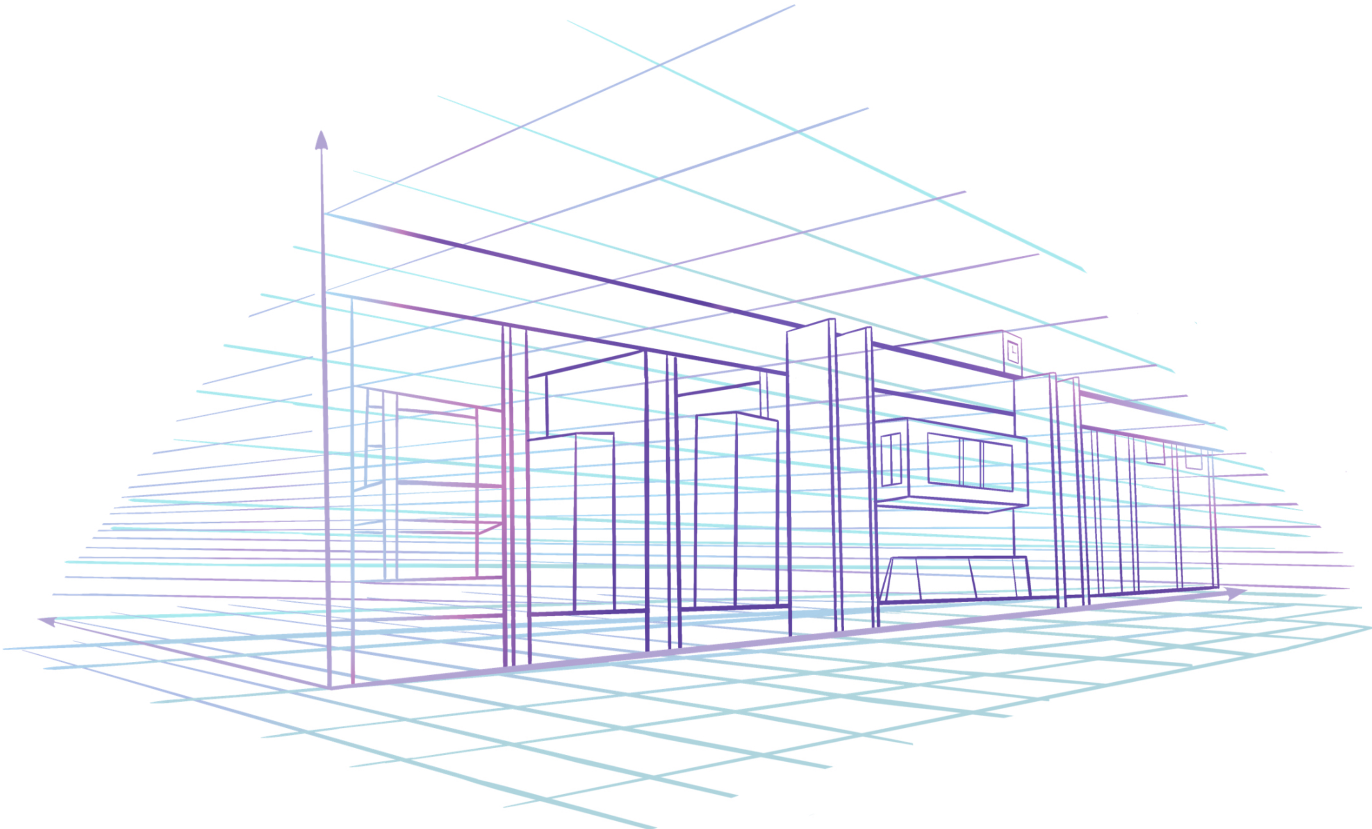
Linear Algebra

線形代数学

G30 Program, Nagoya University



Lecturer: Henrik Bachmann
Written by: Henrik Bachmann, Tue Tai Nguyen, Vic Austen
Illustrations: Hannah Patricia Halim

Version 13 (January 25, 2024)

Contents

About this course

These notes are based on the Linear Algebra I & II lectures of the G30 Program at Nagoya University given in the fall and spring semesters of 2019 - 2023. The concept of this course initially evolved from the Linear Algebra I & II given by Erik Darpö in the years before, which was based on the book [B]. The course is anticipated for two semesters (Linear Algebra I - fall semester, Linear Algebra II - spring semester) with 16 lectures of 90 minutes each. This includes a midterm and final exam each semester. In the year 2020, this course was (due to the pandemic) given online, and therefore, recordings of the lectures exist. The content and notation in the lecture videos are similar, but not exactly the same, as in these notes. However, the videos can be used for students who missed a lecture or who want to recall the content on their own. A possible schedule, together with links to the corresponding sections & videos, is given as follows:

Linear Algebra I

Week	**Content**	Section	Lecture video
01	Introduction & Linear systems	Introduction Chapter 1	Video 1
02	Matrices and vectors	Chapter 2	Video 2
03	Sets and functions	Chapter 3	Video 3
04	Linear maps	Chapter 4	Video 4
05	Linear maps in geometry	Chapter 5	Video 5
06	Matrix multiplication	Chapter 6	Video 6
07	*Midterm Exam*	Chapter 14	
08	The inverse of a linear map	Chapter 7	Video 7
09	Subspaces, Kernel & Image	Chapter 8	Video 8
10	Subspaces, Kernel & Image II	Chapter 8	Video 9
11	Linear independence & Bases I	Chapter 9 Chapter 10	Video 10
12	Bases II & Dimension	Chapter 10	part of Video 10
13	Coordinates & Orthogonal bases	Chapter 11 Chapter 12	Video 11
14	Orthogonal bases & The Gram-Schmidt algorithm	Chapter 12	Video 12
15	Orthogonal projection, Least square approximation	Chapter 13	Video 13
16	*Final Exam*	Chapter 14	

Linear Algebra II

Week	**Content**	Section	Lecture video
01	Recall Linear Algebra I & Overview	Introduction	Video 1.1
02	Vector spaces	Chapter 14	Video 1.2
03	Linear maps	Chapter 15	Video 2
04	The matrix of a linear map	Chapter 16	Video 3
05	Determinants & Mathematical induction	Chapter 17	Video 4
06	Properties of the determinant I	Chapter 17	Video 5
07	*Midterm exam*		
08	Properties of the determinant II	Chapter 17	Video 6
09	Eigenvalues and eigenvectors I	Chapter 18	Video 7
10	Eigenvalues and eigenvectors II	Chapter 18	Video 8
11	Eigenvalues and eigenvectors III (Spectral Theorem)	Chapter 18	Video 9
12	Applications	Chapter 19	Video 10
13	Continuous dynamical systems	Chapter 20	Video 11
14	Linear differential equations I	Chapter 20	Video 12
15	Linear differential equations II	Chapter 20	Video 13
16	*Final Exam*		Video 14 (Review)

A more visually pleasant overview of this course and its timeline is given by the Linear Algebra road map on the next page.


LINEAR ALGEBRA Road Map
LINEAR ALGEBRA I
ようこそ!
START!
2x₁ + x₂ = 1
-x₁ + 3x₂ = -2
R^n
A∪B
LINEAR SYSTEM MATRICES, VECTORS
SETS & MAPS
-Week- 3
Week 1-2
A∩B
LINEAR MAPS
Week 4-6
F: R^n → R^m
① F(u+v)=F(u)+F(v)
② F(λu)=λF(u)
F(v)=Av.
Week 9-10
[F]=A∈R^{m×n}
SUBSPACES
U={(x₁, x₂)∈R² | x₁+x₂=0}
IMAGE
KERNEL
MIDTERM EXAM
W8. INVERSE of linear maps
F: R^n → R^m
Ker F
im F
W11-12
Bases, linear independence
Winter vacation
ORTHOGONALITY
week 14-16
u·v=0
Week 13.
COORDINATES
FINAL EXAM
LINEAR ALGEBRA II
VECTOR SPACES
WEEK 1-2
GOLDEN week
F(R,R)
P
C
R^n
C^∞(R,R)
D: f ↦ f'
LINEAR MAPS
W2-3
F: V → W
im(F) ⊂ W
Ker(F) ⊂ V
Spring break
MIDTERM
WEEK 4-6
determinants
F_n = F_{n-1} + F_{n-2}
Eigenvalues
Eigenvectors
F(v) = λv
Week 8-10
Dynamical Systems
x_{t+1} = Ax_t
f''(t)+f(t)=2t
θ
FINISH
FINALS
Week 11
Applications
19.9%
4.2%
10.1%
LINEAR DIFFERENTIAL EQUATIONS

Linear Algebra I

Introduction

In the realm of numbers and vectors we wade,
Linear Algebra, where scientific problems are laid.
Automotive, physics, chemistry, and bio too,
All fields that this essential course will accrue.

Linear systems we'll study, matrices we'll discern,
Through sets and functions, much knowledge we'll earn.
Linear maps transform, in spaces they unwind,
Subspaces and kernels, in invertible maps they're confined.

Independence and bases, dimensions we'll chart,
Each concept a masterpiece, a mathematical art.
As coordinates shift, new perspectives arise,
With Gram-Schmidt's process, orthonormal bases materialize.

Projections orthogonal, and least squares approximation,
Tools for data fitting, across each student's vocation.
"Linear Algebra II," a deeper journey we embark,
Where vector spaces and linear maps leave their mark.

Determinants and eigenvalues, the story further unfurls,
With linear differential equations, like precious pearls.
This course, a bridge from theory to practical narration,
Equipping students with robust mathematical foundation.

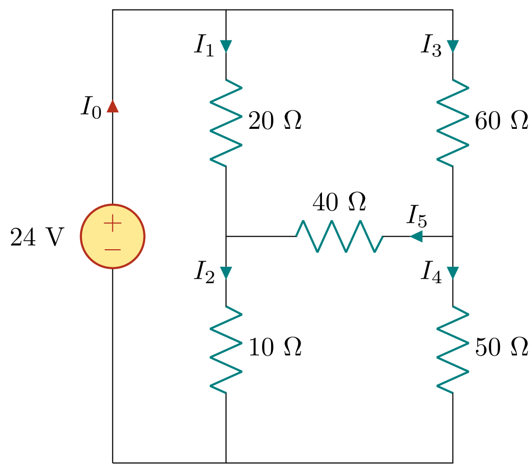
– ChatGPT

Linear Algebra, a fundamental branch of mathematics, offers essential tools and language to articulate and solve a broad array of problems across diverse scientific disciplines. For those studying automotive engineering, physics, chemistry, and bio-agriculture, the principles embedded within this field become invaluable in comprehending and modeling complex systems pertinent to their areas of study.

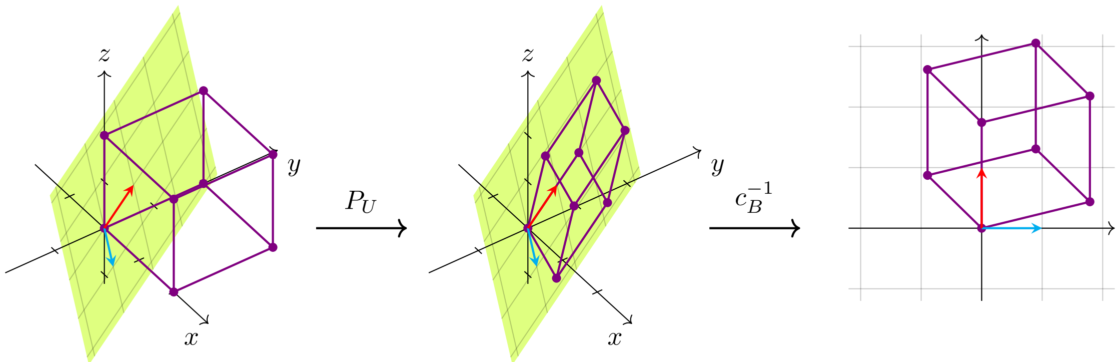
In this course, we will study vector spaces and linear transformations between these spaces, often referred to as "flat spaces" in a geometric sense. In these spaces, the basic operations of addition and scalar multiplication behave in a "linear" way, meaning they satisfy properties like distributivity, commutativity, and associativity. Unlike curved spaces, which are studied in differential geometry, linear spaces are characterized by the absence of curvature, making them easier to analyze and understand.

The study of linear spaces is crucial for various applications in science and engineering, as they often serve as good approximations for more complex structures. Whether it is solving systems of equations, analyzing data sets, or transforming shapes in computer graphics, the principles of linear algebra are foundational.

For example, in electrical circuits, solving linear systems of equations is often essential for applying Kirchhoff's laws, which govern the conservation of charge and energy in the circuit. Kirchhoff's current law states that the sum of currents entering a junction must equal the sum of currents leaving it, while Kirchhoff's voltage law states that the sum of the voltages around any closed loop in a circuit must be zero. These laws can be translated into a system of linear equations where the unknowns are the currents or voltages in the circuit componentes. In the diagram on the right, you see an example of an electrical circuit with unknown currents $I_0, \dots, I_5$. In Example 7, we show how to calculate these by solving a linear system.



After studying linear systems and matrices and vectors, we will talk about linear maps. Often these maps have some kind of geometric interpretation, and their applications are endless. For example, if you want to program a 3D-game engine, you need to project 3-dimensional objects onto a 2-dimensional space (your monitor). This will be done by a linear map. Later in this course we will illustrate this by explaining how a 3-dimensional cube can be displayed in the plane after choosing a certain viewing angle and we will give a explanation of the following illustration:



1

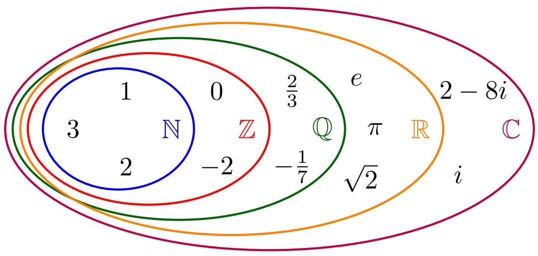
Linear systems

In linear algebra, linear systems play a crucial role in understanding and solving various real-world problems. A linear system is a collection of linear equations involving the same variables. These equations represent relationships between variables and can be used to model various scenarios from diverse fields such as physics, economics, and engineering.

We will denote the set of **real numbers** by $\mathbb{R}$. $\mathbb{R}$ contains all numbers usually considered in high school, such as $1, -1, 0, 2, 3, \frac{3}{8}, \pi, \sqrt{2}, e, \dots$. There are rigorous definitions of the real numbers, which would be part of a pure mathematics lecture[1]. But in this course, we just assume that they exist and that everyone is familiar with them.

♡ *Remark.* Even though we will focus on real numbers most of the time in this course, we will also mention the usual notations for certain subsets of them.

The **natural numbers** $\mathbb{N}$ consists of the number $1, 2, 3, \dots$. Allowing 0 and negative numbers leads to the integers $\mathbb{Z} = \{0, -1, 1, 2, -2, \dots\}$. The set of fractions $\frac{a}{b}$ with $a \in \mathbb{Z}, b \in \mathbb{N}$ are called **rational numbers**, denoted by $\mathbb{Q}$. But not all numbers appearing "naturally" can be written as fractions. For example, the diagonal of a square of side-length one is, by Pythagoras, the positive solution of $x^2 = 1 + 1 = 2$, i.e. $x = \sqrt{2}$. But one can show that $\sqrt{2}$ is not rational.



The "completion" of rational numbers, which can be think of as filling up the missing gaps, leads then to the real numbers $\mathbb{R}$ mentioned above. The story does not end here, since there is an even bigger class of numbers which are often of interest. For example, one might be interested in solutions of the equation $x^2 = -1$, which does not exist if one just allows real numbers. This leads to the notion of **complex numbers** $\mathbb{C}$, which are numbers of the form $a + bi$ with $a, b \in \mathbb{R}$ and i being a new symbol satisfying $i^2 = -1$.

A linear equation can be written in the form:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, a_2, \dots, a_n \in \mathbb{R}$ and $b \in \mathbb{R}$ are constants, and $x_1, x_2, \dots, x_n$ are variables.

Let us first consider a real-life situation that can be described by a linear system.

Example 1 (Cakes) A chocolate-obsessed patisserie only sells two types of cake: chocolate tart (henceforth, 'tarts') and chocolate cake (henceforth, 'cake'). To create one cake, 2 bars of chocolate, 3 tablespoons of sugar, and four eggs are required. For one tart, 3 bars of chocolate,

[1] See for example `https://en.wikipedia.org/wiki/Construction_of_the_real_numbers` for an overview of the "construction" of real numbers.

5 tablespoons of sugar, and 7 eggs are required.
On one particular day, the patisserie used 77 bars of chocolate, 124 tablespoons of sugar, and 137 eggs was used up. We want to find out how many cakes and tarts they made.
If we denote the number of cakes by x_1, and the number of tarts by x_2, then we can write three equations relating the number of cakes and tarts with the number of bars of chocolate, tablespoons of sugar, and eggs used:

$$\begin{aligned} 2x_1 + 3x_2 &= 77 \\ 3x_1 + 5x_2 &= 124 \\ 4x_1 + 7x_2 &= 171 \end{aligned}$$

To solve such simultaneous equations, one can use either substitution or elimination method. In this example, let us use the elimination method (which entails 'eliminating' variables).
First, we multiply the first equation by 2 to obtain:

$$4x_1 + 6x_2 = 154$$

By subtracting this from the third equation, we eliminate the variable x:

$$4x_1 + 7x_2 - (4x_1 + 6x_2) = 171 - 154 \Longrightarrow x_2 = 17$$

We can then find the value of x_1 by plugging $x_2 = 17$ into the first equation:

$$2x_1 + 3 \cdot 17 = 77 \Longrightarrow x_1 = \frac{1}{2}(77 - 51) = 13$$

As such, by solving the system of linear equations, we know that on that day, the patisserie made $x_1 = 13$ cakes and $x_2 = 17$ tarts.

No one reading these notes will probably own a patisserie or will be interested in solving problems as in the above example. But we will see that linear systems, and in particular the study of their solutions, arise in various serious real life applications. First let us fix the notation as follows:

Definition 1.1 (i) For real numbers $a_1, a_2, \ldots, a_n, b \in \mathbb{R}$ an equation of the form

$$a_1 x_1 + \cdots + a_n x_n = b$$

is called a **linear equation**.
(ii) A finite collection of linear equations is called a **linear system**.
(iii) A **solution of a linear system** is a simultaneous solution for all of its equations.

<u>Goal:</u> Given a linear system we want to find <u>all</u> of its solutions.

While the method used in Example 1 is fine for smaller systems, larger, more complicated systems would be a nightmare to solve using such crude methods. As such, we must develop a more systematic method of solving such linear systems. One way is to add multiples of one equation to another one, or multiply an equation with a non-zero number. By doing this correctly, a new linear system with clearer solutions may be obtained from the original linear system.

Example 2 (Unique solution) Consider and solve the following linear system

$$2\begin{cases} x_1 + 3x_2 = 1 \\ -2x_1 + x_2 = 2 \end{cases} \Rightarrow \quad \begin{matrix} \\ \frac{1}{7} \end{matrix}\begin{cases} x_1 + 3x_2 = 1 \\ \qquad 7x_2 = 4 \end{cases} \Rightarrow \quad \begin{matrix} \\ -3 \end{matrix}\begin{cases} x_1 + 3x_2 = 1 \\ \qquad x_2 = \frac{4}{7} \end{cases} \Rightarrow \quad \begin{cases} x_1 \qquad = -\frac{5}{7} \\ \qquad x_2 = \frac{4}{7} \end{cases}$$

Is $x_1 = -\frac{5}{7}$, $x_2 = \frac{4}{7}$ a solution to the original linear system? The answer is yes because the operations work also in reverse.

$$\begin{cases} x_1 = -\frac{5}{7} \\ x_2 = \frac{4}{7} \end{cases} \Rightarrow \begin{cases} x_1 + 3x_2 = 1 \\ \quad\;\; x_2 = \frac{4}{7} \end{cases} \Rightarrow \begin{cases} x_1 + 3x_2 = 1 \\ \quad\; 7x_2 = 4 \end{cases} \Rightarrow \begin{cases} \quad\; x_1 + 3x_2 = 1 \\ -2x_1 + \;\; x_2 = 2 \end{cases}$$

Therefore,

$$\begin{cases} \quad\; x_1 + 3x_2 = 1 \\ -2x_1 + \;\; x_2 = 2 \end{cases} \iff \begin{cases} x_1 = -\frac{5}{7} \\ x_2 = \frac{4}{7} \end{cases}$$

This linear system has exactly one solution.

Remark. For anyone who is unfamiliar with the notations of implication "$\Rightarrow$" and equivalence "$\Leftrightarrow$", consider two statements denoted by p and q. In mathematics, statements are sentences or equations or inequalities, which are true or false, no ambivalence. Given the statements p and q, we can form new statements $p \Rightarrow q$ and $p \Leftrightarrow q$.

When p is true, $p \Rightarrow q$ is true if q is true and false if q is false. Otherwise, when p is false, $p \Rightarrow q$ is true regardless of q. The statement $p \Rightarrow q$ can be read as "the statement p implies the statement q" or "if p, then q". The meaning is that the truth of p leads to the truth of q, which makes $p \Rightarrow q$ true. For example, if $x = 2$, then $x^2 = x \cdot x = 2 \cdot 2 = 4$. For that, we can write $x = 2 \Rightarrow x^2 = 4$ and this statement is true. For another example, the statement $x < 0 \Rightarrow x^3 > 0$ is false because if $x = -1 < 0$, then $x^3 = (-1)^3 = -1 < 0$. We can sometimes see true statements $p \Rightarrow q$ where p is always false such as $x^2 < 0 \Rightarrow x = 100$. In these cases, we say that they are vacuously true.

The statement $p \Leftrightarrow q$ is a combination of $p \Rightarrow q$ and $q \Rightarrow p$. This statement is true when p, q are both true or both false. Therefore, the truth of $p \Leftrightarrow q$ means that p is equivalent to q and vice versa. It can be read as "The statement p is equivalent to the statement q" or "p if and only if q". For example, $x = 1 \Leftrightarrow x + 1 = 2$ is true because if $x = 1$ then $x + 1 = 1 + 1 = 2$, and if $x + 1 = 2$ then $x = (x + 1) - 1 = 2 - 1 = 1$. For another example, the statement $x = 2 \Leftrightarrow x^2 = 4$ is false because when $x = -2$, we have $x^2 = (-2)^2 = 4$ and hence, $x^2 = 4$ does not imply $x = 2$.

In Examples 1 and 2, the linear system has a unique solution. However, not all linear systems have unique solutions; some have no solutions (for example, the system $x_1 = 2$, $x_1 = 3$ has no solution), and some have infinitely many solutions. An example of such a linear system is given below.

Example 3 (Infinitely many solutions) Consider the following linear system:

$$\begin{cases} x_1 - 9x_2 - 3x_3 + x_4 = 4 \\ 3x_1 - 2x_2 + x_3 - 2x_4 = 2 \\ 2x_1 + 7x_2 + 4x_3 - 3x_4 = -2 \\ \qquad 25x_2 + 9x_3 - 2x_4 = -4 \end{cases} \Leftrightarrow \begin{cases} x_1 - 9x_2 - 3x_3 + x_4 = 4 \\ \qquad 25x_2 + 10x_3 - 5x_4 = -10 \\ \qquad 25x_2 + 10x_3 - 5x_4 = -10 \\ \qquad 25x_2 + 9x_3 - 2x_4 = -4 \end{cases}$$

$$\Leftrightarrow \begin{cases} x_1 - 9x_2 - 3x_3 + x_4 = 4 \\ \qquad 25x_2 + 10x_3 - 5x_4 = -10 \\ \qquad\qquad 0 = 0 \\ \qquad\qquad x_3 - 3x_4 = 6 \end{cases} \Leftrightarrow \begin{cases} x_1 - 9x_2 \qquad - 8x_4 = -14 \\ \qquad 25x_2 \qquad + 25x_4 = 50 \\ \qquad\qquad -x_3 + 3x_4 = 6 \end{cases}$$

row-reduced echelon form

$$\Leftrightarrow \quad 9 \begin{cases} x_1 - 9x_2 - 8x_4 = -14 \\ x_2 + x_4 = 2 \\ x_3 - 3x_4 = -6 \end{cases} \Leftrightarrow \boxed{\begin{cases} x_1 + x_4 = 4 \\ x_2 + x_4 = 2 \\ x_3 - 3x_4 = -6 \end{cases}} \quad (*)$$

$$\Leftrightarrow \begin{cases} x_1 = 4 - x_4 \\ x_2 = 2 - x_4 \\ x_3 = -6 + 3x_4 \end{cases}$$

The variable x_4 is arbitrary.

We set $x_4 = t$ for some $t \in \mathbb{R}$.

All solutions are given by $\begin{cases} x_1 = 4 - t \\ x_2 = 2 - t \\ x_3 = -6 + 3t \\ x_4 = t \end{cases}$ for $t \in \mathbb{R}$. This means that this linear system has infinitely many solutions.

In the linear system $(*)$:

- Each equation contains a variable that occurs in no other equation: (x_1, x_2, x_3), called **pivot variables**.
- The other variables (x_4) are called **free variables**.

A linear system of this shape said to be on **row-reduced echelon form**. In general, this means that the following three conditions are satisfied:

(i) The first (that is, the leftmost) variable in each equation has coefficient 1.

(ii) If x_i is the first variable in one of the equations, then it does not occur in any other equation in the system.

(iii) If x_i is the first variable in one equation, then the equations below it do not contain any of the variables $x_1, x_2, \ldots, x_{i-1}$.

As we saw in the above example we only need three different operations to bring any linear system to row-reduced echelon form:

Definition 1.2 The following operations on a linear system are called **elementary row operations**.
(R1) Add a multiple of an equation to another.
(R2) Multiply an equation with a non-zero number.
(R3) Change the order of the equations.

Since all elementary row operations work in reverse (i.e., all elementary row operations can be undone),

Proposition 1.3 *Applying an elementary row operation to a linear system does not change the set of all solutions of said linear system.*

To bring an arbitrary linear system onto row-reduced echelon form we can use the following algorithm.

Algorithm 1.4 (Gaussian elimination / Row reduction) Given a linear system

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

The procedure for bringing this linear system to its row-reduced echelon form is as follows:

I. Downwards:

1) Make the first equation contain the first variable by using (R3).
2) Make the coefficient of this variable equal to 1 by using (R2).
3) Eliminate this variable from all other equations by using (R1).
4) Iterate with the first occurring variable in the remaining equations.

II. Upwards

1) Let x_i be the first variable in the last equation. Eliminate x_i from all other equations by using (R1).
2) Go to previous equations and iterate.

One example of the usage of Gaussian elimination to solve a linear system is given below:

Example 4 Consider the linear system below. First, the 'downwards' part:

(R3)

$$\begin{cases} \quad\quad\;\; - x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 - x_4 = 0 \\ x_1 + x_2 - x_3 \quad\;\; = 1 \end{cases} \Leftrightarrow$$

(R1)

$$-1 \begin{cases} x_1 + x_2 + x_3 - x_4 = 0 \\ \quad\quad\;\; - x_3 + x_4 = 0 \\ x_1 + x_2 - x_3 \quad\;\; = 1 \end{cases}$$

(R2)

$$\Leftrightarrow \quad -1 \begin{cases} x_1 + x_2 + x_3 - x_4 = 0 \\ \quad\quad\;\; - x_3 + x_4 = 0 \\ \quad\quad - 2x_3 + x_4 = 1 \end{cases} \Leftrightarrow$$

(R1)

$$2 \begin{cases} x_1 + x_2 + x_3 - x_4 = 0 \\ \quad\quad\quad\;\; x_3 - x_4 = 0 \\ \quad\quad - 2x_3 + x_4 = 1 \end{cases}$$

(R2)

$$\Leftrightarrow \quad \begin{cases} x_1 + x_2 + x_3 - x_4 = 0 \\ \quad\quad\quad\;\; x_3 - x_4 = 0 \\ \quad\quad\quad\quad\;\; - x_4 = 1 \end{cases} -1 \quad \Leftrightarrow$$

From here on upwards

(R1)

$$1 \quad 1 \begin{cases} x_1 + x_2 + x_3 - x_4 = 0 \\ \quad\quad\quad\;\; x_3 - x_4 = 0 \\ \quad\quad\quad\quad\quad\; x_4 = -1 \end{cases}$$

(R1)

$$\Leftrightarrow \quad -1 \begin{cases} x_1 + x_2 + x_3 \quad\;\; = -1 \\ \quad\quad\quad\;\; x_3 \quad\;\; = -1 \\ \quad\quad\quad\quad\quad\; x_4 = -1 \end{cases} \Leftrightarrow \begin{cases} x_1 + x_2 \quad\quad\quad\;\; = 0 \\ \quad\quad\quad\;\; x_3 \quad\;\; = -1 \\ \quad\quad\quad\quad\quad\; x_4 = -1 \end{cases}$$

pivot variables
free variables

The free variables can be chosen arbitrarily. We set $x_2 = t$ with an arbitrary $t \in \mathbb{R}$.

All solutions are given by $\begin{cases} x_1 = -t \\ x_2 = t \\ x_3 = -1 \\ x_4 = -1 \end{cases}$ for $t \in \mathbb{R}$.

So far, all linear systems considered had always one (or infinitely many) solutions. One example of using Gaussian elimination to prove that a linear system has no solution is given below.

Example 5 (No solution) Consider the linear system below.

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ x_1 + 3x_2 + 4x_3 = 3 \\ x_1 + 4x_2 + 5x_3 = 4 \end{cases} \Leftrightarrow \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ \quad x_2 + x_3 = 2 \\ \quad 2x_2 + 2x_3 = 3 \end{cases} \Leftrightarrow \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ \quad x_2 + x_3 = 2 \\ \quad 0 = -1 \end{cases}$$

This shows that the original linear system has no solution because it is equivalent to a linear system containing a contradictory equation (of course, $0 \neq -1$!).

Another type of problem that can be solved using this method involves a linear system that is parameterized (i.e., some part of the linear system is determined by a parameter). Usually, one will be asked to consider how the solution to the linear system look with different values of the parameter. One typical problem of this kind is given below:

Example 6 (Parameterized) Consider the following linear system with a parameter $a \in \mathbb{R}$:

$$\begin{cases} (a-1)x_1 + 3x_2 = 2 \\ \quad x_1 - x_2 = 1 \end{cases}.$$

We want to determine for which real numbers a the linear system has solutions and find all the solutions in these cases. To find the solutions of this linear system, we try to bring it on row-reduced echelon form.

$$\begin{cases} (a-1)x_1 + 3x_2 = 2 \\ \quad x_1 - x_2 = 1 \end{cases} \Leftrightarrow \begin{cases} x_1 - x_2 = 1 \\ (a-1)x_1 + 3x_2 = 2 \end{cases} \Leftrightarrow \begin{cases} x_1 - x_2 = 1 \\ (a+2)x_2 = 3-a \end{cases}$$

Now we would like to divide by $2+a$, but this is not possible if $a = -2$ (division by zero). Therefore, we assume that $a \neq -2$ and consider the $a = -2$ case separately.

- Case $a \neq -2$:

$$\begin{cases} x_1 - x_2 = 1 \\ (a+2)x_2 = 3-a \end{cases} \Leftrightarrow \begin{cases} x_1 - x_2 = 1 \\ x_2 = \frac{3-a}{2+a} \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{5}{2+a} \\ x_2 = \frac{3-a}{2+a} \end{cases}$$

  For the case $a \neq -2$ there is exactly one solution given by $\begin{cases} x_1 = \frac{5}{2+a} \\ x_2 = \frac{3-a}{2+a} \end{cases}$.

- Case $a = -2$:

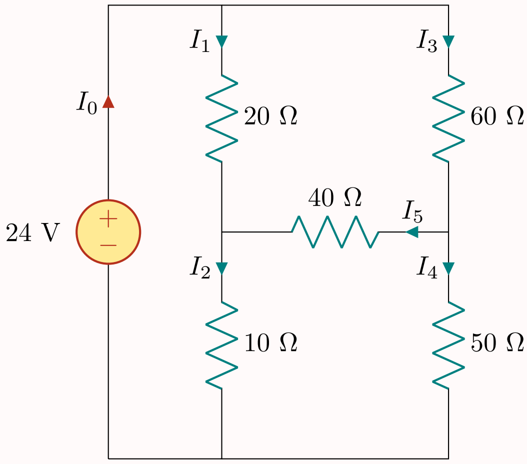
$$\begin{cases} x_1 - x_2 = 1 \\ (a+2)x_2 = 3-a \end{cases} \Leftrightarrow \begin{cases} x_1 - x_2 = 1 \\ 0 = 5 \end{cases}$$

  There are no solutions in the case $a = -2$.

Example 7 (Circuit analysis) Given the following circuit, we want to determine all currents through resistors.

