Elements of Functional Analysis

May 11, 2006

it is a draft of Lecture Notes of H.M. Khudaverdian.
Manchester, Spring 2006

Contents

1 Weierstraß and Stone—Weierstraß Approximations Theorems

Preliminary notes

Consider a metric space X. (A metric $\rho(x, y)$ defines a topology on X.

One can consider the space $B(X)$ of bounded real-valued (or complex) valued functions on X.

An arbitrary real (complex) valued function f is bounded (it belongs to the space $B(X)$) if for the function f there exists $M > 0$ such that $|f(x)| < M$ for every point x in the space X. (As always $|a|$ is absolute value of real number or module of complex number).

One can define a metric on the space $B(X)$:

$$d(f, g) = \sup_{x \in X} |f(x) - g(x)| \tag{1}$$

One can see that metric is well-defined by this expression, i.e.

1) $d(f, g) \geq 0$ and $d(f, g) = 0$ iff $f \equiv g$

2) $d(f, g) + d(g, h) \geq d(f, h)$ (triangle inequality)

Exercise Prove it.

Now consider the case when X is a compact space [1].

In this case every continuous function is bounded on X and attaints its supremum and inferum .

The space $C(X)$ of continuous functions is subspace of $B(X)$.

Two words about uniform convergence:

f_n tends uniformly to f ($f_n \overset{U}{\to} f$) if for every $\varepsilon > 0$ there exists N such that **for every** x $|f_n(x) - f(x)| < \varepsilon$, i.e. closedness is irrelevant to x. It is easy to see that $f_n \overset{U}{\to} f$ means that f_n tends to f in the metric $d(f, g)$.

Counterexample

Consider $f_n = x^n$ on $[0, 1]$ and discontinuous function $f = 0$ if $0 \leq x < 1$, $f(1) = 1$. Prove that for every x the sequence $\{f_n\}$ tends to f, but this sequence does not tend uniformly to f.

[1] The topological space is compact if every covering by open sets contains finite subcovering. Every infinite set in compact space X has a condensation point. This condition is equivalent to compactness condition for metric space. The compact subspace of metric space is closed and bounded. Closed and bounded subspaces of **finite**-dimensional euclidean space are compact.(Condition of finite-dimensionality is essential!) In particular a cube in $\mathbf{R}^n$ is compact (Heine-Borel lemma)

1.1 Weierstraß Approximation Theorem: Approximation of continuous function by polynomials

Consider first very special case when $X = [a, b]$ is a compact interval of $\mathbf{R}$ and $C(X)$ be a space of continuous functions on $[a, b]$.

Practically it is convenient to work with polynomials. Weierstraß Approximation Theorem states:

Every continuous function on the closed interval $[a, b]$ *can be uniformly approximated by polynomials,i.e.for every continuous function* f *on the closed interval* $[a, b]$ *there exists a sequence* $\{P_n(x)\}$ *of polynomials such that* $P_n(x) \stackrel{U}{\to} f$

In detail:

For every continuous function f on the closed interval $[a, b]$ there exists a sequence $\{P_n(x)\}$ of polynomials such that

$$\forall \varepsilon \quad \exists N\colon \quad \forall n > N \text{ and } \forall x \in [a, b] \ |f(x) - P_n(x)| < \varepsilon \tag{2}$$

One can reformulate it in the following way: $\forall \varepsilon > 0$ *there exists polynomial* P *such that* $d(P, f) < \varepsilon$, where $d(f, g) = \sup |f(x) - g(x)|$

Or:

The space of polynomials is dense in the space of continuous functions on the closed interval $[a, b]$ *in the topology induced by the metric* d.

All these formulations are equivalent.

To prove Theorem we construct explicitly for every continuous function f the sequence of polynomials $P_n^{(f)}(x)$ such that this sequence tends uniformly to f. We mention that there are many different sequences of polynomials tending to the given function.

Let f be an arbitrary continuous function on $[a, b]$. Without loss of generality we can assume [2] that $[a, b] = [0, 1]$ and consider the following polynomials

$$B_n^{(f)}(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) C_n^k x^k (1 - x)^{n-k}\,, \tag{3}$$

where $C_n^k = \frac{n!}{k!(n-k)!}$ sometimes it is denoted by $\binom{n}{k}$

These polynomials are called *Bernstein-Weierstraß polynomials.* n-th Bernstein-Weierstraß polynomial has an order n. It is defined by $n + 1$

[2] see in more detail the beginnign of subsection 1.3

values of function f. [3].

We will prove in next subsections that the sequence $\{B_n^{(f)}(x)\}$ tends uniformly to the function f for an every function f which is continuous on the closed interval $[0,1]$

Try to calculate explicitly these polynomials in the case if f is a simple polynomial:

It is evident that

$$B_n^{(f=1)}(x) = \sum_{k=0} C_n^k x^k (1-x)^{n-k} = (x+(1-x))^n = 1\,, \tag{5}$$

$$B_n^{(f=x)}(x) = \sum_{k=0}^{n} \left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k} = \frac{1}{n}\sum_{k=0}^{n} k C_n^k x^k (1-x)^{n-k} = \tag{6}$$

$$\sum_{k=0}^{n} C_n^{k-1} x^k (1-x)^{n-k} = x\sum_{k=1}^{n} C_n^{k-1} x^{k-1} (1-x)^{n-k} = x\,. \tag{7}$$

and finally for $f = x^2$

$$B_n^{(f=x^2)}(x) = \sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 C_n^k x^k (1-x)^{n-k} = x^2 + \frac{1}{n}(x-x^2) \tag{8}$$

because

$$\left(\frac{k}{n}\right)^2 C_n^k = \frac{k(k-1)+k}{n^2} C_n^k = \frac{k(k-1)}{n^2}\frac{n!}{k!(n-k)!} + \frac{k}{n^2}\frac{n!}{k!(n-k)!} =$$
$$\frac{n-1}{n}\frac{(n-2)!}{(k-2)!(n-k)!} + \frac{(n-1)!}{n(k-1)!(n-k)!} = \frac{n-1}{n}C_{n-2}^{k-2} + \frac{1}{n}C_{n-1}^{k-1}$$

Hence

$$B_n^{(f=x^2)}(x) = \frac{x^2(n-1)}{n}\sum_k C_{n-2}^{k-2}x^{k-2}(1-x)^{n-k} + \frac{x}{n}\sum_k C_{n-1}^{k-1}x^k(1-x)^{n-k} = \frac{x^2(n-1)}{n} + \frac{x}{n}$$

One can analogously perform explicit calculations for arbitrary polynomial $f = x^r$. But in the general case straightforward integration does not lead to an answer. We need an independent proof.

[3]Naively it seems more appropriate instead polynomial above to consider as $P_n(x)$ n-th degree polynomial which coincides with function f in these $n+1$ points, so called Lagrange interpolation polynomials

$$P_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\frac{Q_n(x)}{Q'\left(\frac{k}{n}\right)\left(x-\frac{k}{n}\right)} \tag{4}$$

. But these polynomials *do not tend uniformly to f!?*

1.2 Heuristic arguments leading to Bernstein polynomials

The construction of polynomials (3) belongs to mathematicien S.Bernstein (1880—1968) who was specialist in probability theory.

Consider heuristic arguments based on probability theory which lead to the construction (3).

Consider a false coin such that if we toss it then we see "head" with probability x and "tail" with probability $1-x$. More formally we say that we consider random variable ξ such that it takes value 1 with probability x and value 0 with probability $1-x$.

Now toss this coin n times. One can see *the probability that exactly k times we have "head" is equal to*

$$p_k(x) = C_n^k x^k (1-x)^{n-k} \,.$$

On the other hand heuristically it is evident that if n tends to infinity then a sequence $\{p_k\}$ is concentrated around $x \approx k/n$: $p_k \approx 0$ if $k \not\approx xn$, i.e. for arbitrary continuous function f

$$f(x) = f(x)\sum_{k=0}^{\infty} p_k \approx \sum_{k\colon |k/N-x|\approx 0} f(x)p_k \approx \sum_{k\colon |k/N-x|\approx 0} f\left(\frac{k}{n}\right) p_k \approx \sum_{k=0}^{\infty} f\left(\frac{k}{n}\right) p_k$$

We see that the left hand of this expression is nothing but Bernstein-Weierstraß polynomial for f. This gives us heuristic motivation that a sequence $\{B_n^f\}$ of Bernstein-Weierstraß polynomials tends to continuous function f. [4]

1.3 Rigid proof

First of all show that it is enough to prove the Theorem for continuous functions on $[0,1]$.

[4]More formally consider random variable

$$\xi = \frac{\xi_1 + \ldots \xi_n}{n}$$

where all ξ_k are random variables such that it takes value 1 with probability x and value 0 with probability $1-x$.Then for continuous function $f(x)$ the mean value $< f(\xi) >$ of random variable $f(\xi)$ is just $n-th$ Bernstein-Weierstraß polynomial (3). On the other hand it is equal approximately to $f(x)$.

Indeed suppose f is continuous on the closed interval $[a, b]$. Consider a transformation

$$x \mapsto (b-a)x + a \tag{9}$$

This is a linear one-one transformation of $[0, 1]$ onto $[a, b]$. Hence for every $f \in C([a, b])$ consider a function

$$\tilde{f}(x) = f\left((b-a)x + a\right) \tag{10}$$

It is evident that the function $\tilde{f}$ is continuous on $[0, 1]$ as well as f is continuous on $[a, b]$. Consider a polynomial $\tilde{P}(x) = \sum a_n x^n$ such that it approximates the function f up to ε:

$$\forall x \in [0,1] \quad |\tilde{f}(x) - \tilde{P}(x)| < \varepsilon, \quad \text{i.e. } d(\tilde{f}, \tilde{P}) < \varepsilon,$$

(We suppose that Approximation Theoremisa is already proved or $C([0, 1])$) Denote by $y = (b-a)x + a$. Then $x = \frac{y-a}{b-a}$ We have:

$$|\tilde{f}(x) - \tilde{P}(x)| = |f\left((b-a)x + a\right) - \tilde{P}(x)| = |f(y) - \tilde{P}\left(\frac{y-b}{b-a}\right)| < \varepsilon \tag{11}$$

Thus we come to the polynomial $P(y) = \tilde{P}(\frac{y-a}{b-a})$ such that it approximates a function f up to ε: $d(f, P) < \varepsilon$ ■

So it remains to prove Weierstraß Approximation theorem for $[0, 1]$.

Proof Consider arbitrary $\varepsilon > 0$. The function f is continuous on closed interval $[0, 1]$. Hence it is bounded and uniformly continuous; i.e. there exists δ such that

$$\forall x, y \in [0,1]\colon \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon\,. \tag{12}$$

and there exists M such that

$$\forall x \in [0,1], \quad |f(x)| < M\,. \tag{13}$$

Now estimate the dependance on n, M, ε and δ of distance between function f and polynomial $B_n(x)$:

$$d(B_n^{(f)}, f) = \max_x |B_n^{(f)}(x) - f(x)|\,.$$

Estimate $|B_n^{(f)}(x) - f(x)|$ for arbitrary $x \in [0, 1]$.

Note that according to (5)

$$f(x) = \sum_{k=0}^{n} f(x) C_n^k x^k (1-x)^{n-k} .$$

Then

$$B_n^{(f)}(x) - f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k} - f(x) =$$

$$\sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right) \right) C_n^k x^k (1-x)^{n-k} =$$

$$\sum_{k:\ |k/n-x|<\delta}^{n} \left(f(x) - f\left(\frac{k}{n}\right) \right) C_n^k x^k (1-x)^{n-k} +$$

$$\sum_{k:\ |k/n-x|\geq\delta}^{n} \left(f(x) - f\left(\frac{k}{n}\right) \right) C_n^k x^k (1-x)^{n-k} .$$

Denote the first term in the last expression by $I_1(x, \varepsilon)$ and second one by $I_2(x, \varepsilon)$. According triangle inequality we have that

$$|f(x) - B_n| \leq |I_1(x, \varepsilon)| + |I_2(x, \varepsilon)| .$$

Thus estimate $|I_1(x, \varepsilon)|$ and $|I_2(x, \varepsilon)|$.

Using the fact that the continuous function f is uniformly continuous on the closed interval $[0, 1]$ (see (12) we see that

$$|I_1(x, \varepsilon)| < \sum_{k:\ |k/n-x|<\delta}^{n} \varepsilon C_n^k x^k (1-x)^{n-k} \leq \varepsilon \sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} = \varepsilon .$$

(here all terms are positive we omit symbols "| |") Now using the fact that the continuous function f is bounded on the closed interval $[0, 1]$ (see (139)) we see that

$$|f(x) - f(k/n)| < 2M .$$

Hence

$$|I_2(x, \varepsilon)| < 2M \sum_{k:\ |k/n-x|\geq\delta}^{n} C_n^k x^k (1-x)^{n-k} .$$

(here all terms are positive and we omit symbols "| |")

Now we use very simple but very effective trick[5]:

$$\sum_{k:\ |k/n-x|\geq\delta}^{n} C_n^k x^k (1-x)^{n-k} \leq \frac{1}{\delta^2} \sum_{k:\ |k/n-x|\geq\delta}^{n} \left(\frac{k}{n}-x\right)^2 C_n^k x^k (1-x)^{n-k} \leq$$

$$\frac{1}{\delta^2} \sum_{k=0}^{n} \left(\frac{k}{n}-x\right)^2 C_n^k x^k (1-x)^{n-k} .$$

On the other hand according to (5), (6) and (8)

$$\sum_{k=0}^{n} \left(\frac{k}{n}-x\right)^2 C_n^k x^k (1-x)^{n-k} = \sum_{k=0}^{n} \left(\frac{k}{n}^2 - 2x\frac{k}{n} + x^2\right) C_n^k x^k (1-x)^{n-k} =$$

$$B_n^{f=x^2} - 2xB_n^{f=x} + x^2 B_n^{f=1} =$$

$$x^2 + \frac{1}{n}(x-x^2) - 2x^2 + x^2 = \frac{1}{n}(x-x^2)$$

And we finally come to

$$I_2(x,\varepsilon) \leq \frac{2M}{n\delta^2}|x-x^2|$$

Collecting all results above we see that

$$|f(x) - B_n(x)| \leq \varepsilon + \frac{2M}{n\delta^2}$$

Hence it is easy to see that if we choose

$$N > \frac{2M}{\varepsilon\delta^2}$$

then for every $n > N$ and for every x $|f(x) - B_n(x)| \leq 2\varepsilon$, i.e. $d|f, B_n| < 2\varepsilon$

■

At what extent are important the condition that interval is closed?

Exercise.

Consider function $f = 1/x$ which is continuous on the *open* interval $(0, 1)$.

1. Prove that it is not possible to find a sequence of polynomials $\{P_n\}$ such that $\{P_n\}$ tends uniformly to $f = \frac{1}{x}$

2. Check at what step it fails the procedure of constructing of Bernstein-Weierstraß polynomials for this case.

[5]it is nothing but famous Tchebychev inequality

1.4 Stone-Weierstraß Approximation Theorem

Note that space of continuous functions is vector spaces, moreover it is an algebra w.r.t. point-wise multiplication. The subspace of polynomials in the space of continuous functions is

1) subalgebra ($f, g \in A \Rightarrow$, $af + bg, fg \in A$ for every $a, b \in A$)

2) it contains a unity (the function $f \equiv 1$)

3) this algebra separates points of interval $[0, 1]$

We can formulate more general result.

Stone-Weierstraß Approximation Theorem

Let $C(X)$ be a metric space of continuous functions on the **compact** *metric space X. (Metrics on $C(X)$ is defined as usual by* (1)*) Let A be a subspace of $C(X)$ such that*

1) A is subalgebra of $C(X)$ (i.e. A is closed under operations of addition, multiplication)

2) this subalgebra contains unity

3) this subalgebra separates points of X, i.e. for every points a, b $a \neq b$ there exists a function $f \in A$ such that $f(a) \neq f(b)$.

Then A is dense in $C(X)$.

Comment little bit all these conditions:

1) if X is not compact then one can consider continuous functions which goes to infinity on borders. These functions have little chance to be approximated (Consider the following counterexample: $X = (0, 1)$. Then function $f = 1/x$ has no chance to be approximated by polynomials. (Why?)

The condition that A have to be subalgebra is very important. To see its importance consider the following counterexample: consider on compact space $[0, 1]$ a space of polynomials $a + bx$ where a, b are arbitrary real coefficients. This is a vector space but not algebra (Why?). All conditions but the condition 1) are obeyed. It is easy to see that a function $y = x^2$ cannot be approximated by these polynomials.

To see the importance of condition that A contains unit element consider algebra of all polynomials of degree ≥ 1 (generated by polynomials $\{x, x^2, x^3 \ldots\}$) on the closed interval $[0, 1]$. All conditions but the condition 2) are obeyed. One can easy see that the function $f = 1$ cannot be approximated by functions from this subalgebra.

Now consider the importance of the last condition, that A separates points. Intuitively this is most important condition. Suppose it is not the case. Then there exist two points a, b $a \neq b$ such that all functions from A take the same value at these points. Consider continuous function f such that it takes *different* values at the points a, b. Evidently this function cannot be approximated by functions from the algebra A. (E.g. consider the set A of even polynomials on the interval $[-1, 1]$. We see that all conditions but 3-rd are obeyed. But the function $f = x$ cannot be approximated by these polynomials.)

Example of application of Stone-Weierstraß Theorem

Consider the set A of all trigonometric polynomials on $\mathbf{R}$, i.e., the set of functions of the form

$$f(x) = \sum_{n=0}^{N} (a_n \cos nx + b_n \sin nx) \,,$$

where N is an arbitrary non-negative integer and a_n, b_n are arbitrary real numbers.

One can easy to prove that

1) A is an algebra (Check straightforwardly using trigonometric identities).

2) This algebra contains unity ($1 = \cos nx$ if $n = 0$),

3)this algebra separates points of the interval $[0, 2\pi)$. (To prove this it is enough to consider two functions $(\cos x, \sin x)$)

Now try to apply Stone-WeierstraßTheorem. But $[0, 2\pi)$ is not compact. Not so good. Try to consider $[0, 2\pi]$. Another problem: A does not separate points of $[0, 2\pi]$.

But one can look on the case from little bit different angle:

Consider continuous function on the circle S^1 ($x = \varphi$ angle, $0 \leq \varphi < 2\pi$), i.e. continuous *periodic* functions ($f(\varphi + 2\pi) = f(\varphi)$). S^1 is compact. Conditions 1) 2) 3) means that we can apply Stone-WeierstraßTheorem. Thus

Algebra A is dense in the space of continuous functions on S^1, or in other words every continuous periodic function on S^1 can be uniformly approximated by a trigonometric polynomial.

This statement is equivalent to the statement:

Every continuous periodic function on $\mathbf{R}$ with period 2π can be uniformly approximated by a trigonometric polynomial.

1.5 Proof of Stone-Weierstraß Theorem

Let A be an algebra of functions in $C(X)$ obeying conditions that this algebra possesses unity and it separates the points of X.

Consider a set $\bar{A}$—a closure of algebra A. ($f \in \bar{A}$ if there is at least one point of A in any neighborhood of f, i.e. for every $\varepsilon > 0$ the ball $B_\varepsilon(f)$ contains at least one point from A)

Prove that $\bar{A} = C(X)$, i.e. A is dense in $C(X)$.

Lemma 1. For every pair of distinct points $x_1, x_2 \in X$ and for every two numbers y_1, y_2 there exists a function $f \in A$ such that

$$f(x_1) = y_1, f(x_2) = y_2 \tag{14}$$

Lemma 2 The function $|f|$ belongs to $\bar{A}$ if $f \in \bar{A}$

Lemma 3 If functions $f_1, \dots, f_n$ belong to $\bar{A}$ then functions $F = \min\{f_1, \dots, f_n\}$ and $G = \max\{f_1, \dots, f_n\}$ also belong to $\bar{A}$.

Lemma 4 for every function $f \in C(X)$, for every $x \in X$ and for every $\varepsilon > 0$ there exist a function h
$barA$ such that

$$h(t)\big|_{t=x} = f(x) \,\text{and}\quad h(t) > f(t) - \varepsilon \tag{15}$$

Proof of the Lemma 1. The algebra A separates points. Hence there exists a function $g \in A$ such that $g(x_1) \neq g(x_2)$. A is an algebra with unity. Hence a function

$$f(x) = y_1 \frac{g(x) - g(x_2)}{g(x_1) - g(x_2)} + y_2 \frac{g(x) - g(x_1)}{g(x_2) - g(x_1)} \tag{16}$$

is well-defined, it belongs to A and $f(x_1) = y_1, f(x_2) = y_2$

Proof of the Lemma 2. Suppose $f \in A$. f is bounded (because X is compact.) Suppose $|f| < M$. Using Weierstraß Approximation Theorem for the closed interval $[-M, M]$ consider for an arbitrary ε a polynomial $P(x) = a_n x^n$ such that it is approximation of the function $|x|$ on the interval $[-M, M]$ up to ε:

$$\forall x \in [-M, M], \quad ||x| - P(x)| < \varepsilon \tag{17}$$

Hence we see that

$$\forall x \in [-M, M], \quad ||f| - \sum a_n f^n| < \varepsilon \tag{18}$$

where function $\sum a_n f^n$ belongs to the algebra A.

