Linear Algebra: An Abstract Approach for Beginners

Nicolas Martinet

(draft)

Contents

Chapter 1

Introduction

1.1 Fundamental Concepts and Notations

Set

A set is a collection of distinct objects. Sets are wrote using curly brackets like this: $\{a, b, c\}$. The set with no elements is called the empty set and is denoted as $\varnothing$.

Finite and Infinite

Note that the order doesn't matter i.e $\{a, b, c\} = \{b, a, c\}$

Familiar sets include:

- The set of natural numbers, $\mathbb{N}$:

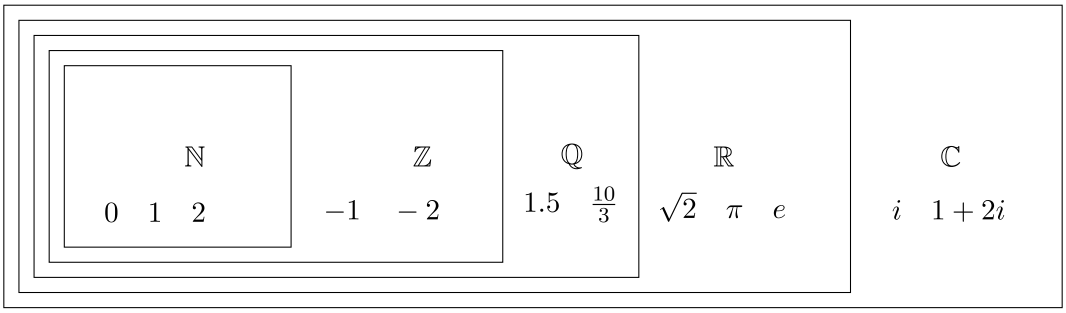
$$\mathbb{N} = \{0, 1, 2, 3, \ldots\}$$

- The set of integers, $\mathbb{Z}$, extends the natural numbers to include negative numbers:

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

- The set of rational numbers, $\mathbb{Q}$, is defined as the set of all numbers that can be expressed in the form $\frac{a}{b}$, where $a \in \mathbb{Z}$ is the numerator and $b \in \mathbb{Z}^*$ is the denominator (note $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$ is the set of all non-zero integers).
- The set of real numbers, $\mathbb{R}$, is perhaps the most widely known. It also includes numbers that cannot be expressed as fractions, such as $\sqrt{2}$ or π.

- The most intricate but also profoundly significant set is the set of complex numbers, $\mathbb{C}$. A typical complex number is written in the form $a + ib$, where $a \in \mathbb{R}$ is the real part, $b \in \mathbb{R}$ is the imaginary part, and i is the imaginary unit with the property $i^2 = -1$.



It's worth noting that:

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

More broadly, a set can contain any type of object, not just numbers. For instance, one could conceptualize a set S as:

$$S = \{\blacksquare, \bigstar, \blacklozenge, \blacktriangle\}$$

Some Notation

- The symbol $\in$ is used to represent ”belongs to” or ”is an element of.” It is used to indicate that an object or element is a member of a particular set or collection.
- The symbol $\subseteq$ is to represent the concept of ”is a subset of” or ”is contained in.” It indicates that one set is entirely contained within another set, including the possibility that they may be equal. In other words, if set A is a subset of set B, it means that every element of A is also an element of B.
- The symbol $\iff$ is used in mathematics to represent ”if and only if.” It is a biconditional symbol that indicates a two-way implication, meaning that a statement on the left side of $\iff$ is true if and only if the statement on the right side is also true.

$$A \subseteq B \iff \forall x,\ x \in A \Rightarrow x \in B$$

The symbol $\forall$ is a notation used in mathematics that stands for ”for all” or ”for every.” It's a way of specifying that the statement that follows it holds true for every element in a certain set or for every instance of a certain condition.

Let $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$

$$\forall x \in S, \quad x > 0$$

The symbol $\exists$ is used in mathematics to mean "there exists" or "there is at least one." It is used to assert the existence of at least one element that satisfies a certain condition within a set or a domain.

$$\exists x \in S, \quad x > 5$$

Symbols can also be combined

$$\forall x \in S, \exists y \in S, \quad x + y = 10$$

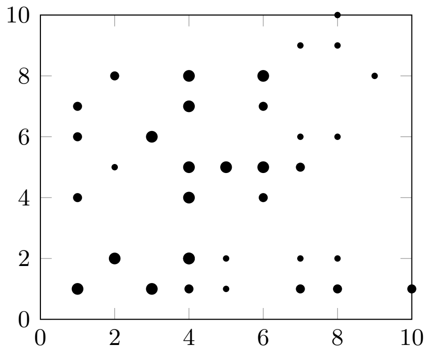
Given that, once you have chosen x, there exists a **unique** y that satisfies $x + y = 10$, we can actually rewrite above equation to be even more accurate:

$$\forall x \in S, \exists! y \in S, \quad x + y = 10$$

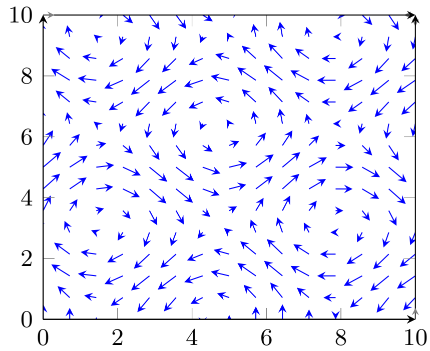
1.2 Vector Space basics

In physics, you may have heard of vectors. Vectors are a way of representing quantities that have both magnitude (size) and direction. Unlike regular numbers (scalars) that have only magnitude, vectors give us more information, like telling us not just how much, but also in which direction.

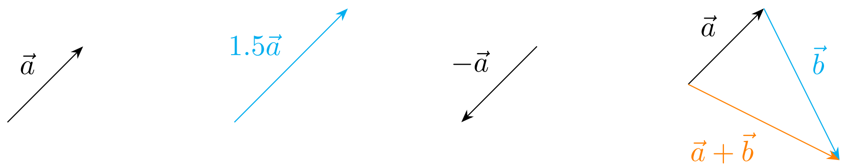
Scalar (like temperature)



Vector (like wind)



Vectors can be combined with another vector to form new vectors. This process is known as vector addition. A vector can also be scaled by a number, more precisely known as a scalar. Scaling a vector means changing its magnitude (size) without altering its direction. When you multiply a vector by a scalar, you make the vector longer or shorter. If the scalar is negative, it also reverses the vector's direction.

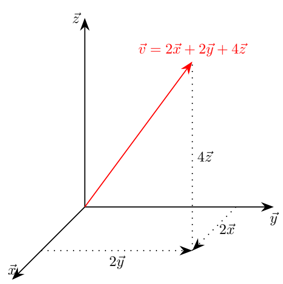


Vectors live in what we call a vector space. A vector space is a mathematical structure where vectors can be added together and multiplied (scaled) by scalars. The key property of a vector space is that any vector in it can be scaled and combined with another vector in the space, and the resulting vector will still belong to the same vector space. We say that vector spaces are "closed" under these operations, meaning that the operations of vector addition and scalar multiplication will always yield a vector that remains within the same vector space.

Using mathematical notation, we can define the vector space V as the set of all vectors that belong to the vector space, and $\mathbb{K}$ as the set of scalars, more precisely called a field . The closure property can be mathematically expressed as follows:

$$\forall (v_1, v_2) \in V^2, \forall k \in \mathbb{K}, \quad (kv_1 + v_2) \in V$$

In vector spaces, we can also introduce the concept of a basis. A basis is a set of vectors that, when combined, can create every possible vector in that space. We say that the basis span the entire space. In an n-dimensional vector space, a basis will consist of n such vectors. The number of vectors in the basis is by definition the dimension of the vector space.



A simple example is our 3-dimensional world. Here, the standard basis is often represented by the unit vectors along the x, y, and z axes. Every point or vector $\mathbf{v}$ in three-dimensional space can be described as a combination of these three vectors.

$$v = a\mathbf{x} + b\mathbf{y} + c\mathbf{z}$$

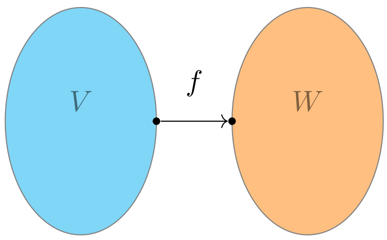
where a, b, and c are scalar coefficients corresponding to the x, y, and z coordinates, respectively.

we can therefore represent $\mathbf{v}$ with the vector column notation using a coordinate system derived from that basis.

$$\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

We will often encounter the concept vector subspaces. A vector subspace is essentially a smaller vector space that sits inside a larger one. It follows all the rules of a vector space but is confined to a portion of the larger space. For example, the 3-dimensional space we defined above with x, y, and z contains vector subspace like the plane formed by x and y. Every vector in this 2D plane (vector subspace) also belongs to the larger 3D space (vector space).

