Metric spaces. Differential calculus of functions of several real variables.

Lecturer Aleksandr Rotkevich, fall 2023.

1 Metric spaces.

1.1 Definitions and examples.

Definition 1.1. *A metric space is a pair* (X, ρ)*, where* X *is a set, and a function* $\rho : X \times X \to [0, +\infty)$ *satisfies the following three axioms. For every* $x, y, z \in X$

1. $\rho(x, y) = 0$ *if and only if* $x = y$ ***[Identity axiom]****;*
2. $\rho(x, y) = \rho(y, x)$ ***[Symmetry axiom]****;*
3. $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ ***[Triangle inequality]****.*

Elements of X *are usually points. The function* ρ *that satisfies* $(1-3)$ *is called metric. The metric defines a concept of distance between any two points of* X*.*

Examples 1.2. 1. *Let* X *be any set. The function*

$$\rho(x, y) = \begin{cases} 1, & x \neq y \\ 0, & x = y \end{cases}$$

is called a discrete metric. And defines a discrete space (X, ρ)*.*

2. $X = \mathbb{R}$ *or* $X = \mathbb{C}$*,* $\rho = |x - y|$ *.*

3. $X = \mathbb{R}^n$ *ore* $X = \mathbb{C}^n$, $\rho(x,y) = \rho_2(x,y) = \sqrt{\sum\limits_{k=1}^{n}(x_k - y_k)^2}$ - *is a real (complex) euclidean space.*

4. $X = \mathbb{R}^n$ *or* $X = \mathbb{C}^n$, $\rho_p(x,y) = \left(\sum\limits_{k=1}^{n} |x_k - y_k|^p\right)^{1/p}, \ 1 \le p < \infty.$

5. $X = \mathbb{R}^n$ *or* $X = \mathbb{C}^n$, $\rho_\infty(x,y) = \max\limits_{1\le k\le n} |x_k - y_k|.$

6. *Let* $X = \mathcal{B}[a,b]$ *denote the set of bounded functions on the interval* $[a,b]$ *or* $X = C[a,b]$. *For* $f, g \in \mathcal{B}[a,b]$ *let*

$$\rho(f,g) = \sup_{x\in[a,b]} |f(x) - g(x)|.$$

Then ρ *is the metric on* X.

7. *Let* $X = R[a,b]$ *be the set of Riemann integrable functions or* $X = C[a,b]$ *then*

$$\rho(f,g) = \left(\int_a^b |f-g|^p\right)^{1/p}, \ p \ge 1,$$

defines a metric on X.

1.2 Open and closed sets.

Definition 1.3. *Let* (X, ρ) *be a metric space,* $D \subset X$, $a \in X$. *Let* $r > 0$. *Then*

- $B(a,r) = V_a(r) = V_a = \{x \in X : \rho(a,x) < r\}$ *is the open ball (neighborhood) of radius* r *around* a;

- $\overline{B}(a,r) = \{x \in X : \rho(a,x) \leq r\}$ *is the closed ball of radius* r *around* a;

- $S(a,r) = \{x \in X : \rho(a,x) = r\}$ *is the sphere of radius* r *around* a.

Definition 1.4. *A point* a *is called* ***interior point*** *of a set* D *if there exists an open ball centered at* a *which is completely contained in* D.
A set D *is called* ***open*** *if every point of* D *is interior.*

Examples 1.5. X, $\varnothing$, $B(a,r)$ *are open sets.*

Theorem 1.6 (Properties of open sets.)**.** *1. The union of any family of open sets is open.*

2. The intersection of a finite number of open sets is open.

3. The intersection of an infinite number of open sets is not always open.

Proof. **1.** Let $\{U_\alpha\}_{\alpha\in A}$ be a family of open sets and $U = \bigcup\limits_{\alpha\in A} U_\alpha$. Consider $p \in U$ then there exists $\alpha \in A$ such that $p \in U_\alpha$. Since U_α is open there exists a neighborhood V_p of point p such that $p \in V_p \subset U_\alpha \subset U$. Consequently p is the interior point of U.
2. Let $\{U_k\}_{k=1}^n$ be a finite family of open sets and $U = \bigcap\limits_{k=1}^{n} U_k$. Consider $p \in U$ then $p \in U_k$ for every $k = 1..n$ and for every k there exists

a neighborhood $V_p(r_k)$ of a point p contained in U_k. Consequently, $p \in V_p(r) = \bigcap_{k=1}^{n} V_p(r_k)$, where $r = min\{r_k : 1 \leq k \leq n\} \subset U$ and p is interior point of U.
3. Consider

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}.$$

□

Definition 1.7. *A set*

$$\text{int}\, D = D^{\circ} = \{p \in D : p \textit{ is interior point of} D\}$$

is called the interior of D.

Theorem 1.8.
- int$D$ *is the largest open subset of D.*
- int$D$ *is the union of all open subsets of D.*
- *A set is open if and only if it coincides with its interior.*

Proof. **1.** Let U be open set such that int$D \subset U \subset D$. Let $p \in U$ then there exist a neighborhood V_p of a point p contained in U which is a subset of D. Consequently p is the interior point of D and $p \in$ int$D$. Thus $U =$ int$D$.
2. Let U be open subset of D then $U \subset$ int$D$. Also int$D$ is itself open, consequently, int$D$ is a subset of union of open subset of D and also contains this union.
3. If D is open then it is the largest open subset of itself, consequently $D =$ int$D$. If $D =$ int$D$ then it is open since int$D$ is open. □

Definition 1.9. *A point a is called a* ***limit point*** *(or* ***cluster point*** *or* ***accumulation point****) of a set D if every neighborhood of a contains a point of D other than a itself that is $D \cap V_a \setminus \{a\} = D \cap \dot{V}_a \neq \varnothing$ for every V_a.*
A set $\dot{V}_a = V_a \setminus \{a\}$ is called a ***punctured neighborhood*** *of a.*

Lemma 1.10. *Let a be a limit point of a set D, then the intersection of any its (punctured) neighborhood with a set D is infinite.*

Proof. Let a be a limit point. Then for $r = 1$ there exists $x_1 \in D \cap \dot{V}_a(1)$. Let $r_2 = \rho(a, x_1)$. Then there exists $x_2 \in D \cap \dot{V}_a(r_2)$. Assume that $x_1, \ldots, x_n$ are already chosen. Let $r_n = \rho(a, x_n)$ and choose $x_{n+1} \in D \cap \dot{V}_a(r_2)$. Consequently, $x_n \neq x_m$ if $m \neq n$ and every neighborhood of a contains infinitely many points of a set D. □

Definition 1.11. *A set is* ***closed*** *if it contains all its limit points.*

Examples 1.12. $X,\ \varnothing,\ \{a\},\ \overline{B}(a, r)$.

Theorem 1.13. *A set D is open if and only if $X \setminus D$ is closed.*

Proof. Let D be closed and consider $a \in X \setminus D$. Then a is not a limit point of D and there exists a neighborhood V_a that doesn't contains points of D. Consequently, $V_a \subset X \setminus D$.

Assume that $X \setminus D$ is open. Let $a \in X$ be a limit point of D. Then every neighborhood of a contains points of D. Consequently $a \notin X \setminus D$ and $a \in D$. This implies that D is a closed set. □

Lemma 1.14. *Properties of closed sets.*

1. *The intersection of any family of closed sets is closed.*

2. *The union of a finite number of closed sets is closed. The union of an infinite number of closed sets is not always closed.*

Proof. The proof follows from the De Morgan's Law

$$\left(\bigcap_{\alpha\in A} G_\alpha\right)^c = \bigcup_{\alpha\in A} G_\alpha^c$$

and properties of open sets (Lemma 1.6). □

Definition 1.15. *The* ***closure of a set*** *is a set that consists of points of a set and of the limit points of a set. Notation:* $\mathrm{cl}(D)$ *or* $\overline{D}$.

Lemma 1.16.

- $\mathrm{cl}\, D$ *is the smallest closed set that contains* D.
- $\mathrm{cl}\, D$ *is the intersection of all closed sets that contain* D.
- *A set is closed if and only if it coincides with its closure.*

Definition 1.17. *A point* a *is a* ***boundary point*** *of a set* D *if any neighborhood of* a *contains both point of* D *and a point of* $X \setminus D$. *A set of boundary points of* D *is called* ***a boundary of*** D *and is denoted by* ∂D.

Exercises.

1. $\partial D = \mathrm{cl}(D) \setminus \mathrm{int}(D)$
2. ∂D is closed.
3. $\mathrm{int}(D^c) = \mathrm{cl}(D^c)$.
4. Provide an example of a metric space (X, d) and a set $D \subset X$ such that $\mathrm{int}\, D = \varnothing$, $\mathrm{cl}\, D = X$.

2 Limit of a sequence.

Definition 2.1. *A point $a \in X$ is called* ***a limit of a sequence*** $\{x_n\}_{n=1}^{\infty}$ *(in metric space (X, d)), if* $\lim\limits_{n\to\infty} d(x_n, a) \to 0$.

- A sequence that has a limit is called **convergent**. If a sequence has no limit it is called **divergent**.

- Notations: $a = \lim\limits_{n\to\infty} x_n = \lim x_n,\ x_n \to a \ldots$

- A limit is unique, that is if $a = \lim\limits_{n\to\infty} x_n$ and $b = \lim\limits_{n\to\infty} x_n$ then $a = b$.

  Proof. $0 \le d(a, b) \le d(a, x_n) + d(x_n, b) \to 0$. □

- A convergent sequence is bounded (a set is bounded if it is contained in some ball).

  Proof. Let $d(x_n, a) \to 0$. Then a sequence $\{d(x_n, a)\}$ is bounded and $r = 2 \sup d(a, x_n) < \infty$. Finally,

  $$\{x_n\} \subset B(a, r), \quad r = 2 \sup d(a, x_n) < \infty.$$

  □

Lemma 2.2. *Let (X, d) be a metric space, $\{x_n\}$ be a sequence of points of X and $a \in X$. The following assertions are equivalent.*

1. $a = \lim x_n$;

2. *For every $\varepsilon > 0$ there exists number $N \in \mathbb{N}$ such that $d(x_n, a) < \varepsilon$ for every $n \ge N$.*

3. *For every neighborhood V_a of a point a there exists number $N \in \mathbb{N}$ such that $x_n \in V_a$ for every $n \geq N$.*

Proof. $1 \Leftrightarrow 2$ since 2 is the reformulation of the condition $d(x_n, a) \to 0$.

$2 \Rightarrow 3$ Let $V_a = B(a, \varepsilon)$. Then there exists number $N \in \mathbb{N}$ such that $d(x_n, a) < \varepsilon$ for every $n \geq N$. Consequently, $x_n \in V_a$ for every $n \geq N$.

$3 \Rightarrow 2$ Let $\varepsilon > 0$. Then there exists number $N \in \mathbb{N}$ such that $x_n \in V_a = B(a, \varepsilon)$ for every $n \geq N$. Consequently, $d(x_n, a) < \varepsilon$ for every $n \geq N$. □

Lemma 2.3. *Let (X, d) be a metric space, $D \subset X$, $a \in X$. A point a is a limit point of D if and only if there exists a sequence $\{x_n\} \subset D \setminus \{a\}$ that converges to a.*

Proof. $\Rightarrow$. Let a be a limit point of D. Then for every $n \in \mathbb{N}$ there exists a point $x_n \in D \setminus \{a\}$ such that $x_n \in B\left(a, \frac{1}{n}\right) \cap D \setminus \{a\}$. Consequently, $d(a, x_n) < \frac{1}{n}$ and $x_n \to a$.
$\Leftarrow$. Assume that there exists a sequence $\{x_n\} \subset D \setminus \{a\}$ that converges to a. Let V_a be a neighborhood of a. Then there exists $N \in \mathbb{N}$ such that $x_n \in \dot{V}_a \cap D$ for every $n \geq N$. Consequently, $\dot{V}_a \cap D \neq \varnothing$ and a is a limit point. □

3 Complete metric spaces

Definition 3.1. *A sequence $\{x_n\}_{n=1}^{\infty}$ is a* ***Cauchy sequence*** *(**fundamental**) if for every $\varepsilon > 0$ there exists a natural number N such that $d(x_n, x_m) < \varepsilon$ for every $n, m > N$.*

Notice that every convergent sequence is a Cauchy sequence while the converse is not true.

Definition 3.2. *A metric space (X,d) is* ***complete*** *if every Cauchy sequence of points of X is convergent.*

Examples 3.3. *1. A discrete space is always complete.*

2. $X=\mathbb{R}$ with metrics $d=|x-y|$ is complete.

3. $X=\mathbb{Q}$ with metrics $d=|x-y|$ is not complete.

4. $X=\mathbb{R}^n$ or $X=\mathbb{C}^n$ with metrics $d_p(x,y)=\left(\sum\limits_{k=1}^{n}|x_k-y_k|^p\right)^{1/p}$, $1\leq p\leq\infty$ is complete.

5. $X=C[a,b],\ d(f,g)=\max\limits_{x\in[a,b]}|f(x)-g(x)|$ is complete.

6. $X=C[a,b],\ d_1(f,g)=\int\limits_a^b|f-g|$ is not complete.

7. For every metric space (X,d) there exists a complete metric space (X^*,d^*) such that (X,d) is a subspace of (X^*,d^*) and X is dense in X^*. The ***completion*** *(X^*,d^*) is unique up to isometry.*

Definition 3.4. *We say that (X,d) is a subspace of (X^*,d^*) if $X\subset X^*$ and d is a restriction of d^* to $X\times X$.*

Theorem 3.5. *The Euclidean space $\mathbb{R}^n$ with metric d_p, $1\leq p\leq\infty$ is complete.*

Proof. We will consider the case when $1\leq p<\infty$. Consider Cauchy sequence $\{x^j\}_{j=1}^{\infty}$ in $\mathbb{R}^n$ with metric $d_p(x,y)=\left(\sum\limits_{k=1}^{n}|x_k-y_k|^p\right)^{1/p}$.

Then $\left|x_k^j - x_k^i\right| \leq d_p(x^j, x^i)$ and sequences of coordinates are Cauchy in $\mathbb{R}$ and, consequently, are convergent.

Let $a_k = \lim\limits_{j\to\infty} x_k^j$ and $a = (a_1, \ldots, a_n)$. Finally,

$$d(x^j, a) = \left(\sum_{k=1}^{n} \left|x_k^j - a_k\right|^p\right)^{1/p} \to 0,\ j \to +\infty.$$

$\square$

Theorem 3.6. *The set of continuous functions with uniform metric*

$$d(f, g) = \max_{x\in[a,b]} |f(x) - g(x)|$$

is the complete metric space.

Proof. Consider Cauchy sequence $\{f_k\} \subset C[a, b]$. Then for every fixed $t \in [a, b]$ a sequence $\{f_k(t)\}$ is Cauchy since $|f_k(t) - f_j(t)| \leq d(f_k, f_j)$. Consequently, for every $t \in [a, b]$ there exist a limit

$$g(t) = \lim_{k\to\infty} f_k(t).$$

We need to prove continuity of g and $d(f_k, g) \to 0,\ k \to +\infty$.

Let $\varepsilon > 0$. Then $\exists N > 0 :\ \forall k, j > N\ d(f_k, f_j) < \varepsilon$. Consequently,

$$|f_k(t) - f_j(t)| \leq d(f_k, f_j)\varepsilon,\ t \in [a, b]$$

and letting $j \to +\infty$ we obtain $|f_k(t) - g(t)| \leq \varepsilon,\ t \in [a, b]$. Considering $sup_{t\in[a,b]}$, we get $d(f_k, g) < \varepsilon$.

Let's prove the continuity of g. Let $\varepsilon > 0$. By previous there exists a function f_k such that $d(g, f_k) < \varepsilon/3$. The uniform continuity of f_k implies the existence of such $\delta > 0$ that

$$|x - y| < \delta \Rightarrow |f_k(x) - f_k(y)| < \varepsilon/3.$$

Then for $|x-y|<\delta$

$$|g(x)-g(y)| \le |g(x)-f_k(x)| + |f_k(x)-f_k(y)| + |f_k(y)-g(y)| < \varepsilon.$$

$\square$

Theorem 3.7. *The space $C[a,b]$ with metric*

$$d_1(f,g) = \int_a^b |f-g|$$

is not complete.

Proof. Consider a sequence of continuous on $[-1,1]$ functions

$$f_n(x) = \begin{cases} -1, & x \in [-1,-1/n], \\ nx, & x \in (-1/n, 1/n], \\ 1, & x \in [1/n, 1] \end{cases}.$$

Then

$$d_1(f_n, f_m) \le \int_{-1/n}^{1/n} n\,|x|\,dx + \int_{-1/m}^{1/m} m\,|x|\,dx = \frac{1}{n} + \frac{1}{m}$$

and the sequence $\{f_n\}$ is Cauchy in metric d_1. We will prove that this sequence has no limit.
Let $g(x) = \operatorname{sign}(x)$. Then $\int_{-1}^{1} |f_n - g| \to 0,\ n \to +\infty$ and for every continuous function f we have $\int_{-1}^{1} |f-g| > 0$. At the same time

$$\int_{-1}^{1} |f-g| \le \int_{-1}^{1} |f-f_n| + \int_{-1}^{1} |f_n - g|,$$

Consequently, $\int_{-1}^{1} |f-f_n|$ can not tend to 0. $\square$

3.1 Principle of Nested Balls

Lemma 3.8. *1. A Cauchy sequence is bounded.*
2. If a Cauchy sequence has a convergent subsequence then it is convergent.

Theorem 3.9. *1. Assume that a sequence $\{x_n\}$ is Cauchy. Then there exists a number $N \in \mathbb{N}$ such that $d(x_n, x_m) < 1$ for $n, m \geq N$. Then $d(x_n, x_N) < 1$ for every $n \geq N$. Let*

$$r = \max(d(x_1, x_N), \ldots, d(x_{N-1}, x_N), 1).$$

Then $d(x_n, x_N) < r + 1$ for ever $n \in \mathbb{N}$.
2. Assume that $\{x_n\}$ is Cauchy sequence that has a partial limit $a = \lim x_{n_k}$. Let $\varepsilon > 0$. Then there exist $K, N \in \mathbb{N}$ such that $d(a, x_{n_k}) < \varepsilon/2$ for $k \geq K$ and such that $d(x_n, x_m) < \varepsilon/2$ for $n, m \geq N$. Let $M = max(n_K, N)$. Then for every $m > M$

$$d(x_m, a) \leq d(x_m, x_{n_M}) + d(x_{n_M}, a) < \varepsilon.$$

and $a = \lim x_n$.

Theorem 3.10. *A metric space is complete if and only if the intersection of arbitrary sequence of closed nested balls is not empty. (A sequence of closed balls $\{\overline{B}_n\}_{n=1}^{\infty}$ is nested if $\overline{B}_{n+1} \subset \overline{B}_n$).*

Proof. **Necessity**. Without loss of generality we assume that radiuses of balls tend to 0. Let $x_n \in \overline{B}(a_n, r_n)$. Then $d(x_n, x_m) \leq max(2r_n, 2r_m)$ and a sequence $\{x_n\}$ is Cauchy and, by completeness, has a limit $a = \lim x_n$. Then $a \in \bigcap \overline{B}(a_n, r_n)$ since $x_m \in \overline{B}(a_n, r_n)$ for $m > n$ and a is a limit point of a ball $\overline{B}(a_n, r_n)$.