By Kirchhoff's laws, we can derive the following linear system:

$$\begin{cases} I_1 - I_2 + I_5 = 0 \\ I_3 - I_4 - I_5 = 0 \\ 20I_1 + 10I_2 = 24 \\ -20I_1 + 60I_3 + 40I_5 = 0 \\ -10I_2 + 50I_4 - 40I_5 = 0 \end{cases}.$$



$$20 \quad -20 \quad \begin{cases} I_1 - I_2 + I_5 = 0 \\ I_3 - I_4 - I_5 = 0 \\ 20I_1 + 10I_2 = 24 \\ -20I_1 + 60I_3 + 40I_5 = 0 \\ -10I_2 + 50I_4 - 40I_5 = 0 \end{cases}$$

$$\Leftrightarrow \frac{1}{30} \begin{cases} I_1 - I_2 + I_5 = 0 \\ I_3 - I_4 - I_5 = 0 \\ 30I_2 - 20I_5 = 24 \\ -20I_2 + 60I_3 + 60I_5 = 0 \\ -10I_2 + 50I_4 - 40I_5 = 0 \end{cases} \Leftrightarrow \begin{cases} I_1 - I_2 + I_5 = 0 \\ I_3 - I_4 - I_5 = 0 \\ I_2 - \frac{2}{3}I_5 = \frac{4}{5} \\ -20I_2 + 60I_3 + 60I_5 = 0 \\ -10I_2 + 50I_4 - 40I_5 = 0 \end{cases}$$

$$\Leftrightarrow 10 \quad 20 \begin{cases} I_1 - I_2 + I_5 = 0 \\ I_2 - \frac{2}{3}I_5 = \frac{4}{5} \\ I_3 - I_4 - I_5 = 0 \\ -20I_2 + 60I_3 + 60I_5 = 0 \\ -10I_2 + 50I_4 - 40I_5 = 0 \end{cases}$$

$$\Leftrightarrow -60 \begin{cases} I_1 - I_2 + I_5 = 0 \\ I_2 - \frac{2}{3}I_5 = \frac{4}{5} \\ I_3 - I_4 - I_5 = 0 \\ 60I_3 + \frac{140}{3}I_5 = 16 \\ 50I_4 - \frac{140}{3}I_5 = 8 \end{cases} \Leftrightarrow \frac{1}{60} \begin{cases} I_1 - I_2 + I_5 = 0 \\ I_2 - \frac{2}{3}I_5 = \frac{4}{5} \\ I_3 - I_4 - I_5 = 0 \\ 60I_4 + \frac{320}{3}I_5 = 16 \\ 50I_4 - \frac{140}{3}I_5 = 8 \end{cases}$$

$$\Leftrightarrow -50 \begin{cases} I_1 - I_2 + I_5 = 0 \\ I_2 - \frac{2}{3}I_5 = \frac{4}{5} \\ I_3 - I_4 - I_5 = 0 \\ I_4 + \frac{16}{9}I_5 = \frac{4}{15} \\ 50I_4 - \frac{140}{3}I_5 = 8 \end{cases} \Leftrightarrow -\frac{9}{1220} \begin{cases} I_1 - I_2 + I_5 = 0 \\ I_2 - \frac{2}{3}I_5 = \frac{4}{5} \\ I_3 - I_4 - I_5 = 0 \\ I_4 + \frac{16}{9}I_5 = \frac{4}{15} \\ -\frac{1220}{9}I_5 = -\frac{16}{3} \end{cases}$$

$$\Leftrightarrow \quad -1 \quad \tfrac{2}{3} \quad 1 \quad -\tfrac{16}{9} \quad \left\{\begin{array}{ccccccccc} I_1 & - & I_2 & & & & + & I_5 & = & 0 \\ & & I_2 & & & & - & \frac{2}{3}I_5 & = & \frac{4}{5} \\ & & & & I_3 & - & I_4 & - & I_5 & = & 0 \\ & & & & & & I_4 & + & \frac{16}{9}I_5 & = & \frac{4}{15} \\ & & & & & & & & I_5 & = & \frac{12}{305} \end{array}\right. \Leftrightarrow \quad 1 \quad 1 \quad \left\{\begin{array}{ccccccc} I_1 & - & I_2 & & & & = & -\frac{12}{305} \\ & & I_2 & & & & = & \frac{252}{305} \\ & & & I_3 & - & I_4 & = & \frac{12}{305} \\ & & & & & I_4 & = & \frac{12}{61} \\ & & & & & I_5 & = & \frac{12}{305} \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{cccccccc} I_1 & & & & & = & \frac{48}{61} \\ & I_2 & & & & = & \frac{252}{305} \\ & & I_3 & & & = & \frac{72}{305} \\ & & & I_4 & & = & \frac{12}{61} \\ & & & & I_5 & = & \frac{12}{305} \end{array}\right.$$

In addition, we can determine the current coming directly from the source.

$$I_0 = I_1 + I_3 = I_2 + I_4 = \frac{312}{305} \text{ A.}$$

Exercises

Exercise 1. Which of the following linear systems are on row-reduced echelon form? For those that are not, find an equivalent system (i.e. one which has the same solutions) that is on row-reduced echelon form. For each system, find all solutions.

(i) $\left\{\begin{array}{ccccccc} x_1 & + & 4x_2 & + & 7x_3 & = & 1 \\ 2x_1 & + & 5x_2 & + & 8x_3 & = & 2 \\ 3x_1 & + & 6x_2 & + & 10x_3 & = & 1 \end{array}\right.$

(ii) $x_1 + 2x_2 + 3x_3 + 4x_4 = 2022$

(iii) $\left\{\begin{array}{ccccccccc} x_1 & + & x_2 & + & x_3 & + & 2x_4 & = & 0 \\ & & x_2 & & & + & x_4 & = & 0 \end{array}\right.$

(iv) $\left\{\begin{array}{ccccc} x_1 & + & 2x_2 & = & 3 \\ 4x_1 & + & 8x_2 & = & 16 \end{array}\right.$

(v) $\left\{\begin{array}{ccc} x_1 & = & 6 \\ x_2 & = & 9 \\ x_3 & = & 1 \end{array}\right.$

Exercise 2. Decide for which real numbers $a \in \mathbb{R}$ the following linear system has solutions. Give all the solutions in these cases.

$$\left\{\begin{array}{ccccccc} (a-1)^2x_1 & + & x_2 & + & ax_3 & = & 0 \\ x_1 & + & x_2 & & & = & 0 \\ 2x_1 & + & 2x_2 & + & x_3 & = & a \end{array}\right. .$$

Exercise 3. Let $a, b \in \mathbb{R}$ be two arbitrary real numbers. Consider the following linear system:

$$\left\{\begin{array}{ccccc} x_1 & + & x_2 & = & 2 \\ ax_1 & + & 2x_2 & = & b \end{array}\right.$$

Find all the solutions to this linear system depending on a and b.
(Hint: You need to consider different special cases of a and b separately)

Exercise 4. Decide for which real numbers $a \in \mathbb{R}$ the following linear system has solutions. Give all the solutions in these cases.

$$\left\{\begin{array}{rcrcrcr} 2x_1 & + & 12x_2 & + & 7x_3 & = & 12a+7 \\ 2x_1 & + & 4x_2 & + & 2x_3 & = & 12a \\ x_1 & + & 10x_2 & + & 6x_3 & = & 7a+8 \end{array}\right. .$$

Exercise 5. A ramen store in Sakae offers three types of ramen: Miso ramen (price for one portion: 700¥), Taiwan ramen (800¥), and Tonkotsu ramen (850¥). For one portion of Miso ramen one needs 3 tablespoons (tbsp) of salt, one clove of garlic and no chili. One portion of Taiwan ramen needs 2 tbsp. of salt, 2 cloves of garlic and 4 tbsp. of chili. For one portion of Tonkotsu ramen 2 tbsp. of salt, 3 cloves of garlic and one tbsp. of chili is needed.[a] In one day the store uses 142 tbsp. of salt, 146 cloves of garlic, and 152 tbsp. of chili.
How much money (in ¥) did the store earn on this day? Describe this problem by using a linear system and then solve it.

[a] These amounts are made up and should probably not be used to make tasty ramen.

Exercise 6. (8 Points) A Japanese restaurant in 八事(Yagoto, a neighbourhood in Nagoya) is holding an Ebi Festival, and thus is only selling three types of dishes: Ebi Sushi (¥370), Ebi Tempura Don (¥590), and Ebi Fry Bentō (¥830).
One serving of Ebi Sushi requires 3 ounces of shrimp, 1 cup of rice, and 3 tablespoon of shouyu. 5 ounces of shrimp, 4 cups of rice, and $\frac{5}{2}$ tablespoons of shouyu are needed for one portion of Ebi Tempura Don. For one serving of Ebi Fry Bento, 8 ounces of shrimp, 3 cups of rice, and $\frac{1}{2}$ tablespoons of shouyu are needed. In one certain day, the store expended 1000 ounces of shrimp, 500 cups of rice, and 500 tablespoons of shouyu.
The market prices are: ¥50 per ounce of shrimp, ¥30 per cup of rice, and ¥5 per tablespoon of shouyu. Given all these information, how much profit did the restaurant make on this certain day? Describe this problem by using a linear system, bring the linear system on row-reduced echelon form and solve it.

2

Matrices & Vectors

In the previous chapter, we studied linear systems and learned how to solve them using elementary operations. However, as the number of variables and equations increases, these methods become increasingly cumbersome. To overcome this difficulty, we introduce vectors and matrices in this chapter, which allow us to write linear systems in a more concise and elegant manner. Vectors and matrices are fundamental tools in linear algebra, and we will explore their properties and operations in detail.

Definition 2.1

(i) A $m \times n$-**matrix** is given by an array (m rows, n columns) of numbers $a_{ij} \in \mathbb{R}$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} .$$

Notation: We often just write $A = (a_{ij})$ if the size of A, i.e. m and n, are known from context.
By $\mathbb{R}^{m \times n}$ we denote the set all of all $m \times n$-matrices.

(ii) A (column-) **vector** of size n is a $n \times 1$-matrix

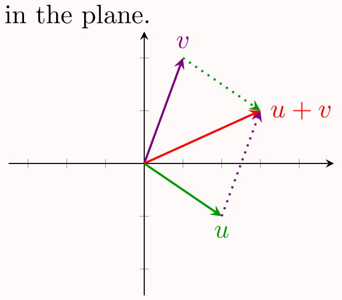
$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

and the set of all vectors of size n is denoted by $\mathbb{R}^n = \mathbb{R}^{n \times 1}$.

Example 8 For $n = 2$ we can visualize vectors in the plane.

$$v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$u = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$



We can also add vectors e.g. $u + v = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

In general the sum of matrices is defined by just adding each entry.

Definition 2.2 For matrices $A = (a_{ij}), B = (b_{ij}) \in \mathbb{R}^{m\times n}$ and a real number $\lambda \in \mathbb{R}$ we define

$$\begin{aligned} A + B &= (a_{ij} + b_{ij}) \in \mathbb{R}^{m\times n} && \text{(Sum of two matrices)}\,, \\ \lambda A &= (\lambda a_{ij}) \in \mathbb{R}^{m\times n} && \text{(Scalar multiplication)}\,. \end{aligned}$$

In the case $\lambda = -1$ we write $(-1)A = -A$ and $A - B$ means $A + (-1)B$.

The matrices A and B need to be of the same size, otherwise the sum $A + B$ is not defined. A special case of the addition of matrices is given by the addition of vectors. For $u, v \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we have

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad u + v = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix}, \quad \lambda v = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}.$$

Definition 2.3 The product of a matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ and a vector $v \in \mathbb{R}^n$ is defined by

$$Av = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{pmatrix} \in \mathbb{R}^m.$$

We have: $(m\times n\text{-matrix}) \cdot (\text{vector of size } n) = (\text{vector of size } m)$.

Example 9 Here is the product of a 3×2 matrix and a vector in $\mathbb{R}^2$:

$$\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1\cdot(-1) + 4\cdot 3 \\ 2\cdot(-1) + 5\cdot 3 \\ 3\cdot(-1) + 6\cdot 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 13 \\ 15 \end{pmatrix},$$

where we get a vector in $\mathbb{R}^3$.

This product of a matrix and a vector satisfies the following rules.

Proposition 2.4 *We have for $A \in \mathbb{R}^{m\times n}$, $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$*

(i) $A(x + y) = Ax + Ay$,

(ii) $A(\lambda x) = \lambda(Ax)$.

Proof. This is Exercise 7. $\square$

Example 10 Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 4 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Find all $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ with $Ax = b$.

$$Ax = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + 3x_3 \\ x_1 + x_2 + 4x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = b \quad \Leftrightarrow \quad \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ x_1 + x_2 + 4x_3 = 2 \end{cases}.$$

This is a linear system. We also call $Ax = b$ a linear system because it gives us a linear system.

Solving:

$$\begin{matrix} -1 \\ \end{matrix} \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ x_1 + x_2 + 4x_3 = 2 \end{cases} \Leftrightarrow \quad \begin{matrix} \\ 2 \end{matrix} \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ \quad - x_2 + x_3 = 1 \end{cases}$$

$$\Leftrightarrow \quad -1 \begin{cases} x_1 \quad\quad + 5x_3 = 3 \\ \quad - x_2 + \quad x_3 = 1 \end{cases} \Leftrightarrow \quad \begin{cases} \overset{\text{pivot variables}}{x_1} \quad\quad + 5x_3 = 3 \\ x_2 - \underset{\text{free variable}}{x_3} = -1 \end{cases}$$

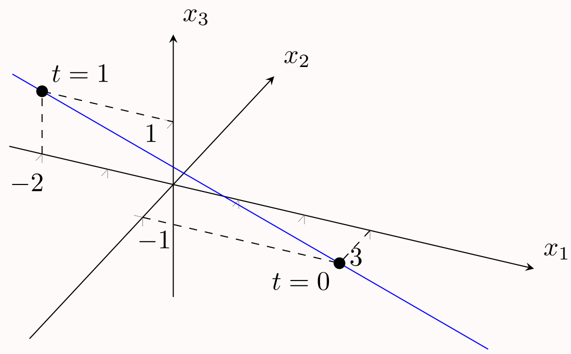
Solution: $\begin{cases} x_1 = 3 - 5t \\ x_2 = -1 + t \\ x_3 = t \end{cases}$ for $t \in \mathbb{R}$. Using the vector notation, this can be written as:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 - 5t \\ -1 + t \\ t \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -5 \\ 1 \\ 1 \end{pmatrix}.$$

Plotting the above solution for all possible values of t gives a line in $\mathbb{R}^3$.

$t = 0 : x = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$

$t = 1 : x = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} -5 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$



We can also use the matrix notation when solving a linear system to avoid having to write the symbols x_i of variables all the time as follows.

Definition 2.5 For a matrix $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ and a vector $b \in \mathbb{R}^m$ the matrix

$$(A \mid b) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right) \in \mathbb{R}^{m \times (n+1)}$$

is called the **augmented matrix** of the linear system $Ax = b$.

The augmented matrix $(A \mid b)$ is just the matrix A where we append the vector b as a column. The line $\mid$ is a useful notation to distinguish between the left- and right-hand side of the corresponding linear system but it has no mathematical meaning. We will view $(A \mid b)$ as a usual matrix with m rows and $n + 1$ columns.

Definition 2.6 The following operations on a matrix are called **elementary row operations**.
(R1) Add a multiple of one row to another row.
(R2) Multiply a row with a <u>non-zero</u> number.
(R3) Interchange two rows.

Applying a row operation to a linear system (Definition 1.2) corresponds to the same row operation (Definition 2.6) on the corresponding augmented matrix of this linear system.

Definition 2.7 Two matrices A and B are called **row equivalent**, if B can be obtained from A by elementary row operations. In this case we write

$$A \sim B.$$

Notice that if $A \sim B$, then also $B \sim A$, i.e. A can be obtained from B by elementary row operations. In Example 10, we can use matrix notation for solving the linear system:

$$(A \mid b) = \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 1 & 1 & 4 & 2 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -1 & 1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 5 & 3 \\ 0 & -1 & 1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 5 & 3 \\ 0 & 1 & -1 & -1 \end{array}\right).$$

Each row operation creates two equivalent linear systems, which results in the following proposition.

Proposition 2.8 *Let $A, B \in \mathbb{R}^{m\times n}$ and $b, c \in \mathbb{R}^m$. If $(A \mid b) \sim (B \mid c)$ then the linear systems $Ax = b$ and $Bx = c$ have the same solutions.*

The final result after applying row operations which helps us directly obtain the solution of a linear system is defined as follows.

Definition 2.9 A matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ is on **row-reduced echelon form** if

(i) The first non-zero element on each row (if any) is equal to 1.

(ii) If there is a leading 1 in a row, then all rows above contain a leading 1 further to the left.

(iii) If a_{ij} is the first non-zero element in row i, then there are no other non-zero elements in the j-th column.

The first non-zero element in a row of a matrix in row-reduced echelon form is called **pivot element**.

In Example 10, we obtain the row-reduced echelon form as follows:

$$(A \mid b) = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 1 & 1 & 4 & 2 \end{pmatrix} \sim \left(\begin{array}{ccc|c} 1 & 0 & 5 & 3 \\ 0 & 1 & -1 & -1 \end{array}\right).$$

Is the row-reduced echelon form unique for every matrix? The following theorem will help us answer this question.

Theorem 2.10 *Every matrix A is row equivalent to a unique matrix B on row-reduced echelon form and we write*

$$B = \operatorname{rref}(A)\,.$$

Proof. We prove this theorem by induction on the number of columns of A (see Chapter 17 or `https://en.wikipedia.org/wiki/Mathematical_induction` if you are not familiar with this concept). If A has only one column, then there are only two different row-reduced echelon forms:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = e_1 \quad \text{or} \quad \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{0}.$$

Indeed, A cannot be row equivalent to both of these forms because there is no sequence of row operations leading one form to another (the zero vector is only row equivalent to itself). Hence, we already proved the base step (1 column).

Now, let $n > 1$ and assume that every matrix with $n-1$ columns is equivalent to a unique matrix on row-reduced echelon form (RREF). Furthermore, assume A has m rows and n columns. Suppose that A is row equivalent to B and C, which are both on RREF. Let A_1, B_1, and C_1 be matrices formed by the first $n-1$ columns of A, B, and C respectively. Since B_1 and C_1 are both row equivalent to A_1 and on RREF, they are equal by the induction hypothesis. Suppose for a contradiction that $B \neq C$. Then, there exists an index j such that $b_{jn} \neq c_{jn}$.

From Proposition 2.8, the linear systems $Ax = \mathbf{0}$, $Bx = \mathbf{0}$, and $Cx = \mathbf{0}$ all have the same solutions. Let $v \in \mathbb{R}^n$ be any solution to $Ax = \mathbf{0}$. Then, we have $Bv = \mathbf{0}$ and $Cv = \mathbf{0}$, and therefore $(B - C)v = \mathbf{0}$.

Since the first $n-1$ columns of $B-C$ are zeroes (due to $B_1 = C_1$), we get the j^{th} row of the system: $(b_{jn} - c_{jn})v_n = 0$, so $v_n = 0$. Hence, every solution to $Ax = \mathbf{0}$ has zero at the last entry.

The matrix B_1 on RREF has some nonzero rows and then some zero rows. Say there are k zero rows. We can then write B_1 in the form

$$\begin{pmatrix} D \\ \mathbf{0}_{k\times(n-1)} \end{pmatrix},$$

where $D \in \mathbb{R}^{(m-k)\times(n-1)}$ with no zero rows and $\mathbf{0}_{k\times(n-1)}$ is the $k \times (n-1)$ zero matrix. Then, B and C have the form

$$B = \begin{pmatrix} D & b \\ \mathbf{0}_{k\times(n-1)} & t \end{pmatrix}, \qquad C = \begin{pmatrix} D & c \\ \mathbf{0}_{k\times(n-1)} & u \end{pmatrix},$$

where $b, c \in \mathbb{R}^{m-k}$ and $t, u \in \mathbb{R}^k$. Suppose that $t = \mathbf{0}$. The first $m-k$ rows of B all have left-most elements equal to 1. For $1 \le i \le m-k$, let such element in row i of B occur in column c_i. Also, let $b = \begin{pmatrix} b_1 \\ \vdots \\ b_{m-k} \end{pmatrix}$. Then, the linear system $Bx = \mathbf{0}$ has a solution with the c_i^{th} element equal to b_i, the last element equal to -1, and zeroes elsewhere. This contradicts to that every solution to $Ax = \mathbf{0}$ has zero at the last entry. Thus, $t \neq \mathbf{0}$.

Since B is on RREF, $t = e_1$ and $b = \mathbf{0}$. The same argument applies to C, so $u = e_1$ and $c = \mathbf{0}$. Hence, $B = C$, which is a contradiction to the assumption that $B \neq C$. Therefore, $B = C$, which completes the induction step. $\square$

For a general matrix $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ the row-reduced echelon form $(B \mid c)$ of the augmented matrix $(A \mid b)$ has the following shape

$$\mathrm{rref}(A \mid b) = (B \mid c) = m\left(\begin{array}{cccccccccccccccccc|c}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * & 0 & & & * \\
 & & & & & & & 1 & * & \cdots & * & 0 & * & \cdots & * & \vdots & & & \vdots \\
 & & & & & & & & & & & 1 & * & \cdots & * & \vdots & & & \vdots \\
 & & & & & & & & & & & & & & \ddots & \vdots & & & \vdots \\
 & & & & & & & & & & & & & & & 0 & & & \vdots \\
 & & & & & & & & & & & & & & & 1 & * & \cdots & * \\
 & & & & & & & & & & & & & & & & & & \color{red}{\lambda} \\
 & & & & & & & & & & & & & & & & & &
\end{array}\right)\begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array}\right\} c' \\ \\ \\ \end{array}$$

$$\xleftarrow{\qquad\qquad\qquad n \qquad\qquad\qquad} \qquad \color{red}{\lambda \in \{0,1\}}$$

We can read off solutions for linear system $Ax = b$ after finding $\mathrm{rref}(A \mid b)$ as follows:

$$(A \mid b) \sim (B \mid c) = \mathrm{rref}(A \mid b), \text{ for } A, B \in \mathbb{R}^{m\times n}, \text{ and } b, c \in \mathbb{R}^m$$

1) If the last column contains a pivot element $\color{red}{(\lambda = 1)}$: **No solutions** (since $0 \neq 1$).
   Else $\color{red}{(\lambda = 0)}$:
2) If every column of B contains a pivot element then we have the **unique solution** $x = c'$.
3) Some columns of B do not contain pivot elements: **Infinitely many solutions**.

From the above discussion, we see that the number of pivot elements and their location are important, which leads us to introduce the following definition.

> **Definition 2.11** Let $A \in \mathbb{R}^{m\times n}$ be a matrix. The **rank** $\mathrm{rk}(A)$ of A is the number of pivot elements in $\mathrm{rref}(A)$.

Following the discussion of three situations arising when we analyze the row-reduced echelon form of a matrix, we summarize them by using the notion of rank in the following proposition.

Proposition 2.12 *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. The solution of $Ax = b$ depend on $\operatorname{rk}(A \mid b)$ and $\operatorname{rk}(A)$ as follows:*

*(i) If $\operatorname{rk}(A \mid b) > \operatorname{rk}(A)$ then $Ax = b$ has **no solutions**.*

*(ii) If $\operatorname{rk}(A \mid b) = \operatorname{rk}(A) = n$ then $Ax = b$ has a **unique solution**.*

*(iii) If $\operatorname{rk}(A \mid b) = \operatorname{rk}(A) < n$ then $Ax = b$ has **infinitely many solutions**.*

Proof. (i) If $\operatorname{rk}(A \mid b) > \operatorname{rk}(A)$, then the last column of $\operatorname{rref}(A \mid b)$ contains a pivot element. Hence, the linear system $Ax = b$ has no solution.

(ii) If $\operatorname{rk}(A \mid b) = \operatorname{rk}(A) = n$, then the last column of $\operatorname{rref}(A \mid b)$ does not contain a pivot element and every column of $\operatorname{rref}(A)$ contains a pivot element. Thus, the linear system $Ax = b$ has a unique solution given by $x = c'$.

(iii) If $\operatorname{rk}(A \mid b) = \operatorname{rk}(A) < n$, then the last column of $\operatorname{rref}(A \mid b)$ does not contain a pivot element but some columns of $\operatorname{rref}(A)$ do not contain pivot elements. Therefore, the linear system $Ax = b$ has infinitely many solutions.

$\square$

Exercises

Exercise 7. Show that for all $A \in \mathbb{R}^{m\times n}$, $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we have

(i) $A(x + y) = Ax + Ay$,

(ii) $A(\lambda x) = \lambda(Ax)$.

(Without using Proposition 2.4).

Exercise 8. We define the following matrices and vectors:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & 0 \\ 2 & 4 \\ 0 & -3 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 8 & 0 \\ 1 & 2 & -1 \end{pmatrix}, \quad E = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix},$$

$$t = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad u = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} 3 \\ -4 \end{pmatrix}, \quad w = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}.$$

(i) Decide which of the following expressions are defined. Evaluate them if possible.

$$At, \quad Au, \quad wA, \quad 2A, \quad A + B, \quad A + C, \quad A + D, \quad \frac{3}{4}Bt, \quad Bu, \quad B + B, \quad Dw,$$
$$Cv, \quad t + u, \quad tu, \quad -v, \quad u + w, \quad t - u, \quad \frac{1}{2}w, \quad C + w, \quad Et, \quad Ev \quad E\,(Ev)\,.$$

(ii) Draw the following vectors in $\mathbb{R}^2$

$$t, \quad ,v, \quad -2t, \quad t - \frac{1}{2}v, \quad v + t, \quad t + v, \quad Et, \quad Ev, \quad E\,(Ev)\,, \quad Bt, \quad Bv\,.$$

Can you guess what happens in general to a vector in $\mathbb{R}^2$ when you multiply it with B or E? Try to give a geometric interpretation. (without a proof)

Exercise 9. Let $a, b, c, d \in \mathbb{R}$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(i) Show that $\operatorname{rk}(A) = 2$ if and only if $ad - bc \neq 0$.

(ii) We define the following subset of $\mathbb{R}^2$

$$L = \{x \in \mathbb{R}^2 \mid x = Av \text{ for some } v \in \mathbb{R}^2\}.$$

How does L look like if $\mathrm{rk}(A) = 1$? How does it look like if $\mathrm{rk}(A) = 2$?

Exercise 10. Give examples of matrices $A, B, C \in \mathbb{R}^{3\times 3}$, which are all <u>not</u> on row-reduced echelon form, such that $\mathrm{rk}(A) = 1$, $\mathrm{rk}(B) = 2$, $\mathrm{rk}(C) = 3$.

Exercise 11. Let $A \in \mathbb{R}^{3\times 3}$ be a matrix.

(i) Show that if $\mathrm{rk}(A) = 3$ then there exists just one vector $x \in \mathbb{R}^3$ with $Ax = 0$.

(ii) Show that if $\mathrm{rk}(A) \leq 2$ then there exist infinitely many vectors $x \in \mathbb{R}^3$ with $Ax = 0$.

Exercise 12. Let $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ be a polynomial of degree 3 with real coefficients $a_0, a_1, a_2, a_3 \in \mathbb{R}$. For this polynomial p we define the vector v_p by

$$v_p = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \in \mathbb{R}^4 .$$

Find a matrix $D \in \mathbb{R}^{4\times 4}$, such that $v_{p'} = D v_p$, where p' denotes the derivative of the polynomial p with respect to x. What is the rank of D?

3

Sets & Functions

A **set** is a collection of distinct objects, grouped together as a single entity. It is precisely, but not necessarily explicitly, defined. The objects that belong to a set are called its elements. If a set has finitely many elements we call it a finite set and otherwise infinite set. We have already seen examples of infinite sets: $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$.

Example 11
1) $\{2, 4, \pi\}$ is a finite set.
2) $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$ is the set of natural numbers.
3) $\mathbb{Q}$ is the set of rational numbers.
4) $\emptyset = \{\}$ is the empty set, which has no element.

Given a set A, we write "$a \in A$" if a is an element of A and "$a \notin A$" if a is not an element of A.
A set A is a subset of another set B when every element of A belongs to B. That is, if $a \in A$, then $a \in B$. In this case, we write $A \subset B$. The empty set $\emptyset$ is subset of any other set.

Example 12
1) $2 \in \mathbb{N}$, $\frac{1}{2} \notin \mathbb{N}$, $\pi \in \mathbb{R}$, $\pi \notin \mathbb{Q}$.
2) $\emptyset \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.
3) $\{1, 2, 3\} \subset \mathbb{N}$.

From a set A, we can define a subset of A that contains all elements of A satisfying a condition and we write it in the format $\{a \in A \mid \text{condition}\}$.

Example 13
1) $\{m \in \mathbb{N} \mid m \text{ is even}\}$ is the set of all even numbers.
2) Let H be the set of all humans. $\mathrm{NU} = \{h \in \mathrm{H} \mid h \text{ is a student at Nagoya University}\} \subset \mathrm{H}$.
3) $\{x \in \mathbb{R}^n \mid Ax = b\}$ is the set of all solutions of $Ax = b$.

We can create new sets from given sets by operations on sets. Given two sets A and B, we define the following operations (that produce new sets):

(i) **Union:** $A \cup B$ is the set of all objects that are elements of A or B or both.

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

(ii) **Intersection:** $A \cap B$ is the set of all objects that are elements of both A and B.

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

(iii) **Difference:** $A \setminus B$ is the set of all objects that are elements of A but not B.

$$A \setminus B = \{x \in A \mid x \notin B\}$$

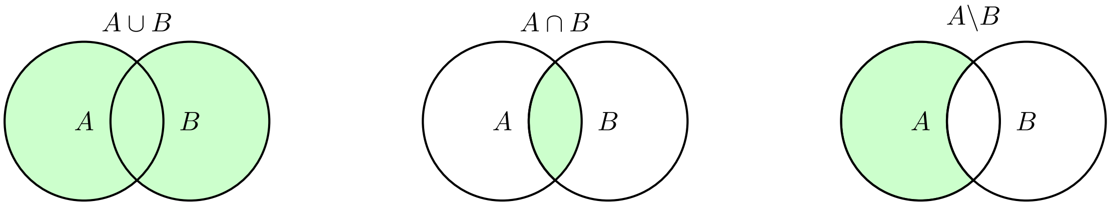


Figure 3.1: Visualization of the union, intersection and difference of two sets A and B.

Example 14 For the sets $A = \{-1, 2, 3\}$ and $B = \mathbb{N} = \{1, 2, 3, \dots\}$, we have
1) $A \cup B = \{-1, 1, 2, 3, \dots\}$,
2) $A \cap B = \{2, 3\}$,
3) $A \setminus B = \{-1\}$.

In order to introduce later the notion of **linear maps** in Chapter 4, which will be an important concept in linear algebra, now we want to examine the notion of **function** in general.

Definition 3.1 Let X and Y be two sets.

(i) A **function** $f : X \to Y$ is a rule, assigning to each element $x \in X$ an element $f(x) \in Y$. This is also denoted by

$$\begin{aligned} f : X &\longrightarrow Y \\ x &\longmapsto f(x)\,. \end{aligned}$$

(ii) For $f : X \to Y$, the set X is called the **domain of** f and Y is called the **codomain of** f.

A function is also sometimes called a **map**. These two names (for the exact same mathematical object) are used interchangeably in literature.

Definition 3.2 For a function $f : X \to Y$, the **image of** f is defined by

$$\operatorname{im}(f) = \{y \in Y \mid \exists x \in X : y = f(x)\}\,.$$

Another notation is $\operatorname{im}(f) = f(X)$. The image is a subset of the codomain, i.e., $\operatorname{im}(f) \subset Y$.

A fundamental operation for functions that allows us to create a new function from one function or several functions is as follows.

Definition 3.3 **Composition of functions:** For two functions $f : X \to Y$ and $g : Y \to Z$, the composition $g \circ f$ of f and g is defined by

$$\begin{aligned} g \circ f : X &\longrightarrow Z \\ x &\longmapsto (g \circ f)(x) = g(f(x))\,. \end{aligned}$$

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & Y \arrow[r, "g"] & Z \end{tikzcd}$$

The following notions are characteristics of certain special types of functions.

Definition 3.4 A function $f : X \to Y$ is called

(i) **injective** if $f(x_1) \neq f(x_2)$ whenever $x_1 \neq x_2$. $(x_1, x_2 \in X)$

(ii) **surjective** if $\operatorname{im}(f) = Y$.

(iii) **bijective** if it is both injective and surjective.

Given a set X, we define the **identity function** on X as follows:

$$\begin{aligned} \mathrm{id}_X : X &\longrightarrow X, \\ x &\longmapsto x\,. \end{aligned}$$

If $f : X \to Y$ is bijective or, equivalently, for *every* $y \in Y$ there exists a *unique* $x \in X$ with $f(x) = y$, then we can define $g : Y \to X$ by $g(y) = x$ for $f(x) = y$. In that case, we have

$$\begin{aligned} (g \circ f)(x) &= g(f(x)) = g(y) = x \quad \forall x \in X \quad \Rightarrow \quad g \circ f = \mathrm{id}_X, \\ (f \circ g)(y) &= f(g(y)) = f(x) = y \quad \forall y \in Y \quad \Rightarrow \quad f \circ g = \mathrm{id}_Y\,. \end{aligned}$$

For any function f, when there exists a function g which satisfies the above two conditions, we say that f is **invertible** and g is the **inverse** of f. Usually, we denote the inverse as $g = f^{-1}$. Hence, a bijective function is also invertible and vice versa.

Example 15

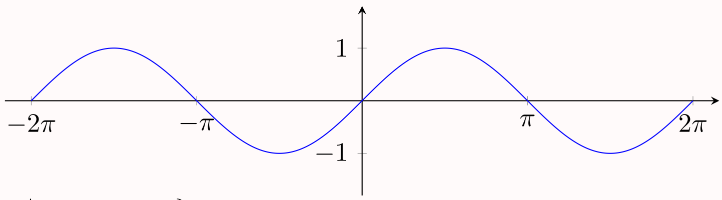
1) Let H be the set of all humans. $\mathrm{NU} = \{h \in \mathrm{H} \mid h \text{ is a student at Nagoya University}\} \subset \mathrm{H}$.

$$\begin{aligned} f_1 : \mathrm{NU} &\longrightarrow \mathbb{N}, \\ s &\longmapsto \text{ Student ID of student } s. \end{aligned}$$

f_1 is injective because there are no two students with the same ID.
f_1 is not surjective because not every natural number is the student ID of a student.

2) Now consider the sine function with domain and codomain $\mathbb{R}$:

$$\begin{aligned} f_2 : \mathbb{R} &\longrightarrow \mathbb{R}, \\ x &\longmapsto \sin(x)\,. \end{aligned}$$

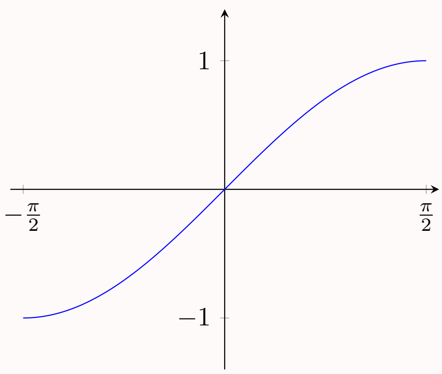


$\operatorname{im}(f_2) = [-1, 1] = \{x \in \mathbb{R} \mid -1 \leq x \leq 1\}$.
f_2 is not surjective because $2 \in \mathbb{R}$ but $2 \notin \operatorname{im}(f_2)$.
f_2 is not injective because $f_2(0) = f_2(2\pi) = 0$ but $0 \neq 2\pi$.

3) If we consider the sine function where we restrict the domain and codomain the situation changes:

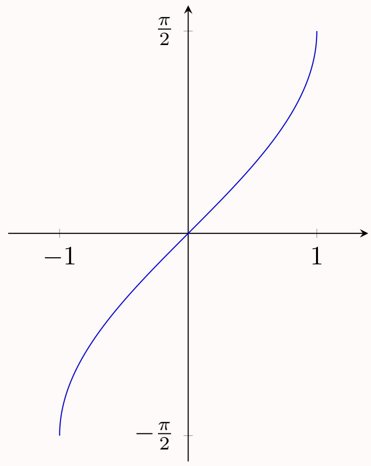
$$\begin{aligned} f_3 : \left[-\frac{\pi}{2}, \frac{\pi}{2} q\right] &\longrightarrow [-1, 1], \\ x &\longmapsto \sin(x). \end{aligned}$$

$\operatorname{im}(f_3) = [-1, 1] \;\Rightarrow\; f_3$ is surjective.
f_3 is also injective because, for each $y \in [-1, 1]$, there exists *exactly one* $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ with $f_3(x) = y$.



