UCT Mathematics

Václav Mácha

These lecture notes contain mathematical knowledge needed to pass through math exam at the University of Chemistry and Technology. They are released online and they are available for free.
On the other hand, my work on this text is still not finished and the lecture notes will be updated several times during the semester. Thus, this text may contain mistakes. In case you find any, let me know.

Contents

1 Numbers, sets, and functions

1.1 Logic

A *proposition* is such sentence that we can decide about its correctness, i.e., whether it is true or false. For example:
– 'three plus four' is not a proposition,
– 'three plus four is six' is a proposition (obviously wrong). This proposition is atomic (elementary) – it cannot be decomposed. – 'three plus four is seven and one plus one is three' is a proposition as well, however, this proposition is not atomic since it can be decomposed into a proposition 'three plus four is seven', into another proposition 'one plus one is three' and a connective 'and'.

How to make non-atomic propositions
Propositions may be joined into new proposition by using logical connectives:

- conjunction - and - &: 'three plus four is seven and one plus one is three' is an example of a conjunction of two propositions, proposition $A =$'three plus four is seven' and proposition $B =$'one plus one is three'. It may be written as $A\&B$. The whole conjunction is false. Nevertheless, if we replace B by $C =$'one plus one is two', then $A\&C$ will be true – the conjunction is true only if both propositions are true.

- disjunction - or - $\vee$: Using the same notation as above, we understand $A \vee B$ as 'three plus four is seven or one plus one is three'. This time, the proposition $A \vee B$ is true – the disjunction is true once there is at least one true proposition.

- implication - if ... then - $\Rightarrow$: 'if sun shines then it is hot' – here we have two elementary propositions $D =$'sun shines' and $E =$'it is hot'. The implication $D \Rightarrow E$ is false only in case the sun shines and, simultaneously, it is not hot. The implication is true in all other cases.

- equivalence - if and only if - $\Leftrightarrow$

- negation - it is not true that ... - $\neg$: 'it is not true that sun shines', or, with the above notation, $\neg D$. Note that this particular negation might be abbreviated as 'the sun does not shine'. It holds that $\neg\neg A = A$.

The summary is provided by the following table

A	B	$A\&B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$	$\neg A$
true	true	true	true	true	true	false
true	false	false	true	false	false	false
false	true	false	true	true	false	true
false	false	false	false	true	true	true

- Some rules (including the De Morgan laws):
$A \Rightarrow B$ is the same as $(\neg A) \vee B$

$\neg(A \vee B)$ is the same as $(\neg A)\&(\neg B)$
$\neg(A\&B)$ is the same as $(\neq A) \vee (\neg B)$
as a result of the previous lines we deduce that $\neg(A \Rightarrow B)$ is $A\&\neg B$.

Quantifiers

Existential quantifier $\exists$ is read as 'there is' or 'there exists'. For example, 'there exists a natural number n such that $2n = 5$' can be written by use of symbols as $\exists n \in \mathbb{N}, 2n = 5$ (here $\mathbb{N}$ denotes a set of all natural numbers – see the next subsection). We just remark that this statement is false.
Universal quantifier $\forall$ is read as 'for all' or 'every'. For example 'every unicorn can breath under water'. The first two words of this sentence might be shortened to $\forall u \in U$ where U denotes a set of all unicorns. Let us remark that the above statement is true – every proposition is true assuming it tackles all individuals from an empty set, here we tacitly assume that unicorns do not exist.
Example

- $\forall x \in \mathbb{N}\ \exists y \in \mathbb{N}\ 2y = x$ reads as 'For every natural number x there is a natural number y such that $2y = x$. This claim is obviously wrong as taking $x = 5$ would yield $y = 2.5$ which is clearly not natural.
- $\forall x \in \mathbb{X}, \sqrt{x} \in \mathbb{R}$ reads as 'For every real number x, its square root is real' or, in a better way, 'The square root of every real number is real'. Obviously, this claim is false as the square root of a negative number is not real.
- $\exists x \in \mathbb{N}\ 2x = 6$ reads as 'there is a natural number x such that $2x = 6$'. This sentence is valid since the natural number $x = 3$ fits into the claim.
- $\forall x \in \mathbb{R}\ x^2 \in \mathbb{R}$ means 'The square of every real number is real'. This claim is true.

1.2 Sets

The sets are given by one of the following ways:

- list of elements: $M = \{1, 2, 3, 4\}$ is a set containing numbers 1, 2, 3 and 4.
- a condition (or more conditions): $M = \{w,\ w \text{ is a word containing exactly five letters}\}$ or $M = \{w,\ w \text{ is a word containing exactly five letters}, w \text{ is a noun}\}$.
- semantic form: M is a set of first five even natural numbers.

Definition 1.1. *Let X and Y be two sets. By $X \cup Y$ we denote a* union *of sets X and Y which is a set containing elements of both sets, i.e.,*

$$X \cup Y = \{x,\ (x \in X) \vee (x \in Y)\}.$$

By $X \cap Y$ we denote an intersection *of sets X and Y which is a set consisting of elements belonging simultaneously to both sets, i.e.,*

$$X \cap Y = \{x,\ (x \in X)\&(x \in Y)\}.$$

The Cartesian product *$X \times Y$ is a set of all ordered couples such that the first component belongs to X and the second to Y. Namely,*

$$X \times Y = \{\langle x, y\rangle,\ (x \in X)\&(y \in Y)\}.$$

We say that X is a subset *of Y if every element of X is in Y. The notation is $X \subset Y$ and we may write*

$$X \subset Y \Leftrightarrow ((x \in X) \Rightarrow (x \in Y)).$$

Sets X and Y are equal *if $X \subset Y$ and simultaneously $Y \subset X$.*

Let $X \subset Y$. By $Y \setminus X$ we understand a set of all elements in Y which are not in X, i.e.,

$$Y \setminus X = \{(y \in Y)\&(y \notin X)\}.$$

Hereinafter, the empty set is denoted by $\emptyset$.

Example:

- Let $X = \{1, 2, 3\}$ and $Y = \{2, 3, 4\}$. Then $X \cap Y = \{2, 3\}$, $X \cup Y = \{1, 2, 3, 4\}$, $X \setminus \{1\} = \{2, 3\}$ and it holds that $(X \setminus \{1\}) \subset Y$. Further,

$$X \times Y = \{\langle 1, 2\rangle, \langle 1, 3\rangle, \langle 1, 4\rangle, \langle 2, 2\rangle, \langle 2, 3\rangle, \langle 2, 4\rangle, \langle 3, 2\rangle, \langle 3, 3\rangle, \langle 3, 4\rangle\}.$$

1.3 Numbers

We will use the following notation for numbers: $\mathbb{N}$ stands for natural numbers, $\mathbb{Z}$ denotes integers, $\mathbb{Q}$ is a set of all rational numbers and $\mathbb{R}$ denotes the set of all real numbers. Namely:

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

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

$$\mathbb{Q} = \left\{\ldots, \frac{1}{2}, \frac{-3}{2}, \frac{5}{2}, \ldots\right\}$$

$\mathbb{Q}$ is a field. Namely, there are two operations $+$ and $\cdot$ fulfilling

- $\forall x, y, z \in \mathbb{Q}$, $x + (y + z) = (x + y) + z$, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ (associativity)
- $\forall x, y \in \mathbb{Q}$, $x + y = y + x$, $x \cdot y = y \cdot x$ (commutativity)
- $\exists 0 \in \mathbb{Q}$, $\forall x \in \mathbb{Q}$, $x + 0 = x$ (there is null)
- $\exists 1 \in \mathbb{Q}$, $\forall x \in \mathbb{Q}$, $x \cdot 1 = x$ (there is one)
- $\forall x \in \mathbb{Q}$, $\exists -x \in \mathbb{Q}$, $x + (-x) = 0$ (there is an opposite number)
- $\forall x \in \mathbb{Q} \setminus \{0\}$, $\exists x^{-1}$, $x \cdot x^{-1} = 1$ (there is an inverse number)
- $\forall x, y, z \in \mathbb{Q}$, $x \cdot (y + z) = x \cdot y + x \cdot z$ (distributivity)

The set $\mathbb{R}$ is also a field, which is totally ordered and it containes supremum and infimum of every of its subset. Therefore, to properly states basic properties of all real numbers $\mathbb{R}$ we first define a totally ordered set as well as supremum and infimum:

Definition 1.2. *We say that a set X is* totally ordered *if there is a relation $\leq$ fulfilling*

- *$\forall x, y \in X,\ (x \leq y) \vee (y \leq x)$.*
- *$\forall x, y \in X,\ ((x \leq y)\&(y \leq x)) \Rightarrow x = y$.*

- $\forall x, y, z \in X,\ ((x \leq y) \& (y \leq z)) \Rightarrow (x \leq z)$.

Further, we define relation $<$ *as* $x < y \Leftrightarrow (x \leq y \& x \neq y)$.

$\mathbb{R}$ is a totally ordered field which, moreover, satisfies

- $\forall x, y, z \in \mathbb{R}$, $(x < y) \Rightarrow (x + z < y + z)$
- $\forall x, y \in \mathbb{R}$ and $z > 0$, $(x < y) \Rightarrow (z \cdot x < z \cdot y)$

Remark 1.1. *Simply,* $x \geq y$ *is the same as* $y \leq x$ *and* $x > y$ *is the same as* $y < x$.

Definition 1.3. *Let* $A \subset \mathbb{R}$. *We define a* supremum *(or least upper bound, abbreviated as LUB) of* A*,* $\sup A$*, as a number* $M \in \mathbb{R}$ *fulfilling*

$$\forall x \in A,\ (x \leq M) \& (\forall \varepsilon > 0, \exists x \in A,\ x + \varepsilon > M)$$

Similarly, we define infimum *(or greatest lower bound, abbreviated as GLB) of* A*,* $\inf A$*, as a number* $m \in \mathbb{R}$ *fulfilling*

$$\forall x \in A,\ (x \geq m) \& (\forall \varepsilon > 0, \exists x \in A,\ x - \varepsilon < m).$$

Example:

- The supremum of $M = \left\{\frac{1}{n},\ n \in \mathbb{N}\right\}$ is clearly 1: It is contained in M and, simultaneously, it is greater or equal to all the members of M. Further, $\inf M = 0$ because every element of M is greater than zero, on the other hand, for every $\varepsilon > 0$ there is an element $x \in M$ fulfilling $x < 0 + \varepsilon$. Indeed,

$$\frac{1}{n} < \varepsilon$$
$$\frac{1}{\varepsilon} < n$$

  and thus taking $n = \lceil \frac{1}{\varepsilon} \rceil + 1$ gives the demanded element.

Define $\inf \emptyset = +\infty$ and $\sup \emptyset = -\infty$. Previous definition allows to state the last property of real numbers which is: $\forall A \subset \mathbb{R}, \exists M \in \mathbb{R}^*$, $M = \sup A$. That is the way how we get the *extended real numbers* – denoted by $\mathbb{R}^*$ – since we add also numbers $+\infty$ and $-\infty$ – however that was not intended. In order to get the demanded field of numbers we remove $+\infty$ and $-\infty$ as the very last step. Thus $\mathbb{R} = \mathbb{R}^* \setminus \{+\infty, -\infty\}$.

It is worth to mention that rational numbers do not posses this last property. Namely, $\sqrt{2}$ is real number, since it can be defined as

$$\sqrt{2} = \sup\{x, x^2 \leq 2\}.$$

On the other hand, $\sqrt{2}$ is not a rational number. Indeed, let $\sqrt{2} = \frac{p}{q}$ for some $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that p and q do not have a common divisor (and thus it cannot be simplified). Then $(\frac{p}{q})^2 = 2$ which implies $p^2 = 2q^2$ and 2 is a divisor of p which can be written as $p = 2l$ for some $l \in \mathbb{Z}$. We put it into the last equality to get $4l^2 = 2q^2$ yielding $2l^2 = q^2$ and 2 is a divisor of q. Thus p and q have a common divisor 2 which is a contradiction with our assumption.

Definition 1.4. *Let* $a,\ b \in \mathbb{R}^*$*,* $a < b$. *An* open interval (a, b) *is defined as* $(a, b) = \{x \in \mathbb{R}, a < x < b\}$. *Let* $a,\ b \in \mathbb{R}$*,* $a < b$. *A* closed interval $[a, b]$ *is defined as* $[a, b] = \{x \in \mathbb{R}, a \leq x \leq b\}$. *Further, we define* half-open interval *as follows: Let* $a \in \mathbb{R}$ *and* $b \in \mathbb{R}^*$ *be such that* $a < b$. *Then* $[a, b) = \{x \in \mathbb{R}, a \leq x < b\}$. *Let* $a \in \mathbb{R}^*$ *and* $b \in \mathbb{R}$. *Then* $(a, b] = \{x \in \mathbb{R}, a < x \leq b\}$.

1.4 Complex numbers

Definition 1.5. *Let state $i = \sqrt{-1}$. A complex number z is such number that there exists $a, b \in \mathbb{R}$ and $z = a + bi$. The field of complex numbers is denoted by $\mathbb{C}$.*

Operations over $\mathbb{C}$ are as follows. First, let $z_1 = a_1 + b_1 i,\ z_2 = a_2 + b_2 i$ then

- $z_1 + z_2 = a_1 + a_2 + (b_1 + b_2)i$
- $z_1 z_2 = a_1 a_2 - b_1 b_2 + (a_1 b_2 + b_1 a_2)i$
- $\overline{z}_1 = a_1 - b_1 i$, $|z_1|^2 = z_1 \overline{z}_1$
- $\frac{z_1}{z_2} = \frac{z_1 \overline{z}_2}{|z_2|^2}$

Examples:

- $i^2 = -1$, $i^3 = -i$, $i^{29} = i$, $i^{-8} = -1$
- $2 + 3i + 3 - 2i = 5 + i$
- $(1 - 2i)(2 + i) = 4 - 3i$
- $\frac{3-i}{4+i} = \frac{(3-i)(4-i)}{(4+i)(4-i)} = \frac{11-7i}{17}$
- $|3 + 4i| = 5$

1.5 Few words about proofs

A mathematical theorem (lemma, observation) are usually of the form $A \Rightarrow B$ where A denotes the assumptions of a theorem and B denotes the claims of the theorem. The methods of proof of such implication is the following:

- Direct. To prove $A \Rightarrow B$, we present a set of implications which starts from A and end up in B.
  Example: Let $a > 1$, then $a^2 > 1$. Proof: $(a > 1) \Rightarrow (a > 0) \Rightarrow (a^2 > a > 1) \Rightarrow (a^2 > 1)$.
- Indirect. Rather than proving $A \Rightarrow B$, we prove $\neg B \Rightarrow \neg A$.
  Example: Let $a, b \in \mathbb{R}$ and let $ab = 0$. Then either $a = 0$ or $b = 0$. Proof: we show that $(a \neq 0)\&(b \neq 0)$ implies $ab \neq 0$. Let $a > 0$ and $b > 0$. Then $ab > 0$. In other cases we proceed similarly, for example if $a < 0$ and $b > 0$, then we use the previous argument for $-a$ and b.
- Contradiction. Instead of proving $A \Rightarrow B$, we show that $A\&\neg B$ yields contradiction and thus cannot occur. For example a claim $\sqrt{2} \notin \mathbb{Q}$ which was presented in the previous subsection.
- Mathematical Induction – special kind of proof, the rest of this subsection is devoted to this.

Mathematical induction
is a method how to prove an assertion $V(n)$ for every $n \in \mathbb{N}$. (For example, let $n \in \mathbb{N}$ and let $V(n)$ be 'it holds that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$'.)
Math induction helps to prove that $V(n)$ holds for every $n \in \mathbb{N}$. It consists of two steps:

1. $V(1)$ holds.

2. for every $k \in \mathbb{N}$ it holds that $V(k) \Rightarrow V(k+1)$.

Example:
We prove that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$. First, we show the validity of this equality for $n = 1$. In this case we have

$$L = 1 = \frac{1 \cdot 2}{2} = R.$$

To verify the second step, assume that for an arbitrary $k \in \mathbb{N}$ the assertion holds true. We intent to prove that

$$\left(\sum_{i=1}^{k} i = \frac{k(k+1)}{2} \right) \Rightarrow \left(\sum_{i=1}^{k+1} i = \frac{(k+1)(k+2)}{2} \right)$$

To prove the last equality, let start with its left hand side and show it is equal to the right hand side (by using the assumption). We have

$$L = \sum_{i=1}^{k+1} i = \sum_{i=1}^{k} i + k + 1 = \frac{k(k+1)}{2} + k + 1 = \frac{k(k+1)}{2} + \frac{2(k+1)}{2} = \frac{(k+2)(k+1)}{2} = R.$$

2 Linear Algebra

2.1 Vector spaces

Definition 2.1. *A set V endowed with operations $+$ (sum) and $.$ (multiplication by a real number) which satisfy $u + v \in V$ for all $u, v \in V$ and $\alpha.u \in V$ for all $u \in V$ and $\alpha \in \mathbb{R}$ is called* vector space *(or a linear space) if the following properties are true:*

i) $u + v = v + u$ *for all* $u, v \in V$,

ii) $u + (v + w) = (u + v) + w$ *for all* $u, w \in V$,

iii) $\exists 0 \in V$ *for which it holds that* $0 + v = v$ *for all* v,

iv) *for all* v *there is an element* $-v$ *such that* $v + (-v) = 0$,

v) $\alpha.(\beta.v) = (\alpha.\beta).v$ *for all* $\alpha, \beta \in \mathbb{R}$ *and for all* $v \in V$,

vi) $1.v = v$ *for all* $v \in V$,

vii) $(\alpha + \beta).v = \alpha.v + \beta.v$ *for all* $\alpha, \beta \in \mathbb{R}$ *and for all* $v \in V$,

viii) $\alpha.(v + w) = \alpha.v + \alpha.w$ *for all* $\alpha \in \mathbb{R}$ *and for all* $v, w \in V$.

An element of the vector space is called vector.

Remark 2.1 (on notation)**.** *It is customary to denote vectors either by bold letters (i.e.,* $\mathbf{v} \in V$*) or by letters with an arrow (i.e.,* $\vec{v} \in V$*). Hereinafter we use non-bold and non-arrowed letters to denote vectors (i.e.,* $v \in V$*). This does not cause any misunderstandings. In case we work with a group of vectors* $v_i \in \mathbb{R}^n$, $i \in \{1, \ldots, d\}$ *and we need to highlight the* $k-th$ *component, we use* $(v_i)_k$.

Examples:

- The space of ordered pairs of real numbers $(u, v) \in \mathbb{R}^2$ with summation and product defined as
$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2), \quad \alpha(u_1, v_1) = (\alpha u_1, \alpha v_1)$$
for all $(u_1, v_1), (u_2, v_2) \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$ is a vector space.

- In general, all ordered $n-$tuples of real numbers $(u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$ for $n \in \mathbb{N}$ form a vector space.

- The set S of all $(x, y) \in \mathbb{R}^2$ satisfying
$$x + 2y = 0 \tag{1}$$
is a vector space. Since this is a subset of the vector space mentioned above, it is enough to verify that $((x_1, y_1), (x_2, y_2) \in S) \Rightarrow (x_1 + x_2, y_1 + y_2) \in S$ and $(\alpha \in \mathbb{R} \& (x, y) \in S) \Rightarrow (\alpha x, \alpha y) \in S$. So let (x_1, y_1) and (x_2, y_2) satisfy (1). Then $(x_1 + x_2, y_1 + y_2)$ also satisfies (1) since
$$x_1 + x_2 + 2(y_1 + y_2) = x_1 + 2y_1 + x_2 + 2y_2 = 0.$$
Next, let $\alpha \in \mathbb{R}$ be arbitrary number and let (x, y) satisfies (1). Then
$$\alpha x + 2\alpha y = \alpha(x + 2y) = 0$$
and $(\alpha x, \alpha y) \in S$.

- On the other hand, the set S of all pairs $(x, y) \in \mathbb{R}^2$ satisfying
$$x + 2y = 1$$
is not a vector space. For example, a zero vector $(0, 0)$ does not belong to S and the third property from the definition of vector space is not fulfilled.

- The set of polynomials is a vector space.

- The set of polynomials of degree 2 is not a vector space. In particular, a zero polynomial does not belong to this set as the zero polynomial has not degree 2.

- On the other hand, the set of polynomials of degree 0, 1 or 2 is a vector space.

Definition 2.2. *Let V be a vector space and let $S \subset V$ be such that*

i) $\forall s_1, s_2 \in S$, $s_1 + s_2 \in S$ *and*

ii) $\forall \alpha \in \mathbb{R}$ *and* $\forall s \in S$ *we have* $\alpha s \in S$.

Then S itself is a vector space and we say that S is a subspace *of V. If S is nonempty and $S \neq V$ then we will say that S is a* proper subspace.

Examples:

- A subset $S = \{(x, y, 0) \in \mathbb{R}^3\}$ of $V = \mathbb{R}^3$ is a proper subspace.

- All (x, y) solving $x + 2y = 0$ form a subspace of $V = \mathbb{R}^2$ see also one of the previous examples.

