Discrete mathematics

Bernadett Aradi

2019 Fall

Information on the course, teaching materials:
https://arato.inf.unideb.hu/aradi.bernadett/discretemath.html

Table of contents

Introduction: sets

- Set, element of a set (notation: $\in$, negation: $\notin$): basic concepts.
- Defining a set: by enumeration, e.g., $\{1, 2, 3\}$,
  or with the help of a defining property T concerning the elements of a given set S in the way $\{x \in S \mid T(x)\}$, e.g.,

  $$\{x \in \mathbb{N} \mid 1 \leq x \leq 5\}.$$

- Emptyset: the unique set, that doesn't have any element.
  Notation: $\emptyset$.
- Notation of the subset relation: $\subset$.
- Two sets are equal or coincide if their elements are the same.
  Equivalently, if they are each others' subsets:

  $$A = B \quad \Longleftrightarrow \quad A \subset B \text{ and } B \subset A.$$

Cardinality of sets, power set

Definition

The power set of a given set S is the set of all subsets of S. Notation: $\mathcal{P}(S)$ or 2^S.

E.g., in the case of $S = \{0, 1, 2, 3\}$:

$$\begin{aligned}\mathcal{P}(S) =&\{\emptyset, \{0\}, \{1\}, \{2\}, \{3\}, \{0,1\}, \{0,2\}, \{0,3\}, \{1,2\}, \{1,3\}, \{2,3\},\\ &\{0,1,2\}, \{0,1,3\}, \{0,2,3\}, \{1,2,3\}, S\}\end{aligned}$$

Definition

If a set has a finite number of elements, then this number is called the cardinality of the set. Notation for a given set S: $\#S$.
In this case we say that S is a finite set.

Theorem

If S has cardinality of n, then the power set of S has cardinality of 2^n, that is

$$\#(\mathcal{P}(S)) = 2^{\#S}.$$

Fundamental operations on sets

- The complement of a set A: $\overline{A}$.
- The union of two sets: $A \cup B$.
- The intersection of two sets: $A \cap B$.
- The (set-theoretic) difference of two sets: $A \setminus B$.
- The symmetric difference of two sets, notation: $\triangle$.

$$A \triangle B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$$

  E.g., if $A = \{0, 1, 2, 3, 4\}$, $B = \{2, 4, 6, 8, 10\}$ what is $A \triangle B =$?

- The Cartesian product of two sets, notation: $\times$.

$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

  E.g., if $A = \{0, 1, 2\}$, $B = \{1, 2\}$ what is $A \times B =$?

Theorem – De Morgan's laws

If A and B are arbitrary sets, then

$$\overline{(A \cup B)} = \overline{A} \cap \overline{B} \quad \text{and} \quad \overline{(A \cap B)} = \overline{A} \cup \overline{B}.$$

Furthermore, these identities hold for arbitrary number of sets.

Notation

Special sets of numbers:

- $\mathbb{N} = \{1, 2, 3, \dots\}$: the set of natural numbers (to be defined later)
- $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$: the set of integers
- $\mathbb{Q}$: the set of rational numbers
- $\mathbb{R}$: the set of real numbers
- $\mathbb{C}$: the set of complex numbers (to be defined later)

Quantifiers:

- $\exists$: 'there exists' (existential quantifier)
- $\forall$: 'for all' (universal quantifier)

E.g., $\exists n \in \mathbb{N} : 2n = 6$, but $\nexists n \in \mathbb{N} : 2n = 7$

$\forall m \in \mathbb{N} : m \in \mathbb{Z}$, but $\not\forall m \in \mathbb{Z} : m \in \mathbb{N}$

Introduction: functions

Function: an association rule, assignment or correspondence $x \mapsto f(x)$

If the function f accomplishes a correspondence between the set D (the domain of the function) and the set R (the range of the function), then we can view the function as pairs $(x, f(x))$, where $x \in D$ and $f(x) \in R$.

$$f\colon D \to R,\ x \mapsto f(x)$$

That is, the function is a subset of the Cartesian product $D \times R$, such that if

$$f\colon x \mapsto y_1 \quad \text{and} \quad f\colon x \mapsto y_2,$$

then necessarily $y_1 = y_2$.

Examples of functions

- $x \in \mathbb{R}$, $x \mapsto f(x) := x^2$
- $x \in \mathbb{R}^+$, $x \mapsto f(x) :=$ {a number with square x}

  Not a function!
- $n \in \mathbb{N}$, $n \mapsto f(n) :=$ {an odd number such that it's a divisor of n}

  Not a function!
- $n \in \mathbb{N}$, $n \mapsto f(n) :=$ {the greatest positive divisor of n}

  Function!

Notation

The meaning of $:=$ is: definition, prescribing a value, 'let it be equal with'

Notation

The meaning of different arrows: $\to$, $\mapsto$, $\Rightarrow$, $\Leftrightarrow$

Basic functions

- constant: $f(x) = c$
- first order (linear): $f(x) = mx + b$
- second order: $f(x) = ax^2 + bx + c \quad (a \neq 0)$
  factored form: $f(x) = a \cdot \left(x - \frac{-b+\sqrt{b^2-4ac}}{2a}\right) \cdot \left(x - \frac{-b-\sqrt{b^2-4ac}}{2a}\right)$
- polynomial
- exponential: $f(x) = a^x \quad (a > 0,\ a \neq 1)$
- logarithmic: $f(x) = \log_a x \quad (a > 0,\ a \neq 1)$
- trigonometric functions
- absolute value function
- sign function or signum function

Properties of functions

Let us consider an arbitrary function

$$f\colon D \to R,\ x \mapsto f(x).$$

Definition

The function f is injective if $f(a) = f(b)$ implies $a = b$.

That is, in this case the function f assigns a *different* value to each element.

Definition

The function f is surjective if for every element y in R there exists an element $x \in D$ such that $f(x) = y$.

That is, f is surjective if all elements of R become an image of an element.

Definition

The function f is bijective if it is injective and surjective.

Reasoning with mathematical induction

Let us assume that we want to prove a proposition (for example, the relation below) for *all natural numbers*:

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2, \quad \forall n \in \mathbb{N}.$$

Then we can use the following reasoning:

(1) We prove the proposition for $n = 1$. (By trial and error.)

(2a) We assume that the proposition is true for an arbitrary natural number k,

(2b) then we prove it for the natural number $k + 1$.

(2a): inductive hypothesis

The set of natural numbers

For the axiomatic introduction of this set we use the so-called Peano axioms.

Definition – Peano axioms

(P1) 1 is a natural number.

(P2) For every natural number n there exists uniquely a successor natural number.

(P3) There is no natural number whose successor is 1.

(P4) If two natural numbers have the same successors, then the two natural numbers coincide.

(P5) Axiom of induction: if A is a set such that

- it contains the natural number 1,
- for every element of A its successor is also in A,

then A contains all the natural numbers.

The conditions (P1)–(P5) uniquely determine a set, which is called the set of natural numbers. Notation: $\mathbb{N}$.

Remarks on the Peano axioms

(P2) For every natural number n it is possible to provide a 'greater by 1' natural number, which is called the successor of n.
$\rightsquigarrow n+1$, $S(n)$ (S: successor function)

(P4) If two natural numbers have the same successors, then the two natural numbers coincide.
In other words: the successor function is injective.

(P5) Axiom of induction: if A is a set such that

- it contains the natural number 1,
- for every element of A its successor is also in A,

then A contains all the natural numbers.
In other words: A is an inductive set. $\Rightarrow \mathbb{N}$ is the smallest inductive set.
Another example for inductive sets: the set of positive numbers ($\mathbb{R}^+$).

The Peano axioms with mathematical formalism

Definition – Peano axioms

Let $\mathbb{N}$ be a set satisfying the following conditions:

(P1) $1 \in \mathbb{N}$

(P2) $\forall n \in \mathbb{N} : \exists S(n) \in \mathbb{N},\ S(n) =: n+1$
or: $\exists S : \mathbb{N} \to \mathbb{N}$ so-called successor function

(P3) $\nexists n \in \mathbb{N} : S(n) = 1$

(P4) $n, m \in \mathbb{N} : S(n) = S(m) \Rightarrow n = m$

(P5)

$$\left.\begin{array}{l} 1 \in A \\ n \in A \Rightarrow S(n) \in A \end{array}\right\} \Longrightarrow \mathbb{N} \subset A$$

Then $\mathbb{N}$ is uniquely determined, and it is called the set of natural numbers.

Proof by induction

Based on the definition the elements of $\mathbb{N}$ are:

$$1,\ S(1),\ S(S(1)),\ S(S(S(1))),\dots,S(S(\dots(S(1))\dots)),\dots$$
$$S(1) = 1 + 1 =: 2$$
$$S(S(1)) = S(1) + 1 =: 3$$

The axiom of induction expresses that all the natural numbers can be given with the help of the special natural number 1 and the successor function S. Thus, if we want to prove a proposition (for example, a relation below) for *all natural numbers*, then we can apply the reasoning of mathematical induction:

(1) We prove the proposition for $n = 1$. (By trial and error.)

