Matematička teorija kernela

Duje Jerić-Miloš

29. listopada 2023.

Sadržaj

1 Osnovni teorijski koncepti

U ovom poglavlju cilj je definirati posebnu podklasu klase normiranih vektorskih prostora, tzv. Hilbertove prostore, koji će biti jedan od ključnih koncepata u matematičkoj teoriji kernela.

1.1 Vektorski prostori

Prisjetimo se, **vektorski prostor** (nad $\mathbb{R}$) je neprazan skup X zajedno s preslikavanjima $+ : X \times X \to X$ i $\cdot : \mathbb{R} \times X \to X$ koja zadovoljavaju sljedeća svojstva:

(V1) $(x + y) + z = x + (y + z)$, za sve $x, y, z \in X$.

(V2) $x + y = y + x$, za sve $x, y \in X$.

(V3) Postoji $\mathbf{0} \in X$ takav da za sve $x \in X$ vrijedi $x + \mathbf{0} = x$.

(V4) Za svaki $x \in X$ postoji jedinstveni $-x \in X$ takav da vrijedi $x + (-x) = \mathbf{0}$.

(V5) Za svaki $\alpha, \beta \in \mathbb{R}$ i svaki $x \in X$ vrijedi $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$

(V6) Za svaki $\alpha, \beta \in \mathbb{R}$ i svaki $x \in X$ vrijedi $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$

(V7) Za svaki $\alpha \in \mathbb{R}$ i sve $x, y \in X$ vrijedi $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$.

(V8) Za svaki $x \in X$ vrijedi $1 \cdot x = x$.

Podskup S vektorskog prostora X je **linearno nezavisan skup** ako se nijedan vektor $s \in S$ ne može prikazati kao konačna linearna kombinacija preostalih vektora iz S. Podskup G vektorskog prostora X je **skup generatora** (ili **izvodnica**) ako se svaki vektor iz X može prikazati kao konačna linearna kombinacija elemenata iz S. **Algebarska (Hamelova) baza** B vektorskog prostora X je linearno nezavisan skup generatora. Može se pokazati da svaki vektorski prostor ima Hamelovu bazu te da svake dvije baze vektorskog prostora imaju istu kardinalnost pa kardinalnost proizvoljne baze vektorskog prostora nazivamo njegovom **algebarskom (Hamelovom) dimenzijom**.

Za vektorske prostore koji imaju konačnu algebarsku bazu kažemo da su **konačnodimenzionalni**, a u suprotnom kažemo da su **beskonačnodimenzionalni**.

Napomenimo da radimo sa realnim vektorskim prostorima, makar je situacija potpuno analogna i za kompleksne (ali npr. ne moramo se zamarati kompleksnim konjugacijama i sl.).

Primjer 1. Skup $\mathbb{R}^n$ s preslikavanjima $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ i $\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ definiranim po koordinatama, tj.

$$(x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1, ..., x_n + y_n), \quad \alpha(x_1, ..., x_n) = (\alpha x_1, ..., \alpha x_n)$$

je vektorski prostor (nad $\mathbb{R}$), pri čemu je $\mathbf{0} = (0, ..., 0)$ i za svaki $(x_1, ..., x_n)$ je $-(x_1, ..., x_n) = (-x_1, ..., -x_n)$.

Primjer 2. Neka je X neprazan skup i neka je $\mathcal{F}$ skup svih funkcija $f : X \to \mathbb{R}$. Definiramo preslikavanja $+ : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ i $\cdot : \mathbb{R} \times \mathcal{F} \to \mathcal{F}$ na sljedeći način. Za funkcije $f, g \in \mathcal{F}$ te skalar $\alpha \in \mathbb{R}$ definiramo funkcije $f + g, \alpha f : X \to \mathbb{R}$ pravilom

$$(f + g)(x) = f(x) + g(x), \quad (\alpha \cdot f)(x) = \alpha f(x), \quad \forall x \in X.$$

Skup $\mathcal{F}$ uz ovako definirana preslikavanja $+, \cdot$ je vektorski prostor (nad $\mathbb{R}$). Pri tome je $\mathbf{0}$ nulfunkcija $0 : X \to \mathbb{R}$, $0(x) = 0$, za svaki $x \in X$, a za $f \in \mathcal{F}$ je $-f$ funkcija $-f : X \to \mathbb{R}$ funkcija definirana pravilom $(-f)(x) = -f(x)$, za svaki $x \in X$.

1.2 Normirani prostori

Definicija 3. Neka je X vektorski prostor (s obzirom na preslikavanja $+ : X \times X \to X$ i $\cdot : \mathbb{R} \times X \to X$). Za funkciju $\|\cdot\| : X \to \mathbb{R}$, $\|\cdot\| : x \mapsto \|x\|$ kažemo da je **norma na prostoru** X ako vrijede sljedeća svojstva:

(N1) $\|x\| \geq 0$, za svaki $x \in X$.

(N2) $\|x\| = 0$ ako i samo ako je $x = \mathbf{0}$.

(N3) $\|\alpha \cdot x\| = |\alpha| \, \|x\|$

(N4) $\|x + y\| \leq \|x\| + \|y\|$

Vektorski prostor X zajedno s normom $\|\cdot\|$ naziva se **normirani prostor**. Za $x \in X$, nenegativan realan broj $\|x\|$ nazivamo **normom vektora** x. Za vektor $x \in X$ kažemo da je **normiran** ako mu je norma jednaka 1.

Primjer 4. Funkcija $\|\cdot\| : X \to \mathbb{R}$, $\|(x_1, ..., x_n)\| = \sqrt{\sum_{i=1}^n x_i^2}$, je norma na vektorskom prostoru $\mathbb{R}^n$ koju nazivamo **euklidska (standardna)** norma na $\mathbb{R}^n$. Ispostavlja se da su sve norme $\mathbb{R}^n$ na određeni način ekvivalentne (vidi dodatak A), tj. induciraju istu topologiju.

Primjer 5. Neka je $\Omega \subseteq \mathbb{R}^n$ neprazan skup i $\mathcal{BC}$ skup svih omeđenih i neprekidnih funkcija $f : \Omega \to \mathbb{R}$. Skup Ω s obzirom na standardno definirano zbrajanje te množenje funkcije skalarom ima strukturu vektorskog prostora. Definirajmo funkciju $\|\cdot\| : \mathcal{BC} \to \mathbb{R}$,

$$\|f\| = \sup\{f(x) : x \in \Omega\}.$$

Funkcija $\|\cdot\|$ je norma na vektorskom prostoru $\mathcal{BC}$ koju nazivamo **supremum normom** na prostoru $\mathcal{BC}$.

Intuitivno, dvije su funkcije blizu (ne razlikuju se više od nekog $\varepsilon > 0$) u prostoru $\mathcal{BC}$ u supremum normi akko su im slike ($f(x)$, $g(x)$) blizu svugdje na domeni.

Primjer 6. Neka su X i Y normirani prostori s normama $\|\cdot\|_X$ i $\|\cdot\|_Y$. Za linearni operator $A : X \to Y$ kažemo da je **ograničen** ako je $\sup\{\|A(x)\|_Y : \|x\|_X = 1\} < \infty$. Neka je $B(X, Y)$ skup svih ograničenih linearnih operatora $A : X \to Y$. Skup $B(X, Y)$ je vektorski prostor (koji nazivamo prostorom ograničenih linearnih operatora), a s

$$A \mapsto \|A\| := \sup\{\|A(x)\| : \|x\| = 1\}$$

je na njemu definirana norma koju nazivamo **operatorskom normom**. Specijalno, prostor $B(X, \mathbb{R})$ nazivamo **topološkim dualom** prostora X i označavamo X^*. Nije teško pokazati da je operator ograničen akko je neprekidan, stoga je X^* samo prostor neprekidnih linearnih funkcionala.

Primijetimo da je s $d(x,y) = \|x-y\|$ definirana funkcija $d : X \times X \to \mathbb{R}$ koja je metrika na normiranom prostoru X (tj. $d(x,y)$ predstavlja *udaljenost* vektora x i y). Specijalno, ako uzmemo $y = 0$, dobivamo da je $d(x,0) = \|x\|$, tj. norma vektora x je njegova udaljenost od nulvektora.

Definicija 7. Neka su X i Y normirani prostori s normama $\|\cdot\|_X$ i $\|\cdot\|_Y$. Kažemo da je funkcija $f : X \to Y$ **izometrija** ako za sve $x, x' \in X$ vrijedi

$$\|f(x) - f(x')\|_Y = \|x - x'\|_X .$$

Imajući na umu da je s $\|x - x'\|$ definirana udaljenost među vektorima x i x', možemo reći da je izometrija funkcija koja čuva udaljenost među vektorima, tj. udaljenost slika u prostoru Y jednaka je udaljenosti polaznih vektora u prostoru X. Ako postoji bijektivna linearna izometrija f između normiranih prostora X i Y (koja svakoj točki $x \in X$ pridruži njoj ekvivalentnu točku $f(x) \in Y$ i obratno) onda kažemo da su X i Y **izometrični** i pišemo $X \simeq Y$

Definicija 8. Neka je $x_0 \in X$ proizvoljan vektor i $r > 0$ proizvoljan realan broj. Skup svih vektora $x \in X$ koji su x_0 udaljeni za manje od r nazivamo **kuglom sa središtem u** x_0 **radijusa** r i označavamo $B(x,r)$, tj. $B(x,r) = \{x \in X : \|x - x_0\| < r\}$.

Budući da je sada na X definirana metrika, ima smisla promatrati topološke pojmove u prostoru X (otvorenost, zatvorenost, konvergenciju, ...).