Definition 2.3. *Let V be a vector space, $n \in \mathbb{N}$ and $\{u_i\}_{i=1}^n \subset V$. Their* linear combination *is any vector w of the form*

$$w = \sum_{i=1}^{n} \alpha_i u_i$$

where α_i are real numbers.

Examples:

- Consider a vector space $\mathbb{R}^3$. The vector $(2,5,3)$ is a linear combination of $(1,1,0)$ and $(0,1,1)$ because
$$(2,5,3) = 2(1,1,0) + 3(0,1,1).$$
- On the other hand, $(0,-2,1)$ is not a linear combination of $(1,1,0)$ and $(0,1,1)$. Indeed, if it was, then there would be two numbers α and β such that
$$(0,-2,1) = \alpha(1,1,0) + \beta(0,1,1).$$
This equation can be rewritten as a system
$$\begin{aligned} 0 &= \alpha \\ -2 &= \alpha + \beta \\ 1 &= \beta \end{aligned}$$
and we deduce that it is impossible to find α and β such that these equations are fulfilled.

Definition 2.4. *The set of all linear combinations of $v_1,\ v_2, \ldots,\ v_n$ is called a* linear span *of a set $\{v_1, v_2, \ldots, v_n\}$. Precisely,*

$$\text{span}\ \{v_1, v_2, \ldots, v_n\} = \left\{ \sum_{i=1}^{n} \alpha_i v_i,\ \alpha_i \in \mathbb{R} \right\}.$$

Lemma 2.1. *Linear span is a vector space.*

Examples:

- The set $\{(x,y,z) \in \mathbb{R}^3,\ 2x+y+z=0\}$ contains a span of $v_1 = (1,-2,0)$ and $v_2 = (0,1,1)$ (or, for example, $w_1 = (1,0,2)$ and $w_2 = (1,1,3)$).
- Exercise: try to prove that $\{(x,y,z) \in \mathbb{R}^3,\ 2x+y+z=0\} = \text{span}\{(1,-2,0),\ (0,1,1)\}$.

Definition 2.5. *Vectors $v_1, v_2, \ldots, v_n \in V$ are said to be* linearly dependent *if the equation*

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$$

has a nontrivial solution (i.e. a solution $\alpha_1,\ \alpha_2, \ldots,\ \alpha_n$ where at least one coefficient is zero).
Vectors $v_1, v_2, \ldots, v_n \in V$ are linearly independent *if the equation*

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$$

has only solution $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$.

Examples:

- Vectors $(1, 0)$, $(0, 1)$ and $(-2, 3)$ are linearly dependent since

$$2.(1, 0) + (-3).(0, 1) + 1.(-2, 3) = (0, 0).$$

- Vectors $(1, 1, 0)$, $(2, 2, 0)$ and $(-1, 0, 1)$ are linearly dependent since

$$16.(1, 1, 0) + (-8).(2, 2, 0) + 0.(-1, 0, 1) = (0, 0, 0).$$

- Vectors $(2, 3, 1, 0)$, $(1, 0, -1, 0)$ and $(0, 1, 0, -1)$ are linearly independent. Indeed, the equation

$$\alpha(2, 3, 1, 0) + \beta(1, 0, -1, 0) + \gamma(0, 1, 0, -1) = (0, 0, 0, 0)$$

  necessarily yields $\alpha = \beta = \gamma = 0$.

Definition 2.6. *Let $V = \text{span}\,\{v_1, v_2, \dots, v_n\}$. Then we say that $\{v_1, v_2, \dots, v_n\}$* generates *V and the vectors $\{v_1, v_2, \dots, v_n\}$ are* generators *of V.*

Observation 2.1. *Let $v_1, v_2, \dots, v_n$ be linearly dependent. Then one of the vectors is a linear combination of the remaining vectors. Precisely, there is $i \in \{1, \dots, n\}$ such that $v_i \in \text{span}\,\{\{v_1, v_2, \dots, v_n\} \setminus \{v_i\}\}$.*

Proof. According to assumptions, there is $i \in \{1, \dots, n\}$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$$

has a solution with $\alpha_i \neq 0$. Assume, without lost of generality, that $i = 1$. We may rearrange the equation as

$$v_1 = -\frac{\alpha_2}{\alpha_1} v_2 - \frac{\alpha_3}{\alpha_1} v_3 - \ldots - \frac{\alpha_n}{\alpha_1} v_n.$$

□

Corollary 2.1. *Let $v_1 \in \text{span}\,\{v_2, \dots, v_n\}$. Then*

$$\text{span}\,\{v_2, \dots, v_n\} = \text{span}\,\{v_1, v_2, \dots, v_n\}.$$

Proof. Clearly, $\text{span}\,\{v_2, \dots, v_n\} \subset \text{span}\,\{v_1, v_2, \dots, v_n\}$. Next, let

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Since $v_1 = \sum_{i=2}^{n} \beta_i v_i$ for some $\beta_i \in \mathbb{R}$, we get

$$v = \sum_{i=2}^{n} (\alpha_i + \alpha_1 \beta_i) v_i$$

and $v \in \text{span}\,\{v_2, \dots, v_n\}$. □

Definition 2.7. *Let $\{v_1, \dots, v_n\}$ be a set of linearly independent vectors that generates V. Then $\{v_1, \dots, v_n\}$ is a* basis *of V.*

Theorem 2.1. *Every two basis of a vector space V has the same number of elements.*

Definition 2.8. *We say that V is of dimension $n \in \mathbb{N}$ iff every basis has n elements.*

Examples:

- The set $\{(1,0),(0,1)\} \subset \mathbb{R}^2$ is a basis. Indeed, every vector $(a,b) \in \mathbb{R}^2$ can be written as $a(1,0)+b(0,1)$. Moreover, the vectors are linearly independent since $\alpha_1(1,0)+\alpha_2(0,1)=0$ has only the trivial solution. Thus, the dimension of $\mathbb{R}^2$ is 2.

- Vectors $\{1,x,x^2\}$ form a basis of a vector space containing polynomials of degree at most two. The dimension of this vector space is thus 3.

Definition 2.9. *Let $\{v_i, i=1,\dots,n\}$ be independent vectors and let $v \in span\{v_i, i=1,\dots,n\}$. Then the numbers $\alpha_i,\ i=1,\dots,n$ satisfying*

$$v = \sum_{i=1}^{n} \alpha_i v_i$$

are determined uniquely and they are called **coordinates** *of v with respect to the given basis.*

Examples

- The coordinates of $(0,1)$ with respect to $(3,2)$ and $(4,3)$ are $(-4,3)$. Indeed, $-4(3,2)+3(4,3)=(0,1)$.

- The coordinates of $P(x)=x^2+3x+4$ with respect to $Q(x)=x^2+2$ and $R(x)=\frac{3}{2}x+1$ are $(1,2)$.

2.2 Matrices

Definition 2.10. *A* matrix *is a table of numbers arranged in rows and columns. Namely, let m,n be natural numbers. Then*

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

The matrix A has $m-$rows and $n-$columns. The matrix A is said to be of type (m,n).

Example A matrix

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

has two rows and three columns and it is of type $(2,3)$ (or it is of type two by three).

Operations with matrices Let $A=(a_{ij})_{i=1,j=1}^{m,n}$ and $B=(b_{ij})_{i=1,j=1}^{m,n}$ be two matrices of the same type. Then we define

$$A+B=(a_{ij}+b_{ij})_{i=1,j=1}^{m,n}.$$

Let $\alpha \in \mathbb{R}$. Then $\alpha A=(\alpha a_{ij})_{i=1,j=1}^{m,n}$.
For a matrix $A=(a_{ij})_{i=1,j=1}^{m,n}$ we define a transpose matrix A^T as

$$A^T=(a_{ji})_{j=1,i=1}^{n,m}$$

Let A be of type (m,n) and B be of type (n,p). Then $C := AB$ of type (m,p) is defined as

$$C = (c_{ij})_{i=1,j=1}^{m,p}$$

where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Example

- $$\begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 0 & 1 & -2 \end{pmatrix} + \begin{pmatrix} 2 & 2 & 2 & -5 \\ 1 & 1 & -3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 4 & -5 \\ 1 & 1 & -2 & 2 \end{pmatrix}.$$

- $$3 \begin{pmatrix} 1 & \frac{1}{2} \\ 2 & 2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 3 & \frac{3}{2} \\ 6 & 6 \\ -9 & 3 \end{pmatrix}$$

- $$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

  or

  $$\begin{pmatrix} 1 & 1 & 3 \\ 2 & -1 & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ 1 & -1 \\ 3 & 1 \end{pmatrix}$$

  or

  $$\begin{pmatrix} 3 & -1 & -1 & 0 \end{pmatrix}^T = \begin{pmatrix} 3 \\ -1 \\ -1 \\ 0 \end{pmatrix}$$

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

Remark 2.2. *Matrices of a given type (m,n) forms a vector space of dimension nm.*

Remark 2.3. *Warning:*

$$AB \neq BA.$$

Definition 2.11. *A matric A is called* **symmetric** *if $A = A^T$.*

Definition 2.12. *A* rank of matrix *A is a dimension of vector space generated by its rows. It is denoted by rank$A$.*

Observation 2.2. *It holds that rank$A$ = rankA^T.*

Definition 2.13. *An* **elementary transformation** *of a matrix is*

- *scaling the entire row with a nonzero real number or*
- *interchanging two rows within a matrix or*
- *adding $\alpha-$multiple of one row to another for an arbitrary $\alpha \in \mathbb{R}$.*

Let A arise from B by one or more elementary transformations. Then we write $A \sim B$.

Example

$$\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \sim \begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix} \sim \begin{pmatrix} -2 & 4 \\ 2 & 1 \end{pmatrix} \sim \begin{pmatrix} -2 & 4 \\ 6 & -7 \end{pmatrix}.$$

Definition 2.14. *A* leading coefficient *of a row is the first non-zero coefficient in that row. We say that a matrix A is in an* echelon form *if the leading coefficient (also called a* pivot*) of a nonzero row is always strictly to the right of the leading coefficient of the row above it.*

Example Consider the following matrices:

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

The matrix A is in echelon form whereas the matrix B is not in echelon form.

Observation 2.3. *Let A be in echelon form. Then its rank is equal to the number of non-zero rows.*

Proof. Let $v_1, \ldots, v_n$ denote the non-zero rows. It suffices to show that these vectors are linearly independent. Let solve the equation

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0. \tag{2}$$

Let $p_1 \in \mathbb{N}$ be the position of the leading coefficient of v_1. Then the above equation yields

$$\alpha_1 (v_1)_{p_1} = 0$$

and $\alpha_1 = 0$. Therefore, the equation is simplified to

$$\alpha_2 v_2 + \alpha_3 v_3 + \ldots + \alpha_n v_n = 0.$$

Similarly as above, let $p_2 \in \mathbb{N}$ be the position of the leading coefficient of v_2. Then we deduce

$$\alpha_2 (v_2)_{p_2} = 0$$

and $\alpha_2 = 0$. The same can be deduced for every $\alpha_i,\ i \in \mathbb{N}$ and, consequently, there is only a trivial solution to (2) ☐

The Gauss elimination method

The Gauss elimination method is a sequence of elementary transformations which transform a given matrix A into an echelon form. As an example, we take a matrix

$$A = \begin{pmatrix} 2 & 2 & -2 \\ 4 & 1 & 0 \\ 5 & 2 & -1 \end{pmatrix}.$$

In the first step, we use elementary transformations in order to get rid of 4 in the second row and 5 in the last row. So we add (-1) times the first row to the second and $-5/2$ times the first row to the last one. We get

$$A \sim \begin{pmatrix} 2 & 2 & -2 \\ 0 & -3 & 4 \\ 0 & -3 & 4 \end{pmatrix}.$$

Next, we want to eliminate the second element in the last row. In order to do so, we add (-1) times the second row to the last one to get

$$\begin{pmatrix} 2 & 2 & -2 \\ 0 & -3 & 4 \\ 0 & -3 & 4 \end{pmatrix} \sim \begin{pmatrix} 2 & 2 & -2 \\ 0 & -3 & 4 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 2 & 2 & -2 \\ 0 & -3 & 4 \end{pmatrix}.$$

Here we use the fact that the zero row can be omitted without any serious consequence.

Notice that A has a rank two and that means that the vectors $(2,2,-2)$, $(4,1,0)$ and $(5,2,-1)$ are linearly dependent.

2.3 Systems of linear equations

Systems of equations

We are going to deal with system of m linear equations with n unknowns $x_1, \ x_2, \dots, \ x_n$.

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

We use notation $x = (x_1, x_2, \dots, x_n)$, $b = (b_1, b_2, \dots, b_n)$ and $A = (a_{ij})_{i=1,j=1}^{mn}$. Then the above system may be rewritten as

$$Ax^T = b^T.$$

The system of equations will be represented by an augmented matrix – i.e. a matrix $(A|b^T)$ where $A = (a_{i,j})_{i=1,j=1}^{mn}$ and b^T is the column on the right hand side. For example, a system of equations

$$\begin{aligned} 2x + 5y &= 10 \\ 3x + 4y &= 24 \end{aligned}$$

is represented by an augmented matrix

$$\begin{pmatrix} 2 & 5 & | & 10 \\ 3 & 4 & | & 24 \end{pmatrix}.$$

Such matrix consists of two parts – matrix $A = \begin{pmatrix} 2 & 5 \\ 3 & 4 \end{pmatrix}$ and a vector of right hand side $b = (10, 24)$. Let solve the system by Gauss elimination:

$$\begin{pmatrix} 2 & 5 & | & 10 \\ 3 & 4 & | & 24 \end{pmatrix} \sim \begin{pmatrix} 6 & 15 & | & 30 \\ 3 & 4 & | & 24 \end{pmatrix} \sim \begin{pmatrix} 6 & 15 & | & 30 \\ 6 & 8 & | & 48 \end{pmatrix} \sim \begin{pmatrix} 6 & 15 & | & 30 \\ 0 & -7 & | & 18 \end{pmatrix}$$

The last row of the last matrix represent an equation

$$-7y = 18 \ \Rightarrow \ y = -\frac{18}{7}.$$

The first row of the last matrix represent

$$6x + 15y = 30$$

and once we plug there $y = -\frac{18}{7}$ we deduce

$$x = \frac{80}{7}.$$

Theorem 2.2 (Frobenius)**.** *A system of linear equations has solution if and only if* $rank A = rank(A|b^T)$.

Example: Solve

$$\begin{aligned} -x+y+z &= 0 \\ 2y+x+z &= 1 \\ 2z+3y &= 2. \end{aligned}$$

We have

$$\begin{pmatrix} -1 & 1 & 1 & | & 0 \\ 1 & 2 & 1 & | & 1 \\ 0 & 3 & 2 & | & 2 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 1 & | & 0 \\ 0 & 3 & 2 & | & 1 \\ 0 & 3 & 2 & | & 2 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 1 & | & 0 \\ 0 & 3 & 2 & | & 1 \\ 0 & 0 & 0 & | & 1 \end{pmatrix}$$

and, according to the Frobenius theorem, there is no solution to the given system. Let us emphasize that the last row represents an equation

$$0x+0y+0z = 1.$$

Example Let find all solutions to the system

$$\begin{aligned} 2x+y-z &= 3 \\ x-2y+3z &= -1 \end{aligned}$$

We use the Gauss elimination in order to deduce

$$\begin{pmatrix} 2 & 1 & -1 & | & 3 \\ 1 & -2 & 3 & | & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 3 & | & -1 \\ 2 & 1 & -1 & | & 3 \end{pmatrix} \sim \begin{pmatrix} \color{red}{1} & -2 & 3 & | & -1 \\ 0 & \color{red}{5} & -7 & | & 5 \end{pmatrix}$$

The red terms are the leading terms. The corresponding unknowns should be expressed by others. The unknown which does not have a corresponding leading term should be chosen as a parameter. Here we take $z = t$ where $t \in \mathbb{R}$ is a parameter. The last row of the last matrix yields $5y - 7t = 5$ and thus $y = \frac{7}{5}t + 1$. We deduce from the first row that $x = 1 - \frac{1}{5}t$. All solutions are of the form

$$(x,y,z) = (1,1,0) + t\left(-\frac{1}{5}, \frac{7}{5}, 1\right).$$

Exercise

- Solve

$$\begin{aligned} -x+py+pz &= 1 \\ x+y+pz &= 2 \\ px+y+2pz &= 5-2x \end{aligned}$$

  where p is a real parameter.

2.4 Square matrices

Definition 2.15. *Matrices of type* (n,n) *where* $n \in \mathbb{N}$ *are called* square matrices.

Definition 2.16. *A matrix* I *of type* (n,n) *is called an* identity matrix *if* $I = (a_{ij})_{i=1,j=1}^{nn}$, $a_{ii} = 1$ *for all* $i \in \{1,\ldots,n\}$ *and* $a_{ij} = 0$ *whenever* $i \neq j$.

For example,

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

for $n = 3$. It holds that $AI = IA = A$ for every matrix A of type (n, n).

Definition 2.17. *Let A by a matrix of type (n, n). If there is a matrix B of type (n, n) such that*

$$AB = BA = I$$

then B will be called an inverse matrix *to A and we use notation $B = A^{-1}$.*

The Gauss elimination may be used to find A^{-1}. In particular, one has to write down an augmented matrix $(A|I)$ and use elementary transformations to get (I, B). If this is possible, then $B = A^{-1}$.

Example Find A^{-1} to $A = \begin{pmatrix} 2 & -1 \\ 3 & -3 \end{pmatrix}$:

$$\begin{pmatrix} 2 & -1 & | & 1 & 0 \\ 3 & -3 & | & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 & | & 1 & 0 \\ 1 & -2 & | & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & | & -1 & 1 \\ 2 & -1 & | & 1 & 0 \end{pmatrix}$$
$$\sim \begin{pmatrix} 1 & -2 & | & -1 & 1 \\ 0 & 3 & | & 3 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & | & -1 & 1 \\ 0 & 1 & | & 1 & -\frac{2}{3} \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & | & 1 & -\frac{1}{3} \\ 0 & 1 & | & 1 & -\frac{2}{3} \end{pmatrix}$$

Consequently, $A^{-1} = \begin{pmatrix} 1 & -\frac{1}{3} \\ 1 & -\frac{2}{3} \end{pmatrix}$.

Definition 2.18. *A* square matrix *is a matrix of type (n, n) for some $n \in \mathbb{N}$.*
A square matrix A is called regular *if there is A^{-1}. Otherwise it is called* singular.

Observation 2.4. *Let A be a regular matrix. Then a system $Ax^T = b^T$ has a unique solution.*

Proof. Indeed, it suffices to apply A^{-1} from the left side on both sides of equation

$$Ax^T = b^T$$

to obtain

$$x^T = A^{-1}b^T.$$

□

Example The above proof describes another way how to solve a system of equations. Namely, we can first find A^{-1} and then $x^T = A^{-1}b^T$. Let solve the following two systems

$$\begin{aligned} 2x + y + z &= 3 \\ x + 3z &= -7 \\ 2x + y &= 1 \end{aligned}$$

and

$$\begin{aligned} 2x + y + z &= 0 \\ x + 3z &= 3 \\ 2x + y &= -1. \end{aligned}$$

Note that the matrix A of the systems (without the right hand side) is always the same. We compute A^{-1} as follows

$$\begin{pmatrix} 2 & 1 & 1 & | & 1 & 0 & 0 \\ 1 & 0 & 3 & | & 0 & 1 & 0 \\ 2 & 1 & 0 & | & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 3 & | & 0 & 1 & 0 \\ 2 & 1 & 1 & | & 1 & 0 & 0 \\ 2 & 1 & 0 & | & 0 & 0 & 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 3 & | & 0 & 1 & 0 \\ 0 & 1 & -5 & | & 1 & -2 & 0 \\ 0 & 1 & -6 & | & 0 & -2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 3 & | & 0 & 1 & 0 \\ 0 & 1 & -5 & | & 1 & -2 & 0 \\ 0 & 0 & -1 & | & -1 & 0 & 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 3 & | & 0 & 1 & 0 \\ 0 & 1 & -5 & | & 1 & -2 & 0 \\ 0 & 0 & 1 & | & 1 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & | & -3 & 1 & 3 \\ 0 & 1 & 0 & | & 6 & -2 & -5 \\ 0 & 0 & 1 & | & 1 & 0 & -1 \end{pmatrix}$$

Thus, the first system has a solution

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3 & 1 & 3 \\ 6 & -2 & -5 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ -7 \\ 1 \end{pmatrix} = \begin{pmatrix} -13 \\ 27 \\ 2 \end{pmatrix}$$

and the second system has a solution

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3 & 1 & 3 \\ 6 & -2 & -5 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

2.5 Determinant

Definition 2.19. *Let A be a square matrix of type $(1,1)$ – i.e., $A = (a)$ for some $a \in \mathbb{R}$. The* determinant *of such matrix A is $\det A = a$.*
Let $A = (a_{i,j})$ be a square matrix of type (n,n). We denote by M_{ij} the determinant of a matrix $(n-1, n-1)$ which arises from A by leaving out the $i-$th row and $j-$th column. Choose $k \in \{1, \ldots, n\}$. Then

$$\det A = (-1)^{k+1} a_{k1} M_{k1} + (-1)^{k+2} a_{k2} M_{k2} + \ldots + (-1)^{k+n} a_{kn} M_{kn} = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} M_{kj}.$$

Examples:
Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Then $\det A = a_{11}a_{22} - a_{12}a_{21}$.
Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Then

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Observation 2.5. *Let A be a square matrix. Then*

- *if B arises from A by multiplying one row by a real number α, then* $\det B = \alpha \det A$.
- *If B arises from A by switching two rows, then* $\det B = -\det A$.
- *If B arises from A by adding $\alpha-$multiple of one row to another one, then* $\det B = \det A$.

