

Richard Wong

Honors Multivariable Calculus

Math 32A/BH, UCLA

September 20, 2022

Preface

These lecture notes are written for Math 32A/BH, the honors multivariable calculus sequence at UCLA. Multivariable calculus is the mathematical language that allows us to describe the geometry of the physical world around us, such as the motion of planets in orbit, the behavior of electromagnetic forces, or the path of steepest ascent through the hills of Los Angeles.

The honors sequence differs from Math 32A/B in that it explores the topics of multivariable calculus with more mathematical rigor. In particular, in this course, you will explore not just the geometric meaning behind these definitions and theorems, but you will also learn why these definitions and theorems are true. I also plan to help you develop the reasoning and questioning skills needed to explore these mathematical ideas. Some questions you might see include the following:

What changes, and what stays the same when we move from single variable calculus to multivariable calculus? What does it mean to take a derivative of a multivariable function?

Moreover, this course builds the foundation for more advanced topics in mathematics, such as linear algebra, real analysis, complex analysis, and differential geometry. This course is recommended for students interested in learning about advanced mathematics.

—

These lecture notes are intended to supplement the course lecture slides, and they will provide a mix of both computational and proof-based questions. They draw on material from the following texts:

- *Calculus and Analysis in Euclidean Space*, by Shurman.
- *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach*, by Hubbard and Hubbard.
- *Multivariable Calculus*, by Rogawski and Adams.

You might find it convenient to obtain copies of these books for reference.

—

Please let me know if you notice any typos or errors, or if you have any comments or suggestions!

Contents

Preface	iii
1 Linear Algebra on \mathbb{R}^n	1
2 Analysis on \mathbb{R}^n	34
3 The Multivariable Derivative	35
4 Integration on \mathbb{R}^n	36
5 Integration on manifolds	37
6 Divergence theorem	38
Hints and Solutions	39

Chapter 1

Linear Algebra on \mathbb{R}^n

"Mathematics is the art of reducing any problem to linear algebra" - William Stein

Linear algebra is one of the foundational concepts in mathematics, and is essential to properly understanding multivariable calculus.

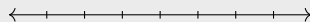
We will begin by studying certain properties of \mathbb{R}^n , and we will see how \mathbb{R}^n and its properties can be generalized to vector spaces. We will then study how vector spaces can be compared to each other, through linear maps.

1.1 An introduction to \mathbb{R}^n

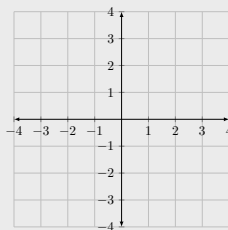
Definition 1.1.1. Given a positive integer n , ***n -dimensional Euclidean space*** (denoted \mathbb{R}^n) is the set of all ordered n -tuples of real numbers. That is,

$$\mathbb{R}^n := \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R}\}$$

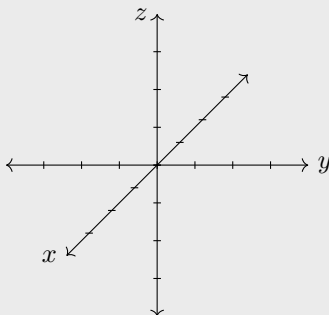
We are familiar with \mathbb{R}^1 (often described as the real line \mathbb{R}),



as well as \mathbb{R}^2 (often described as the Euclidean plane), which has the x and y axes.



\mathbb{R}^3 is often described as Euclidean space, and has 3 axes (denoted the x , y , and z axes).



We can study and understand n -dimensional Euclidean space in many different ways by imposing different *structures* on the set \mathbb{R}^n . For example, we might wish to consider \mathbb{R}^n geometrically (more precisely, as a topological space).

Thinking geometrically, an element $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ can be thought of as describing *coordinates* that tell us precisely where \mathbf{x} is located inside of \mathbb{R}^n . In this framework, we will call elements of \mathbb{R}^n points.

Definition 1.1.2. A **point in \mathbb{R}^n** (often denoted $\mathbf{x} \in \mathbb{R}^n$) is an element of \mathbb{R}^n , considered as a topological space. That is,

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

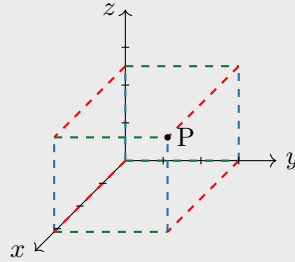
We say that x_1 is the **first coordinate**, x_2 is the **second coordinate**, and so on, until x_n (which is the n -th coordinate).

The point $O = (0, 0, \dots, 0) \in \mathbb{R}^n$ is called the **origin** (in \mathbb{R}^n).

We can also consider the point $P = (\pi, 3, 2.5) \in \mathbb{R}^3$.

The first coordinate tells us that along the x -axis, we are π units in the positive direction from the origin; the second coordinate tells us we are 3 units in the positive direction along the y -axis, and the third coordinate tells us we are 2.5 units in the positive direction along the z -axis.

We can visualize P by drawing a box such that one corner is at the origin, the edges are aligned with the axes in \mathbb{R}^3 , and the dimensions of the box correspond with the coordinates. P is then the corner of the box that is furthest from the origin.



On the other hand, we are not required to think of \mathbb{R}^n in this way. Indeed, the structures that we impose on \mathbb{R}^n should reflect how we want to think.

For example, we can instead think of an element $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$ not as describing coordinates, but rather, as describing **displacement**. In other words, given a starting point P and an ending point Q , we can measure the *change* in each coordinate. In this framework, we will call elements of \mathbb{R}^n vectors:

Definition 1.1.3. A **vector in \mathbb{R}^n** (often denoted $\mathbf{v} \in \mathbb{R}^n$) is an element of \mathbb{R}^n , considered as a vector space. That is,

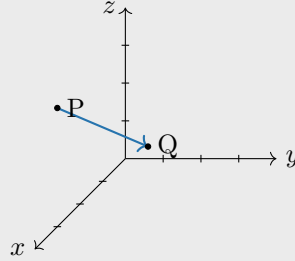
$$\mathbf{v} = \langle v_1, \dots, v_n \rangle$$

We say that v_1 is the **first component**, v_2 is the **second component**, and so on, until v_n (which is the n -th component).

Suppose that you started at the point $P = (3, 0, \pi) \in \mathbb{R}^3$, and ended up at the point $Q = (4, 3, e) \in \mathbb{R}^3$. Then your total displacement along the x -axis is $4 - 3 = 1$ unit; $3 - 0 = 3$ units along the y -axis; and $e - \pi$ units along the z -axis. Then your total displacement can be described as a vector

$$\mathbf{PQ} = \langle 1, 3, e - \pi \rangle \in \mathbb{R}^3$$

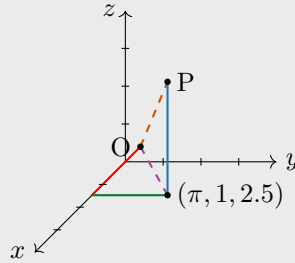
We can visualize the vector \mathbf{PQ} as an arrow in \mathbb{R}^3 , starting at P and ending at Q :



We say that the terminal point Q is the **head** of the vector \mathbf{PQ} , and that the initial point P is the **tail** of the vector \mathbf{PQ} .

Given a vector \mathbf{PQ} , we can recover information about the relationship between the points P and Q . For example, we can calculate the distance between two points P and Q in \mathbb{R}^3 by iterating the Pythagorean theorem twice:

The distance between $P = (1, 1, 1)$ and $Q = (\pi, 3, 2.5)$ is $\sqrt{(\pi - 1)^2 + (3 - 1)^2 + (2.5 - 1)^2}$



Observe that the points P , $(\pi, 1, 1)$, and $(\pi, 1, 2.5)$ form a right triangle, whose hypotenuse has length $\sqrt{(\pi - 1)^2 + (2.5 - 1)^2}$. Similarly, the points P , $(\pi, 1, 2.5)$, and Q also form a right triangle, where the length of the hypotenuse is precisely the distance between P and Q .

Observe that the distance between $P = (1, 1, 1)$ and $Q = (\pi, 3, 2.5)$ is precisely the square root of the sum of the squares of the components of the vector \mathbf{PQ} . That is, if we write the components of $\mathbf{PQ} = \langle \pi - 1, 2, 1.5 \rangle = \langle v_1, v_2, v_3 \rangle$, then the distance between P and Q is

$$\sqrt{v_1^2 + v_2^2 + v_3^2}$$

This observation generalizes to \mathbb{R}^n , as we would need to iterate the Pythagorean theorem $n - 1$ times. If we consider two points $P = (x_1, \dots, x_n)$ and $Q = (y_1, \dots, y_n)$, and write the components of the vector $\mathbf{PQ} = \langle x_1 - y_1, \dots, x_n - y_n \rangle = \langle v_1, \dots, v_n \rangle$, we again find that the distance between P and Q is

$$\sqrt{v_1^2 + \dots + v_n^2}$$

Thus, we should give this quantity a name:

Definition 1.1.4. Given a vector $\mathbf{v} = \langle v_1, \dots, v_n \rangle$, its **magnitude** is given by

$$\|\mathbf{v}\| = \sqrt{\sum_{i=1}^n v_i^2}$$

Proposition 1.1.5. The distance between two points $P = (x_1, \dots, x_n)$ and $Q = (y_1, \dots, y_n)$ is precisely $\|\mathbf{PQ}\|$.

The vector $\mathbf{PQ} = \langle 1, 3, e - \pi \rangle \in \mathbb{R}^3$ has magnitude

$$\|\mathbf{PQ}\| = \sqrt{1^2 + 3^2 + (e - \pi)^2}$$

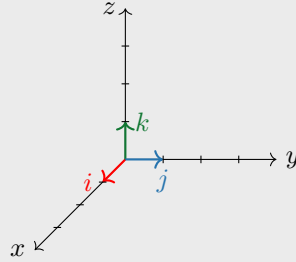
Definition 1.1.6. A **unit vector** in \mathbb{R}^n is a vector $\mathbf{u} \in \mathbb{R}^n$ such that $\|\mathbf{u}\| = 1$.

In \mathbb{R}^3 , there are special unit vectors, called the **standard basis vectors** of \mathbb{R}^3 :

$$\mathbf{i} = \langle 1, 0, 0 \rangle$$

$$\mathbf{j} = \langle 0, 1, 0 \rangle$$

$$\mathbf{k} = \langle 0, 0, 1 \rangle$$



The vector $\mathbf{e}_{\mathbf{PQ}} = \left\langle \frac{1}{\sqrt{(e-\pi)^2+10}}, \frac{3}{\sqrt{(e-\pi)^2+10}}, \frac{e-\pi}{\sqrt{(e-\pi)^2+10}} \right\rangle \in \mathbb{R}^3$ is a unit vector.

