Linear Algebra

Analysis Series

M. Winklmeier

Chigüiro Collection

Work in progress. Use at your own risk.

Contents

Chapter 1

Introduction

In this chapter chapter we will start with one of the main topics of linear algebra: The solution of systems of linear equations. We are not only interested an efficient way to find its solutions, but we also want to understand what is possible for the solutions and how we can say something about their structure. To do so, it will be crucial to find a geometric interpretation of systems of linear equations.

A linear system is a set of equations for a number of unknowns which have to be satisfied simultaneously and where the unknowns appear only linearly. Typically the unknowns are called x, y, z or $x_1, x_2 \dots, x_n$. The following is an example of a linear system of 3 equations for 5 unknowns:

$$x_1 + x_2 + x_3 + x_4 + x_5 = 3, \quad 2x_1 + 3x_2 - 5x_3 + x_4 = 1, \quad 3x_1 - 8x_5 = 0.$$

An example of a non-linear system is

$$x_1 x_2 + x_3 + x_4 + x_5 = 3, \quad 2x_1 + 3x_2 - 5x_3 + x_4 = 1, \quad 3x_1 - 8x_5 = 0$$

because in the first equation we have a product of two of the unknowns. Also things like x^2, $\sqrt[3]{x}$, xyz, x/y or $\sin x$ would make a system non-linear.

Now let us briefly discuss the simplest non-trivial case: A system consisting of one linear equation for one unknown x. Its most general form is

$$ax = b.$$

where a and b are given constants. We want to find all $x \in \mathbb{R}$ which satisfy this equation. The solution to this problem depends on the coefficients a and b. We have to distinguish several cases.

Case 1. $a \neq 0$. In this case, there is only one solution, namely $x = b/a$.
Case 2. $a = 0$, $b \neq 0$. In this case, there is no solution because whatever value we choose for x, the left hand side ax will always be zero and therefore cannot be equal to b.
Case 3. $a = 0$, $b = 0$. In this case, there are infinitely many solutions. In fact, every $x \in \mathbb{R}$ solves the equation.

So we see that already in this simple case we have three very different structures for the solution of he system.

Now let us look at a system of one linear equation for two unknowns x, y. Its most general form is

$$ax + by = c.$$

Here, a, b, c are given constants and we want to find all pairs x, y so that the equation is satisfied. For example, if $a = b = 0$ and $c \neq 0$, then the system has no solution, whereas if for example $a \neq 0$, then there are infinitely many solutions because no matter how we choose y, we can always satisfy the system by taking $x = \frac{1}{a}(c - y)$.

qu:01:01

Question 1.1

Is it possible that the system has exactly one solution?

Remark. Come back to this question after you have studied Chapter 3.

The general form of a system of two linear equations for one unknown is

$$a_1 x = b_1, \quad a_2 x = b_2$$

and that of a system of two linear equations for two unknowns is

$$a_{11}x + a_{12}y = c_1, \quad a_{21}x + a_{22}y = c_2$$

where a_1, a_2, b_1, b_2, respectively $a_{11}, a_{12}, a_{21}, a_{22}, c_1, c_2$ are constants and x, respectively x, y are the unknowns.

qu:01:02

Question 1.2

Can you find find examples for the coefficients such that the systems have

(i) no solution,
(ii) exactly one solution,
(iii) exactly two solutions,
(iv) infinitely many solutions?

Can you maybe even give a general rule for when which behaviour occurs?

Remark. Come back to this question after you have studied Chapter 3.

In this chapter we will define what a linear system is and we will analyse in detail the case of a 2×2 system of two equations for two unknowns.

1.1 Examples of systems of linear equations; coefficient matrices

Let us start with a few examples of linear systems of linear equations.

Example 1.1. Assume that a car dealership sells motorcycles and cars. Altogether they have 25 vehicles in their shop with a total of 80 wheels. How many motorcycles and cars are in the shop?

Solution. First, we give names to the quantities we want to calculate. So let M = number of motorcyles, C = number of cars in the zoo. If we write the information given in the exercise in formulas, we obtain

$$\begin{aligned} &① & M + C &= 25, & &\text{(total number of vehicles)}\\ &② & 2M + 4C &= 80, & &\text{(total number of wheels)} \end{aligned}$$

since we assume that every motorcycle has 2 wheels and every car has 2 wheels. Equation ① tells us that $M = 25 - C$. If we insert this into equation ②, we find

$$80 = 2(25 - C) + 4C = 50 - 2C + 4C = 50 + 2C \quad\Longrightarrow\quad 2C = 30 \quad\Longrightarrow\quad C = 15.$$

This implies that $M = 25 - C = 25 - 15 = 10$. Note that in our calculations and arguments, all the implication arrows go “from left to right”, so what we can conclude at this instance is that *the system has only one possible candidate for a solution and this candidate is* $M = 10, C = 15$. We have *not* (yet) show that it really *is* a solution. However, inserting these numbers in the original equation shows that this is indeed a solution.

So the answer is: There are 10 motorcycles and 15 cars (and there is no other possibility). ◇

Let us put one more equation into the system.

Example 1.2. Assume that a car dealership sells motorcycles and cars. Altogether they have 28 vehicles in their shop with a total of 80 wheels. Moreover, the shop arranges them in 7 distinct areas of the shop so that in each area there are either 3 cars or 5 motorcycles. How many motorcycles and cars are in the shop?

Solution. Again, let M = number of motorcyles, C = number of cars. The information of the exercise gives the following system of equations:

$$\begin{aligned} &① & M + C &= 25, & &\text{(total number of vehicles)}\\ &② & 2M + 4C &= 80, & &\text{(total number of wheels)}\\ &③ & M/5 + C/3 &= 7. & &\text{(total number of areas)} \end{aligned}$$

As in the previous exercise, we obtain from that $M = 10, C = 15$. Clearly, this also satisfies equation ③. ◇

Example 1.3. Assume that a car dealership sells motorcycles and cars. Altogether they have 28 vehicles in their shop with a total of 80 wheels. Moreover, the shop arranges them in 5 distinct areas of the shop so that in each area there are either 3 cars or 5 motorcycles. How many motorcycles and cars are in the shop?

Solution. Again, let M = number of motorcycles, C = number of cars. The information of the exercise gives the following equations:

$$\begin{aligned} &① & M + C &= 25, & &\text{(total number of vehicles)}\\ &② & 2M + 4C &= 80, & &\text{(total number of wheels)}\\ &③ & M/5 + C/3 &= 5. & &\text{(total number of areas)} \end{aligned}$$

As in the previous exercise, we obtain that $M = 10$, $C = 15$ using only equations ① and ②. However, this does not satisfy equation ③; so there is no way to choose M and C such that all three equations are satisfied simultaneously. Therefore, a shop as in this example does not exist. $\diamond$

Example 1.4. Assume that a zoo has birds and cats. The total count of legs of the animals is 60. Feeding a bird takes 5 minutes, feeding a cat takes 10 minutes. The total time to feed the animals is 150 minutes. How many birds and cats are in the zoo?

Solution. Let B = number of birds, C = number of cats in the zoo. The information of the exercise gives the following equations:

$$\begin{array}{lrl} ① & 2B + 4C = 60, & \text{(total number of legs)} \\ ② & 5B + 10C = 150, & \text{(total time for feeding)} \end{array}$$

The first equation gives $B = 30 - 2C$. Inserting this into the second equation, gives

$$150 = 5(30 - 2C) + 10C = 150 - 10C + 10C = 150$$

which is always true, independently of the choice of B and C. Indeed, for instance $B = 10, C = 10$ or $B = 14, C = 8$, or $B = 0, C = 15$ are solutions. We conclude that the information given in the exercise it no sufficient to calculate the number of animals in the zoo. $\diamond$

Remark. The reason for this is that both equations ① and ② are basically the same equation. If we divide the first one by 2 and the second one by 5, then we end up in both cases with the equation $B + 2C = 30$, so both equations contain exactly the same information.

We give a few more examples.

Example 1.5. Find a polynomial P of degree at most 3 with

$$P(0) = 1, \quad P(1) = 7, \quad P'(0) = 3, \quad P'(2) = 23. \tag{1.1}$$

Solution. A polynomial of degree at most 3 is known, if we know its 4 coefficients. In this exercise, the unknowns are the coefficients of the polynomial P. If we write $P(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta$, then we have to find $\alpha, \beta, \gamma, \delta$ such that (1.1) is satisfied. Note that $P'(x) = 3\alpha x^2 + 2\beta x + \gamma$. Hence (1.1) is equivalent to the following system of equations:

$$\left.\begin{array}{r} P(0) = 1, \\ P(1) = 7, \\ P'(0) = 3, \\ P'(2) = 23. \end{array}\right\} \iff \left\{\begin{array}{lr} ① & \delta = 1, \\ ② & \alpha + \beta + \gamma + \delta = 7, \\ ③ & \gamma = 3, \\ ④ & 24\alpha + 8\beta + 2\gamma + \delta = 23. \end{array}\right.$$

Clearly, $\delta = 1$ and $\gamma = 3$. If we insert this in the remaining equations, we obtain a system of two equations for the two unknowns α, β:

$$\begin{array}{lr} ②' & \alpha + \beta = 3, \\ ④' & 24\alpha + 8\beta = 16. \end{array}$$

From (2') we obtain $\beta = 4-\alpha$. If we insert this into (4'), we get that $16 = 24\alpha + 8(4-\alpha) = 16\alpha + 32$, that is, $\alpha = (32-16)/16 = 1$. So the only possible solution is

$$\alpha = 1, \quad \beta = 2, \quad \gamma = 3, \quad \delta = 1.$$

It is easy to verify that the polynomial $P(x) = x^3 + 2x^2 + 3x + 1$ has all the desired properties. $\diamond$

Example 1.6. A pole is 5 metres long and shall be coated with varnish. There are two types of varnish available: The green one adds 3 g per 50 cm to the pole, the red one adds 6 g per meter to the pole. Is it possible to coat the pole in a combination of the varnishes so that the total weight added is

(a) 35 g? (b) 30 g?

Solution. (a) We call g the length of the pole which will be covered in green and r the length of the pole which will be covered in red. Then we obtain the system of equations

$$\begin{array}{lrl} (1) & g + r = 5 & \text{(total length)} \\ (2) & 6g + 6r = 35 & \text{(total weight)} \end{array}$$

The first equation gives $r = 5-g$. Inserting into the second equation yields $35 = 6g + 6(5-g) = 30$ which is a contradiction. This shows that there is no solution.
(b) As in (a), we obtain the system of equations

$$\begin{array}{lrl} (1) & g + r = 5 & \text{(total length)} \\ (2) & 6g + 6r = 30 & \text{(total weight)} \end{array}$$

Again, the first equation gives $r = 5-g$. Inserting into the second equation yields $30 = 6g+6(5-g) = 30$ which is always true, independently of how we choose g and r as long as (1) is satisfied. This means that in order to solve the system of equations, it is sufficient to solve only the first equation since then the second one is automatically satisfied. So we have infinitely many solutions. Any pair g, r such that $g + r = 5$ gives a solution. So for any g that we choose, we only have to set $r = 5-g$ and we have a solution of the problem. Of course, we could also fix r and then choose $g = 5-r$ to obtain a solution.
For example, we could choose $g = 1$, then $r = 4$, or $g = 0.00001$, then $r = 4.99999$, or $r = -2$ then $g = 7$. Clearly, the last example does not make sense for the problem at hand, but it still does satisfy our system of equations. $\diamond$

All the examples were so-called linear systems of linear equations. Let us give a precise definition of what we mean by this.

Definition 1.7. A $m \times n$ *system of linear equations* (or simply a *linear system*)is a system of m linear equations for n unknowns of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad\qquad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1.2}$$

The unknowns are $x_1, \dots, x_n$. The numbers a_{ij} and b_i ($i = 1, \dots, m$, $j = 1, \dots, n$) are given. The numbers a_{ij} are called the *coefficients of the linear system* and the numbers $b_1, \dots, b_n$ are called the *right side of the linear system.*

In the special case when all b_i are equal to 0, the system is called a *homogeneous*; otherwise it is called *inhomogeneous.*
The system (1.2) is called *consistent* if it has at least one solution. It is called *inconsistent* if it has no solution.

The *coefficient matrix* A of the system is the collection of all coefficients a_{ij}

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \tag{1.3}$$

The *augmented coefficient matrix* A of the system is the collection of all coefficients a_{ij} and the right hand side

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{array}\right). \tag{1.4}$$

The coefficient matrix is nothing else than the collection of the coefficients a_{ij} ordered in some sort of table or rectangle such that the place of the coefficient a_{ij} is in the ith row of the jth column. The augmented coefficient matrix contains additionally the constants from the right hand side.

Important observation. There is one-to-one correspondence between linear systems and augmented coefficient matrices: Given a linear system, it is easy to write down its augmented coefficient matrix. On the other hand, given an augmented coefficient matrix, it is easy to reconstruct the corresponding linear system.

Let us write down the coefficient matrices of our examples.

Example 1.1: This is a 2×2 system with coefficients $a_{11} = 1$, $a_{11} = 1$, $a_{21} = 2$, $a_{22} = 4$ and right hand side $b_1 = 60, b_2 = 200$. The system has a unique solution. The coefficient matrix and the augmented coefficient matrix are

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix}, \qquad (A|b) = \left(\begin{array}{cc|c} 1 & 1 & 60 \\ 2 & 4 & 200 \end{array}\right).$$

Example 1.2: This is a 3×2 system with coefficients $a_{11} = 1$, $a_{11} = 1$, $a_{21} = 2$, $a_{22} = 4$, $a_{31} = 2$, $a_{32} = 3$, and right hand side $b_1 = 60, b_2 = 200$, $b_3 = 140$. The system has a unique solution. The coefficient matrix and the augmented coefficient matrix are

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 4 \\ 2 & 3 \end{pmatrix}, \qquad (A|b) = \left(\begin{array}{cc|c} 1 & 1 & 60 \\ 2 & 4 & 200 \\ 2 & 3 & 140 \end{array}\right),$$

Example 1.3: This is a 3×2 system with coefficients $a_{11} = 1$, $a_{11} = 1$, $a_{21} = 2$, $a_{22} = 4$, $a_{31} = 2$, $a_{32} = 3$, and right hand side $b_1 = 60, b_2 = 200$, $b_3 = 100$. The system has no solution. The coefficient matrix is the same as in Example 1.2, the augmented coefficient matrix is

$$(A|b) = \left(\begin{array}{cc|c} 1 & 1 & 60 \\ 2 & 4 & 200 \\ 2 & 3 & 100 \end{array}\right),$$

Example 1.5: This is a 4×4 system with coefficients $a_{11} = 0$, $a_{12} = 0$, $a_{13} = 0$, $a_{14} = 1$, $a_{21} = 1$, $a_{22} = 1$, $a_{23} = 1$, $a_{24} = 1$, $a_{31} = 0$, $a_{32} = 0$, $a_{33} = 1$, $a_{34} = 0$, $a_{41} = 24$, $a_{42} = 8$, $a_{43} = 2$, $a_{44} = 1$, and right hand side $b_1 = 1, b_2 = 7$, $b_3 = 3$, $b_4 = 23$. The system has a unique solution. The coefficient matrix and the augmented coefficient matrix are

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 24 & 8 & 2 & 1 \end{pmatrix}, \quad A = \left(\begin{array}{cccc|c} 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 7 \\ 0 & 0 & 1 & 0 & 3 \\ 24 & 8 & 2 & 1 & 23 \end{array}\right).$$

We saw that Examples 1.1, 1.2, 1.5, 1.6 (b) have unique solutions. In Example 1.6 (b) the solution is not unique (Very much on the contrary. They even have infinitely many solutions!). Examples 1.3 and 1.6(a) do not admit solutions. So given an $m \times n$ system of linear equations, two important questions arise naturally:

- *Existence*: Does the system have a solution?
- *Uniqueness*: If the system has a solution, is it unique?

More generally, we would like to be able so say something about the structure of solutions of linear systems. For example, is it possible that there is only one solution? That there are exactly two solutions? That there are infinite solutions? That there is is no solution? Can we give criteria for existence and/or uniqueness of solutions? Can we give criteria for existence of infinite solutions? Is there an efficient way to calculate all the solutions of a linear system?

(Spoiler alert: *A system of linear equations has either no or exactly one or infinite solutions. It is not possible that it has, e.g., exactly* 7 *solutions.* This will be discussed in detail in Chapter **??**.**??**)

Before answering these questions for general $m \times n$ systems, we will have a closer look at 2×2 systems in the next section.

You should now have understood

- what a linear system is,
- what a coefficient matrix and an augmented coefficient matrix is,
- its relation with linear systems,
- that a linear system can have different types of solutions.

You should now be able to

- pass easily from a linear $m \times n$ system to its (augmented) coefficient matrix and back,
- ...

1.2 Linear 2×2 systems

Let us come back to the equation from Example 1.1. For convenience, we write now x instead of B and y instead of C. Recall that the system of equations that we are interested in solving is

$$\begin{aligned} &\textcircled{1} \qquad x + y = 60, \\ &\textcircled{2} \qquad 2x + 4y = 200. \end{aligned} \tag{1.5}$$

We want to give a **geometric** meaning to this system of equations. To this end we think of pairs x, y as points (x, y) in the plane. Let us forget about equation $\textcircled{2}$ for a moment and concentrate only on $\textcircled{1}$. Clearly, there are infinitely many solutions. If we choose an arbitrary x, we can always find y such that $\textcircled{1}$ satisfied (just take $y = 60 - x$). Similarly, if we choose any y, then we only have to take $x = 60 - y$ and we obtain a solution of $\textcircled{1}$.

Where in the xy-plane lie *all* solutions of $\textcircled{1}$? Clearly, $\textcircled{1}$ is equivalent to $y = 60 - x$ which we easily identify as the equation of the line L_1 in the xy-plane which passes through $(0, 60)$ and has slope -1. In summary, a pair (x, y) is a solution of $\textcircled{1}$ if and only if it lies on the line L_1.

If we apply the same reasoning to $\textcircled{2}$, we find that a pair (x, y) satisfies $\textcircled{2}$ if and only if (x, y) lies on the line L_2 in the xy-plane given by $y = \frac{1}{4}(200 - 2x)$ (this is the line in the xy-plane passing through $(9, 50)$ with slope $-\frac{1}{2}$).

Now it is clear that a pair (x, y) satisfies both $\textcircled{1}$ and $\textcircled{2}$ if and only if it lies both on L_1 and L_2. So finding the solution of our system (1.5) is the same as finding the intersection of the two lines L_1 and L_2. From elementary geometry we know that there are exactly three possibilities:

(i) L_1 and L_2 are not parallel. Then they intersect in exactly one point.

(ii) L_1 and L_2 are parallel and not equal. Then they do not intersect.

(iii) L_1 and L_2 are parallel and equal. Then $L_1 = L_2$ and they intersect in infinite points (they intersect in every point of $L_1 = L_2$).

In our example we know that the slope of L_1 is -1 and that the slope of L_2 is $-\frac{1}{2}$, so they are not parallel and therefore intersect in exactly one point. Consequently, the system (1.5) has exactly one solution, see Figure 1.1.

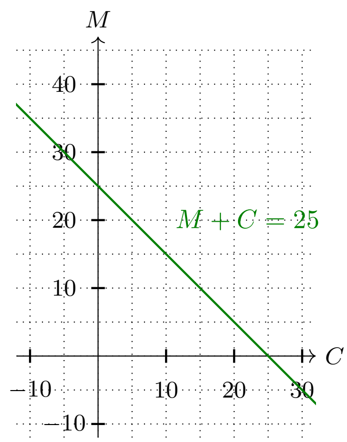


FIGURE 1.1: Graphs of the lines L_1, L_2 which represent the equations from the system (1.5) (see also Example 1.1). Their intersection represents the unique solution of the system.

If we look again at Example 1.6, we see that in Case (a) we have to determine the intersection of the lines

$$L_1 : y = 5 - x, \qquad L_2 : y = \frac{35}{6} - x.$$

Both lines have slope -1 so they are parallel. Since the constant terms in both lines are not equal, they never intersect, showing that the system of equations has no solution, see Figure 1.2.
In Case (b), the two lines that we have to intersect are

$$G_1 : y = 5 - x, \qquad G_2 : y = 5 - x.$$

We see that $G_1 = G_2$, so every point on G_1 (or G_2) is solution of the system and therefore we have infinite solutions, see Figure 1.2.

Important observation. If the solution of the system is unique or not, has nothing to do with the right hand side of the system because this only depends on whether the two lines are parallel or not, and this in turn depends only on the coefficients on the left hand side.

Now let us consider the general case.

One linear equation with two unknowns

The general form of one linear equation with two unknowns is

$$\alpha x + \beta y = \gamma. \tag{1.6}$$

For the set of solutions, there are three possibilities:

(i) *The set of solutions forms a line.* This happens if at least one of the coefficients α or β is different from 0. If $\beta \neq 0$, then set of all solutions is equal to the line $L : y = -\frac{\alpha}{\beta}\, x + \frac{\gamma}{\beta}$ which

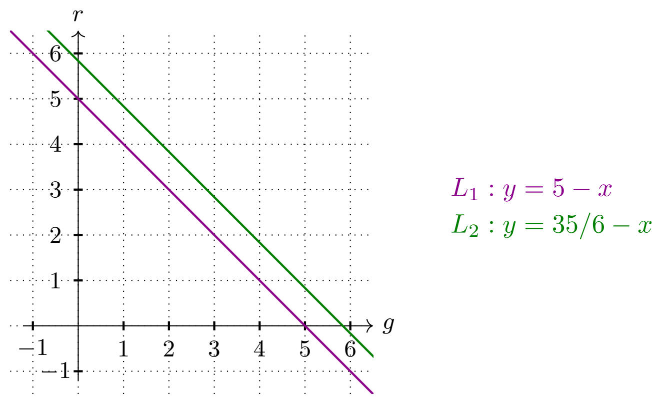


FIGURE 1.2: Example 1.6. Graphs of L_1, L_2.

is a line with slope $-\frac{\alpha}{\gamma}$. If $\beta = 0$ and $\alpha \neq 0$, then the set of solutions of (1.6) is a line parallel to the y-axis passing through $(\frac{\gamma}{\alpha})$.

(ii) *The set of solutions is all of the plane.* This happens if $\alpha = \beta = \gamma = 0$. In this case, clearly every pair (x, y) is a solution of (1.6).

(iii) *The set of solutions is empty.* This happens if $\alpha = \beta = 0$ and $\gamma \neq 0$. In this case, no pair (x, y) can be a solution of (1.6) since the left hand side is always 0.

In the first two cases, (1.6) has infinitely many solutions, in the last case it has no solution.

Two linear equations with two unknowns

The general form of one linear equation with two unknowns is

$$\begin{aligned} &① \quad Ax + By = U \\ &② \quad Cx + Dy = V. \end{aligned} \tag{1.7}$$

We are using the letters A, B, C, D instead of $a_{11}, a_{12}, a_{21}, a_{22}$ in order to make the calculations more readable. If we interprete the system of equations as intersection of two geometrical objects, we already know how the possible solutions will be

(i) *a point* if ① and ② describe two non-parallel lines.

(ii) *a line* if ① and ② describe the same line; or if one of the equations is a plane and the other one is a line.

(iii) *a plane* if both equations describe a plane.

(iv) *the empty set* if the two equations describe parallel but different lines; or if one of the equations has no solution.

In case (i), the system has exactly one solution, in cases (ii) and (iii) the system has infinitely many solutions and in case (iv) the system has no solution.

In summary, we have the following very important observation.

Remark 1.8. The system (1.7) has either exactly one solution or infinitely many solutions or no solution.

It is not possible to have for instance exactly 7 solutions.

qu:1:03

Question 1.3

What is the geometric interpretation of

(i) a system of 3 linear equations for 2 unknowns?

(ii) a system of 2 linear equations for 3 unknowns?

What can be said about the structure of its solutions?

Algebraic proof of Remark 1.8. Now we want to prove the Remark 1.8 algebraically and we want to find a criteria on a, b, c, d which allows us to decide easily how many solutions there are. Let us look at the different cases.

