Linear Algebra, Set theory and Numbers (LALG 119) course notes

Bachelor in Computer Science and Engineering

20/21 B1 S1

Rix Silverith

Source code on GitHub

UNIVERSIDAD AUTONOMA DE MADRID

Contents

Chapter 1

Set theory and functions

The axiomatic formulation of set theory is pretty complicated (maybe because the concept of set is one of the most basics), therefore in this course we are only considering the intuitive idea of what a set is.

Definition 1.1 (Set). A set is a collection of objects about which is possible to determine weather or not a particular object is a member of the set.

Note. It's worth considering also the set with no elements, the **empty set**, which is denoted by $\emptyset$.

Sets are usually denoted by capital letters, and the objects within them are referred to as **elements**, which are denoted by lowercase letters.

A set can be described in two similar ways. On the one hand, the explicit way, by giving a list of their elements. One the other hand, the implicit way, by using the so called **set-builder notation**, which uses braces to enclose a property that is the qualification for membership in the set.

Example. Let A and B be two sets such that

$$A = \{x \mid x \text{ is a natural number and } x^2 - 1 = 0\} = \{-1, 1\} \tag{1.1}$$

$$B = \{x \mid x \text{ is an even natural number }\} = \{x : 2 \mid x \text{ and } x \text{ is natural }\} \tag{1.2}$$

Notation. In this previous example, both symbols $\mid$ and : are used to denote *such as*.

We write $a \in A$ if a is an element in the set A. Otherwise, we write $a \notin A$ to mean that a is not an element in the set A. For instance, $2 \in \mathbb{N}$ but $-2 \notin \mathbb{N}$.

Definition 1.2 (Empty set). We use $\emptyset$ to refer to the set with no elements, denominated the **empty set**.

Definition 1.3 (Subset). Let A and B be sets. We say that A is a **subset** of B if every element in A is an element in B. If A is a subset of B we write $A \subset B$ and we say *A is contained in B* or *B contains A*. Otherwise we write $A \not\subset B$.

Remark. Let X be any set. Then, the empty set is a subset of X, $\emptyset \subset X$.

Definition 1.4 (Equal sets). Two sets A and B are equal if $A \subset B$ and $B \subset A$; i.e. if they have the same elements.

Definition 1.5 (Properly contained sets). We say that a set A is **properly contained** in the set B if $A \subset B$ but $A \neq B$.

For example, $\mathbb{N}$ is properly contained in $\mathbb{Z}$ because $\mathbb{N} \subset \mathbb{Z}$, but $\mathbb{N} \neq \mathbb{Z}$.

Notation. We use $:=$ to mean *by definition*. For example, $\mathbb{N} := \{0, 1, 2 \ldots\}$.

Proposition 1.6. Let A and B be two sets. If $A \subset B$ and $B \subset A \implies A = B$.

> *Proof.* Definitions of $B \subset A$ and $A \subset B$ respectively indicate that $x \in B \implies x \in A$ and that $x \in A \implies x \in B$, thus $x \in A \iff x \in B$ and therefore $A = B$. ■

Definition 1.7 (Power set). Let A be a set. The power set of A is a set whose elements are all the subsets of A, and it's denoted by $\mathcal{P}(A)$.

Example. Let $A = \{a, b, c\}$, then the power set of A is

$$\mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}. \tag{1.3}$$

In this previous example, we must keep notice that $\emptyset \in \mathcal{P}(A)$ and $\emptyset \subset \mathcal{P}(A)$ are not exactly the same thing. In the first one we are refering to the $\emptyset$ element in the set $\mathcal{P}(A)$; while in the second, $\emptyset$ is the empty set, which as seen previously is a subset of any set. $\{\emptyset\} \subset \mathcal{P}(A)$ has also a different meaning, as $\emptyset$ in this case is the element contained in $\mathcal{P}(A)$. We couldn't write $\{\emptyset\} \in \mathcal{P}(A)$ because the element $\{\emptyset\}$ is not contained in $\mathcal{P}(A)$.

1.1 Unions and intersections of sets

Definition 1.8 (Union of sets). Let A and B be two sets. The union of A and B is another set whose elements are the elements in A and the elements in B. We denote this set by $A \cup B$.

$$A \cup B \stackrel{\text{def}}{=} \{x \mid x \in A \text{ or } x \in B\}. \tag{1.4}$$

Definition 1.9 (Intersection of sets). Let A and B be two sets. The intersection of A and B is another set whose elements are the elements that A and B have in common. We denote this set by $A \cap B$.

$$A \cap B \stackrel{\text{def}}{=} \{x \mid x \in A \text{ and } x \in B\} \tag{1.5}$$

Definition 1.10 (Disjoint sets). Let A and B be two sets. If $A \cap B = \emptyset$ it's said that A and B are disjoint.

Let A, B and C be three non empty sets. The following properties are hold related to the union and intersection operations between those sets.

Property name	**Properties**
Commutativity	$A \cap B = B \cap A$
Commutativity	$A \cup B = B \cup A$
Idempotency	$A \cup A = A$
Idempotency	$A \cap A = A$
Disjoint sets	$A \cap B = \emptyset$
Associativity	$(A \cap B) \cap C = A \cap (B \cap C)$
Associativity	$(A \cup B) \cup C = A \cup (B \cup C)$
Distributivity	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
Distributivity	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
Cancelation	$A \cup (B \cap A) = A$
Cancelation	$A \cap (B \cup A) = A$

Table 1.1: *Some properties of unions and intersections of sets.*

1.2 Universal and complementary sets

Sometimes it's convenient to assume that the sets that we are considering are subsets of a larger one, $\mathcal{U}$, denominated the *universal set*.

Remark. For any set A we have that $A \cap \mathcal{U} = A$ and $\mathcal{U} \cup A = \mathcal{U}$.

Definition 1.11 (Complement of a set). The complement of a set A is another set such that

$$A^C = A' \overset{\text{def}}{=} \{a \in \mathcal{U} \mid a \notin A\}. \tag{1.6}$$

Proposition 1.12 (De Morgan's laws). et A and B be two sets in $\mathcal{U}$. Then, the following equalities are hold.

$$1)\ (A \cup B)^C = A^C \cap B^C \qquad 2)\ (A \cap B)^C = A^C \cup B^C \tag{1.7}$$

Proof. In order to prove the De Morgan's laws we need first to proof that one side of the equation is contained in the other and viceversa. Only then we can say that the equality holds.

1) Let $x \in (A \cup B)^C \iff x \notin A \cup B \iff x \notin A$ and $x \notin B \iff x \in A^C$ and $x \in B^C \iff x \in A^C \cap B^C$. ■

$$\text{Let } a \in (A \cup B)^C \implies a \in \mathcal{U} \text{ and } a \notin A \cup B \implies \tag{1.8}$$
$$\implies a \in \mathcal{U} \text{ and } a \notin A \text{ and } a \notin B \implies \tag{1.9}$$
$$\implies a \in A^C \text{ and } a \in B^C \implies \tag{1.10}$$
$$\implies a \in A^C \cap B^C \tag{1.11}$$

So, we can afirm that $(A \cup B)^C \subset A^C \cap B^C$.

$$\text{Now, let } b \in A^C \cap B^C \implies b \in A^C \text{ and } b \in B^C \implies \tag{1.12}$$
$$\implies b \in \mathcal{U} \text{ but } a \notin A \text{ and } a \notin B \implies \tag{1.13}$$
$$\implies b \in \mathcal{U} \text{ and } b \notin A \cup B \implies \tag{1.14}$$
$$\implies b \in (A \cup B)^C \tag{1.15}$$

So, we can afirm that $A^C \cap B^C \subset (A \cup B)^C$. Therefore,

$$(A \cup B)^C = A^C \cap B^C \tag{1.16}$$

Proof. *(Second De Morgan's law)*. In order to proof the second De Morgan's law we need first to proof that $(A \cap B)^C \subset A^C \cup B^C$ and then $A^C \cup B^C \subset (A \cap B)^C$. Only then we can say $(A \cap B)^C = A^C \cup B^C$.

$$\text{Let } a \in (A \cap B)^C \implies a \in \mathcal{U} \text{ and } a \notin A \cap B \implies \tag{1.17}$$
$$\implies a \in \mathcal{U} \text{ and } a \notin A \text{ or } a \notin B \implies \tag{1.18}$$
$$\implies a \in A^C \text{ or } a \in B^C \implies \tag{1.19}$$
$$\implies a \in A^C \cup B^C \tag{1.20}$$

So, we can afirm that $(A \cap B)^C \subset A^C \cup B^C$.

$$\text{Now, let } b \in A^C \cup B^C \implies b \in A^C \text{ or } b \in B^C \implies \tag{1.21}$$
$$\implies b \in \mathcal{U} \text{ but } a \notin A \text{ or } a \notin B \implies \tag{1.22}$$
$$\implies b \in \mathcal{U} \text{ and } b \notin A \cap B \implies \tag{1.23}$$
$$\implies b \in (A \cap B)^C \tag{1.24}$$

So, we can afirm that $A^C \cup B^C \subset (A \cap B)^C$. Therefore,

$$(A \cap B)^C = A^C \cup B^C \tag{1.25}$$

■

1.3 Partitions of sets

Definition 1.13 (Partition). A partition of a non-empty set A is a separation of A into mutually disjoint non-empty subsets, A_α, such that $A_\alpha \neq A_\beta$ and $\cup A_\alpha = A$.

Remark. Note that if A is a finite set then to give a partition is equivalent to writting A as $A_1 \cup A_2 \cup \ldots \cup A_n$ with $A_i \neq \emptyset$ and disjoint two by two.

1.4 Other operations with sets

Definition 1.14 (Difference). Let A and B be two sets. The difference of A and B is another set, $A \backslash B$, whose elements are the elements in A which are not contained in B.

$$A \backslash B \stackrel{\text{def}}{=} \{a \in A \mid a \notin B\}. \tag{1.26}$$

Definition 1.15 (Symmetric difference). Let A and B be two sets. The symmetric difference of A and B is another set, $A \triangle B$, whose elements in A that are not contained in B and the elements in B that are not contained in A.

$$A \triangle B \stackrel{\text{def}}{=} \{a \in A \mid a \notin B\} \cup \{b \in B \mid b \notin A\}. \tag{1.27}$$

Remark. $A \triangle B = A \cup B \backslash A \cap B$.

Example. Let $A = \{1,2,3,4\}$ and $B = \{3,5,7\}$.

$$A \backslash B = \{1,2,4\}. \tag{1.28}$$

$$A \triangle B = \{1,2,4\} \cup \{5,7\} = \{1,2,4,5,7\}. \tag{1.29}$$

Definition 1.16 (Cartesian product). Let A and B be two sets. The cartesian product of A and B is the set of the ordered pairs of the form (a,b) where $a \in A$ and $b \in B$.

$$A \times B \stackrel{\text{def}}{=} \{(a,b) \mid a \in A,\ b \in B\}. \tag{1.30}$$

Example. Let $A = \{a,b,c\}$ and $B = \{c,3\}$.

$$A \times B = \{(a,c), (a,3), (b,c), (b,3), (c,c), (c,3)\}. \tag{1.31}$$

Note. In general, $B \times A \neq A \times B$.

