4 Przestrzenie Hilberta (wstęp)

Zanim przejdziemy do iloczynów skalarnych, podam jeszcze jedno z zastosowań dualnego wzoru na normę (czyli twierdzenia "o wydobywaniu normy").

Prawdopodobnie znają Państwo pojęcie uzupełnienia $(\tilde{X}, \tilde{d})$ przestrzeni metrycznej (X, d) i jego konstrukcję. W przypadku przestrzeni unormowanych można krok po kroku wykazywać, że działania wektorowe z X mają jednoznaczne ciągłe przedłużenia do działań na $\tilde{X}$. W zbiorze wszystkich ciągów Cauchy'ego (x_n) w przestrzeni X wprowadzamy relację równoważności: $(x_n) \sim (y_n)$ gdy $\|x_n - y_n\| \to 0$ przy $n \to \infty$. $\tilde{X}$ jest przestrzenią ilorazową względem tej relacji, z normą $\|[(x_n)]\| := \lim \|x_n\|$. Izometryczne zanurzenie $X \to \tilde{X}$ przyporządkowuje wektorowi $x \in X$ klasę równoważności ciągu stałego $(x, x, x, \ldots)$. Każda przestrzeń unormowana jest więc izometryczna z gęstym podzbiorem przestrzeni Banacha.

Zamiast tej konstrukcji -możemy stosunkowo łatwo zanurzyć izometrycznie X w przestrzeń zupełną, a następnie zdefiniować $\tilde{X}$ jako domknięcie obrazu X w tym zanurzeniu. Każda przestrzeń dualna do przestrzeni unormowanej jest przestrzenią Banacha, bo wiemy już, że przestrzeń $\mathcal{B}(X, Y)$ jest zupełna, o ile Y jest zupełna. Tutaj $Y = \mathbb{K}$, czyli ciało skalarów $\mathbb{R}$ lub $\mathbb{C}$, które jest zupełne, stąd mamy zupełność $X^* = \mathcal{B}(X, \mathbb{K})$. Ale mamy zanurzenie kanoniczne

$$\iota : X \to X^{**}, \quad \iota(\phi) : x \mapsto \phi(x),$$

które jest (na mocy dualnego wzoru na normę) izometryczne.

Wniosek. Domknięcie $\iota(X)$ w X^{**} jest więc uzupełnieniem X.

Gdy X jest przestrzenią zupełną, to jej kanoniczny obraz $\iota(X)$ jest domknięty, jako zupełna podprzestrzeń w X^{**}. Podprzestrzeń ta może (ale nie musi) być równa X^{**}.

Definicja. Przestrzeń Banacha X jest refleksywna, gdy $\iota(X) = X^{**}$.

Na przykład, wszystkie przestrzenie skończenie wymiarowe, przestrzenie $L^p(\mu), \ell^p$ dla $1 < p < \infty$ są refleksywne. Z kolei, żadna z przestrzeni: ℓ^1, $\ell^\infty, L^1(\mu), L^\infty(\mu), C[a, b]$, $C^k[a, b]$ – nie jest refleksywna. Jednym z warunków równoważnych refleksywności jest, by każdy funkcjonał liniowy ciągły ϕ na X osiągał na jakimś z wektorów sfery jednostkowej wartość równą $\|\phi\|$. Np. na przestrzeni $C[0, 2]$ funkcjonał $\phi(f) := \int_0^1 f(t)\, dt - \int_1^2 f(t)\, dt$ nie osiąga swojej normy =2 na sferze. Można też podać warunki wystarczające dla refleksywności w terminach tzw. ścisłej wypukłości kuli jednostkowej. Nawet gdy przestrzeń X jest refleksywna, czyli "jest równa swojej przestrzeni drugiej dualnej", nie musi być ona izometryczna z pierwszą dualną. Wyjątek stanowi jedna z najbardziej regularnych -klasa przestrzeni Hilberta. Zacznijmy od przestrzeni z iloczynem skalarnym, których zupełności nie zakładamy. Tym razem będziemy głównie rozważać przypadek skalarów zespolonych (chociaż spora część wyników ma swój odpowiednik w przypadku przestrzeni nad ciałem $\mathbb{R}$).

Rozważać będziemy odwzorowania $\omega : X \times X \to \mathbb{C}$ (nazywane też *formami* na przestrzeni X). Przez $\overline{\omega(z, y)}$ oznaczamy liczbę sprzężoną do $\omega(z, y)$.