$\boxed{\text{Case 1. } B \neq 0.}$ In this case we can solve ① for y and obtain $y = \frac{1}{B}(U - Ax)$. In ② this gives $Cx + \frac{D}{B}(U - Ax) = V$. If we put all terms with x on one side and all other terms on the other side, we obtain

$$②' \quad (AD - BC)x = DU - BV$$

(i) If $\boxed{AD - BC \neq 0}$ then there is at most one solution, namely $x = \frac{DU-BV}{AD-BC}$ and consequently $y = \frac{1}{B}(U - Ax) = \frac{AV-CU}{AD-BC}$. Inserting these expressions for x and y in our system of equations, we see that they indeed solve the system (1.7), so that we have exactly one solution.

(ii) If $\boxed{AD - BC = 0}$ then equation ②' reduces to $0 = DU - BV$. This equation has either no solution (if $DU - BV \neq 0$) or infinite solutions (if $DU - BV = 0$). Since ① has infinite solutions, it follows that the system (1.7) has either no solution or infinite solutions.

$\boxed{\text{Case 2. } D \neq 0.}$ This case is analogous to Case 1. In this case we can solve ② for y and obtain $y = \frac{1}{D}(V - Cx)$. In ② this gives $Ax + \frac{B}{D}(V - Cx) = U$. If we put all terms with x on one side and all other terms on the other side, we obtain

$$②' \quad (AD - BC)x = DU - BV$$

We have the same subcases as before:

(i) If $\boxed{AD - BC \neq 0}$ then there is exactly one solution, namely $x = \frac{DU-BV}{AD-BC}$ and consequently $y = \frac{1}{B}(U - Ax) = \frac{AV-CU}{AD-BC}$.

(ii) If $\boxed{AD - BC = 0}$ then equation 2' reduces to $0 = DU - BV$. This equation has either no solution (if $DU - BV \neq 0$) or infinite solutions (if $DU - BV = 0$). Since 2 has infinite solutions, it follows that the system (1.7) has either no solution or infinite solutions.

$\boxed{\text{Case 3. } B = 0 \text{ and } D = 0.}$ Observe that in this case $\boxed{AD - BC = 0}$. In this case the system (1.7) reduces to

$$Ax = U, \quad Cx = V. \tag{1.8}$$

We see that the system no longer depends on y. So, if the system (1.8) has at least one solution, then we automatically have infinite solutions since we can choose y freely. If the system (1.8) has no solution, then the original system (1.7) cannot have a solution either.

Note that there are no other cases for the coefficients than these three cases. □

In summary, we proved the following theorem.

Theorem 1.9. *Let us consider the linear system*

$$\begin{aligned} &1 \quad Ax + By = U \\ &2 \quad Cx + Dy = V. \end{aligned} \tag{1.9}$$

(i) *The system* (1.9) *has exactly one solution if and only if* $\boxed{AD - BC \neq 0}$. *In this case, the solution is*

$$x = \frac{DU - BV}{AD - BC}, \quad y = \frac{AV - CU}{AD - BC}. \tag{1.10}$$

(ii) *The system* (1.9) *has no solution or infinitely many solutions if and only if* $\boxed{AD - BC \neq 0}$.

Definition 1.10. The number $\boxed{d = AD - BC}$ is called the *determinant* of the system (1.9).

In Chapter **??** we will generalise this concept to $n \times n$ systems for $n \geq 3$.

Remark 1.11. Let us see how this connects to our geometric interpretation of the system of equations. Assume that $B \neq 0$ and $D \neq 0$. Then we can solve 1 and 2 for y to obtain equations for a pair of lines

$$L_1: \quad y = -\frac{A}{B}x + \frac{1}{B}U, \qquad L_2: \quad y = -\frac{C}{D}x + \frac{1}{D}V.$$

The two lines intersect in exactly one point if and only if they have different slopes, i.e., if $-\frac{A}{B} \neq -\frac{C}{D}$. After multiplication by $-BD$ we see that this is the same as $AD \neq BC$, or $AD - BC \neq 0$.
On the other hand, the lines are parallel (hence they are either equal or they have no intersection) if $-\frac{A}{B} \neq -\frac{C}{D}$. This is the case if and only if $AD = BC$, or in other word, if $AD - BC = 0$.

qu:1:04

Question 1.4

Consider the cases when $B = 0$ or $D = 0$ and make the connection between Theorem 1.9 and the geometric interpretation of the system of equations.

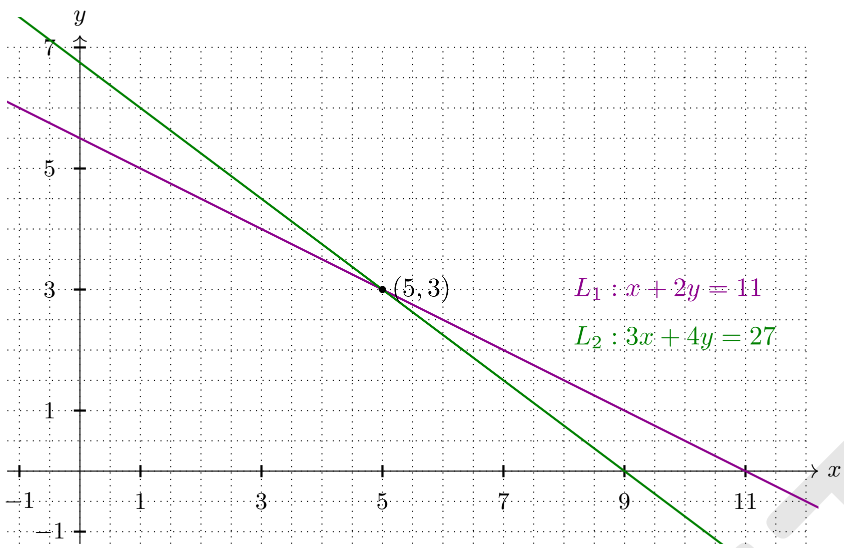


FIGURE 1.3: Example 1.12(a). Graphs of L_1, L_2 and their intersection $(5, 3)$.

Let us consider some more examples.

Examples 1.12. (a)

$$\begin{aligned} ① \quad x + 2y &= 11 \\ ② \quad 3x + 4y &= 27. \end{aligned}$$

Clearly, the determinant is $d = 4 - 6 = -2 \neq 0$. So the system has *exactly one solution*.

We can check this easily: The first equation gives $x = 11 - 2y$. Inserting this into the second equations leads to

$$3(11 - 2y) + 4y = 27 \quad \Longrightarrow \quad -2y = -6 \quad \Longrightarrow \quad y = 3 \quad \Longrightarrow \quad x = 11 - 2 \cdot 3 = 5.$$

So the solution is $x = 5, y = 3$. (If we did not have Theorem 1.9, we would have to check that this is not only a candidate for a solution, but indeed is one.)

Check that the formula (1.10) is satisfied.

(b)

$$\begin{aligned} ① \quad x + 2y &= 1 \\ ② \quad 2x + 4y &= 5. \end{aligned}$$

Here, the determinant is $d = 4 - 4 = 0$, so we expect *either no solution or infinite solutions*. The first equations gives $x = 1-2y$. Inserting into the second equations gives $2(1-2y)+4y = 5$. We see that the terms with y cancel and we obtain $2 = 5$ which is a contradiction. Therefore, the system of equations has *no solution*.

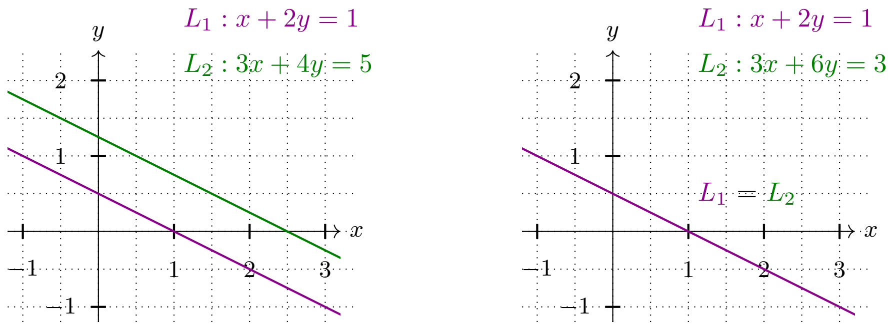


FIGURE 1.4: Picture on the left: The lines L_1, L_2 from Example 1.12(b) are parallel and do not intersect. Therefore the linear system has no solution.
Picture on the right: The lines L_1, L_2 from Example 1.12(c) are equal. Therefore the linear system has infinitely many solutions.

(c)
$$\begin{aligned} &① \quad x + 2y = 1 \\ &② \quad 3x + 6y = 3. \end{aligned}$$

The determinant is $d = 6 - 6 = 0$, so again we expect *either no solution or infinite solutions*. The first equations gives $x = 1-2y$. Inserting into the second equations gives $3(1-2y)+6y = 3$. We see that the terms with y cancel and we obtain $3 = 3$ which is true. Therefore, the system of equations has *infinite solutions* given by $x = 1 - 2y$.

Remark. This was somewhat clear since we can obtain the second equation from the first one by multiplying both sides by 3 which shows that both equations carry the same information and we loose nothing if we simply forget about one of them.

Exercise 1.13. Find all $k \in \mathbb{R}$ such that the system

$$\begin{aligned} &① \quad kx + (15/2 - k)y = 1 \\ &② \quad 4x + \qquad 2ky = 3 \end{aligned}$$

has exactly one solution.

Solution. We only need to calculate the determinant and find all k such that it is different from zero. So let us start by calculating

$$d = k \cdot 2k - (15/2 - k) \cdot 4 = 2k^2 + 4k - 30 = 2(k^2 + 2k - 15) = 2[(k + 1)^2 - 16].$$

So we see that there are exactly two values for k where $d = 0$, namely $k = -1 \pm 4$, that is $k_1 = 3$, $k_2 = -5$. For all other k, we have that $d \neq 0$.
So the answer is: The system has exactly one solution if and only if $k \in \mathbb{R} \setminus \{-5, 3\}$. ◇

Remark 1.14. 1. Note that the answer does not depend on the right hand side of the system of the equation. Only the coefficients on the left hand side determine if there is exactly one solution or not.

2. If we wanted to, we could also calculate the solution x, y in the case $k \in \mathbb{R} \setminus \{-5, 3\}$. We could do it by hand or use (1.10). Either way, we find

$$x = \frac{1}{d}[2k - 3(15/2 - k)] = \frac{5k - 45/2}{2k^2 + 4k - 30}, \qquad y = \frac{1}{d}[6k - 4] = \frac{6k - 4}{2k^2 + 4k - 30}.$$

   Note that the denominators would become 0 if $k = -5$ or $k = 3$.

3. What happens if $k = -5$ or $k = 3$? In both cases, $d = 0$, so we will either have no solution or infinite solutions.

   If $k = -5$, then the system becomes $-5x + 25/2y = 1, \quad 4x - 10y = 3$.
   Multiplying the first equation by $-4/5$ and not changing the second equation, we obtain

$$4x - 10 = -\frac{4}{5}, \qquad 4x - 10y = 3$$

   which clearly cannot be satisfied simultaneously.

   If $k = 3$, then the system becomes $3x - 9/2y = 1, \quad 4x + 6y = 3$.
   Multiplying the first equation by $4/3$ and not changing the second equation, we obtain

$$4x - 6y = \frac{4}{3}, \qquad 4x - 6y = 3$$

   which clearly cannot be satisfied simultaneously.

You should have understood

- the geometric interpretation of a linear $m \times 2$ system and how this helps to understand the qualitative structure of solutions,
- how the determinant helps to decide whether a linear 2×2 system has a unique solution or not,
- that it depends only on the coefficients of the system if its solution is unique; it does *not* depend on the right side of the equation (the actual values of the solutions of course do depend on the right side of the equation),
- ...

You should now be able to

- pass easily from a linear $m \times 2$ system to its geometric interpretation and back,
- calculate the determinant of a linear 2×2 system,
- determine if a linear 2×2 system has a unique, no or infinitely many solutions and calculate them,
- give criteria for existence/uniqueness of solutions,
- ...

1.3 Exercises

1. Para los siguientes sistemas, si es posible,

   (i) escribe la matriz de coeficientes y la matriz aumentada,

   (ii) calcule el determinante y concluya sobre existencia y unicidad de soluciones,

   (iii) encuentre todas las soluciones,

   (iv) haga un dibujo.

   (a) $-3x + 2y = 18, \quad x + 2y = 2.$

   (b) $2x + 8y = 6, \quad 3x + 12y = 2.$

   (c) $2x - 4y = 6, \quad -x + 2y = -1.$

   (d) $3x - 2y = -1, \quad x + 3y = 18, \quad 2x - 5y = -8.$

   (e) $x - y = 5, \quad -3x + 2y = 3, \quad 2x + 3y = 14.$

2. Encuentre todas las soluciones de los siguientes sistemas y visualice las ecuaciones y las soluciones en el plano.

   (a) $3x + 5y = 7, \quad -9x - 15y = 10,$

   (b) $2x + 5y = 10, \quad x + 2y + 3 = 0,$

   (c) $2x + y = 4, \quad 3x - 2y = -1, \quad 5x + 3y = 7,$

   (d) $x + 5y = 3, \quad -3x + 2y = 8, \quad 2x + 3y = -1.$

3. (a) Encuentre todos los números k tal que es siguiente sistema de ecuaciones tiene exactamente una solución y calcule esta solución. ¿Qué pasa para los otros k?

   $$kx + 5y = 0, \qquad 3x + (2 + k)y = 0.$$

   (b) Haga los mismo para el sistema

   $$kx + 5y = 5, \qquad 3x + (2 + k)y = -3.$$

4. (a) Encuentre todos los números k tal que es siguiente sistema de ecuaciones tiene exactamente una solución y calcule esta solución. ¿Qué pasa para los otros k?

   $$kx + 2y = 0, \qquad 2x - (3 + k)y = 0.$$

   (b) Haga los mismo para el sistema

   $$kx + 2y = 6, \qquad 2x - (3 + k)y = -3.$$

5. (a) Encuentre un polinomio P de grado 3 con

   $$P(1) = 2, \quad P(-1) = 6, \quad P'(1) = 8, \quad P(0) + 4P'(0) = 0.$$

(b) ¿Existe un polinomio de grado 2 que satisface lo de arriba? De ser así, ¿cu'antos hay? Justifique su respuesta.

(c) ¿Existe un polinomio de grado 4 que satisface lo de arriba De ser así, ¿cu'antos hay? Justifique su respuesta.

6. En una bodega hay soluciones de un cierto químico con concentraciones de 1% y de 13%. ¿Cuántos mililitros de cada una de las soluciónes disponibles se requieren para obtener 500 ml de una solucón de este químico con contentración de 5%?

7. Considere la ecuación

$$3x + 4y = 5. \tag{1.11}$$

(a) ¿Existe otra ecuación lineal tal que la solución del sistema de (1.11) y la nueva ecuación es $(3, -1)$? Encuentre tal ecuación o diga por qué no existe.

(b) ¿Existen otras dos ecuaciones lineales tal que la solución del sistema de (**??**) y las nuevas ecuaciones es $(3, -1)$? Encuentre tales ecuaciones o diga por qué no existen.

(c) ¿Existe otra ecuación lineal tal que la solución del sistema de (1.11) y la nueva ecuación es $(2, -3)$? Encuentre tal ecuación o diga por qué no existe.

(d) ¿Existen otras dos ecuaciones lineales tal que la solución del sistema de (1.11) y las nuevas ecuaciones es $(2, -3)$? Encuentre tales ecuaciones o diga por qué no existen.

(e) Encuentre otra ecuación lineal tal que el sistema de (1.11) y la nueva ecuación no tenga solución.

(f) Encuentre otra ecuación lineal tal que el sistema de (1.11) y la nueva ecuación tenga infinitas soluciones.

Chapter 2

$\mathbb{R}^2$ and $\mathbb{R}^3$

2.1 Vectors in $\mathbb{R}^2$

Recall that the xy-plane is the set of all pairs (x, y) with $x, y \in \mathbb{R}$. We will denote it by $\mathbb{R}^2$.
Maybe you already encountered vectors in a physics lecture. For instance velocities and forces are described by vectors. The velocity of a particle says how fast and in which direction the particle moves. Usually, velocities are represented by an arrow which points in the direction in which the particle moves and whose length is proportional to the magnitude of the velocity.
A force has strength and a direction so it is represented by an arrow which point in the direction in which it acts and with length proportional to its strength.
Observe that it is not important where in the space $\mathbb{R}^2$ or $\mathbb{R}^3$ we put the arrow. As long it points in the same direction and has the same length, it is considered the same vector. We call two arrows *equivalent* if they have the same direction and the same length. A *vector* is the set of all arrows which are equivalent to a given arrow. Each specific arrow in this set is called a *representation* of the vector. A special representation is the arrow that starts in the origin $(0, 0)$. Vectors are usually denoted by a small letter with an arrow on top, for example $\vec{v}$.

Given two points P, Q in the xy-plane, we write $\overrightarrow{PQ}$ for the vector which is represented by the arrow that starts in P and ends in Q. For example, let $P(1, 1)$ and $Q(3, 4)$ be points in the xy-plane. Then the arrow from P to Q is $\overrightarrow{PQ} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.
We can identify a point $P(p_1, p_2)$ in the xy-plane with the vector starting in $(0, 0)$ and ending in P. We denote this vector by $\overrightarrow{0P}$ or $\begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$ or sometimes by $(p_1,\, p_2)^t$ in order to save space (the subscript ^{t} stands for "transposed"). p_1 is called the *x-coordinate* or the *x-component* of $\vec{v}$ and p_2 is called the *y-coordinate* or the *y-component* of $\vec{v}$.

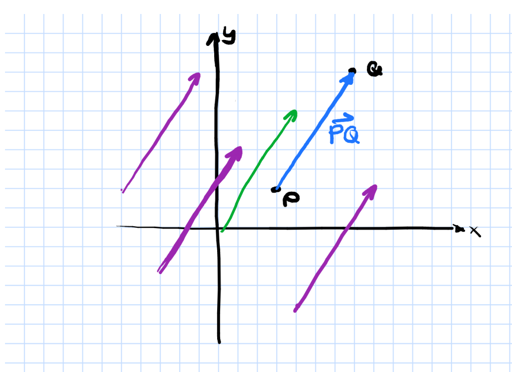


FIGURE 2.1: The vector $\vec{v}$ and several of its representations. The green arrow is the special representation whose initial point is in the origin.

On the other hand, every vector $\begin{pmatrix} a \\ b \end{pmatrix}$ describes a unique point in the xy-plane, namely the tip of the arrow which represents the given vector and starts in the origin. Clearly its coordinates are (a, b). Therefore we can identify the set of all vectors in $\mathbb{R}^2$ with $\mathbb{R}^2$ itself.

Observe that the slope of the arrow $\vec{v} = (a, b)$ is $\frac{b}{a}$ if $a \neq 0$. If $a = 0$, then we obtain a vector which is parallel to the y-axis.

For example, the vector $\vec{v} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$, can be represented as an arrow whose initial point is in the origin and its tip is at the point $(2, 5)$. If we put its initial point anywhere else, then we find the tip by moving 2 units to the right (parallel to the x-axis) and 5 units up (parallel to the y-axis).

A very special vector is the zero vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Is is usually denoted by $\vec{0}$.

We call numbers in $\mathbb{R}$ *scalars* in order to distinguish them from vectors.

Algebra with vectors

If we think of a force and we double its strength then the corresponding vector should be twice as long. If we *multiply* the force by 5, then the length of the corresponding vector should be 5 times as long, that is, if for instance a force $\vec{F} = (3, 4)$ is given, then $5\vec{F}$ should be $(5 \cdot 3,\ 5 \cdot 4) = (15, 20)$.

In general, if a vector $\vec{v} = (a, b)$ and a scalar c are given, then $c\vec{v} = (ca, cb)$. Note that the resulting vector is always parallel to the original one. If $c > 0$, then the resulting vector points in the same direction as the original one, if $c < 0$, then it points in the opposite direction, see Figure 2.2.

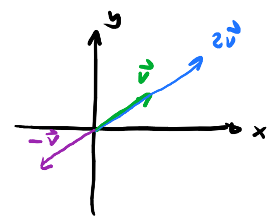


FIGURE 2.2: Multiplication of a vector by a scalar.

> Given two points $P(p_1, p_2)$, $Q(q_1, q_2)$ in the xy-plane. Convince yourself that $\overrightarrow{PQ} = -\overrightarrow{QP}$.

How should we *sum* two vectors? Again, let us think of forces. Assume we have two forces $\vec{F}_1$ and $\vec{F}_2$ both acting on the same particle. Then we get the resulting force by drawing the arrow representing $\vec{F}_1$ and at its tip put the initial point of the arrow representing $\vec{F}_2$. The total force is then represented by the arrow starting in the initial point of $\vec{F}_1$ and ending in the tip of $\vec{F}_2$.

> Convince yourself that we obtain the same result if we start with $\vec{F}_2$ and put the initial point of $\vec{F}_1$ at the tip of $\vec{F}_2$.

We could also think of the sum of velocities. For example, if the have a train with velocity $\vec{v}_t$ and on the train a passenger is moving with relative velocity $\vec{v}_p$, then the total velocity is the vector sum of the two.

Now assume that $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ and $\vec{w} = \begin{pmatrix} p \\ q \end{pmatrix}$. Algebraically, we obtain the components of their sum by summing the components: $\vec{v} + \vec{w} = \begin{pmatrix} a+p \\ b+q \end{pmatrix}$, see Figure 2.3.

When you do vector sums, you should always think in triangles (or polygons if you sum more than two vectors).

Given two points $P(p_1, p_2)$, $Q(q_1, q_2)$ in the xy-plane. Convince yourself that $\overrightarrow{0P} + \overrightarrow{PQ} = \overrightarrow{0Q}$ and consequently $\overrightarrow{PQ} = \overrightarrow{0Q} - \overrightarrow{0P}$.
How could you write $\overrightarrow{QP}$ in terms of $\overrightarrow{0P}$ and $\overrightarrow{0Q}$? What is its relation with $\overrightarrow{PQ}$?

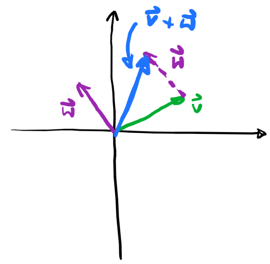


FIGURE 2.3: Sum of two vectors.

Our discussion of how the product of a vector and a scalar and how the sum of two vectors should be, leads us to the following formal definition.

Definition 2.1. Let $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$, $\vec{w} = \begin{pmatrix} p \\ q \end{pmatrix} \in \mathbb{R}^2$, $c \in \mathbb{R}$. Then:

Vector sum: $$\vec{v} + \vec{w} = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} a+p \\ b+q \end{pmatrix},$$

Product with a scalar: $$c\vec{v} = c\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ca \\ cb \end{pmatrix}.$$

It is easy to see that the vector sum satisfies what one expects from a sum: $(\vec u + \vec v) + \vec w = \vec u + (\vec v + \vec w)$ (associativity) and $\vec v + \vec w = \vec w + \vec v$ (commutativity). Moreover, we have the distributivity laws $(a+b)\vec v = a\vec v + b\vec v$ and $a(\vec v + \vec w) = a\vec v + a\vec w$. Let verify for example associativity. To this end, let $\vec u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $\vec v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, $\vec w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$. Then

$$\begin{aligned}
(\vec u + \vec v) + \vec w &= \left[\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right] + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} (u_1 + v_1) + w_1 \\ (u_2 + v_2) + w_2 \end{pmatrix} \\
&= \begin{pmatrix} u_1 + (v_1 + w_1) \\ u_2 + (v_2 + w_2) \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} (v_1 + w_1) \\ (v_2 + w_2) \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \left[\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right] \\
&= \vec u + (\vec v + \vec w).
\end{aligned}$$

In the same fashion, verify that commutativity and distributivity holds.

We can take these properties and define an *abstract vector space*. We shall call a set of things, called *vectors*, with a "well-behaved" sum of its elements and a "well-behaved" product of its elements with scalars a *vector space*. The precise definition is the following.

Vector Space Axioms. Let V be a set. Then V is called an $\mathbb{R}$-vector space and its elements are called *vectors* if

(a) **Associativity:** $(u+v)+w = u+(v+w)$ for every $u, v, w \in V$.