Another important concept is that of linear maps (or functions). These are special kinds of functions that take vectors from one vector space V and map them to another vector space W. These linear maps have unique properties: they preserve the operations of vector addition and scalar multiplication.



The linearity property of $f : V \to W$ can be written:

$$\forall (v_1, v_2) \in V^2, \forall k \in \mathbb{K}, \quad f(k \cdot v_1 + v_2) = kf(v_1) + f(v_2)$$

It's important to understand that all of these explanations are simplifications of what fundamentally is algebra. That's why we will delve deeper into the concepts and applications of algebraic structures, vector space and morphisms (generalization of function).

We aim to bridge the gap between basic understanding and more advanced concepts, offering a comprehensive guide that gradually builds from foundational principles to more complex theories.

1.3 Matrix

The column vectors we defined before is actually a specific case of what we commonly refer to as a matrix.

Matrices

A matrix is essentially a rectangular array of numbers, often used for representing linear transformations or systems of linear equations. Matrices are denoted as:

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1p} \\ m_{21} & m_{22} & \cdots & m_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{np} \end{bmatrix}$$

where M is a matrix of size $n \times p$ over a field $\mathbb{K}$, typically denoted as $M_{n,p}(\mathbb{K})$.

In essence, a column vector is a matrix with a single column and belong to $M_{n,1}(\mathbb{K})$. We can only add and substracts matrices if both are of the same dimensions.

$$\begin{bmatrix} -4 & 3 & 2 \\ 7 & 0 & 1 \end{bmatrix} + \begin{bmatrix} -2 & 7 & 12 \\ -1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -6 & 10 & 14 \\ 6 & 0 & 1 \end{bmatrix}$$

We can also multiply matrices by a scalar.

$$k \begin{bmatrix} -4 & 3 \\ 7 & 0 \end{bmatrix} = \begin{bmatrix} -4k & 3k \\ 7k & 0 \end{bmatrix}$$

A square matrix is a special type of matrix where the number of rows and columns are equal. It is denoted as $M_n(\mathbb{K})$, where n indicates the size of the matrix.

Transpose of a Matrix

Given a matrix M, its transpose, denoted as M^T, is defined by flipping M over its diagonal. Formally, if M^T is the transpose of $M \in M_{n,p}(\mathbb{K})$, then $M^T \in M_{p,n}(\mathbb{K})$ and $m^T_{i,j} = m_{j,i}$.

For a non-square matrix:

$$M = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}, \quad M^T = \begin{bmatrix} a & c & e \\ b & d & f \end{bmatrix}$$

For a square matrix:

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}, \quad M^T = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$$

Matrix Multiplication

Let $A \in M_{m,n}(\mathbb{K})$ and $B \in M_{n,p}(\mathbb{K})$. The product, denoted as AB, is a matrix in $M_{m,p}(\mathbb{K})$ where :

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \quad i = 1, \ldots, m \text{ and } j = 1, \ldots, p$$

.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}, \quad AB = \begin{bmatrix} ae + bg & \\ & \end{bmatrix}$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}, \quad AB = \begin{bmatrix} ae + bg & af + bh \\ & \end{bmatrix}$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}, \quad AB = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & \end{bmatrix}$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}, \quad AB = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

Consider the matrices $A \in M_{1,3}(\mathbb{K})$ and $B \in M_{3,1}(\mathbb{K})$. The result is a scalar in $\mathbb{K}$:

$$A = \begin{bmatrix} a & b & c \end{bmatrix}, \quad B = \begin{bmatrix} d \\ e \\ f \end{bmatrix}, \quad AB = ad + be + cf$$

Consider the matrices $A \in M_{n,1}(\mathbb{K})$ and $B \in M_{1,n}(\mathbb{K})$. The resulting matrix is in $M_n(\mathbb{K})$:

$$A = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad B = \begin{bmatrix} d & e & f \end{bmatrix}, \quad AB = \begin{bmatrix} ad & ae & af \\ bd & be & bf \\ cd & ce & cf \end{bmatrix}$$

Note: Matrix multiplication is not commutative. This means that for two matrices A and B, in general, $AB \neq BA$.

Identity Matrix

The identity matrix of size n, usually denoted by I_n, is a square matrix in which all the elements of the principal diagonal are ones and all other elements are zeros. The identity matrix for $M_n(\mathbb{K})$ is denoted by I_n and is defined as:

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

It acts as the multiplicative identity in matrix multiplication, meaning that for any matrix $A \in M_n(\mathbb{K})$, $AI_n = I_nA = A$.

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

This is a special way of forming an identity matrix.

Chapter 2

Algebraic Structure

2.1 Simple Algebraic Structures

Having understood the concept of a set, we now introduce the notion of a binary operation. Let a and b be two elements of a set S. More formally, this means $(a, b) \in S^2$.

Binary Operation

A binary operation, denoted by $\bullet$ is a mapping (or a function) from $S \times S$ to S.

To put it in simple terms, if you take two elements a and b from S, you can combine them to produce another element of S. Familiar binary operations include addition $(+)$ and multiplication $(\times)$ on $\mathbb{N}$. For instance, the operation $1 + 2$ is a binary operation from $\mathbb{N} \times \mathbb{N}$ that yields $3 \in \mathbb{N}$.

With a grasp on sets and binary operations, we can delve into the main topic of this section: algebraic structures. An algebraic structure $(S, (f_i)_{i \in I})$ comprises a set S from which any two elements can be combined using the functions f_i, which must adhere to certain properties.

This might sound abstract, so let's explore some fundamental algebraic structures.

Magma

A Magma $(M, \bullet)$ is the simplest algebraic structure. It consists of a set M and a binary operation $\bullet$. No specific properties are imposed on the operation.

Here, with just one operation, the function set $(f_i)_{i \in I}$ is essentially reduced to $\bullet$. In essence, a magma ensures that any two elements in M, say $(a, b) \in M^2$, can be combined using the operation $\bullet$, and the result will still belong to M. Consider the set $M = \{\blacksquare, \bigstar, \blacklozenge\}$. This can be visualized with the following table:

$\bullet$	$\blacksquare$	$\bigstar$	$\blacklozenge$
$\blacksquare$	$\blacksquare$	$\blacksquare$	$\bigstar$
$\bigstar$	$\blacklozenge$	$\bigstar$	$\blacksquare$
$\blacksquare$	$\blacksquare$	$\blacklozenge$	$\bigstar$

where for instance $(\blacksquare \bullet \bigstar) = \blacksquare \neq (\bigstar \bullet \blacksquare) = \blacklozenge$ (not commutative)

Note: The crucial aspect here is that the operation $\bullet$ can combine any two elements in M, and the resultant will always be an element of M. This is called the **closure** property of a binary operation.

Many operations and sets we encounter form a magma. For instance, $(\mathbb{N}, +)$, $(\mathbb{N}, \times)$, $(\mathbb{R}, +)$, $(\mathbb{R}, \times)$ are all examples of magmas. However, $(\mathbb{N}, -)$ is not a magma because the operation $1 - 2$ yields -1, and $-1 \notin \mathbb{N}$. On the other hand, if we consider the set of integers, $\mathbb{Z}$, instead of the natural numbers $\mathbb{N}$, then $(\mathbb{Z}, -)$ does form a magma. Also $(\mathbb{N}, \div)$,$(\mathbb{Z}, \div)$,$(\mathbb{Q}, \div)$,$(\mathbb{R}, \div)$ are not magma because division by 0 is not defined. Also in the case of $\mathbb{N}$ and $\mathbb{Z}$, the operation $\frac{a}{b}$ is not closed. $(\mathbb{Z}^*, +)$ is not a magma because for instance $2 + (-2) = 0 \notin \mathbb{Z}$.

Having established an understanding of the most rudimentary algebraic structures, we can now delve into more advanced structures, which can be thought of as magmas but with additional properties.