Definicje. Odwzorowanie ω nazywamy półtoraliniowym, (ang. *sesquilinear*) gdy jest ono liniowe względem pierwszej zmiennej i "anty-liniowe" względem drugiej, czyli gdy $\forall_{z \in X}$ odwzorowania $X \ni x \to \omega(x, z)$ oraz $X \ni y \to \overline{\omega(z, y)}$ są liniowe.

ω jest skośnie symetryczne, gdy $\forall_{x,y \in X}\, \omega(y, x) = \overline{\omega(x, y)}$. (Odwzorowania skośnie symetryczne i liniowe wzgl. pierwszej zmiennej są więc półtoraliniowe.)

Nieujemna (odp. dodatnia) określoność oznacza, że $\forall_{x \in X \setminus \{0\}}\, \omega(x, x) \geqslant 0$ (odp. $\omega(x, x) > 0$).

Iloczynem skalarnym nazywamy odwzorowanie półtoraliniowe, skośnie symetryczne i dodatnio określine. Używamy wówczas symbolu $\langle x, y \rangle$ zamiast $\omega(x, y)$. Przestrzeń z iloczynem skalarnym nazywamy przestrzenią unitarną (lub pre-hilbertowską). Jeśli dodatkowo względem normy $\|x\| := \sqrt{\langle x, x \rangle}$, to X nazywamy przestrzenią Hilberta.

Stosowanego w geometrii analitycznej w $\mathbb{R}^d$ symbolu $\vec{x}\cdot\vec{y}$ dla iloczynu skalarnego nie używamy, bo zazwyczaj wektory x, y będą funkcjami i symbol $x\cdot y$ oznacza ich "zwykły", a nie skalarny iloczyn.

Symbol $\Re\lambda$ oznacza część rzeczywistą liczby λ, czyli $\frac{1}{2}(\lambda+\bar{\lambda})$. (Nieco większa ogólność poniższego twierdzenia będzie nam później potrzebna.)

Twierdzenie. Niech $\Omega(x) := \omega(x,x)$ dla $x \in X$. Jeśli ω jest półtoraliniowe, to

1. $\Omega(x+y) = \Omega(x) + \omega(x,y) + \omega(y,x) + \Omega(y)$.
2. $\Omega(x+y) + \Omega(x-y) = 2(\Omega(x) + \Omega(y))$ (=tożsamość równoległoboku)
3. Gdy ponadto ω jest nieujemnie określone i skośnie symetryczne, to

   $$|\omega(x,y)|^2 \leqslant \Omega(x)\Omega(y). \text{ W szczególności, } |\langle x,y\rangle| \leqslant \|x\|\|y\|.$$

4. Gdy $\mathbb{K} = \mathbb{C}$, to $4\omega(x,y) = \Omega(x+y) - \Omega(x-y) + i\Omega(x+iy) - i\Omega(x-iy)$
5. Odwzorowanie $\|\cdot\|$ określone wzorem $\|x\| = \sqrt{\langle x,x\rangle}$ jest normą.

Uwagi: Gdy ω jest iloczynem skalarnym, to $\Omega(x) = \|x\|^2$ i teza (1) przybiera postać:

$$(1') \qquad \|x+y\|^2 = \|x\|^2 + 2\Re\langle x,y\rangle + \|y\|^2,$$

zaś teza (2) odpowiada klasycznej tożsamości równoległoboku dla normy euklidesowej w $\mathbb{R}^2$. Teza (3) dla iloczynu skalarnego, to tzw. Nierówność Schwarza (lub: Cauchy'ego-Buniakowskiego-Schwarza, skrótowo "CBS"):

$$(3') \qquad |\langle x,y\rangle| \leqslant \|x\|\|y\|.$$

Dopiero z niej wynika

$$\texttt{nierówność trójkąta:} \qquad \|x+y\| \leqslant \|x\| + \|y\|.$$

Teza (4) oznacza tzw. Wzór Polaryzacyjny. Ponieważ $i^2 = -1$, wzór polaryzacyjny zapisywany jest też w postaci

$$\omega(x,y) = \frac{1}{4}\sum_{k=1}^{4} i^k \Omega(x + i^k y).$$

Uwaga: W przestrzeniach nad ciałem $\mathbb{R}$ zachodzi on **jedynie dla form dwuliniowych symetrycznych**, przyjmując następujący kształt:

$$4\omega(x,y) = \Omega(x+y) - \Omega(x-y).$$

Dowód twierdzenia. Wzór (1) jest oczywistą konsekwencja dwuliniowości wzgl. skalarów rzeczywistych. Dodając go stronami do wzoru (1) dla $x-y$ otrzymamy (2).