Primjerice, niz vektora (x_n) u X konvergira prema vektoru $x \in X$ ako i samo ako za svaku (po volji malenu) grešku $\varepsilon > 0$ možemo pronaći neki dovoljno daleki član niza $N \in \mathbb{N}$ nakon kojega svi članovi $n \in \mathbb{N}, n \geq N$ ne izlaze iz kugle radijusa ε centrirane u x, tj. akko

$$(\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall n \in \mathbb{N}, n \geq N)\, \|x_n - x\| < \varepsilon$$

Nadalje, sljedeće definicije ćemo također koristiti u nastavku:

- Skup $A \subset X$ je **zatvoren** ako svaki niz (x_n) u A, konvergira u A.
- **Zatvorenje skupa** A, $\overline{A}$ unija je skupa A te skupa koji tvore limesi svih konvergentnih nizova s članovima u A. $\overline{A}$ je naravno zatvoren.
- $A \subset X$ **gust** u X ako je svaka točka $x \in X$ limes nekog niza točaka $x_n \in A$. Ovo nam govori da se bilo koja točka može aproksimirati pomoću nekih (prebrojivo mnogo) točaka iz A. Ekvivalentno, $A \subset X$ je gust u X ako $\overline{A} = X$.

Definicija 9. Neka je (x_n) niz u normiranom prostoru X. **Niz parcijalnih suma niza** (x_n) je niz (s_n), gdje je $s_n = x_1 + x_2 + ... + x_n$. Par $((x_n), (s_k))$ niza (x_n) i niza (s_k) njegovih parcijalnih suma nazivamo **redom vektora u normiranom prostoru** X i označavamo $\sum x_n$. Ukoliko niz (s_n) konvergira prema nekom vektoru $x \in X$, onda kažemo da **red** $\sum x_n$ **konvergira** i pišemo

$$\sum_{n \in \mathbb{N}} x_n = x.$$

Ako je $(x_i, i \in I)$ proizvoljna familija vektora iz X te $J = \{i \in I : x_i \neq 0\}$ najviše prebrojiv skup te skup te red $\sum_J x_j$ konvergira, onda kažemo da **generalizirani red** $\sum_I x_i$ konvergira i definiramo

$$\sum_{i \in I} x_i = \sum_{j \in J} x_j.$$

Javljat će nam se sume pozitivnih realnih brojeva po proizvoljnom indeksnom skupu i korisno je dati ekvivalentnu karakterizaciju takve sume koja je jednostavnija za korištenje.

Neka je $f : I \to [0, \infty]$ preslikavanje na proizvoljnom skupu I. Sumu $\sum_{i \in I} f(i)$ definiramo kao supremum po svim konačnim sumama $\sup\{\sum_{i \in F} f(i) : F \subset I, |F| < \infty\}$ [1] Nije teško pokazati da je $\sum_{i \in I} f(i)$ konačan jedino kada je $f(i) = 0$ za sve osim prebrojivo mnogo $i \in I$ pa zapravo nismo dobili nikakav generalniji pojam sumacije. Naime, ako je A_n skup svih i za koje $|f(i)| > 1/n$, onda ako bi čitav $A = \{i : f(i) \neq 0\} = \bigcup_{n \in \mathbb{N}} A_n$ bio neprebrojiv, to znači da je barem jedan A_k neprebrojiv, stoga je $\sum_{i \in F} f(i) > |F|/k$ za svaki konačni poskup $F \subset A_k$ pa $\sum_{i \in A_k} f(i) = \infty$, stoga ni suma po $I \supset A_k$ ne može biti konačna. Zapravo ovaj argument pokazuje da je f sumabilna jedino ako su svi A_n konačni.

Sa druge strane je jasno da je limes parcijalnih suma pozitivnih realnih brojeva dan supremumom istih (a zbog apsolutne konvergencije nije bitan redoslijed članova reda pa je taj supremum isti kao i supremum po svim konačnim podskupovima od $\mathbb{N}$).

[1] Ovo je samo integral funkcije f po diskretnoj mjeri/mjeri za brojanje na skupu I.

1.3 Unitarni prostor

Definicija 10. Neka je X vektorski prostor (nad $\mathbb{R}$). Za funkciju $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{R}$ kažemo da je **unitarni/unutrašnji/skalarni produkt** ako zadovoljava sljedeća svojstva:

(U1) $\langle\cdot|\cdot\rangle$ je simetrična: $\langle x|y\rangle = \langle y|x\rangle$ za sve $x, y \in X$

(U2) $\langle\cdot|\cdot\rangle$ je linearna u prvom argumentu: $\langle \alpha x + \beta z|y\rangle = \alpha\langle x|y\rangle + \beta\langle z|y\rangle$ za sve $x, y, z \in X$ i $\alpha, \beta \in \mathbb{R}$

(U3) za sve $x \in X$, $\langle x|x\rangle \geq 0$ i $\langle x|x\rangle = 0$ samo za $x = 0$.

Primijetimo da je po (U1) $\langle\cdot|\cdot\rangle$ linearna i u drugom argumentu, tj. *bilinearna* je.

Vektorski prostor X zajedno sa unutrašnjim produktom $\langle\cdot|\cdot\rangle : X \times X \to \mathbb{R}$ zovemo **unitarnim** prostorom.
Na svakom unitarnom prostoru može se pokazati da funkcija $\|x\| = \sqrt{\langle x|x\rangle}$ definira normu (jedino je nejednakost trokuta netrivijalna te slijedi primjenom Cauchy-Schwarzove nejednakosti - vidi ispod).

Primjer 11. Na vektorskom prostoru $\mathbb{R}^n$ možemo definirati tzv. **standardni skalarni umnožak** . Za $x = (x_1, ..., x_n)$, $y = (y_1, ..., y_n) \in \mathbb{R}$ definiramo

$$\langle x|y\rangle = \sum_{i=1}^{n} x_i y_i.$$

Norma inducirana ovim skalarnim produktom je standardna (euklidska) norma iz primjera 4.

Primjer 12. Označimo s $C([-1,1])$ skup svih neprekidnih realnih funkcija $f : [-1,1] \to \mathbb{R}$. Skup $C([-1,1])$ je vektorski prostor (s obzirom na standardno definirano zbrajanje funkcija te množenja funkcija skalarom). Ako su $f, g \in C[-1,1]$ neprekidne funkcije, onda je i funkcija $fg : [-1,1] \to \mathbb{R}, (fg)(t) = f(t)g(t)$, za sve $t \in [-1,1]$ neprekidna funkcija (a svaka neprekidna funkcija na segmentu $[-1,1]$ je Riemann integrabilna) pa definiramo skalarni produkt $\langle\cdot|\cdot\rangle : C([-1,1]) \times C([-1,1]) \to \mathbb{R}$, $\langle f|g\rangle = \int_{-1}^{1} f(t)g(t)dt$.

Primijetimo da su bilinearnost i pozitivnost trivijalne. Jedino netrivijalno svojstvo je $\langle f|f\rangle = 0 \implies f = 0$. Ovo slijedi jer ako je $f^2 > 0$ u nekoj točki, onda po neprekidnosti $f^2 > 0$ na čitavom intervalu oko te točke pa je nužno integral funkcije f^2 po tom intervalu (pa i po čitavom $[-1,1]$) veći od 0.

Definicija 13. Reći ćemo da su elementi x, y unitarnog prostora X **ortogonalni** (ili **okomiti**) ako $\langle x|y\rangle = 0$. Za skup $S \subseteq X$ kažemo da je **ortogonalan** ako su svaka dva različita vektora iz S međusobno ortogonalna. Ortogonalan skup $\mathcal{O} \subset X$ u kojem je svaki vektor normiran nazivamo **ortonormiranim skupom**.

Primijetimo da je svaki ortogonalan (a onda i ortonormiran) skup vektora linearno nezavisan skup.

Primjer 14. Prostor $C([-1,1])$ sa prethodno definiranim skalarnim umnoškom ima za ortogonalni skup Legendreove polinome P_n. Eksplicitni izrazi ovih polinoma su (do na normalizaciju) u potpunosti određeni ovim zahtjevom. Tradicionalno ih je normalizirati tako da $P_n(1) = 1$. Alternativno, možemo ih normalizirati tako da tvore ortonormirani skup.

Propozicija 15 (Besselova nejednakost)**.** Neka je $\{e_1, ..., e_n\}$ konačan skup ortonormiranih vektora iz unitarnog prostora X, onda za svaki $x \in X$

$$\sum_i \langle x|e_i\rangle^2 \leq \|x\|^2$$

Dokaz.

$$0 \leq \langle x - \sum_{i=1}^{n} \langle x|e_i\rangle e_i | x - \sum_{i=1}^{n} \langle x|e_i\rangle e_i\rangle = \langle x|x\rangle - 2\sum_i \langle x|e_i\rangle^2 + \sum_{i,j} \langle x|e_i\rangle\langle x|e_j\rangle\langle e_i|e_j\rangle$$

Kako $\langle e_i|e_j\rangle$ iznosi 0 za sve $i \neq j$, a 1 kada $i = j$, imamo:

$$0 \leq \langle x|x\rangle - 2\sum_i \langle x|e_i\rangle^2 + \sum_i \langle x|e_i\rangle^2 = \|x\|^2 - \sum_i \langle x|e_i\rangle^2$$

$\square$

Uzimajući supremum po konačnim sumama, vidimo da ista nejednakost vrijedi i za sume po proizvoljnim indeksnim skupovima te da jedino prebrojivo mnogo $\langle x|e_i\rangle$ može biti različito od 0.