Sufficiency. Let $\{x_n\}$ be a Cauchy sequence. We choose a subsequnce $\{x_{n_j}\}_{j=1}^{\infty}$ so that $d(x_n, x_{n_j}) < 2^{-j-1}$ for $n > n_j$. Balls $\overline{B}(x_{n_j}, 2^{-j})$ are nested and have a common point a. Then $x_{n_j} \to a$ and by previous lemma a sequence x_n is convergent. □

Definition 3.11. *A set A is called **dense in a set** B if $\overline{A} = B$. A set A is called **(everywhere) dense** if $\overline{A} = X$. A set A is called **nowhere dense** if the interior of a closure of A is empty, that is A is not dense in any ball. A metric space that contains a countable dense subset is called **separable**.*

Remark. Most spaces defined above are separable. (For example, a set of polynomials with rational coefficients is dense in $C[a, b]$, $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$ with Euclidean metric.)

Theorem 3.12 (Baire's theorem)**.** *A complete metric space can not be expressed as the union of a countable number of nowhere dense sets.*

Proof. Assume the converse. Let (X, d) be a complete metric space such that $X = \bigcup\limits_{k=1}^{\infty} A_k$ while A_k are nowhere dense sets. Since A_1 is nowhere dense we can find a closed ball $\overline{B}_1 \subset X \setminus A_1$. To prove that consider some point $a_1 \in X \setminus \overline{A}_1$ and its neighborhood $V_{a_1} = B(a_1, r_1)$ that is contained in $X \setminus \overline{A}$. Now we can choose $\overline{B_1} = \overline{B}(a_1, r_1/2)$.

Further A_2 is nowhere dense and we similarly can choose a ball $B_2 \subset B_1 \setminus A_2$. (Again consider some point of $B_1 \setminus \overline{A}_2$ and its neighborhood $B(a_2, r_2)$ that is contained in $B_1 \setminus \overline{A}_2$ and $\overline{B}_2 = \overline{B}(a_2, r_2/2)$.) Continuing this argument we obtain a sequence of balls B_n such that $B_n \subset B_{n-1} \setminus A_n$.

This construction implies that $\overline{B_n} \subset X \setminus \left(\bigcup_{k=1}^{n} A_k \right)$. Consequently,

$$\bigcap_{n=1}^{\infty} \overline{B_n} \subset X \setminus \left(\bigcup_{k=1}^{\infty} A_k \right) = \varnothing,$$

which contradicts the nested ball property of a complete space (X, d). $\square$

4 Compactness.

4.1 Definition and properties of compact sets

Definition 4.1. *A* ***cover*** *of of a set K is a family of sets $C = \{G_\alpha\}_{\alpha \in A}$ such that $K \subset \bigcup_{\alpha \in A} G_\alpha$. We say that C is* ***open cover*** *if each of its members is an open set. A* ***subcover*** *of C is a subset of C that still covers K.*

Definition 4.2. *A subset K of a metric space is* ***compact*** *if every open cover of K has* ***a finite subcover*** *of K. A metric space (X, d) is* ***compact*** *if X is compact.*

Examples 4.3. *A closed segment $[a, b]$ is compact by Heine-Borel theorem.*

Theorem 4.4. *1. A compact set is bounded and closed.*

2. A closed subset of a compact set is compact.

Proof. **1.** Let K be compact. Let $a \in X$. Then $K \subset X = \bigcup_{n=1}^{\infty} B(a, n)$. The open cover $\{B(a, n)\}$ contains a finite subcover and, consequently, $K \subset B(a, N)$ for some N.

Let's prove that K is closed. Let p be a limit point of K such that $p \notin K$. Then $K \subset X \setminus \{p\} = \bigcup_{n=1}^{\infty} X \setminus \overline{B}(p, 1/n)$. The open cover $\{X \setminus \overline{B}(p, 1/n)\}$ contains a finite subcover and, consequently, $X \setminus \overline{B}(p, 1/N)$ for some N and $B(p, 1/N) \cap K = \varnothing$ and p is not a limit point of K.
2. Let K be compact and K_1 be a closed subset of K. If $C = \{G_\alpha\}$ is an open cover of K_1, then with an open set $X \setminus K$ we get an open cover of K which has a finite subcover C_1 of K. Consequently, C_1 without $X \setminus K$ covers K_1. $\square$

4.2 Sequential compactness.

Definition 4.5. *A set K is sequentially compact if every sequence of points of K has a convergent subsequence.*

Theorem 4.6 (The Heine-Borel Theorem)**.** *A metric space is compact if and only if it is sequentially compact.*

Proof. **Necessity**. Let (X, d) be compact. Assume that it is not sequentially compact. Then there exists a sequence $M = \{x_n\} \subset X$ that has no convergent subsequence. Consequently M has no limit points, is closed and, as a closed subset of a compact set, is compact.

For every j there exists a neighborhood V_{x_j} such that $M \bigcap V_{x_j} = \{x_j\}$. A family of open balls $\{V_{x_j}\}_{j=1}^{\infty}$ covers M and has no finite subcovers.

Sufficiency. Let X be sequentially compact but not compact. Then there exists an open cover $C = \{G_\alpha\}_{\alpha \in A}$ that has no finite subcovers.

Let $r(x) = \sup\{r > 0 : \exists G_\alpha : B(x, r) \subset G_\alpha\}$.

It is clear that $r(x) > 0$. We will prove that $\inf\limits_{x \in X} r(x) = 2r_0 > 0$.

Assume the converse and let $\inf\limits_{x \in X} r(x) = 0$. Then there exists a sequence $\{x_n\}$ such that $r(x_n) \to 0$. By sequential compactness we can choose a convergent subsequence of $\{x_n\}$. Without loss of generality we assume that $\{x_n\}$ has a limit a. Consequently, there exist a number N such that $d(x_n, a) < r(a)/2$ for $n > N$. Then $x_n \in B(a, r(a)/2) \subset G_\alpha$ and $r(x_n) \geq r(a)/2 > 0$, which contradicts the assumption $r(x_n) \to 0$.

Let $x_1 \in X$ and $B(x_1, r_0) \subset G_{\alpha_1}$. There exists $x_2 \in X \setminus G_{\alpha_1}$ and α_2 such that $B(x_2, r_0) \subset G_{\alpha_2}$. Continuing this way we construct a sequence $\{x_k\}$ such that $d(x_k, x_j) \geq r_0$. Consequently every subsequence of $\{x_k\}$ is not Cauchy and, since, divergent. This contradicts the sequential compactness.

$\square$

4.3 Compactness in $\mathbb{R}^n$

Remark 4.7. *Notice that the convergence in $\mathbb{R}^n$ is equivalent to the convergence of coordinates*

$$(x_1^{(k)}, \ldots, x_n^{(k)}) \to (a_1, \ldots, a_n) \iff x_j^{(k)} \to a_j,\ 1 \leq j \leq n.$$

Lemma 4.8. *A closed cube $[a, b]^n$ is (sequentially) compact in $\mathbb{R}^n$.*

Proof. Induction by dimension n. Base case $n = 1$. A closed segment $[a, b]$ is compact by Heine-Borel lemma.

Induction step $(n \to n + 1)$. Let $\{x^k\} \subset [a, b]^{n+1}$. Then $x^k = (y^k, x_{n+1}^k)$, where $y^k = (x_1^k, \ldots, x_n^k) \in [a, b]^n$ and $x_{n+1}^k \in [a, b]$. By induction assumption there exists a convergent subsequence y^{k_j}. We also can choose a convergent subsequence $x_{n+1}^{k_{j_i}}$ from a sequence $x_{n+1}^{k_j}$. Then a sequence $x^{k_{j_i}}$ is convergent.

$\square$

Theorem 4.9 (A characterization of compacts in $\mathbb{R}^n$.)**.** *Let $K \subset \mathbb{R}^n$. The following assertions are equivalent:*

1. K is bounded and closed.

2. K is compact

3. K sequentially compact.

Proof. Assertions (2) and (3) are equivalent by Heine-Borel theorem. Assertion (2) implies (3), since every compact is bounded and closed.

Let $K' \subset \mathbb{R}^n$ be bounded and closed. Then K is a subset of some closed cube and is compact as a closed subset of a compact. $\square$

5 A limit of a function.

Definition 5.1. *Let (X, d_X) and (Y, d_Y) be metric spaces, $D \subset X$, $f : D \to Y$, $a \in X$ be a limit point of D. We say that $A \in Y$ is a limit of function f at point a if one of the following equivalent assertions is satisfied.*
1. Cauchy definition.

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D \setminus \{a\} : d_X(x, a) < \delta \Rightarrow d_Y(f(x), A) < \varepsilon.$$

2. Definition by neighborhoods.

$$\forall V_A\ \exists V_a : f(\dot{V}_a \cap D) \subset V_A.$$

3. The definition by Heine (via sequences)

$$\forall \{x_n\} :\ x_n \in D \setminus \{a\},\ x_n \to a \Rightarrow f(x_n) \to A.$$

Notations: $\lim\limits_{x \to a} f(x) = A,\ f(x) \underset{x \to a}{\longrightarrow} A.$

5.1 Properties of a limit.

- Definitions (1-3) of a limit are equivalent.

  Proof. $1 \Rightarrow 2$. Let $V_A = B(A, \varepsilon)$ be a neighborhood of A. applying Cauchy's definition we see that

  $$\exists \delta > 0 \, \forall x \in D \setminus \{a\} : d_X(x, a) < \delta \Rightarrow d_Y(f(x), A) < \varepsilon$$

  and let $V_a = B(a, \delta)$. Consequently, for every $x \in \dot{V}_a \cap D$ $f(x) \in V_A$, that means that $f(\dot{V}_a \cap D) \subset V_A$.
  $2 \Rightarrow 3$. Consider a sequence $\{x_n\} \subset D \setminus \{a\}$ such that $x_n \to a$. Let V_A be a neighborhood of A then

  $$\exists V_a : f(\dot{V}_a \cap D) \subset V_A.$$

  Since $x_n \to a$ there exists $N \in \mathbb{N}$ such that $x_n \in V_a$ for every $n \geq N$ and, consequently, $f(x_n) \in V_A$.
  $3 \Rightarrow 1$. Assume the converse:

  $$\exists \varepsilon_0 > 0 : \; \forall \delta > 0 \, \exists x \in D \setminus \{a\} : d_X(x, a) < \delta \text{ and } d_Y(f(x), A) > \varepsilon_0.$$

  Then for every $\delta_n = \frac{1}{n}$ there exists $x_n \in D \setminus \{a\}$ such that $d_X(x_n, a) < \delta$ and $d_Y(f(x), A) > \varepsilon_0$. So $x_n \to$ and $f(x_n) \not\to A$, which contradicts Heine definition. □

- The limit of a function is unique.

  Proof. Let A and B be the limits of function f at a. Consider a sequence $\{x_n\} \in D \setminus \{a\}$ that converges to a. then by Heine definition

  $$f(x_n) \to A, \quad f(x_n) \to B.$$

Then, by uniqueness of a limit of a sequence $A = B$. $\square$

- A function that has a limit at a is bounded in some neighborhood of a, that is there exists a neighborhood V_a of a such that $f(V_a)$ is bounded.

  Proof. The proof directly follows from the definition by neighborhoods. $\square$

Theorem 5.2 (Cauchy's criterion for functions)**.** *Let (X, d_X) and (Y, d_Y) be two metric spaces and $f : D \subset X \to Y$, a be a limit point of D. Assume that Y is complete. Then function f has a limit at a if and only if*

$$\forall \varepsilon > 0\ \exists V_a\ \forall u, v \in \dot{V}_a \cap D\ d_Y(f(u), f(v)) < \varepsilon.$$

Completeness of Y is important!!!.

5.2 Limit of a composition

Theorem 5.3 (On a limit of a composition)**.** *Let X, Y, Z be metric spaces, $f : D \subset X \to Y$, $g : E \subset Y \to Z$, $f(D) \subset E$, a be a limit point of D. Assume that*

1. *there exists a limit $A = \lim\limits_{x \to a} f(x)$;*

2. *$f(x) \neq A$ in some neighborhood of a;*

3. *there exists a limit $\lim\limits_{y \to A} g(y) = B$.*

Then $\lim\limits_{x \to a} g \circ f = B$.

Proof. We will prove this theorem applying the definition by Heine. Let $x_n \to a,\ x_n \neq a$. Then $y_n = f(x_n) \to A$ and starting from some number $f(x_n) \neq A$. Consequently, $g(f(x_n)) = g(y_n) \to B$.

$\square$

5.3 Limit and iterated limit of a function.

Let $D_1, D_2 \subset \mathbb{R}$, a and b to be limit points of D_1 and D_2 respectively, $(D_1 \setminus \{a\}) \times (D_2 \setminus \{b\}) \subset D$, $f : D \to \mathbb{R}$ or $\mathbb{C}$.

1. Assume that for every $x \in D_1 \setminus \{a\}$ the limit $\varphi(x) = \lim\limits_{y \to b} f(x, y)$ exists. Then the limit of function φ at a is called an **iterated limit** of function f at (a, b) and

$$\lim_{x \to a} \varphi(x) = \lim_{x \to a} \lim_{y \to b} f(x, y).$$

2. Assume that for every $y \in D_1 \setminus \{b\}$ the limit $\psi(x) = \lim\limits_{x \to a} f(x, y)$ exists. Then the limit of function ψ at a is called an **iterated limit** of function f at (a, b) and

$$\lim_{y \to b} \psi(y) = \lim_{y \to b} \lim_{x \to a} f(x, y).$$

3. Recall that A is a limit of a function f at (a, b), if

$$\forall V_A\ \exists V_a, V_b :\ (x, y) \in (V_a \times V_b) \setminus \{(a, b)\} \Rightarrow f(x, y) \in V_A,$$

$$A = \lim_{\substack{x \to a \\ y \to b}} f(x, y) = \lim_{(x,y) \to (a,b)} f(x, y).$$

Theorem 5.4. *Let $D_1, D_2 \subset \mathbb{R}$, a and b are limit points D_1 and D_2 respectively, $(D_1 \setminus \{a\}) \times (D_2 \setminus \{b\}) \subset D$, $f : D \to \mathbb{R}$ or $\mathbb{C}$. Assume that*

- *There exists finite or infinite limit* $A = \lim\limits_{\substack{x\to a\\ y\to b}} f(x,y)$.
- $\forall x \in D_1 \setminus \{a\}\ \exists \varphi(x) = \lim\limits_{y\to b} f(x,y) < \infty$.

Then the iterated limit $\lim\limits_{x\to a} \varphi(x)$ *exists and is equal to* A.

Proof. We will prove the theorem in cases when A is finite. Let $\varepsilon > 0$ then there exist neighborhoods V_a and V_b such that

$$|f(x,y) - A| < \varepsilon$$

for every $x \in \dot{V}_a \cap D_1$ and $y \in \dot{V}_b \cap D_2$. Considering limits at b we see that

$$|\varphi(x) - A| = \lim_{y\to b} |f(x,y) - A| \leq \varepsilon$$

for every $x \in \dot{V}_a \cap D_1$. Consequently, $\lim\limits_{x\to a} \varphi(x) = A$. $\square$

Exercise. Consider f, g, h and find their limits and iterated limits $(0,0)$.

$$f(x,y) = \frac{x^2 - y^2}{x^2 + y^2}; \quad g(x,y) = \frac{xy}{x^2 + y^2}; \quad h(x,y) = x \sin\frac{1}{y} + y \sin\frac{1}{x}.$$

Remark. If $f = (f_1, \ldots, f_n) : X \to \mathbb{R}^n$ then $f \to A = (A_1, \ldots, A_n) \iff f_j \to A_j$.

Remark. If (X, d_X), (Y, d_Y) are metric spaces the function

$$d_{X\times Y}((x_1, y_1), (x_2, y_2)) = \sqrt{d_X(x_1, x_2)^2 + d_Y(y_1, y_2)^2}$$

defines metric on a set $X \times Y$.

6 Continuity.

6.1 Definition of continuity.

Definition 6.1. *Let (X, d) be a metric space, $D \subset X$. A point $a \in D$ is the* ***isolated point of*** *D if there exists a neighborhood V_a of a such that $V_a \cap D = \{a\}$.*

Notice that any point of a set is either the limit point or isolated point.

Definition 6.2. *Assume that (X, d_X) and (Y, d_Y) are metric spaces, $D \subset X$, $f : D \to Y$. A function f id called* ***continuous*** *at $a \in D$ if one of the following four equivalent assertions is satisfied.*

1. *Either a is an isolated point or $\lim\limits_{x \to a} f(x) = f(a)$.*

2. *Cauchy definition.*

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in D(d_X(x, a) < \delta \Rightarrow d(f(x), f(a)) < \varepsilon).$$

3. *The definition by neighborhoods.*

$$\forall V_{f(a)} \ \exists V_a : f(V_a \cap D) \subset V_{f(a)}.$$

4. *Heine definition.*

$$\forall \{x_n\} : \ x_n \in D, \ x_n \to a \Rightarrow f(x_n) \to f(a).$$

If the function f is not continuous at a then we say that f is **discontinuous** at a.

Remark 6.3. *The equivalence of definitions in case when a is a limit point follows from equivalence of different definitions of the limit. The case of isolated point is rather trivial. If a is isolated point then for some $\delta > 0$ the only point that satisfies inequality $d_X(x, a) < \delta$ is the point $x = a$ and in this case $d(f(x), f(a)) = 0 < \varepsilon$ for any $\varepsilon > 0$. Also for $V_a = V_a(\delta)$ the intersection $V_a \cap D = \{a\}$ and $f(V_a \cap D) = \{f(a)\} \subset V_{f(a)}$. If $x_n \to a$ and $x_n \in D$ then starting from some number $x_n = a$ so $f(x_n) = f(a) \to f(a)$.*

6.2 Properties of continuous functions.

If a is an isolated point of D then every function defined on D is continuous at a. In particular, all functions on discrete spaces are continuous.

Theorem 6.4 (On continuity of composition)**.** *Assume that X, Y, Z are metric spaces, $f : D \subset X \to Y$, $g : E \subset Y \to Z$, $f(D) \subset E$, f is continuous at $a \in D$, g is continuous at $f(a)$. Then the composition $g \circ f$ is continuous at a.*

Definition 6.5.

- *We say that f is continuous on D if it is continuous at every point of D.*

- *Assume that $f : X \to Y$ is bijection of two metric spaces. Then there exists an inverse function $f^{-1} : Y \to X$. If functions f and f^{-1} are continuous then f is called homeomorphism, and spaces X, Y are homeomorphic.*

- *A bijection $f : X \to Y$ is called isometry, and spaces $(X d_X)$ and (Y, d_Y) are called isometric if*

$$d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2).$$

Theorem 6.6 (Characterization of continuity by pre-images)**.** *Assume that X, Y are two metric spaces. A function $f : X \to Y$ is continuous on X if and only if pre-image of every open set in Y is open in X.*

Proof. **Necessity.** Let $f : X \to Y$ be continuous and $G \subset Y$ be open. If $a \in f^{-1}(G)$ then $f(a) \in G$ and there exists a neighborhood $V_{f(a)} \subset G$. There exists neighborhood V_a of a point a such that $f(V_a) \subset V_{f(a)} \subset G$. Then $a \in V_a \subset f^{-1}(G)$.