What are the differences between the points perspective and the vectors perspective?

Though the two perspectives might seem similar, there are important differences:

For example, from the points perspective, we should think of elements in \mathbb{R}^n as being analogous to *locations* on a map. It **doesn't make sense** to add the location of Los Angeles, California (say, coordinates with latitude 34° N, longitude 118° W) to the location of Austin, Texas (say, coordinates 30° N, 98° W).

On the other hand, from the vectors perspective, we should think of elements in \mathbb{R}^n as being analogous to *displacements*. In this case, let us consider two displacements: The displacement from Los Angeles to Austin means travelling 4° south, and travelling 20° east. Similarly, the displacement

from Austin to New York City (coordinates 41° N, 74° W) means travelling 11° north, and travelling 24° east.

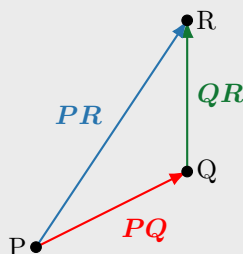
It **does make sense** to add these two displacements together, which entails travelling a total of 7° north, and 44° east. Observe that the result of this addition is the **total displacement** from Los Angeles to New York City (travelling 7° north, and 44° east).

We can formulate this in terms of vectors as follows:

Definition 1.1.7 (Vector addition). *Given two vectors $\mathbf{u} = \langle u_1, \dots, u_n \rangle$, $\mathbf{v} = \langle v_1, \dots, v_n \rangle$, we can add two vectors together to form the sum $\mathbf{u} + \mathbf{v} \in \mathbb{R}^n$. In terms of components,*

$$\mathbf{u} + \mathbf{v} = \langle u_1 + v_1, \dots, u_n + v_n \rangle$$

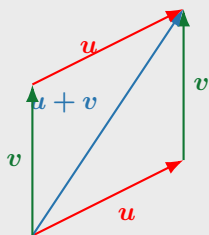
There is a nice geometric picture that accompanies vector addition: Given two vectors \mathbf{PQ} and \mathbf{QR} , then the sum of these two vectors is the vector \mathbf{PR} (that is, the vector with tail P , and head R)



That is, we add vectors "tip to tail".

Proposition 1.1.8. *Vector addition is commutative. That is, the vectors $\mathbf{u} + \mathbf{v}$ and $\mathbf{v} + \mathbf{u}$ are the same vector.*

This follows immediately from the definition in terms of components. However, there is a nice geometric picture that accompanies this property:



That is, if we move our vectors freely, the $\mathbf{u} + \mathbf{v}$ is the diagonal of the parallelogram spanned by \mathbf{u} and \mathbf{v} .

Let us be more precise about moving vectors freely: we can combine the points perspective with the vector perspective. Using our previous analogies, it makes sense to add the *displacement* from Los Angeles to Austin to the *location* New York City.

We should interpret this as saying to start at the coordinates 41° N, 74° W, and to then travel 4° south, and 20° east. The end result is a **location** with coordinates 37° N, 54° W (somewhere in the middle of the Atlantic Ocean!).

We can formulate this in terms of vectors as follows:

Definition 1.1.9. In terms of coordinates, given a point $P = (x_1, \dots, x_n) \in \mathbb{R}^n$, and a vector $\mathbf{v} = \langle v_1, \dots, v_n \rangle \in \mathbb{R}^n$, then we can define the **point** $P + \mathbf{v}$ to be the point with coordinates

$$P + \mathbf{v} = (x_1 + v_1, \dots, x_n + v_n)$$

Informally, the point $P + \mathbf{v}$ is the tip of the vector that starts at P , and travels along the length of \mathbf{v} . We can also make this mathematically precise through vector addition:

Definition 1.1.10. A **position vector** in \mathbb{R}^n is a vector whose tail (or initial point) is the origin.

If the head of the position vector is the point P , then we often denote the position vector as the vector \mathbf{OP} .

Proposition 1.1.11. Given a point $P \in \mathbb{R}^n$, and a vector $\mathbf{v} \in \mathbb{R}^n$, the point $P + \mathbf{v}$ has the same coordinates as the head (or terminal point) of the vector $\mathbf{OP} + \mathbf{v}$.

We are now able to describe various vectors at different locations in \mathbb{R}^n . We have seen cases where two vectors indeed start and end at the same point, but in definition 1.1.9, we now see that it is possible for "the same vector" to have different starting points. This leads us to the following question:

When are two vectors the same?

We first must define what we mean for two vectors to be the same:

Definition 1.1.12. A vector $\mathbf{w} = \mathbf{AB}$ is said to be **equivalent** to a vector $\mathbf{v} = \mathbf{PQ}$ if the terminal point B is the same as the point $A + \mathbf{v}$

We sometimes also say that \mathbf{w} is a **translation** of \mathbf{v} .

Suppose that we have a vector $\mathbf{v} = \mathbf{PQ} \in \mathbb{R}^n$, and suppose that the components of \mathbf{v} are $\mathbf{v} = \langle v_1, \dots, v_n \rangle$.

Consider the point $P = (v_1, \dots, v_n)$. Then \mathbf{v} is equivalent to the position vector \mathbf{OP} , which starts at the origin O , and ends at the point P .

The example above is the key idea of the proof of the following proposition:

Proposition 1.1.13. Consider the points $A, B, P, Q \in \mathbb{R}^n$. The vectors \mathbf{AB} and \mathbf{PQ} are equivalent if and only if they have the same components.

We now turn to a second way to distinguish vectors:

Definition 1.1.14. A **scalar** (often denoted $\lambda \in \mathbb{R}$) is an element of the field of real numbers, \mathbb{R} .

Definition 1.1.15. Given a vector $\mathbf{v} = \langle v_1, \dots, v_n \rangle \in \mathbb{R}^n$ and a scalar $\lambda \in \mathbb{R}$, the **scalar multiple** $\lambda \mathbf{v}$ is the vector

$$\lambda \mathbf{v} = \langle \lambda v_1, \dots, \lambda v_n \rangle$$

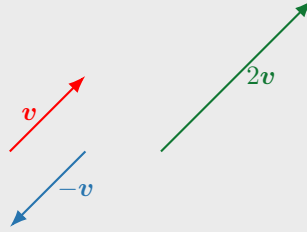
That is, we multiply λ with each component.

We can interpret scalar multiples geometrically through the following calculation:

Proposition 1.1.16. *The magnitude of a scalar multiple $\lambda \mathbf{v}$ is the same as the magnitude of \mathbf{v} times the absolute value of λ . That is,*

$$\|\lambda \mathbf{v}\| = |\lambda|(\|\mathbf{v}\|)$$

This tells us that scalar multiplication changes the magnitude of a vector (that is, it *scales* it by a factor of $|\lambda|$).



Definition 1.1.17. *Two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are said to be **parallel** if they are scalar multiples of each other. That is, there exists a scalar $\lambda \in \mathbb{R}$ such that*

$$\mathbf{v} = \lambda \mathbf{w}$$

Consider the points $P = (3, -2, 1)$ and $Q = (2, -1, 1)$ in \mathbb{R}^3 , and consider the vectors $\mathbf{u} = \langle -1, 1, 0 \rangle$, $\mathbf{v} = \langle -2, 2, 0 \rangle$, $\mathbf{w} = \langle 1, -1, 0 \rangle$, $\mathbf{a} = \langle 1, -1, 1 \rangle$.

- The vector \mathbf{PQ} is equivalent to \mathbf{u} .
- The vector \mathbf{PQ} is parallel to \mathbf{v} , with $\lambda = 2$.
- The vector \mathbf{PQ} is parallel to \mathbf{w} , with $\lambda = -1$.
- The vector \mathbf{PQ} is **not** parallel to \mathbf{a} . That is, there is no λ such that $\mathbf{PQ} = \lambda \mathbf{a}$.

If such a λ , existed, that would require that the equations $-1 = \lambda 1$ and $0 = \lambda 1$ to both be true. But that would be a contradiction, since λ cannot be both -1 and 0 .

Parallel vectors are so named because they give rise to parallel lines. See definition 1.1.22 in the exercises to learn how we describe lines in \mathbb{R}^n , and exercise 1.12.

Observe that in the example above, we saw that two vectors can be parallel, yet point in different directions. We can make this idea precise:

Definition 1.1.18. *Given a non-zero vector $\mathbf{v} \in \mathbb{R}^n$, its **direction vector** \mathbf{e}_v is the unit vector such that*

$$\|\mathbf{v}\| \mathbf{e}_v = \mathbf{v}$$

Note that we used the word "the" in the definition above - in order for this definition to make sense, we should make sure to prove that the direction vectors is *unique*.

Proposition 1.1.19. *Given a non-zero vector $\mathbf{v} \in \mathbb{R}^n$, its **direction vector** \mathbf{e}_v is unique. That is,*

$$\mathbf{e}_v = \frac{1}{\|\mathbf{v}\|} \mathbf{v}$$

Thus, we can now algebraically explain the difference between the vectors PQ and QP :

Definition 1.1.20. Suppose that the vector w is parallel to v , with $v = \lambda w$. If $\lambda > 0$, then we say w has the **same direction** as v . If $\lambda < 0$, then w has the **opposite direction** as v .

Proposition 1.1.21. Two vectors are equivalent if and only if they have the same magnitude and the same direction vector.

In other words, if two vectors have different magnitudes, then they must be different vectors. Similarly, if two vectors have different direction vectors, then they must be different vectors.

To summarize, we have learned the following:

A vector in \mathbb{R}^n is determined equivalently by:

a. an initial point and a terminal point.

$$v = PQ$$

b. its components.

$$v = \langle x_1, \dots, x_n \rangle$$

c. a direction and magnitude.

$$v = \|v\|e_v$$

Exercises

1.1 Explain the relationship between points and vectors.

More concretely, given a point in \mathbb{R}^n , is there a natural way to obtain a vector in \mathbb{R}^n ? Conversely, given a vector in \mathbb{R}^n , is there a natural way to obtain a point in \mathbb{R}^n ?

1.2 Sketch and find the components of the unit position vector in \mathbb{R}^2 that makes an angle of 30 degrees counterclockwise from the positive x -axis.

1.3 Sketch and find the components of the unit position vector in \mathbb{R}^2 that makes an angle of 30 degrees counterclockwise from the negative x -axis.

a. Is this vector equivalent to the vector in exercise 1.2?

b. Is this vector parallel to the vector in exercise 1.2?

1.4 Find the components of the unit position vector in \mathbb{R}^2 that makes an angle of θ degrees counterclockwise from the negative x -axis.

