. Let

$$\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

be a basis for V, so that $\mathrm{Dim}(V) = n$. We just need to show that W is finite-dimensional, i.e. that there exists a basis

$$\mathcal{B} = \{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for W. For then $\mathcal{B}$ will be a list of k linearly independent vectors which live in W (and hence also in V) and hence we must have $k \leq n$ by Proposition 2.2.10, as $\mathcal{C}$ spans V.

We show that W is finite-dimensional as follows.

If W is the zero vector space $\{\mathbf{0}\}$, then W is finite-dimensional by definition.

If W is not the zero vector space, then there exists a nonzero vector $\mathbf{w}_1 \in W$. Consider the list $\mathcal{B}_1 = \{\mathbf{w}_1\}$. Note that $\mathcal{B}_1$ is linearly independent, by Item 3 of Proposition 2.2.8. So, if $\mathcal{B}_1$ spans W, then it is a basis for W, and so W is finite-dimensional and we are done.

If $\mathcal{B}_1$ does not span W, then there exists a vector $\mathbf{w}_2 \in W$ which is not a scalar multiple of $\mathbf{e}_1$. Now consider the list $\mathcal{B}_2 = \{\mathbf{e}_1, \mathbf{e}_2\}$. Once again, $\mathcal{B}_2$ is linearly independent, by Item 3 of Proposition 2.2.8. So, if $\mathcal{B}_2$ spans W, then it is a basis for W, and we are done.

If $\mathcal{B}_2$ does not span W, then there exists a vector $\mathbf{w}_3 \in W$ which is not a linear combination of $\mathbf{w}_1$ and $\mathbf{w}_2$. Now consider the list $\mathcal{B}_3 = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$. Again, $\mathcal{B}_3$ is linearly independent, by Item 3 of Proposition 2.2.8. If it does not span W, then there exists a vector $\mathbf{w}_4 \in W$ which is not a linear combination of $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$. So consider the list $\mathcal{B}_4 = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4\}$.

This process must terminate for some $k \leq n$. If not, then it will produce a list $\mathcal{B}_{n+1} = \{\mathbf{w}_1, \ldots, \mathbf{w}_{n+1}\}$. This would be a linearly independent list of $n+1$ vectors from V. But $\mathrm{Dim}\, V = n$, so this is impossible, by Proposition 2.2.10. Hence for some $k \leq n$ we must have that $\mathcal{B}_k$ is a basis for W, and we are done. ∎

2.3.4 Infinite-dimensional vector spaces

It is good to have an example of an infinite-dimensional vector space.

Proposition 2.3.21 Poly *is infinite-dimensional.*

Proof. Suppose Poly is finite-dimensional. This means there exists a finite collection of polynomials $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ which spans Poly. But, let d be the highest degree of all the polynomials in the list $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$. Then $\mathbf{p} := x^{d+1}$ is a polynomial which is not in the span of $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$, since adding polynomials together and multiplying them by scalars can never increase the degree. We have arrived at a contradiction. So our initial assumption cannot be correct, i.e. Poly cannot be finite-dimensional. ∎

Example 2.3.22 We will not prove this here, but the following vector spaces are also infinite-dimensional:

- $\mathbb{R}^\infty$,
- $\mathrm{Fun}(X)$ where X is an infinite set,
- $\mathrm{Cont}(I)$ for any nonempty interval I, and
- $\mathrm{Diff}(I)$ for any open interval I

- Poly^k

□

2.3.5 The sifting algorithm and its uses

If we consider the proof of Proposition 2.2.10 (the 'Bumping off' Proposition) carefully, we find that it makes use of a *sifting algorithm*. This algorithm can actually be applied to *any* list of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in a vector space. Consider each vector $\mathbf{v}_i$ in the list consecutively. If $\mathbf{v}_i$ is the zero vector, or if it is a linear combination of the preceding vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}$, remove it from the list.

Example 2.3.23 Sift the following list of vectors in $\mathbb{R}^3$:

$$\begin{aligned} \mathbf{v}_1 &= (1,2,-1), & \mathbf{v}_2 &= (0,0,0), & \mathbf{v}_3 &= (3,6,-3) \\ \mathbf{v}_4 &= (1,0,5), & \mathbf{v}_5 &= (5,4,13), & \mathbf{v}_6 &= (1,1,0). \end{aligned}$$

We start with $\mathbf{v}_1$. Since it is not the zero vector, and is not a linear combination of any preceding vectors, it remains. We move on to $\mathbf{v}_2$, which is zero, so we remove it. We move on to $\mathbf{v}_3$, which by inspection is equal to $3\mathbf{v}_1$, so we remove it. We move on to $\mathbf{v}_4$. It is not zero, and cannot be expressed as a multiple of $\mathbf{v}_1$ (check this), so it remains. We move on to $\mathbf{v}_5$. We check if it can be written as a linear combination

$$\mathbf{v}_5 = a\mathbf{v}_1 + b\mathbf{v}_4$$

and find the solution $a = 2, b = 3$ (check this), so we remove it. Finally we move on to $\mathbf{v}_6$. We check if it can be written as a linear combination

$$\mathbf{v}_6 = a\mathbf{v}_1 + b\mathbf{v}_4$$

and find no solutions (check this), so it remains. Our final sifted list is

$$\mathbf{v}_1, \mathbf{v}_4, \mathbf{v}_6.$$

□

Checkpoint 2.3.24 Do the three 'check this' calculations above.

Sifting is a very useful way to construct a basis of a vector space!

Lemma 2.3.25 *If a list of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ spans a vector space V, then sifting the list will result in a basis for V.*

Proof. At each step, the vector that is removed from the list is either the zero vector, or a linear combination of the vectors before it. So if we remove this vector, the resulting list will still span V. Thus by the end of the process, the final sifted list of vectors still spans V.

To see that the final sifted list is linearly independent, we can apply Proposition 2.2.8. By construction, no vector in the final sifted list is a linear combination of the preceding vectors (if it was, it would have been removed!). Hence the final sifted list is not linearly dependent, so it must be linearly independent! ■

Corollary 2.3.26 *Any linearly independent list of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in a finite-dimensional vector space V can be extended to a basis of V.*

Proof. Since V is finite-dimensional, it has a basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Now consider the list

$$L : \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$$

which clearly spans V. By sifting this list, we will arrive at a basis for V, by Lemma 2.3.25. Some of the **e**-vectors may have been removed. But none of the **v**-vectors will have been removed, since that would mean some $\mathbf{v}_i$ is a linear combination of the preceding vectors $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}$, which is impossible, as $\mathbf{v}_1, \dots, \mathbf{v}_k$ is linearly independent list. Hence after sifting the list L we indeed extend our original list $\mathbf{v}_1, \dots, \mathbf{v}_k$ to a basis of V. ■

Corollary 2.3.27 *If $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ is a linearly independent list of n vectors in an n-dimensional vector space V, then it is a basis.*

Proof. By Corollary 2.3.26, we can extend $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ to a basis for V. But V has dimension n, so the basis must contain only n vectors by Theorem 2.3.2 (Invariance of Dimension). So we have not added any vectors at all! Hence our original list was already a basis. ■

Example 2.3.28 In Example 2.2.7 we showed that the list of polynomials

$$\mathbf{q}_0(x) := 1, \;\; \mathbf{q}_1(x) := x, \;\; \mathbf{q}_2(x) := 2x^2 - 1, \;\; \mathbf{q}_3(x) := 4x^3 - 3x$$

is linearly independent in Poly_3. Since $\mathrm{Dim}\,\mathrm{Poly}_3 = 4$, we see that it is a basis of Poly_3.

In Exercise 2.1.5, you showed that $\mathbf{q}_0, \dots, \mathbf{q}_3$ is a basis for Poly_3 by 'brute force'. This new method is *different*!

□

2.3.6 Exercises

1. Sift the list of vectors

$$\mathbf{v}_1 = (0,0,0), \qquad \mathbf{v}_2 = (1,0,-1), \qquad \mathbf{v}_3 = (1,2,3)$$
$$\mathbf{v}_4 = (3,4,5), \qquad \mathbf{v}_5 = (4,8,12), \qquad \mathbf{v}_6 = (1,1,0).$$

2. Let V be a vector space of dimension n. State whether each of the following statements is true or false. If it is true, prove it. If it is false, give a counterexample.

(a) Any linearly independent list of vectors in V contains at most n vectors.

(b) Any list of vectors which spans V contains at least n vectors.

3. Complete the proof of the following lemma.

Lemma. Suppose V is a vector space of dimension n. Then any linearly independent set of n vectors in V is a basis for V.

Proof. Let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a linearly independent set of vectors in V.

Suppose that $\mathcal{B}$ is *not* a basis for V.
Therefore, $\mathcal{B}$ does not span V, since ... (a)
Therefore, there exists $\mathbf{v} \in V$ such that ... (b)
Now, add $\mathbf{v}$ to the list $\mathcal{B}$ to obtain a new list $\mathcal{B}' :=$ ... (c)
The new list $\mathcal{B}'$ is linearly independent because ... (d)
This is a contradiction because ... (e)
Hence, $\mathcal{B}$ must be a basis for V.

4. Use Exercise 2.3.6.2(a) to show that the list of matrices in $\mathrm{Mat}_{2,2}$ in Exercise 2.2.2 is linearly dependent.

5. In each case, use the results in Exercise 2.3.6.2 and Exercise 2.3.6.3 to determine if $\mathcal{B}$ is a basis for V.

(a) $V = \mathrm{Poly}_2$, $\mathcal{B} = \{2 + x^2,\, 1 - x,\, 1 + x - 3x^2,\, x - x^2\}$

(b) $V = \mathrm{Mat}_{2,2}$,

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 3 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right\}$$

(c) $V = \mathrm{Trig}_2$, $\mathcal{B} = \{\sin^2 x,\, \cos^2 x,\, 1 - \sin 2x,\, \cos 2x + 3\sin 2x\}$

6. Let $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ be a linearly independent list of vectors in a vector space V. State whether each of the following statements is true or false. If it is true, prove it. If it is false, give a counterexample. (Hint: Use the definition of linear independence.)

(a) The list $\{\mathbf{u} + \mathbf{v},\, \mathbf{v} + \mathbf{w},\, \mathbf{u} + \mathbf{w}\}$ is linearly independent.

(b) The list $\{\mathbf{u} - \mathbf{v},\, \mathbf{v} - \mathbf{w},\, \mathbf{u} - \mathbf{w}\}$ is linearly independent.

7. For each of the following, show that V is a subspace of Poly_2, find a basis for V, and compute $\mathrm{Dim}\, V$.

(a) $V = \{p \in \mathrm{Poly}_2 : p(2) = 0\}$

(b) $V = \{p \in \mathrm{Poly}_2 : xp'(x) = p(x)\}$

8. Prove or disprove: there exits a basis $\{p_0, p_1, p_2, p_3\}$ of Poly_3 such that none of the polynomials p_0, p_1, p_2, p_3 have degree 2.

9. Prove or disprove: if U and W are distinct subspaces of V with $U \neq V$ and $W \neq V$, then $\dim(U + V) = \dim(U) + \dim(V)$. (Recall the definition of the sum of two subspaces from Exercise 1.6.4.9.)

2.3.7 Solutions

2.4 Coordinate vectors

There is a more direct way to think about a basis of a vector space.

Proposition 2.4.1 Bases give coordinates. *Let $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be a list of vectors in a vector space V. Then the following statements are equivalent:*

1. *$\mathcal{B}$ is a basis for V.*
2. *Every vector $\mathbf{v} \in V$* can *be written as a linear combination*

$$\mathbf{v} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n \tag{2.4.1}$$

in precisely one *way. (That is, for each $\mathbf{v} \in V$ there* exist *scalars a_1, a_2, $\ldots$, a_n satisfying (2.4.1), and that moreover these scalars are* unique*).*

It is important to understand the mathematical phrase 'There exists a unique X satisfying Y'. It means two things. Firstly, that *there does exist an X* which satisfies Y. And secondly, that there is *no more than* *one* X which satisfies Y.

Proof.
$1 \Rightarrow 2$**.** Suppose that the list of vectors$\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis for V. Suppose $\mathbf{v} \in V$. Since the list of vectors $\mathcal{B}$ spans V, we know that we *can* write $\mathbf{v}$ as a linear combination of the vectors in the list in at least one way,

$$\mathbf{v} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n. \tag{2.4.2}$$

We need to show that this is the *only* way to express $\mathbf{v}$ as a linear combination of the vectors $\mathbf{e}_i$. Indeed, suppose that we also have

$$\mathbf{v} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + \cdots + b_n\mathbf{e}_n. \tag{2.4.3}$$

Subtracting these two equations gives

$$\mathbf{0} = (a_1 - b_1)\mathbf{e}_1 + (a_2 - b_2)\mathbf{e}_2 + \cdots + (a_n - b_n)\mathbf{e}_n.$$

Since the list of vectors $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ is linearly independent, we conclude that

$$a_1 - b_1 = 0,\ a_2 - b_2 = 0,\ \cdots,\ a_n - b_n = 0.$$

That is, $a_1 = b_1$, $a_2 = b_2$, and so on up to $a_n = b_n$, and hence the expansion (2.4.2) is unique.
$2 \Leftarrow 1$**.** Conversely, suppose that every vector $\mathbf{v}$ can be written as a unique linear combination

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

The fact that each $\mathbf{v}$ *can* be written as a linear combination of the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ means that $\mathcal{B}$ spans V. We still need to show that this list $\mathcal{B}$ is linearly independent. So, suppose that there exist scalars b_1, b_2, …, b_n such that

$$b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + \cdots + b_n\mathbf{e}_n = \mathbf{0}. \tag{2.4.4}$$

We need to show that all the b_i must equal zero. We already know *one* possible solution of (2.4.4) : simply set each $b_i = 0$. But we are told that each vector (in particular, the vector $\mathbf{0}$) can be expressed as a linear combination of the $\mathbf{e}_i$ in exactly one way. Hence this must be the only solution, i.e. we must have $b_1 = b_2 = \cdots = b_n = 0$, and so the list $\mathcal{B}$ is linearly independent. ■

Definition 2.4.2 Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ be a basis for a vector space V, and let $\mathbf{v} \in V$. Write

$$\mathbf{v} = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + \mathbf{a}_n\mathbf{b}_n\,.$$

The scalars a_i appearing in the above expansion are called the **coordinates of the vector v with respect to the basis** $\mathcal{B}$. The column vector

$$[\mathbf{v}]_{\mathcal{B}} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathrm{Col}_n$$

is called the **coordinate vector of v with respect to the basis** $\mathcal{B}$. ♢

> I indicate that a collection of things is a *list* (where the order matters) and not merely a *set* (where the order does not matter) using my own home-made symbols $\{\}$. A basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is a list of vectors. The order of the vectors matters because it affects the coordinate vector $[\mathbf{v}]_{\mathcal{B}}$.

Example 2.4.3 Find the coordinate vector of $\mathbf{p} = 2x^2 - 2x + 3$ with respect to the basis $\mathcal{B} = \left\{1 + x, x^2 + x - 1, x^2 + x + 1\right\}$ of Poly_3.

Solution. We need to write $\mathbf{p}$ as a linear combination of polynomials from the basis $\mathcal{B}$:

$$2x^2 - 2x + 3 = a_1(1 + x) + a_2(x^2 + x - 1) + a_3(x^2 + x + 1)$$

Collecting powers of x^2, x and 1 on the right hand side gives:

$$2x^2 - 2x + 3 = (a_2 + a_3)x^2 + (a_1 + a_2 + a_3)x + (a_1 - a_2 + a_3)1$$

This translates into the equations:

$$\begin{aligned} a_2 + a_3 &= 2 \\ a_1 + a_2 + a_3 &= -2 \\ a_1 - a_2 + a_3 &= 3 \end{aligned}$$

We can solve these equations by hand, or we can use SageMath:

```
var('a1␣a2␣a3')

show(solve((a2 + a3 == 2,
      a1 + a2 + a3 == -2,
      a1 - a2 + a3 == 3), (a1, a2, a3)))
```

We compute the coordinates of p as $a_1 = -4$, $a_2 = -\frac{5}{2}$, $a_3 = \frac{9}{2}$. In other words,

$$2x^2 - 2x + 3 = -4(1 + x) - \frac{5}{2}(x^2 + x - 1) + \frac{9}{2}(x^2 + x + 1)$$

Therefore,

$$[\mathbf{p}]_{\mathcal{B}} := \begin{bmatrix} -4 \\ -\frac{5}{2} \\ \frac{9}{2} \end{bmatrix}$$

□

Example 2.4.4 Find the coordinate vectors of $\mathbf{v}$ and $\mathbf{w}$ in Figure 2.4.5 with respect to the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

Solution. By inspection, we see that $\mathbf{v} = 2\mathbf{b}_1 - \mathbf{b}_2$, so that

$$[\mathbf{v}]_{\mathcal{B}} := \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Also by inspection, we see that $\mathbf{w} = -3\mathbf{b}_1 + 2\mathbf{b}_2$, so that

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

□

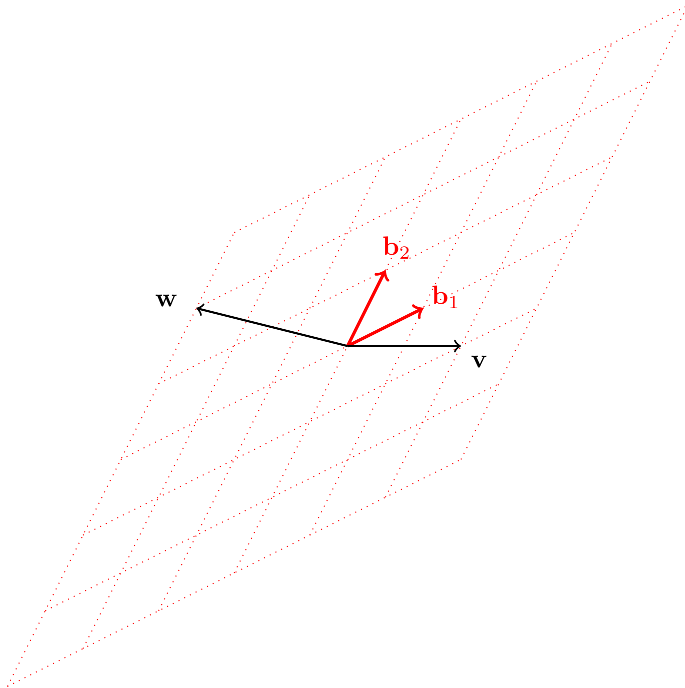


Figure 2.4.5: The basis B for $\mathbb{R}^2$.

Example 2.4.6 Find the coordinate vector of the function $\mathbf{f}$ given by

$$\mathbf{f}(x) = \sin^2 x - \cos^3 x$$

with respect to the standard basis

$$\mathcal{S} = \{1, \cos x, \sin x, \cos 2x, \sin 2x, \cos 3x, \sin 3x\}$$

of Trig_3.

Solution. Using the addition formulae for sin and cos as in Exercise 1.6.23, we compute:

$$\sin^2 x - \cos^3 x = \frac{1}{2} - \frac{3}{4}\cos x - \frac{1}{2}\cos 2x - \frac{1}{4}\cos 3x. \tag{2.4.5}$$

We could also do this in SageMath as follows:

```
x=var('x')

f = sin(x)^2 - cos(x)^3

show(f.reduce_trig())
```

Hence

$$[\mathbf{f}]_{\mathcal{S}} = \begin{bmatrix} \frac{1}{2} \\ -\frac{3}{4} \\ 0 \\ -\frac{1}{2} \\ 0 \\ -\frac{1}{4} \\ 0 \end{bmatrix}$$

□

Checkpoint 2.4.7 Check the expansion (2.4.5) by hand.

Solution.

$$\cos 2x = 1 - 2\sin^2 x \implies \sin^2 x = \frac{1}{2} - \frac{1}{2}\cos 2x$$

$$\cos^3 x = \cos x \cos^2 x = \cos x(1-\sin^2 x) = \cos x(\frac{1}{2}+\frac{1}{2}\cos 2x) = \frac{1}{2}\cos x+\frac{1}{2}\cos x \cos 2x$$

$$\cos x \cos 2x = \cos 3x+\sin x \sin 2x = \cos 3x+2\sin^2 x \cos x = \cos 3x+(1-\cos 2x)\cos x = \cos 3x+\cos x-\cos 2x \cos x$$

Thus

$$\sin^2 x-\cos^3 x = \frac{1}{2}-\frac{1}{2}\cos 2x-(\frac{1}{2}\cos x+\frac{1}{2}\cos x\cos 2x) = \frac{1}{2}-\frac{1}{2}\cos 2x-\frac{1}{2}\cos x-\frac{1}{2}(\frac{1}{2}\cos 3x+\frac{1}{2}\cos x) = \frac{1}{2}-\frac{3}{4}\cos$$

Lemma 2.4.8 *Let $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be a basis for a vector space V. Then for all vectors $\mathbf{v}, \mathbf{w} \in V$ and all scalars k we have*

1. $[\mathbf{v}+\mathbf{w}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}} + [\mathbf{w}]_{\mathcal{B}}$
2. $[k\mathbf{v}]_{\mathcal{B}} = k[\mathbf{v}]_{\mathcal{B}}$

Proof. (a) Suppose that

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

and

$$\mathbf{w} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + \cdots + b_n\mathbf{e}_n\,.$$

Then, using the rules of a vector space, we compute

$$\mathbf{v}+\mathbf{w} = (a_1+b_1)\mathbf{e}_1 + (a_2+b_2)\mathbf{e}_2 + \cdots + (a_n+b_n)\mathbf{e}_n\,.$$

From this we read off that

$$[\mathbf{v}+\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{bmatrix}.$$

The proof of (b) is similar. ∎

Exercises

1. Prove Lemma 2.4.8(b) in the case where V is two-dimensional, so that $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2\}$. Justify each step using the rules of a vector space.

Solution. As before, suppose that

$$\mathbf{v} = a_1\mathbf{e}_1 + a_2$$

Which gives

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

Then

$$\begin{aligned} k\mathbf{v} &= k(a_1\mathbf{e}_1 + a_2\mathbf{e}_2) \\ &= k(a_1\mathbf{e}_1) + k(a_2\mathbf{e}_2) && \text{(R4)} \\ &= (ka_1)\mathbf{e}_1 + (ka_2)\mathbf{e}_2 && \text{(R6)} \end{aligned}$$

Reading off the coefficients, we obtain

$$k\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} ka_1 \\ ka_2 \end{bmatrix},$$

as desired.

2. Let $\mathcal{B}$ be the basis of $\mathrm{Mat}_{2,2}$ given by

$$\mathsf{B}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathsf{B}_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \mathsf{B}_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \mathsf{B}_4 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Determine $[\mathsf{A}]_{\mathcal{B}}$, where

$$\mathsf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Solution. To determine $[\mathsf{A}]_{\mathcal{B}}$, we must find the scalars a_1, a_2, a_3, a_4 satisfying

$$a_1\mathsf{B}_1 + a_2\mathsf{B}_2 + a_3\mathsf{B}_3 + a_4\mathsf{B}_4 = \mathsf{A}.$$

This results in a system of 4 linear equations in 4 variables, one equation for each entry in A:

$$\begin{aligned} a_1 + a_2 + a_3 &= 1 \\ a_3 + a_4 &= 2 \\ a_3 - a_4 &= 3 \\ a_1 - a_2 + a_3 &= 4 \end{aligned}$$

Solving this equation, we get

$$a_1 = 0, \quad a_2 = -\frac{3}{2}, \quad a_3 = \frac{5}{2}, \quad a_4 = -\frac{1}{2}.$$

Hence

$$[\mathsf{A}]_{\mathcal{B}} = \begin{bmatrix} 0 \\ -\frac{3}{2} \\ \frac{5}{2} \\ -\frac{1}{2} \end{bmatrix}$$

3.

(a) Find a basis $\mathcal{B}$ for the vector space

$$V := \{p \in \mathrm{Poly}_2 : p(2) = 0\}.$$

(b) Consider $p(x) = x^2 + x - 6$. Show that $p \in V$.

(c) Determine the coordinate vector of p with respect to your basis $\mathcal{B}$, i.e. determine $[p]_{\mathcal{B}}$.

Solution.

(a)

$$p(x) = x^2 + x - 6 = (x+3)(x-2)$$

and so $p(2) =$ and hence $p \in V$.

(b) Recall that $\mathcal{B} = \{x - 2, x(x-2)\}$.

$$p(x) = x^2 + x - 6 = (x+3)(x-2) = 3(x-2) + x(x-2)$$

and so

$$[p]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

4. Find the coordinate representation of $\mathbf{p}(x) = 3x^3 - 7x + 1$ with respect to your basis in Exercise 2.3.6.8.

Solution. We shall use the basis $\mathcal{B} = \{x^3, x^3 + x^2, x, 1\}$. Since $\mathbf{p}(x)$ has no degree 2 term, we know immediately that

$$\mathbf{p}(x) = ax^3 + 0(x^3 + x^2) + cx + d.$$

Reading off the rest of the coefficients, we see that $a = 3, c = -7, d = 1$. Hence

$$[p]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 0 \\ -7 \\ 1 \end{bmatrix}.$$

5. Consider the vector space W from Example 2.3.15,

$$W = \{(w_1, w_2, w_3) \in \mathbb{R}^3 : w_1 + 2w_2 + 3w_3 = 0\},$$

and the following bases for W:

$$\mathcal{B} = \{\mathbf{a}, \mathbf{b}\}, \quad \mathcal{C} = \{\mathbf{u}, \mathbf{v}\}$$

where

$$\mathbf{a} = (1, 0, -\frac{1}{3}), \qquad \mathbf{b} = (0, 1, -\frac{2}{3})$$
$$\mathbf{u} = (1, 2, -\frac{5}{3}), \qquad \mathbf{v} = (-4, 2, 0)$$

Consider the vector $\mathbf{w} = (-2, 4, -2) \in W$. Compute $[\mathbf{w}]_{\mathcal{B}}$ and $[\mathbf{w}]_{\mathcal{C}}$.