Definition 1.17 (Cardinality). Let A and B be two sets. The cardinality of A, $\text{card}\,(A) = |A|$, is the number of elements in A.

$$\text{If card}\,(A) < \infty \text{ and card}\,(B) < \infty \Longrightarrow \text{card}\,(A \times B) = \text{card}\,(A) \cdot \text{card}\,(B)\,. \tag{1.32}$$

1.5 Definition of function and related concepts

Definition 1.18 (Function). A **function**, or **map**, from a non-empty set X to a non-empty set Y is a subset f of $X \times Y$ such that $\forall x \in X$ that appears as part of a pair in f, there is one, and only one $y \in Y$ such that $(x, y) \in f$, and we write $f : X \longrightarrow Y$.

In other terms, a function between two sets X and Y is just a way to assign an element of X an element of Y.

Definition 1.19. If $f : X \longrightarrow Y$ and $C \subset X$, $D \subset Y$, the following sets are defined.

$$f\,(C) = \{y \in Y \mid y = f\,(x) \text{ for some } x \in C\} \tag{1.33}$$

$$f^{-1}\,(D) = \{x \in X \mid f\,(x) \in D\}. \tag{1.34}$$

Definition 1.20. Let $f : X \longrightarrow Y$ and let $g : X \longrightarrow Y$ be two functions. We say that $f = g$ if $\forall x \in X,\ f\,(x) = g\,(x)$.

Definition 1.21 (Domain). If $f : X \longrightarrow Y$, the set X is the **domain** of the function, which contains all the input values of the function.

Remark. Since a function is defined on its entire domain, its domain coincides with its domain of definition.

Definition 1.22 (Codomain). If $f : X \longrightarrow Y$, the set Y is the **codomain** of the function, which contains all of the output values of the function.

Definition 1.23 (Image). If $f : X \longrightarrow Y$, the **image** (or **range**) of f is the subset range$\subset Y$ such that

$$\text{range} := \{y \in Y \mid \exists,\ x \in X \text{ with } f\,(x) = y\}. \tag{1.35}$$

Notation. The range of a function f is often denoted by $f\,(A)$ or by Imf, which stands for *image of* f.

Definition 1.24 (Image of an element). If $x \in X$, then the image of x under f, denoted $f\,(x)$, is the value of f when applied to x.

Note. $f(x)$ is alternatively known as the output of f for argument x.

Definition 1.25 (Image of a subset). The image of a subset $S \subset X$ under f, denoted $f(S)$, is the subset of Y such that

$$f(S) \stackrel{\text{def}}{=} \{f(s) \mid x \in S\}. \tag{1.36}$$

Definition 1.26 (Image of a function). The **image** of a function is the image of its entire domain, also known as the range of the function.

1.6 Surjective, injective and bijective functions

Definition 1.27 (Surjective function). A function f from a set X to a set Y is **surjective** (also known as **onto**), if for every element y in the codomain Y of f, there is at least one element x in the domain X of f such that $f(x) = y$. In other words, a surjective function is a function whose image is equal to its codomain. Symbolically,

$$\text{If } f : X \to Y\text{, then } f \text{ is said to be surjective if } \forall\, y \in Y,\ \exists\, x \in X \mid f(x) = y. \tag{1.37}$$

Remark. It's not required that x be unique; the function f may map one or more elements of X to the same element of Y.

Note. The French word *sur* means *over* or *above*, and relates to the fact that the image of the domain of a surjective function completely covers the function's codomain.

Notation. If $f : X \longrightarrow Y$ is such that $f(x) = y$ we say that y is the image of x by f and we'll say that x is the preimage of y by f.

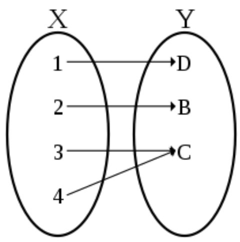

Figure 1.1: *A surjective function from domain X to codomain Y.*

Definition 1.28 (Injective function). Let f be a function whose domain is a set X. The function f is said to be **injective** (or **one-to-one**), provided that for all a and b in X, whenever $f(a) = f(b)$, then $a = b$. Symbolically,

$$\text{If } f : X \longrightarrow Y\text{, then } f \text{ is injective if } \forall\, a, b \in X,\ f(a) = f(b) \Longrightarrow a = b. \tag{1.38}$$

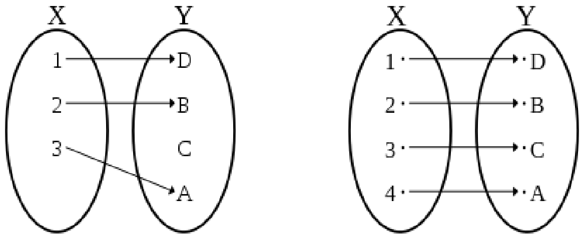

Figure 1.2: *Injective non-surjective function (left) / Injective surjective function (right).*

Definition 1.29 (Bijective function). Let f be a function from a set X to a set Y. The function f is said to be **bijective** if it's both surjective and injective. In other words, each element of one set is paired with exactly one element of the other set, and each element of the other set is paired with exactly one element of the first set.

Remark. If X and Y are finite sets, then the existence of a bijection means they have the same number of elements.

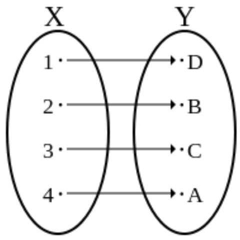

Figure 1.3: *A bijective function, $f : X \longrightarrow Y$.*

A bijective function from the set X to the set Y has an **inverse function** from Y to X.

Following these definitions, a function $f : X \longrightarrow Y$ is surjective if and only if all the elements of X are map to an element of Y, and it's injective if two different elements of X are always applied on two different elements of Y; thus if we want to prove that certain f is not surjective we should find an element of Y that is not in the image of f, and if we want to prove that is not injective we should find two different elements of X whose images by f are the same.

1.7 Composition of functions

Definition 1.30 (Composite function). Let $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$, then we define the **composition** of g and f as the function $g \circ f : X \longrightarrow Z$ such that $(g \circ f)(x) = g(f(x))$.

Definition 1.31 (Identity function). The function $f : X \longrightarrow X$ that leaves all of the elements invariant, this is $f(x) = x,\ \forall x \in X$, is called **identity function**, and is usually denoted by Id_X.

Definition 1.32 (Inverse function). Given $f : X \longrightarrow Y$, it's said that the function $g : Y \longrightarrow X$ is the **inverse** of f, denoted by $g = f^{-1}$, if $g \circ f = \mathrm{Id}_X$ and $f \circ g = \mathrm{Id}_Y$.

Note that in the previous definition it's not enough to prove one condition for the other to automatically hold. For instance, if $f : X \longrightarrow \mathbb{R}$ and $g : \mathbb{R} \longrightarrow X$ where $X = \{x \in \mathbb{R} \mid x \geqslant 0\}$ are given by $g(x) = x^2$ and $f(x) = +\sqrt{x}$, $(g \circ f)(x) = \left(+\sqrt{x}\right)^2 = x$ holds $\forall x \in X$. However, $(f \circ g)(x) = x,\ \forall x \in X$ is not hold as negative numbers don't hold $+\sqrt{x^2} = x$.

Intuitively, the inverse of a function from X to Y is simply considering it on the opposite way, from Y to X. Thsi requires every element of Y to have exactly one preimage; i.e. the function must be injective.

Proposition 1.33. A function $f : X \longrightarrow Y$ is invertible if and only if f is bijective. In addition, $\left(f^{-1}\right)^{-1} = f$.

Proposition 1.34. If a function f is invertible, then it's inverse, f^{-1}, is unique, which means that there's exactly one function f^{-1} satisfying this property.

Example. Let $f : \mathbb{Q} \longrightarrow \mathbb{Q}$, $f(x) = 2x$ be a bijective function. To compute it's inverse suppose $x = f^{-1}(y)$, then $f\left(f^{-1}(y)\right) = 2f^{-1}(y)$, and therefore $f^{-1}(y) = \frac{y}{2}$.

Remark. It's easy to notice that the process to compute the inverse of $y = f(x)$ is reduced to solving for y in $x = f(y)$, getting $y = f^{-1}(x)$.

Chapter 2

Equivalence relations

In mathematics sometimes there are situations in which it is convenient to establish relations among the elements of a set. Defining a relation on a set means we have a way to compare the elements in that set. Under certain conditions (equivalence relations) this will one to subdivide the elements of a set into different groups that share similar properties. [Cha02]

Definition 2.1 (Relation). A relation on a set A is a non-empty subset, $\mathcal{R}$, of $A \times A$. If $x, y \in A$ holds that the ordered pair $(a, b) \in \mathcal{R}$, it is said that a **is related to** b, usually written $a\mathcal{R}b$.

So technically any subset of $A \times A$ is a relation on A. However, it should posses certain characteristics in order to be kind of interesting. Relations are generally used to compare two elements in some way. That is, we use them to determine whether two elements are *related* in the manner specified.

Example. Let the relation $x\mathcal{R}y \iff xy$ with $x, y \in \mathbb{N}$. Then, that relation is defined by the following set:

$$\mathcal{R} = \{n, m : n \mid m\}. \tag{2.1}$$

Remark. For any set A, both $\emptyset$ and $A \times A$ are relations on A. The $\emptyset$ relation doesn't relate elements to anything, not even themselves. On the other hand, the relation $A \times A$ relates every element to every element of A. Actually, this two relations are pretty useless, but worth mentioning.

Definition 2.2. Let $\mathcal{R}$ be a relation defined on a set A. Then, $\mathcal{R}$ is

- **Reflexive,** if $\forall x \in A$, $x\mathcal{R}x$.
- **Symmetric,** if $\forall x, y \in A$, $x\mathcal{R}y \implies y\mathcal{R}x$.
- **Antisymmetric,** if $\forall x, y \in A$, $x\mathcal{R}y$ and $y\mathcal{R}x \implies x = y$.
- **Transitive,** if $\forall x, y, z \in A$, $x\mathcal{R}y$ and $y\mathcal{R}z \implies x\mathcal{R}z$.

2.1 Equivalence and order relations

Taking Definition 2.2 into account, is convenient to point out two main types of relations.

Definition 2.3 (Equivalence relations). An *equivalence relation*, denoted by $\sim$, is a relation that holds the **reflexive**, **symmetric** and **transitive** properties.