Hence, f_3 is bijective with the inverse

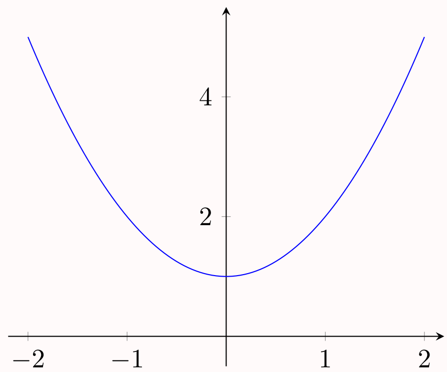
$$f_3^{-1} = \arcsin : [-1, 1] \longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$



4) Consider the following function:

$$f_4 : \mathbb{R} \longrightarrow \mathbb{R}.$$
$$x \longmapsto x^2 + 1,$$

$\operatorname{im}(f_4) = [1, \infty) = \{x \in \mathbb{R} \mid x \geq 1\} \neq \mathbb{R}.$
$f_4(-x) = x^2 = f_4(x) \quad \forall x \in \mathbb{R}.$
Hence, f_4 is neither injective nor surjective.



5) Consider the following function:

$$f_5 : \mathbb{R}^2 \longrightarrow \mathbb{R}^2,$$
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Ax,$$

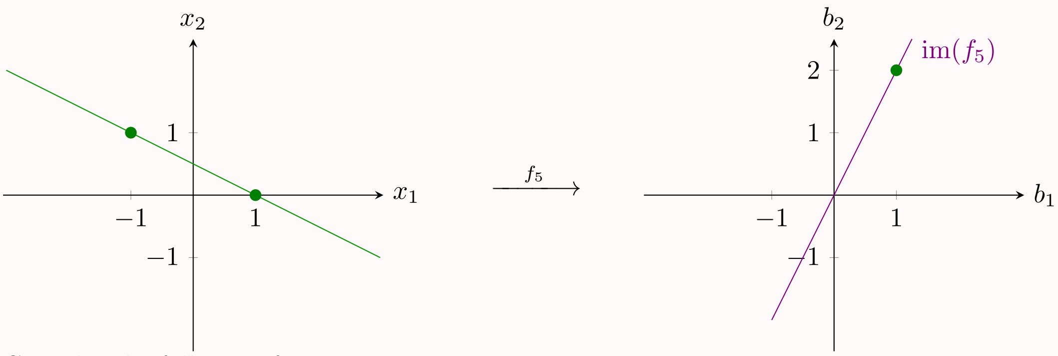
where $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ and $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. To calculate the image of f_5 we need to find all $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2$ such that there exists a $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ with $f_5(x) = Ax = b$, which means that for each $b \in \mathbb{R}^2$ we need to solve the linear system $Ax = b$.

$$(A \mid b) = \overset{-2}{\hookrightarrow} \left(\begin{array}{cc|c} 1 & 2 & b_1 \\ 2 & 4 & b_2 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & 0 & b_2 - 2b_1 \end{array}\right).$$

Hence, $Ax = b$ has solutions when $b_2 = 2b_1$. Thus, $\operatorname{im}(f_5) = \left\{ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2 \,\middle|\, b_2 = 2b_1 \right\} \neq \mathbb{R}^2$.
Therefore, f_5 is not surjective.
If $b_2 = 2b_1$, we have the solution $\begin{cases} x_1 = b_1 - 2t \\ x_2 = t \end{cases}$, for $t \in \mathbb{R}$.
For $b_1 = 1$:

$$\begin{cases} x_1 = 1 - 2t \\ x_2 = t \end{cases} \quad \text{and } f_5 \begin{pmatrix} 1 - 2t \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \forall t \in \mathbb{R}.$$

Hence, f_5 is not injective since $f_5 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = f_5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ but $\begin{pmatrix} -1 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.



6) Consider the following function:

$$\begin{aligned} f_6 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2, \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Ax, \end{aligned}$$

where $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. In order to find

$$\mathrm{im}(f_6) = \{y \in \mathbb{R}^2 \mid \exists x \in \mathbb{R}^2 : f_6(x) = y\} = \{y \in \mathbb{R}^2 \mid Ax = y \text{ has a solution } x \in \mathbb{R}^2\},$$

we need to understand the solutions of $Ax = y$.

$$\begin{aligned} (A \mid y) = \begin{array}{c} -3 \\ \end{array} \left(\begin{array}{cc|c} 1 & 2 & y_1 \\ 3 & 4 & y_2 \end{array}\right) &\sim \begin{array}{c} \\ -\frac{1}{2} \end{array} \left(\begin{array}{cc|c} 1 & 2 & y_1 \\ 0 & -2 & y_2 - 3y_1 \end{array}\right) \\ \sim \begin{array}{c} \\ -2 \end{array} \left(\begin{array}{cc|c} 1 & 2 & y_1 \\ 0 & 1 & \frac{3}{2}y_1 - \frac{1}{2}y_2 \end{array}\right) &\sim \left(\begin{array}{cc|c} 1 & 0 & -2y_1 + y_2 \\ 0 & 1 & \frac{3}{2}y_1 - \frac{1}{2}y_2 \end{array}\right) \end{aligned}$$

This shows that the linear system $Ax = y$ has a unique solution

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2y_1 + y_2 \\ \frac{3}{2}y_1 - \frac{1}{2}y_2 \end{pmatrix}$$

for every $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$. Hence, $\mathrm{im}(f_6) = \mathbb{R}^2$ and f_6 is bijective. In addition, the inverse $f_6^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ of f_6 is defined as follows:

$$f_6^{-1} = \begin{pmatrix} -2x_1 + x_2 \\ \frac{3}{2}x_1 - \frac{1}{2}x_2 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Bx,$$

where $B = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$.
Later we will see that $B = A^{-1}$ is the inverse of A (Chapter 7).

7) Consider the following function:

$$\begin{aligned} f_7 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} x_1 \\ x_1 + x_2 \\ x_2 - x_1 \end{pmatrix} \end{aligned}$$

We want to calculate $\mathrm{im}(f_7) = \{y \in \mathbb{R}^3 \mid \exists x \in \mathbb{R}^2 : f_7(x) = y\}$. For this, we first rewrite

$$f_7(x) = f_7 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 + x_2 \\ x_2 - x_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Ax,$$

where $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ -1 & 1 \end{pmatrix}$. Hence, $\operatorname{im}(f_7) = \{y \in \mathbb{R}^2 \mid Ax = y \text{ has a solution } x \in \mathbb{R}^2\}$. Therefore, we want to understand the solutions of $Ax = y$.

$$(A \mid y) = \left(\begin{array}{cc|c} 1 & 0 & y_1 \\ 1 & 1 & y_2 \\ -1 & 1 & y_3 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & y_1 \\ 0 & 1 & y_2 - y_1 \\ 0 & 1 & y_3 + y_1 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & y_1 \\ 0 & 1 & y_2 - y_1 \\ 0 & 0 & 2y_1 - y_2 + y_3 \end{array}\right)$$

This shows that $Ax = y$ has a solution if and only if $2y_1 - y_2 + y_3 = 0$.
Thus, $\operatorname{im}(f_7) = \left\{ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; 2y_1 - y_2 + y_3 = 0 \right\}$ and f_7 is not surjective since $\operatorname{im}(f_7) \neq \mathbb{R}^3$.
For $y \in \operatorname{im}(f_7)$, the system $Ax = y$ has a unique solution because there are no free variables. Therefore, f_7 is injective (but not surjective).

Exercises

Exercise 13. Let X be a finite set. Show that a function $f : X \to X$ is injective if and only if it is surjective.

Exercise 14. Let $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ and $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$ denote the set of natural numbers (together with zero) and the integers. Decide if the following function is injective and/or surjective:

$$\begin{aligned} g : \mathbb{N}_0 &\longrightarrow \mathbb{Z} \\ n &\longmapsto \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ -\frac{n+1}{2} & \text{if } n \text{ is odd} \end{cases} . \end{aligned}$$

Exercise 15. Which of the following functions are injective and/or surjective?

$$\begin{aligned} f_1 : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto e^x , \end{aligned}$$

$$\begin{aligned} f_2 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + 4x_2 \\ x_1 - x_2 \end{pmatrix} , \end{aligned}$$

$$\begin{aligned} f_3 : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto 3x + 2 , \end{aligned}$$

$$\begin{aligned} f_4 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} x_1 + 2x_2 \\ x_1 x_2 \end{pmatrix} . \end{aligned}$$

4

Linear maps

In this chapter, we discuss linear maps, which serve as a fundamental concept in linear algebra. Up until now, we have explored the properties of vectors in $\mathbb{R}^n$ and their various operations. Linear maps, also known as linear transformations, are special functions that map one vector space to another while preserving the underlying structure and operations (vector addition and scalar multiplication) of the original space. As we have not yet formally introduced the notion of vector spaces (Chapter 15), it is essential to understand that these abstract mathematical structures provide a broader and more general framework for linear maps.

Let's consider a simple example of a linear map using matrix multiplication. Suppose we have an arbitrary $m \times n$ matrix A and a vector $x \in \mathbb{R}^n$. The linear map F represented by the matrix A can be defined as the function $F(x) = Ax$, which maps a vector x in $\mathbb{R}^n$ to a vector in $\mathbb{R}^m$. This linear map satisfies by Proposition 2.4 the property $F(x+y) = A(x+y) = Ax+Ay = F(x)+F(y)$ for any $x, y \in \mathbb{R}^n$ and $F(\lambda x) = A(\lambda x) = \lambda Ax = \lambda F(x)$ for $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. This property will be the definition of a linear map and we will see, that all linear maps indeed come from a matrix in the above way.

As we explore the properties and applications of linear maps, you will see that such transformations play a significant role in various areas of mathematics and real-world applications, including computer graphics, data analysis, and engineering.

Definition 4.1 A function $F : \mathbb{R}^n \to \mathbb{R}^m$ is a **linear map** if for all $u, v \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ we have

(i) $F(u + v) = F(u) + F(v)$,
(ii) $F(\lambda u) = \lambda\, F(u)$.

Example 16

1) For any matrix $A \in \mathbb{R}^{m\times n}$, the function

$$\begin{aligned} F : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longmapsto Ax \end{aligned}$$

is a linear map. This follows from Proposition 2.4 as follows: for any $u, v \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

(i) $F(u + v) = A(u + v) = Au + Av = F(u) + F(v)$,
(ii) $F(\lambda u) = A(\lambda u) = \lambda(Au) = \lambda F(u)$.

Special case: When $n = m$ and $A = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \overset{\text{def}}{=} I_n$, called identity matrix, we have $F(x) = x, \quad \forall x \in \mathbb{R}^n$. In this case, $F = \mathrm{id}_{\mathbb{R}^n}$.

2) The function

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} x_1 x_2 \\ x_1 \end{pmatrix} \end{aligned}$$

is not a linear map. For $\lambda = 2$ and $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$,

$$\begin{aligned} f(\lambda u) &= f\left(2\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = f\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2\cdot 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \\ \lambda f(u) &= 2f\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1\cdot 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}. \end{aligned}$$

Therefore, $f(\lambda u) \neq \lambda f(u)$ for the case $\lambda = 2$.

In fact we will see now that any linear map is given by a function like in Example 16 1).

Theorem 4.2 *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Then there exists a unique matrix $[F] \in \mathbb{R}^{m\times n}$, such that for all $x \in \mathbb{R}^n$ we have*

$$F(x) = [F]x\,.$$

Here the left-hand side is the evaluation of the function F at x and the right-hand side is the multiplication of the matrix $[F]$ with the vector x.

Proof. For $1 \le j \le n$, we consider vectors $e_j \in \mathbb{R}^n$ such that

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{the j}^{\text{th}} \text{ entry}$$

Every $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ can be uniquely written as $x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$.
Since F is linear, we have

$$\begin{aligned} F(x) &= F(x_1 e_1 + x_2 e_2 + \ldots + x_n e_n) \\ &\overset{\text{(i)}}{=} F(x_1 e_1) + F(x_2 e_2 + \ldots + x_n e_n) = \ldots = F(x_1 e_1) + F(x_2 e_2) + \ldots + F(x_n e_n) \\ &\overset{\text{(ii)}}{=} x_1 F(e_1) + x_2 F(e_2) + \ldots + x_n F(e_n) \end{aligned}$$

Now set $[F] = \begin{pmatrix} | & | & & | \\ F(e_1) & F(e_2) & \cdots & F(e_n) \\ | & | & & | \end{pmatrix} \in \mathbb{R}^{m\times n}$. With this, we have

$$[F]x = [F]\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 F(e_1) + x_2 F(e_2) + \ldots + x_n F(e_n),$$

And therefore, $F(x) = [F]x$. $\square$

Definition 4.3 The matrix $[F]$ in Theorem 4.2 is called **the matrix of F**.

Remark. If F is a linear map and we know the values $F(e_j)$ $(1 \leq j \leq n)$, then we know the value of $F(x)$ for any x.

Example 17

1) If $F : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear map with $F\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix}$ and $F\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}3\\4\end{pmatrix}$, then

$$F\begin{pmatrix}-1\\3\end{pmatrix} = F\left(-1\begin{pmatrix}1\\0\end{pmatrix} + 3\begin{pmatrix}0\\1\end{pmatrix}\right) = -F\begin{pmatrix}1\\0\end{pmatrix} + 3F\begin{pmatrix}0\\1\end{pmatrix} = -\begin{pmatrix}1\\2\end{pmatrix} + 3\begin{pmatrix}3\\4\end{pmatrix} = \begin{pmatrix}8\\10\end{pmatrix}.$$

In general,

$$F\begin{pmatrix}x_1\\x_2\end{pmatrix} = F\left(x_1\begin{pmatrix}1\\0\end{pmatrix} + x_2\begin{pmatrix}0\\1\end{pmatrix}\right) = x_1F\begin{pmatrix}1\\0\end{pmatrix}+x_2F\begin{pmatrix}0\\1\end{pmatrix} = x_1\begin{pmatrix}1\\2\end{pmatrix}+x_2\begin{pmatrix}3\\4\end{pmatrix} = \begin{pmatrix}1&3\\2&4\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}.$$

The matrix of F is $[F] = \begin{pmatrix}1&3\\2&4\end{pmatrix}$.

2) This works in more general cases. Assume that $F : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear map with

$$F\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\\-1\end{pmatrix}, \qquad F\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\1\\2\end{pmatrix}.$$

What is $F\begin{pmatrix}x_1\\x_2\end{pmatrix}$ for any x_1, x_2? We have

$$\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_2\begin{pmatrix}1\\1\end{pmatrix} + (x_1 - x_2)\begin{pmatrix}1\\0\end{pmatrix}.$$

Hence,

$$\begin{aligned}F\begin{pmatrix}x_1\\x_2\end{pmatrix} &= F\left(x_2\begin{pmatrix}1\\1\end{pmatrix} + (x_1 - x_2)\begin{pmatrix}1\\0\end{pmatrix}\right) = x_2F\begin{pmatrix}1\\1\end{pmatrix} + (x_1 - x_2)F\begin{pmatrix}1\\0\end{pmatrix}\\ &= x_2\begin{pmatrix}0\\1\\2\end{pmatrix} + (x_1 - x_2)\begin{pmatrix}1\\1\\-1\end{pmatrix} = \begin{pmatrix}x_1 - x_2\\x_1\\-x_1 + 3x_2\end{pmatrix} = \begin{pmatrix}1&-1\\1&0\\-1&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}.\end{aligned}$$

The matrix of F is $[F] = \begin{pmatrix}1&-1\\1&0\\-1&3\end{pmatrix}$.

Remark. From the above examples, we see that, in order to know the value of a linear map F at any x, it suffices to know the value $F(v_1)$, …, $F(v_n)$ where v_1, …, v_n are vectors such that we can write for any x as

$$x = \alpha_1v_1 + \ldots + \alpha_nv_n,$$

for some $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$.

How to check if a given function $F : \mathbb{R}^n \to \mathbb{R}^m$ is linear or not?

- To show that F is linear, one can either show that:
  - There exists a matrix $A \in \mathbb{R}^{m\times n}$ with $F(x) = Ax$ (and hence, $A = [F]$), **or**

– **For all** $u, v \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ that $F(u+v) = F(u) + F(v)$ and $F(\lambda u) = \lambda F(u)$.

- To show that F is not linear, it suffices to give **one** example of $u, v \in \mathbb{R}^n$ with $F(u+v) \neq F(u) + F(v)$, or **one** example of $u \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ with $F(\lambda u) \neq \lambda F(u)$.

Example 18 The function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto e^x \end{aligned}$$

is not linear because the **one explicit example** with $u = v = 0$ gives:

$$\begin{aligned} f(u+v) = f(0+0) = f(0) = e^0 = 1, \\ f(u) + f(v) = f(0) + f(0) = e^0 + e^0 = 2. \end{aligned}$$

Therefore, $f(u+v) \neq f(u) + f(v)$ for the case $u = v = 0$.

What you should not do, when proving that this function is not linear, is to write "It is $e^{u+v} \neq e^u + e^v$ and therefore f is not linear". Even though $e^{u+v} \neq e^u + e^v$ is true for almost all $u, v \in \mathbb{R}$, there are cases where it is not true. For example, for $u = 1$ and $v = 0.4586\ldots$,

$$e^{u+v} = 4.3002\ldots = e^u + e^v.$$

Exercises

Exercise 16. Which of the following functions are linear maps?

$$\begin{aligned} f_1 : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto e^x\,, \end{aligned} \qquad \begin{aligned} f_2 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + 4x_2 \\ x_1 - x_2 \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} f_3 : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto 3x + 2\,, \end{aligned} \qquad \begin{aligned} f_4 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} x_1 + 2x_2 \\ x_1 x_2 \end{pmatrix}. \end{aligned}$$

Exercise 17. We define the following four functions:

$$\begin{aligned} f_1 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} 2x_1 + x_2 \\ x_1 x_2 \end{pmatrix}, \end{aligned} \qquad \begin{aligned} f_2 : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{2x}{x^2+4}\,, \end{aligned}$$

$$\begin{aligned} f_3 : \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto \begin{pmatrix} 3\cos(x) \\ 2\sin(x) \end{pmatrix}, \end{aligned} \qquad \begin{aligned} f_4 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} 2x_1 - x_2 \\ x_1 - 3x_2 \\ x_1 - x_2 \end{pmatrix}. \end{aligned}$$

(i) Calculate the image of each function, i.e. describe $\mathrm{im}(f_j)$ for $j = 1, 2, 3, 4$ as explicit as possible. If you can not find a mathematical description try to describe the elements of the image in words.
(ii) Decide for each function if it is injective and/or surjective and/or bijective.
(iii) Decide which of the above functions are linear maps.

Justify your answers in (ii) and (iii).

Exercise 18. Show that there exist a unique linear map $G : \mathbb{R}^2 \to \mathbb{R}^3$ with the property

$$G\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \qquad G\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

What is the value of $G(x)$ for an arbitrary $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$? Determine the matrix of G.

Exercise 19. Let $F : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear map. Show that F can not be injective.

Exercise 20. Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Show that the following two statements are equivalent:

(i) F is injective.

(ii) The only solution to $F(x) = 0$ is $x = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

To show that both statements are equivalent, you need to show that (i) implies (ii) and (ii) implies (i).

5

Linear maps in geometry

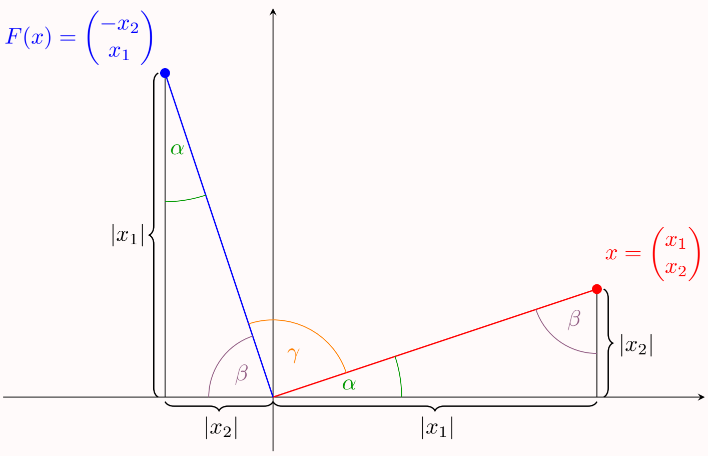
In the previous chapter, we introduced the concept of linear maps, which form the foundation of linear algebra and are essential tools for studying geometry. In this chapter, we will delve deeper into the topic of linear maps and discuss certain classes of them which have a geometric interpretation.

Example 19 Consider the function

$$F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2.$$
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

This function is indeed a linear map because

$$F\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$



We have

$$\begin{cases} \alpha + \beta + \gamma = 180^\circ \\ \alpha + \beta = 90^\circ \end{cases} \quad \Rightarrow \quad \gamma = 90^\circ.$$

Hence, F rotates x by 90° counterclockwise.

In Example 19, we say that x and $F(x)$ are "orthogonal" (meaning perpendicular) to each other. How to check if $x, y \in \mathbb{R}^2$ are orthogonal in general? What about $x, y \in \mathbb{R}^n$?

Definition 5.1 Let $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$, $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$.

(i) The **dot product** of u and v is defined by

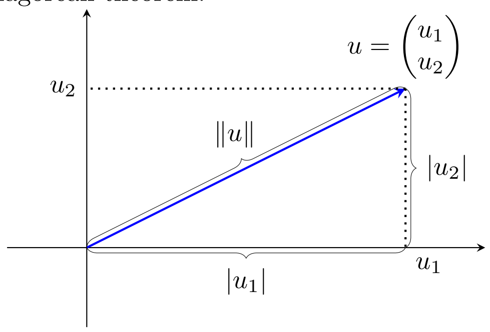
$$u \bullet v = u_1 v_1 + \cdots + u_n v_n \,.$$

(ii) u and v are **orthogonal** (to each other) if $u \bullet v = 0$.

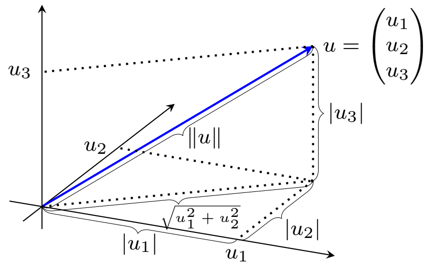
(iii) The **norm** of u is defined by

$$\|u\| = \sqrt{u \bullet u} = \sqrt{u_1^2 + \cdots + u_n^2}\,.$$

Remark. (i) The dot product (also called scalar product or inner product) allows us to speak about length and angle in $\mathbb{R}^n$. For $n = 2$ and $n = 3$, the norm of a vector is equal to its length by Pythagorean theorem:

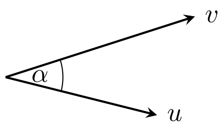


$$\|u\| = \sqrt{u_1^2 + u_2^2}$$



$$\|u\| = \sqrt{u_1^2 + u_2^2 + u_3^2}$$

Also, one can show that, for any non-zero vectors u and v,



$$u \bullet v = \|u\| \, \|v\| \cos(\alpha)$$
$$\Leftrightarrow \cos(\alpha) = \frac{u \bullet v}{\|u\| \, \|v\|}.$$

For $n > 3$, this gives the definition of an angle between u and v.

(ii) Write $u^T = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \in \mathbb{R}^{1 \times n}$ (T stands for "transpose", for which we will give a proper definition later) , then

$$u \bullet v = u^T v = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + \cdots + u_n v_n,$$

where $u^T v$ is the product of the matrix u^T and the vector v.

Proposition 5.2 *The dot product satisfies the following properties for all $u, v, w \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$:*

(i) $u \bullet v = v \bullet u$,

(ii) $u \bullet (v + w) = u \bullet v + u \bullet w$,

(iii) $u \bullet (\lambda v) = \lambda(u \bullet v)$.

Proof. (i) $u \bullet v = u_1 v_1 + \cdots + u_n v_n = v_1 u_1 + \cdots + v_n u_n = v \bullet u.$

(ii)
$$\begin{aligned} u \bullet (v+w) &= u_1(v_1+w_1) + \cdots + u_n(v_n+w_n) \\ &= (u_1v_1 + u_1w_1) + \cdots + (u_nv_n + u_nw_n) \\ &= (u_1v_1 + \cdots + u_nv_n) + (u_1w_1 + \cdots + u_nw_n) \\ &= u \bullet v + u \bullet w. \end{aligned}$$

(iii) $u \bullet (\lambda v) = u_1(\lambda v_1) + \cdots + u_n(\lambda v_n) = \lambda(u_1v_1 + \cdots + u_nv_n) = \lambda(u \bullet v).$ □

In the following, we will give examples of linear maps which have geometric interpretations.

1. Scaling

Let $\lambda > 0$ and define the linear map

$$\begin{aligned} h_\lambda : \mathbb{R}^n &\longrightarrow \mathbb{R}^n. \\ x &\longmapsto \lambda x \end{aligned}$$

This map is indeed linear because it has the matrix

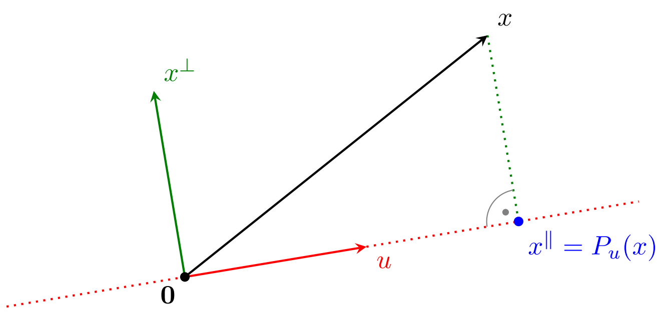
$$[h_\lambda] = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = \lambda I_n,$$

where I_n is the $n \times n$ identity matrix, the matrix of the identity map $\mathrm{id}_{\mathbb{R}^n}$. When we apply the map h_λ to a vector, we scale the norm (length) of this vector by a factor λ.

2. Orthogonal projection

Let $u \in \mathbb{R}^n$ with $u \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{0}$. We want to define a map $P_u : \mathbb{R}^n \to \mathbb{R}^n$ that sends a vector $x \in \mathbb{R}^n$ to another vector $x^\parallel$ such that $x = x^\perp + x^\parallel$, where

$$\begin{aligned} x^\parallel &= \lambda u \qquad \text{for some unknown } \lambda \in \mathbb{R}, \\ x^\perp \bullet u &= 0. \end{aligned}$$



To find λ, we do the following calculation

$$u \bullet x = u \bullet (x^\perp + x^\parallel) = u \bullet x^\perp + u \bullet x^\parallel = 0 + u \bullet (\lambda u) = \lambda(u \bullet u).$$

Hence, $\lambda = \dfrac{u \bullet x}{u \bullet u}$. Observe that $\mathrm{im}(P_u) = \{x \in \mathbb{R}^n \mid \exists \lambda \in \mathbb{R} : x = \lambda u\} = \{\lambda u \in \mathbb{R}^n \mid \lambda \in \mathbb{R}\}$.

Definition 5.3 Let $u \in \mathbb{R}^n$ with $u \neq \mathbf{0}$. We define the **orthogonal projection** P_u onto the line spanned by u as

$$\begin{aligned} P_u : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto \frac{u \bullet x}{u \bullet u} u\,. \end{aligned}$$

It is in our interest that the orthogonal projection is indeed a linear map.

Proposition 5.4 *$P_u : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map.*

Proof. For $x, y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$, using Proposition 5.2, we have

$$\begin{aligned} P_u(x+y) &= \frac{u \bullet (x+y)}{u \bullet u} u = \frac{u \bullet x + u \bullet y}{u \bullet u} u \\ &= \frac{u \bullet x}{u \bullet u} u + \frac{u \bullet y}{u \bullet u} u = P_u(x) + P_u(y), \end{aligned}$$

$$P_u(\lambda x) = \frac{u \bullet (\lambda x)}{u \bullet u} u = \frac{\lambda (u \bullet x)}{u \bullet u} u = \lambda P_u(x).$$

As a result, P_u is a linear map. □

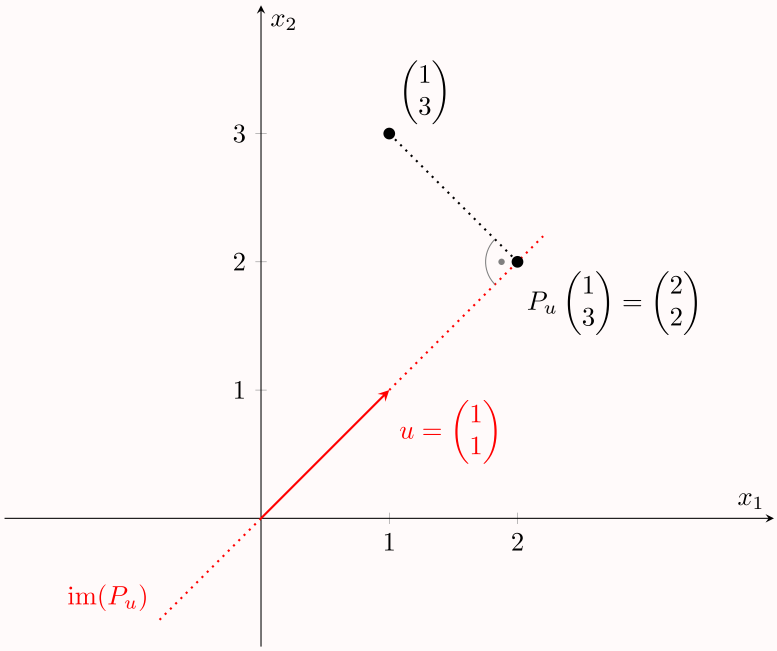
Example 20 For $n = 2$, consider $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. We have

$$\begin{aligned} u \bullet u &= 1 \cdot 1 + 1 \cdot 1 = 2, \\ u \bullet x &= 1 \cdot x_1 + 1 \cdot x_2 = x_1 + x_2. \end{aligned}$$

Then,

$$\begin{aligned} P_u(x) &= \frac{u \bullet x}{u \bullet u} u = \frac{x_1 + x_2}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{x_1+x_2}{2} \\ \frac{x_1+x_2}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = [P_u]x, \end{aligned}$$

where $[P_u] = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.



For example, we have

$$P_u\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix}\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}2\\2\end{pmatrix}.$$

In the case $n = 1$, the dot product is just the multiplication of real numbers, and we have

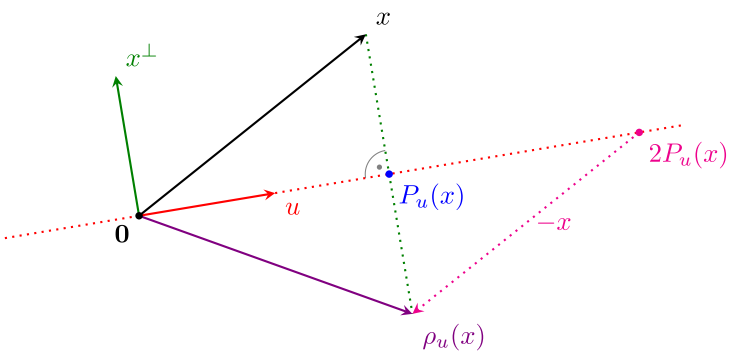
$$\begin{aligned} P_u : \mathbb{R} &\longrightarrow \mathbb{R}. \\ x &\longmapsto \frac{ux}{uu}u = x \end{aligned}$$

Hence, $P_u = \mathrm{id}_{\mathbb{R}}$ in the case $n = 1$. For $n > 1$, $P_u : \mathbb{R}^n \to \mathbb{R}^n$ is not injective and not surjective (check it yourself), and

$$\mathrm{im}(P_u) = \{\lambda u \mid \lambda \in \mathbb{R}\}.$$

3. Reflections

Now we want to define a map $\rho_u : \mathbb{R}^n \to \mathbb{R}^n$, which reflects $x \in \mathbb{R}^n$ along the line spanned by u.



To do that, we proceed as follows:

$$\begin{aligned}\rho_u(x) = x - 2x^{\perp} &= x - 2(x - P_u(x)) \\ &= 2P_u(x) - x \\ &= 2\frac{u \bullet x}{u \bullet u} u - x\end{aligned}$$

Definition 5.5 Let $u \in \mathbb{R}^n$ with $u \neq \mathbf{0}$. We define the **reflection** ρ_u along the line spanned by u as

$$\begin{aligned}\rho_u : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto 2\frac{u \bullet x}{u \bullet u} u - x\,.\end{aligned}$$

Again, this map is also a linear map.

Proposition 5.6 *$\rho_u : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a linear map.*

Proof. This is Exercise 21 □

Example 21 For $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, the ρ_u is the reflection along the diagonal. What is the matrix of ρ_u? We have

$$[\rho_u] = \begin{pmatrix} | & | \\ \rho_u(e_1) & \rho_u(e_2) \\ | & | \end{pmatrix},$$

where

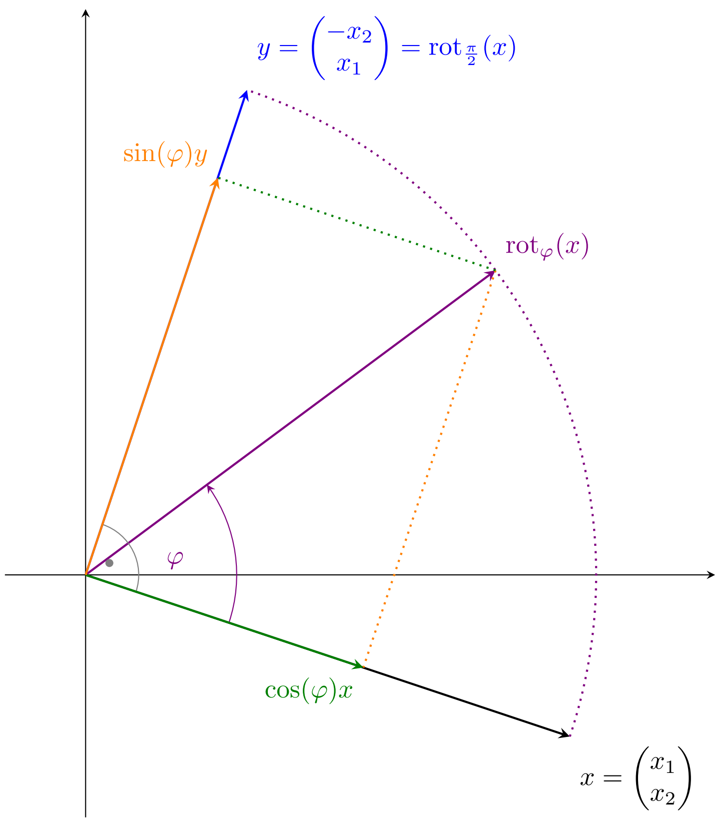
$$\begin{aligned}\rho_u(e_1) &= 2\frac{u \bullet e_1}{u \bullet u} u - e_1 = \left(2 \cdot \frac{1 \cdot 1 + 1 \cdot 0}{1 \cdot 1 + 1 \cdot 1}\right)\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ \rho_u(e_2) &= 2\frac{u \bullet e_2}{u \bullet u} u - e_2 = \left(2 \cdot \frac{1 \cdot 0 + 1 \cdot 1}{1 \cdot 1 + 1 \cdot 1}\right)\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.\end{aligned}$$

Hence, $[\rho_u] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

For any $n \geq 1$ and any $u \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, the map ρ_u is bijective with $\rho_u^{-1} = \rho_u$ because $\rho_u \circ \rho_u = \mathrm{id}_{\mathbb{R}^n}$ (this is supported with our common sense).