(2a) We assume that the proposition is true for an arbitrary natural number k,

(2b) then we prove it for the natural number $k + 1$.

(2a): inductive hypothesis

Examples for proof by induction

1. The sum of the first n natural numbers is $\frac{n(n+1)}{2}$. We can apply induction here. ✓
2. $x + \frac{1}{x} \geq 2,\ \forall x > 0.$ We cannot apply induction for this!
3. Prove that

$$1 + 3 + 5 + \cdots + (2n-1) = n^2, \quad \forall n \in \mathbb{N}.$$

4. Prove that

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}, \quad n \in \mathbb{N}.$$

Notation

$$\sum_{i=1}^{n} \text{sum}, \qquad \prod_{i=1}^{n} \text{product}$$

The set of integers

We introduce the set of integers with the help of the already defined set of natural numbers. All integers can be written as the difference of two natural numbers:

$$\mathbb{Z} = \{n - m \mid n, m \in \mathbb{N}\}.$$

The integers are: classes of these types of differences, e.g.,

- 3 is represented by the class $\{4-1, 5-2, 6-3, \ldots, 72-69, \ldots\}$
- 0 is represented by the class $\{1-1, 2-2, 3-3, \ldots, 51-51, \ldots\}$
- -5 is represented by the class $\{1-6, 2-7, 3-8, \ldots, 100-105, \ldots\}$

Divisors, divisibility

Let $a, b \in \mathbb{Z}$.

Definition

We say that b is a divisor of a, or a is a multiple of b, or a is divisible by b if there exists $c \in \mathbb{Z}$ such that $a = b \cdot c$.
Notation: $b|a$

Theorem – the properties of divisibility

1. $\forall a \neq 0, a \in \mathbb{Z} : a|0,\ 1|a,\ a|a$
2. If $a|b$ and $c \in \mathbb{Z}$, then $a|bc$. $(a|b \wedge c \in \mathbb{Z} \Rightarrow a|bc)$
3. If $a|b_1$ and $a|b_2$, then $a|(b_1 + b_2)$.
4. If $a|b$ and $b|c$, then $a|c$.
5. If $a|b$ and $b|a$, then $a = \pm b$.

2 és 3 $\Rightarrow$ If $a|b_i$, $i = 1, 2, \ldots, n$ and $c_1, c_2, \ldots, c_n \in \mathbb{Z}$, then

$$a|(b_1 c_1 + b_2 c_2 + \cdots + b_n c_n).$$

Divisibility rules

$A \in \mathbb{N} \Rightarrow$

$$A = a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + \cdots + a_2 \cdot 10^2 + a_1 \cdot 10 + a_0,$$

$$a_i \in \{0, 1, \ldots, 9\},\ a_n \neq 0.$$

- Divisibility by 2

$$A = (a_n \cdot 10^{n-1} + a_{n-1} \cdot 10^{n-2} + \cdots + a_2 \cdot 10 + a_1) \cdot 10 + a_0$$

  $2|10$, thus if $2|a_0$, then $2|A$
- Divisibility by 5: $A =$as above
  $5|10$, thus if $5|a_0$, then $5|A$
- Divisibility by 4: $4 \not| 10$, but $4|100$

$$A = (a_n \cdot 10^{n-2} + a_{n-1} \cdot 10^{n-3} + \cdots + a_2) \cdot 100 + a_1 \cdot 10 + a_0$$

  $4|100$, so if $4|(a_1 \cdot 10 + a_0)$, then $4|A$
- Divisibility by 25: analogously to 4, since $25|100$.

Divisibility rules

- Divisibility by 8: $8 \not| 100$, however $8|1000 \Rightarrow$
$$8|A \iff 8|(100a_2 + 10a_1 + a_0)$$
$100a_2 + 10a_1 + a_0$ is the remainder when dividing A by 1000.
- Divisibility by 3 and 9: $10^k - 1 = 99\dots9 \Rightarrow 3|(10^k - 1)$, $9|(10^k - 1)$
$$A = a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + \dots + a_2 \cdot 10^2 + a_1 \cdot 10 + a_0 =$$
$$= a_n(10^n - 1) + a_{n-1}(10^{n-1} - 1) + \dots + a_1(10 - 1) +$$
$$+ a_n + a_{n-1} + \dots + a_1 + a_0$$
$\Rightarrow A$ is divisible by 3 or 9 if the sum of its digits is divisible by 3 or 9
- Divisibility by 11: $10^1 + 1 = 11$, $10^2 - 1 = 99$, $10^3 + 1 = 1001$, $10^4 - 1 = 9999$, ... We can prove that
$$11|(10^k + 1) \text{ if } k \text{ is odd and } 11|(10^k - 1) \text{ if } k \text{ is even.}$$
$$A = a_0 + a_1(10^1 + 1) - a_1 + a_2(10^2 - 1) + a_2 + \dots =$$
$$= (a_1(10^1 + 1) + a_2(10^2 - 1) + \dots) + (a_0 - a_1 + a_2 - a_3 + \dots)$$

$\Rightarrow A$ is divisible by 11 if the alternating sum of its digits is divisible by 11

Definition

We say that $d \in \mathbb{N}$ is the greatest common divisor of the integers a and b

- $d|a$ and $d|b$,
- for all $\bar{d} \in \mathbb{N}$ such that $\bar{d}|a$ and $\bar{d}|b$, the relation $\bar{d}|d$ also holds.

Notation: $d = \gcd(a, b)$.

Furthermore $d \in \mathbb{N}$ is the greatest common divisor of $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ if

- $d|a_i$, $i \in \{1, \ldots, n\}$,
- for every $\bar{d} \in \mathbb{N}$ such that $\bar{d}|a_i$ $(i \in \{1, \ldots, n\})$, the relation $\bar{d}|d$ also holds.

Definition

The integers a and b are called relatively prime or coprime numbers if $\gcd(a, b) = 1$.

Definition

We say that $k \in \mathbb{N}$ is the least common multiple of $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ if

- $a_i|k$, $i \in \{1, \ldots, n\}$,
- for all $\bar{k} \in \mathbb{N}$ such that $a_i|\bar{k}$ $(i \in \{1, \ldots, n\})$, the property $k|\bar{k}$ also holds.

Notation: $k = \mathrm{lcm}(a_1, a_2, \ldots, a_n)$.

The Euclidean algorithm

Theorem – Euclidean division

Given arbitrary $a, b \in \mathbb{Z}$, $b \neq 0$ numbers there uniquely exist integers $q, r \in \mathbb{Z}$ such that

$$a = b \cdot q + r, \quad 0 \leq r < |b|.$$

The Euclidean algorithm (or Euclid's algorithm)

$a, b \in \mathbb{Z}$, $b \neq 0$, theorem above $\Rightarrow q, r \in \mathbb{Z}$, let us denote them by q_0, r_0 this time:

$$a = b \cdot q_0 + r_0$$

Let us repeat the Euclidean division with b and $r_0 \Rightarrow q_1, r_1 \in \mathbb{Z}$, then with r_0 and r_1 ($\Rightarrow q_2, r_2 \in \mathbb{Z}$):

$$b = r_0 \cdot q_1 + r_1$$

$$r_0 = r_1 \cdot q_2 + r_2.$$

By continuing the procedure in this manner (each time with the obtained remainders) we finish in finite steps, since

$$|b| > r_0 > r_1 > r_2 > \cdots > r_i > \cdots \geq 0.$$

Theorem

When applying the Euclidean algorithm for the integers a and $b \neq 0$, the last non-zero remainder is the greatest common divisor of a and b. Furthermore, if $d := \gcd(a, b)$, then the equation

$$ax + by = d$$

can be solved among integers. That is, there exist $x, y \in \mathbb{Z}$ solutions.

Example: $\gcd(1227, 216) = ?$, $\gcd(-1227, -216) = ?$

Definition

Equations of the form $ax + by = c$ (where $a, b, c \in \mathbb{Z}$ are known, $x, y \in \mathbb{Z}$ are unknown) are called linear Diophantine equations.

Theorem

The linear Diophantine equation $ax + by = c$ is solvable if, and only if, $\gcd(a, b) | c$.

Example: Solve the Diophantine equation $147x + 69y = 3$.

Prime numbers

Every $n > 1$, $n \in \mathbb{N}$ has two positive divisors: 1 and n, these are called the trivial divisors of n. All the other divisors are called non-trivial divisors.

Definition

Natural numbers which are greater than 1 and has only trivial divisors are called prime numbers or primes. Natural numbers with also non-trivial divisors are called composite numbers. 1 is a unit.

Theorem

An integer $p > 1$ is prime if, and only if, $p|ab$ implies $p|a$ or $p|b$.

Example: $15|60$

Theorem – the fundamental theorem of arithmetic (also called unique-prime-factorization theorem)

Every natural number greater than 1 is either a prime itself or is the product of prime numbers. Furthermore, this product is unique up to the order of the factors. The obtained unique product is called the canonical representation or the standard form of n, which is $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$, where $p_1, p_2, \dots, p_r$ are pairwise different primes, $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{N}$.