Dowód(3) polega na zastosowaniu stale nieujemnego trójmianu kwadratowego $f(t)$ zmiennej $t \in \mathbb{R}$, gdzie $f(t) = \Omega(x+ty)$. Ze względu na skośną symetrię, stosując (1) mamy $f(t) = \Omega(x) + 2t\Re\omega(x,y) + t^2\Omega(y)$, co daje niedodatniość jego wyróżnika: $\Delta = (2\Re\omega(x,y))^2 - 4\Omega(x)\Omega(y) \leqslant 0$. Po elementarnym przekształceniu, mamy $|\Re\omega(x,y)| \leqslant \sqrt{\Omega(x)\Omega(y)}$. Dla pewnej liczby λ o module 1 mamy $|\omega(x,y)| = \lambda\omega(x,y) = \omega(\lambda x, y)$, natomiast $\Omega(\lambda x) = \Omega(x)$. Z ostatniej nierówności z λx wstawionym w miejsce x wynika więc (3). Dowód (4), choć prosty, jest dość żmudny. Dodajemy 4 składniki prawej strony (4), każdy wyrażony przez prawą stronę (1). Ponieważ $\Omega(\pm iy) = \Omega(y)$, składniki zawierające Ω wyzerują się. $\Omega(x+y) - \Omega(x-y)$ daje wartość $2\omega(x,y) + 2\omega(y,x)$, co wyjaśnia kłopoty w niesymetrycznym przypadku rzeczywistym. Analogicznie, $i\Omega(x+iy) - i\Omega(x-iy) = i(\omega(x,iy) + \omega(iy,x)) - i(\omega(x,-iy) + \omega(-iy,x)) = -i^2\omega(x,y) + i^2\omega(y,x) - i^2\omega(x,y) + i^2\omega(y,x) = 2\omega(x,y) - 2\omega(y,x)$. Dodając ten wynik do poprzednio wyliczonej sumy otrzymamy $4\omega(x,y)$.

Nierówność trójkąta dla normy, podniesiona stronami do kwadratu, to nierówność $\|x+y\|^2 \leqslant \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2$, której lewa strona jest równa $\|x\|^2 + 2\Re\langle x, y\rangle + \|y\|^2$, więc wystarczy zastosować nierówność "CBS", by porównać środkowe składniki. $\square$

Wniosek. W przestrzeni unitarnej X wzór $\phi_z(x) := \langle x, z\rangle$ określa funkcjonał liniowy, ciągły. Jego norma jest równa $\|z\|$. Odwzorowanie $X \ni z \to \phi_z \in X^*$ jest anty-liniową izometrią.
Dowód. Z nierówności Schwarza otrzymujemy wprost: $|\phi_z(x)| \leqslant \|z\|\|x\|$, stąd $\|\phi\| \leqslant \|z\|$. Nierówność nie może być tu ostra, gdyż dla $x = z$ mamy $\phi_z(z) = \|z\|^2 \leqslant \|\phi\|\|z\|$. Addytywna zależność jest oczywista, zaś
$\phi_{\lambda z}(x) = \langle x, \lambda z\rangle = \bar{\lambda}\phi_z(x)$. $\square$

(Wkrótce wykażemy, że na przestrzeni Hilberta każdy funkcjonał liniowy ciągły jest tej postaci. Aby to wykazać potrzebujemy pewnej konstrukcji: rzutu prostopadłego.)

4.1 Rzut prostopadły

Definicja. Jeśli M jest wypukłym, domkniętym podzbiorem przestrzeni unitarnej X, to wektor $y \in X$ jest rzutem prostopadłym wektora w na zbiór M, co oznaczamy $y = P_M w$, gdy $y \in M$ oraz $\|w - y\| = \mathrm{dist}(w, M)$.

Twierdzenie. Rzut na zbiór wypukły domknięty M w przestrzeni Hilberta zawsze istnieje i jest wyznaczony jednoznacznie. (Wystarczy tu jedynie założyć zupełność M jako podprzestrzeni metrycznej w przestrzeni unitarnej X, nawet bez zakładania zupełności X.)