6. Let V be the vector space of solutions to the differential equation

$$y'' + y = 0. \tag{2.4.6}$$

(a) Show that $\mathcal{B} = \{\cos x, \sin x\}$ is a basis for V.

(b) Let $y \in V$ be defined as the unique solution to the differential equation in (2.4.6) satisfying

$$y(\frac{\pi}{6}) = 1, \quad y'(\frac{\pi}{6}) = 0.$$

(Note that we can indeed define y uniquely in this way due to Theorem 2.3.18.) Compute $[y]_{\mathcal{B}}$.

(c) Let $z(x) = \cos\left(x - \frac{\pi}{3}\right)$.

i. Show that $z \in V$ by checking that it solves the differential equation (2.4.6).

ii. Determine $[z]_{\mathcal{B}}$.

7. Let V be the vector space of solutions to the differential equation

$$(1 - x^2)y'' - xy' + 4y = 0, \quad x \in (-1, 1). \tag{2.4.7}$$

(a) Show that y_1 and y_2 are elements of V, where

$$y_1(x) = 2x^2 - 1, \quad y_2(x) = x\sqrt{1-x^2}.$$

(b) Show that $\mathcal{B} = \{\mathbf{y}_1, \mathbf{y}_2\}$ is a basis for V.

(c) Let $y \in V$ be defined as the unique solution to the differential equation in (2.4.7) satisfying

$$y(\frac{1}{2}) = 1, \quad y'(\frac{1}{2}) = 0.$$

(Note that we can indeed define y uniquely in this way due to Theorem 2.3.18.) Compute $[y]_{\mathcal{B}}$.

Solutions

· Exercises

2.4.1. Solution. As before, suppose that

$$\mathbf{v} = a_1\mathbf{e}_1 + a_2$$

Which gives

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

Then

$$\begin{aligned} k\mathbf{v} &= k(a_1\mathbf{e}_1 + a_2\mathbf{e}_2) \\ &= k(a_1\mathbf{e}_1) + k(a_2\mathbf{e}_2) && \text{(R4)} \\ &= (ka_1)\mathbf{e}_1 + (ka_2)\mathbf{e}_2 && \text{(R6)} \end{aligned}$$

Reading off the coefficients, we obtain

$$k\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} ka_1 \\ ka_2 \end{bmatrix},$$

as desired.

2.4.2. Solution. To determine $[\mathsf{A}]_{\mathcal{B}}$, we must find the scalars a_1, a_2, a_3, a_4 satisfying

$$a_1\mathsf{B}_1 + a_2\mathsf{B}_2 + a_3\mathsf{B}_3 + a_4\mathsf{B}_4 = \mathsf{A}.$$

This results in a system of 4 linear equations in 4 variables, one equation for each entry in A:

$$\begin{aligned} a_1 + a_2 + a_3 &= 1 \\ a_3 + a_4 &= 2 \\ a_3 - a_4 &= 3 \\ a_1 - a_2 + a_3 &= 4 \end{aligned}$$

Solving this equation, we get

$$a_1 = 0, \quad a_2 = -\frac{3}{2}, \quad a_3 = \frac{5}{2}, \quad a_4 = -\frac{1}{2}.$$

Hence

$$[\mathsf{A}]_{\mathcal{B}} = \begin{bmatrix} 0 \\ -\frac{3}{2} \\ \frac{5}{2} \\ -\frac{1}{2} \end{bmatrix}$$

2.4.3. Solution.

(a)

$$p(x) = x^2 + x - 6 = (x+3)(x-2)$$

and so $p(2) =$ and hence $p \in V$.

(b) Recall that $\mathcal{B} = \{x-2, x(x-2)\}$.

$$p(x) = x^2 + x - 6 = (x+3)(x-2) = 3(x-2) + x(x-2)$$

and so

$$[p]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

2.4.4. Solution. We shall use the basis $\mathcal{B} = \{x^3, x^3 + x^2, x, 1\}$. Since $\mathbf{p}(x)$ has no degree 2 term, we know immediately that

$$\mathbf{p}(x) = ax^3 + 0(x^3 + x^2) + cx + d.$$

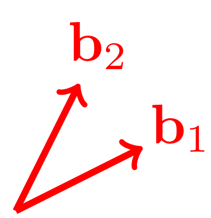
Reading off the rest of the coefficients, we see that $a = 3, c = -7, d = 1$. Hence

$$[p]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 0 \\ -7 \\ 1 \end{bmatrix}.$$

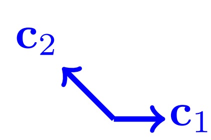
2.5 Change of basis

2.5.1 Coordinate vectors are different in different bases

Suppose that $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ are two different bases for $\mathbb{R}^2$, shown below:



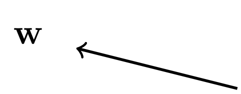
(a) The basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$



(b) The basis $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$.

Figure 2.5.1: Two different bases for $\mathbb{R}^2$

Suppose we are given a vector $\mathbf{w} \in \mathbb{R}^2$:



We would like to compute the coordinate vector of the *same* vector $\mathbf{w}$ with respect to the two different bases $\mathcal{B}$ and $\mathcal{C}$.

For this particular $\mathbf{w}$, from Figure 2.5.2, we see that in the basis $\mathcal{B}$, we have

$$\mathbf{w} = -3\mathbf{b}_1 + 2\mathbf{b}_2 \qquad \therefore\ [\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}. \tag{2.5.1}$$

On the other hand, in the basis $\mathcal{C}$, we see from Figure 2.5.3 that

$$\mathbf{w} = \mathbf{c}_1 - 3\mathbf{c}_2 \qquad \therefore\ [\mathbf{w}]_{\mathcal{C}} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}. \tag{2.5.2}$$

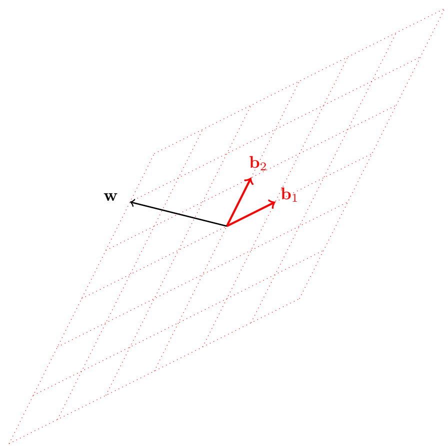


Figure 2.5.2: $\mathbf{w} = -3\mathbf{b}_1 + 2\mathbf{b}_2$

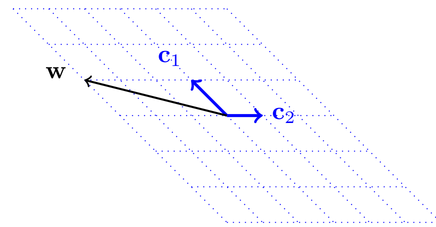


Figure 2.5.3: $\mathbf{w} = \mathbf{c}_1 - 3\mathbf{c}_2$

So, the *same* vector $\mathbf{w}$ has different coordinate vectors $[\mathbf{w}]_{\mathcal{B}}$ and $[\mathbf{w}]_{\mathcal{C}}$ with respect to the bases $\mathcal{B}$ and $\mathcal{C}$!

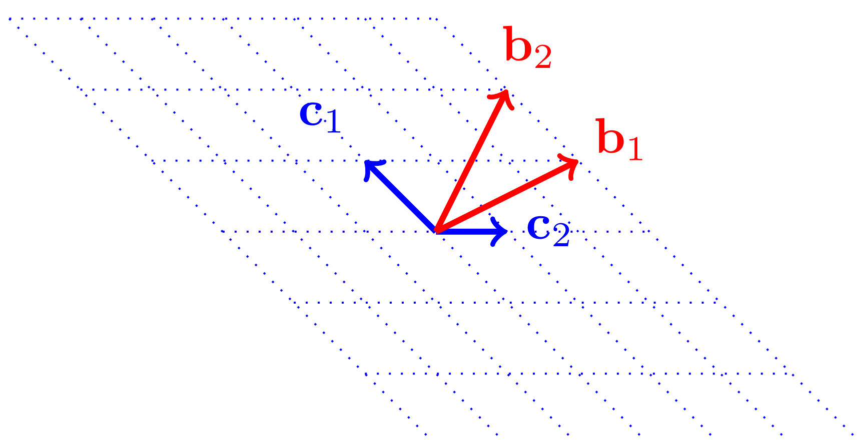
2.5.2 Changing from one basis to another

Now, suppose we know the bases $\mathcal{B}$ and $\mathcal{C}$, and we know the coordinate vector $[\mathbf{w}]_{\mathcal{B}}$ of $\mathbf{w}$ in the basis $\mathcal{B}$,

$$[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 2 \end{bmatrix},$$

that is, $\mathbf{w} = -3\mathbf{b}_1 + 2\mathbf{b}_2$. How can we compute $[\mathbf{w}]_{\mathcal{C}}$, the coordinate vector of $\mathbf{w}$ in the basis $\mathcal{C}$?

The best way is to express each vector in the basis $\mathcal{B}$ as a linear combination of the basis vectors in $\mathcal{C}$. In the next figure, the vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ are displayed against the background of integral linear combinations of the basis $\mathcal{C}$:



We read off that:

$$\mathbf{b}_1 = \mathbf{c}_1 + 3\mathbf{c}_2 \tag{2.5.3}$$
$$\mathbf{b}_2 = 2\mathbf{c}_1 + 3\mathbf{c}_2 \tag{2.5.4}$$

Therefore, we compute:

$$\begin{aligned} \mathbf{w} &= -3\mathbf{b}_1 + 2\mathbf{b}_2 \\ &= -3(\mathbf{c}_1 + 3\mathbf{c}_2) + 2(2\mathbf{c}_1 + 3\mathbf{c}_2) \\ &= \mathbf{c}_1 - 3\mathbf{c}_2 \end{aligned}$$

From this we read off that

$$[\mathbf{w}]_{\mathcal{C}} = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \tag{2.5.5}$$

which is the right answer, as we know from (2.5.2).

In fact, this calculation can be phrased in terms of matrices.

Definition 2.5.4 Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be bases for a vector space V. The **change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$** is the $n \times n$ matrix $\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ whose columns are the coordinate vectors $[\mathbf{b}_1]_{\mathcal{C}}, \ldots, [\mathbf{b}_n]_{\mathcal{C}}$:

$$\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}} := \left[\begin{bmatrix} \\ \mathbf{b}_1 \\ \\ \end{bmatrix}_{\mathcal{C}} \quad \begin{bmatrix} \\ \mathbf{b}_2 \\ \\ \end{bmatrix}_{\mathcal{C}} \quad \cdots \quad \begin{bmatrix} \\ \mathbf{b}_n \\ \\ \end{bmatrix}_{\mathcal{C}} \right].$$

$\Diamond$

Example 2.5.5 In our running example, we see from (2.5.3) and (2.5.4) that

$$[\mathbf{b}_1]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad [\mathbf{b}_2]_{\mathcal{C}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Hence the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$ is

$$\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 3 & 3 \end{bmatrix}.$$

$\square$

Before we move on, we need to recall something about matrix multiplication. Suppose you collect together m column vectors to form a matrix:

$$\left[\begin{bmatrix} \\ \mathrm{C}_1 \\ \\ \end{bmatrix} \quad \begin{bmatrix} \\ \mathrm{C}_2 \\ \\ \end{bmatrix} \quad \cdots \quad \begin{bmatrix} \\ \mathrm{C}_m \\ \\ \end{bmatrix} \right]$$

(For instance, our change-of-basis matrix $\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ was formed in this way.) Then the product of this matrix with a column vector can be computed as follows:

$$\left[\begin{bmatrix} \\ \mathrm{C}_1 \\ \\ \end{bmatrix} \quad \begin{bmatrix} \\ \mathrm{C}_2 \\ \\ \end{bmatrix} \quad \cdots \quad \begin{bmatrix} \\ \mathrm{C}_m \\ \\ \end{bmatrix} \right] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = a_1 \begin{bmatrix} \\ \mathrm{C}_1 \\ \\ \end{bmatrix} + a_2 \begin{bmatrix} \\ \mathrm{C}_2 \\ \\ \end{bmatrix} + \cdots + a_m \begin{bmatrix} \\ \mathrm{C}_m \\ \\ \end{bmatrix}. \tag{2.5.6}$$

Checkpoint 2.5.6 Prove the above formula!

Solution. We check the i^{th} entry of the LHS of (2.5.6) using just the definition of matrix multiplation.

$$(\text{LHS})_i = (\text{C}_1)_i a_1 + \ldots + (\text{C}_n)_i a_n = (\text{RHS})_i$$

and we're done!

We can now prove the following theorem.

Theorem 2.5.7 Change of basis. *Suppose that $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ are bases for a vector space V, and let $\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ be the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$. Then for all vectors $\mathbf{v}$ in V,*

$$[\mathbf{v}]_{\mathcal{C}} = \mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}. \tag{2.5.7}$$

Proof. Let $\mathbf{v} \in V$. Expand it in the basis $\mathcal{B}$:

$$\mathbf{v} = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_n\mathbf{b}_n, \text{i.e. } [\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Then,

$$\begin{aligned}
{}_{\mathcal{C}} &= [a_1\mathbf{b}_1 + \cdots + a_n\mathbf{b}_n]_{\mathcal{C}} \\
&= a_1[\mathbf{b}_1]_{\mathcal{C}} + \cdots + a_n[\mathbf{b}_n]_{\mathcal{C}} && \text{(Lemma 2.4.8)} \\
&= \left[\begin{bmatrix} \\ \mathbf{b}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \mathbf{b}_2 \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \mathbf{b}_n \\ \\ \end{bmatrix}\right] \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} && \text{(2.5.6)} \\
&= \mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}\,.
\end{aligned}$$

■

Example 2.5.8 In our running example, the theorem says that for *any* vector $\mathbf{v} \in \mathbb{R}^2$,

$$[\mathbf{v}]_{\mathcal{C}} = \begin{bmatrix} 1 & 2 \\ 3 & 3 \end{bmatrix} [\mathbf{v}]_{\mathcal{B}}.$$

In particular, this must hold for our vector $\mathbf{w}$, whose coordinate vector in the basis $\mathbf{B}$ was:

$$[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}.$$

So in this case, the theorem is saying that

$$[\mathbf{w}]_{\mathcal{C}} = \begin{bmatrix} 1 & 2 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

which agrees with our previous calculation (2.5.5)! □

2.5.3 Exercises

1. This is a continuation of Exercise 2.4.2. Consider the following two bases for $\text{Mat}_{2,2}$:

$$\mathcal{B} = \left\{ \text{B}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{B}_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \text{B}_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \text{B}_4 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$$

$$\mathcal{C} = \left\{ \mathrm{C}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \mathrm{C}_2 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \mathrm{C}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \mathrm{C}_4 = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \right\}$$

(a) Determine the change-of-basis matrices $\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ and $\mathsf{P}_{\mathcal{B}\leftarrow\mathcal{C}}$.

(b) Determine $[\mathsf{A}]_{\mathcal{B}}$ and $[\mathsf{A}]_{\mathcal{C}}$ where

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

(c) Check that $[\mathsf{A}]_{\mathcal{C}} = \mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathsf{A}]_{\mathcal{B}}$ and that $[\mathsf{A}]_{\mathcal{B}} = \mathsf{P}_{\mathcal{B}\leftarrow\mathcal{C}}[\mathsf{A}]_{\mathcal{C}}$.

2. Compute the change-of-basis matrix $\mathsf{P}_{\mathcal{B}\leftarrow\mathcal{S}}$ from the standard basis

$$\mathcal{S} = \{1, \cos x, \sin x, \cos 2x, \sin 2x\}$$

of Trig_2 to the basis

$$\mathcal{B} = \left\{1, \cos x, \sin x, \cos^2 x, \sin^2 x\right\}.$$

3. Figure 2.5.9 displays a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ for $\mathbb{R}^2$, a background of integral linear combinations of $\mathbf{b}_1$ and $\mathbf{b}_2$, and a certain vector $\mathbf{w} \in \mathbb{R}^2$. Similarly, Figure 2.5.10 displays another basis $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ for $\mathbb{R}^2$, a background of integral linear combinations of $\mathbf{c}_1$ and $\mathbf{c}_2$, and the same vector $\mathbf{w} \in \mathbb{R}^2$.

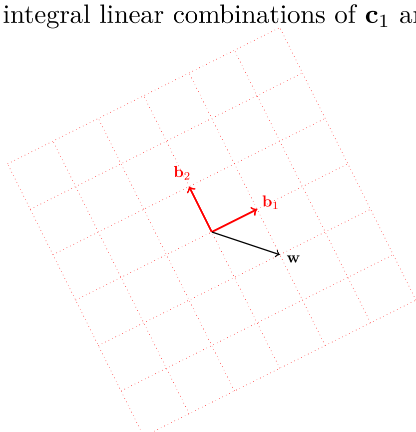


Figure 2.5.9: The vector $\mathbf{w}$ against a background of integral linear combinations of the basis vectors from $\mathcal{B}$.

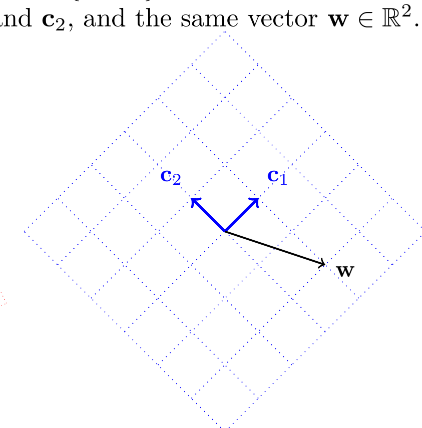
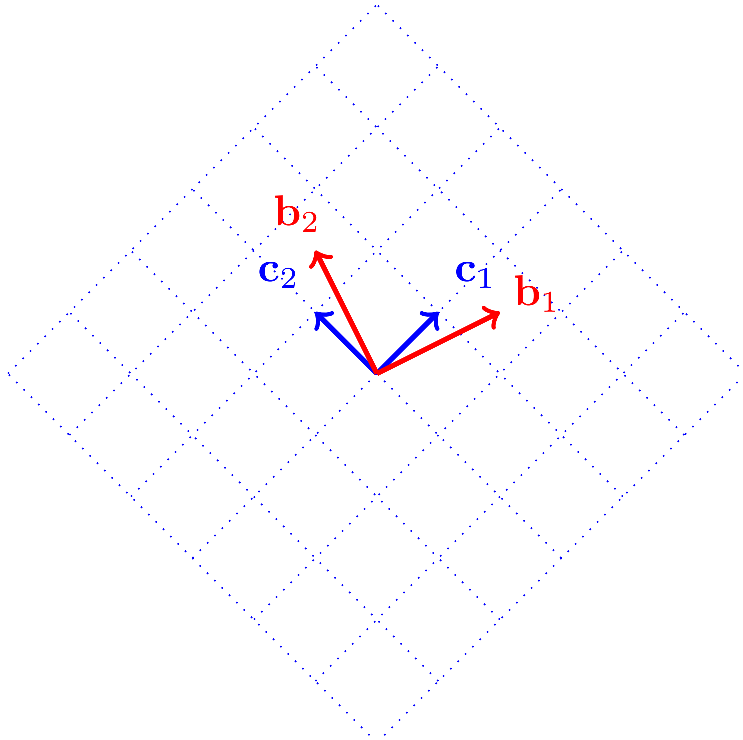


Figure 2.5.10: The vector $\mathbf{w}$ against a background of integral linear combinations of the basis vectors from $\mathcal{C}$.

(a) Determine $[\mathbf{w}]_{\mathcal{B}}$, directly from Figure 2.5.9.

(b) Determine $[\mathbf{w}]_{\mathcal{C}}$, directly from Figure 2.5.10.

(c) The following figure displays the $\mathcal{B}$ basis against a background of integral linear combinations of the $\mathcal{C}$ basis:



Determine the change-of-basis matrix $\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$. (You may assume that all coefficients are either integers and half-integers.)

(d) Multiply the matrix you computed in (c) with the column vector you computed in (a). That is, compute the product $\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathsf{w}]_{\mathcal{B}}$. Is your answer the same as what you obtained in (b)?

2.5.4 Solutions

Chapter 3

Linear maps

3.1 Definitions and Examples

Recall that a *function* (or a *map*) $f : X \to Y$ from a set X to a set Y is simply a rule which assigns to each element of X an element $f(x)$ of Y. We write $x \mapsto f(x)$ to indicate that an element $x \in X$ maps to $f(x) \in Y$ under the function f. See Figure 3.1.1. Two functions $f, g : X \to Y$ are *equal* if $f(x) = g(x)$ for all x in X.

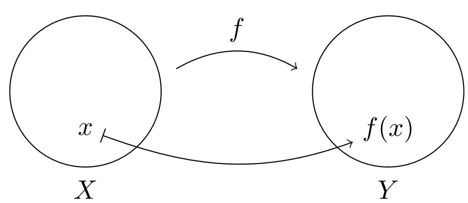


Figure 3.1.1: A function $f : X \to Y$.

Definition 3.1.2 Let V and W be vector spaces. A **linear map** from V to W is a function $T : V \to W$ satisfying:

- $T(\mathbf{v} + \mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}')$ for all vectors $\mathbf{v}, \mathbf{v}' \in V$.
- $T(k\mathbf{v}) = kT(\mathbf{v})$ for all vectors $\mathbf{v} \in V$ and scalars $k \in \mathbb{R}$.

$\Diamond$

Another name for a linear map is a *linear transformation*.

Example 3.1.3 Identity map. Let V be a vector space. The function

$$\begin{aligned} \mathrm{id}_V : V &\to V \\ \mathbf{v} &\mapsto \mathbf{v} \end{aligned}$$

is called the *identity map* on V. It is clearly a linear map, since

$$\begin{aligned} \mathrm{id}_V(\mathbf{v} + \mathbf{w}) &= \mathbf{v} + \mathbf{w} \\ &= \mathrm{id}_V(\mathbf{v}) + \mathrm{id}_V(\mathbf{w}) \end{aligned}$$

and

$$\mathrm{id}_V(k\mathbf{v}) = k\mathbf{v}$$

$$= k\,\mathrm{id}_V(\mathbf{v}).$$

□

Example 3.1.4 Projection. The function

$$\begin{aligned} T : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto x \end{aligned}$$

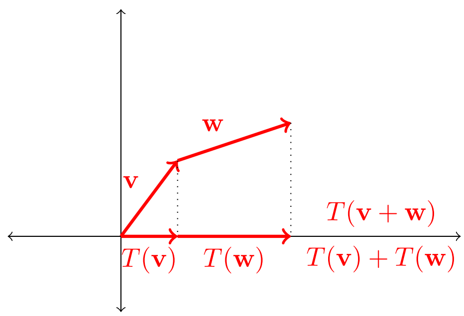
which projects vectors onto the x-axis is a linear map.

Let us check additivity algebraically:

$$T\left((x_1, y_1)\right) + \left((x_2, y_2)\right) \stackrel{?}{=} T\left((x_1, y_1)\right) + T\left((x_2, y_2)\right)$$

$$\begin{aligned} \text{LHS} &= T\left((x_1 + x_2, y_1 + y_2)\right) \\ &= x_1 + x_2 \end{aligned} \qquad\qquad \text{RHS} = x_1 + x_2$$

Here is a a graphical version of this proof:



□

Checkpoint 3.1.5 Prove algebraically that we also have $T(k\mathbf{v}) = kT(\mathbf{v})$, so that T is a linear map.

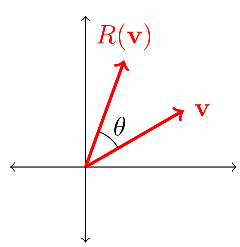
Solution.

$$\begin{aligned} T(k\mathbf{v}) &= T(k(x, y)) \\ &= T((kx, ky)) \\ &= kx = kT((x, y)) \\ &= kT(\mathbf{v}) \end{aligned}$$

Example 3.1.6 Rotation. Fix an angle θ. The function

$$\begin{aligned} R : \mathbb{R}^2 &\to \mathbb{R}^2 \\ \mathbf{v} &\mapsto \text{rotation of } \mathbf{v} \text{ counterclockwise through angle } \theta \end{aligned}$$

is a linear map, by a similar graphical argument as in Example 3.1.4.