4. Rotations in $\mathbb{R}^2$

We want to define a map $\mathrm{rot}_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ that describes a counterclockwise rotation with angle $\varphi \in \mathbb{R}$.



To do that, we proceed as follows:

$$\begin{aligned}
\mathrm{rot}_\varphi(x) &= \cos(\varphi)x + \sin(\varphi)y \\
&= \cos(\varphi)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \sin(\varphi)\begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \\
&= \begin{pmatrix} \cos(\varphi)x_1 - \sin(\varphi)x_2 \\ \cos(\varphi)x_2 + \sin(\varphi)x_1 \end{pmatrix} \\
&= \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= [\mathrm{rot}_\varphi]x,
\end{aligned}$$

where $[\mathrm{rot}_\varphi] = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$.

Definition 5.7 For $\varphi \in \mathbb{R}$ the counterclockwise **rotation by an angle** φ (in $\mathbb{R}^2$) is given by

$$\begin{aligned}
\mathrm{rot}_\varphi : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\
x &\longmapsto \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} x\,.
\end{aligned}$$

We have $\mathrm{rot}_{\varphi_1} \circ \mathrm{rot}_{\varphi_2} = \mathrm{rot}_{\varphi_1+\varphi_2}$, i.e. rot_φ is invertible with inverse $\mathrm{rot}_{-\varphi}$ because

$$\mathrm{rot}_\varphi \circ \mathrm{rot}_{-\varphi} = \mathrm{rot}_0 = \mathrm{id}_{\mathbb{R}^2}\,.$$

Exercises

Exercise 21. Let $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^n$ be with $u \neq \mathbf{0}$.

(i) Show that the reflection $\rho_u : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a linear map.

(ii) Show that the matrix of the projection $P_u : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is given by

$$[P_u] = \frac{1}{u \bullet u} u u^T \in \mathbb{R}^{n \times n},$$

where $u^T = (u_1\, u_2\, \ldots\, u_n) \in \mathbb{R}^{1 \times n}$. Use this to give an expression for $[\rho_u]$.

Exercise 22. Show that for all $u \in \mathbb{R}^n$ with $u \neq \mathbf{0}$ the projection P_u and the reflection ρ_u satisfy for all $x \in \mathbb{R}^n$ the following two properties:

(i) $P_u(P_u(x)) = P_u(x)$.

(ii) $\rho_u(\rho_u(x)) = x$.

6

Composition of linear maps & Matrix multiplication

Linear maps are functions so we can compose them

$$
\begin{tikzcd}
\mathbb{R}^n \arrow[r, "F"] \arrow[rr, bend right, "G\circ F = GF"'] & \mathbb{R}^m \arrow[r, "G"] & \mathbb{R}^l.
\end{tikzcd}
$$

If F, G are linear maps, then does GF inherit the linearity property from them? The answer is yes by the following theorem.

Theorem 6.1 *If $F : \mathbb{R}^n \to \mathbb{R}^m$ and $G : \mathbb{R}^m \to \mathbb{R}^l$ are linear, then $GF : \mathbb{R}^n \to \mathbb{R}^l$ is linear.*

Proof. Since F and G are linear, we have for $x, y \in \mathbb{R}^n$,

$$GF(x+y) = G\big(F(x+y)\big) = G\big(F(x)+F(y)\big) = G\big(F(x)\big) + G\big(F(y)\big) = GF(x) + GF(y).$$

Also, we have for $\lambda \in \mathbb{R}, x \in \mathbb{R}^n$,

$$GF(\lambda x) = G\big(F(\lambda x)\big) = G\big(\lambda F(x)\big) = \lambda G\big(F(x)\big) = \lambda GF(x).$$

Therefore, GF is also a linear map. $\square$

A natural question follows: What is the matrix of GF?

Example 22 We consider the following linear maps

$$F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad x \longmapsto \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix} x,$$

$$G : \mathbb{R}^2 \longrightarrow \mathbb{R}^3, \quad x \longmapsto \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} x.$$

We have the matrices of the map F and G as follows:

$$[F] = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}, \qquad [G] = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

We want to calculate the matrix of $GF : \mathbb{R}^2 \to \mathbb{R}^3$. We have

$$[GF] = \begin{pmatrix} | & | \\ GF(e_1) & GF(e_2) \\ | & | \end{pmatrix},$$

where

$$GF(e_1) = G\big(F(e_1)\big) = G\left(F\begin{pmatrix}1\\0\end{pmatrix}\right) = G\begin{pmatrix}1\\3\end{pmatrix} = [G]\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}-2\\3\\1\end{pmatrix},$$

$$GF(e_2) = G\big(F(e_2)\big) = G\left(F\begin{pmatrix}0\\1\end{pmatrix}\right) = G\begin{pmatrix}2\\-1\end{pmatrix} = [G]\begin{pmatrix}2\\-1\end{pmatrix} = \begin{pmatrix}3\\-1\\2\end{pmatrix}.$$

Hence,

$$[GF] = \begin{pmatrix} | & | \\ GF(e_1) & GF(e_2) \\ | & | \end{pmatrix} = \begin{pmatrix} | & | \\ [G]\left(\begin{smallmatrix}1\\3\end{smallmatrix}\right) & [G]\left(\begin{smallmatrix}2\\-1\end{smallmatrix}\right) \\ | & | \end{pmatrix} = \begin{pmatrix} -2 & 3 \\ 3 & -1 \\ 1 & 2 \end{pmatrix}.$$

We see that in order to find $[GF]$, we multiply $[G]$ with each of the columns of $[F]$ to get each of the corresponding columns of $[GF]$, which motivates us to define the matrix multiplication as follows.

Definition 6.2 Let $A \in \mathbb{R}^{l\times m}$ and $B \in \mathbb{R}^{m\times n}$, where B has the columns $v_1, \dots, v_n \in \mathbb{R}^m$, i.e.

$$B = \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix}.$$

Then the **product of A and B** is the $l \times n$-matrix with columns $Av_1, \dots, Av_n \in \mathbb{R}^l$, i.e.

$$A \cdot B = \begin{pmatrix} | & & | \\ Av_1 & \dots & Av_n \\ | & & | \end{pmatrix} \in \mathbb{R}^{l\times n}.$$

Example 23 1) $A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}$ then

$$A \cdot B = \begin{pmatrix} | & | \\ A\left(\begin{smallmatrix}1\\3\end{smallmatrix}\right) & A\left(\begin{smallmatrix}2\\-1\end{smallmatrix}\right) \\ | & | \end{pmatrix} = \begin{pmatrix} -2 & 3 \\ 3 & -1 \\ 1 & 2 \end{pmatrix}.$$

Compare this with Example 22, where $[G] = A$, $[F] = B$ and $[GF] = A \cdot B$.

2) Consider the following matrix multiplication:

$$\begin{pmatrix} 0 & 3 & 0 \\ 1 & -1 & 1 \\ 1 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 1 & 1 \\ 3 & 2 \end{pmatrix}.$$

3) Consider the following matrix multiplication:

$$\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In the next lecture, we will learn that $\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$ is the inverse of $\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}$.

We already know that each matrix corresponds to a linear map. The following theorem confirms that the matrix multiplication is indeed equivalent to the composition of the corresponding linear maps.

Theorem 6.3 *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ and $G : \mathbb{R}^m \to \mathbb{R}^l$ be linear maps. Then the matrix of GF is given by the product of the matrices of G and F, i.e.*

$$[GF] = [G] \cdot [F] \,.$$

Proof. We have $[F] = \begin{pmatrix} | & & | \\ F(e_1) & \cdots & F(e_n) \\ | & & | \end{pmatrix}$ and

$$\begin{aligned} [G] \cdot [F] &= \begin{pmatrix} | & & | \\ [G]F(e_1) & \cdots & [G]F(e_n) \\ | & & | \end{pmatrix} \\ &= \begin{pmatrix} | & & | \\ G\big(F(e_1)\big) & \cdots & G\big(F(e_n)\big) \\ | & & | \end{pmatrix} \\ &= \begin{pmatrix} | & & | \\ GF(e_1) & \cdots & GF(e_n) \\ | & & | \end{pmatrix} \\ &= [GF]. \end{aligned}$$

Hence, we have proved the theorem. □

Example 24 1) Given the linear map

$$\begin{aligned} F : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2, \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} 2x_1 - x_2 \\ x_1 + 3x_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \end{aligned}$$

what is the matrix of $F \circ F$?

- By hand:

$$\begin{aligned} F \circ F(x) = F\big(F(x)\big) = F \begin{pmatrix} 2x_1 - x_2 \\ x_1 + 3x_2 \end{pmatrix} &= \begin{pmatrix} 2(2x_1 - x_2) - (x_1 + 3x_2) \\ (2x_1 - x_2) + 3(x_1 + 3x_2) \end{pmatrix} \\ &= \begin{pmatrix} 3x_1 - 5x_2 \\ 5x_1 + 8x_2 \end{pmatrix} = \begin{pmatrix} 3 & -5 \\ 5 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &\Rightarrow [FF] = \begin{pmatrix} 3 & -5 \\ 5 & 8 \end{pmatrix}. \end{aligned}$$

- Using Theorem 6.3:

$$[FF] = [F] \cdot [F] = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & -5 \\ 5 & 8 \end{pmatrix}.$$

2) In the last lecture, we define the rotation by angle φ as the linear map

$$\begin{aligned} \mathrm{rot}_\varphi : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2, \\ x &\longmapsto \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} x. \end{aligned}$$

The composition $\mathrm{rot}_{\varphi_1} \circ \mathrm{rot}_{\varphi_2}$ has the meaning as the rotation by φ_2 and then by φ_1. We also mention in the last lecture that

$$\mathrm{rot}_{\varphi_1} \circ \mathrm{rot}_{\varphi_2} = \mathrm{rot}_{\varphi_1+\varphi_2}\,. \tag{$*$}$$

Using Theorem 6.3, we have

$$\begin{aligned}
[\mathrm{rot}_{\varphi_1} \circ \mathrm{rot}_{\varphi_2}] &= [\mathrm{rot}_{\varphi_1}] \cdot [\mathrm{rot}_{\varphi_2}] \\
&= \begin{pmatrix} \cos(\varphi_1) & -\sin(\varphi_1) \\ \sin(\varphi_1) & \cos(\varphi_1) \end{pmatrix} \begin{pmatrix} \cos(\varphi_2) & -\sin(\varphi_2) \\ \sin(\varphi_2) & \cos(\varphi_2) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\varphi_1)\cos(\varphi_2) - \sin(\varphi_1)\sin(\varphi_2) & -\cos(\varphi_1)\sin(\varphi_2) - \sin(\varphi_1)\cos(\varphi_2) \\ \sin(\varphi_1)\cos(\varphi_2) + \cos(\varphi_1)\sin(\varphi_2) & -\sin(\varphi_1)\sin(\varphi_2) + \cos(\varphi_1)\cos(\varphi_2) \end{pmatrix}.
\end{aligned}$$

In addition, $[\mathrm{rot}_{\varphi_1+\varphi_2}] = \begin{pmatrix} \cos(\varphi_1+\varphi_2) & -\sin(\varphi_1+\varphi_2) \\ \sin(\varphi_1+\varphi_2) & \cos(\varphi_1+\varphi_2) \end{pmatrix}$. By using $(*)$, we obtain the angle sum identities:

$$\begin{aligned}
\cos(\varphi_1+\varphi_2) &= \cos(\varphi_1)\cos(\varphi_2) - \sin(\varphi_1)\sin(\varphi_2), \\
\sin(\varphi_1+\varphi_2) &= \sin(\varphi_1)\cos(\varphi_2) + \cos(\varphi_1)\sin(\varphi_2).
\end{aligned}$$

The matrix multiplication has the following properties.

Proposition 6.4 *For all $A \in \mathbb{R}^{l\times m}, B, D \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{n\times p}$ and $\lambda \in \mathbb{R}$ we have*

(i) $A \cdot I_m = I_l \cdot A = A$, *where I_m denotes the $m \times m$-identity matrix.*

(ii) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

(iii) $A \cdot (B + D) = A \cdot B + A \cdot D$.

(iv) $(B + D) \cdot C = B \cdot C + D \cdot C$.

(v) $\lambda(A \cdot B) = (\lambda A) \cdot B = A \cdot (\lambda B)$.

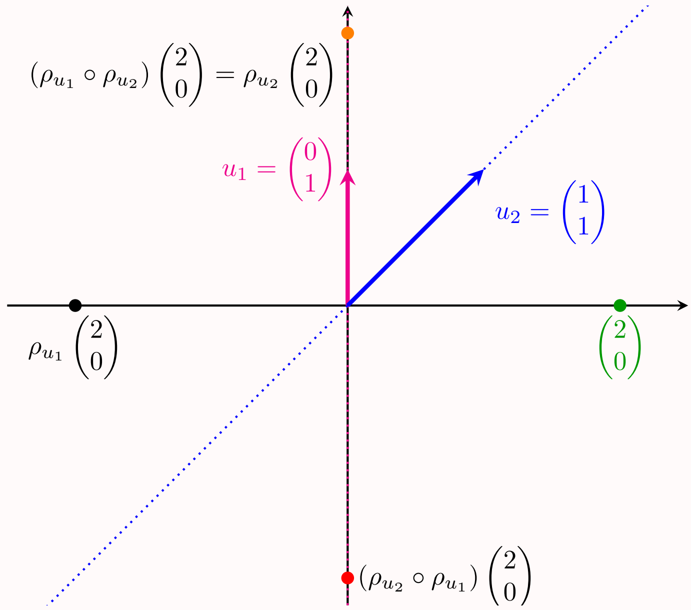
Proof. Check by yourself. Similar to Proposition 2.4. □

Remark. If A, $B \in \mathbb{R}^{n\times n}$ then in general we have $A \cdot B \neq B \cdot A$. For example,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ but } \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Example 25 In the last lecture, for $u \neq \mathbf{0}$, ρ_u denotes the reflection along the line spanned by u. Consider $u_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $u_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and their corresponding reflections with their matrices:

$$[\rho_{u_1}] = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$[\rho_{u_2}] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$



We have

$$\rho_{u_1}\left(\rho_{u_2}\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) = \rho_{u_1}\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$
$$\rho_{u_2}\left(\rho_{u_1}\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) = \rho_{u_2}\begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$$

Hence, $\rho_{u_1}\left(\rho_{u_2}\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) \neq \rho_{u_2}\left(\rho_{u_1}\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right)$. Generally, $\rho_{u_1}(\rho_{u_2}(x)) \neq \rho_{u_2}(\rho_{u_1}(x))$ for any vector $x \in \mathbb{R}^2$ because

$$[\rho_{u_1} \circ \rho_{u_2}] = [\rho_{u_1}] \cdot [\rho_{u_2}] = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$
$$[\rho_{u_2} \circ \rho_{u_1}] = [\rho_{u_2}] \cdot [\rho_{u_1}] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

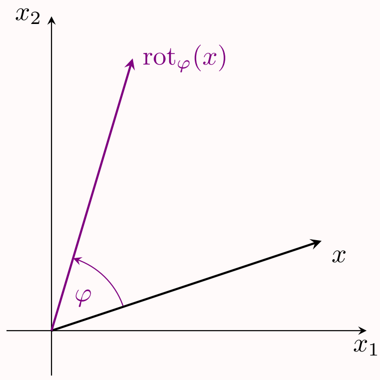
$$\Rightarrow [\rho_{u_1} \circ \rho_{u_2}] \neq [\rho_{u_2} \circ \rho_{u_1}].$$

Therefore, reflecting first along u_2 and then u_1 is different to first reflecting along u_1 and then u_2.

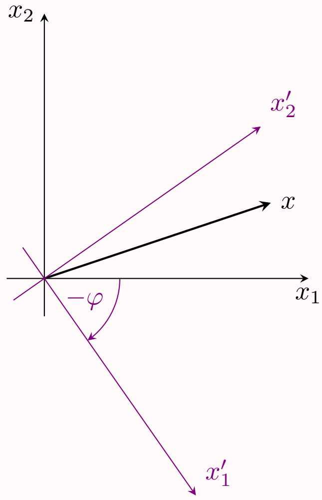
Remark. Notice that sometimes (really rare) we have $A \cdot B = B \cdot A$. For example,

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

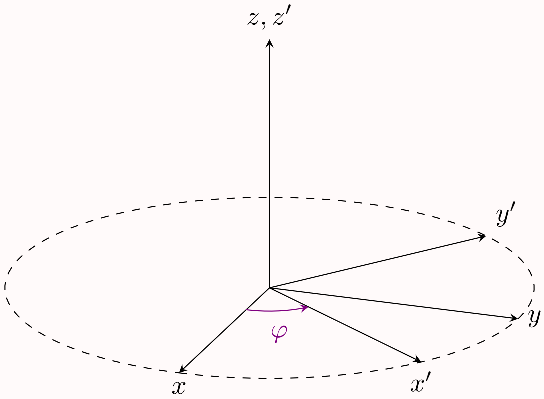
Example 26 (Rotations in $\mathbb{R}^3$) There are two interpretations of rotation maps. Take the two-dimensional rotation rot_φ with $\varphi \in \mathbb{R}$ for example. For any vector $x \in \mathbb{R}^2$, the vector $\mathrm{rot}_\varphi(x)$ can be interpreted in two ways. In Chapter 5, we let the coordinate system be unchanged so $\mathrm{rot}_\varphi(x)$ is a new vector in the same coordinate system. In this case, the map rot_φ is called a **active transformation**.



On the other hand, the vector x can be kept unchanged and the coordinate system is rotated instead. As a result, the components of $\mathrm{rot}_\varphi(x)$ are the coordinates of x in this new coordinate system. In this case, rot_φ is called a **passive transformation**.



Hence, for any $\varphi > 0$, the map rot_φ corresponds to a counterclockwise rotation by the angle φ when applied to the vector x. On the other hand, it also corresponds to a clockwise rotation by the angle φ when applied to the coordinate system. The same duality of roles often occurs with many transformations in physics.
For the following discussion of rotations in three dimension, we will interpret those rotations as passive transformations. The simplest three-dimensional rotations are rotations around any coordinate axes. For example, consider the following rotation.



Consider a vector $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$. This rotation is similar to a rotation in $\mathbb{R}^2$ except that there is one more component which is left unchanged. Hence, the new coordinates are as follows.

$$\begin{aligned} x' &= \cos(\varphi)x + \sin(\varphi)y \\ y' &= -\sin(\varphi)x + \cos(\varphi)y \\ z' &= z \end{aligned}$$

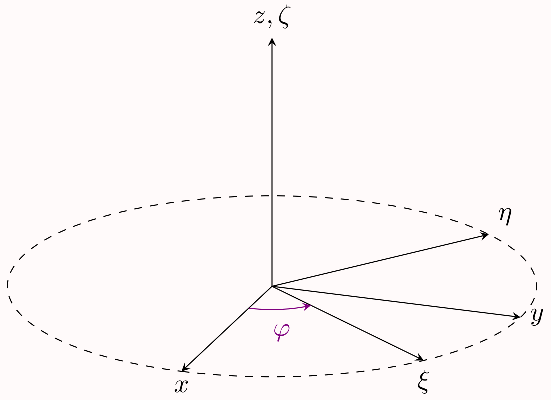
The new coordinates can be written as $v' = R_z(-\varphi)v$, where the matrix $R_z(\theta)$ is defined for any $\theta \in \mathbb{R}$ as

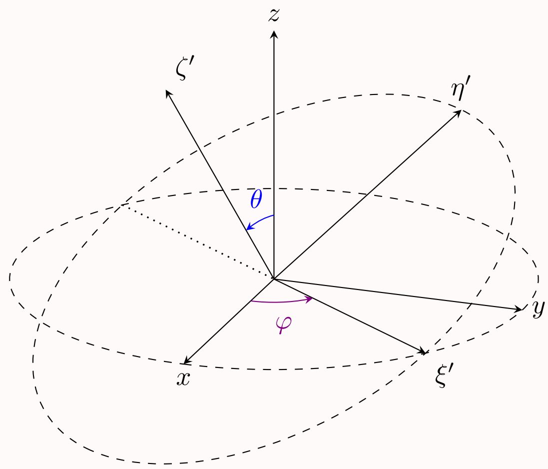
$$R_z(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

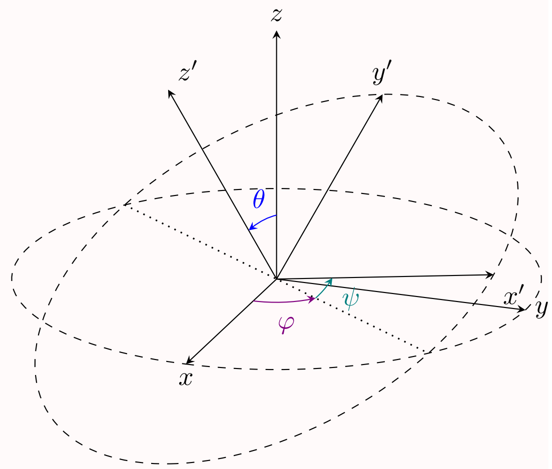
Similarly, we can define the two matrices $R_x(\theta)$ and $R_y(\theta)$ from the rotations around x- and y-axes as follows:

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}, \qquad R_y(\theta) = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix}.$$

Recall that every rotations rot_φ in $\mathbb{R}^2$ is specified by the angle φ. In $\mathbb{R}^3$, we need three parameters to describe any rotations. There are many sets of parameters that can be used but the most common and useful ones are **Euler angles** or **Eulerian angles**. The idea behind this is that every rotations can be decomposed into three successive rotations each of which is about one of the axes, with the condition that no two successive rotations can be about the same axis. Hence, there are totally 12 possible conventions in defining the Euler angles (in a right-handed coordinate system). We want to introduce 3 convenctions which are widely used in physics and engineering.
The first convention here is used widely in celestial mechanics, applied mechanics, and frequently in molecular and solid-state physics.







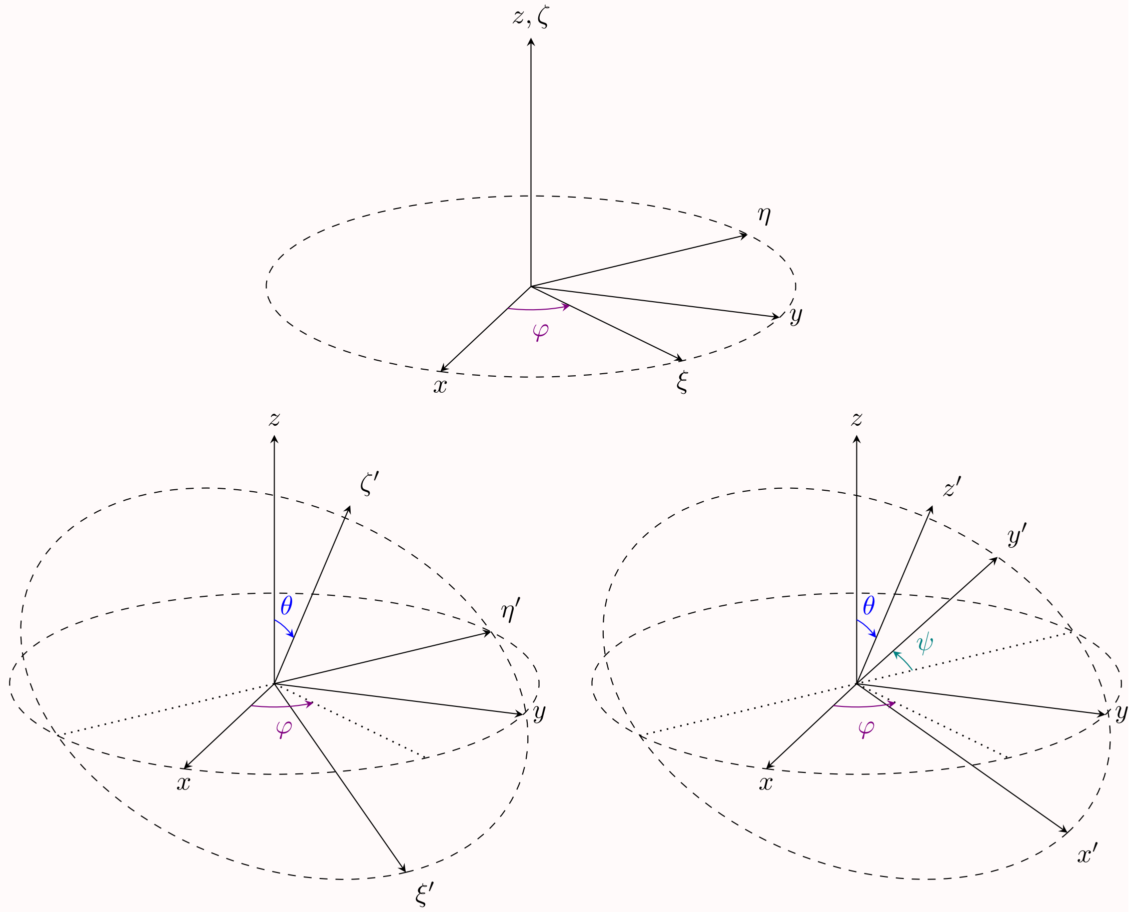
Consider any $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$. We get angles $u = \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}$, $u' = \begin{pmatrix} \xi' \\ \eta' \\ \zeta' \end{pmatrix}$, $v' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \in \mathbb{R}^3$ after each rotations as follows: $u = R_z(-\varphi)v$, $u' = R_x(-\theta)u$, and $v' = R_z(-\psi)u'$. Hence, we get the rotated vector v' from v as follows:

$$v' = R_z(-\psi)u' = R_z(-\psi)R_x(-\theta)u = R_z(-\psi)R_x(-\theta)R_z(-\varphi)v.$$

As a result, we get the matrix $R(\varphi, \theta, \psi)$ describing any general three-dimensional rotations as follows:

$$\begin{aligned}
&R(\varphi, \theta, \psi) \\
=&R_z(-\psi)R_x(-\theta)R_z(-\varphi) \\
=&\begin{pmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\varphi) & \sin(\varphi) & 0 \\ -\sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
=&\begin{pmatrix} \cos(\psi)\cos(\varphi) - \cos(\theta)\sin(\varphi)\sin(\psi) & \cos(\psi)\sin(\varphi) + \cos(\theta)\cos(\varphi)\sin(\psi) & \sin(\psi)\sin(\theta) \\ -\sin(\psi)\cos(\varphi) - \cos(\theta)\sin(\varphi)\cos(\psi) & -\sin(\psi)\sin(\varphi) + \cos(\theta)\cos(\varphi)\cos(\psi) & \cos(\psi)\sin(\theta) \\ \sin(\theta)\sin(\varphi) & -\sin(\theta)\cos(\varphi) & \cos(\theta) \end{pmatrix}.
\end{aligned}$$

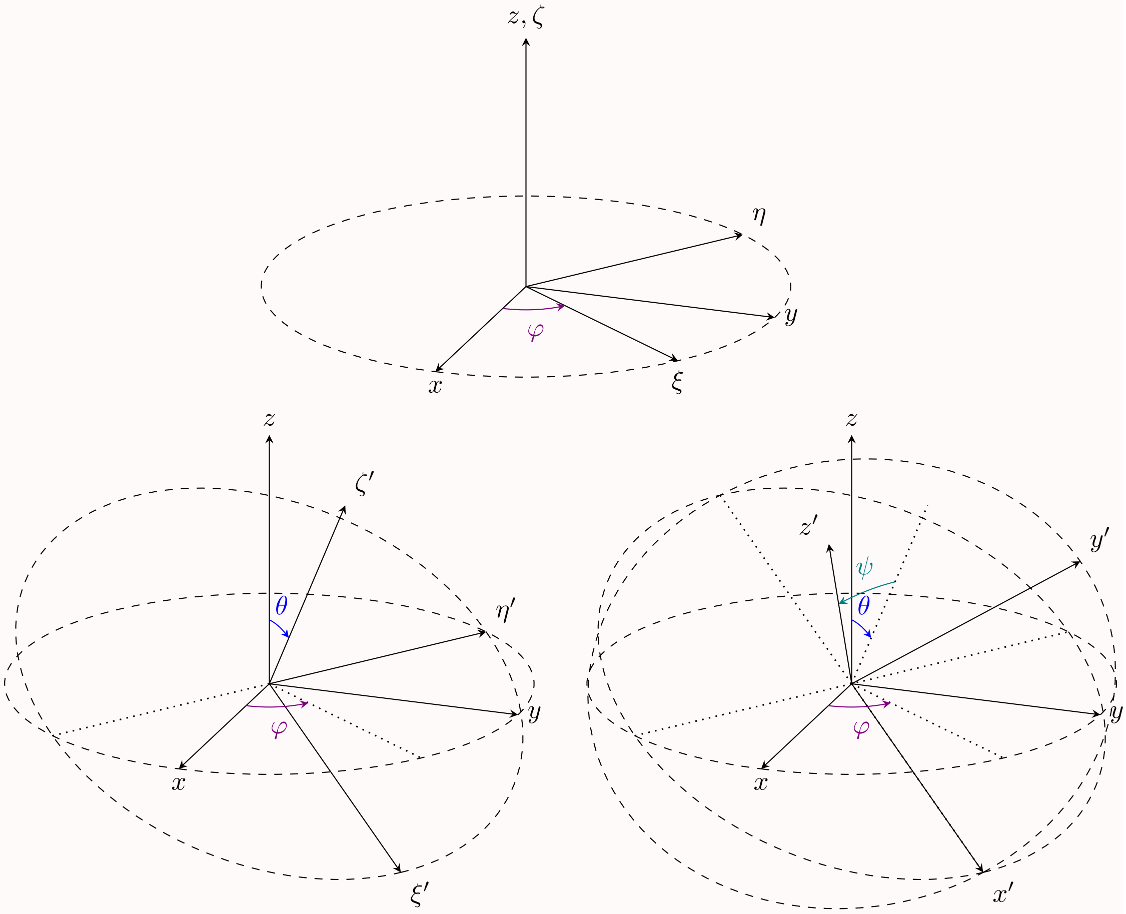
The second convention is used widely in quantum mechanics, nuclear physics, and particle physics.



Similarly, we get

$$
\begin{aligned}
&R(\varphi,\theta,\psi)\\
=&R_z(-\psi)R_y(-\theta)R_z(-\varphi)\\
=&\begin{pmatrix}\cos(\psi) & \sin(\psi) & 0\\ -\sin(\psi) & \cos(\psi) & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\cos(\theta) & 0 & -\sin(\theta)\\ 0 & 1 & 0\\ \sin(\theta) & 0 & \cos(\theta)\end{pmatrix}\begin{pmatrix}\cos(\varphi) & \sin(\varphi) & 0\\ -\sin(\varphi) & \cos(\varphi) & 0\\ 0 & 0 & 1\end{pmatrix}\\
=&\begin{pmatrix}-\sin(\psi)\sin(\varphi)+\cos(\theta)\cos(\varphi)\cos(\psi) & \sin(\psi)\cos(\varphi)+\cos(\theta)\sin(\varphi)\cos(\psi) & -\cos(\psi)\sin(\theta)\\ -\cos(\psi)\sin(\varphi)-\cos(\theta)\cos(\varphi)\sin(\psi) & \cos(\psi)\cos(\varphi)-\cos(\theta)\sin(\varphi)\sin(\psi) & \sin(\psi)\sin(\theta)\\ \sin(\theta)\cos(\varphi) & \sin(\theta)\sin(\varphi) & \cos(\theta)\end{pmatrix}.
\end{aligned}
$$

The last convention, also called **Tait-Bryan angles**, is widely used in engineering applications relating to the orientation of moving vehicles such as aircraft and satellites.



In a similar fashion, we get

$$
\begin{aligned}
&R(\varphi,\theta,\psi)\\
=&R_x(-\psi)R_y(-\theta)R_z(-\varphi)\\
=&\begin{pmatrix}1 & 0 & 0\\ 0 & \cos(\psi) & \sin(\psi)\\ 0 & -\sin(\psi) & \cos(\psi)\end{pmatrix}\begin{pmatrix}\cos(\theta) & 0 & -\sin(\theta)\\ 0 & 1 & 0\\ \sin(\theta) & 0 & \cos(\theta)\end{pmatrix}\begin{pmatrix}\cos(\varphi) & \sin(\varphi) & 0\\ -\sin(\varphi) & \cos(\varphi) & 0\\ 0 & 0 & 1\end{pmatrix}\\
=&\begin{pmatrix}\cos(\theta)\cos(\varphi) & \cos(\theta)\sin(\varphi) & -\sin(\theta)\\ \sin(\psi)\sin(\theta)\cos(\varphi)-\cos(\psi)\sin(\varphi) & \sin(\psi)\sin(\theta)\sin(\varphi)+\cos(\psi)\cos(\varphi) & \cos(\theta)\sin(\psi)\\ \cos(\psi)\sin(\theta)\cos(\varphi)+\sin(\psi)\sin(\varphi) & \cos(\psi)\sin(\theta)\sin(\varphi)-\sin(\psi)\cos(\varphi) & \cos(\theta)\cos(\psi)\end{pmatrix}
\end{aligned}
$$

In this case, the three parameters have their names: the angle φ of the rotation about the vertical axis (z- or ζ-axis) is the **heading** or **yaw** angle; the angle θ of the rotation around a perpendicular axis (y- or η-axis) fixed in the vehicle and normal to the figure axis (x- or ξ-axis) is the **pitch** or *attitude* angle; the angle ψ of the rotation about the figure axis of the vehicle is the **roll** or **bank** angle.
Furthermore, there is another useful set of parameters called the **Cayley-Klein parameters**. This set comprises 4 parameters that are better than the Euler angles to use in numerical computation due to the large number of trigonometric functions involved when using the Euler angles. Besides, the four-parameter sets are also useful in branches of physics, wherever rotations or rotational symmetry are involved. However, we will not introduce it here because complex numbers are involved. For more details, see [G].

Exercises

Exercise 23. Let $u = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $d = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $x = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$.

(i) Calculate the matrices $[P_u]$ and $[\rho_u]$ in this special case.