Sufficiency. Let $f : X \to Y$ be such $f^{-1}(G)$ is open in X for every $G \subset Y$. Then the preimage of every neighborhood $V_{f(a)}$ is open and contains some neighborhood of a. Consequently there exists a neighborhood V_a such that $V_a \subset f^{-1}(V_{f(a)})$. Hence $f(V_a) \subset V_{f(a)}$. □

Theorem 6.7. *A continuous image of a compact is compact.*

Proof 1. Let X be compact and $f : X \to Y$ be continuous. Consider an open cover $\{G_\alpha\}_{\alpha \in A}$ of $f(X)$. Then the compact space X is covered by open sets $f^{-1}(G_\alpha)$ and there exists a finite subcover $X \subset \bigcup_{j=1}^{n} f^{-1}(G_{\alpha_j})$. Then

$$f(X) \subset f\left(\bigcup_{j=1}^{n} f^{-1}(G_{\alpha_j})\right) \subset \bigcup_{j=1}^{n} G_{\alpha_j}.$$

Proof 2. We know that compactness in metric spaces is equivalent to sequential compactness. Let $y_n \in f(X)$. Then $x_n = f^{-1}(y_n) \in X$ and we can choose a convergent subsequence $\{x_{n_j}\}$. Then by continuity of f the sequence $y_{n_j} = f(x_{n_j})$ is convergent.

Corollary 6.7.1. *1. A continuous image of compact set is bounded and closed.*

2. Weierstrass' theorems on continuous functions. The function is continuous on a compact is bounded and obtains its maximal and minimal value.

Definition 6.8. *Assume that (X, d_X) and (Y, d_Y) are metric spaces. The function $f : X \to Y$ is uniformly continuous if*

$$\forall \varepsilon > 0 \ \exists \delta > 0 : d_X(x_1, x_2) < \delta \Rightarrow \ d_Y(f(x_1), f(x_2)) < \varepsilon.$$

Theorem 6.9 (G. Cantor)**.** *A function continuous on a compact is uniformly continuous.*

Proof. We will prove this theorem by contradiction. Assume that there exist two sequences x_n, y_n and a number $\varepsilon_0 > 0$ such that $d_X(x_n, y_n) < \frac{1}{n}$, but $d_Y(f(x_n), f(y_n)) > \varepsilon_0$. Then $\exists x_{n_k} \to a$. Consequently, $y_{n_k} \to a$, $f(x_{n_k}), f(y_{n_k}) \to f(a)$ and $d_Y(f(x_{n_k}), f(y_{n_k})) \to 0$. This leads to a contradiction. □

7 Normed vector spaces.

Definition 7.1. *Assume that K is a field and X be some set and there are defined two functions addition $+ : X \times X \to X$ and (scalar) multiplication $\cdot : K \times X \to X$ that satisfy following properties.*

1. $(x + y) + z = x + (y + z), \quad x, y, z \in X;$
2. $x + y = y + x, \quad x, y \in X;$
3. $\exists \theta \in X : \ \forall x \in X \ 0 \cdot x = \theta;$
4. $(\lambda + \mu)x = \lambda x + \mu x, \quad x \in K, \ \lambda, \mu \in K;$

5. $(\lambda\mu)x = \lambda(\mu x), \quad x \in K,\ \lambda, \mu \in K;$

6. $1 \cdot x, \quad x \in X.$

Then X is called a vector space over the field K.
Elements of K are called scalars and elements of X are called vectors.
An element θ is called a zero (zero vector) of a space.
$-y := (-1)y$

Examples 7.2. *1. $X = \mathbb{R}^n$ or $X = \mathbb{C}^n$ with the field of scalars $K = \mathbb{R}$ or $K = \mathbb{C}$. The addition and multiplication are defined by coordinates.*

$$x = (x_1, \dots, x_n), \quad y = (y_1, \dots, y_n);$$
$$(x + y) = (x_1 + y_1, \dots, x_n + y_n), \quad \lambda x = (\lambda_1 x_1, \dots, \lambda_n x_n).$$

Zero vector is defined by $\mathbb{O} = (0, \dots, 0)$.

2. *Similarly we can define a vector space on set of real or complex sequences or on spaces of sequences that are preserved under addition and multiplication by scalar.*

3. *If X is some set of real or complex valued functions $f : D \to \mathbb{R}$ or $\mathbb{C}$ (for example, a set of continuous functions). The addition and multiplication by scalar (real or complex) are defined by*

$$(f + g)(x) = f(x) + g(x), \quad (\lambda f)(x) = \lambda f(x).$$

Definition 7.3. *Let X be a vector space with $\mathbb{R}$ or $\mathbb{C}$ as a field of scalars. A function $p : X \to [0, +\infty)$ is called **a norm** if the following properties are satisfied (axioms of norm).*

1. *Positive definiteness.* $p(x) = 0$ *if and only if* $x = \mathbb{O}$.

2. *Positive homogeneity.* $p(\lambda x) = |\lambda|\, p(x)$.

3. *Triangle inequality (semi-additivity).* $p(x + y) \leq p(x) + p(y)$.

Notations. $p(x) = \|x\|$. *An ordered pair* $(X, \|\|)$ *is called a normed (vector) spaces.*
A function $p : X \to [0, +\infty)$ *that satisfies* $(2 - 3)$ *is called seminorm.*

Lemma 7.4 (Properties of seminorm.)**.**

$$1.p\left(\sum_{k=1}^{n} \lambda_k x_k\right) \leq \sum_{k=1}^{n} |\lambda_k|\, p(x_k); \quad 2.p(\theta) = 0;$$
$$3.p(-x) = p(x); \quad 2. |p(x) - p(y)| \leq p(x - y).$$

1. $X = \mathbb{R}^n$ or $X = \mathbb{C}^n$, $\quad \|x\| = \sqrt{\sum\limits_{k=1}^{n} |x_k|^2}$

2. $X = \mathbb{R}^n$ or $X = \mathbb{C}^n$, $1 \leq p < \infty$. $\|x\|_p = \left(\sum\limits_{k=1}^{n} |x_k|^p\right)^{1/p}$.

3. $X = \mathbb{R}^n$ or $X = \mathbb{C}^n$. $\|x\|_\infty = \max\limits_{1 \leq k \leq n} |x_k|$.

4. $X = \ell^\infty$, $\quad \|x\|_\infty = \sup\limits_{k \geq 1} |x_k|$.

5. $X = C[a, b]$ $\|f\| = \max\{|f(x)|, x \in [a, b]\}$.

6. $X = C[a, b]$, $1 \leq p < \infty$, $\|f\|_p = \left(\int\limits_a^b |f(x)|^p\right)^{1/p}$.

A function $d(x, y) = \|x - y\|$ is called the metric generated by the norm. Convergence with respect to the norm is convergence with respect to this metric.

If metric is associated by the norm then $\|x\| = d(0, x)$. But not for every metric space a function $d(0, x)$ is a norm.

Examples 7.5. *Consider the metric $d(x, y) = \frac{|x-y|}{1+|x-y|}$ on $\mathbb{R}$. Then $d(0, x)$ is not a norm.*

Definition 7.6. *A normed space that is complete in the metric generated by the norm is called a **Banach space**.*

Theorem 7.7 (Arithmetic properties of convergent sequences in a normed space.)**.** *Let $(X, \|\|)$ be a real or complex normed space. Let $\{x_n\}, \{y_n\}$ be two sequences X, $\{\lambda_n\}$ be a number sequence such that $x_n \to x_0,\ y_n \to y_0,\ \lambda_n \to \lambda_0$. Then*

$$1. x_n \pm y_n \to x_0 \pm y_0 \quad 2. \lambda_n x_n \to \lambda_0 x_0 \quad 3. \|x_n\| \to \|y_n\|$$

The proof is analogous to the case of number sequences.

Corollary 7.7.1. *Theorem 7.7 in fact proves that addition $+ : X \times X \to X$, multiplication by scalar $\cdot K \times X \to X$ and a norm $\|\| : X \to [0, +\infty)$ are continuous (see the remark on metric in $X \times Y$.*

Corollary 7.7.2. *Assume that (X, d) is a metric space, $(Y, \|\|)$ is a normed space, $f, G : D \subset X \to Y$, $\lambda : D \to \mathbb{R}$ or $\mathbb{C}$, a is a limit point of D. Let $A, B \in X,\ \lambda_0 \in \mathbb{R}$ $(\lambda_0 \in \mathbb{C})$ and $f \underset{x\to a}{\longrightarrow} A$, $g \underset{x\to a}{\longrightarrow} B$, $\lambda \underset{x\to a}{\longrightarrow} \lambda_0$. Then*

$$1. f \pm g \underset{x\to a}{\longrightarrow} A \pm B \quad 2. \lambda f \to \lambda_0 A \quad 3. \|f\| \to \|A\| .$$

Corollary 7.7.3. *Assume that (X, d) is a metric space, $(Y, \|\|)$ is a normed space, $f, G : D \subset X \to Y$, $\lambda : D \to \mathbb{R}$ or $\mathbb{C}$. If functions f, g, λ are continuous at a. Then functions $f \pm g$, λf, $\|f\|$ are also continuous at a.*

8 Contraction mapping theorem.

Definition 8.1. *Let (X, d) be a metric space and $E \subset X$. A **contraction mapping**, or **contraction** or **contractor** is a function $f : E \to X$ with the property that there exists a number $0 < q < 1$ such that*

$$d(f(x), f(y)) \leq qd(x, y), \ x, y \in E. \tag{1}$$

*The smallest such value of q is called the **Lipschitz constant** or the **coefficient of contraction** of f.*

A contraction is uniformly continuous on E.

Definition 8.2. *A point x is a fixed point of function f if $x = f(x)$.*

Theorem 8.3 (Banach fixed-point theorem.)**.** *Assume that a metric space (X, d) is complete and a set $E \subset X$ is closed. Then the contraction $f : E \to E$ has unique fixed point.*

Proof. **Uniqueness.** Assume that $x_1, x_2 \in E$ are fixed points of contraction f. Then

$$d(x_1, x_2) = d(f(x_1), f(x_2)) \leq qd(x_1, x_2).$$

Consequently, $d(x_1, x_2) = 0$ and $x_1 = x_2$.

Existence. Let $x_0 \in E$ and $x_n = f(x_{n-1}),\ n \in \mathbb{N}$. We will prove that a sequence $\{x_n\}_{n=1}^{\infty}$ is Cauchy. Let $n < m,\ k = m - n$. Then

$$d(x_m, x_n) \leq qd(f(x_{m-1}), f(x_{n-1})) = \ldots = q^n d(x_{m-n}, x_0) \leq \frac{d(x_1, x_0)}{1-q} q^n$$

since

$$d(x_k, x_0) \leq d(x_k, x_{k-1}) + d(x_{k-1}, x_{k-2}) + \ldots + d(x_1, x_0) \leq$$
$$(q^{k-1} + q^{k-2} + \ldots + 1) d(x_1, x_0) \leq \frac{d(x_1, x_0)}{1-q}.$$

A Cauchy sequence x_n has some limit $a \in E$ since X is complete and E is closed. Then

$$a = \lim x_{n+1} = \lim f(x_n) = f(a).$$

□

Examples 8.4. ***1.*** *Assume that $f : [a, b] \to [a, b]$ is a Lipschitz function*

$$|f(x_1) - f(x_2)| \leq K\, |x_1 - x_2|$$

with constant $K < 1$. Then f is contraction and a sequence $x_0 \in [a, b],\ x_1 = f(x_0),\ x_2 = f(x_1), \ldots$ converges to the unique solution of equation $x = f(x)$.
2. *The solution of the equation $\varphi(x) = 0$ where $\varphi(a) < 0 < \varphi(b)$ and $0 < K_1 \leq \varphi'(x) \leq K_2$ on $[a, b]$ can be considered as the solution of equation*

$$x = f(x) = x - \lambda\varphi(x).$$

Choosing λ is sufficiently small we get into the context of the previous example since $-1 < 1 - \lambda K_2 \leq f'(x) \leq 1 - \lambda K_1 < 1$.

3. *Two previous examples can be easily generalized to Banach (complete normed) space. So we can approximate solution of equation $F(x) = x$, when $F : X \to X$ is Lipschitz with constant less than* 1*, or the solution of equation $\Phi(x) = \mathbb{O}$ by consideration of additional equation $x = F(x) = x - \lambda\Phi(x)$.*

8.1 Approximations of solutions of linear systems.

Let
$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}; \quad b = \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix};$$
Consider a system of linear equations

$$Ax = b \iff Ax - b = 0 \iff x = (A + I)x - b = \tilde{A}x - b = \Phi(x),$$

The condition that implies the contraction property of Φ depends on the choice of the metric in $\mathbb{R}^n$.

Assume that we consider $\mathbb{R}^n$ with Euclidean. Then by Cauchy-Bunyakowski-Schwarz inequality

$$d(\Phi(x'), \Phi(x'')) = \sum_i \Big(\sum_j \tilde{a}_{ij}(x'_j - x''_j) \Big)^2 \le \Big(\sum_{i,j=1}^{n} \tilde{a}_{ij}^2 \Big) d^2(x', x'').$$

And it is sufficient that
$$\sum_{i,j=1}^{n} \tilde{a}_{ij}^2 < 1. \qquad (2)$$

for Φ to be contraction.

Exercise.

1. Find sufficient (or necessary and sufficient, if possible) conditions on Φ to be contraction in $\mathbb{R}^n$ considering metrics

$$d_\infty(x', x'') = \max_{1 \le k \le n} |x'_k - x''_k| \, ;$$

$$d_1(x', x'') = \sum_{k=1}^{n} |x'_k - x''_k| \, .$$

2. Prove that condition (2) is not necessary for Φ to be contraction in $\mathbb{R}^n$ with Euclidean. What is necessary and sufficient condition.

8.2 Existence of solution to Cauchy problem.

Definition 8.5. *Let $G \subset \mathbb{R}^2$ be open. $f : G \to \mathbb{R}$. The equation $x'(t) = f(t, x(t))$ is a Ordinary Differential equation (ODE). (We are looking for a function $x \in C^1(\Delta)$ such that $\Delta \subset \mathbb{R}$ is a segment and $(t, x(t)) \in G$ for all $t \in \Delta$.)*

Examples.

1. $G = \mathbb{R}^2,\ f(t, s) = t,\ x'(t) = t,\ x : \mathbb{R} \to \mathbb{R}, x(t) = \frac{t^2}{2} + C;$

2. $G = \mathbb{R}^2,\ f(t, s) = s,\ x'(t) = x(t),\ x : \mathbb{R} \to \mathbb{R}, x(t) = ce^t.$

Definition 8.6. *Let $G \subset \mathbb{R}^2$ be open, $f \in C(G), (t_0, x_0) \in G$. A function $x \in C^1(\Delta)$ is a solution of Cauchy problem with initial value condition (t_0, x_0) if $t_0 \in \Delta^0$ and*

$$\begin{cases} (t, x(t)) \in G; \\ x'(t) = f(t, x(t)),\ t \in \Delta; \\ x(t_0) = x_0. \end{cases} \tag{3}$$

Theorem 8.7 (Picard's theorem)**.** *Let $G \subset \mathbb{R}^2$ be open, $f : G \to \mathbb{R}$ is Lipschitz with respect to the second variable, that is there exists $K > 0$*

$$|f(t,u) - f(t,\tilde{u})| \leq K\,|u - \tilde{u}|\,,\ (t,u),(t,\tilde{u}) \in G.$$

Then for every initial value $(t_0, x_0) \in G$ there exists $\delta > 0$ and unique function $x \in C^1(\Delta)$, where $\Delta = (t_0 - \delta, t_0 + \delta)$ that solves Cauchy problem (3).

We can generalize this theorem to the system of ODE's. Let $G \subset \mathbb{R}^{m+1}$ be open, $f = (f_1, \ldots, f_m) \in C(G \to \mathbb{R}^m)$ be Lipschitz in the last m variables, that is there exists $K > 0$ such that

$$\|f(t,u_1,\ldots,u_m) - f(t,\tilde{u}_1,\ldots,\tilde{u}_m)\| \leq K\,\|u - \tilde{u}\|\,,\ (t,u),(t,\tilde{u}) \in G.$$

Then the Cauchy problem $\begin{cases} x_1' = f_1(t, x_1(t), \ldots, x_m(t)), \\ \ldots \\ x_m' = f_m(t, x_1(t), \ldots, x_m(t)) \end{cases}$ with initial value $\begin{cases} x_1(t_0) = x_1^0, \\ \ldots \\ x_m(t_0) = x_m^0 \end{cases}$ has a unique solution.

Proof of Picard's theorem. A function f is continuous. Consequently, $\exists r, c > 0$ s.t. $|f| \leq c$ in $B((t_0, x_0), r)$. Let $\delta = \min\left\{\frac{1}{2L}, \frac{r}{\sqrt{1+c^2}}\right\}$ and $\Delta = (t_0 - \delta, t_0 + \delta)$.

1. We will formulate the Cauchy problem in equivalent way:

$$\begin{cases} x \in C^1(\Delta); \\ (t, x(t)) \in G; x'(t) = f(t, x(t)); \\ x(t_0) = x_0; \end{cases} \tag{4}$$

$$\Longleftrightarrow$$

$$\begin{cases} x \in C(\Delta); \\ (t, x(t)) \in G; x(t) = \Phi(x)(t) = x_0 + \int\limits_{t_0}^{t} f(s, x(s))ds. \end{cases} \tag{5}$$

2. A space $X = C(\Delta)$ with uniform metric $d(x, \tilde{x}) = \max\limits_{s\in\Delta} |x(s) - \tilde{x}(s)|$ is complete. Let

$$E = \{x \in C(\Delta) : |x(t) - x_0| \leq c\delta,\ t \in \Delta\}.$$

A solution of (5) is a unique fixed point of the operator Φ, which exists if two following properties are satisfied

1. $\Phi(E) \subset E$;

2. $\Phi : E \to E$ is contractor.

Let $x \in E$. Then

$$|s - t_0|^2 + |x(s) - x_0|^2 \leq \delta^2 + c^2\delta^2 \leq r^2 \frac{1 + c^2}{1 + c^2} = r^2$$

and $(s, x(s)) \in B((t_0, x_0), r)$.

1. Let $x \in E$. We will prove that $y = \Phi(x) \in E$. The continuity of y follows from the theorem on the continuity of integral with variable upper

limit. Moreover,

$$|y(t) - x_0| \le \int_{t_0}^{t} |f(s, x(s))| \, ds \le |t - t_0| \max_s |f(s, x(s))| \le c\delta,$$

2. To prove the contraction condition consider

$$|\Phi(x)(t) - \Phi(\tilde{x})(t)| \le \int_{t_0}^{t} |f(s, x(s)) - f(s, \tilde{x}(s))| \, ds \le$$
$$|t - t_0| \max_s |f(s, x(s)) - f(s, \tilde{x}(s))| \le$$
$$L\delta \max_s |x(s) - \tilde{x}(s)| \le \frac{1}{2} d(x, \tilde{x})$$

□

9 Linear operators in Euclidean spaces.

Let X be a vector space with the scalar field K and zero vector θ_X.