2.2 Monoid

Monoid

A Monoid $(M, \bullet)$ is a magma that satisfies both **associativity** and the existence of an **identity element**. The properties are:

Associativity:

$$\forall (a, b, c) \in M^3, \quad (a \bullet b) \bullet c = a \bullet (b \bullet c)$$

Identity:

$$\exists e \in M, \quad \forall a \in M, \quad a \bullet e = e \bullet a = a$$

Here, e is called the identity element. Two straightforward examples of identity elements are 0 for addition $(+)$ and 1 for multiplication $(\times)$ on $\mathbb{R}$. This is evident as $1 \times a = a$ and $a + 0 = a$ (with a a number). Similarly, the property of associativity can be illustrated by the $1 + (2 + 3) = (1 + 2) + 3$ and $1 \times (2 \times 3) = (1 \times 2) \times 3$.

By analogy, this algebraic structure complements the previous ones by incorporating the operation of addition.

2.3 Group

Group

A Group $(G, \bullet)$ is a commutative monoid, meaning it possesses the properties of **associativity**, **identity**, and additionally, the **inverse**.
Inverse:

$$\forall a \in G, \exists b \in G \text{ such that } a \bullet b = e$$

Here b is the inverse of a. Every element in G as an inverse. It's worth noting that each successive algebraic structure we introduce brings with it new properties. These added constraints reduce the number of structures that satisfy it. It ensure that we study only the most interesting ones. For instance $(\mathbb{R}, \times)$ that could previously be considered a monoid cannot be a group because 0 is not invertible. We must therefore consider $(\mathbb{R}^*, \times)$ instead.

While $(\mathbb{N}, +)$ is a monoid, it isn't a group because $\forall a \in \mathbb{N}$, the value b that satisfies $a + b = 0$ is $-a$, but $-a \notin \mathbb{N}$. Thus, transitioning from a monoid to a group necessitates evolving structures like $(\mathbb{N}, +)$ to others like $(\mathbb{Z}, +)$ to ensure the inverse property holds.

To put it another way, the inverse property of a group guarantees that the binary operation $\bullet$ is always indirectly complemented by an "inverse" operation. In the context of addition $(+)$, this complementary operation is subtraction $(-)$. For multiplication $(\times)$, it's division $(\div)$.

> "Dividing is equivalent to multiplying by the inverse, and subtracting is equivalent to adding the inverse (commonly called the opposite)"

Before delving into the algebraic structure of a field, denoted as $\mathbb{K}$, it is crucial to understand another complex structure called a ring. Unlike simpler structures, both fields and rings incorporate two binary operations, denoted $+$ and $\times$.

2.4 Ring

Definition of a Ring

A ring $(R, +, \times)$ is an algebraic structure satisfying the following properties:

- $(R, +)$ forms a commutative group (with identity 0).
- $(R, \times)$ forms a monoid (with identity 1).
- **Commutativity of** $(R, +)$**:**

$$\forall (a, b) \in R^2, \quad a + b = b + a$$

- **Distributivity of** $+$ **with respect to** $\times$**:**

$$\forall (a, b, c) \in R^3, \quad a \times (b + c) = a \times b + a \times c \quad \text{(Left Distributivity)}$$

$$\forall (a, b, c) \in R^3, \quad (a + b) \times c = a \times c + b \times c \quad \text{(Right Distributivity)}$$

Note that we've switched from the generic operation symbol $\bullet$ to $+$ and $\times$, which are more familiar. This change is motivated by the realization that the structures satisfying ring axioms are inherently numerical.

In everyday language, when we express "there are 2 apples and 3 apples," it is equivalent to saying "there are $2 + 3 = 5$ apples." This translation is a manifestation of the inherent counting property of numbers. In a similar vein, the expression $(ac + bc) = (a + b)c$ in algebra is analogous to the previous phrasing. It becomes evident that this algebraic property, known as **distributivity**, emerges from the fundamental capability of numbers to quantify. This is their primary function.

In the definition of a ring, $(R, +, \times)$, we notice that the sole set in play is R. This differs from the apple example where a distinction is made between the numbers (2 and 3) and the objects being counted (apples). In the context of rings, the entities we're counting with are from the same set as the entities we're counting.

Prior to introducing the property of commutativity, our algebraic structures could encompass a broader range of sets, including sets of functions. For instance, consider the set of functions mapping $\mathbb{R}$ to $\mathbb{R}$, denoted as $\mathbb{R}^{\mathbb{R}}$. Our previous constraints (like the one of a group) did not prevent us from considering such sets, such as $(\mathbb{R}^{\mathbb{R}}, +)$. Yet, the imposition of commutativity dramatically narrows this scope. To illustrate, consider the functions $f(x) = 2x$ and $g(x) = x^2$, both of which belong to $\mathbb{R}^{\mathbb{R}}$. Although individually they fit within our structure, their composition lacks commutativity, evidenced by $f \circ g \neq g \circ f$. Therefore sets shuch as $\mathbb{R}^{\mathbb{R}}$ cannot verify **commutativity** property of a ring

Those observations underscores that the constraints imposed on R to satisfy ring properties essentially ensure that R must be a set of numbers.

By analogy, rings complements groups (that already have addition and subtraction) by incorporating the operation of multiplication ($\times$).

Binomial Theorem

Let $(a, b) \in R^2$ that commute ($a \times b = b \times a$)

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

where $\binom{n}{k}$ is the binomial coefficient, calculated as $\frac{n!}{k!(n-k)!}$.

Proof:

1. **Base Case:** For $n = 0$,

$$(a+b)^0 = 1$$

$$\sum_{k=0}^{0} \binom{0}{k} a^{0-k} b^k = \binom{0}{0} a^0 b^0 = 1.$$

Hence, the base case is true.

2. **Inductive Step:** Assume the statement is true for some $n \geq 0$, i.e.,

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

We need to show it holds for $n + 1$.

3. **Proving for $n+1$:**

$$
\begin{aligned}
(a+b)^{n+1} &= (a+b)^n \times (a+b) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k\right) \times (a+b) \quad \text{(by induction hypothesis)} \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \times a + \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \times b \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k+1} \quad \text{(using commutativity of } a \text{ and } b\text{)} \\
&= a^{n+1} + \sum_{k=1}^{n} \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^{n-1} \binom{n}{k} a^{n-k} b^{k+1} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n} \binom{n}{k} a^{n+1-k} b^k + \sum_{k=1}^{n} \binom{n}{k-1} a^{n+1-k} b^k + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1}\right] a^{n+1-k} b^k + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} a^{n+1-k} b^k + b^{n+1} \quad \text{(using Pascal's rule)} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k
\end{aligned}
$$

□

2.5 Field

Field

A Field $(\mathbb{K}, +, \times)$ is defined as a ring where:

- $(\mathbb{K}, +)$ forms a commutative group (with identity 0).
- $(\mathbb{K}^*, \times)$ forms a **commutative group** (with identity 1)

The distributive rules of $+$ with respect to $\times$ also apply.

By analogy, this algebraic structure complements the previous ones by incorporating the operation of division ($\div$).

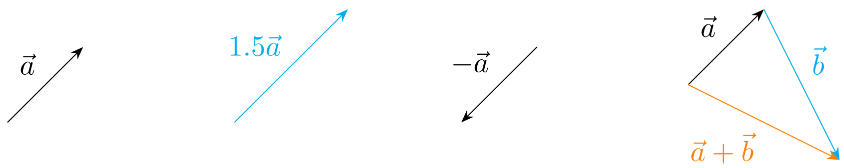
Add explanation k*

The notation $\mathbb{K}$ originates from the German word "Körper", which translates to "body" in English. This notation was introduced in 1871 by the German mathematician Richard Dedekind 22 years before the term "field" introduced by the American mathematician E. H. Moore in 1893.

Note that in $\mathbb{K}^*$, we specifically exclude the element 0 due to the requirement of inverse in group structures. By analogy, this final algebraic structure complements the previous ones by incorporating the operation of division.