1.5 Are these vectors parallel to the vector $v = \langle 2, -1, 1 \rangle$? Are they equivalent?

a. PQ , where $P = (1, 1, 0)$, and $Q = (3, 0, 1)$.

b. QP , where $P = (1, 1, 0)$, and $Q = (3, 0, 1)$.

c. AB , where $A = (1, 1, 0)$, and $B = (5, -1, 2)$.

d. $PQ + 2AB$, where PQ and AB are as above.

e. $PQ + 2\langle 1, 0, 0 \rangle$, where PQ is as above.

1.6 Let $u = \langle 1, 1, 0 \rangle$, and let $v = \langle 0, 0, 1 \rangle$.

- a. Compute $\mathbf{u} - \mathbf{v}$.
- b. Sketch \mathbf{u} , \mathbf{v} , and $\mathbf{u} - \mathbf{v}$.
- c. In general, what is the relationship between $\mathbf{u} - \mathbf{v}$ and $\mathbf{v} - \mathbf{u}$?

1.7 Given four points, A , B , C , and D in \mathbb{R}^3 , we obtain a quadrilateral $ABCD$. If $ABCD$ is a parallelogram, and the coordinates of the corners are $A = (1, 0, 1)$, $B = (3, 3, 2)$, and $C = (2, 4, 5)$, then what are the coordinates of D ?

(**Hint:** Recall that $ABCD$ is given by a line from A to B to C to D , and then back to A . Use vectors to solve this problem!).

1.8 Consider the parallelogram $ABCD$. Prove that the diagonals BC and AD intersect each other at their midpoints.

(**Hint:** Describe the diagonals as vectors).

In the following exercises, we will describe lines in \mathbb{R}^n .

Definition 1.1.22. The line \mathcal{L} in \mathbb{R}^n , passing through the point $P = (x_1, \dots, x_n)$, in the direction of the vector $\mathbf{v} = \langle v_1, \dots, v_n \rangle$, can be described by the vector-valued function

$$\mathbf{r}(t) = \mathbf{r}_0 + t\mathbf{v}$$

where \mathbf{r}_0 is the vector $\mathbf{r}_0 = \mathbf{OP} = \langle x_1, \dots, x_n \rangle$.

We call $\mathbf{r}(t)$ the **vector parametrization of \mathcal{L}** .

Observe that by definition, the vector $\mathbf{r}(t)$ is a position vector for all t . The tips of these position vectors are precisely the points on the line \mathcal{L} .

1.9 Recall that we can describe a line in \mathbb{R}^2 as the set of points $L = \{(x, y) \mid y = mx + b\}$. Find a vector parametrization of the line L .

- a. What is the relationship between the constant b and the vector \mathbf{r}_0 ?
- b. What is the relationship between the slope m and the direction vector \mathbf{v} ?

1.10 Find the vector parametrization of the line that passes through the points $P = (1, 0, 2)$ and $Q = (2, 5, -1)$.

1.11 Given a point P and a direction vector \mathbf{v} , is the vector parametrization $\mathbf{r}(t) = \mathbf{r}_0 + t\mathbf{v}$ of a line \mathcal{L} unique?

- a. Is there a vector parametrization of \mathcal{L} using the same direction vector \mathbf{v} but a different starting point Q ? If so, find the coordinates of a possible Q .
- b. Is there a vector parametrization of \mathcal{L} using the same starting point P , but a different direction vector, \mathbf{w} ? If so, find the components of a possible \mathbf{w} .
- c. Is there a vector parametrization of \mathcal{L} using a different direction vector \mathbf{w} and a different starting point Q ? If so, find a possible Q and \mathbf{w} .

1.12 Suppose that $\mathbf{r}_1(t) = \mathbf{r}_0 + t\mathbf{v}$ and $\mathbf{r}_2(s) = \mathbf{r}_0 + s\mathbf{w}$ are two parametrizations of the same line. What is the relationship between \mathbf{v} and \mathbf{w} ?

1.13 When are $\mathbf{r}_1(t) = \mathbf{OP} + t\mathbf{v}$ and $\mathbf{r}_2(s) = \mathbf{OQ} + s\mathbf{v}$ are two parametrizations of the same line?

Definition 1.1.23. We say that two lines $\mathbf{r}_1(t)$ and $\mathbf{r}_2(s)$ **intersect** if there is a point P lying on both lines.

1.14 Determine whether the following two lines intersect. If they do, find the common point of intersection.

$$\mathbf{r}_1(t) = \langle 3, 2, 6 \rangle + t\langle 1, 0, -1 \rangle$$

$$\mathbf{r}_2(s) = \langle 4, 12, 24 \rangle + s\langle -1, 10, 20 \rangle$$

(**Hint:** If a point lies on both lines, then that point must satisfy both equations).

1.15 Determine whether the following two lines intersect. If they do, find the common point of intersection.

$$\mathbf{r}_1(t) = \langle 3, 2, 6 \rangle + t\langle 1, 0, -1 \rangle$$

$$\mathbf{r}_2(s) = \langle 4, 5, 1 \rangle + s\langle -1, 3, 3 \rangle$$

1.16 Suppose that you have two lines with parametrizations $\mathbf{r}_1(t)$ and $\mathbf{r}_2(s)$, respectively. What is the difference between solving the vector equation $\mathbf{r}_1(t) = \mathbf{r}_2(s)$ versus solving the vector equation $\mathbf{r}_1(t) = \mathbf{r}_2(t)$?

Definition 1.1.24. We say that two distinct lines are **skew**, if they do not intersect each other, and they are not parallel to each other.

1.17 Write down parametrizations of two distinct lines that are skew to each other.

1.2 Vector spaces

How can we generalize and abstract these properties of \mathbb{R}^n ?

In these lecture notes, we will only study vector spaces whose field of scalars is \mathbb{R} - that is, we are only considering *real* vector spaces¹. So in these notes we will omit the adjective real and simply refer to them as vector spaces.

Definition 1.2.1. A **vector space** V is a set with a distinguished element $\mathbf{0} \in V$, together with two operations, $+$: $V \times V \rightarrow V$, and \cdot : $\mathbb{R} \times V \rightarrow V$ satisfying the following axioms for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, and for all $\lambda, \alpha \in \mathbb{R}$.

(i) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ (Additive associativity)

(ii) $\mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v} = \mathbf{v}$ (Additive identity)

(iii) For all $\mathbf{v} \in V$, there exists $\mathbf{w} \in V$ such that $\mathbf{v} + \mathbf{w} = \mathbf{0}$ (Additive inverse)

(iv) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (Additive commutativity)

(v) $\lambda(\alpha\mathbf{v}) = (\lambda\alpha)\mathbf{v}$ (Scalar associativity)

(vi) $1\mathbf{v} = \mathbf{v}$ (Scalar identity)

(vii) $(\lambda + \alpha)\mathbf{u} = \lambda\mathbf{u} + \alpha\mathbf{u}$ (Distribution over scalar addition)

¹That is, scalars must be elements of the field \mathbb{R} (as in definition 1.1.14). In other classes, such as linear algebra, you will see that we can generalize by replacing the field \mathbb{R} with any other field \mathbb{F} , and a lot of the theory of vector spaces and linear algebra still work over an arbitrary field \mathbb{F} .

(viii) $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$ (*Distribution over vector addition*)

The elements of V are called *vectors*, and elements of \mathbb{R} are called *scalars*^a.

^asee footnote 1

When faced with an abstract definition like this, it's especially important to first find a few concrete examples to keep in mind. Luckily, from our discussion in the previous section, we have in fact checked most of these axioms:

Theorem 1.2.2. \mathbb{R}^n is a vector space.

\mathbb{R}^n is the prototypical example of a vector space². What this means is that whenever you are trying to determine whether or not something is true about vector spaces, you should try to see if it is true for \mathbb{R}^n . Then, you should try to translate the proof in the \mathbb{R}^n case to the general case, purely using the axioms and properties of abstract vector spaces.

However, you should also be careful not to trick yourself into thinking that every vector space behaves like \mathbb{R}^n ! Below is an example of a vector space that is quite different from \mathbb{R}^n : [add problem references](#)

Theorem 1.2.3. The set of $\mathcal{F}(\mathbb{R})$ of functions $f : \mathbb{R} \rightarrow \mathbb{R}$ is a vector space with pointwise addition and pointwise scalar multiplication.

$$(f + g)(x) := f(x) + g(x) \quad (\lambda f)(x) := \lambda(f(x))$$

With these examples in mind, we can now look back at the axioms of a vector space (definition 1.2.1) and see *why* these axioms are important.

Axioms (i) through (iv) describe how to *add* vectors using the operation $+$: $V \times V \rightarrow V$. From previous experiences with addition, we should require that $+$ should be (i) associative and (iv) commutative. Moreover, we also require that the zero vector $\mathbf{0} \in V$ should behave the way that $0 \in \mathbb{R}$ does, with regards to (ii) adding zero, as well as (iii) the existence of additive inverses (e.g. negative numbers in \mathbb{R}).

Axioms (v) through (vii) describe how to *scalar multiply* vectors using the operation \cdot : $\mathbb{R} \times V \rightarrow V$. The intuition for these axioms comes from our knowledge of multiplication for real numbers, which is (v) associative and (vii) distributes over addition of real numbers. Moreover, the element $1 \in \mathbb{R}$ is the *multiplicative identity* - that is, for any element $a \in \mathbb{R}$, we have that $1(a) = a$. This property of $1 \in \mathbb{R}$ is mirrored in axiom (vi).

Warning: It is important to note, however, that scalar multiplication is **not** multiplication of vectors. Indeed, scalar multiplication is a function whose input is precisely one scalar and one vector, and whose output is a vector. You will contrast this with multiplication in [problem reference](#).

In later sections, we will discuss operations whose inputs are two vectors, and whose output is a vector.

Vector subspaces

Let us now turn to other examples of vector spaces.

²In fact, in a later exercise, you can show that any **finite dimensional** vector space is isomorphic to \mathbb{R}^n .

Consider the set

$$W = \{(x, 0, 0) \in \mathbb{R}^3\}$$

The first observation to make is that W is a subset of \mathbb{R}^3 - that is, every element of W is also contained inside of \mathbb{R}^3 .

Now, if we take any two elements in $(x_1, 0, 0), (x_2, 0, 0) \in W$, we can then add them together as elements of the vector space \mathbb{R}^3 to get the vector $(x_1 + x_2, 0, 0) \in \mathbb{R}^3$. Observe that the sum $(x_1 + x_2, 0, 0)$ is still an element of W - thus the vector addition in \mathbb{R}^3 naturally defines a vector addition on W .

Furthermore, if we take an element $(x_1, 0, 0) \in W$ and a scalar $\lambda \in \mathbb{R}$, we can again treat $(x_1, 0, 0)$ as an element in \mathbb{R}^3 , and form the scalar multiple $\lambda((x_1, 0, 0)) = (\lambda x_1, 0, 0)$. Once again, the scalar multiple $\lambda((x_1, 0, 0))$ is still an element of W - thus the scalar multiplication in \mathbb{R}^3 naturally defines a scalar multiplication on W .