Propozicija 16 (Cauchy-Schwarzova nejednakost)**.** Neka je X unitarni prostor, onda za sve $x, y \in X$ vrijedi:

$$|\langle x|y\rangle| \leq \|x\| \, \|y\|$$

Dokaz. Za $y = 0$ tvrdnja je trivijalna, a za $y \neq 0$ primijenimo Besselovu nejednakost na jednočlanu ortonormiranu familiju $y/\|y\|$ □

Sada iz Cauchy-Schwarzove nejednakosti imamo: $\|x+y\|^2 = \|x\|^2 + 2\langle x|y\rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x|y\rangle| + \|y\|^2 \leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$. Korjenovanjem, vidimo da $\|x\| = \sqrt{\langle x|x\rangle}$ zadovoljava nejednakost trokuta

Propozicija 17. Na svakom Unitarnom prostoru vrijedi tzv. pravilo paralelograma

$$\|x+y\| + \|x-y\| = 2(\|x\|^2 + \|y\|^2)$$

Dokaz. Ovo lagano vidimo zbrajanjem identiteta $\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x|y\rangle + \|y\|^2$. □

1.4 Potpunost. Hilbertov prostor

Definicija 18. Neka je X normirani vektorski prostor. Za niz (x_n) u X kažemo da je **Cauchyjev** ako

$$(\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall m, n \geq N)\, \|x_m - x_n\| < \varepsilon$$

Grubo govoreći, *kasniji* članovi niza postaju međusobno sve bliži i bliži. Jasno je da je svaki konvergentan niz Cauchyjev, ali obrat ne mora vrijediti.

Teorem 19. Neka je X unitarni prostor. Sljedeće je ekvivalentno:

1. (**Potpunost**) Svaki Cauchyjev niz u X konvergira.
2. Neka je $A \subset X$ *zatvoreni* potprostor te $A^\perp = \{x : \langle x|y\rangle = 0 \ \forall y \in A\}$ vektorski prostor vektora koji su ortogonalni na svaki vektor iz A. Onda $X = A \oplus A^\perp$ (pa je $A^\perp$ ortogonalni komplement od A). Eksplicitno, svaki $x \in X$ se može zapisati na jedinstven način kao suma $x = a + a^\perp$, gdje su a i $a^\perp$ jedinstvene točke u A i $A^\perp$ respektivno čije su udaljenosti od točke x minimalne.
3. (**Rieszov teorem**) X je refleksivan: $X \simeq X^*$, gdje je ekvivalencija u smislu izometrije, a X^* smo snabdjeli operatorskom normom. Eksplicitno, izometrija je dana sa $x \mapsto \langle x|\cdot\rangle$.

Unitarni prostor koji zadovoljava jedan od ovih ekvivalentnih uvjeta zovemo **Hilbertovim prostorom**.

Za fizičare je primjerice posljednji uvjet (Rieszov teorem) posebno značajan. Elementi Hilbertovog prostora X se još zovu ket vektori, a funkcional $\langle x|\cdot\rangle$ se označava samo sa $\langle x|$ i zove pripadajućim bra vektorom ket vektora x. Same ket vektore je onda tradicionalno zapisati kao $|x\rangle$. Posljedica ovakve notacije je da djelovanje bra vektora $\langle x|$ na ket vektor $|y\rangle$ onda samo postane $\langle x|(|y\rangle) = \langle x|y\rangle$

Prvo valja dokazati lemu:

Lema 20. Neka su V i W normirani prostori i $B(V,W)$ prostor svih omeđenih operatora sa operatorskom normom. Ako je W potpuni normirani prostor, onda je i $B(V,W)$ isto potpun.

Dokaz. Neka je $T_n \in B(V,W)$ Cauchyjev niz. Tada je očito za svaki $x \in X$ niz $(T_n(x))$ Cauchyjev u W pa konvergira prema nekom vektoru iz W, označimo ga $T(x)$. Prvo primijetimo da je operator $T : x \mapsto T(x)$ linearan. Naime, $T_n(x+y) = T_n(x) + T_n(y)$ za svaki $n \in \mathbb{N}$ pa prelaskom na limes slijedi $T(x+y) = T(x) + T(y)$. Isto tako $T_n(\alpha x) = \alpha T_n(x)$ za svaki $n \in \mathbb{N}$ pa prelaskom na limes slijedi $T(x) + T(y)$.

Dokažimo još da je operator T omeđen, tj. $T \in B(V,W)$ te da niz (T_n) konvergira prema $T \in B(V,W)$. Primijetimo prvo, kako je niz (T_n) Cauchyjev, postoji $n_0 \in \mathbb{N}$ takav da za sve $k, n \geq n_0$ vrijedi $\|T_n - T_k\| \leq \varepsilon$ pa za proizvoljan $x \in V$ vrijedi $\|(T_n - T_k)(x)\| \leq \|T_n - T_k\|\,\|x\| \leq \varepsilon\,\|x\|$. Uzimanjem limesa kad $k \to \infty$ slijedi da je $\|(T_n - T)(x)\| \leq \varepsilon\,\|x\|$, za sve $x \in V$, pa je $T_n - T$ ograničen, a kako je $B(V,W)$ vektorski prostor i $T_n \in B(V,W)$, slijedi da je i $T = T_n - (T_n - T) \in B(V,W)$. Konačno, $\|T_n - T\| = \sup\{\|(T_n - T)(x)\| : \|x\| \leq 1\} \leq \sup\{\varepsilon\,\|x\| : \|x\| \leq 1\} = \varepsilon$, za sve $n \geq n_0$, čime je pokazano da (T_n) konvergira prema T. Dakle, $B(V,W)$ je potpun.

□

Dokaz teorema 19. Pokazat ćemo da $1 \implies 2 \implies 3 \implies 1$

- (**1** $\implies$ **2**) Za $x \in X$ neka je $\delta = \inf\{\|x-y\| : y \in A\}$ i neka je $y_n \in A$ tako da $\|x - y_n\| \to \delta$. Naime, za svaki n odaberemo po jedan $y_n \in A$ tako da $\|x - y_n\| \leq \delta + 1/n$, što možemo po definiciji supremuma i aksiomu (prebrojivog) odabira.

  Pokažimo da je niz tako konstruirani niz y_n Cauchyjev. Po pravilu paralelograma imamo

  $$\|y_n - y_m\|^2 + \|y_n + y_m - 2x\|^2 = 2(\|y_n - x\|^2 + \|y_m - x\|^2)$$

  Stoga,

  $$\|y_n - y_m\|^2 = 2(\|y_n - x\|^2 + \|y_m - x\|^2) - 4\left\|\frac{1}{2}(y_n + y_m) - x\right\|^2$$

  Kako je $1/2(y_n + y_m) \in A$, imamo $\|y_n - y_m\|^2 \leq 2(\|y_n - x\|^2 + \|y_m - x\|^2) - 4\delta$. Konačno, za $n, m \to \infty$ desna strana teži u 0 pa je y_n zaista Cauchyjev.

  Neka je sada $y = \lim_{n\to\infty} y_n$. Kako je $\delta \leq \|x - y\| = \delta \leq \|x - \lim_{n\to\infty} y_n\| = \lim_{n\to\infty} \|x - y_n\| \leq \lim_{n\to\infty} (\delta + 1/n) = \delta$, slijedi da se promatrani infum postiže kao minimum u točki y. Stavimo $z = x - y$. Kako je A zatvoren i $y_n \in A$, imamo $y \in A$, a trebamo pokazati da i $z \in A^\perp$. Naime, neka je $u \in A$ proizvoljan i promotrimo funkciju $f : \mathbb{R} \to \mathbb{R}$, $f(t) = \|z + tu\|^2 = \|z\|^2 + 2\langle z|u\rangle t + t^2 \|u\|^2$. f ima minimum u $t = 0$ jer $z + tu = x - (y - tu)$ $(y - tu \in A)$, a $\|x - v\|$ za $v \in A$ minimum poprima u $v = y$ te iznosi δ. Dakle, $0 = f'(0) = 2\langle z|u\rangle$ i vidimo da zaista $z \in A^\perp$.

  Kako $x = y + z$, pokazali smo da $X = A + A^\perp$, a da $X = A \oplus A^\perp$ slijedi odmah jer $A \cap A^\perp = \{0\}$. Naime, za $y \in A \cap A^\perp$ imamo posebno $\langle y|y\rangle = 0$ pa $y = 0$.

- (**2** $\implies$ **3**) Prvo primijetimo da je $x \mapsto \langle x|\cdot\rangle$ nužno injekcija $X \to X^*$ i izometrija. Naime, jezgra istog preslikavanja je $\{x : \langle x|y\rangle = 0\ \forall y \in X\} = \{0\}$ (jer posebno $\langle x|x\rangle = 0$ pa $x = 0$) pa je injektivno. S druge strane, $\|\langle x|\cdot\rangle\| = \sup\{\langle x|y\rangle : \|y\| \leq 1\} = \|x\|$ budući da se supremum postiže kao maksimum u $y = x/\|x\|$ (po Cauchy-Schwarzovoj nejednakosti) Dakle, potrebno je jedino pokazati surjektivnost.

  Ako je $f \in X^*$ nulfunkcional, onda očito možemo uzeti samo $x_f = 0$, inače stavimo $A = \{y \in X : f(y) = 0\}$. Kako je f neprekidna, ovo je zatvoreni pravi potprostor od X $(f \neq 0)$ pa $A^\perp \neq \{0\}$. Sada odaberemo jedinični $z \in A^\perp$ i stavimo za prizvoljni $y \in X$, $u = f(y)z - f(z)y$. Kako je očito $f(u) = 0$, imamo $u \in A$ pa $0 = \langle u|z\rangle = f(y)\|z\|^2 - f(z)\langle y|z\rangle = f(y) - \langle y|f(z)z\rangle$. Stoga, $f(y) = \langle f(z)z|y\rangle$ i $x \mapsto \langle x|\cdot\rangle$ je surjektivna.