(ii) Calculate the following vectors and draw them in one picture together with u, d and x

$$P_u(x), \qquad \rho_u(x), \qquad (P_u \circ P_d)(x), \qquad \mathrm{rot}_{\frac{\pi}{2}}(x),$$
$$(P_u \circ \mathrm{rot}_{\frac{\pi}{2}})(x), \quad (\mathrm{rot}_{\frac{\pi}{2}} \circ P_u)(x), \quad (P_d \circ \mathrm{rot}_{\frac{\pi}{2}} \circ P_u)(x)\,.$$

7

The inverse of a linear map

In Chapter 3, we learned that a function $f : X \to Y$ is invertible if there exists a function $g : Y \to X$ such that for every $x \in X$ and $y \in Y$, $g(f(x)) = x$ and $f(g(y)) = y$. Then, $g = f^{-1}$ and it is called the inverse of f.

We also saw that invertibility is equivalent to bijectivity. A function f is bijective if for every $y \in Y$, there uniquely exists $x \in X$ such that $y = f(x)$. Equivalently, a function is bijective when it is injective and surjective.

In this chapter, we want to answer the following questions:

(a) When is a linear map invertible?

(b) Is the inverse also linear?

(c) How to calculate the inverse?

Example 27 Consider the linear map

$$\begin{aligned} F : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2, \\ x &\longmapsto \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} x \,. \end{aligned}$$

Is F invertible?

Take $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ and check if $F(x) = y$ has a unique solution.

$$-2 \left(\begin{array}{cc|c} 1 & 3 & y_1 \\ 2 & 4 & y_2 \end{array}\right) \sim -\tfrac{1}{2} \left(\begin{array}{cc|c} 1 & 3 & y_1 \\ 0 & -2 & y_2 - 2y_1 \end{array}\right) \sim -3 \left(\begin{array}{cc|c} 1 & 3 & y_1 \\ 0 & 1 & y_1 - \frac{1}{2}y_2 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & -2y_1 + \frac{3}{2}y_2 \\ 0 & 1 & y_1 - \frac{1}{2}y_2 \end{array}\right)$$

We get the unique solution

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2y_1 + \frac{3}{2}y_2 \\ y_1 - \frac{1}{2}y_2 \end{pmatrix} = \begin{pmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} .$$

Hence, F is invertible and its inverse is

$$\begin{aligned} F^{-1} : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2, \\ y &\longmapsto \begin{pmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} y. \end{aligned}$$

The **rank of a linear map** $F : \mathbb{R}^n \to \mathbb{R}^m$ is defined by the rank of its matrix, i.e. $\mathrm{rk}(F) := \mathrm{rk}([F])$. The following theorem answers the question about when a linear map is invertible.

Theorem 7.1 *A linear map $F : \mathbb{R}^n \to \mathbb{R}^m$ is invertible if and only if $m = n = \mathrm{rk}(F)$.*

Remark. This theorem is in the format "Statement 1 if and only if Statement 2", which means that both the statements are either true or false. In addition, it also means both "Statement 1 implies Statement 2" and "Statement 2 implies Statement 1". To prove this theorem, we need to prove both the implications.

Proof. F is invertible $\iff$ $[F]x = y$ has a unique solution for all $y \in \mathbb{R}^m$.

$$\begin{array}{cc} & \xleftarrow{\quad n+1 \quad} \\ \updownarrow m & \Big([F] \mid y\Big) \end{array} \sim \ldots \sim \big(B \mid z\big) = \mathrm{rref}\Big([F] \mid y\Big),$$

where $B = \mathrm{rref}\big([F]\big)$ and $z \in \mathbb{R}^m$ can be an arbitrary vector depending on y.

In one direction, suppose F is invertible. We want to show that $n = m = \mathrm{rk}(F) = \mathrm{rk}\big([F]\big)$.
If $\mathrm{rk}(F) < m$, then

$$\big(B \mid z\big) = \left(\begin{array}{c|c} * & \\ & z \\ \hline 0 & \end{array}\right) \updownarrow m$$

Hence, no solution for some z (and thus, some y). Therefore, $\mathrm{rk}(F) = m$.
If $\mathrm{rk}(F) < n$, then

$$\begin{array}{cc} & \xleftarrow{\qquad\qquad n \qquad\qquad} \\ \big(B \mid z\big) = & \left(\begin{array}{cccccc|c} 1 & & & * & & & \\ & \ddots & & * & & & \\ & & 1 & * & & & z \\ & & & 0 & 1 & & \\ & & & \vdots & & \ddots & \\ & & & 0 & & & \end{array}\right) \end{array}$$

In this case, there are columns without pivot elements so there is no unique solution (infinitely many solutions). Therefore, $\mathrm{rk}(F) = n$.

Conversely, if $m = n = \mathrm{rk}(F)$, then $B = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = I_n$. Hence, $[F]x = y$ has a unique solution for all $y \in \mathbb{R}^m$; therefore, F is invertible. $\square$

Next, the following proposition confirms that the inverse of a linear map is also linear.

Proposition 7.2 *If a linear map $F : \mathbb{R}^n \to \mathbb{R}^n$ is invertible, then its inverse F^{-1} is also linear.*

Proof. Let $u, v \in \mathbb{R}^n$. Set $x = F^{-1}(u)$ and $y = F^{-1}(v)$, i.e. $F(x) = u$ and $F(y) = v$.
Then

$$\begin{aligned} F^{-1}(u) + F^{-1}(v) = x + y = \mathrm{id}(x+y) &= F^{-1}F(x+y) \\ &= F^{-1}\big(F(x) + F(y)\big) \qquad \text{(because F is linear)} \\ &= F^{-1}(u+v). \end{aligned}$$

In addition, for $\lambda \in \mathbb{R}$,

$$\lambda F^{-1}(u) = \lambda x = F^{-1}F(\lambda x) = F^{-1}\big(\lambda F(x)\big) = F^{-1}(\lambda u).$$

Hence, F^{-1} is linear. □

Since each linear map corresponds to a matrix, we can define the inverse of a matrix corresponding to a invertible linear map as follows.

Definition 7.3 If $A \in \mathbb{R}^{n\times n}$ is the matrix of an invertible linear map $F : \mathbb{R}^n \to \mathbb{R}^n$ (i.e. $A = [F]$), then we define the **inverse of** A by $A^{-1} := \big[F^{-1}\big]$.

Naturally, we can ask when a matrix is invertible, which is answered by the following theorem.

Theorem 7.4 *The inverse of $A \in \mathbb{R}^{n\times n}$ exists (A is invertible) if and only if $\mathrm{rref}(A) = I_n$.*

Proof. This theorem follows from the proof of Theorem 7.1. □

The inverses of matrices have the following properties.

Proposition 7.5 *If $A, B \in \mathbb{R}^{n\times n}$ are invertible we have*
(i) $AA^{-1} = A^{-1}A = I_n$,
(ii) $(BA)^{-1} = A^{-1}B^{-1}$.

Proof. Suppose A and B are matrices of linear maps $F : \mathbb{R}^n \to \mathbb{R}^n$ and $G : \mathbb{R}^n \to \mathbb{R}^n$, respectively.

(i) Using Theorem 6.3, we have

$$\begin{aligned} AA^{-1} &= [F] \cdot \big[F^{-1}\big] = \big[F \circ F^{-1}\big] = [\mathrm{id}_{\mathbb{R}^n}] = I_n, \\ A^{-1}A &= \big[F^{-1}\big] \cdot [F] = \big[F^{-1} \circ F\big] = [\mathrm{id}_{\mathbb{R}^n}] = I_n. \end{aligned}$$

Hence, $AA^{-1} = A^{-1}A = I_n$.

(ii) We have $BA = [G] \cdot [F] = [G \circ F]$ and $(G \circ F)^{-1} = F^{-1} \circ G^{-1}$. Therefore,

$$(BA)^{-1} = \big[(G \circ F)^{-1}\big] = \big[F^{-1} \circ G^{-1}\big] = \big[F^{-1}\big] \cdot \big[G^{-1}\big] = A^{-1}B^{-1}.$$ □

In Example 27, we determined the inverse of $A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$ by solving the linear system $Ax = y$. In general, if we want to determine the inverse of $A \in \mathbb{R}^{n\times n}$, we can use the following algorithm (a variant of Gaussian elimination).

Algorithm 7.6 (Gauss-Jordan elimination) Given a matrix $A \in \mathbb{R}^{n\times n}$, use the Gaussian elimination to bring the augmented matrix $(A \mid I_n)$ to the row-reduced echelon form $(B \mid C)$. There are 2 scenarios:

- If $B \neq I_n$, then A is not invertible.
- If $B = I_n$, then A is invertible and $C = A^{-1}$. That is, we get

$$(A \mid I_n) \sim \ldots \sim \left(I_n \mid A^{-1}\right) = \operatorname{rref}(A \mid I_n).$$

Example 28 Determine the inverse of $A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$.

$$(A \mid I_2) = \begin{matrix} -2 \\ \hookrightarrow \end{matrix} \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array}\right) \sim \begin{matrix} \\ -\frac{1}{2} \end{matrix} \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & -2 & -2 & 1 \end{array}\right)$$

$$\sim \begin{matrix} \rightarrow \\ -3 \end{matrix} \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & -\frac{1}{2} \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & -2 & \frac{3}{2} \\ 0 & 1 & 1 & -\frac{1}{2} \end{array}\right) = \left(I_2 \mid A^{-1}\right).$$

We can check that

$$AA^{-1} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = A^{-1}A.$$

(Definition 7.7 and Theorem 7.8 are just a remark and they are not so important for the rest of this course. They will appear again in detail in Linear Algebra II)

Definition 7.7 For $\lambda \in \mathbb{R}$ with $\lambda \neq 0$ and $1 \leq i, j \leq n$ we define the **elementary matrices** $R_i^{\lambda,j}, R_i^{\lambda}, R_{i,j} \in \mathbb{R}^{n\times n}$ by

$$R_i^{\lambda,j} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & \lambda & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}, \quad R_i^{\lambda} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \lambda & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}, \quad R_{i,j} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & & 1 \\ & & & \ddots & \\ & & 1 & & 0 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}.$$

Here the λ in $R_i^{\lambda,j}$ is in the i-th row and j-th column, in R_i^{λ} it is in the i-th row, and in $R_{i,j}$ the 0 are on the diagonal in the i-th row and j-th column.

Multiplying with an elementary matrix from the left corresponds to the elementary row operations (Definition 2.6)

(R1) Multiyplying with $R_i^{\lambda,j}$: Add λ-times row j to row i.

(R2) Multiyplying with R_i^{λ}: Multiply row j by λ. ($\lambda \neq 0$)

(R3) Multiyplying with $R_{i,j}$: Change row i and j.

Taking a look at Example 28 again, we have

$$A = \begin{matrix} -2 \\ \hookrightarrow \end{matrix} \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \sim \begin{matrix} \\ -\frac{1}{2} \end{matrix} \begin{pmatrix} 1 & 3 \\ 0 & -2 \end{pmatrix} \sim \begin{matrix} \rightarrow \\ -3 \end{matrix} \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2,$$

where each step corresponds to the multiplication with an elementary matrix:

$$R_2^{-2,1}\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix} = \begin{pmatrix}1 & 0\\ -2 & 1\end{pmatrix}\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix} = \begin{pmatrix}1 & 3\\ 0 & -2\end{pmatrix},$$
$$R_2^{-1/2}\begin{pmatrix}1 & 3\\ 0 & -2\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}1 & 3\\ 0 & -2\end{pmatrix} = \begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix},$$
$$R_1^{-3,2}\begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & -3\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

As a result, we get

$$R_1^{-3,2}R_2^{-1/2}R_2^{-2,1}A = I_2$$
$$\Leftrightarrow \quad A = \left(R_1^{-3,2}R_2^{-1/2}R_2^{-2,1}\right)^{-1} = \left(R_2^{-2,1}\right)^{-1}\left(R_2^{-1/2}\right)^{-1}\left(R_1^{-3,2}\right)^{-1} = R_2^{2,1}R_2^{1/2}R_1^{3,2},$$

where we use the fact that all row operations are reversible, which means that all corresponding elementary matrices are invertible. The last result shows that the invertible matrix A can be written as a product of elementary matrices $R_2^{2,1}$, $R_2^{1/2}$, $R_1^{3,2}$. This result holds in general.

Theorem 7.8 *Every invertible matrix is a product of elementary matrices.*

Exercises

Exercise 24. Decide if the following two linear maps are invertible. Determine their inverses if they exist.

$$F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \qquad \begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} \longmapsto \begin{pmatrix}2x_2 + 2x_3\\ x_1 - 4x_2 + 6x_3\\ x_2 + x_3\end{pmatrix},$$

$$G : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \qquad \begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} \longmapsto \begin{pmatrix}10x_1 + x_2 - 26x_3\\ x_1 - 2x_3\\ -x_1 + x_3\end{pmatrix}.$$

8

Subspaces, Kernel & Image

In the previous sections, we considered subsets of $\mathbb{R}^n$ which arose from the study of linear maps. For example, given a linear map $F : \mathbb{R}^n \to \mathbb{R}^m$, we could find its image $\mathrm{im}(F) \subset \mathbb{R}^m$. In the case $m = 3$, we saw that the image could be everything ($\mathbb{R}^3$), a plane, a line, or just contains only one point $\mathbf{0} \in \mathbb{R}^3$.

These sets are examples of subspaces, in which if you take any two vectors from a subspace, their sum and any scalar multiple of them also remain within the same subspace. The definition of subspaces is given as follows.

Definition 8.1 A subset $U \subset \mathbb{R}^n$ is a **subspace** of $\mathbb{R}^n$ if

(i) $\mathbf{0} \in U$,

(ii) for all $u, v \in U$ we have $u + v \in U$,

(iii) for all $u \in U$ and $\lambda \in \mathbb{R}$ we have $\lambda u \in U$.

Example 29 1) $U = \{\mathbf{0}\}$ and $U = \mathbb{R}^n$ are always subspaces of $\mathbb{R}^n$ for all $n \geq 1$.

2) General subspaces of $\mathbb{R}^n$ for $n = 1,\ 2,\ 3$ are given as follows.

$n = 1$: $\{0\}$, $\mathbb{R}$.

$n = 2$: $\{\mathbf{0}\}$, $\mathbb{R}^2$,

$\{\lambda v \in \mathbb{R}^2 \mid \lambda \in \mathbb{R}\}$ for any $v \in \mathbb{R}^2$, $v \neq \mathbf{0}$.

$n = 3$: $\{\mathbf{0}\}$, $\mathbb{R}^3$,

$\{\lambda v \in \mathbb{R}^3 \mid \lambda \in \mathbb{R}\}$ for any $v \in \mathbb{R}^3$, $v \neq \mathbf{0}$,

$\{\lambda_1 v + \lambda_2 u \in \mathbb{R}^3 \mid \lambda_1, \lambda_2 \in \mathbb{R}\}$ for any $u, v \in \mathbb{R}^3$, $u, v \neq \mathbf{0}$.

3) $U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 + x_2 - 3x_3 = 4 \right\} \subset \mathbb{R}^3$ is not a subspace because $\mathbf{0} \notin U$.

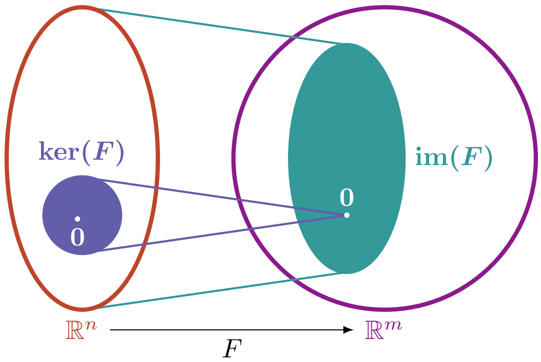
4) $U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; -1 \leq x_1 \leq 1 \right\} \subset \mathbb{R}^3$ is also not a subspace because we have $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in U$ but $2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \notin U$.

A lot of subspaces come from linear maps (actually all of them). We will see that the image of a linear map is a subspace. Another subspace coming from a linear map is its **kernel**.

Definition 8.2 For a linear map $F : \mathbb{R}^n \to \mathbb{R}^m$ the **kernel of** F is defined by

$$\ker(F) = \{x \in \mathbb{R}^n \mid F(x) = \mathbf{0}\}\,.$$

In other words, the kernel of a linear map F is the set of all solutions to the linear system $[F]x = \mathbf{0}$. The following figure illustrates the kernel and the image of a linear map.



Proposition 8.3 *For any linear map $F : \mathbb{R}^n \to \mathbb{R}^m$ we have the following:*
(i) The kernel $\ker(F)$ is a subspace of $\mathbb{R}^n$.
(ii) The image $\operatorname{im}(F)$ is a subspace of $\mathbb{R}^m$.

Proof. To show that a subset U is a subspace, we need to check the 3 conditions from Definition 8.1.

(i) $\ker(F)$ is a subspace of $\mathbb{R}^n$ because it satisfies the following conditions.

(a) $\mathbf{0} \in \ker(F)$ because $F(\mathbf{0}) = [F]\mathbf{0} = \mathbf{0}$.

(b) For any $u, v \in \ker(F)$, we have $F(u + v) = F(u) + F(v) = \mathbf{0} + \mathbf{0} = \mathbf{0}$.
Therefore, $u + v \in \ker(F)$.

(c) For any $u \in \ker(F)$ and $\lambda \in \mathbb{R}$, we have $F(\lambda u) = \lambda F(u) = \lambda \cdot \mathbf{0} = \mathbf{0}$. Thus, $\lambda u \in \ker(F)$.

(ii) $\operatorname{im}(F)$ is a subspace of $\mathbb{R}^m$ because it satisfies the following conditions.

(a) We have $\mathbf{0} \in \operatorname{im}(F)$ since $F(\mathbf{0}) = \mathbf{0}$.

(b) Let $u, v \in \operatorname{im}(F)$. In other words, $u = F(x)$ and $v = F(y)$ for some $x, y \in \mathbb{R}^n$.
Then we have $u + v = F(x) + F(y) = F(x + y)$. Because $x + y \in \mathbb{R}^n$, we have $u + v \in \operatorname{im}(F)$.

(c) Let $u \in \operatorname{im}(F)$, $u = F(x)$ for some $x \in \mathbb{R}^n$, and $\lambda \in \mathbb{R}$.
We have $\lambda u = \lambda F(x) = F(\lambda x)$, which implies that $\lambda u \in \operatorname{im}(F)$. □

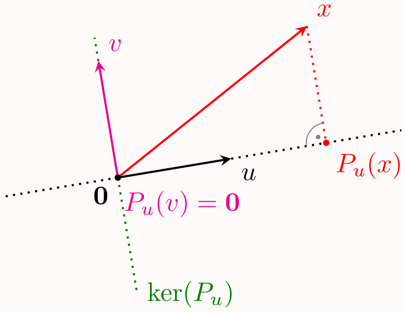
Remark. Actually, every subspace can be written as the kernel and the image of some linear maps. However, we cannot prove this yet.

Example 30 1) Let $u \in \mathbb{R}^n$ and $u \neq \mathbf{0}$. Consider the orthogonal projection $P_u : \mathbb{R}^n \to \mathbb{R}^n$.

The kernel of P_u is given by all vector $v \in \mathbb{R}^n$ such that $u \bullet v = 0$ because $u \neq \mathbf{0}$ and

$$P_u(v) = \frac{u \bullet v}{u \bullet u} u = \mathbf{0} \;\Leftrightarrow\; \frac{u \bullet v}{u \bullet u} = 0 \;\Leftrightarrow\; u \bullet v = 0.$$

Hence, $\ker(P_u) = \{v \in \mathbb{R}^n \mid u \bullet v = 0\}$.
For $n = 2$, $\ker(P_u)$ is a line.
For $n = 3$, $\ker(P_u)$ is a plane.



2) Consider the linear map

$$\begin{aligned} F : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ x &\longmapsto \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 0 & 1 \end{pmatrix} x \end{aligned}$$

<u>Kernel:</u> We have $x \in \ker(F) \Leftrightarrow F(x) = \mathbf{0}$. Therefore, we need to find solutions to the linear system $[F]x = \mathbf{0}$.

$$\big([F] \mid \mathbf{0}\big) = \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Hence, $x = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{0}$ and therefore, $\ker(F) = \{\mathbf{0}\}$.

<u>Image:</u> To calculate the image of F, we need to check for which $y \in \mathbb{R}^3$ there exist at least one $x \in \mathbb{R}^2$ with $F(x) = y$.

$$\big([F] \mid y\big) = \left(\begin{array}{cc|c} 1 & 1 & y_1 \\ 2 & 1 & y_2 \\ 0 & 1 & y_3 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 1 & y_1 \\ 0 & -1 & y_2 - y_1 \\ 0 & 1 & y_3 \end{array}\right)$$

$$\sim \left(\begin{array}{cc|c} 1 & 1 & y_1 \\ 0 & -1 & -2y_1 + y_2 \\ 0 & 0 & -2y_1 + y_2 + y_3 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & -y_1 + y_2 \\ 0 & 1 & 2y_1 - y_2 \\ 0 & 0 & -2y_1 + y_2 + y_3 \end{array}\right)$$

Thus, $\operatorname{im}(F) = \left\{ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; y_3 = 2y_1 - y_2 \right\}$. Notice that this calculation can also be used to calculate the kernel by setting $y_1 = y_2 = y_3 = 0$. Besides, for any $y \in \operatorname{im}(F)$, we can set $y_1 = \lambda_1$, $y_2 = \lambda_2$, and write

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ 2\lambda_1 - \lambda_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Therefore, $\operatorname{im}(F) = \left\{ \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; \lambda_1, \lambda_2 \in \mathbb{R} \right\}$.

3) Consider the linear map

$$\begin{aligned} G : \mathbb{R}^4 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} x \end{aligned}$$

Kernel: We solve the linear system $[G]x = \mathbf{0}$.

$$([G] \mid \mathbf{0}) = \left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right)$$

Hence, the solution is $\begin{cases} x_1 = -2t_1 - t_2 \\ x_2 = t_1 \\ x_3 = t_1 \\ x_4 = t_2 \end{cases}$, for $t_1, t_2 \in \mathbb{R}$.

Another way of writing this solution is

$$x = t_1 \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{for } t_1, t_2 \in \mathbb{R}.$$

Therefore, $\ker(G) = \left\{ t_1 \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^4 \;\middle|\; t_1, t_2 \in \mathbb{R} \right\}$.

Image: We check for which $y \in \mathbb{R}^2$ that the linear system $[G]x = y$ has solutions.

$$([G] \mid y) = \left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & y_1 \\ 1 & 1 & 1 & 1 & y_2 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & y_1 \\ 0 & -1 & 1 & 0 & y_2 - y_1 \end{array}\right)$$
$$\sim \left(\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 3y_1 - 2y_2 \\ 0 & 1 & -1 & 0 & y_1 - y_2 \end{array}\right).$$

Since $\operatorname{rk}\big([G] \mid y\big) = \operatorname{rk}\big([G]\big)$ for any $y \in \mathbb{R}^2$, the linear system $[G]x = y$ has solutions for any $y \in \mathbb{R}^2$. Hence, $\operatorname{im}(G) = \mathbb{R}^2$.

In Example 30, we see that the sets containing all sums of multiples of some vectors appear frequently when we try to determine the kernel and the image of a linear map.

Definition 8.4 (i) A **linear combination** of vectors $v_1, \ldots, v_n \in \mathbb{R}^m$ is a vector of the form

$$u = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \in \mathbb{R}^m$$

for some numbers $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

(ii) The **span** of $v_1, \ldots, v_n \in \mathbb{R}^m$ is the set of all linear combinations

$$\operatorname{span}\{v_1, \ldots, v_n\} = \{\lambda_1 v_1 + \cdots + \lambda_n v_n \in \mathbb{R}^m \mid \lambda_1, \ldots, \lambda_n \in \mathbb{R}\}.$$

Example 31 1) $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ is a linear combination of $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$ since $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = 2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$.

2) Every vector $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ is a linear combination of $e_1, \ldots, e_n$ because

$$x = x_1 e_1 + \ldots + x_n e_n.$$

Therefore, $\mathbb{R}^n = \operatorname{span}\{e_1, \ldots, e_n\}$.

3) From Example 30, we have

$$\operatorname{im}(F) = \operatorname{span}\left\{\begin{pmatrix}1\\0\\2\end{pmatrix}, \begin{pmatrix}0\\1\\-1\end{pmatrix}\right\},$$
$$\ker(G) = \operatorname{span}\left\{\begin{pmatrix}-2\\1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\0\\1\end{pmatrix}\right\}.$$

Remark. Given a linear map

$$F : \mathbb{R}^n \longrightarrow \mathbb{R}^m,$$
$$x \longmapsto Ax,$$

with $A = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$, we can always write $\operatorname{im}(F) = \operatorname{span}\{v_1, \dots, v_n\}$.

Proposition 8.5 *For $v_1, \dots, v_n \in \mathbb{R}^m$ we have the following.*
(i) $\operatorname{span}\{v_1, \dots, v_n\}$ *is a subspace of* $\mathbb{R}^m$.
(ii) If $U \subset \mathbb{R}^m$ is a subspace and $v_1, \dots, v_n \in U$ then $\operatorname{span}\{v_1, \dots, v_n\} \subset U$.

Proof. The proof is left as Exercise 25. □

Recall that a linear map $F : \mathbb{R}^n \to \mathbb{R}^m$ is surjective if and only if $\operatorname{im}(F) = \mathbb{R}^m$. A similar statement exists for injective functions as follows.

Proposition 8.6 *A linear map $F : \mathbb{R}^n \to \mathbb{R}^m$ is injective if and only if* $\ker(F) = \{\mathbf{0}\}$.

Proof. The proof is left as Exercise 20. □

We also know that F is injective if each column of $\operatorname{rref}([F])$ contains a pivot element, i.e., $\operatorname{rk}(F) = n$. Similarly, we know that F is surjective if each row of $\operatorname{rref}([F])$ contains a pivot element, i.e., $\operatorname{rk}(F) = m$. Summarizing everything, we get the following theorem.

Theorem 8.7 *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map.*
(i) We have the following equivalent statements for F being injective:

$$F \text{ is injective} \iff \ker(F) = \{\mathbf{0}\} \iff \operatorname{rk}([F]) = n\,.$$

(ii) We have the following equivalent statements for F being surjective:

$$F \text{ is surjective} \iff \operatorname{im}(F) = \mathbb{R}^m \iff \operatorname{rk}([F]) = m\,.$$

(iii) If $\underline{m = n}$ then the following statements are equivalent:

$$F \text{ is bijective} \iff F \text{ is injective} \iff F \text{ is surjective}\,.$$

Exercises

Exercise 25. Show the following without using Proposition 8.5:

(i) For $v_1, \ldots, v_n \in \mathbb{R}^m$ the set $\operatorname{span}\{v_1, \ldots, v_n\}$ is a subspace of $\mathbb{R}^m$.

(ii) If $U \subset \mathbb{R}^m$ is a subspace and $v_1, \ldots, v_n \in U$ then $\operatorname{span}\{v_1, \ldots, v_n\} \subset U$.

Exercise 26. Which of the following subsets are subspaces? Justify your answers.

$$\begin{aligned}
U_1 &= \left\{x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\right\} , \\
U_2 &= \left\{x \in \mathbb{R}^3 \mid x_1 \cdot x_2 \cdot x_3 = 0\right\} , \\
U_3 &= \{x \in \mathbb{R}^n \mid Ax = Bx\} , \quad \text{where } A, B \in \mathbb{R}^{m \times n}, \\
U_4 &= \left\{x \in \mathbb{R}^2 \mid x_1 \leq x_2\right\} .
\end{aligned}$$

Exercise 27.

(i) Which of the following subsets are subspaces? Justify your answers.

$$\begin{aligned}
U_1 &= \left\{x \in \mathbb{R}^3 \mid x_1 - x_2 = x_3\right\} , \\
U_2 &= \left\{x \in \mathbb{R}^2 \mid x_1^2 - x_2^2 = 0\right\} , \\
U_3 &= \{x \in \mathbb{R}^n \mid Ax = -2x\} , \quad \text{where } A \in \mathbb{R}^{n \times n} \text{ is a fixed matrix}, \\
U_4 &= \{x \in \mathbb{R}^n \mid x \bullet v = 0\} , \quad \text{for a fixed } v \in \mathbb{R}^n.
\end{aligned}$$

(ii) For each subset U in (i) which is a subspace, find numbers $a, b \geq 1$ and a linear map $F : \mathbb{R}^a \to \mathbb{R}^b$ such that $\ker(F) = U$.

Exercise 28.

(i) Which of the following subsets are subspaces? Justify your answers.

$$\begin{aligned}
U_1 &= \left\{x \in \mathbb{R}^3 \mid \|x\| \leq 1\right\} , \\
U_2 &= \left\{x \in \mathbb{R}^3 \mid 2x_1 - x_2 = x_3\right\} , \\
U_3 &= \left\{x \in \mathbb{R}^3 \mid x_1^2 - x_1 = 0\right\} , \\
U_4 &= \left\{x \in \mathbb{R}^4 \mid P_u(x) = \mathbf{0}\right\} , \quad \text{where } u = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix} .
\end{aligned}$$

(ii) For the subspaces U in (i): Find vectors $v_1, \ldots, v_l$ such that $U = \operatorname{span}\{v_1, \ldots, v_l\}$.

(Challenge for (ii): Try to choose the $v_1, \ldots, v_l$ such that they are pairwise orthogonal and all of them have norm 1. Such a basis is called orthonormal basis and we will study them in Chapter 12)

Exercise 29. Consider the following subspace

$$W = \ker(P_u) = \left\{x \in \mathbb{R}^3 \mid P_u(x) = \mathbf{0}\right\} , \quad \text{where } u = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} .$$

(i) Determine vectors $v_1, \ldots, v_m \in \mathbb{R}^3$ with $W = \operatorname{span}\{v_1, \ldots, v_m\}$.

(ii) Find a linear map $H : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ with $\operatorname{im}(H) = W$.

(iii) Calculate $\ker(H)$ and $\ker(P_u \circ H)$.

Exercise 30.
(i) Let $U, V \subset \mathbb{R}^m$ be subspaces. Decide whether the union $U \cup V$ is also a subspace or not.
(ii) Let $U, V \subset \mathbb{R}^m$ be subspaces. Decide whether the intersection $U \cap V$ is also a subspace or not.

Exercise 31.
(i) Decide if the following two linear maps are invertible. Determine their inverses if they exist.

$$F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3\,, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} 2x_1 - x_2 \\ -4x_1 + x_2 - x_3 \\ 6x_1 - 2x_2 + x_3 \end{pmatrix}, \qquad G : \mathbb{R}^3 \longrightarrow \mathbb{R}^3\,, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} -x_1 + x_3 \\ -4x_1 + x_2 - x_3 \\ 6x_1 - 2x_2 + x_3 \end{pmatrix}.$$

(ii) Determine $\ker(F)$ and $\ker(G)$.

Exercise 32. Find an example of a subset $U \subset \mathbb{R}^2$ which is <u>not</u> a subspace, but which
(i) includes $\mathbf{0}$ and which is closed under addition.
(ii) includes $\mathbf{0}$ and which is closed under scalar multiplication.
In other words: Find examples of subsets, which just satisfy 2 of the 3 conditions for subspaces.

9

Linear independence

In Example 30, we considered the linear map

$$\begin{aligned} G : \mathbb{R}^4 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} x \end{aligned}$$

and determined its image

$$\operatorname{im}(G) = \mathbb{R}^2 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

But we also learned that

$$\begin{aligned} \operatorname{im}(G) &= \text{ span of columns of } [G] \\ &= \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

This shows that

$$\operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

From the left-hand side, we just need 2 vectors to span $\operatorname{im}(G)$. Meanwhile, on the right-hand side, there are too many vectors, so we want to remove some extra vectors. In general, given a subspace as a span of some vectors, we may wonder about the minimum number of vectors and which vectors we need to keep to describe this subspace. In order to answer that question, we first need the following lemma.

Lemma 9.1 *Let $v_1, \dots, v_l \in \mathbb{R}^m$. If $v_l \in \operatorname{span}\{v_1, \dots, v_{l-1}\}$ then*

$$\operatorname{span}\{v_1, \dots, v_l\} = \operatorname{span}\{v_1, \dots, v_{l-1}\}.$$

Proof. Set $V = \operatorname{span}\{v_1, \dots, v_l\}$ and $W = \operatorname{span}\{v_1, \dots, v_{l-1}\}$. Clearly, we have $W \subset V$, so we want to show $V \subset W$. If $v \in \operatorname{span}\{v_1, \dots, v_l\} = V$, then there exist $\lambda_1, \dots, \lambda_l \in \mathbb{R}$ with

$$v = \lambda_1 v_1 + \ldots + \lambda_l v_l. \tag{$*$}$$

Since $v_l \in \operatorname{span}\{v_1, \dots, v_{l-1}\}$, there also exist $\alpha_1, \dots, \alpha_{l-1} \in \mathbb{R}$ with

$$v_l = \alpha_1 v_1 + \ldots + \alpha_{l-1} v_{l-1}. \tag{$**$}$$

Combining $(*)$ and $(**)$, we have

$$\begin{aligned} v &= \lambda_1 v_1 + \ldots + \lambda_{l-1} v_{l-1} + \lambda_l(\alpha_1 v_1 + \ldots + \alpha_{l-1} v_{l-1}) \\ &= (\lambda_1 + \lambda_l \alpha_1) v_1 + \ldots + (\lambda_{l-1} + \lambda_l \alpha_{l-1}) v_{l-1} \end{aligned}$$

And therefore $v \in \operatorname{span}\{v_1, \ldots, v_{l-1}\} = W$, i.e. $V \subset W$. As a result, $V = W$. $\square$

Example 32 For the linear map G in Example 30, we get

$$\operatorname{im}(G) = \operatorname{span}\left\{ \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} \right\}.$$

We can show directly that $\operatorname{im}(G) = \mathbb{R}^2$ by using Lemma 9.1 without solving the linear system $[G]x = y$. Since $\begin{pmatrix}2\\1\end{pmatrix} = 2\begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}0\\1\end{pmatrix} \in \operatorname{span}\left\{ \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\}$ and also $\begin{pmatrix}1\\1\end{pmatrix} \in \operatorname{span}\left\{ \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\}$, by Lemma 9.1, we get

$$\operatorname{im}(G) = \operatorname{span}\left\{ \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\}.$$

Since $\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}0\\1\end{pmatrix}$, we get (again by Lemma 9.1)

$$\operatorname{im}(G) = \operatorname{span}\left\{ \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\}.$$

Since $\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}$, we get (again by Lemma 9.1)

$$\operatorname{im}(G) = \operatorname{span}\left\{ \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\} = \mathbb{R}^2.$$

As a result,

$$\operatorname{span}\left\{ \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} \right\} = \operatorname{span}\left\{ \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\}.$$

General question: When is it possible to remove elements from $\operatorname{span}\{v_1, \ldots, v_l\}$ without changing it?