Chapter 3

Vector Space



What is a vector ? In both physics and geometry, a vector is often described as an arrow with both direction and magnitude. One property is that given any two vectors $\vec{a}$ and $\vec{b}$, a new vector $\vec{c}$ can be formed by combining $\vec{a}$ and $\vec{b}$.

$$\vec{c} = \vec{a} + \vec{b}$$

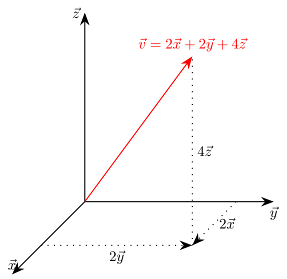
Another important property is the ability to *scale* a vector $\vec{v}$ by a scalar k. The operation of scaling adjusts the length of the vector without altering its direction.

$$k\vec{v}$$

In classical Newtonian physics, our world is described as a three-dimensional space. Objects can move [forward/backward], [left/right], and [up/down]. These directions can be denoted by the basis vectors $\vec{x}$, $\vec{y}$, and $\vec{z}$. Any point in this 3D space can be represented as a linear combination of these basis vectors. We say $\{\vec{x}, \vec{y}, \vec{z}\}$ forms a *spanning set* of the space. This means every vector $\vec{v}$ in the space can be represented as a combination of $\vec{x}$, $\vec{y}$ and $\vec{z}$.

$$\forall \vec{v} \in E, \;\; \exists (a, b, c) \in \mathbb{R}^3, \quad \vec{v} = a\vec{x} + b\vec{y} + c\vec{z}$$

"for all vector v in the vector space E, there exists real scalars (a,b,c) such that v is a linear combination of the vectors $\vec{x}$, $\vec{y}$, and $\vec{z}$ "



here $a = 2$, $b = 2$, and $c = 4$.

This is an intuitive but restricted understanding of what vectors can be.

Algebra broadens the concept of vectors far beyond just arrows in space. In mathematics, a vector is essentially any object that can be added together and multiplied by scalars, according to certain rules. This generalized definition encompasses a wide range of mathematical objects including matrices, functions, series, and polynomials. All of these can be scaled and combined just like the geometric vectors, making vectors a versatile and central concept in mathematics.

say we stop arrow on vector and explain scalar is K field

Vector Space over a Field $\mathbb{K}$

Let $(E, +)$ be a commutative group and $(\mathbb{K}, +, \times)$ be a field. We say that E is vector space over $\mathbb{K}$, if there exists an external binary operation from $\mathbb{K} \times E$ to E, denoted as $a \cdot x$, satisfying the following axioms:

- $\forall x \in E, 1_{\mathbb{K}} \cdot x = x$
- $\forall a, b \in \mathbb{K}, \forall x \in E, (a + b) \cdot x = a \cdot x + b \cdot x$
- $\forall a, b \in \mathbb{K}, \forall x \in E, (a \times b) \cdot x = a \cdot (b \cdot x)$
- $\forall a \in \mathbb{K}, \forall x, y \in E, a \cdot (x + y) = a \cdot x + a \cdot y$

In this context, the elements of E are called vectors and are denoted by $x, y, z, \ldots$, while the elements of $\mathbb{K}$ are called scalars and are denoted by $a, \beta, \gamma, \lambda, \ldots$.

This definition may seem daunting at first, but it is not introducing any fundamentally new concepts. Here's a breakdown:

- The **commutative group** $(E, +)$ ensures that vectors can be combined through addition to form a new vector in E, and that an identity element (the null vector), often denoted 0_E, exists in E.

- The **field** $(\mathbb{K}, +, \times)$ serves as the scalar field. It ensures that any vector in E can be multiplied by any scalar in $\mathbb{K}$. For example, the opposite vector $-v$ of v, or kv where k can be as large as you desire, will still belong to E.

- The **external binary operation** from $\mathbb{K} \times E$ to E, along with the four axioms, ensures the interactions between the scalars $\mathbb{K}$ and the vectors E. It ensures that the scaling of any vector in E by any scalar in $\mathbb{K}$ will still be in the vector space E

If the notion of a scalar field $\mathbb{K}$ feels too abstract, it may be comforting to know that in most cases, $\mathbb{K}$ is either $\mathbb{R}$ (the set of real numbers) or $\mathbb{C}$ (the set of complex numbers). If you're not comfortable with complex numbers, you can essentially think of $\mathbb{K}$ as being $\mathbb{R}$, the set of all real numbers.

Vector Sub-space, Spanning Set, linearly independent Set

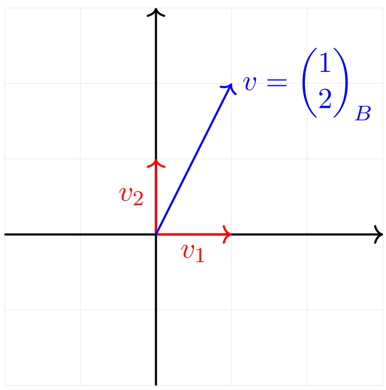
Until now, we have been representing vectors in an abstract manner without specific coordinates. However, in practical applications like physics, coordinates play a crucial role in vector manipulation. To assign coordinates to a vector, we introduce the concept of a *basis*.

Consider a basis $B = \{v_1, v_2\}$ for a given vector space. A vector v in this space can be expressed as a linear combination of the vectors in B:

$$v = av_1 + bv_2 \quad \text{where} \quad (a, b) \in \mathbb{K}^2$$

This allows us to represent the vector v in terms of its coordinates with respect to the basis B as :

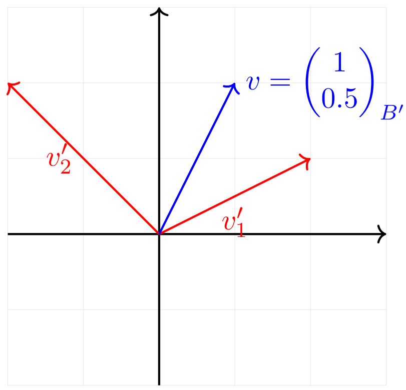
$$v = \begin{pmatrix} a \\ b \end{pmatrix}_B$$



"to obtain vector v, you need to scale vector v_1 by $a = 1$ and vector v_2 by $b = 2$".

It is essential to understand that even though we are now representing v using numbers, the vector itself is not a collection of numbers. The abstract nature of a vector remains intact because our coordinate representation of v is based not on numbers, but on the two vectors of the basis which are scaled by scalars from the vector field $\mathbb{K}$ (in this case $\mathbb{R}$).

If we consider a different basis $B' = \{v_1', v_2'\}$, the representation $v = \begin{pmatrix} a \\ b \end{pmatrix}_{B'}$ would differ from our previous representation.



It's important to emphasize that v does not inherently belong to $\mathbb{R}^n$ (or $\mathbb{K}^n$) as previously explained. However, for simplification in some contexts, we might assume a standard basis and start representing v directly as if $v \in \mathbb{R}^n$. Such a representation is commonly referred to as a *column vector* but is actually a special case of a matrix. Given the focus of this book on the theoretical aspects of algebra, we will try to avoid using the simplified *column vector* representation.

3.0.1 Families of Vectors

In earlier discussions, we often referred to sets of vectors or elements without explicitly indexing them. However, using indexed families offers a clearer and more structured way to refer to and manipulate elements within these sets.

Indexed Family

Let I and X be sets, and let f be a function such that

$$\begin{array}{rcccc} f & : & I & \rightarrow & X \\ & & i & \mapsto & f(i) = x_i \end{array}$$

The function f thus establishes an indexing of elements in X denoted by $(x_i)_{i \in I}$. I is the index set and X is the indexed set.

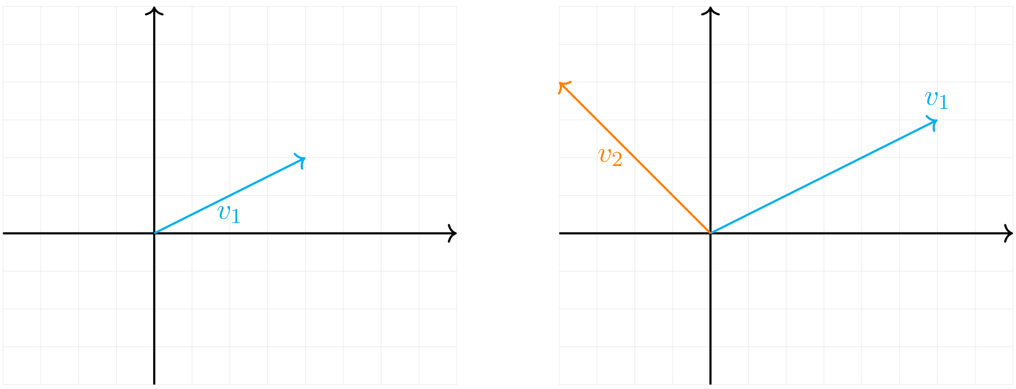
Add infinite and family

Definition of a Spanning Family

Let K be a field, and let E be a vector space over K. A finite family $(v_1, v_2, \ldots, v_n)$ of elements of E (vectors) is said to be a generating set of E if:

$$\forall x \in E, \exists (\lambda_1, \lambda_2, \ldots, \lambda_n) \in K^n, x = \sum_{k=1}^{n} \lambda_k v_k.$$

This means every vector x in E can be expressed as a linear combination of the vectors in the family.