Observation 2.6. *Let A be a square matrix having zeros under the main diagonal (i.e., $a_{ij} = 0$ for $i > j$). Then* $\det A = a_{11}a_{22}a_{33}\dots a_{nn}$.

Example Compute $\det A$ for

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

According to the rules for transformations, we have

$$\begin{aligned} \det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & 3 & -3 & 1 \\ 2 & -3 & 0 & 2 \\ 0 & 0 & 3 & -1 \end{pmatrix} &= \det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & 3 & -3 & 1 \\ 0 & -1 & 0 & 6 \\ 0 & 0 & 3 & -1 \end{pmatrix} \\ &= -\det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & -1 & 0 & 6 \\ 0 & 3 & -3 & 1 \\ 0 & 0 & 3 & -1 \end{pmatrix} = -\det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & -1 & 0 & 6 \\ 0 & 0 & -3 & 19 \\ 0 & 0 & 3 & -1 \end{pmatrix} \\ &= -\det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & -1 & 0 & 6 \\ 0 & 0 & -3 & 19 \\ 0 & 0 & 0 & 18 \end{pmatrix} = 54. \end{aligned}$$

Theorem 2.3. *Let A be $n \times n$ matrix. Statements following are equivalent:*

- $\det A = 0$.
- $Ax^T = 0$ *has a nontrivial solution.*
- *A is a singular matrix matrix.*
- $rank A = n$.
- *Rows of A are linearly dependent vectors.*
- *Columns of A are linearly dependent vectors.*

Theorem 2.4 (the Cramer rule)**.** *Consider a system $Ax^T = b^T$. Assume that A is a regular n by n matrix. Let $j \in \{1, \dots, n\}$ and denote by A_j a matrix arising from A by replacing $j-$th column by a vector b^T. Then*

$$x_j = \frac{\det A_j}{\det A}.$$

Example We use the Cramer rule to solve

$$\begin{aligned}3x-2y+4z&=3\\-2x+5y+z&=5\\x+y-5z&=0\end{aligned}$$

We have $A=\begin{pmatrix}3&-2&4\\-2&5&1\\1&1&-5\end{pmatrix}$ and $\det A=-88$.

Further, $A_x=\begin{pmatrix}3&-2&4\\5&5&1\\0&1&-5\end{pmatrix}$ and $\det A_x=-108$. Consequently, $x=\frac{-108}{-88}=\frac{27}{22}$.

Next, $A_y=\begin{pmatrix}3&3&4\\-2&5&1\\1&0&-5\end{pmatrix}$ and $\det A_y=-122$. Consequently $y=\frac{-122}{-88}=\frac{61}{44}$.

Finally, $A_z=\begin{pmatrix}3&-2&3\\-2&5&5\\1&1&0\end{pmatrix}$ and $\det A_z=-46$. Consequently $z=\frac{-46}{-88}=\frac{23}{44}$.

2.6 Eigenvalues and eigenvectors

Definition 2.20. *Let A be a square matrix. We are looking for λ for which there is a nontrivial solution to*

$$Ax^T=\lambda x^T.$$

Such number λ is called eigenvalue.

This means that

$$(A-\lambda I)x^T=0.$$

This equation has a nontrivial solution only if $A-\lambda I$ is a singular matrix. Consequently, λ is an eigenvalue if and only if

$$\det(A-\lambda I)=0.$$

Definition 2.21. *The polynomial* $\det(A-\lambda I)$ *is called a* characteristic polynomial.

Definition 2.22. *Let λ be an eigenvalue of A. A vector v solving*

$$(A-\lambda I)v=0$$

is called an eigenvector *corresponding to λ.*

Remark 2.4. *If v is an eigenvector then tv is also an eigenvector for all $t\in\mathbb{R}$.*
Let v and w be eigenvectors corresponding to the same eigenvalue. Then $tv+sw$ is also an eigenvector for all $t,s\in\mathbb{R}$.
Generally, let $u_i,\ i=\{1,\dots,k\}$ be eigenvectors corresponding to λ. Then all their linear combinations are also eigenvectors corresponding to λ.
In what follows, if we say that there is only one eigenvector v, we mean that there is just one-dimensional space of eigenvectors spanned by v. If we say that there are two eigenvectors v,w, we mean that there is two-dimensional space of eigenvectors spanned by v,w. And so on.

Example Find all eigenvalues and eigenvectors to $A = \begin{pmatrix} 5 & 1 \\ 4 & 5 \end{pmatrix}$.
First, we find eigenvalues by solving

$$0 = \det\left(\begin{pmatrix} 5 & 1 \\ 4 & 5 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det \begin{pmatrix} 5-\lambda & 1 \\ 4 & 5-\lambda \end{pmatrix}$$
$$= 25 - 10\lambda + \lambda^2 - 4 = \lambda^2 - 10\lambda + 21.$$

We obtain

$$\lambda_1 = 3,\ \lambda_2 = 7.$$

Consider first $\lambda_1 = 3$. Then we have to solve $\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$. We have

$$\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 \end{pmatrix}$$

and we take $y = t$ and $x = -\frac{t}{2}$. Thus $(x, y) = t(-1/2, 1)$ and $v_1 = (-1/2, 1)$ is an eigenvector related to $\lambda = 3$.
Consider $\lambda_2 = 7$. Then

$$\begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix} \sim \begin{pmatrix} -2 & 1 \end{pmatrix}.$$

and we take $y = t$ and $x = \frac{t}{2}$. Consequently, $v_2 = (1/2, 1)$ is an eigenvector related to the eigenvalue $\lambda = 7$.

Exercise: Find eigenvalues and eigenvectors to $A = \begin{pmatrix} 10 & -9 \\ 4 & -2 \end{pmatrix}$.
First, we have to solve

$$0 = \det \begin{pmatrix} 10-\lambda & -9 \\ 4 & -2-\lambda \end{pmatrix} = \lambda^2 - 8\lambda + 16.$$

This yields the only solution $\lambda_1 = 4$. To find an eigenvector we solve

$$\begin{pmatrix} 6 & -9 \\ 4 & -6 \end{pmatrix} \sim \begin{pmatrix} 2 & -3 \end{pmatrix}.$$

Thus, $(3/2, 1)$ is an eigenvector.

Exercise: Find eigenvalues and eigenvectors to $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
Solve

$$0 = \det \begin{pmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2.$$

We get $\lambda = 1$. To find eigenvalues we have to solve

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

The solutions are of the form $s(1, 0) + t(0, 1)$ for all real numbers $s, t \in \mathbb{R}$.

Definition 2.23. *A* generalized eigenvector w *corresponding to an eigenvalue* λ *is a vector satisfying*

$$(A - \lambda I)w^T = v^T$$

where v *is an eigenvector corresponding to* λ*.*

Lemma 2.2. *Let* λ *be a double root of the characteristic polynomial. Assume, moreover, that there is just one corresponding eigenvector. Then there is a generalized eigenvector corresponding to* λ*.*

Exercise: Consider again the matrix $\begin{pmatrix} 10 & -9 \\ 4 & -2 \end{pmatrix}$. We already know that $\lambda = 4$ is the only eigenvalue and, consequently, the matrix $A - \lambda I$ has the form

$$\begin{pmatrix} 6 & -9 \\ 4 & -6 \end{pmatrix}$$

and the corresponding eigenvector is $\begin{pmatrix} 3 & 2 \end{pmatrix}$. We look for a vector $w = (x, y)$ solving

$$\begin{pmatrix} 6 & -9 \\ 4 & -6 \end{pmatrix} w^T = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

By the Gauss elimination

$$\begin{pmatrix} 6 & -9 & | & 3 \\ 4 & -6 & | & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & -3 & | & 1 \end{pmatrix}$$

Here $y = t$, $t \in \mathbb{R}$ is free and we have $2x - 3t = 1$ and, therefore, $x = \frac{1}{2} - \frac{3}{2}t$. Every vector of the form $\left(\frac{1}{2} - \frac{3}{2}t, t\right)$ is the generalized eigenvector – for example a vector $(-1, 1)$.

2.7 Definiteness

Definition 2.24. *Let* A *be an* n *by* n *symmetric matrix. The mapping*

$$Q : \begin{array}{l} \mathbb{R}^n \to \mathbb{R} \\ v \mapsto vAv^T \end{array}$$

is called a quadratic form.

Examples:

- Quadratic form given by a matrix $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ is

  $$(x, y) \mapsto \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^2 - 2xy + y^2$$

  and we write $Q(x, y) = x^2 - 2xy + y^2$.

- A matrix A associated with the quadratic form

  $$Q(x, y, z) = x^2 - 3xz + y^2 - z^2$$

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

- A quadratic form given by $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 2 & 1 & -2 \end{pmatrix}$ is

$$Q(x, y, z) = x^2 - y^2 - 2z^2 + 4xz + 2yz.$$

Definition 2.25. *A quadratic form Q is*

- positive-definite *if $Q(v) > 0$ for every $v \in \mathbb{R}^n \setminus \{0\}$*
- *positive-semidefinite if $Q(v) \geq 0$ for every $v \in \mathbb{R}^n$*
- negative-definite *if $Q(v) < 0$ for every $v \in \mathbb{R}^n \setminus \{0\}$*
- *negative-semidefinite if $Q(v) \leq 0$ for every $v \in \mathbb{R}^n$*
- *indefinite if there are $v_1, v_2 \in \mathbb{R}$ such that $Q(v_1) < 0 < Q(v_2)$*

Examples:

- $Q(x, y) = x^2 - 2xy + y^2$ is positive-semidefinite since $Q(x, y) = (x - y)^2 \geq 0$. Note that Q is not positive-definite as $Q(1, 1) = 0$.
- $Q(x, y) = x^2 - y^2$ is indefinite because $Q(1, 0) = 1 > 0$ and $Q(0, 1) = -1 < 0$.
- $Q(x, y) = x^2 + 2xy + 2y^2$ is positive-definite because $Q(x, y) = (x + y)^2 + y^2$ and this is always non-negative and $Q(x, y) = 0$ if and only if $x = y = 0$.
- $Q(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ is indefinite. Indeed, $Q(x, y) = x^2 + 2xy = x(x + 2y)$ and, clearly, $Q(1, 0) = 1 > 0$ and $Q(1, -1) = -1 < 0$.

Definition 2.26. *The definiteness of a symmetric matrix A is inherited from the associated quadratic form.*

Theorem 2.5 (Sylvester rule)**.** *Let A be n by n matrix. Denote $D_0 = 1$, $D_1 = \det(a_{11})$, $D_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \ldots, D_n = \det A$ and assume $D_0, D_1, \ldots, D_n \neq 0$. If all products $D_0 \cdot D_1,\ D_1 \cdot D_2, \ldots, D_{n-1}D_n$ are positive, A is a positive-definite matrix. If all the products are negative, A is a negative-definite matrix.*

Examples:

- $Q(x, y) = x^2 + 2xy + 2y^2$ is positive-definite (we already know it). Nevertheless, let verify it by the Sylvester rule. The associated symmetric matrix is $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ and we have $D_0 = 1$, $D_1 = 1$, $D_2 = 1$ and Q is indead positive-definite.
- Consider $Q(x, y) = -x^2 - y^2$. We have $Q(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ and, therefore, $D_0 = 1$, $D_1 = -1$, and $D_2 = 1$. Consequently, the Sylvester rule yields that Q is negative-definite.

3 Functions

3.1 Mappings, introduction

Definition 3.1. *Let $f \subset (X \times Y)$ be a subset which fulfills for every $x \in X$ and $y_1,\ y_2 \in Y$ that*

$$((\langle x, y_1\rangle \in f)\&(\langle x, y_2\rangle \in f)) \Rightarrow (y_1 = y_2).$$

Then we say that f is a mapping *which maps X to Y. We write $f : X \to Y$. A usual notation for $\langle x, y\rangle \in f$ is $f(x) = y$ or $f : x \mapsto y$.*

A domain *is a set of all $x \in X$ for which there exists y such that $f(x) = y$. The domain of f is denoted by* $\operatorname{Dom} f$. *The set of all $y \in Y$ for which there exists $x \in X$ such that $f(x) = y$ is called* range. *It is denoted by* $\operatorname{Ran} f$.

Let $A \subset \operatorname{Dom} f$. An image of A (denoted by $f(A)$) is a set in $\operatorname{Ran} f$ *defined as*

$$f(A) = \{y \in Y,\ \exists x \in A,\ y = f(x)\}.$$

Let $B \subset \operatorname{Ran} f$. A preimage of B (denoted by $f^{-1}(B)$) is a set in $\operatorname{Dom} f$ *defined as*

$$f^{-1}(B) = \{x \in X,\ \exists y \in B,\ y = f(x)\}.$$

Remark 3.1. *Usually, if X and Y are number sets ($\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$), then f is called a* function. *Nevertheless, we will often use the term 'function' also for mappings.*

Example: Let $f = \{\langle 3, 1\rangle, \langle 1, 2\rangle, \langle 2, 2\rangle\}$. We have $\operatorname{Dom} f = \{1, 2, 3\}$, $\operatorname{Ran} f = \{1, 2\}$. On the other hand, let $g = \{\langle 1, 3\rangle, \langle 2, 1\rangle, \langle 2, 2\rangle\}$. Now g is not a function because we have one value of x ($x = 2$) which is mapped to two different values of y (either $y = 1$ or $y = 2$). This contradicts the very first property of the definition.

Observation 3.1. *For every $A,\ B \subset \operatorname{Dom} f$ it holds that*

$$f(A \cup B) = f(A) \cup f(B)$$

Proof. It holds that

$$\begin{aligned}(y \in f(A \cup B)) \Rightarrow (\exists x \in (A \cup B),\ y = f(x)) \Rightarrow ((\exists x \in A,\ y = f(x)) \vee (\exists x \in B,\ y = f(x)))\\ \Rightarrow ((y \in f(A)) \vee (y \in f(B))) \Rightarrow (y \in f(A) \cup f(B))\end{aligned}$$

and we have just proven that $f(A \cup B) \subset (f(A) \cup f(B))$.
On the other hand

$$\begin{aligned}(y \in f(A) \cup f(B)) \Rightarrow ((y \in f(A)) \vee (y \in f(B))) \Rightarrow ((\exists x \in A,\ y = f(x)) \vee (\exists x \in B,\ y = f(x)))\\ \Rightarrow (\exists x \in (A \cup B),\ y = f(x)) \Rightarrow (y \in f(A \cup B))\end{aligned}$$

which yields $(f(A) \cup f(B)) \subset f(A \cup B)$. This concludes the proof. □

Definition 3.2. *A function $f : X \mapsto Y$ is said to be*

- injective *if* $\forall x_1,\ x_2 \in \operatorname{Dom} f,\ f(x_1) = f(x_2) \Rightarrow x_1 = x_2$,
- surjective *if* $\operatorname{Ran} f = Y$,
- bijective *if it is* surjective *and* injective.

We use a term injection *(resp.* surjection *or* bijection*) for injective (resp. surjective of bijective) function.*

Example: Let consider the mapping from the previous example, i.e., $f = \{\langle 3,1\rangle, \langle 1,2\rangle, \langle 2,2\rangle\}$. This function is not injective since $f(1) = 2$ as well as $f(2) = 2$. On the other hand, when taking $Y = \{1, 2\}$, then f is sufjective.

Definition 3.3. *Let $f : X \to Y$ and let $g : Y \to Z$ be such that* Ran $f \subset$ Dom g*. Then a* composition *of functions g and f is a function $g \circ f : X \to Z$ defined as*

$$(g \circ f)(x) = g(f(x)).$$

If there is a function $g : Y \to X$ such that Dom f = Ran g, Dom g = Ran f*, $(g \circ f)(x) = x$ for all $x \in$* Dom f *then g is called an* inverse function *to f and we denote it by f^{-1}. An* invertible *function is a function for which there exists the inverse function.*

Example: Take the function f from the previous example and consider a function h given as

$$h = \{\langle 1,5\rangle, \langle 2,8,\rangle\}.$$

Since Dom $h = \{1, 2\}$ = Ran f, we may write down a function $h \circ f$ (or, equivalently $h(f(x))$. We have

$$h(f(3)) = 5,\ h(f(1)) = 8,\ h(f(2)) = 8.$$

The function f from the previous example is not invertible since it is not injective. Take a function j defined as

$$j = \{\langle 1,4\rangle, \langle 2,1\rangle, \langle 3,7\rangle, \langle 4,10\rangle\}.$$

The function h is injective and it is surjective assuming $Y = \{1, 4, 7, 10\}$. Thus there exists j^{-1} and it is a function

$$j^{-1} = \{\langle 1,2\rangle, \langle 4,1\rangle, \langle 7,3\rangle, \langle 10,4\rangle\}.$$

Observation 3.2. *It holds that* Dom f = Ran f^{-1} *and* Ran f = Dom f^{-1} *whenever f is an invertible function.*

Proof. Obvious. □

Recall that a function $f(x) = x$ is often called *identity* and that not every function has its inverse. Moreover, $f \circ g$ is also an identity.

Observation 3.3. *Let $f : X \to Y$,* Dom $f = X$*. The inverse function f^{-1} exists if and only if f is injective.*

Proof. Let f be injective. Thus, for every $y \in$ Ran f there exists only one $x \in$ Dom f such that $y = f(x)$. It suffices to define $f^{-1}(y) = x$.

Let f be not injective. There exist $x_1,\ x_2 \in$ Dom f and $y \in$ Ran f such that $x_1 \neq x_2$ and $f(x_1) = f(x_2) = y$. Let $f^{-1}(y) = x_1$ – this is necessary to have $f^{-1}(f(x_1)) = x_1$. Then $f^{-1}(f(x_2)) = f^{-1}(y) = x_1 \neq x_2$ and thus f^{-1} is not an inverse function. □

Definition 3.4. *An* indicator function *of a set $A \subset X$ is a function $f : X \mapsto \{0, 1\}$,* Dom $f = X$ *fulfilling $f(x) = 1$ if and only if $x \in A$. Such function is denoted by χ_A.*

Definition 3.5. *We say that $f : X \mapsto \mathbb{R}$ is* bounded from above *if there is $M \in \mathbb{R}$ such that $f(x) \leq M$ for each $x \in$* Dom f*. It is* bounded from below *if there is $m \in \mathbb{R}$ such that $f(x) \geq m$ for every $x \in$* Dom f*. We say that f is* bounded *if f is bounded from above and from below.*

3.2 Exercises

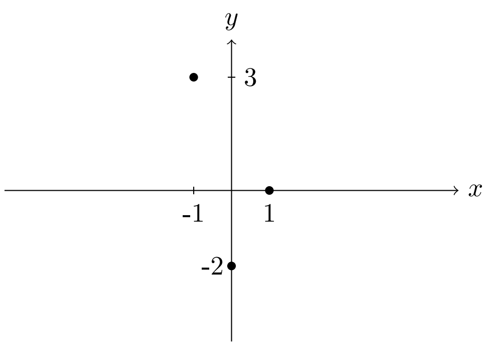
1. Show that $1 > 0$.
2. Show that $\sup(0,2) = \sup[0,2] = 2$.
3. Find $\sup A$ and $\inf A$ of $A = \left\{\frac{1}{n}\right\}_{n=1}^{\infty}$ (i.e., a set $A = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots\right\}$).
4. Which of these subsets of $\mathbb{N} \times \mathbb{N}$ is a function?
   (a) $f = \{\langle 1,5\rangle, \langle 2,4\rangle, \langle 1,3\rangle\}$
   (b) $g = \{\langle 1,2\rangle, \langle 5,3\rangle, \langle 10,1\rangle\}$
   (c) $h = \{\langle 3,3\rangle, \langle 4,3\rangle, \langle 7,7\rangle, \langle 10,3\rangle\}$
5. Consider a function h defined in the previous exercise. Write Dom h and Ran h.
6. Does the following modification of Observation 3.1
$$\forall A, B \subset \text{Dom } f,\ f(A \cap B) = f(A) \cap f(B)$$
hold? If yes, prove it. If no, try to think for which functions does it hold.
7. Let $f,\ g : \mathbb{N} \mapsto \mathbb{N}$ be defined as
$$f = \{\langle 2,2\rangle,\ \langle 3,2\rangle,\ \langle 4,6\rangle,\ \langle 1,3\rangle\}$$
$$g = \{\langle 2,3\rangle,\ \langle 3,2\rangle,\ \langle 6,2\rangle\}$$
Write $f \circ g$ and $g \circ f$.
8. Let f be an invertible function. Show that f^{-1} is determined uniquely.