Number of divisors

Theorem

The number of positive divisors of a natural number $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ is

$$d(n) = (\alpha_1 + 1)(\alpha_2 + 1)\dots(\alpha_r + 1).$$

Example: $1{,}455{,}300 = 2^2 \cdot 3^3 \cdot 5^2 \cdot 7^2 \cdot 11$ and $185{,}130 = 2 \cdot 3^2 \cdot 5 \cdot 11^2 \cdot 17$

Theorem

There are infinitely many prime numbers.

Proof: Suppose that there are only finitely many prime numbers, let them be $p_1, p_2, \dots, p_k$. Consider the number $b = p_1 \cdot p_2 \cdot \dots \cdot p_k + 1$. Then $b \neq 1$ and b is a composite number, thus for some index $i \in \{1, 2, \dots, k\}$ we have $p_i | b$. But $p_i | \prod p_j$ as well, thus $p_i | 1$, which is a contradiction.

Remark

The integers a and b are coprime numbers if there are no common prime factors in their canonical representation.

Congruence

Let $a, b \in \mathbb{Z}$, $m \in \mathbb{N}$.

Definition

We say that a and b are congruent modulo m if $m|(a-b)$.
Notation: $a \equiv b \pmod{m}$, m: is the modulus of the congruence.

Example: for $m = 4$ we have $3 \equiv 11 \pmod{4}$

The integers $a, b \in \mathbb{Z}$ are congruent modulo m if they provide the same remainder when divided by m.

Theorem

The congruence modulo m is a so-called equivalence relation:
reflexive, symmetric, transitive.

Definition

Let us consider the class of integers which are congruent with each other modulo m. The obtained classes are called the congruence classes or residue classes modulo m. The residue classes are represented by the integers $0, 1, \ldots, m-1$. Thus, there are m residue classes modulo m.

The properties of congruence

Proposition – the properties of congruence

Let $m \in \mathbb{N}$ $(m \geq 2)$ and $a, b, c, d \in \mathbb{Z}$.

1. If $a \equiv b$ and $c \equiv d \pmod{m}$, then
$$a \pm c \equiv b \pm d \pmod{m} \quad \text{and} \quad a \cdot c \equiv b \cdot d \pmod{m}.$$
2. If $a \cdot c \equiv b \cdot c \pmod{m}$ and $\gcd(c, m) = 1$, then $a \equiv b$.

Example: $15 \equiv 63 \pmod{8}$ and $10 \equiv 18 \pmod{8}$

Definition

Any set of m integers, no two of which are congruent modulo m, is called a complete residue system modulo m. The set of integers $\{0, 1, 2, \ldots, m-1\}$ is called the least residue system modulo m.

Example: for $m = 5$ the set $\{5, 6, 12, 28, 9\}$ is a complete residue system, while $\{0, 1, 2, 3, 4\}$ is the least residue system.

Proposition

If $a \equiv b \pmod{m}$, then $\gcd(a, m) = \gcd(b, m)$.

Reduced residue system

Definition

A residue class is a member of the reduced residue system if its members are coprime to the modulus. Notation: the number of elements of a reduced residue system modulo m is denoted by $\varphi(m)$. That is

$$\varphi(m) = \#\{a \in \{1, \dots, m\} \mid \gcd(a, m) = 1\}.$$

The name of the function φ: Euler's φ function or Euler's totient function.

By definition, $\varphi(1) = 1$.

Examples: cardinality of the reduced residue system:

m	complete	reduced	$\varphi(m)$
$m = 2$	0,1	1	$\varphi(2) = 1$
$m = 3$	0,1,2	1,2	$\varphi(3) = 2$
$m = 4$	0,1,2,3	1,3	$\varphi(4) = 2$
$m = 5$	0,1,2,3,4	1,2,3,4	$\varphi(5) = 4$
$m = 6$	0,1,2,3,4,5	1,5	$\varphi(6) = 2$
$m = 7$	0,1,2,3,4,5,6	1,2,3,4,5,6	$\varphi(7) = 6$

Euler's φ function

Proposition

If p is a prime, then $\varphi(p) = p - 1$.

Theorem

The value of Euler's φ function can be calculated by the formula

$$\varphi(m) = m \cdot \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right),$$

where m has canonical representation $m = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$.

Example: $m = 24$, $\varphi(24) = ?$

Theroem – Euler's theorem

If $\gcd(a, m) = 1$, then $a^{\varphi(m)} \equiv 1 \pmod{m}$.

Corollary – Fermat's little theorem

If p is a prime and $p \nmid a$, then $a^{p-1} \equiv 1 \pmod{p}$.

Example: what is the remainder when dividing 2^{2019} by 15?

Congruence equations

Theorem

The (linear) congruence equation $ax \equiv b \pmod{m}$ is solvable among integers if, and only if, $\gcd(a, m) | b$.

Proof: we can derive a Diophantine equation from the congruence equation:

$$ax \equiv b \pmod{m} \;\Leftrightarrow\; m|(ax-b) \;\Leftrightarrow$$
$$\Leftrightarrow\; \exists y \in \mathbb{Z}:\; my = ax - b \;\Leftrightarrow\; ax - my = b$$

Remark: if $c \in \mathbb{Z}$ is a solution, then so is $c + km$.

Example: $13x \equiv 5 \pmod{29}$

Complex numbers

Looking for solutions of equations in different sets:

- $\mathbb{N}$: $5 + x = 3 \Rightarrow$ not solvable
- $\mathbb{Z}$: $5 \cdot x = 3 \Rightarrow$ not solvable
- $\mathbb{Q}$: $x^2 = 3 \Rightarrow$ not solvable
- $\mathbb{R}$: $x^2 = -3 \Rightarrow$ not solvable

$\rightsquigarrow$ Let's „extend" $\mathbb{R}$ with $\sqrt{-1}$.

Notation, definition

The symbol $i := \sqrt{-1}$ is the imaginary unit.

Definition

Numbers of the form $a + bi$ where $a, b \in \mathbb{R}$ and $i^2 = -1$, are called complex numbers. The set of complex numbers is denoted by $\mathbb{C}$.

Let $z = a + bi \in \mathbb{C}$. This is called the algebraic form of z.

$a = \Re(z)$: real part of z

$b = \Im(z)$: imaginary part of z.

Operations with complex numbers, visual representation

Operations in $\mathbb{C}$

If $z = a + bi$ and $w = c + di$ are complex numbers, then

$$z + w := (a + c) + (b + d)i,$$

$$z \cdot w := (ac - bd) + (ad + bc)i.$$

Visual representation: on the complex plane.

A complex number $z = a + bi$ is uniquely determined by a and b, but by two other values as well:

- the absolute value of z: $r := |z| := \sqrt{a^2 + b^2}$
- the argument of z: φ. We choose this such that it satisfies

$$a = r \cdot \cos\varphi,$$

$$b = r \cdot \sin\varphi.$$

With the help of these we can write z as

$$z = r \cdot (\cos\varphi + i \cdot \sin\varphi),$$

which form is unique if $r > 0$ and $\varphi \in [0, 2\pi[$. This is called the trigonometric form of z.

Operations with the trigonometric form

Examples:

1. What is the algebraic form of the complex number which has absolute value 3 and argument $\frac{\pi}{4}$?
2. What is the trigonometric form of $z = \sqrt{3} - i$?

Definition

The conjugate of $z = a + bi$ is $\bar{z} = a - bi$.

Then $z \cdot \bar{z} = (a + bi)(a - bi) = a^2 + b^2 = |z|^2$.

Let

$$z_1 = r_1 \cdot (\cos\varphi_1 + i \cdot \sin\varphi_1) \quad \text{and} \quad z_2 = r_2 \cdot (\cos\varphi_2 + i \cdot \sin\varphi_2).$$

- Multiplication: $z_1 \cdot z_2 = r_1 r_2 \cdot \big(\cos(\varphi_1 + \varphi_2) + i \cdot \sin(\varphi_1 + \varphi_2)\big)$
- Division: $\frac{z_1}{z_2} = \frac{r_1}{r_2} \cdot \big(\cos(\varphi_1 - \varphi_2) + i \cdot \sin(\varphi_1 - \varphi_2)\big)$
- Powers: if $z = r \cdot (\cos\varphi + i \cdot \sin\varphi)$ and $n \in \mathbb{Z}$, then

$$z^n = r^n \cdot \big(\cos(n\varphi) + i \cdot \sin(n\varphi)\big) \qquad \text{(Moivre's formula).}$$

Example: $z = \sqrt{3} - i$, $z^{60} = ?$

Determining the nth roots in $\mathbb{C}$

$$\sqrt[n]{z} = ?$$

The equation $x^n = z$ (where x is the unknown parameter) has solutions as the nth roots of z, and there are n of them. $\rightsquigarrow w_0, w_1, \ldots, w_{n-1}$
If $w = \varrho \cdot (\cos\psi + i \cdot \sin\psi)$ is an nth root of z, then

$$\begin{aligned} \varrho^n = r \quad &\Rightarrow \quad \varrho = \sqrt[n]{r} \quad \text{(uniquely determined positive real number)},\\ n\psi \approx \varphi \quad &\Rightarrow \quad n\psi = \varphi + 2k\pi. \end{aligned}$$

Theorem – nth roots of a complex number

If $z = r \cdot (\cos\varphi + i \sin\varphi)$ and $n \in \mathbb{N}$, then the equation $x^n = z$ has exactly n solutions, these are

$$w_k = \sqrt[n]{r} \cdot \left(\cos\frac{\varphi + 2k\pi}{n} + i \cdot \sin\frac{\varphi + 2k\pi}{n} \right), \quad k = 0, 1, \ldots, n-1.$$

Example: $z = \sqrt{3} - i$, $\sqrt[3]{z} = ?$

Roots of unity

Definition

The *n*th roots of the (complex) number 1 are called the *n*th roots of unity. Thus, these are the solutions of the equation $x^n = 1$.