(b) **Commutativity:** $v+w = w+v$ for every $u, v \in V$.

(c) **Identity element of addition:** There exists an element $\mathbb{O} \in V$, called the *additive identity* such that for every $v \in V$, we have $\mathbb{O} + v = v + \mathbb{O} = v$.

(d) **Inverse element:** For all $v \in V$, we have an inverse element v' such that $v + v' = \mathbb{O}$.

(e) **Identity element of multiplication by scalar:** For every $v \in V$, we have that $1v = v$.

(f) **Compatibility:** For every $v \in V$ and $\lambda, \mu \in \mathbb{R}$, we have that $(\lambda\mu)v = \lambda(\mu v)$.

(g) **Distributivity laws:** For all $v, w \in V$ and $\lambda, \mu \in \mathbb{R}$, we have

$$(\lambda + \mu)v = \lambda v + \mu v \quad \text{and} \quad \lambda(v + w) = \lambda v + \lambda w.$$

These axioms are fundamental for linear algebra and we will come back to them in Chapter **??**.

Check that $\mathbb{R}^2$ is a vector space, that its additive identity is $\mathbb{O} = \vec{0}$ and that for every vector $\vec{v} \in \mathbb{R}^2$, its additive inverse is $-\vec{v}$.

It is important to note that there are vector spaces that do not look like $\mathbb{R}^2$ and that we cannot alwauays write vectors as columns. For instance, the set of all polynomials form a vector space (we can add them, the sum is additive and commutative; the additive identity is the zero polinomial and for every polynomial p, its additive inverse is the polynomial $-p$; we can multiply polyonmials with scalars and obtain another polynomial, etc.). The vectors in this case are polynomials and it does not make sense to speak about its components or coordintates. (We will however learn how to represent certain subspaces of the space of polynomials as subspaces of some $\mathbb{R}^n$ in Chapter **??**.)

After this brief excursion about abstract vector spaces, let us return to our familiar $\mathbb{R}^2$. We know that it can be identified with the xy-plane. This means that $\mathbb{R}^2$ has more structure than only being a vector space. For example, we can measure angles and lenghts. Observe that these concepts do *not* appear in the definition of a vector space. They are something in addition to the the vector space properties.
Let us now look at some more geometric properties of vectors in $\mathbb{R}^2$. Clearly a vector is known if we know its length and its angle with the x-axis. From the Pythagoras theorem it is clear that the length of a vector $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ is $\sqrt{a^2 + b^2}$.

Definition 2.2 (Norm of a vector in $\mathbb{R}^2$). The *length* of $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ is denoted by $\|\vec{v}\|$. It is given by

$$\|\vec{v}\| = \sqrt{a^2 + b^2}\ .$$

Other names for the length of $\vec{v}$ are *magnitude of* $\vec{v}$ or *norm of* $\vec{v}$.

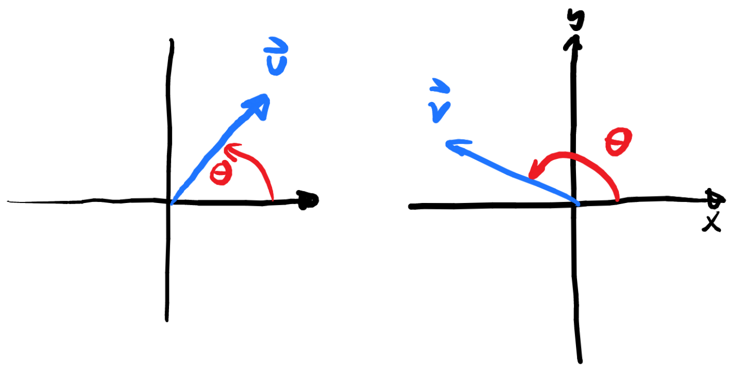


FIGURE 2.4: Length and angle of a vector.

As already mentioned earlier, the slope of vector $\vec{v}$ is $\frac{b}{a}$ if $a \neq 0$. If φ is the angle of the vector $\vec{v}$ with the x-axis then $\tan\varphi = \frac{b}{a}$ if $a \neq 0$. If $a = 0$, then $\varphi = 0$ or $\varphi = \pi$. Recall that the range of arctan is $(-\pi/2,\ \pi/2)$, so we cannot simply take arctan of the fraction $\frac{a}{b}$ in order to obtain φ. Observe that $\arctan\frac{b}{a} = \arctan\frac{-b}{-a}$, however the vectors $\begin{pmatrix} a \\ b \end{pmatrix}$ and $\begin{pmatrix} -a \\ -b \end{pmatrix}$ are parallel but point in opposite directions, so they do *not* have the same angle with the x-axis. From elementary geometry, we find

$$\tan\varphi = \frac{b}{a} \text{ if } \ a \neq 0 \qquad \text{and} \qquad \varphi = \begin{cases} \arctan\frac{b}{a} & \text{if } a > 0, \\ \pi - \arctan\frac{b}{a} & \text{if } a < 0, \\ \pi/2 & \text{if } a = 0,\ b > 0, \\ -\pi/2 & \text{if } a = 0,\ b < 0. \end{cases}$$

Note that this formula gives angles with values $[-\pi/2,\ 3\pi/2)$.

Proposition 2.3 (Properties of the norm). *Let $\lambda \in \mathbb{R}$ and $\vec{v}$, $\vec{w} \in \mathbb{R}^2$. Then the following is true:*

(i) $\|\lambda\vec{v}\| = |\lambda|\|\vec{v}\|$,

(ii) $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$,

(iii) $\|\vec{v}\| = 0$ *if and only if* $\vec{v} = \vec{0}$.

Proof. Let $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$, $\vec{w} = \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$.

(i)
$$\begin{aligned} \|\lambda\vec{v}\| &= \left\|\lambda\begin{pmatrix} a \\ b \end{pmatrix}\right\| = \big\|(\lambda a,\ \lambda b)\big\| = \sqrt{(\lambda a)^2 + (\lambda b)^2} = \sqrt{\lambda^2(a^2+b^2)} = |\lambda|\sqrt{a^2+b^2} \\ &= |\lambda|\|\vec{v}\|. \end{aligned}$$

(ii) This will be shown later in **??**.

(iii) Since $\|\vec{v}\| = \sqrt{a^2 + b^2}$ it follows that $\|\vec{v}\| = 0$ if and only if $a = 0$ and $b = 0$. This is the case if and only if $\vec{v} = \vec{0}$. □

Definition 2.4. A vector $\vec{v} \in \mathbb{R}^2$ is called a *unit vector* if $\|\vec{v}\| = 1$.

Note that every vector $\vec{v} \neq \vec{0}$ defines a unit vector pointing in the same direction as itself by $\|\vec{v}\|^{-1}\vec{v}$.

Remark 2.5. (i) The tip of every unit vector lies on the unit circle, and every vector whose initial point is the origin and whose tip lies on the unit circle is a unit vector.

(ii) Every unit vector is of the from $\begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$ where φ is its angle with the positive x-axis.

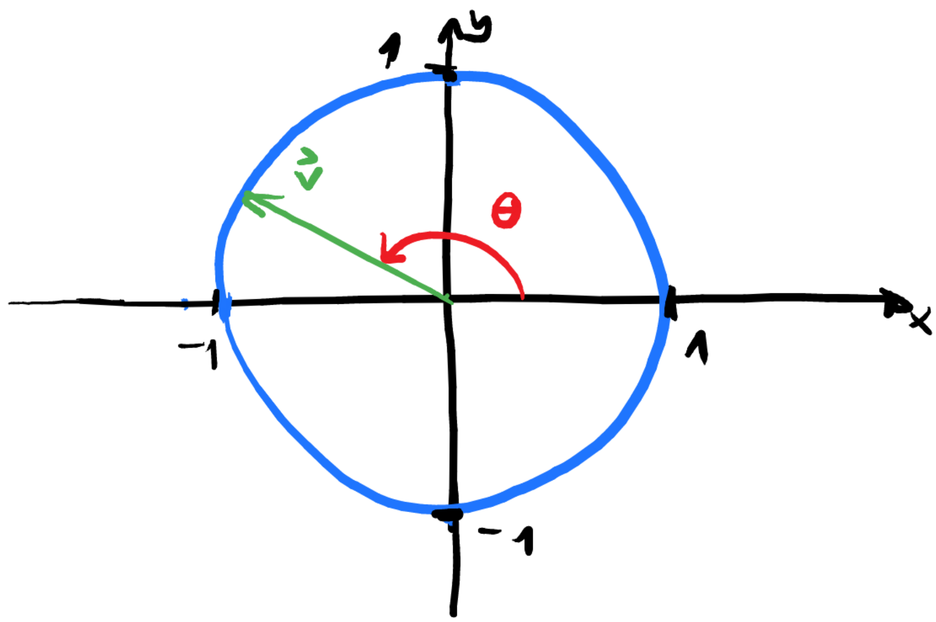


FIGURE 2.5: Unit vectors.

Finally, we define two very special unit vectors:

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Clearly, $\vec{e}_1$ is parallel to the x-axis, $\vec{e}_2$ is parallel to the y-axis and $\|\vec{e}_1\| = \|\vec{e}_2\| = 1$.

Remark 2.6. Every vector $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ can be written as

$$\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ b \end{pmatrix} = a\vec{e}_1 + b\vec{e}_2.$$

Remark 2.7. Another notation for $\vec{e}_1$ and $\vec{e}_2$ is $\hat{\imath}$ and $\hat{\jmath}$.

You should have understood

- the concept of an abstract vector space and vectors,
- the vector space $\mathbb{R}^2$ and how to calculate with vectors in $\mathbb{R}^2$,
- the difference between a point $P(a, b)$ in $\mathbb{R}^2$ and a vector $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ in $\mathbb{R}^2$,
- geometric concepts (angles, length of a vector),
- ...

You should now be able to

- perform algebraic operations in the vector space $\mathbb{R}^2$ and visualize them in the plane,
- calculate lengths and angles,
- calculate unit vectors, scale vectors,
- perform simple abstract proofs (e.g., prove that $\mathbb{R}^2$ is a vector space).
- ...

2.2 Inner product in $\mathbb{R}^2$

In this section we will explore further geometric properties of $\mathbb{R}^2$ and we will introduce the so-called inner product. Many of thess properties carry over almost literally to $\mathbb{R}^3$ and more generally, to $\mathbb{R}^n$. Let us start with a definition.

Definition 2.8 (Inner product). Let $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \vec{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ be vectors in $\mathbb{R}^2$. The *inner product* of $\vec{v}$ and $\vec{w}$ is

$$\langle \vec{v}\,, \vec{w} \rangle := v_1 w_1 + v_2 w_2.$$

The inner product is also called *scalar product* or *dot product* and it can also be denoted by $\vec{v} \cdot \vec{w}$.

We usually prefer the notation $\langle \vec{v}\,, \vec{w} \rangle$ since this notation is used frequently in physics and extends naturally to abstract vector spaces with an inner product. Moreover, the the notation with the dot seems to suggest that the dot product behaves like a usual product, but it does not, see Remark 2.11.

Before we give properties of the inner product and explore what it is good for, we first calculate a few examples to familiarize ourselves with it.

Examples 2.9.

(i) $\left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \end{pmatrix} \right\rangle = 2 \cdot (-1) + 3 \cdot 5 = -2 + 15 = 13.$

(ii) $\left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\rangle = 2^2 + 3^2 = 4 + 9 = 13.$ Observe that this is equal to $\left\| \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\|^2$.

(iii) $\left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle = 2, \quad \left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = 3.$

(iv) $\left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \end{pmatrix} \right\rangle = 0.$

Proposition 2.10 (Properties of the inner product). *Let $\vec{u}$, $vecv, \vec{w} \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$. Then the following holds.*

(i) $\langle \vec{v}, \vec{v} \rangle = \|\vec{v}\|^2$. *In dot notation:* $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$.

(ii) $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$. *In dot notation:* $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$.

(iii) $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$. *In dot notation:* $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$.

(iv) $\langle \lambda\vec{u}, \vec{v} \rangle = \lambda \langle \vec{u}, \vec{v} \rangle$. *In dot notation:* $(\lambda\vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v})$.

Proof. Let $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$.

(i) $\langle \vec{v}, \vec{v} \rangle = v_1^1 + v_2^2 = \|\vec{v}\|^2$.

(ii) $\langle \vec{u}, \vec{v} \rangle = u_1 v_1 + u_2 v_2 = v_1 u_1 + v_2 u_2 = \langle \vec{v}, \vec{u} \rangle$.

(iii)
$$\begin{aligned} \langle \vec{u}, \vec{v} + \vec{w} \rangle &= \left\langle \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix} \right\rangle \\ &= u_1(v_1 + w_1) + u_2(v_2 + w_2) = u_1 v_1 + u_2 v_2 + u_1 w_1 + u_2 w_2 \\ &= \left\langle \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right\rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle. \end{aligned}$$

(iv) $\langle \lambda\vec{u}, \vec{v} \rangle = \left\langle \begin{pmatrix} \lambda u_1 \\ \lambda u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\rangle = \lambda u_1 v_1 + \lambda u_2 v_2 = \lambda(u_1 v_1 + u_2 v_2) = \lambda \langle \vec{u}, \vec{v} \rangle.$ $\square$

Remark 2.11. Observe that the proposition shows that the inner product is commutative and distributive, so it has some properties of the "usual product" that we are used to from the product in $\mathbb{R}$ or $\mathbb{C}$, but there are some properties that show that the inner product is **not** a product.

(a) The inner products takes two vectors and gives back a number, so it gives back an object that is **not** of the same type as the two things we put in.

(b) In Example 2.9(iv) we saw that it may happen that $\vec{v} \neq \vec{0}$ and $\vec{w} \neq \vec{0}$ but still $\langle \vec{v}\,, \vec{w} \rangle = 0$ which is impossible for a "decent" product.

(c) Given a vector $\vec{v} \neq 0$ and a number $c \in \mathbb{R}$, there are many solutions of the equation $\langle \vec{v}\,, \vec{x} \rangle = c$ for the vector $\vec{x}$, in stark contrast to the usual product in $\mathbb{R}$ or $\mathbb{C}$. As an example, look at Example 2.9(i) and (ii). Therefore it makes no sense to write something like $\vec{v}^{-1}$.

(d) There is no such thing as a neutral element for scalar multiplication.

Now let us see what the inner product is good for. We will see that the inner product between two vectors is connected to the angle between them and it will help us to define orthogonal projections of one vector onto another.

Let us start with a definition.

Definition 2.12. Let $\vec{v}, \vec{w}$ be vectors in $\mathbb{R}^2$. The *angle between $\vec{v}$ and $\vec{w}$* is the smallest nonnegative angle between them, see Figure 2.6. It is denoted by $\sphericalangle(\vec{v}, \vec{w})$.

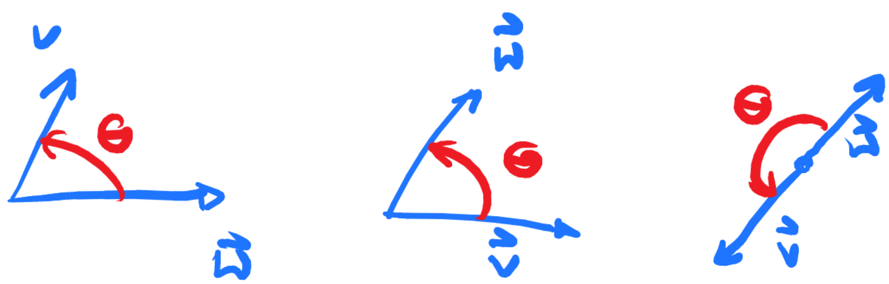


FIGURE 2.6: Angle between two vectors.

The following properties of the angle are easy to see.

Proposition 2.13. (i) $\sphericalangle(\vec{v}, \vec{w}) \in [0, \pi]$ *and* $\sphericalangle(\vec{v}, \vec{w}) = \sphericalangle(\vec{w}, \vec{v})$.

(ii) *If* $\lambda > 0$, *then* $\sphericalangle(\lambda\vec{v}, \vec{w}) = \sphericalangle(\vec{v}, \vec{w})$.

(iii) *If* $\lambda < 0$, *then* $\sphericalangle(\lambda\vec{v}, \vec{w}) = \pi - \sphericalangle(\vec{v}, \vec{w})$.

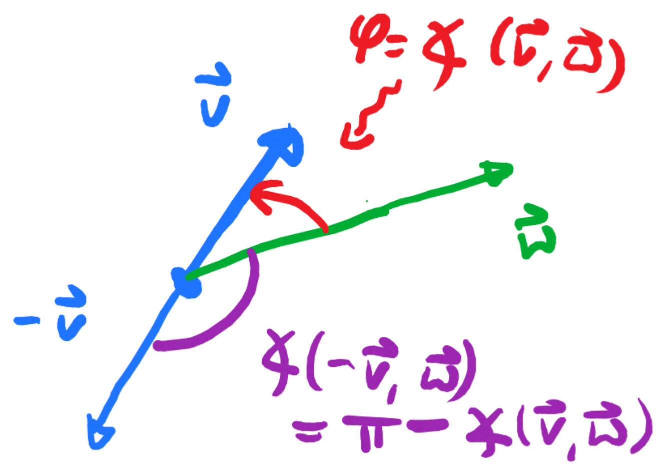


FIGURE 2.7: Angle between vectors $\vec{v}$ and $\vec{w}$.

Definition 2.14. (a) Two non-zero vectors $\vec{v}$ and $\vec{w}$ are called *parallel* if $\sphericalangle(\vec{v}, \vec{w}) = 0$ or π. In this case we use the notation $\vec{v} \parallel \vec{w}$. In addition, the vector $\vec{0}$ is parallel to every vector.

(b) Two non-zero vectors $\vec{v}$ and $\vec{w}$ are called *orthogonal* (or *perpendicular*) if $\sphericalangle(\vec{v}, \vec{w}) = \pi/2$. In this case we use the notation $\vec{v} \perp \vec{w}$. In addition, the vector $\vec{0}$ is perpendicular to every vector.

The following properties should be known from geometry. We will prove them after Theorem 2.18.

Proposition 2.15. *Let $\vec{v}, \vec{w}$ be vectors in $\mathbb{R}^2$. Then:*

(i) *If $\vec{v} \parallel \vec{w}$ and $\vec{v} \neq \vec{0}$, then there exists $\lambda \in \mathbb{R}$ such that $\vec{w} = \lambda \vec{v}$.*

(ii) *If $\vec{v} \parallel \vec{w}$ and $\lambda, \mu \in \mathbb{R}$, then also $\lambda \vec{v} \parallel \mu \vec{w}$.*

(iii) *If $\vec{v} \perp \vec{w}$ and $\lambda, \mu \in \mathbb{R}$, then also $\lambda \vec{v} \perp \mu \vec{w}$.*

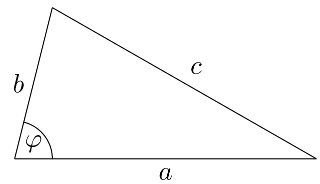
Remark 2.16. (i) Observe that (i) is wrong if we do not assume that $\vec{v} \neq \vec{0}$ because if $\vec{v} = \vec{0}$, then it is parallel to every vector $\vec{w}$ in $\mathbb{R}^2$, but there is no $\lambda \in \mathbb{R}$ such that $\lambda \vec{v}$ could ever become different from $\vec{0}$.

(ii) Observe that the reverse direction in (ii) is true only if $\lambda \neq 0$ and $\mu \neq 0$.

Without proof, we state the following theorem which should be known.

Theorem 2.17 (Cosine Theorem). *Let a, b, c be the sides or a triangle and let φ be the angle between the sides a and b. Then*

$$c^2 = a^2 + b^2 - 2ab\cos\varphi. \tag{2.1}$$



Theorem 2.18. *Let $\vec{v}, \vec{w} \in \mathbb{R}^2$ and let $\varphi = \sphericalangle(\vec{v}, \vec{w})$. Then*

$$\boxed{\langle \vec{v}\,, \vec{w}\rangle = \|\vec{v}\|\|\vec{w}\|\cos\varphi.}$$

Proof.

The vectors $\vec{v}$ and $\vec{w}$ define a triangle in $\mathbb{R}^2$, see Figure 2.8. Now we apply the cosine theorem with $a = \|\vec{v}\|$, $b = \|\vec{w}\|$, $c = \|\vec{v} - w\|$. We obtain

$$\|\vec{v} - \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\|\vec{v}\|\|\vec{w}\|\cos\varphi. \tag{2.2}$$

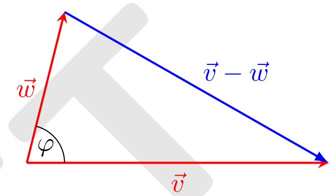


FIGURE 2.8: Triangle given by $\vec{v}$ and $\vec{w}$.

On the other hand,

$$\begin{aligned}\|\vec{v} - \vec{w}\|^2 &= \langle \vec{v} - \vec{w}\,, \vec{v} - \vec{w}\rangle = \langle \vec{v}\,, \vec{v}\rangle - \langle \vec{v}\,, \vec{w}\rangle - \langle \vec{w}\,, \vec{v}\rangle + \langle \vec{w}\,, \vec{w}\rangle = \langle \vec{v}\,, \vec{v}\rangle - 2\langle \vec{v}\,, \vec{w}\rangle + \langle \vec{w}\,, \vec{w}\rangle \\ &= \|\vec{v}\|^2 - 2\langle \vec{v}\,, \vec{w}\rangle + \|\vec{w}\|^2. \end{aligned} \tag{2.3}$$

Comparison of (2.2) and (2.3) shows that

$$\|\vec{v}\|^2 + \|\vec{w}\|^2 - 2\|\vec{v}\|\|\vec{w}\|\cos\varphi = \|\vec{v}\|^2 - 2\langle \vec{v}\,, \vec{w}\rangle + \|\vec{w}\|^2,$$

which gives the claimed formula. □

A very important consequence of this theorem is that we can now determine if two vectors are parallel or perpendicular to each other by simply calculating their inner product as can be seen from the following corollary.

Corollary 2.19. *Let $\vec{v}, \vec{w} \in \mathbb{R}^2$ and $\varphi = \sphericalangle(\vec{v}, \vec{w})$. Then:*

(i) $|\langle \vec{v}\,, \vec{w}\rangle| \le \|\vec{v}\|\,\|\vec{w}\|$.

(ii) $\vec{v} \parallel \vec{w} \quad \iff \quad \|\vec{v}\|\,\|\vec{w}\| = |\langle \vec{v}\,, \vec{w}\rangle|$.

(iii) $\vec{v} \perp \vec{w} \quad \iff \quad \langle \vec{v}\,, \vec{w}\rangle = 0$.

Proof. (i) By Theorem 2.18 we have that $|\langle \vec{v}\,, \vec{w}\rangle| = \|\vec{v}\|\,\|\vec{w}\|\cos\varphi \le \|\vec{v}\|\,\|\vec{w}\|$ since $0 \le \cos\varphi \le 1$ for $\varphi \in [0, \pi]$.

The claims in (ii) and (iii) are clear if one of the vectors is equal to $\vec{0}$ since the zero vector is parallel and orthogonal to every vector in $'R^2$. So let us assume now that $\vec{v} \ne \vec{0}$ and $\vec{w} \ne \vec{0}$.

(ii) From Theorem 2.18 we have that $|\langle \vec{v}\,, \vec{w}\rangle| = \|\vec{v}\|\,\|\vec{w}\|$ if and only if $\cos\varphi = 1$. This is the case if and only if $\varphi = 0$ or π, that is, if and only if $\vec{v}$ and $\vec{w}$ are parallel.

(iii) From Theorem 2.18 we have that $|\langle \vec{v}\,, \vec{w}\rangle| = 0$ if and only if $\cos\varphi = 0$. This is the case if and only if $\varphi = \pi/2$, that is, if and only if $\vec{v}$ and $\vec{w}$ are perpendicular. ☐

Prove Proposition 2.15(ii) and (iii).