Definition 2.4 (Order relation). An *order relation* is relation that holds the **reflexive**, **antisymmetric** and **transitive** properties. Moreover, if $\forall x, y$ is hold $x\mathcal{R}y$ or $y\mathcal{R}x$, then it is said that it is a **total order relation**. Otherwise, it is said that it is of **partial order**.

Example. Prove or disprove that $x\mathcal{R}y \iff x \leqslant y$ with $x, y \in \mathbb{N}$ is an equivalence relation.

So, in order to check if this is an equivalence relation we should first check if it satisfies the reflexive, symmetric and transitive properties.

(a) $x\mathcal{R}x \iff x \leqslant x,\ \forall x \in \mathbb{N}$, then the relation $\mathcal{R}$ is reflexive.

(b) If $x\mathcal{R}y$ then $y\mathcal{R}x$? No. Just take $2\mathcal{R}4 \iff 2 \leqslant 4$ but $4\mathcal{R}2 \iff 4 \nleqslant 2$. Then, the relation is not symmetric, and therefore it is not an equivalence relation.

(c) If $x\mathcal{R}y$ and $y\mathcal{R}z \implies x\mathcal{R}z$?

$$\left.\begin{array}{rcl} x\mathcal{R}y & \implies & x \leqslant y \\ y\mathcal{R}z & \implies & y \leqslant z \end{array}\right\} \implies x \leqslant z \iff x\mathcal{R}z \implies \mathcal{R} \text{ is transitive.} \tag{2.2}$$

Because the reflexive and the transitive conditions are met, but not the symmetric, the $\mathcal{R}$ relation is not an equivalence relation. ■

2.2 Equivalence classes and the quotient set

Sometimes when working with elements in a set, it is convenient to consider that some of them are *equal*, even though they are not. To declare two elements as equal we define a relation in the set and, to make sure that we don't get something ilogical, we need this relation to be an equivalence relation.

Definition 2.5 (Equivalence class). Given an equivalence relation $\sim$ on a set A and an element $a \in A$, the **equivalence class** of a is the set

$$[a] = \overline{a} \overset{\text{def}}{=} \{x \in Aa \sim x\} \tag{2.3}$$

In other words, we can say that the equivalence class of the element a of a set A is the set of all elements x in the set A which are related to the element a. We should also point out that the equivalence classes correspoding to non-related elements are disjoint and non-empty. Therefore, they define a partition of A.

Definition 2.6 (Quotient set). The quotient set of a set A by the equivalence relation $\sim$, denoted by $A/\sim$, is the set of all the equivalence classes.

Example. Let $\sim$ be an equivalence relation defined on the set $A = \{1, 2, 3, 5, 6, 9\}$ such that $n \sim m \iff 3 \mid n - m$ (3 divides $n - m$), $n, m \in \mathbb{Z}$. Then, we have

$$[1] = \{1\} \qquad [2] = [5] = \{2, 5\} \qquad [3] = [6] = [9] = \{3, 6, 9\}. \tag{2.4}$$

Therefore, we can write the quotient set as $A/\sim = \{[1], [2], [3]\}$.

Chapter 3

The $\mathbb{Z}$ and $\mathbb{Z}_n$ rings

3.1 Elemental algebraic structures

In this section we are defining some structures that frequently appear in mathematics. Although they can seen unnecessary generalizations, they are pretty useful for us to not to prove the same result in different contexts.

Definition 3.1 (Operation). Let $A \subset \mathcal{U}$, an operation in A is a function from $A \times A$ to $\mathcal{U}$. When its image is in A it is said that it is an **internal composition law**, or that is **closed**.

Definition 3.2 (Group). A *group*, G, is a set in which it is defined a closed operation, let's denote it by $*$, such that the **associative property**,

$$g * (h * f) = (g * h) * f, \tag{3.1}$$

is satisfied, there exists an **identity element**,

$$\exists\, e \in G \forall g \in G, e * g = g * e = g, \tag{3.2}$$

and there exists an **inverse element**,

$$\forall g \in G, \exists\, h \in G h * g = g * h = e. \left(h = g^{-1}\right). \tag{3.3}$$

Moreover, if $*$ is a **commutative operation** $(g * h = h * g)$ it is said that G is an **abelian**, or **commutative group**.

Definition 3.3 (Ring). A *ring*, X, is a set in which there are defined two closed operations, $\oplus$ and $\otimes$ (addition and product), that satisfy the following properties:

- X is an abelian group with respect to $\oplus$.
- $\otimes$ is an associative operation on X.
- The distributive laws are hold (from the left) $(a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$ and (from the right) $c \otimes (a \otimes b) = (c \otimes a) \oplus (c \otimes b)$.

Note. If $\otimes$ is commutative, it is said that X is a *commutative ring*, and if $\otimes$ has an identity element (multiplicative identity) it is said that the ring is *unitary*, or *with identity*.

Definition 3.4 (Field). A *field*, F, is a commutative ring with identity such that $F - \{0\}$ is an abelian group with respect to the product, $\otimes$.

In a not so formal way, an abelian group is a set in which we can add or substract (add the inverse), while in a commutative ring we can also multiply; and in a field, divide (except for 0); i.e. every element in a field has a multiplicative inverse.

When it is not clear what operations are being considered in a set, they are usually indicated explicitly next to set, all inside parenthesis. Thus, for instance $(\mathbb{Z}, +)$, which is an abelian group is the set of integers with the addition operation.

Example. $(\mathbb{Z}, \cdot)$ and $(\mathbb{N}, +)$ are not abelian groups because, for example, 3 does not have an inverse in any of the two sets.

Example. $(\mathbb{N}, +, \cdot)$ is not a ring. $(\mathbb{N}, +)$ is a binary closed operation that satisfies the associative property, has 0 as an indentity element, but it doesn't have an additive inverse for each $n \in \mathbb{N}$. Therefore, $(\mathbb{N}, +)$ does not form a group.

However, $(\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$ and $(\mathbb{C}, +, \cdot)$ are all commutative rings with identity.

Example. $(\mathbb{R}, +, \cdot)$ and $(\mathbb{C}, +, \cdot)$ are examples of fields.

3.2 Greatest common divisor. Euclid's algorithm

Theorem 3.1 (Division algorithm on $\mathbb{Z}$). *Let $a, b \in \mathbb{Z}$ with $b > 0$. Then, there exists unique integers q, r (called quotient and remainder) such that*

$$a = bq + r \quad \text{with } 0 \leqslant r < |b|\,. \tag{3.4}$$

Proof. Let r be least positive value that $a - bq$ takes when $q \in \mathbb{Z}$, then $r < |b|$ ya que si $r \geqslant |b|$, aumentando o disminuyendo (si b es negativo) q en una unidad obtendríamos un valor de r menor, el cociente y el resto son únicos porque $a = q_1 + r_1, a = bq_2 + r_2$ implica $b\,(q_2 - q_1) = r_1 - r_2$ lo que contradice que $0 \leqslant r_1, r_2 < |b|$ excepto en el caso trivial $q_2 - q_1 = r_1 - r_2 = 0$. ∎

If in theorem 3.1 $r = 0$, then b divides a.

Definition 3.5 (Divisibility). Given two integers a and b with $b \neq 0$, we say that b *divides* a, written $b \mid a$, if there is some $q \in \mathbb{Z}$ such that $a = bq$.

$$b \mid a \Longleftrightarrow \exists\, qa = bq. \tag{3.5}$$

Moreover, we also say that b is a *divisor* of a, or that a is a *multiple* of b. Otherwise, if b does not divide a we write $b \nmid a$.

Remark. 1 and -1 divide all the integers. In other words, all integers are multiple of both 1 and -1 (including 0). If 0 divides an integer b that means by definition that $b = c \cdot 0$ for some $c \in \mathbb{Z}$, and that would imply that $b = 0$. Therefore, 0 only divides 0. Now, 0 is a multiple of all the integers as if $a \in \mathbb{Z} \implies a \mid 0$ because $0 = 0 \cdot a$.

Now let's take a look to some properties on divisibility.

Proposition 3.6. If an integer c divides another two integers a and b, then c also divides any linear combination of a and b. In other words,

$$\text{if } c \mid a \text{ and } c \mid b \implies c \mid \alpha a + \beta b, \quad \forall \alpha, \beta \in \mathbb{Z}. \tag{3.6}$$

Proof. Since $c \mid a \implies \exists\, c' \in \mathbb{Z} a = cc'$, and since $c \mid b \implies \exists\, c'' \in \mathbb{Z} b = cc''$. If we consider $\alpha, \beta \in \mathbb{Z}, \alpha a + \beta b = \alpha\,(cc') + \beta\,(cc'') = c\,(\alpha c' + \beta c'') \implies c \mid \alpha a + \beta b$. ∎

Proposition 3.7. Let $a, b, c \in \mathbb{Z}$. If $a \mid b$ and $b \mid c$ then $a \mid c$.

Proof. If $a \mid b \implies b = a\alpha, \alpha \in \mathbb{Z}$ and if $b \mid c \implies c = b\beta, \beta \in \mathbb{Z}$, then $c = b\beta = a\alpha\beta \implies a \mid c$. ∎

Proposition 3.8. Let $a, b \in \mathbb{Z}$. If $a \mid b$ and $b \mid a \implies |a| = |b|$.

Proof. Since $a \mid b \implies b = a\alpha$ for some $\alpha \in \mathbb{Z}$, and since $b \mid a \implies a = b\beta$ for some $\beta \in \mathbb{Z}$. Then, $b = a\alpha = (b\beta)\,\alpha = b\,(\alpha\beta) \implies b = b\,(\beta\alpha) \implies b - b\,(\beta\alpha) = 0 \implies b\,(1 - (\beta\alpha)) = 0$.

- If $b = 0 \implies a = 0$.
- If $b \neq 0 \implies 1 - \beta\alpha = 0 \implies 1 = \beta\alpha \implies \beta = \alpha = 1$ or $\beta = \alpha = -1$.
- If $\beta = \alpha = 1 \implies a = b$.
- If $\alpha = \beta = -1 \implies a = -b$.

∎

The most remarkable consequence of (3.5) is that we can define the greatest common divisor and that there's a method to compute it called *Euclid's algorithm*. For this reason, when a ring holds an analogue property to Theorem (3.1) it's said that is an Euclidian domain.

Definition 3.9 (Greatest common divisor). An integer d is said to be the greatest common divisor of two integers a and b if $d > 0$, d divides both numbers and if any other integer c divides a and $b \implies c \mid d$.