Since

$$z = 1 = 1 + 0 \cdot i = \cos 0 + i \cdot \sin 0,$$

and the *n*th roots of a complex number are given by the formula

$$\sqrt[n]{r} \cdot \left(\cos \frac{\varphi + 2k\pi}{n} + i \cdot \sin \frac{\varphi + 2k\pi}{n} \right), \quad k = 0, 1, \ldots, n-1,$$

the *n*th roots of unity are

$$\varepsilon_k = \cos \frac{2k\pi}{n} + i \cdot \sin \frac{2k\pi}{n}, \quad k = 0, 1, \ldots, n-1.$$

Remark: $\forall n \in \mathbb{N}$ we have $\varepsilon_0 = 1$.

Examples: what are the *n*th roots of unity in the cases $n = 2$, $n = 3$ and $n = 4$? Plot them on the complex plane.

Polynomials

Definition

Let x be a so-called indeterminate (or variable, a symbol). The expression

$$p(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

where $a_i \in \mathbb{R}$, is called a polynomial.

- The set of polynomials with real coefficients is denoted by $\mathbb{R}[x]$.
- If $a_n \neq 0$, then n is the degree or order of the polynomial. Notation: $\deg(p) = n$.
- The real numbers a_i are called the coefficients of the polynomial.
- If $p(x) = a_0$, then it is a zero-order or constant polynomial.

Examples:

$$p_1(x) = 3 + 2x + x^4 + 3x^5 \ \rightarrow \ \text{degree 5 polynomial}$$
$$p_2(x) = 2 + x^3 + 3x^4 + 0 \cdot x^5 \ \rightarrow \ \text{degree 4 polynomial}$$

Operations with polynomials

Definition

Let us consider the polynomials

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n \quad \text{and} \quad q(x) = b_0 + b_1 x + \cdots + b_m x^m.$$

- The two polynomials are equal if $n = m$ and $a_i = b_i$, $i = 0, 1, \ldots, n$.
- The sum of the two polynomials if, e.g., $n > m$, is

$$\begin{aligned}(p+q)(x) :=& p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + \\ &+ (a_m + b_m)x^m + a_{m+1}x^{m+1} + \cdots + a_n x^n.\end{aligned}$$

- The product of the two polynomials is

$$\begin{aligned}(p \cdot q)(x) :=& p(x) \cdot q(x) = (a_0 \cdot b_0) + (a_0 b_1 + a_1 b_0)x + \\ &+ (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + \cdots + a_n b_m x^{m+n}.\end{aligned}$$

Thus:

$$\deg(p+q) = \max\{\deg(p), \deg(q)\}$$
$$\deg(p \cdot q) = \deg(p) + \deg(q)$$

Euclidean division for polynomials

Theorem – Euclidean division for polynomials

If

$$p(x) = a_n x^n + \cdots + a_1 x + a_0,$$
$$s(x) = b_m x^m + \cdots + b_1 x + b_0,$$

where $a_n \neq 0$, $b_m \neq 0$ and $m < n$, then there exist uniquely polynomials $q(x)$ and $r(x)$ such that

$$p(x) = s(x) \cdot q(x) + r(x), \quad \deg(q) = n - m, \quad \deg(r) < m = \deg(s).$$

Example: $p(x) = x^4 + 3x^2 - 4$, $s(x) = x^2 + 2x$

Definition

If in the previous theorem $p(x) = s(x) \cdot q(x)$, that is, $r(x) = 0$, then $s(x)$ is a divisor of $p(x)$, which we denote by $s(x) | p(x)$.

Example: $p(x) = x^5 - 3x^4 + 4x + 1$, $s(x) = x^2 + x + 1$

The roots of polynomials

Definition

Let $p(x)$ be a polynomial with real (or complex) coefficients, that is, $p(x) \in \mathbb{R}[x]$ (or $p(x) \in \mathbb{C}[x]$). The number $b \in \mathbb{C}$ is a root or solution of $p(x)$ if $p(b) = 0$.

Remark: If b is a root of $p(x)$, then $(x - b)|p(x)$.

- For second-order polynomials:

$$ax^2 + bx + c = a(x - x_1)(x - x_2)$$

- $p(x) = x^3 + x^2 - 2x - 8$, since $p(2) = 0$, $x_0 = 2$ is a root

Definition

The multiplicity of the root b in the polynomial $p(x)$ is k if $(x - b)^k | p(x)$, but $(x - b)^{k+1} \not| \, p(x)$. If $k = 1$, then b is a simple root, if $k > 1$, then it is a multiple root of $p(x)$.

E.g: What is the multiplicity of $x_0 = 1$ in the polynomial below?

$$p(x) = 2x^5 - 4x^4 + 6x^3 - 14x^2 + 16x - 6$$

The Fundamental Thm of Algebra and its consequences

Theorem – the Fundamental Theorem of Algebra

Let $p(x) \in \mathbb{C}[x]$ be a non-constant polynomial. Then $\exists x_0 \in \mathbb{C} : p(x_0) = 0$, that is, $p(x)$ has a complex root.

Remark: Let $p(x) \in \mathbb{C}[x]$ be a polynomial of degree n ($n \geq 1$) and let $x_0 \in \mathbb{C}$ be a root of it. Then $(x - x_0) | p(x)$, thus:

$$p(x) = (x - x_0) \cdot q(x), \quad \text{ahol } \deg(q) = n - 1.$$

But the Fundamental Thm of Algebra holds also for $q(x)$, so there exists a root $x_1 \in \mathbb{C}$ of it, which implies the form $q(x) = (x - x_1) \cdot q_1(x)$, and that

$$p(x) = (x - x_0) \cdot (x - x_1) \cdot q_1(x); \qquad \ldots$$

Corollaries

- A degree n polynomial with complex coefficients has, counted with multiplicity, exactly n complex roots.
- A degree n polynomial with real coefficients has exactly n complex and at most n real roots.

The factored form of polynomials (over $\mathbb{R}$)

$$p(x) = a_n(x - x_1)^{\alpha_1} \dots (x - x_k)^{\alpha_k} \cdot (x^2 + b_1 x + c_1)^{\beta_1} \dots (x^2 + b_l x + c_l)^{\beta_l}$$

where $(x - x_i)^{\alpha_i}$: linear (or first-order) factors
and $(x^2 + b_j x + c_j)^{\beta_j}$: second-order factors,
they cannot be factorized over $\mathbb{R}$
non-real (complex) roots are here,
which are pairwise conjugate

$$n = \deg(p) = \alpha_1 + \dots + \alpha_k + 2\beta_1 + \dots + 2\beta_l$$

If n is odd, then $\exists \alpha_i \neq 0$, so in this case $x_i \in \mathbb{R}$ is a root.

One more corollary of the Fundamental Thm of Algebra

- If $p(x)$ is an odd degree polynomial with real coefficients, then it has a real root.

The factored form over $\mathbb{C}$:

$$p(x) = a_n(x - x_1)^{\alpha_1} \dots (x - x_j)^{\alpha_j}, \quad \text{where } n = \alpha_1 + \dots + \alpha_j.$$

Horner's method

- Horner's method is an efficient algorithm for the evaluation of a polynomial. Efficient = fewer number of arithmetic operations.
- Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0$.
- Note that

$$p(x) = \big(\ldots((a_n x + a_{n-1}) \cdot x + a_{n-2}) \cdot x + \cdots + a_1\big) \cdot x + a_0.$$

Example.

$$p(x) = 2x^5 - 4x^4 + 6x^3 - 14x^2 + 16x - 6, \qquad p(-2) =?, \quad p(1) =?$$

- Number of arithmetic operations to perform without Horner's method: $n - 1 + n = 2n - 1$ multiplications and n additions
- Number of arithmetic operations to perform with Horner's method: n multiplications, n additions

Theorem

The integer roots of a polynomial with real coefficients divide the constant term of the polynomial.

Combinatorics – Permutation of distinct elements

Definition

Let A be a set with n distinct elements. By a permutation of A we mean an arrangement of all the members of A into some sequence.