Thus, we have shown that if we define vector addition and scalar multiplication on W as coming from \mathbb{R}^3 , it then follows that W satisfies the axioms a vector space.

We can now give a name to this phenomenon of a subset of a vector space turning out to be a vector space:

Definition 1.2.4. A nonempty subset W of a vector space V , is a **vector subspace** of V if for all $v, w \in W$ and for all $\lambda \in \mathbb{R}$,

- (i) $v + w \in W$ (closure under addition)
- (ii) $\lambda(v) \in W$ (closure under scalar multiplication)

Again, to keep ourselves from thinking only in terms of \mathbb{R}^n , here is another vector subspace example:

Theorem 1.2.5. The set of $\mathcal{C}(\mathbb{R})$ of continuous functions $f : \mathbb{R} \rightarrow \mathbb{R}$ is a vector space with pointwise addition and pointwise scalar multiplication.

$$(f + g)(x) := f(x) + g(x) \quad (\lambda f)(x) := \lambda(f(x))$$

Theorem 1.2.6. $\mathcal{C}(\mathbb{R})$ is a subspace of $\mathcal{F}(\mathbb{R})$.

Warning: It is not true that every subset of a vector space is a vector space. For a sophisticated example, see exercise ???. For a more immediate example, consider the following:

The set of two elements $A = \{(1, 0), (0, 1)\}$ is a subset of \mathbb{R}^2 , but is not a vector subspace of \mathbb{R}^2 , as it fails to be closed under both addition and scalar multiplication.

In fact, we can generalize this example:

Proposition 1.2.7. If A is a finite set of elements in a vector space contains a non-zero vector, then A cannot be a vector subspace^a.

^aRemember, we are considering only real vector spaces in these notes.

Proof. One proof, using vector addition, is as follows: Suppose that A be a finite set of elements in a vector space with cardinality $|A| = n$, and suppose that $\mathbf{a} \in A$ is a non-zero vector. Then consider the

sums $\sum_{i=1}^j \mathbf{a}$ for $1 \leq j \leq n+1$. Since \mathbf{a} is non-zero, we have described $n+1$ different vectors, and so at least one of the sums $\sum_{i=1}^j \mathbf{a}$ cannot be contained in A . Thus A cannot be a vector subspace.

A variation on this proof can be phrased using scalar multiplication: Let W be a vector subspace of V , and let $\mathbf{w} \in W$ be a non-zero vector. Then since vector subspaces are closed under scalar multiplication, W contains $\lambda \mathbf{w}$ for all $\lambda \in \mathbb{R}$. Moreover, since \mathbf{w} is non-zero, we have shown that W must have at least as many elements as \mathbb{R} . In other words, we have described an (uncountably) infinite number of elements in W , and so any finite set of vectors in a vector space containing a non-zero vector cannot be a vector subspace. \square

To phrase this in another way, a vector subspace with finitely many elements must be the **zero** vector space. **problem**. However, one might wonder if we can fix this example in some way:

Given a (finite) set of elements A in a vector space V , can we produce a vector subspace? Even better, can we produce the **smallest** vector space containing A (that is, without including any extraneous vectors)?

We see from our observations above to construct the smallest vector space containing A , we must first include all scalar multiples of elements in A . Once we have done so, we then need to include any possible sum of scalar multiples. This is called a **linear combination**:

Definition 1.2.8. Let V be a vector space. A **linear combination** of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ is a vector of the form

$$\lambda_1 \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_k$$

where the λ_i are scalars.

Thus, we now have a candidate set for the smallest vector space containing A : the set of all linear combinations of vectors in A .

Definition 1.2.9. Let V be a vector space, and let A be a subset of vectors in V . The set of all linear combinations of vectors in A is called the **span** of A , and is written

$$\text{span}(A)$$

We first must prove that $\text{span}(A)$ is a vector subspace:

Theorem 1.2.10. Let V be a vector space, and let $A = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a set of vectors in V . Then $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a vector subspace of V .

Proof. We will first prove that $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is closed under addition. Let $\mathbf{w}_1, \mathbf{w}_2 \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$. Then by definition,

$$\mathbf{w}_1 = \lambda_1 \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_k \quad \text{and} \quad \mathbf{w}_2 = \alpha_1 \mathbf{v}_1 + \dots + \alpha_k \mathbf{v}_k$$

Then by the vector space axioms,

$$\mathbf{w}_1 + \mathbf{w}_2 = (\lambda_1 + \alpha_1) \mathbf{v}_1 + \dots + (\lambda_k + \alpha_k) \mathbf{v}_k,$$

Hence $\mathbf{w}_1 + \mathbf{w}_2 \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$.

We now prove it is closed under scalar multiplication. Suppose λ is an arbitrary scalar. Then by the vector space axioms,

$$\lambda \mathbf{w}_1 = (\lambda \lambda_1) \mathbf{v}_1 + \dots + (\lambda \lambda_k) \mathbf{v}_k,$$

Hence $\lambda \mathbf{w}_1 \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$. \square

Proposition 1.2.11. *Fix a vector $\mathbf{v} \in V$. Then*

$$\text{span}\{\mathbf{v}\} = \{\lambda\mathbf{v} \mid \lambda \in \mathbb{R}\}$$

Proof. By definition, $\{\lambda\mathbf{v} \mid \lambda \in \mathbb{R}\} \subseteq \text{span}\{\mathbf{v}\}$. On the other hand, since there is only one element in the set $\{\mathbf{v}\}$, the sum of any elements in $\{\mathbf{v}\}$ is again a scalar multiple of \mathbf{v} . Thus we have that $\{\lambda\mathbf{v} \mid \lambda \in \mathbb{R}\} \supseteq \text{span}\{\mathbf{v}\}$. \square

For example, if \mathbf{v} is a non-zero element in \mathbb{R}^n , then by definition, $\text{span}\{\mathbf{v}\}$ is a line passing through the origin.

Let $A = \{\langle 1, 0 \rangle, \langle 0, 1 \rangle\}$ be a subset of \mathbb{R}^2 . Then $\text{span}(A) = \mathbb{R}^2$.

Proof. By definition, $\text{span}(A) \subseteq \mathbb{R}^2$. On the other hand, pick an arbitrary element $\langle x, y \rangle$ in \mathbb{R}^2 . We see from the axioms that $x\langle 1, 0 \rangle + y\langle 0, 1 \rangle = \langle x, y \rangle$. Thus $\text{span}(A) \supseteq \mathbb{R}^2$. \square

We can now prove that $\text{span}(A)$ is the **smallest** vector subspace containing A .

Theorem 1.2.12. *Let V be a vector space, and let $A = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a set of vectors in V . Suppose W is a vector subspace containing A . Then W contains $\text{span}(A)$.*

Proof. **proof** \square

Bases

We have figured out how to construct a vector subspace $\text{span}(A)$ from a given subset A . One natural question to ask is the converse:

Given a vector subspace W , can we construct a subset B such that $W = \text{span}(B)$? Moreover, is it possible to find a smallest such subset B ?

Let us first consider an example:

Consider the vector subspace

$$W = \{(x, y, 0) \in \mathbb{R}^3\}$$

We can write W as the span of several different subsets:

- (i) $W = \text{span}\{(1, 0, 0), (0, 1, 0)\}$
- (ii) $W = \text{span}\{(1, 0, 0), (0, 1, 0), (2, 0, 0)\}$
- (iii) $W = \text{span}\{(1, 0, 0), (0, 1, 0), (1, 1, 0)\}$
- (iv) $W = \text{span}\{(1, 0, 0), (0, 1, 0), (0, 0, 0)\}$
- (v) $W = \text{span}\{(1, 0, 0), (0, 2, 0)\}$
- (vi) $W = \text{span}\{(1, 1, 0), (1, 2, 0)\}$

From this example, we can make a few observations: examples (ii) to (iv) **should really be thought of as example (i)**, but containing extraneous information. In other words, example (ii) contains two vectors that are scalar multiples of each other; example (iii) contains a vector that is a linear

combination of the other two vectors. Example (iv) also contains a vector that is a linear combination of the other two vectors. [reference zero](#)lineardependent

We can make this idea of "extraneous information" rigorous with the following definition:

Definition 1.2.13. A set of vectors $A \subset V$ is said to be **linearly dependent** if for every nonempty finite subset of vectors $\{v_1, \dots, v_k\} \subset A$, there exist scalars a_i , not all zero, such that

$$a_1 v_1 + \dots + a_k v_k = 0$$

Otherwise, the set A is said to be **linearly independent**.

This condition is precisely the same as our linear combination criterion as discussed above.

Proposition 1.2.14. A set of vectors $A \subset V$ is linearly dependent if and only if one of the vectors is a linear combination of the others.

In some sense, the notion of linear independence gives us a partial answer to finding a "smallest subset" that spans W :

Definition 1.2.15. An ordered set of vectors $\mathcal{B} \subset V$ is said to be a **basis** of V if

- $\text{span}(\mathcal{B}) = V$, and
- \mathcal{B} is linearly independent.

That is, a **basis** is a linearly independent spanning set.

Consider the vector subspace

$$W = \{(x, y, 0) \in \mathbb{R}^3\}$$

The following subsets are bases of W :

- (i) $\{(1, 0, 0), (0, 1, 0)\}$
- (ii) $\{(0, 1, 0), (1, 0, 0)\}$
- (iii) $\{(1, 0, 0), (0, 2, 0)\}$
- (iv) $\{(1, 1, 0), (1, 2, 0)\}$

Observe that examples (i) and (ii) are different bases, because the vectors are listed in a different order!

Examples (iii) and (iv) demonstrate the reason that we used quotes in talking about the "smallest subset". There is no meaningful relationship between these two examples and example (i) - neither of these subsets contain the other. This shows us that the set of bases of a subspace are not well-ordered - that is, there is no good notion of "smallest" when talking about bases. So for an arbitrary vector space, there is no reason to prefer any of the bases over the other.

However, specifically when talking about the vector space \mathbb{R}^n , there is a preferred basis, called the standard basis:

Definition 1.2.16. The **standard basis** of \mathbb{R}^n is the collection of vectors

$$\{e_i \mid i = 1, \dots, n\},$$

where \mathbf{e}_i denotes the vector with a 1 in the i -th coordinate and 0 elsewhere.

For example, the standard basis of \mathbb{R}^3 is the ordered set $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

Why is having a basis for a vector space useful?

Theorem 1.2.17. Given a basis $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of V , the linear combination $\mathbf{v} = \sum c_i \mathbf{v}_i$ is **unique**.