Proposition 3.10 (Bézout's identity). If $d = \gcd(a, b)$, then exists $n, m \in \mathbb{Z}$ such that $d = an + bm$. In fact, all the solutions $x, y \in \mathbb{Z}$ to the equation $d = ax + by$ are of the form

$$\begin{cases} x = n - bt/d \\ y = m + at/d \end{cases} \qquad \text{with } t \in \mathbb{Z}. \tag{3.7}$$

Theorem 3.2 (Well ordering principle). *Any non-empty subset of $\mathbb{N}$ has a minimum on $\mathbb{N}$.*

Now, how do we find the $\gcd(a, b)$? We can use the following theorem.

Theorem 3.3. *Let $a, b \in \mathbb{Z}$ where at least one of them is non-zero. Then, exists a greatest common divisor d of a and b. Moreover, d can be written as $d = a\alpha + b\beta$ for some $\alpha, \beta \in \mathbb{Z}$. In fact, d is the smallest positive number that can be written as a linear combination of a and b.*

Proof. If we have two numbers a and b, the theorem says that we can find $\gcd(a, b)$ by writting all the possible linear combinations of a and b.

- If $b = 0$ then $a \neq 0 \implies d = \gcd(a, b) = |a|$
- If $a, b \neq 0$, consider the set $S := \{a\alpha + b\beta \alpha, \beta \in \mathbb{Z}\}$. Consider now $S' := \{a\alpha + b\beta \alpha, \beta \in \mathbb{Z} : a\alpha + b\beta > 0\} \subseteq S$.

  Claim: $S' \neq \emptyset$. By construction, $a, b \in S$ if $a > 0 \implies a \in S'$, otherwise $-a > 0$ and $-a \in S'$. Then, $S' \subseteq \mathbb{N}, S' \neq \emptyset$. Making use of the well ordered principle we can ensure that S' has a minimum.

  Suppose we call $d = \min S'$. In particular, there exists $\alpha, \beta \in \mathbb{Z} d = a\alpha + b\beta$ (by definition, d is the smallest positive number that can be written in this form).

  Claim: $d = \gcd(a, b)$. By construction we have that $d > 0$. Now, we divide a by b using the division algorithm on $\mathbb{Z} \implies \exists\, q, r \in \mathbb{Z} a = qd + r$ where $0 \leqslant r < d$. From this we can solve for r.

  $$r = qd - a = q(a\alpha + b\beta) - a = b\beta + a(q\alpha - 1) \tag{3.8}$$

  Therefore, r can be written as a linear combination of a and b. If $r > 0 \implies r \in S' \implies d \leqslant r$. Here we reach a contradiction unless $r = 0$.

  In the same way, using the same argument we get that $d \mid b$.

  Now let $c \in \mathbb{Z} : c \mid a$. Since $c \mid a \implies a = ca'$ for some $a' \in \mathbb{Z}$, and since $c \mid b \implies b = cb'$ for some $b' \in \mathbb{Z}$, we know that $d = a\alpha + b\beta = (ca')\alpha + (cb')\beta = c(a'\alpha + b'\beta) = d \implies c \mid d$. Therefore, $d = \gcd(a, b)$.

■

Example. Let $a = 2$ and $b = 3$, then $\gcd(2,3) = 1$. The previous theorem basically tells us that we can write the $\gcd(2,3)$ as the linear combination $1 = (-1)\,2 + 3 \cdot 1$. Because of the properties previously seen we know that the linear combination $2\alpha + 3\beta$ is always divisible by 2.

In order to find the linear expression $d = an + bm$ for given a, b we use the euclidean algorithm.

Proposition 3.11. If $c \mid a$ and $c \mid b \implies c \mid \gcd(a,b)$. In other words, $\gcd(a,b)$ is a multiple of any other common divisor of a and b.

Proof. Let $d = \gcd(a,b)$. Suppose that $c \mid a, b$. Because $d = \gcd(a,b) \implies \exists\ \alpha, \beta \in \mathbb{Z} d = \alpha a + \beta b$. Since $c \mid a \implies a = a'c$ for some $a' \in \mathbb{Z}$, and since $c \mid b \implies b = b'c$ for some $b' \in \mathbb{Z}$, then $d = \alpha a + \beta b = \alpha a' c + \beta b' c = c\,(\alpha a' + \beta b')$. ■

Remark. This is just a property of the $\gcd(a,b)$.

Example. Let $a = 12$ and $b = 30$, then $\gcd(a,b) = 6 \implies$ any common divisor of 12 and 30 divides 6. Common divisors of 12 and 30 are 1, 2, 3 and 6. In this case, $1, 2, 3 \mid 6$ but, for instance, 2 does not divide 3.

Proposition 3.12. If $\gcd(a,b) = d$ and if we write $a = da', b = db'$ for some $a', b' \in \mathbb{Z}$, then $\gcd(a', b') = 1$.

Proof. Since $d = \gcd(a,b) \implies d = a\alpha + b\beta = (a'd)\,\alpha + (b'd)\,\beta = (a'\alpha + b'\beta)\,d$. Now, dividing $d = (a'\alpha + b'\beta)\,d$ by d we have $1 = a'\alpha + b'\beta \implies \gcd(a', b') = 1$. ■

Proposition 3.13. Let $a, b, n \in \mathbb{Z}$. Then, $\gcd(na, nb) = n \gcd(a,b)$.

Proof. Let $d = \gcd(a,b)$. Since $d \mid a, b \implies nd \mid na$ and $nd \mid nb \implies nd$ is a common divisor of na and nb, but is $nd = \gcd(na, nb)$?

Again, since $d = \gcd(a,b) \implies d = a\alpha + b\beta$ for some $\alpha, \beta \in \mathbb{Z}$. If we multiply by n we have $nd = na\alpha + nb\beta = \gcd(na, nb)$. ■

Remark. Two numbers a and b can have many common divisors, but among them the only one that can be written as a linear combination of a and b is the $\gcd(a,b)$.

Theorem 3.4 (Euclid's theorem). *Let $a, b \in \mathbb{Z}$. If $a \mid bc$ and $\gcd(a, b) = 1 \implies a \mid c$.*

Proof. Since $\gcd(a, b, c) = 1 \implies \gcd(ca, cb) = c \implies$ since $a \mid ca$ and $a \mid bc \implies a \mid c$. ■

Example. Let $a = 3, b = 4, c = 6$. Then $3 \mid bc = 24$. Since $a = 3$ does not divide $b = 4$, then $a = 3 \mid c = 6$.

3.3 Prime and coprime numbers. Factorization theorem

Definition 3.14 (Coprimes). Let $a, b \in \mathbb{Z}$If $\gcd(a, b) = 1$ we say that a and b are **relatively prime** or **coprime**.

Definition 3.15 (Prime number). An integer p is said to be **prime** if $p > 0$ and the only divisors of p are ± 1 and $\pm p$.

Proposition 3.16. Every positive integer greater than 1 is divisible by a prime number.

Proof. Let $a \in \mathbb{Z}_{>1}$.

- If a is a prime number then the proof is done.
- If a is not a prime number, then there exists by definition a $c \in \mathbb{Z}_{>1} < a$ such that $c \mid a$.

  Let $S := \{c \in \mathbb{Z}_{>1} : c \mid a \text{ and } c < a\}, S \subseteq \mathbb{N}, S \neq \emptyset$. By the wll order principle we know that there is a minimum element in S. So, let $d := \min\{cc \in S\}$. Then, $d > 1, d < a$ and $d \mid a$. If d is the smallest integer in S satisfying these three properties it has to be a prime number.

  If d was not a prime $\implies \exists b > 1, b < d, b \mid d$, but since $d \mid a \implies b \mid a \implies b \in S$, and this is a contradiction. Therefore, d is prime and $d \mid a$.

■

Theorem 3.5 (Euclid's theorem). *There are infinitely many primes.*

Proof. Suppose, to get a contradiction, that there is a finite number of primes. Let $a := 1 + p_1 + \ldots + p_n \in \mathbb{Z}_{>1}$. By the claim, a is divisible by some prime. Hence, $\exists\, i \in \{1 \ldots n\} : p_i \mid a \implies p_i \mid (1 + p_1 + \ldots + p_n) \implies$

$1 = a - p_1 \cdot \ldots \cdot p_n \Longrightarrow p_i \mid 1$, which is a contradiction. ■

Theorem 3.6 (Fundamental theorem of arithmetic). *Every positive integer $n \in \mathbb{Z}_{>1}$ can be written as a product of prime numbers, and this factorization is unique except for the order of the factors.*

Proof. To prove this theorem will are using mathematical induction. Then, the first case that makes sense for the statement of the theorem is $n = 2$, which is a prime. That means this is already the factorization in product of primes.

Induction hypothesis: Assume that the theorem holds for positive integers $< n$, for $n > 2$. Using the induction hypothesis we'll prove that the theorem also holds for n ($n > 2$). Now,

- If n is a prime there's nothing to prove.
- Suppose n is not a prime. We have proven that any integer greater than 1 is always divisible by a prime. Therefore, there is some prime p that divides n: $n = pn_1$, with $p > 1$. We know that $n_1 < n$. We can apply the induction hypothesis to n_1:

$$n_1 = p_1 \ldots p_s \Longrightarrow n = p \cdot p_1 \ldots p_s \tag{3.9}$$

is a factorization of n into primes.

So we have proven that any positive number $n > 1$ can be written as a product of primes. Now we want to show that this factorization is unique. We will prove this by contradiction.

Suppose that for a given $n > 1$ we have two factorizations into primes:

$$n = p_1 \cdot p_2 \ldots p_s = q_1 \cdot q_2 \ldots q_\ell \tag{3.10}$$

where p_i, q_j are primes from $i = \ell \ldots s, j = 1 \ldots \ell$. Since p_1 is a prime and since $p_1 \mid n \Longrightarrow p_1 \mid q_1 \cdot \ldots \cdot q_\ell \Longrightarrow p_1 \mid q_j$ for some $j \in \{1 \ldots \ell\}$ (Whenever a prime divides a product it has to divide one of the factors). So we can cancel p_1 and q_j in the expression for n and repeat to conclude that $s = \ell$ and $\{p_1 \ldots p_s\} = \{q_1 \ldots q_\ell\}$.

■

3.4 Linear Diophantine equations

Definition 3.17 (Diophantine equation). A *Diophantine equation* is a polynomial equation whose solutions are restricted to integers.

Definition 3.18 (Linear Diophantine equation). A *linear Diophantine equation* is a first-degree Diophantine equation.

These type of equations are named after the ancient Greek mathematician Diophantus, and they are important when a problem requires a solution in whole amounts.

The study of problems that require integer solutions is often referred to as *Diophantine analysis*. Although the practical applications of Diophantine analysis have been somewhat limited in the past, this kind of analysis has become much more important in the digital age, as it is very important in the study of public-key cryptography, for example.

Proposition 3.19. A Diophantine equation of the form $ax + by = c$ with $a, b, c \in \mathbb{Z}$ has integer solutions if and only if $\gcd(a, b) \mid c$. In other words, there's no integer solutions if $\gcd(a, b)$ does not divide c.