□

Example 3.1.7 (Cross product with a fixed vector) Fix a vector $\mathbf{w} \in \mathbb{R}^3$. The function

$$\begin{aligned} C : \mathbb{R}^3 &\to \mathbb{R}^3 \\ \mathbf{v} &\mapsto \mathbf{w} \times \mathbf{v} \end{aligned}$$

is a linear map because of the properties of the cross-product,

$$\begin{aligned} \mathbf{w} \times (\mathbf{v}_1 + \mathbf{v}_2) &= \mathbf{w} \times \mathbf{v}_1 + \mathbf{w} \times \mathbf{v}_2 \\ \mathbf{w} \times (k\mathbf{v}) &= k\mathbf{w} \times \mathbf{v}. \end{aligned}$$

□

Example 3.1.8 (Dot product with a fixed vector) Fix a vector $\mathbf{u} \in \mathbb{R}^3$. The function

$$\begin{aligned} D : \mathbb{R}^3 &\to \mathbb{R} \\ \mathbf{v} &\mapsto \mathbf{u} \cdot \mathbf{v} \end{aligned}$$

(here $\cdot$ is the dot product of vectors, not scalar multiplication!) is a linear map, because of the properties

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v}_1 + \mathbf{v}_2) &= \mathbf{u} \cdot \mathbf{v}_1 + \mathbf{u}.\mathbf{v}_2 \\ \mathbf{u}.(k\mathbf{v}) &= k\mathbf{u} \cdot \mathbf{v}. \end{aligned}$$

We will soon see that *all* linear maps $\mathbb{R}^3 \to \mathbb{R}$ (indeed all linear maps $\mathbb{R}^n \to \mathbb{R}$) are of this form.

□

Example 3.1.9 (Differentiation) The operation 'take the derivative' can be interpreted as a linear map

$$\begin{aligned} D : \mathrm{Poly}_n &\to \mathrm{Poly}_{n-1} \\ \mathbf{p} &\mapsto \mathbf{p}'\,. \end{aligned}$$

For example, $D(2x^3 - 6x + 2) = 6x - 6$. □

Checkpoint 3.1.10 (a) Why is D a map from Poly_n to Poly_{n-1}? (b) Check that D is linear.

Solution.

1. As you know from calculus, taking the derivative of a polynomial decreases every power of x by 1. So if $\mathbf{p}$ is in Poly_n, then $\mathbf{p}$ has degree at most n . Therefore, its image under D has degree at most $n-1$. Thus $\mathbf{p}'$ is in Poly_{n-1}.

2. Let $\mathbf{p}, \mathbf{q}$ be in Poly_n. For concreteness, we write

$$\mathbf{p} = \sum_{j=0}^{n} a_j x^j \mathbf{q} = \sum_{j=0}^{n} b_j x^j.$$

We check additivity of D:

$$D(\mathbf{p}+\mathbf{q}) = D\left(\sum_{j=0}^{n}(a_j+b_j)x^j\right) = \sum_{j=1}^{n}(a_j+b_j)jx^{j-1} = \sum_{j=1}^{n} ja_jx^{j-1}+\sum_{j=0}^{n} jb_jx^{j-1} = D(\mathbf{p})+D(\mathbf{q}).$$

Next we check how D interacts with scalar multiplication:

$$D(k\mathbf{p}) = D\left(\sum_{j=0}^{n} ka_jx^j\right) = \sum_{j=1}^{n} ka_jx^{j-1} = k\left(\sum_{j=1}^{n} a_jx^{j-1}\right) = kD(\mathbf{p}).$$

And so we conclude that $D$is linear.

Example 3.1.11 (Antiderivative) The operation 'find the unique antiderivative with zero constant term' can be interpreted as a linear map

$$\begin{aligned} A : \mathrm{Poly}_n &\to \mathrm{Poly}_{n+1} \\ \mathbf{p} &\mapsto \int_0^x \mathbf{p}(t)\,dt \end{aligned}$$

For example, $A(2x^3 - 6x + 2) = 4x^4 - 3x^2 + 2x$. □

Checkpoint 3.1.12 (a) Check that A is indeed a linear map. (b) Why is A a map from Poly_n to Poly_{n+1}? (c) Check that A is linear.

Solution.

1.

2. You know from calculus that the antiderivative of a polynomial $\mathbf{p}$ must always have degree one greater than $\mathbf{p}$. Hence A maps Poly_n to Poly_{n+1}.

3. Let $\mathbf{p}, \mathbf{q}$ be in Poly_n. Using the usual properties of the integral, we compute

$$A(\mathbf{p}+\mathbf{q}) = \int_0^x \mathbf{p}(t)+\mathbf{q}(t)\,dt = \int_0^x \mathbf{p}(t)\,dt + \int_0^x \mathbf{q}(t)\,dt = A(\mathbf{p}) + A(\mathbf{q}).$$

Similarly,

$$A(k\mathbf{p}) = k\int_0^x \mathbf{p}(t)\,dt = \int_0^x k\mathbf{p}(t)\,dt = A(k\mathbf{p}).$$

Example 3.1.13 Shift map. Define the 'shift forward by 1' map

$$\begin{aligned} S : \mathrm{Poly}_n &\to \mathrm{Poly}_n \\ \mathbf{p} &\mapsto S(\mathbf{p}) \end{aligned}$$

by $S(\mathbf{p})(x) = \mathbf{p}(x-1)$.

Consider the case $n = 3$. In terms of the standard basis

$$\mathbf{p}_0(x) = 1,\ \mathbf{p}_1(x) = x,\ \mathbf{p}_2(x) = x^2,\ \mathbf{p}_3(x) = x^3$$

of Poly_3, we have:

$$\begin{aligned} S(\mathbf{p}_0) &= \mathbf{p}_0 \\ S(\mathbf{p}_1) &= \mathbf{p}_1 - \mathbf{p}_0 \\ S(\mathbf{p}_2) &= \mathbf{p}_2 - 2\mathbf{p}_1 + \mathbf{p}_0 \end{aligned}$$

$$S(\mathbf{p}_3) = \mathbf{p}_3 - 3\mathbf{p}_2 + 3\mathbf{p}_1 - \mathbf{p}_0$$

□

Checkpoint 3.1.14 Check that S is a linear map.

Solution. Let

$$\mathbf{p} = \sum_{j=0}^{n} a_j x^j \mathbf{q} = \sum_{j=0}^{n} b_j x^j$$

$$S(k\mathbf{p}) = S\left(\sum_{j=0}^{n} k a_j x^j\right) = \sum_{j=0}^{n} k a_j (x-1)^j = k \sum_{j=0}^{n} a_j (x-1)^j = kS(\mathbf{p}).$$

$$S(\mathbf{p}+\mathbf{q}) = S\left(\sum_{j=0}^{n} (a_j + b_j) x^j\right) = \sum_{j=0}^{n} (a_j+b_j)(x-1)^j = \sum_{j=0}^{n} a_j (x-1)^j + \sum_{j=0}^{n} b_j (x-1)^j = S(\mathbf{p})+S(\mathbf{q}).$$

Checkpoint 3.1.15 Check this.

Solution. $S(\mathbf{p}_0) = \mathbf{p}_0$ is trivial.

$$S(\mathbf{p}_1) = x-1 = \mathbf{p}_1 - \mathbf{p}_0 S(\mathbf{p}_2) = (x-1)^2 = x^2-2x+1 = \mathbf{p}_2 - 2\mathbf{p}_1 + \mathbf{p}_0 S(\mathbf{p}_3) = (x-1)^3 = x^3-3x^2+3x-1 = \mathbf{p}_3 - 3\mathbf{p}$$

Example 3.1.16 Matrices give rise to linear maps. Every $n \times m$ matrix A induces a *linear map*

$$\begin{aligned} T_{\mathsf{A}} : \mathrm{Col}_m &\to \mathrm{Col}_n \\ \mathbf{v} &\mapsto \mathsf{A}\mathbf{v}\,. \end{aligned}$$

That is, $T_{\mathsf{A}}(\mathbf{v}) := \mathsf{A}\mathbf{v}$ is the matrix product of A with the column vector $\mathbf{v}$. The fact that T_{A} is indeed a linear map follows from the linearity of matrix multiplication (Proposition A.0.4 parts 2 and 3).

> Note that an $n \times m$ matrix gives a linear map from Col_m to Col_n!

□

Lemma 3.1.17 *Suppose $T : V \to W$ is a linear map. Then*

1. $T(\mathbf{0}_V) = \mathbf{0}_W$.
2. $T(-\mathbf{v}) = -T(\mathbf{v})$ *for all vectors* $\mathbf{v} \in V$.

Proof.

$$\begin{aligned} \text{(i)} T(\mathbf{0}_V) &= T(0\mathbf{0}_V) && (\text{R8 applied to } \mathbf{v} = \mathbf{0}_V \in V) \\ &= 0T(\mathbf{0}_V) && (\text{T is linear}) \\ &= \mathbf{0}_W && (\text{R8 applied to } \mathbf{v} = T(\mathbf{0}_V) \in W) \end{aligned}$$

$$\begin{aligned} \text{(ii)} T(-\mathbf{v}) &= T((-1)\mathbf{v}) && (\text{defn of } -\mathbf{v} \text{ in } V) \\ &= (-1)T(\mathbf{v}) && (\text{T is linear}) \\ &= -T(\mathbf{v}) && (\text{defn of } -T(\mathbf{v}) \text{ in } W) \end{aligned}$$

■

The next result is very important. It tells us that if we know how a linear map T acts on a basis, then we know how it acts on the whole vector space (this is the 'uniqueness' part). Moreover, we are free to write down any willy-nilly formula for what T does on the basis vectors, and we are guaranteed that this will always extend to a linear map defined on the whole vector space (this is the 'existence' part).

Proposition 3.1.18 Sufficient to Define a Linear Map on a Basis. *Suppose $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ is a basis for V and $\mathbf{w}_1, \ldots, \mathbf{w}_m$ are vectors in W. Then there exists a unique linear map $T : V \to W$ such that*

$$T(\mathbf{e}_i) = \mathbf{w}_i, i = 1 \ldots m.$$

Proof. Proof of existence. To define a linear map T, we must define $T(\mathbf{v})$ for each vector $\mathbf{v}$. We can write $\mathbf{v}$ in terms of its coordinate vector $[\mathbf{v}]_{\mathcal{B}}$ with respect to the basis $\mathcal{B}$ as

$$\mathbf{v} = [\mathbf{v}]_{\mathcal{B},1}\mathbf{e}_1 + \cdots + [\mathbf{v}]_{\mathcal{B},m}\mathbf{e}_m \tag{3.1.1}$$

where $[\mathbf{v}]_{\mathcal{B},i}$ is the entry at row i of the coordinate vector $[\mathbf{v}]_{\mathcal{B}}$. We define

$$T(\mathbf{v}) := [\mathbf{v}]_{\mathcal{B},1}\mathbf{w}_1 + [\mathbf{v}]_{\mathcal{B},2}\mathbf{w}_2 + \cdots [\mathbf{v}]_{\mathcal{B},m}\mathbf{w}_m. \tag{3.1.2}$$

We clearly have $T(\mathbf{e}_i) = \mathbf{w}_i$. To complete the proof of existence, we must show that T is linear:

$$\begin{aligned} T(\mathbf{v} + \mathbf{v}') &= [\mathbf{v} + \mathbf{v}']_{\mathcal{B},1}\mathbf{w}_1 + \cdots + [\mathbf{v} + \mathbf{v}']_{\mathcal{B},m}\mathbf{w}_m \\ &= \big([\mathbf{v}]_{\mathcal{B},1} + [\mathbf{v}']_{\mathcal{B},1}\big)\mathbf{w}_1 + \cdots + \big([\mathbf{v}]_{\mathcal{B},m} + [\mathbf{v}']_{\mathcal{B},m}\big)\mathbf{w}_m \quad ([\mathbf{v} + \mathbf{v}']_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}} + [\mathbf{v}']_{\mathcal{B}}) \\ &= [\mathbf{v}]_{\mathcal{B},1}\mathbf{w}_1 + \cdots + [\mathbf{v}]_{\mathcal{B},m}\mathbf{w}_m + \\ &\qquad [\mathbf{v}']_{\mathcal{B},1}\mathbf{w}_1 + \cdots + [\mathbf{v}']_{\mathcal{B},m}\mathbf{w}_m \\ &= T(\mathbf{v}) + T(\mathbf{v}'). \end{aligned}$$

Similarly, one can check that $T(k\mathbf{v}) = kT(\mathbf{v})$, which completes the proof of existence.

Proof of uniqueness. Suppose that $S, T : V \to W$ are linear maps with

$$S(\mathbf{e}_i) = \mathbf{w}_i, \text{ and } T(\mathbf{e}_i) = \mathbf{w}_i, i = 1 \ldots m. \tag{3.1.3}$$

Then,

$$\begin{aligned} S(\mathbf{v}) &= S\big([\mathbf{v}]_{\mathcal{B},1}\mathbf{e}_1 + \cdots + [\mathbf{v}]_{\mathcal{B},m}\mathbf{e}_m\big) \\ &= [\mathbf{v}]_{\mathcal{B},1}S(\mathbf{e}_1) + \cdots + [\mathbf{v}]_{\mathcal{B},m}S(\mathbf{e}_m) && (S \text{ is linear}) \\ &= [\mathbf{v}]_{\mathcal{B},1}\mathbf{w}_1 + \cdots + [\mathbf{v}]_{\mathcal{B},m}\mathbf{w}_m && (\text{since } S(\mathbf{e}_i) = \mathbf{w}_i) \\ &= [\mathbf{v}]_{\mathcal{B},1}T(\mathbf{e}_1) + \cdots + [\mathbf{v}]_{\mathcal{B},m}T(\mathbf{e}_m) && (\text{since } T(\mathbf{e}_i) = \mathbf{w}_i) \\ &= T([\mathbf{v}]_{\mathcal{B},1}\mathbf{e}_1 + \cdots + [\mathbf{v}]_{\mathcal{B},m}\mathbf{e}_m) && (T \text{ is linear}) \\ &= T(\mathbf{v}). \end{aligned}$$

Hence $S = T$, in other words the linear map satisfying (3.1.3) is unique. ■

Example 3.1.19 As an application of Proposition 3.1.18, we can define a linear map

$$T : \text{Col}_2 \to \text{Fun}(\mathbb{R})$$

simply by defining its action on the standard basis of Col_2. For instance, we may set

$$\mathsf{e}_1 \to f_1$$

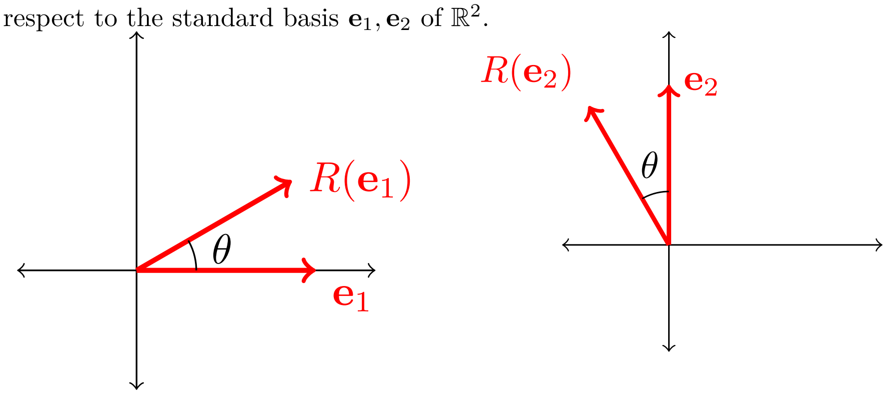
$$\mathsf{e}_2 \to f_2$$

The point is that we are free to send e_1 and e_2 to *any functions* f_1 and f_2 we like, and we are assured that this will give a well-defined linear map $T : \mathrm{Col}_2 \to \mathrm{Fun}(\mathbb{R})$. For instance, we might set $f_1(x) = \sin x$ and $f_2(x) = |x|$. Then the general formula for T is

$$\left(T \begin{bmatrix} a \\ b \end{bmatrix}\right)(x) = a \sin x + b|x|$$

$\square$

Example 3.1.20 Rotation map on standard basis. Let us compute the action of the 'counterclockwise rotation by θ' map R from Example 3.1.6 with respect to the standard basis $\mathbf{e}_1, \mathbf{e}_2$ of $\mathbb{R}^2$.



From the figure, we have:

$$R(\mathbf{e}_1) = (\cos\theta, \sin\theta) \qquad\qquad R(\mathbf{e}_2) = (-\sin\theta, \cos\theta)$$

so that

$$R(\mathbf{e}_1) = \cos\theta\, \mathbf{e}_1 + \sin\theta\, \mathbf{e}_2, R(\mathbf{e}_2) = -\sin\theta\, \mathbf{e}_1 + \cos\theta\, \mathbf{e}_2$$

Now that we know the action of R on the standard basis vectors, we can compute its action on an arbitrary vector $(x, y) \in \mathbb{R}^2$:

$$\begin{aligned} R\left((x, y)\right) &= R(x\mathbf{e}_1 + y\mathbf{e}_2) \\ &= xR(\mathbf{e}_1) + yR(\mathbf{e}_2) \\ &= x(\cos\theta, \sin\theta) + y(-\sin\theta, \cos\theta) \\ &= (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta). \end{aligned}$$

$\square$

Exercises

1. Let V be a vector space, and let $\mathbf{a} \neq \mathbf{0}$ be a fixed vector. Let Define the map T as follows:

$$\begin{aligned} T : V &\to V \\ \mathbf{v} &\mapsto \mathbf{a} + \mathbf{v} \end{aligned}$$

(a) Is T a linear map? (Yes or no)

(b) Prove your assertion from (a).

Solution.

(a) No.

(b) According to Lemma 3.1.17, $T(0) = 0$ is a necessary condition for T to be linear. However,

$$T(\mathbf{0}) = \mathbf{a} + 0 = \mathbf{a} \neq \mathbf{0}.$$

Hence T cannot be linear.

2. Consider the map $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by $T\left((x, y, z)\right) = (z, x, y)$.

(a) Is T a linear map? (Yes or no)

(b) Prove your assertion from (a).

Solution.

(a) Yes.

(b)

$$T\left((x, y, z) + (a, b, c)\right) = T\left((x + a, y + b, z + c)\right) = (z+c, x+a, y+b) = (z, x, y)+(c, a, b) = T\left((x, y, z)\right)$$

$$T(k(x, y, z)) = T(kx, ky, kz) = (kz, kx, ky) = k(z, x, y) = kT((x, y, z)).$$

3. Define the 'multiplication by x^2' map

$$\begin{aligned} M : \mathrm{Poly}_n &\to \mathrm{Poly}_{n+2} \\ \mathbf{p} &\mapsto M(\mathbf{p}) \end{aligned}$$

where $M(\mathbf{p})(x) = x^2\mathbf{p}(x)$.

(a) Why does M map from Poly_n to Poly_{n+2}?

(b) Prove that M is linear.

(c) Compute the action of M in the standard basis for Poly_3, as in Example 3.1.13.

Solution.

(a) If $p(x)$ has degree at most n, then $\mathbf{q}(x) := x^2\mathbf{p}(x)$ has degree at most $n + 2$.

(b) The proof is simple and follows from the usual properties of polynomials. Using the fact that multiplication of polynomials distributes over addition we compute

$$M(\mathbf{p}+\mathbf{q})(x) = x^2(\mathbf{p}(x)+\mathbf{q}(x)) = x^2\mathbf{p}(x)+x^2\mathbf{q}(x) = M(\mathbf{p})(x)+M(\mathbf{q})(x).$$

We may consider the scalar k as a constant polynomial. Thus, using the commutativity and associativity of polynomial multiplication we compute

$$M(k\mathbf{p})(x) = x^2k\mathbf{p}(x) = k(x^2\mathbf{p}(x)) = kM(\mathbf{p})(x).$$

(c)

$$M(\mathbf{p}_0) = x^21 = x^2 = \mathbf{p}_2 M(\mathbf{p}_1) = x^2x = x^3 = \mathbf{p}_3 M(\mathbf{p}_2) = x^2x^2 = x^4 = \mathbf{p}_4 M(\mathbf{p}_3) = x^2x^3 = x^5 = \mathbf{p}$$

4. Define the 'integrate over the interval $[-1,1]$' map

$$
\begin{aligned}
I : \mathrm{Poly}_n &\to \mathbb{R} \\
\mathbf{p} &\mapsto \int_{-1}^{1} \mathbf{p}(x)\, dx
\end{aligned}
$$

(a) Prove that I is linear.

(b) Compute the action of I with respect to the standard basis $\mathbf{p}_0, \ldots, \mathbf{p}_3$ for Poly_3.

(c) Compute the action of I with respect to the basis $\mathbf{q}_0, \ldots, \mathbf{q}_3$ for Poly_3 from Example 2.2.7.

Solution.

(a) We use the properties of the integral.

$$I(\mathbf{p}+\mathbf{q}) = \int_{-1}^{1} (\mathbf{p}(x) + \mathbf{q}(x))\ dx = \int_{-1}^{1} \mathbf{p}(x)\, dx + \int_{-1}^{1} \mathbf{q}(x)\, dx = I(\mathbf{p}) + I(\mathbf{q}).$$

$$I(k\mathbf{p}) = \int_{-1}^{1} k\mathbf{p}(x)\, dx = k \int_{-1}^{1} \mathbf{p}(x)\, dx = kI(\mathbf{p})$$

(b)

$$I(\mathbf{p}_0) = \int_{-1}^{1} 1dx = x|_{-1}^{1} = 2I(\mathbf{p}_1) = \int_{-1}^{1} xdx = \frac{1}{2}x^2|_{-1}^{1} = 0I(\mathbf{p}_2) = \int_{-1}^{1} x^2dx = \frac{1}{3}x^3|_{-1}^{1} = \frac{2}{3}I(\mathbf{p}_3) =$$

(c) We use the results above:

$$I(\mathbf{q}_0) = \int_{-1}^{1} 1dx = x|_{-1}^{1} = 2I(\mathbf{q}_1) = \int_{-1}^{1} xdx = \frac{1}{2}x^2|_{-1}^{1} = 0I(\mathbf{q}_2) = \int_{-1}^{1} 2x^2-1dx = 2\int_{-1}^{1} x^2dx - \int_{-1}^{1}$$

5. Compute the action of the differentiation map $D : \mathrm{Poly}_4 \to \mathrm{Poly}_3$ from Example 3.1.9 with respect to the standard bases of these two vector spaces.

Solution.

$$D(\mathbf{p}_0) = D(1) = 0D(\mathbf{p}_1) = D(x) = 1 = \mathbf{p}_0D(\mathbf{p}_2) = D(x^2) = 2x = 2\mathbf{p}_1D(\mathbf{p}_3) = D(x^3) = 3x^2 = 3\mathbf{p}_2D(\mathbf{p}_4)$$

6. Consider the cross-product linear map $C : \mathbb{R}^3 \to \mathbb{R}^3$ from Example 3.1.7 in the case $\mathbf{w} = (1,2,3)$. Compute the action of C with respect to the standard basis of $\mathbb{R}^3$.

Solution.

$$C((1,0,0)) = (1,2,3)\times(1,0,0) = (0,3,-2)C((0,1,0)) = (1,2,3)\times(0,1,0) = (-3,0,1)C((0,0,1)) = (1,2,3$$

Solutions

· Exercises

3.1.1. Solution.

(a) No.

(b) According to Lemma 3.1.17, $T(0) = 0$ is a necessary condition for T to be linear. However,

$$T(\mathbf{0}) = \mathbf{a} + 0 = \mathbf{a} \neq \mathbf{0}.$$

Hence T cannot be linear.

3.1.2. Solution.

(a) Yes.

(b)

$$T\left((x,y,z)+(a,b,c)\right) = T\left((x+a,y+b,z+c)\right) = (z+c,x+a,y+b) = (z,x,y)+(c,a,b) = T\left((x,y,z)\right)+T$$

$$T(k(x,y,z)) = T(kx,ky,kz) = (kz,kx,ky) = k(z,x,y) = kT((x,y,z)).$$

3.1.3. Solution.

(a) If $p(x)$ has degree at most n, then $\mathbf{q}(x) := x^2\mathbf{p}(x)$ has degree at most $n+2$.