Hence we prove that for an arbitrary $\varepsilon > 0$ there exists a function (it is a function $\sum a_n f^n$) in the algebra A such that it is on the distance not far than ε. Hence $|f|$ belongs to the closure of A: $f \in \bar{A}$ ■

Proof of the Lemma 3. The proof follows from the following obvious relations:

$$\min\{f, g\} = \frac{f+g}{2} - \frac{|f-g|}{2}, \quad \max\{f, g\} = \frac{f+g}{2} + \frac{|f-g|}{2}$$

and from the previous lemma ■

Proof of the Lemma 4. Fix a point $x \in X$.
Now consider for every point $y \in X$ a function $H_y(t)$ in A such that a function $H_y(t)$ coincides with a function $f(t)$ in two points x and y. (There exists such function in algebra A according to the Lemma 1). Then because of continuouity a $\delta(x)$ such that $H_y(t) > f(t) - \varepsilon$ for all $t \in B_\delta(y)$. X is compact. Hence there exist finite numbers of points $\{y_1, \ldots, y_n\}$ and $\{\delta_1, \ldots, \delta_n\}$ such that balls $\{B_{\delta_k}\}$ cover X. The function

$$F = \max\{H_{y_1}, \ldots, H_{y_n}\}$$

belongs to the algebra $\bar{A}$ (according to the Lemma 3). It is equal to f at the point x because all functions $\{H_1, \ldots, H_k\}$ are equal to $f(x)$ at the point x. It obeys condition $F(x) > f(x) - \varepsilon$ because at least one of the functions from the set $\{H_1, \ldots, H_k\}$ obeys this condition. ■

Now we are ready to prove Stone-Weierstraß Theorem.

Consider arbitrary continuous function f on compact metric space X. Fix $\varepsilon > 0$. According to the Lemma 4 consider for every point x a function $G_x(t)$ such that $G_x(t) > f(x) - \varepsilon$. $G_x(t)\big|_{t=x} = f(x)$. Hence there exists $\delta(x)$ such that $|f - G_x| < \varepsilon$ for every t belonging to the ball $B_\delta(x)$. X is compact. Hence there exist finite numbers of points $\{x_1, \ldots, x_k\}$ and $\{\delta_1, \ldots, \delta_k\}$ such that balls $\{B_{\delta_k}\}(x_k)$ cover X. The function

$$F = \min\{G_{x_1}, \ldots, G_{x_k}\}$$

belongs to the algebra $\bar{A}$ (according to the Lemma 3) and it obeys the condition $|f - F| < \varepsilon$ for all $x \in X$ ■

2 Normed spaces

2.1 Recalling vector spaces

Basic definitions of vector spaces. Special attention on the conception of linear independence and definition of infinite-dimensional vector space.

Vectors $\mathbf{x_1}, \ldots, \mathbf{x_m}$ in the vector space V are called linear dependent if there exist coefficients $\lambda_1, \ldots, \lambda_m$ such that *not all coefficients are equal to zero* and linear combination $\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \cdots + \lambda_m\mathbf{x}_m$ is equal to zero.

Or equivalently: Vectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$ in the vector space V are called linear dependent if one of these vectors can be expressed as linear combination of other vectors:

$$\exists m_0\colon \quad 1 \leq m_0 \leq m,\ \mathbf{x}_{m_0} = \sum_{i \neq m_0} \lambda_i \mathbf{x_i}$$

If vector space V is *finite-dimensional* then there exists a number N such that the following condition is satisfied:

Vectors $\mathbf{x}_1, \ldots, \mathbf{x_m}$ in the vector space V are called linear *independent* if these vectors are not linear dependent.

Dimension of vector space. Vector space has dimension N if there exist N *linear independent* vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ and an arbitrary set of $N+1$ vectors $\{\mathbf{v}_1, \ldots \mathbf{v}_N, \mathbf{v}_{N+1}\}$ in V is the set of *linear dependent* vectors.

A set $\{\mathbf{e}_i\}$ of N linear independent vectors in N-dimensional space is a *basis* of V: for every vector $\mathbf{x} \in \mathbf{V}$, $\mathbf{x} = \sum_{i=1}^N x^i \mathbf{e}_i$, where components x^i are defined uniquely.

Definition Vector space is called infinite-dimensional if for every N there exist N linear independent vectors.

In other words infinite-dimensional vector space for every N possesses N-dimensional finite vector subspace.

Examples of infinite dimensional vector spaces:

1) Space of polynomials (Check that polynomials $\{1, x, x^2, \ldots, x^N\}$) are linear independent for an arbitrary N

2) Space of functions.

3) Consider the vector space l of all infinite sequences. "Vectors" of l are sequences $\mathbf{x} = \{x_n\}$ $(n = 1, 2, 3, \ldots)$ of real numbers, $x_n \in \mathbf{R}$. The vector space operations can be introduced naturally: If $\mathbf{x} = \{x_n\}$, $\mathbf{y} = \{y_n\}$ are two vectors then the sequence $\mathbf{z} = \mathbf{x} + \mathbf{y}$ is defined as $\mathbf{z} = \{z_n\}$: $z_n = x_n + y_n$, the sequence $\lambda\mathbf{x}$ is defined as $\lambda\mathbf{x} = \{\lambda x_1, \lambda x_2, \ldots, \}$

The space l contains N-dimensional spaces $\mathbf{R}^N$ for every N (R^N is a space of sequences $\{x_1, x_2, \ldots, x_N\}$) containing N elements.

This infinite-dimensional space contains following very important infinite-dimensional subspaces:

a) subspace l_{fin} of finite sequences, i.e. for every sequence $\mathbf{x} = \{x_n\} \in l_{\text{fin}}$ there exists N such that for every $n > N$

$$x_n = 0\,, \tag{19}$$

Remark Every element $\mathbf{x} = \{x_1, x_2, \ldots, x_n, 0, 0, 0, \ldots\}$ of l_{fin} can be considered as an element in some $\mathbf{R}^n$ but these $\mathbf{R}^n$ can have arbitrary large dimension.

b) subspace l_1 of sequences $\mathbf{x} = \{x_n\}$ such that

$$\sum_n |x_n| < \infty\,, \tag{20}$$

c) subspace l_2 of sequences $\mathbf{x} = \{x_n\}$ such that

$$\sum_n |x_n|^2 < \infty\,, \tag{21}$$

d) subspace l_∞ of sequences $\mathbf{x} = \{x_n\}$ such that

$$\sup_n \{|x_n|\} < \infty\,. \tag{22}$$

e) One can consider in general for every $p > 0$ vector space l_p as a space of sequences $\mathbf{x} = \{x_n\}$ such that

$$\sum_n |x_n|^p < \infty\,, \tag{23}$$

$l_p = l_1$ if $p = 1$, $l_p = l_2$ if $p = 2$. It is convenient to consider $l_p = l_\infty$ for $p = \infty$.

One can see that l_1, l_2, l_∞ (and in general all l_p) are vector subspaces of l. To show it we have to check that for arbitrary $\mathbf{x}, \mathbf{y} \in l_p$ $\mathbf{x} + \mathbf{y} \in l_p$. E.g. consider l_2. If $\mathbf{x}, \mathbf{y} \in l_2$, then $\sum_{n=1}^\infty |x_n|^2 < \infty, \sum_{n=1}^\infty |y_n|^2 < \infty$. Consider inequality:

$$(x_n + y_n)^2 = 2x_n^2 + 2y_n^2 - (x_n - y_n)^2 \leq 2x_n^2 + 2y_n^2.$$

Due to this inequality for arbitrary N

$$\sum_{n=1}^N (x_n + y_n)^2 \leq 2\sum_{n=1}^N x_n^2 + 2\sum_{n=1}^N y_n^2 < \infty.$$

Hence taking limit $N \to \infty$ we arrive at $\sum_{n=1}^\infty |x_n + y_n|^2 < \infty$.

In a general case if $\mathbf{x}, \mathbf{y} \in l_p$, then $\sum_{n=1}^\infty |x_n|^p < \infty$, $\sum_{n=1}^\infty |y_n|^p < \infty$ then the condition $\sum_{n=1}^\infty |x_n + y_n|^p < \infty$, is satisfied due to the inequality:

$$|x_n + y_n|^p \leq \left(2\left(\max\{|x_n|, |y_n|\}\right)\right)^p \leq 2^p\left(|x_n|^p + |y_n|^p\right)$$

The condition that l_p is vector space for $p = \infty$ has to be checked independently: If $\mathbf{x} = \{x_1, \dots, x_n, \dots\}$ belongs to l_∞ then the sequence $\{|x_1|, \dots, |x_n|, \dots\}$ is bounded, i.e. there exists M_1 such that for every k, $|x_k| < M_1$. Analogously if $\mathbf{y} = \{x_1, \dots, x_n, \dots\} \in l_\infty$ then there exists M_2 such that for every

k', $|x_{k'}| < M_2$. Hence for every k, $|x_k + y_k| \leq M_1 + M_2$. Thus sequence $\{|x_1 + y_1|, \ldots, |x_n + y_n|, \ldots\}$ is bounded. Hence $\mathbf{x} + \mathbf{y} \in l_\infty$.

Remark Later we consider l_p only for $p \geq 1$. In this case one can define a norm on these spaces (see later).

Intuitively we feel that in examples above are considered very different spaces in spite of the fact that all the spaces have the same dimension $= \infty$.

For finite-dimensional vector spaces dimension in some sense completely characterizes vector space. (Two finite-dimensional vector spaces are isomorphic iff their dimensions coincide)

For infinite-dimensional vector spaces situation is different.

There are "different infinities" The space of all polynomials, the set of all continuous functions, the set of all bounded functions —these spaces are very different.

The same is with spaces l_p.

Try to understand little bit more rigorously what happened.

Theorem. *The vector space l_1 is a proper subspace of l_2, l_2 is a proper subspace of l_2 and l_2 is a proper subspace of l_∞, i.e.*

$$l_1 \subset l_2 \subset l_\infty \,, \text{but } \; l_1 \neq l_2, \; , \text{and } \; l_2 \neq l_\infty \,. \tag{24}$$

More generally for every p, q such that $1 \leq p < q \leq \infty$

$$l_p \subset l_q \subset l_\infty \,, l_p \neq l_q, \text{if } \; p \neq q \,. \tag{25}$$

Proof of Theorem.

Prove that $l_1 \subset l_2$. In general the idea of the proof is the same for all embeddings $l_p \subset l_q$.

First proof

Suppose $\mathbf{x} = \{x_k\} \in l_1$. Then $\sum_{k=0}^\infty |x_k| < \infty$. To prove that $\sum |x_k|^2 < \infty$ do the following steps:

1) $\sum |x_k| < \infty \Rightarrow$ there exist N such that for all $k \geq n$ $|x_k| < 1$. Hence $x_k^2 < |x_k|$ if k is enough big ($k \geq N$). Hence

$$\sum_{k=1}^\infty x_k^2 = \sum_{k=1}^N x_k^2 + \sum_{k=N+1}^\infty x_k^2 < \sum_{k=1}^N x_k^2 + \sum_{k=N+1}^\infty |x_k| < \sum_{k=1}^N x_k^2 + \sum_{k=1}^\infty |x_k| < \infty$$

One can prove analogously that

$$l_p \subseteq l_q \subseteq l_\infty\,, \text{ if } \quad p < q. \tag{26}$$

Another proof

The embedding $l_1 \subseteq l_2$ follows in a very simple way from the following inequality: If $\sum_{n=1}^N |x_n| < M$ then $\left(\sum_{n=1}^N |x_n|\right)^2 < M$. Hence $\sum_{n=1}^N x_n^2 < M^2$. Thus $l_1 \subseteq l_2$.

To prove that $l_p \neq l_q$ if $p \neq q$ it is enough to consider counterexample. There are plenty counterexamples. E.g. for $p = 1, q = 2$ consider the harmonic sequence $\mathbf{x} = \{x_k\}$ where

$$x_k = \frac{1}{k} \tag{27}$$

It is easy to see considering integrals $\int_1^N \frac{dx}{x^\alpha}$ that $\mathbf{x} \in l_2$ and $\mathbf{x} \notin l_1$ if $\frac{1}{2} < \alpha \leq 1$ (see in more detail the Example in the end of the following subsection). To prove that $l_2 \neq l_\infty$ one can consider sequence $\mathbf{x}$ such that all its terms are equal to c, where $c \neq 0$. This sequence belongs to l_∞ and it does not belong to l_2 (as well as to arbitrary l_p for $1 \leq p < \infty$).

Analogously for arbitrary $1 \leq p < q \leq \infty$ consider sequence $\mathbf{x} = \{x_k\}$ where

$$x_k = \left(\frac{1}{k}\right)^{\frac{1}{p}} \tag{28}$$

It is easy to see that $\mathbf{x} \in l_q$ and $\mathbf{x} \notin l_p$.

The Theorem states that *vector spaces l_p are different for different p.*

To analyze different infinite dimensional spaces one need additional structures of vector spaces.

We consider the structure of *norm* on the vector spaces.

2.2 Norms on vector spaces

Norm it is "length" of the vector. We will see that in some sense all norms in finite-dimensional space are equivalent each other and it is not the case in infinite-dimensional space.

Definition The norm $||\cdot||$ of vectors in vector space V is the function on vectors which takes values in real numbers and obey the following properties:

$$||\mathbf{v}|| \geq 0 \text{ and } ||\mathbf{v}|| = 0 \iff \mathbf{v} = 0, \tag{29}$$

,

2) for every $\lambda \in \mathbf{R}$,

$$||\lambda \mathbf{v}|| = |\lambda| ||\mathbf{v}|| ., \tag{30}$$

3) for every two vectors $\mathbf{v}$, $\mathbf{u}$

$$||\mathbf{v} + \mathbf{u}|| \leq ||\mathbf{v}|| + ||\mathbf{u}|| . \qquad \text{(triangle inequality)} \tag{31}$$

We see that really norm can be viewed intuitively as a length.

Definition The vector space V with a norm $|| \cdot ||$ on it is called normed vector space $(V, || \cdot ||)$.

Examples of normed vector spaces.

1. Consider on the vector space l_p *in the case if* $p \geq 1$ the following norm[6]: for every $\mathbf{x} = \{x_k\} \in l_p$

$$||\mathbf{x}||_p = \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{\frac{1}{p}} , \quad \text{if} \quad 1 \leq p < \infty \tag{32}$$

and

$$||\mathbf{x}||_\infty = \sup_k (|x_k|) , \quad \text{if} \quad p = \infty \tag{33}$$

For example for the case $p = 1$

$$||\mathbf{x}||_1 = \left(\sum_{k=1}^{\infty} |x_k| \right) , \tag{34}$$

for $p = 2$

$$||\mathbf{x}||_2 = \sqrt{\sum_{k=1}^{\infty} |x_k|^2} , \quad \text{if} \quad 1 \leq p < \infty \tag{35}$$

Thus we come to normed vector spaces $(l_p, || \cdot ||_p)$ for $p \geq 1$.

The conditions (29) and (30) can be checked easily. To check the triangle inequality (31) for the general case (32) is a hard work. It is easy to check this condition for $p = 1$ (it is just triangle inequality for numbers). One can easy to check it for $p = \infty$ using considerations above when we proved

[6]here and below I often give definitions for general case where $p \geq 1$ or $p = \infty$. You can consider only cases $p = 1, 2, \infty$ and do not worry about general case.

that l_∞ is a vector space. For the case $p=2$ triangle condition follows from Cauchy-Bunyakovsky inequality [7]. Give a brief proof for this case.

Let $\mathbf{x},\mathbf{y}\in l_2$, $\mathbf{x}=\{x_i\},\mathbf{y}=\{y_i\}$ Using Cauchy-Bunyakovsky identity we have that for every N: $\sum_{k=1}^N (x_k+y_k)^2=$

$$=\sum_{k=1}^N x_k^2+2\sum_{k=1}^N x_k y_k+\sum_{k=1}^N y_k^2\leq \sum_{k=1}^N x_k^2+2\sqrt{\left(\sum_{k=1}^N x_k^2\right)\left(\sum_{k=1}^N y_k^2\right)}+\sum_{k=1}^N y_k^2 \tag{36}$$

Taking square root of right and left hand sides of this inequality we come to

$$\sqrt{\sum_{k=1}^N (x_k^2+y_k^2)}\leq \sqrt{\sum_{k=1}^N x_k^2}+\sqrt{\sum_{k=1}^N y_k^2} \tag{37}$$

Now taking limit $n\to\infty$ we come to the triangle inequality for $||\cdot||_2$

If $\mathbf{x}\in l_2$ but $\mathbf{x}\not\in l_1$ then $||\mathbf{x}||_1=\infty$. In the same way if $\mathbf{x}\in l_\infty$ but $\mathbf{x}\not\in l_2$ then $||\mathbf{x}||_2=\infty$. The following very important inequality holds:

$$||\mathbf{x}||_1\geq ||\mathbf{x}||_2\geq ||\mathbf{x}||_\infty \tag{38}$$

In the general case

$$||\mathbf{x}||_p\geq ||\mathbf{x}||_q \text{ if } p\leq q\,. \tag{39}$$

Prove inequality (38):

Consider arbitrary $\mathbf{x}\in l_1$. Then it evident that for arbitrary N

$$(|x_1|+|x_2|+\cdots+|x|_N)^2\geq |x_1|^2+|x_2|^2+\cdots+|x|_N^2\,. \quad \text{(Open the brackets!)}$$

Thus taking $N\to\infty$ we come to $||\mathbf{x}||_1^2\geq ||\mathbf{x}||_2^2$.

To prove the second inequality consider

$$|x_1|^2+|x_2|^2+\cdots+|x|_N^2\geq (\sup_k\{|x_1|,|x_2|,\ldots,|x|_N\})^2$$

Hence taking $N\to\infty$ we come to $||\mathbf{x}||_2^2\geq ||\mathbf{x}||_\infty^2$.

[7] In the case of sequences Cauchy-Bunyakovsky inequality states that $\sum a_k^2\sum b_k^2\geq (\sum a_k b_k)^2$. If we remember the formula expressing the angle between two vectors via scalar product ($(\mathbf{a},\mathbf{b})=|\mathbf{a}||\mathbf{b}|\cos\alpha$), then we see that Cauchy-Bunyakovsky inequality can be interpreted as a condition that cosinus of the angle between two vectors is less or equal to 1.

Note that the inequality above gives another proof the of Theorem (24) (E.g. suppose $\mathbf{x} \in l_1$, i.e. $\sum_{n=1}^{\infty} |x_n| < \infty$. Then by the inequality above $||\mathbf{x}||_2 = \sqrt{\sum_{n=1}^{\infty} x_n^2} < \infty$. Hence $\sum_{n=1}^{\infty} x_n^2 < \infty$, i.e. $\mathbf{x} \in l_2$).

One can discuss the meaning of inequality (38) drawing the pictures (see e.g. Coursework).

Two interesting questions remain:

1) Why $p \geq 1$?

2) What are relations between the definition of norm for $p < \infty$ and $p = \infty$. How the limit is preformed?

We try to give some hints of answering these questions in the next subsection.

Finally consider an important example

Example Consider the sequence $\mathbf{x} = x_k$ where $x_k = \frac{1}{k}$ (Harmonic sequence). Show in detail that this sequence belongs to l_2 and it does not belong to l_1.

Consider little bit more general sequence $\mathbf{x} = x_k$ where $x_k = \frac{1}{k^{1+s}}$ $(s \geq 0)$ and estimate $\sum x_k$ via integral $\int \frac{dx}{x^{1+s}}$: It is easy to see that

$$\int_k^{k+1} dx/x^{1+s} < 1/k^{1+s} < \int_{k-1}^{k} dx/x^{1+s}$$

Making summation by k we come to

$$1 + \sum_{k=2}^{N} \left(\int_k^{k+1} dx/x^{1+s} \right) < 1 + \sum_{k=2}^{N} 1/k^{1+s} < 1 + \sum_{k=2}^{N} \left(\int_{k-1}^{k} dx/x^{1+s} \right),$$

$$1 + \int_2^{N+1} dx/x^{1+s} < \sum_{k=1}^{N} 1/k^{1+s} < \int_1^{N} dx/x^{1+s}$$

Note that $1 > \int_1^2 dx/x^{1+s}$. Hence

$$\int_1^{N+1} \frac{dx}{x^{1+s}} < \sum_1^N \frac{1}{k^{1+s}} < 1 + \int_1^N \frac{dx}{x^{1+s}} \tag{40}$$

The integrals can be easily calculated: $\int \frac{dx}{x^{1+s}} = -\frac{1}{sx^s} + C$ for $s \neq 0$ and $\int \frac{dx}{x^{1+s}} = \log x + C$. We come to inequalities:

$$\log(N+1) < \sum_1^N \frac{1}{k} < 1 + \log N \tag{41}$$

and for $s > 0$:

$$\frac{1}{s} - \frac{1}{s(N+1)^s} < \sum_{1}^{N} \frac{1}{k^{1+s}} < 1 + \frac{1}{s} - \frac{1}{N^s} \tag{42}$$

It follows from these inequalities that the sequence $\mathbf{x} = x_k = \frac{1}{k^2}$ (s=1) belongs to l_1 and $1 \leq ||\mathbf{x}||_{\mathbf{1}} \leq \mathbf{2}$. (One can show that for this sequence $||\mathbf{x}||_{\mathbf{1}} = \frac{\pi^{\mathbf{2}}}{\mathbf{6}}$).