Remark. The previous proposition is just a consequence of the Bézout's identity.

Example. For instance, the Diophantine equation $6x + 20y = 7$ has no integer solutions as $\gcd(6, 20) = 2$ but 2 does not divide 7.

3.4.1 Initial solution to a linear Diophantine equation

Finding solutions to linear Diophantine equations involves finding an initial solution, and then altering that solution is some way to find the remaining solutions. When finding this initial solution is important to recognize first if the equation we are dealing with has or not solutions in $\mathbb{Z}$. As it is stated in proposition 3.19, one can determine if solutions exist or not by computing the GCD of the coefficients of the variables, and then determining if the constant term can be divided by that GCD.

If solutions do exist, then there is an efficient method to find an initial solution. The extended version of the Euclidean algorithm seen previously will give us both the GCD of the coefficients and an initial solution.

So, given an equation $ax + by = n$, we will use the Euclidean algorithm to compute $\gcd(a, b) = d$ and determine if there are any solutions. The extended version of this algorithm consists on solving the equations used to compute the GCD for the remainders, and using substitution, go through the steps of the Euclidean algorithm to find a solution to the equation $ax_i + by_i = d$. Then, the initial solution to the equation $ax + by = n$ is the ordered pair

$$\left(x_i \cdot \frac{n}{d}, y_i \cdot \frac{n}{d} \right). \tag{3.11}$$

Example. **Find a solution to the Diophantine equation** $6x + 10y = 20$**.**

Since $\gcd(6, 10) = 2$ and 2 divides 20 $\implies$ the equation $6x + 10y = 20$ has a solution in $\mathbb{Z}$. To make the problem easier we can simplify the equation by dividing by the $\gcd(6, 10)$, which yields

$$3x + 5y = 10, \tag{3.12}$$

which has the exact same solutions as the initial equation. Now, let the equation

$$3x_i + 5y_i = 1. \tag{3.13}$$

Since $\gcd(3, 5) = 1$, all solutions to (3.13) multiplied by 10 are solutions to (3.12). In this case, there is no need to use the Euclidean algorithm as it is pretty easy to see that $x_i = 2$ and $y_i = -1$ verifies (3.13) and therefore, the ordered pair $(20, -10)$ is an initial solution to the equation $6x + 5y = 20$.

3.4.2 General solution to linear Diophantine equations

3.5 Congruences

If we fix an integer $n \in \mathbb{Z}$ then we can define an equivalence relation $\sim$ on $\mathbb{Z}$ by saying that $x \sim y \iff x - y$ is a multiple of n; i.e. if $n \mid x - y$.

Definition 3.20 (Congruence). Let $n \in \mathbb{N}$, $a, b \in \mathbb{Z}$ and let $\sim_n$ be an equivalence relation on $\mathbb{Z}$ such that $a \sim_n b \iff n \mid a - b$. If $a \sim_n b$ it's said that a and b are *congruent modulo n*, denoted by $a \equiv b \mod n$.

Lemma 3.21. Two numbers are congruent modulo n if and only if when divided by n they leave the same remainder.

Proof. Suppose

$$a = nq_1 + r_1 \qquad b = nq_2 + r_2 \tag{3.14}$$

Then, $a \equiv b \mod n \implies (a - nq_1) - (b - nq_2) \implies n \mid r_1 - r_2$, and as $0 < r_1, r_2 < |n|$, this implies $r_1 = r_2$. ■

Definition 3.22. For $x \in \mathbb{Z}$ it is defined the equivalence class of x with respect to the congruece equivalence relation $\equiv \mod n$ by

$$[x] \stackrel{\text{def}}{=} \{a \in \mathbb{Z} a \equiv x \mod n\} \tag{3.15}$$

Example. Take $n = 3$ and $x = 0, 1, 2$. So this yields the equivalence classes

$$[0] = \{a \in \mathbb{Z} a \equiv 0 \mod 3\} = \{0, \pm 3, \pm 6, \dots\} \tag{3.16}$$
$$[1] = \{a \in \mathbb{Z} a \equiv 1 \mod 3\} = \{\dots, -5, -2, 1, 4, 7, 10, \dots\} \tag{3.17}$$
$$[2] = \{a \in \mathbb{Z} a \equiv 2 \mod 3\} = \{\dots, -4, -1, 2, 8, 11, \dots\} \tag{3.18}$$

From the previous example we work out that $\equiv \mod n$ divides $\mathbb{Z}$ in n equivalence classes or partitions that correspond to the n possible values of the remainder ($r = 0, 1, 2, \dots, |n| - 1$).

Definition 3.23. Fixed n, the set of least residues is given by $\{0, 1, \dots, n-1\}$.

Therefore, for all $a \in \mathbb{Z}$, a is congruent to exactly one of the least residues modulo n.

Proof. Use the division algorithm with a and n. This led us to

$$a = nq + r \quad \text{with } 0 \leqslant r \leqslant n - 1. \tag{3.19}$$

From this it follows that $a - r = nq \implies a \equiv r \mod n$. ■

Proposition 3.24. If $a \equiv b \mod n$ and $c \equiv d \mod n$ then $a + c \equiv (b + d) \mod n$ and $ac \equiv bd \mod n$.

Proposition 3.25. If $a \equiv b \mod m$ and $n \mid m$ then $a \equiv b \mod n$.

Definition 3.26. We say that $\mathbb{Z}_n$ is the quotient set of the relation of congruence modulo n by $\mathbb{Z}$.

Proposition 3.27. If we define in $\mathbb{Z}_n$ the sum and multiplication operations

$$[a] + [b] = [a + b] \quad \text{and} \quad [a] \cdot [b] = [a \cdot b] \tag{3.20}$$

then, in general, $(\mathbb{Z}, +, \cdot)$ is a commutative ring, as both operations have same properties they have on $\mathbb{Z}$.

Proposition 3.28. $(\mathbb{Z}_n, +, \cdot)$ is a field $\iff$ n is prime.

Proof. If n is not prime, $n = ab$ with $0 < |a|, |b| < |m|$, then $[a], [b] \neq [0]$ and $[a] \cdot [b] = [m] = [0]$. Therefore, $[b]$ can't have a multiplicative inverse because

$$[b] \cdot [c] = [1] \implies [a] \cdot [b] \cdot [c] = [0] \tag{3.21}$$

and this contradicts our hypothesis about a. Then, $\mathbb{Z}_n$ is not a field.

On the other hand, if n is prime, then any a with $1 \leqslant a < p$ holds that $(a, p) = \pm 1$, and this implies that exists some integers x amd y such that $ax + py = 1$. Once we know this integers we have

$$[ax + py] = [1] \implies [ax] + [py] = [1] \implies [ax] = [1]. \tag{3.22}$$

Therefore, the class $[a]$ has an inverse (which is $[x]$). ∎

Proposition 3.29. In general, an equivalence class $[a]$ has a multiplicative inverse in $\mathbb{Z}_n \iff a$ and n are coprimes; i.e. $\gcd(a, n) = 1$.

Proof. Since $\gcd(a, n) = 1$, and applying the Bezout's identity (3.10), we have that $\exists m, \ell \in \mathbb{Z} am + b\ell = 1 \iff am \equiv 1 \mod n$. ∎

Notation. Usually we write $\mathbb{Z}_n^*$ to design the set of equivalence classes of $\mathbb{Z}_n$ that have a multiplicative inverse.

Proposition (3.29) has several consequences:

- Let p be a prime number. Then, we can define the set

$$\mathbb{Z}_p^* := \{[1], \ldots, [p-1]\} \tag{3.23}$$

  in which all elements but zero have a multiplicative inverse; i.e. they are units. Therefore, by (3.28) $(\mathbb{Z}_p, +, \cdot)$ is always a field.

- Let $n \in \mathbb{Z}_{>0}$, which is not a prime number. Then, we can define the set

$$\mathbb{Z}_n^* := \{[a] \in \mathbb{Z}_n \gcd(a, n) = 1\} \tag{3.24}$$

  which contains the equivalence classes that have a multiplicative inverse in $\mathbb{Z}_n$. Now, we will use the *Euler's φ function* to represent the cardinality of (3.24).

$$\varphi(n) := |\mathbb{Z}_n^*| = (\mathbb{Z}_n^*) \tag{3.25}$$

  A question arise here. What are the values of $\varphi(n)$? If $n = p$, where p is a prime, we have that $\varphi(n) = \left|\mathbb{Z}_p^*\right| = p - 1$.

  However, if $n = p^k$, where p is prime and $k \in \mathbb{Z}_{\geqslant 1}$, to find a value of $\varphi(p^k)$ we should write a list of the elements in $\mathbb{Z}_{p^k}$ and remove the classes that come from multiples of p. In this case, there are p^{k-1} elements that come from multiples of p. Thus, we get $\varphi(p^k) = p^k - p^{k-1}$.

  This result let us know how many elements in a ring with $n = p$ have a multiplicative inverse. Together with the following lemma we can determine, for instance, $|\mathbb{Z}_{12}^*|$.

Lemma 3.30. Suppose n and m in $\mathbb{Z}$ are coprime. Then $\varphi(nm) = \varphi(n) \cdot \varphi(m)$.

Example. To compute $|\mathbb{Z}_{12}^*|$ we know that $12 = 3 \cdot 4$ and $\gcd(3,4) = 1$. Then, by lemma 3.30 we get

$$\varphi(12) = \varphi(3) \cdot \varphi(4) = 2 \cdot \left(2^2 - 2\right) = 4. \tag{3.26}$$

Proposition 3.31 (Cancelation law). For a, b, $c \in \mathbb{Z}$ we have $ca \equiv cb \mod n \iff a \equiv b \mod \frac{n}{\gcd(n,c)}$.

Proof. Let us prove $\implies$ first. Suppose $ca \equiv cb \mod n \implies n \mid c(a-b) \implies c(a-b) = nk, k \in \mathbb{Z}$. Let $d = \gcd(c,n)$. Note that $\frac{c}{d}(a-b) = \frac{n}{d}k$, then $\gcd\left(\frac{c}{d}, \frac{n}{d}\right) = 1 \implies a - b = \frac{n}{d}k', k' \in \mathbb{Z}$. Finally, $a \equiv b \mod \frac{n}{d}$. The proof for the other direction is pretty easy. ∎

Theorem 3.7 (Euler's theorem). *Let $a, n \in \mathbb{Z}$. If $\gcd(a,n) = 1$, then $[a]^{\varphi(n)} \equiv [1] \mod n$.*

Proof. To be done. ∎

3.5.1 Linear congruences

Theorem 3.8. *Consider an equation of the form $[a][x] \equiv [b] \mod n$. Then,*

- *if $\gcd(a,n) = d \nmid b \implies$ the equation has no solutions.*
- *if $\gcd(a,n) = d \mid b \implies$ the equation has d different solutions in $\mathbb{Z}_n^*$.*

Proof. The equation $[a][x] \equiv [b] \mod n$ has a solution if and only if $ax + ny = b$ has integer solutions, and this happens when $\gcd(a,n) \mid b$. This proves the first statement of the theorem.