Theorem

The number of all permutations of a set of n distinct elements is
$P_n = n! = n(n-1)(n-2)\ldots 2 \cdot 1$.
Other notation: $P(n, n) = n!$.

Examples:

(1) There are 10 participants at a running competition. How many different orders can they finish in?

(2) How many 5 digit numbers can be formed from the digits 3,4,5,7,9, if each digit can be listed only once?

What if we consider the digits 2,2,2,7,7?

Permutation of multisets or ordering with identical items

How many 5 digit numbers can be formed from the digits 2,2,2,7,7?
Solution: If we treat all the 2's and 7's as different numbers, then we had 5! possible orders. But by changing the orders of the 2's, for example, we get the same 5-digit number. $\Rightarrow$ We have to divide 5! by the possible orders of identical elements, so the result is:

$$\frac{5!}{2! \cdot 3!} = 10.$$

Theorem

If we consider n elements of k type, ℓ_1 from the first type, ℓ_2 from the second type, etc. (so $\ell_1 + \ell_2 + \dots + \ell_k = n$), then the number of all permutations of these n elements is

$$P_n^{\ell_1,\dots,\ell_k} = \frac{n!}{\ell_1! \dots \ell_k!}$$

Example: We have 2 red, 1 orange and 3 yellow flowers, which we want to put in our window. How many possibilities do we have for the order of the flowers?

Partial permutations or k-permutations of n

Definition and theorem

In the case of a k-permutation of n or partial permutation we consider arrangements of a fixed length k of elements taken from a given set of size n. Here each element can occur at most once. The number of these arrangements is

$$P(n,k) = nPk = \frac{n!}{(n-k)!} = n \cdot (n-1) \dots (n-k+1).$$

Here necessarily $n \geq k$.

So we choose k elements out of n and arrange them into some order
$\rightsquigarrow$ ordered selections without repetition

Examples: (1) There are 10 participants at a running competition. How many possibilities are there for the podium (that is, for the first 3 places)?
(2) There is a game, where there are 5 different prizes and they choose the winners from 200 participants randomly. How many possibilities are there for choosing the winners if everyone can win at most 1 prize?
What if the participants can be chosen more than once?

Permutations with repetition

Definition and theorem

By permutations with repetition we mean ordered arrangements of the elements of a set of size n of length k where repetition is allowed. These are also called k-tuples. The number of all arrangements of this type is

$$P(n,k)^{\text{rep}} = n^k.$$

So we choose k elements and arrange them into some order, the elements can occur more than once. Thus here $n < k$ is possible as well.

$\rightsquigarrow$ ordered selections with repetition

Examples: (1) In how many ways can one fill a toto coupon? (14 matches, 3 possible results: 1, 2, or X)

(2) How many subsets does a set with n elements have?

The subsets are in a one-to-one correspondence with binary sequences of length n: $100101\ldots110$.

$$P(2,n)^{\text{rep}} = 2^n \text{ possibilities.}$$

Combination without repetition

Definition and theorem

A combination is a way of selecting items from a collection, such that (unlike permutations) the order of selection does not matter. So we choose a subset of k elements of a set with n elements. The number of all ways to select k items out of n elements without regard to order of selection is:

$$C(n,k) = nCk = \frac{n!}{k!(n-k)!} =: \binom{n}{k}.$$

By definition $0! = 1$.
Here necessarily $n \geq k$.

Examples:

(1) Find the number of possible fillings of a lottery coupon (5 numbers from 90).

(2) There is a game, where there are 5 alike prizes and they choose the winners from 200 participants randomly. How many possibilities are there for choosing the winners if everyone can win at most 1 prize?
What if the participants can be chosen more than once?

Combination with repetition

Definition and theorem

A k-combination with repetitions, or k-multicombination, or multisubset of size k from a set S is given by a sequence of k not necessarily distinct elements of S, where order is not taken into account. The number of all ways to select k items out of n elements without regard to order of selection is:

$$C(n,k)^{\text{rep}} = \left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{k}.$$

Here $n < k$ is possible as well.

Examples: (1) In how many ways can we distribute 10 (similar) apples among 4 children?

(2) If we roll 3 (alike) dice, how many possible ways are there for the results (distribution of the thrown numbers)?

Proposition

Let $k, n \in \mathbb{N} \cup \{0\}$, $n \geq k$. Then

$$\binom{n}{k} = \binom{n}{n-k}.$$

Binomial theorem or binomial expansion

Theorem – binomial theorem

Let $x, y \in \mathbb{C}$, $n \in \mathbb{N}$. Then

$$(x+y)^n = \binom{n}{n}x^n + \binom{n}{n-1}x^{n-1}y + \binom{n}{n-2}x^{n-2}y^2 + \\ + \cdots + \binom{n}{1}xy^{n-1} + \binom{n}{0}y^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}.$$

Definition

The expression $\binom{n}{k}$ is called a binomial coefficient.

Proposition

For all $n \in \mathbb{N}$, $0 < k < n$, we have

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Corollaries: Pascal's triangle; number of subsets of a set with n elements.

Linear algebra, vector spaces

Definition

A non-empty set V is called a vector space over $\mathbb{R}$ and the elements of V are called vectors if there are two operations

- vector addition: $V \times V \to V$, $(v, w) \mapsto v + w$,
- scalar multiplication: $\mathbb{R} \times V \to V$, $(\lambda, v) \mapsto \lambda v$,

satisfying the conditions below:

Vector addition:

(a) commutativity, that is $\forall v, w \in V$: $v + w = w + v$;
(b) associativity, that is $\forall u, v, w \in V$: $(u + v) + w = u + (v + w)$;
(c) there exists a zero vector: $0 \in V$, such that $v + 0 = v$ $(\forall v \in V)$;
(d) $\forall v \in V$ there exists a so-called additive inverse,
a vector denoted by $-v$, such that $v + (-v) = 0$.

Scalar multiplication:

(a) $\forall \lambda, \mu \in \mathbb{R},\ v \in V$: $(\lambda + \mu)v = \lambda v + \mu v$;
(b) $\forall \lambda \in \mathbb{R},\ v, w \in V$: $\lambda(v + w) = \lambda v + \lambda w$;
(c) $\forall \lambda, \mu \in \mathbb{R},\ v \in V$: $\lambda(\mu v) = (\lambda\mu)v$;
(d) $\forall v \in V$: $1v = v$.

Examples for vector spaces, linear subspace

Examples:

1. $\mathbb{R}^2$: the vectors of the plane. Elements: ordered pairs (x, y), $x, y \in \mathbb{R}$.
2. $\mathbb{R}^n$, its elements: ordered n-tuples: $(x_1, x_2, \ldots, x_n)$, $x_i \in \mathbb{R}$.

Definition

A non-empty subset W of the vector space V is called a linear subspace (or simply a subspace) of V if it is a vector space itself, that is, W is closed under vector addition and scalar multiplication.

Examples:

1. $\{0\}$ and V are linear subspaces of V, called trivial subspaces.
2. In $\mathbb{R}^2$ the elements of the form $(x, 0)$ constitute a subspace $(x \in \mathbb{R})$.
3. If v is a fixed vector is $\mathbb{R}^2$, then $W = \{\lambda v \in V \mid \lambda \in \mathbb{R}\}$ is a subspace.

Linear combination

Definition

Let $v_1, v_2, \dots, v_n$ be vectors in V. The linear combinations of them are vectors of the form

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n; \qquad \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}.$$

Remark: The zero vector can always be obtained as a linear combination. This is called the trivial linear combination.

Examples:

1. $V = \mathbb{R}^2$, $v = (2,1)$, $w = (0,3)$.
   Which vectors in $\mathbb{R}^2$ can be obtained as linear combinations of v and w?
2. Let us fix a vector $v \neq 0$, linear combinations: vectors of the form λv.

Theorem and definition

Let $v_1, v_2, \dots, v_n$ be vectors in V. Then the set of all linear combinations of these vectors form a linear subspace of V, called the subspace spanned or generated by $v_1, v_2, \dots, v_n$.

Notation: $\mathcal{L}(v_1, \dots, v_n)$.

Linearly dependent and independent vectors

Definition

Let $v_1, v_2, \ldots, v_n$ be vectors in V. We say that these vectors are linearly dependent if there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ *not all 0*, such that

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0.$$

(Thus if the zero vector can be obtained as a non-trivial linear combination of the vectors.) Otherwise we say that the vectors are linearly independent.

Remark: So in case of linear independence the condition

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$$

implies that $\lambda_i = 0$, $\forall i \in \{1, \ldots, n\}$.

Example: $V = \mathbb{R}^2$, $v = (2, 1)$, $w = (0, 3)$, are v and w linearly independent?

Proposition

A set of vectors is linearly dependent if, and only if, some of the vectors can be obtained as a linear combination of the rest of the vectors.

Proposition

Consider a fixed set of vectors in a vector space.