See

Definition 1.2.18. Let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be a basis of V . The **coordinates of a vector** $\mathbf{v} = \sum c_i \mathbf{v}_i$ is defined as the ordered set

$$[\mathbf{v}]_{\mathcal{B}} := \{c_1, \dots, c_n\}$$

Consider the vector $\mathbf{v} = \langle x, y, z \rangle \in \mathbb{R}^3$, and consider the standard basis in \mathbb{R}^3 . Then we can write \mathbf{v} as

$$\mathbf{v} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$$

The coordinates of \mathbf{v} in the standard basis is the ordered set (x, y, z) .

Consider the vector subspace $W = \{(x, y, 0) \in \mathbb{R}^3\}$ with basis $\{(1, 1, 0), (1, 2, 0)\}$. This

To find the coordinates of an arbitrary vector $\mathbf{v} = \langle x, y, 0 \rangle \in W$, we first observe that W is a two dimensional vector space, so the coordinates will be an ordered set of two elements, say (a, b) .

We must solve the linear system of equations

$$a(1, 1, 0) + b(1, 2, 0) = (x, y, 0)$$

So we see have two equations, $a + b = x$, and $a + 2b = y$, thus subtracting the first from the second, we see that $b = y - x$. This in turn tells us that $a + (y - x) = x$, so $a = 2x - y$.

Thus the coordinates for $\mathbf{v} = \langle x, y, 0 \rangle \in W$ is the ordered set $(2x - y, y - x)$.

The concept of a basis also helps us distinguish vector spaces.

Theorem 1.2.19. If $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for V , then any other basis has cardinality n .

Lemma 1.2.20. Suppose that the set $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for a vector space V . Then any subset of V that contains more than n vectors is linearly dependent.

Theorem 1.2.19 tells us that the dimension of a vector space V is an invariant. **todo**

Definition 1.2.21. If V has a basis of n elements, then the **dimension** of V is n .

The dimension of \mathbb{R}^n is n .

Corollary 1.2.22. Let V be a vector space of dimension n , and W is a vector space of dimension m . Then if $n \neq m$, then V and W cannot be the same vector space.

Proposition 1.2.23. *Let V be a vector space of dimension n . Then any set of n linearly independent vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset V$ forms a basis for V .*

Proposition 1.2.24. *Let V be a vector space of dimension n . Then any set of n vectors $\{\mathbf{y}_1, \dots, \mathbf{y}_n\} \subset V$ that spans V also forms a basis for V .*

Exercises

1.18 Let $\mathbf{u} = \frac{1}{\sqrt{2}}\langle 1, 1 \rangle$ and $\mathbf{v} = \frac{1}{\sqrt{2}}\langle 1, -1 \rangle$ be vectors in \mathbb{R}^2 . Write $\mathbf{w} = \langle 20, 4 \rangle \in \mathbb{R}^2$ as a linear combination of \mathbf{u} and \mathbf{v} .

1.19 Consider the set of complex numbers

$$\mathbb{C} := \{a + bi \mid a, b \in \mathbb{R}\}$$

a. Show that \mathbb{C} is a vector space with the usual addition and scalar multiplication.

b. Find a basis for \mathbb{C} as a vector space.

1.20 Prove that for any given vector space V , its distinguished element $\mathbf{0}$ must be unique.

1.21 Given a vector space V , prove that for any vector $\mathbf{v} \in V$, the scalar multiple of $0\mathbf{v}$ is the zero vector. That is,

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

1.22 Given a vector space V , prove that the set $\{\mathbf{0}\}$ is a vector subspace of V .

1.23 What is the span of the set $\{\mathbf{0}\} \in \mathbb{R}^n$?

1.24 For a fixed vector $\langle a_1, \dots, a_n \rangle \in \mathbb{R}^n$, is the set $W = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_i^n a_i x_i = 0\}$ a subspace of \mathbb{R}^n ?

1.25 For a fixed vector $\langle a_1, \dots, a_n \rangle \in \mathbb{R}^n$, is the set $W = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_i^n a_i x_i = 1\}$ a subspace of \mathbb{R}^n ?

1.26 For a fixed vector $\langle a_1, \dots, a_n \rangle \in \mathbb{R}^n$, is the set

$$W = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_i^n a_i (x_i)^2 = 0\}$$

a subspace of \mathbb{R}^n ?

1.27 Suppose that U_1 and U_2 are subspaces of a vector space V . Prove that their intersection,

$$U_1 \cap U_2 := \{\mathbf{v} \in V \mid \mathbf{v} \in U_1 \text{ and } \mathbf{v} \in U_2\}$$

is also a subspace of V .

1.28 Is the set $\{(1, 0, 1), (2, 1, 0)\}$ a basis of \mathbb{R}^3 ?

1.29 Is the set $\{(1, 0, 1), (2, 1, 0), (1, 1, -1)\}$ a basis of \mathbb{R}^3 ?

1.30 Is the set $\{(1, 0, 1), (2, 1, 0), (1, 1, 1)\}$ a basis of \mathbb{R}^3 ?

1.31 Is the set $\{(1, 0, 0), (0, 1, 0), (0, 0, 1), (0, 0, 0)\}$ a basis of \mathbb{R}^3 ?

1.32 Prove that any subset of a vector space that contains the zero vector is linearly dependent.

1.33 Show that the collection P_n of all polynomials (with real coefficients) of degree less than or equal to n is a **subspace** of the vector space $\mathcal{P}(\mathbb{R})$, and that it has $\{1, x, x^2, \dots, x^n\}$ as a basis.

1.34 Given a basis $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ of V , prove that the linear combination $\mathbf{v} = \sum c_i \mathbf{v}_i$ is unique.

1.3 Maps of vector spaces

Now that we have a good grasp on what vector spaces are, we can now study how they relate to each other. In other words, we would like to study certain kinds of maps of vector spaces. For example, in this section, we will see what it means for two vector spaces to be the same!

Let us first consider two examples of maps:

Fix a vector $\mathbf{a} = \langle a, b, c \rangle$, and consider the map $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that sends a vector $\mathbf{v} = \langle x, y, z \rangle$ to the vector $\mathbf{v} + \mathbf{a} = \langle x + a, y + b, z + c \rangle$.

Observe that the map f satisfies the following properties:

- (i) For any two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, we have that $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$.
- (ii) For any scalar $\lambda \in \mathbb{R}$, and for any vector $\mathbf{v} \in V$, we have that $f(\lambda \mathbf{u}) = \lambda f(\mathbf{u})$.

Consider the function $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that sends a vector $\langle x, y, z \rangle$ to the vector $\langle x^2, y^2, z^2 \rangle$.

Observe that the map g satisfies neither of the two properties

- (i) For any two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, we have that $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$.
- (ii) For any scalar $\lambda \in \mathbb{R}$, and for any vector $\mathbf{v} \in \mathbb{R}^3$, we have that $f(\lambda \mathbf{u}) = \lambda f(\mathbf{u})$.

See the problems [scalarspreserving](#) for examples of maps that preserve one, but not the other.

Why are properties (i) and (ii) important?

Observe that if a map f satisfies properties (i) and (ii), then it preserves the vector space structure on \mathbb{R}^n . That is, it preserves vector space addition and scalar multiplication.

Definition 1.3.1. A map of vector spaces $T : V \rightarrow W$ is called a **linear map** if

$$T\left(\sum_i^k \alpha_i x_i\right) = \sum_i^k \alpha_i T(x_i)$$

for all $k \in \mathbb{N}$, for all $\alpha_i \in \mathbb{R}$, and all vectors $x_i \in V$.

Proposition 1.3.2. A map of vector spaces $f : V \rightarrow W$ is a linear map if and only if

- (i) For any two vectors $\mathbf{u}, \mathbf{v} \in V$, we have that $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$.
- (ii) For any scalar $\lambda \in \mathbb{R}$, and for any vector $\mathbf{v} \in V$, we have that $f(\lambda \mathbf{u}) = \lambda f(\mathbf{u})$.

Is the map $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = c$ linear? For what c ?

Is the map $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = bx + c$ linear? For what b and c ?

Is the map $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = ax^2 + bx + c$ linear? For what a , b , and c ?

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the map that sends x to ax . Observe that f is linear.

Show that the map $i : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ defined by $i(x, y) = (x, y, 0)$ is linear.

Show that the map $p : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ defined by $p(x, y, z) = (x, y)$ is linear.

We can prove a lot of properties about linear maps of vector spaces simply using definition 1.3.1. Just as when we studied vector spaces, whenever you are trying to determine whether or not something is true about linear maps of vector spaces, you should try to see if it is true for linear maps in \mathbb{R}^n . Then, you should try to translate the proof to the general case, purely using the definition of a linear map.

Proposition 1.3.3. *Let $T : V \rightarrow W$ be a linear map. Then T sends the 0 object in V to the 0 object in W . That is,*

$$T(\mathbf{0}) = \mathbf{0}$$

Proposition 1.3.4. *Let $T_1, T_2 : V \rightarrow W$ be two linear maps. Prove that the sum of two linear maps $T_1 + T_2$ is a linear map.*

Proposition 1.3.5. *Let $T : V \rightarrow W$ be a linear map. Prove that for any scalar $\lambda \in \mathbb{R}$, the map λT is a linear map.*

Proposition 1.3.6. *Given two linear maps $S : V \rightarrow W$ and $T : U \rightarrow V$, prove that the composite $S \circ T : U \rightarrow W$ is also linear.*

more properties, abstractly

Definition 1.3.7. *Let $f : V \rightarrow W$ be a map of vector spaces. We say that f is **injective** or **one-to-one** (or sometimes, f is an **injection**) if the following holds:*

$$\text{For all } \mathbf{v}_1, \mathbf{v}_2 \in V, \text{ if } f(\mathbf{v}_1) = f(\mathbf{v}_2), \text{ then } \mathbf{v}_1 = \mathbf{v}_2.$$

*That is, any element in the codomain of f is the image of **at most** one element in its domain.*

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the map that sends x to ax . Show that f is injective.

Show that the map $i : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ defined by $i(x, y) = (x, y, 0)$ is injective.

Equivalently, the contrapositive statement of injectivity states that f is **injective (one-to-one)** if the following holds:

$$\text{For all } \mathbf{v}_1, \mathbf{v}_2 \in V, \text{ if } \mathbf{v}_1 \neq \mathbf{v}_2, \text{ then } f(\mathbf{v}_1) \neq f(\mathbf{v}_2).$$

Definition 1.3.8. *Let $f : V \rightarrow W$ be a map of vector spaces. We say that f is **surjective** or **onto** (or sometimes, f is a **surjection**) if the following holds:*

$$\text{For all } \mathbf{w} \in W, \text{ there exists a } \mathbf{v} \in V \text{ such that } f(\mathbf{v}) = \mathbf{w}.$$

*That is, any element in the codomain of f is the image of **at least** one element in its domain.*

Show that the map $p : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ defined by $p(x, y, z) = (x, y)$ is surjective.