Now, let's prove the second statement. If $\gcd(a,n) = d \mid b$, then to solve $[a][x] \equiv [b] \mod n$ in $\mathbb{Z}_n$ we look for the integer solutions to $ax + ny = b$, which are of the form

$$\begin{cases} x = x_0 + \frac{n}{d}\ell \\ y = y_0 - \frac{a}{d}\ell \end{cases} \quad \text{with } \ell \in \mathbb{Z}, \tag{3.27}$$

where (x_0, y_0) is a particular solution. This gives us that the solutions to $[a][x] \equiv [b] \mod n$ are of the form

$$[x] = [x_0] + \frac{[n]}{d}[\ell] \mod n. \tag{3.28}$$

Taking $\ell = 0, 1, 2, \ldots, d-1$ gives us d different solutions in $\mathbb{Z}_n$. ■

Chapter 4

Matrices and linear systems

4.1 Definition of a matrix

Definition 4.1 (Matrix). A *matrix* is a collection of numbers ordered in rows and columns.

$$\begin{bmatrix} 2 & 2 & 1 \\ -1 & 0 & 4 \end{bmatrix} \tag{4.1}$$

is an example of a 2×3 matrix.

4.2 Operations with matrices

4.3 Determinants

4.4 Solving linear systems of equations using matrices

Chapter 5

Vector spaces

Definition 5.1 (Vector space). A vector space over a field F is a set V together with two operations, $+ : V \times V \longrightarrow V$ and $\cdot : F \times V \longrightarrow V$, such that $(V, +)$ is an abelian group, and that for any $\mathbf{v_1}$, $\mathbf{v_2} \in V$, α_1, $\alpha_2 \in F$ the following properties are verified:

$$1 \cdot \mathbf{v_1} = \mathbf{v_1}, \qquad (\alpha_1 + \alpha_2) \cdot \mathbf{v_1} = \alpha_1 \mathbf{v_1} + \alpha_2 \mathbf{v_1},$$
$$\alpha_1 \cdot (\alpha_2 \cdot \mathbf{v_1}) = (\alpha_1 \alpha_2) \cdot \mathbf{v_1}, \qquad \alpha_1 \cdot (\mathbf{v_1} + \mathbf{v_2}) = \alpha_1 \mathbf{v_1} + \alpha_1 \mathbf{v_2}.$$

Notation. Elements in F are usually called *scalars* to differentiate them from the ones in V, that are *vectors*.

Example. $\mathbb{R}^n$, with $n \in \mathbb{Z}^+$, is a vector space over $\mathbb{R}$ defined by

$$\mathbb{R}^n \stackrel{\text{def}}{=} \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n \text{ times}} = \{(\alpha_1, \ldots, \alpha_n) \mid \alpha_i \in \mathbb{R}\} \tag{5.1}$$

with addition and scalar multiplication

$$(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) \tag{5.2}$$
$$\lambda (\alpha_1, \ldots, \alpha_n) = (\lambda \alpha_1, \ldots, \lambda \alpha_n). \tag{5.3}$$

This is the main vector space in which we will work in this topic.

From any field F, a new vector space can be defined over F in the same manner as the previous example.

Example. $m \times n$ matrices with coefficients in a field F, form the vector space

$$\mathcal{M}_{m \times n}(F) = \left\{ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \mid a_{ij} \in F,\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n \right\}. \tag{5.4}$$

Example. The solutions $\mathbf{x} = (x_1, \ldots, x_n)$ of the homogeneous system $A\mathbf{x} = B$ with $A = (a_{ij}) \in F$ form a vector space over F.

5.1 Vector subspaces

Definition 5.2 (Vector subspace). Let V be a vector space over a field F. A nonempty subset $W \subset V$ defines a subspace of $V \iff \forall \lambda, \mu \in F$ and $\mathbf{w_1}, \mathbf{w_2} \in W \Longrightarrow \lambda \mathbf{w_1} + \mu \mathbf{w_2} \in W$.

To check if some subset W of a vector space V is a subspace of V it's enough to see if W accomplish definition 5.2. However, in many cases is easier to check that the following properties are verified:

$$\mathbf{w_1}, \mathbf{w_2} \in W \implies \mathbf{w_1} + \mathbf{w_2} \in W, \tag{5.5}$$

$$\mathbf{w} \in W,\ \lambda \in F \implies \lambda \mathbf{w} \in W. \tag{5.6}$$

Example. $W = \{(x, y, 0) \in \mathbb{R}^3 \mid x, y \in \mathbb{R}\}$ is a vector subspace of $\mathbb{R}^3$, while $U = \{(x, y, 1) \in \mathbb{R}^3 \mid x, y \in \mathbb{R}\}$ is not. Note that, for instance, $(1, 1, 1) \in U$ but $2 \cdot (1, 1, 1) \notin U$, which contradicts property 5.5.

Example. $W = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0, 2x - 3y - z = 0\}$ is a vector subspace of $\mathbb{R}^3$, but also of $V = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$.

In conclusion, whenever we set linear and homogeneous conditions in $\mathbb{R}^n$, a subspace is obtained. This is closely related to the description of a vector space as the solutions of an homogeneous system.

Definition 5.3 (Linear combination). Let $\mathbf{v_1}, \ldots, \mathbf{v_n} \in V$ be vectors within a vector space V over some field F. A linear combination of $\mathbf{v_1}, \ldots, \mathbf{v_n}$ is an expression of the form $\alpha_1 \mathbf{v_1} + \ldots + \alpha_n \mathbf{v_n}$, where $\alpha_1, \ldots, \alpha_n \in F$.

Example. In $\mathbb{R}^2$ every vector $\mathbf{v} = (x, y)$ is a linear combination of $\mathbf{v_1} = (1, 0)$ and $\mathbf{v_2} = (0, 1)$ since $\mathbf{v} = x\mathbf{v_1} + y\mathbf{v_2}$.

Proposition 5.4. In a vector space, the identity element for vector addition, $\mathbf{0} = (0, \ldots, 0)$, is a linear combination of any vector $\mathbf{v}$, since $\mathbf{0} = 0 \cdot \mathbf{v}$.

Definition 5.5 (Subspace generated by a set). Let C be a subset of a vector space. The vector subspace generated by C, denoted by $\langle C \rangle$ or $\mathcal{L}(C)$, is the set of all the possible vector linear combinations of C.

Proposition 5.6. If $C \subseteq V$, $\langle C \rangle$ is a vector subspace of V.

Example. In $\mathbb{R}^2$, $\langle (1, 0), (0, 1) \rangle = \{\alpha_1 (1, 0) + \alpha_2 (0, 1)\} = \mathbb{R}^2$, since we have already seen in a previous example that every vector of $\mathbb{R}^2$ is a linear combination of $(1, 0)$ and $(0, 1)$.

Example. Let $C = \{(1, 1, 2), (0, 2, -1), (1, 3, 1) \subset \mathbb{R}^3\}$, then

$$\langle C \rangle = \{\lambda (1, 1, 2) + \mu (0, 2, -1) + \nu (1, 3, 1)\} \tag{5.7}$$

$$= \{(\lambda + \nu, \lambda + 2\mu + 3\nu, 2\lambda - \mu + \nu) \mid \lambda, \mu, \nu \in \mathbb{R}\}, \tag{5.8}$$

since $(1, 3, 1)$ is a linear combination of $(1, 1, 2)$ and $(0, 2, -1)$, more concisely

$$(1, 3, 1) = (1, 1, 2) + (0, 2, -1), \tag{5.9}$$

the vector $(1,3,1)$ can be omitted; i.e., $\langle C\rangle = \langle (1,1,2),(0,2,-1)\rangle$.

The previous example suggest defining the concept of some dependence between vectors in a set.

Definition 5.7 (Linear independence). A set of vectors $\{\mathbf{v_1},\ldots,\mathbf{v_n}\}$ from a vector space V is **linearly independent** $\iff$ the equation $\alpha_1\mathbf{v_1}+\ldots+\alpha_n\mathbf{v_n}=\mathbf{0}$ can only be satisfied by $\alpha_i = 0$ for $i=1,\ldots,n$. Otherwise, the set of vectors is **linearly dependent**.

Remark. Definition 5.7 implies that no vector in a set of linear independent vectors can be written as a linear combination of the remaining vectors in the set. In other words, even more concisely, a set of vectors is linear independent $\iff$ $\mathbf{0}$ can be represented as a linear combination of its vectors in a unique way.

Note. Many times it's said that several vectors are linearly independent (or dependent), meaning that they form a linearly independent (or dependent) set, in this case we take as finite subset the one formed by themselves.

Example. Vectors $(1,1,0)$, $(2,1,1)$, $(5,3,2)\in\mathbb{R}^3$ are not linearly independent. Writing the linear combination

$$\alpha_1\,(1,1,0)+\alpha_2\,(2,1,1)+\alpha_3\,(5,3,2)=(0,0,0) \tag{5.10}$$

we reach the system

$$\left.\begin{aligned}\alpha_1+2\alpha_2+5\alpha_3&=0\\ \alpha_1+\alpha_2+3\alpha_3&=0\\ \alpha_2+2\alpha_3&=0\end{aligned}\right\}\quad\Longrightarrow\quad \alpha_2=-2\alpha_3,\ \alpha_1=-\alpha_3 \tag{5.11}$$

and since solutions to this system depends on a parameter there are infinite solutions. Therefore, the vectors are linearly dependent.

Proposition 5.8. Any set of vectors containing the null vector, $\mathbf{0}=(0,\ldots,0)$ is always linearly dependent.

Proposition 5.9. In general, a non null vector $\mathbf{v}\in\mathbb{R}^n$ is linearly independent.

Proof. The only solution to the equation $\alpha\mathbf{v}=\mathbf{0}$, $\alpha\in\mathbb{R}$ is $\alpha=0$. Suppose there is another solution with $\alpha\neq 0$. This implies $\exists\alpha^{-1}\in\mathbb{R}$, which multiplied in both sides of the equation yields

$$\alpha\mathbf{v}=\mathbf{0}\quad\iff\quad\alpha^{-1}\alpha\mathbf{v}=\alpha^{-1}\mathbf{0}\quad\iff\quad\mathbf{v}=\mathbf{0}, \tag{5.12}$$

and this is a contradiction. Then, if $\mathbf{v}\neq\mathbf{0}$, vector $\mathbf{v}$ is always linearly independent. ■

5.2 Basis and dimension

Taking about vector spaces it's covenient considering subsets that generate all the space and that have the minimum possible number of vectors as, in some way, the *size* of these sets determines the size of the vector space.