For instance, the family $\{v_1\}$ is not spanning the two-dimensional vector space because it consists of only one vector. However, the family $\{v_1, v_2\}$ on the right is spanning the two-dimensional vector space, as any vector in the space can be reached by combining v_1 and v_2 with scalars from the field.

Definition of Linearly Independent Family

Let $(v_i)_{i\in I}$ be a family (finite or infinite) of vectors in a vector space E. This family is said to be linearly independent if the only linear combination of the vectors v_i equal to the zero vector 0_E is the one where all coefficients are zero. For a finite family $(v_i)_{1\leq i\leq n}$, this condition is expressed as:

$$\forall(a_1,\dots,a_n)\in K^n,\quad \left(\sum_{i=1}^{n} a_i v_i = 0_E \Rightarrow a_1 = a_2 = \cdots = a_n = 0_K\right).$$

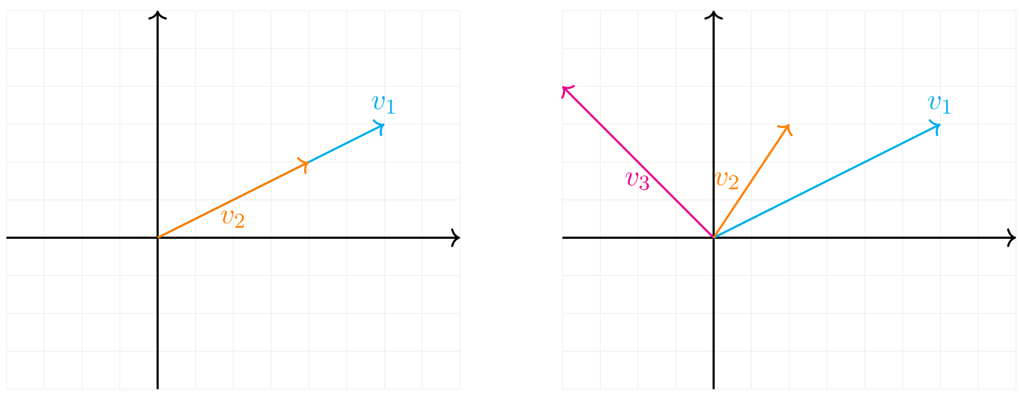
If one of the vectors in the family is the null vector 0_E, then the family of vectors is automatically linearly dependent.

Equivalently, a family of vectors $(v_i)_{i\in I}$ in a vector space E is linearly independent if no vector in the family can be expressed as a linear combination of the others:

$$\forall(a_1,\dots,a_n)\in K^n,\quad \left(v_j a_j = \sum_{\substack{i=1\\ i\neq j}}^{n} -a_i v_i \Rightarrow a_1 = a_2 = \cdots = a_n = 0_K\right).$$

This implies that if one of the vectors can be represented as a linear combination of the others, then the family $(v_i)_{i\in I}$ is linearly dependent:

$$\forall v\in(v_i)_{i\in I}, \exists(a_1,\dots,a_{n-1})\in K^{n-1},\ \text{not all } a_i = 0_K,\quad v = \sum_{\substack{i=1\\ v_i\neq j}}^{n} a_i v_i$$



On the left, the family $\{v_1, v_2\}$ is clearly dependent because v_1 can be expressed as a scaling of v_2. On the right, the family $\{v_1, v_2, v_3\}$ is also linearly dependent because each of the vectors can be represented by a linear combination of the other two. Note that $\{v_1, v_2, v_3\}$ is spanning the two-dimensional vector space. If you take only two elements of the family, like $\{v_2, v_3\}$ or $\{v_1, v_3\}$, then the family is linearly independent and still spans the vector space. If you take only one element like $\{v_1\}$, then the family is independent but not spanning the vector space.

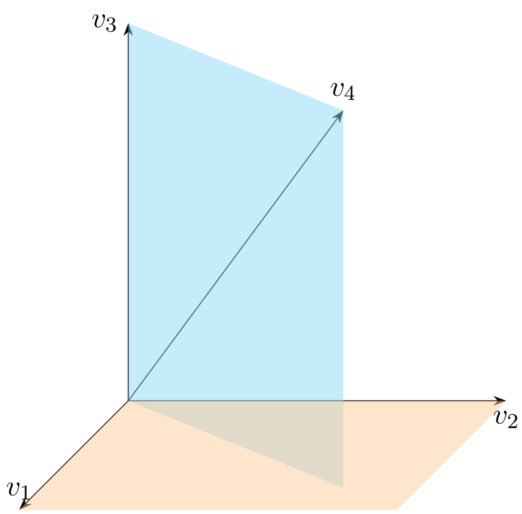
Here, you may gain the intuition that families such as $\{v_1, v_2\}$ or $\{v_1, v_3\}$ have just what we need to to span the vector space, without any redundant vectors. This is the essence of what we call a **basis**.

Basis in Vector Space

A basis B in a vector space E is a family of vectors $\{v_1, ..., v_n\}$ where:

- The family is **linearly independent**, meaning no vector in the family can be expressed as a linear combination of the others.
- The family **spans the space**, meaning any vector in E can be expressed as a linear combination of the vectors in the family.

Without the spanning property, the basis would be unable to cover certain portions of the vector space. The requirement for a basis to be a set a linearly independent family ensure that we do not include redundant vectors in the basis.



We can summarize with this three-dimensional vector space representation where $\{v_1, v_2, v_3\}$ is a basis for the vector space. $\{v_1, v_2\}$ is not a basis because it does not span the vector space. $\{v_1, v_2, v_3, v_4\}$ is not a basis because it is linearly dependent. There are many other bases of the vector space like $\{v_1, v_2, v_4\}$ or $\{v_2, v_3, v_4\}$.

Even though $\{v_1, v_2\}$ is not a basis for the vector space, it is a basis for the subspace represented by the orange plane. Similarly, $\{v_3, v_4\}$ is a basis for the subspace represented by the blue plane.

3.1 Examples of Vector Spaces

3.1.1 Numbers

Real Numbers Numbers, such as those in $\mathbb{R}$, may seem quite obvious as vector spaces. For any number $x \in \mathbb{R}$, the additive inverse is simply $-x$. A basis for $\mathbb{R}$ as a vector space over itself could be as simple as $\{1\}$. Every real number can be multiplied by another real number, and the result is still a real number, so the field in this context is $\mathbb{R}$ itself. This makes $\mathbb{R}$ one of the most trivial vector spaces over the field $\mathbb{R}$.

Complex Numbers Similarly, complex numbers $x \in \mathbb{C}$ can form a vector space. In this case, possible basis could be $\{1\}$ or $\{i\}$. Complex numbers form a vector space over $\mathbb{C}$

3.1.2 Tuples

Tuples

A **tuple** is an ordered list of elements. It can be defined as a finite sequence $(x_1, x_2, \ldots, a_x)$.

Examples:

1. **Multi-Dimension Real Vector Space:** $\mathbb{R}^3$ is a typical example of vector space. One vector in this space can be represented such as (x, y, z). This is the example we introduced at the beginning of the Vector Space chapter. In two dimension on a plane, $\mathbb{R}^2$ works too. More generally $\mathbb{R}^n$ over the field $\mathbb{R}$ is a vector space.

2. **Multi-Dimension Complex Vector Space:** $\mathbb{C}^n$ over the fields $\mathbb{R}$ is a vector space. or $\mathbb{C}$.

3.1.3 Polynomials

Polynomials

Polynomials are expressions of the form $P(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots + a_1 X + a_0$, where a_i are coefficients from a field $\mathbb{K}$, and n is a non-negative integer representing the degree of the polynomial. More compactly, a polynomial can be expressed as $P(X) = \sum_{i=0}^{n} a_i X^i$. The set of all such polynomials is denoted by $\mathbb{K}[X]$, which forms a ring. In the context of vector spaces, $\mathbb{K}[X]$ can be considered a vector space over $\mathbb{K}$, with the vectors being the polynomials themselves and scalar multiplication and addition defined in the usual manner of polynomial arithmetic.