(b) The proof is simple and follows from the usual properties of polynomials. Using the fact that multiplication of polynomials distributes over addition we compute

$$M(\mathbf{p}+\mathbf{q})(x) = x^2(\mathbf{p}(x)+\mathbf{q}(x)) = x^2\mathbf{p}(x)+x^2\mathbf{q}(x) = M(\mathbf{p})(x)+M(\mathbf{q})(x).$$

We may consider the scalar k as a constant polynomial. Thus, using the commutativity and associativity of polynomial multiplication we compute

$$M(k\mathbf{p})(x) = x^2k\mathbf{p}(x) = k(x^2\mathbf{p}(x)) = kM(\mathbf{p})(x).$$

(c)

$$M(\mathbf{p}_0) = x^21 = x^2 = \mathbf{p}_2 M(\mathbf{p}_1) = x^2x = x^3 = \mathbf{p}_3 M(\mathbf{p}_2) = x^2x^2 = x^4 = \mathbf{p}_4 M(\mathbf{p}_3) = x^2x^3 = x^5 = \mathbf{p}_5.$$

3.1.4. Solution.

(a) We use the properties of the integral.

$$I(\mathbf{p}+\mathbf{q}) = \int_{-1}^{1} (\mathbf{p}(x)+\mathbf{q}(x))\ dx = \int_{-1}^{1} \mathbf{p}(x)\,dx+\int_{-1}^{1} \mathbf{q}(x)\,dx = I(\mathbf{p})+I(\mathbf{q}).$$

$$I(k\mathbf{p}) = \int_{-1}^{1} k\mathbf{p}(x)\,dx = k\int_{-1}^{1} \mathbf{p}(x)\,dx = kI(\mathbf{p})$$

(b)

$$I(\mathbf{p}_0) = \int_{-1}^{1} 1dx = x|_{-1}^{1} = 2I(\mathbf{p}_1) = \int_{-1}^{1} xdx = \frac{1}{2}x^2|_{-1}^{1} = 0I(\mathbf{p}_2) = \int_{-1}^{1} x^2dx = \frac{1}{3}x^3|_{-1}^{1} = \frac{2}{3}I(\mathbf{p}_3) = \int_{-1}^{1}$$

(c) We use the results above:

$$I(\mathbf{q}_0) = \int_{-1}^{1} 1dx = x|_{-1}^{1} = 2I(\mathbf{q}_1) = \int_{-1}^{1} xdx = \frac{1}{2}x^2|_{-1}^{1} = 0I(\mathbf{q}_2) = \int_{-1}^{1} 2x^2-1dx = 2\int_{-1}^{1} x^2dx-\int_{-1}^{1} 1 =$$

3.1.5. Solution.

$$D(\mathbf{p}_0) = D(1) = 0D(\mathbf{p}_1) = D(x) = 1 = \mathbf{p}_0 D(\mathbf{p}_2) = D(x^2) = 2x = 2\mathbf{p}_1 D(\mathbf{p}_3) = D(x^3) = 3x^2 = 3\mathbf{p}_2 D(\mathbf{p}_4) = D$$

3.1.6. Solution.

$$C((1,0,0)) = (1,2,3)\times(1,0,0) = (0,3,-2)C((0,1,0)) = (1,2,3)\times(0,1,0) = (-3,0,1)C((0,0,1)) = (1,2,3)\times(0$$

3.2 Composition of linear maps

Definition 3.2.1 If $S : U \to V$ and $T : V \to W$ are linear maps, then the **composition of T with S** is the map $T \circ S : U \to W$ defined by

$$(T \circ S)(\mathbf{u}) := T(S(\mathbf{u}))$$

where $\mathbf{u}$ is in U. $\Diamond$

See Figure 3.2.2. tikzalign

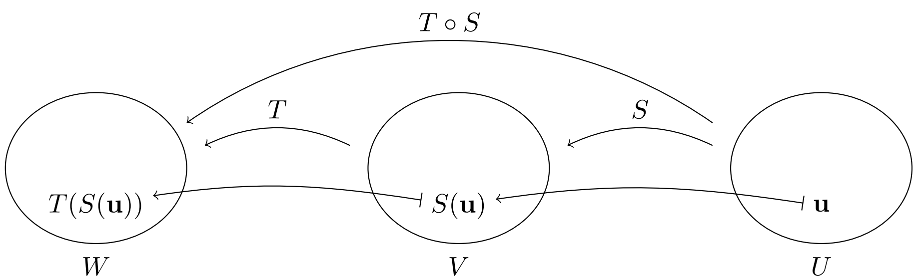


Figure 3.2.2: Composition of linear maps.

The standard convention in mathematics is to write the evaluation of a function from right-to-left, eg. $f(x)$. That is, you start with the right hand symbol x, then you apply f. So the most natural way to draw these pictures is from right-to-left!

Example 3.2.3 Let $S : \mathbb{R}^3 \to \mathrm{Poly}_2$ and $T : \mathrm{Poly}_2 \to \mathrm{Poly}_4$ be the linear maps defined by

$$S((a,b,c)) := ax^2 + (a-b)x + c, T(p(x)) = x^2 p(x)$$

Then $T \circ S$ can be computed as follows:

$$\begin{aligned}(T \circ S)((a,b,c)) &= T(S((a,b,c))) \\ &= T(ax^2 + (a-b)x + c) \\ &= x^2(ax^2 + (a-b)x + c) \\ &= ax^4 + (a-b)x^3 + cx^2.\end{aligned}$$

$\square$

Proposition 3.2.4 *If $S : U \to V$ and $T : V \to W$ are linear maps, then $T \circ S : U \to W$ is also a linear map.*

Proof. Let $\mathbf{u_1}, \mathbf{u}_2 \in U$. Then:

$$\begin{aligned}(T \circ S)(\mathbf{u}_1 + \mathbf{u}_2) &= T(S(\mathbf{u}_1 + \mathbf{u}_2)) && \text{(defn of } T \circ S\text{)} \\ &= T(S(\mathbf{u}_1) + S(\mathbf{u}_2)) && (S \text{ is linear})\end{aligned}$$

$$
\begin{aligned}
&= T(S(\mathbf{u}_1)) + T(S(\mathbf{u}_2)) && (T \text{ is linear}) \\
&= (T \circ S)(\mathbf{u}_1) + (T \circ S)(\mathbf{u}_2) && (\text{defn of } T \circ S)
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
(T \circ S)(k\mathbf{u}) &= T(S(k\mathbf{u})) && (\text{defn of } T \circ S) \\
&= T(kS(\mathbf{u})) && (S \text{ is linear}) \\
&= kT(S(\mathbf{u})) && (T \text{ is linear}) \\
&= k(T \circ S)(\mathbf{u}) && (\text{defn of } T \circ S)
\end{aligned}
$$

■

Example 3.2.5 Consider the antiderivative (A) and derivative (D) linear maps

$$
\begin{aligned}
A &: \mathrm{Poly}_n \to \mathrm{Poly}_{n+1} \\
D &: \mathrm{Poly}_{n+1} \to \mathrm{Poly}_n .
\end{aligned}
$$

Is $D \circ A = \mathrm{id}_{\mathrm{Poly}_n}$?**Solution**. We compute the action of $D \circ A$ on the basis x^k, $k = 0 \ldots n$ of Poly_n:

$$x^k \overset{A}{\mapsto} \frac{x^{k+1}}{k+1} \overset{D}{\mapsto} \frac{k+1}{k+1} x^k = x^k$$

Hence for $k = 0 \ldots n$,

$$
\begin{aligned}
(D \circ A)(x^k) &= x^k \\
&= \mathrm{id}_{\mathrm{Poly}_n}(x^k).
\end{aligned}
$$

Since $D \circ A$ and $\mathrm{id}_{\mathrm{Poly}_n}$ agree on a basis for Poly_n, they agree on all vectors $\mathbf{p} \in \mathrm{Poly}_n$ by Proposition 3.1.18. Hence $D \circ A = \mathrm{id}_{\mathrm{Poly}_n}$.

> In fact, the statement that $D \circ A = \mathrm{id}_{\mathrm{Poly}_n}$ is precisely Part I of the Fundamental Theorem of Calculus, applied to the special case of polynomials!

□

Checkpoint 3.2.6 Is $A \circ D = \mathrm{id}_{\mathrm{Poly}_{n+1}}$? If it is, prove it. If it is not, give an explicit counterexample.

Solution. The statement is not true! For example, let $\mathbf{p}(x) = x + 1$. Then

$$(A \circ D)(x + 1) = A(1) = x \neq x + 1.$$

Exercises

1. Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be the 'rotation by θ' map from Example 3.1.20,

$$R((x, y)) = (x \cos\theta - y \sin\theta,\ x \sin\theta + y \cos\theta).$$

Check algebraically that $R_\phi \circ R_\theta = R_{\phi+\theta}$ by computing the action of the linear maps on both sides of this equation on an arbitrary vector $(x, y) \in \mathbb{R}^2$.

Solution.

$$R_\phi R_\theta(x, y) = R_\phi(x \cos\theta - y \sin\theta, x \sin\theta + y \cos\theta)$$

$$
\begin{aligned}
&= ((x\cos\theta - y\sin\theta)\cos\phi - (x\sin\theta + y\cos\theta)\sin\phi,\ (x\cos\theta - y\sin\theta)\sin\phi + (x\sin\theta + y\cos\theta) \\
&= (x(\cos\theta\cos\phi - \sin\theta\sin\phi) - y(\sin\theta\cos\phi + \cos\theta\sin\phi),\ x(\cos\theta\sin\phi + \sin\theta\cos\phi) + y(\cos\theta \\
&= \big(x\cos(\theta+\phi) - y\sin(\theta+\phi),\ x\sin(\theta+\phi) + y\cos(\theta+\phi_{)}\big) \\
&= R_{\phi+\theta}(x,y)
\end{aligned}
$$

2. Let $M : \mathrm{Poly}_3 \to \mathrm{Poly}_4$ be the 'multiplication by x' map, $M(p)(x) = xp(x)$. Let $S : \mathrm{Poly}_4 \to \mathrm{Poly}_4$ be the map $S(p(x)) = p(x-1)$. Similarly let $T : \mathrm{Poly}_3 \to \mathrm{Poly}_3$ be the map $T(p(x)) = p(x-1)$. Compute $S \circ M$ and $M \circ T$. Are they equal?

Solution.

$$(S \circ M)(p(x)) = S(M(p(x))) = S(xp(x)) = (x-1)p(x-1)$$

whereas

$$(M \circ S)(p(x)) = M(p(x-1)) = xp(x-1).$$

Thus $S \circ M \neq M \circ S$.

Solutions

· Exercises

3.2.1. Solution.

$$
\begin{aligned}
R_\phi R_\theta(x,y) &= R_\phi(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) \\
&= ((x\cos\theta - y\sin\theta)\cos\phi - (x\sin\theta + y\cos\theta)\sin\phi,\ (x\cos\theta - y\sin\theta)\sin\phi + (x\sin\theta + y\cos\theta)\cos\phi \\
&= (x(\cos\theta\cos\phi - \sin\theta\sin\phi) - y(\sin\theta\cos\phi + \cos\theta\sin\phi),\ x(\cos\theta\sin\phi + \sin\theta\cos\phi) + y(\cos\theta\cos\phi \\
&= \big(x\cos(\theta+\phi) - y\sin(\theta+\phi),\ x\sin(\theta+\phi) + y\cos(\theta+\phi_{)}\big) \\
&= R_{\phi+\theta}(x,y)
\end{aligned}
$$

3.2.2. Solution.

$$(S \circ M)(p(x)) = S(M(p(x))) = S(xp(x)) = (x-1)p(x-1)$$

whereas

$$(M \circ S)(p(x)) = M(p(x-1)) = xp(x-1).$$

Thus $S \circ M \neq M \circ S$.

3.3 Isomorphisms of vector spaces

Suppose you have two sets,

$$A = \{\mathrm{bird}, \mathrm{eye}, \mathrm{person}\} \mathrm{and} B = \{, , \}.$$

The elements of A and B are not *the same*, so A is not *equal* to B. But this is unsatisfactory — clearly the elements of A are just English versions of the Chinese symbols in B. How can we make this mathematically precise?

We could define two maps, say

$$
\begin{aligned}
S : A &\to B \\
\mathrm{bird} &\mapsto \\
\mathrm{eye} &\mapsto
\end{aligned}
$$

$$\text{cross} \mapsto$$

and

$$\begin{aligned} T : B &\to A \\ &\mapsto \text{bird} \\ &\mapsto \text{eye} \\ &\mapsto \text{cross.} \end{aligned}$$

Then we observe that

$$T \circ S = \mathrm{id}_A \,\text{and} S \circ T = \mathrm{id}_B \,. \tag{3.3.1}$$

A pair of maps $S : A \to B$ and $T : B \to A$ satisfying (3.3.1) is called an *isomorphism of sets* between A and B. If you like, you can rename T as S^{-1} since $S^{-1} \circ S = \mathrm{id}_A$ and $S \circ S^{-1} = \mathrm{id}_B$. (Calling T by the name S^{-1} from the beginning would have been presumptive of me. I needed to first define it, and then check that it satisfied (3.3.1). Only then did I have the right to call it S^{-1}!)

Perhaps you are somewhat of a penny-pincher. You see the need for the English-to-Chinese map S, but not the need for a Chinese-to-English map T. After all, you say, since no two different English symbols in A get mapped to the same Chinese symbol in B ('S is one-to-one') and every Chinese symbol $y \in B$ is equal to $S(x)$ for some $x \in A$ ('S is onto'), we have no need for T. It is an extravagance!

To this I respond: you are right, but is it not useful to have the explicit Chinese-to-English map T? In bookshops, cross-language dictionaries like this most often come bundled as a pair, in a single volume. After all, if one needs to look up the English word for , it is a nuisance to have to traverse through the entire English-to-Chinese dictionary, trying to find the English word which translates to !

This motivates the following definition.

Definition 3.3.1 We say that a linear map $S : V \to W$ is an **isomorphism** if there exists a linear map $T : W \to V$ such that

$$T \circ S = \mathrm{id}_V \,\text{and} S \circ T = \mathrm{id}_W \,. \tag{3.3.2}$$

$\Diamond$

Lemma 3.3.2 Uniqueness of Inverses. *If $S : V \to W$ is an isomorphism and $T, T' : W \to V$ both satisfy*

$$\begin{aligned} T \circ S &= \mathrm{id}_V, & S \circ T &= \mathrm{id}_W \\ T' \circ S &= \mathrm{id}_V, & S \circ T' &= \mathrm{id}_W \end{aligned}$$

then $T = T'$.

This lemma justifies us calling T *the* inverse of S (as opposed to *an* inverse of S), and also justifies us writing $T = S^{-1}$.

Proof. To show that $T = T'$, we must show that for all $\mathbf{w} \in W$, $T(\mathbf{w}) = T'(\mathbf{w})$. Indeed:

$$T(\mathbf{w}) = T(\mathrm{id}_W(\mathbf{w})) \qquad \text{(Defn of } \mathrm{id}_W)$$

$$
\begin{aligned}
&= T((S \circ T')(\mathbf{w})) && (S \circ T' = \mathrm{id}_W)\\
&= T(S(T'(\mathbf{w}))) && (\text{Defn of } S \circ T')\\
&= (T \circ S)(T'(\mathbf{w})) && (\text{Defn of } T \circ S)\\
&= \mathrm{id}_V(T'(\mathbf{w})) && (T \circ S = \mathrm{id}_V)\\
&= T'(\mathbf{w}) && (\text{Defn of } \mathrm{id}_V).
\end{aligned}
$$

■

Definition 3.3.3 We say that two vector spaces V and W are **isomorphic** if there exists an isomorphism between them. ♢

Example 3.3.4 Show that $\mathbb{R}^n$ is isomorphic to Poly_{n-1}

Solution. We define a pair of linear maps

$$S : \mathbb{R}^n \rightleftarrows \mathrm{Poly}_{n-1} : T$$

as follows:

$$
\begin{aligned}
(a_1, a_2, \ldots, a_n) &\overset{S}{\mapsto} a_1 + a_2 x + \cdots + a_n x^{n-1}\\
(a_1, a_2, \ldots, a_n) &\overset{T}{\hookleftarrow} a_1 + a_2 x + \cdots + a_n x^{n-1}
\end{aligned}
$$

We clearly have $T \circ S = \mathrm{id}_{\mathbb{R}^n}$ and $S \circ T = \mathrm{id}_{\mathrm{Poly}_{n-1}}$. □

Checkpoint 3.3.5 Check that these maps are linear.

We will now show that up to isomorphism, there is only one vector space of each dimension!

Theorem 3.3.6 *Two finite-dimensional vector spaces V and W are isomorphic if and only if they have the same dimension.*

Proof. $\Rightarrow$. Suppose V and W are isomorphic, via a pair of linear maps $S : V \rightleftarrows W : T$. Let $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ be a basis for V. Then I claim that $\mathcal{C} = \{S(\mathbf{e}_1), \ldots, S(\mathbf{e}_m)\}$ is a basis for W. Indeed, the list of vectors $\mathcal{C}$ is linearly independent, since if

$$a_1 S(\mathbf{e}_1) + a_2 S(\mathbf{e}_2) + \cdots + a_m S(\mathbf{e}_m) = \mathbf{0}_W,$$

then applying T to both sides we obtain

$$
\begin{aligned}
T(a_1 S(\mathbf{e}_1) + a_2 S(\mathbf{e}_2) + \cdots + a_m S(\mathbf{e}_m)) &= T(\mathbf{0}_W)\\
\therefore a_1 T(S(\mathbf{e}_1)) + a_2 T(S(\mathbf{e}_2)) + \cdots + a_m T(S(\mathbf{e}_m)) &= \mathbf{0}_V && (T \text{ is linear})\\
\therefore a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \cdots + a_m \mathbf{e}_m &= \mathbf{0}_V && (T \circ S = \mathrm{id}_V)
\end{aligned}
$$

which implies that $a_1 = a_2 = \cdots = a_m = 0$, since $\mathcal{B}$ is linearly independent. Moreover, the list of vectors $\mathcal{C}$ spans W, for if $\mathbf{w} \in W$, then applying T, we can write

$$T(\mathbf{w}) = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \cdots + a_m \mathbf{e}_m$$

for some scalars a_i since $\mathcal{B}$ spans V. But then

$$
\begin{aligned}
\mathbf{w} &= S(T(\mathbf{w})) && (\text{since } S \circ T = \mathrm{id}_W)\\
&= S(a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \cdots + a_m \mathbf{e}_m)\\
&= a_1 S(\mathbf{e}_1) + a_2 S(\mathbf{e}_2) + \cdots + a_m S(\mathbf{e}_m) && (S \text{ is linear})
\end{aligned}
$$

so that $\mathcal{C}$ spans W. Hence $\mathcal{C}$ is a basis for V, so $\operatorname{Dim} V =$ number of vectors in $= m$, while $\operatorname{Dim} W =$ number of vectors in $= m$.

$\Leftarrow$. Suppose $\operatorname{Dim} V = \operatorname{Dim} W$. Let $\mathbf{e}_1, \ldots, \mathbf{e}_m$ be a basis for V, and let

$\mathbf{f}_1, \ldots, \mathbf{f}_m$ be a basis for W. (We know that the number of basis vectors is the same since $\operatorname{Dim} V = \operatorname{Dim} W$.)

To define linear maps

$$S : V \rightleftarrows W : T$$

it is sufficient, by Proposition 3.1.18 (Sufficient to Define a Linear Map on a Basis), to define the action of S and T on the basis vectors. We set:

$$\mathbf{e}_i \overset{S}{\mapsto} \mathbf{f}_i$$
$$\mathbf{e}_i \overset{T}{\leftarrow\!\shortmid} \mathbf{f}_i$$

Clearly we have $T \circ S = \mathrm{id}_V$ and $S \circ T = \mathrm{id}_W$. ■

Example 3.3.7 Show that $\mathrm{Mat}_{n,m}$ is isomorphic to $\mathbb{R}^{mn}$.

Solution. We simply observe that by Example 2.3.11, $\operatorname{Dim} \mathrm{Mat}_{n,m} = mn$ while from Example 2.3.4, $\operatorname{Dim} \mathbb{R}^{mn}$ is also equal to mn. □

There is one very important isomorphism which we will use over and over. Let V be a vector space with basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$. Consider the map

$$\begin{aligned} {}_{\mathcal{B}} : V &\to \mathrm{Col}_m \\ \mathbf{v} &\mapsto [\mathbf{v}]_{\mathcal{B}} \end{aligned}$$

which sends a vector $\mathbf{v} \in V$ to its corresponding coordinate vector $[\mathbf{v}]_{\mathcal{B}} \in \mathrm{Col}_m$. Lemma 2.4.8 says precisely that $[.]_{\mathcal{B}}$ is a linear map. We will now describe its inverse.

Definition 3.3.8 Let V be an m-dimensional vector space with basis $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$. Let $\mathsf{c} \in \mathrm{Col}_m$ be an m-dimensional column vector. Then the **vector in V corresponding to c with respect to the basis** $\mathcal{B}$ is

$$\mathbf{vec}_{V,\mathcal{B}}(\mathsf{c}) := \mathsf{c}_1\mathbf{e}_1 + \mathsf{c}_2\mathbf{e}_2 + \cdots + \mathsf{c}_m\mathbf{e}_m.$$

◊

Example 3.3.9 The polynomials $\mathcal{B} = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ where

$$\mathbf{p}_1 := 1 + x, \mathbf{p}_2 := 1 + x + x^2, \mathbf{p}_3 := 1 - x^2$$

are a basis of Poly_2 (check this). Then, for instance,

$$\begin{aligned} \mathbf{vec}_{\mathrm{Poly}_3,\mathcal{B}}\left(\begin{bmatrix} 2 \\ -3 \\ 3 \end{bmatrix}\right) &= 2(1+x) - 3(1+x+x^2) + 3(1-x^2) \\ &= 2 - x - 6x^2 \in \mathrm{Poly}_3 \,. \end{aligned}$$

□

Checkpoint 3.3.10 Show that:

1. $\mathbf{vec}_{V,\mathcal{B}}(\mathsf{c} + \mathsf{c}') = \mathbf{vec}_{V,\mathcal{B}}(\mathsf{c}) + \mathbf{vec}_{V,\mathcal{B}}(\mathsf{c}')$
2. $\mathbf{vec}_{V,\mathcal{B}}(k\mathsf{c}) = k\,\mathbf{vec}_{V,\mathcal{B}}(\mathsf{c})$.

This means that $\mathbf{vec}_{V,\mathcal{B}} : \mathrm{Col}_m \to V$ is a linear map.

Solution.

1.

$$\begin{aligned}\mathbf{vec}_{V,\mathcal{B}}(\mathsf{c}+\mathsf{c}') &= \mathbf{vec}_{V,\mathcal{B}}\left(\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} c_1' \\ \vdots \\ c_n' \end{bmatrix}\right) \\ &= \mathbf{vec}_{V,\mathcal{B}}\left(\begin{bmatrix} c_1 + c_1' \\ \vdots \\ c_n + c_n' \end{bmatrix}\right) \\ &= (c_1 + c_1')\mathbf{e}_1 + \cdots + (c_n + c_n')\mathbf{e}_n \\ &= (c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n) + (c_1'\mathbf{e}_1 + \cdots + c_n'\mathbf{e}_n) \\ &= \mathbf{vec}_{V,\mathcal{B}}(\mathsf{c}) + \mathbf{vec}_{V,\mathcal{B}}(\mathsf{c}')\end{aligned}$$

2.

$$\begin{aligned}\mathbf{vec}_{V,\mathcal{B}}(k\mathsf{c}) &= \mathbf{vec}_{V,\mathcal{B}}\left(k\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}\right) \\ &= \mathbf{vec}_{V,\mathcal{B}}\left(\begin{bmatrix} kc_1 \\ \vdots \\ kc_n \end{bmatrix}\right) \\ &= (kc_1\mathbf{e}_1 + \cdots + kc_n\mathbf{e}_n) \\ &= k(c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n) \\ &= k\,\mathbf{vec}_{V,\mathcal{B}}(\mathsf{c})\end{aligned}$$

Theorem 3.3.11 *Let V be a vector space with basis $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$. The maps*

$$\begin{aligned}{}_{\mathcal{B}} : V &\rightleftarrows \mathrm{Col}_m : \mathbf{vec}_{V,\mathcal{B}} \\ \mathbf{v} &\mapsto [\mathbf{v}]_{\mathcal{B}} \\ \mathbf{vec}_{V,\mathcal{B}}(\mathsf{c}) &\leftarrow\!\shortmid \mathsf{c}\end{aligned}$$

are an isomorphism of vector spaces.

Proof. Given $\mathbf{v} \in V$, expand it in the basis $\mathcal{B}$:

$$\mathbf{v} = a_1\mathbf{e}_1 + \ldots + a_m\mathbf{e}_m.$$

Then

$$\begin{aligned}(\mathbf{vec}_{V,\mathcal{B}} \circ [\cdot]_{\mathcal{B}})(\mathbf{v}) &= \mathbf{vec}_{V,\mathcal{B}}([\mathbf{v}]_{\mathcal{B}}) \\ &= \mathbf{vec}_{V,\mathcal{B}}\left(\begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}\right) \\ &= a_1\mathbf{e}_1 + \cdots + a_m\mathbf{e}_m \\ &= \mathbf{v}\end{aligned}$$

so that $\mathbf{vec}_{V,\mathcal{B}} \circ [\cdot]_{\mathcal{B}} = \mathrm{id}_V$. Conversely, given

$$\mathsf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \in \mathrm{Col}_m\,,$$

we have

$$\begin{aligned}\left([.]_{\mathcal{B}} \circ \mathbf{vec}_{V,\mathcal{B}}\right)(\mathsf{c}) &= \left[\mathbf{vec}_{V,\mathcal{B}}(\mathsf{c})\right]_{\mathcal{B}} \\ &= \left[\mathsf{c}_1\mathbf{e}_1 + \cdots + \mathsf{c}_m\mathbf{e}_m\right] \\ &= \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \\ &= \mathsf{c}\end{aligned}$$

where the second last step uses the *definition* of the coordinate vector of $\mathbf{v} = \mathsf{c}_1\mathbf{e}_1 + \cdots + \mathsf{c}_m\mathbf{e}_m$. Hence $[\cdot]_{\mathcal{B}} \circ \mathbf{vec}_{V,\mathcal{B}} = \mathrm{id}_{\mathrm{Col}_m}$. ■

> The result above is very important in linear algebra. It says that, once we have chosen a basis for an abstract finite-dimensional vector space V, we can treat the elements of V as if they were column vectors!

Exercises

1. Are the following vector spaces isomorphic?

$$\begin{aligned}V &= \left\{\mathsf{v} \in \mathrm{Col}_4 : \begin{bmatrix} 1 & 2 & 0 & -1 \\ -1 & 1 & 1 & 0 \end{bmatrix} \mathsf{v} = 0\right\} \\ W &= \left\{p \in \mathrm{Poly}_2 : \int_0^2 p(x)dx = 0\right\}.\end{aligned}$$

If they are, construct an explicit isomorphism between them. If not, prove that they are not isomorphic.