Definition 5.10 (Basis). A set B of vectors is a basis of a vector space $V \iff B$ is both linearly independent and a system of generators, this is, the set B generates the vector space; i.e. $\langle B \rangle = V$.

Remark. For a given vector space there are multiple choices of basis, but all of them have the same cardinality.

Example. **Prove that** $B = \{(1,1), (1,-1)\}$ **is a basis of** $\mathbb{R}^2$**.**

1. To see if B is system of generators one should study if
$$(x,y) = \lambda(1,1) + \mu(1,-1) \tag{5.13}$$
always has solutions λ, μ, for any $(x,y) \in \mathbb{R}^2$. This lead us to
$$\left.\begin{aligned} \lambda + \mu = x \\ \lambda - \mu = y \end{aligned}\right\} \iff \lambda = \frac{x+y}{2},\ \mu = \frac{x-y}{2}. \tag{5.14}$$
Since there is always a solution, B is a system of generators; i.e. B generates $\mathbb{R}^2$.
2. For checking if B is linearly independent we consider
$$(0,0) = \lambda(1,1) + \mu(1,-1), \tag{5.15}$$
that can only be solved with $\lambda = \mu = 0$.

Example. Is $B = \{1, \sin x, x\} \subset \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ a basis of $\langle B \rangle$? Note that, by definition, B is a system of generators, therefore it's only necessary to prove that it is linearly independent. This is equivalent to prove that if λ, μ, ν verifies
$$\lambda + \mu \sin x + \nu x = 0 \tag{5.16}$$
for all x, then $\lambda = \mu = \nu = 0$. Differentiating several times we obtain
$$\lambda + \mu \sin x + \nu x = 0, \tag{5.17}$$
$$\mu \cos x + \mu = 0, \tag{5.18}$$
$$-\mu \sin x = 0. \tag{5.19}$$
The last equation implies $\mu = 0$ and from the others we get $\lambda = \nu = 0$.

Definition 5.11 (Dimension). A vector space has a **finite dimension** if it has a basis with a finite number of elements, and it corresponds to the cardinality of any of its basis. Otherwise, it has **infinite dimension**.

Remark. If S is finite, then $V = \langle S \rangle \implies V$ has finite dimension.

Example. $B = \{(1,0,0,\ldots,0), (0,1,0,\ldots,0), (0,0,1,\ldots,0), \ldots, (0,0,0,\ldots,1)\}$ is a basis of $\mathbb{R}^n$, usually called the canonical basis.

Example. The space of all real functions, $\mathcal{F} = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ has infinite dimension.

Let's see now how to compute the basis of a vector space.

Example. **Find a basis for** $V = \{(x,y,z) \in \mathbb{R}^3 \mid x+y+z=0\}$**.**

In this kind of problems it is better to express the conditions that define the subspace in terms of some variable that can take arbitrary values. In this case, note that

$$\mathbf{v} \in V \iff \mathbf{v} = (-y-z, y, z) \tag{5.20}$$

where y and z can take any real value. Therefore,

$$\begin{aligned} \mathbf{v} \in V &\implies \mathbf{v} = y\,(-1,1,0) + z\,(-1,0,1) && (5.21)\\ &\implies \mathbf{v} \in \langle (-1,1,0), (-1,0,1) \rangle && (5.22) \end{aligned}$$

which proves that $B = \{(-1,1,0), (-1,0,1)\}$ is a system of generators. It is easy to see that B is also linearly independent, then, B is a basis of V.

Theorem 5.1 (Steinitz's theorem). *Let $B = \{\mathbf{u_1}, \ldots, \mathbf{u_n}\}$ a basis of a vector space V, and let $\mathbf{v_1}, \ldots, \mathbf{v_m}$ be m linearly independent vectors with $m \leqslant n$, then there exists m vectors of B that can be replaced by $\mathbf{v_1}, \ldots, \mathbf{v_m}$, obtaining a new basis.*

Example. Consider the canonical basis of $\mathbb{R}^3$ $B = \{(1,0,0), (0,1,0), (0,0,1)\}$ and the set of linearly independent vectors $S = \{(0,1,0), (1,2,1)\}$. Following theorem 5.1 we have $n = 3$, $m = 2$ and vectors

$$\mathbf{u_1} = (1,0,0), \qquad \mathbf{u_2} = (0,1,0), \qquad \mathbf{u_3} = (0,0,1) \tag{5.23}$$

$$\mathbf{v_1} = (0,1,0), \qquad \mathbf{v_2} = (1,2,1). \tag{5.24}$$

Note that $\mathbf{u_2}$ and $\mathbf{u_3}$ can be replaced by $\mathbf{v_1}$ and $\mathbf{v_2}$ obtaining the basis of $\mathbb{R}^3$ given by $\hat{B} = \{(1,0,0), (0,1,0), (1,2,1)\}$, but if we replace $\mathbf{u_1}$ and $\mathbf{u_3}$ by $\mathbf{v_1}$ and $\mathbf{v_2}$ we don't get a new basis.

Corollary 5.12. In a vector space of finite dimension all basis have the same number of vectors.

Proof. If B and B' are basis with $m = |B| < |B'| = n$, then, following Steinitz's theorem 5.1 we could get a new basis B'' such that $B''B$ with $B'' \neq B$, but since $\langle B \rangle$ is the whole space, this would imply that vectors in $B'' - B$ linearly dependent on the ones in B and therefore B'' is not linearly independent. ∎

Definition 5.13 (Dimension). The dimension of a given vector space with finite dimension corresponds to the cardinality of any of its basis.

Note. Usually, this definition for dimension is completed by stating that the trivial vector space $V = \{\mathbf{0}\}$ has dimension zero.

The following corollary prevents us from checking if a possible basis is a system of generators if we know beforehand the dimension of the vector space.

Corollary 5.14. In a vector space of dimension n, whatever n linearly independent vectors form a basis.

Example. Prove that $B = \{(1, 2, 1), (1, 2, 0), (0, 1, 1)\}$ is a basis of $\mathbb{R}^3$.

Using the previous corollary is enough to see that the vectors in B are linearly independent, since $\dim \mathbb{R}^3 = 3$.

Knowing that any vector can be written as a linear combination of the elements of a basis, it is worth defining the numbers that appear as coefficients, which characterises the vector.

Definition 5.15 (Coordinates). Let V be a vector space of finite dimension and $B = \{\mathbf{u_1}, \ldots, \mathbf{u_n}\}$ one of its basis. Then, the coordinates of certain vector $\mathbf{v} \in V$ in basis B are $(\lambda_1, \ldots, \lambda_n)$, if $\mathbf{v} = \lambda_1 \mathbf{u_1} + \ldots + \lambda_n \mathbf{u_n}$.

Note. Many times it is written $\mathbf{v} = (\lambda_1, \ldots, \lambda_n)$. When using this notation it must be clear which is the basis we are considering. Also, it is worth noting that coordinates depend on the order of the elements which form the basis. Therefore, more consisely, coordinates are not only associated to the basis, but also to the way its elements are ordered.

5.3 Operations with vector subspaces

Proposition 5.16. If W, Z are two vector subspaces of a vector space V, the sum of W and Z, denoted by $W + Z$, is the smallest vector subspace of V that contains both W and Z. In other words,

$$W + Z \overset{\text{def}}{=} \{\mathbf{x} \mid \mathbf{x} = \mathbf{w} + \mathbf{z},\ \mathbf{w} \in W,\ \mathbf{z} \in Z\}. \tag{5.25}$$

Proposition 5.17. If $W, Z \subseteq V$ are vector subspaces of a vector space V, the intersection of both subspaces, denoted by $W \cap Z$, is

$$W \cap Z \stackrel{\text{def}}{=} \{x \in \mathbb{R} \mid x \in W \wedge x \in Z\}. \tag{5.26}$$

Remark. Note that the two subspace are never disjoint since the zero vector $\mathbf{0}$ is within every subspace.

Lemma 5.18. If W and Z are subspaces of a vector space V, then $W + Z$ and $W \cap Z$ are also vector spaces.

Proof. To prove that the intersection $W \cap Z$ is a vector subspace of $\mathbb{R}^n$, we check the following subspace criteria:

1. The subspace $W \cap Z \neq \emptyset$; i.e. the zero vector $\mathbf{0}$ of $\mathbb{R}^n$ is in $W \cap Z$.

   As W and Z are subspaces of $\mathbb{R}^n$, the zero vector $\mathbf{0}$ is in both W and Z. Therefore, $\mathbf{0} \in W \cap Z \implies W \cap Z \neq \emptyset$.

2. For all $\mathbf{x}, \mathbf{y} \in W \cap Z \implies \mathbf{x} + \mathbf{y} \in W \cap Z$.

   Suppose $\mathbf{x}, \mathbf{y} \in W \cap Z$. Since $\mathbf{x}, \mathbf{y} \in W \cap Z \implies \mathbf{x}, \mathbf{y} \in W$ and $\mathbf{x}, \mathbf{y} \in Z$. Hence both Z and W are vector subspaces it follows that $\mathbf{x} + \mathbf{y} \in W$ and $\mathbf{x} + \mathbf{y} \in Z \implies \mathbf{x} + \mathbf{y} \in W \cap Z$.

3. For all $\mathbf{x} \in W \cap Z, \alpha \in \mathbb{R} \implies \alpha \mathbf{x} W \cap Z$.

   Since $\mathbf{x} \in W \cap Z \implies \mathbf{x} \in W$ and $\mathbf{x} \in Z$, and since W and Z are vector subspaces, $\alpha \mathbf{x} \in W$ and $\alpha \mathbf{x} \in Z \implies \alpha \mathbf{x} \in W \cap Z$.

■

Proposition 5.19 (Grassman's formula). If W and Z are two subspaces of a vector space with finite dimension, then

$$\dim (W + Z) = \dim W + \dim Z - \dim (W \cap Z). \tag{5.27}$$

Chapter 6

Linear transformations

One of the main goals of linear algebra is the characterization of the solutions to the set of m linear equations in n unknowns $x_1, \ldots, x_n$. Linear transformations, also called linear maps, and their properties give us a lot of insight into the characteristics of the solutions to a system of linear equations.