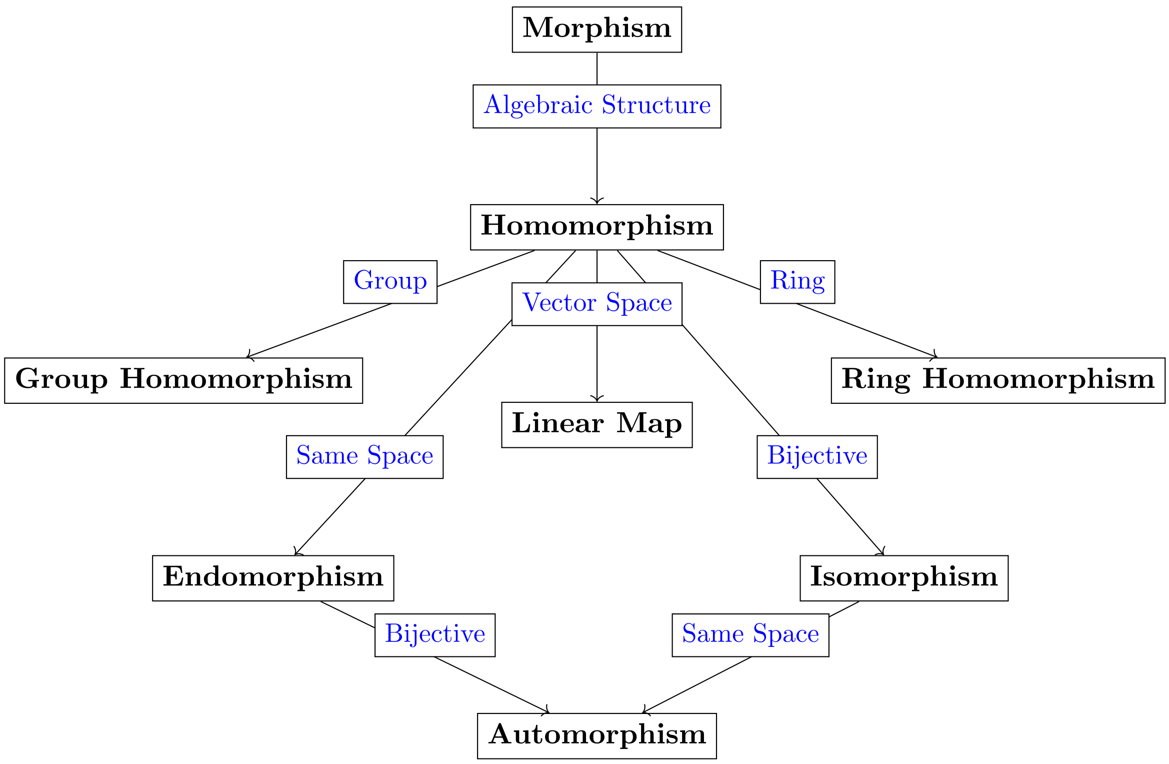
Examples:

1. **Polynomials over Real Numbers:** Consider the set of all polynomials with real coefficients, denoted as $\mathbb{R}[x]$. This set forms a vector space over the field $\mathbb{R}$. A typical vector (polynomial) in this space might be $p(x) = 2x^2 + 3x + 4$.

2. **Space of Polynomials up to Degree n:** The set of all polynomials with degree less than or equal to n over a field $\mathbb{K}$, denoted as $\mathbb{K}[x]_{\leq n}$, is a vector space. For instance, in $\mathbb{R}[x]_{\leq 2}$, a vector might be $q(x) = a_0 + a_1 x + a_2 x^2$, where a_0, a_1, and a_2 are real numbers.

Chapter 4

Morphism



4.1 Map (Function)

Map (Function)

A map f (or function) is a binary relation between two sets, where each element of the first set X (called the domain) is associated with **exactly one** element of the second set Y (called the codomain).

$$f : X \to Y$$

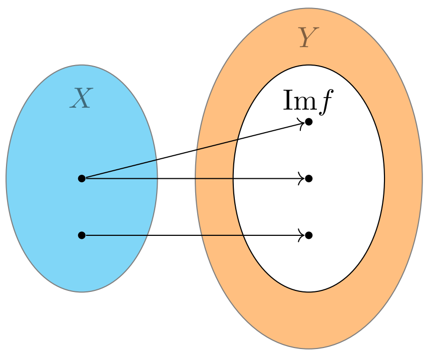
- X is the **domain**
- Y is the **codomain**

Image

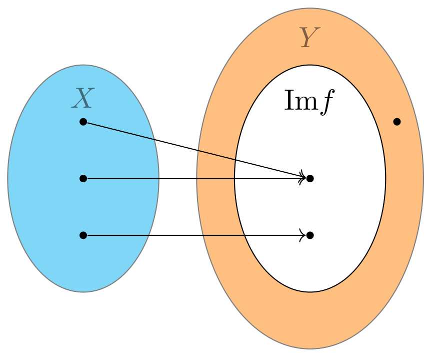
The **image** of f, denoted by $\mathrm{Im}(f)$, is the set of all output values that f may produce from the elements of its domain X.

$$\mathrm{Im}(f) = \{f(x) \mid x \in X\}$$

$$\mathrm{Im}(f) \subseteq Y$$



Is not a Function



Is a Function

- On the left, f **is not a function** because some points in set X are associated with multiple points in $f(X)$. This violates the definition of a function, where each element in the domain must map to exactly one element in the codomain.
- On the right, f **is a function** since every point in set X is associated with exactly one point in $f(X)$. Notes that some point in the codomain may not belong to $\mathrm{Im}(f)$.

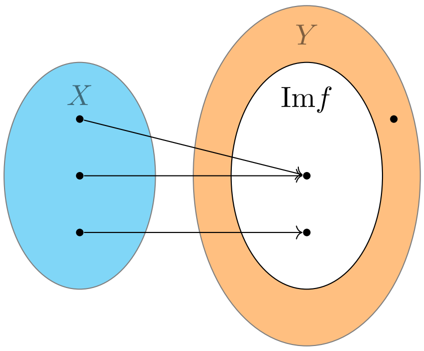
It is possible for one element in Y to be the result of mapping from two (or more) different elements in X.

Injective Function

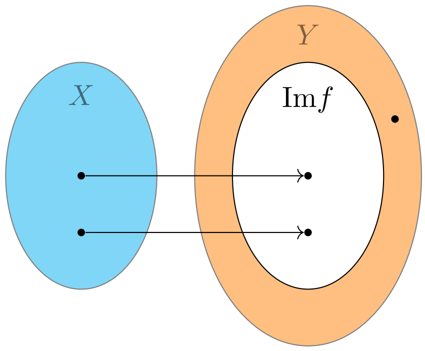
A function $f : X \to Y$ is said to be injective if for every element y in Y, there is **at most** one element x in X such that $f(x) = y$. Equivalently, this can be expressed as:

$$\forall (x_1, x_2) \in X^2, \quad f(x_1) = f(x_2) \Rightarrow x_1 = x_2.$$

This condition ensures that no two distinct elements in X map to the same element in Y.



Is not an Injection



Is an Injection

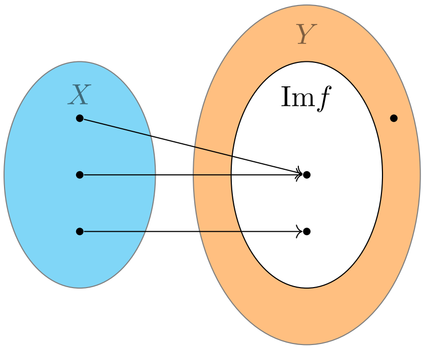
Surjective Function

A function $f : X \to Y$ is said to be surjective if for every element y in Y, there is **at least** one element x in X such that $f(x) = y$. Equivalently, this can be expressed as:

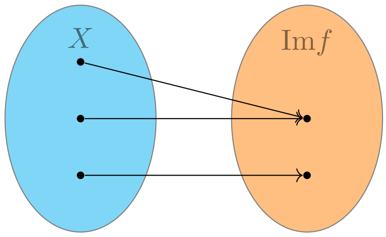
$$\forall y \in Y, \exists x \in X, \quad y = f(x)$$

This condition ensures that the codomain is equal to the image of f.

$$\mathrm{Im}(f) = Y$$



Is not a Surjection

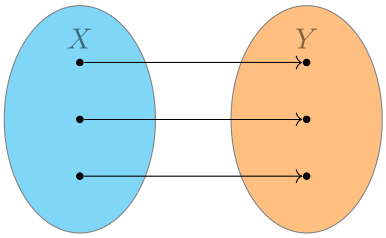


Is a Surjection

Bijective Map

A map $f : X \to Y$ is said to be bijective if it is both **injective** and **surjective**. This means that for every element y in Y, there is **exactly** one element x in X such that $f(x) = y$. Equivalently, this can be expressed as:

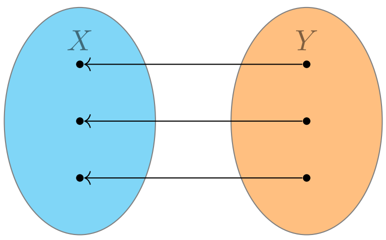
$$\forall y \in Y, \exists! x \in X, \quad f(x) = y.$$



Inverse Function of a Bijection

$f : X \to Y$ is bijective $\Leftrightarrow$ $f^{-1} : Y \to X$, the inverse function of f exists.