The harmonic sequence $x_k = \frac{1}{k}$ does not belong to l_1 but belongs to l_2.

2.3 Norm and convexity

Recall that the domain D of a affine space is called *convex* if for every two points $a, b \in D$ the interval $[a, b]$ belongs to D too.

Exercise In analysis the function $f(x)$ is called convex on the closed interval $[a, b]$ if its second derivative is positive. Prove that this is equivalent to the condition that the domain $\{(x, y)\colon\ x \in [a, b], y \geq f(x)\}$ is a convex domain on $\mathbf{R}^2$. (See in details Appendix 1)

Let $||\ ||$ be a norm on a vector space V. Consider the unit ball B_1 of this norm, i.e. all vectors such that their norm is less or equal to 1.

Lemma The unit ball of the norm is convex.

The proof is not difficult. Consider vectors $a, b \in B_1$. (We identify points with vectors which "connect" them with origin)

Then their length is less or equal than 1. Consider arbitrary point on the interval which connects points a, b. It is easy to see that it is equal to

$$x = \lambda a + \mu b\,, \quad \text{where } \lambda, \mu \geq 0, \text{ and } \lambda + \mu = 1\,.$$

Then using triangle inequality it is easy to see that

$$||x|| = ||\lambda a + \mu b|| = ||\lambda a + (1-\lambda)b|| \leq \lambda ||a|| + (1-\lambda)||b|| \leq \lambda + (1-\lambda) = 1\,. \tag{43}$$

This lemma allows us to study norms drawing he pictures.

E.g. consider unit sphere (boundary) for unit ball for norms $||\,,\,||_p$ with $p = 1, 2, \infty$.

Study for simplicity the case $V = \mathbf{R}^2$. let x, y be cartesian coordinates on $\mathbf{R}^2$.

For $p = 1$ the unit "circle"it is a square $|x| + |y| = 1$. Its vertices belong to axis and axis are diagonals parallel to axis. For $p = 2$ it is a circle $x^2 + y^2 = 1$. For $p = \infty$ the "circle" is a square $\max |x|, |y| = 1$ with edges parallel to axis.

Inequality (38) follows from the picture that the "circle" for $p = 1$ is in the circle for $p = 2$ and the circle $p = 2$ is in the "circle" $p = 1$.

Drawing these pictures we can answer on questions why for $p = 1$ the formula (32) does not give the norm. Indeed if we draw the picture of unit ball for $p < 1$ we will see that it is not convex.

Another interesting analysis using pictures: Drawing unit "circles" for $p \to \infty$ we will see that they tend to the "circle" for $p = \infty$.

2.4 Regression: Norms ⇒ Metrics ⇒ Topology

We just remind here very briefly

1) A norm $||\cdot||$ on vector space V induces metrics $d_{||\,||}$ on this space:

$$d_{||\,||}(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}|| \tag{44}$$

The open ball $B_\varepsilon(\mathbf{x})$ in this metrics is just set of the vectors $\mathbf{y}$ such that $||\mathbf{x} - \mathbf{y}|| < \varepsilon$.

Before defining topology induced by norm recall the definition of topology on set X:

Definition The collection $\mathcal{F}$ of sets in space X (subset in the space $\mathcal{P}(X)$ of all the sets of X) defines topology if the following conditions are satisfied:

- empty set and whole set belong to $\mathcal{F}$: $\emptyset \in \mathcal{F}$, $X \in \mathcal{F}$
- union of arbitrary sets in $\mathcal{F}$ is set in $\mathcal{F}$: if $V_\alpha \in \mathcal{F}$ for $\alpha \in I$ then

$$\cup_{\alpha \in I} V_\alpha \in \mathcal{F}$$

- If two sets V_1, V_2 belong to $\mathcal{F}$ then their intersection belongs to $\mathcal{F}$ too: if $V_1 \in \mathcal{F}$ and $V_2 \in \mathcal{F}$ then

$$V_1 \cap V_2 \in \mathcal{F}$$

Subsets belonging to $\mathcal{F}$ are called *open subsets*.

But only intersection of finite number of open subsets is open.

The space X with topology $\mathcal{F}$ on it is called topological space.

Example. Consider $\mathbf{R}$. We call the set $V \subset \mathbf{R}$ open if every point x in the set V is interior, i.e. there exist δ such that interval $(x - \delta, x + \delta)$ belongs to V. In other words set is open if it is a union of open intervals. It is easy to see that the collection of open sets define topology on X.

The concept of topological space is most abstract concept to formulate continuity[8]:

Now return to metric spaces and to normed vector spaces. Let (X, d) be a metric space with d metric on it (in the case of normed vector space $(V, ||\cdot||)$ metric is defined by (44)).

Definition In a metric space (X, d) a point x is called an an internal point of a set U ($x \in U$) if there exists a ball $B_\varepsilon(x)$ which belongs to U:

$$x \in U \text{ is an internal point in } U \text{ if } \quad \exists \varepsilon > 0\colon \quad d(y,x) < \varepsilon \Rightarrow y \in U$$

Definition A set U in the metric space (X, d) is called *open set* if its all points are internal.

It is easy to see that this definition of openness obey all axioms above.

The topology on the normed vector space $(V, ||\cdot||)$ is defined by the set of all open sets.

2.5 Equivalence of norms.

Definition Two norms $||\cdot||, ||\cdot||'$ defined on vector space V are equivalent if there exist constants $c_1 > 0$, $c_2 > 0$ such that for every vector $\mathbf{x} \in V$

$$c_1||\mathbf{x}|| \leq ||\mathbf{x}||' \leq c_2||\mathbf{x}|| \,. \tag{45}$$

We will see that

- Equivalent norms lead to the same class of continuous functions (see the first Theorem after Lemma and remarks at the end of the Section)
- This concept is "empty" in finite-dimensional space. All the norms are equivalent in finite dimensional vector space: see the second Theorem after the Lemma

Now we formulate the following important lemma:

Lemma

[8] *The function f is called continuous if the preimage of every open set is continuous.* One can see that this definition is substraction of definitions of continuous functions.

Let V be a vector space with two norms $||\cdot||$, $||\cdot||'$ defined on it. Then the following conditions are equivalent:
a) There exists a positive constant $c > 0$ such that for every $\mathbf{x} \in V$

$$||\mathbf{x}|| \leq c||\mathbf{x}||', \tag{46}$$

b) The function $f(x) = ||\mathbf{x}||$ is continuous in the topology induced by the norm $||\cdot||'$,

c) The topology $\mathcal{F}$ induced by the norm $||\cdot||$ belongs to the topology $\mathcal{F}'$ induced by the norm $||\cdot||'$, $\mathcal{F} \subseteq \mathcal{F}'$, i.e. if a set U is an open set in the topology induced by the norm $||\cdot||$ then this set is an open set in the topology induced by the norm $||\cdot||'$.

The lemma states that:

$$(a) \Leftrightarrow (b) \Leftrightarrow (c) \tag{47}$$

We give the proof of this Lemma after formulation and proving two very important Theorems which follow from this lemma:

Theorem

Equivalent norms define on a vector space the same topology.

Theorem

All norms on a finite-dimensional vector space are equivalent.

The first Theorem is simply a corollary of the Lemma.
Indeed consider on normed vector space two equivalent norms (45). To prove that these norms define the same topology we have to prove that every set U is open in the topology $\mathcal{F}'$ iff it is open in the topology $\mathcal{F}$. And it is just the statement (47).

Remark This means that every notion of continuity in both norms coincide: function is continuous with respect to the norm $||\cdot||$ iff it is continuous with respect to the norm $||\cdot||'$, because continuity is defined in terms of open sets. (See also Remarks in the end of the Section)

The second Theorem justifies the point of view that conception of norm is not essential in finite-dimensional case (from the point of view of defining topology).

To prove the second Theorem prove that in finite-dimensional space every norm is equivalent to the standard norm $|| \quad \cdot ||_2$. ($||\mathbf{x}||_2 = \sqrt{x_1^2 + \cdots + x_n^2}$.)

Consider the function $f = ||\mathbf{x}||'$, where $|| \cdot ||'$ is an arbitrary norm on $\mathbf{R}^n$.

First prove that this function is continuous w.r.t. the standard norm (i.e. topology induced by standard norm $|| \cdot || = || \cdot ||_2$). This follows from the inequality:

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| = \big| ||\mathbf{x}||' - ||\mathbf{x}_0||' \big| \leq ||\mathbf{x} - \mathbf{x}_0||' =$$

$$|| \sum_{k=1}^{n} (x^k - x_0^k)\mathbf{e}_k||' \leq \sum_{k=1}^{n} |x^k - x_0^k| \, ||\mathbf{e}_k||' \leq M \sum_{k=1}^{n} |x^k - x_0^k| \leq M\sqrt{n}||\mathbf{x} - \mathbf{x}_0||_2 \, .$$

Here $\mathbf{e}_n$ is a basis in $\mathbf{R}^n$, $M = \max\{||\mathbf{e}_1||', \ldots, ||\mathbf{e}_n||'\}$. The last inequality is due to Cauchy-Bunyakovsky-Schwarz inequality:

$$\sum_{i=1}^{n} |a_i| \leq \sqrt{n}\sqrt{a_1^2 + \cdots + a_n^2} \, . \tag{48}$$

Consider the sphere of radius 1 with respect to the standard norm $|| \quad ||_2$: $S = \{x \colon \quad ||x||_2 = \sqrt{x_1^2 + \cdots + x_n^2} = 1\}$. The function $f = ||\mathbf{x}||'$ is continuous on Σ^n and sphere S^n is compact, hence function f is bounded on S^n: $m \leq f(x) \leq M$. (Moreover it attaints maximal and minimal values m, M.) Thus we see that for every $\mathbf{x} \in S^n$ ($||\mathbf{x}||_\mathbf{2} = \mathbf{1}$) $m \leq ||\mathbf{x}||' \leq M$:

$$m \leq ||\mathbf{x}||' \leq M \quad \text{if } ||\mathbf{x}||_2 = 1$$

Hence for every vector $\mathbf{x}$

$$m||\mathbf{x}||_2 \leq ||\mathbf{x}||' \leq M||\mathbf{x}||_2$$

Remark Pay attention: we use essentially finite-dimensionality of the space $\mathbf{R}^n$ when we proved continuity of the function $f = ||\mathbf{x}||'$ in topology induced by the norm $|| \quad ||_2$ and when we consider S^n as compact! (Look on the inequality (48)) To understand it more properly solve the following exercises:

Exercise: Prove that function $f = ||\mathbf{x}||_1$ is *not continuous* function on the space l_1 with respect to the norm $|| \cdot ||_2$.

To do it consider for example the sequence $\{\mathbf{x}^{(n)}\}$ of the following elements (sequences): each $\mathbf{x}^{(n)}$ is a sequence such that its first n terms are equal to $1/n$ and other terms are equal to 0:

$$\mathbf{x}^{(n)} = \underbrace{\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}}_{\text{first } n \text{ terms}}, 0, 0 \dots, 0, \dots \tag{49}$$

Then it is evident that $||\mathbf{x}^{(n)}||_2 \to 0$ and $||\mathbf{x}^{(n)}||_1 = 1$. The sequence $\mathbf{x}^{(n)}$ tends to zero in the topology $\mathcal{F}_2$ but the value of the function $f = ||\mathbf{x}||_1$ on this sequence is equal to 1. On the other hand the value of the function $f = ||\mathbf{x}||_1$ on $\mathbf{x} = 0$ is equal to zero. Hence f is not continuous function in the topology $\mathcal{F}_2$ induced by the norm $|| \cdot ||_2$ at least at the point 0.

Another example: Function $f = ||\mathbf{x}||_2 = \sqrt{\sum_k x_k^2}$ is not continuous function on the space $(l_2, || \, ||_\infty)$, (i.e. on l_2 with respect to the topology induced by the norm $||\mathbf{x}||_\infty = \sup_k\{x_k\}$). To see it consider a sequence $\{\mathbf{x}^{(\mathbf{n})}\}$ of elements (sequences) such that first n^2 terms of the sequence $\mathbf{x}^{(\mathbf{n})}$ are equal to $1/n$ and all other terms of this sequence are equal to zero. Then it is easy to see that for all these sequences their $|| \cdot ||_2$ norm is equal to 1 and their $|| \cdot ||_\infty$ norm tends to zero. The function $f = ||\mathbf{x}||_2$ takes constant non-zero value 1 on the sequence of points tending to zero in the topology $\mathcal{F}_\infty$. This just means that this function is not continuous in the topology induced by the norm $|| \cdot ||_\infty$ at least at the zero.

2.6 Proof of the Lemma (47)

.

We prove that

$$(b) \Rightarrow (a) \Rightarrow (c) \Rightarrow (b)$$

1) Proof that $(b) \Rightarrow (a)$.

Let $f(\mathbf{x}) = ||\mathbf{x}||$ be a function continuous in topology $\mathcal{F}'$. $f(\mathbf{0}) = 0$. Condition of continuity at the point $\mathbf{x} = \mathbf{0}$ is following: for all $\varepsilon > 0$ there exists $\delta > 0$ such that the image of the ball $B_\delta^{(\prime)}(0)$ belongs to the interval $(-\varepsilon, \varepsilon)$, i.e.

$$||\mathbf{x}||' < \delta \Rightarrow |f(\mathbf{x})| = ||\mathbf{x}|| < \varepsilon \, .$$

Pick any $\varepsilon > 0$ and consider corresponding δ. We will prove that for arbitrary $\mathbf{x}$

$$||\mathbf{x}|| \leq \frac{2\varepsilon}{\delta}||\mathbf{x}||' , \tag{50}$$

i.e. the condition (46) is satisfied for $c = 2\varepsilon/\delta$.

It is evident for $\mathbf{x} = 0$. Consider arbitrary $\mathbf{x} \neq 0$. Suppose $||\mathbf{x}||' = r$.

Then $f(\delta\mathbf{x}/2r) = ||\delta\mathbf{x}/2r||' = \frac{\delta}{2r}r = \delta/2 < \delta$. Hence $||\frac{\delta\mathbf{x}}{2r}|| < \varepsilon$ Thus we come to (50):

$$||\mathbf{x}|| = \frac{2r}{\delta}||\frac{\delta\mathbf{x}}{2r}|| < \frac{2r}{\delta}\varepsilon = \frac{2\varepsilon}{\delta}||\mathbf{x}||' .$$

2) Proof that $(a) \Rightarrow (c)$.

We have to prove the following: it follows from the condition (46) that $\mathcal{F} \subseteq \mathcal{F}'$, i.e. set U is open in topology $\mathcal{F}'$ if it is open in topology $\mathcal{F}$.

Suppose U is an arbitrary open set in the topology $\mathcal{F}$. Consider an arbitrary point $a \in U$ and prove that this point belongs to some ball $B^{(\prime)}_\varepsilon(a)$ such that $B^{(\prime)}_\varepsilon(a) \subseteq U$ [9]. Thus we will prove that the set U is open in the topology $\mathcal{F}'$.

A point a belongs to the open set U in the topology $\mathcal{F}$. Hence there exists a ball $B_\varepsilon(a)$ such that this ball belongs to U, i.e. all points $\mathbf{x}$ such that $||\mathbf{x} - a|| < \varepsilon$ belong to U.

Consider the ball $B^{(\prime)}_{\frac{\varepsilon}{c}}(a)$ and prove that this ball belongs to the ball $B_\varepsilon(a)$. If $\mathbf{x}$ belongs to the ball $B^{(\prime)}_{\frac{\varepsilon}{c}}(a)$ then $||\mathbf{x} - a||' < \frac{\varepsilon}{c}$. Hence according to the condition (46) $||\mathbf{x} - a|| < \varepsilon$. This means that $B^{(\prime)}_{\varepsilon/c}(a) \subseteq B_\varepsilon(a)$. Hence the ball $B^{(\prime)}_{\varepsilon/c}(a)$ belongs to U. Thus we prove that a is an internal point in topology $\mathcal{F}'$.

3) Proof that $(c) \Rightarrow (b)$.

Let $\mathcal{F} \subseteq \mathcal{F}'$. Prove that the function $f(\mathbf{x}) = ||\mathbf{x}||$ is continuous in topology $\mathcal{F}'$. Take arbitrary point $\mathbf{x}_0$. Consider for arbitrary ε the open ball $B_\varepsilon(\mathbf{x}_0)$:

$$B_\varepsilon(\mathbf{x}_0) = \{\mathbf{x}: \qquad ||\mathbf{x} - \mathbf{x}_0|| < \varepsilon\} .$$

The ball $B_\varepsilon(\mathbf{x}_0)$ belongs to $\mathcal{F}$, i.e. it is an open set in the topology $\mathcal{F}$. Hence the ball $B_\varepsilon(\mathbf{x}_0)$ belongs to $\mathcal{F}'$, i.e. it is an open set in the topology $\mathcal{F}$, because

[9] as always a ball $B^{(\prime)}_\varepsilon(a)$, $(B_\varepsilon(a))$ is a set of points such that the distance between these points in metric induced by $||\cdot||'$ (in metric induced by $|| \quad ||$) is less that ε.

$\mathcal{F} \subseteq \mathcal{F}'$. Hence the point $\mathbf{x}_0$ is an internal point of this ball in the topology $\mathcal{F}'$, i.e. there exist a δ such that

$$B'_\delta(\mathbf{x}_0) \subseteq B_\varepsilon(\mathbf{x}_0), \text{ i.e. } \quad \forall \mathbf{x} \ \ ||\mathbf{x}-\mathbf{x}_0||' < \delta \Rightarrow ||\mathbf{x}-\mathbf{x}_0|| < \varepsilon$$

Thus we come to inequality which proves continuity of f in the point $\mathbf{x}_0$:

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| = |||\mathbf{x}|| - ||\mathbf{x}_0||| \leq ||\mathbf{x}-\mathbf{x}_0|| < \varepsilon \,, \text{if } ||\mathbf{x}-\mathbf{x}_0||' < \delta$$

Remark Indeed one can prove in a way similar to the proof above that from condition c) it follows that *every* function which is continuous in topology $\mathcal{F}$ will be continuous in topology $\mathcal{F}'$.

One can give very short proof of this fact based on the statement that f *is continuous function iff the preimage of an arbitrary open set is open.*

Indeed let f be continuous function in topology $\mathcal{F}$. (We consider an arbitrary continuous function from normed vector space V to $\mathbf{R}$, not necessarily the function $f = ||\mathbf{x}||$). Prove that it is continuous in topology $\mathcal{F}'$ too. Let U be an open set in $\mathbf{R}$. Then its preimage $f^{-1}(U)$ be an open set in topology $\mathcal{F}$. Hence It will be open in topology $\mathcal{F}'$. Hence f is continuous function in topology $\mathcal{F}'$

Remark We proved that $b) \Rightarrow a) \Rightarrow c) \Rightarrow b)$

It is very instructive to prove straightforwardly that condition a) follows from the condition c), i.e. if every set open in $\mathcal{F}$ is open in $\mathcal{F}'$ then the condition (46) is satisfied. Prove it. Consider the ball of the radius 1 with the centre in origin in topology $\mathcal{F}$, $B_1(0) = \{\mathbf{x}\colon\ ||\mathbf{x}|| < 1\}$. It is open set in $\mathcal{F}$. Hence it is open set in $\mathcal{F}'$. In particularly the origin, the point 0 is internal point. This means that there exists ε such that a ball $B_\varepsilon(0)'$ belongs to the ball $B_1(0)$. Thus we prove that there exists ε such that

$$||\mathbf{x}||' < \varepsilon \Rightarrow ||\mathbf{x}|| < 1$$

This is just the condition (46). Namely take an arbitrary $\mathbf{x}$. Suppose that $||\mathbf{x}||' = r$. Then for the vector $\frac{\varepsilon \mathbf{x}}{2r}$ we have that $||\frac{\varepsilon \mathbf{x}}{2r}||' < \varepsilon$. Hence $||\frac{\varepsilon \mathbf{x}}{2r}|| < 1$ and $||\mathbf{x}|| < \frac{2r}{\varepsilon}$. We proved that $||\mathbf{x}|| < \frac{2r}{\varepsilon}$ in the case if $||\mathbf{x}||' = r$. Hence we proved the condition (46) for $c = \frac{2}{\varepsilon}$.