Definition 6.1 (Linear transformation). Let V and W be two vector spaces over the same field F. A function $T : V \longrightarrow W$ is a *linear transformation*, or *linear map*, if it preserves addition and scalar multiplication between the two vector spaces; in other words, if

$$T(\mathbf{v}+\mathbf{u}) = T(\mathbf{v}) + T(\mathbf{u}) \qquad \forall\, \mathbf{v},\mathbf{u} \in V, \tag{6.1}$$

$$T(\lambda\mathbf{v}) = \lambda T(\mathbf{v}) \qquad \forall \lambda \in F,\ \mathbf{v} \in V. \tag{6.2}$$

Equations (6.1) and (6.2) can be summarized into

$$T(\lambda\mathbf{v}+\mu\mathbf{u}) = \lambda T(\mathbf{v}) + \mu T(\mathbf{u}) \qquad \forall\, \lambda,\mu \in F,\ \mathbf{v},\mathbf{u} \in V, \tag{6.3}$$

which, for any vectors $\mathbf{v_1}, \ldots, \mathbf{v_n} \in V$ and scalars $\lambda_1, \ldots, \lambda_n$, can be written in a more general as

$$T\left(\sum_{i=1}^{n} \lambda_i \mathbf{v_i}\right) = \sum_{i=1}^{n} \lambda_i T(\mathbf{v_i}). \tag{6.4}$$

Thus, a linear transformation is said to be *operation preserving*. That is, it doesn't matter whether the linear transformation is applied before or after the operations of addition and scalar multiplication, and therefore, linear combinations are preserved between the two vector spaces.

Definition 6.2. A linear transformation $T : V \longrightarrow W$ is

- a *monomorphism* if it is injective; i.e. $T(\mathbf{v}) = T(\mathbf{u}) \Longleftrightarrow \mathbf{v} = \mathbf{u}$.
- an *epimorphism* if it is surjective; i.e. $\forall\, \mathbf{w} \in W, \exists\, \mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{u}$.
- an *isomorphism* if it is bijective; i.e. $\forall\, \mathbf{w} \in W, \exists\,!\ \mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{u}$.

Example. For all vector space V the indentity transformation

$$\mathrm{Id}_V : V \longrightarrow V, \qquad \mathrm{Id}_V(\mathbf{v}) = \mathbf{v}, \quad \forall \mathbf{v} \in V \tag{6.5}$$

is linear. More generally, if V is a vector space over a field F and we take a fixed scalar $\lambda \in F$, the transformation

$$T_\lambda : V \longrightarrow V \qquad \mathbf{v} \longmapsto T_\lambda(\mathbf{v}) = \lambda\mathbf{v} \tag{6.6}$$

is linear. In particular, the transformation that maps any vector of V to the vector $\mathbf{0} \in W$ is linear, and is called *zero map*. We call *homothety* in V any linear transformation from V to V of the form $T_\lambda(\mathbf{v}) = \lambda\mathbf{v}$, where λ is the *ratio* of the homothety.

Example. If A is a real $m \times n$ matrix, then A defines a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ by sending the column vector $\mathbf{x} \in \mathbb{R}^n$ to the column vector $A\mathbf{x} \in \mathbb{R}^m$.

Example. Differentiation defines a linear map from the space of all differentiable functions to the space of all functions.

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{i=1}^{n} c_i f_i(x)\right) = \sum_{i=1}^{n} c_i \frac{\mathrm{d}f_i(x)}{\mathrm{d}x}. \tag{6.7}$$

6.1 Representation in terms of matrices

Linear transformations are mostly commonly written in terms of matrix multiplication. If V and W are two finite-dimensional vector spaces and a basis is defined for each of them, then every linear transformation from V to W can be represented by a matrix. This fact is useful to perform concrete computations.

Proposition 6.3. Fixed basis B_V and B_W for two finite-dimension vector spaces V and W, respectively, over a field F, any linear transformation $T : V \longrightarrow W$ can be written as $T(\mathbf{v}) = A\mathbf{v}$, where $A \in \mathcal{M}_{m\times n}(F)$ with $m = \dim W$ and $n = \dim V$. Columns of A are vectors $T(\mathbf{v_i})$ where vectors $\mathbf{v_i} \in B_V$.

Proof. Let $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ be a basis for V. Then every vector $\mathbf{v} \in V$ is uniquely determined by the coefficients $\lambda_1, \ldots, \lambda_n$ in the field F. If $T : V \longrightarrow W$ is a linear transformation,

$$T\left(\sum_{i=1}^{n} \lambda_i \mathbf{v_i}\right) = \sum_{i=1}^{n} \lambda_i T(\mathbf{v_i}), \tag{6.8}$$

which implies that the transformation T is entirely determined by vectors $T(\mathbf{v_1})$, ..., $T(\mathbf{v_i})$. Now let $\{\mathbf{w_1}, \ldots, \mathbf{w_m}\}$ be a basis for W. Then we can represent each vector $T(\mathbf{v_j})$ as

$$T(\mathbf{v_j}) = \sum_{i=1}^{m} a_{ij}\mathbf{w_i}. \tag{6.9}$$

Thus, the transformation T is entirely determined by the values of a_{ij}. If we put these values into an $m \times n$ matrix A, then we can conveniently use it to compute the vector output of T for any vector in V. To get A, every column j of A is a vector $T\left(\mathbf{v_j}\right)$ ■

In other words, fixed basis B_V and B_W of V and W respectively, coordinates of a vector $\mathbf{v}$ in basis B_V and the ones of $T(\mathbf{v})$ in B_W are related by the multiplication of a certain matrix, whose columns are the coordinates in B_W of the images of the elements in B_V.

Example. The linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ defined by $T(x,y,z) = (x-y, y-z)$ is given by the matrix

$$M = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}. \tag{6.10}$$

So T can also be defined for vectors $\mathbf{v} = (x,y,z)^T$ by the matrix product

$$T(\mathbf{v}) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \tag{6.11}$$

Here, the dimension of the initial vector space corresponds to the number of columns in the matrix, while the dimension of the target vector space is the number of rows in the matrix.

6.2 Kernel and image. Rank-nullity theorem

Related to a linear transformation $T : V \longrightarrow W$ from a vector space V to a vector space W, two subspaces of V and W, the *kernel* and the *image* of T respectively, are defined.

Definition 6.4 (Kernel). Let $T : V \longrightarrow W$ be a linear transformation, then the kernel of T is given by

$$\operatorname{Ker} T \overset{\text{def}}{=} \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}. \tag{6.12}$$

Definition 6.5 (Image). Let $T : V \longrightarrow W$ be a linear transformation, then the image of T is

$$\operatorname{Im} T \overset{\text{def}}{=} \{\mathbf{y} \in W \mid T(\mathbf{x}) = \mathbf{y} \text{ with } \mathbf{x} \in V\}. \tag{6.13}$$

Proposition 6.6. Let $T : V \longrightarrow W$ be a linear transformation. Then the image of T is a subspace of W and the kernel of T is a subspace of V.

Proof. We have to show that both the kernel and the image are closed under addition and scalar multiplication.

For the kernel, if $\mathbf{v}, \mathbf{u} \in \operatorname{Ker} T$; i.e. $T(\mathbf{v}) = T(\mathbf{u}) = \mathbf{0}$, since T is linear,

$$T(\mathbf{v}+\mathbf{u}) = T(\mathbf{v}) + T(\mathbf{u}) = \mathbf{0}+\mathbf{0} = \mathbf{0} \in \operatorname{Ker} T, \tag{6.14}$$
$$T(\lambda \mathbf{v}) = \lambda T(\mathbf{v}) = \lambda \mathbf{0} = \mathbf{0} \in \operatorname{Ker} T, \quad \forall \lambda \in F. \tag{6.15}$$

Therefore, $\operatorname{Ker} T \subset V$ is a vector subspace.

For the image, let $\mathbf{w_1}, \mathbf{w_2} \in \operatorname{Im} T$ and $\lambda \in F$. By definition, there exist $\mathbf{u_1}, \mathbf{u_2} \in V$ such that $\mathbf{w_1} = T(u_1)$ and $\mathbf{w_2} = T(\mathbf{u_2})$. Since T is linear,

$$T(\mathbf{u_1}+\mathbf{u_2}) = T(\mathbf{u_1}) + T(\mathbf{u_2}) = \mathbf{w_1}+\mathbf{w_2}, \tag{6.16}$$
$$T(\lambda \mathbf{u_1}) = \lambda T(\mathbf{u_1}) = \lambda \mathbf{w_1}. \tag{6.17}$$

Therefore, $\operatorname{Im} T$ is a vector subspace of W. ■

Proposition 6.7. Let $T : V \longrightarrow W$ be a linear transformation. Then, the following properties hold.

- T is a monomorphism $\Longleftrightarrow \operatorname{Ker} T = \{\mathbf{0}\}$.
- T is an epimorphism $\Longleftrightarrow \operatorname{Im} T = W$.
- T is an isomorphism $\Longleftrightarrow \operatorname{Im} T = W$ and $\operatorname{Ker} T = \{\mathbf{0}\}$.

Proof. To prove the first point is enough to note that $T(\mathbf{v_1}) = T(\mathbf{v_2})$ with different $\mathbf{v_1}, \mathbf{v_2} \in V$ implies $T(\mathbf{v}) = \mathbf{0}$ with $\mathbf{v} = \mathbf{v_1} - \mathbf{v_2} \neq \mathbf{0}$. The other way around, if $\mathbf{v} \neq \mathbf{0} \in \operatorname{Ker} T$ we could take $\mathbf{v_1} = \mathbf{v}$, $\mathbf{v_2} = \mathbf{0}$.

The last two points are obvious. ■

Definition 6.8. Let the *rank* of T be $(T) = \dim(\operatorname{Im} T)$ and the *nullity* of T be $\nu(T) = \dim(\operatorname{Ker} T)$.

Theorem 6.1 (Rank-Nullity Theorem). *Let $T : V \longrightarrow W$ be a linear transformation. Then $(T) + \nu(T) = \dim V$.*

In order to compute the dimension of these vector subspaces we need first to know the range of the matrix of T independent of basis.

Proposition 6.9. Let A be the matrix of a linear transformation $T : V \longrightarrow W$ for a finite number of basis of V and W, then

$$\dim \operatorname{Ker} T = \dim V - \operatorname{Rg} A \qquad \dim \operatorname{Im} T = \operatorname{Rg} A. \tag{6.18}$$

Corollary 6.10. Suppose $\dim V = \dim W = n < \infty$. A linear transformation $T : V \longrightarrow W$ is bijective $\Longleftrightarrow$ its matrix A verifies $\operatorname{Rg} A = n$.

6.3 Change of basis

Appendix A

Problem Sets

A.1 Problem Set 1: Set theory and functions

A.2 Problem Set 2: Equivalence relations

A.3 Problem Set 3: The $\mathbb{Z}$ and $\mathbb{Z}_n$ rings

A.4 Problem Set 4: Congruences

A.5 Problem Set 5: Polynomial rings

A.6 Problem Set 6: Elemental group theory

A.7 Problem Set 7: Matrices, vectors and linear systems of equations

A.8 Problem Set 8: Vector spaces and linear transformations

Bibliography

[Cha02] F. Chamizo. Álgebra i. primer curso de ingeniería informática. matematicas.uam.es/~fernando.chamizo/libreria/fich/APalgebraIinf96.pdf, 2002.