Definition 9.2 (i) Vectors $v_1, \ldots, v_l \in \mathbb{R}^m$ are called **linearly independent** if the equation

$$\lambda_1 v_1 + \cdots + \lambda_l v_l = 0 \tag{9.0.1}$$

with $\lambda_1, \ldots, \lambda_l \in \mathbb{R}$ just has the unique solution $\lambda_1 = \cdots = \lambda_l = 0$.

(ii) If there exist another solution of (9.0.1), i.e. where at least for one $j = 1, \ldots, l$ we have $\lambda_j \neq 0$, then the vectors $v_1, \ldots, v_l$ are called **linearly dependent**.

Example 33 Are the vectors $v_1 = \begin{pmatrix}1\\1\\2\end{pmatrix}$, $v_2 = \begin{pmatrix}1\\-1\\1\end{pmatrix}$, and $v_3 = \begin{pmatrix}-1\\5\\1\end{pmatrix}$ linearly independent?

The equation $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = \mathbf{0}$ is equivalent to

$$\begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix} \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 5 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

To solve this linear system, we need to find the row-reduced echelon form of the augmented matrix:

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 1 & -1 & 5 & 0 \\ 2 & 1 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -2 & 6 & 0 \\ 0 & -1 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 \end{array}\right)$$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Hence, the solution is $\begin{cases} \lambda_1 = -2t \\ \lambda_2 = 3t \\ \lambda_3 = t \end{cases}$, for $t \in \mathbb{R}$. Therefore, v_1, v_2, v_3 are linearly dependent since there are non-zero solutions.
For $t = 1$, we get $\lambda_1 = -2$, $\lambda_2 = 3$, $\lambda_3 = 1$. Thus, we get

$$-2v_1 + 3v_2 + v_3 = -2\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + 3\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 5 \\ 1 \end{pmatrix} = \mathbf{0}$$

$$\Rightarrow v_3 = 2v_1 - 3v_2.$$

Therefore, $v_3 \in \operatorname{span}\{v_1, v_2\}$ and by Lemma 9.1 $\operatorname{span}\{v_1, v_2, v_3\} = \operatorname{span}\{v_1, v_2\}$. On the other hand, v_1 and v_2 are linearly independent since

$$\left(\begin{array}{cc|c} | & | & 0 \\ v_1 & v_2 & 0 \\ | & | & 0 \end{array}\right) = \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 2 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right),$$

which shows that $\mathbf{0}$ is the only solution and hence, $\lambda_1 v_1 + \lambda_2 v_2 = \mathbf{0} \iff \lambda_1 = \lambda_2 = 0$. As a result, we cannot remove v_1 or v_2 without changing the span.

We can summarize in the following theorem what we have just done in Example 33.

Theorem 9.3 *Let $v_1, \ldots, v_l \in \mathbb{R}^m$. The following statements are equivalent:*

(i) $v_1, \ldots, v_l$ are linearly dependent.

(ii) There exists a $j = 1, \ldots, l$, such that v_j is a linear combination of the other vectors.

(iii) There exists a $j = 1, \ldots, l$ with

$$\operatorname{span}\{v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_l\} = \operatorname{span}\{v_1, \ldots, v_l\}.$$

Proof. The statement (ii) $\Rightarrow$ (iii) is Lemma 9.1. Now we prove the statement (iii) $\Rightarrow$ (ii) as follows:

$$v_j \in \operatorname{span}\{v_1, \ldots, v_l\} \overset{\text{(iii)}}{=} \operatorname{span}\{v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_l\} \quad \Rightarrow \quad v_j \in \operatorname{span}\{v_1, \ldots, \cancel{v_j}, \ldots, v_l\}.$$

Next, for proving the statement (ii) $\Rightarrow$ (i), suppose $v_j = \lambda_1 v_1 + \ldots + \lambda_{j-1} v_{j-1} + \lambda_{j+1} v_{j+1} + \ldots + \lambda_l v_l$. Then,

$$0 = \lambda_1 v_1 + \ldots + \lambda_{j-1} v_{j-1} - 1 \cdot v_j + \lambda_{j+1} v_{j+1} + \ldots + \lambda_l v_l.$$

Hence, $v_1, \ldots, v_l$ are linearly dependent.

Finally, for proving the statement (i) $\Rightarrow$ (ii), suppose $\lambda_1 v_1 + \ldots + \lambda_l v_l = 0$ with $\lambda_j \neq 0$. Then,

$$v_j = \left(\frac{\lambda_1}{\lambda_j}\right) v_1 + \ldots + \left(\frac{\lambda_{j-1}}{\lambda_j}\right) v_{j-1} + \left(\frac{\lambda_{j+1}}{\lambda_j}\right) v_{j+1} + \ldots + \left(\frac{\lambda_l}{\lambda_j}\right) v_l.$$

Thus, $v_j \in \operatorname{span}\{v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_l\}$. □

The following lemma shows that in a subspace, we cannot have more linearly independent vectors than the vectors spanning the subspace.

Lemma 9.4 *Let $V \subset \mathbb{R}^n$ be a subspace, $v_1, \ldots, v_l \in V$ linearly independent and $V = \operatorname{span}\{w_1, \ldots, w_m\}$ for some $w_1, \ldots, w_m \in \mathbb{R}^n$. Then we have $l \leq m$.*

Proof. The proof is left as Exercise 33. □

The following lemma shows how to make a set of linearly independent vectors bigger.

Lemma 9.5 *If $v_1, \ldots, v_l \in \mathbb{R}^n$ are linearly independent and $w \in \mathbb{R}^n$ with $w \notin \operatorname{span}\{v_1, \ldots, v_l\}$ then $v_1, \ldots, v_l, w$ are linearly independent.*

Proof. Assume that $\lambda_1 v_1 + \ldots + \lambda_l v_l + \mu w = 0$. If $\mu \neq 0$, then

$$w = \left(-\frac{\lambda_1}{\mu}\right) v_1 + \ldots + \left(-\frac{\lambda_l}{\mu}\right) v_l \in \operatorname{span}\{v_1, \ldots, v_l\},$$

which contradicts to the assumption. Hence, $\mu = 0$, which gives $\lambda_1 v_1 + \ldots + \lambda_l v_l = 0$.

Then, $\lambda_1 = \ldots = \lambda_l = 0$ because $v_1, \ldots, v_l$ are linearly independent. Thus, $v_1, \ldots, v_l$, and w are linearly independent. □

Exercises

Exercise 33. Let $V \subset \mathbb{R}^n$ be a subspace, $v_1, \ldots, v_l \in V$ linearly independent and $V = \operatorname{span}\{w_1, \ldots, w_m\}$ for some $w_1, \ldots, w_m \in \mathbb{R}^n$. Show that we have $l \leq m$. (Without using Lemma 9.4)
In other words: Show that a subspace spanned by m vectors can not contain more than m linearly independent vectors.

10

Bases & dimensions

We saw that if $v_1, \dots, v_l \in \mathbb{R}^n$ are linearly dependent, then there exist a $1 \le j \le l$ such that

$$\operatorname{span}\{v_1, \dots, v_l\} = \operatorname{span}\{v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_l\}.$$

Therefore, we will be just interested in the case when $v_1, \dots, v_l$ are linearly independent.

Definition 10.1 Let $V \subset \mathbb{R}^n$ be a subspace. Vectors $v_1, \dots, v_l \in V$ form a **basis of V** if
(i) $V = \operatorname{span}\{v_1, \dots, v_l\}$,
(ii) $v_1, \dots, v_l$ are linearly independent.
In this case we also say that $\{v_1, \dots, v_l\}$ is a basis of V.

Later we will also be interested in the order of the v_j and write a basis as a tuple $(v_1, \dots, v_l)$.

Example 34 1) $\{e_1, e_2\}$, $\left\{e_1, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$, $\left\{e_2, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ are three different bases of $\mathbb{R}^2$.

2) Consider $U = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \\ 1 \end{pmatrix}\right\} = \operatorname{span}\{v_1, v_2, v_3\}$. We want to find a basis for the subspace U. We have seen in Example 33:
(a) v_1, v_2, v_3 are linearly dependent because $-2v_1 + 3v_2 + v_3 = \mathbf{0}$.
(b) $v_3 \in \operatorname{span}\{v_1, v_2\}$ and hence, $U = \operatorname{span}\{v_1, v_2\}$ by Lemma 9.1.
(c) v_1 and v_2 are linearly independent.
Therefore, $\{v_1, v_2\}$ is a basis of U.

3) For any $n \ge 1$, $\{e_1, \dots, e_n\}$ is a basis of $\mathbb{R}^n$. This basis is called the **standard basis**.

We will see that a basis of a subspace is a convenient tool for us to study the subspace because we can represent every vector of the subspace in terms of vectors in the basis. By working with bases, we can focus on understanding a smaller set of vectors rather than dealing with the entire space, enabling us to analyze, manipulate, and comprehend subspaces with greater ease. Fortunately, the following theorem shows that bases always exist.

Theorem 10.2 *For any subspace $V \subset \mathbb{R}^n$ we have the following:*
(i) V has a basis.
(ii) All bases of V have the same number of elements.
(iii) If $v_1, \dots, v_l \in V$ are linearly independent then there exist $u_{l+1}, \dots, u_t \in V$, such that $\{v_1, \dots, v_l, u_{l+1}, \dots, u_t\}$ is a basis of V.
(iv) If $V = \operatorname{span}\{w_1, \dots, w_m\}$ then there exists a subset $\{u_1, \dots, u_t\} \subset \{w_1, \dots, w_m\}$ such that $\{u_1, \dots, u_t\}$ is a basis of V.

Proof. (ii) Let $\{v_1, \ldots, v_l\}$ and $\{w_1, \ldots, w_m\}$ be bases of V. Then, we have by Lemma 9.4,

$$v_1, \ldots, v_l \in V \text{ are linearly independent and } V = \text{span}\{w_1, \ldots, w_m\} \quad \Rightarrow \quad l \leq m,$$

$$w_1, \ldots, w_m \in V \text{ are linearly independent and } V = \text{span}\{v_1, \ldots, v_l\} \quad \Rightarrow \quad m \leq l.$$

Therefore, we have $l = m$.

(iv) 1. If $w_1, \ldots, w_m$ are linearly independent, then $\{u_1, \ldots, u_t\} = \{w_1, \ldots, w_m\}$ is a basis of V.

2. Otherwise, if $\{w_1, \ldots, w_m\}$ are linearly dependent, then By Theorem 9.3 there exist a $j = 1, \ldots, m$ with $\text{span}\{w_1, \ldots, w_{j-1}, w_{j+1}, \ldots, w_m\} = \text{span}\{w_1, \ldots, w_m\} = V$.

Now repeat 1. and 2. with $\text{span}\{w_1, \ldots, w_{j-1}, w_{j+1}, \ldots, w_m\}$, i.e. remove vectors like w_j until when the remaining vectors are linearly independent. Eventually, we will get a basis $\{u_1, \ldots, u_t\}$ of V.

(iii) Assume that $v_1, \ldots, v_l \in V$ are linearly independent.

1. If $\text{span}\{v_1, \ldots, v_l\} = V$, then $\{v_1, \ldots, v_l\}$ is a basis of V.

2. Otherwise, if $\text{span}\{v_1, \ldots, v_l\} \neq V$, then there exists $u \in V$ with $u \notin \text{span}\{v_1, \ldots, v_l\}$. In that case, we set $u_{l+1} = u$. By Lemma 9.5, $v_1, \ldots, v_l, u_{l+1}$ are linearly independent.

Repeat 1. and 2. for $\{v_1, \ldots, v_l, u_{l+1}\}$ until when $V = \text{span}\{v_1, \ldots, v_l, u_{l+1}, \ldots, u_t\}$.

(i) If $V = \{\mathbf{0}\}$ then $\emptyset = \{\}$ is a basis because $\text{span}(\emptyset) = \{\mathbf{0}\}$ by convention. Otherwise, we can construct a basis using (iii). □

Definition 10.3 Let $V \subset \mathbb{R}^n$ be a subspace. The **dimension of V**, denoted by $\dim(V)$, is the number of elements in a basis of V.

Example 35 1) $\dim(\mathbb{R}^n) = n$ because $\{e_1, \ldots, e_n\}$ is a basis of $\mathbb{R}^n$.

2) The dimension of $U = \text{span}\left\{\begin{pmatrix}1\\1\\2\end{pmatrix}, \begin{pmatrix}1\\-1\\1\end{pmatrix}, \begin{pmatrix}-1\\5\\1\end{pmatrix}\right\} = \text{span}\{v_1, v_2, v_3\}$ is $\dim(U) = 2$ because $\{v_1, v_2\}$ is a basis.

Corollary 10.4 *Let $V \subset \mathbb{R}^n$ be a subspace with $\dim(V) = m$ and $v_1, \ldots, v_m \in V$. Then the following statements are equivalent:*
(i) $v_1, \ldots, v_m$ *are linearly independent.*
(ii) $V = \text{span}\{v_1, \ldots, v_m\}$.
(iii) $\{v_1, \ldots, v_m\}$ *is a basis of* V.

Proof. **(i) ⇒ (ii):** If $V \neq \text{span}\{v_1, \ldots, v_m\}$, then by Theorem 10.2 (iii), there exists a basis with more than m elements, which contradicts to the assumption that $\dim(V) = m$. Hence, $V = \text{span}\{v_1, \ldots, v_m\}$.

(ii) ⇒ (i): If $v_1, \ldots, v_m$ are linearly dependent, then by Theorem 9.3 and Theorem 10.2 (iv), there exists a basis with less than m elements, which contradicts to the assumption that $\dim(V) = m$. Thus, $v_1, \ldots, v_m$ are linearly independent.

(i) + **(ii)** $\iff$ **(iii)** by definition 10.1. $\square$

Example 36 Determine bases for $\ker(F)$ and $\mathrm{im}(F)$ of the following linear map:

$$F : \mathbb{R}^4 \longrightarrow \mathbb{R}^3,$$
$$x \longmapsto \begin{pmatrix} 1 & -2 & 1 & 2 \\ 2 & -4 & 1 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix} x.$$

Kernel: We need to find $\mathrm{rref}([F] \mid \mathbf{0})$ to solve the linear system $[F]x = \mathbf{0}$. Because the column corresponding to $\mathbf{0}$ does not change after row operations, we just need to find $\mathrm{rref}([F])$.

$$\begin{pmatrix} 1 & -2 & 1 & 2 \\ 2 & -4 & 1 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 & 2 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \mathrm{rref}([F]).$$

Hence,

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \ker(F) \iff [F]x = \mathbf{0} \iff \begin{cases} x_1 = 2t_1 - t_2 \\ x_2 = t_1 \\ x_3 = -t_2 \\ x_4 = t_2 \end{cases}, \quad \text{for } t_1, t_2 \in \mathbb{R}$$

$$\iff x = t_1 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix} = t_1 v_1 + t_2 v_2, \quad \text{for } t_1, t_2 \in \mathbb{R},$$

where $v_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ and $v_2 = \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$. Therefore, $\ker(F) = \mathrm{span}\{v_1, v_2\}$.

Next, we need to check whether v_1 and v_2 are linearly independent or not.

$$\mathbf{0} = t_1 v_1 + t_2 v_2 = \begin{pmatrix} 2t_1 - t_2 \\ t_1 \\ -t_2 \\ t_2 \end{pmatrix} \implies t_1 = t_2 = 0.$$

Thus, v_1 and v_2 are linearly independent and $\{v_1, v_2\}$ is a basis of $\ker(F)$.

Image: We have $[F] = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \right\} = \mathrm{span}\{u_1, u_2, u_3, u_4\}$.

$$\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 + \lambda_4 u_4 = \mathbf{0} \iff [F] \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} = \mathbf{0} \iff \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} \in \ker(F).$$

Based on the result we got above,

$$\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 + \lambda_4 u_4 = \mathbf{0} \implies \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} = t_1 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \quad \text{for } t_1, t_2 \in \mathbb{R}.$$

When $t_1 = 1$ and $t_2 = 0$, we have

$$2u_1 + u_2 = \mathbf{0} \iff u_2 = -2u_1 \implies u_2 \in \operatorname{span}\{u_1, u_3\}.$$

When $t_1 = 0$ and $t_2 = 1$,

$$-u_1 - u_3 + u_4 = \mathbf{0} \iff u_4 = -u_1 + u_3 \implies u_3 \in \operatorname{span}\{u_1, u_3\}.$$

Hence, $\operatorname{im}(F) = \operatorname{span}\{u_1, u_3\}$. In addition, u_1 and u_3 are linear independent because when $\lambda_1 u_1 + \lambda_3 u_3 = \mathbf{0}$:

$$\left(\begin{array}{cc|c} | & | & 0 \\ u_1 & u_3 & 0 \\ | & | & 0 \end{array}\right) = \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right) \implies \lambda_1 = \lambda_3 = 0.$$

Thus, $\{u_1, u_3\}$ is a basis of $\operatorname{im}(F)$.

From Example 36, we can summarize the general calculation of bases for $\ker(F)$ and $\operatorname{im}(F)$ as follows:

Consider a linear map

$$\begin{aligned} F : \mathbb{R}^n &\longrightarrow \mathbb{R}^m, \\ x &\longmapsto Ax. \end{aligned}$$

- Let $\operatorname{rref}(A)$ have pivot elements in columns $c_1, \ldots, c_r$. Then, the columns $c_1, \ldots, c_r$ in the original matrix A form a basis of $\operatorname{im}(F)$. Hence, the dimension of $\operatorname{im}(F)$ is equal to the number of pivot elements in $\operatorname{rref}(A)$ or equal to $\operatorname{rk}(F)$.

$$\dim(\operatorname{im}(F)) = \text{ the number of pivot elements in } \operatorname{rref}(A).$$

- The vectors obtained in the "standard parametrization" (i.e. for each free variable x_i there is a parameter t_j) of the solutions to $F(x) = \mathbf{0}$ form a basis of $\ker(F)$. Thus, the dimension of $\ker(F)$ is equal to the number of free variables.

$$\dim(\ker(F)) = \text{ the number of free variables.}$$

As a consequence, we get the following theorem:

Theorem 10.5 *For a linear map $F : \mathbb{R}^n \to \mathbb{R}^m$ we have*

$$n = \dim(\ker(F)) + \dim(\operatorname{im}(F)).$$

Proof. The statement follows from the following facts:

$$\begin{aligned} n &= \text{ the number of columns of } [F], \\ \dim(\ker(F)) &= \text{ the number of columns without pivot elements.} \\ \dim(\operatorname{im}(F)) &= \text{ the number of columns with pivot elements.} \end{aligned}$$

□

Exercises

Exercise 34. Determine a basis of the following subspace

$$U = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 8 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}.$$

Exercise 35. Determine bases for the kernel and the image of the following linear map

$$\begin{aligned} F : \mathbb{R}^5 &\longrightarrow \mathbb{R}^3 \\ x &\longmapsto \begin{pmatrix} 2 & -6 & -1 & 5 & 3 \\ -1 & 3 & 2 & -4 & -3 \\ 1 & -3 & -1 & 3 & 4 \end{pmatrix} x \,. \end{aligned}$$

Exercise 36. Let $U, V \subset \mathbb{R}^n$ be two subspaces. We define their sum by

$$U + V := \{x \in \mathbb{R}^n \mid \text{there exist } u \in U, v \in V \text{ with } x = u + v\}\,.$$

(i) Show that $U + V$ is a subspace of $\mathbb{R}^n$.
(ii) Show that we have

$$\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V)\,.$$

Exercise 37. For $t \in \mathbb{R}$ we define

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 2 \\ t \end{pmatrix}, \quad v_3 = \begin{pmatrix} t \\ 4 \\ (t-2)^2 \end{pmatrix}$$

and set $V = \operatorname{span}\{v_1, v_2, v_3\}$. For each $t \in \mathbb{R}$ determine a basis of V and calculate its dimension.

Exercise 38. The following exercise is intended to show the basic idea of 3D computer graphics, by showing how to get a 2-dimensional picture (to be shown on a 2-dimensional monitor) from an 3-dimensional object.

(i) We define the corners of a cube with side length 18 in $\mathbb{R}^3$ by the following set of 8 points:

$$W = \left\{ \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \in \mathbb{R}^3 \mid w_1, w_2, w_3 \in \{0, 18\} \right\}.$$

Make a drawing of a cube with side length 18 in $\mathbb{R}^3$, i.e. draw the 8 points in the set W and connect two points if they differ just by one entry.

(This just means that you draw a cube like you would usually draw it. "Differ by one entry" just means that these points are on the same edge of the cube.)

(ii) Show that $D = (d_1, d_2, d_3)$ is a basis of $\mathbb{R}^3$, where

$$d_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad d_2 = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}, \quad d_3 = \begin{pmatrix} 3 \\ -6 \\ 3 \end{pmatrix}.$$

(iii) Write each $x \in W$ as a linear combination in the basis D, i.e. for each x find $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$

with

$$x = \lambda_1 d_1 + \lambda_2 d_2 + \lambda_3 d_3 \,.$$

(iv) For each $x \in W$ draw the points (λ_1, λ_2) in $\mathbb{R}^2$. Connect two points if the corresponding elements in W just differ by one entry.

Explanation: What you should get in iv) is a drawing of the 3-dimensional cube in 2 dimensions. The basis D somehow describes from which direction you look at the cube. If you replaced the D by the standard basis (e_1, e_2, e_3), you would get a picture of the cube from the top (i.e., just a square). The λ_3, which you did not use for the drawing, describes the distance in the viewing direction.

11

Coordinates

From now on we will consider ordered bases, which means that we will write $(b_1, \dots, b_n)$ (a tuple) for a basis instead of $\{b_1, \dots, b_n\}$ (a set). The difference is, that we care about the order now. For example, the two sets $\{b_1, b_2\} = \{b_2, b_1\}$ are the same, but $(b_1, b_2) \neq (b_2, b_1)$.

Definition 11.1 Let $B = (b_1, \dots, b_m)$ be a basis of a subspace $V \subset \mathbb{R}^n$. We define the **coordinate map** by

$$\begin{aligned} c_B : \mathbb{R}^m &\longrightarrow \mathbb{R}^n \\ \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} &\longmapsto \lambda_1 b_1 + \cdots + \lambda_m b_m \,. \end{aligned}$$

Clearly, the map c_B is a linear map. In addition, it is also a bijection, as stated in the following theorem.

Theorem 11.2 *Let $B = (b_1, \dots, b_m)$ be a basis of a subspace $V \subset \mathbb{R}^n$.*
(i) The coordinate map $c_B : \mathbb{R}^m \longrightarrow V$ is bijective.
(ii) For all $x \in V$ there exist unique $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ such that

$$x = \lambda_1 b_1 + \cdots + \lambda_m b_m \,.$$

Proof. (i) The map $c_B : \mathbb{R}^m \longrightarrow V$ is surjective since $\mathrm{im}(c_B) = \mathrm{span}\{b_1, \dots, b_m\} = V$.

Additionally,

$$\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} \in \ker(c_B) \quad \Longleftrightarrow \quad c_B(\lambda) = \lambda_1 b_1 + \ldots + \lambda_m b_m = \mathbf{0} \quad \Longrightarrow \quad \lambda = \mathbf{0},$$

where the implication comes from the assumption that $b_1, \dots, b_m$ are linearly independent. Hence, $\ker(c_B) = \{\mathbf{0}\}$ and, by Theorem 8.7, c_B is injective.

As a result, the map c_B is bijective.

(ii) This is just a reformulation of (i). □

As a consequence of Theorem 11.2, we can define **coordinates** and **coordinate vector** of any vectors in a subspace V as follows.

Definition 11.3 Let $B = (b_1, \dots, b_m)$ be a basis of a subspace $V \subset \mathbb{R}^n$ and $x \in V$ with

$$x = \lambda_1 b_1 + \cdots + \lambda_m b_m .$$

(i) The numbers $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ are the **coordinates of x (in the basis B)**.

(ii) The **coordinate vector** of x (with respect to B) is given by

$$[x]_B = c_B^{-1}(x) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} .$$

Example 37 1) $B = (e_1, \dots, e_n)$ is a basis of $\mathbb{R}^n$. For all $x \in \mathbb{R}^n$, we have $[x]_B = x$.

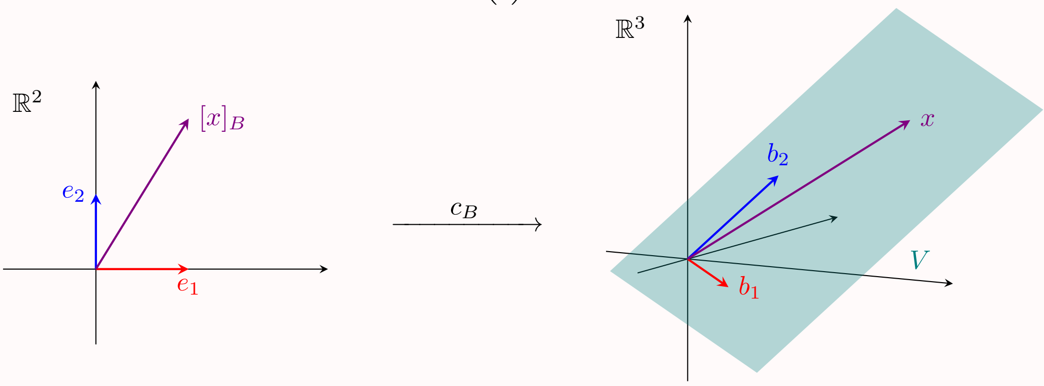
2) Consider $b_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ and $b_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$. Clearly, b_1 and b_2 are linearly independent; therefore,

$B = (b_1, b_2)$ is a basis of $V = \operatorname{span}\{b_1, b_2\}$. Is it true that $x = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} \in V$? What is $[x]_B$?

In order to answer those questions, we need to solve $\lambda_1 b_1 + \lambda_2 b_2 = x$.

$$\left(\begin{array}{cc|c} | & | & | \\ b_1 & b_2 & x \\ | & | & | \end{array}\right) = 1 \left(\begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 2 & 4 \\ -1 & 3 & 5 \end{array}\right) \sim \tfrac{1}{2} \; -2 \left(\begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 2 & 4 \\ 0 & 4 & 8 \end{array}\right) \sim -1 \left(\begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right)$$

Hence, $x = 1 \cdot b_1 + 2 \cdot b_2 \quad \Longrightarrow \quad [x]_B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.



Coordinate vectors have the following properties.

Proposition 11.4 *Let $B = (b_1, \dots, b_m)$ be a basis of a subspace $V \subset \mathbb{R}^n$. Then we have for all $x, y \in V$ and $\mu \in \mathbb{R}$*

(i) $[x + y]_B = [x]_B + [y]_B$,

(ii) $[\mu x]_B = \mu [x]_B$,

(iii) $[0]_B = 0$.

Proof. Firstly, we have $[x]_B = c_B^{-1}(x)$. In addition, c_B^{-1} is linear since c_B is linear (Proposition 7.2). Hence, all the properties follow from the linearity of c_B^{-1}. $\square$

Since c_B is a linear map, we can consider its matrix as follows.

Definition 11.5 Let $B = (b_1, \dots, b_n)$ be a basis of $\mathbb{R}^n$. The **change-of-basis matrix** associated with B is

$$S_B = [c_B] = \begin{pmatrix} | & & | \\ b_1 & \dots & b_n \\ | & & | \end{pmatrix}.$$

Remark. (a) S_B is invertible for any basis B of $\mathbb{R}^n$ since $S_B \in \mathbb{R}^{n\times n}$ and the linear system $S_B\lambda = \mathbf{0}$ or $\lambda_1 b_1 + \ldots + \lambda_n b_n = \mathbf{0}$ has a unique solution $x = \mathbf{0}$.

(b) For any $x \in \mathbb{R}^n$,

$$S_B[x]_B = c_B\big([x]_B\big) = c_B\big(c_B^{-1}(x)\big) = x.$$

The following definition may be a little complicated and confusing at first reading, but it will be useful later for the study of linear maps.

Definition 11.6 Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, B_1 be a basis of $\mathbb{R}^n$ and B_2 be a basis of $\mathbb{R}^m$. The **matrix of F with respect to B_1 and B_2** is the matrix

$$[F]_{B_1}^{B_2} := [c_{B_2}^{-1} \circ F \circ c_{B_1}].$$

In the case $n = m$ and $B_1 = B_2$ we just write $[F]_{B_1} := [F]_{B_1}^{B_1}$.

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "F"] & \mathbb{R}^m \arrow[d, "c_{B_2}^{-1}"] \\ \mathbb{R}^n \arrow[u, "c_{B_1}"] \arrow[r, "c_{B_2}^{-1}\circ F\circ c_{B_1}"'] & \mathbb{R}^m \end{tikzcd}$$

With this definition, we get the following proposition, which gives more insight into the definition of a matrix with respect to some bases.

Proposition 11.7 *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, B_1 be a basis of $\mathbb{R}^n$ and B_2 be a basis of $\mathbb{R}^m$.*

(i) We have

$$[F]_{B_1}^{B_2} = S_{B_2}^{-1}[F]S_{B_1}.$$

(ii) If $B_1 = (b_1, \dots, b_n)$ then

$$[F]_{B_1}^{B_2} = \begin{pmatrix} | & & | \\ [F(b_1)]_{B_2} & \dots & [F(b_n)]_{B_2} \\ | & & | \end{pmatrix}.$$

Proof. (i) $[F]_{B_1}^{B_2} = [c_{B_2}^{-1} \circ F \circ c_{B_1}] = [c_{B_2}^{-1}][F][c_{B_1}] = S_{B_2}^{-1}[F]S_{B_1}$.

(ii) The i^{th} column of $[F]_{B_1}^{B_2}$ is

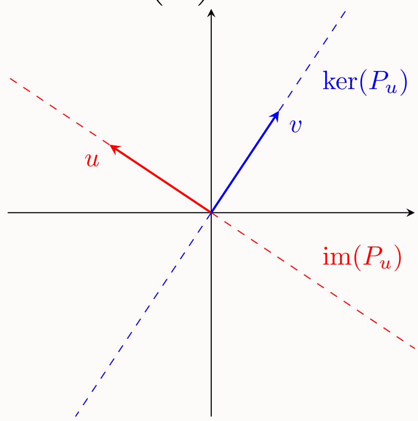
$$[F]_{B_1}^{B_2} e_i = c_{B_2}^{-1} \circ F \circ c_{B_1}(e_i) = c_{B_2}^{-1}\Big(F\big(c_{B_1}(e_i)\big)\Big) = c_{B_2}^{-1}\big(F(b_i)\big) = [F(b_i)]_{B_2}. \qquad \square$$

In many cases, we can write down $[F]_B$ with respect to some bases B much easier than $[F]$. After that, we can use Proposition 11.7 to obtain $[F]$.

Example 38

1) Consider $B_1 = (e_1, \dots, e_n)$ in $\mathbb{R}^n$ and $B_2 = (e_1, \dots, e_m)$ in $\mathbb{R}^m$.
   Then, $[c_{B_1}] = I_n$ and $[c_{B_2}] = I_m$. For any linear map $F : \mathbb{R}^n \to \mathbb{R}^m$, it is always true that
$$[F]_{B_1}^{B_2} = [F].$$

2) Consider the orthogonal projection $P_u : \mathbb{R}^2 \to \mathbb{R}^2$ with $u = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$.



We have
$$\begin{aligned} \ker(P_u) &= \{x \in \mathbb{R}^2 \mid u \bullet x = 0\} \\ &= \operatorname{span}\left\{\begin{pmatrix} 2 \\ 3 \end{pmatrix}\right\} = \operatorname{span}\{v\}, \end{aligned}$$
$$\operatorname{im}(P_u) = \operatorname{span}\{u\}.$$

We also have
$$\lambda_1 u + \lambda_2 v = \mathbf{0} \quad \Rightarrow \quad \begin{cases} u \bullet (\lambda_1 u + \lambda_2 v) = 0 \\ v \bullet (\lambda_1 u + \lambda_2 v) = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} \lambda_1 (u \bullet u) = 0 \\ \lambda_2 (v \bullet v) = 0 \end{cases} \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 0.$$

Hence, u and v are linearly independent, so $B = (u, v)$ is a basis of $\mathbb{R}^2$. Notice: this technique will be used again in the next chapter.

We have $P_u(u) = u$ and $P_u(v) = \mathbf{0}$, so
$$[P_u]_B = [P_u]_B^B = \begin{pmatrix} | & | \\ [P_u(u)]_B & [P_u(v)]_B \\ | & | \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

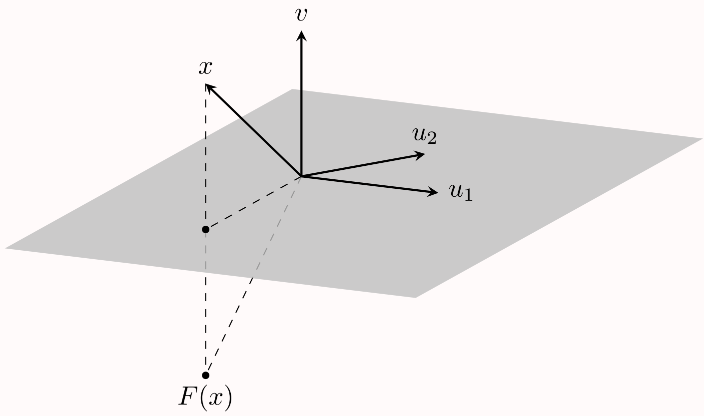
Change-of-basis matrix is
$$S_B = \begin{pmatrix} | & | \\ u & v \\ | & | \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 2 & 3 \end{pmatrix}$$
and its inverse is
$$S_B^{-1} = \frac{1}{13}\begin{pmatrix} -3 & 2 \\ 2 & 3 \end{pmatrix}.$$

By Proposition 11.7 we have $[P_u]_B = [P_u]_B^B = S_B^{-1}[P_u]S_B$. Hence,
$$[P_u] = S_B [P_u]_B S_B^{-1} = \begin{pmatrix} -3 & 2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\frac{1}{13}\begin{pmatrix} -3 & 2 \\ 2 & 3 \end{pmatrix} = \frac{1}{13}\begin{pmatrix} 9 & -6 \\ -6 & 4 \end{pmatrix}.$$

3) Let $F : \mathbb{R}^3 \to \mathbb{R}^3$ be the reflection through the plane
$$U = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right\} = \operatorname{span}\{u_1, u_2\}.$$

We want to determine the matrix $[F]$. Ideas for how to solve this problem are as follows.

(a) Find a "good" basis B where we can write down $[F]_B$ directly.
(b) For this, try to find $v \in \mathbb{R}^3$ which is orthogonal to u_1, u_2, and set $B = (u_1, u_2, v)$. Notice that u_1 and u_2 are linearly independent (check it!).