Definition 9.1. *Let $x_1, \dots, x_n \in X,\ \lambda_1, \dots, \lambda_n \in K$. A vector $\sum_{k=1}^{n} \lambda_k x_k$ is called a **linear combination** of vectors $x_1, \dots, x_n \in X$. We say that a linear combination is nontrivial if at least one of the scalars $\lambda_1, \dots, \lambda_n \in K$ is not zero.*

*A set of linear combinations of vectors from a set E is called a **linear span** of a set E and is denoted by* span$E$.

Definition 9.2. *Vectors $x_1, \dots, x_n \in X$ are called **linearly dependent** if there exists a nontrivial linear combination of the vectors that is equal to zero. Otherwise the are called linearly independent.*

A set of vectors $E \subset X$ is called linearly independent if every nontrivial linear combination of vectors of set E is not zero.

Definition 9.3. *Maximal number of linearly independent vectors is called* ***dimension*** *of a vector space X.*

Definition 9.4. *Linearly independent family of vectors $E \subset X$ such that $X = \text{span}E$ is a* ***basis*** *of X.*

Every vector $x \in X$ is a unique linear combination of vectors of basis. In $n-$dimensional space every set of n linearly independent vectors is a basis.

1. Dimension of $\mathbb{R}^n$ is equal to n. The standard basis is defined by standard unit vectors $\{e^k\}_{k=1}^n$, where

$$e_j^k = \begin{cases} 0, & k \neq j, \\ 1, & k = j \end{cases}.$$

   Then $x = (x_1, \ldots, x_n) = \sum_{k=1}^{n} x_k e^k$ where $x_1, \ldots, x_n$ are scalar components of vector x.

2. The dimension of $\mathbb{C}^n$ over $\mathbb{C}$ is n.

3. The dimension of $\mathbb{C}^n$ over $\mathbb{R}$ is $2n$.

4. The dimension of $C[a, b]$ is infinite since $\{1, x, x^2, \ldots\}$ is linearly independent.

9.1 Linear operators.

"Operator" is a synonym to function, the operator with values in $\mathbb{R}$ or $\mathbb{C}$. We us notation $Ax = A(x)$.

Definition 9.5. *Let X, Y be two vector spaces with the same scalar field K. An operator $K : X \to Y$ is linear if*

$$A(\lambda x + \mu y) = \lambda Ax + \mu Ay, \quad x, y \in X, \ \lambda, \mu \in K.$$

Lemma 9.6 (Properties of linear operators.)**.** *Let $A : X \to Y$ be linear operator.*

1. $A\left(\sum\limits_{k=1}^{N} \lambda_k x_k\right) = \sum\limits_{k=1}^{N} \lambda_k A x_k.$

2. $A\theta_X = \theta_Y.$

3. *A set of linear operators from X to Y is a vector space.*

Proof. **1.** This assertion can be checked by induction. For $N = 1$ and $N = 2$ it follows from the definition. Assume that it is true for some $N \in \mathbb{N}$. Then by the linearity of A and by inductional assumption

$$A\left(\sum_{k=1}^{N+1} \lambda_k x_k\right) = A\left(\sum_{k=1}^{N} \lambda_k x_k\right) + \lambda_{N+1} A x_{N+1} = \sum_{k=1}^{N+1} \lambda_k A x_k.$$

2. $A(\theta_X) = A(\theta_X - \theta_X) = A(\theta_X) - A(\theta_X) = \theta_Y$
3. Define the addition and multiplication by the scalar as following

$$(A + B)(x) = A(x) + B(x), \quad (\lambda A)(x) = \lambda A(x).$$

These operations define the structure of a vector space with scalar field K on a set of linear operations since Y is a vector space. (Exercise: check the axioms of the vector space). $\square$

Examples 9.7. *1. Every matrix of size $m \times n$ is a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$.*

2. Every linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$ is defined by matrix. Sometimes we'll denote the matrix of operator $A : \mathbb{R}^n \to \mathbb{R}^m$ by (A).

Denote by A_i the coordinate functions of the operator $A : X \to \mathbb{R}^m$. It is clear that coordinate functions of linear operator are linear operators as well.

Every matrix (A) if the size $m \times n$ defines the linear operator A from $\mathbb{R}^n$ to $\mathbb{R}^m$ that acts as multiplication of vector-column $\mathbb{R}^n$ by the matrix A (the rule of multiplication of row by the column):

$$Ax = (A) \cdot x.$$

Let a_{ik} be elements of the matrix $(A)(i \in [1 : m], k \in [1 : n])$. Then

$$A_i x = \sum_{k=1}^{n} a_{ik} x_k.$$

We will prove that the inverse is also true, every linear operator A from $\mathbb{R}^n$ to $\mathbb{R}^m$ is defined by the unique matrix of size $m \times n$. Indeed, let $a_{ik} = A_i e^k$. Then

$$A_i x = A_i \left(\sum_{k=1}^{n} x_k e^k \right) = \sum_{k=1}^{n} x_k A_i e^k = \sum_{k=1}^{n} a_{ik} x_k.$$

3. A function $A : C[a, b] \to \mathbb{R}$ defined by

$$A(f) = f(x_0)$$

.

4. *A function* $A : C[a,b] \to \mathbb{R}$ *defined by*

$$A(f) = \int_a^b f.$$

5. *An operator* $A : C[a,b] \to C[a,b]$ *defined by*

$$A(f)(x) = \int_a^x f, \ x \in [a,b].$$

Theorem 9.8. *Let* $A : \mathbb{R}^n \to \mathbb{R}^n$ *be a linear operator. TFAE*

1. *Operator* A *is invertible;*

2. $A(\mathbb{R}^n) = \mathbb{R}^n$*;*

3. $\det A \neq 0$.

Proof. The invertibility implies bijectivity. Consequently, $1 \Rightarrow 2$.

If $A(\mathbb{R}^n) = \mathbb{R}^n$, then the system of equations $Ax = y$ has a solution for every $y \in \mathbb{R}^n$, and consequently, $\det A \neq 0$ and $2 \Rightarrow 3$.

If $\det A \neq 0$, then the matrix (A) is invertible and a matrix $(A)^{-1}$ is a matrix of operator that is inverse to A. Consequently, $3 \Rightarrow 2$. □

Definition 9.9. *Let* X, Y *be real or complex normed spaces and* $A : X \to Y$ *be linear. A* ***norm of operator*** A *is a value*

$$\|A\| = \|A\|_{X \to Y} = \sup_{\|x\|_X \leq 1} \|Ax\|_Y . \tag{6}$$

Definition 9.10. *The operator is called bounded if $\|A\| < +\infty$. A set of bounded operators from X to Y is denoted by $\mathcal{L}(X \to Y)$ or $\mathcal{L}(X, Y)$.*

Theorem 9.11. *Let X, Y be real or complex normed spaces, $X \neq \{\theta\}$ and $A : X \to Y$ be linear. Then*

$$\sup_{\|x\|_X \leq 1} \|Ax\|_Y = \sup_{\|x\|_X < 1} \|Ax\|_Y = \sup_{\|x\|_X = 1} \|Ax\|_Y = \sup_{x \neq \theta} \frac{\|Ax\|_Y}{\|x\|_X} = \inf \left\{ C \in \mathbb{R} : \|Ax\|_y \leq C \|x\|_X \right\}. \quad (7)$$

Proof. Denote five values the equality of which we want to prove as N_1, N_2, N_3, N_4, N_5. We will prove that

$$N_3 \leqslant N_1 \leqslant N_5 \leqslant N_4 \leqslant N_2 \leqslant N_3.$$

The inequality $N_3 \leqslant N_1$ is obvious since the $\sup E \leq \sup F$ if $E \subset F$.

Now we will check that $N_1 \leqslant N_5$. If $N_5 = +\infty$ then the inequality is trivial. If $N_5 < +\infty$ then $\|Ax\| \leqslant (N_5 + \varepsilon) \|x\|$ for every $x \in X$ and $\varepsilon > 0$. Since ε is arbitrary then $\|Ax\| \leqslant N_5 \|x\|$ for every $x \in X$. Consequently, $\|Ax\| \leqslant N_5$ for every x such that $\|x\| \leq 1$. It remains to take supremum over x.

Now we prove that $N_5 \leqslant N_4$. If $N_4 = +\infty$ then the inequality is trivial. If $N_4 < +\infty$ then $\frac{\|Ax\|}{\|x\|} \leqslant N_4$ for every $x \neq \theta$. Hence, $\|Ax\| \leqslant N_4 \|x\|$ for every $x \in X$ and by the definition of infimum $N_5 \leqslant N_4$.

To check that $N_4 \leqslant N_2$ let $\varepsilon > 0$. Consider $x \neq \theta$ and denote $x_\varepsilon = \frac{x}{(1+\varepsilon)\|x\|}$. Consequently, $\|x_\varepsilon\| = \frac{1}{1+\varepsilon} < 1$ and

$$\frac{\|Ax\|}{\|x\|} = (1 + \varepsilon) \|Ax_\varepsilon\| \leqslant (1 + \varepsilon) N_2.$$

Considering supremum of the left-hand side we see that $N_4 \leqslant (1+\varepsilon)N_2$ and $N_4 \leqslant N_2$ since ε is arbitrary.

To prove that $N_2 \leqslant N_3$ we will check that $\|Ax\| \leqslant N_3$ for every $x \in X$ such that $\|x\| < 1$. If $x = \theta$ the inequality is trivial. If $x \neq \theta$ we let $\widetilde{x} = \frac{x}{\|x\|}$. Then $\|\widetilde{x}\| = 1$ and

$$\|Ax\| = \|x\| \|A\widetilde{x}\| \leqslant \|A\widetilde{x}\| \leqslant N_3.$$

Considering supremum of the left-hand side we obtain the estimate. $\square$

Remark 9.12. *For every $x \in X$*

$$\|Ax\| \leq \|A\| \|x\| .$$

Remark 9.13. *If for some $C > 0$ every $x \in X$*

$$\|Ax\| \leq C \|x\|$$

then $\|A\| \leq C$.

If for some $c > 0$ and some $x^* \in X$

$$\|Ax^*\| \geq c \|x^*\|$$

then $\|A\| \geq C$.

Corollary 9.13.1. *Let X, Y be real or complex normed spaces, $X \neq \{\theta\}$ and $A : X \to Y$ be linear. TFAE*

1. *A bounded;*

2. *A continuous at zero;*

3. *A continuous;*

4. A uniformly continuous.

5. An image of every bounded set is bounded;

6. An image of the unit ball $B(\theta, 1)$ is bounded;

Proof. Implications $4 \Rightarrow 3 \Rightarrow 2$, $5 \Rightarrow 6$ are obvious.

We'll check that $1 \Rightarrow 4$. If A obtains only zero values the the assertion is trivial. Assume that $\|A\| \neq 0$. Let $\varepsilon > 0$ and $\delta = \frac{\varepsilon}{\|A\|}$. Then for every $x, y \in X$ such that $\|x - y\| < \delta$

$$\|Ax - Ay\| = \|A(x - y)\| \leq \|A\| \, \|x - y\| < \varepsilon$$

and this proves the uniform continuity.

To prove $2 \Rightarrow 1$ let $\varepsilon = 1$ and apply the definition of continuity. Then there exists $\delta > 0$ such that $\|Ax\| < 1$ for every $x \in B(\theta, \delta)$. Consequently,

$$\|Ax\| = \frac{1}{\delta} \, \|A(\delta x\| < \frac{1}{\delta}$$

for every $x \in B(\theta, 1)$ and $\|A\| \leq \frac{1}{\delta}$.

$1 \Leftrightarrow 6$ Since $\|A\| = \sup\limits_{\|x\|<1} \|Ax\| = \sup\limits_{x \in B(\theta,1)} \|Ax\|$ then the boundedness of the image of the unit ball is equivalent to the finitenes of the left-hand side of this identity then boundedness of the image of the limit ball is equivalent to boundedness of the operator.

It remains to prove that $1 \Rightarrow 5$. Let $E \subset X$ be bounded. Then $E \subset B(\theta, R)$ for some $R > 0$. Consequently,

$$\|Ax\| \leq \|A\| \, \|x\| \leq \|A\| \, R$$

for every $x \in E$ and $A(E)$ is a subset of $B(\theta_Y, \|A\| \, R)$. $\square$

Examples.

- Linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$ is always continuous.
- Every linear operator defined on the finitely dimensional space is bounded.
- The operator of differentiation $D : C^1[a,b] \to C[a,b]$ is linear but discontinuous, if we define norm in spaces $C^1[a,b]$ and $C[a,b]$ by $\|f\| = \max\limits_{[a,b]} |f|$. (Exercise. consider a sequence $f_n(x) = \frac{x^n}{\sqrt{n}}$.)

Theorem 9.14 (On the estimate of the norm of linear operator in Euclidean spaces)**.** *Let* $A : \mathbb{R}^n \to \mathbb{R}^m$ *be linear with matrix* $\{a_{ik}\}_{i=1,}^{m}{}_{k=1}^{n}$. *Then*

$$\|A\|^2 \leq \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik}^2. \tag{8}$$

Proof. By the definition of Euclidean norm and Cauchy-Bunyakowsky-Schwartz inequality we see that

$$\|Ax\|^2 = \sum_{i=1}^{m} (A_i x)^2 = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} x_k \right)^2 \leqslant$$

$$\sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik}^2 \right) \left(\sum_{j=1}^{n} x_j^2 \right) = \left(\sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik}^2 \right) \|x\|^2 .$$

□

Corollary 9.14.1. *1. If* $A \in \mathcal{L}(\mathbb{R}^n \to \mathbb{R}^1)$ *then matrix* A *is a row vector* $(a_1, \ldots, a_n)$ *and* $\|A\|^2 = \sum\limits_{k=1}^{n} a_k^2$.

2. If $A \in \mathcal{L}(\mathbb{R}^1 \to \mathbb{R}^n)$ then matrix A is a column $(a_1, \ldots, a_m)^T$ and $\|A\|^2 = \sum\limits_{i=1}^{m} a_i^2$.

Definition 9.15. *A product of operators $A : X \to Y$ and $B : Y \to Z$ is the operator $BA : X \to Z$ defined by composition $BA(x) = B(A(x))$.*

Examples 9.16. *A product of linear operators $A : \mathbb{R}^n \to \mathbb{R}^m$ and $B : \mathbb{R}^m \to \mathbb{R}^l$ is linear operator $BA : \mathbb{R}^n \to \mathbb{R}^l$ whose matrix is a product of matrices of operators A and B, that is*

$$(BA) = (B)(A).$$

Remark 9.17. *Assume that X, Y, Z are normed spaces, $A : X \to Y$ and $B : Y \to Z$are linear operators. Then*

$$\|BA\| \leq \|B\| \|A\|.$$

In particular, the product of bounded operators is bounded

Proof. $\|B(A(x))\| \leq \|B\| \|A(x)\| \leq \|B\| \|A\| \|x\|$. □

Remark 9.18. *1. The identical operator in $\mathbb{R}^n$ is denoted by I or I_n. The diagonal elements of the matrix of this operator are equal to 1 and off-diagonal are 0.*
2. A matrix of operator A^{-1} that is inverse to A is inverse to the matrix of operator A, that is $(A^{-1}) = (A)^{-1}$, while $AA^{-1} = A^{-1}A = I$.

9.2 Equivalence of norms in $\mathbb{R}^n$.

Definition 9.19. *Let X be real or complex vector space with norms p_1 and p_2. We say that norms p_1 and p_2 are equivalent if there exist* *constants $c, C > 0$ such that*

$$cp_1(x) \le p_2(x) \le Cp_1(x), \;\; x \in X. \tag{9}$$

The convergence with respect to equivalent norms p_1, p_2 is equivalent, that is $p_1(x_n - a) \to 0 \iff p_2(x_n - a) \to 0$.

Theorem 9.20. *In $\mathbb{R}^n$ all norms are equivalent*

Proof. It is enough to prove that all norms are equivalent or Euclidean norm. A function p is continuous in Euclidean norm since

$$|p(x) - p(y)| \le p(x - y) = p(\sum_{k=1}^{n}(x_k - y_k)e^k) \le \\ \sum_{k=1}^{n} |x_k - y_k|\, p(e^k) \le M \left\| x - y \right\|,$$

where $M^2 = \sum\limits_{k=1}^{n} p(e^k)^2$.

Consequently, it by the Weierstrass theorem it obtains its maximal and minimal values on the unit sphere.

Let

$$c = \min_{\|x\|=1} p(x) \text{ and } C = \max_{\|x\|=1} p(x).$$

Then

$$0 < c \le p\left(x/\|x\|\right) \le C < +\infty$$

and

$$c\left\| x \right\| \le p(x) \le C \left\| x \right\|.$$

□

10 Differential calculus in Euclidean spaces.

Definition 10.1. *Assume that $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ and x is the interior point of D. A function f is **differentiable** at x if there exists a linear operator $A \in \mathcal{L}(\mathbb{R}^n \to \mathbb{R}^m)$ such that*

$$f(x+h) = f(x) + Ah + o(h), \; h \to \mathbb{O}_n, \tag{10}$$

*Operator A is a **differential** or **total derivative** of f at x and is denoted by $f'(x)$.*
*A matrix $(f'(x))$ of operator $f'(x)$ is called **Jacobi matrix** of f at x.*

Remark 10.2.

1. *By $\varphi(h) = o(h), \; h \to \mathbb{O}_n$ we mean that $\alpha(h) = \frac{\varphi(h)}{\|h\|} \to 0, \; h \to \mathbb{O}_n$. We also can assume that $\alpha(\mathbb{O}_n) = 0$.*

2. *If f is differentiable at a then it is continuous at x.*

3. *Sometimes we use notations $df(x,h)$ or $d_x f(h)$ for the value $f'(x)h$ of differential on vector h. In some applications we use notation dx for the vector and $df(x,dx) = f'(x)dx$.*

4. *Jacobi matrix of function $f : \mathbb{R}^n \to \mathbb{R}^m$ has m rows and n columns.*

Remark 10.3. *The differential is unique.*

Proof. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at x. Then for every $h \in \mathbb{R}^n$

$$f(x+th) = f(x) + tAh + o(t), \; t \to 0.$$

Consequently,

$$Ah = \lim_{t\to 0} \frac{f(x+th) - f(x)}{t}$$

and the value of operator A on the vector h is defined uniquely. $\square$

If $f : \mathbb{R}^n \to \mathbb{R}$ then its Jacobi matrix is $(A) = (\alpha_1, \ldots, \alpha_n)$ while $Ah = \langle \alpha,\ h \rangle = \sum\limits_{k=1}^{n} \alpha_k h_k$ is a **scalar product** of vectors $\alpha, h \in \mathbb{R}^n$.

Definition 10.4. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$. A function f is differentiable at $x \in \operatorname{int} D$ if there exists a vector $\alpha \in \mathbb{R}^n$ such that*

$$f(x+h) = f(x) + \langle \alpha,\ h \rangle + o(h), \quad h \to \mathbb{O}_n,$$

A vector α is called a gradient vector of f at x.