3.3 Real functions

By a *real function* we mean a function $f : \mathbb{R} \mapsto \mathbb{R}$.

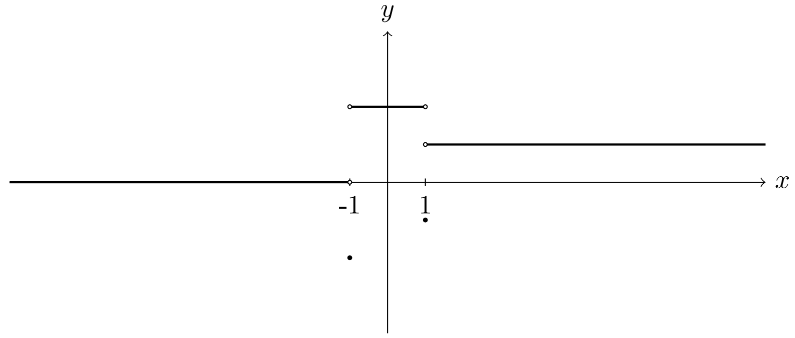
Definition 3.6. *A* graph *of a function f is a subset of plane consisting of points $\langle x, f(x)\rangle$ where $x \in$ Dom f.*

Consider a function $f = \{\langle 1,0\rangle,\ \langle -1,3\rangle,\ \langle 0,-2\rangle\}$. Its graph looks as follows



It is worth pointing out that Dom $f = \{-1,0,1\}$ and Ran $f = \{-2,0,3\}$.

A graph of function $f = 2\chi_{(-1,1)} - 2\chi_{\{-1,1\}} + \chi_{[1,\infty)}$ is



Definition 3.7. *Let $f : \mathbb{R} \mapsto \mathbb{R}$ and $I \subset \mathrm{Dom}\ f$ be an interval. We say that f is on I*

- increasing *if* $\forall x_1,\ x_2 \in I,\ (x_1 < x_2) \Rightarrow f(x_1) < f(x_2)$,
- decreasing *if* $\forall x_1,\ x_2 \in I,\ (x_1 < x_2) \Rightarrow f(x_1) > f(x_2)$,
- non-decreasing *if* $\forall x_1,\ x_2 \in I,\ (x_1 < x_2) \Rightarrow f(x_1) \leq f(x_2)$,
- non-increasing *if* $\forall x_1,\ x_2 \in I,\ (x_1 < x_2) \Rightarrow f(x_1) \geq f(x_2)$.

If f posses one of these properties we will say that f is monotone.

Definition 3.8. *A function $f : \mathbb{R} \mapsto \mathbb{R}$ is called* periodic*, if there is a number $l > 0$ such that $f(x) = f(x + l)$ for all $x \in \mathbb{R}$. The least number l with that property is called a* period *of a function f and f is then $l-$periodic.*

Definition 3.9. *A function $f : \mathbb{R} \mapsto \mathbb{R}$ is said to be* continuous *at a point $x_0 \in \mathrm{Dom}\ f$ if*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in (x_0 - \delta, x_0 + \delta) \cap \mathrm{Dom}\ f,\ |f(x) - f(x_0)| < \varepsilon.$$

A function $f : \mathbb{R} \mapsto \mathbb{R}$ is said to be left-continuous *(resp.* right-continuous*) at a point $x_0 \in \mathrm{Dom}\ f$ if*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in (x_0 - \delta, x_0) \cap \mathrm{Dom}\ f,\ |f(x) - f(x_0)| < \varepsilon$$
$$(\textit{resp. } \forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in (x_0, x_0 + \delta) \cap \mathrm{Dom}\ f,\ |f(x) - f(x_0)| < \varepsilon)$$

We say that f is continuous on a set $S \subset \mathbb{R}$ if it is continuous at all of its points.

We define a function $\mathrm{sgn}\ (x) = \chi_{[0,\infty)} - \chi_{(-\infty,0)}$ – the function is called 'signum' or 'sign', it is equal to -1 for x negative and 1 otherwise. This function is not continuous at $x_0 = 0$. However, it is right-continuous at 0. Indeed, let $\varepsilon = \frac{1}{2}$. Then for every $\delta > 0$, $-\frac{\delta}{2} \in (-\delta, \delta)$ and

$$\left| \mathrm{sgn}\ \left(-\frac{\delta}{2}\right) - \mathrm{sgn}\ (0) \right| = |-1-1| = 2 > \frac{1}{2}.$$

On the other hand, for every $\varepsilon > 0$ we can state (for example) $\delta = \varepsilon$ and then for every $x \in (0, \delta)$ it holds that $\mathrm{sgn}\ (x) = 1 = \mathrm{sgn}\ (0)$ and thus $|\mathrm{sgn}\ (x) - \mathrm{sgn}\ (0)| = 0 < \varepsilon$.

Before we go on let us recall the triangle inequality

$$|a + b| \leq |a| + |b|$$

which holds true for all $a, b \in \mathbb{R}$. We immediately deduce that, also,

$$|a| - |b| \leq |a - b|.$$

Observation 3.4. *Let f and g be functions continuous at x_0. Then also $f(x) \pm g(x)$ and $f(x) \cdot g(x)$ are continuous at x_0. Moreover, if $g(x_0) \neq 0$ then $\frac{f(x)}{g(x)}$ will be continuous at x_0.*

Proof. Proof: We prove it for $f + g$ as $f - g$ can be done similarly. Due to continuity we have $\forall \varepsilon > 0$ $\exists \delta_1 > 0$ and $\delta_2 > 0$ such that $|f(x) - f(x_0)| < \frac{\varepsilon}{2}$ and $|g(x) - g(x_0)| < \frac{\varepsilon}{2}$ whenever $|x - x_0| < \delta$. But this means that (due to the triangle inequality)

$$|f(x) + g(x) - (f(x_0) + g(x_0))| < |f(x) - f(x_0)| + |g(x) - g(x_0)| < \varepsilon.$$

Now we turn our attention to the product rule. First of all, since $f(x_0)$ is real and the function is continuous, there exists $\delta_1 > 0$ and $M_1 > 0$ such that $|f(x)| < M_1$ whenever $x \in (x_0 - \delta_1, x_0 + \delta_1) \cap \mathrm{Dom}\, f$ (see exercises at the end of this section). Similarly, there exists $\delta_2 > 0$ and $M_2 > 0$ such that $|g(x)| < M_2$ whenever $x \in (x_0 - \delta_2, x_0 + \delta_2) \cap \mathrm{Dom}\, f$. Due to continuity, for all $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(x_0)| < \frac{\varepsilon}{2M_2}$ and $|g(x) - g(x_0)| < \frac{\varepsilon}{2M_1}$ for all $x \in (x_0 - \delta, x_0 + \delta)$. We may moreover assume that $\delta < \min\{\delta_1, \delta_2\}$. Then we have

$$\begin{aligned} |f(x)g(x) - f(x_0)g(x_0)| &= |f(x)(g(x) - g(x_0)) + g(x_0)(f(x) - f(x_0))| \\ &\leq |f(x)||g(x) - g(x_0)| + |g(x_0)||f(x) - f(x_0)| < \varepsilon \end{aligned}$$

for all $x \in (x_0 - \delta, x_0 + \delta)$.

To prove the last claim it suffices to show that $\frac{1}{g}$ is continuous at x_0 and to use the just proven product rule. Without loss of generality, assume that $g(x_0) > 0$ and denote its value by $y_0 = g(x_0)$. Then, due to the continuity of g, there exists $\delta_1 > 0$ such that $g(x) > \frac{y_0}{2}$ for all $x \in (x_0 - \delta_1, x_0 + \delta_1) \cap \mathrm{Dom}\, g$. Further, for each $\varepsilon > 0$ there exists $\delta > 0$ such that $|g(x) - g(x_0)| < y_0^2 \frac{\varepsilon}{2}$ for each $x \in (x_0 - \delta, x_0 + \delta)$ and, moreover, we assume that $\delta < \delta_1$. Then we have

$$\left| \frac{1}{g(x)} - \frac{1}{g(x_0)} \right| = \left| \frac{g(x_0) - g(x)}{g(x)g(x_0)} \right| \leq \frac{|g(x_0) - g(x)|}{y_0 \frac{y_0}{2}} < \varepsilon$$

for each $x \in (x_0 - \delta, x_0 + \delta) \cap \mathrm{Dom}\, g$. □

It is easy to deduce that $f(x) \equiv c$ and $f(x) = x$ are continuous on $\mathbb{R}$.

Definition 3.10. *We say that f is an* odd function *if*

$$\forall x \in \mathrm{Dom}\, f,\ -x \in \mathrm{Dom}\, f \text{ and } f(-x) = -f(x).$$

We say that f is an even function *if*

$$\forall x \in \mathrm{Dom}\, f,\ -x \in \mathrm{Dom}\, f \text{ and } f(-x) = f(x).$$

3.4 Further comments on continuous functions

This section is devoted to advanced properties of continuous functions. They will be mentioned without a proof which is usually not elementary.

Before that, we introduce a notion of a maximum and minimum of set $A \subset \mathbb{R}$.

Definition 3.11. *Let* $\sup A$ *be an element of $A \subset \mathbb{R}$. Then* $\sup A$ *is the highest number of A (or a maximum of A) and we write* $\sup A = \max A$. *Similarly, if* $\inf A$ *is an element of A, then* $\inf A$ *will be the lowest number of A (or a minimum of A) and we write* $\inf A = \min A$.

The minimum and maximum does not necessarily exists for a general set $A \subset \mathbb{R}$. For example, $A = \left\{\frac{1}{n},\ n \in \mathbb{N}\right\}$ has maximum 1, however, minimum does not exists. The infimum 0 is not contained in this set.

Note also that every set $A \subset \mathbb{R}$ with finitely many elements has its maximum and minimum.

Definition 3.12. *Let f be continuous on an interval $I \subset \mathbb{R}$. Then we write $f \in \mathcal{C}(I)$.*

Theorem 3.1 (Weierstrass)**.** *Let $f \in \mathcal{C}([a,b])$. Then f is bounded and there exists $t, u \in [a,b]$ such that $f(u) \leq f(x) \leq f(t)$ for all $x \in [a,b]$.*

Actually, the previous theorem states that every function which is continuous on a closed interval attains its maximum and minimum value.

Theorem 3.2 (Bolzano)**.** *Let $f \in \mathcal{C}([a,b])$ and $f(a)f(b) < 0$. Then there is $\eta \in (a,b)$ such that $f(\eta) = 0$.*

Lemma 3.1. *Let f be an odd function and $(-a,a) \subset \operatorname{Dom} f$ for some $a > 0$. Then $f(0) = 0$.*

3.5 Elementary functions

Now we are in position where we can define and state basic properties of functions which will be of use hereinafter.

3.5.1 Polynomials

Polynomials are function which arises from a constant function $f \equiv c$, $c \in \mathbb{R}$ and an identity function $f(x) = x$ by finite number of multiplication and additions. In particular, every polynomial is of the form

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x^1 + a_0,$$

where $n \in \mathbb{N}$ and $a_0, \ldots, a_n \in \mathbb{R}$. The numbers $a_0, \ldots, a_n$ are called coefficients. The degree of $p(x)$ is n in a case $a_n \neq 0$ and we write $\deg p = n$. The term $a_n x^n$ is called a *leading term*. Recall that $p(x) = x^n$ is odd function for odd n and it is an even function for n even. The maximal domain of $p(x)$ is always $\mathbb{R}$. All x such that $p(x) = 0$ are called *roots* of polynomial p. Let x_0 be a root of $p(x)$. Then $p(x) = (x - x_0)q(x)$ where $q(x)$ is a polynomial and it holds that $\deg p(x) = \deg q(x) + 1$.

3.5.2 Rational functions

A *rational function* is a fraction whose nominator and denominator are polynomials. I.e., a rational function f is of the form

$$f(x) = \frac{p(x)}{q(x)}.$$

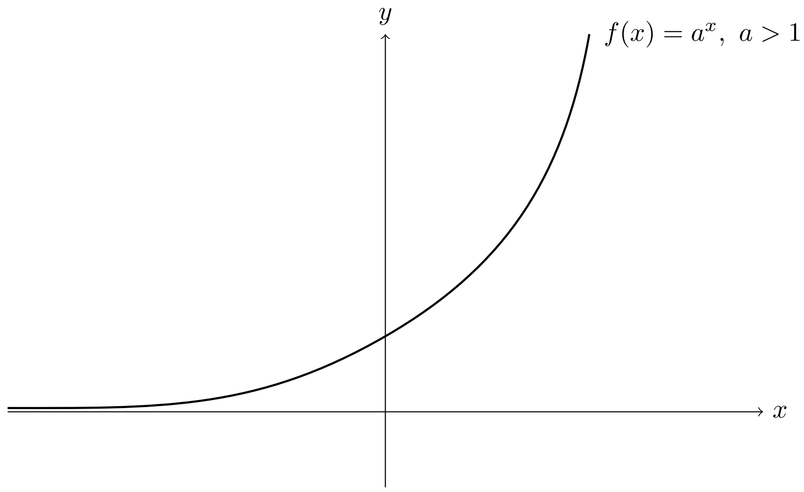
The domain of f is all real numbers except roots of $q(x)$.

3.5.3 Exponential and logarithm

Consider a number $a > 0$. Let $n \in \mathbb{N}$, we define $a^n = a \cdot a \cdot \ldots \cdot a$ where a appears n times on the right hand side. Further, we define $a^{\frac{1}{n}}$ as such number b that $b^n = a$. This allows to define a^r for all rational numbers $r \in Q$. Namely, let $r > 0$, we define $a^r = a^{\frac{p}{q}} = (a^p)^{\frac{1}{q}}$. For $r < 0$ we take $a^r = \frac{1}{a^{-r}}$. Finally, we are allowed to define uniqely a continuous function

$$f(x) = a^x \tag{3}$$

whose values are prescribed in the aforementioned way for all rational inputs. Since the function is constant for $a \equiv 1$, we remove this particular base from our definition and we consider the relation (3) only for $a \in (0,1) \cup (1,\infty)$. It holds that Dom $f = \mathbb{R}$ and Ran $f = (0,\infty)$. Further, $f(0) = 1$ (roughly speaking, every number powered to 0 equals one). The function is strictly increasing for $a > 1$ and strictly decreasing for $a < 1$. The picture below is a graph of a function $f(x) = a^x$ for some $a > 0$.

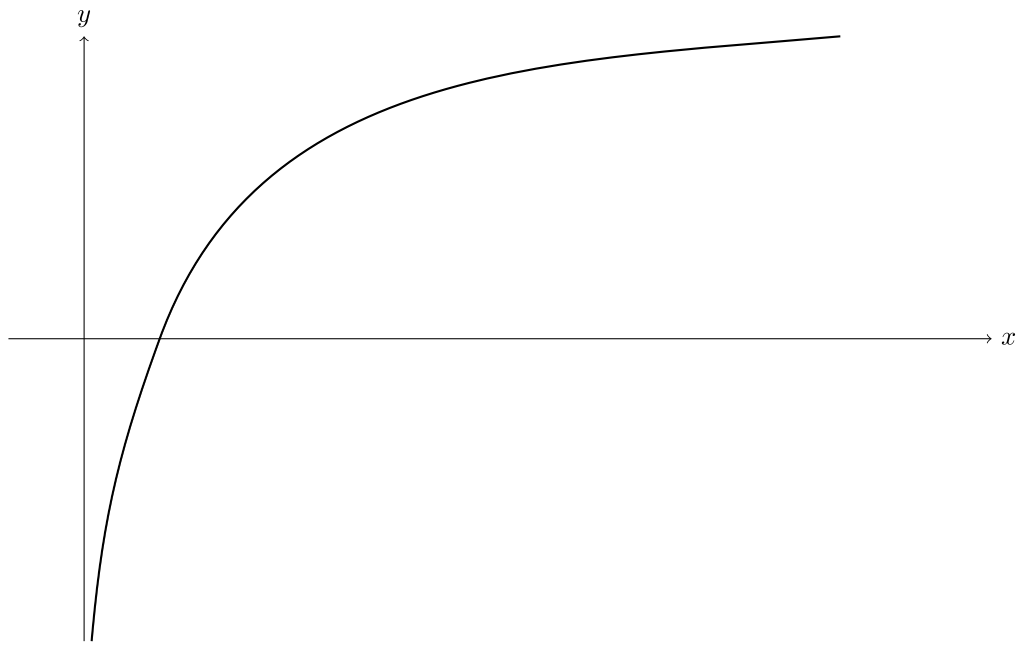


Since $x \mapsto a^x$ is injective there exists an inverse function. We will denote it by $\log_a$ and it is called logarithm to base a. In particular

$$\log_a y = x \quad \Leftrightarrow \quad a^x = y.$$

Recall that $a \in (0,1) \cup (1,\infty)$ and, due to the properties of the inverse functions, Dom $\log_a = (0,\infty)$ and Ran $\log_a = \mathbb{R}$. Recall also, that since $a^0 = 1$, we have $\log_a 1 = 0$ for every $a \in (0,1) \cup (1,\infty)$.

The graph of $f(x) = \log_a(x)$, $a > 1$ is the following



Let e be Euler's number (for its definition see relation (5)). The logarithm to base e is called *natural logarithm* and, because of its importance, we omit the index e in its notation (i.e. $\log x = \log_e x$).

3.5.4 Irrational functions

Next, we define *nth root* $f(x) = \sqrt[n]{x}$ as an inverse to $g(x) = x^n$. Recall that g is invertible for n odd and Dom g = Ran $g = \mathbb{R}$. Thus, Dom $\sqrt[n]{x}$ = Ran $\sqrt[n]{x} = \mathbb{R}$ for n odd.

However, g is not invertible for n even. In that case we have to restrict the domain of g to $[0, \infty)$ in order to have an injective function. The range of this restricted function is also $[0, \infty)$. As a consequence, Dom $\sqrt[n]{x}$ = Ran $\sqrt[n]{x} = [0, \infty)$ for n even.

The nth root is always an increasing function.

3.5.5 Trigonometric functions

There is just one pair of continuous functions $s(x)$ and $c(x)$ with the following properties

- $s(x)^2 + c(x)^2 = 1$
- $s(x + y) = s(x)c(y) + c(x)s(y)$
- $c(x + y) = c(x)c(y) - s(x)s(y)$
- $0 < xc(x) < s(x) < x$ for all $x \in (0, 1)$.

The function s is called sinus and the function c is called cosine. We also introduce notation $\sin x = s(x)$ and $\cos x = c(x)$. These functions have the following properties:

- Dom $\sin x$ = Dom $\cos x = \mathbb{R}$, Ran $\sin x$ = Ran $\cos x = [-1, 1]$.
- $\sin x$ is an odd function, $\cos x$ is an even function.

- $\sin x$ and $\cos x$ are 2π periodic function.

There are several 'known' values of sin and cos:

$x=$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	π	$\frac{3}{2}\pi$
$\sin x$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	0	-1
$\cos x$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	-1	0

Besides, we define a function $\tan x = \frac{\sin x}{\cos x}$ (tangens) and a function $\cot x = \frac{\cos x}{\sin x}$ (cotangens). These functions are $\pi-$periodic, their range is $\mathbb{R}$ and

$$\text{Dom } \tan x = \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z} \right\},\ \text{Dom } \cot x = \mathbb{R} \setminus \{k\pi, k \in \mathbb{Z}\}.$$

3.5.6 Cyclometric functions

Roughly speaking, cyclometric functions are inverse functions to the aforementioned trigonometric functions. However, every trigonometric function is periodic and thus it is not one-to-one. To obtain the inverse function, we have to restrict the domain of every trigonometric function. In particular, we define functions $\sin_r$, $\cos_r$, $\tan_r$ and $\cot_r$ as follows

$$\begin{aligned}
\sin_r x = \sin x,\ \text{Dom } \sin_r = [-\pi_2, \pi_2] \\
\cos_r x = \cos x,\ \text{Dom } \cos_r = [0, \pi] \\
\tan_r x = \tan x,\ \text{Dom } \tan_r = [-\pi_2, \pi_2] \\
\cot_r x = \cot x,\ \text{Dom } \cot_r = [0, \pi]
\end{aligned}$$

Now, since these functions are injective, we may define

$$\begin{aligned}
\text{arcsin} &= \sin_r^{-1} \\
\text{arccos} &= \cos_r^{-1} \\
\text{arctan} &= \tan_r^{-1} \\
\text{arccot} &= \cot_r^{-1}
\end{aligned}$$

Let write down several properties of each function:

- Dom $\arcsin = [-1, 1]$, Ran $\arcsin = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, arcsin is an increasing function and $\arcsin(-1) = -\frac{\pi}{2}$, $\arcsin(0) = 0$ and $\arcsin(1) = \frac{\pi}{2}$
- Dom $\arccos = [-1, 1]$, Ran $\arccos = [0, \pi]$, arccos is a decreasing function and $\arccos(-1) = \pi$, $\arccos(0) = \frac{\pi}{2}$ and $\arccos(1) = 0$.
- Dom $\arctan = \mathbb{R}$, Ran $\arctan = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, arctan is an increasing function and $\arctan(0) = 0$.
- Dom $\text{arccot} = \mathbb{R}$, Ran $\text{arccot} = (0, \pi)$, arccot is a decreasing function and $\text{arccot}(0) = \frac{\pi}{2}$.