To find $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in \mathbb{R}^3$ such that $v \bullet u_1 = v \bullet u_2 = 0$, we need to solve $\begin{cases} v_1 + v_2 = 0 \\ v_2 + v_3 = 0 \end{cases}$.

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right)$$

Hence, all solutions are given by $t \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ for $t \in \mathbb{R}$. We choose $t = 1$, i.e. set $v = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

We also have

$$\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 v = \mathbf{0} \quad \Rightarrow \quad \begin{cases} \lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 v = \mathbf{0} \\ v \bullet (\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 v) = 0 \end{cases} \quad \Rightarrow \begin{cases} \lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 v = \mathbf{0} \\ \lambda_3 (v \bullet v) = 0 \end{cases}$$

$$\Rightarrow \quad \begin{cases} \lambda_1 u_1 + \lambda_2 u_2 = \mathbf{0} \\ \lambda_3 = 0 \end{cases} \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \lambda_3 = 0.$$

Thus, u_1, u_2, v are linearly independent, so $B = (u_1, u_2, v)$ is a basis of $\mathbb{R}^3$.
We have

$$F(u_1) = u_1 \quad \Rightarrow \quad [F(u_1)]_B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$F(u_2) = u_2 \quad \Rightarrow \quad [F(u_2)]_B = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$F(v) = -v \quad \Rightarrow \quad [F(v)]_B = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

$$\Rightarrow \quad [F]_B = \begin{pmatrix} | & | & | \\ [F(u_1)]_B & [F(u_2)]_B & [F(v)]_B \\ | & | & | \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

In addition, the change-of-basis matrix is

$$S_B = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

and its inverse is

$$S_B^{-1} = \frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ -1 & 1 & 2 \\ 1 & -1 & 1 \end{pmatrix}.$$

Therefore,

$$[F] = S_B[F]_B S_B^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ -1 & 1 & 2 \\ 1 & -1 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}.$$

Exercises

Exercise 39. Let $U = \operatorname{span}\{u_1, u_2, u_3, u_4\} \in \mathbb{R}^4$, where

$$u_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 2 \end{pmatrix}, \quad u_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

(i) Determine a basis $B = (b_1, \ldots, b_m)$ of U.

(ii) Calculate the coordinate vectors $[u_j]_B \in \mathbb{R}^m$ for $j = 1, 2, 3, 4$.

(iii) For which values of $a \in \mathbb{R}$ does the vector $x = \begin{pmatrix} 3 \\ 3+a \\ 3+a \\ 2+2a \end{pmatrix}$ belong to U? For such a determine the coordinate vector $[x]_B$.

12

Orthonormal bases & Gram-Schmidt algorithm

In this chapter, we will discuss a special type of basis for a subspace. Before introducing any concepts in this chapter, let us recall some notations that were introduced in Definition 5.1 of Chapter 5:

(i) If $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$, then the **dot product** of u and v is defined by

$$u \bullet v = u_1 v_1 + \cdots + u_n v_n \,.$$

(ii) $u, v \in \mathbb{R}^n$ are called **orthogonal** if $u \bullet v = 0$.

(iii) The **norm** of $u \in \mathbb{R}^n$ is defined by $\|u\| = \sqrt{u \bullet u}$.

In addition, the norm and the dot product have some useful properties as stated in the following proposition.

Proposition 12.1 *Let $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$.*
(i) $\|\lambda x\| = |\lambda| \cdot \|x\|$.
(ii) $|x \bullet y| \leq \|x\| \cdot \|y\|$ *(Cauchy-Schwartz inequality).*
The equality " $=$ " occurs when x, y are linearly dependent.
(iii) $\|x + y\| \leq \|x\| + \|y\|$ *(Triangle inequality).*
The equality " $=$ " also occurs for this inequality when x, y are linearly dependent.

Proof. Using the properties of the dot product from Proposition 5.2, we have

(i) $\|\lambda x\| = \sqrt{(\lambda x) \bullet (\lambda x)} = \sqrt{\lambda^2 (x \bullet x)} = |\lambda| \sqrt{x \bullet x} = |\lambda| \cdot \|x\|$.

(ii) For any $\mu \in \mathbb{R}$, we have

$$\begin{aligned} (x + \mu y) \bullet (x + \mu y) \geq 0 \quad &\Leftrightarrow \quad x \bullet x + (\mu y) \bullet x + x \bullet (\mu y) + (\mu y) \bullet (\mu y) \geq 0 \\ &\Leftrightarrow \quad \|x\|^2 + 2(x \bullet (\mu y)) + \|\mu y\|^2 \geq 0 \\ &\Leftrightarrow \quad \|x\|^2 + 2\mu(x \bullet y) + \mu^2 \|y\|^2 \geq 0. \end{aligned}$$

If $y = \mathbf{0}$, then the statement is trivial because $|x \bullet \mathbf{0}| = \|x\| \cdot \|\mathbf{0}\| = 0$. Therefore, we assume $y \neq \mathbf{0}$. Choosing $\mu = -(x \bullet y)/\|y\|^2$, we have

$$\|x\|^2 - 2\frac{x \bullet y}{\|y\|^2}(x \bullet y) + \frac{(x \bullet y)^2}{\|y\|^4}\|y\|^2 \geq 0 \quad \Leftrightarrow \quad \|x\|^2 \|y\|^2 \geq (x \bullet y)^2 \quad \Leftrightarrow \quad \|x\| \cdot \|y\| \geq |x \bullet y|.$$

(iii) We have

$$\|x + y\|^2 = (x + y) \bullet (x + y) = \|x\|^2 + 2(x \bullet y) + \|y\|^2 .$$

Using the Cauchy-Schwartz inequality, we get

$$\begin{aligned} & \|x + y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \\ \Leftrightarrow \quad & \|x + y\| \leq \|x\| + \|y\| , \end{aligned}$$

where the equivalence comes from the fact that $\|x + y\|$ and $\|x\| + \|y\|$ are both non-negative. □

Definition 12.2 (i) A vector $u \in \mathbb{R}^n$ is called a **unit vector** if $\|u\| = 1$. (i.e. $u \bullet u = 1$)

(ii) Every vector $u \in \mathbb{R}^n$ with $u \neq \mathbf{0}$ can be normalized by

$$\widehat{u} = \frac{1}{\|u\|} u .$$

The vector $\hat{u}$ is a unit vector and shows in the same direction as u.

(iii) Vectors $u_1, \ldots, u_l \in \mathbb{R}^n$ are called **orthonormal** if for $1 \leq i, j \leq l$,

$$u_i \bullet u_j = \begin{cases} 1 & \text{, if } i = j \\ 0 & \text{, if } i \neq j \end{cases} .$$

Equipped with these definitions, we can define the main object of this chapter as follows.

Definition 12.3 A basis $B = (b_1, \ldots, b_m)$ of a subspace U is called an **orthonormal basis (ONB)** of U if $b_1, \ldots, b_m$ are orthonormal.

Example 39 For the subspace $U = \mathbb{R}^n$, the standard basis $B = (e_1, \ldots, e_n)$ is an orthonormal basis.

The following proposition demonstrates some useful properties of orthonormal bases.

Proposition 12.4 *(i) If $v_1, \ldots, v_m \in \mathbb{R}^n$ are orthonormal (ON), then they are linearly independent.*

(ii) Let $B = (v_1, \ldots, v_m)$ be an ONB of $V \subset \mathbb{R}^n$ and $u \in V$. Then

$$[u]_B = \begin{pmatrix} u \bullet v_1 \\ \vdots \\ u \bullet v_m \end{pmatrix} \in \mathbb{R}^m ,$$

i.e. $u = \sum_{i=1}^m (u \bullet v_i) v_i$.

(iii) If $B = (v_1, \ldots, v_m)$ is an ONB of $V \subset \mathbb{R}^n$ and $u, w \in V$, then

$$u \bullet w = [u]_B \bullet [w]_B .$$

Proof. (i) Assume that $v_1, \ldots, v_m$ are ON and $\lambda_1 v_1 + \ldots + \lambda_m v_m = \mathbf{0}$. For any $1 \leq j \leq m$, we have

$$v_j \bullet (\lambda_1 v_1 + \ldots + \lambda_m v_m) = 0 = \lambda_1 (v_j \bullet v_1) + \ldots + \lambda_m (v_j \bullet v_m) = \lambda_j .$$

Hence, $\lambda_1 = \ldots = \lambda_m = 0$, which shows that $v_1, \ldots, v_m$ are linearly independent.

(ii) Since $u \in V$, we can write $u = \lambda_1 v_1 + \ldots + \lambda_m v_m$. Doing the same calculations as in (i), we get

$$\text{For } 1 \le j \le m, \quad u \bullet v_j = \lambda_j \quad \Rightarrow \quad u = (u \bullet v_1)v_1 + \ldots + (u \bullet v_m)v_m \quad \Rightarrow \quad [u]_B = \begin{pmatrix} u \bullet v_1 \\ \vdots \\ u \bullet v_m \end{pmatrix}.$$

(iii) Let $[u]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, $[w]_B = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$. We have

$$u \bullet w = (x_1 v_1 + \ldots + x_m v_m) \bullet (y_1 v_1 + \ldots + y_m v_m) = x_1 y_1 + \ldots + x_m y_m = [u]_B \bullet [w]_B. \qquad \square$$

Now we want to introduce another important concept in the following definition.

Definition 12.5 For a subspace $U \subset \mathbb{R}^n$ we define the **orthogonal complement of U in $\mathbb{R}^n$** by

$$U^\perp = \{x \in \mathbb{R}^n \mid x \bullet u = 0 \text{ for all } u \in U\} .$$

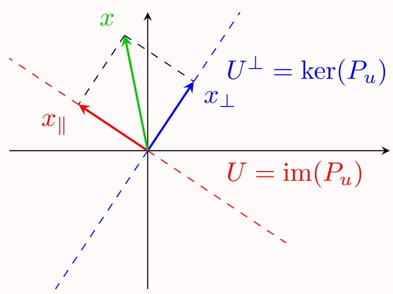
Example 40 Consider the orthogonal projection $P_u : \mathbb{R}^2 \to \mathbb{R}^2$ with $u = \frac{1}{\sqrt{13}}\begin{pmatrix} -3 \\ 2 \end{pmatrix}$.
Recall that $P_u(x) = x_\parallel = \frac{u \bullet x}{u \bullet u} u$ for any $x \in \mathbb{R}^2$.
In this case, $\|u\| = 1$ so we have

$$P_u(x) = x_\parallel = (u \bullet x)u,$$

for any $x \in \mathbb{R}^2$. Clearly, $\{u\}$ is an ONB of the subspace $U = \mathrm{im}(P_u)$. We also know that $\ker(P_u) = \{x \in \mathbb{R}^2 \mid x \bullet u = 0\}$. Fix an element $v \in \ker(P_u)$. For any $w \in U$, we have $w = \lambda u$ for some $\lambda \in \mathbb{R}$ and

$$v \bullet w = v \bullet (\lambda u) = \lambda(v \bullet u) = 0,$$

which implies that $v \in U^\perp$. Thus, $\ker(P_u) \subset U^\perp$. Furthermore, it is clear that $U^\perp \subset \ker(P_u)$. As a result, $\ker(P_u) = U^\perp$, which shows that $U^\perp$ is a subspace.



Motivated by this example, we get the following lemma.

Lemma 12.6 *Let $U \subset \mathbb{R}^n$ be a subspace.*

(i) $U^\perp \subset \mathbb{R}^n$ is a subspace.

(ii) We have $U \cap U^\perp = \{\mathbf{0}\}$.

(iii) If $(u_1, \ldots, u_r)$ is a basis of U, $x \in \mathbb{R}^n$, then

$$x \in U^\perp \iff x \bullet u_1 = \cdots = x \bullet u_r = 0 .$$

(iv) Let $(f_1, \ldots, f_r)$ be an ONB of U and $x \in \mathbb{R}^n$. Then

$$x = x_\parallel + x_\perp ,$$

where

$$x_{\|} = \sum_{i=1}^{r}(x \bullet f_i)f_i \in U$$
$$x_{\perp} = x - x_{\|} \in U^{\perp}.$$

Proof. (i) Clearly, $\mathbf{0} \in U^{\perp}$ since $\mathbf{0} \bullet u = 0$ for any $u \in U$. Given any $x, y \in U^{\perp}$ and $\lambda \in \mathbb{R}$, we have for all $u \in U$,

$$(x+y) \bullet u = x \bullet u + y \bullet u = 0 + 0 = 0 \quad \Rightarrow \quad x + y \in U^{\perp},$$

$$(\lambda x) \bullet u = \lambda(x \bullet u) = \lambda \cdot 0 = 0 \quad \Rightarrow \quad \lambda x \in U^{\perp}.$$

Therefore, $U^{\perp}$ is a subspace of $\mathbb{R}^n$.

(ii) If $x \in U \cap U^{\perp}$, then $x \bullet x = 0 = x_1^2 + \ldots + x_n^2 \quad \Rightarrow \quad x_1 = \ldots = x_n = 0 \quad \Rightarrow \quad x = \mathbf{0}$.

(iii) " $\Longrightarrow$ " is clear.

" $\Longleftarrow$ " : For all $w \in U$, we have $w = \lambda_1 u_1 + \ldots + \lambda_m u_m$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$. Then, we have

$$x \bullet w = x \bullet (\lambda_1 u_1 + \ldots + \lambda_m u_m) = \lambda_1(x \bullet u_1) + \ldots + \lambda_m(x \bullet u_m) = 0 + \ldots + 0 = 0.$$

Hence, $x \in U^{\perp}$.

(iv) $x_{\|} = \sum_{i=1}^{r}(x \bullet f_i)f_i \in U$ since $(f_1, \ldots, f_r)$ is a basis of U. We want to show that $x_{\perp} \in U^{\perp}$.

For all $1 \leq j \leq r$:

$$f_j \bullet x_{\perp} = f_j \bullet (x - x_{\|}) = f_j \bullet x - f_j \bullet \sum_{i=1}^{r}(x \bullet f_i)f_i = f_j \bullet x - x \bullet f_j = 0.$$

Hence, $x_{\perp} \in U^{\perp}$ by using (iii). □

We see that orthonormal bases are extremely useful for many calculations, so we may be concerned about how to get them. Fortunately, the following algorithm allows us to obtain an orthonormal basis from an arbitrary basis of any subspace.

Algorithm 12.7 (Gram-Schmidt algorithm (GSA)) Let $B = (b_1, \ldots, b_m)$ be an arbitrary basis of a subspace $U \subset \mathbb{R}^n$. The GSA constructs an orthonormal basis $F = (f_1, \ldots, f_m)$ of U out of the basis B in the following m steps:

Step 1: Set $f_1 = \widehat{b_1} = \frac{1}{\|b_1\|} b_1$.

Step l $(2 \leq l \leq m)$: We have constructed orthonormal vectors $f_1, \ldots, f_{l-1}$ in the steps before. Now set

$$w_l = b_l - (b_l \bullet f_1)f_1 - \cdots - (b_l \bullet f_{l-1})f_{l-1} = b_l - \sum_{i=1}^{l-1}(b_l \bullet f_i)f_i$$

and define $f_l = \widehat{w_l} = \frac{1}{\|w_l\|} w_l$.

Example 41 Consider $b_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $b_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$, $b_3 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$. $B = (b_1, b_2, b_3)$ is a basis of $\mathbb{R}^3$. We will construct an ONB $F = (f_1, f_2, f_3)$ by GSA and demonstrate why it works.

Step 1: Set $f_1 = \widehat{b_1} = \frac{1}{\|b_1\|} b_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $U_1 = \operatorname{span}\{f_1\} = \{b_1\}$.

Step 2: We want to find a vector $f_2 \in \operatorname{span}\{f_1, b_2\}$ such that f_2 is orthogonal to f_1. From Lemma 12.6 (iv), we have

$$b_2 = b_{2\|} + b_{2\perp} = (b_2 \bullet f_1) f_1 + b_{2\perp},$$

where $b_{2\|} \in U_1$ and $b_{2\perp} \in U_1^{\perp}$.

Set $w_2 = b_{2\perp} = b_2 - (b_2 \bullet f_1) f_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \frac{3}{\sqrt{3}} f_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$.

Then, set $f_2 = \widehat{w_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$ and $U_2 = \operatorname{span}\{f_1, f_2\} = \operatorname{span}\{b_1, b_2\}$.

Hence, (f_1, f_2) is an ONB of U_2.

Step 3: Now, we want to find a vector $f_3 \in \operatorname{span}\{f_1, f_2, b_3\}$ such that $f_3 \in U_2^{\perp}$. Again, from Lemma 12.6 (iv), we have

$$b_3 = b_{3\|} + b_{3\perp} = (b_3 \bullet f_1) f_1 + (b_3 \bullet f_2) f_2 + b_{3\perp},$$

where $b_3 = b_{3\|} \in U_2$ and $b_{3\perp} \in U_2^{\perp}$.

Set $w_3 = b_{3\perp} = b_3 - (b_3 \bullet f_1) f_1 - (b_3 \bullet f_2) f_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \frac{5}{6} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$.

Then, set $f_3 = \widehat{w_3} = \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$ and

$$U_3 = \operatorname{span}\{f_1, f_2, f_3\} = \operatorname{span}\{b_1, b_2, f_3\} = \operatorname{span}\{b_1, b_2, b_3\}.$$

Hence, $B = (f_1, f_2, f_3)$ is an ONB of $U_3 = \mathbb{R}^3$ from that basis.

As a consequence of the Gram-Schmidt algorithm, we get the following theorem.

Theorem 12.8 *Every subspace of $\mathbb{R}^n$ has an ONB.*

Proof. According to Theorem 10.2 (i), every subspace has a basis. Using GSA, we get an ONB. □

Corollary 12.9 *Let $U \subset \mathbb{R}^n$ be a subspace. For all $x \in \mathbb{R}^n$ there exist unique $x_\| \in U$ and $x_\perp \in U^\perp$ with*

$$x = x_\| + x_\perp \,.$$

Proof. **Existence:** By Theorem 12.8, there exists an ONB $(f_1, \ldots, f_m)$ of U. And by Lemma 12.6 (iv), we get $x_\|$ and $x_\perp$.

Uniqueness: Let $x = x_\parallel + x_\perp = y_\parallel + y_\perp$ for $x_\parallel, y_\parallel \in U$ and $x_\perp, y_\perp \in U^\perp$. Then, we have

$$U \ni x_\parallel - y_\parallel = y_\perp - x_\perp \in U^\perp .$$

Hence, $x_\parallel - y_\parallel \in U \cap U^\perp$ and $y_\perp - x_\perp \in U \cap U^\perp$. However, since $U \cap U^\perp = \{\mathbf{0}\}$ by Lemma 12.6 (ii), we have

$$x_\parallel - y_\parallel = y_\perp - x_\perp = \mathbf{0} \quad \Longrightarrow \quad x_\parallel = y_\parallel \text{ and } y_\perp = x_\perp . \qquad \square$$

Exercises

Exercise 40. Let $U \subset \mathbb{R}^n$ be a subspace with orthonormal basis $(f_1, \dots, f_r)$. We define the orthogonal projection onto U by

$$\begin{aligned} P_U : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto \sum_{i=1}^{r} (x \bullet f_i) f_i \, . \end{aligned}$$

Show the following properties of P_U:

(i) If $U = \operatorname{span}\{u\}$ with $u \in \mathbb{R}^n$ and $u \neq \mathbf{0}$ then P_U is the projection P_u we defined in Chapter 5.

(ii) P_U is a linear map.

(iii) $P_U \circ P_U = P_U$.

(iv) $\operatorname{im} P_U = U$ and $\ker(P_U) = U^\perp$, where $U^\perp$ is the orthogonal complement of U defined by

$$U^\perp = \{x \in \mathbb{R}^n \mid x \bullet u = 0 \text{ for all } u \in U\} \, .$$

Exercise 41. We define the following vectors

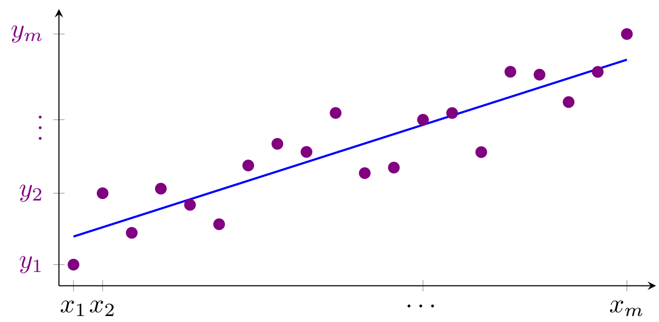
$$b_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \qquad b_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \qquad b_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \end{pmatrix} .$$

These form a basis $B = (b_1, b_2, b_3)$ of the subspace $U = \operatorname{span}\{b_1, b_2, b_3\} \subset \mathbb{R}^4$ (You do not need to show this). Use the Gram-Schmidt algorithm to construct an orthonormal basis $F = (f_1, f_2, f_3)$ of U from B.

13

Orthogonal Projection & Least squares

Assume you measure some data $(x_1, y_1), \ldots, (x_m, y_m)$, and you want to find a *line* which interpolates these points in the best possible way. If all points would lie on a line $\ell(x) = ax + b$, then they would satisfy



$$\begin{cases} ax_1 + b = y_1 \\ ax_2 + b = y_2 \\ \quad\vdots \\ ax_m + b = y_m \end{cases} \iff \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_m & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \iff A \begin{pmatrix} a \\ b \end{pmatrix} = y \quad (*).$$

However, if they are not on one line (like in the picture), then the linear system $(*)$ has no solutions because $y \notin \operatorname{im}(A)$. Nevertheless, in the picture, we see that there might be a "best possible" line.

This chapter aims to explain how to obtain this "best-fit line." In general, for some matrix $A \in \mathbb{R}^{m \times n}$ and some vector $y \in \mathbb{R}^m$, we want to find a vector $x \in \mathbb{R}^n$ such that Ax is the closest point to y. The main idea is to project y onto the image of A and then obtain a linear system we can solve for x. Later, we will see that x can be obtained by solving the normal equation

$$A^T Ax = A^T y.$$

From Corollary 12.9, for a subspace $U \subset \mathbb{R}^n$ and a vector $x \in \mathbb{R}^n$, there uniquely exist $x_\perp \in U^\perp$ and $x_\parallel \in U$ with $x = x_\perp + x_\parallel$. Hence, we have the following map.

Definition 13.1 Let $U \subset \mathbb{R}^n$ be a subspace. The map

$$\begin{aligned} P_U : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto x_\parallel \end{aligned}$$

is the **orthogonal projection** onto U.

Remark. This generalizes the P_u for $u \in \mathbb{R}^n,\ u \neq 0$ we defined before by setting $U = \operatorname{span}\{u\}$.

Proposition 13.2 *Let $U \subset \mathbb{R}^n$ be a subspace.*

(i) P_U *is a linear map.*

(ii) $P_U^2 = P_U$.

(iii) $\ker(P_U) = U^\perp$ *and* $\operatorname{im} P_U = U$.

(iv) If $(f_1, \ldots, f_m)$ *is an ONB of* U*, then*

$$P_U(x) = (x \bullet f_1)f_1 + \cdots + (x \bullet f_m)f_m\,.$$

Proof. (iv) is exactly Lemma 12.6 (iv). Using that, we can prove the other statements.

(i) P_U is linear because for any $x,\ y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, we have

$$P_U(x+y) = \sum_{i=1}^m \big((x+y) \bullet f_i\big)f_i = \sum_{i=1}^m (x \bullet f_i)f_i + \sum_{i=1}^m (y \bullet f_i)f_i = P_U(x) + P_U(y)$$

$$P_U(\lambda x) = \sum_{i=1}^m \big((\lambda x) \bullet f_i\big)f_i = \lambda \sum_{i=1}^m (x \bullet f_i)f_i = \lambda P_U(x)$$

(ii) For any $x \in \mathbb{R}^n$ and any j such that $1 \leq j \leq m$, we have

$$P_U(x) \bullet f_j = f_j \bullet P_U(x) = f_j \bullet \sum_{i=1}^m (x \bullet f_i)f_i = \sum_{i=1}^m (x \bullet f_i)(f_j \bullet f_i) = x \bullet f_j.$$

Hence, for any $x \in \mathbb{R}^n$, we have

$$P_U \circ P_U(x) = P_U\big(P_U(x)\big) = \sum_{i=1}^m \big(P_U(x) \bullet f_i\big)f_i = \sum_{i=1}^m (x \bullet f_i)f_i = P_U(x).$$

Thus, $P_U^2 = P_U \circ P_U = P_U$.

(iii) For the kernel, we have

$$\begin{aligned}
x \in \ker(P_U) \quad &\Leftrightarrow \quad P_U(x) = \mathbf{0} \\
&\Leftrightarrow \quad \sum_{i=1}^m (x \bullet f_i)f_i = \mathbf{0} \\
&\Leftrightarrow \quad x \bullet f_1 = \ldots = x \bullet f_m = 0 && (f_1, \ldots, f_m \text{ are linearly independent}) \\
&\Leftrightarrow \quad x \in U^\perp && (\text{by Lemma 12.6 (iii)})
\end{aligned}$$

Hence, $\ker(P_U) = U^\perp$.

For the image, if $u \in \operatorname{im}(P_U)$, then clearly $u \in U$ because $u \in \operatorname{span}\{f_1, \ldots, f_m\} = U$. Otherwise, if $u \in U$, then we have by Proposition 12.4 (ii),

$$u = \sum_{i=1}^m (u \bullet f_i)f_i = P_U(x),$$

which implies that $u \in \operatorname{im}(P_U)$. Therefore, $u \in \operatorname{im}(P_U)$ if and only if $u \in U$, i.e. $\operatorname{im}(P_U) = U$. $\square$

Proposition 13.3 *Let $U \subset \mathbb{R}^n$ be a subspace and $x \in \mathbb{R}^n$. Then for all $u \in U$ we have*

$$\|x - P_U(x)\| \leq \|x - u\| \ .$$

We just have equality in the case when $u = P_U(x)$. In other words, if x is outside of U, then $P_U(x)$ is the closest point to x which is in U.

Proof. For any $u \in U$, doing the same calculation as the one in the proof of Proposition 12.1 (ii), we get

$$\begin{aligned} \|x - u\|^2 &= \|(x - P_U(x)) + (P_U(x) - u)\|^2 \\ &= \|x - P_U(x)\|^2 + 2\big((x - P_U(x)) \bullet (P_U(x) - u)\big) + \|P_U(x) - u\|^2 \end{aligned}$$

From Lemma 12.6 (iv), we have $(x - P_U(x)) = x_\perp \in U^\perp$. Because $(P_U(x) - u) \in U$, we have

$$(x - P_U(x)) \bullet (P_U(x) - u) = 0.$$

Therefore,

$$\|x - u\|^2 = \|x - P_U(x)\|^2 + \|P_U(x) - u\|^2 \geq \|x - P_U(x)\|^2 \quad \Leftrightarrow \quad \|x - u\| \geq \|x - P_U(x)\| \, .$$

The equality occurs when $\|P_U(x) - u\| = 0 \quad \Leftrightarrow \quad u = P_U(x)$. ☐

Definition 13.4 The **transpose** of a matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ is the matrix $A^T = (a_{ji}) \in \mathbb{R}^{n\times m}$.

Example 42 Given the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathbb{R}^{2\times 3},$$

its transpose is

$$A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

Proposition 13.5 *(i) For $A, B \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}$ we have*

$$(A + B)^T = A^T + B^T \, , \qquad (\lambda A)^T = \lambda A^T \, .$$

(ii) For $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times l}$ we have

$$(AB)^T = B^T A^T \in \mathbb{R}^{l\times m} \, .$$

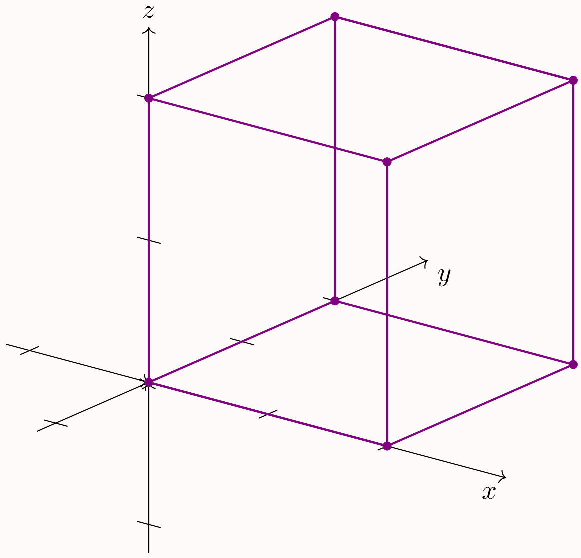
(iii) For $x, y \in \mathbb{R}^n$ we have $x \bullet y = x^T y$.

Proof. This can be checked by direct calculations. ☐

Example 43 (Basics behind 3D-Graphics) In this example, we want to project a cube in $\mathbb{R}^3$ to $\mathbb{R}^2$. This has the natural application of visualizing 3D-Graphics (in our case a cube) on a monitor. Our cube will have side-length 2 and its vertices are given by the following 8 points:

$$C = \left\{ \begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}2\\0\\0\end{pmatrix}, \begin{pmatrix}0\\2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\2\end{pmatrix}, \begin{pmatrix}2\\2\\0\end{pmatrix}, \begin{pmatrix}2\\0\\2\end{pmatrix}, \begin{pmatrix}0\\2\\2\end{pmatrix}, \begin{pmatrix}2\\2\\2\end{pmatrix} \right\}.$$

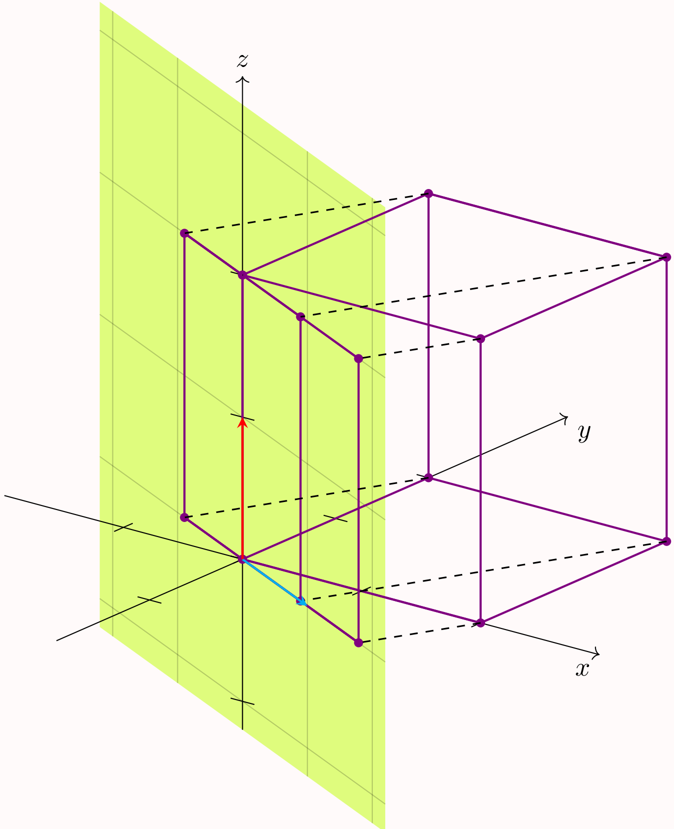
To make it look like a cube we connect two points if they share a facet, i.e. if their coordinates just differ by one entry. We obtain the following picture:

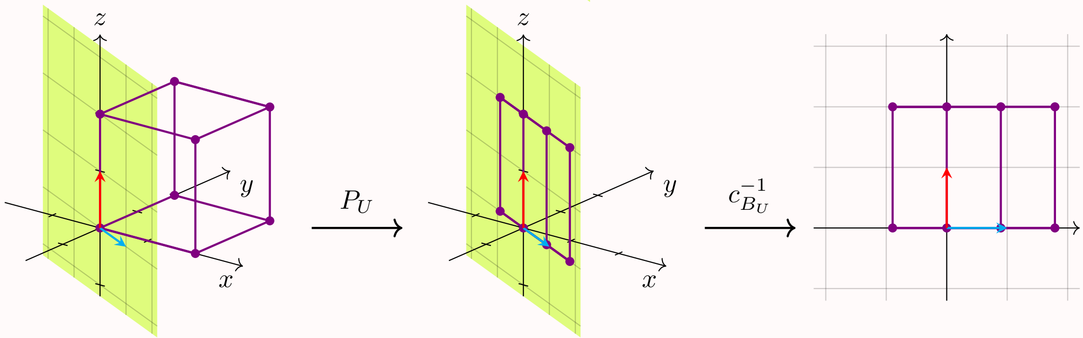


Now we want to project this cube onto a plane, which determines the viewing angle onto the scene. Let $U = \operatorname{span}\{f_1, f_2\} \subset \mathbb{R}^3$ with

$$f_1 = \frac{1}{\sqrt{5}}\begin{pmatrix}2\\-1\\0\end{pmatrix} \approx \begin{pmatrix}0.89\\-0.45\\0\end{pmatrix}, \qquad f_2 = \begin{pmatrix}0\\0\\1\end{pmatrix}.$$

We want to project the points in C onto the plane U and then calculate the coordinates with respect to the orthonormal basis $B_U = (f_1, f_2)$. For each $x \in C$ we now want to calculate $[P_U(x)]_{B_U}$. By Proposition 13.2 we have $P_U(x) = (x \bullet f_1)f_1 + (x \bullet f_2)f_2$.