Notations: $\alpha = f'(x) = (f'(x)) = \operatorname{grad} f(x) = \nabla f(x)$.

Lemma 10.5. *Differentiability of a function $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is equivalent to differentiability of its coordinate functions. The rows of Jacobi matrix of f are gradients of coordinate functions.*

Proof. Assume that $f = (f_1, \ldots, f_m)$ is differentiable at x and write down the definition of differentiablity in coordinate functions:

$$f_i(x+h) = f_i(x) + A_i h + \alpha_i(h)\,|h|\,, \quad 1 \le i \le m. \tag{11}$$

Coordinate functions A_i of linear operator A are linear and the continuity and equality to zero of coordinate functions α_i is equivalent to the same property of function α. Consequently, f_i is differentiable.

To prove the inverse assertion assume that all coordinate functions are differentiable. Then for every $1 \le i \le m$ there exists linear function A_i and function α_i that is continuous an is equal to zero such that (11) holds. Consequently f satisfies the definition of differentiablity with $A = (A_1, \ldots, A_m)$ and $\alpha = (\alpha_1, \ldots, \alpha_m)$. $\square$

Remark 10.6. *The proof of the lemma implies that rows of Jacobi matrix of function f are gradients of coordinate functions of f, that is*

$$(f'(x)) = \begin{pmatrix} f_1'(x) \\ \cdots \\ f_m'(x) \end{pmatrix}$$

Definition 10.7. *Let*

$$D_1 = \{x \in \operatorname{int} D : f \text{ is differentiable at } x\}.$$

A function

$$f' : D_1 \to \mathcal{L}(\mathbb{R}^n \to \mathbb{R}^m),$$

which value at $x \in D_1$ is the operator $f'(x)$ is sometimes also called a derivative of f.

Examples 10.8. *1. The constant function is differentiable and its differential is zero.*
Indeed if $f(x) = C \in \mathbb{R}^m$ then for every $x, h \in \mathbb{R}$

$$f(x+h) = f(x) + \Theta h + \mathbb{O}_m,$$

where Θ is $m \times n$ matrix with all elements equal to 0. 2. The derivative of linear operator is this linear operator, that is $(Ax)' = A$.
This follows from identity

$$A(x+h) = Ax + Ah + \mathbb{O}_m.$$

1. Linearity of differentiation. If $f, g : D \subset \mathbb{R}^n \to \mathbb{R}^m$ are differentiable at $x \in \operatorname{int} D$, $\lambda \in \mathbb{R}$ then $f+g$, λf are differentiable at x and

$$(f+g)'(x) = f'(x) + g'(x),$$
$$(\lambda f)'(x) = \lambda f'(x).$$

Proof. Functions f and g are differentiable, consequently,

$$f(x+h) = f(x) + f'(x)h + o(h), \quad h \to \mathbb{O}_n, \tag{12}$$
$$g(x+h) = g(x) + g'(x)h + o(h), \quad h \to \mathbb{O}_n. \tag{13}$$

Summarizing (12) and (13) we see that

$$(f+g)(x+h) = (f+g)(x) + \left(f'(x) + g'(x)\right)h + o(h), \quad h \to \mathbb{O}_n.$$

Consequently $f+g$ satisfies the definition of differentiability at x and the differential operator is equal to $f'(x)+g'(x)$. The proof for λf is obtained by multiplication of (12) by λ. □

2. Derivative of composition. Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $x \in \operatorname{int} D$, $f(x) \in \operatorname{int} E$ and $g : E \subset \mathbb{R}^m \to \mathbb{R}^l$ be differentiable at $f(x)$. Then the composition $g \circ f$ be differentiable at x and

$$(g \circ f)(x) = g'(f(x)) \cdot f'(x). \tag{14}$$

Proof. Let $y = f(x)$ and apply the definition of differentiability of f at x and g at y

$$f(x+h) = f(x) + f'(x)h + \alpha(h)|h|,$$
$$g(y+k) = g(y) + g'(y)k + \beta(k)|k|,$$

where functions α and β are continuous and equal to 0 at 0. Let $k = f'(x)h + \alpha(h)|h| = \varkappa(h)$. Then

$$g(f(x+h)) = g(f(x)) + g'(f(x))\left(f'(x)h + \alpha(h)|h|\right) + \beta(\varkappa(h))|\varkappa(h)| =$$
$$= g(f(x)) + g'(f(x))f'(x)h + \gamma(h)|h|,$$

where

$$\gamma(h) = g'(y)\alpha(h) + \beta(\varkappa(h))\frac{|\varkappa(h)|}{|h|}.$$

We need to prove that $\gamma(h) \underset{h\to \mathrm{O}_n}{\longrightarrow} \mathbb{O}_l$. This will imply differentiability of $g \circ f$ at x with differential operator $g'(f(x))f'(x)$.

Since the differential operator $g'(y)$ os linear and continuous the first term is infinitely small as $h \to \mathbb{O}_n$. The composition $\beta \circ \kappa$ is equal to zero at zeroand continuous as the result of algebraic operation that involve continuous functions. Since

$$\frac{|\varkappa(h)|}{|h|} \leqslant \|f'(x)\| + |\alpha(h)|,$$

this ratio is bounded in some punctured neighbourhood of 0. consequently the second term is als infinitely small $\square$

3. Differentiation of a product of a vector and scalar function Assume that $D \subset \mathbb{R}^n$, functions $\lambda : D \to \mathbb{R}$ and $f : D \to \mathbb{R}^m$ are differntiable at $x \in \operatorname{int} D$. Then λf be differentiable at x and

$$(\lambda f)'(x)h = (\lambda'(x)h)f(x) + \lambda(x)(f'(x)h),\ h \in \mathbb{R}^n. \tag{15}$$

Proof. Considering coordinate function we may assume that $m = 1$. Let $h \in \mathbb{R}_n$ then by differentiability of λ and f we obtain

$$\begin{aligned}(\lambda f)(x+h)-(\lambda f)(x) = (\lambda(x+h)-\lambda(x))f(x+h)+\lambda(x)(f(x+h)-f(x)) = \\ = (\lambda'(x)h + o(h))\, f(x+h) + \lambda(x)\,(f'(x)h + o(h)) = \\ = (\lambda'(x)h)\, f(x) + \lambda(x)\,(f'(x)h) + \varphi(h),\end{aligned}$$

where

$$\varphi(h) = (\lambda'(x)h + o(h))\,(f(x+h) - f(x)) + o(h)f(x) + \lambda(x)o(h).$$

The differentiability of f implies continuity at x. Consequently, $\varphi(h) = o(h)$ as $h \to \mathbb{O}_n$. This implies that λf satisfies the definition of differentiabily at x with differential operator defined by (15). $\square$

4. Derivative of a scalar product. Let $f, g : D \subset \mathbb{R}^n \to \mathbb{R}^m$. The **scalar product** of f and g is the function defined as

$$\langle f, g\rangle = \sum_{i=1}^{m} f_i g_i : D \to \mathbb{R} \tag{16}$$

Let $f, g : D \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $x \in \operatorname{Int} D$ then $\langle f, g\rangle$ is differentiable at x and

$$\langle f, g\rangle'(x)h = \langle f'(x)h, g(x)\rangle + \langle f(x), g'(x)h\rangle\,, \quad h \in \mathbb{R}^n. \tag{17}$$

Proof. The rules for derivative of a sum and a multipllication of scalar function imply that $\langle f, g\rangle$ is differentiable at x and

$$\begin{aligned}\langle f, g\rangle'(x)h = \sum_{i=1}^{m} (f_i g_i)'(x)h = \sum_{i=1}^{m} (f_i'(x)hg_i(x) + f_i(x)g_i'(x)h) = \\ = \langle f'(x)h, g(x)\rangle + \langle f(x), g'(x)h\rangle\,.\end{aligned}$$

for every $h \in \mathbb{R}^n$ $\square$

Remark 10.9. *If $n = 1$ then formula 17 can be written as*

$$\langle f, g\rangle'(x) = \langle f'(x), g(x)\rangle + \langle f(x), g'(x)\rangle\,,$$

where $\langle f'(x), g(x)\rangle$ is the scalar product of $f'(x)$ and $g(x)$ as of the vectors in $\mathbb{R}^m$, and analogously for $\langle f(x), g'(x)\rangle$.

Remark 10.10. *The rules of differentiation (1-4) can be expressed in terms of Jacobi matrices as follows*

$$((\lambda f + \mu g)'(x)) = \lambda (f'(x)) + \mu (g'(x))$$
$$((g \circ f)'(x)) = (g'(f(x))) (f'(x))$$
$$((\lambda f)'(x)) = f(x) (\lambda'(x)) + \lambda(x) (f'(x))$$
$$(\langle f, g \rangle'(x)) = (g(x))^T (f'(x)) + (f(x))^T (g'(x)) \, .$$

10.1 Lagrange's theorem for vector functions

Theorem 10.11 (Lagrange's theorem for vector functions)**.** *Let $f : [a, b] \to \mathbb{R}^n$ be differentiable on $[a, b]$. Then there exists a point $c \in [a, b]$ such that*

$$|f(b) - f(a)| \leq |f'(c)| \, |b - a| \, . \tag{18}$$

Proof. Let $\varphi(x) = \langle f(x), f(b) - f(a) \rangle$. Then φ is continuous on $[a, b]$ and differentiable on (a, b). By the Lagrange's theorem applied to φ there exists a point $c \in (a, b)$ such that

$$\varphi(b) - \varphi(a) = \varphi'(c)(b - a).$$

By the definition of φ and the rule of differentiation of scalar product we see that

$$\varphi(b) - \varphi(a) = \langle f(b) - f(a), f(b) - f(a) \rangle = |f(b) - f(a)|^2,$$
$$\varphi'(c) = \langle f'(c), f(b) - f(a) \rangle \, .$$

Then, by Cauchy-Bunyakowsky-Schwartz inequality we see that

$$|f(b) - f(a)|^2 = \langle f'(c), f(b) - f(a) \rangle \, (b - a) \leqslant |f'(c)| \, |f(b) - f(a)| (b - a).$$

If $f(b) = f(a)$ then the assertion of the theorem is obvious. If $f(b) \neq f(a)$, when it is enough to divide the obtained inequality by $|f(b) - f(a)|$.

$\square$

Remark 10.12. *Lagrange's theorem for real-valued function*

$$f(b) - f(a) = f'(c)(b-a) \tag{19}$$

is not true for vector-functions in this form.

To see this consider a vector-function $f(x) = (\cos x, \sin x), x \in [0, 2\pi]$ *that describes the motion on the circle. Then*

$$f(2\pi) - f(0) = (0,0), \quad f'(x) = (-\sin x, \cos x),$$
$$|f'(x)| = \sqrt{(-\sin x)^2 + (\cos x)^2} = 1$$

for every x*. Consequently the identity (19) is not satisfied for any* $c \in [0, 2\pi]$.

Corollary 10.12.1 (Lagrange's theorem for mappings)**.** *Assume that* D *is open in* R^n, $f : D \to \mathbb{R}^m$ *is differentiable on* D*, while the segment that connects points* $a, b \in \mathbb{R}^n$ *is contained in* D*. Then there exists* $\theta \in (0,1)$ *such that* $|f(b) - f(a)| \leq |f'(a + \theta(b-a))| \, |b-a|$ *.*

Proof. Let

$$F(t) = f(a + t(b-a)), \quad t \in [0,1].$$

This definition is correct since $\overline{a,b} \subset D$. Then $F : [0,1] \to \mathbb{R}^m$ is differentiable on $[0,1]$. Moreover, $F(1) = f(b)$, $F(0) = f(a)$ and by the rule of the differentiation of the composition

$$F'(t) = f'(a + t(b-a))(b-a).$$

Consequently, by Theorem 10.11 there exists such $\theta \in (0,1)$ that

$$|F(1) - F(0)| \leqslant |F'(\theta)| \cdot 1.$$

Noticing that

$$|F'(\theta)| \leqslant \|f'(a + \theta(b-a))\| \, |b-a|$$

we conclude the proof. $\square$

Corollary 10.12.2. *Assume that in Corrollary 10.12.1* $\|f'(u)\| \leq M$ *on* $\overline{a,b}$. *Then*

$$|f(b) - f(a)| \leq M \, |b-a| \, .$$

10.2 Partial derivatives.

Definition 10.13. *Let* $f : D \subset \mathbb{R}^n \to \mathbb{R}, x \in \operatorname{int} D, \ h \in \mathbb{R}^n$. *The directional derivative of a function* f *along the vector* h *at point* x *is a limit*

$$\lim_{t \to 0} \frac{f(x+th) - f(x)}{t}. \tag{20}$$

If the directional derivative of a function f *along the vector* h *at point* x *exists then the function is differentiable at* x *along* h.
Notations: $D_h f(x) = \frac{\partial f}{\partial h}(x)$.

Remark 10.14. $D_h f(x) = F_h'(0)$, *where* $F_h'(t) = f(x+th)$.

Theorem 10.15. *If* $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *is differentiable at* $x \in \operatorname{int} D$ *then* f *is differentiable at* x *along every vector* h *and*

$$D_h f(x) = f'(x) h.$$

Proof. The rule for differentiation of composition implies differentiability of F_h at zero and that

$$D_h f(x) = F_h'(0) = f'(x)h.$$

□

A value of differential $f'(x)h$ can be expressed as a scalar product

$$f'(x)h = df(x,h) = \langle \operatorname{grad} f(x),\ h \rangle . \tag{21}$$

Corollary 10.15.1. *If $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable at $x \in \operatorname{int} D$,* $\operatorname{grad} f(x) \neq \mathbb{O}_n$*. Then for every direction $h \in \mathbb{R}^n$ (such that $\|h\| = 1$)*

$$-\left|\operatorname{grad} f(x)\right| \leq D_h f(x) \leq \left|\operatorname{grad} f(x)\right| .$$

The equality is obtained only for $h = \pm \frac{\operatorname{grad} f(x)}{|\operatorname{grad} f(x)|}$.

Proof. Theorem 10.15 and Cauchy-Bunyakowsky-Schwartz inequality imply that

$$|D_h f(x)| = |\langle \operatorname{grad} f(x), h \rangle| \leqslant |\operatorname{grad} f(x)| \cdot |h| = |\operatorname{grad} f(x)|.$$

and the inequality turns into identity if and only if vectors $\operatorname{grad} f(x)$ and h are collinear. □

Definition 10.16. *A partial derivative f along the $k-$th variable at x is a directional derivative*

$$D_k(f) = D_{x_k} f(x) = f'_{x_k}(x) = \frac{\partial f}{\partial x_k}(x) = \frac{\partial f}{\partial e^k}(x).$$

That means that

$$D_k f(x) = \lim_{t \to 0} \frac{f(x + te^k) - f(x)}{t} == F_k'(0),$$

where $F_k(t) = f(x + te^k)$ and $\varphi(u) = f(x_1, \ldots, u, \ldots, x_n)$.

Corollary 10.16.1. *Assume that $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable at $x \in \operatorname{int} D$. Then for every $1 \leq k \leq n$ there exists $D_k f(x)$ and*

$$D_k f(x) = f'(x)e^k = \left\langle \operatorname{grad} f(x),\ e^k \right\rangle . \tag{22}$$

Using the coordinates of the arguments we can write this definitions in the following form

$$D_k f(x) = \lim_{t \to 0} \frac{f(x_1, \ldots, x_k + t, \ldots, x_n) - f(x_1, \ldots, x_k, \ldots, x_n)}{t}.$$

In practice the calculation of partial derivative is reduced to the calculation of derivative of a function of one variable. Consider the auxiliary function (of one variable)

$$\varphi_k(u) = f(x_1, \ldots, u, \ldots, x_n)$$

(we fix all arguments of f except the k-th). Since $x \in \operatorname{Int} D$ the function φ_k is defined in some neighbourhood of x_k. By the definition of the derivative

$$D_k f(x) = \lim_{t \to 0} \frac{\varphi_k(x_k + t) - \varphi_k(x_k)}{t} = \varphi_k'(x_k) .$$

Notice that the notation $\frac{\partial f}{\partial x_k}$ has to be concerned as a single; x_k in denominator (as the index x_k in f'_{x_k} and $D_{x_k} f$) hos no relation with the k-th coordinate of point x at which the derivative is considered and only indicates the index of the variable. Symbol ∂ unlike the symbol d is used to denote the partial derivative or for derivative by the vector (don't confuse them!). If we consider a function with not so many variables we usually use different letters do denote them, for example, $f'_x(x,y,z)$ or $\frac{\partial f}{\partial z}(x,y,z)$.

Corollary 10.16.2 (Jacobi matrix and gradiaent in terms ofpartial derivatives)**.** *Assume that $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x \in \operatorname{int} D$. Then the elements of Jacobi matrix are partial derivatives of coordinate functions, that is*

$$(f'(x)) = (D_k f_i(x))_{i=1,k=1}^{m\quad n} = \begin{pmatrix} D_1 f_1(x) & \dots & D_n f_1(x) \\ \dots & \dots & \dots \\ D_1 f_m(x) & \dots & D_n f_m(x) \end{pmatrix}.$$

In particular for $m = 1$

$$\operatorname{grad} f(x) = (D_1 f(x), \dots, D_n f(x)).$$

10.3 Derivative of composition in coordinates

Assume that $f = (f_1, \dots, f_m) : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x \in \operatorname{int} D$, $f(x) \in \operatorname{int} E$ and $g = (g_1, \dots, g_l) : E \subset \mathbb{R}^m \to \mathbb{R}^l$ is differentiable at $f(x)$. Then $g \circ f$ is differentiable at x and

$$D_k (g \circ f)_j(x) = \sum_{i=1}^{m} D_i g_j(f(x)) D_k f_i(x). \tag{23}$$

Examples 10.17. ***1.*** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $x, h \in \mathbb{R}^n$. Then*

$$F'(t_0) = f'(x + t_0 h)h$$

2.

$$(f(r\cos t, r\sin t))'_r = \qquad D_1 f(r\cos t, r\sin t)\cos t + D_2 f(r\cos t, r\sin t)\sin t$$
$$(f(r\cos t, r\sin t))'_t = D_1 f(r\cos t, r\sin t)(-r\sin t) + D_2 f(r\cos t, r\sin t)(r\cos t)$$

10.4 Differentiability of a function with continuous partial derivatives.

Examples 10.18. ***1.*** *Let*

$$f(x,y) = \begin{cases} 1, & y = x^2,\ x > 0; \\ 0, & \textit{otherwise.} \end{cases}$$

Then $D_h f(0,0) = 0$ *for every* $h \in \mathbb{R}^2$ *while* f *is not differentiable (and even not continuous) at* $(0,0)$.
2. *Let*

$$f(x,y) = \begin{cases} 1, & xy = 0; \\ 0, & xy \neq 0. \end{cases}$$

Then

$$f'_x(0,0) = \lim_{t \to 0} \frac{f(t,0) - f(0,0)}{t} = 0.$$

Analogously, $f'_y(0,0) = 0$*. If* $h \in \mathbb{R}^2$ *is not collinear to* e^1 *and to* e^2 *then the function* $F(t) = f(th)$ *is discontinuous at zero and* $D_h f(0,0)$ *doesn't exist.*

Theorem 10.19. *Assume that* $f : D \subset \mathbb{R}^n \to \mathbb{R},\ x \in \operatorname{int} D$*, and all partial derivatives of* f *exist in a neighbourhood of* x *and are continuous at* x*. Then* f *is differentiable at* x*.*

Examples 10.20. ***2.*** *Let* $f(x,y) = x^2 + y^2$ *if only one of the numbers* x, y *is rational and* $f(x,y) = 0$ *otherwise. Then* f *is differentiable at zero,* $\operatorname{grad} f(0,0) = (0,0)$*, while* f *is discontinuous and has no partial derivatives at other points.*
3. *Let* $f(x,y) = (x^2+y^2)\sin\frac{1}{x^2+y^2}$ *for* $(x,y) \neq (0,0)$ *and* $f(0,0) = 0$*. Then* f *is differentiable* $\mathbb{R}^2$*, while* f'_x *and* f'_y *are discontinuous at* $(0,0)$*.*

Proof of Theorem 10.19. Let's do induction by $n \in \mathbb{N}$. For $n = 1$ the theorem is obvious and it is enough to prove induction step from $n-1$ to n.