Solution. V consists of all vectors (x, y, z, w) satisyfing the linear equations

$$\begin{aligned}x + 2y - w &= 0 \\ -x + y + z &= 0.\end{aligned}$$

We are free to choose x and y arbitrarily, but then (1) above fixes w and (2) fixes z. Hence V is a 2 dimensional subspace.

Onto W. Let $p(x) = ax^2 + bx + c$. For $p(x)$ to be in W, $p(x)$ must satisfy the following equation:

$$\int_0^2 ax^2 + bx + c\,dx = \frac{8a}{3} + 2b + 2c = 0.$$

Thus, for any choice of a and b, c is uniquely determined. Hence W is a 2 dimensional subspace.

Since both V and W are 2 dimensional vector spaces, they are isomorphic by Theorem 3.3.6. To exhibit an explicit isomorphism between the V and W we shall need find bases for both spaces.

Since V is 2 dimensional, a basis for V consists of any two non-zero vectors in V that are not scalar multiples of one another. By inspection, we find the basis $\mathcal{B}_V = \{(1, 0, 1, 1), (0, 1, -1, 2)\}$. By similar reasoning, we find a basis $\mathcal{B}_W = \{\frac{3}{8}x^2 - \frac{1}{2}, x - 1\}$. By Proposition 3.1.18, there is a unique linear map $T : V \to W$ such that

$$T((1, 0, 1, 1)) = \frac{3}{8}x^2 - \frac{1}{2}$$

$$T((0,1,-1,2)) = x - 1.$$

This map is an isomorphism, as demonstrated by the proof of Theorem 3.3.6.

2. Are the following vector spaces isomorphic?

$$V = \{\mathbf{v} \in \mathbb{R}^3 : \mathbf{v} \times (1,2,3) = \mathbf{0}\}$$

$$W = \left\{M \in \mathrm{Mat}_{2,2} : M^T = -M\right\}.$$

If they are, construct an explicit isomorphism between them. If not, prove that they are not isomorphic.

Solution. We use some geometry to find the dimension of V. $\mathbf{v} \in V$ if and only if

$$|\mathbf{v}||(1,2,3)|\sin\theta = 0$$

where θ is the angle between $\mathbf{v}$ and $(1,2,3)$. Thus V consists of $\mathbf{0}$ as well as all those vectors parallel to $(1,2,3)$. But this set is precisely all vectors of the form $k(1,2,3)$ with $k \in \mathbb{R}$. Hence V is 1 dimensional with basis $\{(1,2,3)\}$.

W consists of all matrices (a_{ij}) satisfying

$$\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = \begin{bmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{bmatrix}.$$

Thus $a_{11} = a_{22} = 0$ and $b = -c$. So W consists of all those matrices of the form

$$\begin{bmatrix} 0 & k \\ -k & 0 \end{bmatrix}.$$

This also shows that

$$\left\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$$

is a basis for W. Hence V and W are isomorphic with the isomorphism given by the unique linear map $V \to W$ satisfying

$$(1,2,3) \mapsto \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Solutions

· Exercises

3.3.1. Solution. V consists of all vectors (x,y,z,w) satisyfing the linear equations

$$x + 2y - w = 0$$
$$-x + y + z = 0.$$

We are free to choose x and y arbitrarily, but then (1) above fixes w and (2) fixes z. Hence V is a 2 dimensional subspace.

Onto W. Let $p(x) = ax^2 + bx + c$. For $p(x)$ to be in W, $p(x)$ must satisfy the following equation:

$$\int_0^2 ax^2 + bx + c\,dx = \frac{8a}{3} + 2b + 2c = 0.$$

Thus, for any choice of a and b, c is uniquely determined. Hence W is a 2 dimensional subspace.

Since both V and W are 2 dimensional vector spaces, they are isomorphic by Theorem 3.3.6. To exhibit an explicit isomorphism between the V and W we shall need find bases for both spaces.

Since V is 2 dimensional, a basis for V consists of any two non-zero vectors in V that are not scalar multiples of one another. By inspection, we find the basis $\mathcal{B}_V = \{(1,0,1,1),(0,1,-1,2)\}$. By similar reasoning, we find a basis $\mathcal{B}_W = \{\frac{3}{8}x^2 - \frac{1}{2}, x-1\}$. By Proposition 3.1.18, there is a unique linear map $T : V \to W$ such that

$$\begin{aligned} T((1,0,1,1)) &= \frac{3}{8}x^2 - \frac{1}{2} \\ T((0,1,-1,2)) &= x - 1. \end{aligned}$$

This map is an isomorphism, as demonstrated by the proof of Theorem 3.3.6.

3.3.2. Solution. We use some geometry to find the dimension of V. $\mathbf{v} \in V$ if and only if

$$|\mathbf{v}||(1,2,3)|\sin\theta = 0$$

where θ is the angle between $\mathbf{v}$ and $(1,2,3)$. Thus V consists of $\mathbf{0}$ as well as all those vectors parallel to $(1,2,3)$. But this set is precisely all vectors of the form $k(1,2,3)$ with $k \in \mathbb{R}$. Hence V is 1 dimensional with basis $\{(1,2,3)\}$.

W consists of all matrices (a_{ij}) satisfying

$$\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = \begin{bmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{bmatrix}.$$

Thus $a_{11} = a_{22} = 0$ and $b = -c$. So W consists of all those matrices of the form

$$\begin{bmatrix} 0 & k \\ -k & 0 \end{bmatrix}.$$

This also shows that

$$\left\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$$

is a basis for W. Hence V and W are isomorphic with the isomorphism given by the unique linear map $V \to W$ satisfying

$$(1,2,3) \mapsto \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

3.4 Linear maps and matrices

Definition 3.4.1 Let $T : V \to W$ be a linear map from a vector space V to a vector space W. Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ be a basis for V and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be a basis for W. The **matrix of T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$** is defined as the $n \times m$ matrix whose columns are the coordinate vectors of $T(\mathbf{b}_i)$ with respect to the basis $\mathcal{C}$:

$$[T]_{\mathcal{C}\leftarrow\mathcal{B}} := \begin{bmatrix} \begin{bmatrix} T(\mathbf{b}_1) \end{bmatrix}_{\mathcal{C}} & \begin{bmatrix} T(\mathbf{b}_2) \end{bmatrix}_{\mathcal{C}} & \ldots & \begin{bmatrix} T(\mathbf{b}_m) \end{bmatrix}_{\mathcal{C}} \end{bmatrix}$$

$\Diamond$

Do you understand why $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ is an $n \times m$ matrix?

Example 3.4.2 Example from class! □

Theorem 3.4.3 *Let $T : V \to W$ be a linear map from a vector space V to a vector space W. Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ be a basis for V and $\mathcal{C}$ be a basis for W. Then for all vectors $\mathbf{v}$ in V,*

$$[T(\mathbf{v})]_{\mathcal{C}} = [T]_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} \tag{3.4.1}$$

where the right hand side is the product of the matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ with the column vector $\mathbf{v}]_{\mathcal{B}}$.

Proof. Similar to the proof of the Change-Of-Basis Theorem (Theorem 2.5.7). Let $\mathbf{v} \in V$. Expand it in the basis $\mathcal{B}$:

$$\mathbf{v} = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_m\mathbf{b}_m, \text{i.e. } [\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}.$$

Then,

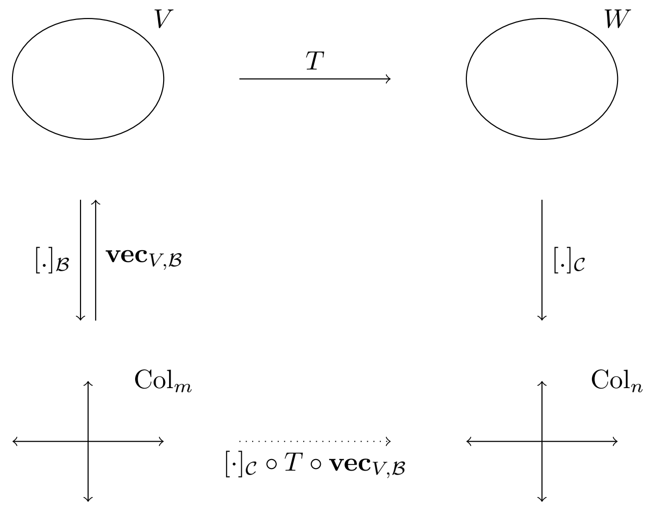
$$\begin{aligned}
{}_{\mathcal{C}} &= [T(a_1\mathbf{b}_1 + \cdots + a_m\mathbf{b}_m)]_{\mathcal{C}} \\
&= [a_1T(\mathbf{b}_1) + \cdots + a_nT(\mathbf{b}_m)]_{\mathcal{C}} && (T \text{ is linear}) \\
&= a_1[T(\mathbf{b}_1)]_{\mathcal{C}} + \cdots + a_n[T(\mathbf{b}_m)]_{\mathcal{C}} && (\text{Lemma 2.4.8}) \\
&= \begin{bmatrix} \begin{bmatrix} \\ T(\mathbf{b}_1) \\ \\ \end{bmatrix}_{\mathcal{C}} & \begin{bmatrix} \\ T(\mathbf{b}_2) \\ \\ \end{bmatrix}_{\mathcal{C}} & \cdots & \begin{bmatrix} \\ T(\mathbf{b}_m) \\ \\ \end{bmatrix}_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \\
&= [T]_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}.
\end{aligned}$$

■

If we could give this theorem a name, it would be "The Relationship Between Linear Maps, Coordinate Vectors, and Matrix Multiplication of Column Vectors Theorem"!

Example 3.4.4 Continuation of example from class! □

We can interpret Theorem 3.4.3 in a more abstract way as follows. We have the following diagram of linear maps of vector spaces:



The map at the top is the linear map $T : V \to W$. The map on the left from V to Col_m is the coordinate vector map $[.]_{\mathcal{B}}$ associated to the basis $\mathcal{B}$. Its inverse map $\mathbf{vec}_{V,\mathcal{B}} : \mathrm{Col}_m \to V$ is also drawn. The map on the right is the coordinate vector map $[\cdot]_{\mathcal{C}}$ from W to Col_n associated to the basis C. The dotted arrow on the bottom is the composite map, and can be computed explicitly as follows.

Lemma 3.4.5 *The composite map*

$$[\cdot]_{\mathcal{C}} \circ T \circ \mathbf{vec}_{V,\mathcal{B}} : \mathrm{Col}_m \to \mathrm{Col}_n$$

is multiplication by the matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$. *That is, for all column vectors* u *in* Col_m*,*

$$\left([\cdot]_{\mathcal{C}} \circ T \circ \mathbf{vec}_{V,\mathcal{B}}\right)(\mathsf{u}) = [T]_{\mathcal{C}\leftarrow\mathcal{B}}\,\mathsf{u}.$$

Proof. Let u be a column vector in Col_m. Define $\mathbf{v} := \mathbf{vec}_{V,\mathcal{B}}(\mathsf{u})$. Then $\mathbf{v}$ is the vector in V whose coordinate vector with respect to the basis $\mathcal{B}$ is u. That is, $\mathsf{u} = [\mathbf{v}]_{\mathcal{B}}$. So,

$$\begin{aligned}
\left([\cdot]_{\mathcal{C}} \circ T \circ \mathbf{vec}_{V,\mathcal{B}}\right)(\mathsf{u}) &= [.]_{\mathcal{C}}\left(T\left(\mathbf{vec}_{V,\mathcal{B}}(\mathsf{u})\right)\right) && \text{(Defn of composite map)}\\
&= [\cdot]_{\mathcal{C}}(T(\mathbf{v})) && \text{(Defn of } \mathbf{v}\text{)}\\
&= [T(\mathbf{v})]_{\mathcal{C}} && \text{(Defn of } [\cdot]_{\mathcal{C}}\text{)}\\
&= [T]_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} && \text{(Theorem 3.4.3).}
\end{aligned}$$

■

Before we move on, we need to recall another thing about matrices. Suppose A is a matrix with n rows. Let

$$\mathsf{e}_1 = \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\ \mathsf{e}_2 = \begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix},\ \ldots,\ \mathsf{e}_n = \begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$$

be the standard basis for Col_n. Then the ith column of A can be obtained by multiplying A with e_i:

$$i^{\text{th}} \text{ column of } \mathsf{A} = A\mathsf{e}_i. \tag{3.4.2}$$

Checkpoint 3.4.6 Check this!

Now we can prove the following important Theorem.

Theorem 3.4.7 Functoriality of the Matrix of a Linear Map. *Let $S : U \to V$ and $T : V \to W$ be linear maps between finite-dimensional vector spaces. Let $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ be bases for U, V and W respectively. Then*

$$[T \circ S]_{\mathcal{D}\leftarrow\mathcal{B}} = [T]_{\mathcal{D}\leftarrow\mathcal{C}}[S]_{\mathcal{C}\leftarrow\mathcal{B}}$$

where the right hand side is the product of the matrices $[T]_{\mathcal{D}\leftarrow\mathcal{C}}$ and $[S]_{\mathcal{C}\leftarrow\mathcal{B}}$.

Proof. We have:

$$\begin{aligned}
&i\text{th column of } [T \circ S]_{\mathcal{D}\leftarrow\mathcal{B}} \\
&\qquad = [(T \circ S)(\mathbf{b}_i)]_{\mathcal{D}} && (\text{Defn of } [T \circ S]_{\mathcal{D}\leftarrow\mathcal{B}}) \\
&\qquad = [T(S(\mathbf{b}_i))]_{\mathcal{D}} && (\text{Defn of } T \circ S) \\
&\qquad = [T]_{\mathcal{D}\leftarrow\mathcal{C}}\,[S(\mathbf{b}_i)]_{\mathcal{C}} && (\text{Theorem 3.4.3}) \\
&\qquad = [T]_{\mathcal{D}\leftarrow\mathcal{C}}\,[S]_{\mathcal{C}\leftarrow\mathcal{B}}\,[\mathbf{b}_i]_{\mathcal{B}} && (\text{Theorem 3.4.3}) \\
&\qquad = [T]_{\mathcal{D}\leftarrow\mathcal{C}}\,[S]_{\mathcal{C}\leftarrow\mathcal{B}}\,\mathsf{e}_i && (\text{ since } [\mathbf{b}_i]_{\mathcal{B}} = \mathsf{e}_i\ ) \\
&\qquad = i\text{th column of } [T]_{\mathcal{D}\leftarrow\mathcal{C}}[S]_{\mathcal{C}\leftarrow\mathcal{B}} && (3.4.2).
\end{aligned}$$

■

Corollary 3.4.8 *Let $T : V \to W$ be a linear map, and suppose $\mathcal{B}$ is a basis for V, and $\mathcal{C}$ is a basis for W. Then*

$$T \textit{ is an isomorphism} \iff [T]_{\mathcal{C}\leftarrow\mathcal{B}} \textit{ is invertible.}$$

Proof. $\Rightarrow$. Suppose the linear map T is an isomorphism. This means there exists a linear map $S : W \to V$ such that

$$S \circ T = \mathrm{id}_V \ \text{ and } T \circ S = \mathrm{id}_W$$

Therefore,

$$[S \circ T]_{\mathcal{B}\leftarrow\mathcal{B}} = [\mathrm{id}_V]_{\mathcal{B}\leftarrow\mathcal{B}} \text{ and } [T \circ S]_{\mathcal{C}\leftarrow\mathcal{C}} = [\mathrm{id}_W]_{\mathcal{C}\leftarrow\mathcal{C}}.$$

Therefore, by the Functoriality of the Matrix of a Linear Map (Theorem 3.4.7),

$$[S]_{\mathcal{B}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}} = I \text{ and } [T]_{\mathcal{C}\leftarrow\mathcal{B}}[S]_{\mathcal{B}\leftarrow\mathcal{C}} = I$$

Therefore the matrix $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ is invertible, with inverse given by

$$[T]^{-1}_{\mathcal{C}\leftarrow\mathcal{B}} = [S]_{\mathcal{B}\leftarrow\mathcal{C}}.$$

$\Leftarrow$. Suppose the matrix $[T] \equiv [T]_{\mathcal{C}\leftarrow\mathcal{B}}$ is invertible. Define the linear map

$$S : W \to V$$

by firstly defining it on the basis vectors in $\mathcal{C}$ by

$$S(\mathbf{c}_i) := \sum_{p=1}^{\dim V} [T]^{-1}_{pi}\mathbf{b}_p$$

and then extending to all of W by linearity. Then we have

$$\begin{aligned}
(T \circ S)(\mathbf{c}_i) &= T(S(\mathbf{c}_i)) \\
&= T\left(\sum_{p=1}^{\dim V} [T]^{-1}_{pi}\mathbf{b}_p\right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{p=1}^{\dim V} \sum_{q=1}^{\dim W} [T]^{-1}_{pi} [T]_{qp} \mathbf{c}_q \\
&= \sum_{q=1}^{\dim W} \left(\sum_{p=1}^{\dim V} [T]_{qp} [T]^{-1}_{pi} \right) \mathbf{c}_q \\
&= \sum_{q=1}^{\dim W} \left([T][T]^{-1} \right)_{qi} \mathbf{c}_q \\
&= \sum_{q=1}^{\dim W} I_{qi} \mathbf{c}_q \\
&= \sum_{q=1}^{\dim W} \delta_{qi} \mathbf{c}_q \\
&= \mathbf{c_i}.
\end{aligned}
$$

Therefore, $T \circ S = \mathrm{id}_W$. In a similar way, we can prove that $S \circ T = \mathrm{id}_V$. Therefore the linear map T is an isomorphism, with inverse map $T^{-1} = S$. ■

We can refine this a bit further. Explicitly, 'the inverse of the matrix of a linear map equals the matrix of the inverse of the linear map'.

Corollary 3.4.9 *Suppose $\mathcal{B}$ and $\mathcal{C}$ are bases for vector spaces V and W respectively. Suppose a linear map $T : V \to W$ has inverse $T^{-1} : W \to V$. Then*

$$[T]^{-1}_{\mathcal{C}\leftarrow\mathcal{B}} = [T^{-1}]_{\mathcal{B}\leftarrow\mathcal{C}}.$$

Proof. We have

$$
\begin{aligned}
{}_{\mathcal{C}\leftarrow\mathcal{B}}[T^{-1}]_{\mathcal{B}\leftarrow\mathcal{C}} &= [T \circ T^{-1}]_{\mathcal{C}\leftarrow\mathcal{C}} && (\text{Theorem 3.4.7}) \\
&= [\mathrm{id}_W]_{\mathcal{C}\leftarrow\mathcal{C}} && (T \circ T^{-1} = \mathrm{id}_W) \\
&= I
\end{aligned}
$$

and

$$
\begin{aligned}
{}_{\mathcal{B}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}} &= [T^{-1} \circ T]_{\mathcal{B}\leftarrow\mathcal{B}} && (\text{Theorem 3.4.7}) \\
&= [\mathrm{id}_V]_{\mathcal{B}\leftarrow\mathcal{B}} && (T^{-1} \circ T = \mathrm{id}_V) \\
&= I.
\end{aligned}
$$

■

The next Lemma says that the 'change-of-basis matrix' from Section 2.5 is just the matrix of the identity linear map with respect to the relevant bases.

Lemma 3.4.10 *Let $\mathcal{B}$ and $\mathcal{C}$ be bases for an m-dimensional vector space V. Then*

$$\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = [\mathrm{id}]_{\mathcal{C}\leftarrow\mathcal{B}}.$$

Proof.

$$
\begin{aligned}
\mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}} &= [[\mathbf{b}_1]_{\mathcal{C}} \cdots [\mathbf{b}_m]_{\mathcal{C}}] && (\text{Defn of } \mathsf{P}_{\mathcal{C}\leftarrow\mathcal{B}}) \\
&= [[\mathrm{id}(\mathbf{b}_1)]_{\mathcal{C}} \cdots [\mathrm{id}(\mathbf{b}_m)]_{\mathcal{C}}] \\
&= [\mathrm{id}]_{\mathcal{C}\leftarrow\mathcal{B}}. && (\text{Defn of } [\mathrm{id}]_{\mathcal{C}\leftarrow\mathcal{B}})
\end{aligned}
$$

■

The next Theorem tells us how the matrix of a linear operator changes when we change the bases used in computing it.

Theorem 3.4.11 *Let $\mathcal{B}$ and $\mathcal{C}$ be bases for a vector space V, and let $T : V \to V$ be a linear operator on V. Then*

$$[T]_{\mathcal{C}\leftarrow\mathcal{C}} = \mathsf{P}^{-1}[T]_{\mathcal{B}\leftarrow\mathcal{B}}P$$

where $P \equiv P_{\mathcal{B}\leftarrow\mathcal{C}}$.

Proof.

$$\begin{aligned}
\text{RHS} &= \mathsf{P}^{-1}[T]_{\mathcal{B}\leftarrow\mathcal{B}}P \\
&= [\text{id}]^{-1}_{\mathcal{B}\leftarrow\mathcal{C}}[T]_{\mathcal{B}\leftarrow\mathcal{B}}[\text{id}]_{\mathcal{B}\leftarrow\mathcal{C}} && \text{(Lemma 3.4.10)} \\
&= [\text{id}]_{\mathcal{C}\leftarrow\mathcal{B}}[T]_{\mathcal{B}\leftarrow\mathcal{B}}[\text{id}]_{\mathcal{B}\leftarrow\mathcal{C}} && \text{(Corollary 3.4.9)} \\
&= [\text{id} \circ T \circ \text{id}]_{\mathcal{C}\leftarrow\mathcal{C}} && \text{(Theorem 3.4.7)} \\
&= [T]_{\mathcal{C}\leftarrow\mathcal{C}} \\
&= \text{LHS}.
\end{aligned}$$

■

Exercises

1. Let

$$T : \text{Trig}_1 \to \text{Trig}_2$$

be the 'multiply with $\sin x$' linear map, $T(f)(x) = \sin x f(x)$. Compute $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ with respect to the standard basis $\mathcal{B}$ of Trig_1 and $\mathcal{C}$ of Trig_2.

Solution. Recall the standard double angle formulae:

$$\begin{aligned}
\sin(2x) &= 2\sin x \cos x \\
\cos(2x) &= \cos^2 x - \sin^2 x = 2\cos^2 x - 1 = 1 - 2\sin^2 x
\end{aligned}$$

With these in mind, we compute:

$$\begin{aligned}
T(T_0) &= \sin x = T_2 \\
T(T_1) &= \sin x \cos x = \frac{1}{2}\sin(2x) = \frac{1}{2}T_4 \\
T(T_2) &= \sin x \sin x = \frac{1}{2} - \frac{1}{2}\cos(2x) = \frac{1}{2}T_0 - \frac{1}{2}T_3
\end{aligned}$$

Thus

$$[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}.$$

2. Let

$$S : \text{Trig}_2 \to \text{Trig}_2$$

be the 'shift by $\frac{\pi}{6}$' map, $S(f)(x) = f(x - \frac{\pi}{6})$. Compute $[S]_{\mathcal{C}\leftarrow\mathcal{C}}$ with respect to the standard basis $\mathcal{C}$ of Trig_2.

Solution. In this exercise, we shall use the standard angle addition formulae for trigonometric functions:

$$\begin{aligned}
\cos(x - y) &= \cos x \cos y + \sin x \sin y \\
\sin(x - y) &= \sin x \cos y - \cos x \sin y
\end{aligned}$$

We compute

$$\begin{aligned}
S(T_0) &= 1 = T_0 \\
S(T_1) &= \cos(x - \frac{\pi}{6}) = \frac{\sqrt{3}}{2}\cos x + \frac{1}{2}\sin x = \frac{\sqrt{3}}{2}T_1 + \frac{1}{2}T_2 \\
S(T_2) &= \sin(x - \frac{\pi}{6}) = -\frac{1}{2}\cos x + \frac{\sqrt{3}}{2}\sin x = -\frac{1}{2}T_1 + \frac{\sqrt{3}}{2}T_2 \\
S(T_3) &= \cos(2x - \frac{\pi}{6}) = -\frac{1}{2}\sin^2 x + \frac{\sqrt{3}}{2}\cos^2 x + \sin x \cos x \\
&= -\frac{\sqrt{3}}{2}(\frac{1}{2} - \frac{1}{2}\cos x) + \frac{\sqrt{3}}{2}(\frac{1}{2} + \frac{1}{2}\cos 2x) + \frac{1}{2}\sin 2x \\
&= \frac{\sqrt{3}}{2}\cos 2x + \frac{1}{2}\sin 2x = \frac{\sqrt{3}}{2}T_3 + \frac{1}{2}\sin 2x.
\end{aligned}$$

Similarly,

$$S(T_4) = \sin(2x - \frac{\pi}{6}) = -\frac{1}{2}\cos 2x + \frac{\sqrt{3}}{2}\sin 2x = -\frac{1}{2}T_3 + \frac{\sqrt{3}}{2}T_4.$$

Hence

$$[S]_{\mathcal{C}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}.$$

3. Verify Theorem 3.4.3 for the linear map $S : \mathrm{Mat}_{2,2} \to \mathrm{Mat}_{2,2}$ given by $S(\mathsf{M}) = \mathsf{M}^T$, for the vector $\mathsf{A} \in \mathrm{Mat}_{2,2}$ given by

$$\mathsf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

and using the following bases of $\mathrm{Mat}_{2,2}$:

$$\mathcal{B} = \mathcal{C} = \{\mathsf{M}_1 = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \mathsf{M}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \mathsf{M}_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \mathsf{M}_4 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\}.$$

4. Check that the linear maps T and S from Exercises 3.4.1 and Exercise 3.4.2 satisfy $[S \circ T]_{\mathcal{C}\leftarrow\mathcal{B}} = [S]_{\mathcal{B}\leftarrow\mathcal{B}}[T]_{\mathcal{C}\leftarrow\mathcal{B}}$.