- (**3** $\implies$ **1**) Kako je $X^* = B(X, \mathbb{R})$ i $\mathbb{R}$ je potpun, ovo je samo poseban slučaj prethodne leme [2].

$\square$

Primjer 21. Može se pokazati da je svaki konačnodimenzionalni (normirani/unitarni) prostor potpun.

U nastavku ćemo vidjeti da postoje potpuni beskonačnodimenzionalni prostori te da nije svaki beskonačnodimenzionalni prostor potpun, tj. da vrijedi: konačnodimenzionalni unitarni/Hilbertovi prostori $\subsetneqq$ Hilbertovi prostori $\subsetneqq$ unitarni prostori

Primjer 22. Neka je ℓ^2 skup svih kvadratno sumabilnih nizova realnih brojeva, tj. $\ell^2 = \{x = (\alpha_n) : \mathbb{N} \to \mathbb{R} : \text{red } \sum_{n\in\mathbb{N}} (\alpha_n)^2 \text{ konvergira}\}$. ℓ^2 ima strukturu vektorskog prostora s obzirom na standardno definirane operacije zbrajanja nizova te množenja niza skalarom. Nadalje, red realnih brojeva $\sum \alpha_n\beta_n$ konvergira (apsolutno). Stoga je dobro definirana funkcija $\langle\cdot|\cdot\rangle : \ell^2 \times \ell^2 \to \mathbb{R}$, $\langle(\alpha_n)|(\beta_n)\rangle = \sum_{n\in\mathbb{N}} \alpha_n\beta_n$ koja je očito skalarni produkt na ℓ^2. Može se pokazati da je ℓ^2 uz ovako definiran skalarni produkt potpun unitaran (tj. Hilbertov) prostor.

Primjer 23. Neka je $C([-1, 1])$ unitarni prostor iz Primjera 12, onda se može pokazati da $C([-1, 1])$ s **nije** potpun. Štoviše, čak i ako $C([-1, 1])$ proširimo do prostora Riemann integrabilnih funkcija (sa istim skalarnim produktom[3]) nećemo dobiti potpuni prostor. Problem je dakle u samom pojmu integrabilnosti - potrebno je koristiti Lebesgueov integral.

[2] Prostor izometričan potpunom prostoru nužno i sam mora biti potpun.

[3] Ovdje se javljaju problemi sa zahtijevom da $\langle f|f\rangle = 0 \implies f = 0$ pa je potrebno raditi sa klasama ekvivalencije kao u primjeru 24

Neka je X normirani prostor s normom $|\cdot|$. **Upotpunjenjem** prostora X nazivamo par $(\tilde{X}, A)$ koji se sastoji od normiranog prostora $\tilde{X}$ s normom $\|\cdot\|$ te preslikavanja $A : X \to \tilde{X}$ sa sljedećim svojstvima:

1. Preslikavanje $A : X \to \tilde{X}$ je linearno
2. Preslikavanje $A : X \to \tilde{X}$ je izometrija
3. Skup $A(X)$ je gust na $\tilde{X}$.
4. Normirani prostor $\tilde{X}$ s normom $\|\cdot\|$ je potpun.

Primjer 24. Neka je (X, Σ, μ) prostor mjere i $\mathcal{L}^2(X, \Sigma, \mu)$ skup svih Σ-izmjerivih funkcija $f : X \to \mathbb{R}$ sa svojstvom da je funkcija $f^2 : X \to \mathbb{R}$ Lebesgue integrabilna. Može se pokazati da je $\mathcal{L}^2(X, \Sigma, \mu)$ realan vektorski prostor. Definirajmo na $\mathcal{L}^2(X, \Sigma, \mu)$ relaciju $\sim$ na sljedeći način: $f \sim g \Leftrightarrow_{\text{def}} f = g(s.s)$ (funkcije f i g su jednake skoro svuda ako je skup svih točaka gdje se one razlikuju zanemariv). Relacija $\sim$ je relacija ekvivalencije na skupu $\mathcal{L}^2(X, \Sigma, \mu)$, a dvije funkcije pripadaju istoj klasi ako i samo ako se podudaraju svugdje osim eventualno na nekom zanemarivom podskupu od X. Kvocijentni skup $\mathcal{L}^2(X, \Sigma, \mu)/\sim$ označavamo $L^2(X, \Sigma, \mu)$, a kada je jasno o kojem se skupu X, σ-algebri Σ i mjeru μ radi, jednostavno pišemo L^2. U ovom primjeru od sada koristimo skraćenu notaciju, tj. pišemo jednostavno L^2. Na kvocijentnom skupu možemo definirati operacije zbrajanja klasa i množenja klasa skalarom na sljedeći način. Ako su $[f], [g] \in L^2$ i $\alpha \in \mathbb{R}$, onda definiramo $[f] + [g]$ kao onu klasu kojoj pripada funkcija $f + g$ i definiramo $\alpha[f]$ kao onu klasu kojoj pripada funkcija αf, tj. $[f] + [g] =_{\text{def}} [f + g]$ i $\alpha[f] =_{\text{def}} [\alpha f]$. Lako je pokazati da ovako definirano zbrajanje i množenje skalarom ne ovise o izboru predstavnika klasa te da s obzirom na ovako definirana preslikavanja, prostor L^2 ima strukturu realnog vektorskog prostora. Definirajmo funkciju $\langle\cdot|\cdot\rangle : L^2 \times L^2 \to \mathbb{R}$, $\langle [f]|[g] \rangle = \sqrt{\int fg d\mu}$. Lako se pokaže da je ovako definirano preslikavanje dobro definirano i skalarni produkt na L^2 - simetričnost je očita, linearnost slijedi iz linearnosti integrala, a pozitivna definitnost slijedi iz poznate činjenice o Lebesgueovom integralu: ako je f izmjeriva funkcija, onda je $\int f d\mu = 0$ ako i samo ako je $f = 0(s.s)$, odnosno ako i samo ako je $[f] = [0]$. Neka je $\|\cdot\|$ norma na L^2 inducirana gore navedenim skalarnim produktom, tj. $\|[f]\| = \int f^2 d\mu$. Može se pokazati da je ovako dobiveni prostor L^2 potpun, tj. L^2 je Hilbertov (i Banachov) prostor. Također, može se pokazati da ukoliko uzmemo na $X = [0, 2]$ Lebesgueovu mjeru, dobiveni prostor L^2 bit će upotpunjenje prostora $C([0, 2])$ s obzirom na skalarni produkt definiran s $\langle x|y \rangle = \int_{[0,2]} x(t)y(t)dt$, pri čemu je za navedeni integral svejedno je li Lebesgueov ili Riemannov budući da se na neprekidnim funkcijama na segmentu podudaraju.

1.5 Koordinate i baze

Definicija 25. Neka je X Hilbertov prostor i $\mathcal{O} = \{e_i : i \in I\} \subseteq X$ ortonormirani skup. Kažemo da je $\mathcal{O}$ **Hilbertova baza** prostora X ako za svaki vektor $x \in X$ postoji jedinstvena familija $(\alpha_i)_{i \in I}$ realnih brojeva (tzv. koordinata od x) od kojih je najviše prebrojivo mnogo različito od 0 te vrijedi

$$x = \sum_{i \in I} \alpha_i e_i$$

Valja primijetiti da iz bilinearnosti i neprekidnosti skalarnog umnoška odmah imamo $\alpha_i = \langle x|e_i \rangle$ pa se gornja suma može pisati i u obliku $x = \sum_{i \in I} \langle x|e_i \rangle e_i$.

Lema 26. Neka je X Hilbertov prostor, neka je $\mathcal{O} = (e_i)_{i \in I}$ ortonormirani skup vektora i $\mathrm{Span}(\mathcal{O})$ skup svih konačnih linearnih kombinacija vektora e_i. Sjedeće je ekvivalentno:

1. $\mathrm{Span}(\mathcal{O})$ je gust u X, tj. $\overline{\mathrm{Span}(\mathcal{O})} = X$
2. Ako $\langle x|e_i \rangle = 0$ za sve e_i, onda $x = 0$ (x je određen svojim koordinatama)
3. $(e_i)_{i \in I}$ je Hilbertova baza.
4. (**Parsevalova jednakost**) $\|x\|^2 = \sum_{i \in I} |\langle x|e_i \rangle|^2$

Dokaz. Pokazat ćemo da $2 \Longrightarrow 3 \Longrightarrow 4 \Longrightarrow 2$ tako da su 2, 3 i 4 ekvivalentni. Trivijalno je da $3 \Longrightarrow 1$ i još ćemo pokazati da $1 \Longrightarrow 2$.

- (**1** $\implies$ **2**) Ako je $\overline{\mathrm{Span}(\mathcal{O})} = X$, onda za svaki $x \in X$ postoji niz $(x_n)_{n\in\mathbb{N}}$ konačnih linearnih kombinacija koji konvergira u x. Neka je $\langle x|e_i\rangle = 0$ za sve $i \in I$. Po neprekidnosti skalarnog umnoška, $\langle x|x\rangle = \lim_{n\to\infty}\langle x_n|x\rangle$. Kako je svaki x_n konačna linearna kombinacija vektora e_i, iz ovoga slijedi da $\langle x|x_n\rangle = 0$ pa nužno $\langle x|x\rangle = 0$, tj. $x = 0$.