Let $h = (h_1, \ldots, h_{n-1}, h_n) = h' + h_n e^n = \in \mathbb{R}^n$ and

$$R(h) = f(x+h) - f(x) - \sum_{k=1}^{n} D_k f(x) h_k =$$

$$(f(x+h'+h_n e^n) - f(x+h') - D_n f(x) h_n) + (f(x+h') - f(x) - \sum_{k=1}^{n-1} D_k f(x) h_k)$$

By induction assumption

$$\left| f(x+h') - f(x) - \sum_{k=1}^{n-1} D_k f(x) h_k \right| = \alpha_1(h') \left| h' \right|,$$

where $\alpha_1(h') \to 0$ as $|h'| \to 0$.

Let

$$\varphi(t) = f(x + h' + t h_n e^n)$$

then by Lagrange's theorem and formula for the derivative of composition there exists $c(h) \in (0,1)$ such that

$$f(x+h'+h_n e^n) - f(x+h') = \varphi(1) - \varphi(0) = \varphi'(c) = D_n f(x+h'+c(h) h_n e^n) h_n.$$

Consequently, by the continuity of $D_n f$ we obtain

$$|f(x+h'+h_n e^n) - f(x+h') - D_n f(x) h_n| =$$
$$|D_n f(x+h'+c(h)h_n e^n) - D_n f(x)|\,|h_n| = \alpha_2(h)\,|h_n| = o(h),$$

where $\alpha_2(h) \to 0$ as $|h| \to 0$. Finally, $R(h) = o(h)$ as $|h| \to 0$. $\square$

10.5 Higher order partial derivatives.

Definition 10.21. *Let $r \in \mathbb{N}$. Assume that partial derivatives of order $r-1$ are already defined. Let $1 \le i_1, \ldots, i_r \le n$, $f : D \subset \mathbb{R}^n \to \mathbb{R}$, $x \in D$. A partial derivative of order r with respect to the variables $i_1, \ldots, i_r$ at x is defined by*

$$D^r_{i_1 \ldots i_r} f(x) = D_{i_r}(D^{r-1}_{i_1 \ldots i_{r-1}} f)(x),$$

(if the right part of the equality exists).

Notations. $D^r_{i_1 \ldots i_r} f(x) = D^r_{x_{i_1} \ldots x_{i_r}} f(x) = f^{(r)}_{x_{i_1} \ldots x_{i_r}}(x) = \frac{\partial^r f}{\partial x_{i_1} \ldots \partial x_{i_r}}$.

If $i_1 = \ldots = i_r$ then the partial derivative

$$D^2_{x_i, x_i} f(x) = D^2_{x_i^2} f(x) = f''_{x_i^2}(x) = \frac{\partial^2 f}{\partial x_i^2}$$

is called "own" partial derivative with respect to variable x_i. Otherwise it is called the cross partial derivative.

Examples 10.22. ***1.*** *Let $f(x,y) = x^y$. Then*

$$f'_x(x,y) = yx^{y-1}, \quad f''_{xy}(x,y) = x^{y-1} + yx^{y-1} \ln x;$$

$$f'_y(x,y) = x^y \ln x, \quad f''_{yx}(x,y) = \frac{x^y}{x} + yx^{y-1} \ln x;$$

Here $f''_{xy} = f''_{yx}$.

2. *Let*

$$f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2}, & (x,y) \neq (0,0); \\ 0, & (x,y) = (0,0. \end{cases}$$

Then

$$f'_x(0,0) = \lim_{x\to 0} \frac{f(x,0) - f(0,0)}{x} = 0$$

and for $(x,y) \neq (0,0)$

$$f'_x(x,y) = y\frac{x^2-y^2}{x^2+y^2} + 2x^2y\frac{x^2+y^2-(x^2-y^2)}{(x^2+y^2)^2} = y\frac{x^4-y^4+4x^2y^2}{(x^2+y^2)^2}.$$

Consequently,

$$f''_{xy}(0,0) = \lim_{y\to 0} \frac{f(0,y) - f(0,0)}{y} = \lim_{y\to 0} \frac{-y^5}{y^5} = -1.$$

Analogously, $f'_y(0,0) = 0$ *and for* $(x,y) \neq (0,0)$

$$f'_y(x,y) = x\frac{x^4-y^4+4x^2y^2}{(x^2+y^2)^2}.$$

Thus $f''_{yx}(0,0) = 1 \neq f''_{xy}(0,0) = -1$.

Theorem 10.23 (Schwarz's theorem)**.** *Let* $f : D \subset \mathbb{R}^2 \to \mathbb{R}$, $(x^0, y^0) \in \operatorname{int} D$. *Assume that mixed partial derivatives* f''_{xy} *and* f''_{yx} *exist in a neighbourhood* V *of a point* (x^0, y^0) *and are continuous at* (x^0, y^0). *Then*

$$f''_{xy}(x^0, y^0) = f''_{yx}(x^0, y^0).$$

Proof. Let $h, k \in \mathbb{R}^n$ be such that $(x^0+h, y^0+k) \in V$. We will find two approximation of value

$$\Delta = f(x^0+h, y^0+k) - f(x^0+h, y^0) - f(x^0, y^0+k) + f(x^0, y^0).$$

in terms of mixed partial derivatives f''_{xy} and f''_{yx}.

1. Let $\varphi(s) = f(x^0+sh, y^0+k) - f(x^0+sh, y^0),\ s \in [0,1]$. Applying Lagrange's theorem we have $\theta_1 \in (0,1)$ such that

$$\Delta = \varphi(1) - \varphi(0) = \varphi'(\theta_1) = (f'_x(x^0+\theta_1 h, y^0+k) - f'_x(x^0+\theta_1 h, y^0))h.$$

Let $\tilde{\varphi}(t) = f_x(x^0+\theta_1 h, y^0+tk),\ t \in [0,1]$. Then there exists $\theta_2 \in (0,1)$ such that

$$\Delta = \tilde{\varphi}(1) - \tilde{\varphi}(0) = \tilde{\varphi}'(\theta_2) = f''_{xy}(x^0+\theta_1 h, y^0+\theta_2 k)hk.$$

2. Analogously $\Delta = f''_{yx}(x^0+\theta_3 h, y^0+\theta_4 k)hk$ for some $\theta_3, \theta_4 \in (0,1)$. Consequently,

$$f''_{yx}(x^0+\theta_3 h, y^0+\theta_4 k) = f''_{xy}(x^0+\theta_1 h, y^0+\theta_2 k).$$

and the continuity of f''_{xy} and f''_{yx} implies $f''_{xy}(x^0, y^0) = f''_{yx}(x^0, y^0)$. $\square$

Definition 10.24. *Assume that $r \in \mathbb{N}$ and $D \subset \mathbb{R}^n$ is open. A function $f : D \to \mathbb{R}$ is $r-$**smooth or of class** C^r on D if all its derivatives of order r exist and are continuous on D.*
*A function $f : D \to \mathbb{R}^m$ is $r-$**smooth or of class** C^r on D all its coordinate functions are of class C^r.*
A set of functions r-smooth on D is denoted by $C^r(D)$. $C^0(D) := C(D)$.
A class $C^\infty(D) := \bigcap\limits_{r=0}^{+\infty} C^r(D)$ is a set of all (infinitely) smooth functions

Theorem 10.25. *Let $f \in C^r(D)$. Assume that a tuple $(j_1, \ldots, j_r)$ is obtained from a tuple $(i_1, \ldots, i_r)$ by rearrangement. Then*

$$D^r_{j_1,\ldots,j_r} f(x) = D^r_{i_1,\ldots,i_r} f(x).$$

Proof. From the course of Algebra we know that any rearrangement is a composition of finite family of elementary transpositions (that rearrange two neighbor elements). Consequently, it is enough to prove that partial derivative doesn't change by elementary transposition of indexes:

$$D^r_{i_1,\ldots,i_{k-1},i_k,i_{k+1},i_{k+2}\ldots i_r} f(x) = D^r_{i_1,\ldots,i_{k-1},i_{k+1},i_k,i_{k+2}\ldots i_r} f(x), \quad i_k \neq i_{k+1}.$$

Let $x \in D$ and define

$$\begin{aligned}\varphi(s,t) = D^{k-1}_{i_1,\ldots,i_{k-1}} f\left(x + (s - x_{i_k})e^{i_k} + (t - x_{i_{k+1}})e^{i_{k+1}}\right) \\ = D^{k-1}_{i_1,\ldots,i_{k-1}} f(x_1, \ldots, s, \ldots, t, \ldots, x_n).\end{aligned}$$

Consequently, $\varphi \in C^2\left(V_{(x_{i_k}, x_{i_{k+1}})}\right)$ and by previous theorem

$$\varphi''_{st}(x_{i_k}, x_{i_{k+1}}) = \varphi''_{ts}(x_{i_k}, x_{i_{k+1}}),$$

that is

$$D^r_{i_1,\ldots,i_{k-1},i_k,i_{k+1},i_{k+2}\ldots i_r} f(x) = D^r_{i_1,\ldots,i_{k-1},i_{k+1},i_k,i_{k+2}\ldots i_r} f(x).$$

□

Remark 10.26. *Notice that $C^{r+1} \subset C^r$ and $C^{r+1} \neq C^r$.*

10.6 Multidimensional Taylor's formula.

- A **multiindex** is a tuple $k = (k_1, \ldots, k_n) \in \mathbb{N}_0^n$.
- A value $|k| = k_1 + \ldots + k_n$ is a **height of multiindex** k.
- $k! = k_1! \ldots k_n!$, $h^k = h_1^{k_1} \ldots h_n^{k_n}$, $h \in \mathbb{R}^n$;
- $f^{(k)}(x) = D^k f(x) = f^{(k_1,\ldots,k_n)}(x) = \frac{\partial^{|k|} f(x)}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}}$.

Theorem 10.27 (Taylor-Lagrange theorem for multivariate functions)**.** *Assume that $D \subset \mathbb{R}^n$ is open, $f \in C^{r+1}(D)$, $a, x \in \mathbb{R}^n$, $\overline{a, x} \subset D$. Then there exist $\theta \in (0,1)$ such that*

$$f(x) = \sum_{|k| \le r} \frac{f^{(k)}(a)}{k!}(x-a)^k + \sum_{|k|=r+1} \frac{f^{(k)}(a + \theta(x-a))}{k!}(x-a)^k. \quad (24)$$

A polynomial $T_{r,a} f(x) = \sum\limits_{|k| \le r} \frac{f^{(k)}(a)}{k!}(x-a)^k$ is a **Taylor's polynomial** of degree r of function f at a. The difference $R_{r,a} f(x) = f(x) - T_{r,a} f(x)$ is a **residue or residual term** in Taylor's-Lagrange's formula.

Lemma 10.28. *Let $f \in C^r(D)$ and assume that a segment that connects points $x, x+h$ lies in D. Let*

$$F(t) = f(x + th), \ t \in [0,1].$$

Then $F \in C^r[0,1]$ and

$$F^{(l)}(t) = \sum_{|k|=l} \frac{l!}{k!} f^{(k)}(x+th) h^k, \ 0 \le l \le r. \quad (25)$$

Proof. Denote $g(t) = x + th$. Then $g \in C^\infty([0,1] \to D)$ and if $F = f \circ g$ then $F \in C^r[0,1]$. We will prove the equality (10.28) using induction by l.

Base. If $l = 0$ then the equation is obvious.

Induction step. Let $r \in \mathbb{N}$ and assume that the equality (10.28) holds for some $l \in [0 : r-1]$. By the inductional assumption we see that

$$F^{(l+1)}(t) = (F^{(l)})'(t) = \sum_{|k|=l} \frac{l!}{k!} \frac{d}{dt} f^{(k)}(x+th)h^k.$$

Applying the chain rule we see that

$$\frac{d}{dt} f^{(k)}(x+th) = \sum_{i=1}^{n} D_i(k)(x+th)h_i = \sum_{i=1}^{n} f^{(k+e^i)}(x+th)h_i.$$

Let $p = k + e^i$ as the new (multi)index of summation. Then the condition $|k| = l$ implies that $|p| = l$ and $p_i \geq 1$. However the term $k_i + 1 = p_i$ in nominator allows us to omit the condition $p_i \geq 1$ since the corresponding term vanishes when $p_i = 0$. Changing the order of summation after this change we obtain

$$F^{(l+1)}(t) = \sum_{i=1}^{n} \sum_{|p|=l+1} \frac{l! p_i}{p!} f^{(p)}(x+th)h^p = \sum_{|p|=l+1} \frac{l!}{p!} f^{(p)}(x+th)h^p \sum_{i=1}^{n} p_i =$$

$$= \sum_{|p|=l+1} \frac{l!}{p!} f^{(p)}(x+th)h^p(l+1) = \sum_{|p|=l+1} \frac{(l+1)!}{p!} f^{(p)}(x+th)h^p.$$

$\square$

Remark 10.29. *Denote points as a and $x = a + h$ (assuming that the closed segment that connects points a and $a+h$ is a subset of D).*

Then the Taylor-Lagranges formula has the following form

$$f(a+h)=\sum_{|k|\leqslant r}\frac{f^{(k)}(a)}{k!}h^k+\sum_{|k|=r+1}\frac{f^{(k)}(a+\theta h)}{k!}h^k.$$

We will use this reformulation to prove the theorem.

Proof of Theorem 10.27. Let

$$F(t)=f(a+th),\quad t\in[0,1].$$

Then $F\in C^{(r+1)}[0,1]$ by Lemma 10.28 and we can apply Taylor-Lagranges formula for the function of one variable t. There exists such $\theta\in(0,1)$ that

$$F(1)=\sum_{l=0}^{r}\frac{F^{(l)}(0)}{l!}(1-0)^l+\frac{F^{(r+1)}(\theta)}{(r+1)!}(1-0)^{r+1}.$$

Applying formula 25 to calculate derivatives of F we see that

$$f(x+h)=\sum_{l=0}^{r}\frac{1}{l!}\sum_{|k|=l}\frac{l!}{k!}f^{(k)}(a)h^k+\frac{1}{(r+1)!}\sum_{|k|=r+1}\frac{(r+1)!}{k!}f^{(k)}(a+\theta h)h^k=$$
$$=\sum_{|k|\leqslant r}\frac{f^{(k)}(a)}{k!}h^k+\sum_{|k|=r+1}\frac{f^{(k)}(a+\theta h)}{k!}h^k.$$

□

Corollary 10.29.1 (Polynomial formula)**.**

$$\left(\sum_{k=1}^{n}x_k\right)^r=\sum_{|k|=r}\frac{r!}{k!}x^k,\ r\in\mathbb{N}_0,\ x=(x_1,\ldots,x_n)\in\mathbb{R}^n.$$

Corollary 10.29.2 (Taylor-Peano theorem for multivariate functions). *Assume that $D \subset \mathbb{R}^n$ is open, $r \in \mathbb{N}_0$, $f \in C^r(D)$, $a \in D$. Then*

$$f(a+h) = \sum_{|k| \leq r} \frac{f^{(k)}(a)}{k!} h^k + o(|h|^r), \;\; h \to \mathbb{O}. \tag{26}$$

Proof. Since $f \in C^r$ we can apply the Taylor's-Lagrange's formula for a polynomial of degree $r-1$

$$f(x+h) = \sum_{|k| \leqslant r-1} \frac{f^{(k)}(x)}{k!} h^k + \sum_{|k|=r} \frac{f^{(k)}(x+\theta h)}{k!} h^k = \sum_{|k| \leqslant r} \frac{f^{(k)}(x)}{k!} h^k + R(h),$$

where

$$R(h) = \sum_{|k|=r} \frac{f^{(k)}(x+\theta h) - f^{(k)}(x)}{k!} h^k, \quad \theta \in (0,1).$$

Notice that $|h_i| \leqslant |h|$. Then, for every multiindex $(k) = r$, we have the following estimate

$$\frac{|h^k|}{|h|^r} = \frac{|h_1|^{k_1}}{|h|^{k_1}} \cdots \frac{|h_n|^{k_n}}{|h|^{k_n}} \leqslant 1.$$

Consequently,

$$\frac{|R(h)|}{|h|^r} \leqslant \sum_{|k|=r} \frac{\left| f^{(k)}(x+\theta h) - f^{(k)}(x) \right|}{k!},$$

that tends to 0 as $h \to \mathbb{O}_n$ by continuity of all derivatives of order r. $\square$

A function $d^l f(a, \cdot) : \mathbb{R}^n \to \mathbb{R}$ defined by

$$d^l f(a,h) = \sum_{|k|=l} \frac{l!}{k!} f^{(k)}(a) h^k$$

is the **differential of order** l of function f at a. With this notation we can rewrite Taylor-Lagrange's formula in Theorem 10.27) as

$$f(a+h) = \sum_{l=0}^{r} \frac{1}{l!} d^l f(a,h) + \frac{1}{(r+1)!} d^{r+1} f(a+\theta h, h). \qquad (27)$$

10.7 Extremal points of multivariate function. Unconditional optimization.

Definition 10.30. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R},\ p \in D$. Assume that there exists a neighborhood V_p of p such that*

- $\forall x \in V_p \cap D\ f(x) \leq f(p)$. *Then p is* ***a (local) maximum point*** *of a function f;*
- $\forall x \in \dot{V}_p \cap D\ f(x) < f(p)$. *Then p is* ***a (local) strict maximum point*** *of a function f;*
- $\forall x \in V_p \cap D\ f(x) \geq f(p)$. *Then p is* ***a (local) minimum point*** *of a function f;*
- $\forall x \in \dot{V}_p \cap D\ f(x) > f(p)$. *Then p is* ***a (local) strict minimum point*** *of a function f.*

If p is a (strict) minimum or maximum point then p is a (strict) extremum point.