1. If two (or more) of the vectors are the same, then the vectors are linearly dependent.
2. If one of the vectors is a scalar multiple of another vector, then the vectors are linearly dependent.
3. If the zero vector is among the vectors, then the vectors are linearly dependent. That is, a linearly independent set of vectors cannot contain the zero vector.
4. If a subset of the vectors is linearly dependent, then the entire set is linearly dependent.

Basis

Definition

Let $\mathcal{G}$ be a set of vectors of V. We say that $\mathcal{G}$ generates the vector space V if the spanned subspace of $\mathcal{G}$ is the whole vector space. In this case all vectors of V can be obtained as a linear combination of elements of $\mathcal{G}$.

Example: $V = \mathbb{R}^2$, $v = \binom{2}{1}$, $w = \binom{0}{3}$. Then $\{v, w\}$ generates $\mathbb{R}^2$. Let $u = \binom{1}{0}$. Then $\{u, v, w\}$ generates $\mathbb{R}^2$ as well, however, this set is linearly dependent, since $6u - 3v + w = 0$. $\Rightarrow$ A vector of $\mathbb{R}^2$ can be expressed as a linear combination of $\{u, v, w\}$ in more than one ways, e.g.,

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} = v + w = 2u + \frac{4}{3}w.$$

Definition

A basis of V is a linearly independent set of vectors which generate V.

Basis, dimension

Basis: linearly independent set of vectors spanning the whole vector space.

- If $\mathcal{B}$ is a basis, then all elements of V can be uniquely expressed as a linear combination of the elements of $\mathcal{B}$.
- There are infinitely many bases of V.

Theorem and definition

Given a vector space V, all of its bases have the same cardinality (consist of the same number of vectors). This number is called the dimension of the vector space. Notation: $\dim(V)$. Remark: if $V = \{0\}$, then $\dim(V) = 0$.

Examples:

1. $\mathbb{R}^n$: vector space of n-tuples. (Vector addition, scalar multiplication element-wise.) A basis: $\{e_1, e_2, \dots, e_n\}$, it is called the natural (or canonical) basis. $\Rightarrow \dim(\mathbb{R}^n) = n$.
2. $\mathbb{R}^2$: special case of previous example ($n = 2$). Natural basis: $\{e_1, e_2\}$, where $e_1 = \binom{1}{0}$, $e_2 = \binom{0}{1}$. Another basis: $\{v, w\}$, $v = \binom{2}{1}$, $w = \binom{0}{3}$.

Coordinates with respect to a basis

Theorem

If we find n linearly independent vectors in an n-dimensional vector space, then it is a basis.

Definition

Let V be a vector space, $\mathcal{B} = \{b_1, \ldots, b_n\}$ is one of its basis. Then all $v \in V$ can be uniquely expressed as a linear combination of the elements of $\mathcal{B}$, thus there exist unique scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$, such that

$$v = \lambda_1 b_1 + \cdots + \lambda_n b_n.$$

These scalars are called the coordinates of v with respect to the basis $\mathcal{B}$. Then in the basis $\mathcal{B}$ the vector v has the form:

$$v = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

The rank of a set of vectors

Definition

Let $\mathcal{A}$ be a set of vectors. The rank of $\mathcal{A}$ is the dimension of the generated vector space:

$$\mathrm{rank}(\mathcal{A}) = \dim(\mathcal{L}(\mathcal{A})).$$

Example: $V = \mathbb{R}^3$, let $\mathcal{A} = \{u, v, w\}$, where

$$u = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \quad w = \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix}.$$

Since $w = 2u + v$, w is in the subspace spanned by u and v. But u and v are linearly independent, thus $\mathrm{rank}(\mathcal{A}) = 2$.

Remark: Let V be an n-dimensional vector space, $\mathcal{A} = \{v_1, \ldots, v_m\} \subset V$. Then $\mathrm{rank}(\mathcal{A}) \leq n$ and $\mathrm{rank}(\mathcal{A}) \leq m$.

Theorem

The rank of a set of vectors doesn't change if we add the linear combination of some of the vectors to a vector.

Matrices

Definition

A matrix is a rectangular array of numbers. An $m \times n$ (m-by-n) matrix has m rows and n columns.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \qquad \text{elements of } A: \ a_{ij} \qquad A = (a_{ij})$$

The set of all $m \times n$ matrices is denoted by $\mathcal{M}_{m \times n}$.

Definition

- If $n = m$, then the matrix is a square matrix.
- The main diagonal of a matrix is formed by the elements $(a_{11}, a_{22}, a_{33}, \dots)$.
- The identity matrix of size n is the $n \times n$ matrix such that the elements on the main diagonal are equal to 1 and all other elements are zero. Notation: I_n.

Matrix operations

1. Matrix addition

We can only add matrices of the same type.
If $A = (a_{ij})$ and $B = (b_{ij})$ are $m \times n$ matrices, then we calculate the sum entrywise: $C = A + B$, where $c_{ij} = a_{ij} + b_{ij}$; $i = 1, \ldots, m$, $j = 1, \ldots, n$.

2. Scalar multiplication

We do the scalar multiplication entrywise. That is, let $\lambda \in \mathbb{R}$,
$A = (a_{ij}) \in \mathcal{M}_{m \times n}$, $\lambda A = (\lambda a_{ij}) \in \mathcal{M}_{m \times n}$.

3. Matrix multiplication

Let $A = (a_{ij})$ be an $m \times k$ and $B = (b_{ij})$ be a $k \times n$ matrix. Then the product of A and B is the $m \times n$ matrix $C = (c_{ij})$, such that

$$c_{ij} = \sum_{r=1}^{k} a_{ir} b_{rj}.$$

Theorem – properties of matrix multiplication

- If A is an $m \times n$ matrix, then $I_m \cdot A = A$ and $A \cdot I_n = A$.
- If A, B, C are such matrices that AB and BC exist, then $(AB)C = A(BC)$. The matrix multiplication is associative.
- If A and B are of the same size and AC exists, then BC exists as well and $(A + B)C = AC + BC$.
- Matrix multiplication is not commutative, that is, in general $AB \neq BA$.

Definition

Let A be an $m \times n$ matrix. The $n \times m$ matrix, which has rows as the columns of A is denoted by A^T and it is called the transpose of A.

Proposition – properties of transposition

- $(A^T)^T = A$
- Transposition and matrix multiplication: $(AB)^T = B^T \cdot A^T$.

Definition

Let A be a square matrix of order n.

- A is symmetric if $A^T = A$,
- A is skew-symmetric if $A^T = -A$.

Examples:

$$A = \begin{pmatrix} 2 & -3 & 4 \\ -3 & -1 & 7 \\ 4 & 7 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 2 & 1 \\ -2 & 0 & -5 \\ -1 & 5 & 0 \end{pmatrix}$$

Here A is symmetric, B is skew-symmetric.

The inverse of a matrix

Definition

We say that a square matrix A of order n is invertible or that it has an inverse if there exists a square matrix B of order n, such that

$$AB = BA = I_n.$$

Theorem

If A is invertible, then its inverse is uniquely determined. Notation: A^{-1}.

Example:

$$A = \begin{pmatrix} 4 & 3 \\ 7 & 5 \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} -5 & 3 \\ 7 & -4 \end{pmatrix}$$

Proposition – properties of matrix inverse

- If A is invertible, then so is A^{-1}, and $(A^{-1})^{-1} = A$.
- If A and B are invertible and AB exists, then $(AB)^{-1} = B^{-1}A^{-1}$.
- If A is invertible, then so is A^T, and $(A^{-1})^T = (A^T)^{-1}$.

Determinants

Definition

Let $n \in \mathbb{N}$ and let σ denote a permutation of the set $\{1, 2, \ldots, n\}$, that is, let

$$\sigma\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\},\ i \mapsto \sigma(i)$$

be a bijective function. (Here $\sigma(i)$ denotes the ith element in the permutation.) We say that in the permutation σ the elements i and j are in inversion if $i < j$ but $\sigma(i) > \sigma(j)$. The permutation σ is called even if the number of pairs being in inversion in σ is even, and odd if this number is odd.

Examples: $\{1, 2, 3, 4\}$,

$\sigma_1 = (1, 3, 4, 2)$	Number of inversions: 2
$\sigma_2 = (1, 2, 3, 4)$	Number of inversions: 0
$\sigma_3 = (4, 3, 2, 1)$	Number of inversions: 6
$\sigma_4 = (2, 3, 4, 1)$	Number of inversions: 3

Determinants

Definition

Let $A = (a_{ij})$ be a square matrix. Let us choose n elements of A such that we choose exactly one element from each row and each column. The chosen elements:

$$a_{1\sigma(1)}, a_{2\sigma(2)}, \dots, a_{n\sigma(n)}.$$

The determinant of A is

$$\det(A) = |A| = \sum_{\sigma} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}.$$

Here $\varepsilon(\sigma) = \begin{cases} 1, & \text{if } \sigma \text{ is even,} \\ -1, & \text{if } \sigma \text{ is odd.} \end{cases}$

There are $n!$ terms in the sum above.