- (**2** $\implies$ **3**) Neka je $x \in X$. Po Besselovoj nejednakosti znamo da postoji najviše prebrojivo mnogo $i_1, i_2, ...$ za koje $\langle x|e_{i_j}\rangle \neq 0$ i za koje red $\sum_{j\in\mathbb{N}} |\langle x|e_{i_j}\rangle|^2$ konvergira. Sa druge strane (za recimo $m > n$)

$$\left\|\sum_{j=1}^{m}\langle x|e_{i_j}\rangle e_{i_j} - \sum_{j=1}^{n}\langle x|e_{i_j}\rangle e_{i_j}\right\| = \left\|\sum_{j=n}^{m}\langle x|e_{i_j}\rangle e_{i_j}\right\| = \sum_{j=n}^{m}|\langle x|e_{i_j}\rangle|^2$$

  Desna strana konvergira u 0 kako $m, n \to \infty$. Zaključujemo da su parcijalne sume reda $\sum_{i\in I}\langle x|e_i\rangle e_i = \sum_{j\in\mathbb{N}}\langle x|e_{i_j}\rangle e_{i_j}$ Cauchyjev niz u X te red $\sum_{i\in I}\langle x|e_i\rangle e_i$ konvergira.

  Sada kako je $\langle x - \sum_{i\in I}\langle x|e_i\rangle e_i|e_j\rangle = 0$, za sve $j \in I$, po **2** slijedi da je $x - \sum_{i\in I}\langle x|e_i\rangle e_i = 0$, tj. $x = \sum_{i\in I}\langle x|e_i\rangle e_i$.

- (**3** $\implies$ **4**)

  Primijetimo da

$$\|x\|^2 - \sum_{j=1}^{n}|\langle x|e_{i_j}\rangle|^2 = \left\|x - \sum_{j=1}^{n}|\langle x|e_{i_j}\rangle e_{i_j}\right\|^2$$

  Po **3** desna strana konvergira u 0 kako $n \to \infty$, stoga slijedi tvrdnja.

- (**4** $\implies$ **2**) Očito

□

Teorem 27. Svaki Hilbertov prostor ima Hilbertovu bazu.

Dokaz. Ovo je standardna primjena Zornove leme. Familija ortonormirani skupova Ort od X je uređena s obzirom na inkluziju. Kako svaki lanac ima gornju granicu (unija svih njegovih elemenata), Zornova lema daje maksimalni element $\mathcal{O} \in Ort$ (koji nije manji ni od jednog drugog). $\mathcal{O}$ je svakako ortonormirani skup kao član familije Ort. Pokažimo da je i baza. Po prethodnoj lemi je dovoljno pokazati da je skup svih konačnih linearnih kombinacija $\mathrm{Span}(\mathcal{O})$ gust u X. Kako je $\overline{\mathrm{Span}(\mathcal{O})}$ zatvoreni potprostor od X po teoremu 19 imamo $X = \overline{\mathrm{Span}(\mathcal{O})} \oplus \overline{\mathrm{Span}(\mathcal{O})}^{\perp}$. Kada bismo imali $\overline{\mathrm{Span}(\mathcal{O})} \neq X$, onda $\overline{\mathrm{Span}(\mathcal{O})}^{\perp} \neq \{0\}$, ali u tom slučaju samo odaberemo jedinični $y \in \overline{\mathrm{Span}(\mathcal{O})}^{\perp}$ i ubacimo ga u $\mathcal{O}$ tako narušavajući maksimalnost ortonormiranog skupa $\mathcal{O}$. Dakle, $\overline{\mathrm{Span}(\mathcal{O})} = X$.

□

Primjer 28. Neka je A proizvoljan (indeksni) skup. **Generaliziranim nizom** nazivamo svaku funkciju $x : A \to \mathbb{R}$ (za $\alpha \in A$, vrijednost $x(\alpha)$ označavamo s x_α a sam x označavat ćemo (x_α)).

Neka je $\ell^2(A)$ skup svih generaliziranih nizova realnih brojeva koji su kvadratno sumabilni, tj. $\ell^2(A) = \{(x_\alpha) :$ generalizirani red $\sum_A x_\alpha^2$ konvergira$\}$. Skup $\ell^2(A)$ očito ima strukturu vektorskog prostora i lako se pokaže da je preslikavanje $\langle\cdot|\cdot\rangle : \ell^2(A) \times \ell^2(A)\mathbb{R}$, $\langle(x_\alpha)|(y_\alpha)\rangle = \sum_{\alpha\in A} x_\alpha y_\alpha$ dobro definirano i skalarni produkt na $\ell^2(A)$.

Teorem 29 (Riesz-Fischer)**.** Neka je X Hilbertov prostor sa Hilbertovom bazom $(e_i)_{i\in A}$. Onda je preslikavanje $x \mapsto \alpha_i = \langle x|e_i\rangle$ koje šalje $x \in X$ u njegovu koordinatu bijektivna izometrija $X \to \ell^2(A)$.

Ovaj teorem nam kaže da je prethodni primjer na određeni način daje *jedinu* moguću familiju primjera Hilbertovih prostora.

Dokaz. Iz definicije Hilbertove baze i Parsevalove jednakosti je jasno da preslikavanje koje šalje $x \in X$ u njegovu familiju koordinata $(\alpha_i)_{i\in A}$ mora biti injektivna (i linearna) izometrija. Preostaje pokazati jedino surjektivnost istog preslikavanja. Ako je $(\alpha_i)_{i\in A} \in \ell^2(A)$ kvadratno sumabilni niz, onda je $\sum_{i\in A}|\alpha_i|^2 < \infty$ pa ima najviše prebrojivo mnogo članova različito od 0. U dokazu leme 26 (2 $\implies$ 3) smo pokazali da su parcijalne sume reda $\sum_{i\in A}\alpha_i e_i$ Cauchyjev niz u X pa konvergiraju u neki $x \in X$. Taj x ima koordinate $(\alpha_i)_{i\in A}$ i surjektivnost je dokazana. □

Teorem 30 (Separabilni Hilbertov prostor)**.** Sljedeće je ekvivalentno:

1. X ima prebrojivu Hilbertovu bazu
2. X je separabilan, odnosno ima prebrojiv gusti podskup

Dokaz. Pretpostavimo da Hilbertov prostor X ima prebrojivu bazu $\mathcal{O}$. Tada je skup $S = \operatorname{Span}(\mathcal{O}) \subseteq X$ svih konačnih linearnih kombinacija gust podskup od X. Neka je D skup svih konačnih linearnih kombinacija elemenata iz O s racionalnim koeficijentima. Očito je D prebrojiv skup. Dokažimo još da je D gust na X. Neka je $x \in S$ proizvoljan. Tada postoje $e_{n_1}, ..., e_{n_k} \in O$ te skalari $\alpha_1, ..., \alpha_k \in \mathbb{R}$ takvi da je $x = \sum_{i=1}^{k} \alpha_i e_{n_i}$. Za svaki $i \in \{1, 2, ..., k\}$ postoji niz racionalnih brojeva $(q_l^i)_{l\in\mathbb{N}}$ koji konvergira prema α_i. Promotrimo niz (x_l), $x_l = \sum_{i=1}^{k} q_l^i e_{n_i}$. Niz (x_l) je niz u D koji konvergira prema x pa je $x \in \overline{D}$. Time je pokazano da je $S \subseteq \overline{D}$ pa kako je $X = \overline{S} \subseteq \overline{\overline{D}} = \overline{D} \subseteq X$ pa je $\overline{D} = X$, tj. D je gust prebrojiv podskup od X. Time je dokazano da je X separabilan.

Obratno, neka je X separabilan Hilbetov prostor te neka je $A = \{a_n : n \in \mathbb{N}\}$ gust prebrojiv podskup od X. Bez smanjenja općenitosti možemo pretpostaviti da je A indeksiran tako da $a_1 \neq 0$. Induktivno ćemo konstruirati podniz (a_{n_k}) niza (a_n) takav da je skup $\{a_{n_k} : k \in \mathbb{N}\}$ linearno nezavisan te za svaki $k \in \mathbb{N}$ vrijedi $\operatorname{Span}\{a_{n_1}, ..., a_{n_k}\} = \operatorname{Span}\{a_1, ..., a_{n_k}\}$. Budući da je $a_1 \neq 0$, možemo uzeti $a_{n_1} := a_1$. Pretpostavimo da smo za neki $k \in \mathbb{N}$ definirali $a_{n_1}, ..., a_{n_k}$. Prostor razapet tim vektorima je konačnodimenzionalan potprostor od X pa je zatvoren, tj. $X \setminus \operatorname{Span}\{a_{n_1}, ..., a_{n_k}\}$ je neprazan i otvoren podskup pa (po karakterizaciji gustog podskupa) siječe skup A u nekom x_l. Očito je $l > n_k$ (za svaki takav l). Neka je $n_{k_1} := \min\{l \in \mathbb{N} : x_l \in \ X \setminus \operatorname{Span}\{a_{n_1}, ..., a_{n_k}\}$. Po konstrukciji je očito skup $\{a_{n_1}, ..., a_{n_k}, a_{n_{k+1}}\}$ linearno nezavisan, a zbog odabira n_{k+1} kao najmanjeg indkesa l za koji je $x_l \in \ X \setminus \operatorname{Span}\{a_{n_1}, ..., a_{n_k}\}$, slijedi da je $\operatorname{Span}\{a_{n_1}, ..., a_{n_{k+1}}\} = \operatorname{Span}\{a_1, ..., a_{n_{k+1}}\}$. Time je proveden korak induktivne izgradnje. Dobiveni podniz (a_{n_k}) sastoji se od linearno nezavisnih vektora i očito mu je Span jednak $\operatorname{Span}(A)$. Dobiveni podniz ortonormiramo Gramm-Schmidtovim postupkom, tj. definiramo induktivno $e_1 := \frac{x_{n_1}}{\|x_{n_1}\|}$, a za $k \in \mathbb{N}$ stavljamo

$$e_{n_{k+1}} := \frac{1}{\left\| x_{n_{k+1}} - \sum_{i=1}^{n} \langle x_{n_{k+1}} | e_i \rangle e_i \right\|} \left(x_{n_{k+1}} - \sum_{i=1}^{n} \langle x_{n_{k+1}} | e_i \rangle e_i \right).$$

Dobiveni skup $E = \{e_k : k \in \mathbb{N}\}$ je ortonormiran i vrijedi $X = \overline{A} \subseteq \overline{\operatorname{Span}(A)} = \overline{\operatorname{Span}(A)} \subseteq \overline{E} \subseteq \overline{\operatorname{Span}(E)} = X$ pa je $\overline{\operatorname{Span}(E)} = X$ pa je E Hilbertova baza. $\square$

2 Teorija kernela

U ovoj sekciji uzmimo da $X = \mathbb{R}^n$ (ali čitatelj ako hoće, može zamijeniti za X proizvoljni lokalno kompaktni Hausdorffov prostor.)