Remark 10.31. *If $p \in \operatorname{int} D$. Then some neighborhood of p is contained in D and we can omit intersection with D in this definition.*

Remark 10.32. *The function doesn't necessary obtain it's maximal and minimal value at the extremal point, it will be such only for sufficiently near points. That's why points from definition are called local extremal points. Points at which function obtains its maximal and minimal value are called global extremal points.*

Theorem 10.33 (Necessary condition for extremal point)**.** *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$, $k = 1 \dots n$. Assume that $p \in \operatorname{int} D$ is an extremal point of f and the partial derivative $D_k F(p)$ exists. Then $D_k f(p) = 0$.*

Proof. By definition of extremal point there exists $\delta > 0$ such that

$$f(p) = \max_{V_p(\delta)} f \quad \text{or} \quad f(p) = \min_{V_p(\delta)} f.$$

Let

$$F_k(t) = f(p + te^k), \quad |t| < \delta.$$

Then $F_k'(0)$ exists and is equal to $D_k f(p)$, while $t_0 = 0$ is an extremal point of F_k. Consequently, by necessary condition for extremal point of function of one variable,

$$D_k f(p) = F_k'(0) = 0.$$

$\square$

Remark 10.34. *If in conditions of Theorem 10.33 $D_k f(p)$ exists for every $k = 1 \dots n$ then $D_k f(p) = 0$ for every $k = 1 \dots n$. For differentiable function we can write down this system of n variables in the following form*

$$\operatorname{grad} f(p) = \mathbb{O}_n,$$

Definition 10.35. *Let* $f : D \subset \mathbb{R}^n \to \mathbb{R}$*. A point* $p \in \operatorname{int} D$ *is stationary point if* $\operatorname{grad} f(p) = \mathbb{O}_n$*.*

In particular, if function f is differentiable at p then a point p is stationary.

Consequently, extremal points are among stationary points and point at which function is not differentiable.

The extremality of a stationary point may be studied in terms of second differential

$$d^2 f(p, h) = \sum_{k,j=1}^{n} D_{ij}^2 f(p) h_i h_j \tag{28}$$

which is a quadratic form.

Definition 10.36. *Let* K *be a quadratic form in* n *variables.*

1. *A form* K *is* ***positively definite*** *if* $K(h) > 0,\ h \in \mathbb{R}^n \setminus \{\mathbb{O}_n\}$*.*

2. *A form* K *is* ***negatively definite*** *if* $K(h) < 0,\ h \in \mathbb{R}^n \setminus \{\mathbb{O}_n\}$*.*

3. *A form that has values of different signs, that is there exist* $h_1, h_2 \in \mathbb{R}^n$ *such that* $K(h_1)K(h_2) < 0$ *is* ***indefinite****.*

4. *A form is* ***positively semidefinite*** *if* $K(h) \geq 0$ *(*$K(h) \leq 0$*) and there exists* $h \neq \mathbb{O}_n$ *such that* $K(h) = 0$*.*

Remark 10.37. *Assume that* K *is a positively definite quadratic form. Then there exist* $c > 0$ *such that*

$$K(h) \geq c\,|h|^2 .$$

Proof. Since a quadratic form is continuous it obtains its maximal and values on compact sets. Let $c = \min\limits_{|h|=1} K(h) > 0$. Then

$$K(h) = K\left(h/\|h\|\right)\|h\|^2 \geq c\,|h|^2,$$

where $\alpha(h) \to 0,\ \|h\| \to 0.$ $\square$

Theorem 10.38 (Sufficient Condition for extremum of function of several complex variables)**.** *Let D be an open set in $\mathbb{R}^n$, $f \in C^2(D)$, $p \in D$ be a stationary point of function f. Then the following assertions hold.*

1. *If the form $d^2 f(p)$ is positive definite then p is a point of strict minimum of f.*
2. *If the form $d^2 f(p)$ is negative definite then p is a point of strict maximum of f.*
3. *If the form $d^2 f(p)$ is indefinite then p is not extremal point of f.*

Proof. We apply Taylor-Peano formula for function f

$$f(p+h) = f(p) + d^1 f(p,h) + \frac{1}{2}d^2 f(p,h) + \frac{1}{2}\alpha(h)\,|h|^2,$$

where $\alpha(h) \to 0$ as $h \to \mathbb{O}_n$. Let

$$R(h) = 2(f(p+h) - f(p)), \quad K(h) = d^2 f(p,h).$$

Since p is a stationary point then $d^1 f(p,h) = 0$. Consequently,

$$R(h) = K(h) + \frac{1}{2}\alpha(h)\,|h|^2.$$

1. If $K(h)$ is positive definite then there exists $\gamma > 0$ such that

$$K(h) \geq \gamma \left|h\right|^2, \quad h \in \mathbb{R}^n.$$

Since α is infinitesimal there exists $\delta > 0$ such that

$$|\alpha(h)| \leq \frac{\gamma}{2}, \quad 0 < |h| < \delta.$$

Consequently,

$$R(h) \geq \gamma \left|h\right|^2 - \frac{\gamma}{2}\left|h\right|^2 = \frac{\gamma}{2}\left|h\right|^2 > 0, \quad 0 < |h| < \delta,$$

and p is a point of strict minimum of f.
2. Since $d^2(-f)(p,h) = -d^2f(p,h)$ this case can be deduced to the previous by consideration of function f.
3. Since K is no definite there exists $h^1, h^2 \in \mathbb{R}^n \setminus \{\mathbb{O}_n\}$ for which $K(h^1) > 0$ and $K(h^2) < 0$. Since function α is infinitesimal there exists such $\delta > 0$ that

$$\left|\alpha(th^1)\right| < \frac{K(h^1)}{2\left|h^1\right|^2}, \quad t \in (0, \delta).$$

Consequently for every $t \in (0, \delta)$ we have

$$R(th^1) = K(th^1) + \alpha(th^1)\left|th^1\right|^2 > t^2 K(h^1) - t^2\frac{K(h^1)}{2} = t^2\frac{K(h^1)}{2} > 0.$$

Analogously, $R(th^2) < 0$ if t is sufficiently small. This implies that p is not extremal point of f. $\square$

Remark 10.39. *If $d^2f(p)$ is semidefinite Theorem 10.38 doesn't allow do determine the type of function p. For example for functions $f(x,y) = x^4 + y^4$ and $g(x,y) = x^4 - y^4$ the point $(0,0)$ is stationary and the second differential at this point is 0. Nevertheless, $(0,0)$ is a point of local minimum of function f and is not extremal for function g since $g(t,2t) < 0$ and $g(2t,t) > 0$ for every $t \neq 0$.*

10.7.1 Sylvester's criterion

Principal minors of a matrix $A = (a_{ij})_{i,j=1\dots n}$ are defined by

$$\Delta_k = \det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}. \tag{29}$$

Theorem 10.40 (Sylvester's criterion)**.** *Let $\Delta_k,\ k = 1\dots n$, be principal minor of symmetric matrix of a quadratic form K.*

Then K is positively defined iff $\Delta_k > 0,\ k = 1\dots n$.

Then K is negatively defined iff $(-1)^k\Delta_k > 0,\ k = 1\dots n$.

10.8 Inverse function theorem.

Recall that a linear operator $\mathbb{R}^n \to \mathbb{R}^n$ is invertible if and only if it is bijective (and is bijective iff is injective or surjective).

Lemma 10.41. *Let $B \in \mathcal{L}(\mathbb{R}^n)$. Assume that there exists a number $m > 0$ such that*

$$\|Bx\| \geq m\|x\|,\ x \in \mathbb{R}^n.$$

Then the operator B is invertible and $\left\|B^{-1}\right\| \leq m^{-1}$.

Proof. Since $\|Bx\| \neq 0,\ x \neq \mathbb{O}_n$ then the operator B is bijective and, consequently, is invertible

$$m\left\|B^{-1}y\right\| \leq \left\|B(B^{-1}y)\right\| = \|y\|.$$

∎

Theorem 10.42 (An operator that is close to invertible is invertible.). *Assume that the linear operator $A : \mathbb{R}^n \to \mathbb{R}^n$ is invertible and $B \in \mathcal{L}(\mathbb{R}^n)$, $\|B - A\| < \left\|A^{-1}\right\|^{-1}$. Then*

1. *operator B is invertible;*
2. $\left\|B^{-1}\right\| \le \frac{1}{\left\|A^{-1}\right\|^{-1} - \|B-A\|}$;
3. $\left\|B^{-1} - A^{-1}\right\| \le \frac{\left\|A^{-1}\right\| \|B-A\|}{\left\|A^{-1}\right\|^{-1} - \|B-A\|}$.

Proof. The triangle inequality implies that

$$\|Bx\| \ge \|Ax\| - \|(B-A)x\| \ge \left(\left\|A^{-1}\right\|^{-1} - \|B - A\|\right)\|x\|$$

since $\|x\| = \left\|A^{-1}Ax\right\| \le \left\|A^{-1}\right\| \|Ax\|$.

Consequently, operator B is invertible and the estimate from second assertion is satisfied.

$$\left\|B^{-1} - A^{-1}\right\| = \left\|B^{-1}(A-B)A^{-1}\right\| \le \frac{\left\|A^{-1}\right\| \|B-A\|}{\|A^{-1}\|^{-1} - \|B-A\|}$$

□

Corollary 10.42.1. *Assume that the linear operator $A : \mathbb{R}^n \to \mathbb{R}^n$ is invertible and $\|B_k - A\| \to 0$. Then starting from some number operator B_k is invertible and $\left\|B_k^{-1} - A^{-1}\right\| \to 0$.*

Remark 10.43. *Assume that $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ is invertible, $x \in \operatorname{int} D$, $y = f(x) \in \operatorname{int} f(D)$, f is differentiable at x, f^{-1} is differentiable at y. Then the operator $f'(x)$ is invertible and*

$$(f^{-1})'(y) = (f'(x))^{-1}.$$

Lemma 10.44. *Let $D \subset \mathbb{R}^n$ be open, $f \in C^1(D \to \mathbb{R}^n)$, $a \in D$. Let the operator $A = f'(a)$ be invertible and $\lambda = \frac{1}{4}\left\|A^{-1}\right\|^{-1}$. Then there exists a neighbourhood U of a point a such that the following assertions are satisfied.*

1. *For every $x \in U$ operator $f'(x)$ is invertible;*

2. *If $x, x+h \in U$ then*

$$\|f(x+h) - f(x) - Ah\| \leq 2\lambda \|h\|, \tag{30}$$
$$\|f(x+h) - f(x)\| \geq 2\lambda \|h\|. \tag{31}$$

Proof. There exists a neighbourhood U of a point a such that

$$\|f'(x) - A\| < 2\lambda.$$

Then the operator $f'(x)$ is invertible for $x \in V$ by Theorem 10.42.

Let $F(u) = f(u) - Au$. Then

$$\|F'(u)\| = \|f'(u) - A\| < 2\lambda,\ u \in V,$$

And By the Lagrange's theorem for vector-valued functions we obtain

$$\|f(x+h) - f(x) - Ah\| = \|F(x+h) - F(x)\| \leq 2\lambda \|h\|.$$

Finally,

$$\|f(x+h) - f(x)\| \geq \|Ah\| - \|f(x+h) - f(x) - Ah\| \geq$$
$$2\lambda \|h\| \geq 4\lambda h - 2\lambda h = 2\lambda h.$$

□

Theorem 10.45 (Inverse function theorem)**.** *Assume that $D \subset \mathbb{R}^n$ is open, $f \in C^1(D \to \mathbb{R}^n)$, $a \in D$, operator $f'(a)$ is invertible. Then there exists a neighbourhood U of a point a such that*

1. *function $f\big|_U$ is invertible;*
2. *a set $V = f(U)$ is open;*
3. *$f^{-1} \in C^1(V \to U)$;*
4. *$(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1}$ for every $y \in V$.*

Remark 10.46. *If in the statement of the inverse function theorem $f \in C^r$ then $f^{-1} \in C^r$.*

Proof. Let $A = f'(a)$, $\lambda = \frac{1}{4}\left\|A^{-1}\right\|^{-1}$. The neighborhood U is chosen by Lemma 10.44.
1. The function $f\big|_U$ is invertible by estimate (30).
2. Let's prove that $V = f(U)$ is open. Let $y_0 \in V$. Then $\exists x_0 \in U$: $f(x_0) = y_0$. Let $\overline{B}(x_0, r) \subset U$. We will prove that

$$B(y_0, \lambda r) \subset f(B(x_0, r)) \subset f(U) = V.$$

Let $y \in B(y_0, \lambda r)$ and $F(x) = \|f(x) - y\|, x \in \overline{B}(x_0, r)$.

Then $F(x_0) < \lambda r$, function F is continuous on the compact set $\overline{B}(x_0, r)$ and by Weierstrass theorem obtains its minimal value at some point x^*. We will show that $f(x^*) = y$.

If $\|x - x_0\| = r$ then

$$F(x) > F(x) + F(x_0) - \lambda r = \|f(x) - y\| + \|f(x_0) - y\| - \lambda r \geq \\ \|f(x) - f(x_0)\| - \lambda r \geq 2\lambda r - \lambda r = \lambda r > F(x_0).$$

Consequently, $x^* \in B(x_0, r)$.

Consider a function $F^2(x) = \langle f(x) - y,\ f(x) - y \rangle$. Then

$$\operatorname{grad} F^2(x^*) = 2(f(x^*) - y)^T f'(x^*) = \mathbb{O}_n.$$

The operator $f'(x)$ is invertible. Consequently, $f(x^*) = y$ and $y \in f(B(x_0), r)$.

3,4. Let $y = f(x), y + k = f(x + h)$. Then

$$\left\| f^{-1}(y + k) - f^{-1}(y) \right\| = \|h\| \le (2\lambda)^{-1} \|h\|,$$

and this proves the continuity of f^{-1}.

We will prove the differentiability of f^{-1}. Let $y \in V$ and $x = f^{-1}(y)$. Then

$$f(x + h) - f(x) = f'(x)h + \alpha(h) \|h\|, \tag{32}$$

where $\alpha(\mathbb{O}) = \mathbb{O}$ and α is continuous at zero. Let $y + k \in V$ and

$$h = f^{-1}(y + k) - f^{-1}(y) = \tau(k).$$

Consequently, applying $(f'(x))^{-1}$ to equality (32) we obtain

$$(f'(x))^{-1}k = f^{-1}(y + k) - f^{-1}(y) + (f'(x))^{-1}\alpha(h) \|h\|,$$

$$f^{-1}(y + k) - f^{-1}(y) = (f'(x))^{-1}k + \beta(k) \|k\|,$$

where $\beta(k) = -(f'(x))^{-1}\alpha(h)\frac{\|h\|}{\|k\|}$. Finally,

$$\|\beta(k)\| \le (2\lambda)^{-1} \left\| (f'(x))^{-1} \right\| \|\alpha(\tau(k))\| \to 0,\ \|k\| \to 0.$$

$\square$

Theorem 10.47 (Open mapping theorem.)**.** *Let $D \subset \mathbb{R}^n$ be open, $f \in C^1(D \to \mathbb{R}^n)$. If $f'(a)$ is invertible for every $a \in D$ and the mapping f is open, that is $f(G)$ is open for every open set $G \subset D$.*

Proof. Let G be open and $y \in f(G)$. Then there exists point $x \in G$ such that $y = f(x)$. Operator $f'(x)$ is invertible, consequently, by th inverse function theorem there exists a neighborhood U of x such that $U \subset G$ and $V = f(U)$ is open. Consequently, $y \in V \subset G$ is the inner point of $f(G)$. $\square$

Definition 10.48. *Let $U, V \subset \mathbb{R}^n$ be open. A function $f : U \to V$ is a **diffeomorphism** if a function f is bijective and $f, f^{-1} \in C^1$.*

Examples 10.49. *Let $f(x, y) = (e^x \cos y, e^x \sin y)$. Then $f \in C^\infty(\mathbb{R}^2 \to \mathbb{R}^2)$ and*

$$\det f'(x, y) = \det \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = e^{2x} \neq 0.$$

A function f is locally invertible but not globally since

$$f(x, y + 2\pi) = f(x, y).$$

10.9 Implicit function.

Let

$$x = (x_1, \ldots, x_n) \in \mathbb{R}^n, \quad y = (y_1, \ldots, y_m) \in \mathbb{R}^m,$$
$$(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^{n+m}.$$

Assume that a function $\Phi : \mathbb{R}^{n+m} \to \mathbb{R}^m$ is differentiable at a point $a = (x^0, y^0)$. We define

$$\Phi'_x(a)h = \Phi'(a)(h, \mathbb{O}_m),\ h \in \mathbb{R}^n; \Phi'_y(a)k \quad = \Phi'(a)(\mathbb{O}_n, k),\ k \in \mathbb{R}^n.$$

Then $\Phi'(x, y)(h, k) = \Phi'_x(a)h + \Phi'_y(a)k$.

Consider the equation $\Phi(x, y) = \mathbb{O}_m$ which we can consider as a system of equations

$$\begin{cases} \Phi_1(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \Phi_m(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0; \end{cases} \tag{33}$$

Definition 10.50. *Let $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ be open, $U \times V \subset D$. If for every $x \in V$ there exists a unique $y \in V$ such that $\Phi(x, y) = \mathbb{O}_m$, then we say that an **implicit function** $\varphi : U \to V$ is declared.*

And a the same time $\Phi(x, \varphi(x)) = 0$, $x \in U$.

Theorem 10.51 (Implicit function theorem)**.** *Let $D \subset \mathbb{R}^{n+m}$ be open, $\Phi \in C^1(D \to \mathbb{R}^m)$, $(x^0, y^0) \in D$, $\Phi(x^0, y^0) = 0$. Assume that operator $\Phi'_y(x^0, y^0)$ is invertible. Then there exist neighborhoods U, V of points x_0, y_0 such that*

1. *there exists an implicit function $\varphi : U \to V$ such that*

$$\Phi(x, \varphi(x)) = \mathbb{O}_m, \ x \in U.$$

2. *$\varphi \in C^1(U \to V)$.*
3. *$\varphi'(x) = -(\Phi'_y(x, y))^{-1}\Phi'_x(x, y)$, $x \in U$.*

Proof. **1.** Consider an auxiliary function

$$F(x, y) = (x, \Phi(x, y)), \quad (x, y) \in D.$$

First n coordinate functions of F are the same as for identical map and the last m same as for Φ. Consequently, $F \in C^1(D \to \mathbb{R}^{n+m})$ and Jacobi matrix of F has is as following

$$(F'(x,y)) = \begin{pmatrix} \mathbb{1}_{n\times n} & \mathbb{O}_{n\times m} \\ \Phi'_x(x,y) & \Phi'_y(x,y) \end{pmatrix},$$

where $\mathbb{1}_{n\times n}$ is identity matrix of size $n \times n$ and $\mathbb{O}_{n\times m}$ is zero matrix of size $n \times m$. Notice that

$$\det F'(x,y) = \det \Phi'_y(x,y), \quad (x,y) \in D.$$

Consequently, $\det F'(x^0,y^0) \neq 0$ and operator $F'(x^0,y^0)$ is invertible. By the inverse function theorem there exists $r > 0$ such that the restriction of F onto the ball $W = B_{n+m}((x^0,y^0),r)$ is invertible and for every $(x,y) \in W$ the operator $F'(x,y)$ is invertible. Let

$$U_1 = B_n\left(x^0, r/\sqrt{2}\right), \quad V = B_m\left(x^0, r/\sqrt{2}\right).$$

Then for $U_1 \times V \subset W$. Since $U_1 \times V$ is open its image $W_1 = F(U_1 \times V)$ is also open by the open mapping theorem. Since $(x^0, \mathbb{O}_m) = F(x^0,y^0) \in W_1$ there exists an open neighborhood U of x^0 that $(x, \mathbb{O}_m)$ for every $x \in U$.