3 Banach spaces

Definition Normed vector space $(V, ||\cdot||)$ is called *Banach space* if it is complete (with respect to the metric induced by the norm $||\cdot||$), i.e. if every Cauchy sequence has a limit in this space.

A short Remaining:

The sequence x_n of points in metric space (X, d) is called fundamental sequence (Cauchy sequence) if for every $\varepsilon > 0$ there exists N such that for every $m, n > N$ $d(x_n, x_m) < \varepsilon$.

If sequence has limit then it is fundamental. In general case not every fundamental sequence has a limit.

Metric space (X, d) is called complete metric space if every fundamental sequence has a limit.

Metric space (X, d) is complete if every fundamental sequence (Cauchy sequence) has a limit in it.

E.g. $\mathbf{R}$ with standard metric is a complete metric space. Its subset of rational points is not complete metric space: for example the sequence of rational points: $x_n = 1, \frac{3}{2}, \frac{7}{5}, \frac{17}{12}, \frac{41}{29}, \frac{99}{70}, \frac{239}{169}, \dots$ has not limit in $\mathbf{Q}$ (its limit is $\sqrt{2} \in \mathbf{R}$)

Before considering examples of Banach space we formulate in the next subsection very important Lemma.

3.1 Lemma: Space of bounded functions on arbitrary set is Banach space

Let S be an arbitrary set. We emphasize the fact that S is arbitrary set, for example the set $\{\dagger, *, \&\}$ of three elements.

Define $C(S)$ as a space of bounded real-valued (or complex-valued) functions on this space.

For example the space $C(S)$ for $S = \{\dagger, *, \&\}$ is nothing but the set of triples $\{a, b, c\}$ where $a, b, c \in \mathbf{R}$. (If S is finite set then the condition to be bounded is automatically obeyed.)

The space $C(S)$ for $S = \{$ the set of positive integers $\}$ is nothing but the set l_∞ of bounded sequences $\{a_1, a_2, a_3, \dots\}$ where $a_1, a_2, a_3, \dots \in \mathbf{R}$.

The space $C(S)$ for $S = [0, 1]$ is nothing but the set of bounded functions on $[0, 1]$.

$C(S)$ is linear space and one can consider the following norm: for every $f\colon\ S \to \mathbf{R}$

$$||f|| = \sup_{x \in S} |f(x)| \tag{51}$$

Not difficult to prove that it is the norm.

(Note that if $S = \{1, 2, 3, 4, \dots\}$ then the norm above is nothing but the standard norm $||, ||_\infty$ on the l_∞.)

Now we formulate and prove very simple but very important lemma:.

Lemma (52)

For every set S the normed linear space $C(S)$ of bounded functions on S with the norm (51) *is Banach space.*

Proof.

Consider arbitrary fundamental sequence $\{f_n(x)\}$, $(x \in S)$.

Take arbitrary $x_0 \in S$. $\{f_n(x)\}$ is an fundamental sequence of functions on S. Hence in particular $\{f_n(x_0)\}$ is *fundamental sequence* of real numbers!:

From condition

$$\forall \varepsilon\, \exists N : \qquad \forall m, n > N\,, ||f_n - f_m|| < \varepsilon \tag{53}$$

follows the condition:

$$\forall \varepsilon\, \exists N : \qquad \forall m, n > N\,, |f_n(x_0) - f_m(x_0)| < \varepsilon \tag{54}$$

But $\mathbf{R}$ is complete. Hence there exists a limit of the sequence $\{f_n(x_0)\}$. Denote this limit by $g(x_0)$.

Thus we construct a new function such that for every $x \in S$ $\lim_{n\to\infty} f_n(x) = g(x)$. We will prove now that $g(x)$ is bounded function and $f_n \to g$ (in the norm (51)).

Prove first that $g(x)$ is bounded. Take arbitrary ε. Then according to (53) choose N such that for all $n, m > N$

$$|f_n(x) - f_m(x)| < \varepsilon \tag{55}$$

for all $x \in S$. Take $m \to \infty$ in this relation we come to

$$|f_n(x) - g(x)| \le \varepsilon \tag{56}$$

Now using triangle inequality we see that $|g(x)| \le |f_n(x) - g(x)| + |f_n(x)| \le |f_n(x)| + \varepsilon$. So it is bounded.

The fact that $f_n \to g$ immediately follows from relations (55) and (56).

Thus we prove the Lemma (52).

3.2 Basic examples of Banach spaces: $l_1, l_2, l_\infty, C([a,b])$.

Theorem

l_p with norm $||\cdot||_{(p)}$ is Banach space if $p = 1, 2, \infty$.

Note that for $p = \infty$ the statement of Theorem is reformulation of the Lemma above, because l_∞ is just the space of bounded functions on the set S of positive integers. (See examples above). In the case $l = l_1$, l_2 we need little more work.

Give a brief proof for $p = 1$. Let $\mathbf{x^{(n)}}$ be Cauchy sequence in l_1. Then due to embedding $l_1 \subset l_\infty$ and relation $||\mathbf{x}||_\infty \leq ||\mathbf{x}||_\mathbf{1}$ it is fundamental sequence in l_∞ which is Banach space (w.r.to the norm $||\cdot||_\infty$). Let $\mathbf{y}$ be a limit of the fundamental (Cauchy) sequence in l_∞: $||\mathbf{x^{(n)}} - \mathbf{y}||_\infty \to 0$. Prove that $\mathbf{y} \in l_1$ and the sequence of the sequences $\mathbf{x^{(n)}}$ tends to the sequence $\mathbf{y}$ with respect to the norm $||\cdot||_1$.

For every $\varepsilon > 0$ there exists N such that for every $n, m > N$ $||\mathbf{x^{(n)}} - \mathbf{x^{(m)}}|| = \sum_{i=1}^\infty |x_i^{(n)} - x_i^{(m)}| < \varepsilon/2$. Thus for any k

$$\sum_{i=1}^k |x_i^{(n)} - x_i^{(m)}| < \varepsilon/2 \Rightarrow \lim_{m\to\infty} \sum_{i=1}^k |x_i^{(n)} - x_i^{(m)}| = \sum_{i=1}^k |x_i^{(n)} - y_i| \leq \varepsilon/2$$

Thus $\lim_{k\to\infty} \sum_{i=1}^k |x_i^{(n)} - y_i| = \sum_{i=1}^\infty |x_i^{(n)} - y_i| = ||\mathbf{x}^{(n)} - \mathbf{y}||_1 < \varepsilon$ We proved that for every $\varepsilon > 0$ there exists N such that for every $n > N$ $||\mathbf{x}^{(n)} - \mathbf{y}||_1 < \varepsilon$. This means that $\mathbf{y}$ is a limit of the sequence $\{\mathbf{x}^{(n)}\}$ with respect to the norm $||\cdot||_1$ ■

In a similar way one can prove that l_2 is a Banach space (with respect to the norm $||\cdot||_2$)

Now we consider another very important basic example of Banach space:

The space $C([a,b])$ of continuous functions on the closed interval $[a,b]$ with norm:

$$||f|| = \max_{x\in[a,b]} |f(x)| \tag{57}$$

It is easy to see that the space $C([a,b])$ is a vector space and the relation above defines norm correctly.

Theorem

The space of continuous functions on $[a, b]$ *with norm* $||f|| = \max_{x\in[a,b]} |f(x)|$ *is Banach space.*

The proof of this Theorem is essentially founded on the Lemma(52).

Proof of the Theorem.

Consider vector space $C([a, b])$ with the norm $||f|| = \sup_x |f(x)|$. This norm is well-defined on $C([a, b])$ because $[a, b]$ is compact. Let $\{f_n(x)\}$ be a fundamental sequence in $C([a, b])$.

It follows immediately from lemma that the fundamental sequence $\{f_n(x)\}$ of continuous functions tends uniformly to the bounded function. Namely $C([a, b])$ is the subspace of the space $B([a, b])$ of bounded functions on the closed interval $[a, b]$. (This is because every continuous function on the closed interval is bounded.) Hence it follows from Lemma (52) that the space $B([a, b])$ is Banach space. Hence every fundamental sequence $\{f_n(x)\}$ of continuous functions has a limit in $B([a, b])$ the bounded function $F(x)$. $f_n \to F$ (in topology induced by the norm[10] (57).

It remains to prove that this bounded function $F(x)$ is continuous too, i.e. it belongs to the space $C([a, b])$. Prove it. For arbitrary x and arbitrary ε take n such such that for function f_n, $||F - f_n|| < \frac{\varepsilon}{3}$. Take δ such that $|f_n(x) - f_n(y)| < \frac{\varepsilon}{3}$ for every $y \in (x - \delta, x + \delta)$. This is due to continuity of the function f_n at the point x. Then $|F(x) - F(y)| \le |F(x) - f_n(x)| + |f_n(x) - f_n(y)| + |F(y) - f_n(y)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$. Thus F is continuous ■

3.3 Completion of normed spaces

Definition The metric space $(\tilde{X}, \tilde{d})$ is completion of the metric space (X, d) ($\tilde{d}$, d are metrics on X, $\tilde{X}$ respectively) if

- $X \subseteq \tilde{X}$,
- metric $\tilde{d}$ coincides with a metric d on X: $\tilde{d}\big|_X = d$
- $(\tilde{X}, \tilde{d})$ is a complete metric space.

[10] It is not worthless to repeat the proof again just for this case: The fundamental sequence $\{f_n\}$ defines the fundamental sequence of numbers $\{f_n(x_0)\}$ at the every point $x_0 \in [a, b]$. $\mathbf{R}$ is complete. Consider $F(x) = \lim_{n\to\infty}\{f_n(x_0)\}$. If $||f_n - f_m|| < \varepsilon$ then taking $n \to \infty$ we see that $||F - f_m|| \le \varepsilon$. Hence F is bounded function and $f_n \to F$ in $||\cdot||$.

- X is dense in $\tilde{X}$

A very important Hausdorff Theorem states that every metric space (X, d) can be completed and all completions are isometric. The proof of this Theorem is based on the following consideration:

Consider the space $\mathcal{X}$ of *all fundamental sequences of points in* X. Then consider the following equivalence relation: we say that two elements in $\mathcal{X}$,i.e. two sequences $\{x_n\}$, $\{y_n\}$ in X are equivalent if the sequence $\{z_n\}$, such that $z_{2k} = x_k, z_{2k+1} = y_k$ is fundamental sequence too. It is easy to see that the set of equivalence classes of $\mathcal{X}$ is naturally provided with metric and it will be a completion of (X, d). (It is just the construction of real numbers from rationals)[11].

We study some examples of non-complete spaces and consider their completions.

3.4 l_1 as non-complete subspace of the Banach space $(l_2, || \,||_2)$. Banach space $(l_2, || \,||_2)$ is a completion of the normed space $(l_1, || \,||_2)$.

We know that $l_1 \neq l_2$. (E.g. consider harmonic sequence).

Now consider l_1 as a subspace of normed vector space $(l_2, || \cdot ||_2)$,

We prove now that

- normed vector space $(l_1, || \cdot ||_2)$ is not complete, it is not closed subspace of the Banach space $(l_2, || \cdot ||_2)$.

- arbitrary element $\mathbf{y}$ of l_2 can be considered as limit of sequence of elements $\mathbf{x^{(n)}}$ of l_1 w.r.t. the norm $|| \cdot ||_2$.

  Thus we will prove what we claimed in the title of the subsection:

  Banach vector space $((l_2, || \cdot ||_2))$ *is a completion of the normed vector space* $(l_1, || \cdot ||_2)$

Proof:

Namely consider arbitrary element $\mathbf{y} = \{y_1, y_2, y_3 \dots\}$ in l_2.

We find a sequence $\mathbf{x}^{(n)}$ of elements in l_1 such that they tend to $\mathbf{y}$ with respect to the norm $|| \cdot ||_2$.

[11]The molto of this proof is: the realm of ideas is absolute

Namely define for an arbitrary n an element $\mathbf{x}^{(n)}$ in the following way: The first n terms of the sequence $\mathbf{x}^{(n)}$ coincide with first n elements of the sequence $\mathbf{y}$ and all other elements of the sequence $\mathbf{x}^{(n)}$ are equal to zero:

$$\mathbf{x}^{(n)} = \{x_k^{(n)}\}\colon \begin{cases} x_k^{(n)} = y_k & \text{if} \quad k \leq n \\ x_k^{(n)} = 0 & \text{if} \quad k \geq n+1 \end{cases}$$

Then one can see that

1) all sequences $\mathbf{x}^{(n)}$ belong to l_1 (because they are finite).

2)$\mathbf{x}^{(n)} \to \mathbf{y}$ in the topology induced by the norm $|| \, ||_2$:

$$||\mathbf{x}^{(n)} - \mathbf{y}||_2 = \sqrt{y_{n+1}^2 + y_{n+2}^2 + \dots} \to 0$$

because $\sqrt{\sum_{k=1}^{\infty} y_k^2}$ is convergent. Hence $\mathbf{x}^{(n)}$ is fundamental sequence in l_1 which tends to $\mathbf{y}$. If $\mathbf{y} \notin l_1$ then this fundamental sequence has not limit in l_1. We proved that $(l_1, || \cdot ||_2)$ is not complete space (is not Banach space) and $(l_2, || \cdot ||_2)$ is its completion ∎

Exercise Show that l_2 as a non-complete subspace of the Banach space $(l_\infty, || \, ||_\infty)$. (See solutions of Coursework)

Is it true that Banach space $(l_\infty, || \, ||_\infty)$ is a completion of the normed space $(l_2, || \, ||_\infty)$? No it is not right. E.g. consider the element $\mathbf{y} = (1, 1, 1, \dots,)$ in l_∞ One can show that there is no a sequence $\mathbf{x}^{(n)}$ in l_2 which tends to $\mathbf{y}$. Try to prove it. (Use the fact that for any element $\mathbf{x} = \{x_1, x_2, \dots \dots\}$ in l_2 $x_n \to 0$.)

3.5 Banach space of continuous functions on $[a, b]$ as a completion of the space $C^\infty([a, b])$ of smooth functions. Another glance on the Weierstraß Theorem.

We know very well that according to Weierstraß Approximation theorem the space $P([a, b])$ of polynomials is dense in the Banach space $C([a, b])$.

Look on this statement from another point of view:

Proposition The vector space of polynomials with norm (57) is non-complete normed space. Its completion is Banach space $C([a, b])$.

It is another reformulation of Weierstraß Approximation Theorem. Indeed the fact that every continuous function can be uniformly approximated by polynomials means that every continuous function is limit of fundamental sequence of polynomials.

From this statement immediately follows that the vector space $C^\infty([a,b])$ of smooth functions with norm (57) is not complete normed space and its completion is Banach space of continuous functions. Prove this statement.

Arbitrary continuous function on $[a,b]$ can be uniformly approximated by polynomials (according WeierstraßApproximation Theorem). But polynomials are smooth functions!. Hence we proved that $C([a,b])$ is completion of $C^\infty([a,b])$.

Polynomials are smooth functions. Hence Banach space $C([0,1])$ is a completion of the subspace $C^\infty([0,1])$ of smooth functions. ■

3.6 Another important examples of subspaces of $C([a,b])$

The space $C([a,b])$ with metric induced by the norm $||f|| = \max_{x\in[a,b]} |f(x)|$ is very important example of Banach space. In the previous subsection we showed examples of normed but non-complete subspaces of this Banach space. Now we consider two very important examples of subspaces of $C([a,b])$ where we will consider norm on subspaces different from the initial norm $||f|| = \max_{x\in[a,b]} |f(x)|$.

Example 1 Consider subspace functions in $C([a,b])$ which have continuous first derivative. The subspace of this functions we denote by $C^1(a,b)$. It is easy to see that this subspace *is not complete* with respect to the norm $||f|| = \max_{x\in[a,b]} |f(x)|$. Indeed consider e.g. the sequence $\{P_n(x)\}$ of polynomials which according to Weierstraß Approximation Theorem tends to the function $f = |x|$ on the interval $[0,1]$. This function is continuous but it is not differentiable at the point $x = 0$. On the other hand all polynomials belong to $C^1([-1,1])$. Hence the $\{P_n(x)\}$ is Cauchy sequence in $C^1([-1,1])$ such that its limit, the function $f = |x|$ is outside of the space $C^1([-1,1])$.

One can consider another norm on the space $C^1([a,,b])$ which makes it Banach space. To construct this metric note that for every $f \in C^1([-1,1])$ the derivative f' is continuous function as well as the function f. Hence f' as well as f have well-defined norm in $C([a,,b])$. Consider a new norm

$$||f||^1 = ||f|| + ||f'|| = \max_x(|f(x)|) + \max_x(|f'(x)|) \tag{58}$$

where $||f|| = \max_x(|f(x)|)$ is a standard norm on $C([a,b])$.

It is easy to see that this relation indeed defines a norm on $C^1(a,b)$.

Exercise Prove that the space $C^1([a,b])$ with the norm (58) is Banach space.

Solution To show that it is complete space it is suffice to show that it is completion of the space of polynomials with respect to the norm (58). It means that we have to prove that every function $f \in C^1(a,b)$ can be approximated by polynomials with respect to the norm (58). Indeed consider polynomials $\{P_n\}$ which tend according to Weierstraß Approximation Theorem to the derivative f' (w.r.t. to the usual norm). Then consider polynomials

$Q_n(x) = \int_a^x P_n(t)dt$ It is easy to see that polynomials $Q_n(x)$ tend to the function f with respect to the usual norm. Hence the sequence $\{Q_n(x)\}$ tends to function f with respect to the norm (58).

Example 2 Consider on the space of continuous functions a norm:

$$||f||' = \sqrt{\int_a^b f^2(x)dx} \tag{59}$$

This norm usually is denoted by $||f||_{(2)}$ (compare with notation for norm in l_2)

$$||f||' = ||f||_{(2)} = \sqrt{\int_a^b f^2(x)dx} \tag{60}$$

Exercise Prove that it is a norm.

Proof The conditions $||f|| = 0 \Leftrightarrow f \equiv 0$ and $||\lambda f|| = |\lambda| ||f||$ are obeyed. It is trivial.

Triangle inequality follows from Cauchy-Bunyakovsky inequality. [12] Indeed according to this inequality

$$\left(||f+g||_{(2)}\right)^2 = \int_a^b (f+g)^2 dx = \int_a^b f^2 dx + 2\int fg dx + \int_a^b g^2 dx$$

$$\leq \int_a^b f^2 dx + 2\sqrt{\int_a^b f^2 dx \int_a^b g^2 dx} + \int_a^b g^2 dx = \left(||f||_{(2)} + ||g||_{(2)}\right)^2$$

[12] It is not worthless to repeat this very improtant inequality for the case of functions space: Cauchy-Bunyakovsky inequality is:

$$\left(\int_a^b f(x)g(x)dx\right)^2 \leq \int_a^b f^2(x)dx \int g^2(x)dx \tag{61}$$

. To prove it consider polynomial $P(t) = \int_a^b (tf(x) - g(x))^2 dx = At^2 - 2ABt + B^2$. For all real t $P(t) \geq 0$. Hence its discriminant $B^2 - AC$ is negative. This is just (61)

Hence

$$||f+g||_{(2)} \le ||f||_{(2)} + ||g||_{(2)} \tag{62}$$

The space $C([a,b])$ with the norm (60) is not complete space. Its completion is so called space L_2 which plays the fundamental role in Quantum mechanics (Note that the metric (60) is natural generalisation of Pythagorean metrics on the space of functions.)

How to see that it is not a complete space with respect to the norm (60)? The norm (60) measures the "average" distance.

Consider a sequence $\{f_n\}$ of continuous functions such that the limit of this sequence (with respect to the norm (60)) is not continuous function. This will prove that $\{f_n\}$ is Cauchy sequence which has not limit in the space $C([a,b])$. E.g. one can take as $f_n(x)$ a function which is equal to 0 on the interval $\left(a, \frac{a+b}{2} - \frac{1}{n}\right)$, is equal to 1 on the interval $\left(\frac{a+b}{2}, b\right)$ and is linear on the interval $\left(\frac{a+b}{2} - \frac{1}{n}, \frac{a+b}{2}\right)$. Then it is easy to see that $\{f_n(x)\}$ is Cauchy sequence which *does not tend* to continuous function. (Show it!)