Example 2.20. Theorem 2.18 allows us to calculate the angle of a given vector with the x-axis easily (see Figure 2.9):

$$\cos\varphi_x = \frac{\langle \vec{v}\,, \vec{e}_1\rangle}{\|\vec{v}\|\|\vec{e}_1\|}, \qquad \cos\varphi_y = \frac{\langle \vec{v}\,, \vec{e}_2\rangle}{\|\vec{v}\|\|\vec{e}_2\|}.$$

If we now use that $\|\vec{e}_1\| = \|\vec{e}_2\| = 1$ and that $\langle \vec{v}\,, \vec{e}_1\rangle = v_1$ and $\langle \vec{v}\,, \vec{e}_2\rangle = v_2$, then

$$\cos\varphi_x = \frac{v_1}{\|\vec{v}\|}, \qquad \cos\varphi_y = \frac{v_2}{\|\vec{v}\|}.$$

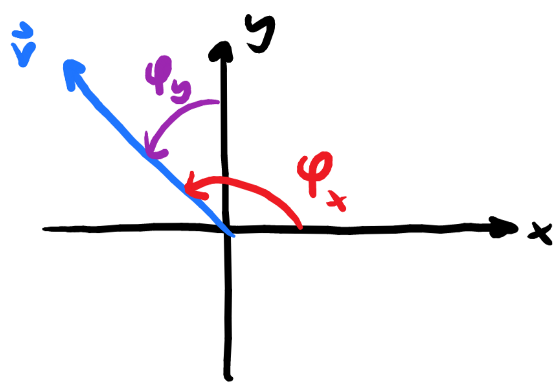


FIGURE 2.9: Angle of $\vec{v}$ with the axes.

You should have understood

- the concepts of being parallel and of being perpendicular,
- the relation of the inner product with the length of a vector and the angle between two vectors,
- that the inner product is commutative and associative, but that it is not a product,
- ...

You should now be able to

- calculate the inner product of two vectors,
- use the inner product to calculate angles between vectors
- use the inner product to determine if two vectors are parallel, perpendicular or neither,
- ...

2.3 Orthogonal Projections in $\mathbb{R}^2$

Let $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{R}^2$ and $\vec{w} \neq \vec{0}$. Geometrically, we have an intuition of what the orthogonal projection of $\vec{v}$ onto $\vec{w}$ should be and that we should be able to construct it as described in the following procedure: We move $\vec{v}$ such that its initial point coincides with that of $\vec{w}$. Then we extend $\vec{w}$ to a line and construct a line that passes through the tip of $\vec{v}$ and is perpendicular to $\vec{w}$. The vector from the initial point to the intersection of the two lines should then be the *orthogonal projection of* $\vec{v}$ *onto* $\vec{w}$. see Figure 2.10

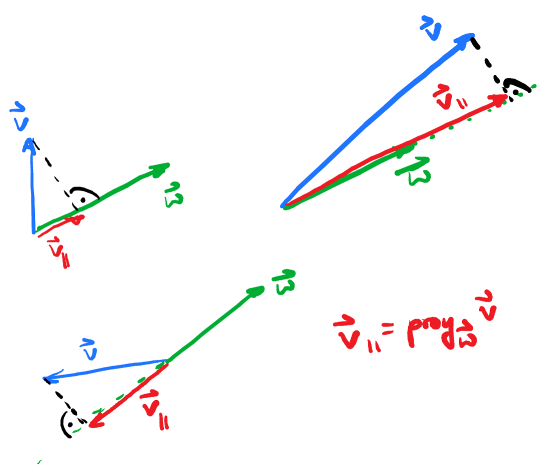


FIGURE 2.10: Orthogonal projections in $\mathbb{R}^2$.

What we did was to decompose the vector $\vec{v}$ in part parallel to $\vec{w}$ and a part perpendicular to $\vec{w}$ so that their sum gives us back $\vec{v}$. The parallel part is the orthogonal projection.
In the following theorem we give the precise meaning of orthogonal projection, we show that a decomposition as describe above always exists and that we even give a formula for orthogonal projection.

Theorem 2.21 (Orthogonal projection). *Let $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{R}^2$ and $\vec{w} \neq \vec{0}$. Then there exist uniquely determined vectors $\vec{v}_\|$ and $\vec{v}_\perp$ such that*

$$\vec{v}_\| \parallel \vec{w}, \quad \vec{v}_\perp \perp \vec{w} \qquad \textit{and} \qquad \vec{v} = \vec{v}_\| + \vec{v}_\perp. \tag{2.4}$$

The vector $\vec{v}_\|$ is called the orthogonal projection of $\vec{v}$ onto $\vec{w}$ *and it is given by*

$$\vec{v}_\| = \frac{\langle \vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}\, \vec{w}. \tag{2.5}$$

Proof. Assume we have vectors $\vec{v}_\|$ and $\vec{v}_\perp$ satisfying (2.4). Since $\vec{v}_\|$ and $\vec{w}$ are parellel by definition and $\vec{w} \neq \vec{0}$, there exists $\lambda \in \mathbb{R}$ such that $\vec{v}_\| = \lambda\vec{w}$, so in order to find $\vec{v}_\|$ it is suffiecient to determine λ. For this, we notice that $\vec{v} = \lambda\vec{w} + \vec{v}_\perp$ by (2.4). Taking the inner product on both sides with $\vec{w}$ leads to

$$\begin{aligned} \langle \vec{v}\,, \vec{w}\rangle = \langle \lambda\vec{w} + \vec{v}_\perp\,, \vec{w}\rangle = \langle \lambda\vec{w}\,, \vec{w}\rangle + \underbrace{\langle \vec{v}_\perp\,, \vec{w}\rangle}_{=0 \text{ since } \vec{v}_\perp \perp \vec{w}} = \langle \lambda\vec{w}\,, \vec{w}\rangle = \lambda\langle \vec{w}\,, \vec{w}\rangle = \lambda\|\vec{w}\|^2 \\ \implies \qquad \lambda = \frac{\langle \vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}. \end{aligned}$$

So if a sum representatio of $\vec{v}$ as in (2.4) exists, then the only possibility is

$$\vec{v}_\| = \lambda\vec{w} = \frac{\langle \vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}\vec{w} \qquad \text{and} \qquad \vec{v}_\perp = \vec{v} - \vec{v}_\| = \vec{v} - \frac{\langle \vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}\vec{w}.$$

This already proves uniqueness of the vectors $\vec{v}_\|$ and $\vec{v}_\perp$. It remains to show that they indeed have the properties that we want. Clearly, by construction $\vec{v}_\|$ is parallel to $\vec{w}$ and $\vec{v} = \vec{v}_\| + \vec{v}_\perp$ since we defined $\vec{v}_\perp = \vec{v} - \vec{v}_\|$. Finally, we verify that $\vec{v}_\perp$ is orthogonal to $\vec{w}$:

$$\langle \vec{v}_\perp\,, \vec{w}\rangle = \left\langle \vec{v} - \frac{\langle \vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}\vec{w}\,, \vec{w}\right\rangle = \langle \vec{v}\,, \vec{w}\rangle - \left\langle \frac{\langle \vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}\vec{w}\,, \vec{w}\right\rangle = \langle \vec{v}\,, \vec{w}\rangle - \frac{\langle \vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}\,\langle \vec{w}\,, \vec{w}\rangle = 0$$

where in the last step we used that $\langle \vec{w}\,, \vec{w}\rangle = \|\vec{w}\|^2$. $\square$

Notation 2.22. Instead of $\vec{v}_\|$ we often write $\operatorname{proj}_{\vec{w}} \vec{v}$, in particular when we want to emphasise onto which vector we are projecting.

Remark 2.23. (i) $\operatorname{proj}_{\vec{w}} \vec{v}$ depends only on the direction of $\vec{w}$. It does **not** depend on its length.

Proof. By our geometric intuition, this should be clear. But we can see this also from the formula. Suppose we want to project $\vec{v}$ onto $c\vec{w}$ for some $c \in \mathbb{R} \setminus \{0\}$. Then

$$\operatorname{proj}_{c\vec{w}} \vec{v} = \frac{\langle \vec{v}\,, c\vec{w}\rangle}{\|c\vec{w}\|^2}(c\vec{w}) = \frac{c\langle \vec{v}\,, \vec{w}\rangle}{c^2\|\vec{w}\|^2}(c\vec{w}) = \frac{\langle \vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}\vec{w} = \operatorname{proj}_{\vec{w}} \vec{v}. \qquad \square$$

Convince yourself graphically it should not matter if we project $\vec{v}$ on $\vec{w}$ or on $5\vec{w}$ or on $-\frac{7}{5}\vec{w}$; only the direction of $\vec{w}$ matters, not its length.

(ii) For every $c \in \mathbb{R}$, we have that $\operatorname{proj}_{\vec{w}}(c\vec{v}) = c \operatorname{proj}_{\vec{w}} \vec{v}$.

Proof. Again, by geometric considerations, this should be clear. The corresponding calculation is

$$\operatorname{proj}_{\vec{w}}(c\vec{v}) = \frac{\langle c\vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}\vec{w} = \frac{c\langle \vec{v}\,, \vec{w}\rangle}{\|\vec{w}\|^2}\vec{w} = c \operatorname{proj}_{\vec{w}} \vec{v}. \qquad \square$$

(iii) As special cases of the above, we find $\operatorname{proj}_{\vec{w}}(-\vec{v}) = -\operatorname{proj}_{\vec{w}} \vec{v}$ and $\operatorname{proj}_{-\vec{w}} \vec{v} = \operatorname{proj}_{\vec{w}} \vec{v}$.

(iv) $\vec{v} \parallel \vec{w} \quad \Longrightarrow \quad \operatorname{proj}_{\vec{w}} \vec{v} = \vec{v}$.

(v) $\vec{v} \perp \vec{w} \quad \Longrightarrow \quad \operatorname{proj}_{\vec{w}} \vec{v} = \vec{0}$.

(vi) $\operatorname{proj}_{\vec{w}} \vec{v}$ is the unique vector in $\mathbb{R}^2$ such that

$$(\vec{v} - \operatorname{proj}_{\vec{w}} \vec{v}) \perp \vec{v} \quad \text{and} \quad \operatorname{proj}_{\vec{w}} \vec{v} \parallel \vec{w}.$$

We end this section with some examples.

Example 2.24. Let $\vec{u} = 2\vec{\mathrm{e}}_1 + 3\vec{\mathrm{e}}_2$, $\vec{v} = 4\vec{\mathrm{e}}_1 - \vec{\mathrm{e}}_2$.

(i) $\operatorname{proj}_{\vec{\mathrm{e}}_1} \vec{u} = \frac{\langle \vec{u}\,, \vec{\mathrm{e}}_1\rangle}{\|\vec{\mathrm{e}}_1\|^2}\vec{\mathrm{e}}_1 = \frac{2}{1^2}\vec{\mathrm{e}}_1 = 2\vec{\mathrm{e}}_1$.

(ii) $\operatorname{proj}_{\vec{\mathrm{e}}_2} \vec{u} = \frac{\langle \vec{u}\,, \vec{\mathrm{e}}_2\rangle}{\|\vec{\mathrm{e}}_2\|^2}\vec{\mathrm{e}}_2 = \frac{3}{1^2}\vec{\mathrm{e}}_2 = 3\vec{\mathrm{e}}_2$.

(iii) Similarly, we can calculate $\operatorname{proj}_{\vec{\mathrm{e}}_1} \vec{v} = 4\vec{e}_1$, $\operatorname{proj}_{\vec{\mathrm{e}}_2} \vec{v} = -\vec{e}_2$.

(iv) $\operatorname{proj}_{\vec{u}} \vec{v} = \frac{\langle \vec{u}\,, \vec{v}\rangle}{\|\vec{u}\|^2}\vec{u} = \dfrac{\left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ -1 \end{pmatrix} \right\rangle}{\|\vec{u}\|^2}\vec{u} = \frac{8-3}{2^2+3^2}\vec{u} = \frac{5}{13}\vec{u} = \frac{5}{13}\begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

(v) $\operatorname{proj}_{\vec{v}} \vec{u} = \frac{\langle \vec{v}\,, \vec{u}\rangle}{\|\vec{v}\|^2}\vec{v} = \dfrac{\left\langle \begin{pmatrix} 4 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\rangle}{\|\vec{u}\|^2}\vec{u} = \frac{8-3}{4^2+(-1)^2}\vec{v} = \frac{5}{17}\vec{v} = \frac{5}{17}\begin{pmatrix} 4 \\ -1 \end{pmatrix}$.

Example 2.25 (Angle with coordinate axes). Let $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 \setminus \{\vec{0}\}$.

Then $\cos \sphericalangle(\vec{v}, \vec{\mathrm{e}}_1) = \frac{a}{\|\vec{v}\|}$, $\cos \sphericalangle(\vec{v}, \vec{\mathrm{e}}_2) = \frac{b}{\|\vec{v}\|}$, hence

$$\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix} = \|\vec{v}\| \begin{pmatrix} \cos \sphericalangle(\vec{v}, \vec{\mathrm{e}}_1) \\ \cos \sphericalangle(\vec{v}, \vec{\mathrm{e}}_2) \end{pmatrix}.$$

qu:2:01

Question 2.1

Let $\vec{w}$ be a vector in $\mathbb{R}^2 \setminus \{\vec{0}\}$.

(i) Can you describe geometrically all the vectors $\vec{v}$ whose projection onto $\vec{w}$ is equal to $\vec{0}$?

(ii) Can you describe geometrically all the vectors $\vec{v}$ whose projection onto $\vec{w}$ have length 2?

(iii) Can you describe geometrically all the vectors $\vec{v}$ whose projection onto $\vec{w}$ have length $3\|\vec{w}\|$?

You should have understood

- the concept of orthogonal projections in $\mathbb{R}^2$,
- ...

You should now be able to

- calculate the projection of a given vector onto another vector,
- ...

2.4 Vectors in $\mathbb{R}^n$

In this section we extend our calculations from $\mathbb{R}^2$ to $\mathbb{R}^n$. If $n = 3$, then we obtain $\mathbb{R}^3$ which usually serves as model for our everyday physcial world and which you probably already are familiar with from physics lectures.
First, let us define $\mathbb{R}^n$.

Definition 2.26. For $n \in \mathbb{N}$ we define the set

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in \mathbb{R} \right\}.$$

Again we can think of vectors as arrows. As in $\mathbb{R}^2$, we can identify every point in $\mathbb{R}^n$ with the arrow that starts in the origin of coordinate system and ends in the given point. The set of all arrows with the same length and the same direction is called a vector in $\mathbb{R}^n$. So every point $P(p_1, \dots, p_n)$ defines a vector $\vec{v} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$ and vice versa. As before, we sometimes denote vectors as $(p_1, \dots, p_n)^t$ in order to save (vertical) space. The superscript t stands for "transposed".

$\mathbb{R}^n$ becomes a vector space with the operations

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n,\ \vec{v} + \vec{w} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}, \qquad \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n,\ c\vec{v} = \begin{pmatrix} cv_1 \\ \vdots \\ cv_n \end{pmatrix}. \tag{2.6}$$

Show that $\mathbb{R}^n$ is a vector space. That is, you have to show that the vector space axioms on page 25 hold.

As in $\mathbb{R}^2$, we can define the norm of a vector, the angle between two vectors and an inner product. Note that the definition of a the angle between two vectors is not different from that one in $\mathbb{R}^2$ since when we are given two vectors, thay always lie in a common plane which we can imagine as some sort of rotated $\mathbb{R}^2$.
Let us give the formal definitions.

Definition 2.27 (Inner product; norm of a vector). For vectors $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ and $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ the *inner product* (or *scalar product* or *dot product*) is defined as

$$\langle \vec{v}\,, \vec{w} \rangle = \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \right\rangle = v_1 w_1 + \cdots + v_n w_n.$$

The *length* of $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$ is denoted by $\|\vec{v}\|$ and it is given by

$$\|\vec{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}\ .$$

Other names for the length of $\vec{v}$ are *magnitude of* $\vec{v}$ or *norm of* $\vec{v}$.

As in $\mathbb{R}^2$, we have the following properties:

(i) *Symmetry of the inner product:* For all vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$, we have that $\langle \vec{v}\,, \vec{w} \rangle = \langle \vec{w}\,, \vec{v} \rangle$.

(ii) *Bilinearity of the inner product:* For all vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ and all $c \in \mathbb{R}$, we have that $\langle \vec{u}\,, \vec{v} + c\vec{w} \rangle = \langle \vec{u}\,, \vec{v} \rangle + c\langle \vec{u}\,, \vec{w} \rangle$.

(iii) *Relation of the inner product with the angle between vectors:* Let $\vec{v}, \vec{w} \in \mathbb{R}^n$ and let $\varphi = \sphericalangle(\vec{v}, \vec{w})$. Then
$$\langle \vec{v}\,, \vec{w} \rangle = \|\vec{v}\|\,\|\vec{w}\| \cos\varphi.$$

Remark 2.28. Actually, the inner product usually is used to *define* the angle between two vectors by the formula above.

In particular, we have (cf. Proposition 2.15):

$$
\begin{array}{llllll}
\text{(a)} & \vec{v} \parallel \vec{w} & \iff & \sphericalangle(\vec{v}, \vec{w}) \in \{0, \pi\} & \iff & |\langle \vec{v}\,, \vec{w} \rangle| = \|\vec{v}\|\,\|\vec{w}\| \\
\text{(b)} & \vec{v} \perp \vec{w} & \iff & \sphericalangle(\vec{v}, \vec{w}) = \pi/2 & \iff & \langle \vec{v}\,, \vec{w} \rangle = 0.
\end{array}
$$

(iv) *Relation of norm and inner product:* For all vectors $\vec{v} \in \mathbb{R}^n$, we have that $\|\vec{v}\|^2 = \langle \vec{v}\,, \vec{v} \rangle$.

(v) *Properties of the norm:* For all vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$ and scalars $c \in \mathbb{R}$, we have that $\|c\vec{v}\| = |c|\|\vec{v}\|$ and $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$.

(vi) *Orthogonal projections of one vector onto another:* For all vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$ with $\vec{w} \neq \vec{0}$ the *orthogonal projection of $\vec{v}$ onto $\vec{w}$* is

$$
\operatorname{proj}_{\vec{w}} \vec{v} = \frac{\langle \vec{v}\,, \vec{w} \rangle}{\|\vec{w}\|^2}\, \vec{w}.
$$

As in $\mathbb{R}^3$, we have three “special vectors” which are parallel to the coordinate axes:

$$
\vec{\mathrm{e}}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{\mathrm{e}}_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \vec{\mathrm{e}}_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.
$$

In the special case $n = 3$, the vectors $\vec{\mathrm{e}}_1$, $\vec{\mathrm{e}}_2$ and $\vec{\mathrm{e}}_3$ are sometimes denoted by $\hat{\imath}, \hat{\jmath}, \hat{k}$.
For a given vector $\vec{v} \neq \vec{0}$, we can now easily determine its angle with the coordinate axes. Let φ_j be the angle between $\vec{v}$ and the x_j-axis. Then

$$
\varphi_j = \sphericalangle(\vec{v}, \vec{\mathrm{e}}_j) \quad \Longrightarrow \quad \cos\varphi_x = \frac{\langle \vec{v}\,, \vec{e}_j \rangle}{\|\vec{v}\|\,\|\vec{\mathrm{e}}_j\|} = \frac{v_j}{\|\vec{v}\|}.
$$

From this we see that $\vec{v} = \|\vec{v}\| \begin{pmatrix} \cos\varphi_1 \\ \vdots \\ \cos\varphi_n \end{pmatrix}$. Sometimes the notation

$$
\hat{v} := \frac{\vec{v}}{\|\vec{v}\|} = \|\vec{v}\| \begin{pmatrix} \cos\varphi_1 \\ \vdots \\ \cos\varphi_n \end{pmatrix}
$$

is used. Clearly $\|\hat{v}\| = 1$ because $\|\hat{v}\| = \|\|\vec{v}\|^{-1}\vec{v}\| = \|\vec{v}\|^{-1}\|\vec{v}\| = 1$. Therefore $\hat{v}$ is a unit vector pointing in the same direction as the original vector $\vec{v}$.

You should have understood

- the vector space $\mathbb{R}^n$ and vectors in $\mathbb{R}^n$,
- geometric concepts (angles, length of a vector) in $\mathbb{R}^n$,
- that $\mathbb{R}^2$ from chapter 2.1 is a special case of $\mathbb{R}^n$ from this section,
- ...

You should now be able to

- perform algebraic operations in the vector space $\mathbb{R}^3$ and, in the case $n = 3$, visualize them in space,
- calculate lengths and angles,
- calculate unit vectors, scale vectors,
- perform simple abstract proofs (e.g., prove that $\mathbb{R}^n$ is a vector space).
- ...

2.5 Vectors in $\mathbb{R}^3$ and the cross product

The space $\mathbb{R}^3$ is very important since it is used in mechanics to model the space we live in. On $\mathbb{R}^3$ we can define an additional operation on vectors, the so-called *cross product*. Another name for it its *vector product*. It takes two vectors and gives back two vectors. It does have several properties which makes it look like a product, however we will see that it is **not** a product. Here is the definition.

Definition 2.29 (Cross product). Let $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \in \mathbb{R}^3$. Their *cross product* (or *vector product* or *wedge product*) is

$$\vec{v} \times \vec{w} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} := \begin{pmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{pmatrix}.$$

Another notation for the cross product is $\vec{v} \wedge \vec{w}$.

The cross product is defined only in $\mathbb{R}^3$!

Before we collect some easy properties of the cross product, let us calculate a few examples.

Examples 2.30. Let $\vec{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}$.

- $\vec{u} \times \vec{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 2 \cdot 7 - 3 \cdot 6 \\ 3 \cdot 5 - 1 \cdot 7 \\ 1 \cdot 6 - 2 \cdot 5 \end{pmatrix} = \begin{pmatrix} 14 - 18 \\ 15 - 7 \\ 6 - 10 \end{pmatrix} = \begin{pmatrix} -4 \\ 8 \\ -4 \end{pmatrix}$,

- $\vec{v} \times \vec{u} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \times \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \cdot 3 - 7 \cdot 2 \\ 7 \cdot 1 - 3 \cdot 5 \\ 5 \cdot 2 - 6 \cdot 1 \end{pmatrix} = \begin{pmatrix} 18 - 14 \\ 7 - 15 \\ 10 - 6 \end{pmatrix} = \begin{pmatrix} 4 \\ -8 \\ 4 \end{pmatrix}$,
- $\vec{v} \times \vec{e}_1 = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \cdot 0 - 7 \cdot 0 \\ 7 \cdot 0 - 7 \cdot 1 \\ 5 \cdot 0 - 6 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -7 \\ -6 \end{pmatrix}$,
- $\vec{v} \times \vec{v} = \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \times \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} = \begin{pmatrix} 6 \cdot 7 - 7 \cdot 6 \\ 7 \cdot 5 - 5 \cdot 7 \\ 5 \cdot 6 - 6 \cdot 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Proposition 2.31 (Properties of the cross product). *Let $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$ and let $c \in \mathbb{R}$. Then:*

(i) $\vec{u} \times \vec{0} = \vec{0} \times \vec{u} = \vec{0}$.

(ii) $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$.

(iii) $\vec{u} \times (\vec{v} + \vec{w}) = (\vec{u} \times \vec{v}) + (\vec{u} \times \vec{w})$.

(iv) $(c\vec{u}) \times \vec{v} = c(\vec{u} \times \vec{v})$.

(v) $\vec{u} \parallel \vec{v} \quad \Longrightarrow \quad \vec{u} \times \vec{v} = \vec{0}$. *In particular,* $\vec{v} \times \vec{v} = \vec{0}$.

(vi) $\langle \vec{u}\,, \vec{v} \times \vec{w} \rangle = \langle \vec{u} \times \vec{v}\,, \vec{w} \rangle$.

(vii) $\langle \vec{u}\,, \vec{u} \times \vec{v} \rangle = 0$ *and* $\langle \vec{v}\,, \vec{u} \times \vec{v} \rangle = 0$, *in particular*

$$\boxed{\vec{v} \perp \vec{v} \times \vec{u}, \quad \vec{u} \perp \vec{v} \times \vec{u}}$$

that means that the vector $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$.

Proof. The proofs of the formulas (i) to (v) are easy calculations (you should do them!).