Solutions

· Exercises

3.4.1. Solution. Recall the standard double angle formulae:

$$\begin{aligned}
\sin(2x) &= 2\sin x \cos x \\
\cos(2x) &= \cos^2 x - \sin^2 x = 2\cos^2 x - 1 = 1 - 2\sin^2 x
\end{aligned}$$

With these in mind, we compute:

$$\begin{aligned}
T(T_0) &= \sin x = T_2 \\
T(T_1) &= \sin x \cos x = \frac{1}{2}\sin(2x) = \frac{1}{2}T_4 \\
T(T_2) &= \sin x \sin x = \frac{1}{2} - \frac{1}{2}\cos(2x) = \frac{1}{2}T_0 - \frac{1}{2}T_3
\end{aligned}$$

Thus

$$[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}.$$

3.4.2. Solution. In this exercise, we shall use the standard angle addition formulae for trigonometric functions:

$$\cos(x-y) = \cos x \cos y + \sin x \sin y$$
$$\sin(x-y) = \sin x \cos y - \cos x \sin y$$

We compute

$$\begin{aligned} S(T_0) &= 1 = T_0 \\ S(T_1) &= \cos(x - \frac{\pi}{6}) = \frac{\sqrt{3}}{2}\cos x + \frac{1}{2}\sin x = \frac{\sqrt{3}}{2}T_1 + \frac{1}{2}T_2 \\ S(T_2) &= \sin(x - \frac{\pi}{6}) = -\frac{1}{2}\cos x + \frac{\sqrt{3}}{2}\sin x = -\frac{1}{2}T_1 + \frac{\sqrt{3}}{2}T_2 \\ S(T_3) &= \cos(2x - \frac{\pi}{6}) = -\frac{1}{2}\sin^2 x + \frac{\sqrt{3}}{2}\cos^2 x + \sin x \cos x \\ &= -\frac{\sqrt{3}}{2}(\frac{1}{2} - \frac{1}{2}\cos x) + \frac{\sqrt{3}}{2}(\frac{1}{2} + \frac{1}{2}\cos 2x) + \frac{1}{2}\sin 2x \\ &= \frac{\sqrt{3}}{2}\cos 2x + \frac{1}{2}\sin 2x = \frac{\sqrt{3}}{2}T_3 + \frac{1}{2}\sin 2x. \end{aligned}$$

Similarly,

$$S(T_4) = \sin(2x - \frac{\pi}{6}) = -\frac{1}{2}\cos 2x + \frac{\sqrt{3}}{2}\sin 2x = -\frac{1}{2}T_3 + \frac{\sqrt{3}}{2}T_4.$$

Hence

$$[S]_{\mathcal{C}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}.$$

3.5 Kernel and range of a linear map

Definition 3.5.1 Let $T : V \to W$ be a linear map between vector spaces V and W. The **kernel** of T, written $\mathrm{Ker}(T)$, is the set of all vectors $\mathbf{v} \in V$ such that are mapped to $\mathbf{0}_W$ by T. That is,

$$\mathrm{Ker}(T) := \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}_W\}.$$

The **image** of T, written $\Im(T)$, is the set of all vectors $\mathbf{w} \in W$ such that $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v} \in V$. That is,

$$\Im(T) := \{\mathbf{w} \in W : \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in V\}$$

$\Diamond$

See Figure 3.5.2 and Figure 3.5.3 for a schematic representation.

Sometimes, to be absolutely clear, I will put a subscript on the zero vector to indicate which vector space it belongs to, eg. $\mathbf{0}_W$ refers to the zero vector in W, while $\mathbf{0}_V$ refers to the zero vector in V.

Another name for the kernel of T is the *nullspace* of T, and another name for the image of T is the *range* of T.

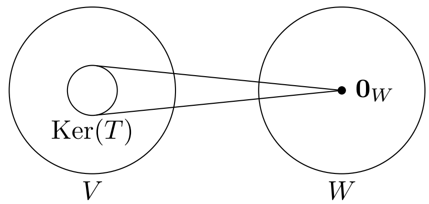


Figure 3.5.2: Ker(T)

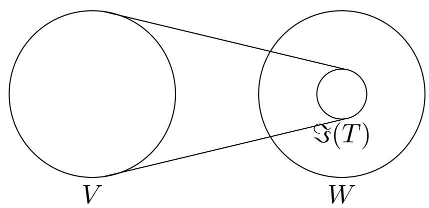


Figure 3.5.3: $\Im(T)$

Lemma 3.5.4 *Let $T : V \to W$ be a linear map. Then:*

1. Ker(T) *is a subspace of* V
2. $\Im(T)$ *is a subspace of* W

Proof. (i) We must check the three requirements for being a subspace.

1. Ker(T) is closed under addition. Suppose $\mathbf{v}$ and $\mathbf{v}'$ are in Ker(T). In other words, $T(\mathbf{v}) = \mathbf{0}$ and $T(\mathbf{v}') = \mathbf{0}$. We need to show that $\mathbf{v} + \mathbf{v}'$ is in Ker(T), in other words, that $T(\mathbf{v} + \mathbf{v}') = \mathbf{0}$. Indeed,
$$T(\mathbf{v} + \mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}') = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$
2. $\mathbf{0}_V \in$ Ker(T). To show that $\mathbf{0}_V$ is in Ker(T), we need to show that $T(\mathbf{0}_V) = \mathbf{0}_W$. Indeed, this is true since T is a linear map, by Lemma 3.1.17.
3. Ker(T) is closed under scalar multiplication. Suppose $\mathbf{v} \in$ Ker(T) and $k \in \mathbb{R}$ is a scalar. We need to show that $k\mathbf{v} \in$ Ker(T), that is, we need to show that $T(k\mathbf{v}) = \mathbf{0}$. Indeed,
$$T(k\mathbf{v}) = kT(\mathbf{v}) = k\mathbf{0} = \mathbf{0}.$$

(ii) Again, we must check the three requirements for being a subspace.

1. $\Im(T)$ is closed under addition. Suppose $\mathbf{w}$ and $\mathbf{w}'$ are in $\Im(T)$. In other words, there exist vectors $\mathbf{v}$ and $\mathbf{v}'$ in V such that $T(\mathbf{v}) = \mathbf{w}$ and $T(\mathbf{v}') = \mathbf{w}'$. We need to show that $\mathbf{w} + \mathbf{w}'$ is also in $\Im(T)$, in other words, that

there exists a vector $\mathbf{u}$ in V such that $T(\mathbf{u}) = \mathbf{w} + \mathbf{w}'$. Indeed, set $\mathbf{u} := \mathbf{v} + \mathbf{v}'$. Then,

$$T(\mathbf{u}) = T(\mathbf{v} + \mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}') = \mathbf{w} + \mathbf{w}'.$$

2. $\mathbf{0}_W \in \Im(T)$. To show that $\mathbf{0}_W \in \Im(T)$, we need to show that there exists $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{0}_W$. Indeed, choose $\mathbf{v} = \mathbf{0}_V$. Then $T(\mathbf{v}) = T(\mathbf{0}_V) = \mathbf{0}_W$ by Lemma 3.1.17.

3. $\Im(T)$ is closed under scalar multiplication. Suppose $\mathbf{w} \in \Im(T)$ and k is a scalar. We need to show that $k\mathbf{w} \in \Im(T)$. The fact that $\mathbf{w}$ is in $\Im(T)$ means that there exists a $\mathbf{v}$ in V such that $T(\mathbf{v}) = \mathbf{w}$. We need to show that there exists a $\mathbf{u} \in V$ such that $T(\mathbf{u}) = k\mathbf{w}$. Indeed, set $\mathbf{u} := k\mathbf{v}$. Then
$$T(\mathbf{u}) = T(k\mathbf{v}) = kT(\mathbf{v}) = k\mathbf{w}.$$

■

Now that we know that the kernel and image of a linear map are subspaces, and hence vector spaces in their own right, we can make the following definition.

Definition 3.5.5 Let $T : V \to W$ be a linear map from a finite-dimensional vector space V to a vector space W. The **nullity** of T is the dimension of $\mathrm{Ker}(T)$, and the **rank** of T is the dimension of $\Im(T)$:

$$\begin{aligned} \mathrm{Nullity}(T) &:= \mathrm{Dim}(\mathrm{Ker}(T)) \\ \mathrm{Rank}(T) &= \mathrm{Dim}(\Im(T)) \end{aligned}$$

◊

The ‘dimension of Ker(T)’ makes sense because Ker(T) is a subspace of a finite-dimensional vector space V, and hence is finite-dimensional by Proposition 2.3.20. We do not yet know that $\Im(T)$ is finite-dimensional, but this will follow from the Rank-Nullity Theorem (Theorem 3.5.9).

Example 3.5.6 Let $\mathbf{a} \in \mathbb{R}^3$ be a fixed nonzero vector. Consider the ‘cross product with $\mathbf{a}$’ linear map from Example 3.1.7,

$$\begin{aligned} C : \mathbb{R}^3 &\to \mathbb{R}^3 \\ \mathbf{v} &\mapsto \mathbf{a} \times \mathbf{v} \end{aligned}$$

Determine the kernel, image, nullity and rank of C.

Solution. The kernel of C is the subspace of $\mathbb{R}^3$ consisting of all vectors $\mathbf{v} \in V$ such that $\mathbf{a} \times \mathbf{v} = \mathbf{0}$. From the geometric formula for the cross-product,

$$|\mathbf{a} \times \mathbf{v}| = |\mathbf{a}||\mathbf{v}| \sin\theta$$

where θ is the angle from $\mathbf{a}$ to θ, we see that

$$\mathbf{a} \times \mathbf{v} = \mathbf{0} \Leftrightarrow \mathbf{v} = \mathbf{0} \text{ or } \theta = 0 \text{ or } \theta = \pi.$$

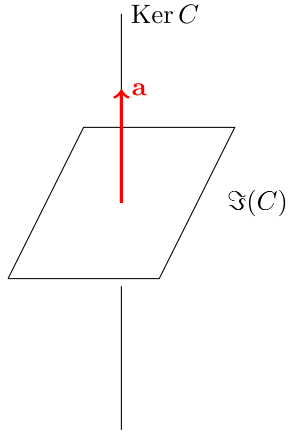
In other words, $\mathbf{v}$ must be a scalar multiple of $\mathbf{a}$. So,

$$\mathrm{Ker}(C) = \{k\mathbf{a}, k \in \mathbb{R}\}.$$

I claim that the *image* of C is the subspace of *all* vectors perpendicular to $\mathbf{a}$, i.e.

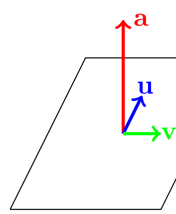
$$\Im(C) := \{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}. \tag{3.5.1}$$

If you believe me, then the picture is as follows:



Let me prove equation (3.5.1). By definition, the image of C is the subspace of $\mathbb{R}^3$ consisting of all vectors $\mathbf{w}$ of the form $\mathbf{w} = \mathbf{a} \times \mathbf{v}$ for some $\mathbf{v} \in \mathbb{R}^3$. This implies that $\mathbf{w}$ is perpendicular to $\mathbf{a}$. This was the 'easy' part. The 'harder' part is to show the converse. That is, we need to show that if $\mathbf{u}$ is perpendicular to $\mathbf{a}$, then $\mathbf{u}$ is in the image of C, i.e. there exists a vector $\mathbf{v}$ such that $C(\mathbf{v}) = \mathbf{u}$.

Indeed, we can choose $\mathbf{v}$ to be the vector obtained by rotating $\mathbf{u}$ by 90 degrees clockwise in the plane I, and scaling it appropriately:



In terms of a formula, we have

$$\mathbf{v} = \frac{|\mathbf{u}|}{|\mathbf{a}|}\mathbf{u} \times \mathbf{a}.$$

Note that this is not the *only* vector $\mathbf{v}$ such that $C(\mathbf{v}) = \mathbf{u}$. Indeed, if we add to $\mathbf{v}$ any vector that lies on the line through $\mathbf{a}$, the resulting vector

$$\tilde{\mathbf{v}} = \mathbf{v} + k\mathbf{a}$$

also satisfies $C(\tilde{\mathbf{v}}) = \mathbf{u}$, since

$$C(\tilde{\mathbf{v}}) = C(\mathbf{v} + k\mathbf{a}) = C(\mathbf{v}) + C(k\mathbf{a}) = \mathbf{u} + \mathbf{0} = \mathbf{u}.$$

□

Example 3.5.7 Determine the kernel, image, nullity and rank of the linear map

$$\begin{aligned} I : \mathrm{Trig}_2 &\to \mathbb{R} \\ f &\mapsto \int_0^{\pi} f(x)dx. \end{aligned}$$

Solution. The kernel of I consists of all degree 2 trigonometric polynomials

$$f(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x$$

such that

$$\int_0^\pi (a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x)\, dx = 0.$$

Performing the integrals, this becomes the equation

$$\pi a_0 + 2b_1 = 0$$

with no constraints on the other constants a_1, a_2, b_2. In other words,

$$\mathrm{Ker}(I) = \Big\{ \text{ all trigonometric polynomials of the form } (a_0(1 - \frac{\pi}{2} \sin x) + a_1 \cos x + a_2 \cos 2x + b_2 \sin 2x), \text{ wher}$$

Hence $\mathrm{Nullity}(I) = \mathrm{Dim}(\mathrm{Ker}(I)) = 4$.

The image of I consists of all real numbers $p \in \mathbb{R}$ such that there exists a $f \in \mathrm{Trig}_2$ such that $I(f) = p$. I claim that

$$\Im(I) = \mathbb{R}.$$

Indeed, given $p \in \mathbb{R}$, we may choose $f(x) = \frac{p}{2} \sin x$, since

$$I(f) = \frac{p}{2} \int_0^\pi \sin x\, dx = p.$$

Hence $\Im(I) = \mathbb{R}$, and $\mathrm{Rank}(I) = 1$.

Note that the choice of $f(x) = \frac{p}{2} \sin(x)$ satisfying $I(f) = p$ is not unique. We could set $\tilde{f} = f + g$ where $g \in \mathrm{Ker}(I)$ and we would still have $I(\tilde{f}) = p$:

$$I(\tilde{f}) = I(f + g) = I(f) + I(g) = p + 0 = p.$$

$\square$

Example 3.5.8 Consider the function

$$\begin{aligned} T : \mathrm{Poly}_2 &\to \mathbb{R}^2 \\ p &\mapsto (p(1), p'(1)). \end{aligned}$$

Show that T is a linear map, and determine its kernel, image, rank and nullity.

Solution. We first show that T is a linear map. Let $p, q \in \mathrm{Poly}_2$. Then

$$\begin{aligned} T(p+q) &= ((p+q)(1),\, (p+q)'(1)) && \text{(Defn of } T) \\ &= (p(1) + q(1),\, (p+q)'(1)) && \text{(Defn of the function } p+q) \\ &= (p(1) + q(1),\, (p' + q')(1)) && ((p+q)' = p' + q') \\ &= (p(1) + q(1),\, p'(1) + q'(1)) && \text{(Defn of the function } p' + q') \\ &= (p(1),\, p'(1)) + (q(1),\, q'(1)) && \text{(Defn of addition in } \mathbb{R}^2) \\ &= T(p) + T(q). \end{aligned}$$

The proof of $T(kp) = kT(p)$ is similar.

The kernel of T is the set of all polynomials

$$p(x) = a_0 + a_1 x + a_2 x^2$$

such that $T(p) = (0, 0)$. This translates into the equation

$$(a_0 + a_1 + a_2,\ a_1 + 2a_2) = (0, 0).$$

This in turn translates into two equations:

$$\begin{aligned} a_0 + a_1 + a_2 &= 0 \\ a_0 + a_1 + a_2 &= 0 \\ a_1 + 2a_2 &= 0 \end{aligned}$$

whose solution is $a_2 = t$, $a_1 = -2t$, $a_0 = -t$, where $t \in \mathbb{R}$. Hence

$$\operatorname{Ker}(T) = \left\{ \text{ all polynomials of the form } -t - 2tx + tx^2 \text{ where } t \in \mathbb{R} \right\}.$$

Hence $\operatorname{Nullity}(T) = 1$.

The image of T is the set of all $(v, w) \in \mathbb{R}^2$ such that there exists a polynomial $p = a_0 + a_1 x + a_2 x^2$ in Poly_2 such that $T(p) = (v, w)$. So, (v, w) is in the image of T if and only if we can find a polynomial $p = a_0 + a_1 x + a_2 x^2$ such that

$$(a_0 + a_1 + a_2, a_1 + 2a_2) = (v, w).$$

In other words, (v, w) is in the image of T if and only if the equations

$$\begin{aligned} a_0 + a_1 + a_2 &= v \\ a_1 + 2a_2 &= w \end{aligned}$$

have a solution for some a_0, a_1, a_2. But these equations *always* have a solution, for *all* $(v, w) \in \mathbb{R}^2$ is. For instance, one solution is

$$a_2 = 0,\ a_1 = w,\ a_0 = v - w$$

which corresponds to the polynomial

$$p(x) = v - w + wx. \tag{3.5.2}$$

Note that indeed $T(p) = (v, w)$. Hence,

$$\Im(T) = \{\text{all } (v, w) \in \mathbb{R}^2\} = \mathbb{R}^2.$$

Hence $\operatorname{Rank}(T) = \operatorname{Dim}(\Im(T)) = 2$.

Note that the choice of the polynomial $p(x) = v - w + wx$ from (3.5.2) which satisfies $T(p) = (v, w)$ is not the *only* possible choice. Indeed, any polynomial of the form $\tilde{p} = p + q$ where $q \in \operatorname{Ker}(T)$ will also satisfy $T(\tilde{p}) = (v, w)$, since

$$T(\tilde{p}) = T(p + q) = T(p) + T(q) = (v, w) + (0, 0) = (v, w).$$

□

Theorem 3.5.9 Rank-nullity theorem. *Let $T : V \to W$ be a linear map from a finite-dimensional vector space V to a vector space W. Then*

$$\operatorname{Nullity}(T) + \operatorname{Rank}(T) = \operatorname{Dim}(V).$$

Proof. Let $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ be a basis for $\operatorname{Ker}(T)$. Since $\mathcal{B}$ is a list of linearly independent vectors in V, we can extend it to a basis $\mathcal{C} = \{\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{f}_1, \ldots, \mathbf{f}_p\}$ for V, by Corollary 2.3.26. I claim that

$$\mathcal{D} := \{T(\mathbf{f}_1), \ldots, T(\mathbf{f}_p)\}$$

is a basis for $\Im(T)$. If I can prove that, then we are done, since then we have

$$\begin{aligned}\text{Nullity}(T) + \text{Rank}(T) &= k + p \\ &= \text{Dim}(V).\end{aligned}$$

Let us prove that $\mathcal{D}$ is a basis for $\Im(T)$.

$\mathcal{D}$ is linearly independent. Suppose

$$b_1 T(\mathbf{f}_1) + \cdots + b_p T(\mathbf{f}_p) = \mathbf{0}_W.$$

We recognize the left hand side as $T(b_1\mathbf{f}_1 + \cdots + b_p\mathbf{f}_p)$. Hence

$$b_1\mathbf{f}_1 + \cdots + b_p\mathbf{f}_p \in \text{Ker}(T)$$

which means we can write it as a linear combination of the vectors in $\mathcal{B}$,

$$b_1\mathbf{f}_1 + \cdots + b_p\mathbf{f}_p = a_1\mathbf{e}_1 + \cdots + a_k\mathbf{e}_k.$$

Bringing all terms onto one side, this becomes the equation

$$-a_1\mathbf{e}_1 - \cdots - a_k\mathbf{e}_k + b_1\mathbf{f}_1 + \cdots + b_p\mathbf{f}_p = \mathbf{0}_V.$$

We recognize the left hand side as a linear combination of the $\mathcal{C}$ basis vectors. Since they are are linearly independent, all the scalars must be zero. In particular, $b_1 = \cdots = b_p = 0$, which is what we wanted to prove.

$\mathcal{D}$ spans W. Suppose $\mathbf{w} \in \Im(T)$. We need to show that $\mathbf{w}$ is a linear combination of the vectors from $\mathcal{D}$. Since $\mathbf{w}$ is in the image of T, there exists $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$. Since $\mathcal{C}$ is a basis for V, we can write

$$\mathbf{v} = a_1\mathbf{e}_1 + \cdots + a_k\mathbf{e}_k + b_1\mathbf{f}_1 + \cdots + b_p\mathbf{f}_p$$

for some scalars $a_1, \ldots, a_k, b_1, \ldots, b_p$. Then

$$\begin{aligned}\mathbf{w} &= T(\mathbf{v}) \\ &= T(a_1\mathbf{e}_1 + \cdots + a_k\mathbf{e}_k + b_1\mathbf{f}_1 + \cdots + b_p\mathbf{f}_p) \\ &= a_1T(\mathbf{e}_1) + \cdots + a_kT(\mathbf{e}_k) + b_1T(\mathbf{f}_1) + \cdots + b_pT(\mathbf{f}_p) \\ &= b_1T(\mathbf{f}_1) + \cdots + b_pT(\mathbf{f}_p) && (\mathbf{e}_i \in \text{Ker}(T))\end{aligned}$$

so that $\mathbf{w}$ is indeed a linear combination of the vectors from $\mathcal{D}$. ■

Exercises

1. Verify the Rank-Nullity theorem for the following linear maps. That is, for each map T, (a) determine $\text{Ker}(T)$ and $\Im(T)$ explicitly, (b) determine the dimension of $\text{Ker}(T)$ and $\Im(T)$, (c) check that these numbers satisfy the Rank-Nullity theorem.

(a) The identity map $\text{id}_V : V \to V$ on a finite-dimensional vector space V.

(b) The zero map

$$\begin{aligned}Z : V &\to V \\ \mathbf{v} &\mapsto \mathbf{0}\end{aligned}$$

on a finite-dimensional vector space V.

(c) The map

$$\begin{aligned} T : \mathrm{Poly}_3 &\to \mathrm{Col}_3 \\ p &\mapsto \begin{bmatrix} p(1) \\ p(2) \\ p(3) \end{bmatrix} \end{aligned}$$

(d) The map

$$\begin{aligned} S : \mathrm{Trig}_2 &\to \mathrm{Col}_2 \\ f &\mapsto \begin{bmatrix} \int_0^\pi f(x)\cos x dx \\ \int_0^\pi f(x)\sin x dx \end{bmatrix} \end{aligned}$$

Solution.

(a) id_V sends only a single vector to 0, namely the vector $\mathbf{0}_v$. Thus $\mathrm{Ker}(\mathrm{id}_v) = \{\mathbf{0}_v\}$ and hence $\mathrm{Nullity}(\mathrm{id}_v) = 0$. $\Im(\mathrm{id}_V) = V$ and so $\mathrm{Rank}(\mathrm{id}_V) = \mathrm{Dim}(V)$. The equation

$$\mathrm{Dim}\, V + 0 = \mathrm{Dim}\, V.$$

verifies the Rank-Nullity Theorem for this example.

(b) The zero map spends every element in V to $\mathbf{0}$. Hence $\mathrm{Ker}(Z) = V$ and so $\mathrm{Nullity}(Z) = \mathrm{Dim}\, V$. For the same reason $\Im(Z) = \{0\}$ and thus $\mathrm{Rank}(Z) = 0$. The equation

$$\mathrm{Dim}\, V + 0 = \mathrm{Dim}\, V.$$

verifies the Rank-Nullity Theorem for this example.

(c) $\mathbf{p}(x) \in \mathrm{Ker}(T)$ if and only if

$$p(1) = p(2) = p(3) = 0.$$

But any degree 3 polynomial with $1, 2$ and 3 as roots must be of the form

$$a(x-1)(x-2)(x-3)$$

where $a \in \mathbb{R}$. Conversely, any element of this form is in $\mathrm{Ker}\, T$. Hence

$$\mathrm{Ker}(T) = \{a(x-1)(x-2)(x-3) \in \mathrm{Poly}_3 \,:\, a \in \mathbb{R}\}$$

and thus $\mathrm{Nullity}(T) = 1$.