3.6 Exercises

1. Try to think about the following statement: If both functions $f(x)$ and $g(x)$ are not monotone on $\mathbb{R}$, then their sum $f(x) + g(x)$ is not monotone on $\mathbb{R}$.

   Prove if it is true, find a counterexample if it is false.

2. If a function is not monotone, then it does not have an inverse function. It is true or false? And why?
3. Let f be increasing invertible function. Show that f^{-1} is also increasing. Consider also the case of decreasing invertible function.
4. Use a definition of continuity in order to proof that a function $f(x) = x^2\chi_{(-1,1)} + \chi_{[1,3]}$ is continuous in $x_0 = 1$.
5. Determine all points of continuity of a function $f(x) = x\chi_{\mathbb{Q}} - x\chi_{\mathbb{R}\setminus\mathbb{Q}}$.
6. Find all roots of $p(x) = x^3 - 6x^2 + 11x - 6$.
7. Let f be continuous at x_0. Then there exists $\delta > 0$ and $M > 0$ such that $|f(x)| < M$ for every $x \in (x_0 - \delta, x_0 + \delta) \cap \mathrm{Dom}\, f$. Prove or disprove this claim.
8. Deduce the values of $\sin x$ and $\cos x$ for $x = \frac{2}{3}\pi,\ \frac{3}{4}\pi,\ \frac{5}{6}\pi,\ \frac{7}{6}\pi,\ \frac{5}{4}\pi,\ \frac{4}{3}\pi,\ \frac{5}{3}\pi,\ \frac{7}{4}\pi,\ \frac{11}{6}\pi$.

3.7 Limits of functions

Definition 3.13. *A* limit point *of a set $S \subset \mathbb{R}$ is every point $x_0 \in \mathbb{R}$ such that for every $\delta > 0$ it holds that $((x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)) \cap S \neq \emptyset$.*

Consider, for example, $S = (0,1) \cup \{2\}$. The set of all its limit point is a closed interval $[0,1]$. We are ready to define a limit of a function. First, we consider finite limits.

Definition 3.14. *Let $f : \mathbb{R} \mapsto \mathbb{R}$ and let x_0 be a limit point of* Dom f. *We say, that $A \in \mathbb{R}$ is a* limit of f at x_0 *if*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in ((x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)) \cap \mathrm{Dom}\, f,\ |f(x) - A| < \varepsilon.$$

We write

$$\lim_{x \to x_0} f(x) = A$$

Observation 3.5. *Once the limit exists, it is determined uniquely.*

Proof. Let $\lim_{x\to x_0} f(x) = A$ and $\lim_{x\to x_0} f(x) = B$ for some different $A, B \in \mathbb{R}$. Take $\varepsilon = \frac{1}{3}|B - A|$. According to the definition of a limit, there exists $\delta > 0$ such that $|f(x) - A| < \varepsilon$ and, simultaneously, $|f(x) - B| < \varepsilon$ for some $x \in (x_0 - \delta, x_0 + \delta)$. We use the triangle inequality to deduce

$$|A - B| = |A - f(x) + f(x) - B| \leq |A - f(x)| + |f(x) - B| \leq \frac{2}{3}|A - B|.$$

□

Thus, the definition of the limit is correct.

Observation 3.6. *Let f be a function continuous in a limit point x_0 of* Dom f . *Then*

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Proof. Let $\varepsilon > 0$ be arbitrary. As f is continuous, there exists $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ whenever $|x - x_0| < \varepsilon,\ x \in \mathrm{Dom}\, f$. But that is exactly that δ which suits the definition of a limit. □

Here we would like to emphasize that every elementary function from the previous chapter is continuous on its domain.

This is the first tool which allows a computation. For example

$$\lim_{x\to 3} x-5=-2.$$

Ok, that was too easy. Anyway, we may use it to simplify fractions. Consider for example a function $f(x)=\frac{x^2+4x+3}{x^2-1}$. This function is clearly not defined at points -1 and 1 and is continuous everywhere else. Anyway, we may compute

$$\lim_{x\to -1}\frac{x^2+4x+3}{x^2-1}=\lim_{x\to -1}\frac{(x+1)(x+3)}{(x-1)(x+1)}=\lim_{x\to -1}\frac{x+3}{x-1}=-1$$

Definition 3.15. *Let x_0 be a limit point of* Dom f*. We say that $A\in\mathbb{R}$ is a* left-sided limit *of f at x_0 (resp.* right-sided limit *of f in x_0) if*

$$\forall\varepsilon>0,\ \exists\delta>0,\ \forall x\in(x_0-\delta,x_0)\cap\mathrm{Dom}\ f,\ |f(x)-A|<\varepsilon.$$

(resp.

$$\forall\varepsilon>0,\ \exists\delta>0,\ \forall x\in(x_0,x_0+\delta)\cap\mathrm{Dom}\ f,\ |f(x)-A|<\varepsilon.)$$

We write

$$\lim_{x\to x_0-} f(x)=A \quad (\textit{resp.}\ \lim_{x\to x_0+} f(x)=A).$$

A special case of the one-sided limit is a limit at infinity. This is defined as follows

Definition 3.16. *Let for all $K\in\mathbb{R}$ there be $x\in$ Dom f such that $x>K$. We say that $A\in\mathbb{R}$ is a limit of f at ∞ if*

$$\forall\varepsilon>0,\ \exists K\in\mathbb{R},\ \forall x>K,\ x\in\mathrm{Dom}\ f,\ |f(x)-A|<\varepsilon.$$

We write $\lim_{x\to\infty} f(x)=A$.
We say that A is a limit of $f(x)$ at $-\infty$ if A is a limit of $f(-x)$ at ∞. We write $\lim_{x\to-\infty} f(x)=A$.

Besides that, we define also infinite limits

Definition 3.17. *Let x_0 be a limit point of* Dom f*. We say that $+\infty$ is a limit of f at a point x_0 if*

$$\forall K>0,\ \delta>0,\ \forall x\in((x_0-\delta,x_0)\cup(x_0,x_0+\delta))\cap\mathrm{Dom}\ f,\ f(x)>K.$$

We write $\lim_{x\to x_0} f(x)=+\infty$.
We say that $-\infty$ is a limit of f at x_0 if $\lim_{x\to x_0} -f(x)=+\infty$*. We write* $\lim_{x\to x_0} f(x)=-\infty$.

Of course, one can define also one-sided infinite limits and infinite limits in infinity. We left it to reader as an exercise.

The following observation is one of the most crucial tool for the computation of limits. We call it 'arithmetic of limits'.

Lemma 3.2 (Arithmetic of limits)**.** *Let $f,g:\mathbb{R}\mapsto\mathbb{R}$ and let x_0 be a limit point of* Dom f *and* Dom g*. Let, moreover, $c\in\mathbb{R}$. Then*

$$\begin{aligned}\lim_{x\to x_0}(f(x)\pm g(x))&=\lim_{x\to x_0} f(x)\pm\lim_{x\to x_0} g(x)\\ \lim_{x\to x_0} cf(x)&=c\lim_{x\to x_0} f(x)\\ \lim_{x\to x_0}(f(x)g(x))&=\lim_{x\to x_0} f(x)\lim_{x\to x_0} g(x)\\ \lim_{x\to x_0}\frac{f(x)}{g(x)}&=\frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)}\end{aligned}\tag{4}$$

assuming the right hand side has meaning.

The right hand side is meaningful once we do not divide by zero and if there does not appear *indefinite terms*, i.e.,

$$\frac{0}{0},\ \frac{\infty}{\infty},\ \infty-\infty, 1^{\infty}$$

Note that the arithmetic of limits holds also for the one-sided limits.

Example: Let compute a limit $\lim_{x\to\infty}\frac{x-1}{x-2}$. According to arithmetic of limits $\lim_{x\to\infty} x-1=\infty$ and $\lim_{x\to\infty} x-2=\infty$. However, we cannot write that

$$\lim_{x\to\infty}\frac{x-1}{x-2}=\frac{\infty}{\infty}$$

as we get an indefinite term. The solution makes use of $\lim_{x\to\infty}\frac{1}{x}=0$. This particular limit is left as an exercise. So we compute

$$\lim_{x\to\infty}\frac{x-1}{x-2}=\lim_{x\to\infty}\frac{1-\frac{1}{x}}{1-\frac{2}{x}}=\frac{1-0}{1-2\cdot 0}=1$$

where we first multiply the numerator and denominator by $\frac{1}{x}$ and, second, we use the arithmetic of limits.

Observation 3.7. *Let* $\lim_{x\to x_0} f(x)=A$ *for some* $x_0\in\mathbb{R}$ *and* $A\in\mathbb{R}^*$. *Then also* $\lim_{x\to x_0-} f(x)=A$ *and* $\lim_{x\to x_0+} f(x)=A$.

Once again, the proof of this observation is postponed to the next section.

Let consider $\lim_{x\to 0}\frac{1}{x}$. We are going to show that $\lim_{x\to 0-}\frac{1}{x}=-\infty$ and $\lim_{x\to 0+}\frac{1}{x}=+\infty$. In such case, $\lim_{x\to 0}\frac{1}{x}$ does not exist according to the just mentioned observation.

Let $K>0$. We take $\delta=\frac{1}{K}$ and, consequently, for all $x\in(0,\delta)$ it holds that $f(x)=\frac{1}{x}>\frac{1}{\delta}=K$ and $\lim_{x\to 0+}\frac{1}{x}=\infty$.

Similarly, for all $x\in(-\delta,0)$ it holds that $f(x)=\frac{1}{x}<\frac{1}{\delta}=-K$ and thus $\lim_{x\to 0-}\frac{1}{x}=-\infty$.

Example:

- Let compute $\lim_{x\to 0}\frac{1}{x}$. First, if $x\to 0+$ then $f(x)\to\infty$. Indeed, for every $M>0$ we take $\delta<\frac{1}{M}$ and $\frac{1}{x}>M$ whenever $x\in(0,\delta)$. Similarly, one can deduce that $\lim_{x\to 0-}\frac{1}{x}=-\infty$. Consequently,
$$\lim_{x\to 0}\frac{1}{x} \text{ does not exist.}$$

- How about $\lim_{x\to 0}\frac{1}{x^2}$? In this case we have
$$\lim_{x\to 0+}\frac{1}{x^2}\stackrel{AL}{=}\left(\lim_{x\to 0+}\frac{1}{x}\right)^2=\infty\cdot\infty=\infty$$
and
$$\lim_{x\to 0-}\frac{1}{x^2}\stackrel{AL}{=}\left(\lim_{x\to 0-}\frac{1}{x}\right)^2=-\infty\cdot-\infty=\infty.$$
We deduce that
$$\lim_{x\to 0}\frac{1}{x^2}=\infty.$$

3.8 Advanced limits

There is precisely one real number e such that

$$\lim_{x\to 0}\frac{e^x-1}{x}=1. \tag{5}$$

This number is called Euler's number, it is irrational and its value is approximately 2.72.

Thus we also get

$$\lim_{x\to 0}\frac{\log(x+1)}{x}=1$$

The definition of $\sin x$ allows to deduce

$$\lim_{x\to 0}\frac{\sin x}{x}=0. \tag{6}$$

Lemma 3.3 (Limit of composed function)**.** *Let* $\lim_{x\to x_0} g(x)=A$ *and* $\lim_{y\to A} f(y)=B$*. Then*

$$\lim_{x\to x_0} f(g(x))=B,$$

if at least one of the following is true:

1. *f is continuous at the point A or*
2. *there is δ such that for all $x\in(x-\delta,x_0)\cup(x_0,x+\delta)$ it holds that $g(x)\neq A$.*

Now we are allowed to deduce further limits which will be used without any further explanation (here note that the inner function $g(x)=\frac{x}{2}$ is injective and thus the second assumption of the previous lemma is fulfilled):

$$\lim_{x\to 0}\frac{1-\cos x}{x^2}=\lim_{x\to 0}\frac{\sin^2\left(\frac{x}{2}\right)+\cos^2\left(\frac{x}{2}\right)-\cos^2\left(\frac{x}{2}\right)+\sin^2\left(\frac{x}{2}\right)}{x^2}=\lim_{x\to 0}\frac{2\sin^2\left(\frac{x}{2}\right)}{4\left(\frac{x}{2}\right)^2}$$
$$=\frac{1}{2}\lim_{x\to 0}\frac{\sin\left(\frac{x}{2}\right)}{\frac{x}{2}}\frac{\sin\left(\frac{x}{2}\right)}{\frac{x}{2}}=\frac{1}{2}\lim_{x\to 0}\frac{\sin\left(\frac{x}{2}\right)}{\frac{x}{2}}\lim_{x\to 0}\frac{\sin\left(\frac{x}{2}\right)}{\frac{x}{2}}=\frac{1}{2}$$

Lemma 3.4 (Sandwich Lemma)**.** *Let $x_0\in\mathbb{R}$ and let there is $\delta>0$ such that*

$$f(x)\le g(x)\le h(x),\ \forall x\in(x_0-\delta,x_0)\cup(x_0,x_0+\delta).$$

Then $\lim_{x\to x_0} f(x)=\lim_{x\to x_0} h(x)=A$ *implies* $\lim_{x\to x_0} g(x)=A$*.*

Further limits of elementary functions:

- $\lim_{x\to\infty} a^x=\infty$ for $a>1$,
- $\lim_{x\to\infty}\log_a x=\infty$ for $a>1$,
- $\lim_{x\to 0+}\log_a x=-\infty$ for $a>1$,
- $\lim_{x\to\frac{\pi}{2}-}\tan x=\infty$,
- $\lim_{x\to\infty}\arctan x=\frac{\pi}{2}$,

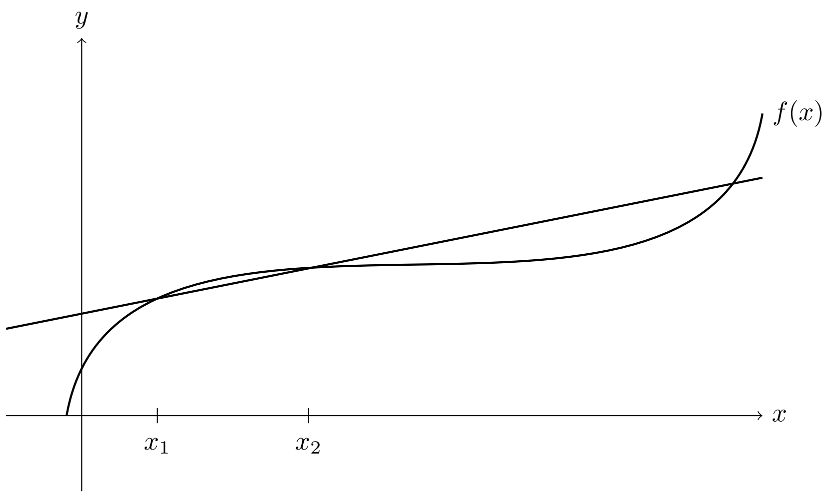
- $\lim_{x\to\infty} \operatorname{arccot} x = 0$,
- $\lim_{x\to-\infty} \operatorname{arccot} x = \pi$.

Examples

- Is $f(x) = \left(\frac{1}{x}\right)\chi_{[1,\infty)} + \left(\frac{(2x+2)(x-1)}{(x+2)(x-1)}\right)\chi_{(-\infty,1)}$ continuous? First, it is clearly continuous in ever point of $(-\infty,-2)$, $(-2,1)$, and $(1,\infty)$. It is not continuous at $x=-2$ since this point does not belong to the domain of f. Next, $f(1)=1$ and $\lim_{x\to 1-} f(x) = \frac{4}{3}$ and thus the function is not continuous at $x=1$.
- And how about $f(x) = e^x\chi_{(-\infty,0]} + \left(\frac{\sin(4x)-\sin(3x)}{4x-3x}\right)\chi_{(0,\infty)}$? In this case the function is clearly continuous everywhere on $\mathbb{R}\backslash\{0\}$. In zero we have $f(0)=e^0=1$ and $\lim_{x\to 0+} f(x) = 1$ and, therefore, the function is continuous even there.

3.9 Derivative

Consider a graph of a function $f(x)$, for example, of the following form



The equation of the line passing through point $\langle x_1, f(x_1)\rangle$ and $\langle x_2, f(x_2)\rangle$ is

$$y = \frac{f(x_2)-f(x_1)}{x_2-x_1}(x-x_1)+f(x_1).$$

How to make a tangent line? Just simply tend with x_2 to x_1. So the tangent line has equation

$$y = k(x-x_1)+f(x_1)$$

where

$$k = \lim_{x_2\to x_1} \frac{f(x_2)-f(x_1)}{x_2-x_1}$$

assuming the limit exists. We denote $h := x_2 - x_1$ and then we may write

$$k = \lim_{h\to 0} \frac{f(x_1+h)-f(x_1)}{h}.$$

Definition 3.18. *Let $f : \mathbb{R} \mapsto \mathbb{R}$. We define*

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}.$$

We say that $f'(x)$ is a derivative *of f at point x.*

In particular, a derivative of f in a point x is a slope of the tangent line passing through $\langle x, f(x)\rangle$.

Let emphasize that f' does not exist for every function.

Observation 3.8. *Let $f'(x_0)$ is real. Then f is continuous at x_0.*

Proof. Indeed, it is enough to compute

$$\lim_{x\to x_0} f(x) - f(x_0) = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}(x-x_0) = f'(x_0)\cdot 0 = 0.$$

Consequently, $\lim_{x\to x_0} f(x) = f(x_0)$ and the function is continuous at x_0. □

Let compute several derivatives of elementary functions. First of all, since

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \ldots + ab^{n-2} + b^{n-1}),$$

we get for $f(x) = x^n$

$$\begin{aligned}\lim_{h\to 0} \frac{(x+h)^n - x^n}{h} &= \lim_{h\to 0} \frac{h((x+h)^{n-1} + (x+h)^{n-2}x + \ldots + (x+h)x^{n-2} + x^{n-1})}{h} \\ &= \lim_{h\to 0} (x+h)^{n-1} + (x+h)^{n-2}x + \ldots + (x+h)x^{n-2} + x^{n-1} = nx^{n-1}.\end{aligned}$$

Thus, $(x^n)' = nx^{n-1}$.

Take $f(x) = e^x$.

$$\lim_{h\to 0} \frac{e^{x+h} - e^x}{h} - \lim_{h\to 0} \frac{e^x(e^h-1)}{h} = e^x \lim_{h\to 0} \frac{e^h - 1}{h} = e^x.$$

Consequently, $(e^x)' = e^x$. Consider $f(x) = \sin x$. We compute

$$\begin{aligned}\lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} &= \lim_{h\to 0} \frac{\sin x \cos h + \sin h \cos x - \sin x}{h} \\ = \lim_{h\to 0} \left(\frac{\sin h \cos x}{h} - \frac{\sin x(1-\cos h)}{h} \right) &\overset{AL}{=} \cos x \lim_{h\to 0} \frac{\sin h}{h} - \sin x \lim_{h\to 0} \frac{1-\cos h}{h^2} h \overset{AL}{=} \cos x\end{aligned}$$

and we deduced that $(\sin x)' = \cos x$.

Let compute derivative of $f(x) = \cos x$:

$$\begin{aligned}\lim_{h\to 0} \frac{\cos(x+h) - \cos x}{h} &= \lim_{h\to 0} \frac{\cos x \cos h - \sin x \sin h - \cos x}{h} \\ &= \lim_{h\to 0} \left(\frac{\cos x(\cos h - 1)}{h} + \frac{-\sin x \sin h}{h} \right).\end{aligned}$$

Similarly as before we deduce that

$$(\cos x)' = -\sin x.$$

Finally, let compute a derivative of $\log x$. We have

$$\lim_{h\to 0} \frac{\log(x+h) - \log x}{h} = \lim_{h\to 0} \frac{\log\left(\frac{x+h}{x}\right)}{h} = \lim_{h\to 0} \frac{\log\left(1+\frac{h}{x}\right)}{h} = \lim_{h\to 0} \frac{\log\left(1+\frac{h}{x}\right)}{\frac{h}{x}} \frac{1}{x} \overset{LOCF}{=} \frac{1}{x}.$$

Consequently,

$$(\log x)' = \frac{1}{x}.$$

Lemma 3.5. *Let f and g be differentiable functions. Then*

$$\begin{aligned} (f(x) \pm g(x))' &= f'(x) \pm g'(x) \\ (f(x)g(x))' &= f'(x)g(x) + f(x)g'(x) \\ \left(\frac{f(x)}{g(x)}\right)' &= \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2} \end{aligned}$$

if both sides have sense.