4 Linear functionals on normed vector spaces

The linear functional L on the normed space V it is a linear function on vectors with values in real numbers[13]. i.e.

$$\text{for every } \lambda, \mu \in \mathbf{R} \text{ and } f, g \in V \ , L(\lambda f + \mu g) = \lambda L(f) + \mu L(g)\,,, \tag{63}$$

In the case if V is finite-dimensional case then the space of linear functionals is has the same dimension. One can define isomorphism between these spaces (depending on the choice of basis).

It is easy to see that every linear functional is bounded and continuous.

It is not the case if $\dim V = \infty$

In the case of infinite-dimensional case we consider *only bounded functionals.*

4.1 Bounded linear functionals .

A linear functional L on the normed vector space $(V, ||\ ||)$ is called *bounded* if there exists a positive constant C such that for every vector $\mathbf{x} \in V$

$$|L(\mathbf{x})| \le C||\mathbf{x}|| \tag{64}$$

[13] In the case if normed space is defined over field $\mathbf{C}$ of complex numbers then functional takes values in $\mathbf{C}$.

or in other words (more formally):

$$\sup_{||\mathbf{x}||\neq 0} \frac{|L(\mathbf{x})|}{||\mathbf{x}||} < \infty \tag{65}$$

These both equivalent definitions (64) and (65) can be formulated on unit sphere:

A linear functional L on the normed vector space $(V, ||\ ||)$ is called *bounded* if there exist a positive constant C such that for every vector $\mathbf{x} \in V$ with norm 1: $||\mathbf{x}|| = 1$

$$|L(\mathbf{x})| \leq C\,, \tag{66}$$

or in other words (more formally):

$$\sup_{||\mathbf{x}||=1} |L(\mathbf{x})| < \infty \tag{67}$$

It is easy to see that space of linear bounded functionals is linear space. (Little bit not trivial is checking the fact that sum of bounded functionals is bounded)

We denote by V^* the space of linear bounded functionals over normed space $(V, ||\ ||)$

Let $V = C([0,1])$ Banach space of continuous functions with standard norm.

Consider examples and counterexamples:

- $$L(f) = f\left(\frac{1}{2}\right) \qquad \text{(bounded linear functional)}$$
- $$L(f) = \int x f(x) dx \text{ (bounded linear functional)} \tag{68}$$
- $$L(f) = \int g(x) f(x) dx \text{ (bounded linear functional)} \tag{69}$$
- $$L(f) = |f\left(\frac{1}{2}\right)| \qquad \text{(non-linear functional)}$$

- $$L(f) = \int f^2(x)dx \quad \text{(non-linear functional)}$$

- $$L(f) = f'(0) \quad \text{(linear non-bounded functional!)} \tag{70}$$

  The last example is linear non-bounded functional which is defined by the condition above only on the continuous functions which have derivative at the point 0. Prove that it is unbounded: consider its values on the functions $f_w = \sin wx$. It is evident that $||f_w|| = 1$ and $|Lf_w| = |w|.$, $||f_w|| = 1$ and $|Lf_w| = |w| \to \infty$ if $w \to \infty$. Another way to see that this functional is unbounded: Consider its values on polynomials $P_n = x^n$. The norm of these polynomials is equal to 1 and values on functionals on these polynomials tends to infinity.

  We will see below that unboundness of linear functionals makes them discontinuous.

4.2 Norm on the space of bounded functionals

One can consider the following norm on the space V^* of bounded functionals:

$$||L|| = \sup_{||\mathbf{x}|| \neq 0} \frac{|L(\mathbf{x})|}{||\mathbf{x}||}, \tag{71}$$

or in other words:

$$||L|| = \sup_{||\mathbf{x}||=1} |L(\mathbf{x})|. \tag{72}$$

One can easy prove that the relation above is well-defined, i.e. this definition obeys all axioms of norms.

Example 1

The norm of the functional $L(f) = f(\frac{1}{2})$ is equal to 1. Indeed consider arbitrary function $f \in C([0,1])$ with $||f|| = 1$. Then $f(\frac{1}{2}) \leq 1$. Hence $|L(f)| \leq 1$ for $||f|| = 1$. On the other hand $L(f) = 1$ if $f \equiv 1$. Hence $||L|| = \sup_{||f||=1} |L(f)| = 1$.

Examples

Consider examples of calculating norms of functionals related with integrals (70).

Example 2 Calculate the norm of functional $L(f) = \int xf dx$.

We solve a more general problem

Proposition *Suppose g is positive continuous function on the closed interval $[a,b]$. Then the norm of linear bounded functional $L(f) = \int_a^b g(x)f(x)dx$ on the Banach space $C([0,1])$ of continuous functions is equal to the number $c = \int_a^b g(x)dx$.*

Proof:
Use Mean Value Theorem:

$$L(f) = \int_a^b g(x)f(x)dx = f(\xi)\int_a^b g(x)dx\,, \quad \text{where } \xi \in [a,b] \qquad (73)$$

Hence $|L(f)| \leq \int_a^b g(x)dx$. On the other hand it is evident that $|L(f)| = \int_a^b g(x)dx$ in the case if $f \equiv 1$. We come to the conclusion that $||L|| = \int_a^b g(x)dx$

In particular the norm of the functional $L(f) = \int_0^1 xf(x)dx$ is equal to $\int_0^1 xdx = \frac{1}{2}$. ($L(f) = \frac{1}{2}$ if $f \equiv 1$ and $|L(f)| \leq \frac{1}{2}||f||$ for every $f \in C([0,1])$.)

Example 3 Calculate the norm of the functional

$$L_a\colon L_a(f) = \int_{-1}^1 \frac{af(x)dx}{1+a^2x^2}$$

Solution:

$$\int_{-1}^1 \frac{af(x)dx}{1+a^2x^2} = f(\xi)\int_{-1}^1 \frac{dx}{1+a^2x^2} = f(\xi)\int_{-a}^a \frac{du}{1+u^2} = 2f(\xi)\text{arctan}\, a, \quad -1 \leq \xi \leq 1$$

One can see that this functional is bounded and its norm is equal to $||L_a|| = \int_{-a}^a \frac{du}{1+u^2} = 2\text{arctan}\, a$. Indeed for every $f\colon ||f|| \leq 1$, $|L(f)| = |f(\xi)|2\text{arctan}a \leq 2\text{arctan}a$ and moreover $L_a(f) = 2\text{arctan}\, a$ if $f \equiv 1$.

One can show that the functional $L_a(f) \to \pi f(0)$ in the case if $a \to \infty$. To understand why it happens draw the graph of the function $f(x) = \frac{a}{1+a^2x^2}$. You see that it will be the "bell" which becomes narrower when $a \to \infty$ (the width of this "bell" is of order $\frac{1}{a}$). Exact calculations are little bit tricky: Consider

$$L_a(f) = \int_{-1}^1 \frac{af(x)dx}{1+a^2x^2} = \int_{-a}^a \frac{f(u/a)du}{1+u^2} = \int_{-\sqrt{a}}^{\sqrt{a}} \frac{f(u/a)du}{1+u^2} + 2\int_{\sqrt{a}}^a \frac{f(u/a)du}{1+u^2}$$

By Mean value theorem the first integral tends to $\pi f(0)$ because $\int_{-\sqrt{a}}^{\sqrt{a}} \frac{f(u/a)du}{1+u^2} = f(\frac{\xi}{a}) \int_{-\sqrt{a}}^{\sqrt{a}} \frac{du}{1+u^2} \to f(\frac{\xi}{a}) \int_{-\infty}^{\infty} \frac{du}{1+u^2} \to f(0)\pi$. The second integral tends to zero.

Remark We see that for every function $f \in C([0,1])$ $L_a(f) \to \pi f(0)$. But it does not mean that functional L_a tends to functional $L_0(f) = \pi f(0)$ if $a \to infty$ in the standard norm of functionals [14].

Consider a sequence of functions f_n such that $f_n(x) = n|x|$ if $0 \le x \le 1/n$ and $f = 1$ if $|x| \ge \frac{1}{n}$. All functions $f_n(x)$ are continuous and $||f_n|| = 1$. Take any $a > 0$. It is easy to see that $L_a(f_n) \to L_a(f)|_{f=1} = 2\arctan a$ if $n \to \infty$. On the other hand $L_0(f_n) = 0$ for all n. Thus we see that $||L_a - L|| \ge 2\arctan a$. Hence $||L_a - L_0||$ does not tend to zero if $a \to \infty$.

In all examples considered above the function g is positive (see Proposition above).

At what extent is important this condition? To understand it consider the following example

Example 4 Calculate the norm of the functional $L(f) = \int_{-1}^{1} xf(x)dx$ on the space $C([-1,1])$. Here we cannot use the proposition above. (Note that $\int_{-1}^{1} xdx = 0$.) Nevertheless it is evident that $|\int_{-1}^{1} xf(x)dx| \le \int_{-1}^{1} |x||f(x)|dx = 2\int_0^1 x|f(x)|dx$ is bounded and $|L(f)| = |\int_{-1}^{1} xf(x)dx| \le 2\int_0^1 x|f(x)|dx \le 2\int_0^1 xdx = 1$ if $||f|| = 1$. One can see that $L(f) = 0$ for $f \equiv 1$ but norm of this functional still is equal to 1. To show it consider the sequence of functions $\{f_n(x)\}$ where the $f_n(x)$ is defined in the following way:

$$f(x) = \begin{cases} 1 & \text{if } x > 1/n \\ nx & \text{if } -1/n \le x \le 1/n \\ -1 & \text{if } x < 1/n \end{cases} \tag{74}$$

This sequence of functions tends to discontinuous function. On the other hand $\lim_{n\to\infty} L(f_n) = 1$. Hence $||L|| = 1$ in spite of the fact that in this example the maximum value is not attained on any continuous function.

4.3 Bounded functionals are continuous

It turns out that normed space of functionals is Banach space:

[14] It is like uniform and point-wise convergense: if sequence of functions tend to the function point-wise it does not mean that this sequence has to tend uniformly!

Theorem

The vector space V^* of bounded functionals on arbitrary normed vector space with standard norm (72) *is a Banach space (even in the case if normed vector* $(V, ||\cdot||)$*is not complete)*

Proof is based on the Lemma (52).

Let $\{L_n\}$ be a fundamental sequence of bounded linear functionals on the normed vector space $(V, ||\cdot||)$. It means that for every $\varepsilon > 0$ there exists N such that for every $m, n > N$ $||L_n - L_m|| = \sup_{||\mathbf{x}||=1} |L_n(\mathbf{x}) - L_m(\mathbf{x})| < \varepsilon$. Consider $L(\mathbf{x}) = \lim_{n\to\infty} L_n(\mathbf{x})$. Using arguments similar to that of Lemma (52) one can see that this relation defines bounded linear operator on V such that $||L_n - L|| \to 0$

It is very instructive to apply this Theorem to the space of linear bounded functionals on non-complete normed space, e.g. space of smooth functions, or space of polynomials

Example Consider the normed space of bounded functionals on the space of polynomials in $C([a, b])$ (i.e. we consider bounded functionals not on all continuous functions but only on polynomials) The space of polynomials is not complete: the Weierstraß Approximation Theorem tells us that this space is dense in the space $C([a, b])$ of all continuous functions. (See the Subsection: Another glance on Weierstraß Theorem)

Nevertheless the space of bounded functionals on the space of polynomials is complete normed space, i.e. this space is Banach space. Moreover this space of functionals is isomorphic to the space of functionals on the space $C([a, b])$ of all continuous functions on $[a, b]$. Every bounded functional L defined on polynomials can be defined by continuity on continuous function: Let $f \in C([a, b])$ and $\{P_n\}$ be a sequence of polynomials which tends uniformly to f. Consider the sequence $\{L(P_n)\}$. If $c = ||L||$ is a norm of bounded functional L then $|L(P_n)| \leq c||P_n||$ and $|L(P_n) - L(P_m)| \leq c||P_n - P_m||$. This implies that $|L(P_n)|$ is a Cauchy (fundamental) sequence of numbers. Hence this sequence has a limit. One can define the value of L on continuous functions f as a limit:

$$L(f) \colon = \lim_{n\to\infty} L(P_n) \qquad (P_n \to f) \tag{75}$$

If $\{\tilde{P}_n\}$ is a sequence of another polynomials which tend to the same function f then it is easy to see that both sequences $\{L(P_n)\}, \{L(\tilde{P}_n)\}$ have the same limit.

One can see that the map (75) indeed establishes the isomorphism between the bounded functionals on $C([a,b])$ and bounded functionals on the subspace of polynomials. Indeed the map (75) is a linear map. If $L \neq L'$ on polynomials then they are not equal on $C([a,b])$ too. Every bounded functional on $C([a,b])$ can be uniquely defined by its values on polynomials: if two functionals L, L' have the same values on polynomials then by (75) they are the same on arbitrary continuous function.

At what extent is important the condition that functionals are bounded? We used it in our proof when we proved that sequence $L(P_n)$ is a fundamental sequence of numbers. Is it really necessary? Yes, See an example (76) after the next Theorem.

Remark One can consider instead a space of polynomials a space of smooth functions on the closed interval $[a,b]$. Space of smooth functions as well as space of polynomials is dense in the space of continuous functions (see subsection 3.5). By the same reasoning a space of bounded linear functionals on smooth functions is isomorphic to space of bounded linear functionals on continuous functions.

Now formulate a simple but important Theorem:

Theorem

A linear functional L on the normed vector space is bounded iff it is continuous.

This theorem states that condition for functional to be bounded and to be continuous are equivalent.

Proof Suppose L is bounded functional on V: for every f, $|L(f)| < M||f||$. The continuity follows from the fact that $|L(g)-L(f)| = |L(g-f)| < M\delta$ in the case if $||g-f|| < \delta$. Inverse implication: Suppose L is continuous. Consider the continuity condition at the point $f = 0$. For every $\varepsilon > 0$ there exists δ such that $|L(f)| < \varepsilon$ if $||f|| < \delta$. Then for $f\colon ||f|| = \frac{\delta}{2}\ L||(f)|| < \varepsilon$. Hence $|L(f)| < \frac{2\varepsilon}{\delta}||f||$. Hence $||L|| < \frac{2\varepsilon}{\delta}$.

Note that this Theorem is trivial for finite-dimensional case: If L is linear functional on n-dimensional space then it is evidently linear and continuous. The story becomes interesting for infinite-dimensional cases. It is like for norms: all norms were equivalent in finite-dimensional normed spaces and examples of non-equivalent norms exist only in infinite-dimensional normed spaces. (See the subsection Equivalence of norms in the second Section)

Example In the analysis of the previous Theorem we see that every bounded functional on polynomials (which are dense in the space of continuous functions on $[a, b]$) can be defined by continuity on the whole space of continuous functions. The condition that polynomial is bounded is crucial: E.g. consider unbounded functional on the polynomials: e.g. the following functional:

$$L(P) = P'(1) \tag{76}$$

The value of this functional on the polynomial $L = x^n$ is equal to nx^{n-1}. We see that L is unbounded. One can define a functional by the formula above on functions which have derivative at the point 1. But is it possible to define a continuous functional on the space $C([-1, 1])$ of *all* continuous functions such that it is equal to functional (76)?

No. According the Theorem we see that one *cannot define a continuous functional* on the space $C[-1, 1]$ of all continuous functions such that it is equal to (76): continuous functional has to be bounded, but the functional (76) is unbounded on the subspace of polynomials[15] on polynomials.

4.4 Spaces l_1^*, l_2^*

In finite-dimensional case a dual space is isomorphic to a linear space. Situation is different in infinite-dimensional case.

Study spaces of linear bounded functionals for Banach spaces l_1, l_2.

1. Case of l_1. Show that $l_1^* = l_\infty$.

Consider arbitrary sequence $\mathbf{y} = \{y_n\} \in l_\infty$: $\mathbf{y} = \{y_1, y_2, \dots\}$ such that $\sup_n |y|_n < \infty$.

Assign to this sequence the functional $\mathbf{L_y}$ such that its value on an every sequence $\mathbf{x} = \{y_n\} \in l_1$ is given by scalar product:

$$\mathbf{y} \mapsto \mathbf{L}_y: \ \mathbf{L}_y(\mathbf{x}) = \sum y_n x_n$$

It is easy to see that this functional is well-defined linear bounded functional:

$$\text{If } |y_n| \le M, \text{ then } \quad |\mathbf{L}_y(\mathbf{x})| \le M \sum |x_n| = M ||\mathbf{x}||_1, \ i.e. \tag{77}$$

$$||\mathbf{L}_y|| \le ||\mathbf{y}||_\infty$$

[15] What about to construct unbounded discontinuous functional on all continuous functions which coincides with (76) on polynomials? It is curious to note that we have to use Choice Axiom for this purpose

We see that every element of l_∞ defines linear bounded functional on l_1: we constructed the linear map, injection from l_∞ to l_1.

Show that this map is one-one correspondence, i.e. every bounded functional $L \in l_1^*$ defines an element $\mathbf{y}$ such that $L = \mathbf{L_y}$

Denote by $\mathbf{e}_n$ the sequence such that its n-th term is equal to 1 and all other terms are equal to zero and assign to every linear bounded functional $L \in l_1^*$ the sequence $\mathbf{y}$ such that $k - th$ term of this sequence is equal to the value of the functional L on the sequence $\mathbf{e}_k$:

$$\mathbf{y}: \quad y_k = L(\mathbf{e}_k)$$

The norm of the sequence $\mathbf{e}_n$ is equal to 1: $||\mathbf{e}_n||_1 = 1$. Hence if norm of functional L is equal to M, then $|y_k| \leq ||L|| = M$. Hence $\mathbf{y} = \{y_n\} \in l_\infty$. It is easy to see that $L = \mathbf{L_y}$.

Thus we established isomorphism between l_1^* and l_∞

Analogously one can show that l_2^* is just isomorphic to itself: $l_2^* = l_2$

Consider arbitrary sequence $\mathbf{y} = \{y_n\} \in l_2$: $\mathbf{y} = \{y_1, y_2, \dots\}$ such that $\sum_n y_n^2 < \infty$.

Assign to this sequence the functional $\mathbf{L_y}$ such that its value on an every sequence $\mathbf{x} = \{y_n\} \in l_1$ is given by scalar product:

$$\mathbf{y} \mapsto \mathbf{L}_y: \ \mathbf{L}_y(\mathbf{x}) = \sum y_n x_n$$

It is easy to see using Cauchy-Bunyakovsky inequality this functional is well-defined linear bounded functional.

$$|\mathbf{L}_y(\mathbf{x})|^2 \leq \left(\sum_k y_k^2\right)\left(\sum_n x_n^2\right), \ i.e. \tag{78}$$

$$|\mathbf{L}_y(\mathbf{x})| \leq ||\mathbf{x}||_2 ||\mathbf{y}||_2 \Rightarrow ||\mathbf{L}_y|| \leq ||\mathbf{y}||_2$$

Every element of l_2 defines linear bounded functional on l_2. Assigning to every bounded functional $L \in l_2^*$ a sequence $\mathbf{y}$: $\quad y_k = L(\mathbf{e}_k)$ we see that $L = \mathbf{L_y}$. Thus we established isomorphism between l_2^* and l_2

In the next section we will see that Banach space l_2 is Hilbert space. It is why its dual coincides with it.

5 Hilbert spaces

Hilbert spaces it is Banach spaces where we can consider not only the length (norm) of the functions also the angle between functions because the norm in Hilbert space is defined via scalar product.

Remember that in finite-dimensional case Euclidean space $\mathbf{R}^n$ is distinguished by having an inner (scalar) product $\langle , \rangle$ defined in the way such that the norm $||x|| = \sqrt{\langle x, x\rangle}$, $\langle x, y\rangle = x^1y^1 + \dots x^ny^n$.

We study now vector spaces equipped with inner product then consider norms generated by this inner product.

Remark Before we consider real normed spaces. We did not consider complex vector spaces, i.e. vector spaces over $\mathbf{C}$. If $\mathbf{u}_1, \dots, \mathbf{u}_n$ are elements of vector space then their linear combination is λ^1 Elements of complex vector space are linear combinations with

5.1 Vector spaces with scalar product

We say that vector space V is equipped with scalar (inner) product $\langle x, x\rangle$ if $\langle x, x\rangle$ defines positively defined Hermitian form on V, i.e.:

-
$$\langle \lambda y + \mu z, x\rangle = \lambda\langle y, x\rangle + \mu\langle z, x\rangle\,, \text{(linearity on the first argument)} \quad (79)$$

- $\langle x, y\rangle$ is a complex number conjugated to the complex number $\langle y, x\rangle$;
$$\langle x, y\rangle = \overline{\langle y, x\rangle} \quad (80)$$

-
$$\langle x, \lambda y + \mu z\rangle = \bar\lambda\langle x, y\rangle + \bar\mu\langle x, z\rangle\,, \text{(antilinearity on the second argument)} \quad (81)$$

  The real number $\langle x, x\rangle$ is not negative:
$$\langle x, x\rangle \geq 0 \text{ and } \langle x, x\rangle = 0 \text{ iff } x = 0\,. \quad (82)$$

Note that condition (81) follows from (79) and (80).The fact that $\langle x, x\rangle$ is a real number follows from the condition (80):

In the case of real spaces (vector space over real numbers) the conditions above are simplified: $\langle\, ,\, \rangle$ defines the usual symmetric positive form:

-
$$\langle x, y\rangle = \langle y, x\rangle \text{ is a real number,} \quad (83)$$
$$\langle x, \lambda y + \mu z\rangle = \langle \lambda y + \mu z, x\rangle = \lambda\langle x, y\rangle + \mu\langle x, z\rangle\,, \quad (84)$$

  (linearity condition)

- $$\langle x, x\rangle \geq 0\,,\ \langle x, x\rangle = 0 \iff x = 0\,. \tag{85}$$

One can define a norm in a vector space with scalar product in the following way:

$$||x|| = \sqrt{\langle x, x\rangle} \tag{86}$$

Prove that this equation indeed defines a norm.