I claim that $\Im(T) = \mathrm{Col}_3$. That is, for any

$$\begin{bmatrix} s \\ t \\ u \end{bmatrix} \in \mathrm{Col}_3,$$

we can find a polynomial $\mathbf{p}$, of degree 3 or less, such that $p(1) = s, p(2) = t, p(3) = u$. To show this, you could set up the system of linear equations

$$a + b + c + d = s$$

$$8a + 4b + 2c + d = t$$
$$27b + 9b + 3c + d = u$$

and solve for the coefficients of $\mathbf{p}$. What is perhaps more elegant is to use the theory of Lagrange interpolation polynomials [1]. Given any 3 distinct points in $\mathbb{R}^2$, we can always find a degree two polynomial going through these 3 points. In our case, let the points be $(1, s), (2, t), (3, u)$. The degree two Lagrange polynomial going through these points is

$$\mathbf{p}(x) = u\left[\frac{(x-2)(x-3)}{(1-2)(1-3)}\right] + s\left[\frac{(x-1)(x-3)}{(2-1)(2-3)}\right] + t\left[\frac{(x-1)(x-2)}{(3-1)(3-2)}\right].$$

It is easy to check that $p(1) = s, p(2) = t, p(3) = u$. Hence $\Im(T) = \mathrm{Col}_3$ and so $\mathrm{Rank}(T) = 3$.

$$\mathrm{Rank}(T) + \mathrm{Nullity}(T) = 1 + 3 = 4 = \mathrm{Dim}(\mathrm{Poly}_3).$$

(d) It will be useful to have the following integrals at hand:

$$\int_0^\pi \cos(x) = 0$$
$$\int_0^\pi \sin(x) = 0$$
$$\int_0^\pi \cos^2(x) = \pi/2$$
$$\int_0^\pi \sin^2(x) = \pi/2$$
$$\int_0^\pi \sin(x)\cos(x) = 0$$
$$\int_0^\pi \cos(2x)\cos(x) = 0$$
$$\int_0^\pi \sin(2x)\cos(x) = \frac{4}{3}$$
$$\int_0^\pi \cos(2x)\sin(x) = -\frac{2}{3}$$
$$\int_0^\pi \sin(2x)\sin(x) = 0$$

Now suppose

$$\mathbf{f}(x) = a + b\cos(x) + c\sin(x) + d\cos(2x) + e\sin(2x) \in \mathrm{Ker}(S).$$

Then

$$\int_0^\pi f(x)\cos(x)\,dx = b\frac{\pi}{2} + e\frac{4}{3} = 0 \tag{3.5.3}$$
$$\int_0^\pi f(x)\sin(x)\,dx = c\frac{\pi}{2} - d\frac{2}{3} = 0. \tag{3.5.4}$$

Conversely, any f satisfying the linear equations above is certainly in $\mathrm{Ker}(S)$. Hence

$$\mathrm{Ker}(S) = \{a+b\cos(x)+c\sin(x)+d\cos(2x)+e\sin(2x)\ :\ b\frac{\pi}{2}+e\frac{4}{3} = 0 \text{ and } c\frac{\pi}{2}-d\frac{2}{3} = 0\}$$

We can freely choose a since the constant will not affect the integrals. We are free to choose b, but then e is fully determined by (1) above. Similarly, we can freely choose c but then d is determined by (2). Hence

$$\mathrm{Nullity}(S) = 3.$$

We claim that $\Im(S) = \mathrm{Col}_2$. To see this, suppose

$$\begin{bmatrix} s \\ t \end{bmatrix} \in \mathrm{Col}_2 .$$

Now consider

$$f(x) = \frac{2s}{\pi}\cos(x) + \frac{2t}{\pi}\sin(x).$$

Using the table of integrals above, we see that

$$S(T) = \begin{bmatrix} s \\ t \end{bmatrix}.$$

Hence

$$\mathrm{Rank}(T) = 2.$$

Since $\mathrm{Dim}(\mathrm{Trig}_2) = 5$, the Rank-Nullity theorem is verified for S.

2. Given an example of a linear map $T : \mathrm{Col}_4 \to \mathrm{Col}_4$ such that $\mathrm{Rank}(T) = \mathrm{Nullity}(T)$.

Solution. Define T by

$$T : \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \\ 0 \\ 0 \end{bmatrix}.$$

$$\mathrm{Ker}(T) = \left\{ \begin{bmatrix} 0 \\ 0 \\ z \\ w \end{bmatrix} \in \mathrm{Col}_4 \right\}.$$

and

$$\Im(T) = \left\{ \begin{bmatrix} x \\ y \\ 0 \\ 0 \end{bmatrix} \in \mathrm{Col}_4 \right\}.$$

Both $\mathrm{Ker}(T)$ and $\Im(T)$ are isomorphic to Col_2 and hence

$$\mathrm{Rank}(T) = \mathrm{Nullity}(2) = 2.$$

3. For each of the following assertions, state whether it is *true* or *false*. If it is true, prove it. If it is false, prove that it is false.

(a) There exists a linear map $T : \mathbb{R}^5 \to \mathbb{R}^2$ such that

$$\mathrm{Ker}(T) = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 : x_1 = 3x_2 \text{ and } x_3 = x_4 = x_5\}.$$

(b) There exists a linear map $F : \mathrm{Trig}_3 \to \mathrm{Trig}_3$ such that $\mathrm{Rank}(T) = \mathrm{Nullity}(T)$.

Solution.

[1] https://en.wikipedia.org/wiki/Lagrange_polynomial

(a) The statement is false. To see this, notice that if such a map were to exist then its kernel would be 2-dimensional since any choice of x_1 and x_3 uniquely determines an element in $\mathrm{Ker}(T)$. But then by the Rank-Nullity theorem, $\mathrm{Rank}(T) = 3$ since $\mathbb{R}^5$ is 5 dimensional. But $\Im(T)$ is a subspace of $\mathbb{R}^2$ - which is absurd, since $\mathbb{R}^2$ itself is 2-dimensional.

(b) The statemnt is false. Suppose such a map were to exist. Recall that $\mathrm{Dim}(\mathrm{Trig}_3) = 7$. Then by the Rank-Nullity theorem,

$$7 = \mathrm{Rank}(T) + \mathrm{Nullity}(T) = 2\,\mathrm{Rank}(T).$$

But 7 is odd, so we have a contradiction! Thus no such map can exist.

4. Let $f(x, y, z)$ be a function on $\mathbb{R}^3$ and fix a point $\mathbf{p} = (x_0, y_0, z_0) \in \mathbb{R}^3$. For each vector $\mathbf{u} \in \mathbb{R}^3$, we can regard the derivative of f in the direction of $\mathbf{u}$ at $\mathbf{p}$ as a map

$$\begin{aligned} D_{\mathbf{p}} : \mathbb{R}^3 &\to \mathbb{R} \\ \mathbf{u} &\mapsto (\nabla f)(\mathbf{p}) \cdot \mathbf{u}. \end{aligned}$$

(a) Show that $D_{\mathbf{p}}$ as defined above is a linear map.

(b) Consider the example of $f(x, y, z) = x^2 + y^2 + z^2$. Determine $\mathrm{Ker}(D_{\mathbf{p}})$ for all points $\mathbf{p} \in \mathbb{R}^3$.

Solution.

(a) $D_{\mathbf{p}}$ being a linear maps follows from the usual properties of the dot product:

$$\begin{aligned} D_{\mathbf{p}}(\mathbf{u} + \mathbf{v}) &= \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot (\mathbf{u} + \mathbf{v}) \\ &= \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{u} + \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{v} \\ &= D_{\mathbf{p}}(\mathbf{u}) + D_{\mathbf{p}}(\mathbf{v}) \end{aligned}$$

Similarly,

$$\begin{aligned} D_{\mathbf{p}}(k\mathbf{u}) &= \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{k}u \\ &= \begin{bmatrix} k\frac{\partial f}{\partial x}(\mathbf{p}) \\ k\frac{\partial f}{\partial y}(\mathbf{p}) \\ k\frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{u} \\ &= k \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{u} \\ &= kD_{\mathbf{p}} \cdot \mathbf{u} \end{aligned}$$

And so $D_{\mathbf{p}}$ is linear.

(b) $\mathbf{u} = (x_1, y_1, z_1) \in \mathrm{Ker}(D_{\mathbf{p}})$ if and only if

$$D_{\mathbf{p}}(\mathbf{u}) = 0 \tag{3.5.5}$$

$$\iff \left(\frac{\partial f}{\partial x}(\mathbf{p}), \frac{\partial f}{\partial y}(\mathbf{p}), \frac{\partial f}{\partial z}(\mathbf{p})\right) \cdot (u_0, u_1, u_2) = 0 \tag{3.5.6}$$

$$\iff 2x_0x_1 + 2y_0y_1 + 2z_0z_1 = 0. \tag{3.5.7}$$

Geometrically, Ker($D_{\mathbf{p}}$ consists of all vectors $\mathbf{v}$ that lie tangent to a circle of radius $|\mathbf{p}|$ centred at the origin at the point $\mathbf{p}$.

5. Using the Rank-Nullity theorem, give a different proof of the fact that the image of the map C in Example 3.5.6 is $\{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}$.

Solution. For reference, we reproduce a portion of the proof in in Example 3.5.6:

"The kernel of C is the subspace of $\mathbb{R}^3$ consisting of all vectors $\mathbf{v} \in V$ such that $\mathbf{a} \times \mathbf{v} = \mathbf{0}$. From the geometric formula for the cross-product,

$$|\mathbf{a} \times \mathbf{v}| = |\mathbf{a}||\mathbf{v}| \sin\theta$$

where θ is the angle from $\mathbf{a}$ to θ, we see that

$$\mathbf{a} \times \mathbf{v} = \mathbf{0} \Leftrightarrow \mathbf{v} = \mathbf{0} \text{ or } \theta = 0 \text{ or } \theta = \pi.$$

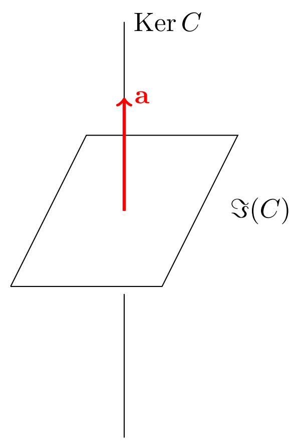
In other words, $\mathbf{v}$ must be a scalar multiple of $\mathbf{a}$. So,

$$\mathrm{Ker}(C) = \{k\mathbf{a}, k \in \mathbb{R}\}.$$

I claim that the *image* of C is the subspace of *all* vectors perpendicular to $\mathbf{a}$, i.e.

$$\Im(C) := \{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}. \tag{3.5.8}$$

If you believe me, then the picture is as follows:



Let me prove equation (3.5.1). By definition, the image of C is the subspace of $\mathbb{R}^3$ consisting of all vectors $\mathbf{w}$ of the form $\mathbf{w} = \mathbf{a} \times \mathbf{v}$ for some $\mathbf{v} \in \mathbb{R}^3$. This implies that $\mathbf{w}$ is perpendicular to $\mathbf{a}$.""

And thus we know that

$$\Im(C) \subset \{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}$$

We shall use the Rank-Nullity theorem to show the converse in a fantastically succint way. By the Rank-Nullity theorem

$$\mathrm{Dim}(\mathbb{R}^3) = \mathrm{Nullity}(C) + \mathrm{Rank}(C)$$

$$\implies 3 = 1 + \operatorname{Rank}(C).$$

And so we also know the $\Im(C)$ is a 2-dimensional subspace of $\mathbb{R}^3$. Of course,

$$\{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}$$

is also 2-dimensional. But now, if one 2-dimensional subspace is contained in another 2-dimensional subspace then the two subspaces must necessarily be the same! Hence

$$\Im(C) = \{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}$$

(By using the Rank-Nullity theorem, we managed to bypass the trickest part of Example 3.5.6!)

Solutions

· Exercises

3.5.1. Solution.

(a) id_V sends only a single vector to 0, namely the vector $\mathbf{0}_v$. Thus $\operatorname{Ker}(\mathrm{id}_v) = \{\mathbf{0}_v\}$ and hence $\operatorname{Nullity}(\mathrm{id}_v) = 0$. $\Im(\mathrm{id}_V) = V$ and so $\operatorname{Rank}(\mathrm{id}_V) = \operatorname{Dim}(V)$. The equation

$$\operatorname{Dim} V + 0 = \operatorname{Dim} V.$$

verifies the Rank-Nullity Theorem for this example.

(b) The zero map spends every element in V to $\mathbf{0}$. Hence $\operatorname{Ker}(Z) = V$ and so $\operatorname{Nullity}(Z) = \operatorname{Dim} V$. For the same reason $\Im(Z) = \{0\}$ and thus $\operatorname{Rank}(Z) = 0$. The equation

$$\operatorname{Dim} V + 0 = \operatorname{Dim} V.$$

verifies the Rank-Nullity Theorem for this example.

(c) $\mathbf{p}(x) \in \operatorname{Ker}(T)$ if and only if

$$p(1) = p(2) = p(3) = 0.$$

But any degree 3 polynomial with $1, 2$ and 3 as roots must be of the form

$$a(x-1)(x-2)(x-3)$$

where $a \in \mathbb{R}$. Conversely, any element of this form is in $\operatorname{Ker} T$. Hence

$$\operatorname{Ker}(T) = \{a(x-1)(x-2)(x-3) \in \mathrm{Poly}_3 \ : \ a \in \mathbb{R}\}$$

and thus $\operatorname{Nullity}(T) = 1$.

I claim that $\Im(T) = \mathrm{Col}_3$. That is, for any

$$\begin{bmatrix} s \\ t \\ u \end{bmatrix} \in \mathrm{Col}_3,$$

we can find a polynomial $\mathbf{p}$, of degree 3 or less, such that $p(1) = s, p(2) = t, p(3) = u$. To show this, you could set up the system of linear equations

$$a + b + c + d = s$$

$$8a + 4b + 2c + d = t$$
$$27b + 9b + 3c + d = u$$

and solve for the coefficients of $\mathbf{p}$. What is perhaps more elegant is to use the theory of Lagrange interpolation polynomials [2]. Given any 3 distinct points in $\mathbb{R}^2$, we can always find a degree two polynomial going through these 3 points. In our case, let the points be $(1, s), (2, t), (3, u)$. The degree two Lagrange polynomial going through these points is

$$\mathbf{p}(x) = u\left[\frac{(x-2)(x-3)}{(1-2)(1-3)}\right] + s\left[\frac{(x-1)(x-3)}{(2-1)(2-3)}\right] + t\left[\frac{(x-1)(x-2)}{(3-1)(3-2)}\right].$$

It is easy to check that $p(1) = s, p(2) = t, p(3) = u$. Hence $\Im(T) = \mathrm{Col}_3$ and so $\mathrm{Rank}(T) = 3$.

$$\mathrm{Rank}(T) + \mathrm{Nullity}(T) = 1 + 3 = 4 = \mathrm{Dim}(\mathrm{Poly}_3).$$

(d) It will be useful to have the following integrals at hand:

$$\int_0^\pi \cos(x) = 0$$
$$\int_0^\pi \sin(x) = 0$$
$$\int_0^\pi \cos^2(x) = \pi/2$$
$$\int_0^\pi \sin^2(x) = \pi/2$$
$$\int_0^\pi \sin(x)\cos(x) = 0$$
$$\int_0^\pi \cos(2x)\cos(x) = 0$$
$$\int_0^\pi \sin(2x)\cos(x) = \frac{4}{3}$$
$$\int_0^\pi \cos(2x)\sin(x) = -\frac{2}{3}$$
$$\int_0^\pi \sin(2x)\sin(x) = 0$$

Now suppose

$$\mathbf{f}(x) = a + b\cos(x) + c\sin(x) + d\cos(2x) + e\sin(2x) \in \mathrm{Ker}(S).$$

Then

$$\int_0^\pi f(x)\cos(x)\,dx = b\frac{\pi}{2} + e\frac{4}{3} = 0 \tag{3.5.9}$$
$$\int_0^\pi f(x)\sin(x)\,dx = c\frac{\pi}{2} - d\frac{2}{3} = 0. \tag{3.5.10}$$

Conversely, any f satisfying the linear equations above is certainly in $\mathrm{Ker}(S)$. Hence

$$\mathrm{Ker}(S) = \{a+b\cos(x)+c\sin(x)+d\cos(2x)+e\sin(2x)\ : b\frac{\pi}{2}+e\frac{4}{3} = 0 \text{ and } c\frac{\pi}{2}-d\frac{2}{3} = 0\}$$

We can freely choose a since the constant will not affect the integrals. We are free to choose b, but then e is fully determined by (1) above. Similarly, we can freely choose c but then d is determined by (2). Hence

$$\text{Nullity}(S) = 3.$$

We claim that $\Im(S) = \text{Col}_2$. To see this, suppose

$$\begin{bmatrix} s \\ t \end{bmatrix} \in \text{Col}_2\,.$$

Now consider

$$f(x) = \frac{2s}{\pi}\cos(x) + \frac{2t}{\pi}\sin(x).$$

Using the table of integrals above, we see that

$$S(T) = \begin{bmatrix} s \\ t \end{bmatrix}.$$

Hence

$$\text{Rank}(T) = 2.$$

Since $\text{Dim}(\text{Trig}_2) = 5$, the Rank-Nullity theorem is verified for S.

3.5.2. Solution. Define T by

$$T : \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \\ 0 \\ 0 \end{bmatrix}.$$

$$\text{Ker}(T) = \left\{ \begin{bmatrix} 0 \\ 0 \\ z \\ w \end{bmatrix} \in \text{Col}_4 \right\}.$$

and

$$\Im(T) = \left\{ \begin{bmatrix} x \\ y \\ 0 \\ 0 \end{bmatrix} \in \text{Col}_4 \right\}.$$

Both $\text{Ker}(T)$ and $\Im(T)$ are isomorphic to Col_2 and hence

$$\text{Rank}(T) = \text{Nullity}(2) = 2.$$

3.5.3. Solution.

(a) The statement is false. To see this, notice that if such a map were to exist then its kernel would be 2-dimensional since any choice of x_1 and x_3 uniquely determines an element in $\text{Ker}(T)$. But then by the Rank-Nullity theorem, $\text{Rank}(T) = 3$ since $\mathbb{R}^5$ is 5 dimensional. But $\Im(T)$ is a subspace of $\mathbb{R}^2$ - which is absurd, since $\mathbb{R}^2$ itself is 2-dimensional.

(b) The statemnt is false. Suppose such a map were to exist. Recall that $\text{Dim}(\text{Trig}_3) = 7$. Then by the Rank-Nullity theorem,

$$7 = \text{Rank}(T) + \text{Nullity}(T) = 2\,\text{Rank}(T).$$

But 7 is odd, so we have a contradiction! Thus no such map can exist.

3.5.4. Solution.

(a) $D_{\mathbf{p}}$ being a linear maps follows from the usual properties of the dot product:

$$\begin{aligned} D_{\mathbf{p}}(\mathbf{u}+\mathbf{v}) &= \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot (\mathbf{u}+\mathbf{v}) \\ &= \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{u} + \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{v} \\ &= D_{\mathbf{p}}(\mathbf{u}) + D_{\mathbf{p}}(\mathbf{v}) \end{aligned}$$

Similarly,

$$\begin{aligned} D_{\mathbf{p}}(k\mathbf{u}) &= \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{k}u \\ &= \begin{bmatrix} k\frac{\partial f}{\partial x}(\mathbf{p}) \\ k\frac{\partial f}{\partial y}(\mathbf{p}) \\ k\frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{u} \\ &= k \begin{bmatrix} \frac{\partial f}{\partial x}(\mathbf{p}) \\ \frac{\partial f}{\partial y}(\mathbf{p}) \\ \frac{\partial f}{\partial z}(\mathbf{p}) \end{bmatrix} \cdot \mathbf{u} \\ &= kD_{\mathbf{p}} \cdot \mathbf{u} \end{aligned}$$

And so $D_{\mathbf{p}}$ is linear.

(b) $\mathbf{u} = (x_1, y_1, z_1) \in \mathrm{Ker}(D_{\mathbf{p}})$ if and only if

$$D_{\mathbf{p}}(\mathbf{u}) = 0 \tag{3.5.11}$$

$$\iff \left(\frac{\partial f}{\partial x}(\mathbf{p}), \frac{\partial f}{\partial y}(\mathbf{p}), \frac{\partial f}{\partial z}(\mathbf{p})\right) \cdot (u_0, u_1, u_2) = 0 \tag{3.5.12}$$

$$\iff 2x_0x_1 + 2y_0y_1 + 2z_0z_1 = 0. \tag{3.5.13}$$

Geometrically, $\mathrm{Ker}(D_{\mathbf{p}}$ consists of all vectors $\mathbf{v}$ that lie tangent to a circle of radius $|\mathbf{p}|$ centred at the origin at the point $\mathbf{p}$.

3.5.5. Solution. For reference, we reproduce a portion of the proof in in Example 3.5.6:

"The kernel of C is the subspace of $\mathbb{R}^3$ consisting of all vectors $\mathbf{v} \in V$ such that $\mathbf{a} \times \mathbf{v} = \mathbf{0}$. From the geometric formula for the cross-product,

$$|\mathbf{a} \times \mathbf{v}| = |\mathbf{a}||\mathbf{v}| \sin\theta$$

where θ is the angle from $\mathbf{a}$ to θ, we see that

$$\mathbf{a} \times \mathbf{v} = \mathbf{0} \Leftrightarrow \mathbf{v} = \mathbf{0} \text{ or } \theta = 0 \text{ or } \theta = \pi.$$

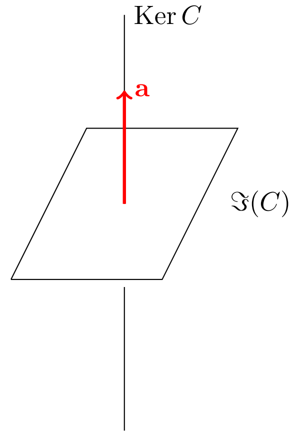
In other words, $\mathbf{v}$ must be a scalar multiple of $\mathbf{a}$. So,

$$\mathrm{Ker}(C) = \{k\mathbf{a}, k \in \mathbb{R}\}.$$

I claim that the *image* of C is the subspace of *all* vectors perpendicular to $\mathbf{a}$, i.e.

$$\Im(C) := \{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}. \tag{3.5.14}$$

If you believe me, then the picture is as follows:



Let me prove equation (3.5.1). By definition, the image of C is the subspace of $\mathbb{R}^3$ consisting of all vectors $\mathbf{w}$ of the form $\mathbf{w} = \mathbf{a} \times \mathbf{v}$ for some $\mathbf{v} \in \mathbb{R}^3$. This implies that $\mathbf{w}$ is perpendicular to $\mathbf{a}$.""

And thus we know that

$$\Im(C) \subset \{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}$$

We shall use the Rank-Nullity theorem to show the converse in a fantastically succint way. By the Rank-Nullity theorem

$$\begin{aligned} &\mathrm{Dim}(\mathbb{R}^3) = \mathrm{Nullity}(C) + \mathrm{Rank}(C) \\ \implies &3 = 1 + \mathrm{Rank}(C). \end{aligned}$$

And so we also know the $\Im(C)$ is a 2-dimensional subspace of $\mathbb{R}^3$. Of course,

$$\{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}$$

is also 2-dimensional. But now, if one 2-dimensional subspace is contained in another 2-dimensional subspace then the two subspaces must necessarily be the same! Hence

$$\Im(C) = \{\mathbf{u} \in \mathbb{R}^3 : \mathbf{u} \cdot \mathbf{a} = 0\}$$

(By using the Rank-Nullity theorem, we managed to bypass the trickest part of Example 3.5.6!)

3.6 Injective and surjective linear maps

Definition 3.6.1 A function $f : X \to Y$ from a set X to a set Y is called **one-to-one** (or **injective**) if whenever $f(x) = f(x')$ for some $x, x' \in X$ it necessarily holds that $x = x'$. The function f is called **onto** (or **surjective**) if for all $y \in Y$ there exists an $x \in X$ such that $f(x) = y$. $\Diamond$

If f is a *linear map* between vector spaces (and not just an arbitrary function between sets), there is a simple way to check if f is injective.

Lemma 3.6.2 *Let $T : V \to W$ be a linear map between vector spaces. Then:*

$$T \text{ is injective} \iff \mathrm{Ker}(T) = \{\mathbf{0}_V\}.$$

Proof. $\Rightarrow$. Suppose $T : V \to W$ is one-to-one. We already know one element in $\mathrm{Ker}(T)$, namely $\mathbf{0}_V$, since $T(\mathbf{0}_V) = \mathbf{0}_W$ as T is linear. Since T is one-to-one, this must be the only element in $\mathrm{Ker}(T)$.

$\Leftarrow$. Suppose $\mathrm{Ker}(T) = \{\mathbf{0}_V\}$. Now, suppose that

$$T(\mathbf{v}) = T(\mathbf{v}')$$

for some vectors $\mathbf{v}, \mathbf{v}' \in V$. Then we have $T(\mathbf{v}) - T(\mathbf{v}') = \mathbf{0}_W$, and since T is linear, this means $T(\mathbf{v} - \mathbf{v}') = \mathbf{0}_W$. Hence $\mathbf{v} - \mathbf{v}' \in \mathrm{Ker}(T)$, and so $\mathbf{v} - \mathbf{v}' = \mathbf{0}_V$, in other words, $\mathbf{v} = \mathbf{v}'$, which is what we wanted to show. ∎

Another simplification occurs if T is a linear map from a vector space V to itself (i.e. T is a linear operator on V), and V is finite-dimensional.

Lemma 3.6.3 *Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V. Then:*

$$T \text{ is injective} \iff T \text{ is surjective.}$$

Proof. $\Rightarrow$. Suppose T is injective.

$$\begin{aligned}
&\therefore \mathrm{Ker}(T) = \{\mathbf{0}_V\} && \text{(Lemma 3.6.2)}\\
&\therefore \mathrm{Nullity}(T) = 0 &&\\
&\therefore \mathrm{Rank}(T) = \mathrm{Dim}(V) && \text{(by Rank-Nullity Theorem)}\\
&\therefore \Im(T) = V && \text{(Corollary 2.3.27)}
\end{aligned}$$

Hence T is surjective.