Proof. The first relation follows directly from the arithmetic of limits. Indeed,

$$\begin{aligned} (f(x) \pm g(x))' = \lim_{h\to 0} & \frac{f(x+h) \pm g(x+h) - (f(x) \pm g(x))}{h} \\ &= \lim_{h\to 0} \frac{(f(x+h) - f(x)) \pm (g(x+h) - g(x))}{h} \\ &\overset{AL}{=} \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \pm \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} = f'(x) \pm g'(x). \end{aligned}$$

Further,

$$\begin{aligned} (f(x)g(x))' &= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\ &= \lim_{h\to 0} \frac{f(x+h)(g(x+h) - g(x)) + g(x)(f(x+h) - f(x)}{h} \\ &\overset{AL}{=} \lim_{h\to 0} \frac{f(x+h)(g(x+h) - g(x))}{h} + \lim_{h\to 0} \frac{g(x)(f(x+h) - f(x))}{h} \\ &= f(x)g'(x) + f'(x)g(x) \end{aligned}$$

which is a proof of the second relation.

Finally,

$$\begin{aligned} \left(\frac{f(x)}{g(x)}\right)' &= \lim_{h\to 0} \frac{1}{h}\left(\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}\right) = \lim_{h\to 0} \frac{1}{h}\left(\frac{f(x+h)g(x) - g(x+h)f(x)}{g(x+h)g(x)}\right) \\ &= \lim_{h\to 0} \frac{1}{g(x+h)g(x)} \frac{(f(x+h) - f(x))\,g(x) - f(x)(g(x+h) - g(x))}{h} \\ &= \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2} \end{aligned}$$

which proves the last relation. $\square$

Let compute
$$(\tan x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}$$
and, similarly, we may deduce $(\cot x)' = -\frac{1}{\sin^2 x}$.

We present the following lemma without a proof. It concern the derivative of composed functions.

Lemma 3.6. *Let f and g be differentiable functions and let $b = f(a)$. Then*
$$(g \circ f)'(a) = g'(b)f'(a) = g'(f(a))f'(a).$$

So we may use this to compute the derivative of a^x:
$$(a^x)' = \left(e^{x \log a}\right)' = e^{x \log a}(x \log a)' = \log a \; e^{x \log a} = \log a \; a^x.$$

Finally, we may also compute remaining derivatives of elementary functions:
$$1 = (x)' = (\arctan \circ \tan x)' = \arctan'(\tan x) \tan' x.$$
We thus deduce that $\arctan'(\tan x) = \frac{1}{\tan'(x)}$ and thus
$$\arctan'(\tan x) = \cos^2 x = \frac{\cos^2 x}{\sin^2 x + \cos^2 x} = \frac{1}{1 + \frac{\sin^2 x}{\cos^2 x}} = \frac{1}{1 + \tan^2 x}$$
which yield
$$\arctan'(x) = \frac{1}{1 + x^2}.$$
The similar computation may be performed also for other cyclometric functions. To sum up, we present the following table:

$f(x)$	$f'(x)$	conditions
x^n	nx^{n-1}	$n \in \mathbb{R}$, x as usual
e^x	e^x	$x \in \mathbb{R}$
a^x	$\log a \; a^x$	$a \in (0,1) \cup (1,\infty)$, $x \in \mathbb{R}$
$\log x$	$\frac{1}{x}$	$x \in (0,\infty)$
$\sin x$	$\cos x$	$x \in \mathbb{R}$
$\cos x$	$-\sin x$	$x \in \mathbb{R}$
$\tan x$	$\frac{1}{\cos^2 x}$	$x \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi, k \in \mathbb{Z}\}$
$\cot x$	$-\frac{1}{\sin^2 x}$	$x \in \mathbb{R} \setminus \{k\pi, k \in \mathbb{Z}\}$
$\arctan x$	$\frac{1}{1+x^2}$	$x \in \mathbb{R}$
$\operatorname{arccot} x$	$-\frac{1}{1+x^2}$	$x \in \mathbb{R}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$x \in (-1,1)$
$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$	$x \in (-1,1)$

3.9.1 Mean-value theorems

Lemma 3.7. *Let f be defined on an interval (a,b) let it attain its maximum (resp. minimum) in a point $x_0 \in (a,b)$, and let $f'(x_0)$ exist. Then $f'(x_0) = 0$.*

Proof. Let x_0 be a point of maximum. For contradiction let $f'(x_0) \neq 0$ and without loss of generality assume $f'(x_0) > 0$. But then there is $\delta > 0$ such that $\frac{f(x)-f(x_0)}{x-x_0} > 0$ for all $x \in (x_0, x_0+\delta)$. But this means that $f(x) > f(x_0)$ which is in contradiction with the very first assumption. $\square$

Lemma 3.8 (Rolle)**.** *Let $f \in \mathcal{C}([a,b])$ and let f' exist for all $x \in (a,b)$. Moreover, let $f(a) = f(b)$. Then there exists a point $\zeta \in (a,b)$ such that $f'(\zeta) = 0$.*

Proof. For f constant it is enough to take any $x \in (a,b)$. Once f is not constant, there is a point $\zeta \in (a,b)$ where this function attains its maximum or minimum. According to the previous lemma, $f'(\zeta) = 0$. $\square$

Lemma 3.9 (Lagrange)**.** *Let $f \in \mathcal{C}([a,b])$ and let f' exists for all $x \in (a,b)$. Then there exists a point $\zeta \in (a,b)$ such that*

$$f'(\zeta)(b-a) = f(b) - f(a)$$

Proof. Consider a function $F(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a)$. This function satisfies all assumption of the previous lemma ($F(a) = F(b) = 0$) and thus there is ζ such that $F'(\zeta) = 0$. This might be rewritten as

$$0 = f'(\zeta) - \frac{f(b)-f(a)}{b-a}$$

which is the desired equality. $\square$

3.9.2 The course of function

The derivative helps to further analyze the function. This is the main content of this section. First of all, the sign of derivative is in correspondence with the monotonicity of function.

Observation 3.9. *Let $f \in \mathcal{C}(I)$ for some interval $I \subset \mathbb{R}$. Assume that $f'(x)$ is exists for all $x \in I$.*

1. *If $f'(x) > 0$ for all $x \in I$, $f(x)$ is increasing on I.*
2. *If $f'(x) < 0$ for all $x \in I$, $f(x)$ is decreasing on I.*

Proof. We prove just the first part as the second part is just an easy modification. Let $x, y \in I$ be arbitrary points such that $x < y$. According to mean value theorem, there is $\zeta \in (x,y)$ such that $f'(\zeta)(x-y) = f(x) - f(y)$. As $f'(\zeta)$ is positive we get $f(x) < f(y)$ which implies the desired claim. $\square$

Definition 3.19. *We say that $x_0 \in \mathrm{Dom}\ f$ is a local maximum of f if there exists $\delta > 0$ such that $f(x) \leq f(x_0)$ for every $x \in (x_0-\delta, x_0+\delta)$. It is a local minimum of f if there exists $\delta > 0$ such that $f(x) \geq f(x_0)$ for every $x \in (x_0-\delta, x_0+\delta)$.*

We define one additional qualitative property of function:

Definition 3.20. *We say that $f : \mathbb{R} \mapsto \mathbb{R}$ is convex on a set $I \subset \mathrm{Dom}\ f$ if for all $x, y, z \in I$, $x < y < z$ it holds that*

$$\frac{f(y)-f(x)}{y-x} < \frac{f(z)-f(y)}{z-y}.$$

We say that f is concave on I if $-f$ is convex on I.

Observation 3.10. *Let $f \in \mathcal{C}(I)$ for some interval $I \subset \mathbb{R}$. Assume that $f''(x)$ exists for all $x \in I$.*

1. *If $f''(x) > 0$ for all $x \in I$ then f is convex on I.*
2. *If $f''(x) < 0$ for all $x \in I$ then f is concave on I.*

Proof. It is enough to show that f' increasing implies f convex as then the claim follows from Observation 3.9. Take $x, y, z \in I$, $x < y < z$. According to mean value theorem there exist $\eta \in (x, y)$ and $\zeta \in (y, z)$ such that $f'(\eta) = \frac{f(y)-f(x)}{y-x}$ and $f'(\zeta) = \frac{f(z)-f(y)}{z-y}$. But since f' is increasing and $\eta < \zeta$ we get $f'(\eta) < f'(\zeta)$ which implies the convexity of f.

If $f''(x) < 0$ we get $(-f)''(x) > 0$ and according to the first part $-f$ is convex. This gives the second claim. $\square$

Definition 3.21. *We say that $x \in \mathbb{R}$ is a* point of inflection *of $f : \mathbb{R} \mapsto \mathbb{R}$ if f is continuous at x and there is $\delta > 0$ such that one of the following appears*

1. *f is concave on $(x - \delta, x)$ and convex on $(x, x + \delta)$*
2. *f is convex on $(x - \delta, x)$ and concave on $(x, x + \delta)$.*

Roughly speaking, the point x is a point of inflection if f changes from convex to concave or vice versa at point x.

Now we are ready to describe the problem of the course of function. The task 'examine the course of the following function' consists of the following sub-tasks:

1. To find out the domain, to determine whether the function is even, odd or periodic.
2. To find intersections with axes.
3. To examine the behavior of the function at the edges of the domain.
4. To derive function, to determine sets where the function is increasing and decreasing, to determine extremes.
5. To differentiate the function for the second time, to determine sets where the function is concave, convex, to determine points of inflection.
6. To sketch a graph of the function.

Let me comment each of this sub-tasks and let me use a function $f(x) = \frac{x^2+3}{x-1}$ as an example:

1. To determinate the domain one has to be sure that there is no division by 0, that the square root is taken from the non-negative number and that the argument of logarithm is positive. In case of the exemplary function we have to exclude the possibility of $x - 1 = 0$ which means that $\mathrm{Dom}\, f = (-\infty, 1) \cup (1, \infty)$. Directly from the domain one may deduce that this function cannot be even, odd or periodic.
2. The intersections with axis are point of form $\langle 0, f(0)\rangle$ and $\langle x, 0\rangle$ where x solves $f(x) = 0$. In our case we obtain $\langle 0, -3\rangle$ and since

$$0 = \frac{x^2 + 3}{x - 1}$$

has no solution there is no intersection with axis x.

3. We have to evaluate limits on the edges of the domain. Let turn attention to our example. Since the domain is of the form $(-\infty, 1) \cup (1, \infty)$ we have to compute the following four limits:

$$\lim_{x\to-\infty} \frac{x^2+3}{x-1} = -\infty$$
$$\lim_{x\to1-} \frac{x^2+3}{x-1} = -\infty$$
$$\lim_{x\to1+} \frac{x^2+3}{x-1} = \infty$$
$$\lim_{x\to\infty} \frac{x^2+3}{x-1} = \infty$$

Besides, we have to examine asymptotes.

Definition 3.22. *Let* $\lim_{x\to\infty} \frac{f(x)}{x} = k_+ \in \mathbb{R}$ *and let* $\lim_{x\to\infty} f(x) - k_+x = q_+$. *Then an* a*symptote at* ∞ *is a line with equation* $y = k_+x + q_+$.

Let $\lim_{x\to-\infty} \frac{f(x)}{x} = k_- \in \mathbb{R}$ *and let* $\lim_{x\to-\infty} f(x) - k_-x = q_-$. *Then an* a*symptote at* $-\infty$ *is a line with equation* $y = k_-x + q_-$.

I our particular case we have:

$$\lim_{x\to\infty} \frac{x^2+3}{x-1}\frac{1}{x} = 1$$
$$\lim_{x\to\infty} \frac{x^2+3}{x-1} - x = 1$$
$$\lim_{x\to-\infty} \frac{x^2+3}{x-1}\frac{1}{x} = 1$$
$$\lim_{x\to-\infty} \frac{x^2+3}{x-1} - x = 1.$$

So there is only line which represents asymptote at ∞ as well as at $-\infty$ and the equation of that line is

$$y = x + 1.$$

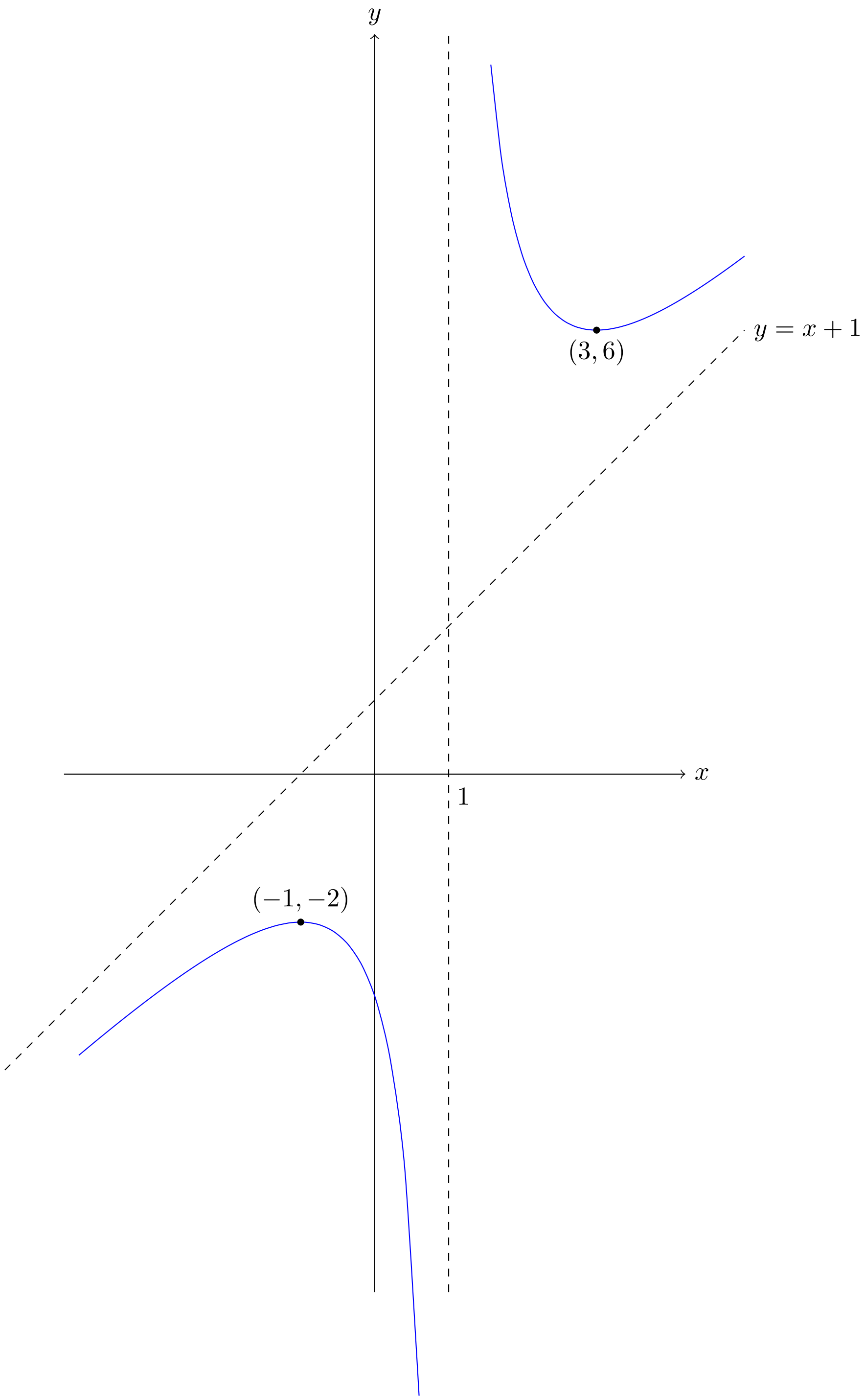
4. We have to differentiate the function and then we have to find all x such that $f'(x) > 0$ and all x for which $f'(x) < 0$. The points where the monotonicity of the function changes are extremal points.

   Take our exemplary function. We have $f'(x) = \frac{x^2-2x-3}{(x-1)^2}$. Consequently, $f'(x) > 0$ whenever $x \in (-\infty, -1)$ and $x \in (3, \infty)$. Moreover, $f'(x) < 0$ for $x \in (-1, 3) \setminus \{1\}$. Thus, f is increasing on $(-\infty, -1)$, f is decreasing on $(-1, 1)$, once again it is decreasing on $(1, 3)$ and f is increasing on $(3, \infty)$. We deduce that the local maximum is at point $x = -1$, its value is -2, the local minimum is at point $x = 3$, its value is 6.

5. We do the same as in the previous step but for the second derivative.

   Consider our exemplary function. We have $f''(x) = \frac{-4}{(x-1)^3}$. Consequently, $f''(x) < 0$ for $x \in (-\infty, 1)$ and $f''(x) > 0$ for $x \in (1, \infty)$ and f is concave on $(-\infty, 1)$ and convex on $(1, \infty)$. If 1 was a point of continuity of f, it would be a point of inflection. However, 1 does not belong to Dom f and thus there is no point of inflection.

6. Now we are ready to draw a graph using all the information we deduced.



3.9.3 Further use

The derivatives may be further used for computation of approximate values and for computation of limits. Without a proof, we state here two important concepts (this time we do not provide any proof):

Lemma 3.10 (l'Hospital)**.** *Let f and g have finite derivatives for all $x \in (a,b) \subset \mathbb{R}$. Assume $g'(x) \neq 0$ and*

$$\lim_{x\to a+} \frac{f'(x)}{g'(x)} = A \in \mathbb{R}^*.$$

Let moreover one of the following is true:

1. $\lim_{x\to a+} f(x) = 0$ *and* $\lim_{x\to a+} g(x) = 0$ *or*
2. $\lim_{x\to a+} |g(x)| = \infty$.

Then

$$\lim_{x\to a+} \frac{f(x)}{g(x)} = A.$$

Definition 3.23 (Taylor's sum)**.** *Let f be $n-$times differentiable at point x_0. Then a polynomial of the form*

$$T_{f,x_0,n}(x) := f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots + \frac{f^{(n)}}{n!} = \sum_{j=0}^{n} \frac{f^{(i)}}{i!}(x-x_0)^i$$

is called the Taylor polynomial *to f at point x_0 of degree n.*

Lemma 3.11. *Assume that f is $(n+1)-$times differentiable at x_0. Let $x \in \mathbb{R}$ be arbitrary and let f is $(n+1)-$times differentiable on a closed interval I with edges at x_0 and x. Then there is ζ in between of x and x_0 such that*

$$f(x) - T_{f,x_0,n}(x) = \frac{f^{(n+1)}(\zeta)}{(n+1)!}(x-x_0)^{n+1}.$$

4 Functions of multiple variables

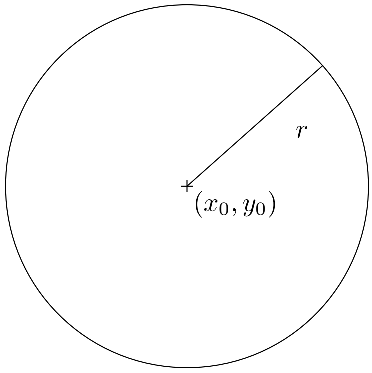
4.1 Few words about topology

Definition 4.1. *An open ball centered at $(x_0,y_0) \in \mathbb{R}^2$ with radius $r \in (0,\infty)$ is a set*

$$B_r(x_0,y_0) = \{(x,y) \in \mathbb{R}^2, \|(x,y)-(x_0,y_0)\| < r\}.$$

Definition 4.2. *A set $M \subset \mathbb{R}^2$ is* open *if for every $(x_0,y_0) \in \mathbb{M}$ there is $r > 0$ such that $B_r(x_0,y_0) \subset M$.*
A set M is called closed *if $\mathbb{R}^2 \setminus M$ is open.*

Example A set $M := (0,1)\times(0,1)$ is open. Indeed, let $(a,b) \in M$. Define $\delta = \min\{a,b,1-a,1-b\}$. Since $a \in (0,1)$ and $b \in (0,1)$ we have $\delta > 0$. Necessarily, $B_{\delta/2}(a,b) \subset M$. On the other hand, a set $M := [0,1]\times(0,1)$ is not open. Consider for example a point $(1,1/2) \in M$. Then every ball $B_r(1,1/2)$ contains a point $(1+r/2,1/2)$ which is outside of M. Note that M is not closed. Why?



Remark 4.1.

- *$\emptyset$ and $\mathbb{R}^2$ are open sets (and closed sets as well),*
- *a union of open sets is an open set,*
- *an intersection of two open sets is an open set,*
- *a union of two closed sets is a closed set,*
- *an intersection of closed sets is a closed set.*

Observation 4.1. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a continuous function. Then $f^{-1}(A)$ is an open set for every $A \subset \mathbb{R}$ open. Similarly, $f^{-1}(B)$ is a closed set for every $B \subset \mathbb{R}$ closed.*

Question What is a *continuous function*? We will see later.
For now: A projection $p : \mathbb{R}^2 \to \mathbb{R}$, $p(x, y) = x$ is a continuous function (as well as projection $q(x, y) = y$). A sum, difference and product of two continuous functions are continuous functions. A quotient of two continuous function is again a continuous function. A composition of two continuous function is a continuous function.