It is evident from (81) that r.h.s of (86) is well-defined. The condition (30) ($||\lambda \mathbf{v}|| = |\lambda|||\mathbf{v}||$) follows from (79) and (81).

It remains to prove triangle inequality (31). Take arbitrary $x, y \in V$. According to well-known Cauchy-Bunyakovsky identity (see also the next subsection)

$$\big|\langle x, y\rangle\big| \leq ||x||||y||\,. \tag{87}$$

Hence

$$||x+y|| = \sqrt{\langle x+y, x+y\rangle} \leq \sqrt{||x||^2 + ||y||^2 + 2||x||||y||} = ||x|| + ||y||$$

because $|\langle x, y\rangle + \langle y, x\rangle| = 2|\mathrm{Re}\,\langle x, y\rangle| \leq 2|\langle x, y\rangle|$ The Cauchy-Bunyakovsky inequality plays the crucial role in Hilbert spaces. It is not worthless to look once more on the proof of this identity (especially for complex case.).

5.2 Proof of Cauchy-Bunyakovsky inequality (complex case)

We formulated proved and used this inequality many times. Now formulate and prove it for complex case. *For every two elements (vectors)* $\mathbf{x}$, $\mathbf{y}$ *in complex vector space* V *with scalar product* $\langle\,,\,\rangle$

$$|\langle \mathbf{x}, \mathbf{y}\rangle|^2 \leq \langle \mathbf{x}, \mathbf{x}\rangle \cdot \langle \mathbf{y}, \mathbf{y}\rangle, \text{ i.e. } \langle \mathbf{x}, \mathbf{y}\rangle \cdot \langle \mathbf{y}, \mathbf{x}\rangle \leq \langle \mathbf{x}, \mathbf{x}\rangle \cdot \langle \mathbf{y}, \mathbf{y}\rangle \tag{88}$$

" $\leq$ " *becomes* " $=$ " *if and only if* $\mathbf{x}$ *and* $\mathbf{y}$ *are collinear:* $\mathbf{x} = \lambda \mathbf{y}$

To prove this inequality consider complex number μ such that $|\mu| = 1$ and $\mu\langle \mathbf{x}, \mathbf{y}\rangle$ is real. (If $\langle \mathbf{x}, \mathbf{y}\rangle = a + bi = \rho e^{i\varphi}$ we take $\mu = e^{-i\varphi}$).

Now consider the following function $P(t) =< t\mu\mathbf{x} - \mathbf{y}, t\mu\mathbf{x} - \mathbf{y} >$ of real variable t. This is the square of the length of the vector $t\mu\mathbf{x} - \mathbf{y}$: $P(t) = ||t\mu\mathbf{x} - \mathbf{y}||^2$. This means that for every t, $P(t) \geq 0$ and $P(t)$ reaches zero iff vectors $\mathbf{x}$ and $\mathbf{y}$ are colinear. On the other hand $P(t)$ is quadratic polynomial:

$$P(t) = \bar{\mu}\mu < \mathbf{x}, \mathbf{x} > t^2 - (\mu\langle \mathbf{x}, \mathbf{y}\rangle + \bar{\mu}\langle \mathbf{y}, \mathbf{x}\rangle)\, t + < \mathbf{y}, \mathbf{y} >$$

Denote by $A =< \mathbf{x}, \mathbf{x} >$, $C =< \mathbf{y}, \mathbf{y} >$ and $B = \mu < \mathbf{x}, \mathbf{y} >$. Numbers A, B, C are real numbers. In particular B is equal to the modulus of the complex number $< \mathbf{x}, \mathbf{y} >$: $B = \mu < \mathbf{x}, \mathbf{y} >= e^{-i\varphi} < \mathbf{x}, \mathbf{y} >= | < \mathbf{x}, \mathbf{y} > |$. Polynomial $P(t) = At^2 - 2Bt + C$ has at most one root if $\mathbf{x}$ and $\mathbf{y}$ are colinear. Hence $B^2 \leq AC$ and $B^2 = AC$ if vectors are colinear. This is just inequality (88) required ■

5.3 Hilbert space. Definition

Definition

Let H be vector space equipped with scalar product. (Scalar product is Hermitian positively defined form, i.e. form on two vectors obeying conditions (79),(80), (81),(82)) (see subsection 5.1 "Vector spaces with scalar product")

The vector space equipped with scalar product $(H, \langle\, , \,\rangle)$ is called *Hilbert* space if normed space $(H, || \cdot ||)$ with a norm $|| \cdot ||$ generated by the scalar product $\langle\, , \,\rangle$ ($||f|| = \sqrt{\langle f, f \rangle}$) is Banach space (complete space), i.e. every fundamental sequence of vectors has limit in this space.

Examples

.

Example 1 (trivial) Finite-dimensional space $\mathbf{R}^n$ with scalar product $\langle \mathbf{x}\mathbf{y} \rangle = x^1 y^1 + \cdots + x^n y^n$ is trivial example of inner product space which is Hilbert space because it is evidently Banach space.

Example 2 Non-trivial example of Hilbert space: it is l_2 with scalar product

$$\langle \mathbf{x}\mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n + \dots \tag{89}$$

This inner product defines just the norm $||\mathbf{x}||_2 = \sqrt{x_1^2 + \cdots + x_n^2 + \dots}$. We come to Banach space $(l_2, || \cdot ||_2)$.

Remark If we consider Banach spaces $(l_p, || \cdot ||_p)$ $(p \geq 1)$ then it is only $(l_2, || \cdot ||_2)$ will be Hilbert space: norm in $(l_2, || \cdot ||_2)$ is generated by scalar product.

The advantage of Hilbert space is that we can measure not only length of the vectors but also "angles" using scalar product (see below).

Remark How to prove that norm in l_p is not generated by scalar product if $p \neq 2$? For this purpose solve the following exercise:

Exercise: If norm is $|| \, ||$ generated by scalar product $\langle \, \rangle$ then it satisfies so called parallelogram identity:

$$2||\mathbf{x}||^2 + 2||\mathbf{y}||^2 = ||\mathbf{x} + \mathbf{y}||^2 + ||\mathbf{x} - \mathbf{y}||^2 \tag{90}$$

This can be proved by simple checking [16].

It follows from this exercise the answer on question: it is easy to check that norm in l_p obeys parallelogram identity iff $p = 2$.

Example 3

Now we consider one of the very important examples of Hilbert space.

Consider the space $C([a, b])$ of continuous functions with a norm $||f|| = \max_x |f(x)|$. We know that it is Banach space. Using (90) one can see that this norm is not generated by any scalar product.

On the space $C([a, b])$ of continuous functions consider the scalar (inner) product

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx \tag{91}$$

One can prove that this relation indeed defines inner product.

Before we considered $C([a, b])$ with norm $||f|| = \max_x |f(x)|$ as a Banach space and functions as vectors in it. Now this inner (scalar) product leads us to different norm:

$$||f||' = \sqrt{\left(\int_a^b f^2(x)dx\right)}$$

The space of continuous functions with this norm defined by the scalar (inner) product(91) has disadvantage with respect to the Banach space $C([a, b])$. It is not complete! (We studied this norm constructing normed not completed vector spaces: see (60) above).

We note that on the other hand the vector space of functions with norm generated by scalar product (91) has an essential advantage: we can consider not only distance between functions, but an "angle" between functions too. E.g. functions are orthogonal each other if $\int fg dx = 0$. Another example: the cosines of the angle between functions $f = 1$ and $g = x$ in the space $C([0, 1])$ is equal to the

$$\cos\alpha = \frac{\langle f, g \rangle}{||f||||g||} = \frac{\langle 1, x \rangle}{||1||||x||} = \frac{\int_0^1 x dx}{\sqrt{\int_0^1 1^2 \cdot dx}\sqrt{\int_0^1 x^2 dx}} =$$

[16] One can prove that inverse implication is true too: $B(x, y)$ such that $4B(x, y) = ||x + y||^2 - ||x - y||$ is bilinear form which defines scalar product.

$$\frac{\frac{1}{2}}{\sqrt{1}\sqrt{\frac{1}{3}}} = \frac{1}{2\sqrt{3}}\,. \qquad (92)$$

Now return to the problem that the space of continuous functions with norm generated by scalar product is not complete.

The completion of the space of continuous functions $C([a,b])$ with respect of inner product (91) is the space $L_2([a,b])$. It is Hilbert space of equivalence classes of Lebesgue measurable real functions on the segment $[a,b]$ such that $\int f^2 dx < \infty$ with scalar product (91). Two measurable functions f, g are considered equal (belong to the same equivalence class) if Lebesgue integral $\int (f-g)^2 dx = 0$[17].

Example Consider the function f such that it is equal to zero if $x \in [0,1/2]$ and it is equal to 1 if $x \in (1/2,1]$. This is measurable square integrable function. It is not continuous function, but one can easy to find a fundamental sequence of continuous functions f_n which tends to the function f in the norm generated by (91). The equivalence class of the function f it is the set $[f]$ of all measurable square integrable functions which obey the condition:

$$[f] = \{\tilde{f}\colon \quad \int_0^1 (\tilde{f} - f)^2 dx = 0\} \qquad (93)$$

E.g. consider the function g such that it is equal to zero if $x \in [0,1/2)$ and it is equal to 1 if $x \in [1/2,1]$ It differs from f at the point $1/2$. But it belong to equivalence class $[f]$. We do not distinguish functions g, f in $L_2[0,1]$.

In our further considerations in this subsection sometimes we consider the inner product space of continuous functions $C([a,b])$ instead its completion the space $L_2([a,b])$

Hilbert space $L_2[a,b]$ is isomorphic to l_2 (See the subsection: "Isomorphism of Hilbert spaces")

[17]The function f is called Lebesgue measurable if for every a the sets $R_a = \{x\colon f(x) < a\}$ and sets $R_a = \{x\colon f(x) > a\}$ are meausrable. Lebesgue measurable functions f and g are considered as equivalent if $f(x) = g(x)$ except of a set of points of measure zero [18]. Very important Lusin's Theorem states that *Let f be a real measurable function on the closed interval $[a,b]$. Then given $\varepsilon > 0$ there exists a continuous function g such that $f(x) = g(x)$ except on a set of points of measure less than ε.* It follows from Lusin's Theorem that Hilbert space $L^2([a,b])$ is indeed completion of the space of continuous functions with respect to the norm induced by the inner product (91).

5.4 Topological basis in Hilbert space

We give first definition of topological basis in arbitrary Banach space V.

Suppose the Banach space V contains the countable set

$$\{\mathbf{e}_n\} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k, \dots\} \tag{94}$$

of vectors such that for every vector $\mathbf{x} \in V$ there exists a unique set of coefficients

$$\{\lambda_n\} = \{\lambda_1, \lambda_2, \dots, \lambda_k, \dots\}$$

such that

$$\mathbf{x} = \sum_{k=1}^{\infty} \lambda^k \mathbf{e}_k \,. \tag{95}$$

The last relation means that

$$\mathbf{x} = \lim_{N\to\infty} \sum_{k=1}^{N} \lambda^k \mathbf{e}_k \,, \text{ i.e. } \lim_{N\to\infty} ||\mathbf{x} - \left(\sum_{k=1}^{N} \lambda^k \mathbf{e}_k\right) || = 0$$

In this case we say that the set $\{\mathbf{e}_n\}$ is a topological basis in V [19].

Remark We need the conception of topological basis in infinite-dimensional case because we deal with infinite sums. Note that we define the sum w.r.t.the norm, topology.

Example Consider in the Banach space l_2 the set $\{\mathbf{e}_n\}$ of vectors where vector $\mathbf{e}_n$ is following sequence: its all terms except the n-th term are equal to 0 and n-th term is equal to 1:

$$\mathbf{e}_n = \underbrace{0, \dots, 0, 1, 0, \dots, 0 \dots}_{\text{the 1 on the } n\text{-th place}} \tag{96}$$

It is easy to see that thus defined $\{\mathbf{e}_n\}$ is the basis: for every sequence $\mathbf{x} = \{x_1, x_2, \dots, x_k, \dots\}$ $\mathbf{x} = \sum_{k=1}^{\infty} x_k \mathbf{e}_k$. Note that the sum of infinite number of terms has sense only with respect to the norm $|| \quad ||_2$:

$$\mathbf{x} = \sum_{k=1}^{\infty} x_k \mathbf{e}_k \text{ means that } ||\mathbf{x} - \sum_{k=1}^{N} x_k \mathbf{e}_k||_2 \to 0 \text{ if } N \to \infty,$$

Now return to Hilbert space.

[19] We do not consider a case when topological basis contains incountable number of vectors

5.5 Orthonormal system and orthonormal basis

In Hilbert space we have concept of orthogonality. So why not to try to consider orthonormal basis.

Definition The system $\{\mathbf{e}_n\}$ of vectors in Hilbert space H is called orthonormal system if all vectors have length 1 and they are orthogonal to each other:

$$\{\mathbf{e}_1, \dots, \mathbf{e}_n, \dots, \}: \quad \langle \mathbf{e}_m, \mathbf{e}_n \rangle = \delta_{mn} = \begin{cases} 0, & \text{if} \quad m \neq n \\ 1, & \text{if} \quad m = n \end{cases} \tag{97}$$

It is evident that in finite-dimensional Hilbert space of dimension n $\mathbf{R}^n$ the orthonormal system containing n elements is a basis.

What to do in infinite-dimensional case? We can consider the orthonormal system $S = \{\mathbf{e}_1, \dots, \mathbf{e}_n, \dots, \}$ containing infinite number of vectors. But there are different infinities, infinity minus N is again infinity. Infinity divided by two is again infinity. E.g.if S if orthonormal system containing infinite number of vectors then the following systems

$$S' = \{\mathbf{e}_5, \mathbf{e}_6, \mathbf{e}_7, \dots, \mathbf{e}_n, \dots, \} \quad \tilde{S} = \{\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_5, \dots, \mathbf{e}_{2n+1}, \dots, \}$$

are orthonormal and they contain infinite number of vectors. But it is evident that $S', \tilde{S}$ are not bases. E.g. the vector $\mathbf{x} = \mathbf{e}_2$ is orthogonal to all the vectors from $\tilde{S}$ and S'.

Hence condition of orthonormality is not sufficient. We have to put also condition that this system is topological basis.

Definition

A system $\{\mathbf{e}_n\}$ of vectors in Hilbert space $(H, \langle\ \rangle$ $(n = 1, 2, 3, \dots)$ is called *orthonormal basis* if

- this system is orthonoromal $\langle \mathbf{e}_n, \mathbf{e}_k \rangle = \delta_{nk}$ (condition (97))
- this system is a topological basis, i.e. for every vector $\mathbf{x} \in V$ there exist a unique set of coefficients $\{\lambda_n\} = \{\lambda_1, \lambda_2, \dots, \lambda_k, \dots\}$ such that
$$\mathbf{x} = \sum_{k=1}^{\infty} \lambda^k \mathbf{e}_k\,. \tag{98}$$

The last relation means that

$$\mathbf{x} = \lim_{N\to\infty} \sum_{k=1}^{N} \lambda^k \mathbf{e}_k\,, \text{ i.e. } \lim_{N\to\infty} ||\mathbf{x} - \left(\sum_{k=1}^{N} \lambda^k \mathbf{e}_k\right)|| = 0$$

Later we will show that if $\{\mathbf{e}\}_{\mathbf{n}}$ is an orthonormal basis then coefficients γ^k of expansion are equal to

$$\gamma^k = \langle \mathbf{x}, \mathbf{e_k} \rangle \tag{99}$$

(See equation (112) in the next subsection)

Remark Note that condition of uniqueness can be omitted (orthogonal vectors are linear independend): If $\mathbf{x} = \sum_{k=1}^{\infty} \lambda^k \mathbf{e}_k = \sum_{k=1}^{\infty} \tilde{\lambda}^k \mathbf{e}_k$ then $\sum_{k=1}^{\infty} (\lambda^k - \tilde{\lambda}_k)\mathbf{e}_k = 0$ and one can see that $\sum_{k=1}^{\infty} |\lambda^k - \tilde{\lambda}_k|^2 = 0$, i.e. $\lambda_i = \tilde{\lambda}_i$.

Remark We consider only Hilbert spaces which have countable topological basis , see also subsection "Isomorphism of Hilbert spaces")

Remark Of course not every orthonormal system is a basis (see example above. Interesting examples see in the end of the subsection "Important calculations")

Now very improtant example of canonical orthonormal basis:

Example Consider in Hilbert space l_2, with scalar product (89) ($\mathbf{x}, \mathbf{y} = \sum x_k y_k$) the following system of vectors;

$$\mathbf{e}_1 = (1, 0, 0, \dots,),\ \mathbf{e}_2 = (0, 1, 0, \dots,),\ \mathbf{e}_3 = (0, 0, 1, \dots,), \dots \tag{100}$$

It is easy to see that it is orthonormal system and this system is a basis: every vector $\mathbf{x} \in l_2$, $\mathbf{x} = (x_1, \dots, x_n)$ is equal to

$$\mathbf{x} = \sum_{k=1}^{\infty} x_k \mathbf{e}_k \tag{101}$$

5.6 Orthonormal systems and orthonormal bases, complete system, closed system. Bessel inequality and Parseval equality

We perform all claculations for real Hilbert spaces. They can be easily generalised for complex case.

Let $\mathbf{x}$ be a vector in Euclidean space (finite-dimensional Hilbert space) and $\mathbf{e}_n$ be an orthonormal basis. If

$$\mathbf{x} = \sum_{m=1}^{n} x^m \mathbf{e}_m \tag{102}$$

then to calculate coefficients x^k one have to take scalar product of the equation above on the basic vector:

$$\langle \mathbf{x}, \mathbf{e_k} \rangle = \langle \sum_{m=1}^{n} x^m \mathbf{e}_m, \mathbf{e}_k \rangle = \sum_{m=1}^{n} x^m \langle \mathbf{e}_m, \mathbf{e}_k \rangle = x^k \tag{103}$$

It is well-known formula: k-th component of the vector is equal to the projetion of vector on the basic vector $\mathbf{e}_k$ and it is equal to the sclarar product of the vector on the basis vector $\mathbf{e}_k$.

Consider infinite-dimensional case. Let $\{\mathbf{e}_n\}$ $(n = 1, 2, 3, \dots)$ be an arbitrary orthonormal system in Hilbert space $\mathcal{H}$, not necessarily orthonormal basis! For arbitrary vector $\mathbf{x} \in \mathcal{H}$ consider its projections on vectors $\mathbf{e}_k$:

The projection on the plane H_N spanned by the vectors $\{\mathbf{e_1}, \mathbf{e_2}, \dots, \mathbf{e_N}\}$ is a vector

$$\mathbf{x}_N = \sum_{k=1}^{N} \alpha_k \mathbf{e_k}, \ \ \alpha_k = \langle \mathbf{x}, \mathbf{e}_k \rangle \tag{104}$$

$\mathbf{x}_N$ is projection of vector $\mathbf{x}$ on the plane H_N spanned by the first vectors $\{\mathbf{e_1}, \mathbf{e_2}, \dots, \mathbf{e_N}\}$. Indeed if $\mathbf{x} = \mathbf{x}_N + \mathbf{h}_N$ then vector $\mathbf{h}_N$ is orthogonal to every vector $\mathbf{e}_i$ for every $i, 1 \leq i \leq N$: $\langle \mathbf{h}_N, \mathbf{e}_i \rangle = \langle f - f_N, \mathbf{e}_i \rangle = \langle \mathbf{x} - \sum_{k=1}^{N} \alpha_k \mathbf{e_k}, \mathbf{e}_i \rangle = \alpha_i - \alpha_i = 0$. Hence $\mathbf{h}_N$ is orthogonal to $\mathbf{x}_N$.