Example:

1. $n = 2$: $\det(A) = |A| = a_{11}a_{22} - a_{12}a_{21}$.
2. $n = 3$: $\det(A) = \dots$.

Theorem

If A and B are square matrices of the same size, then

$$\det(AB) = \det(A) \cdot \det(B).$$

Determinant of matrices of special form

Proposition

For any $n \in \mathbb{N}$ the determinant of the identity matrix is 1.

$$\det(I_n) = 1$$

Proposition

Let A be an upper triangular matrix, that is, a square matrix with zeros underneath its main diagonal:

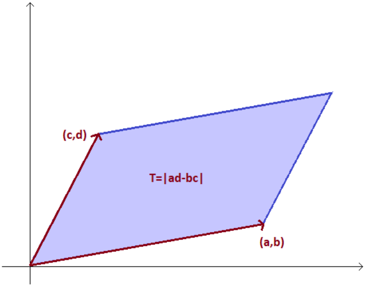
$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Then the determinant of A is the product of the elements in the main diagonal.

Geometric meaning of the determinant

- 2-by-2 determinants: it's absolute value is the area of the parallelogram determined by the rows of the determinant, as vectors

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



- 3-by-3 determinants: it's absolute value is the volume of the parallelepiped determined by the rows of the determinant, as vectors

Proposition – properties of the determinant

- $\det(A) = \det(A^T)$
- If A has a row full of zeros, then $\det(A) = 0$.
- If we interchange 2 rows of A, then the sign of the determinant changes.
- If one row of A is a scalar multiple of another row, then $\det(A) = 0$.
- If we multiply a row of A by a real number λ, then the obtained matrix has determinant $\lambda \cdot \det(A)$.
- If we multiply each row of A by a real number λ, then the obtained matrix has determinant $\lambda^n \cdot \det(A)$.
- The determinant doesn't change if we add a scalar multiple of a row to another row.
- If a row of A is the linear combination of the other rows, then $\det(A) = 0$.
- The properties above are true if we consider columns instead of rows.

Corollary

If $\det(A) \neq 0$, then the rows (or columns) of A are linearly independent vectors. Then is A is of size $n \times n$: its rows form a basis of $\mathbb{R}^n$.

Determinant and matrix inverse

Definition

We say that the square matrix A is regular if $\det(A) \neq 0$.
Otherwise A is said to be singular.

Theorem

A matrix is invertible if, and only if, it is regular. (That is, it's determinant is non-zero.)

Proposition

If A is invertible, then

$$\det(A)^{-1} = \det(A^{-1}).$$

How to calculate the determinant?

1. The rule of Sarrus: only for 2 × 2 and 3 × 3 determinants
2. Gaussian elimination (or row reduction): The modifications below doesn't change the determinant. With the help of them we try to make our determinant to be upper triangular, then the determinant is the product of the elements in the main diagonal.
   - If we multiply the determinant by a non-zero scalar, instead of multiplying all elements of a fixed row by the same scalar.
   - If we add a scalar multiple of a a row to another row.
   - If we interchange two rows, the determinant changes sign.
3. Laplace expansion: We choose an arbitrary row or column of the determinant. E.g., if we choose the i^{th} row, then

$$\det(A) = |A| = \sum_{j=1}^{n} a_{ij} C_{ij}, \quad \text{where}$$

   - C_{ij} is the cofactor of A corresponding to the element a_{ij}, that is,
$$C_{ij} = (-1)^{i+j} A_{ij},$$
   - A_{ij} is the $(n-1) \times (n-1)$ determinant obtained from A by deleting the i^{th} row and j^{th} column of A.

Systems of linear equations

Definition

The system of equations

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

where the real numbers a_{ij} $(i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\})$ and b_k $(k \in \{1, \ldots, m\})$ are known, the variables $x_1, \ldots, x_n$ are unknown, is called a system of linear equations.

- a_{ij}: the coefficients of the system of linear equations
- b_k: the constant terms
- the coefficient matrix and the augmented matrix:

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \quad \text{and} \quad A|b = \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ a_{21} & \ldots & a_{2n} & b_2 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{array}\right)$$

Solvability of systems of linear equations

The corresponding matrix equation: $Ax = b$.

Definition

The system of linear equations is

- solvable if there exists at least one solution, that is, an $x \in \mathbb{R}^n$ such that $Ax = b$ holds;
  - determined if there is exactly 1 solution;
  - undetermined if there are more than 1 solutions;
- overdetermined if it doesn't have a solution.

Definition

The rank of a matrix is the rank of the system of column vectors of the matrix. Notation: $\mathrm{rank}(A)$.

Theorem – condition on solvability

- A system of lin. eq.s is solvable if, and only if $\mathrm{rank}(A) = \mathrm{rank}(A|b)$.
- If it is solvable and $\mathrm{rank}(A) = n$ (where n is the number of unknown parameters), then the system is determined, if $\mathrm{rank}(A) < n$, then undetermined.

Solutions of a system of linear equations

Definition

A system of linear equations is homogeneous if $b = 0$, thus then the matrix equation has the form $Ax = 0$. Otherwise it's called nonhomogeneous.

Remark: 0 is a solution of any homogenous system of linear equations.

Proposition – solutions of a homogeneous system of linear equations

The solutions of a homogeneous system of linear equations form a vector subspace of $\mathbb{R}^n$ with dimension $n - \text{rank}(A)$.

Proposition – solutions of a nonhomogeneous system of linear equations

The solution set of a (solvable) nonhomogeneous system of linear equations $Ax = b$ is of the form $x_0 + H$, where

- x_0 is a particular solution of the system of linear equations;
- H is the solution set of the corresponding homogeneous system of linear equation, that is $Ax = 0$.

Solving a system of linear equations with Gaussian elimination

The set of solutions of a system of linear equations does not change, if we

- multiply an equation by a nonzero constant;
- add a scalar multiple of an equation to another equation;
- interchange two equations;
- discard an equation which is a scalar multiple of another equation.

We annihilate the numbers under the main diagonal with the modifications above. The resulting system is easier to solve.

- If during the process we obtain a row like $(0 \dots 0| \neq 0)$, then the system of linear equations is overdetermined.
- If at the end of the process there are n number of rows, then the system is determined, if fewer number of rows remains, then undetermined. (Here n is the number of the unknown parameters.)

Linear transformations

Definition

Let V be a vector space. $\varphi\colon V \to V$ is a linear transformation if it is

- additive, that is $\forall u, v \in V$: $\varphi(u+v) = \varphi(u) + \varphi(v)$;
- homogeneous, that is $\forall v \in V$, $\lambda \in \mathbb{R}$: $\varphi(\lambda v) = \lambda\varphi(v)$.

Remark: linear transformations map the zero vector to the zero vector.

Examples:

- Rotations, reflections, uniform scaling.
- Projections, e.g., onto a fixed plane of $\mathbb{R}^3$.
- Identity transformation: $\varphi(v) = v$, $\forall v \in V$.

Proposition

A linear transformation is uniquely determined by its action on a basis of V, that is, if $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ is a basis of V, and $w_1, w_2, \ldots, w_n$ are arbitrary vectors, then there uniquely exists a linear transformation φ such that $\varphi(b_i) = w_i$. Furthermore, if $v = \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n$, then its image by φ is

$$\varphi(v) = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n.$$

The matrix of a linear transformation

Definition

Let V be an n-dimensional vector space, $\mathcal{B} = (b_1, b_2, \dots, b_n)$ a basis of V, and consider a linear transformation $\varphi \colon V \to V$.
Then the matrix of φ with respect to $\mathcal{B}$ is the $n \times n$ matrix, such that in its i^{th} column there are the coordinates of $\varphi(b_i)$ with respect to the basis $\mathcal{B}$.

Example: Let $\varphi \colon \mathbb{R}^2 \to \mathbb{R}^2,\ (x, y) \mapsto \varphi(x, y) = (2x - y, -12x + 3y)$.
The matrix of φ in the natural basis. $\varphi(e_1) = \varphi(1, 0) = (2, -12)$, $\varphi(e_2) = \varphi(0, 1) = (-1, 3)$, thus the matrix of φ in this basis is

$$A_\varphi = \begin{pmatrix} 2 & -1 \\ -12 & 3 \end{pmatrix}$$

The matrix of φ w.r.t. the basis $b_1 = (1, 1)$, $b_2 = (0, -1)$. Then $\varphi(b_1) = (1, -9)$ and $\varphi(b_2) = (1, -3)$. These vectors in the basis (b_1, b_2):

$$\varphi(b_1) = (1, -9) = 1 \cdot b_1 + 10 \cdot b_2, \quad \varphi(b_2) = (1, -3) = 1 \cdot b_1 + 4 \cdot b_2.$$

So the sought-for matrix:

$$[A_\varphi]_{(b_1, b_2)} = \begin{pmatrix} 1 & 1 \\ 10 & 4 \end{pmatrix}.$$

Application of the matrix of a linear transformation

Proposition

The determinant and the rank of the matrix of a linear transformation is independent of the chosen basis.