- $\forall x \in X, \quad f^{-1}(f(x)) = x$
- $\forall y \in Y, \quad f(f^{-1}(y)) = y$



Proof:

1. ($\Rightarrow$) Assume f is bijective (both injective and surjective).
   - Surjectivity of f ensures that for every $y \in Y$, there exists at least one $x \in X$ such that $f(x) = y$.
   - Injectivity of f ensures that this x is unique for each y.
   - Thus, we can define a function $f^{-1} : Y \to X$ where $f^{-1}(y)$ is the unique x for which $f(x) = y$.
   - This f^{-1} satisfies $f^{-1}(f(x)) = x$ and $f(f^{-1}(y)) = y$, making it the inverse of f.
2. ($\Leftarrow$) Assume an inverse $f^{-1} : Y \to X$ of f exists.
   - For f^{-1} to be well-defined, every $y \in Y$ must correspond to exactly one $x \in X$ such that $f(x) = y$. This implies f is surjective.
   - For $f^{-1}(f(x)) = x$ to hold for all $x \in X$, different inputs in X must produce different outputs in Y, implying f is injective.
   - Hence, f is both injective and surjective, i.e., bijective.

As we conclude our exploration of the concept of maps, it's important to introduce the notion of 'Morphism'. While a 'map' typically refers to any function or relation between sets, a 'Morphism' is a more structured concept. A morphism needs to respect and preserve the underlying structures of these sets.

4.2 Homomorphism

We initially considered maps as binary relations between sets. However, a richer understanding emerges when we consider maps between algebraic structures. In this context, the function needs to be between two algebraic structures of the same type and must preserve the structural properties. We refer to such a function as **homomorphisms**.

For example:

- **Group Homomorphism:** Given two groups $(G, \bullet)$ and $(H, *)$, a group homomorphism is a function $f : G \to H$ such that for all $(a, b) \in G^2$,

$$f(a \bullet b) = f(a) * f(b).$$

- **Ring Homomorphism:** Given two rings $(R, +_R, \times_R)$ and $(S, +_S, \times_S)$, a ring homomorphism is a function $f : R \to S$ such that the following properties hold:

1. $\forall (a,b) \in R^2, \quad f(a +_R b) = f(a) +_S f(b)$.
2. $\forall (a,b) \in R^2, \quad f(a \times_R b) = f(a) \times_S f(b)$.
3. $f(1_R) = 1_S$, where 1_R and 1_S are the multiplicative identities in R and S respectively.

In a ring homomorphism $f : R \to S$, the additive identity is preserved. This means $f(0_R) = 0_S$.

Proof:

1. Consider the additivity property of a ring homomorphism: $f(a+_R b) = f(a)+_S f(b)$.
2. Letting $a = b = 0_R$, we obtain $f(0_R+_R 0_R) = f(0_R)+_S f(0_R)$. Since $0_R+_R 0_R = 0_R$, this simplifies to $f(0_R) = f(0_R) +_S f(0_R)$.
3. Subtracting $f(0_R)$ from both sides of the equation using the inverse property of $(S, +_S)$, we get $0_S = f(0_R)$.

$\square$

The particularly important homomorphism we will be focusing on now is the one on vector spaces. This homomorphism is called a Linear Map.

4.3 Vector Space Homomorphism (Linear Map)

Linear Map (Vector Space Homomorphism)

Let V and W be vector spaces over $\mathbb{K}$. A map $f : V \to W$ is called linear if it satisfies the following properties:

- **Additivity:** $\forall (u,v) \in V^2$, $f(u + v) = f(u) + f(v)$.
- **Homogeneity:** $\forall \lambda \in \mathbb{K}, \forall v \in V, f(\lambda v) = \lambda f(v)$.

These properties ensure that f preserves the vector space structure when mapping from V to W. the properties of additivity and homogeneity for a linear map can be condensed into

$$\forall (u,v) \in V^2, \forall \lambda \in \mathbb{K}, f(\lambda u + v) = \lambda f(u) + f(v).$$

The set of all linear maps between two vector spaces V and W is denoted by $L(V, W)$.

Now that we have explored homomorphisms and linear maps, it should becomes clear that the concept of homomorphisms, and more generally a morphisms, is an abstract generalization of the basic concept of linearity $f(a + b) = f(a) + f(b)$.

Kernel of a Linear Map

The kernel of $f \in L(V, W)$, denoted as $\text{Ker}(f)$, is the set of all vectors in V that map to the zero vector in W.

$$\text{Ker}(f) = \{v \in V \mid f(v) = 0_W\}$$

Image of a Linear Map

The image of $f \in L(V, W)$, denoted by $\text{Im}(f)$, is the set of all vectors that f may produce from the vectors in V.

$$\text{Im}(f) = \{f(v) \mid v \in V\}$$

We previously introduced the concept of the image of a function (or morphism). Here, we redefine it in the context of linear maps. Additionally, we introduce the concept of the Kernel. It is important to note that the Kernel could not be defined for just any map, as there was no identity element in a general setting unlike with vector spaces.

In our previous discussions, we explored the concepts of surjectivity and injectivity in the context of maps. These concepts are equally applicable to linear maps. What is particularly fascinating is that, for vector spaces and linear maps, we can establish a direct connection between Surjectivity - Injectivity and the concepts of Images - Kernels. This connection is illuminated by two fundamental properties, which we will now examine in detail.

Property 1

$f \in L(V, W)$ is injective $\Leftrightarrow \text{Ker}(f) = \{0_V\}$.

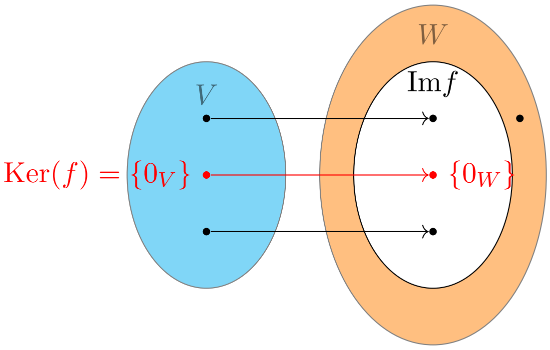
This means being injective is equivalent to the condition where the only element in V that maps to the zero vector 0_W in W is the zero vector 0_V. That's why we usually say that the kernel of a linear map serves as a measure of its non-injectivity. If the kernel contains elements other than 0_V, this indicate that different elements in V are mapped to the same element in W, meaning that f is not injective.

Property 2

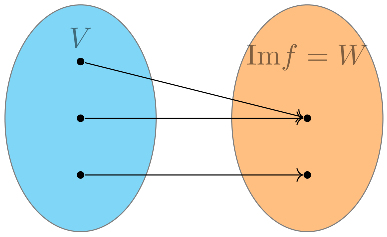
$f \in L(V, W)$ is surjective $\Leftrightarrow \text{Im}(f) = W$.

In other words, f surjective means it covers the entire space W with its outputs.

The surjectivity of f ensures that there are no elements in W that are not mapped to by some element in V.



Property 1 : Injective Vector Space Homomorphism



Property 2 : Surjective Vector Space Homomorphism

Proof Property 1:

($\Rightarrow$)

1. Assume $f \in L(V, W)$ is injective

$$\forall (u, v) \in V^2, \quad f(u) = f(v) \implies u = v.$$

2. Now we fix $u = 0_V$ and because homomorphism preserve identity,

$$\forall v \in V, \quad f(v) = 0_W \implies v = 0_V.$$

3. Therefore, $\text{Ker}(f) = \{0_V\}$

($\Leftarrow$)

1. Assume $\text{Ker}(f) = \{0_V\}$.
2. Consider any $u, v \in V$ such that $f(u) = f(v)$.
3. Then, $f(u) - f(v) = 0_W$.
4. Since f is a linear transformation, $f(u) - f(v) = f(u - v)$.
5. Therefore, $f(u - v) = 0_W$.
6. By our assumption, this implies $u - v \in \text{Ker}(f)$, and hence $u - v = 0_V$.