(vi) The proof is a long but straightforward calculation:

$$\begin{aligned} \langle \vec{u}\,, \vec{v} \times \vec{w} \rangle &= \left\langle \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \times \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - w_3 v_1 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} \right\rangle \\ &= u_1(v_2 w_3 - v_3 w_2) + u_2(v_3 w_1 - v_1 w_3) + u_3(v_1 w_2 - v_2 w_1) \\ &= u_1 v_2 w_3 - u_1 v_3 w_2 + u_2 v_3 w_1 - u_2 v_1 w_3 + u_3 v_1 w_2 - u_3 v_2 w_1 \\ &= u_2 v_3 w_1 - u_3 v_2 w_1 + u_3 v_1 w_2 - u_1 v_3 w_2 + u_1 v_2 w_3 - u_2 v_1 w_3 \\ &= (u_2 v_3 - u_3 v_2) w_1 + (u_3 v_1 - u_1 v_3) w_2 + (u_1 v_2 - u_2 v_1) w_3 \\ &= \langle \vec{u} \times \vec{v}\,, \vec{w} \rangle. \end{aligned}$$

(vii) It follows from (vi) and (v) that

$$\langle \vec{u}\,, \vec{u} \times \vec{v} \rangle = \langle \vec{u} \times \vec{u}\,, \vec{v} \rangle = \langle \vec{0}\,, \vec{v} \rangle = 0. \qquad \square$$

Note that the cross product is distributive but it is not commutative nor associative.

Show the formulas in (i) to (v).

Remark 2.32. The property (vii) explains why the cross product makes sense only in $\mathbb{R}^3$. Given two non-parallel vectors $\vec{v}$ and $\vec{w}$, their cross product is a vector which is orthogonal to both of them and whose length is $\|\vec{v}\|\,\|\vec{w}\|\sin\varphi$ (see Theorem 2.33) and this should define the result uniquely up to a factor ± 1. Here $\varphi = \sphericalangle(\vec{v}, \vec{w})$.

- If we were in $\mathbb{R}^2$, the problem is that "we do not have enough space" because then the only vector orthogonal to $\vec{v}$ and $\vec{w}$ at the same time would be the zero vector $\vec{0}$ and it would not make too much sense to define a product where the result is always $\vec{0}$.

- If we were in some $\mathbb{R}^n$ with $n \leq 4$, the problem is that "we have too much choice". We will see later in Chapter **??** that the orthogonal complement of the plane generated by $\vec{v}$ and $\vec{w}$ has dimension $n-2$ and every vector in the orthogonal complement is orthogonal to both $\vec{v}$ and $\vec{w}$. For example, if we take $\vec{v} = (1,0,0,0)^t$ and $\vec{w} = (0,1,0,0)^t$, then every vector of the form $\vec{a} = (0,0,x,y)^t$ is perpendicular to both $\vec{v}$ and $\vec{w}$ and it easy to find infinitely many vactors of this form which in addition have norm $\|\vec{v}\|\,\|\vec{w}\|\sin\varphi = 1$ ($\vec{a} = (0,0,\sin\vartheta, \pm\cos\vartheta)^t$ for arbitrary $\vartheta \in \mathbb{R}$ works).

Recall that for the inner product we proved the formula $\langle \vec{v}\,, \vec{w}\rangle = \|\vec{v}\|\,\|\vec{w}\|\cos\varphi$ where φ is the angle between the two vectors, see Theorem 2.18. In the next theorem we will prove a similar relation for the cross product.

Theorem 2.33. *Let $\vec{v}, \vec{w}$ be vectors in $\mathbb{R}^3$ and let φ be the angle between them. Then*

$$\boxed{\|\vec{v} \times \vec{w}\| = \|\vec{v}\|\,\|\vec{w}\|\sin\varphi}$$

Proof. A long, but straightforward calculations shows that $\|\vec{v} \times \vec{w}\|^2 = \|\vec{u}\|^2\|\vec{w}\|^2 - \langle \vec{v}\,, \vec{w}\rangle^2$. Now it follows from Theorem 2.18 that

$$\begin{aligned}\|\vec{v} \times \vec{w}\|^2 &= \|\vec{u}\|^2\|\vec{w}\|^2 - \langle \vec{v}\,, \vec{w}\rangle^2 = \|\vec{u}\|^2\|\vec{w}\|^2 - \|\vec{v}\|^2\|\vec{w}\|^2(\cos\varphi)^2 \\ &= \|\vec{u}\|^2\|\vec{w}\|^2(1-(\cos\varphi)^2) = \|\vec{u}\|^2\|\vec{w}\|^2(\sin\varphi)^2.\end{aligned}$$

Observe that $\sin\varphi \geq 0$ because $\varphi \in [0,\pi]$. So if we take the square root we we do not need to take the absolute value and we arrive at the claimed formula. □

Show that $\|\vec{v} \times \vec{w}\|^2 = \|\vec{u}\|^2\|\vec{w}\|^2 - \langle \vec{v}\,, \vec{w}\rangle^2$.

Application: Area of a parallelogram and volume of a parelellepiped

Area of a parallelogram

Let $\vec{v}$ and $\vec{w}$ be two vectors in $\mathbb{R}^3$. Then they define a parallelogram (if the vectors are parallel or one of them is equal to $\vec{0}$, it is a *degenerate parallelogram*).

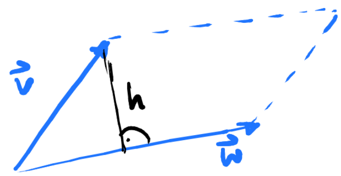


FIGURE 2.11: Parallelogram spanned by $\vec{v}$ and $\vec{w}$.

Proposition 2.34 (Area of a parallelogram). *The area of the parallelogram spanned by the vectors $\vec{v}$ and $\vec{w}$ is*

$$A = \|\vec{v} \times \vec{w}\|. \tag{2.7}$$

Proof. The area of a parallelogram is the product of the length of its base with the height. We can take $\vec{w}$ as base. Let φ be the angle between $\vec{w}$ and $\vec{v}$. Then we obtain that $h = \|\vec{v}\| \sin\varphi$ and therefore, with the help of Theorem 2.33

$$A = \|\vec{w}\|h = \|\vec{w}\|\|\vec{v}\| \sin\varphi = \|\vec{v} \times \vec{w}\|. \qquad \square$$

Note that in the case when $\vec{v}$ and $\vec{w}$ are parallel, this gives the right answer $A = 0$.
Any three vectors in $\mathbb{R}^3$ define a parallelepiped.

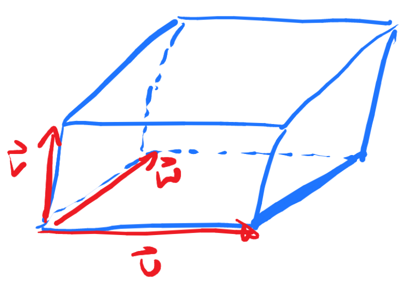


FIGURE 2.12: Parallelepiped spanned by $\vec{u}, \vec{v}, \vec{w}$.

Proposition 2.35 (Volume of a parallelepiped). *The volume of the parallelepiped spanned by the vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ is*

$$V = \|\vec{u}(\vec{v} \times \vec{w})\|. \tag{2.8}$$

Proof. The volume of a parallelepiped is the product of the area of its base with the height. Let us take the parallelogram spanned by $\vec{v}, \vec{w}$ as base. If $\vec{v}$ and $\vec{w}$ are parallel or one or them is equal to $\vec{0}$, then (2.8) is true because $V = 0$ and $\vec{v} \times \vec{w} = \vec{0}$ in this case.
Now let us assume that they are not parallel. By Proposition 2.34 we already know that its base has area $A = \|\vec{v} \times \vec{w}\|$. The height is the length of the orthogonal projection of $\vec{u}$ onto the normal vector of the plane spanned by $\vec{v}$ and $\vec{w}$. We already know that $\vec{v} \times \vec{w}$ is such a normal vector. Hence we obtain that

$$h = \|\operatorname{proj}_{\vec{v}\times\vec{w}} \vec{u}\| = \left\| \frac{\langle \vec{u}\,, \vec{v} \times \vec{w}\rangle}{\|\vec{v} \times \vec{w}\|^2} \vec{v} \times \vec{w} \right\| = \frac{|\langle \vec{u}\,, \vec{v} \times \vec{w}\rangle|}{\|\vec{v} \times \vec{w}\|^2} \|\vec{v} \times \vec{w}\| = \frac{|\langle \vec{u}\,, \vec{v} \times \vec{w}\rangle|}{\|\vec{v} \times \vec{w}\|}.$$

Therefore, the volume of the parallelepiped is

$$V = Ah = \|\vec{v} \times \vec{w}\| \frac{|\langle \vec{u}\,, \vec{v} \times \vec{w}\rangle|}{\|\vec{v} \times \vec{w}\|} = |\langle \vec{u}\,, \vec{v} \times \vec{w}\rangle|.$$

□

Corollary 2.36. *Let $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$. Then*

$$|\langle \vec{u}\,, \vec{v} \times \vec{w}\rangle| = |\langle \vec{v}\,, \vec{w} \times \vec{u}\rangle| = |\langle \vec{w}\,, \vec{u} \times \vec{v}\rangle|.$$

Proof. The formula holds because each of the expressions describes the volume of the parallelepiped spanned by the three given vectors since we can take any of the sides of the parallelogram as its base. □

You should have understood

- the geometric interpreations of the cross product,
- why it exists only in $\mathbb{R}^3$
- ...

You should now be able to

- calculate the cross product,
- use it to say something about the angle between two vectors in $\mathbb{R}^3$,
- use it calculate the area of a parallelogramm and the volume of a parallelepiped,
- ...

2.6 Lines and planes in $\mathbb{R}^3$

In this section we discuss lines and planes, how to describe them and how to calculate, e.g., intersections between them. We work mostly in $\mathbb{R}^3$ and only give some hints on how the concepts discussed here generalise to $\mathbb{R}^n$ with $n \neq 3$. The special case $n = 2$ should be clear.

The formal definition of lines and planes need will be given in Chapter **??** because this requires the concept of linear independence. (for the curious: a *line* is an (affine) one-dimensional subspace of a vector space; a *plane* is an (affine) two-dimensional subspace of a vector space; a *hyperplane* is an (affine) $(n-1)$-dimensional subspace of an n-dimensional vector space). In this section we work with what we know from elemntary geometry.

Lines

Intuitively, it is clear what a line in $\mathbb{R}^3$ should be. In order to describe a line in $\mathbb{R}^3$ completely, it is not necessary to know all its points. It is sufficient to know either

(a) two different points P, Q on the line

or

(b) one point P on the line and the direction of the line.

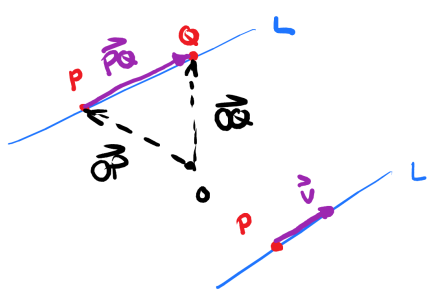


FIGURE 2.13: Line L given (a) by two points P, Q on L, (b) by a point P on L and the direction of L.

Clearly, both descriptions are equivalent because: If we have two different points P, Q on the line L, then its direction is given by the vector $\overrightarrow{PQ}$. If on the other hand we are given a point P on L and a vector $\vec{v}$ which is parallel to L, then we easily get another point Q on L by $\overrightarrow{OQ} = \overrightarrow{0P} + \vec{v}$.

Now we want to give formulas for the line.

Vector equation

Given two points $P(p_1, p_2, p_3)$ and $Q(q_1, q_2, q_3)$ with $P \neq Q$, there is exactly one line L which passes through both points. In formulas, this line is described as

$$L = \left\{ \overrightarrow{0P} + t\overrightarrow{PQ} : t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} p_1 + (q_1 - p_1)t \\ p_2 + (q_2 - p_2)t \\ p_3 + (q_3 - p_3)t \end{pmatrix} : t \in \mathbb{R} \right\} \tag{2.9}$$

If we are given a point $P(p_1, p_2, p_3)$ on L and a vector $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \neq \vec{0}$ parallel to L, then

$$L = \left\{ \overrightarrow{0P} + t\vec{v} : t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} p_1 + v_1 t \\ p_2 + v_2 t \\ p_3 + v_3 t \end{pmatrix} : t \in \mathbb{R} \right\} \tag{2.10}$$

The formulas are easy to understand. They say: In order to trace the line, we first move to an arbitrary point on the line (this is term $\overrightarrow{0P}$) and then we move an amount t along the line. With this procedure we can reach every point on the line, and on the other hand, if we do this, then we are guaranteed to end up on the line.

The formulas (2.9) and (2.10) are called *vector equation* for the line L. Note that they are the same if we set $v_1 = q_1 - p_1$, $v_2 = q_2 - p_2$, $v_3 = q_3 - p_3$. We will mostly use the notation with the v's since it is shorter. The vector $\vec{v}$ is called *directional vector* of the line L.

qu::3

Question 2.2

Is it true that E passes through the origin if and only if $\overrightarrow{0P} = \vec{0}$?

Remark 2.37. It is important to observe that a given line has many different parametrizations.

- For example, the vector equation that we write down depends on the points we choose on L. Clearly, we have infinitely many possibilities to do so.
- Observe that a given line L has many directional vectors. Indeed, if $\vec{v}$ is a directional vector for L, then $c\vec{v}$ is so too for every $c \in \mathbb{R} \setminus \{0\}$.

Check that the following formulas all describe the same line:

(i) $L_1 = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} : t \in \mathbb{R} \right\}$, (ii) $L_2 = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 12 \\ 10 \\ 8 \end{pmatrix} : t \in \mathbb{R} \right\}$,

(ii) $L_3 = \left\{ \begin{pmatrix} 13 \\ 12 \\ 11 \end{pmatrix} + t \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} : t \in \mathbb{R} \right\}$.

qu:2:2

Question 2.3

- How can you see easily if two given lines are parallel or perpendicular to each other?
- How would you define the angle between two lines? Do they have to intersect so that an angle between them can be defined?

Parametric equation

From the formula (2.10) it is clear that a point (x, y, z) belongs to L if and only if there exists $t \in \mathbb{R}$ such that

$$x = p_1 + tv_1, \quad y = p_2 + tv_2, \quad z = p_3 + tv_3. \tag{2.11}$$

If we had started with (2.9), then we had obtained

$$x = p_1 + t(q_1 - p_1), \quad y = p_2 + t(q_2 - p_2), \quad z = p_3 + t(q_3 - p_3) \tag{2.12}$$

The system of equations (2.11) or (2.12) are called the *parametric equations* of L. Here, t is the parameter.

Symmetric equation

Observe that for $(x, y, z) \in L$, the three equations in (2.11) must hold for the same t. So if we assume that $v_1, v_2, v_3 \neq 0$, then we can solve for t and we obtain that

$$\frac{x - p_1}{v_1} = \frac{y - p_2}{v_2} = \frac{z - p_3}{v_3} \tag{2.13}$$

If we use (2.12) then we obtain

$$\frac{x - p_1}{q_1 - p_1} = \frac{y - p_2}{q_2 - p_2} = \frac{z - p_3}{q_3 - p_3}. \tag{2.14}$$

The system of equations (2.13) or (2.6) is called the *symmetric equation* of L.
If for instance, $v_1 = 0$ and $v_2, v_3 \neq 0$, then the line is parallel to the yz-plane and its symmetric equation is

$$x = p_1, \quad \frac{y - p_2}{v_2} = \frac{z - p_3}{v_3}.$$

If $v_1 = v_2 = 0$ and $v_3 \neq 0$, then the line is parallel to the z-axis and its symmetric equation is

$$x = p_1, \quad y = p_2, z \in \mathbb{R}.$$

Representations of lines

In $\mathbb{R}^n$, the vector form of a line is

$$L = \left\{ \overrightarrow{0P} + t\vec{v} : t \in \mathbb{R} \right\}$$

for fixed $P \in L$ and a directional vector $\vec{v}$. Its parametric form is

$$x_1 = p_1 + tv_1, \quad x_2 = p_2 + tv_2, \quad \ldots, \quad x_n = p_n + tv_n, \qquad t \in \mathbb{R},$$

and, assuming that all v_j are different from 0, its symmetric form is

$$\frac{x_1 - p_1}{v_1} = \frac{x_2 - p_2}{v_2} = \cdots = \frac{x_n - p_n}{v_n}.$$

qu:2:4

Question 2.4. Normal form of a line.

In $\mathbb{R}^2$, there is also the *normal form* of a line:

$$L : ax + by = d \tag{2.15}$$

where a, b and d are fixed numbers. This means that L consists of all the points $P(x, y)$ whose coordinates satisfy the equation $ax + by = d$.

(i) Given a line in the form (2.15), find a vector representation.

(ii) Given a line in vector representation, find a normal form (that is, write it as (2.15)).

(iii) What is the geometric interpretation of a, b? (Hint: Draw the line L and the vector $\begin{pmatrix} a \\ b \end{pmatrix}$.

(iv) Can this normal form be extended/generalized to lines in $\mathbb{R}^3$? If it is possible, how can it be done? If it is not possible, why it is not possible.

Planes

In order to know a plane E in $\mathbb{R}^3$ completely, it is sufficient to

(a) three points P, Q on the plane that do not lie on a line,

or

(b) one point P on the plane and two non-parallel vectors $\vec{v}, \vec{w}$ which are both parallel the plane,

or

(c) one point P on the plane and a vector $\vec{n}$ which is perpendicular to the plane,

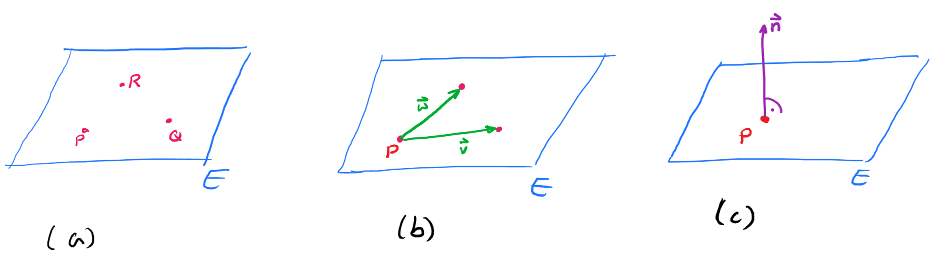


FIGURE 2.14: Plane E given (a) by three points P, Q, R on E, (b) by a point P on E and two vectors $\vec{v}, \vec{w}$ parallel to E. (c) by a point P on E and a vector $\vec{n}$ perpendicular to E.

First, let us see how we can pass from one description to another. Clearly, the descriptions (a) and (b) are equivalent because given three points P, Q, R on E which do not lie on a line, we can form the vectors $\overrightarrow{PQ}$ and $\overrightarrow{PR}$. Theses vectors are then parallel to the plane E but are not parallel to each other. (Of course, we also could have taken $\overrightarrow{QR}$ and $\overrightarrow{QP}$ or $\overrightarrow{RP}$ and $\overrightarrow{RQ}$.) If, on the other hand, we have one point P on E and two vectors $\vec{v}$ and $\vec{w}$, parallel to E and $\vec{v} \not\parallel \vec{w}$, then we can easily get two other points on E, for instance by $\overrightarrow{0Q} = \overrightarrow{0P} + \vec{v}$ and $\overrightarrow{0R} = \overrightarrow{0P} + \vec{w}$. Then the three points P, Q, R lie on E and do not lie on a line.

Vector equation of a plane

In formulas, we can now describe our plane E as

$$E = \left\{ (x, y, z) : \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \overrightarrow{0P} + s\vec{v} + t\vec{w} \quad \text{for some } s, t \in \mathbb{R} \right\}.$$

As in the case of the vector equation of a line, it is easy to understand the formula. We first move to an arbitrary point on the line (this is term $\overrightarrow{0P}$) and then we move parallel to the plane as we please (this is the term $s\vec{v} + t\vec{w}$). With this procedure we can reach every point on the plane, and on the other hand, if we do this, then we are guaranteed to end up on the plane.

qu:2:5

Question 2.5

Is it true that E passes through the origin if and only if $\overrightarrow{0P} = \vec{0}$?

Normal form of a plane

Now we want to use the normal vector of the plane to describe it. Assume that we are given a point P on E and a vector $\vec{n}$ perpendicular to the plane. This means that every vector which is parallel to the plane E must be perpendicular to $\vec{n}$. If we take an arbitrary point $Q(x, y, z) \in \mathbb{R}^3$, then $Q \in E$ if and only if $\overrightarrow{PQ}$ is parallel to E, that means that $\overrightarrow{PQ}$ is orthogonal to $\vec{n}$. Recall that two vectors are perpendicular if and only if their inner product is 0, so $Q \in E$ if and only if

$$\begin{aligned} 0 = \langle n \,, \overrightarrow{PQ} \rangle &= \left\langle \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} , \begin{pmatrix} x - p_1 \\ y - p_2 \\ z - p_3 \end{pmatrix} \right\rangle = n_1(x - p_1) + n_2(y - p_2) + n_3(z - p_3) \\ &= n_1 x + n_2 y + n_3 z - (n_1 p_1 + n_2 p_2 + n_3 - p_3) \end{aligned}$$

If we set $d = n_1 p_1 + n_2 p_2 + n_3 - p_3$, then it follows that a point $Q(x, y, z)$ belongs to E if and only if its coordinates satisfy

$$n_1 x + n_2 y + n_3 z = d. \tag{2.16}$$

Equation (2.16) is called the *normal form* for the plane E and $\vec{n}$ is called a *normal vector* of E.

Notation 2.38. In order to define E, we write $E : n_1 x + n_2 y + n + 3z = d$. As a set, we denote E as $E = \{(x, y, z) : n_1 x + n_2 y + n + 3z = d\}$.

Show that E passes through the origin if and only if $d = 0$.

Remark 2.39. As before, note that the normal equation for a plane is not unique. For instance,

$$x + 2y + 3z = 5 \qquad \text{and} \qquad 2x + 4y + 6z = 10$$

describe the same plane. The reason is that "the" normal vector of a plane is not unique. Given one normal vector $\vec{n}$, then every $c\vec{n}$ with $c \in \mathbb{R} \setminus \{0\}$ is also a normal vector to the plane.

Definition 2.40. The *angle between two planes* is the angle between their normal vectors.

Note that this definition is consistent with the fact that two planes are parallel if and only if their normal vectors are parallel.

Remark 2.41.

- Assume a plane is given as in (b) (that is, we know a point P on E and two vectors $\vec{v}$ and $\vec{w}$ parallel to E but with $\vec{v} \nparallel \vec{w}$). In order to find a description as in (c) (that is one point on E and a normal vector), we only have to find a vector $\vec{n}$ that is perpendicular to both $\vec{v}$ and $\vec{w}$. Proposition 2.31(vii) tells us how to do this: we only need to calculate $\vec{v} \times \vec{w}$.
- Assume a plane is given as in (c) (that is, we know a point P on E and its normal vector). In order to find vectors $\vec{v}$ and $\vec{w}$ as in (b), we can either find two solutions of $\vec{x} \times \vec{n} = 0$ which are not parallel. Or we find only one solution $\vec{v}$ which usually is easy to guess and then calculate $\vec{w} = \vec{v} \times \vec{n}$. This vector is perpendicular to $\vec{n}$ and therefore it is parallel to the plane. It is also perpendicular to $\vec{v}$ and therefore it is not parallel to $\vec{v}$. In total, this vector $\vec{w}$ does what we need.

Representations of planes

In $\mathbb{R}^n$, the vector form of plane is

$$E = \left\{ \overrightarrow{0P} + t\vec{v} + s\vec{w} : t \in \mathbb{R} \right\}$$

for fixed $P \in E$ and a two vectors $\vec{v}$, $\vec{w}$ parallel to the plane but not parellel to each other.

Note that there is no normal form of a plane in $\mathbb{R}^n$ for $n \geq 4$. The reason it that for $n \geq 4$, there are more than just one normal directions to a given plane, so a normal form of a plane E must consist of more than one equations (more precisely, it must consist of $n-2$ equations of the form $n_1x_1 + \dots n_nx_n = d$).

You should have understood

- the concept of lines and planes in $\mathbb{R}^3$,
- how they can be described in formulas,
- ...