Dowód. Dla (dowolnie) ustalonego wektora x istnieje taki ciąg (z_n) w zbiorze M, że dla $\delta := \mathrm{dist}(w, M)$ jest $\|w - z_n\| \to \delta$. Faktycznie, odległość od zbioru, to kres dolny odległości od jego punktów: $\delta = \inf\{\|w - z\| : z \in M\}$. Kres zbioru liczbowego jest zawsze granicą pewnego ciągu elementów tego zbioru. Jeśli istnieje $z_* := \lim z_n$, to z domkniętości M wynika, że $z_* \in M$, a z ciągłości normy -że $\|w - z_*\| = \delta$. Korzystając z tożsamości równoległoboku dla $x = w - z_k, y = w - z_m$ (przenosząc $\|x-y\|^2 = \|z_m - z_k\|^2$ na jedną stronę), mamy

$$\|z_m - z_k\|^2 = 2(\|w - z_k\|^2 + \|w - z_m\|^2) - \|2(w - \frac{z_m + z_k}{2})\|^2.$$

Z wypukłości M wynika, że $\frac{z_m+z_k}{2} \in M$, więc $\|w - \frac{z_m+z_k}{2}\| \geqslant \delta$, a stąd wynika nierówność

$$\|z_m - z_k\|^2 \leqslant 2(\|w - z_k\|^2 + \|w - z_m\|^2) - 4\delta^2 \to 0 \quad \texttt{przy} \quad m, k \to \infty.$$

Prawa strona dąży do zera, gdyż $\|w - z_k\|^2 + \|w - z_m\|^2 \to \delta^2 + \delta^2$. Z ostatniej nierówności (dla z_* w miejscu z_k, $z_\bullet$ w miejscu z_m) wynika też, że gdy dla jakiegoś drugiego wektora $z_\bullet$ mamy również $\|w - z_\bullet\| = \delta$, to $\|z_* - z_\bullet\|^2 \leqslant 4\delta^2 - 4\delta^2 = 0$. Teza o jednoznaczności jest więc wykazana. $\square$

Zauważmy, że istnienie elementów minimalizujących odległość można też wykazać w innych przestrzeniach unormowanych (niekoniecznie unitarnych) -np. gdy zbiór wypukły M jest zwarty. Ale z jednoznacznością mogą być już kłopoty: np. w $\mathbb{R}^2$ z normą "taksówkową"

$$\|(x,y)\|_1 := |x| + |y| \quad \texttt{dla} \quad M = \{(x,y) : x \in [0,1], x + y = 2\}$$

wektor $\vec{w} = (2,2)$ jest w jednakowej odległości $\delta = 2$ od każdego z punktów odcinka M (a nawet od punktów z dwukrotnie dłuższego odcinka odpowiadającego parametrom $x \in [0,2]$)! Zagadnienia minimalizacji odległości bada odgrywająca ważną rolę w matematyce stosowanej tzw. *teoria aproksymacji.* W *przestrzeniach ściśle wypukłych* (czyli takich, gdzie gdzie brzeg kuli nie zawiera żadnego niezdegenerowanego odcinka) można wykazać jednoznaczność elementów minimalizujących odległość. Jednak **najważniejszej** własności rzutu, uzasadniającej nazwę "prostopadły", zawartej w następującym twierdzeniu -już nie da się wykazać:

Wprowadźmy jeszcze oznaczenie dla relacji prostopadłości: $v \perp z$ oznacza $\langle v, z\rangle = 0$, zaś dla zbioru M niech $v \perp M$ oznacza, że $\forall_{z\in M}\ v \perp z$.

Twierdzenie.(CHARAKTERYZACJA RZUTU) Jeśli M jest wypukłym podzbiorem przestrzeni unitarnej X, zaś $z_* \in M$, to

1. $z_* = P_M w \Leftrightarrow \forall_{z\in M}\ \Re\langle w - z_*, z_* - z\rangle \geqslant 0$
2. Jeśli M jest podprzestrzenią liniową, to $z_* = P_M w \Leftrightarrow \forall_{z\in M}\ w - z_* \perp z$.

Innymi słowy, najbliższy punktowi w punkt podprzestrzeni M leży w kierunku prostopadłym do M: $w - P_M w \perp M$. Można też spotkać to twierdzenie pod nazwą "wariacyjna charakteryzacja rzutu" (- ale nie "wariancyjna"!)