Definition 1.3.9. Let $f : V \rightarrow W$ be a map of vector spaces. We say that f is **bijective** (or sometimes, f is a **bijection**) if f is both injective and surjective.

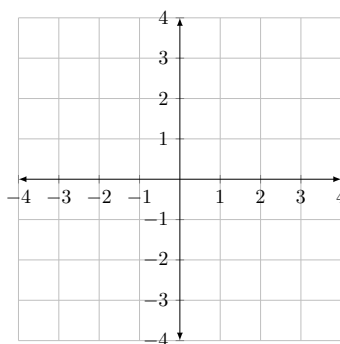
That is, any element in the codomain of f is the image of **exactly** one element in its domain.

Characterising linear maps

What do graphs of linear maps from $\mathbb{R}^n \rightarrow \mathbb{R}^m$ look like?

graph

Observe that not every straight line graph is linear!



Can we describe characterize linear maps $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$?

Let us first start with the simplest case, where $n = m = 1$.

Proposition 1.3.10. Let $T : \mathbb{R} \rightarrow \mathbb{R}$ be a linear map. Observe that if $T(1) = a$, then $T(x) = ax$.

Thus we have shown that linear maps from $\mathbb{R} \rightarrow \mathbb{R}$ must take the form of multiplication by a . From previous examples, we saw that the converse holds as well. Thus, we have classified all linear maps from $\mathbb{R} \rightarrow \mathbb{R}$

Theorem 1.3.11. A map $T : \mathbb{R} \rightarrow \mathbb{R}$ is a linear map if and only if

$$T(x) = ax$$

for some $a \in \mathbb{R}$.

How can we generalize this result?

Observe that the set $\{1\}$ is a basis for the vector space \mathbb{R} . Proposition 1.3.10 tells us that a linear map $T : \mathbb{R} \rightarrow \mathbb{R}$ is determined completely by where T sends the basis 1.

Recall also that given a basis \mathcal{B} of a vector space V , we can uniquely describe any vector $v \in V$ as a linear combination of basis vectors (Theorem 1.2.17). In other words, in terms of its coordinates.

Let us consider the vector space \mathbb{R}^3 , with the standard basis. Then, as we saw before, we can write the vector $v = \langle x, y, z \rangle \in \mathbb{R}^3$ in the standard basis:

$$v = xe_1 + ye_2 + ze_3$$

Let us consider a linear map $T : \mathbb{R}^3 \rightarrow W$, where W is an arbitrary vector space. Observe that using linearity, we have that

$$\begin{aligned} T(\mathbf{v}) &= T(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3) \\ &= xT(\mathbf{e}_1) + yT(\mathbf{e}_2) + zT(\mathbf{e}_3) \end{aligned}$$

Thus, if we know how T acts on the basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, then we can calculate $T(\mathbf{v})$ for any vector $\mathbf{v} \in \mathbb{R}^3$.

This generalizes to the following theorem:

Theorem 1.3.12. *Let V be a vector space with basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, and let W be an arbitrary vector space. A linear map $T : V \rightarrow W$ is determined by what it does on basis vectors.*

Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the linear mapping such that $T\langle 1, 0, 0 \rangle = \langle 1, 0, 1 \rangle$, $T\langle 0, 1, 0 \rangle = \langle 3, 0, 2 \rangle$, and $T\langle 0, 0, 1 \rangle = \langle 4, 2, 1 \rangle$.

Compute $T\langle 3, 3, 2 \rangle$.

Matrices

Is there a more efficient way to describe how a linear transformation acts on a basis?

Let us first set up some machinery:

Definition 1.3.13. *An $m \times n$ (real) matrix A is an array of elements $a_{i,j} \in \mathbb{R}$ with m rows and n columns:*

$$A_{m,n} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

The set of $m \times n$ (real) matrices is denoted $M_{m \times n}(\mathbb{R})$.

Theorem 1.3.14. *The set $M_{m \times n}(\mathbb{R})$, with entry-wise addition and scalar multiplication, is a vector space.*

$$A_{m,1} = \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix}$$

Observe that we can write vectors in \mathbb{R}^m as $m \times 1$ matrices.

An $m \times n$ matrix can act on a vector in \mathbb{R}^n in the following way:

Definition 1.3.15 (Matrix by vector multiplication). *Let $A \in M_{m \times n}(\mathbb{R})$, and let $\mathbf{x} \in \mathbb{R}^n$. The product $A\mathbf{x} \in \mathbb{R}^m$ is defined the vector whose i -th entry is sum $\sum_j a_{i,j}x_j$. That is,*

$$A\mathbf{x} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{bmatrix}$$

Compute $A\mathbf{x}$, where $\mathbf{x} = \langle 7, 8, 9 \rangle$, and

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

Proposition 1.3.16. *Observe that the matrix $A \in M_{m \times n}(\mathbb{R})$ determines a map $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$.*

Thus $m \times n$ matrices are not the same as $n \times m$ matrices, unless $m = n$. If $m = n$, we call $A \in M_{n \times n}(\mathbb{R})$ a **square matrix**.

Theorem 1.3.17. *Let $A \in M_{m \times n}(\mathbb{R})$. The map $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$ is linear.*

Now, we will show the converse:

Definition 1.3.18. *Given a basis $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ for \mathbb{R}^n , the **standard matrix** A of a linear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is given by*

$$[A] = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix} \in M_{m \times n}(\mathbb{R})$$

Thus, we have precisely characterized the linear maps from $\mathbb{R}^n \rightarrow \mathbb{R}^m$ - they correspond exactly to $n \times m$ matrices!

Theorem 1.3.19. *phrasing A linear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ corresponds exactly to a matrix A such that*

$$T(\mathbf{x}) = A\mathbf{x}$$

We call A the **standard matrix** of T .

$$A_{1,n} = [a_{1,1} \ a_{1,2} \ \cdots \ a_{1,n}]$$

Observe that the coordinate vector in \mathbb{R}^n is an $1 \times n$ matrix.

Definition 1.3.20. *If $A = [a_{i,j}]$, the **transpose matrix** is the matrix $A^T = [a_{j,i}]$.*

We previously saw in proposition 1.3.6 that the composition of linear maps is again linear.
What is the standard matrix of the composite?

Let $S : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $T : \mathbb{R}^p \rightarrow \mathbb{R}^n$ be linear maps, with standard matrices $A \in M_{m \times n}(\mathbb{R})$ and $B \in M_{n \times p}(\mathbb{R})$.

Then $S \circ T$ is a map from $\mathbb{R}^p \rightarrow \mathbb{R}^m$, and should have a standard matrix $C \in M_{m \times p}(\mathbb{R})$

Definition 1.3.21 (Matrix multiplication). Let $A \in M_{m \times n}(\mathbb{R})$ and $B \in M_{n \times p}(\mathbb{R})$ be two matrices. Denote the columns of B by the vectors $\mathbf{b}_1, \dots, \mathbf{b}_p$. Then the product of A and B is the matrix

$$AB = \begin{bmatrix} A(\mathbf{b}_1) & A(\mathbf{b}_2) & \cdots & A(\mathbf{b}_p) \end{bmatrix} \in M_{m \times p}(\mathbb{R})$$

In terms of entries, if $A = [a_{i,j}]$ and $B = [b_{i,j}]$, then

$$AB = \left[\sum_k a_{i,k} b_{k,j} \right]$$

Find the matrix AB , where

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 2 & 1 \end{bmatrix}$$

Theorem 1.3.22 (Properties of matrix multiplication). a. $A(BC) = (AB)C$ for $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{n \times p}(\mathbb{R})$, $C \in M_{p \times q}(\mathbb{R})$ (associativity).

b. $A(B + C) = AB + AC$ for $A \in M_{m \times n}(\mathbb{R})$, $B, C \in M_{n \times p}(\mathbb{R})$ (distribution).

c. $(A + B)C = AC + BC$ for $A, B \in M_{m \times n}(\mathbb{R})$, $C \in M_{n \times p}(\mathbb{R})$ (distribution).

d. $\lambda(AB) = (\lambda A)B = A(\lambda B)$ (scalars).

Warning: Matrix multiplication is **not** commutative.

Definition 1.3.23. The $n \times n$ **identity matrix** I_n is the standard matrix of the identity transformation $\text{id} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, which sends a vector \mathbf{v} to itself.

The $n \times n$ identity matrix has the form

$$I_n = [\delta_{ij}]$$

That is, the entry in the i -th row and j -th column is described by the **Kronecker delta** function, which is defined as

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The 3×3 identity matrix is of the form

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Proposition 1.3.24. The identity matrix $I_n \in M_{n \times n}(\mathbb{R})$ satisfies the property that for all $A \in M_{m \times n}(\mathbb{R})$,

$$I_m A = A = A I_n$$

Definition 1.3.25. A matrix $A \in M_{n \times n}(\mathbb{R})$ is **invertible** if there exists a matrix $B \in M_{n \times n}(\mathbb{R})$ such that $AB = BA = I_n$.

B is called the **inverse** of A .

Show that the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$$

has inverse

$$B = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}$$

The matrix

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

is **not** invertible. (**Hint:** use the fact that $A^2 = 0$ to obtain a contradiction)

What does invertibility look like for linear transformations?

Definition 1.3.26. A linear transformation $T : V \rightarrow W$ is **invertible** if there exists a linear transformation $S : W \rightarrow V$ such that $S \circ T = id_V$ and $T \circ S = id_W$.

If this is true, we also call T an **isomorphism of vector spaces**, and V and W are **isomorphic**.

Proposition 1.3.27. If $T : V \rightarrow W$ is linear and invertible, then $T^{-1} : W \rightarrow V$ is also linear and invertible.

The map $\square_S : \mathbb{R}^n \rightarrow M_{n \times 1}$ that sends a point to its coordinate vector (with respect to the standard basis) is an isomorphism.

The vector space $M_{m \times n}(\mathbb{R})$ is isomorphic to the vector space \mathbb{R}^{mn} .

Note that an isomorphism of vector spaces does not preserve anything except the vector space structure.

How do we determine when a linear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is an isomorphism? Equivalently, when is a matrix $A \in M_{m \times n}(\mathbb{R})$ invertible?

From our discussion about dimension earlier, the dimension of a vector space V (that is, the cardinality of any basis of V) is an invariant of V . For example, the dimension of \mathbb{R}^n is n . **Exercise: Show that an isomorphism of vector spaces must preserve dimension.**

Thus we restrict our attention to linear maps $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, or equivalently, square matrices $A \in M_{n \times n}(\mathbb{R})$. For a square matrix, we can develop the theory of **determinants**:

Definition 1.3.28. The *determinant* of a 2×2 matrix is denoted and defined as follows:

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Find the determinant of the matrix

$$A = \begin{bmatrix} 3 & 2 \\ 5 & 4 \end{bmatrix}$$

Definition 1.3.29. The *determinant* of a 3×3 matrix is denoted and defined as follows:

$$\det(A) = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

Find the determinant of the matrix

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 6 & 5 & 4 \\ 2 & 2 & 2 \end{bmatrix}$$

To see what the determinant measures, see [reference to operations](#)

How can we generalize the determinant to $n \times n$ matrices?