(It is useful to view vectors $\mathbf{x}, \mathbf{x}_N$ and $\mathbf{h}_N$ as sides of triangle)

This means that length of the vector $\mathbf{x}_N$ is less or equal to the length of the vector $\mathbf{x}$.

$$||\mathbf{x}||^2 = \langle \mathbf{x}_N + \mathbf{h}_N, \mathbf{x}_N + \mathbf{h}_N \rangle \geq \langle \mathbf{x}_N + \mathbf{h}_N, \mathbf{x}_N + \mathbf{h}_N \rangle = ||\mathbf{x}_N|| \tag{105}$$

Note that $||\mathbf{x}_N|| = \langle \sum_{k=1}^{N} \alpha_k \mathbf{e_k}, \sum_{k=1}^{N} \alpha_k \mathbf{e_k} \rangle = \sum_{k=1}^{N} |\alpha_k|^2$. We come to fundamental inequality:

Let $\mathbf{e}_N$ be a orthonormal system in Hilbert space $\mathcal{H}$. Then for every vector $\mathbf{x} \in H$ the following inequality holds:

$$\sum_{k=1}^{N} |\alpha_k|^2 \leq ||\mathbf{x}_N||^2, \text{ where } a_k = \langle \mathbf{x}, \mathbf{e}_k \rangle \tag{106}$$

Taking limit $N \to \infty$ we come to

$$\sum_{k=1}^{\infty} |\alpha_k|^2 \leq ||\mathbf{x}||^2 \tag{107}$$

This inequality is called Bessel inequality

Geometrical meaning of Bessel inequality is simple: projection of vector has length less or equal than the vector. In finite-dimensional case the Bessel inequality becomes equality if N is dimension of the space. We come to

Definition An orthonormal system $\{\mathbf{e}_n\}$ of vectors in Hilbert space $(H, \langle\ \rangle$ is called *closed* if for every vector $\mathbf{x} \in H$ the Bessel inequality (107) becomes equality:

$$||\mathbf{x}||^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \sum_{k=1}^{\infty} |\alpha_k|^2, \text{ where } a_k = \langle \mathbf{x}, \mathbf{e}_k \rangle, \tag{108}$$

i.e. square of its length is equal to the sum of the squares of its projections on basic vectors (Pythagoras Theorem is obeyed).

The relation (108) is called *Parseval equality.* Parseval equality is nothing but Pythagorean Theorem for infinite-dimensional case!

$$\text{if for arbitrary } n\ \langle \mathbf{y}, \mathbf{y} \rangle = \sum \alpha_n^2, \text{ where } \alpha_n = \sum \langle \mathbf{y}, \mathbf{e}_n \rangle^2 \tag{109}$$

Definition An orthonormal system $\{\mathbf{e}_n\}$ of vectors in Hilbert space $(H, \langle\ \rangle$ is called *complete* if every vector $\mathbf{y} \in H$ which is orthogonal to all vectors from this system is equal to zero:

$$\text{if for arbitrary } n\ \langle \mathbf{y}, \mathbf{e}_n \rangle = 0 \Rightarrow \mathbf{y} = 0\,. \tag{110}$$

It is obvious that in finite-dimensional case conditions that system are a) topological basis b) closed c) complete coincide.

It is true in Hilbert space too:

Theorem.

The following conditions on orthonormal system $\{\mathbf{e_n}\}$ in Hilbert space $\mathcal{H}$
a) it is topological basis
b) it is closed
c) it is complete
are equivalent

Give a short proof of this Theorem.

1. a)$\Rightarrow$ b)

Let an orthonormal system $\mathbf{e}_n$ is topological basis, i.e. for every vector $\mathbf{x}$ $\mathbf{x} = \sum_{k=1}^{\infty} \gamma^k \mathbf{e}_k$. First of all show that all $\gamma_k = \alpha_k = \langle \mathbf{x}, \mathbf{e}_k \rangle$

In finite-dimensional case it is obvious: $\alpha_r = \langle \mathbf{x}, \mathbf{e}_r \rangle = \langle \sum_{k=1}^{n} \gamma_k \mathbf{e}_k, \mathbf{e}_r \rangle = \sum_{k=1}^{n} \gamma_k \langle \mathbf{e}_k, \mathbf{e}_r \rangle = \sum_{k=1}^{n} \gamma_k \delta_{kr} = \gamma_r$.

In infinite-dimensional case we can perform calculations above but we will have delicate problems with infinite sums

Lemma Continuity of scalar product

The function

$$F(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle \tag{111}$$

is continuous function of variable $\mathbf{x}$ *and of variable* $\mathbf{y}$ *with respect to the norm induced by the scalar product.*

This lemma follows immediately from Cauchy-Bunyakovsky inequality.

We use this lemma for dealing with calculations in infinite-dimensional case.

Prove that coefficients $\gamma_r = \alpha_r$ for infinite-dimensional case:

$$\alpha_r = \langle \mathbf{e}_r, \mathbf{x} \rangle = \langle \mathbf{e}_r, \sum_{k=1}^{\infty} \gamma_k \mathbf{e}_k \rangle = \tag{112}$$

$$\text{(by definition of infinite sums)} = \langle \mathbf{e}_r, \lim_{N\to\infty} \sum_{k=1}^{N} \gamma_k \mathbf{e}_k \rangle = \tag{113}$$

$$\text{(by the Lemma about continuity of scalar product)} \lim_{N\to\infty} < \mathbf{e}_r, \sum_{k=1}^{N} \gamma_k \mathbf{e}_k >= \tag{114}$$

$$\lim_{N\to\infty} \left(\sum_{k=1}^{N} \gamma_k \langle \mathbf{e}_r, \mathbf{e}_k \rangle \right) = \gamma_r \,.$$

Now show that Parseval equality (108) (in other words Pyuthagorean Theorem) holds:

$$||\mathbf{x}||^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \langle \sum_{k=1}^{\infty} \alpha_k \mathbf{e}_k, \mathbf{x} \rangle = \langle \lim_{N\to\infty} \sum_{k=1}^{N} \alpha_k \mathbf{e}_k, \mathbf{x} \rangle = \tag{115}$$

using continuity of scalar product

$$\lim_{N\to\infty} \langle \sum_{k=1}^{N} \alpha_k \mathbf{e}_k, \mathbf{x} \rangle = \lim_{N\to\infty} \sum_{k=1}^{N} \alpha_k \langle \mathbf{e}_k, \mathbf{x} \rangle = \lim_{N\to\infty} \left(\sum_{k=1}^{N} \alpha_k \cdot \alpha_k \right) = \sum_{k=1}^{\infty} \alpha_k \cdot \alpha_k \tag{116}$$

Thus we prove that orthonormal system is closed if it is a basis.

2. Proof that b) $\Rightarrow$ c) Now prove that it is complete if it is closed.

Let orthonormal system $\{\mathbf{e}_k\}$ be closed.

Let vector $\mathbf{x} \in \mathcal{H}$ is orthogonal to all vectors $\mathbf{e}_i$. Then all coefficients $\alpha_k = \langle \mathbf{x}, \mathbf{e}_k \rangle$ are equal to zero. Hence by Parseval equality $||\mathbf{x}|| = 0$.

It remains to prove that c)$\Rightarrow$ a)

Let orthonormal system $\{\mathbf{e}_k\}$ be complete. Prove that it is a topological basis.

Take arbitrary vector $\mathbf{x} \in H$. Consider projection of vector $\mathbf{x}$ on the on the plane H_N spanned by vectors $\{\mathbf{e}_1, \dots, \mathbf{e}_N\}$: $\mathbf{x}_N = \sum_{k=1}^N \alpha_k \mathbf{e}_k$, where $\alpha_k = \langle \mathbf{x}, \mathbf{e}_k$—projection of vector $\mathbf{x}$. Consider vectors $h_k = \mathbf{x} - \mathbf{x}_k$. One can see that sequence of these vectors is fundamental (Cauchy) sequence: Indeed by Bessel inequality $\sum_{k=1}^\infty |\alpha_k|^2 \leq ||\mathbf{x}||^2 < \infty$. Hence $||h_n - h_m||^2 = \sum_{k=n}^m |\alpha_k|^2$ tends to zero for arbitrary $m > n$ as $n \to \infty$.

Hence this sequence has limit

$$\lim_{n\to\infty} h_n = h, \text{ i.e } \lim_{n\to infty} ||h - h_n|| = 0$$

It is here where we use that Hilbert space is Banach space!

On the other hand $||h||_n \to 0$. Hence $h = 0$. We proved that $h_k \to 0$, i.e. $\mathbf{x}_n \to \mathbf{x}$. It means that orthonormal system is a basis.

5.7 Important calculations in the Hilbert space of functions.

Return to the space $C([a,b])$ of continuous functions with scalar (inner) product (91). The completion of this scalar (inner) product space is Hilbert space $L^2([a,b])$ considered in the example 3 in the subsection "Hilbert space. Definition and examples"

We see that Pythagoras formulae (115) will lead us to effective method for calculation of many remarkable identities like

$$\sum \frac{1}{n^2} = \frac{\pi^2}{6} \tag{117}$$

.

Consider the different examples of orthonormal bases in $L^2(a,b)$.

Example 1 Consider the set of functions:

$$\left\{\frac{1}{\sqrt{2\pi}}, \sqrt{\frac{1}{\pi}}\cos nx, \sqrt{\frac{1}{\pi}}\sin nx\right\} \tag{118}$$

in the Hilbert space $L^2([0, 2\pi])$ (You can consider for simplicity its non-complete subspace $C([0, 2\pi])$ of real continuous functions on the interval $[0, 2\pi]$).

One can prove that

1) This is orthonormal system (it is easy to prove by direct calculations)

2) This is a basis (One can show it using Weierstraß Approximation Theorem[20]

Of course one can consider another orthonormal bases on the space $L^2([0, 2\pi])$. For example: the system of vectors:

$$\{\mathbf{e}_n\}\} = \left\{\sqrt{\frac{1}{\pi}}\sin\frac{nx}{2}\right\} \tag{119}$$

One can prove that this is orthonormal basis too, i.e.:

$$\langle\sqrt{\frac{1}{\pi}}\sin\frac{nx}{2}, \sqrt{\frac{1}{\pi}}\sin\frac{mx}{2}\rangle = \int_0^{2\pi}\frac{1}{\pi}\sin\frac{nx}{2}\sin\frac{mx}{2}dx = \delta_{mn} \tag{120}$$

and for every continuous function f (more precisely measurable function) there exist γ_n such that

$$f(x) = \sum_{n=1}^{\infty}\gamma_n\sqrt{\frac{1}{\pi}}\sin\frac{rx}{2}$$

This equation where one have to understand in the norm in Hilbert space induced by the inner product(91), i.e.

$$||f - \sum_{n=1}^{\infty}\gamma_r\sqrt{\frac{1}{\pi}}\sin\frac{rx}{2}|| == \sqrt{\int\left(f - \sum_{n=1}^{\infty}\gamma_r\sqrt{\frac{1}{\pi}}\sin\frac{rx}{2}\right)^2 dx} = 0 \tag{121}$$

[20]The proof follows from the following remark: trigonometric polynomials are dense in the subspace of continuous functions which take the same values at the points $1, -1$ with respect to uniform norm $||\ ||$. Hence they are dense in the space of all continuous functions with respect to the norm $||\ ||_2$.

Functions f and $\sum_{n=1}^{\infty}\gamma_r\sqrt{\frac{1}{\pi}}\sin\frac{rx}{2}$ coincide almost elsewhere.

In the school geometry we expand arbitrary vector $\mathbf{a}$ in three-dimensional space over base vectors $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$

$$\mathbf{a} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$$

and consider Pythagoras identity:

$$\mathbf{a}^2 = x^2 + y^2 + z^2$$

We can perform analogous calculations in the Hilbert space $L^2([a,b])$

Example 2 Consider the function $f = x$ in the Hilbert space $L^2([0,2\pi])$ and expand it over basis (118). We will come to the formula:

$$x = a_0\sqrt{\frac{1}{2\pi}} + \sum_{n=1}^{\infty} a_m\sqrt{\frac{1}{\pi}}\cos mx + \sum_{n=1}^{\infty} b_n\sqrt{\frac{1}{\pi}}\sin rx \tag{122}$$

One can calculate a_0, a_n, b_n according to general formulae (104) and (112):

$$a_0 = \int_0^{2\pi} x\sqrt{\frac{1}{2\pi}}dx\,,\ a_m = \int_0^{2\pi} x\sqrt{\frac{1}{\pi}}\cos mxdx\,,\ b_n = \int_0^{2\pi} x\sqrt{\frac{1}{\pi}}\sin nxdx\,, \tag{123}$$

Now apply Pythagoras Theorem we come to identity:

$$\langle x, x\rangle = \int_0^{2\pi} x^2dx = \frac{8\pi^3}{3} = a_0^2 + \sum_{m=1}^{\infty} a_m^2 + \sum_{m=1}^{\infty} b_n^2$$

Making all calculaltions till end we will come to the identity(117).

We can perform all this calculations and come to the identity (117) in a little bit different way:

Example 3 Consider Hilbert space $L^2[0,\pi]$ of Lebesgue measurable real functions on the segment $[0,\pi]$ with the scalar product (91)

The following system of vectors (functions):

$$\{\mathbf{e}_n\} = \left\{\sqrt{\frac{2}{\pi}}\sin nx\right\}\,, \quad n = 1, 2, 3, \dots\,.$$

is an orthonormal basis.

One can calculate the coefficients $\{\gamma_n\}$ of the expansion over the basis $\{\mathbf{e}_n\}$ of the function $f(x) = x$:

$$x = \sum_{n=1}^{\infty} \alpha_n \mathbf{e}_n .$$

Now using Pythagoras identity we come to the relation:

$$\text{square length of the "vector" (function) } x = \langle x, x \rangle = \int_0^\pi x^2 dx = \frac{\pi^3}{3} =$$

$$\text{sum of the squares of components} = \sum_{m=1}^{\infty} \alpha_m^2 \tag{124}$$

On the other hand components $\{\alpha_n\}$ are defined by projections of the vector x on the orthonormal basis:

$$\alpha_n = \langle x, \mathbf{e}_n \rangle = \langle x, \sqrt{\frac{2}{\pi}} \sin nx \rangle = \int_0^\pi \sqrt{\frac{2}{\pi}} x \sin nx dx = \sqrt{\frac{2}{\pi}} \frac{\pi(-1)^{n+1}}{n}$$

Hence:

$$\sum_{n=1}^{\infty} \alpha_n^2 = \sum_{n=1} \frac{2\pi}{n^2} = \langle x, x \rangle = \frac{\pi^3}{3}$$

We come to the famous identity:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \tag{125}$$

Example 4-exercise Repeat the calculations above for the "vector" function $f \equiv 1$ and come to the identity $1 + \frac{1}{9} + \frac{1}{25} + \cdots = ?$

Counterexample Pay attention that the system $\{\sqrt{\frac{2}{\pi}} \sin nx\}$, orthonormal basis in the space $L^2([0,\pi])$, but this same system $\{\sqrt{\frac{1}{\pi}} \sin nx\}$ is only orthonormal system in the space $L^2([0,2\pi])$ but *not the orthonormal* basis: all functions from this system are orthogonal to the function $f = 1$: $\langle 1, \sin nx \rangle = \int_0^{2\pi} \sin nx = 0$.

5.8 Isomorphism of Hilbert spaces

All finite-dimensional spaces are isomorphic if their dimension is the same. It is easy to see that two finite-dimensional Hilbert spaces are also isomorphic if their dimension is the same: It means the following: If $(V_1, \langle\ ,\ L\rangle_1)$, $(V_2, \langle\ ,\ L\rangle_2)$ are two Hilbert spaces and

$$\dim V_1 = \dim V_2 < \infty$$

then there exist linear bijection $A\colon\ V_1 \to V_2$ which preserves scalar products: for every $x, y \in V_1$

$$\langle x, y\rangle_1 = \langle\ Ax, Ay\rangle_2 \tag{126}$$

What about for infinite-dimensional Hilbert spaces.

Infinite-dimensional Banach spaces are not isomorphic: (e.g. there is no linear bijection of $l_1, ||\ ||_1)$ on $l_2, ||\ ||_2$ which preserves the length).

We consider Hilbert spaces which possess countable topological basis. (So called separable Hilbert space) [21]. It is easy to see that in this case Hilbert space has orthonormal countable basis[22].

All infinite-dimensional separable Hilbert spaces are isomorphic to Hilbert space $l_2, ||\ ||_2$. Let H be arbitrary infinite-dimensional separable Hilbert space,i.e. it has orthonormal basis $\{\mathbf{f}_n\}$ $(n = 1, 2, \dots)$. Consider Hilbert space l_2 with canonical orthonormal basis $\{\mathbf{e}_k\}$ (see (100)) Then assign to every $\mathbf{x} = \sum x_k \mathbf{f}_k$ in H the vector $\sum x_k \mathbf{e}_k$ in l_2 This map establishes the isomorphism.

5.9 Linear functionals on Hilbert space

In finite-dimensional case linear space of functionals on linear space V is dual to this space. Euclidean structure (scalar product) on V establishes one-one correspondence between these spaces.

The picture is analogous in Hilbert case. Just instead linear functionals one has to consider bounded linear functionals. We know already that linear functional is bounded if and only if it is continuous (See the subsection 4.3). We often use this fact in this subsection.

[21] The metric space X is called separable if it possesses countable subset dense in X.

[22] examples of non-separable Hilbert spaces are rather exotic. E.g. consider the vector space of *all* functions on $[0, 1]$ which are equal to zero at all the points except finite number of points. Consider the scalar product in this space: $(f, g) = \sum_x f(x)g(x)$. This is non-separable Hilbert space.

For any element $f \in H$ of Hilbert space $\mathcal{H}$ one can consider linear functional

$$F_f(\mathbf{x}) = \langle \mathbf{x}, f \rangle$$

The linear functional is bounded because $|F_f(\mathbf{x})| = |\langle \mathbf{x}, \mathbf{f} \rangle| \le ||\mathbf{f}||||\mathbf{x}||$ and hence it is continuous.

Show that all linear bounded functionals are described by this formula. This establish isomorphism between Hilbert space and its dual.

Theorem let F be an arbitrary continuous (bounded) linear functional on Hilbert space $\mathcal{H}$. Then there exists an unique element $f \in \mathcal{H}$ such that for any $\mathbf{x} \in H$

$$F(\mathbf{x}) = F_f(\mathbf{x}) = \langle \mathbf{x}, f \rangle \tag{127}$$

Proof. Consider kernel V_F of the functional F:

$$V_f = \{\mathbf{x} \in \mathcal{H} \colon \quad F(\mathbf{x}) = 0\}$$

V_f is closed subspace because F is continuous functional. If $V_F \equiv \mathcal{H}$ then nothing to prove: $f = 0$. Let $V_F \ne \mathcal{H}$. Then there exist $\mathbf{y} \in \mathcal{H}$ such that $F(\mathbf{y}) \ne 0$. Consider the component $\mathbf{y}'$ of vector $\mathbf{y}$ which is orthogonal to V_F. Without loss of generality suppose that $F(\mathbf{y}') = 1$. Consider an arbitrary $\mathbf{x} \in H$. If $F(\mathbf{x}) = \alpha$ then $F(\mathbf{x} - \alpha\mathbf{y}') = 0$. Hence the vector $\mathbf{x} - \alpha\mathbf{y}'$ belongs to V_F. This means that scalar product of this vector on the vector $\mathbf{y}'$ is equal to zero: $\langle \mathbf{x} - \alpha\mathbf{y}', \mathbf{y}' \rangle = \langle \mathbf{x}, \mathbf{y}' \rangle - \alpha||\mathbf{y}'||^2 = 0$. Hence $F \equiv F_f$ where $f = \frac{\mathbf{y}'}{||\mathbf{y}'||^2}$

$$F(\mathbf{x}) = \langle \mathbf{x}, f \rangle = \langle \mathbf{x}, \frac{\mathbf{y}'}{||\mathbf{y}'||^2} \rangle = \alpha$$

Proof of uniqueness: If $F(\mathbf{x}) = \langle \mathbf{x}, f \rangle = \langle \mathbf{x}, f' \rangle$. Then $\langle \mathbf{x}, f - f' \rangle = 0$. In particular $\langle f - f', f - f' \rangle$. Thus $f = f'$.

Note that in this proof we did not use topological basis of Hilbert space.

In the case if Hilbert space is separable, i.e. it has countable orthonormal topological basis $\{\mathbf{e}_n\}$ then we can give another proof and present explicit expression for functional and corresponding function f.