Dowód. Wektor $z_t := z_* + t(z - z_*)$ należy do zbioru M dla $t \in [0, 1]$. (Gdy M jest podprzestrzenią liniową, to $\forall_{t\in\mathbb{R}}\ z_t \in M$.) Z definicji rzutu wynika, że odległość w od M jest osiągnieta w punkcie $z_* = z_0$. Innymi słowy, funkcja $\psi(t) := \|w - z_t\|^2$ osiąga wartość najmniejszą dla $t = 0$. Ze wzoru (1') wynika, że ta funkcja (jako trójmian kwadratowy zmiennej t) jest różniczkowalna. Skoro osiąga ona wartość najmniejszą w lewym końcu przedziału [0,1], to $\psi'(0) \geqslant 0$. W przypadku podprzestrzeni, jest to minimum na całej prostej $\mathbb{R}$ i wówczas $\psi'(0) = 0$. Więc wyliczmy:

$$\psi(t) = \|(w - z_*) + t(z_* - z)\|^2 = \|w - z\|^2 + 2t\Re\langle w - z_*, z_* - z\rangle + t^2\|z - z^*\|^2.$$

Stąd $\psi'(0) = 2\Re\langle w - z_*, z_* - z\rangle$, co wobec wiedzy o znaku $\psi'(0)$ daje pierwszą tezę. W przypadku podprzestrzeni liniowej, gdy z przebiega M, to również $u := z_* - z$ reprezentuje dowolny wektor z podprzestrzeni M. Ale funkcjonał $\mathbb{R}$-liniowy $M \ni u \to \Re\langle w - z_*, u\rangle$ jeśli nie zmienia znaku na podprzestrzeni M, to musi się tam zerować. W przypadku zespolonym - z zerowania się części rzeczywistej funkcjonału $\mathbb{C}$-liniowego $u \mapsto \langle w - z_*, u\rangle$ na M wynika zerowanie się tego funkcjonału na M, czyli prostopadłość wektora $w - z_*$ do M.

Implikację w przeciwnym kierunku wystarczy wykazać dla zbioru wypukłego. Minimalizacja odległości w punkcie z_* nastąpi, jeśli wykażemy, że dla dowolnie ustalonego $z \in M$ dla funkcji ψ zdefiniowanej powyżej będzie $\psi(0) = \|w - z_*\|^2 \leqslant \psi(1) = \|w - z\|^2$. Jeśli pochodna będzie nieujemna w (0,1), to ψ będzie niemalejąca i to zakończy dowód. Ale $\psi'(0) \geqslant 0$, zaś ψ'' jest stałą nieujemną, więc ψ' jest też niemalejąca. $\square$.

Wynika też stąd następujące

Twierdzenie.(O ROZKŁADZIE ORTOGONALNYM) Jeśli M jest domkniętą podprzestrzenią przestrzeni Hilberta X, to jej dopełnienie ortogonalne: $M^\perp := \{y \in X : y \perp M\}$ jest domkniętą podprzestrzenią, tworzącą wraz z M sumę prostą równą X. Innymi słowy, każdy wektor $x \in X$ ma dokładnie jeden rozkład w postaci $x = z + y$ dla pewnego $z \in M$, $y \perp M$. Oczywiście, $z = P_M x$.

(Dla dowodu wystarczy sprawdzić, czy wektor $y := x - P_M x$ jest prostopadły do M (to już wiemy) oraz, czy gdy dla innej pary wektorów $z_1 \in M, y_1 \perp M$ mamy również dla tego samego x rozkład $x = z_1 + y_1$, to $z_1 = P_M x, y_1 = y$. Ale $v \perp v \Rightarrow v = 0$, a taka relacja zachodzi dla $v = z - z_1 = w_1 - w \in M \cap M^\perp$. (Domkniętość M i to, że jest ona podprzestrzenią -wynika z faktu, że $M^\perp$ jest przecięciem (branym po $z \in M$) jąder ciągłych funkcjonałów liniowych ϕ_z, gdzie $\phi_z(x) = \langle x, z\rangle$.) $\square$

Jak się wkrótce okaże, zarówno istnienie rzutu na każdą domkniętą podprzestrzeń, jak i twierdzenie o rozkładzie ortogonalnym -a nawet istnienie przynajmniej jednego niezerowego wektora prostopadłego do M dla każdej podprzestrzeni domkniętej Ms o kowymiarze 1 w przestrzeni unitarnej X implikują już zupełność X.