$\Leftarrow$. Suppose T is surjective.

$$\begin{aligned}
&\therefore \Im(T) = V &&\\
&\therefore \mathrm{Rank}(T) = \mathrm{Dim}(V) &&\\
&\therefore \mathrm{Nullity}(T) = 0 && \text{(by Rank-Nullity Theorem)}\\
&\therefore \mathrm{Ker}(T) = \{\mathbf{0}_V\} &&\\
&\therefore T \text{ is injective.} && \text{(Corollary 2.3.27)}
\end{aligned}$$

∎

Proposition 3.6.4 *A linear map $T : V \to W$ is an isomorphism if and only if T is injective and surjective.*

Proof. $\Rightarrow$. Suppose V and W are isomorphic. That is, there exists a pair of linear maps $T : V \rightleftarrows W : S$ such that $T \circ S = \mathrm{id}_W$ and $S \circ T = \mathrm{id}_C$. We will show that T is injective and surjective.

$$\begin{aligned}
\text{Suppose that } & T(\mathbf{v}_1) = T(\mathbf{v}_2).\\
\therefore\ & S(T(\mathbf{v}_1)) = S(T(\mathbf{v}_2))\\
\therefore\ & \mathrm{id}_V(\mathbf{v}_1) = \mathrm{id}_V(\mathbf{v}_2) (\text{using } S \circ T = \mathrm{id}_V)\\
\therefore\ & \mathbf{v}_1 = \mathbf{v}_2
\end{aligned}$$

which shows that T is injective. To show that T is surjective, let $\mathbf{w} \in W$. We must show that there exists $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$. Indeed, set $\mathbf{v} := S(\mathbf{w})$. Then

$$T(\mathbf{v}) = T(S(\mathbf{w}))$$

$$= \mathrm{id}_W(\mathbf{w})(\text{using } T \circ S = \mathrm{id}_W)$$
$$= \mathbf{w}.$$

$\Leftarrow$. Suppose that there exists a linear map $T : V \to W$ which is injective and surjective. We will show that there exists a linear map $S : W \to V$ such that $S \circ T = \mathrm{id}_V$ and $T \circ S = \mathrm{id}_W$, which will prove that V and W are isomorphic.

We define the inverse map S as follows:

$$\begin{aligned} S : W &\to V \\ \mathbf{w} &\mapsto \text{the unique } \mathbf{v} \in V \text{ such that } T(\mathbf{v}) = \mathbf{w}. \end{aligned}$$

This map is well-defined. Indeed, given $\mathbf{w} \in W$, the fact that T is surjective means there does exist *some* $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$. The fact that T is injective implies that this $\mathbf{v}$ is unique. For if there exists another $\mathbf{v}' \in V$ with $S(\mathbf{v}') = \mathbf{w}$, then we have $\mathbf{v}' = \mathbf{v}$ since T is injective.

At this point we have a well-defined *function* $S : W \to V$ which satisfies $T \circ S = \mathrm{id}_W$ and $S \circ T = \mathrm{id}_V$. We only need to check that S is linear.

Let $\mathbf{w}_1, \mathbf{w}_2 \in W$. Then

$$\begin{aligned} S(a\mathbf{w}_1 + b\mathbf{w}_2) &= S\big(aT(S(\mathbf{w}_1)) + bT(S(\mathbf{w}_2))\big) && (\text{using } T \circ S = \mathrm{id}_W) \\ &= S\big(aT(\mathbf{v}_1) + bT(\mathbf{v}_2)\big) && (\text{setting } \mathbf{v}_1 := S(\mathbf{w}_1), \mathbf{v}_2 := S(\mathbf{w}_2))) \\ &= S\big(T(a\mathbf{v}_1 + b\mathbf{v}_2)\big) && (T \text{ is linear}) \\ &= a\mathbf{v}_1 + b\mathbf{v}_2 && ((S \circ T = \mathrm{id}_V) \end{aligned}$$

Hence S is linear, which completes the proof. ■

Proposition 3.6.5 *Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V. The following statements are equivalent:*

- *T is injective.*
- *T is surjective.*
- *T is an isomorphism.*

Proof. (1) is equivalent to (2) by Lemma 3.6.3. On the other hand, (1) and (2) is equivalent to (3) by Proposition 3.6.4. ■

Solutions

Chapter 4

Tutorials

4.1 W214 Linear Algebra 2019, Tutorial 1

Tutorial 1 covers Section 1.1 up until the end of Section 1.5. The following exercises have been selected.

Exercises

1. Prove that set C from Section 1.1 together with the addition operation (1.1.6), the zero vector (1.1.9) and the scalar multiplication operation (1.1.12) forms a vector space.

Solution. Firstly, we note that the addition operator, the zero vector, and scalar multiplication are all well-defined. We are required to check R1-R8.

R1: Let

$$\begin{aligned} p &= a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 \\ q &= b_4x^4 + b_3x^3 + b_2x^2 + b_1x + b_0. \end{aligned}$$

Then

$$\begin{aligned} p + q &= (a_4 + b_4)x^4 + \ldots + (a_0 + b_0) && \text{(defn of addition in C)} \\ &= (b_4 + a_4)x^4 + \ldots (b_0 + a_0) && \text{(s+t = t+s for real numbers s,t)} \\ &= q + p && \text{(defn of addition in C)} \end{aligned}$$

R2: Let $r = c_4x^4 + \ldots c_0$ Then

$$\begin{aligned} (p + q) + r &= \big[a_4 + b_4)x^4 + \ldots + (a_0 + b_0)\big] + (c_4x^4 + \ldots c_0) && \text{(defn of addition in} \\ &= ((a_4 + b_4) + c_4)x^4 + \ldots + ((a_0 + b_0) + c_4) && \text{(defn of addition in} \\ &= (a_4 + (b_4 + c_4)x^4 + \ldots + (a_0 + (b_0 + c_4)) && \text{((s+t)+u = s +(t+u) for real numbers s,t} \\ &= \big[(a_4x^4 + \ldots a_0\big] + \big[(b_4 + c_4)x^4 + \ldots (b_0 + c_0)\big] && \text{(defn of addition in} \\ &= p + (q + r) && \text{(defn of addition in} \end{aligned}$$

R3: Let $z = 0x^4 \ldots 0$. R3a:

$$\begin{aligned} z + p &= (0x^4 \ldots 0) + (a_4x^4 + \ldots a_0) && \text{(defn of addition in C)} \\ &= (0 + a_4)x^4 + \ldots (0 + a_0) && \text{(defn of addition in C)} \\ &= a_4x^4 + \ldots a_0 && \text{(0 + t = t for real numbers t)} \end{aligned}$$

$$= p$$

The rest of the rules are similar.

2. Define the set C' consisting of all polynomials of degree *exactly* 4. Show that if C' is given the addition operation (1.1.6), the zero vector (1.1.9) and the scalar multiplication operation (1.1.12) then C' *does not* form a vector space.

Hint. Give a counterexample!

Solution. We shall show that C' is not closed under addition. Let

$$p = x^4$$
$$q = -x^4 + x^3.$$

Then

$$p + q = x^3.$$

But x^3 is not in C'. Hence C' is not closed under addition and so cannot be a vector space.

3. Consider the set

$$X := \{(a_1, a_2) \in \mathbb{R}^2 : a_1 \geq 0, a_2 \geq 0\}$$

equipped with the same addition operation (1.1.4), zero vector (1.1.8) and scalar multiplication operation (1.1.10) as in A. Does X form a vector space? If not, why not?

Solution. X is not a vector space since scalar multiplication is not defined! For example, consider $(1, 1)$. $(1, 1) \in X$ but $(-1).(1, 1) = (-1, -1)$ is not.

4. Notation quiz! Say whether the following combination of symbols represents a real number or a function.

(a) f

(b) $f(x)$

(c) $k.f$

(d) $(k.f)(x)$

Solution.

(a) Function

(b) Real Number

(c) Function

(d) Real Number

5. Let $X = \{a, b, c\}$.

(a) Write down three different functions f, g, h in $\mathrm{Fun}(X)$.

(b) For each of the functions you wrote down in Item 4.1.5.a, calculate (i) $f + g$ and (ii) $3.h$.

Solution.

(a)

$$f(a) = 4$$
$$f(b) = 0$$
$$f(c) = 2$$

$$g(a) = 1$$
$$g(b) = 1$$
$$g(c) = 1$$

$$h(a) = 0$$
$$h(b) = 3$$
$$h(c) = 0$$

(b)

$$(f+g)(a) = 5$$
$$(f+g)(b) = 1$$
$$(f+g)(c) = 3$$

$$(3.h)(a) = 0$$
$$(3.h)(b) = 9$$
$$(3.h)(c) = 0$$

6. Define a strange new addition operation $\hat{+}$ on $\mathbb{R}$ by

$$x \hat{+} y := x - y, \qquad x, y \in \mathbb{R}.$$

Does $\hat{+}$ satisfy R2? If it does, prove it. If it does not, give a counterexample.

Solution. No, for example:

$$(1 \hat{+} 2) \hat{+} 3 = (1-2) - 3 = -4.$$

But

$$1 \hat{+} (2 \hat{+} 3) = 1 - (2-3) = 2.$$

7. Construct an operation $\boxplus$ on $\mathbb{R}$ satisfying R1 but not R2.

Hint. Try adjusting the formula from Exercise 4.1.6.

Solution. Define $x \boxplus y = |x - y|$. R1 is satisfied since $x \boxplus y = |x-y| = |y-x| = y \boxplus x$. However, R2 is not satisfied since $(1 \boxplus 2) \boxplus 3 = ||1-2|-3| = 2$ but $1 \boxplus (2 \boxplus 3) = |1 - |2-3|| = 0$

8. Prove that for all vectors $\mathbf{v}$ in a vector space, $-(-\mathbf{v}) = \mathbf{v}$.

Solution. The proof is as follows:

$$\begin{aligned}
-(-v) &= (-1).\left((-1).v\right) && \text{(defn of -v applied twice)} \\
&= \left((-1)(-1)\right).v && \text{(R6)} \\
&= 1.v \\
&= v && \text{(R7)}
\end{aligned}$$

9. Let V be a vector space. Suppose that a vector $\mathbf{v} \in V$ satisfies

$$5.\mathbf{v} = 2.\mathbf{v}. \tag{4.1.1}$$

Prove that $\mathbf{v} = \mathbf{0}$.

Solution.

$$\begin{aligned} & 5.v = 2.v \\ & \implies 5.v + (-2).v = 2.v + (-2).v \\ & \implies (5-2).v = (2-2).v \\ & \implies 3.v = 0.v \\ & \implies (\frac{1}{3}3).v = (\frac{1}{3}0)v \\ & \implies 1v = 0v \\ & \implies v = 0 \end{aligned}$$

10. If $k.\mathbf{v} = \mathbf{0}$ in a vector space, then it necessarily follows that $k = 0$.

Solution. False. Take $\mathbb{R}^2$ as an example. If $v = (0, 0)$ then $2.(0, 0) = (0, 0)$ but, of course, $2 \neq 0$.

11. The empty set can be equipped with data D1 , D2 , D3 satisfying the rules of a vector space.

Solution. False. In order for the empty set to be a vector space, it must have a zero vector. That is, we must be able to find some $v \in$ the empty set satisfying the axioms for the zero vector. However, since the empty set has no elements in it, by definition, we cannot ever hope to find such a v. Hence the empty set can never be a vector space.

12. Rule R7 of a vector space follows automatically from the other rules.

Solution. False. Let V be a non-zero vector space (such as $\mathbb{R}^2$). Redefine scalar multiplication as follows

$$k.v := 0 \text{ for all scalars k and all vectors v.}$$

Then V will satisfy all the rules of a vector space except R7. Thus it is not the case that R7 follows from the other rules.

4.2 W214 2019, Linear Algebra Tutorial 2

Note: You are welcome to use SageMath to help you solve some of the problems below. You can either just type into one of the provided Sage cells, or you can use the SageMath cell server.

Exercises

1.6 Subspaces.

1. Read through the webpage version of Subsection 1.6.3 (Solutions to homogenous linear differential equations), which is new and contains a lot of SageMath examples.

2. Show that the set

$$V := \{(a, -a, b, -b) : a, b \in \mathbb{R}\}$$

is a subspace of $\mathbb{R}^4$.

3. Consider the set

$$V := \{\mathbf{f} \in \text{Diff}((-1, 1)) : f'(0) = 2\}$$

Is V a subspace of $\text{Diff}((-1, 1))$? If you think it is, *prove* that it is. If you think it is not, *prove* that it is not!

4. Is $\mathbb{R}^+ := \{\mathbf{x} \in \mathbb{R} : \mathbf{x} \geq 0\}$ a subspace of $\mathbb{R}$? If you think it is, *prove* that it is. If you think it is not, *prove* that it is not!

5. Give an example of a nonempty subset V of $\mathbb{R}^2$ which is closed under scalar multiplication, but V is not a subspace of $\mathbb{R}^2$.

2.1 Linear Combinations and Span.

6. Can the polynomial $p = x^3 - x + 2 \in \text{Poly}_3$ be expressed as a linear combination of

$$p_1 = 1 + x, p_2 = x^3 + x^2 + x - 1, p_3 = x^3 - x^2 + 1 \ ?$$

Setup the appropriate system of simultaneous linear equations. Then solve these by hand, or using SageMath, as in Example Example 2.1.4.

7. Carrying on from the previous question, can the same polynomial $p = x^3 - x + 2 \in \text{Poly}_3$ be expressed as a linear combination of

$$p_1 = 1 + x, p_2 = x^3 + x^2 + x - 1, p_3 = x^3 - x^2 + 1, p_4 = 1 - x \ ?$$

Setup the appropriate system of simultaneous linear equations. Then solve these by hand, or using SageMath, as in Example Example 2.1.4.

8. Show that the polynomials

$$p_1 = 1 + x, p_2 = x^3 + x^2 + x - 1, p_3 = x^3 - x^2 + 1, p_4 = 1 - x$$

from the previous question span Poly_3. Setup the appropriate system of simultaneous linear equations. Then solve these by hand, or using SageMath, as in Example Example 2.1.8.

2.2 (Linear Independence).

9. Read through the webpage version of Section 2.1. I have added some new material, and gave examples of how to use SageMath to solve systems of linear equations.

10. $\mathcal{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} V \mathcal{S} V w V \mathcal{S}' = \{\mathbf{w}, \mathbf{v}_1, \ldots, \mathbf{v}_n\} V$

Consider the following list of matrices, thought of as vectors in $\text{Mat}_{2,2}$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix}, \mathbf{v}_5 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

11. Show that the list is linearly dependent. You are welcome to use SageMath (you will first need to setup the appropriate system of linear equations).

12. Go through the same steps as in Example 2.2.9 to find the first vector in the list which is either the zero vector or a linear combination of the preceding vectors. You are welcome to use SageMath at the points in your calculation when you need to solve a system of simultaneous linear equations.

4.3 W214 2019, Linear Algebra Tutorial 3

Exercises

1.6 Subspaces.

1. Prove or disprove: The set

$$V := \{p \in \text{Poly}_2 : p(3) = 1\}$$

is a subspce of Poly_2.

2.2 Linear independence.

2. Consider the vector space $V = \text{Fun}([0,1])$ of functions on the closed unit interval. Write down a linearly independent list containing 4 vectors in V.

2.3 Basis and Dimension.

3. Prove or disprove: there exits a basis $\{p_0, p_1, p_2, p_3\}$ of Poly_3 such that none of the polynomials p_0, p_1, p_2, p_3 have degree 2.

4. Let $W \subset \mathbb{R}^3$ be the plane orthogonal to the vector $\mathbf{v} = (1,2,3)$, as in Example 1.6.13 and Example 2.3.15. Show that $\{\mathbf{a}, \mathbf{b}\}$ is a basis for W, where

$$\mathbf{a} = (1, 0, -\frac{1}{3}), \quad \mathbf{b} = (0, 1, -\frac{1}{2}).$$

5. For each of the following, show that V is a subspace of Poly_2, find a basis for V, and compute $\text{Dim}\, V$.

(a) $V = \{p \in \text{Poly}_2 : p(0) = 0, p(2) = 0\}$

(b) $V = \{p \in \mathrm{Poly}_2 : \int_0^1 p(t)dt = 0\}$

6. Sift the list of vectors

$$\begin{aligned} &\mathbf{v}_1 = (0,0,0), && \mathbf{v}_2 = (1,0,-1), && \mathbf{v}_3 = (1,2,3) \\ &\mathbf{v}_4 = (3,4,5), && \mathbf{v}_5 = (4,8,12), && \mathbf{v}_6 = (1,1,0). \end{aligned}$$

7. Complete the following 'alternative' proof of Corollary 2.3.27.

Lemma. Suppose V is a vector space of dimension n. Then any linearly independent set of n vectors in V is a basis for V.

Proof. Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a linearly independent set of vectors in V.

Suppose that $\mathcal{B}$ is *not* a basis for V.
Therefore, $\mathcal{B}$ does not span V, since ... (a)
Therefore, there exists $\mathbf{v} \in V$ such that ... (b)
Now, add $\mathbf{v}$ to the list $\mathcal{B}$ to obtain a new list $\mathcal{B}' :=$ ... (c)
The new list $\mathcal{B}'$ is linearly independent because ... (d)
This is a contradiction because ... (e)
Hence, $\mathcal{B}$ must be a basis for V.

8. Use the Bumping Off Proposition or the Invariance of Dimension Theorem to determine if $\mathcal{B}$ is a basis for V.

(a) $V = \mathrm{Poly}_2$, $\mathcal{B} = \{2 + x^2,\ 1 - x,\ 1 + x - 3x^2,\ x - x^2\}$

(b) $V = \mathrm{Mat}_{2,2}$,

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 3 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right\}$$

(c) $V = \mathrm{Trig}_2$, $\mathcal{B} = \{\sin^2 x,\ \cos^2 x,\ 1 - \sin 2x,\ \cos 2x + 3\sin 2x\}$

(d) $V = \mathrm{Mat}_{2,2}$, $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix}, \mathbf{v}_5 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

Appendix A

Reminder about matrices

Let us recall a few things about matrices, and set up our notation.

An $n \times m$ *matrix* A is just a rectangular array of numbers, with n rows and m columns:

$$\mathsf{A} = \begin{bmatrix} \mathsf{A}_{11} & \mathsf{A}_{12} & \cdots & \mathsf{A}_{1m} \\ \mathsf{A}_{21} & \mathsf{A}_{22} & \cdots & \mathsf{A}_{2m} \\ \vdots & & \ddots & \vdots \\ \mathsf{A}_{n1} & \mathsf{A}_{n2} & \cdots & \mathsf{A}_{nm} \end{bmatrix}$$

> I will always write matrices in 'sans serif' font, eg. A. It is difficult to 'change fonts' in handwritten text, but I encourage you to at least reserve the letters A, B, C, etc. for matrices, and S, T, etc. for linear maps!

Two $n \times m$ matrices A and B can be added, to get a new $n \times m$ matrix $\mathsf{A} + \mathsf{B}$:

$$(\mathsf{A} + \mathsf{B})_{ij} := \mathsf{A}_{ij} + \mathsf{B}_{ij}$$

There is the *zero $n \times m$ matrix*:

$$\mathsf{0} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

You can also multiply an $n \times m$ matrix A by a scalar k, to get a new $n \times m$ matrix $k\mathsf{A}$:

$$(k\mathsf{A})_{ij} := k\mathsf{A}_{ij}$$

Lemma A.0.1

1. *Equipped with these operations, the set* $\mathrm{Mat}_{n,m}$ *of all* $n \times m$ *matrices is a vector space.*
2. *The dimension of* Mat_{nm} *is* nm*, with basis given by the matrices*

$$\mathsf{E}_{ij}, i = 1 \ldots n, \ j = 1 \ldots m$$

which have a 1 *in the* i*th row and* j*th column and zeroes everywhere else.*

Proof. Left as an exercise. ■

Example A.0.2 $\mathrm{Mat}_{2,2}$ has basis

$$\mathsf{E}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathsf{E}_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathsf{E}_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \mathsf{E}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

□

Usually A is a matrix, and A_{ij} is the element of the matrix at position (i, j). But now E_{ij} is a matrix in its own right! Its element at position (k, l) will be written as $(\mathsf{E}_{ij})_{kl}$. I hope you don't find this too confusing. In fact, we can write down an elegant formula for the elements of E_{ij} using the Kronecker delta symbol:

$$(\mathsf{E}_{ij})_{kl} = \delta_{ik}\delta_{jl} \tag{A.0.1}$$

Check that (A.0.1) is indeed the correct formula for the matrix elements of E_{ij}.

Example A.0.3 We will write Col_n for the vector space $\mathrm{Mat}_{n,1}$ of n-dimensional *column vectors*, and we will write the standard basis vectors E_{i1} of Col_n more simply as $\mathbf{e}_i$:

$$\mathbf{e}_1 := \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 := \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Vectors in Col_n will be written in bold sans-serif font eg. $\mathbf{v} \in \mathrm{Col}_n$. □

Equipped with these operations, the set $\mathrm{Mat}_{n,m}$ of all $n \times m$ matrices is a vector space with dimension nm . We write Col_n for the vector space $\mathrm{Mat}_{n,1}$ of n-dimensional *column vectors*.

The most important operation is *matrix multiplication*. An $n \times k$ matrix A can be multiplied from the right with a $k \times m$ matrix B to get a $n \times m$ matrix AB,

by defining the entries of AB to be

$$(\mathsf{AB})_{ij} := \mathsf{A}_{i1}\mathsf{B}_{1j} + \mathsf{A}_{i2}\mathsf{B}_{2j} + \cdots + \mathsf{A}_{ik}\mathsf{B}_{kj}.$$

Proposition A.0.4 *The above operations on matrices satisfy the following rules whenever the sums and products involved are defined:*

1. $(\mathsf{A} + \mathsf{B})\mathsf{C} = \mathsf{AC} + \mathsf{BC}$
2. $\mathsf{A}(\mathsf{B} + \mathsf{C}) = \mathsf{AB} + \mathsf{AC}$
3. $(k\mathsf{A})\mathsf{B} = \mathsf{A}(k\mathsf{B}) = k(\mathsf{AB})$
4. $(\mathsf{AB})\mathsf{C} = \mathsf{A}(\mathsf{BC})$

Proof. The proofs of (1) - (3) are all routine checks which you have hopefully done before. Let us prove (4), to practice using Σ-notation! Suppose A, B and C have sizes $n \times k$, $k \times r$ and $r \times m$ respectively, so that the matrix products make sense. Then:

$$\begin{aligned}
((\mathsf{AB})\mathsf{C})_{ij} &= \sum_{p=1}^{r} (\mathsf{AB})_{ip}\mathsf{C}_{pj} \\
&= \sum_{p=1}^{r} \left(\sum_{q=1}^{k} \mathsf{A}_{iq}\mathsf{B}_{qp} \right) \mathsf{C}_{pj} \\
&= \sum_{p,q} \mathsf{A}_{iq}\mathsf{B}_{qp}\mathsf{C}_{pj} \\
&= \sum_{q=1}^{k} \mathsf{A}_{iq} \left(\sum_{p=1}^{r} \mathsf{B}_{qp}\mathsf{C}_{pj} \right) \\
&= \sum_{q=1}^{k} \mathsf{A}_{iq}(\mathsf{BC})_{qj} \\
&= (\mathsf{A}(\mathsf{BC}))_{ij}.
\end{aligned}$$

■

I hope the Σ-notation does not confuse you in the above proof! Let me write out the exact same proof *without* Σ-notation, in the simple case where A, B and C are all 2×2 matrices, and we are trying to calculate, say, the entry at position 11.

$$\begin{aligned}
((\mathsf{AB})\mathsf{C})_{11} &= (\mathsf{AB})_{11}\mathsf{C}_{11} + (\mathsf{AB})_{12}\mathsf{C}_{21} \\
&= (\mathsf{A}_{11}\mathsf{B}_{11} + \mathsf{A}_{12}\mathsf{B}_{21})\mathsf{C}_{11} + (\mathsf{A}_{11}\mathsf{B}_{12} + \mathsf{A}_{12}\mathsf{B}_{22})\mathsf{C}_{21} \\
&= \mathsf{A}_{11}\mathsf{B}_{11}\mathsf{C}_{11} + \mathsf{A}_{12}\mathsf{B}_{21}\mathsf{C}_{11} + \mathsf{A}_{11}\mathsf{B}_{12}\mathsf{C}_{21} + \mathsf{A}_{12}\mathsf{B}_{22}\mathsf{C}_{21} \\
&= \mathsf{A}_{11}(\mathsf{B}_{11}\mathsf{C}_{11} + \mathsf{B}_{12}\mathsf{C}_{21}) + \mathsf{A}_{12}(\mathsf{B}_{21}\mathsf{C}_{11} + \mathsf{B}_{22}\mathsf{C}_{21}) \\
&= \mathsf{A}_{11}(\mathsf{BC})_{11} + \mathsf{A}_{12}(\mathsf{BC})_{21} \\
&= (\mathsf{A}(\mathsf{BC}))_{11}.
\end{aligned}$$

Do you understand the Σ-notation proof now? The crucial step (going from the second to the fourth lines) is called *exchanging the order of summation.*

The *transpose* of an $n \times m$ matrix A is the $m \times n$ matrix A^T whose entries are given by

$$(\mathsf{A}^T)_{ij} := \mathsf{A}_{ji}.$$

$A \in \mathrm{Mat}_{2,2}$

$$AB = BA$$

$B \in \mathrm{Mat}_{2,2}A$

$$\mathsf{A} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}.$$

Bibliography