2.1 Kerneli, Mercerov teorem

Definicija 31. Neka je $X = \mathbb{R}^n$ i $k \in L^2(X \times X)$ kvadratno integrabilna funkcija dvije varijable. Reći ćemo da je k **kernel** ako postoji preslikavanje (tzv. *feature map*) $\varphi : X \to \mathbb{H}$ u neki separabilni Hilbertov prostor $\mathbb{H}$ tako da $k(x,y) = \langle \varphi(x) | \varphi(y) \rangle$, odnosno ako sljedeći dijagram komutira:

$$\begin{tikzcd}
X \times X \arrow[d, "\phi\times\phi"'] \arrow[dr, "k"] & \\
H \times H \arrow[r, "\langle\cdot|\cdot\rangle"'] & \mathbb{R}
\end{tikzcd}$$

Primjer 32 (Polinomni kernel)**.** Neka je $c \geq 0$ realan broj. Definiramo preslikavanje $k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ pravilom

$$k(x,y) = (x^T y + c)^d.$$

Pokažimo da je ovo kernel, npr. za d=2. Primijetimo da je

$$k(x,y) = \left(\sum_{i=1}^{n} x_i y_i + c\right)^2 = \sum_{i=1}^{n} x_i^2 y_i^2 + \sum_{i=2}^{n}\sum_{j=1}^{i-1} (\sqrt{2} x_i x_j)(\sqrt{2} y_i y_j) + \sum_{i=1}^{n} (\sqrt{2c} x_i)(\sqrt{2c} y_i) + c^2,$$

te da u gornjoj sumi ima $n + \sum_{i=2}^{n} (i-1) + n + 1 = \frac{(n+1)(n+2)}{2}$ članova pa za preslikavanje $\varphi : \mathbb{R}^n \to \mathbb{R}^{\frac{(n+1)(n+2)}{2}}$ definirano pravilom

$$\varphi(x) = \left(x_n^2, ..., x_1^2, \sqrt{2} x_n x_1, \sqrt{2} x_{n-1} x_{n-2}), ..., \sqrt{2} x_{n-1} x_1, ..., \sqrt{2} x_2 x_1, \sqrt{2c} x_n, ..., \sqrt{2c} x_1, c\right)$$

vrijedi $k(x,y) = \varphi(x)\varphi(y)^T$ pa zaključujemo da je k kernel.

Primjer 33 (RBF kernel)**.** Definirajmo preslikavanje $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ pravilom

$$k(x,y) = \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right).$$

Preslikavanje k je kernel. Naime, definiramo li

$$\varphi_\tau : \mathbb{R} \to L^2(\mathbb{R}), \quad \varphi_\tau(x) : z \mapsto \frac{\exp(-(x-z)^2)/2\tau^2)}{\sqrt{2\pi}\tau}$$

, onda

$$\langle \varphi_\tau(x) | \varphi_\tau(y) \rangle = \int_{\mathbb{R}} \frac{\exp(-(x-z)^2)/2\tau^2)}{\sqrt{2\pi}\tau} \frac{\exp(-(y-z)^2)/2\tau^2)}{\sqrt{2\pi}\tau}\, dz = \frac{\exp(-(x-y)^2)/4\tau^2)}{2\sqrt{\pi}\tau}$$

Odaberemo li $\tau = \sigma/\sqrt{2}$, imamo (do na normalizaciju) željenu funkciju k.

Općenito, zanima nas koje uvjete mora zadovoljavati funkcija k kako bi bila kernel. Primijetimo, ako je k kernel $k(x,y) = \langle \varphi(x) | \varphi(y) \rangle$, onda je k **pozitivno definitan**, odnosno za svaku konačnu kolekciju $x_1, ..., x_n \in X$ i brojeva $c_1, ..., c_n \in \mathbb{R}$ vrijedi $\sum_{ij}^{n} k(x_i, x_j) c_i c_j \geq 0$

$$\sum_{i,j} k(x_i, x_j) c_i c_j = \sum_{i,j} \langle \varphi(x_i) | \varphi(x_j) \rangle c_i c_j = \langle \sum_i \varphi(x_i) c_i | \sum_j \varphi(x_j) c_j \rangle \geq 0.$$

Pokazat ćemo da je taj uvjet i dovoljan.

Za proizvoljnu kvadratno integrabilnu $k \in L^2(X \times X)$ definiramo tzv. **integralnu transformaciju** $K : L^2(X) \to L^2(X)$ na sljedeći način. Kvadratno integrabilnoj funkciji $\psi \in L^2(X)$ pridružujemo (kvadratno integrabilnu) funkciju $K\psi$,

$$(K\psi)(x) = \int_X k(x,y)\psi(y)dy.$$

Poseban slučaj su npr. Laplaceova $k(x,y) = e^{-xy}$ ili (u kompleksnom slučaju) Fourierova $k(x,y) = e^{2\pi i xy}$).

Definicija 34 (Simetrični operator)**.** Kažemo da je operator $A : \mathbb{H} \to \mathbb{H}$ simetričan ako $\langle Ax|y\rangle = \langle x|Ay\rangle$ za sve $x, y \in \mathbb{H}$.

Ako je $k(x,y) = k(y,x)$, onda je K simetrični operator s obzirom na L^2 normu:

$$\begin{aligned}
\langle\psi|K\phi\rangle &= \int_X \psi(x)\left(\int_X k(x,y)\phi(y)dy\right)dx = \int_X\left(\int_X k(x,y)\psi(x)\phi(y)dy\right)dx \\
&= \int_X\left(\int_X k(x,y)\psi(x)\phi(y)dx\right)dy = \int_X\left(\int_X k(y,x)\psi(x)\phi(y)dx\right)dy \\
&= \int_X \phi(y)\left(\int_X k(y,x)\psi(x)dx\right)dy = \langle K\psi|\phi\rangle
\end{aligned}$$

Prisjetimo se da je podskup K metričkog prostora **kompaktan** akko svaki beskonačan skup A ima gomilište, odnosno točku u koju neki (nestacionarni) niz iz A konvergira. U $\mathbb{R}^2$ se može pokazati da je skup kompaktan akko je omeđen i zatvoren. Dakle u konačnodimenzionalnom slučaju zatvorene kugle su kompaktne.

Definicija 35 (Kompaktni operator)**.** Operator $A : \mathbb{H} \to \mathbb{H}$ je kompaktan ako mu slika jedinične kugle $B \subset \mathbb{H}$ leži unutar nekog kompaktnog skupa K, tj. $A(B) \subset K$.

Usporedimo gornju definiciju sa omeđenim/ograničenim operatorom za kojeg samo zahtijevamo da mu slika jedinične kugle B leži unutar neke omeđene kogle (koja u beskonačnodimenzionalnom slučaju nije kompaktna)

Može se pokazati da je svaki integralni operator je kompaktan.

Teorem 36 (Spektralni teorem za kompaktne operatore)**.** Neka je $\mathbb{H}$ Hilbertov prostor i $A : \mathbb{H} \to \mathbb{H}$ kompaktni simetrični operator. Onda postoji Hilbertova baza ϕ_i ortonormiranih svojstvenih vektora od A sa pridruženim svojstvenim vrijednostima λ_i.

Konkretno, operator A onda ima oblik $A\psi = A\sum_i\langle\psi|\phi_i\rangle\phi_i = \sum_i\langle\psi|\phi_i\rangle A\phi_i = \sum_i\langle\psi|\phi_i\rangle\lambda_i\phi_i$

Teorem 37 (Mercer)**.** Neka je $k \in L^2(X \times X)$ simetrična, neprekidna te iščezava van nekog kompaktnog skupa $C \subset X$, onda je sljedeće ekvivalentno:

1. k je kernel
2. K je pozitivno definitan, odnosno $\langle\psi|K\psi\rangle \geq 0$
3. k je pozitivno definitan, odnosno za svaku konačnu kolekciju $x_1, ..., x_n \in X$ i brojeva $c_1, ..., c_n \in \mathbb{R}$ imamo $\sum_{ij}^n k(x_i, x_j)c_i c_j \geq 0$

Dokaz. Dokazat ćemo da $3 \implies 2 \implies 1 \implies 3$. Izgleda da je ideja standardna, ali autori su se pomogli izvorom [3]

- (**3** $\implies$ **2**) Za neprekidni k i neprekidni ψ parcijalne Riemannove sume moraju zadovoljavati

  $$\sum_{ij} k(x_i, x_j)\psi(x_i)\psi(x_j)\mu(E_i)\mu(E_j) \geq 0$$

  Uzimajući limes vidimo da i

  $$\int_{X\times X} k(x,y)\psi(x)\psi(y)dxdy \geq 0$$

  kako se svaka $\psi \in L^2(X)$ može aproksimirati (u L^2 normi) neprekidnim funkcijama koje iščezavaju van nekog kompaktnog skupa, zaključujemo da ista relacija vrijedi za sve $\psi \in L^2$.