You should now be able to

- pass easily between the different descriptions of lines and planes,
- ...

2.7 Intersections of lines and planes in $\mathbb{R}^3$

Intersection of lines

Given two lines G and L in $\mathbb{R}^3$, there are three possibilities:

(a) The lines intersect in exactly one point. In this case, they cannot be parallel.

(b) The lines intersect in infinitely many points. In this case, the lines have to be equal. In particular the have to be parallel.

(c) The lines do not intersect. Note that in contrast to the case in $\mathbb{R}^2$, the lines do not have to be parallel for this to happen. For example, the line $L : x = y = 1$ is a line parallel to the z-axis passing through $(1, 1, 0)$, and $G : x = z = 0$ is a line parallel to the y-axis passing through $(0, 0, 0)$, The lines do not intersect and they are not parallel.

Example 2.42. We consider four lines $L_j = \{\vec{p}_j + t\vec{v}_j : t \in \mathbb{R}\}$ with

$$\text{(i)}\quad \vec{v}_1 = \begin{pmatrix}1\\2\\3\end{pmatrix},\quad \vec{p}_1 = \begin{pmatrix}0\\0\\1\end{pmatrix},\qquad \text{(ii)}\quad \vec{v}_2 = \begin{pmatrix}2\\4\\6\end{pmatrix}, \vec{p}_2 = \begin{pmatrix}2\\4\\7\end{pmatrix},$$

$$\text{(iii)}\quad \vec{v}_3 = \begin{pmatrix}1\\1\\2\end{pmatrix}, \vec{p}_3 = \begin{pmatrix}-1\\0\\0\end{pmatrix},\qquad \text{(iv)}\quad \vec{v}_4 = \begin{pmatrix}1\\1\\2\end{pmatrix}, \vec{p}_4 = \begin{pmatrix}3\\0\\5\end{pmatrix}.$$

We will calculate their mutual intersections.

$\boxed{L_1 \cap L_2 = L_1}$

Proof. A point $Q(x, y, z)$ belongs to $L_1 \cap L_2$ if and only if it belongs both to L_1 and L_2. This means that there must exist an $s \in \mathbb{R}$ such that $\overrightarrow{0Q} = \vec{p}_1 + s\vec{v}_1$ and there must exist a $t \in \mathbb{R}$ such that $\overrightarrow{0Q} = \vec{p}_2 + t\vec{v}_2$. Note that s and t are different parameters. So we are looking for s and t such that

$$\vec{p}_1 + s\vec{v}_1 = \vec{p}_2 + t\vec{v}_2, \qquad \text{that is} \qquad \begin{pmatrix}0\\0\\1\end{pmatrix} + s\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}2\\4\\7\end{pmatrix} + t\begin{pmatrix}2\\4\\6\end{pmatrix}. \tag{2.17}$$

Once we have solved (2.17) for s and t, we insert them into the equations for L_1 and L_2 respectively, in order to obtain Q. Note that (2.17) in reality is a system of three equations: one equation for each component of the vector equation. Writing it out and solving each equation for s, we obtain

$$\begin{aligned} 0 + s &= 2 + 2t \\ 0 + 2s &= 4 + 4t \\ 1 + 3s &= 7 + 6t \end{aligned} \qquad \iff \qquad \begin{aligned} s &= 2 + 2t \\ s &= 2 + 2t \\ s &= 2 + 2t. \end{aligned}$$

This means that we have infinitely many solutions: Given any point R on L_1, there is a corresponding $s \in \mathbb{R}$ such that $\overrightarrow{0R} = \vec{p}_1 + s\vec{v}_1$. Now if we choose $t = (s-2)/2$, then $\overrightarrow{0R} = \vec{p}_2 + t\vec{v}_2$ holds, hence $R \in L_2$ too. If on the other hand we have a point $R' \in L_2$, then there is a corresponding $t \in \mathbb{R}$ such that $\overrightarrow{0R'} = \vec{p}_2 + t\vec{v}_2$. Now if we choose $s = 2 + 2t$, then $\overrightarrow{0R'} = \vec{p}_1 + t\vec{v}_1$ holds, hence $R' \in L_2$ too. In summary, we showed that $L_1 = L_2$. □

Remark 2.43. We could also have seen that the directional vectors of L_1 and L_2 are parallel. In fact, $\vec{v}_2 = 2\vec{v}_1$. It then suffices to show that L_1 and L_2 have at least one point in common in order to conclude that the lines are equal.

$\boxed{L_1 \cap L_3 = \{(1,2,4)\}}$

Proof. As before, we need to find $s, t \in \mathbb{R}$ such that

$$\vec{p}_1 + s\vec{v}_1 = \vec{p}_3 + t\vec{v}_3, \qquad \text{that is} \qquad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + s\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}. \tag{2.18}$$

We write this as a system of equations, we get

$$\begin{array}{lrcl} (1) & 0 + s &=& -1 + t \\ (2) & 0 + 2s &=& 0 + t \\ (3) & 1 + 3s &=& 0 + 2t \end{array} \qquad \iff \qquad \begin{array}{lrcl} (1) & s - t &=& -1 \\ (2) & 2s - t &=& 0 \\ (3) & 3s - 2t &=& -1 \end{array}$$

From (1) it follows that $s = t - 1$. Inserting in (2) gives $0 = 2(t-1) - t = t - 2$, hence $t = 2$. From (1) we then obtain that $s = 2 - 1 = 1$. Observe that so far we used only equations (1) and (2). In order to see if we really found a solution, we must check if it is consistent with (3). Inserting our candidates for s and t, we find that $3 \cdot 1 - 2 \cdot 2 = -1$ which is consistent with (3).
So $L1$ and L_3 intersect in exactly one point. In order to find it, we put $s = 1$ in the equation for L_1:

$$\overrightarrow{0Q} = \vec{p}_1 + 1 \cdot \vec{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix},$$

hence the intersection point is $Q(1,2,4)$.

In order to check if this result is correct, we can put $t = 2$ in the equation for L_3. The result must be the same. The corresponding calculation is:

$$\overrightarrow{0Q} = \vec{p}_3 + 2 \cdot \vec{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix},$$

which confirms that the intersection point is $Q(1,2,4)$. □

$\boxed{L_1 \cap L_4 = \varnothing}$

Proof. As before, we need to find $s, t \in \mathbb{R}$ such that

$$\vec{p}_1 + s\vec{v}_1 = \vec{p}_4 + t\vec{v}_4, \qquad \text{that is} \qquad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + s\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 5 \end{pmatrix} + t\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}. \tag{2.19}$$

We write this as a system of equations and we get

$$\begin{array}{cr} ① & s = 3 + t \\ ② & 2s = t \\ ③ & 1 + 3s = 5 + 2t \end{array} \iff \begin{array}{cr} ① & s - t = 3 \\ ② & 2s - t = 0 \\ ③ & 3s - 2t = 5 \end{array}$$

From ① it follows that $s = t + 3$. Inserting in ② gives $0 = 2(t + 3) - t = t + 6$, hence $t = -6$. From ① we then obtain that $s = -6 + 3 = -3$. Observe that so far we used only equations ① and ②. In order to see if we really found a solution, we must check if it is consistent with ③. Inserting our candidates for s and t, we find that $3 \cdot (-3) - 2 \cdot (-6) = 3$ which is inconsistent with ③. Therefore we conclude that there is no pair of real numbers s, t which satisfies all three equations ①–③ simultaneously, so the two lines do not intersect. □

Show that $L_3 \cap L_4 = \varnothing$.

Intersection of planes

Given two planes E_1 and E_2 in $\mathbb{R}^3$, there are two possibilities:

(a) The planes intersect. In this case, they necessarily intersect in infinitely many points. The intersection is either a line. In this case E_1 and E_2 are not parallel. Or the intersection is a plane. In this case $E_1 = E_2$.

(b) The planes do not intersect. In this case, the planes must be parallel and not equal.

Example 2.44. We consider the following four planes:

$$E_1 : x + y + 2z = 3, \quad E_2 : 2x + 2y + 4z = 3, \quad E_3 : 2x + 2y + 4z = 6, \quad E_4 : x + y - 2z = 5.$$

We will calculate their mutual intersections.

$\boxed{E_1 \cap E_2 = \varnothing}$

Proof. The set of all points $Q(x, y, z)$ which belong both to E_1 and E_2 is the set of all x, y, z which simultaneously satisfy

$$\begin{array}{cr} ① & x + y + 2z = 3, \\ ② & 2x + 2y + 4z = 3. \end{array}$$

Now clearly, if x, y, z satisfies ①, then it cannot satisfy ② (the right side would be 6). We can see this more formally if we solve ①, e.g., for x and then insert into ②. We obtain from ①: $x = 3 - y - 2z$. Inserting into ② leads to

$$3 = 2(3 - y - 2z) + 2y + 4z = 6,$$

which is absurd. □

Geometrically, this was to be expected. The normal vectors of the planes are $\vec{n}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ and $\vec{n}_2 = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}$ respectively. Since they are parallel, the planes are parallel and therefore they either are equal or they have empty intersection. Now we see that for instance $(3, 0, 0) \in E_1$ but $(3, 0, 0) \notin E_2$, so the planes cannot be equal. Therefore they have empty intersection.

$\boxed{E_1 \cap E_3 = E_1}$

Proof. The set of all points $Q(x, y, z)$ which belong both to E_1 and E_3 is the set of all x, y, z which simultaneously satisfy

$$\begin{array}{cr} ① & x + y + 2z = 3, \\ ② & 2x + 2y + 4z = 6. \end{array}$$

Clearly, both equations are equivalent: if x, y, z satisfies ①, then it also satisfies ② and vice versa. Therefore, $E_1 = E_3$. $\square$

$$\boxed{E_1 \cap E_4 = \left\{ \begin{pmatrix} 4 \\ 0 \\ -\frac{1}{2} \end{pmatrix} + t \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} : t \in \mathbb{R} \right\}.}$$

Proof. First, we notice that the normal vectors $\vec{n}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ and $\vec{n}_4 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ are not parallel, so we expect that the solution is a line in $\mathbb{R}^3$.
The set of all points $Q(x, y, z)$ which belong both to E_1 and E_4 is the set of all x, y, z which simultaneously satisfy

$$\begin{array}{cr} ① & x + y + 2z = 3, \\ ② & x + y - 2z = 5. \end{array}$$

Equation ① shows that $x = 3 - y - 2z$. Inserting into ② leads to $5 = 3 - y - 2z + y - 2z = 3 - 4z$, hence $z = -\frac{1}{2}$. Putting this into ①, we find that $x + y = 3 - 2z = 4$. So in summary, the intersection consists of all points (x, y, z) which satisfy

$$z = -\frac{1}{2}, \quad x = 4 - y \quad \text{with} \quad y \in \mathbb{R} \quad \text{arbitrary},$$

in other words,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 - y \\ y \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ -\frac{1}{2} \end{pmatrix} + \begin{pmatrix} -y \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ -\frac{1}{2} \end{pmatrix} + y \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{with} \quad y \in \mathbb{R} \quad \text{arbitrary}. \qquad \square$$

Intersection of a line with a plane

Finally we want to calculate the intersection of a plane E with a line L. There are three possibilities:

(a) The plane and the line intersect in exactly one point. This happens if and only if L is not parallel to E which is the case if and only if L is not perpendicular to the normal vector of E.

(b) The plane and the line do not intersect. In this case, the E and L must be parallel, that is, L must be perpendicular to the normal vector of E.

(c) The plane and the line intersect in infinitly many points. In this case, L lies in E, that is, E and L must be parallel and they must share at least one point.

As an example we calculate $E_1 \cap L_2$. Since L_2 is clearly not parallel to E_1, we expect that their intersection consists of exactly one point.

$$\boxed{E_1 \cap L_2 = \{(1/9, 2/9, 4/3)\}}$$

Proof. The set of all points $Q(x, y, z)$ which belong both to E_1 and L_2 is the set of all x, y, z which simultaneously satisfy

$$x + y + 2z = 3 \qquad \text{and} \qquad x = 2 + 2t,\ y = 4 + 4t, z = 7 + 6t \ \text{ for some } \ t \in \mathbb{R}.$$

Replacing the expression with t from L_2 into the equation of the plane E_1, we obtain the following euation for t:

$$3 = (2 + 2t) + (4 + 4t) + 2(7 + 6t) = 20 + 18t \qquad \Longrightarrow \qquad t = -17/18.$$

Replacing this t into the equation for L_2 gives the point of intersection $Q(1/9, 2/9, 4/3)$.
In order to check our result, we insert the coordinates in the equation for E_1 and obtain $x+y+2z = 1/9 + 2/9 + 2 \cdot 4/3 = 1/3 + 8/3 = 3$ which shows that $Q \in E_1$. $\square$

Intersection of several lines and planes

If we wanted to intersect for instance, 5 planes in $\mathbb{R}^3$, then we would have to solve a system of 5 equations for 3 unknowns. Or if we wanted to intersect 7 lines in $\mathbb{R}^3$, then we had to solve a system of 3 equations for 7 unknowns. If we solve them as we did here, the process could become quite messy. So the next chapter is devoted to find a systematic and efficient way to solve a system of m linear equations for n unknowns.

You should have understood

- what the possible geometric structures of intersections of lines and planes is and how this depends on their relative oriention,
- the interpretation of a linear system with three unknowns as the intersection of planes in $\mathbb{R}^3$,
- ...

You should now be able to

- calulate the intersection of lines and planes,
- ...

2.8 Exercises

1. Sean $P(2,3)$, $Q(-1,4)$ puntos en $\mathbb{R}^2$ y sea $\vec{v} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$ un vector en $\mathbb{R}^2$.

   (a) Calcule $\overrightarrow{PQ}$.
   (b) Calcule $\overline{PQ}$.
   (c) Calcule $\overrightarrow{PQ} + \vec{v}$.
   (d) Encuentre todos los vectores que son ortogonales a $\vec{v}$.

2. Sea $\vec{v} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \in \mathbb{R}^2$.

   (a) Encuentre todos los vectores unitarios cuya dirección es opuesta a la de $\vec{v}$.
   (b) Encuentre todos los vectores de longitud 3 que tienen la misma dirección que $\vec{v}$.
   (c) Encuentre todos los vectores que tienen la misma dirección que $\vec{v}$ y que tienen doble longitud de $\vec{v}$.
   (d) Encuentre todos los vectores con norma 2 que son ortogonales a $\vec{v}$.

3. Show that the following equations describe the same line:

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} : t \in \mathbb{R} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} 8 \\ 10 \\ 12 \end{pmatrix} : t \in \mathbb{R} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} -4 \\ -5 \\ -6 \end{pmatrix} : t \in \mathbb{R} \right\},$$

$$\left\{ \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix} + t \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} : t \in \mathbb{R} \right\}, \quad \frac{x-1}{4} = \frac{y-2}{5} = \frac{z-3}{6}, \quad \frac{x+3}{4} = \frac{y+3}{5} = \frac{z+3}{6}.$$

   Find at least one more vector equation and one more symmetric equation. Find at least two different parametric equations.

4. Para los siguientes vectores $\vec{u}$ y $\vec{v}$ decida si son ortogonales, paralelos o ninguno de los dos. Calcule el coseno del ángulo entre ellos. Si son paralelos, encuentre números reales λ y μ tales que $\vec{v} = \lambda\vec{u}$ y $\vec{u} = \mu\vec{v}$.

   (a) $\vec{v} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, $\vec{u} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}$,
   (b) $\vec{v} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$, $\vec{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$,
   (c) $\vec{v} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$, $\vec{u} = \begin{pmatrix} -8 \\ 6 \end{pmatrix}$,
   (d) $\vec{v} = \begin{pmatrix} -6 \\ 4 \end{pmatrix}$, $\vec{u} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$.

5. (a) Para las siguientes parejas $\vec{v}$ y $\vec{w}$ encuentre todos los $\alpha \in \mathbb{R}$ tal que $\vec{v}$ y $\vec{w}$ son paralelos:

   (i) $\vec{v} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} \alpha \\ -2 \end{pmatrix}$, (ii) $\vec{v} = \begin{pmatrix} 2 \\ \alpha \end{pmatrix}$, $\vec{w} = \begin{pmatrix} 1+\alpha \\ 2 \end{pmatrix}$, (iii) $\vec{v} = \begin{pmatrix} \alpha \\ 5 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} 1+\alpha \\ 2\alpha \end{pmatrix}$,

(b) Para las siguientes parejas $\vec{v}$ y $\vec{w}$ encuentre todos los $\alpha \in \mathbb{R}$ tal que $\vec{v}$ y $\vec{w}$ son perpendiculares:

(i) $\vec{v} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} \alpha \\ -2 \end{pmatrix}$, (ii) $\vec{v} = \begin{pmatrix} 2 \\ \alpha \end{pmatrix}$, $\vec{w} = \begin{pmatrix} \alpha \\ 2 \end{pmatrix}$, (iii) $\vec{v} = \begin{pmatrix} \alpha \\ 5 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} 1+\alpha \\ 2 \end{pmatrix}$.

6. Sean $\vec{a} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ y $\vec{b} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$.

(a) Calcule $\operatorname{proj}_{\vec{b}} \vec{a}$ y $\operatorname{proj}_{\vec{a}} \vec{b}$.

(b) Encuentre todos los vectors $\vec{v} \in \mathbb{R}^2$ tal que $\|\operatorname{proj}_{\vec{a}} \vec{v}\| = 0$. Describe este conjunto geométricamente.

(c) Encuentre todos los vectors $\vec{v} \in \mathbb{R}^2$ tal que $\|\operatorname{proj}_{\vec{a}} \vec{v}\| = 2$. Describe este conjunto geométricamente.

(d) ¿Existe un vector $\vec{x}$ tal que $\operatorname{proj}_{\vec{a}} \vec{x} \parallel \vec{b}$?
¿Existe un vector $\vec{x}$ tal que $\operatorname{proj}_{\vec{x}} \vec{a} \parallel \vec{b}$?

7. Sean $\vec{a}, \vec{b} \in \mathbb{R}^2$ con $\vec{a} \neq \vec{0}$.

(a) Demuestre que $\|\operatorname{proj}_{\vec{a}} \vec{b}\| \leq \|\vec{b}\|$.

(b) ¿Qué deben cumplir $\vec{a}$ y $\vec{b}$ para que $\|\operatorname{proj}_{\vec{a}} \vec{b}\| = \|\vec{b}\|$?

8. Sean $\vec{a}, \vec{b} \in \mathbb{R}^n$ con $\vec{b} \neq \vec{0}$.

(a) Demustre que $\|\operatorname{proj}_{\vec{b}} \vec{a}\| \leq \|\vec{a}\|$.

(b) Encuentre condiciones para $\vec{a}$ y $\vec{b}$ para que $\|\operatorname{proj}_{\vec{b}} \vec{a}\| = \|\vec{a}\|$.

(c) ¿Es cierto que $\|\operatorname{proj}_{\vec{b}} \vec{a}\| \leq \|\vec{b}\|$?

9. (a) Calcule el área del paralelogramo cuyos vértices adyacentes $A(1,2,3)$, $B(2,3,4)$, $C(-1,2,-5)$ son y calcule el cuarto vértice.

(b) Calcule el área del triángulo con los vértices. $A(1,2,3)$, $B(2,3,4)$, $C(-1,2,-5)$.

(c) Calcule el volumen del paralelepipedo determinado por los vectores
$\vec{u} = \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} -1 \\ 4 \\ 3 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix}$.

10. (a) Demuestre que no existe un elemento neutral para el producto cruz en $\mathbb{R}^3$. Es decir: Demuestre que no existe ningún vector $\vec{v} \in \mathbb{R}^3$ tal que $\vec{v} \times \vec{w} = \vec{w}$ para todo $\vec{w} \in \mathbb{R}^3$.

(b) Sea $\vec{w} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^3$.

(i) Encuentre todos los vectores $\vec{a}, \vec{b} \in \mathbb{R}^3$ tales que $\vec{a} \times \vec{w} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\vec{b} \times \vec{w} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$,

(ii) Encuentre todos los vectores $\vec{v} \in \mathbb{R}^3$ tales que $\langle \vec{v}, \vec{w} \rangle = 4$.

11. Dados líneas L_1 y L_2 y el punto P, determine:

- si L_1 y L_2 son paralelas,
- si L_1 y L_2 tienen un punto de intersección,
- si P pertenece a L_1 y/o a L_2,
- una recta paralela a L_2 que pase por P.

(a) $L_1 : \vec{r}(t) = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$, $\quad L_2 : \frac{x-3}{2} = \frac{y-2}{3} = \frac{z-1}{4}$, $\quad P(5, 2, 11)$.

(b) $L_1 : \vec{r}(t) = \begin{pmatrix} 2 \\ 1 \\ -7 \end{pmatrix} + t \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\quad L_2 : x = t + 1,\ y = 3t - 4,\ z = -t + 2$, $\quad P(5, 7, 2)$.

12. En $\mathbb{R}^3$ considere el plano E dado por $E : 3x - 2y + 4z = 16$.

(a) Encuentre por lo menos tres puntos que pertenecen a E.

(b) Encuentre un punto en E y dos vectores $\vec{v}$ y $\vec{w}$ en E que no son paralelos entre si.

(c) Encuentre un punto en E y un vector $\vec{n}$ que es ortogonal a E.

(d) Encuentre un punto en E y dos vectores $\vec{a}$ y $\vec{b}$ en E con $\vec{a} \perp \vec{b}$.

13. Para los puntos $P(1, 1, 1)$, $Q(1, 0, -1)$ y los siguientes planos E:

- Encuentre la ecuación del plano.
- Determine si P pertenece al plano.
- Encuentre una recta que esté ortogonal a E y que contenga al punto Q.

(i) E es el plano que contiene al punto $A(1, 0, 1)$ y es paralelo a los vectores $\vec{v} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ y $\vec{w} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$.

(ii) E es el plano que contiene los puntos $A(1, 0, 1)$, $B(2, 3, 4)$, $C(3, 2, 4)$.

(iii) E es el plano que contiene el punto $A(1, 0, 1)$ y es ortogonal al vector $\vec{n} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$.

14. Considere el plano $E : 2x - y + 3z = 9$ y la recta $L : x = 3t + 1,\ y = -2t + 3,\ z = 5t$.

(a) Encuentre $E \cap L$.

(b) Encuentre una recta G que no interseque ni al plano E ni a la recta L. Pruebe su afirmación. Cúantas rectas con esta propiedad hay?

15. En $\mathbb{R}^3$ considere el plano E dado por $E : 3x - 2y + 4z = 16$.

(a) Demuestre que los vectores $\vec{a} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}$ y $\vec{v} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$ son paralelos al plano E.

(b) Encuentre números $\lambda, \mu \in \mathbb{R}$ tal que $\lambda\vec{a} + \mu\vec{b} = \vec{v}$.

(c) Demuestre que el vector $\vec{c} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ no es paralelo al plano E y encuentre vectores $c_\parallel$ y $c_\perp$ tal que $c_\parallel$ es paralelo a E, $c_\perp$ es ortogonal a E y $c = c_\parallel + c_\perp$.

16. Sea E un plano en $\mathbb{R}^2$ y sean $\vec{a}$, $\vec{b}$ vectores paralelos a E. Demuestre que para todo $\lambda, \mu \in \mathbb{R}$, el vector $\lambda\vec{a} + \mu\vec{b}$ es paralelo al plano.

17. Sea V un espacio vectorial. Demuestre lo siguiente:

(a) El elemento neutral es único.

(b) $0v = \mathbb{O}$ para todo $v \in V$.

(c) $\lambda\mathbb{O} = \mathbb{O}$ para todo $\lambda \in \mathbb{R}$.

(d) Dado $v \in V$, su inverso $\widetilde{v}$ es único.

(e) Dado $v \in V$, su inverso $\widetilde{v}$ cumple $\widetilde{v} = (-1)v$.

18. De todos los siguientes conjuntos decida si es un espacio vectorial con su suma y producto usual.

(a) $V = \left\{ \begin{pmatrix} a \\ a \end{pmatrix} : a \in \mathbb{R} \right\}$,

(b) $V = \left\{ \begin{pmatrix} a \\ a^2 \end{pmatrix} : a \in \mathbb{R} \right\}$,

(c) V es el conjunto de todas las funciones continuas $\mathbb{R} \to \mathbb{R}$.