Theorem 1.3.30 (Expansion along the i -th row). Let A be an $n \times n$ matrix, and let A_{ij} denote the $(n-1) \times (n-1)$ matrix obtained by deleting row i and column j from A .

$$\det(A) = (-1)^{i+1}a_{i,1}\det(A_{i1}) + \cdots + (-1)^{i+n}a_{i,n}\det(A_{in})$$

Where does this formula come from? See [problem reference](#)

Properties of the determinant

Remark 1.3.31. In general, $\det(A+B) \neq \det(A) + \det(B)$.

Theorem 1.3.32. $\det(AB) = \det(A)\det(B)$.

Theorem 1.3.33. A linear transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ with standard matrix A is an isomorphism if and only if $\det(A) \neq 0$.

Theorem 1.3.34. the determinant of the transpose matrix

$$\det(A^T) = \det(A)$$

[Bucketlist of injective, surjective](#)

Exercises

1.35 Consider the map $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

$$f(x, y) = \begin{cases} x, & \text{if } y = 0 \\ y, & \text{if } y \neq 0 \end{cases}$$

Show that f preserves scalar multiplication, but not vector addition.

1.36 Prove that a map of vector spaces $f : V \rightarrow W$ is a linear map if and only if for all vectors $\mathbf{v}, \mathbf{w} \in V$, and for any scalar $\lambda \in \mathbb{R}$, we have that

$$f(\lambda \mathbf{u} + \mathbf{v}) = \lambda f(\mathbf{u}) + f(\mathbf{v})$$

1.37 Given two linear maps $S : V \rightarrow W$ and $T : U \rightarrow V$, prove that the composite $S \circ T : U \rightarrow W$ is also linear.

1.38 Let $T : V \rightarrow W$ be a linear map. Is the set

$$\text{Ker}(T) := \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}$$

a subspace of V ?

1.39 Let $T : V \rightarrow W$ be a linear map. Is the set

$$\text{Im}(T) := \{T(\mathbf{v}) \mid \mathbf{v} \in V\}$$

a subspace of W ?

1.40 Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the linear mapping such that $T\langle 0, 1, 1 \rangle = \langle 1, 0, 1 \rangle$, $T\langle 1, 0, 1 \rangle = \langle 3, 0, 2 \rangle$, and $T\langle 1, 1, 0 \rangle = \langle 4, 2, 1 \rangle$.

Compute $T(3, 3, 2)$.

1.41 Show that the rotation matrix

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

has inverse

$$S_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

1.42 Show that an isomorphism of vector spaces must preserve dimension. That is, an isomorphism sends a basis of n elements to a basis of n elements.

1.43 Show that if A is invertible, then $\det(A) \neq 0$.

In this set of problems, we will see an abstract way to define the determinant map.

First, we need to define a property of maps of vector spaces:

Definition 1.3.35. A map $M : V_1 \times \cdots \times V_k \rightarrow W$ is said to be **multilinear** if it is linear in each component. That is, for all $1 \leq i \leq k$, and for any fixed $\mathbf{v}_\ell \in V_\ell, \ell \neq i$, the map $V_i \rightarrow W$ given by

$$\mathbf{v} \mapsto M(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k)$$

is linear.

If $k = 2$, then we say M is **bilinear**, instead of multilinear.

1.44 Show that the map $\cdot : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\langle u_1, \dots, u_n \rangle \cdot \langle v_1, \dots, v_n \rangle = \sum u_i v_i$$

is bilinear.

(This map is called the **dot product**, which we will discuss in the next section).

1.45

If we think of a matrix $A = [a_{i,j}] \in M_{m \times n}(\mathbb{R})$ as an array of n column vectors, then the determinant map for 2×2 and 3×3 matrices is multilinear.

Definition 1.3.36. A multilinear map M is said to be **alternating** or **skew-symmetric** if

$$M(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_k) = -M(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_k)$$

for all $1 \leq i < j \leq k$ and for any $\mathbf{v}_\ell \in V_\ell$.

The determinant map for 2×2 and 3×3 matrices is alternating.

The dot product is **not** alternating. It is instead symmetric (commutative).

Definition 1.3.37 (The determinant). There is exists a **unique** map, $\det : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$ satisfying the following:

- \det is multilinear
- \det is alternating (skew-symmetric)
- $\det(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$.

1.4 Operations on vector spaces

So far, we have discussed vector spaces and maps of vector spaces. **Certain operations with geometric interpretations in \mathbb{R}^3**

Dot product

Definition 1.4.1. function $\cdot : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$

Given two vectors $\mathbf{u} = \langle u_1, \dots, u_n \rangle$ and $\mathbf{v} = \langle v_1, \dots, v_n \rangle$ in \mathbb{R}^n , their (**dot product**) is the scalar defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = \sum u_i v_i$$

Observe that we can identify $\mathbf{u} \cdot \mathbf{v}$ as the matrix product $\mathbf{u}^T \mathbf{v}$.

Theorem 1.4.2 (Axioms of the dot product). a. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ (commutativity).

b. $\lambda(\mathbf{u} \cdot \mathbf{v}) = (\lambda \mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot \lambda(\mathbf{v})$ (scalars).

c. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ (distribution).

d. $\mathbf{v} \cdot \mathbf{v} \geq 0$, with equality only when $\mathbf{v} = \mathbf{0}$. (positive definite).

Proposition 1.4.3. Let V be a vector space with inner product $\langle -, - \rangle$. If $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ for all $\mathbf{w} \in V$, then $\mathbf{v} = \mathbf{0}$.

Corollary 1.4.4. If $\mathbf{v} \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w}$ for all $\mathbf{w} \in V$, then $\mathbf{v} = \mathbf{u}$.

Geometry

Proposition 1.4.5. *Observe that $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$.*

We see that the dot product allows us to talk about the notion of distance in \mathbb{R}^n . Thus the dot product has a geometric interpretation!

Proposition 1.4.6. *Let θ be the angle between two non-zero vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n . Then*

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

This follows from trigonometry and the Law of Cosines. Thus, we can use the dot product in \mathbb{R}^n to determine whether the angle between the vectors \mathbf{u} and \mathbf{v} is acute, obtuse, or if \mathbf{u} and \mathbf{v} are orthogonal (that is, if $\theta = \frac{\pi}{2}$).

explicit, proof
define orthonormal

Proposition 1.4.7. *Recall that if $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthonormal basis of V , then for any $\mathbf{v} \in V$,*

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \dots + \langle \mathbf{v}, \mathbf{v}_n \rangle \mathbf{v}_n$$

Theorem 1.4.8 (Properties of the dot product). a. *Cauchy-Schwarz inequality:* $\|\mathbf{u} \cdot \mathbf{v}\| \leq \|\mathbf{u}\| \|\mathbf{v}\|$

b. *Triangle inequality:* $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

Geometric interpretation

Observe that the dot product on \mathbb{R}^n is a function $\cdot : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$. Suppose we fix a vector $\mathbf{n} \in \mathbb{R}^n$. We can then consider the function

$$\cdot \mathbf{n} : \mathbb{R}^n \rightarrow \mathbb{R}$$

defined by

$$\mathbf{v} \mapsto \mathbf{v} \cdot \mathbf{n}$$

of "taking the dot product"
Projection of vectors
orthogonal component

Given a vector $\mathbf{n} \in \mathbb{R}^n$, can we describe the set of vectors $\mathbf{v} \in \mathbb{R}^n$ such that $\mathbf{v} \cdot \mathbf{n} = \mathbf{0}$?

We know that $\mathbf{0}$ is always in this set. We also know from proposition 1.4.6 that this is equivalent to describing all of the non-zero vectors that are *orthogonal* to \mathbf{n} .

Let us

2D
span

3D
span

Proposition 1.4.9. *Given a vector $\mathbf{n} \in \mathbb{R}^n$, the set*

$$\{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} \cdot \mathbf{n} = \mathbf{0}\}$$

is a vector subspace of \mathbb{R}^n of dimension $n - 1$.

From the previous proposition, we can now generalize our examples from \mathbb{R}^2 and \mathbb{R}^3 :

Definition 1.4.10. Let $\mathbf{n} \in V$, with $\mathbf{n} \neq \mathbf{0}$. The **hyperplane** W normal to \mathbf{n} (passing through the origin) is the subspace

$$W = \{\mathbf{v} \in V \mid \mathbf{n} \cdot \mathbf{v} = 0\}$$

We say that \mathbf{n} is a normal vector of W .

Cross product

We saw previously that two non-zero vectors in \mathbb{R}^3 span a plane P in \mathbb{R}^3 . Is there a way to determine the normal vector \mathbf{n} associated to this plane?

In other words, given two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, can we construct a vector $\mathbf{w} \in \mathbb{R}^3$ that is orthogonal to both \mathbf{u} and \mathbf{v} ?

The answer is yes! **We will define the cross product of two vectors in a**

We will first define the cross product in terms of direction and magnitude.

It turns out that in order to describe the direction of the cross product, we will need to make an arbitrary choice. This choice is called the right hand rule:

Definition 1.4.11. We say that an ordered list of three orthogonal vectors \mathbf{u} , \mathbf{v} and \mathbf{w} in \mathbb{R}^3 satisfy the **Right Hand Rule** if the following is true.

- Point the fingers of your right hand in the direction of \mathbf{u}
- Curl your fingers towards \mathbf{v} (maybe rotating your hand!)
- If your thumb points in the direction of \mathbf{w} , then \mathbf{u} , \mathbf{v} and \mathbf{w} satisfy the right hand rule.

Note that the order matters! If the ordered list \mathbf{u} , \mathbf{v} and \mathbf{w} satisfies the right hand rule, then the ordered list \mathbf{v} , \mathbf{u} and \mathbf{w} does not satisfy the right hand rule.

For example, consider the standard basis vectors of \mathbb{R}^3 : $\mathbf{i} = \langle 1, 0, 0 \rangle$, $\mathbf{j} = \langle 0, 1, 0 \rangle$, and $\mathbf{k} = \langle 0, 0, 1 \rangle$.

We see that the ordered list \mathbf{i} , \mathbf{j} , and \mathbf{k} does satisfy the right hand rule, while the ordered list \mathbf{i} , \mathbf{k} , and \mathbf{j} doesn't.

We can now define the cross product in terms of direction and magnitude.