- **(2 $\Longrightarrow$ 1)**
  - Po spektralnom teoremu za kompaktne operatore imamo

    $$K\psi = \sum_{i=0}^{\infty} \lambda_i \langle \phi_i | \psi \rangle \phi_i,$$

    gdje su ϕ_i ortonormirani svojstveni vektori, a λ_i svojstvene vrijednosti operatora K i konvergencija je u L^2 normi.
    Sada vidimo da

    $$\lambda_i = \lambda_i \langle \phi_i | \phi_i \rangle = \langle \phi_i | \lambda_i \phi_i \rangle = \langle \phi_i | K \phi_i \rangle \geq 0$$

  - Primijetimo da je za $\alpha, \beta \in L^2(X)$ (npr. direktno iz Fubini-Tonellijevog teorema) funkcija $\alpha \otimes \beta(x,y) = \alpha(x)\beta(y)$ kvadratno integrabilna na $X \times X$ te imamo $\langle \alpha \otimes \beta | \alpha' \otimes \beta' \rangle = \langle \alpha | \alpha' \rangle \langle \beta | \beta' \rangle$. Nadalje:

    $$\langle k | \alpha \otimes \beta \rangle = \int_{X \times X} k(x,y) \alpha(x) \beta(y) dx dy = \int_X \left(\int_X k(x,y) \alpha(x) dx \right) \beta(y) dy = \langle K\alpha | \beta \rangle$$

  - Kako je sada $\phi_i \otimes \phi_i$ ortonormirani skup u $L^2(X \times X)$, po Besselovoj nejednakosti vidimo da su svojstvene vrijednosti kvadratno sumabilni niz:

    $$\sum_i \lambda_i^2 = \sum_i |\langle \phi_i | K \phi_i \rangle|^2 = \sum_i \langle k | \phi_i \otimes \phi_i \rangle \leq \|k\|^2 < \infty$$

    Dakle, red $\sum_i \lambda_i \phi_i \otimes \phi_i$ konvergira u L^2 normi i to prema k:

    $$\langle \sum_i \lambda_i \phi_i \otimes \phi_i | \alpha \otimes \beta \rangle = \sum_i \lambda_i \langle \phi_i \otimes \phi_i | \alpha \otimes \beta \rangle = \sum_i \lambda_i \langle \phi_i | \alpha \rangle \langle \phi_i | \beta \rangle = \langle K\alpha | \beta \rangle = \langle k | \alpha \otimes \beta \rangle$$

    što pokazuje da k i red imaju iste koordinate pa su jednake (do na skup mjere 0, tj. klasu ekvivalencije)[4].
  - Sada koristimo činjenicu da je k neprekidna i iščezava van kompaktnog skupa u X (do sada smo samo koristili da $k \in L^2(X \times X)$).
    U tom slučaju su svojstveni vektori ϕ_i operatora K neprekidni. Naime iz $\phi_i = \frac{1}{\lambda_i} K \phi_i$ imamo:

    $$\begin{aligned} \phi_i(y) - \phi_i(x) &= \frac{1}{\lambda_i} \left(\int_X k(y,t) \phi_i(t) dt - \int_X k(x,t) \phi_i(t) dt \right) \\ &= \frac{1}{\lambda_i} \int_X (k(x,t) - k(y,t)) \phi_i(t) dt \end{aligned}$$

    Sada:

    $$|\phi_i(y) - \phi_i(x)| \leq \frac{1}{\lambda_i} \int_X |k(x,t) - k(y,t)| \phi_i(t) dt = \frac{1}{\lambda_i} \int_C |k(x,t) - k(y,t)| \phi_i(t) dt$$

    jer $t \mapsto k(x,t)$ iščezava van nekog kompaktnog skupa C pa možemo integraciju urušiti po C.
    Konačno, za proizvoljni $\varepsilon > 0$ možemo pronaći točke x, y toliko blizu da:

    $$\frac{1}{\lambda_i} \int_C |k(x,t) - k(y,t)| \phi_i(t) dt \leq \varepsilon \frac{1}{\lambda_i} \int_C \phi_i(t) dt$$

    Kako se ova veličina može učiniti po volji malom (odabirom dovoljno malog ε), svojstveni vektori su neprekidni.

[4] Valja primijetiti da ako je ϕ_i Hilbertova baza za prostor $L^2(X)$, onda je $\phi_i \otimes \phi_j$ Hilbertova baza za $L^2(X \times X)$ - ovo je jednostavno pokazati ako primijetimo da su linearne kombinacije produkta karakterističnih funkcija guste u $L^2(X \times X)$.

– Sada dokažimo da red $\sum_i \lambda_i \phi_i \otimes \phi_i$ konvergira u k uniformno (pa samim time i po točkama). Naime, prvo primijetimo da je:

$$k_n(x,y) = k(x,y) - \sum_i^n \lambda_i \phi_i(x)\phi_i(y)$$

neprekidan (očito po neprekidnosti ϕ_i i k) i njegova pripadajuća integralna transformacija K_n je pozitivno definitna. Naime, $k_n = k - \sum_{i=1}^n \lambda_i \phi_i \otimes \phi_i = \sum_{i=n+1}^\infty \lambda_i \phi_i \otimes \phi_i$, stoga $\langle K_n \alpha | \alpha \rangle = \langle k_n | \alpha \otimes \alpha \rangle = \sum_{i=n+1}^\infty \lambda_i \langle \phi_i | \alpha \rangle^2 \geq 0$

Iz ovoga možemo zaključiti $k_n(x,x) \geq 0$. Naime, kada bi slučajno za neki $x \in X$ imali $k_n(x,x) < 0$, onda bi iz neprekidnosti $k_n(x,y) < 0$ na čitavoj okolini U od (x,x) pa za npr. $\psi = \chi_U$ (koji iznosi 1 na U, a 0 svugdje drugo) imamo $\langle \psi | K_n \psi \rangle = \int_U k(x,y) dx dy < 0$, što je kontradikcija. Dakle:

$$\sum_i^n \lambda_i \phi_i(x)\phi_i(x) \leq k(x,x)$$

Sada uzimajući supremum po svim $x \in X$ (i limes po n) vidimo da:

$$\sum_i^\infty \lambda_i \phi_i(x)^2 \leq \sup_{x \in X} k(x,x)$$

Konačno, imamo:

$$|\sum_i^\infty \lambda_i \phi_i(x)\phi_i(y)|^2 \leq \left(\sum_i^\infty \lambda_i \phi_i(x)^2 \right) \left(\sum_i^\infty \lambda_i \phi_i(y)^2 \right) \leq (\sup_{x \in X} k(x,x))^2 \leq \infty$$

Iz čega zaključujemo da red $\sum_i \lambda_i \phi_i \otimes \phi_i$ konvergira uniformno u neku (neprekidnu) funkciju, koja mora biti gotovo svugdje jednaka k. Kako su i ta funkcija i k neprekidni, jednakost zapravo vrijedi svugdje.

– Konačno, definirajmo $\varphi_i : x \mapsto \sqrt{\lambda_i}\phi_i(x)$. Onda je za svaki $x \in X$ niz $\varphi(x) = (\varphi_i(x))_{i \in \mathbb{N}}$ kvadratno sumabilan pa je element Hilbertovog prostora ℓ^2. Sada lagano vidimo da

$$k(x,y) = \sum_i \sqrt{\lambda_i}\phi_i(x)\sqrt{\lambda_i}\phi_i(y) = \langle \varphi(x) | \varphi(y) \rangle_{\ell^2}$$

.

- (**1** $\implies$ **4**) Ako je $k(x,y) = \langle \varphi(x) | \varphi(y) \rangle$, onda

$$\sum_{i,j} k(x_i, x_j) c_i c_j = \sum_{i,j} \langle \varphi(x_i) | \varphi(x_j) \rangle c_i c_j = \langle \sum_i \varphi(x_i) c_i | \sum_j \varphi(x_j) c_j \rangle \geq 0$$

□

2.2 RKHS

Definicija 38. Neka je X proizvoljni neprazni skup i $\mathbb{H}$ Hilbertov prostor nekih funkcija $f : X \to \mathbb{R}$. Kažemo da je $\mathbb{H}$ **Hilbertov prostor reproducirajućeg kernela** (eng. *Reproducing kernel Hilbert space* ili **RKHS**) ako je preslikavanje (evaluacija u x ili Diracov delta) $\delta_x : \mathbb{H} \to \mathbb{R}$ definirano sa $f \mapsto f(x)$ neprekidno.

Valja napomenuti da L^2 nije valjani prostor funkcija na kojem je evaluacija δ_x dobro definirana - L^2 je prostor klasa ekvivalencije, sa druge strane ℓ^2 jest.

Ako je $\mathbb{H}$ RKHS, onda po Rieszovom teoremu postoji neki $k_x \in \mathbb{H}$ takav da vrijedi $f(x) = \delta_x(f) = \langle f | k_x \rangle$. Sada imamo $k_x(y) = \delta_y(k_x) = \langle k_x | k_y \rangle =: k(x,y)$. Ovo preslikavanje $k : X \times X \to \mathbb{R}$ zovemo **reproducirajućim kernelom**. Kako je k zaista kernel (u smislu kako smo ga originalno definirali), mora biti simetričan i pozitivno definitan. Vrijedi i svojevrsni obrat:

Teorem 39 (Moore)**.** Svaki simetrični i pozitivno definitni $k : X \times X \to \mathbb{R}$ ima jedinstveni Hilbertov prostor funkcija $f : X \to \mathbb{R}$ za koji je k reproducirajući kernel.