Example Let consider a set

$$M := \{(x, y), x \in (-1, 1),\ y < x^2\}.$$

Is this set open? First, $f(x, y) = |x|$ is a continuous function. Indeed, $f(x, y) = |p(x, y)|$ is a composition of p and $|\cdot|$. Thus, $f^{-1}((-\infty, 1)) = \{(x, y) \in \mathbb{R}^2, x \in (-1, 1)\}$ is an open set.
Next, $g(x, y) = y - x^2$ is a continuous function. Indeed, $g(x, y) = q(x, y) - p(x, y)^2$. Consequently, $f^{-1}((-\infty, 0)) = \{(x, y) \in \mathbb{R}^2, y - x^2 < 0\} = \{(x, y) \in \mathbb{R}^2, y < x^2\}$.
Since $M = f^{-1}((-\infty, 1)) \cap g^{-1}((-\infty, 0))$, we deduce that M is open.

Definition 4.3. *An* interior *of set $M \subset \mathbb{R}^2$ is a set M^0 of all points (x_0, y_0) for which there is $r > 0$ such that $B_r(x_0, y_0) \subset M$. Equivalently, it is the biggest open set contained in M.*
A closure *of a set $M \subset \mathbb{R}^2$ is a set $\overline{M}$ defined as $\overline{M} := \mathbb{R}^2 \setminus (\mathbb{R}^2 \setminus M)^0$. Equivalently, it is the smallest closed set containing M.*
A boundary *of a set M is denoted by ∂M and it is defined as $\overline{M} \setminus M^0$.*

Example Consider $M = [0,1] \times (0,1)$. Then $M^0 = (0,1) \times (0,1)$ and $\overline{M} = [0,1] \times [0,1]$. We deduce that
$$\partial M = \overline{M} \setminus M^0 = ([0,1] \times \{0,1\},\ \{0,1\} \times [0,1])\,.$$

Definition 4.4. *Let $M \subset \mathbb{R}^2$. A point $(x_0, y_0) \in \mathbb{R}^2$ is a* limit point *of M if $B_r(x_0, y_0) \cap M \neq \emptyset$ for every $r > 0$.*
A point $(x_0, y_0) \in M$ is an isolated point *of M if there is $r > 0$ such that $B_r(x_0, y_0) \cap M = \{(x_0, y_0)\}$.*

Example Consider a set $M := \{(x,y) \in \mathbb{R}, y = 0, x = 1/n,\ n \in \mathbb{N}\}$. We claim, that $(0,0)$ is a limit point of M. Indeed, let $r > 0$. Then there is n_r such that $n_r > 1/r$ and, clearly, $(1/n_r, 0) \in M$ is such point that $\|(1/n_r, 0) - (0,0)\| < r$ and thus $B_r(0,0) \cap M = (1/n_r, 0)$.

4.2 Introduction to functions

Definition 4.5. *Let $M \subset \mathbb{R}^n$, $n \in \mathbb{N}$ be a nonempty set. A* real function of multiple variables *defined on a set M is a formula f which assigns a (unique) real number y to every $(x_1, x_2, \ldots, x_n) \in M$. We use the notation*
$$y = f(x_1, x_2, \ldots, x_n).$$
To denote the function itself we use a notation $f : M \to \mathbb{R}$. The set M is called a domain *of f and we write $M = \mathrm{Dom}\ f$.*

Remark 4.2. *In case $n = 2$ or $n = 3$ we use (x,y) or (x,y,z) instead of (x_1, x_2) or (x_1, x_2, x_3).*

Usually, the function will be given only by its formula without any specific domain. In that case, we assume that the domain is a maximal set for which has the formula sense. For example, a function
$$f(x,y) = \log(x+y)$$
is defined on a set
$$\mathrm{Dom}\ f = \{(x,y) \in \mathbb{R}^2,\ x + y > 0\}.$$

Example

- Find (and sketch) a maximal set $M \subset \mathbb{R}^2$ of such pairs (x,y) for which the function
$$f(x,y) = \frac{1}{\sqrt{x^2 + y - 1}}.$$
Necessarily, $\sqrt{x^2 + y - 1} > 0$ and we deduce that the function has sense for all pairs satisfying
$$x^2 + y - 1 > 0$$
which is a part of the plane bounded by certain parabola.

Definition 4.6. *Let $z = f(x,y)$ be a function of two variables. The* graph *of f is a set*
$$\mathit{graph} f = \{(x, y, f(x,y) \in \mathbb{R}^3,\ (x,y) \in \mathrm{Dom}\ f\}.$$

Definition 4.7. *A* contour line *C at height $z_0 \in \mathbb{R}$ is a set*
$$\{(x,y) \in \mathbb{R}^2,\ f(x,y) = z_0\}.$$

Example

- Find contour lines at heights $z_0 = -2, -1, 0, 1, 2$ for a function
$$f(x,y) = \frac{x^2+y^2}{2x}.$$
First of all, the domain of this function does not contain the y axis.
Take $z_0 = -2$. Then $f(x,y) = -2$ yields $(x+2)^2 + y^2 = 4$ and the contour line is the circle centered at $(-2,0)$ with radius $r = 2$ which do not contain the origin (because of the domain of f).
Similarly, For $z_0 = -1$ we get the circle centered at $(-1,0)$ with radius $r = 1$ which do not contain the origin.
The countour line at height $z_0 = 0$ is empty. For $z_0 = 1$ we get the circle centered at $(1,0)$ with radius $r = 1$ not containing the origin.
And finally, for $z_0 = 2$ the contour line is a circle of radius $r = 2$ with center at $(2,0)$ with exception of the origin.

Definition 4.8. *Let $M \subset \mathbb{R}^n$ and $f : M \to \mathbb{R}$. Next, let $\varphi : I \to M$ is a curve ($I \subset \mathbb{R}$ is an interval). Then $f \circ \varphi$ is a* cross-section *of f along φ.*

Examples

- What is the graph of a function
$$f(x,y) = (x+y)^2$$
on a line $p_a : (x,y) = (a,0) + t(1,1)$, $t \in \mathbb{R}$ for some $a \in \mathbb{R}$? And how about lines $q_b : (x,y) = (b,0) + t(1,-1)$, $t \in \mathbb{R}$ for some $b \in \mathbb{R}$?
First,
$$f(a+t,t) = (a+2t)^2$$
and the graph of f along line p_a is a convex parabola with vertex in $t_0 = -\frac{a}{2}$.
Next,
$$f(b+t,-t) = (b)^2$$
and the graph is a horizontal line at height b^2.

- Lets find the graph of a cross-section
$$f(x,y) = \frac{1}{x^2+y^2}$$
along lines
$$(x,y) = t(\cos\alpha, \sin\alpha), t \in (0,\infty)$$
where $\alpha \in [0, 2\pi)$ is a parameter. The function $g = f \circ \varphi$ is given as
$$g(t) = f(t\cos\alpha, t\sin\alpha) = \frac{1}{t^2(\cos^2\alpha + \sin^2\alpha)} = \frac{1}{t^2}.$$
Similarly as above, the sketch of the graph remains as an exercise for the kind reader.

Algebra of functions of two variables:
Sum, product and division is defined 'pointwisely'. Consider, for example, functions $f(x,y) = e^{xy}$ and $g(x,y) = \sqrt{1-x^2-y^2}$. Then

- $(f+g)(x,y) = e^{xy} + \sqrt{1-x^2-y^2}$,
- $(fg)(x,y) = e^{xy}\sqrt{1-x^2-y^2}$,
- $\frac{f}{g}(x,y) = \frac{e^{xy}}{\sqrt{1-x^2-y^2}}$. Beware, here we have to exclude from the domain all points where g equals zero.

Composition of functions: Let $M \subset \mathbb{R}^m$, $f : M \to \mathbb{R}^n$ (this means that there are n functions $f_i : M \to \mathbb{R}$, $i \in \{1, \dots, n\}$) and $g : \mathbb{R}^n \mapsto \mathbb{R}$. Then a *composition* is a function $h = g \circ f$ defined as

$$h(x_1, \dots, x_m) = g(f_1(x_1, \dots, x_m), f_2(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m)).$$

Similarly, if $f : M \mapsto \mathbb{R}$ and $g : \mathbb{R} \mapsto \mathbb{R}$ then $h = g \circ f$ is defined as $h(x_1, \dots, x_m) = g(f(x_1, \dots, x_m))$

We can also introduce the boundedness of a function $f : M \subset \mathbb{R}^n \to \mathbb{R}$. This can be done similarly to the one dimensional case. The precise definition of a bounded function is left as an exercise.

4.3 Continuity

Definition 4.9. *We say that $f : M \mapsto \mathbb{R}$ is continuous at a point $(x_0, y_0) \in M$ if*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall (x,y) \in (M \cap B_\delta(x_0, y_0))\,,\ |f(x,y) - f(x_0, y_0)| < \varepsilon.$$

Let $N \subset M$ and let $f : M \mapsto \mathbb{R}$ be continuous at all points $(x_0, y_0) \in N$. Then we say that f is continuous on N. If f is continuous on Dom f *then we simply say that f is continuous.*

Observation 4.2. *Let f_1 and f_2 be continuous functions. Then*

$$f_1 + f_2,\ f_1 - f_2 \text{ and } f_1 f_2$$

are continuous function. Moreover, $\frac{f_1}{f_2}$ is a continuous function on a set $\{(x,y) \in \mathbb{R}^2, f_2(x,y) \neq 0\}$. Further, $f_1 \circ f_2$ is also a continuous function. We remind that $f(x,y) = x$ and $f(x,y) = y$ are continuous function.

Example

- A function

  $$f(x,y) = \frac{x + \sqrt{x+y}}{1 + \cos^2 x}$$

  wherever it is correctly defined, this means a set

  $$\{(x,y) \in \mathbb{R}^2, y > -x\}.$$

4.4 Limits

Definition 4.10. *Let (x_0, y_0) be a limit point of $M \subset \mathbb{R}^2$ and let $f : M \mapsto \mathbb{R}$. We say that a limit of f at the point (x_0, y_0) is $A \in \mathbb{R}$ if*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall (x,y) \in (M \cap B_\delta(x_0, y_0)),\ |f(x,y) - A| < \varepsilon.$$

We write $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = A$.
We say that a limit of f at the point (x_0, y_0) is ∞ if

$$\forall M > 0,\ \exists \delta > 0,\ \forall (x,y) \in (M \cap B_\delta(x_0, y_0)),\ f(x,y) > M.$$

We write $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = \infty$.
We say that a limit of f at the point (x_0, y_0) is $-\infty$ if $\lim_{(x,y)\to(x_0,y_0)} -f(x,y) = -\infty$.

Observation 4.3 (Arithmetic of limits)**.** *Let f and g be two functions and let (x_0, y_0) be a limit point of* Dom f *and of* Dom g*. Then*

$$\lim_{(x,y)\to(x_0,y_0)} (f+g)(x,y) = \lim_{(x,y)\to(x_0,y_0)} f(x,y) + \lim_{(x,y)\to(x_0,y_0)} g(x,y)$$
$$\lim_{(x,y)\to(x_0,y_0)} fg(x,y) = \lim_{(x,y)\to(x_0,y_0)} f(x,y) \lim_{(x,y)\to(x_0,y_0)} g(x,y)$$
$$\lim_{(x,y)\to(x_0,y_0)} \frac{f}{g}(x,y) = \frac{\lim_{(x,y)\to(x_0,y_0)} f(x,y)}{\lim_{(x,y)\to(x_0,y_0)} g(x,y)}.$$

assuming the right hand side is well defined.

The numbers $\infty - \infty$, $0 \cdot \infty$, $\frac{0}{0}$, $\frac{\infty}{\infty}$ are not well defined (similarly to the one dimensional case).

Observation 4.4. *A function f is continuous at point $(x_0, y_0) \in$* Dom f *if and only if* $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0, y_0)$.

Example

- Consider a function
$$f(x,y) = \frac{x^2y^2}{x^2y^2 + (x-y)^2}.$$
This function is not defined at $(0,0)$. It is possible to define the value $f(0,0)$ in such a way that f is continuous? In particular, does there exists a finite limit
$$\lim_{(x,y)\to(0,0)} f(x,y)?$$
First, we approach $(0,0)$ along the line $y = 0$. We have
$$\lim_{(x,0)\to(0,0)} f(x,0) = \lim_{x\to 0} \frac{0}{x^2} = 0.$$
Next, we approach $(0,0)$ along the line $x = y$. We have
$$\lim_{(x,x)\to(0,0)} f(x,x) = \lim_{x\to 0} \frac{x^4}{x^4} = 1.$$
As a result, $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Lemma 4.1 (Sandwich lemma)**.** *Let f, g, h be three functions defined on $B_\delta(x_0, y_0) \setminus \{(x_0, y_0)\}$ for some $\delta > 0$. Assume*

$$\forall (x,y) \in B_\delta(x_0,y_0) \setminus \{(x_0,y_0)\},\ g(x,y) \le f(x,y) \le h(x,y).$$

If $\lim_{(x,y)\to(x_0,y_0)} g(x,y) = \lim_{(x,y)\to(x_0,y_0)} h(x,y) = A \in \mathbb{R}$ *then also*

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = A.$$

Corollary 4.1. $\lim_{(x,y)\to(x_0,y_0)} |f(x,y)| = 0 \Rightarrow \lim_{(x,y)\to(x_0,y_0)} f(x,y) = 0.$

Example

- Compute
$$\lim_{(x,y)\to(0,0)} \frac{xy}{\sqrt{x^2+y^2}}.$$
We use notation $f(x,y) = \frac{xy}{\sqrt{x^2+y^2}}$. First of all, we have $\lim_{x\to 0} f(x,0) = 0$ and $\lim_{y\to 0} f(0,y) = 0$. Thus, if there is a limit, it is equal to 0. We use the well known AM-GM inequality ($2|xy| \le (x^2+y^2)$) to deduce
$$0 \le \frac{|xy|}{\sqrt{x^2+y^2}} \le \frac{x^2+y^2}{\sqrt{x^2+y^2}} = \sqrt{x^2+y^2} \to 0$$
as $(x,y) \to 0$. The sandwich lemma yields $\lim_{(x,y)\to(0,0)} |f(x,y)| = 0$ and we have just proven that the given limit is equal to 0.

4.5 Derivatives

Definition 4.11. *Let $f : M \subset \mathbb{R}^n \to \mathbb{R}$ and $v \in \mathbb{R}^n$ be such that $\|v\| = 1$. Let $x_0 \in M^0$. The derivative of f with respect to direction v in a point x_0 is*
$$Df(x_0, v) = g'(t)|_{t=0} \text{ where } g(t) = f(x_0 + tv)$$

Remark 4.3. *The direction of an arbitrary vector v is a unit vector $\frac{v}{\|v\|}$.*

Examples

- What is the direction of a line $p : (x,y) = (2,-1) + t(1,3)$? The size of $(1,3)$ is $\sqrt{1^2+3^2} = \sqrt{10}$. Consequently, the direction of the line is $\left(\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\right)$.
- Let $f(x,y) = x^2 e^y$. Let compute $Df\left((1,0), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right)$. The line $p(t)$ passing through $(1,0)$ with the demanded direction has expression
$$p(t) = \left(1 + \frac{t}{\sqrt{2}}, \frac{t}{\sqrt{2}}\right).$$
Thus
$$Df\left((1,0), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right) = \left(\left(1 + \frac{t}{\sqrt{2}}\right)^2 e^{\frac{t}{\sqrt{2}}}\right)'\Bigg|_{t=0} = 1$$

Definition 4.12. *We define* partial derivatives *with respect to x_i as*
$$\frac{\partial f}{\partial x_i}(x_0) = \lim_{h\to 0} \frac{f(x_0 + he_i) - f(x_0)}{h}.$$
where e_i is the vector whose $i-th$ component is 1 and all other components are zero.

Remark 4.4. *It holds that*
$$\frac{\partial f}{\partial x}(x,y) = Df((x,y),(1,0)), \ \frac{\partial f}{\partial y}(x,y) = Df((x,y),(0,1))$$
whenever f is a function of two variables. Similarly, one can deduce the same rule also for a function of n variable.

Definition 4.13. *Let $x_0 \in \mathrm{Dom}\, f \subset \mathbb{R}^n$. A vector of first partial derivatives*

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0) \right)$$

is called the gradient *of f at x_0.*

Example

- Let compute $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ for a function

  $$f(x, y) = 3x^2 y + x^2 + \log(x^2 + y^2).$$

  Let first compute $\frac{\partial f}{\partial x}$. In that case we treat y as a constant and we deduce that

  $$\frac{\partial f}{\partial x} = 6xy + 2x + \frac{2x}{x^2 + y^2}.$$

  In order to compute $\frac{\partial f}{\partial y}$ we treat x as a constant and we get

  $$\frac{\partial f}{\partial y} = 3x^2 + \frac{2y}{x^2 + y^2}.$$

  We remark that in this case we have

  $$\nabla f(x, y) = \left(6xy + 2x + \frac{2x}{x^2 + y^2}, 3x^2 + \frac{2y}{x^2 + y^2} \right).$$

Definition 4.14. *We define second order partial derivatives as follows*

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right), \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right)$$

whenever $i, j \in \{1, \ldots, n\}$, $i \neq j$. Analogously we define the third and higher order partial derivatives. The matrix of second derivatives

$$(\nabla^2 f) = \left(\frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} \right)_{i,j=1}^{n}$$

is called the Hess matrix.

Example

- Let compute the first and second order derivatives for $f(x, y) = \frac{x}{y} - e^{xy}$. We have

  $$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{1}{y} - ye^{xy}, \ \frac{\partial f}{\partial y} = -\frac{x}{y^2} - xe^{xy} \\ \frac{\partial^2 f}{\partial x^2} &= -y^2 e^{xy}, \ \frac{\partial^2 f}{\partial y \partial x} = -\frac{1}{y^2} - e^{xy} - xye^{xy} \\ \frac{\partial^2 f}{\partial y^2} &= 2\frac{x}{y^3} - x^2 e^{xy}, \ \frac{\partial^2 f}{\partial x \partial y} = -\frac{1}{y^2} - e^{xy} - xye^{xy}. \end{aligned}$$

  The corresponding Hess matrix is

  $$\nabla^2 f = \begin{pmatrix} -y^2 e^{xy} & -\frac{1}{y^2} - e^{xy} - xye^{xy} \\ -\frac{1}{y^2} - e^{xy} - xye^{xy} & 2\frac{x}{y^3} - x^2 e^{xy} \end{pmatrix}$$

Observation 4.5. *Let the second order derivative of a function f be continuous. Then*

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Theorem 4.1 (Chain rule – derivative of a composed function)**.** *Let $n = 1$ or 2 and let $f : \mathbb{R}^n \to \mathbb{R}^2$ and $g : \mathbb{R}^2 \to \mathbb{R}$. Then*

$$\frac{\partial (g \circ f)}{\partial x_i} = \frac{\partial g}{\partial y_1}\frac{\partial f_1}{\partial x_i} + \frac{\partial g}{\partial y_2}\frac{\partial f_2}{\partial x_i}, \ i = \{1, n\}.$$

Example

- Let $f(x) = g(\sin x, \cos x)$. Then

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial a}\cos x - \frac{\partial g}{\partial b}\sin x$$

  where we use a notation $g = g(a, b)$.

- Let $f(x, y) = \sqrt{x^2 - y^2}$ and let $x = x(t) = e^{2t}$ and $y = e^{-t}$. Let compute $\frac{\partial f(x(t),y(t))}{\partial t}$:

$$\begin{aligned}\frac{\partial f(x(t), y(t))}{\partial t} &= \frac{\partial f}{\partial x}\bigg|_{(x(t),y(t))}\frac{\partial x(t)}{\partial t} + \frac{\partial f}{\partial y}\bigg|_{(x(t),y(t))}\frac{\partial y(t)}{\partial t} \\ &= \frac{x}{\sqrt{x^2 - y^2}}\Bigg|_{(e^{2t},e^{-t})} 2e^{2t} + \frac{-y}{\sqrt{x^2 - y^2}}\Bigg|_{(e^{2t},e^{-t})}(-e^{-t}) = \frac{2e^{4t} + e^{-2t}}{\sqrt{e^{4t} - e^{-2t}}}.\end{aligned}$$

4.6 Differential

Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$. We try to compute an increment of a function if we move from the point (x_0, y_0) to the point $(x_0 + h, y_0 + k)$, i.e., $\Delta f(x_0, y_0) = f(x_0 + h, y_0 + k) - f(x_0, y_0)$. It can be written as

$$\Delta f(x_0, y_0) = f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) + f(x_0 + h, y_0) - f(x_0, y_0).$$

Assuming $|h|$ and $|k|$ are sufficiently small we can us an approximation

$$\begin{aligned} f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) &\sim \frac{\partial f}{\partial x}(x_0 + h, y_0)k \\ f(x_0 + h, y_0) - f(x_0, y_0) &\sim \frac{\partial f}{\partial y}(x_0, y_0)h \end{aligned}$$

Moreover, $\frac{\partial f}{\partial x}(x_0 + h, y_0) \sim \frac{\partial f}{\partial x}(x_0, y_0)$ if[1] $f \in C^1$. This yields

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) \sim \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k.$$

We denote by $\mathrm{d}x$ the change in the x coordinate and $\mathrm{d}y$ the change in the y coordinate.