Definition 1.4.12 (Magnitude and Direction). Given two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, with an angle θ between them, their **cross product** is the unique vector $\mathbf{u} \times \mathbf{v} \in \mathbb{R}^3$ with the following properties:

- $\mathbf{u} \times \mathbf{v}$ is orthogonal to both \mathbf{u} and \mathbf{v} .
- \mathbf{u} , \mathbf{v} , and $\mathbf{u} \times \mathbf{v}$ satisfies the Right Hand Rule.
- $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin(\theta)$

Note that together, (i) and (ii) specify a unique direction for the cross product - in \mathbb{R}^3 , the set of vectors orthogonal to both \mathbf{u} and \mathbf{v} spans a one-dimensional subspace (a.k.a. a line). Thus, to specify one out of the two directions in this line, we must make an arbitrary choice via the right hand rule. Specifying a magnitude via (3) thus gives us a unique vector. **We will discuss later the choice of magnitude**

What is the cross product $\mathbf{i} \times \mathbf{j}$?

What is the cross product $\mathbf{j} \times \mathbf{k}$?

How can we compute the cross product algebraically? In other words, what are the components of the cross product?

We now give a second, algebraic definition of the cross product. This will lead to a formula for the components of the cross product [link](#)

Definition 1.4.13 (Algebraic characterization). *Given two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, their **cross product** is the unique vector $\mathbf{u} \times \mathbf{v} \in \mathbb{R}^3$ defined by the property:*

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = \det \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{bmatrix} \quad \text{for all } \mathbf{w} \in \mathbb{R}^3$$

This is an interesting perspective on defining a vector in \mathbb{R}^3 : it is determined by how it relates to all other vectors in \mathbb{R}^3 . This kind of idea shows up a lot in higher level mathematics!

First of all, using this perspective, we can prove the following properties of the cross product:

Theorem 1.4.14 (Properties of the cross product). a. $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$ (*anti-commutativity*).

b. $\mathbf{u} \times \mathbf{v}$ is orthogonal to \mathbf{u} and \mathbf{v} .

c. The cross product $\times : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is bilinear.

d. $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ if and only if \mathbf{u} and \mathbf{v} are parallel.

Moreover, we can use the algebraic characterization of the cross product to find the components of the cross product.

Recall from proposition 1.4.7, if we consider the standard basis for \mathbb{R}^3 , $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, we have that

$$\begin{aligned} \mathbf{u} \times \mathbf{v} &= \langle \mathbf{u} \times \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{u} \times \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2 + \langle \mathbf{u} \times \mathbf{v}, \mathbf{e}_3 \rangle \mathbf{e}_3 \\ &= \det \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{e}_1 \end{bmatrix} \mathbf{e}_1 + \det \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{e}_2 \end{bmatrix} \mathbf{e}_2 + \det \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{e}_3 \end{bmatrix} \mathbf{e}_3 \\ &= \left\langle \det \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{e}_1 \end{bmatrix}, \det \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{e}_2 \end{bmatrix}, \det \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{e}_3 \end{bmatrix} \right\rangle \end{aligned}$$

Thus, given the components of two vectors two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, we can compute the components of the cross product $\mathbf{u} \times \mathbf{v}$.

Theorem 1.4.15. *Given two vectors $\mathbf{u} = \langle x_1, y_1, z_1 \rangle$ and $\mathbf{v} = \langle x_2, y_2, z_2 \rangle$, their cross product $\mathbf{u} \times \mathbf{v}$ can be described as:*

$$\begin{aligned} \mathbf{u} \times \mathbf{v} &= \left\langle \det \begin{bmatrix} y_1 & z_1 \\ y_2 & z_2 \end{bmatrix}, -\det \begin{bmatrix} x_1 & z_1 \\ x_2 & z_2 \end{bmatrix}, \det \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \right\rangle \\ &= \langle y_1 z_2 - y_2 z_1, x_2 z_1 - x_1 z_2, x_1 y_2 - x_2 y_1 \rangle \end{aligned}$$

We saw previously that two non-zero vectors in \mathbb{R}^3 span a plane P in \mathbb{R}^3 , and we can now determine the normal vector \mathbf{n} associated to this plane:

Find the equation for the plane P determined by the points $P = (1, 0, -1)$, $Q = (2, 2, 1)$, $R = (4, 1, 2)$.

do this

How are these two definitions of the cross product related to each other?

In other words, (1) what does the right hand rule have to do with determinants? (2) why did we choose the magnitude of $\mathbf{u} \times \mathbf{v}$ to be $\|\mathbf{u}\| \|\mathbf{v}\| \sin(\theta)$?

More generally, what is the geometry of the determinant?

Proposition 1.4.16. *The right hand rule corresponds to having a positive determinant.*

explain

Proposition 1.4.17. *Given two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, with an angle θ between them, then*

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin(\theta)$$

First, observe that if two non-parallel vectors (say, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$) start at the same point, then they uniquely determine a triangle parallelogram.

add picture

Theorem 1.4.18. *Let P be the parallelogram spanned by \mathbf{u} and \mathbf{v} , and T be the triangle spanned by \mathbf{u} and \mathbf{v} . Then*

$$\text{area}(P) = \|\mathbf{u} \times \mathbf{v}\| \quad \text{and} \quad \text{area}(T) = \frac{1}{2} \|\mathbf{u} \times \mathbf{v}\|$$

Without loss of generality, we can rotate our vectors so that they lie in the xy -plane. (Equivalently, we can choose a basis \mathcal{B} so that the coordinates of $\mathbf{u} = \langle a, b, 0 \rangle$, and that $\mathbf{v} = \langle c, d, 0 \rangle$).

Observe that

$$\left| \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right| = \text{area}(P)$$

That is, the absolute value of a 2×2 determinant equals the area of the parallelograms spanned by the rows.

picture

In other words, the determinant of a 2×2 matrix measures the area the matrix maps basis vectors to \mathbf{u} and \mathbf{v} .

What is the volume of the parallelepiped spanned by three vectors \mathbf{u} , \mathbf{v} , and \mathbf{w} ?

picture

Theorem 1.4.19. *Let D be the parallelepiped spanned by \mathbf{u} , \mathbf{v} , and \mathbf{w} , and let θ be the angle between \mathbf{u} and $\mathbf{v} \times \mathbf{w}$. Then*

$$\text{volume}(D) = \|\mathbf{v} \times \mathbf{w}\| \|\mathbf{u}\| |\cos(\theta)|$$

Observe that we can write this formula in terms of a dot product.

Definition 1.4.20. *The scalar $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ is called the scalar triple product.*

Let D be the parallelepiped spanned by $\mathbf{u} = \langle 1, 3, 2 \rangle$, $\mathbf{v} = \langle 6, 5, 4 \rangle$, and $\mathbf{w} = \langle 2, 2, 2 \rangle$. Compute the volume of D .

This computation should seem familiar to you: see example 1.3

Theorem 1.4.21. Let D be the parallelepiped spanned by $\mathbf{u} = \langle u_1, u_2, u_3 \rangle$, $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$, and $\mathbf{w} = \langle w_1, w_2, w_3 \rangle$. Then

$$\text{volume}(D) = |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})| = \left| \det \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{bmatrix} \right|$$

Zero parallelepiped

again, basis, cube mapped to parallelepiped
Generalize to \mathbb{R}^n

Definition 1.4.22. The **standard matrix** A of a linear map $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is given by

$$A = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix} \in M_{n \times n}(\mathbb{R})$$

In other words, the absolute value of $\det A$ is the factor by which T magnifies volume; and $\det A$ is zero if and only if T is not invertible.

"The determinant is astonishing." — Jerry Shurman, Calculus and Analysis in Euclidean Space.

Inner products

How can we generalize the dot product in \mathbb{R}^n to an arbitrary vector space V ?

For \mathbb{R}^n , we saw that the dot product satisfies the property $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$. Thus, our generalization should allow us to define a notion of distance in an abstract vector space.

Definition 1.4.23. Given two vectors $\mathbf{u}, \mathbf{v} \in V$, an **inner product** on V is a map $\langle -, - \rangle : V \times V \rightarrow \mathbb{R}$ that is:

- $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ (symmetric).
- $\langle -, - \rangle$ is bilinear.
- $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, with equality only when $\mathbf{v} = \mathbf{0}$. (positive definite).

Definition 1.4.24. Given an inner product $\langle -, - \rangle$ on V , the **norm** (or **modulus**) of a vector $\mathbf{v} \in V$ is defined as

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Given two vectors $\mathbf{u} = \langle u_1, \dots, u_n \rangle$ and $\mathbf{v} = \langle v_1, \dots, v_n \rangle$ in \mathbb{R}^n , their **dot product** is an inner product.

$$\mathbf{u} \cdot \mathbf{v} = \sum u_i v_i$$

Theorem 1.4.25 (Properties of an inner product). a. *Cauchy-Schwarz inequality:* $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$

b. *Triangle inequality:* $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

What is the geometric interpretation of the inner product for an arbitrary vector space?

Definition 1.4.26. A subset $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} \subseteq V$ is said to be **orthogonal** if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for all $i \neq j$.

Furthermore, if $\|\mathbf{v}_i\| = 1$ for all $1 \leq i \leq k$, we say that the subset $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} \subseteq V$ is **orthonormal**.

Proposition 1.4.27. Equivalently, a set subset $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\} \subseteq V$ is orthonormal if

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}$$

where δ_{ij} is the **Kronecker delta**, which is the symbol

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Exercises

1.46 Find all values of b such that the vectors $\mathbf{u} = \langle b, 3, 2 \rangle$ and $\mathbf{v} = \langle 1, b, 1 \rangle$ are orthogonal.

1.47 Given that $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 3$, and $\mathbf{u} \cdot \mathbf{v} = 2$, evaluate the expression $2\mathbf{u} \cdot (3\mathbf{u} - \mathbf{v})$.

Let A be a nonempty subset of an inner product space V . The **orthogonal complement** of A is the subspace

$$A^\perp := \{\mathbf{v} \in V \mid \langle \mathbf{a}, \mathbf{v} \rangle = 0 \text{ for every } \mathbf{a} \in A\}$$

Prove that A^\perp is a vector subspace.

Let $A = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be k linearly independent vectors in a vector space V of dimension n . Prove that A^\perp is a subspace of dimension $n - k$.

Describe A^\perp is the intersection of k hyperplanes.

1.49 Calculate the cross product of $\mathbf{u} = \langle 1, 1, 0 \rangle$ and $\mathbf{v} = \langle 0, 1, 1 \rangle$

Chapter 2

Analysis on \mathbb{R}^n

Chapter 3

The Multivariable Derivative

Chapter 4

Integration on \mathbb{R}^n

Chapter 5

Integration on manifolds

Chapter 6

Divergence theorem

Hints and Solutions

1.1. Position vectors.