(d) V es el conjunto de todas las funciones f continuas $\mathbb{R} \to \mathbb{R}$ con $f(4) = 0$.

(e) V es el conjunto de todas las funciones f continuas $\mathbb{R} \to \mathbb{R}$ con $f(4) = 1$.

Chapter 3

Linear Systems and Matrices

In this chapter we want to explore how to solve linear systems systematically and efficiently. To this end we will work extensively with matrices We will extend the concept of determinant to $n \times n$ matrices and we will see that as in the 2×2 case, a linear system has a unique solution if and only if its determinant is different from 0.

3.1 Linear systems and Gauß and Gauß-Jordan elimination

We start with a linear system as in Definition 1.7

$$\begin{aligned} a_{11}x_1 + \ a_{12}x_2 + \ \cdots + a_{1n}x_n &= \ b_1 \\ a_{21}x_1 + \ a_{22}x_2 + \ \cdots + a_{2n}x_n &= \ b_2 \\ \vdots \qquad\qquad \vdots \qquad\quad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{3.1}$$

Recall that the system is called consistent if it has at least one solution; otherwise it is called inconsistent. According to (1.3) and (1.4) its associated coefficient matrix and augmented coefficient matrices are

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{3.2}$$

and

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{array}\right). \tag{3.3}$$

Definition 3.1. The set of all matrices with m rows and n columns is denoted by $M(m \times n)$. If we want to emphasise that the matrix has only real entries, then we write $M(m \times n, \mathbb{R})$ or $M_{\mathbb{R}}(m \times n)$.

Another frequently used notations are $M_{m\times n}$. A matrix A is called a *square matrix* if its number of rows is equal to its number of columns.

In order to solve (3.1), we could use the first equation, solve for x_1 and insert this in all the other equations. This gives us a new system with $m-1$ equations for $n-1$ unknowns. Then we solve the next equation for x_2, insert it in the other equations, and we continue like this until we have only one equation left. This of course will fail if for example $a_{11}=0$ because in this case we cannot solve the first equation for x_1. We could save our algorithm by saying: we solve the first equation for the first unknown whose coefficient is different from 0 (or we could take an equation where the coefficient of x_1 is different from 0 and declare this one to be our first equation. After all, we can order the equations as we please). Even with this modification, the process of solving and replacing is error prone.

Another idea is to manipulate the equations. The question is: Which changes to the equations are allowed without changing the information contained in the system? Or, in more mathematical terms, what changes to the equation result in an equivalent system? Here we call two systems equivalent if they have the same set of solutions.

We can check if the new system is equivalent to the original one, if there is a way to restore the original one.

For example, if we exchange the first and the second row, then nothing really happened and we end up with an equivalent system. We can come back to the original equation by simply exchanging again the first and the second row.
If we multiply both sides of first equation on both sides by some factor, let's say, by 2, then again nothing changes. Assume for example that the first equation is $x+3y=7$. If we multiply both sides by two, we obtain $2x+6y=14$. Clearly, if a pair (x,y) satisfies the first equation, then it satisfies also the second one an vice versa. Given the new equation $3x+6y=14$, we can easily restore the old one by simply dividing both sides by 2.

If we take an equation and multiply both of its sides by 0, then we destroy information because we end up with $0=0$ and there is no way to get back the information that was stored in the original equation. So this is not an allowed operation.

Show that squaring both sides of an equation in general does not give an equivalent equation. Are there cases, when it does?

Squaring an equation or taking the logarithm on both sides or other such things usually are not interesting to us because the resulting equation will no longer be a linear equation.

It is more or less clear that the following are the "allowed" operations which do not alter the information contained in a given linear system.
In the following table we write R_j for the jth row. We describe the operation in words, we describe it in shorthand notation and we give its inverse operation (the one that allows us to get back to the unchanged system).

① Swap two rows.	$R_j \leftrightarrow R_k$	$R_j \leftrightarrow R_k$
② Multiply row j by some $\lambda \in \mathbb{R}\setminus\{0\}$	$\lambda R_k \to R_j$	$\frac{1}{\lambda}R_j \to R_j$
③ Replace row k by the sum of row k and a multiple of R_j and keep row j unchanged.	$R_k + \lambda R_j \to R_k$	$R_k - \lambda R_j \to R_k$

Show that instead of the operation 2 we could have taken 2': $R_k \to R_k + R_j$ because 2 can be written and as a composition of various operations of the form 1, 2' and 3. Show how this can be done.

Show that in reality 1 is not necessary since it can be achieved by a composition of operations of the form 2' and 3 (or 2' and 3). Show how this can be done.

From now on, if we speak about row operations, we always refer to the “allowed” operations 1, 2, 3.

Let us see in an example how this works.

Example 3.2.

$$\left.\begin{aligned} x_1 + \;\; x_2 - x_3 &= 1 \\ 2x_1 + 3x_2 + x_3 &= 3 \\ 4x_2 + x_3 &= 7 \end{aligned}\right\} \xrightarrow{R_2 - 2R_2 \to R_2} \left\{\begin{aligned} x_1 + x_2 \;\; - \;\; x_3 &= 1 \\ x_2 \;\; + 3x_3 &= 1 \\ 4x_2 + \;\; x_3 &= 7 \end{aligned}\right\} \xrightarrow{R_3 - 4R_2 \to R_3} \left\{\begin{aligned} x_1 + x_2 - \quad x_3 &= 1 \\ x_2 + \;\; 3x_3 &= 1 \\ -\, 11x_3 &= 3 \end{aligned}\right\}$$

$$\xrightarrow{R_3 - 4R_2 \to R_3} \left\{\begin{aligned} x_1 + x_2 - \;\; x_3 &= \quad 1 \\ x_2 + 3x_3 &= \quad 1 \\ x_3 &= -3/11. \end{aligned}\right.$$

Here we can stop because it is already quite easy to read off the solution. Proceeding from the bottom to the top, we obtain

$$x_3 = -3/11, \qquad x_2 = 1 - 3x_3 = 20/11, \qquad x_1 = 1 + x_3 - x_2 = -12/11.$$

Note that we could continued our row manipulations

$$\cdots \longrightarrow \left\{\begin{aligned} x_1 + x_2 - \quad x_3 &= 1 \\ x_2 + \;\; 3x_3 &= 1 \\ -\, 11x_3 &= 3 \end{aligned}\right\} \xrightarrow{-1/11 R_3 \to R_3} \left\{\begin{aligned} x_1 + x_2 - \;\; x_3 &= \quad 1 \\ x_2 + 3x_3 &= \quad 1 \\ x_3 &= -3/11 \end{aligned}\right\}$$

$$\xrightarrow{R_2 - 3R_3 \to R_2} \left\{\begin{aligned} x_1 + x_2 - x_3 &= \quad 1 \\ x_2 \qquad &= 20/11 \\ x_3 &= -3/11 \end{aligned}\right\} \xrightarrow{R_1 - 1/11 R_3 \to R_1} \left\{\begin{aligned} x_1 + x_2 \qquad &= \;\; 8/11 \\ x_2 + \qquad &= 20/11 \\ x_3 &= -3/11 \end{aligned}\right\}$$

$$\xrightarrow{R_1 - R_2 \to R_1} \left\{\begin{aligned} x_1 + \qquad &= -12/11 \\ x_2 \quad &= \;\; 20/11 \\ x_3 &= \;\; -3/11 \end{aligned}\right.$$

Our strategy was to apply manipulations that successively eliminate the unknowns in the lower equations and we aimed to get to a form of the system of equations where the last one contains the least number of unknowns.

It is important to note that there are infinitely many different routes how to get to the final result, but usually some are quicker than others.

Let us analyse what we did. We looked at the coefficients and we applied transformations such that the coefficients become 0 because by doing so, we eliminate the unknowns from the equations. So in the example above we could just as well delete all the x_j, keep only the augmented coefficient matrix and do the line operations in the matrix. Of course, we have to remember that the numbers in the first columns are the coefficients of x_1, those in the second column are the coefficients of x_2, etc. Then our calculations are translated into the following:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & 3 \\ 0 & 4 & 1 & 7 \end{array}\right) \xrightarrow{R_2 - 2R_2 \to R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 4 & 1 & 7 \end{array}\right) \xrightarrow{R_3 - 4R_2 \to R_3} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -11 & 3 \end{array}\right)$$

$$\xrightarrow{1/11 R_3 \to R_3} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & -3/11 \end{array}\right).$$

If we translate this back into a linear system, we get

$$\begin{aligned} x_1 + x_2 + \ \ x_3 &= \quad 1 \\ x_2 + 3x_3 &= \quad 3 \\ x_3 &= -3/11 \end{aligned}$$

which can be easily solved from the bottom up.
What we did with the matrix was exactly the same as we did with the system of equations but it looks much tidier since we do not have to write down the unknowns all the time.
If we want to solve a linear we write it as an augmented matrix and then we perform row operations until we reach a "nice" form where we can read off the solutions if there are any.
But what is a "nice" form? Remember that if a coefficient is 0, then the corresponding unknown does not show up in the equation.

- In the last equation we want as few unknowns as possible and we want to keep the last unknowns. So as last row we want one that has only zeros in it or one that starts with zeros, until finally we get a non-zero number say in column k. This non-zero number can always be made equal to 1 by dividing the row by it. Now we know how the unknowns $x_k, \dots, x_n$ are related. Note that all the other unknowns $x_1, \dots, x_{k-1}$ have "disappeared" from the equation since their coefficients are 0.

  If $k = n$ as in our example above, then we there is only one solution for x_n.

- The second to last row could also be a zero row or it should start with zeros until we get to a column, say column l, with non-zero entry rich we always can make equal to 1. This column should be to the left of the column k (that is we want $l < k$). Because now we can use what we know from the last row about the unknowns $x_k, \dots, x_n$ to say something about the unknowns $x_l, \dots, x_{k-1}$.

- We continue like this until all rows are as we want them.

Note that the form of such a "nice" looks a bit like it had a triangle consisting of only zeros in its lower left region. There may be zero in the upper right part. If a matrix has the form we just described, we say it is in *row echelon form*. Let us give a precise definition.

Definition 3.3 (Row echelon form). We say that a matrix $A \in M(m \times n)$ is in *row echelon form* if:

- All its zero rows are the last rows.
- The first no-zero entry in a row is 1. It is called the *pivot* of the row.
- The pivot of any row is strictly to the right of the row above.

Definition 3.4 (Reduced row echelon form). We say that a matrix $A \in M(m \times n)$ is in *reduced row echelon form* if:

- A is in row echelon form.
- All the entries in in A which are on top of a pivot are equal to 0.

Let us quickly see some examples.

Examples 3.5.

(a) The following matrices are in reduced row echelon form:

$$\begin{pmatrix} 1 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 6 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(b) The following matrices are in row echelon form but not in reduced row echelon form:

$$\begin{pmatrix} 1 & 6 & 3 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 6 & 3 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 6 & 3 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

(c) The following matrices are not row echelon form:

$$\begin{pmatrix} 1 & 6 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 6 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 6 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 3 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

- Say why the matrices in (b) are not in reduced row echelon form and use row operations to transform them into a matrix in reduced row echelon form.
- Say why the matrices in (c) are not in row echelon form and use row operations to transform them into a matrix in row echelon form. Transform them further to obtain a matrix in reduced row echelon form.

qu:3:01

Question 3.1

If we swap to lines in a matrix this corresponds to writing down the given equations in a different order. What is the effect on a linear system if we swap two columns?

Remember that if we translate a linear system to an augmented coefficient matrix $(A|b)$, perform the row operations to arrive at (reduced) row echelon form $(A'|b')$, and translate back to a linear system, then this new system contains exactly the same information as the original one but it is "tidied up" and it is easy to determine its solution.
The natural question now is: Can we always transform a matrix into one in (reduced) row echelon form? The answer is that this is always possible and that we can given an algorithm how to do so.

Gauß elimination. Let $A \in M(m \times n)$. The Gauß elimination is an algorithm that transforms A into a row echelon form. The steps are as follows:

- Find the first column which does not consist entirely of zeros. Swap rows until the entry in that column in the first row is different from zero.
- Multiply the first row by an appropriate number so that its first non-zero entry is 1.
- Use the first row to eliminate all coefficients below its pivot.
- Now our matrix looks like

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ \vdots & & \vdots & 0 & & & \\ \vdots & & \vdots & \vdots & & \text{A'} & \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix}$$

  where $*$ are arbitrary numbers and A' is a matrix with less columns than A and $m-1$ columns. Now repeat the process for A'. Note that in doing say the first columns to not change since we are only manipulating zeros.

Gauß-Jordan elimination. Let $A \in M(m \times n)$. The Gauß-Jordan elimination is an algorithm that transforms A into a reduced row echelon form. The steps are as follows:

- Use the Gauß elimination to obtain a row echelon form of A.
- Use the pivots to eliminate the non-zero coefficients which are columns above a pivot.

Of course, if we do a reduction by hand, then we do not have the follow the steps of the algorithm strictly if it makes calculations easier. However, these algorithms always work and therefore can be programmed so that a computer can perform them.

Definition 3.6. Two $m \times n$ matrices A and B are called *row equivalent* if there are row operations that transform A into B. (Clearly then B can be transformed by row operations into A.)

Let A be $m \times n$ matrix.

- A can be transformed into infinitely many different row echelon forms.
- There is only one reduced row echelon form that A can be transformed into.

Prove the assertion above.

Before we give examples, we note that if we have transformed an augmented matrix into a row echelon form, then we can immediately say how many solutions the corresponding linear system has.

Theorem 3.7. *Let $(A|b)$ be the augmented coefficient matrix of a linear $m \times n$ system and let $(A'|b')$ be a row reduced form.*

(i) If there is a line of the form $(0 \cdots 0|\beta)$ with $\beta \neq 0$, then the system has no solution.

(ii) If there is no line of the form $(0 \cdots 0|\beta)$ with $\beta \neq 0$, then one of the following holds:

(a) If there is a pivot in every column then the system has exactly one solution.

(b) If there is a column with without a pivot, then the system has infinitely many solutions.

Proof. (i) If $(A'|b')$ has a row of the form $(0 \cdots 0|\beta)$ with $\beta \neq 0$, then the corresponding equation is $0x_1 + \cdots + 0x_n = \beta$ which clearly cannot be satisfied.

(ii) No assume that $(A'|b')$ has no row of the form $(0 \cdots 0|\beta)$ with $\beta \neq 0$. In case (a), the transformed matrix is then of the form

$$\left(\begin{array}{ccccc|c} 1 & a'_{12} & a'_{13} & \cdots & a'_{1n} & b'_1 \\ 0 & 1 & a'_{23} & \cdots & a'_{2n} & b'_2 \\ \vdots & & & \cdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 1 & a'_{(n-1)n} & b'_{n-1} \\ 0 & \cdots & \cdots & \cdots & 1 & b'_n \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \\ \vdots & & & & \vdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & 0 \end{array}\right)$$

Note that the last zero rows appear only if $n < m$. This system clearly has the unique solution

$$x_n = b'_n, \quad x_{n-1} = b'_{n-1} - a_{(n-1)n}x_n, \quad \ldots, \quad x_1 = b'_1 - a_{1n}x_n - \cdots - a_{12}x_2.$$

In case (b), the transformed matrix is then of the form

$$\left(\begin{array}{cccccccccccccccc|c}
0 & \cdots & 0 & 1 & * & \cdots & * & * & * & * & * & \cdots & * & \cdots & * & & b_1' \\
0 & \cdots & & & & \cdots & 0 & 1 & * & \cdots & * & \cdots & * & \cdots & * & & b_2' \\
0 & \cdots & & & & & & & \cdots & 0 & 1 & * & \cdots & \cdots & * & & b_3' \\
\vdots & & & & & & & & & & & & & & & \cdots & \vdots \\
& & & & & & & & & & & & & & & \cdots & \\
0 & \cdots & & & & & & & & & \cdots & 0 & 1 & * & \cdots & * & b_k' \\
0 & \cdots & & & & & & & & & & & & & \cdots & 0 & 0 \\
\vdots & & & & & & & & & & & & & & & \vdots & 0 \\
0 & \cdots & & & & & & & & & & & & & \cdots & 0 & 0
\end{array}\right)$$

where the stars stand for numbers. (If we continue the reduction until we get to the reduced row echelon form, then the number over the 1's must be zeros.) Note that we can choose the unknowns which correspond to the columns without a pivot arbitrarily. They are called the *free variables*. The unknowns which correspond to the columns with pivots can then always be chosen in a unique way such that the system is satisfied. □

We will come back to this theorem later on page **??** (the theorem is stated again in the coloured box).
From the above theorem we get as an immediate consequence the following.

Theorem 3.8. *A linear system has either no, exactly one or infinitely many solutions.*

Now let us see some examples.

Example 3.9 (Example with a unique solution (no free variables)). We consider the linear system

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 12, \\ -x_1 + 2x_2 + 3x_3 &= 15, \\ 3x_1 - 3x_3 &= 1. \end{aligned} \tag{3.4}$$

Solution. We form the augmented matrix and perform row reduction.

$$\left(\begin{array}{rrr|r} 2 & 3 & 1 & 12 \\ -1 & 2 & 3 & 15 \\ 3 & 0 & -3 & 1 \end{array}\right) \xrightarrow{R_1+2R_2 \to R_1} \left(\begin{array}{rrr|r} 0 & 7 & 7 & 42 \\ -1 & 2 & 3 & 15 \\ 3 & 0 & -3 & 1 \end{array}\right) \xrightarrow{R_3+3R_2 \to R_3} \left(\begin{array}{rrr|r} 0 & 7 & 7 & 42 \\ -1 & 2 & 3 & 15 \\ 0 & 6 & 8 & 46 \end{array}\right)$$

$$\xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{array}{rrr|r} -1 & 2 & 3 & 15 \\ 0 & 7 & 7 & 42 \\ 0 & 6 & 8 & 46 \end{array}\right) \xrightarrow[\frac{1}{7}R_2 \to R_2]{-R_1 \to R_1} \left(\begin{array}{rrr|r} 1 & -2 & -3 & -15 \\ 0 & 1 & 1 & 6 \\ 0 & 6 & 8 & 46 \end{array}\right)$$

$$\xrightarrow{R_3-6R_2 \to R_3} \left(\begin{array}{rrr|r} 1 & -2 & -3 & -15 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 2 & 10 \end{array}\right) \xrightarrow{\frac{1}{2}R_3 \to R_3} \left(\begin{array}{rrr|r} 1 & -2 & -3 & -15 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 1 & 5 \end{array}\right).$$

This shows that the system (3.4) is equivalent to the system

$$\begin{aligned} x_1 - 2x_2 - 3x_3 &= -15, \\ x_2 + \ x_3 &= \ \ \ 6, \\ x_3 &= \ \ \ 5 \end{aligned} \tag{3.5}$$

whose solution is easy to write down:

$$x_3 = 5, \quad x_2 = 6 - x_3 = 1, \quad x_1 = -15 + 2x_2 + 3x_3 = 2. \qquad \diamond$$

Remark. If we continue the reduction process until we reach the reduced row echelon form, then we obtain

$$\ldots \longrightarrow \left(\begin{array}{ccc|c} 1 & -2 & -3 & -15 \\ 0 & 1 & 1 & 6 \\ 0 & 0 & 1 & 5 \end{array}\right) \xrightarrow{R_2 - R_3 \to R_2} \left(\begin{array}{ccc|c} 1 & -2 & -3 & -15 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 5 \end{array}\right) \xrightarrow{R_1 + 3R_3 \to R_1} \left(\begin{array}{ccc|c} 1 & -2 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 5 \end{array}\right)$$

$$\xrightarrow{R_1 + 2R_2 \to R_1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 5 \end{array}\right).$$

Therefore the system (3.4) is equivalent to the system

$$\begin{aligned} x_1 \qquad\qquad &= 2, \\ x_2 \qquad &= 1, \\ x_3 &= 5. \end{aligned}$$

whose solution can be read off immediately to be

$$x_3 = 5, \quad x_2 = 1, \quad x_1 = 2.$$

Example 3.10 (Example with two free variables). We consider the linear system

$$\begin{aligned} 3x_1 - 2x_2 + 3x_3 + 3x_4 &= 3, \\ 2x_1 + 6x_2 + 2x_3 - 9x_4 &= 2, \\ x_1 + 2x_3 + \ x_3 - 3x_4 &= 1. \end{aligned} \tag{3.6}$$

Solution. We form the augmented matrix and perform row reduction.

$$\left(\begin{array}{cccc|c} 3 & -2 & 3 & 3 & 3 \\ 2 & 6 & 2 & -9 & 2 \\ 1 & 2 & 1 & -3 & 1 \end{array}\right) \xrightarrow{R_2 - 2R_1 \to R_2} \left(\begin{array}{cccc|c} 3 & -2 & 3 & 3 & 3 \\ 0 & 2 & 0 & -3 & 0 \\ 1 & 2 & 1 & -3 & 1 \end{array}\right) \xrightarrow{R_1 - 3R_3 \to R_1} \left(\begin{array}{cccc|c} 0 & -8 & 0 & 12 & 0 \\ 0 & 2 & 0 & -3 & 0 \\ 1 & 2 & 1 & -3 & 1 \end{array}\right)$$

$$\xrightarrow{R_1 \leftrightarrow R_3} \left(\begin{array}{cccc|c} 1 & 2 & 1 & -3 & 1 \\ 0 & 2 & 0 & -3 & 0 \\ 0 & -8 & 0 & 12 & 0 \end{array}\right) \xrightarrow{R_3 + 4R_2 \to R_3} \left(\begin{array}{cccc|c} 1 & 2 & 1 & -3 & 1 \\ 0 & 2 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{R_1 - R_2 \to R_1} \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The 3rd and the 4th column do not have pivots and we see that the system (3.6) is equivalent to the system

$$\begin{aligned} x_1 \qquad - x_3 \qquad &= 1, \\ x_2 \qquad + x_4 &= 0. \end{aligned}$$

Clearly we can choose x_3 and x_4 (the unknowns corresponding to the columns without a pivot) arbitrarily. We will always be able to then choose x_1 and x_2 such that the system is satisfied. In order to make it clear that x_3 and x_4 are our free variables, we give call them $x_3 = t$ and $x_4 = s$. Then every solution of the system (3.6) is of the form

$$x_1 = 1 + t, \quad x_2 = -s, \quad x_3 = t, \quad x_4 = s, \quad \text{for arbitrary} \quad s, t \in \mathbb{R}.$$ ◇

In *vector form* we can write the solution as follows. A tuple (x_1, x_2, x_3, x_4) is a solution of (3.6) if and only if the corresponding vector is of the form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1+t \\ -s \\ t \\ s \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} \quad \text{for some} \quad s, t \in \mathbb{R}.$$

Remark. Geometrically, the set of all solutions is a plane in $\mathbb{R}^4$.

Example 3.11 (Example with no solution). We consider the linear system

$$\begin{aligned} 2x_1 + x_2 - x_3 &= 7, \\ 3x_1 + 2x_2 - 2x_3 &= 7, \\ -x_1 + 3x_2 - 3x_3 &= 2. \end{aligned} \tag{3.7}$$

Solution. We form the augmented matrix and perform row reduction.

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 7 \\ 3 & 2 & -2 & 7 \\ -1 & 3 & -3 & 2 \end{array}\right) \xrightarrow{R_1+2R_3 \to R_1} \left(\begin{array}{ccc|c} 0 & 7 & -7 & 11 \\ 3 & 2 & -2 & 7 \\ -1 & 3 & -3 & 2 \end{array}\right) \xrightarrow{R_2+3R_3 \to R_2} \left(\begin{array}{ccc|c} 0 & 7 & -7 & 11 \\ 0 & 11 & -11 & 13 \\ -1 & 3 & -3 & 2 \end{array}\right)$$

$$\xrightarrow{R_1 \leftrightarrow R_3} \left(\begin{array}{ccc|c} -1 & 3 & -3 & 2 \\ 0 & 11 & -11 & 13 \\ 0 & 7 & -7 & 11 \end{array}\right) \xrightarrow{11R_3-7R_2 \to R_3} \left(\begin{array}{ccc|c} -1 & 3 & -3 & 2 \\ 0 & 11 & -11 & 13 \\ 0 & 0 & 0 & 30 \end{array}\right).$$

The last line tells us immediately that the system (3.8) has no solution because there is no choice of x_1, x_2, x_3 such that $0x_1 + 0x_2 + 0x_3 = 30$. ◇

You should now have understood

- when two linear systems are equivalent,
- what row operations transform a given system into an equivalent one and why this is so,
- when a matrix is in row echelon and a reduced row echelon form,
- why the linear system associated to a matrix in (reduced) echelon form is easy to solve,
- what the Gauß- and Gauß-Jordan elimination does and why it always works,
- why a given matrix can be transformed into may different row echelon forms, but in only one reduced row echelon form,
- why a linear system always has either no, exactly one or infinitely many solutions,
- ...