[1]Here $f \in C^1$ means that f has continuous first partial derivatives.

Definition 4.15. *Let $f \in C^1$. Then*

$$\mathrm{d}f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\mathrm{d}x + \frac{\partial f}{\partial y}(x_0, y_0)\mathrm{d}y$$

is called the differential *of f at the point (x_0, y_0).*

The differential of a function can be used to determine approximate values. Let for example determine $\sqrt{(0.03)^2 + (2.89)^2}$. Consider a function $f(x, y) = \sqrt{x^2 + y^2}$. We have $\nabla f = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right)$. We choose $x_0 = 0$ and $y_0 = 3$. We have $\mathrm{d}x = 0.03$ and $\mathrm{d}y = -0.11$. It holds that

$$\sqrt{(0.03)^2 + (2.89)^2} \sim \sqrt{0^2 + 3^2} + 0 \cdot 0.03 + 1 \cdot (-0.11) = 2.89$$

Remark 4.5. *It is worth to mention that $\mathrm{d}f(x_0, y_0) = \nabla f(x_0, y_0) \cdot (\mathrm{d}x, \mathrm{d}y)$. This allows to generalize the above notion also for functions of more variables. In particular, if $f : \mathbb{R}^n \to \mathbb{R}$, then*[2]

$$\mathrm{d}f = \nabla f \cdot (\mathrm{d}x_1, \mathrm{d}x_2, \ldots, \mathrm{d}x_n).$$

Definition 4.16. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ have continuous partial derivatives at point (x_0, y_0). Then the* tangent plane *of the graph of f at point (x_0, y_0) is a plane with equation*

$$z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0).$$

Example

- Let compute a tangent plane of the graph of f at point $(1, 2)$ for $f(x, y) = \sqrt{9 - x^2 - y^2}$. We have

  $$\nabla f(x, y) = \left(-\frac{x}{\sqrt{9 - x^2 - y^2}}, -\frac{y}{\sqrt{9 - x^2 - y^2}}\right)$$

  and $\nabla f(1, 2) = (-1/2, -1)$. Thus, the tangent plane is

  $$z = 2 - 1/2(x - 1) - 1(y - 2) = 9/2 - x/2 - y.$$

4.7 The Taylor polynomial

An approximation by a differential is deduced above. In particular

$$f(x, y) \sim f(x_0, y_0) + \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0). \tag{7}$$

Recall that we use it to compute $\sqrt{(0.03)^2 + (2.89)^2}$.
The above considerations leads to the definition of the first-order Taylor polynomial at a point (x_0, y_0) as[3]

$$T_1(x, y) = f(x_0) + \nabla f(x_0) \cdot (x - x_0)$$

whenever $f : \mathbb{R}^n \to \mathbb{R}$. If f is a function of two variables then the graph of T_1 is also a tangent plane to the graph of the function f at the point (x_0, y_0) and it is the only plane which is the best approximation of the function near the point (x_0, y_0).

[2] And here $u \cdot v$ is a scalar multiplication of two vectors with same dimension. It can be understood as a matrix multiplication $u \cdot v^T$.

[3] As above, $\nabla f(x_0) \cdot (x - x_0)$ is a scalar product and it can be seen as a multiplication of two matrices, in particular, $\nabla f(x, y) \cdot (x - x_0)^T$.

Definition 4.17. *Let $f : M \subset \mathbb{R}^n \to \mathbb{R}$ and $x_0 \in M$. We define the second order Taylor polynomial at a point x_0 as*

$$T_2(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0)(\nabla^2 f(x_0))(x - x_0)^T.$$

Let us just remind that the last term is actually the quadratic form considered in Chapter 2.7

Examples

- Let compute the second order Taylor polynomial of $f(x, y) = \sqrt{9 - x^2 - y^2}$ (the function from the previous exercise) at $(1, 2)$. First, we have

$$\nabla^2 f(x,y) = \begin{pmatrix} \frac{-\sqrt{9-x^2-y^2}+\frac{x^2}{\sqrt{9-x^2-y^2}}}{9-x^2-y^2} & -\frac{xy}{\sqrt{9-x^2-y^2}^3} \\ -\frac{xy}{\sqrt{9-x^2-y^2}^3} & \frac{-\sqrt{9-x^2-y^2}+\frac{y^2}{\sqrt{9-x^2-y^2}}}{9-x^2-y^2} \end{pmatrix}.$$

  Therefore

$$\nabla^2 f(1,2) = \begin{pmatrix} \frac{-3}{8} & -1 \\ -1 & 0 \end{pmatrix}$$

  and thus

$$\begin{aligned} T_2(x) &= f(1,2) + \nabla f(1,2) \cdot (x-1, y-2) + \frac{1}{2}(x-1, y-2)(\nabla^2 f(1,2))(x-1, y-2) \\ &= 2 + \left(-\frac{1}{2}, -1\right) \cdot (x-1, y-2) + \frac{1}{2}(x-1, y-2) \begin{pmatrix} \frac{-3}{8} & -1 \\ -1 & 0 \end{pmatrix} (x-1, y-2) \\ &= 2 - \frac{1}{2}x + \frac{1}{2} - y + 2 + \frac{1}{2}\left(-\frac{3}{8}(x-1)^2 - 2(x-1)(y-2)\right). \end{aligned}$$

- We compute an approximate value $\sqrt{(0.03)^2 + (2.89)^2}$ with the help of the second order Taylor polynomial. We choose $(x_0, y_0) = (0, 3)$ and we use notation $f(x, y) = \sqrt{x^2 + y^2}$. We have $\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2+y^2}}$, $\frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2+y^2}}$, $\frac{\partial^2 f}{\partial x^2} = \frac{y^2}{(x^2+y^2)^{3/2}}$, $\frac{\partial^2 f}{\partial y^2} = \frac{x^2}{(x^2+y^2)^{3/2}}$, $\frac{\partial^2 f}{\partial x \partial y} = \frac{-xy}{(x^2+y^2)^{3/2}}$. We deduce that T_2 at $(0, 3)$ is

$$T_2(x,y) = 3 + (y - 3) + \frac{1}{6}x^2$$

  We get $T_2(0.03, 2.89) = 3 + (-0.11) + \frac{1}{6}0.0009 = 2.89015$.

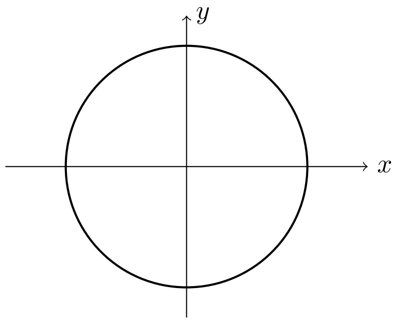
4.8 Implicit functions

Consider a set

$$\{(x,y) \in \mathbb{R}^2,\ x^2 + y^2 = 1\}$$

The equation $x^2 + y^2 = 1$ defines two function $y_1(x)$ and $y_2(x)$ where

$$\begin{aligned} y_1(x) &= \sqrt{1-x^2},\ \text{Dom}\ y_1(x) = [-1, 1], \\ y_2(x) &= -\sqrt{1-x^2},\ \text{Dom}\ y_2(x). \end{aligned}$$



What if it is impossible to express y? Consider an equation

$$f(x, y) = 0.$$

What assumptions should be imposed in order to get uniquely defined function $y(x)$?

Theorem 4.2. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ and $(x_0, y_0) \in \mathbb{R}^2$ be given. If*

i) $f \in C^k$ *for some* $k \in \mathbb{N}$,

ii) $f(x_0, y_0) = 0$,

iii) $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$,

Then there is a uniquely determined function $y(x)$ of class C^k on a neighborhood of point x_0 such that $f(x, y(x)) = 0$ (precisely, there is $\epsilon > 0$ and a function $y(x)$ defined on $(x_0 - \varepsilon, x_0 + \varepsilon)$ such that $f(x, y(x)) = 0$.

Example Consider an equation

$$x^3 + y^3 - 3xy - 3 = 0.$$

Is there a function $y(x)$ determined by the given equation on the neighborhood of a point $(1, 2)$? According to the previous theorem, we have to verify three assumptions:
1, the function $f(x, y) = x^3 + y^3 - 3xy - 3$ should belong (at least) to C^1. That is true since $f(x, y)$ is a polynomial.
2, $f(1, 2)$ should be equal to zero (or, equivalently, the given equation should be satisfied at the given point). This is also true.
3, $\frac{\partial f}{\partial y} = 3y^2 - 3x$ and therefore $\frac{\partial f}{\partial y}(1, 2) = -3 \neq 0$ and the last assumption is also true.
As a result, there is a function $y(x)$ uniquely determined by the given equation in some neighborhood of point $x = 1$, $y = 2$.

Note that the last assumption in the implicit function theorem cannot be omited. Consider the first equation

$$x^2 + y^2 = 1$$

and let decide whether there is a function $y(x)$ given by that equation at the point $(1, 0)$. According to the picture, it is impossible (recall the vertical line test). The theorem may not be applied. Take $f(x, y) = x^2 + y^2 - 1$. We have

$$\frac{\partial f}{\partial y} = 2y, \ \frac{\partial f}{\partial y}(1, 0) = 0$$

and the third assumption is not fulfilled.
Or another example, consider a set

$$\{(x, y) \in \mathbb{R}^2,\ x^2 - y^2 = 0\}.$$

Is this set a graph of some function around a point $(0, 0)$? Once again, we have $f(x, y) = x^2 - y^2$, $\frac{\partial f}{\partial y} = -2y$ and the last assumption of the implicit function theorem is not fulfilled.

Further analysis of the implicitly given function
In order to examine further qualitative properties of the given function we have to compute derivatives at the given points. The easiest method is to differentiate the given equation with respect to x (and to assume that y is in fact a function of x).
Example: Consider an equation

$$e^{2x} + e^y + x + 2y - 2 = 0.$$

This defines on a neighborhood of $(0, 0)$ a function $y(x)$. Indeed, let $f(x, y) = e^{2x} + e^y + x + 2y - 2$. Then f is of class C^k for every $k \in \mathbb{N}$, $f(0, 0) = 0$ and $\frac{\partial f}{\partial y} = e^y + 2$ which yields $\frac{\partial f}{\partial y}(0, 0) = 3 \neq 0$. Let compute $y'''(0)$ (note that the third derivative exists as $f \in C^3$).

Let differentiate the equation with respect to x. We have

$$2e^{2x} + e^y y' + 1 + 2y' = 0$$

and we plug here $x = 0$ and $y = 0$ in order to get

$$2 + y'(0) + 1 + 2y'(0) = 0$$

which yields $y'(0) = -1$.
We differentiate once again with respect to x to get

$$4e^{2x} + e^y y'^2 + e^y y'' + 2y'' = 0$$

and we plug here $x = 0$, $y = 0$ and $y' = -1$. We get

$$4 + 1 + 3y''(0) = 0$$

yielding $y''(0) = -\frac{5}{3}$. We differentiate the equation for the third time in order to get

$$8e^{2x} + e^y y'^3 + e^y 2y'y'' + e^y y'y'' + e^y y''' + 2y''' = 0$$

and once again we plug there $x = 0$, $y = 0$, $y' = -1$ and $y'' = -\frac{5}{3}$. We get

$$8 - 1 + \frac{10}{3} + \frac{5}{3} + 3y''' = 0$$

which gives

$$y'''(0) = -4.$$

In particular, we may write

$$0 = \frac{\partial f(x, y(x))}{\partial x} = \frac{\partial f(x, y)}{\partial x} + \frac{\partial f(x, y)}{\partial y}\frac{\partial y}{\partial x}$$

which gives

$$y'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)}.$$

4.9 Extremes

Similarly to the one-dimensional case, we talk about local and global extremes.

Definition 4.18. *Let $f : M \subset \mathbb{R}^n \to \mathbb{R}$. We say that f attains a* local maximum *at a point $x_0 \in M^0$ if there is $r > 0$ such that $f(x_0) \geq f(x)$ for all $x \in B_r(x_0)$.*
We say that f attains a local minimum *at a point $x_0 \in M^0$ if there is $r > 0$ such that $f(x_0) \leq f(x)$ for all $x \in B_r(x_0)$.*

Definition 4.19. *Let $f : M \subset \mathbb{R}^n \to \mathbb{R}$. We say that f attains its maximum on M at a point $x_0 \in M$ if $f(x_0) \geq f(x)$ for all $x \in M$. Similarly, f attains its minimum on M at a point $x_0 \in M$ if $f(x_0) \leq f(x)$ for all $x \in M$.*

4.9.1 Local extremes

Assume $f \in C^1$. Let f has a local extrem at (x_0, y_0). Then $g(x) = f(x, y_0)$ has also a local extreme at x_0 and, therefore, $g'(x_0) = 0$. Similarly, $h(y) = f(x_0, y)$ has a local extreme at y_0 and thus $h'(y_0) = 0$. This leads to the following observation.

Observation 4.6. *Let $f \in C^1$ have a local extreme at x_0. Then $\nabla f(x_0) = 0$.*

Definition 4.20. *A point $x_0 \in \mathrm{Dom}\ f$ such that $\nabla f(x_0) = 0$ is called a* stationary point.

How to find all local extremes of given function?
Step 1: determine the stationary point.
Step 2: examine the possible extremes in the stationary point.
Reminder: in the one-dimensional case one has to treat the sign of the second derivative in order to decide if there is an extreme in a stationary point.

Example Let find all stationary points of $f(x, y) = x^2 - y^2$. We have $\nabla f(x, y) = (2x, -2y)$ and therefore the only stationary point is $(x_0, y_0) = (0, 0)$. Is there a maximum or minimum?

Observation 4.7. *Let $f \in C^2$ and let x_0 be its stationary point. Then:*

1. *If $\nabla^2 f$ is positive definite, then f attains a local minimum at x_0,*
2. *If $\nabla^2 f$ is negative definite, then f attains a local maximum at x_0.*
3. *If $\nabla^2 f$ is indefinite, then f does not have an extreme at x_0 (saddle point).*
4. *Otherwise, we do not know anything.*

Example: Let go back to $f(x, y) = x^2 - y^2$. We already know that $(x_0, y_0) = (0, 0)$ is a stationary point. We have

$$\nabla^2 f = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Thus $\det \nabla^2 f(0, 0) = -4$ and there is no extreme at $(0, 0)$.

Another example Determine all local extremes of

$$f(x, y) = x^3 + 3xy^2 - 15x - 12y.$$

Step 1, stationary points:

$$\nabla f(x, y) = (3x^2 + 3y^2 - 15, 6xy - 12)$$

and we stationary points are solutions to

$$3x^2 + 3y^2 - 15 = 0$$
$$6xy - 12 = 0$$

which is equivalent to

$$x^2 + y^2 - 5 = 0$$
$$xy = 2.$$

We deduce from the second equation that x and y are different from zero. The second equation yields $x = \frac{2}{y}$. We plug this into the first equation to deduce

$$\frac{4}{y^2} + y^2 - 5 = 0$$

which is equivalent to

$$y^4 - 5y^2 + 4 = 0.$$

We have $y^2 = 4$, $y^2 = 1$ and therefore there are four stationary points

$$A = (-1, -2),\ B = (1, 2),\ C = (2, 1),\ D = (-2, -1).$$

Step 2: We have

$$\nabla f = \begin{pmatrix} 6x & 6y \\ 6y & 6x \end{pmatrix}$$

Further,

$$\nabla^2 f(A) = \begin{pmatrix} -6 & -12 \\ -12 & -6 \end{pmatrix},\ \det \nabla^2 f(A) = -108$$

and A is a saddle point.

$$\nabla^2 f(B) = \begin{pmatrix} 6 & 12 \\ 12 & 6 \end{pmatrix},\ \det \nabla^2 f(B) = -108$$

and B is a saddle point.

$$\nabla^2 f(C) = \begin{pmatrix} 12 & 6 \\ 6 & 12 \end{pmatrix},\ \det \nabla^2 f(C) = 108$$

and C is a point of a local minimum. The value of the local minimum is $f(C) = -28$.

$$\nabla^2 f(D) = \begin{pmatrix} -12 & -6 \\ -6 & -12 \end{pmatrix},\ \det \nabla^2 f(D) = 108$$

and D is a point of a local maximum. The value of the local maximum is $f(D) = 28$.

The least square method

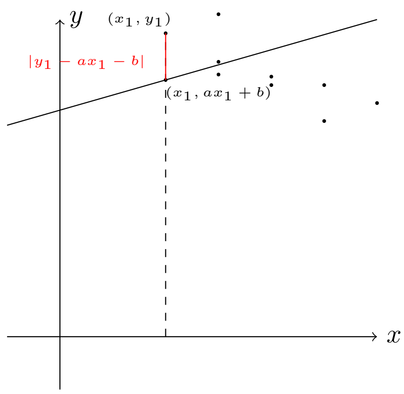
We will solve the following exercise: Assume that the cost of a car (of one given type) depends linearly on its age, i.e.,

$$y = ax + b,\ a, b \in \mathbb{R},$$

where y is the price of a car and x is its age.
Our aim now is to determine this function (constants a and b) from the given sets of data. Below we have a table of particular cars (their price does not follow strictly the above rule since the price come from the free market)

x	2	3	3	3	4	4	5	5	6
y	28.7	24.8	26.0	30.5	23.8	24.6	23.8	20.4	22.1

To find the line which fits best to the given data, we use the least squares method. This means that we are going to minimize the 'distance' between the line $ax + b$ and the given data. We define such distance as sum of squares:



$$|y_1 - ax_1 - b|^2 + |y_2 - ax_2 - b|^2 + \ldots + |y_n - ax_n - b|^2 = \sum_{i=1}^{n} |y_i - ax_i - b|^2.$$

This sum of squares in infact a function f of variables a and b of the form

$$f(a, b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

and we are going to minimize this sum of squares. We compute the partial derivative

$$\frac{\partial f}{\partial a} = -2 \sum_{i=1}^{n} (y_i - ax_i - b)x_i, \quad \frac{\partial f}{\partial b} = -2 \sum_{i=1}^{n} (y_i - ax_i - b).$$

and we deduce that the stationary point of this function has to fulfill

$$\sum_{i=1}^{n} (y_i - ax_i - b)x_i = 0$$
$$\sum_{i=1}^{n} (y_i - ax_i - b) = 0.$$

Recall that unknowns are a and b. We reformulate this into

$$\left(\sum_{i=1}^{n} x_i^2\right) a + \left(\sum_{i=1}^{n} x_i\right) b = \sum_{i=1}^{n} x_i y_i$$
$$\left(\sum_{i=1}^{n} x_i\right) a + nb = \sum_{i=1}^{n} y_i.$$

Recall our example

x	2	3	3	3	4	4	5	5	6
y	28.7	24.8	26.0	30.5	23.8	24.6	23.8	20.4	22.1

where we have

$$n = 9,\ \sum_{i=1}^{9} x_i = 35,\ \sum_{i=1}^{9} x_i^2 = 149,\ \sum_{i=1}^{9} y_i = 224.7,\ \sum_{i=1}^{9} x_i y_i = 848.5.$$

We and up with equation

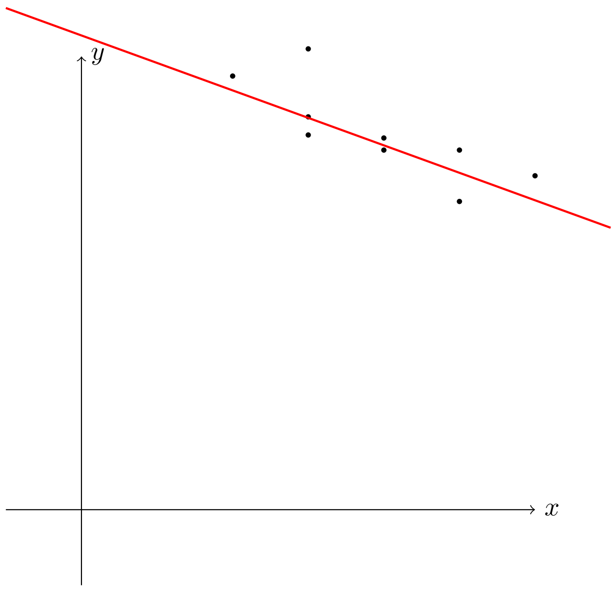
$$149a + 35b = 848.5$$
$$35a + 9b = 224.7$$

which has (approximate) solution

$$a = -2.02, \quad b = 32.8.$$

Thus, the desired line has equation

$$y = -2.02x + 32.8$$


y
x

Index