Proposition

If the matrix of φ with respect to the basis $\mathcal{B}$ is A, then $\varphi(v) = Av$.

Examples: Rotations and reflections in $\mathbb{R}^2$ (with respect to the natural basis):

$$\mathrm{rot}_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \qquad \mathrm{refl}_\alpha = \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix}$$

So if we rotate the vector $v = \binom{2}{6}$ by 60° counter-clockwise about the origin:

$$\begin{pmatrix} \cos 60^\circ & -\sin 60^\circ \\ \sin 60^\circ & \cos 60^\circ \end{pmatrix} \begin{pmatrix} 2 \\ 6 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 - 3\sqrt{3} \\ \sqrt{3} + 3 \end{pmatrix}.$$

Eigenvectors and eigenvalues of a linear transformation

Definition

Let $\varphi\colon V \to V$ be a linear transformation. A non-zero vector $v \in V$ is called the eigenvector of φ if $\exists \lambda \in \mathbb{R}$: $\varphi(v) = \lambda v$. Then λ is the eigenvalue of φ associated with v.

Examples: eigenvectors of rotations, reflections, scalings.

Remarks:

- If v is an eigenvector of φ, then the associated eigenvalue is uniqely determined.
- If λ is an eigenvalue, then the corresponding eigenvectors form a vector subspace of V:

  $L_\lambda := \{v \in V \mid \varphi(v) = \lambda v\}$: the eigenspace of φ associated with λ.

Definition and theorem

The characteristic polynomial of φ is the n-degree polynomial $\det(A - \lambda I_n)$, where n is the dimension of V and A is a matrix of φ with respect to *any* basis. Its roots are just the eigenvalues of φ.

Example to determine the eigenvalues and eigenvectors

Determine the eigenvalues and eigenvectors of the linear transformation below.

$$\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2,\ (x,y) \mapsto \varphi(x,y) = (2x-y, -12x+3y)$$

As we have seen, the matrix of φ w.r.t. the natural basis is $\begin{pmatrix} 2 & -1 \\ -12 & 3 \end{pmatrix}$. Thus the characteristic polynomial of φ is

$$\det(A-\lambda I_n) = \left|\begin{pmatrix} 2 & -1 \\ -12 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right| = \left|\begin{pmatrix} 2-\lambda & -1 \\ -12 & 3-\lambda \end{pmatrix}\right|$$
$$= (2-\lambda)(3-\lambda) - (-1)(-12) = \lambda^2 - 5\lambda - 6 = (\lambda+1)(\lambda-6).$$

So the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 6$. The corresponding eigenvectors of $\lambda_2 = 6$:

$$\begin{cases} 2x - y = 6x \\ -12x + 3y = 6y \end{cases} \Rightarrow \begin{cases} -4x - y = 0 \\ -12x - 3y = 0 \end{cases} \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = t \cdot \begin{pmatrix} 1 \\ -4 \end{pmatrix}, t \in \mathbb{R}.$$

Euclidean vector spaces

Definition

Let V be a vector space over $\mathbb{R}$ and assume there exists a map

$$\langle\, ,\, \rangle \colon V \times V \to \mathbb{R}$$

(thus we assign to each pair of vectors v, w a real number denoted by $\langle v, w \rangle$), such that it is

(a) additive in its first variable: $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$;
(b) homogeneous in its first variable: $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$;
(c) symmetric: $\langle w, v \rangle = \langle v, w \rangle$;
(d) positive definite: $\forall v \in V : \langle v, v \rangle \geq 0$, and $(\langle v, v \rangle = 0 \Leftrightarrow v = 0)$.

Then the number $\langle v, w \rangle$ is called the scalar (or inner) product of v and w. The vector space V endowed with the scalar product $\langle\, ,\, \rangle \colon V \times V \to \mathbb{R}$ is called a Euclidean vector space. Notation: $\mathbb{E} = (V, \langle\, ,\, \rangle)$.

- (a)+(b) ⇒ the scalar product is linear in its first variable
- $\ldots$ +(c) ⇒ the scalar product is linear also in its second variable

Scalar product: positive definite symmetric bilinear form.

Examples of Euclidean vector spaces

(1) $V = \mathbb{R}^2$, $|v|$: the length of v

$$\langle v, w\rangle = |v||w|\cos\angle \quad \Rightarrow \langle\, ,\, \rangle \text{ scalar product on } \mathbb{R}^2$$

(2) $V = \mathbb{R}^n$, let us fix a basis.
Consider $v = (v_1, v_2, \dots, v_n)$, $w = (w_1, w_2, \dots, w_n)$.

$$\langle v, w\rangle = v_1w_1 + v_2w_2 + \cdots + v_nw_n \quad \Rightarrow \text{ scalar product over } \mathbb{R}^n$$

In the case of $n = 2$ and the choice of the natural basis we get (1).

(3) $V = \mathbb{R}^3$, let us fix a basis.
Let $v = (v_1, v_2, v_3)$, $w = (w_1, w_2, w_3)$.

$$\langle v, w\rangle = v_1w_1 + 2v_2w_2 + 3v_3w_3 \quad \Rightarrow \text{ scalar product over } \mathbb{R}^3$$

$\Rightarrow$ There are more possible scalar products on a vector space.

Definition

The scalar product in (2) is called the canonical or natural scalar product of $\mathbb{R}^n$.

The norm of vectors

Definition

Let $\mathbb{E} = (V, \langle\, , \,\rangle)$ be a Euclidean vector space. The norm or length of a vector $v \in V$ is

$$\|v\| := \sqrt{\langle v, v\rangle}.$$

Remark.: positive definiteness makes the square root possible.

Example: for the canonical scalar product of $\mathbb{R}^2$: $\|v\| = \sqrt{v_1^2 + v_2^2} = |v|$.

Theorem – properties of the norm

Let $\mathbb{E} = (V, \langle\, , \,\rangle)$ be a Euclidean vector space, $\|\cdot\|$ is the norm derived from the scalar product. Then the following conditions hold:

- $\forall v \in V$: $\|v\| \geq 0$, furthermore $\|v\| = 0 \Leftrightarrow v = 0$;
- $\|\cdot\|$ is absolute homogeneous: $\forall v \in V$ and $\lambda \in \mathbb{R}$: $\|\lambda v\| = |\lambda| \|v\|$;
- it satisfies the triangle inequality: $\forall v, w \in V$: $\|v + w\| \leq \|v\| + \|w\|$. Here we have equality if, and only if, v and w are non-negative scalar multiples of each other.

Theorem – Cauchy–Schwarz inequality

Let v and w be vectors of the Euclidean space $\mathbb{E} = (V, \langle\ ,\ \rangle)$. Then

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|.$$

Equality holds if, and only if, $v = \lambda w$, $\lambda \in \mathbb{R}$.

Example: for the canonical scalar product of $\mathbb{R}^2$:

$$|v_1 w_1 + v_2 w_2| \leq \sqrt{v_1^2 + v_2^2} \cdot \sqrt{w_1^2 + w_2^2}.$$

Definition

Let v and w be non-zero vectors of $\mathbb{E}$. Then the angle of v and w is

$$\arccos \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|}$$

If v or w is the zero vector, then their angle is $\arccos 0 = \frac{\pi}{2}$ by definition.

Remark.: due to the Cauchy–Schwarz inequality

$$-1 \leq \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \leq 1.$$

Orthogonal vectors

Definition

- We say that v and w are orthogonal if $\langle v, w\rangle = 0$. Notation: $v \perp w$.
- A vector $v \in V$ is called a unit vector if $\|v\| = 1$.

Remark.: $\forall v \in V$, $v \neq 0$ we have that $\frac{v}{\|v\|}$ is a unit vector.

Proposition

Let u be a unit vector and $v \in V$ arbitrary. Then the orthogonal projection of v onto u (or the vector component of v in direction u) is $\langle v, u\rangle u$.

Definition

- A set $\{v_1, v_2, \ldots, v_n\}$ of vectors is called orthogonal if it consists of pairwise orthogonal vectors, that is, $\langle v_i, v_j\rangle = 0$ whenever $i \neq j$.
- The set is orthonormal if it consists of pairwise orthogonal unit vectors.

An example

Let us consider the vectors v and w in $\mathbb{R}^2$:

$$v = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

If we choose the canonical scalar product of $\mathbb{R}^2$ then

$$\langle v, w \rangle = -1 \cdot 0 + 1 \cdot (-1) = -1, \quad \|v\| = \sqrt{2}, \quad \|w\| = 1.$$

However, if we choose the scalar product

$$\langle v, w \rangle = 2 v_1 w_1 + v_1 w_2 + v_2 w_1 + v_2 w_2, \quad \text{for } v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \text{ and } w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix},$$

then

$$\langle v, w \rangle = 2 \cdot (-1) \cdot 0 + (-1)(-1) + 1 \cdot 0 + 1 \cdot (-1) = 0, \quad \|v\| = 1, \quad \|w\| = 1.$$

So $\{v, w\}$ form an orthonormal set in $\mathbb{R}^2$ endowed with the latter scalar product.