Let F be a bounded linear functional on Hilbert space $\mathcal{H}$ provided with orthonormal basis $\{\mathbf{e}_n\}$. Let

$$F(\mathbf{e}_n) = \alpha_n \tag{128}$$

Show that in this case

$$F(\mathbf{x}) = F_f(\mathbf{x}) = \langle \mathbf{x}, f \rangle \text{ where } f = \sum_{n=1}^{\infty} \alpha_n \mathbf{e}_n \tag{129}$$

Consider a sequence of vectors f_N such that $f_N = \sum_{k=1}^N \alpha_k \mathbf{e}_k$. Note that $F(f_N) = \sum_{k=1}^N \alpha_k^2$. If C is a norm of functional F then $|F(f_N)| \le C||f_N||$, i.e. $\sum_{k=1}^N \alpha_k^2 \le C\sqrt{\sum_{k=1}^N \alpha_k^2}$. Hence $||f_N|| = \sqrt{\sum_{k=1}^N \alpha_k^2} \le C$. Thus series $\sum_{k=1}^N \alpha_k^2$ converges. Hence a sequence of vectors f_N tends to the vector $f = \sum_{k=1}^\infty \alpha_k \mathbf{e}_k$. Hence $F(\mathbf{e}_n) = \alpha_n = \langle \mathbf{e}_n, f\rangle$. By linearity and continuity we see that for an arbitrary vector $\mathbf{x}$ the equation (129) holds.

Inverse is true too; to any element $f = \sum_{k=1}^\infty \alpha_k \mathbf{e}_k \in \mathcal{H}$ ($\sum \alpha_k^2 < \infty$) corresponds bounded linear functional F such that $F(\mathbf{x}) = \langle \mathbf{x}, f\rangle$, in particular $F(\mathbf{e}_n) = \alpha_n$. Thus we establish isomorphism between bounded linear functionals on Hilbert space and Hilbert space.

Exercise Let $\{\mathbf{e}_n\}$ be an orthogonal basis in Hilbert space $\mathcal{H}$. Let $\mathbf{z} = \{z_1, z_2, \dots, z_n, \dots\}$ be a sequence in l_2. ($\sum z_k^2 < \infty$). Prove that there exists a linear bounded functional F such that $F(\mathbf{e}_n) = z_n$.

Consider vector $f = \sum z_k \mathbf{e}_k$ in $\mathcal{H}$. This vector is well-defined. Prove it. $\sum z_k^2 < \infty$. Hence $\sum_N^\infty \alpha_k^2 \to 0$ if $N \to \infty$. Hence vectors $f_N = \sum_{k=1}^N z_k \mathbf{e}_k$ form Cauchy sequence and vector $f = \sum_{k=1}^\infty z_k \mathbf{e}_k$ is the limit of this sequence. Formula

$$F(\mathbf{x}) = \langle \mathbf{x}, f\rangle$$

defines linear functional which satisfies the condition $F(\mathbf{e}_n) = z_n$ and it is bounded. Norm of the functional F is equal to the length of the vector f: $|F_f(\mathbf{x})| = |\langle \mathbf{x}, f\rangle| \le ||\mathbf{x}||||f||$ due to CBS inequality.

5.10 Subspaces in Hilbert space

In infinite-dimensional subspace not every linear subspace is closed: e.g. the subspace of polynomials is linear subspace in the Banach space of continuous functions but it is not closed subspace. (Its completion is Banach space of all continuous functions: see subsection 3.5)

The closed subspace of Banach space is Banach space.

The closed linear subspace of Hilbert space is Hilbert subspace.

Consider very important example of the subspace V_f orthogonal to the given element f in the Hilbert space $\mathcal{H}$. Intuitively it is the hyperplane orthogonal to the given vector.

Intuitively it is evident that vector f is a inique (up to a coefficient) vector which is orthogonal to the plane V_f. More precisely consider

$$V_f = \{\text{the set of vectors which are orthogonal to } f\} = \{h \in \mathcal{H}\colon \langle f, h\rangle = 0\} \tag{130}$$

and

$$V_f^{\rm orth} = \{\text{the set of vectors which are orthogonal to } V_f\} = \{g \in \mathcal{H}\colon \langle f, h\rangle = 0 \Rightarrow \langle g, h\rangle = 0\} \tag{131}$$

Proposition *V_f is Hilbert subspace. The vectors orthogonal to V_f are proportional to f, i.e. $V_f^{\rm orth}$ is one-dimensional space spanned by the vector f: $g = \lambda f$ if $g \in V_f^{\rm orth}$*

The proof is very simple if we note that subspace V_f is just a kernel of the functional $F_f(\mathbf{x}) = \langle \mathbf{x}, f\rangle$ considered in the previous subsection. Indeed V_f is Hilbert space as a kernel of bounded functional F_f. Suppose $F_f(f) = 1$. Evidently vector f is orthogonal to V_f. Let vector g be any vector which is orthogonal to V_f. Then vector $g' = g - F_f(g)f$ is orthogonal to V_f too. On the other hand it belongs to V_f. Hence $g' = 0$, i.e. $g = F_f(g)f$ is colinear to f.

Give now more detailed exposition of the proof "independent" of considerations of previous subsection.

Proof': It is obvious that V_f is linear subspace: if $h_1, h_2 \in V_f$ then $\lambda h_1 + \mu h_2 \in V_f$ because $\langle \lambda h_1 + \mu h_2, f\rangle = \lambda\langle h_1, f\rangle + \mu\langle h_2, f\rangle = 0$ if $\langle h_1, f\rangle = \langle h_2, f\rangle = 0$. To prove that V_f is Hilbert space it remains to prove that it is complete (closed), i.e. for every Cauchy sequence $\{h_n\}$ such that all h_n are in V_f its limit is in V_f too. Prove it: If $\{h_n\}$ is a Cauchy sequence in $\mathcal{H}$ then it has limit $h \in \mathcal{H}$ because $\mathcal{H}$ is Hilbert space, hence it is complete by definition. It remains to prove that $h \in V_f$. $\langle h_n, f\rangle = 0$ because $h_n \in V_f$. Hence by continuity of scalar product (see (111))

$$\langle h, f\rangle = \langle \lim_{n\to\infty} h_n, f\rangle = \lim_{n\to\infty}\langle h_n, f\rangle = 0 \Rightarrow h \in V_f$$

Thus we prove that V_f is closed linear subspace of Hilbert space $\mathcal{H}$, i.e. V_f is Hilbert subspace in $\mathcal{H}$.

It remains to prove that every vector g which is orthogonal to V_f is proportional to f.

Suppose that g is orthogonal to V_f.

First of all consider trivial case if $f = 0$. Then $V_f = \mathcal{H}$ and it is easy to see that $g = f = 0$. Indeed If g is orthogonal to V_f then in particularly it is orthogonal to g itself because $g \in V_f = H$. Hence $\langle g, g\rangle = 0$ and $g = 0$

Suppose that $f \neq 0$. Then without loss of generality suppose that $\langle f, f\rangle = 1$ Consider the number $c = \langle f, g\rangle$ and the new vector

$$\tilde{g} = g - cf$$

It is easy to see that by definition of c the vector $\tilde{g} = g - cf$ is orthogonal to f:

$$\langle f, \tilde{g}\rangle = \langle f, g - cf\rangle = \langle f, g\rangle - c\langle f, f\rangle = c - c = 0$$

Hence this vector belongs to V_f. On the other hand the vector $\tilde{g}$ is orthogonal to V_f because g is orthogonal and f is orthogonal to V_f. We come to conclusion that *vector $\tilde{g}$ is orthogonal to itself, because it belongs to V_f and it is orthogonal to V_f.* Hence it is equal to zero: Hence $\langle \tilde{g}, \tilde{g} \rangle = 0 \Rightarrow \tilde{g} = 0$. We see that $\tilde{g} = g - cf = 0$. Hence $g = cf$. We prove that g is proportional to f.

How to find a orthonormal basis in Hilbert space V_f? Let $\{\mathbf{e}\}_\mathbf{n}$ is orthonormal basis in $\mathcal{H}$. If $f = \mathbf{e}_1$ then it is evident that $\{\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \dots\}$ is orthonormal basis in V_f. If e.g. $f = \mathbf{e}_1 + \mathbf{e}_2$, then it is easy to see that

$$\{\frac{\mathbf{e}_1 - \mathbf{e}_2}{\sqrt{2}}, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5, \dots\}$$

is orthonormal basis in V_f.

6 Linear self-adjoint operators in Hilbert space

In this section we consider some basic stuff about linear operators in Hilbert space. We give some general definitions (spectrum of operator, Theorem about inverse operator,e,t,c. for an arbitrary Banach space.

6.1 Linear operators Banach spaces. General properties.

Let $(E_1, ||\cdot||_1, (E_2, ||\cdot||_1)$ be two Banach spaces and A be a linear operator from vector space E_1 to vector space E_2.

We say that operator A is bounded operator if there exists $M > 0$ such that

$$\text{for every } \mathbf{x} \in E_1 \quad ||A\mathbf{x}||_2 \leq M||\mathbf{x}||_1$$

Linear bounded functionals considered before is a special case of linear bounded operator (if $E_2 = \mathbf{R}$).

We consider only bounded linear operators and sometimes omit the prefix bounded.

In finite dimensional case all linear operator are bounded and there are very simple relations between kernel and image of linear operator. They are complimented each other in this case. More in detail, let E_1, E_2 be finite-dimensional vector spaces. Let E_1^*, E_2^* be dual vector space respectively and A^* an operator adjoint to A:

$$A^*\colon\ E_2^* \to E_1^*\colon \quad \langle A^*(x_2), y_1 \rangle \equiv \langle x_2, A^*(y_1) \rangle\,. \tag{132}$$

Then $\ker A$ is orthogonal compliment of $\Im A^*$. In particular if $E_1 = E_2 = V$ then

$$\dim \ker A + \dim \Im A = \dim V \tag{133}$$

It is easy to see from this relation that in finite-dimensional case for an arbitrary number λ

$$\Im(A-\lambda) = V \text{ iff} \ker A = 0 \tag{134}$$

or in the other words for every number λ:

$$\text{Operator } (A-\lambda) \text{ is invertible or } \lambda \text{ is an eigenvalue of } A \tag{135}$$

In infinite-dimensional case the relation between these properties become more complicated.

First of all formulate the fundamental

Theorem (*Banach)* Let A be linear bounded operator from Banach space $(E_1, ||\cdot||_1$ to Banach space $(E_2, ||\cdot||_1)$ such that it establish one-one correspondence between these spaces. Then the map A^{-1} is linear bounded operator.

The fact that A^{-1} is linear operator is almost evident: Indeed note that if $Ax = 0$ then $x = 0$ because $A0 = 0$ and A is bijection. To prove that

$$A^{-1}(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda A^{-1}(\mathbf{x}) + \mu A^{-1}(\mathbf{y}) \tag{136}$$

act on both sides of this expression by operator A.

The proof that operator A^{-1} is bounded is far more complicated[23].

6.2 Spectrum of the operator

Let A be linear bounded operator on the Banach space $(E, ||\cdot||$

Definition We say that number (complex) λ is *a regular value* for the operator A if operator $A-\lambda$ is bijection:

$$\text{Equation } (A-\lambda)x = y \text{ has unique solution for every } y \tag{137}$$

According to Banach Theorem one can say that λ is regular for operator A if there exists on E inverse linear bounded operator $R_\lambda = (A-\lambda)^{-1}$,

Definition We say that number (complex) λ belongs to *the spectrum of the operator* A, $\lambda \in$ *Spectrum* A if λ *is not* regular value for operator A

Definition We say that number (complex) λ is *an eigenvalue* of operator A, if $\ker(A-\lambda) \neq 0$, i.e. if there exists vector $x \neq 0$ (eigenvector) such that

$$Ax = \lambda x$$

It is evident that eigenvalue of A belongs to spectrum of A. Indeed if λ is eigenvalue then $A-\lambda$ is not bijection. Hence $\lambda \in$ *Spectrum* A.

[23] It can be performed in two steps: First considering the sets $M_n = \{y \in E_2\colon \ ||A^{-1}y||_1 \leq n||y||_2\}$ we prove that there exists n such that the set M_n is dense in E_1. Thus we prove that A^{-1} is bounded on dense subset. This implies that A^{-1} is bounded on whole E_2.

As it was noted above In finite-dimensional case every number λ is regular value of operator A or eigenvalue of operator A. In infinite-dimensional a situation is more complicated.

Example Let $L_2[-1,1]$ be a space of Lebesgue square integrable functions (One can view $L_2[-1,1]$ as completion of $C[-1,1]$ of continuous functions with respect to the norm $||f||_2 = \sqrt{\int_{-1}^1 f^2 dx}$, or as equivalence class of measurable functions: two functions f_1, f_2 belong to the same equivalence class if the Lebesgue integral $\int_{-1}^1 (f_1 - f_2)^2 dx = 0$).

Consider the operator

$$A\colon f \mapsto xf \tag{138}$$

The spectrum of this operator is not empty. But on the other hand this operator has no eigenvalues. Explain it.

Operator A is bounded:

$$||Af|| = \sqrt{\int x^2 f^2(x)dx} \leq \sqrt{\int f^2(x)dx} = ||f|| \tag{139}$$

One can see that all $\lambda \notin [-1,1]$ are regular for values of the operator A: Equation

$$(x-\lambda)f = g \tag{140}$$

has a solution

$$f = \frac{g}{x-\lambda} \tag{141}$$

for every function $g \in L_2[-1,1]$. The linear operator $(A-\lambda)^{-1}$ defined by (141) is bounded operator:

$$||(A-\lambda)^{-1}g|| = ||\frac{g}{x-\lambda}|| \leq M||g|| \tag{142}$$

because

$$\int \left(\frac{g(x)}{x-\lambda}\right)^2 dx \leq C \int g^2(x)dx, \quad \text{where } C = \sup_{x\in[-1,1]} \frac{1}{x-\lambda} < \infty,$$

since $\lambda \notin [-1,1]$

Thus we see that points $\lambda \notin [-1,1]$ are regular values of the operator (138).

Consider now $\lambda \in [-1,1]$. Then function $\frac{g}{x-\lambda}$ is not continuous. One can show that in this case the equation (140) has solution in $L_2[-1,1]$ not for all g. E.g. if $g = 1$ and $\lambda \in [-1,1]$ then naively it is evident that equation

$$(x-\lambda)g = 1 \tag{143}$$

has no solution. Naively it is evident that it has no solution. Rigid proof will be given in the next subsection[24]

[24] The relation $f = g$ in Hilbert space $L2[-1,1]$ means that for measurable functions f, g $\int (f-g)^2 dx$. Hence (143) just means that $\int -1^1 \left((x-\lambda)g - 1\right)^2 dx = 0$ for measurable function g

What about eigenvalues?

One can see that this operator *has no eigenvalues*: If $xf = \lambda f$ then $f = 0$ at all points if $\lambda \notin [-1, 1]$. If $\lambda \in [-1, 1]$ then $f = 0$ except all point $x = \lambda$. The equivalence class (see (93)) of the function f in $L_2[-1, 1]$ is equal to zero. ($||f|| = 0$).

All points of the spectrum of operator (138) are not eigenvalues.

We return to this example later when we consider operators in Hilbert space.

6.3 Self-adjoint operators in Hilbert space

Let H be an (separable) Hilbert space, i.e. we suppose that H possesses countable topological basis. It means that H has orthonormal basis $\{\mathbf{e}_n\}$ $(n = 1, 2, \ldots\ldots)$.

Consider linear (bounded) operator in H.

Definition Linear bounded operator A is self-adjoint operator if for every two vectors $x, y \in H$

$$\langle Ax, y \rangle = \langle x, Ay \rangle \tag{144}$$

In the case of finite dimensional case one can easy to describe all self-adjoint operators: If A is self-adjoint operator (symmetric operator) in Euclidean N-dimensional space V then one can find orthonormal basis $\{\mathbf{e}_i\}$ $(i = 1, 2, \ldots, N)$ such that e_n are eigenvectors of A. In other words matrix of operator A is diagonal on the basis $\{\mathbf{e}_n\}$. In the complex case self-adjoint operator (hermitian operator) has real eigenvalues.

It turns out that self-adjoint operator preserves some properties in infinite-dimensional Hilbert space preserves some properties.

In particular

Theorem

1.Eigenvalues of self-adjoint operator are real (even in complex case)

2. Eigenvectors corresponding to different eigenvalues are orthogonal.

Proof

If we consider the complex case then $\langle x, Ax \rangle = \langle Ax, x \rangle$ is real because $\langle a, b \rangle = \overline{\langle b, a \rangle}$. On the other hand in complex case $\langle x, Ax \rangle = \langle x, \lambda x \rangle = \lambda \langle x, x \rangle$, $\langle Ax, x \rangle = \langle \lambda x, x \rangle = \overline{\lambda} \langle x, x \rangle$. Hence $\lambda = \bar{\lambda}$ and λ is real.

Let x, y be two eigenvectors with eigenvalues λ, μ: $Ax = \lambda x$, $Ay = \mu x$ such that $\lambda \neq \mu$. Prove that $\langle x, y \rangle = 0$, i.e. x, y are orthogona, using that $\langle x, Ay \rangle = \langle Ax, y \rangle$

$$\langle x, y \rangle = \frac{1}{\lambda} \langle x, \lambda y \rangle = \frac{1}{\lambda} \langle x, Ay \rangle = \frac{1}{\lambda} \langle Ax, y \rangle = \frac{1}{\lambda} \langle Ax, y \rangle = \frac{\mu}{\lambda} \langle Ax, y \rangle \tag{145}$$

Hence $\langle x, y \rangle = 0$.

In the previous subsection we began to study the difference between eigenvalues and Spectrum for operator in infinite-dimensional case.

Study this problem in detail for self-adjoint operator.

Let A be linear bounded self-adjoint operator A in the Hilbert space H. Denote by G_λ an image of operator $A - \lambda$. $G_\lambda = H$ if λ is regular number for the operator A (see the definition in the previous subsection).

If λ is not regular, i.e. belongs to the spectrum to the operator A then $G_\lambda \neq H$.

It turns out that

Proposition If $\lambda \in$ *Spectrum* A but λ is not an eigenvalue of the operator A then G_λ is dense in H, i.e. $\overline{G}_\lambda = H$.

Prove this Proposition. Suppose that closure of $G_\lambda \neq 0$. Then there exists a vector $x \neq 0$ which is orthogonal to $\overline{G}_\lambda$. In particular vector x is orthogonal to G_λ, i.e. for every y,

$$\langle x, (A-\lambda)y\rangle = \langle (A-\lambda)x, y\rangle = 0$$

Hence $(A-\lambda)x = 0$, i.e. λ is eigenvalue. Contradiction.

Now using this Proposition we can formulate the Theorem which is very useful for finding spectrum of the self-adjoint operator and in particularfor distinguishing between λ which belong to spectrum and λ which are eigenvalues.

Theorem Let A be a linear bounded self-adjoint operator in the Hilbert space H then

Then $\lambda \in$ *SpecA* iff there exists a sequence of vectors $\mathbf{x}^{(n)}$ in Hilbert space H such that

$$||\mathbf{x}^{(n)}|| = 1 \quad \text{and} \quad \lim_{n\to}(A-\lambda)\mathbf{x}^{(n)} = 0 \tag{146}$$

The statement of this Theorem immediately follows from the Proposition above. Pay attention that in the case if λ is in spectrum and moreover it is an eigenvalues then instead condition (146) of the theorem we come to more strong condition:

$$(A-\lambda)\mathbf{x} = 0 \tag{147}$$

If the weaker condition (146) is obeyed but λ is not eigenvalue then λ belongs to spectrum.

Example Return again to our example (138). We proved that all $\lambda \notin [-1,1]$ are regular, all $\lambda \in [-1,1]$ are not eigenvalues. Now using the Theorem we prove that all $\lambda \in [-1,1]$ belong to the spectrum of the operator A in spite of the fact that they are not eigenvalues. (138).

First check that conditions of Theorem is obeyed. We checked already that it is bounded (see (139)). A is self-adjoint operator in the Hilbert space $L_2[-1,1]$ because

$$\langle Af, g\rangle = \int_{-1}^{1} xf(x)g(x)dx = \int_{-1}^{1} f(x)xg(x)dx = \langle f, Ag\rangle$$

Repeat again that all $\lambda \in [-1,1]$ are not eigenvalues: if $xf = (x-\lambda)f$ then function f is equal to zero at all points except may be the point λ (if $\lambda \in [-1,1]$). Hence $f = 0$ in $L_2[-1,1]$. (More precisely it belongs to equivalence class of the function $f \equiv 0$ in $L_2[-1,1]$: see example (93)). So we see that operator $A\colon f \mapsto xf$ has no eigenvalues. Prove that nevertheless all $\lambda \in [-1,1]$ belong to spectrum of the operator (138) using the last Theorem.

Consider e.g. $0 \in [-1,1]$ and the sequence f_n of functions such that f_n is equal to zero at all the points except the interval $\Delta_n = (0 - 1/2n^2, 0 + 1/2n^2)$. Put that function f_n is equal to n on the points of the interval Δ_n. One see that $||f|| = 1$ (Why?). Then consider functions $g_n = Af_n = xf_n$ and check that $||xf_n|| \to 0$. According to Theorem it means that 0 is value of the operator A in the spectrum.

In a similar way one can prove that every $\lambda \in [-1,1]$ belongs to spectrum.