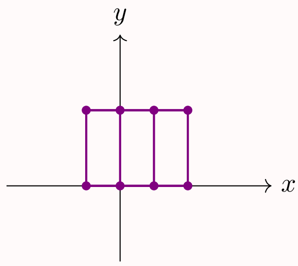
The coordinates are $[P_U(x)]_{B_U} = \begin{pmatrix} x \bullet f_1 \\ x \bullet f_2 \end{pmatrix} = \begin{pmatrix} f_1^T x \\ f_2^T x \end{pmatrix}$. Or in other words, we have

$$[P_U(x)]_{B_U} = (c_{B_U}^{-1} \circ P_U)(x) = \begin{pmatrix} -f_1^T- \\ -f_2^T- \end{pmatrix} x = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix} x.$$

Multiplying each element in C with this matrix gives the following set

$$P = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{4}{\sqrt{5}} \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} \frac{2}{\sqrt{5}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{4}{\sqrt{5}} \\ 2 \end{pmatrix}, \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ 2 \end{pmatrix}, \begin{pmatrix} \frac{2}{\sqrt{5}} \\ 2 \end{pmatrix} \right\}.$$

Drawing these points in $\mathbb{R}^2$ gives a picture of the cube viewed from the side:



Now consider another plane $V = \operatorname{span}\{v_1, v_2\}$ spanned by

$$v_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \qquad v_2 = \frac{1}{\sqrt{21}}\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.$$

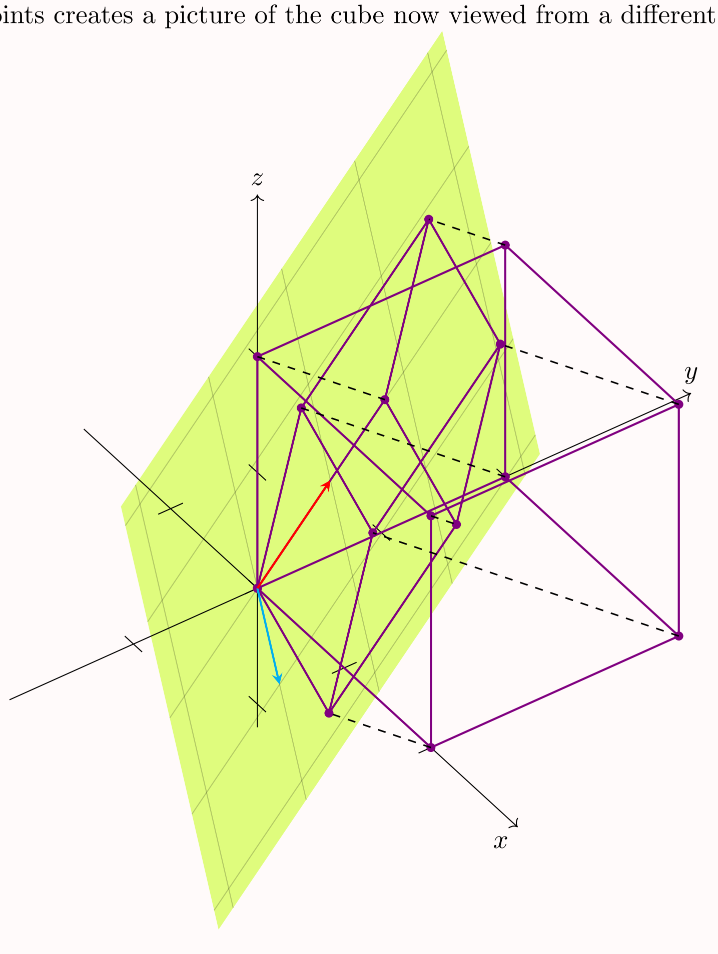
Notice that this again gives an orthonormal basis $B_V = (v_1, v_2)$ of V. As before we get

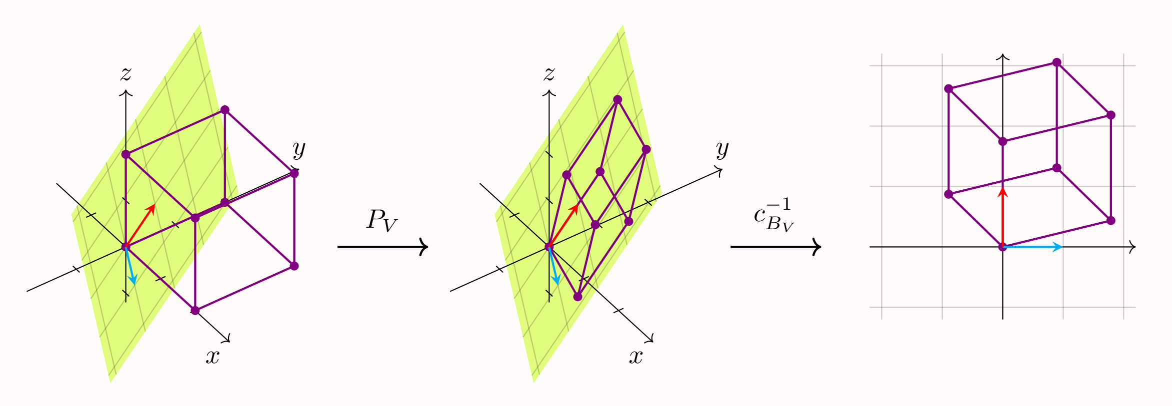
$$[P_V(x)]_{B_V} = (c_{B_V}^{-1} \circ P_V)(x) = \begin{pmatrix} -v_1^T- \\ -v_2^T- \end{pmatrix} x = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{21}} & \frac{2}{\sqrt{21}} & \frac{4}{\sqrt{21}} \end{pmatrix} x.$$

and applying this to each element in C gives

$$Q = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{4}{\sqrt{5}} \\ \frac{2}{\sqrt{21}} \end{pmatrix}, \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ \frac{4}{\sqrt{21}} \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{8}{\sqrt{21}} \end{pmatrix}, \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{6}{\sqrt{21}} \end{pmatrix}, \begin{pmatrix} \frac{4}{\sqrt{5}} \\ \frac{10}{\sqrt{21}} \end{pmatrix}, \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ \frac{12}{\sqrt{21}} \end{pmatrix}, \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{14}{\sqrt{21}} \end{pmatrix} \right\}.$$

Plotting these points creates a picture of the cube now viewed from a different angle:





For any $A \in \mathbb{R}^{m \times n}$, we can define a linear map $F : x \mapsto Ax$. Using that, we define the image and kernel of the matrix A by $\operatorname{im}(A) = \operatorname{im}(F)$ and $\ker(A) = \ker(F)$.

Proposition 13.6 *For all $A \in \mathbb{R}^{m \times n}$ we have $\operatorname{im}(A)^\perp = \ker(A^T)$.*

Proof. Let $x \in \mathbb{R}^n$. Then we have

$$\begin{aligned}
x \in \operatorname{im}(A)^\perp \quad &\Leftrightarrow \quad y \bullet x = 0, \quad \forall y \in \operatorname{im}(A) \\
&\Leftrightarrow \quad (Av) \bullet x = 0, \quad \forall v \in \mathbb{R}^n \\
&\Leftrightarrow \quad (Av)^T x = 0, \quad \forall v \in \mathbb{R}^n && \text{(by Proposition 13.5 (iii))} \\
&\Leftrightarrow \quad v^T A^T x = 0, \quad \forall v \in \mathbb{R}^n && \text{(by Proposition 13.5 (ii))} \\
&\Leftrightarrow \quad v \bullet \left(A^T x\right) = 0, \quad \forall v \in \mathbb{R}^n && \text{(by Proposition 13.5 (iii))} \\
&\Leftrightarrow \quad A^T x = \mathbf{0} \\
&\Leftrightarrow \quad x \in \ker\left(A^T\right).
\end{aligned}$$

□

Corollary 13.7 *Let $A \in \mathbb{R}^{m \times n}$.*

(i) We have $\ker(A^T A) = \ker(A)$.

(ii) We have the following equivalence

$$\ker(A) = \{\mathbf{0}\} \quad \Longleftrightarrow \quad A^T A \in \mathbb{R}^{n \times n} \text{ is invertible}\,.$$

Proof. (i) We have

$$\begin{aligned}
x \in \ker\left(A^T A\right) \quad \Leftrightarrow \quad A^T A x = \mathbf{0} \quad &\Leftrightarrow \quad Ax \in \ker\left(A^T\right) = \operatorname{im}(A)^\perp && \text{(Proposition 13.6)} \\
&\Leftrightarrow \quad Ax \in \operatorname{im}(A) \cap \operatorname{im}(A)^\perp = \{\mathbf{0}\} && \text{(Lemma 12.6 (ii))} \\
&\Leftrightarrow \quad Ax = \mathbf{0} \\
&\Leftrightarrow \quad x \in \ker(A).
\end{aligned}$$

(ii) Since $A^T A$ is a $n \times n$ matrix, we have

$$\begin{aligned}
\ker(A) = \{\mathbf{0}\} \quad &\Longleftrightarrow \quad \ker(A^T A) = \{\mathbf{0}\} \\
&\Longleftrightarrow \quad A^T A \text{ is invertible} && \text{(Theorem 8.7)}
\end{aligned}$$

□

We can now use our results to answer the question at the beginning of this chapter. Here, we will present the **least squares method**.

Problem: Given a linear map $F : \mathbb{R}^n \to \mathbb{R}^m$ and $b \in \mathbb{R}^m$, we want to find $x \in \mathbb{R}^n$ that minimizes the quantity

$$\delta = \|F(x) - b\|.$$

The "least squares" stems from the fact that if $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ and $F(x) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, then

$$\delta = \|F(x) - b\| = \sqrt{(y_1 - b_1)^2 + \ldots + (y_m - b_m)^2}.$$

In other words, the problem is finding $x \in \mathbb{R}^n$ that minimizes the sum of squares of the difference $y_i - b_i$.

Remark. Minimal δ is 0 if and only if $b \in \mathrm{im}(F)$, i.e. $F(x) = b$ has a solution.

By Proposition 13.3, the minimal δ is given in the case $F(x) = P_{\mathrm{im}(F)}(b)$. Writing $[F] = A$, we want to find $x \in \mathbb{R}^n$ such that $Ax = P_{\mathrm{im}(F)}(b)$. We have

$$\begin{aligned} Ax = P_{\mathrm{im}(F)} \quad &\Leftrightarrow \quad (Ax - b) \in (\mathrm{im}(A))^\perp = \ker(A^T) && \text{(by Proposition 13.6)} \\ &\Leftrightarrow \quad A^T(Ax - b) = \mathbf{0} \\ &\Leftrightarrow \quad \underbrace{\boxed{A^T A x = A^T b}}_{\textbf{normal equation}} \end{aligned}$$

Therefore, if $\ker(A) = \{\mathbf{0}\}$ (i.e., the columns of A are linearly independent), then by Corollary 13.7, $A^T A$ is invertible and we get the unique solution to our problem by

$$x = \left(A^T A\right)^{-1} A^T b.$$

Example 44 Find the best possible quadratic polynomial $f(t) = a_0 + a_1 t + a_2 t^2$ to fit the data points $(0, 2),\ (1, 1),\ (2, 2),\ (3, 3)$.
We first translate this problem into linear algebra: We want to minimize

$$(f(0) - 2)^2 + (f(1) - 1)^2 + (f(2) - 2)^2 + (f(3) - 3)^2,$$

so we define the linear map $F : \mathbb{R}^3 \to \mathbb{R}^4$ for $x = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \in \mathbb{R}^3$ by

$$F(x) = \begin{pmatrix} f(0) \\ f(1) \\ f(2) \\ f(3) \end{pmatrix} = \begin{pmatrix} a_0 + a_1 \cdot 0 + a_2 \cdot 0^2 \\ a_0 + a_1 \cdot 1 + a_2 \cdot 1^2 \\ a_0 + a_1 \cdot 2 + a_2 \cdot 2^2 \\ a_0 + a_1 \cdot 3 + a_2 \cdot 3^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = Ax,$$

where $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$. Then, we want to find $x \in \mathbb{R}^3$ such that $\|Ax - b\|$ with $b = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 3 \end{pmatrix}$ is minimal. Therefore, we need to solve the normal equation $A^T A x = A^T b$.

We have

$$A^T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix}, \qquad A^T A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} = \begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix},$$

$$A^T b = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 14 \\ 36 \end{pmatrix}.$$

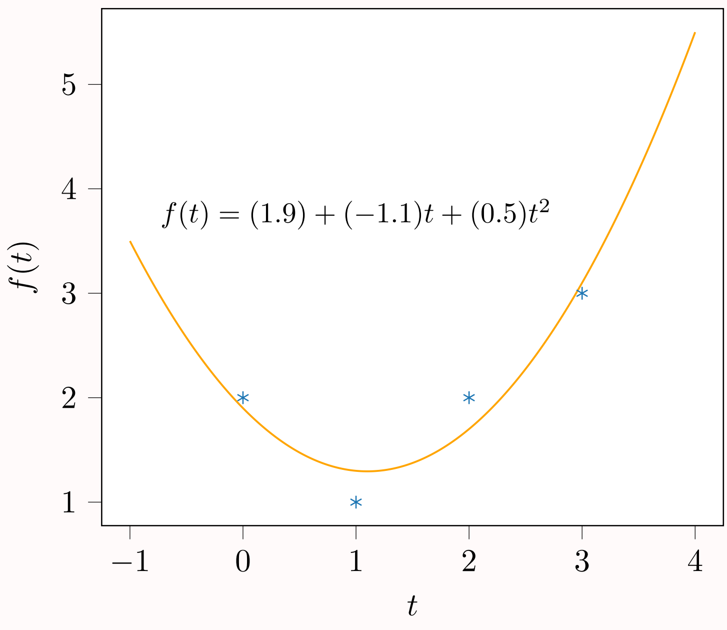
Then, we want solve the linear system

$$\begin{pmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 14 \\ 36 \end{pmatrix}.$$

This linear system has a unique solution $\begin{cases} a_0 = \frac{19}{10} \\ a_1 = -\frac{11}{10} \\ a_2 = \frac{1}{2} \end{cases}$.

Hence, the best fit polynomial is

$$f(t) = \frac{19}{10} - \frac{11}{10}t + \frac{1}{2}t^2.$$



Remark. This method works for arbitrary polynomials (i.e. in particular for lines). In addition, the normal equation always have a unique solution if the columns of A are linearly indepdendent. In applications, this is usually the case since the number of data points (the number m of rows of A) is usually greater than the degree of the polynomial $(n-1)$, where n is the number of columns of A.

Exercises

Exercise 42. Three points (x_1, y_1), (x_2, y_2) and (x_3, y_3) lie on one (non-vertical) line, if there exist $a, b \in \mathbb{R}$ with $ax_j + b = y_j$ for $j = 1, 2, 3$. In other words the linear system

$$\underbrace{\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \end{pmatrix}}_{=A} \begin{pmatrix} a \\ b \end{pmatrix} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}_{=y}$$

has a solution, i.e. $y \in \operatorname{im}(A)$.

(i) Show that the points $(0, 1)$, $(1, 3)$ and $(2, 2)$ do <u>not</u> lie on one line.

For (ii) - (iv) we assume that $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$ and $y = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$.

(ii) Calculate an orthonormal basis $F = (f_1, f_2)$ of $\operatorname{im}(A)$ by using the GSA for the columns of A.
(Hint: The result becomes nicer if you set b_1 = second column of A and b_2 = first column of A.)

(iii) Calculate $z = (y \bullet f_1)f_1 + (y \bullet f_2)f_2$ and show that $z \in \operatorname{im}(A)$.

(iv) Solve the linear system $A \begin{pmatrix} a \\ b \end{pmatrix} = z$ and draw the graph of $f(x) = ax + b$ together with the three points in i). Can you interpret the connection between the graph and the points?

Exercise 43. Assume we have the following data points

i	1	2	3	4
x_i	0	1	2	3
y_i	2	1	3	4

(i) Find the line of best fit for the above data, i.e. find $a, b \in \mathbb{R}$ such that the function $l(x) = ax + b$ minimizes the sum of squares $\sum_{i=1}^{4} (l(x_i) - y_i)^2$.

(ii) Interpolate the data by a quadratic polynomial. For this find $c, d, e \in \mathbb{R}$ such that the function $p(x) = cx^2 + dx + e$ minimizes $\sum_{i=1}^{4} (p(x_i) - y_i)^2$.

(iii) Draw the data points and the graphs of l and p into one diagram.

For both (i) and (ii) solve the exercise by finding the solutions to the normal equation.

14

Midterm & Final exams

In the following you can find the midterms and final exams given in the fall semesters 2019 - 2022 at Nagoya University. Solutions can be found on the corresponding pages of the lectures at `https://www.henrikbachmann.com/teaching.html`.

14.1 Linear Algebra I - Midterm 2019

Exercise 1. (10 Points) Consider the following linear system

$$\left\{\begin{array}{ccccccccc} 2x_1 & + & 3x_2 & + & 4x_3 & + & 5x_4 & = & 6 \\ x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 5 \\ 3x_1 & + & 4x_2 & + & 5x_3 & + & 6x_4 & = & 7 \end{array}\right. .$$

(i) Find a matrix $A \in \mathbb{R}^{3\times 4}$ and and a vector $b \in \mathbb{R}^3$, such that the solutions of the above linear system are given by the vectors $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4$ satisfying $Ax = b$.
(ii) Determine the row-reduced echelon forms of the matrices $(A \mid b)$ and A.
(iii) Find all the solutions to the linear system.
(iv) Calculate the rank of $(A \mid b)$ and A.
(v) Find a vector $c \in \mathbb{R}^3$, such that $Ax = c$ has no solutions. Calculate the rank of $(A \mid c)$.

Exercise 2. (10 Points) Let $u = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ and define the following four functions:

$$\begin{aligned} f_1 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto (u \bullet x)u + x , \end{aligned} \qquad \begin{aligned} f_2 : \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto \begin{pmatrix} 2\cos(x) \\ \sin(x) \end{pmatrix} , \end{aligned}$$

$$\begin{aligned} f_3 : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto \frac{x \bullet x}{u \bullet u} u , \end{aligned} \qquad \begin{aligned} f_4 : \mathbb{R}^3 &\longrightarrow \mathbb{R}^4 \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &\longmapsto \begin{pmatrix} 0 \\ x_1 + 2x_2 + 3x_3 \\ x_1 - x_3 \\ x_2 \end{pmatrix} . \end{aligned}$$

(i) Which of the above functions f_1, f_2, f_3, f_4 are linear maps? For each one that is linear, determine its matrix.
(ii) Draw a picture of the image of f_2. Is f_2 injective and/or surjective?

Exercise 3. (6 Points) Let $G : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map with

$$G\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad G\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} .$$

(i) Determine the matrix of G.
(ii) Determine the matrix of $G \circ G$.

Exercise 4. (6 Points) We define the following linear map

$$\begin{aligned} H : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &\longmapsto \begin{pmatrix} x_1 \\ x_2 + x_3 \\ x_1 + x_2 + x_3 \end{pmatrix} . \end{aligned}$$

(i) Calculate the image of H.
(ii) Decide if H is injective and/or surjective.
(iii) Find all vectors $v \in \mathbb{R}^3$, which are orthogonal to all vectors in the image of H.

14.2 Linear Algebra I - Midterm 2020

Exercise 1. (10 Points) Consider the following linear system

$$\begin{cases} -2x_1 & + & 4x_2 & + & x_3 & + & x_4 & = & 6 \\ -3x_1 & + & 6x_2 & + & x_3 & & & = & 7 \\ x_1 & - & 2x_2 & & & + & x_4 & = & -1 \end{cases} .$$

(i) Find a matrix $A \in \mathbb{R}^{3\times 4}$ and and a vector $b \in \mathbb{R}^3$, such that the solutions of the above linear system are given by the vectors $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4$ satisfying $Ax = b$.
(ii) Determine the row-reduced echelon forms of the matrices $(A \mid b)$ and A.
(iii) Find all the solutions to the linear system.
(iv) Calculate the rank of $(A \mid b)$ and A.
(v) Find all $y \in \mathbb{R}^4$ with $Ay = 2b$ by using your result for iii).

Exercise 2. (8 Points) Let $u = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ and define the following four functions:

$$f_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}^3, \quad x \longmapsto \begin{pmatrix} u \bullet x \\ 0 \\ x \bullet u \end{pmatrix}, \qquad f_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto 2^{x_1+x_2} - 1, \qquad f_3 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} x_1 - 3x_2 \\ 2x_1 + x_2x_3 \end{pmatrix} .$$

(i) Which of the above functions f_1, f_2, f_3 are linear maps? For each one that is linear, determine its matrix.
(ii) Is f_2 injective and/or surjective?

Exercise 3. (8 Points)
(i) Let $G : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map with

$$G\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad G\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} .$$

Determine the matrix of G.
(ii) Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be a function with

$$F\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad F\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad F\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} .$$

Show that F is not a linear map.

Exercise 4. (8 Points) We define the following linear map

$$H : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} x_1 + x_2 \\ x_1 - x_3 \\ x_2 + x_3 \end{pmatrix} .$$

(i) Calculate the image of H.
(ii) Decide if H is injective and/or surjective.
(iii) Find a non-zero vector $v \in \mathbb{R}^3$, such that v is orthogonal to $H(v)$. (Just one explicit vector is enough)

14.3 Linear Algebra I - Midterm 2021

Exercise 1. (10 Points) Consider the following linear system

$$\left\{\begin{array}{rcrcrcrcr} 3x_1 & - & 6x_2 & + & x_3 & + & 5x_4 & = & 5 \\ 2x_1 & - & 4x_2 & + & x_3 & + & 3x_4 & = & 4 \\ -x_1 & + & 2x_2 & - & 2x_3 & & & = & -5 \end{array}\right. .$$

(i) Find a matrix $A \in \mathbb{R}^{3\times 4}$ and and a vector $b \in \mathbb{R}^3$, such that the solutions of the above linear system are given by the vectors $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4$ satisfying $Ax = b$.

(ii) Determine the row-reduced echelon forms of the matrices $(A \mid b)$ and A and calculate their ranks.

(iii) Find all the solutions to the linear system.

(iv) Determine all $x \in \mathbb{R}^4$ which satisfy $Ax = b$ and which are orthogonal to the vector $u = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}$.

Exercise 2. (8 Points) Let $u = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \in \mathbb{R}^2$ and define the following three functions:

$$f_1 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} 2x_1 + 3x_2 \\ x_1 + (u \bullet u)x_3 \end{pmatrix},$$

$$f_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \sin(x_1) + \cos(x_2),$$

$$f_3 : \mathbb{R}^2 \longrightarrow \mathbb{R}^3, \quad x \longmapsto \begin{pmatrix} x \bullet x \\ 0 \\ u \bullet u \end{pmatrix}.$$

(i) Which of the above functions f_1, f_2, f_3 are linear maps? For each one that is linear, determine its matrix.

(ii) Is f_2 injective and/or surjective?

Exercise 3. (8 Points)

(i) Let $G : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map with

$$G\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad G\begin{pmatrix} -2 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \end{pmatrix}.$$

Determine the matrix of G.

(ii) Let $F : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear map with

$$F\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}, \quad F\begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 6 \end{pmatrix}.$$

Show that F is not injective.

Exercise 4. (8 Points) We define the following linear map

$$H : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} x_1 + x_2 - x_3 \\ x_1 + 2x_2 \\ x_2 + x_3 \end{pmatrix}.$$

(i) Calculate the image of H.

(ii) Decide if H is injective and/or surjective.

(iii) Find all vectors $x \in \mathbb{R}^3$ with $H(x) = 2x$.

14.4 Linear Algebra I - Midterm 2022

Exercise 1. (10 Points) Consider the following linear system

$$\left\{\begin{array}{rcrcrcrcl} x_1 & + & 3x_2 & & & + & x_4 & = & 1 \\ & & x_2 & + & 2x_3 & - & 2x_4 & = & 2 \\ 2x_1 & - & 2x_2 & + & x_3 & + & x_4 & = & 3 \end{array}\right. .$$

(i) Find a matrix $A \in \mathbb{R}^{3\times 4}$ and and a vector $b \in \mathbb{R}^3$, such that the solutions of the above linear system are given by the vectors $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4$ satisfying $Ax = b$.
(ii) Determine the row-reduced echelon forms of the matrices $(A \mid b)$ and A and calculate their ranks.
(iii) Find all the solutions to the linear system.
(iv) Determine all $x \in \mathbb{R}^4$ which satisfy $Ax = b$ and which have norm $\|x\| = \sqrt{14}$.

Exercise 2. (8 Points) Let $u = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in \mathbb{R}^2$ and define the following three functions:

$$\begin{aligned} f_1 : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &\longmapsto \begin{pmatrix} (u \bullet u) - 2 \\ x_1 + (u \bullet u)x_3 \end{pmatrix}, \end{aligned} \qquad \begin{aligned} f_2 : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto e^{x_1} - e^{x_2}, \end{aligned} \qquad \begin{aligned} f_3 : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto (x \bullet u)u \,. \end{aligned}$$

(i) Which of the above functions f_1, f_2, f_3 are linear maps? For each one that is linear, determine its matrix.
(ii) Is f_2 injective and/or surjective?

Exercise 3. (8 Points) Let $G : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map with

$$G\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}, \quad G\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} -4 \\ 4 \end{pmatrix}.$$

(i) Determine the matrix of G.
(ii) Find all vectors $x \in \mathbb{R}^2$ such that x is orthogonal to $G(x)$.

Exercise 4. (8 Points) We define the following linear map

$$\begin{aligned} H : \mathbb{R}^3 &\longrightarrow \mathbb{R}^4 \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &\longmapsto \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + 2x_2 + x_3 \\ 2x_1 - 2x_3 \end{pmatrix}. \end{aligned}$$

(i) Calculate the image of H.
(ii) Decide if H is injective and/or surjective.
(iii) Find a linear map $F : \mathbb{R}^2 \to \mathbb{R}^4$ with $\mathrm{im}(F) = \mathrm{im}(H)$.

14.5 Linear Algebra I - Finals 2019

Exercise 1. (12 Points) Let $A = \begin{pmatrix} 0 & 1 & -2 & 3 \\ 1 & -2 & 3 & -4 \\ -2 & 3 & -4 & 5 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}$.

(i) Compute the products AB and BA, or explain why they are not defined.
(ii) Determine whether or not the matrices A and B are invertible and, if they are, compute their inverses.
(iii) Calculate $\text{Im}(B)$ and $\ker(B)$.

Exercise 2. (14 Points) We define the subspace $U = \text{span}\{u_1, u_2, u_3\} \subset \mathbb{R}^3$, where

$$u_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

(i) Determine a basis $B = (b_1, \ldots, b_m)$ of U and calculate its dimension.
(ii) Calculate the coordinate vectors $[u_1]_B$, $[u_2]_B$ and $[u_3]_B$, where B is the basis you determined in i).
(iii) Determine a basis for $U^\perp$.
(iv) Find a linear map $G : \mathbb{R}^2 \to \mathbb{R}^3$ with $\ker(G) = \{0\}$ and $\text{Im}(G) = U$.

Exercise 3. (10 Points) Which of the following subsets of $\mathbb{R}^2$ are subspaces? Justify your answers.

(i) $U_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 + x_2 = x_1 x_2 \right\}$.
(ii) $U_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid 2x_1 = x_1 + x_2 \right\}$.
(iii) $U_3 = \text{span}\left\{ \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\} \bigcup \text{span}\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$.

(Friendly reminder: $\cup$ is the union of two sets)

Exercise 4. (14 Points) We define the following linear map

$$\begin{aligned} T : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} 1 & 2 \\ 0 & 3 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \end{aligned}$$

(i) Calculate an orthonormal basis $F = (f_1, \ldots, f_r)$ for $\text{Im}(T)$.
(ii) Check for which $t \in \mathbb{R}$ the vector $v = \begin{pmatrix} 1 \\ t \\ 1 \end{pmatrix}$ is an element in $\text{Im}(T)$. Determine the coordinate vector $[v]_F$ in this case.
(iii) Find a $w \in \mathbb{R}^3$ with $[w]_F = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.
(iv) Find a $x \in \mathbb{R}^2$ such that $\|T(x) - b\|$ is minimal, where $b = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$.

(In (ii) and (iii) the F is the basis of $\text{Im}(T)$ you calculated in (i)).

14.6 Linear Algebra I - Finals 2020

Exercise 1. (12 Points) Let $A = \begin{pmatrix} 1 & -1 & 0 \\ -2 & 3 & -4 \\ -2 & 3 & -3 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & -1 \\ -2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

(i) Compute the products AB and BA, or explain why they are not defined.
(ii) Determine whether or not the matrices A and B are invertible and, if they are, compute their inverses.
(iii) Find all $x \in \mathbb{R}^3$ with $A^T AA^T AA^T Ax = 0$. Justify you answer.

Exercise 2. (12 Points) We define the subspace $U = \operatorname{span}\{u_1, u_2, u_3, u_4\} \subset \mathbb{R}^3$, where

$$u_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \quad u_4 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

(i) Determine a basis $B = (b_1, \ldots, b_m)$ of U and calculate its dimension.
(ii) Calculate the coordinate vectors $[u_1]_B$, $[u_2]_B$,$[u_3]_B$ and $[u_4]_B$, where B is the basis from i).
(iii) Find a linear map $G : \mathbb{R}^3 \to \mathbb{R}^3$ with $\operatorname{Im}(G) = U$. What is the dimension of $\ker(G)$?

Exercise 3. (12 Points) Set $u = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$. Which of the following subsets of $\mathbb{R}^2$ are subspaces? Justify your answers.

(i) $U_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid 2x_1 - 3x_2 = x_1 \right\}$.
(ii) $U_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 \text{ is an integer, i.e. } x_1 \in \{\ldots, -2, -1, 0, 1, 2, \ldots\} \right\}$.
(iii) $U_3 = \{x \in \mathbb{R}^2 \mid x \notin \operatorname{span}\{u\}\}$.
(iv) $U_4 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \bullet u = x_1 \right\}$.

Exercise 4. (14 Points) We define the following linear map

$$\begin{aligned} H : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\longmapsto \begin{pmatrix} 1 & 1 \\ -2 & 4 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \end{aligned}$$

(i) Show that $\dim(\operatorname{Im}(H)) = 2$.
(ii) Calculate an orthonormal basis (f_1, f_2) for $\operatorname{Im}(H)$.
(iii) Find a vector $v \in \mathbb{R}^3$, such that $B = (f_1, f_2, v)$ is an orthonormal basis for $\mathbb{R}^3$.
(iv) Calculate $[H(x)]_B$ for any $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$.
(v) Find a $x \in \mathbb{R}^2$ such that $\|H(x) - b\|$ is minimal, where $b = \begin{pmatrix} -5 \\ 1 \\ -1 \end{pmatrix}$.

14.7 Linear Algebra I - Finals 2021

Exercise 1. (14 Points) Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -2 \\ 1 & 2 & 5 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & -2 \\ -1 & 2 & 5 \end{pmatrix}$.

(i) Determine whether or not the matrices A and B are invertible and, if they are, compute their inverses.
(ii) Calculate the matrix BA and decide if BA^n is invertible for any integer $n \geq 1$.
(iii) Determine if $\mathrm{Im}(A) \cup \mathrm{Im}(B)$ and $\mathrm{Im}(A) \cap \mathrm{Im}(B)$ are subspaces and, if they are, determine a basis.

Exercise 2. (12 Points) We define the subspace $U = \mathrm{span}\{u_1, u_2, u_3, u_4\} \subset \mathbb{R}^3$, where

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 3 \\ -4 \\ -1 \end{pmatrix}, \quad u_4 = \begin{pmatrix} 4 \\ -2 \\ 2 \end{pmatrix}.$$

(i) Determine a basis $B = (b_1, \ldots, b_m)$ of U and calculate its dimension.
(ii) Calculate the coordinate vectors $[u_1]_B$, $[u_2]_B$,$[u_3]_B$ and $[u_4]_B$, where B is the basis from i).
(iii) Calculate an orthonormal basis $F = (f_1, \ldots, f_m)$ for U and determine $[u_1]_F$, $[u_2]_F$,$[u_3]_F$ and $[u_4]_F$
(iv) Determine a basis for $U^\perp$.

Exercise 3. (12 Points) Set $u = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and let $C \in \mathbb{R}^{2\times 2}$ be an arbitrary matrix. Which of the following subsets of $\mathbb{R}^2$ are subspaces? Justify your answers.

(i) $U_1 = \left\{x \in \mathbb{R}^2 \mid x \bullet x = x \bullet u\right\}$.
(ii) $U_2 = \left\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid 2x_1 - x_2 = 3x_2\right\}$.
(iii) $U_3 = \left\{x \in \mathbb{R}^2 \mid Cx = x\right\}$.
(iv) $U_4 = \left\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 \cdot x_2 \geq x_1\right\}$.

Exercise 4. (12 Points) Assume we have the following data points

i	1	2	3
x_i	1	2	3
y_i	0	-1	-3

(i) Find the line of best fit for the above data, i.e. find $m, n \in \mathbb{R}$ such that the function $l(x) = mx + n$ minimizes the sum of squares $\sum_{i=1}^{3}(l(x_i) - y_i)^2$.
(ii) We define the following linear map

$$H : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$$
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

and set $V = \mathrm{Im}(H)$. Determine a basis of V and for $b = \begin{pmatrix} 0 \\ -1 \\ -3 \end{pmatrix}$ calculate the orthogonal projection $P_V(b)$. Use your result to show that b is not an element in V.

14.8 Linear Algebra I - Finals 2022

Exercise 1. (12 Points) Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 2 & -4 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$ and $C = A^T BA$.

(i) Determine whether or not the matrices A, B, and C are invertible and, if they are, compute their inverses.
(ii) Determine $\operatorname{im}(C)$, $\ker(C)$ and $\operatorname{im}(C) \cap \ker(C)$.
(iii) Give a basis for $\ker(C^n)$ for all $n \geq 1$.

Exercise 2. (14 Points) We define the subspace $U = \operatorname{span}\{u_1, u_2, u_3, u_4\} \subset \mathbb{R}^3$, where

$$u_1 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} -6 \\ -6 \\ -3 \end{pmatrix}, \quad u_3 = \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}, \quad u_4 = \begin{pmatrix} 6 \\ -2 \\ 0 \end{pmatrix}.$$

(i) Determine a basis $B = (b_1, \dots, b_m)$ of U and calculate its dimension.
(ii) Calculate the coordinate vectors $[u_1]_B$, $[u_2]_B$,$[u_3]_B$ and $[u_4]_B$, where B is the basis from i).
(iii) Find a linear map $F : \mathbb{R}^3 \to \mathbb{R}^3$ with $\ker(F) = U^\perp$ and determine a basis of $\operatorname{im}(F)$.

Exercise 3. (12 Points) Let $D \in \mathbb{R}^{2\times 2}$ be an arbitrary matrix. Which of the following sets are subspaces? Justify your answers.

(i) $U_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid 2x_1 - x_2 = -x_3 + x_2 \right\}$.
(ii) $U_2 = \left\{ x \in \mathbb{R}^{2023} \mid x \bullet x \geq -2023 \right\}$.
(iii) $U_3 = \left\{ x \in \mathbb{R}^2 \mid \text{ There exists a } y \in \mathbb{R}^2 \text{ with } Dy = 3x \right\}$.
(iv) $U_4 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 + x_2 = 0 \right\} \cup \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 \cdot x_2 \leq 0 \right\}$.

Exercise 4. (12 Points) We consider the vector $b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, the linear map

$$G : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$$
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 3 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

and define the subspace $U = \operatorname{im}(G)$.

(i) Show that $\dim(U) = 2$ and find an orthonormal basis $F = (f_1, f_2)$ of U.
(ii) Determine the orthogonal projection $y = P_U(b)$ of b onto U and calculate $[y]_F$.
(iii) Find a $x \in \mathbb{R}^2$ such that $\|G(x) - b\|$ is minimal.

Bibliography

[B] O. Bretscher: *Linear Algebra with Applications*, 4th edition, Pearson 2009.
[G] H. Goldstein, C. Poole, J. Safko: *Classical Mechanics*, 3th edition, Addison-Wesley 2001.