Vidimo da nam ovaj teorem daje još jednu karakterizaciju pozitivno definitnog kernela (uz one dane Mercerovim teoremom).

Dokaz.

- Neka je k reproducirajući kernel, onda očito $k(x,y) = \langle k_x | k_y \rangle = \langle k_y | k_x \rangle = k(y,x)$ pa je K simetrična. Također,

$$\sum_{ij} c_i c_j k(x_i, x_j) = \sum_{ij} c_i c_j \langle k_x | k_y \rangle = \langle \sum_i c_i k_{x_i} | \sum_j c_j k_{x_j} \rangle = \left\| \sum_i c_i k_{x_i} \right\|^2 \geq 0.$$

  za sve $c_1, ..., c_n \in \mathbb{R}$ i sve $x_1, ..., x_n \in X$ (i sve $n \in \mathbb{N}$).

- Obratno, ako je k simetrična i pozitivno definitna, onda definiramo $k_x = k(x, \cdot)$ te konstruiramo Hilbertov prostor H kojega funkcije k_x razapinju (ali neće nužno biti baza). Neka prvo $H_0 = \text{span}\{k_x \mid x \in X\}$, što je samo skup svih *konačnih* linearnih kombinacija funkcija k_x. Na H_0 definiramo skalarni umnožak $\langle k_x | k_y \rangle = k(x,y)$ i proširimo po linearnosti, tj. $\langle \sum_i a_i k x_i | \sum_j b_j k_{y_j} \rangle = \sum_{ij} a_i b_j k(x_i, y_j)$. Upotpunimo H_0 do $H = \overline{\text{span}\{k_x\}}$ (s obzirom na normu induciranu skalarnim umnoškom na H_0) i dodefiniramo skalarni umnožak (na jedinstveni način) tako da zadržimo neprekidnost skalarnog umnoška (tj. tražimo bilinearnost na prebrojivim sumama). H je onda Hilbertov prostor i svaki element $f \in H$ se može zadati kao $f = \sum_{i=1}^{\infty} a_i k_{x_i}$ za neke $a_i \in \mathbb{R}$ i $x_i \in X$. Stoga imamo $\langle f | k_x \rangle = \sum_i a_i \langle k_{x_i} | k_x \rangle = \sum_i a_i k_{x_i}(x) = f(x)$.

  Ovdje treba napomenuti da smo iskoristili činjenicu da konvergencija u topologiji RKHS implicira konvergenciju po točkama[5], što vidimo jednostavnom primjenom Cauchy-Schwarzove nejednakosti:

$$|f(x) - f_n(x)| = |\langle f - f_n | k_x \rangle| \leq \|f - f_n\| \, \|k_x\| \, .$$

- Jedinstvenost nije preteška. Recimo da postoji još jedan Hilbertov prostor G za koji je k reproducirajući kernel, onda $\langle k_x | k_y \rangle_H = k(x,y) = \langle k_x | k_y \rangle_G$ pa se po linearnosti G i H moraju slagati i na $H_0 = \text{span}\{k_x\}$. Kako je G potpun i sadržava H_0, mora sadržavati i njegovo upotpunjenje H, tj. $H \subset G$. Konačno, da i $G \subset H$ vidimo jer je $H \subset G$ zatvoren pa $G = H \oplus H^\perp$, tj. $g \in G \implies g = f + f^\perp$ za $f \in H$ i $f^\perp \in H^\perp$. Zapravo, $g(x) = \langle f + f^\perp | k_x \rangle = \langle f | k_x \rangle = f(x)$ jer $f^\perp \perp k_x$ ($k_x \in H$) pa $g = f$ i samim time $g \in H$, tj. $G = H$.

□

Teorem 40 (Teorem o reprezentatoru)**.** Neka je $k : X \times X \to \mathbb{R}$ simetrični i pozitivno definitni kernel sa reproducirajućim prostorom $\mathbb{H}$. Neka su dani podatci (dataset) $(x_1, y_1), ..., (x_n, y_n) \in X \times \mathbb{R}$, strogo rastuća funkcija g (regularizator) i neka funkcija (greške) $E : (X \times \mathbb{R}^2)^n \to \overline{\mathbb{R}} = [-\infty, \infty]$. Minimum tzv. regulariziranog empirijskog funkcionala rizika

$$f \mapsto E(x_1, y_1, f(x_1), ..., x_n, y_n, f(x_n)) + g(|f|)$$

je funkcija f^* oblika konačne sume $f^* = \sum_{i=1}^n \alpha_i k(\cdot, x_i)$ za neke $\alpha_i \in \mathbb{R}$.

Ovo reducira problem minimizacije funkcionala rizika iz beskonačnodimenzionalnog problema u konačnodimenzionalni.

Dokaz. Neka je $\varphi : X \to (X \to \mathbb{R})$, definiran sa $\varphi(x) : X \to \mathbb{R}, \quad \varphi(x) = k(\cdot, x) = k_x$. Onda $\varphi(x)(y) = k(x,y) = \langle \varphi(x) | \varphi(y) \rangle$. Kako je $\text{span}\{\varphi(x_1), ..., \varphi(x_n)\}$ konačnodimenzionalan, zatvoren je pa možemo pisati $f = \sum_i \alpha_i \varphi(x_i) + v$ za $v \in \text{span}\{\varphi(x_1), ..., \varphi(x_n)\}^\perp$. Onda vidimo da $f(x_i) = \langle \sum_j \alpha_j \varphi(x_j) + v | \varphi(x_i) \rangle = \sum_j \alpha_j \langle \varphi(x_i) | \varphi(x_i) \rangle$, što ne ovisi o v. Dakle, ni error funkcija E ne ovisi o v, a regularizacijski član $g(\|f\|)$ je pak manji što je v manji:

$$g\left(\|f\|\right) = g\left(\left\| \sum_i \alpha_i \varphi(x_i) + v \right\|\right) = g\left(\sqrt{\sum_i \alpha_i \varphi(x_i) + |v|^2}\right) \geq g\left(\left\| \sum_i \alpha_i \varphi(x_i) \right\|\right)$$

Dakle, svaki minimizator f^* mora imati $v = 0$, stoga $f^* = \sum_{i=1}^n \alpha_{i=1} \varphi(x_i) = \sum_{i=1}^n \alpha_i k(\cdot, x)$. □

[5] Ovo ne vrijedi općenito u Hilbertovim prostorima!

A Konačnodimenzionalno vs beskonačnodimenzionalno

Topologija konačnodimenzionalnih prostora nije pretjerano interesantna - sve norme induciraju istu topologiju. Naime, imamo:

Propozicija 41 (Ekvivalentnost normi)**.** Neka su $||\cdot||_1$ i $||\cdot||_2$ bilo koje norme na $\mathbb{R}^n$, onda postoje brojevi m i M za koje $m||x||_2 \leq ||x||_1 \leq M||x||_2$ za sve $x \in \mathbb{R}^n$.

Dokaz. Neka je $f(x) = ||x||_1/||x||_2$. f je onda očito dobro definirana na jediničnoj sferi ($||x||_2 = 0$ samo za $x = 0$). Primijetimo da svaka norma mora biti neprekidna u standardnoj euklidskoj topologiji: ako je standardna norma vektora x mala, onda mu je i svaka koordinata mala pa i svaka druga norma po nejednakosti trokuta ($||\sum_i x^i e_i|| \leq \sum_i |x^i|||e_i||$) mora biti mala na x. Dakle, i f je neprekidna. Kako je u $\mathbb{R}^n$ jedinična sfera kompaktna, f na njoj mora poprimiti maksimum M i minimum m. Ovo upravo znači da $m \leq ||x||_1/||x||_2 \leq M$ za svaki jedinični vektor x pa po homogenosti obje norme ($||\alpha x|| = |\alpha|||x||$) ista relacija vrijedi za svaki vektor $x \neq 0$. Iz ovoga sada imamo $m||x||_2 \leq ||x||_1 \leq M||x||_2$ za sve x. $\square$

Ključni dio dokaza je činjenica da je jedinična sfera (i kugla) u konačnodimenzionalnom prostoru kompaktna, tj. da je konačnodimenzionalni prostor lokalno kompaktan (svaka dovoljno mala zatvorena kuglica mora biti kompaktna). Ispostavlja se da nijedan beskonačnodimenzionalni prostor ne može biti lokalno kompaktan pa je lokalna kompaktnost topološki kriterij za konačnudimenzionalnost. U slučaju separabilnog Hilbertovog prostora ovo možemo vidjeti na sljedeći način. Pretpostavimo da imamo ortonormiranu bazu $(e_n)_{n\in\mathbb{N}}$ tako da za svaki $x \in \mathbb{H}$ imamo $x = \sum_n x^i e_i$, gdje je konvergencija u topologiji od $\mathbb{H}$. Sada svaki e_i leži na jediničnoj sferi, ali se e_i ne gomilaju (svaki leži na svojoj osi), stoga smo pronašli niz bez gomilišta pa jedinična sfera (i kugla) u $\mathbb{H}$ ne mogu biti kompaktne.

Literatura

[1] Gerald B. Folland, *Real Analysis: Modern Techniques and Their Applications*, Wiley, 2nd edition, 2007

[2] Walter Rudin, *Real and Complex Analysis*, McGraw-Hill Education, 3rd edition, 1986

[3] C. Ferreira and V. A. Menegatto, *Positive Definitness, Reproducing Kernel Hilbert Spaces and Beyond*, Annals of Functional Analysis, 2013.

[4] N. Aronszajn, *Theory of Reproducing Kernels*, Transactions of the American Mathematical Society, 1950.