We will prove that U and V satisfy the requirements of the theorem. Let $x \in U$. Then $(x,\mathbb{O}_m) \in W_1$ and $F^{-1}(x,\mathbb{O}_m) \in U_1 \times V$. Let $(u,y) = F^{-1}(x,\mathbb{O}_m)$. This equality means that $u = x$ and $\Phi(x,y) = \mathbb{O}_m$. To check uniqueness of y we assume that there exists $\tilde{y} \in V$ such that $\Phi(x,\tilde{y}) = 0$. Then (x,y), $(x,\tilde{y}) \in W$ and $F(x,y) = F(x,\tilde{y}) = (x,\mathbb{O}_m)$. By invertibility of $F|_W$ we see that $y = \tilde{y}$.

2. Consider φ as a composition of three functions

$$\varphi = P \circ F^{-1} \circ Q,$$

where $Qx = (x, \mathbb{O}_m)$ for $x \in \mathbb{R}^n$ and $P(x,y) = y$ for $(x,y) \in \mathbb{R}^{n+m}$. Operators P, Q are linear and, consequently, smooth, while the function F^{-1} is C^1-smooth by the inverse function theorem. Consequently, $\varphi \in C^1(U \to V)$.

3. The equality

$$\Phi(x, \varphi(x)) = \mathbb{O}_m$$

holds on U identically. Considering derivative of this identity and writing down the result in terms of Jacobi matrices we see that

$$\Phi'_x(x, \varphi(x)) + \Phi'_y(x, \varphi(x))\varphi'(x) = \mathbb{O}_{m\times n}.$$

Applying the invertibility of the operator $\Phi'_y(x, \varphi(x))$ we see that

$$\varphi'(x) = -(\Phi'_y(x, \varphi(x)))^{-1}\Phi'_x(x, \varphi(x)).$$

□

Remark 10.52. *If in the conditions of the implicit function theorem $\Phi \in C^r$ the $\varphi \in C^r$.*

Remark 10.53. *If $n = m = 1$ the formula for derivative can be written in the following form*

$$\varphi'(x) = -\left.\frac{\Phi'_x(x,y)}{\Phi'_y(x,y)}\right|_{y=\varphi(x)}. \tag{34}$$

Remark 10.54. *Theorems on the inverse and implicit function guaranties existence of derivatives and for the calculation it is enough to apply the chain rule. For calculation of derivatives of the higher order one can differentiate the formula (34) or differentiate the equality $\Phi(x, \varphi(x)) = 0$ several times. Letting $y = \varphi(x)$, we see that*

$$\Phi'_x(x,y) + \Phi'_y(x,y)y' = 0,$$
$$\Phi''_{x^2}(x,y) + 2\Phi''_{xy}(x,y)y' + \Phi''_{y^2}(x,y)\left(y'\right)^2 + \Phi'_y(x,y)y'' = 0.$$

Since y' is already known from the first identity, y'' can be found from the second one. In this way, the derivatives of an implicit function are calculated sequentially. The case of function of several variables is reduced by consideration of partial derivatives.

10.9.1 Changes of the variables.

Consider a change of variables in expressions that contain partial derivatives. We will restrict ourselves to one example in which we change only the independent variables, viz. transition from Cartesian coordinates to polar coordinates. In more general situation dependent variables may also be changed. The bijectivity and preservation of smoothness are guaranteed by theorems on inverse and implicit functions.

Remark 10.55. *Let $u(x, y)$ be the function of two variables. We will assume that all derivatives in the following formulas exist and are continuous. This assumption implies independence of mixed derivatives from the order of differentiation. Let*

$$x = r\cos\theta = \tilde{x}(r,\theta), \quad u = r\sin\theta = \tilde{y}(r,\theta),$$

$$\tilde{u}(r,\theta) = u(r\cos\theta, r\sin\theta). \tag{35}$$

Invertibility of the map $(\tilde{x}, \tilde{y})$ in neighborhood of any point (r, θ), where $r > 0$ is clear from the geometric reason.

We want to express derivatives of u by x and y in terms of derivatives of $\tilde{u}$ by r and θ. Differentiating equality (35) by r and θ we see that

$$\tilde{u}'_r = \quad u'_x\widetilde{x}'_r + u'_y y'_r = u'_x\cos\theta + u'_y\sin\theta, \tag{36}$$

$$\tilde{u}'_\theta = u'_x\widetilde{x}'_\theta + u'_y\widetilde{y}'_\theta = u'_x(-r\sin\theta) + u'_y r\cos\theta. \tag{37}$$

Consider these two identities as the system of equations regarding u'_x and u'_y. Solving this system we see that

$$u'_x = \cos\theta \tilde{u}'_r - \frac{\sin\theta}{r} u'_\theta; \quad u'_y = \sin\theta \tilde{u}'_r + \frac{\cos\theta}{r} u'_r. \tag{38}$$

To calculate u''_{x^2} we can substitute u by u'_x in the previous formula

$$\begin{aligned} u''_{x^2} &= \cos\theta \left(\widehat{u'_x}\right)'_r - \frac{\sin\theta}{r} \left(\widetilde{u'_x}\right)'_\theta = \cos\theta \left(\cos\theta \tilde{u}''_r + \frac{\sin\theta}{r^2} \tilde{u}'_\theta - \frac{\sin\theta}{r} \tilde{u}''_{\theta T}\right) \\ &- \frac{\sin\theta}{r} \left(- \sin\theta \tilde{u}'_r + \cos\theta \tilde{u}''_{r\theta} - \frac{\cos\theta}{r} \tilde{u}'_\theta - \frac{\sin\theta}{r} \tilde{u}''_{\theta^2}\right) \\ &= \cos^2\theta \widetilde{u}''_{r^2} - \frac{\sin 2\theta}{r} \widetilde{u}''_{r\theta} + \frac{\sin^2\theta}{r^2} \widetilde{u}''_{\theta^2} + \frac{\sin^2\theta}{r} \widetilde{u}'_r + \frac{\sin 2\theta}{r^2} \tilde{u}''_\theta. \end{aligned} \tag{39}$$

Analogously we can deduce formulas for derivatives u''_{y^2} and u''_{xy}.

10.10 Conditional extremum.

In application we often solve the problem of finding of maximum or minimum of a function when the variable is restricted by one or several conditions given by equations.

Definition 10.56. *Let $f : D \subset \mathbb{R}^{n+m} \to \mathbb{R}$, $\Phi : D \to \mathbb{R}^m$, $p \in D$. If $\Phi(p) = \mathbb{O}_m$ and there exists V_p of a point p such that for every $x \in V_p \cap D$ that satisfies the condition $\Phi(x) = \mathbb{O}_m$, one has $f(x) \leq f(p)$ then p is a* ***point of the conditional maximum*** *of function f with restriction $\Phi(x) = \mathbb{O}_m$.*

If, in addition, for any $x \in \dot{V}_p \cap D$ satisfying the connection condition, the inequality $f(x) < f(p)$ holds, then p is called ***a point of a strict conditional maximum****. If one of the inverse inequalities* *holds then p is a point of a* ***conditional minimum*** *and a* ***strict conditional minimum****.*

If p is a point of a (strict) conditional minimum or maximum of function f then p is a ***point of a (strict) conditional extremum*** *of function f.*

Definition 10.57. ***The rank of a matrix*** *is the number of linearly independent rows of a matrix. The rank of a matrix can be determined easily by converting the matrix into row echelon form and counting the number of nonzero rows. It is also equal to the maximal order of nonzero minor.*

Let D be open, $f, \Phi \in C^{(1)}$, $\operatorname{rank} \Phi'(p) = m$. Let's find the sufficient condition for p to be a point of conditional extremum.

In matrix $\Phi'\left(x^0\right)$ there exists a nonzero minor of order m. Without loss of generality we assume that this minor is composed of derivatives by the last m variables (otherwise we can change the numeration of the variables). Denote these variable as $y_1, \dots, y_m$ and first variables by $x_1, \dots, x_n$:

$$\Phi(x^0, y^0) = \mathbb{O}_m, \quad \det \Phi'_y\left(x^0, y^0\right) \neq 0.$$

Then the conditions of implicit function theorem are satisfied. By this theorem there exist neighborhood U and V of points x^0 and y^0 in which connection condition $\Phi(x, y) = \mathbb{O}_m$ can be solved regarding y

$$y = \varphi(x), \quad \varphi \in C^1(U \to V).$$

Substituting $y = \varphi(x)$ in function f we see that x^0 is point of unconditional extremum of function

$$g(x) = f(x, \varphi(x)), \quad x \in U.$$

The necessary condition for extremality of x^0 for function g is

$$g'\left(x^0\right) = \mathbb{O}_n.$$

By the chain rule

$$f'_x(x^0, y^0) + f'_y(x^0, y^0)\varphi'(x^0) = \mathbb{O}_n \tag{40}$$

and by definition of φ

$$\Phi(x, \varphi(x)) = \mathbb{O}_m$$

Differentiating we see that

$$\Phi'_x\left(x^0, y^0\right) + \Phi'_y\left(x^0, y^0\right)\varphi'\left(x^0\right) = \mathbb{O}_{m\times n} \tag{41}$$

Consider the vector-row

$$\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m.$$

Numbers λ_i (whose number coincides with number of connection equations) are called Lagrange's multipliers. Multiplying equation (41) by vector λ from the left we see that

$$\lambda\Phi'_x\left(x^0, y^0\right) + \lambda\Phi'_y\left(x^0, y^0\right)\varphi'\left(x^0\right) = \mathbb{O}_n. \tag{42}$$

Substracting (42) from (40):

$$f'_x\left(x^0, y^0\right) - \lambda\Phi'_x\left(x^0, y^0\right) + \left(f'_y\left(x^0, y^0\right) - \lambda\Phi'_y\left(x^0, y^0\right)\right)\varphi'\left(x^0\right) = \mathbb{O}_n.$$

Now we can choose λ such that the second equation vanishes

$$f'_y\left(x^0, y^0\right) - \lambda\Phi'_y\left(x^0, y^0\right) = \mathbb{O}_m.$$

it is possible since $\Phi'_y\left(x^0, y^0\right)$ is invertible

$$\lambda = f'_y\left(x^0, y^0\right)\left(\Phi'_y\left(x^0, y^0\right)\right)^{-1}.$$

Consequently x^0, y^0 and λ satisfy the following system of equations

$$\begin{cases} f'_x\left(x^0, y^0\right) - \lambda\Phi'_x\left(x^0, y^0\right) = \mathbb{O}_n; \\ f'_y\left(x^0, y^0\right) - \lambda\Phi'_y\left(x^0, y^0\right) = \mathbb{O}_m; \\ \Phi\left(x^0, y^0\right) = \mathbb{O}_m. \end{cases}$$

Theorem 10.58 (Necessary condition for conditional extremum)**.** *Let $D \subset \mathbb{R}^{n+m}$ be open,*

$$f \in C^1(D \to \mathbb{R}), \quad \Phi \in C^1(D \to \mathbb{R}^m),$$

$p \in D$, $\operatorname{rank}\Phi'(p) = m$. *Assume that p is a point of conditional extremum of function f with respect to connection condition $\Phi(p) = \mathbb{O}_m$. Then there exists a vector $\lambda \in \mathbb{R}^m$ such that*

$$\begin{cases} f'(p) - \lambda\Phi'(p) = \mathbb{O}_{n+m} \\ \Phi(p) = \mathbb{O}_m \end{cases}$$

Notice that this system has $n + 2m$ scalar equations and the same number of variables $x_1, \ldots, x_{n+m}, \lambda_1, \ldots, \lambda_m$.

10.11 Examples.

Example 1. Maximal and minimal value of a function on a sphere. Let $\mathbb{S}^{n-1}$ be the unit sphere in $\mathbb{R}^n$, f be a quadratic form with symmetric matrix $A = (a_{ij})_{i,j=1}^n : a_{ij} = a_{ji}$ for every $i, j \in [1:n]$,

$$f(x) = \langle Ax,\ x\rangle = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

The condition of symmetry is not a restriction since every quadratic form can be generated by the symmetric matrix with elements $\frac{a_{ij}+a_{ji}}{2}$. We want to find

$$\max_{x\in\mathbb{S}^{n-1}} \langle Ax, x\rangle \quad \text{and} \quad \min_{x\in\mathbb{S}^{n-1}} \langle Ax, x\rangle.$$

Since the sphere is compact and f is continuous the maximum and minimum exist by the Weierstrass theorem. The sphere $\mathbb{S}^{n-1}$ is defined by connection equation

$$\sum_{i=1}^{n} x_i^2 - 1 \quad 0.$$

Let $\Phi(x) = \sum\limits_{i=1}^{n} x_i^2 - 1$. Since $\operatorname{grad}\Phi(x) = 2x \neq \mathbb{O}_n$ on $\mathbb{S}^{\text{n-1}}$ then $\operatorname{rank}\Phi'(x) = 1$ on $\mathbb{S}^{\text{n-1}}$. Consider the Lagrange's function

$$L(x,\lambda) = \sum_{i,j=1}^{n} a_{ij} x_i x_j - \lambda \left(\sum_{i=1}^{n} x_i^2 - 1 \right)$$

and equate it's derivatives to 0

$$F'_{x_k}(x,\lambda) = 2\left(\sum_{j=1}^{n} a_{kj} x_j - \lambda x_k \right) = 0, \quad k \in [1:n],$$

That is

$$Ax = \lambda x$$

Differentiating we used the symmetry of matrix A. thus λ is the eigenvalue of matrix A and points which are suspicious for extremality are unit eigenvectors of matrix A. Calculating the value of quadratic form at these points we see that

$$\langle Ax, x\rangle = \langle \lambda x, x\rangle = \lambda\langle x, x\rangle = \lambda$$

Theorem 10.59. *The maximal (minimal) value of the quadratic form on the unit sphere is equal to the largest (smallest) eigenvalue of the symmetric matrix of this form and is obtained on the corresponding unit eigenvector.*

From the course of higher algebra we know that eigenvalues of symmetric matrix are real.

Corollary 10.59.1 (Norm of the linear operator in Euclidean spaces)**.** *Let* $A \in \mathcal{L}(\mathbb{R}^n \to \mathbb{R}^m)$*. Then*

$$\|A\| = \max\left\{\sqrt{\lambda} : \lambda \textit{ is an eigenvalue of the matrix } A^T A\right\}.$$

Proof. By the theorem on the calculation of the norm of the linear operator

$$\|A\|^2 = \max_{|x|=1} |Ax|^2 = \max_{|x|=1} \langle Ax, Ax\rangle = \max_{|x|=1} \left\langle A^T Ax, x\right\rangle.$$

Matrix $A^T A$ is symmetric, consequently we can apply Theorem 10.59. □

Example 2. Distance from the point to the hyper-plane Let $a \in \mathbb{R}^n \setminus \{\mathbb{O}_n\}$, $b \in \mathbb{R}$. A **hyperplane** is a set

$$L = \left\{x \in \mathbb{R}^n : \sum_{i=1}^{n} a_i x_i + b = 0\right\}.$$

Let $c \in \mathbb{R}^n$ and set the problem to find a distance from the point c to hyperplane L, that is

$$d(c, L) = \min_{x \in L} \|x - c\|,$$

and the point at which this minimum is obtained. It is clear that minimum of function $x \to \|x - c\|$ exists or doesn't exist at the same time with the function

$$f(x) = \|x - c\|^2 = \sum_{i=1}^{n} (x_i - c_i)^2$$

and is obtained in the same time. Hyper-plane L is defined by the connection condition

$$\sum_{i=1}^{n} a_i x_i + b = 0.$$

Denote $\Phi(x) = \sum_{i=1}^{n} a_i x_i + b$. Since $\operatorname{grad} \Phi(x) = a \neq \mathbb{O}_n$, $\operatorname{rank} \Phi'(x) = 1$ on $\mathbb{R}^n$. Compose the Lagrange's function

$$L(x, \lambda) = \sum_{i=1}^{n} (x_i - c_i)^2 - \lambda \left(\sum_{i=1}^{n} a_i x_i + b \right)$$

and equate its derivatives to zero

$$L'_{x_k}(x, \lambda) = 2(x_k - \lambda_k) - \lambda a_k = 0, \quad k = 1, \ldots, n.$$

To solve this system we multiply the k-th equation by a_k and sum by k from 1 to n :

$$2 \sum_{k=1}^{n} a_k x_k - 2 \sum_{k=1}^{n} a_k c_k - \lambda \sum_{k=1}^{n} a_k^2 = 0.$$

By the connection condition the first sum is equal to $-2b$. Consequently

$$\frac{\lambda}{2} = -\frac{\sum\limits_{k=1}^{n} a_k c_k + b}{\sum\limits_{k=1}^{n} a_k^2}.$$

Applying this expression to derivatives of function F we see that

$$x_k = c_k + \frac{\lambda}{2} a_k = c_k - \frac{\sum\limits_{i=1}^{n} a_i c_i + b}{\sum\limits_{i=1}^{n} a_i^2} a_k.$$

Denoting this point by x^0 and the value $f\left(x^0\right)$ by M:

$$M = f\left(x^0\right) = \sum_{k=1}^{n} \left(x_k^0 - c_k\right)^2 = \frac{\left(\sum\limits_{i=1}^{n} a_i c_i + b\right)^2}{\sum\limits_{i=1}^{n} a_i^2}$$

We will prove that the function f obtains at x^0 the value minimal on L. A function $f(x) \underset{\|x\| \to +\infty}{\longrightarrow} +\infty$ since $\|x - c\|^2 \geqslant (\|x\| - \|c\|)^2 \underset{\|x\| \to +\infty}{\longrightarrow} +\infty$. Consequently, there exists $R > 0$ such that $|f(x)| > M$ for every $x \in \mathbb{R}^n$ such that $\|x\| \geqslant R$. Consider the ball $B = B\left(\mathbb{O}_n, R\right)$. it is clear that $x^0 \in B$. The set $\bar{B} \cap L$ is compact since it is closed and bounded. By the Weierstrass theorem the minimum $\min\limits_{x \in \bar{B} \cap L} f(x)$ exists. Minimum is obtained at some point x^* and is less or equal that M since $f\left(x^0\right) = M$. Moreover, x^* belongs to the open ball B and

$$\min_{x \in L} f(x) = \min_{x \in \bar{B} \cap L} f(x)$$

since $f(x) > M$ in the complement of B. Consequently, x^* satisfies the system whose unique solution is x^0. Thus, $x^* = x^0$, that is

$$\min_{x \in L} f(x) = f\left(x^0\right).$$

Concluding the previous reasoning we see that

$$d(c, L) = \frac{\left|\sum_{i=1}^{n} a_i c_i + b\right|}{\sqrt{\sum_{i=1}^{n} a_i^2}}.$$