7. Thus, $u = v$, proving that f is injective.

Proof Property 2:

($\Rightarrow$)

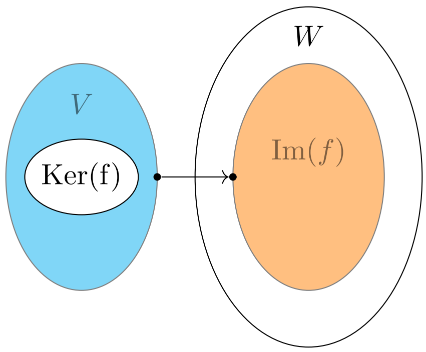
1. Assume $f \in L(V, W)$ is surjective.
2. By definition of surjectivity, $\forall w \in W, \exists v \in V, f(v) = w$.
3. This implies that every element of W is an image of some element in V under f.
4. Therefore, $\text{Im}(f) = W$.

($\Leftarrow$)

1. Assume $\text{Im}(f) = W$.
2. By definition, this means
$$\forall w \in W, \exists v \in V, f(v) = w$$
3. This directly satisfies the definition of surjectivity, where every element in W has a preimage in V under f.
4. Hence, f is surjective.

Rank–Nullity Theorem

Let V and W be vector spaces over $\mathbb{K}$ and $f \in L(V, W)$. Then,
$$\dim(V) = \dim(\text{Ker} f) + \dim(\text{Im} f)$$



$$\dim(\text{Im} f) = \dim(V) - \dim(\text{Ker} f)$$

Proof:

1. Let $\{v_1, v_2, \ldots, v_k\}$ be a basis for $\text{Ker}(f)$. Since $\text{Ker}(f) \subseteq V$, these vectors are in V and are linearly independent.
2. Extend this basis to a basis for V, which includes additional vectors $\{v_{k+1}, v_{k+2}, \ldots, v_n\}$. Thus, the set $\{v_1, v_2, \ldots, v_n\}$ is a basis for V, and $\dim(V) = n$.
3. Consider the images of the basis vectors $\{v_{k+1}, v_{k+2}, \ldots, v_n\}$ under f. These images form a basis for $\text{Im}(f)$ because f is linear and maps linearly independent vectors to linearly independent vectors (except for those in $\text{Ker}(f)$).
4. Therefore, the dimension of $\text{Im}(f)$ is the number of these basis vectors, which is $n - k$.
5. Hence, $\dim(\text{Im} f) = \dim(V) - \dim(\text{Ker} f)$.

Chapter 5

Isomorphism

Chapter 6

Endomorphism (Operator)

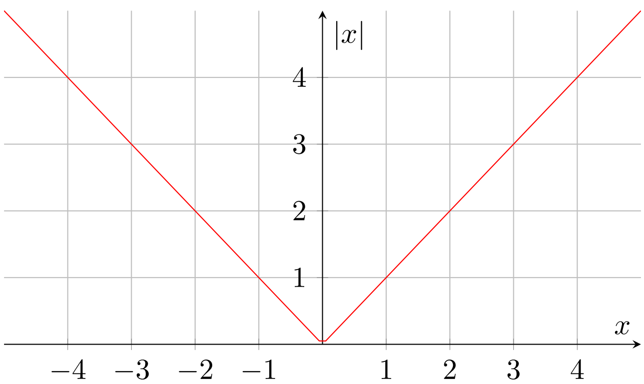
Chapter 7

Normed vector space

7.1 Norm

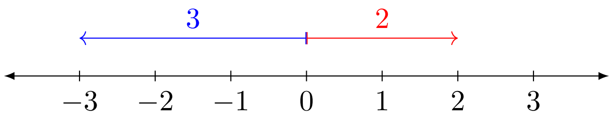
In the initial chapters of this book, we described vectors as entities possessing both magnitude and direction, similar to their representation in physics. However, we have not yet explored the actual measurement of a vector's magnitude or 'length'. A simple way to understand this idea is by thinking about the absolute value $|\cdot|$, which is like measuring how big a number is.

$$|\cdot| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$



Absolute value, at its core, is a measure of magnitude or distance from zero on the number line. For instance, the absolute value of 2 is $|2| = 2$, while that of -3 is $|-3| = 3$.

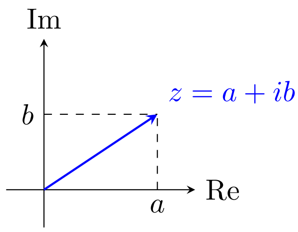
This shows that, in terms of magnitude, -3 is actually greater than 2.



We can generalize absolute value to complex numbers $\mathbb{C}$. With the following definition, the absolute value (or modulus) of a complex number $z = a + ib$ is given by

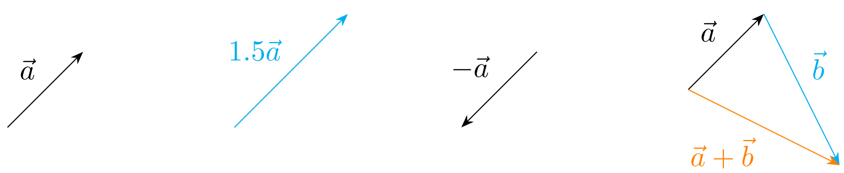
$$|z| = \sqrt{a^2 + b^2}$$

where a and b are real numbers representing the real and imaginary parts of z, respectively. This formula is derived from the Pythagorean theorem and represents the distance of the complex number from the origin in the complex plane.



The limitation of absolute value is that it applies only to numbers, despite $\mathbb{R}$ and $\mathbb{C}$ being vector spaces. To extend this concept to all vector spaces, we introduce the notion of a *vector norm*, which generalizes the idea of absolute value.

One might wonder about the specific equation for a vector norm, similar to that for absolute value. However, vectors can vary significantly in their mathematical structure. Consequently, there isn't a single defining equation for a vector norm. Instead, a function must satisfy a set of properties to be considered a norm. These properties ensure that the function appropriately measures the size or length of a vector in any given vector space.



The properties that define a vector norm include homogeneity, separation, and the triangle inequality. Homogeneity ensures that when scaling a vector by a scalar, the norm scales proportionally. The separation property asserts that the only vector with a norm of zero is the zero vector itself. Lastly, The triangle inequality essentially tells us that the shortest path between two points is a straight line. This means that the direct path from the start of the first vector to the end of the second vector, represented by the sum of the vectors, is always shorter or equal in length to any other path that involves detours or zigzagging.

Norm on a Vector Space

Let V be a vector space over $\mathbb{K}$. A norm on V is any function

$$\mathcal{N} : V \to \mathbb{R}^+$$

satisfying the following three properties:

- **Homogeneity:** $\forall(\lambda, x) \in \mathbb{K} \times V$, $N(\lambda x) = |\lambda|\mathcal{N}(x)$.
- **Separation:** $\forall x \in V$, $\mathcal{N}(x) = 0 \Rightarrow x = 0_V$ (null vector).
- **Triangle inequality:** $\forall(x, y) \in V^2$, $\mathcal{N}(x + y) \leq \mathcal{N}(x) + \mathcal{N}(y)$.

A vector space equipped with a norm is called a *normed vector space* denoted $(V, \|\cdot\|)$. The norm of a vector x in such a space is often denoted by $\|\mathbf{x}\|$.

1. By the properties of **homogeneity**, we have:

$$\| - x\| = | - 1|\|x\| = \|x\|$$

2. By the properties of **homogeneity**, we have:

$$\|0\| = \|0x\| = |0|\|x\| = 0$$

   So the **separation** property could be written with $\iff$ instead of $\Rightarrow$.

3. By the properties of a norm, specifically the **triangle inequality**, we have:

$$\|x - x\| \leq \|x\| + \| - x\| = \|x\| + \|x\|.$$

   This implies $\|0\| \leq 2\|x\|$ and so $0 \leq \|x\|$.

   Therefore, defining the norm as a function to $\mathbb{R}$ instead of $\mathbb{R}^+$ is still valid because **triangle inequality** ensure that the norm is non-negative.

7.2 Banach space

Chapter 8

Pre-Hilbert space

Chapter 9

Hilbert space