You should now be able to

- identify if a matrix is in row echelon or a reduced row echelon form,
- use the Gauß- or Gauß-Jordan elimination to solve linear systems,
- ...

3.2 Homogeneous linear systems

In this short section we deal with the special case homogeneous linear systems. Recall that a linear system (3.1) is called homogeneous if $b_1 = \cdots = b_n = 0$. Such a system has always at least one solution, the so-called *trivial solution* $x_1 = \cdots = x_n = 0$. This also clear from Theorem 3.7 since no matter what row operations we perform, the right side will always remain equal to 0. Note that if we perform Gauß or Gauß-Jordan elimination, there is no need to write down the right hand side since it always will be 0.

If we adapt Theorem 3.7 to the special case of a homogeneous system, we obtain the following.

Theorem 3.12. *Let A be the coefficient matrix of a homogeneous linear $m \times n$ system and let A' be a row reduced form.*

(a) If there is a pivot in every column then the system has exactly one solution, namely the trivial solution.

(b) If there is a column with without a pivot, then the system has infinitely many solutions.

Corollary 3.13. *A homogeneous linear system has either no or infinitely many solutions.*

Let us see an example.

Example 3.14 (Example with infinitely many solutions). We consider the linear system

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 0, \\ 2x_1 + 3x_2 - 2x_3 &= 0, \\ 3x_1 - x_2 - 3x_3 &= 0. \end{aligned} \tag{3.8}$$

Solution. We form the augmented matrix and perform row reduction.

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & -2 \\ 3 & -1 & -3 \end{pmatrix} \xrightarrow{R_2 - 2R_1 \to R_2} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 3 & -1 & -3 \end{pmatrix} \xrightarrow{R_3 - 3R_1 \to R_3} \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 0 & -7 & 0 \end{pmatrix}$$

$$\xrightarrow[\text{the 2nd column}]{\text{use } R_2 \text{to clear}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{-R_2 \to R_2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We see that the third variable is free, so we set $x_3 = t$. The solution is

$$x_1 = t, \quad x_2 = 0, \quad x_3 = t \quad \text{for} \quad t \in \mathbb{R}.$$

or in vector form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{for} \quad t \in \mathbb{R}.$$ ◇

You should now have understood

- why a homogeneous linear system always has either one or infinitely many solutions,
- ...

You should now be able to

- use the Gauß- or Gauß-Jordan elimination to solve homogeneous linear systems,
- ...

3.3 Vectors and matrices; matrices and linear systems

So far we were given a linear system with a specific right hand side and we asked ourselves which x_j do have to feed into the system in order to obtain the given right hand side. Problems of this type are called *inverse problems* since we are given an output (the right hand of the system; the "state" that we want to achieve) and the problem is to find a suitable input in order to obtain the desired output.

Now we change the perspective a bit and we ask ourselves: If we put certain $x_1, \dots, x_n$ into the system, what do we get on the right hand side? To investigate this question, it is very useful to write the system (3.1) in a short form. First note that we can view it as an equality of the two vectors with m components:

$$\begin{pmatrix} a_{11}x_1 + & a_{12}x_2 + & \cdots + a_{1n}x_n \\ a_{21}x_1 + & a_{22}x_2 + & \cdots + a_{2n}x_n \\ \vdots & & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 + & \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}. \tag{3.9}$$

Let A be the coefficient matrix and $\vec{x}$ the vector whose components are $x_1, \ldots, x_n$. Then we write the left hand side of (3.9) as

$$A\vec{x} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + & a_{12}x_2 + & \cdots + a_{1n}x_n \\ a_{21}x_1 + & a_{22}x_2 + & \cdots + a_{2n}x_n \\ \vdots & & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 + & \cdots + a_{mn}x_n \end{pmatrix}. \tag{3.10}$$

With this notation, the linear system (3.1) can be written very short as

$$A\vec{x} = \vec{b}$$

with $\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$. A way to remember the formula is that we "multiply each row of the matrix by the column vector", so we calculate *"row by column"*.

Definition 3.15. The formula in (3.10) is called the multiplication of a matrix and a vector.

A $m \times n$ matrix A takes a vector with n components and gives us back a vector with m components.

Observe that something like $\vec{x}A$ does not make sense!

Remark 3.16. Formula (3.10) can also be interpreted like this. If A is an $m \times n$ matrix and $\vec{x}$ is a vector in $\mathbb{R}^n$, then $A\vec{x}$ is the vector in $\mathbb{R}^m$ which is the sum of the columns of A weighted with coefficients given by $\vec{x}$ since

$$\begin{aligned} A\vec{x} &= \begin{pmatrix} a_{11}x_1 + & a_{12}x_2 + & \cdots + a_{1n}x_n \\ a_{21}x_1 + & a_{22}x_2 + & \cdots + a_{2n}x_n \\ \vdots & & \vdots \\ a_{m1}x_1 + & a_{m2}x_2 + & \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{pmatrix} + \cdots + \begin{pmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{pmatrix} \\ &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \end{aligned} \tag{3.11}$$

Remark 3.17. Recall that $\vec{e}_j$ is the vector which has a 1 is its jth component and has zeros everywhere else. Formula (3.10) shows that for every $j = 1, \dots, n$

$$A\vec{e}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = j\text{th column of } \ A. \tag{3.12}$$

Let us prove some easy properties.

Proposition 3.18. *Let A be an $m \times n$ matrix, $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then*

(i) $A(c\vec{x}) = cA\vec{x}$,

(ii) $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$,

(iii) $A\vec{0} = \vec{0}$.

Proof. The proofs are not difficult. They follow by using the definitions and carry out some straightforward calculations as follows.

(i)

$$A(c\vec{x}) = A\begin{pmatrix} cx_1 \\ \vdots \\ cx_n \end{pmatrix} = \begin{pmatrix} a_{11}cx_1 + & \cdots + a_{1n}cx_n \\ a_{21}cx_1 + & \cdots + a_{2n}cx_n \\ \vdots & \vdots \\ a_{m1}cx_1 + & \cdots + a_{mn}cx_n \end{pmatrix} = c\begin{pmatrix} a_{11}x_1 + & \cdots + a_{1n}x_n \\ a_{21}x_1 + & \cdots + a_{2n}x_n \\ \vdots & \vdots \\ a_{m1}x_1 + & \cdots + a_{mn}x_n \end{pmatrix} = cA\vec{x}.$$

(ii)

$$A(\vec{x} + \vec{y}) = A\begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} = \begin{pmatrix} a_{11}(x_1 + y_1) + & \cdots + a_{1n}(x_n + y_n) \\ a_{21}(x_1 + y_1) + & \cdots + a_{2n}(x_n + y_n) \\ \vdots & \vdots \\ a_{m1}(x_1 + y_1) + & \cdots + a_{mn}(x_n + y_n) \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}x_1 + & \cdots + a_{1n}x_n \\ a_{21}x_1 + & \cdots + a_{2n}x_n \\ \vdots & \vdots \\ a_{m1}x_1 + & \cdots + a_{mn}x_n \end{pmatrix} + \begin{pmatrix} a_{11}y_1 + & \cdots + a_{1n}y_n \\ a_{21}y_1 + & \cdots + a_{2n}y_n \\ \vdots & \vdots \\ a_{m1}y_1 + & \cdots + a_{mn}y_n \end{pmatrix} = A\vec{x} + A\vec{y}.$$

(iii) To show that $A\vec{0} = \vec{0}$, we could simply do the calculation (which is very easy!) or we can use (i):

$$A\vec{0} = A(0\vec{0}) = 0A\vec{0} - \vec{0}.$$

□

Note that in (iii) the $\vec{0}$ on the left hand side is the zero vector in $\mathbb{R}^n$ whereas the $\vec{0}$ on the right hand side is the zero vector in $\mathbb{R}^m$.

Proposition 3.18 gives an important insight in the structure of solutions of linear systems.

Theorem 3.19. (i) *Let $\vec{x}$ and $\vec{y}$ be solutions of the linear system* (3.1)*. Then $\vec{x}-\vec{y}$ is a solution of the associated homogeneous linear system.*

(ii) *Let $\vec{x}$ be a solution of the linear system* (3.1) *and let $\vec{z}$ be a solution of the associated homogeneous linear system. Then $\vec{x}+\vec{z}$ is solution of the system* (3.1)*.*

Proof. Assume that $\vec{x}$ and $\vec{y}$ are solutions of the (3.1), that is

$$A\vec{x} = \vec{b} \quad \text{and} \quad A\vec{y} = \vec{b}.$$

By Proposition 3.18 (i) and (ii) we have

$$A(\vec{x}-\vec{y}) = A\vec{x} + A(-\vec{y}) = A\vec{x} - A\vec{y} = \vec{b} - \vec{b} = \vec{0}$$

which shows that $\vec{x}-\vec{y}$ solves the homogeneous system $A\vec{v} = \vec{0}$ and thereby proves (i).
In order to show (ii), we proceed similarly. If $\vec{x}$ solves the inhomogeneous system (3.1) and $\vec{z}$ solves the associated homogeneous system, then

$$A\vec{x} = \vec{b} \quad \text{and} \quad A\vec{y} = \vec{0}.$$

Now (ii) follows from

$$A(\vec{x}+\vec{z}) = A\vec{x} + A\vec{z} = \vec{b} + \vec{0} = \vec{b}.$$ □

Corollary 3.20. *Let $\vec{x}$ be an arbitrary solution of the inhomogeneous system* (3.1)*. Then the set of all solutions of* (3.1) *is*

$$\{\vec{x}+\vec{z} : \vec{z} \ \textit{ is solution of the associated homogeneous system}\}.$$

This means that in order to find all solutions of an inhomogeneous system it suffices to find one particular solution and all solutions of the corresponding homogeneous system.

We will show later that the set of all solutions of a homogeneous system is a vector space. When you study the set of all solutions of linear differential equations, you will encounter the same structure.

You should now have understood

- that an $m \times n$ matrix can be viewed as an operator that takes vectors in $\mathbb{R}^n$ and spits out a vector in $\mathbb{R}^m$,
- the structure of the set of all solutions of a given linear system,
- ...

You should now be able to

- calculate expressions like $A\vec{x}$,
- relate the solutions of an inhomogeneous system with those of the corresponding homogeneous one,
- ...

3.4 Matrices as functions from $\mathbb{R}^n$ to $\mathbb{R}^m$; composition of matrices

In the previous section we saw that a matrix $A \in M(m \times n)$ takes a vector $\vec{x} \in \mathbb{R}^n$ and gives us back a vector $A\vec{x}$ in $\mathbb{R}^m$. This allows us to view A as a function from $\mathbb{R}^n$ to $\mathbb{R}^m$, and therefore we can define what the sum and composition of two matrices is. Before we do this, let us see a few examples of such matrices. We take examples 2×2 because these can be drawn easily.

Example 3.21. Let us consider $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. This defines a function T_A from $\mathbb{R}^2$ to $\mathbb{R}^2$ by

$$T_A : \mathbb{R}^2 \to \mathbb{R}^2, \quad T_A\vec{x} = A\vec{x}.$$

We write T_A to denote the function induced by A, but sometimes we will also write simply $A : \mathbb{R}^2 \to \mathbb{R}^2$ when it is clear that we consider the matrix A as a function.
We can calculate easily

$$T_A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad T_A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \qquad \text{in general} \quad T_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}.$$

So we see that T_A represents the reflection of a vector $\vec{x}$ about the x-axis.

Example 3.22. Let us consider $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. This defines a function T_B from $\mathbb{R}^2$ to $\mathbb{R}^2$ by

$$T_B : \mathbb{R}^2 \to \mathbb{R}^2, \quad T_B\vec{x} = B\vec{x}.$$

We can calculate easily

$$T_B \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad T_B \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \text{in general} \quad T_B \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}.$$

So we see that T_B represents the projection of a vector $\vec{x}$ onto the y-axis.

Example 3.23. Let us consider $C = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. This defines a function T_C from $\mathbb{R}^2$ to $\mathbb{R}^2$ by

$$T_C : \mathbb{R}^2 \to \mathbb{R}^2, \quad T_C\vec{x} = C\vec{x}.$$

We can calculate easily

$$T_C \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad T_C \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \qquad \text{in general} \quad T_C \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}.$$

So we see that T_C represents the rotation of a vector $\vec{x}$ about $90°$ counterclockwise.

Just as with other functions, we can sum them or compose them. Remember from your calculus classes, that functions are summed "pointwise". That means, if we have two functions $f, g : \mathbb{R} \to \mathbb{R}$, then we their *sum* $f + g$ is a new function which is defined by

$$f + g : \mathbb{R} \to \mathbb{R}, \quad (f + g)(x) = f(x) + g(x). \tag{3.13}$$

The multiplication of a function f with a number c gives the new function cf defined by

$$cf : \mathbb{R} \to \mathbb{R}, \quad (cf)(x) = c(f(x)). \tag{3.14}$$

The composition of functions if defined as

$$f \circ g : \mathbb{R} \to \mathbb{R}, \quad (f \circ g)(x) = f(g(x)). \tag{3.15}$$

Matrix sum

Let us see how this looks like in the case of matrices. Let A and B be matrices. First note that they both must depart from the same space $\mathbb{R}^n$ because we want to apply them to the same $\vec{x}$, that is, both $A\vec{x}$ and $B\vec{x}$ must be defined. Therefore A and B must have the same number of rows because we want to be able to sum $A\vec{x}$ and $B\vec{x}$. So let $A, B \in M(m \times n)$ and let $\vec{x} \in \mathbb{R}$. Then, by definition of the sum of two functions, we have

$$\begin{aligned}
(A+B)\vec{x} = A\vec{x} + B\vec{x} &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\
&= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} + \begin{pmatrix} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n \\ \vdots \\ b_{m1}x_1 + b_{m2}x_2 + \cdots + b_{mn}x_n \end{pmatrix} \\
&= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + b_{m1}x_1 + b_{m2}x_2 + \cdots + b_{mn}x_n \end{pmatrix} \\
&= \begin{pmatrix} (a_{11} + b_{11})x_1 + (a_{12} + b_{12})x_2 + \cdots + (a_{1n} + b_{mn})x_n \\ (a_{21} + b_{11})x_1 + (a_{22} + b_{12})x_2 + \cdots + (a_{2n} + b_{mn})x_n \\ \vdots \\ (a_{m1} + b_{11})x_1 + (a_{m2} + b_{12})x_2 + \cdots + (a_{mn} + b_{mn})x_n \end{pmatrix} \\
&= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{mn} \\ a_{21} + b_{11} & a_{22} + b_{12} & \cdots & a_{2n} + b_{mn} \\ \vdots & \vdots & \vdots & \\ a_{m1} + b_{11} & a_{m2} + b_{12} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}
\end{aligned}$$

We see that $A + B$ is again a matrix of the same size and that the components of this new matrix are just the sum of the corresponding components of the matrices A and B.

Proposition 3.24. *Let $A, B, C \in M(m \times n)$ let $\mathbb{O}$ be the matrix whose entries are all 0 and let $\lambda, \mu \in \mathbb{R}$. Moreover, let $\widetilde{A}$ be the matrix whose entries are the negative entries of A. Then the following is true.*

(i) **Associativity of the matrix sum:** $(A + B) + C = A + (B + C)$.

(ii) **Commutativity of the matrix sum:** $A + B = B + A$.

(iii) **Additive identity:** $A + \mathbb{O} = A$.

(iv) **Additive inverse** $A + \widetilde{A} = \mathbb{O}$.

(v) $1A = A$.

(vi) $(\lambda + \mu)A = \lambda A + \mu A$ *and* $\lambda(A + B) = \lambda A + \lambda B$.

Proof. The claims of the proposition can be proved by straightforward calculations. □

Prove Proposition 3.24.

From the proposition we obtain immediately the following theorem.

Theorem 3.25. *$M(m \times n)$ is a vector space*

Multiplication of a matrix by a scalar

Now let c be a number and let $A \in M(m \times n)$. Then we have

$$(cA)\vec{x} = c(A\vec{x}) = c\left[\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right] = c\begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}$$

$$= \begin{pmatrix} ca_{11}x_1 + \cdots + ca_{1n}x_n \\ \vdots \\ ca_{m1}x_1 + \cdots + ca_{mn}x_n \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

We see that cA is again a matrix and that the components of this new matrix are just the product of the corresponding components of the matrix A with c.

Composition of two matrices

Now let us calculate the composition of two matrices. This is also called the product of the matrices. Assume we have $A \in M(m \times n)$ and we want to calculate AB for some matrix. Note that A describes a function from $\mathbb{R}^n \to \mathbb{R}^m$. In order for AB to make sense, we need that B goes from some $\mathbb{R}^k$ to $\mathbb{R}^n$, that means that $B \in M(n \times k)$. The resulting function AB will then be a map from $\mathbb{R}^k$ to $\mathbb{R}^m$.

So let $B \in M(n \times k)$. Then, by definition of the composition of two functions, we have for every $\vec{x} \in \mathbb{R}^k$

$$(AB)\vec{x} = A(B\vec{x}) = A \left[\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} \right] = A \begin{pmatrix} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k \\ \vdots \\ b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}[b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k] + a_{12}[b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k] + \cdots + a_{1n}[b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k] \\ a_{21}[b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k] + a_{22}[b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k] + \cdots + a_{2n}[b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k] \\ \vdots \\ a_{m1}[b_{11}x_1 + b_{12}x_2 + \cdots + b_{1k}x_k] + a_{m2}[b_{21}x_1 + b_{22}x_2 + \cdots + b_{2k}x_k] + \cdots + a_{mn}[b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nk}x_k] \end{pmatrix}$$

$$= \begin{pmatrix} [a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}]x_1 + [a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1n}b_{n2}]x_2 + \cdots + [a_{11}b_{1k} + a_{12}b_{2k} + \cdots + a_{1n}b_{nk}]x_k \\ [a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1}]x_1 + [a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2n}b_{n2}]x_2 + \cdots + [a_{21}b_{1k} + a_{22}b_{2k} + \cdots + a_{2n}b_{nk}]x_k \\ \vdots \\ [a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1}]x_1 + [a_{m1}b_{12} + a_{m2}b_{22} + \cdots + a_{mn}b_{n2}]x_2 + \cdots + [a_{m1}b_{1k} + a_{m2}b_{2k} + \cdots + a_{mn}b_{nk}]x_k \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1k} + a_{12}b_{2k} + \cdots + a_{1n}b_{nk} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1k} + a_{22}b_{2k} + \cdots + a_{2n}b_{nk} \\ \vdots & & & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1} & a_{m1}b_{12} + a_{m2}b_{22} + \cdots + a_{mn}b_{n2} & \cdots & a_{m1}b_{1k} + a_{m2}b_{2k} + \cdots + a_{mn}b_{nk} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix}$$

We see that AB is a matrix of the size $m \times k$ as was to be expected since the composition function goes from $\mathbb{R}^k$ to $\mathbb{R}^m$. The component jj, ℓ of the new matrix (the entry in lines j and column ℓ) is

$$c_{j\ell} = \sum_{h=1}^{n} a_{jh} b_{h\ell}.$$

So in order to calculate this entry we need from A only its jth row and from B we only need its ℓth column and we multiply them component by component. You can memorise this again as "row by column", more precisely:

$$c_{j\ell} = \text{component in row } j \text{ and column } \ell \text{ of } AB = (\text{row } j \text{ of } A) \times (\text{column } \ell \text{ of } B) \tag{3.16}$$

as in the case of multiplication of a vector by a matrix. Actually, a vector in $\mathbb{R}^n$ can be seen a $n \times 1$ matrix (a matrix with n rows and one column).

Example 3.26. Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 8 & 6 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 7 & 1 & 2 & 3 \\ -2 & 0 & 1 & 4 \\ 2 & 6 & -3 & 0 \end{pmatrix}$. Then

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 & 3 \\ 8 & 6 & 4 \end{pmatrix} \begin{pmatrix} 7 & 1 & 2 & 3 \\ -2 & 0 & 1 & 4 \\ 2 & 6 & -3 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1\cdot 7 + 2\cdot 2 + 3\cdot 2 & 1\cdot 1 + 2\cdot 0 + 3\cdot 6 & 1\cdot 2 + 2\cdot 1 + 3\cdot -3 & 1\cdot 3 + 2\cdot 4 + 3\cdot 0 \\ 8\cdot 7 + 6\cdot 2 + 4\cdot 2 & 8\cdot 1 + 6\cdot 0 + 4\cdot 6 & 8\cdot 2 + 6\cdot 1 + 4\cdot -3 & 8\cdot 3 + 6\cdot 4 + 4\cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} 17 & 19 & -5 & 11 \\ 76 & 32 & 10 & 48 \end{pmatrix} \end{aligned}$$

Les us see some properties of the algebraic operations for matrices that we just introduced.

Proposition 3.27. *Let $A, B, C \in M(m \times n)$ and let $R \in M(m \times k)$ and $S \in M(k \times \ell)$. Then the following is true.*

(i) **Associativity of the matrix product:** $(AB)C = A(BC)$.

(ii) **Distributivity:** $A(B + C) = AB + AC$ *and* $(B + C)A = BA + CA$.

Proof. The claims of the proposition can be proved by straightforward calculations. □

Prove Proposition 3.27.

Very important remark.

The matrix multiplication is not commutative, that is, in general

$$AB \neq BA.$$

That matrix multiplication cannot be commutative is more or less clear since it is the composition of two functions (think of functions that you know from your calculus classes. For example, it does make a difference if you first square a variable and then take the arctan or if you first calculate is arctan and then take the square).

Let us see an example. Let B be the matrix from Example 3.22 and C be the matrix from Example 3.23. Recall that B represents the orthogonal projection onto the y-axis and that C represents counterclockwise rotation by 90°. If we take $\vec{e}_1$ (the unit vector in x-direction), and we first rotate and then project, we get the vector $\vec{e}_2$. If however we project first and rotate then, we get $\vec{0}$. That means, $BC\vec{e}_1 \neq CB\vec{e}_1$, therefore $BC \neq CB$. Let us calculate the products:

$$BC = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad \text{first rotation, then projection,}$$

$$CB = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \qquad \text{first projection, then rotation.}$$

Let A be the matrix from Example 3.21, B be the matrix from Example 3.22 and C the matrix from Example 3.23. Verify that $AB \neq BA$ and $AC \neq CA$ and make understand this result geometrically by following for example where the unit vectors get mapped to.

Note also that usually, when AB is defined, the expression BA is not defined because in general the number of columns of B will be different from the number of rows of A.

We finish this section with the definition of the so-called identity matrix.

Definition 3.28. Let $n \in \mathbb{N}$. Then the $n \times n$ identity matrix is the matrix which has 1s on its diagonal an d has zero everywhere else:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

As notation for the identity matrix, the following symbols are used in the literature: E_n, id_n, Id_n, I_n, $\mathbf{1_n}$, $\mathbb{1}_n$. The subscript n can be omitted if it is clear.

Remark 3.29. It can be easily verified that

$$A\,\mathrm{id}_n = A, \quad \mathrm{id}_n\, B = B, \quad \mathrm{id}_n\, \vec{x} = \vec{x}$$

for every $A \in M(m \times n)$, for every $B \in M(n \times k)$ and for every $\vec{x} \in \mathbb{R}^n$.

You should now have understood

- what the sum and the composition of two matrices is and where the formulas come from,
- why the composition of matrices is not commutative,
- that $M(m \times n)$ is a vector space,
- ...

You should now be able to

- calculate the sum and product (composition) of two matrices,
- ...

Index