

1 Przestrzenie z iloczynem skalarnym

1.1 Iloczyn skalarny

Niech V będzie przestrzenią wektorową nad ciałem \mathbf{K} (zakładamy, że jest to ciało \mathbf{R} lub \mathbf{C}). Odwzorowanie $V \times V \ni (x, y) \rightarrow (x|y) \in \mathbf{K}$ nazywamy *iloczynem skalarnym*, jeżeli dla dowolnych $x, y, y_1, y_2 \in V, \lambda_1, \lambda_2 \in \mathbf{K}$ mamy

1. $(x|\lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 (x|y_1) + \lambda_2 (x|y_2)$ (liniowość w drugim argumencie)

2. $\begin{cases} \text{dla } \mathbf{K} = \mathbf{R} & (x|y) = (y|x) \text{ (symetria)} \\ \text{dla } \mathbf{K} = \mathbf{C} & (x|y) = \overline{(y|x)} \text{ (hermitowskość)} \end{cases}$

3. $(x|x) > 0$ dla $x \neq 0$ (dodatniość).

(zauważmy, że z punktu 2 dla $\mathbf{K} = \mathbf{C}$ wynika, że $(x|x) \in \mathbf{R}$).

Innymi słowy,

- dla $\mathbf{K} = \mathbf{R}$, iloczyn skalarny to forma dwuliniowa symetryczna dodatnia (tzn. o dodatniej formie kwadratowej). Rzeczywista przestrzeń wektorowa wyposażona w iloczyn skalarny nazywa się *przestrzenią euklidesową*.
- dla $\mathbf{K} = \mathbf{C}$, iloczyn skalarny to forma liniowa w drugim argumencie i antyliniowa w pierwszym (takie formy nazywa się *półtoraliniowymi*), hermitowska (w sensie punktu 2) i dodatnia (w sensie punktu 3). Zespólna przestrzeń wektorowa wyposażona w iloczyn skalarny nazywa się *przestrzenią unitarną*.

Uwaga. Odwzorowanie $F : V \rightarrow W$ przestrzeni wektorowych nazywa się *antyliniowym*, jeśli jest addytywne i *anty-jednorodne*, tzn. $F(\lambda x) = \bar{\lambda} F(x)$.

Przykłady.

• $V = \mathbf{R}^n$; $(x|y) = x^1 y^1 + \dots + x^n y^n$ (kanoniczny iloczyn skalarny w \mathbf{R}^n)

• $V = \mathbf{C}^n$; $(x|y) = x^1 y^1 + \dots + x^n y^n$ (kanoniczny iloczyn skalarny w \mathbf{C}^n)

• $V = \mathbf{R}_n[\cdot]$ $(u|v) = \int_a^b u(x)v(x)dx$ (tu i poniżej $a, b \in \mathbf{R}, a < b$)

• $V = \mathbf{R}_n[\cdot]$ $(u|v) = \int_{-\infty}^{\infty} u(x)v(x)e^{-x^2} dx$

• $V = \mathbf{C}_n[\cdot]$ $(u|v) = \int_a^b \bar{u}(x)v(x)dx$

Iloczyn skalarny w przestrzeni V określa w naturalny sposób iloczyn skalarny w każdej podprzestrzeni V .

1.2 Norma określona przez iloczyn skalarny

Wielkość

$$\|x\| := \sqrt{(x|x)} \quad (1)$$

nazywamy *długością* wektora x . Długość posiada własności: $\|x\| = 0 \iff x = \mathbf{0}$, $\|\lambda x\| = |\lambda| \cdot \|x\|$.

Stwierdzenie Nierówność Schwarz. Dla $x, y \in V$ mamy:

$$|(x|y)| \leq \|x\| \cdot \|y\|, \quad (2)$$

przy czym równość ma miejsce wtedy i tylko wtedy, gdy x, y są proporcjonalne (tzn. $x = \alpha y$ dla pewnej liczby α).

Dowód: Jeśli $x = \mathbf{0}$, to teza jest prawdziwa. Załóżmy, że $x \neq \mathbf{0}$. Dowody dla $\mathbf{K} = \mathbf{R}$ i dla $\mathbf{K} = \mathbf{C}$ nieco się różnią; weźmy najspierw przypadek

$\mathbf{K} = \mathbf{R}$ Określmy funkcję f zmiennej rzeczywistej t następująco:

$$f(t) = (tx - y|tx - y) = t^2(x|x) - 2t(x|y) + (y|y)$$

f jest funkcją kwadratową w t , przyjmującą wartości nieujemne (bo iloczyn skalarny jest nieujemny). Zatem wyróżnik trójmianu kwadratowego f jest niedodatni, mamy więc:

$$\Delta = 4(x|y)^2 - 4(x|x)(y|y) \leq 0,$$

czyli nierówność (2). Równość $\Delta = 0$ oznacza istnienie pierwiastka (podwójnego) $t_0 \in \mathbf{R}$, czyli mamy: $f(t_0) = \|t_0 x - y\|^2 = 0$, skąd wynika, że $y = t_0 x$.

$\mathbf{K} = \mathbf{C}$ Zapiśmy najpierw liczbę $(x|y)$ w postaci trygonometrycznej: $(x|y) = |(x|y)|e^{i\phi}$, czyli

$$(e^{i\phi} x|y) = |(x|y)|.$$

(zastąpiliśmy więc wektor x przez wektor obrócony $e^{i\phi} x$.) Ważni słowa, stabilizatory PRZEKREŚT, w sję daty i y plowu. Bo teraz rozpatrzmy nieujemną funkcję zmiennej rzeczywistej t

$$\begin{aligned} f(t) &= (te^{i\phi} x - y|te^{i\phi} x - y) = t^2(x|x) - 2t\Re(x|y) + (y|y) \\ &= t^2(x|x) - 2t|(x|y)| + (y|y). \end{aligned}$$

Dalszy ciąg rozumowania jest analogiczny jak poprzednio.

Wniosek: $\|x + y\| \leq \|x\| + \|y\|$.

Dowód:

$$\begin{aligned} \|x + y\|^2 &= (x + y|x + y) = (x|x) + 2\Re(x|y) + (y|y) \leq \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

(powyższa nierówność wynika z nierówności Schwarz i nierówności dla liczb zespolonych: $\Re z \leq |z|$).

Jeśli ktoś zapamiętał z analizy, co to jest metryka, to łatwo zobaczy, że norma określa metrykę w przestrzeni V : $d(x, y) = \|x - y\|$. A jeśli ktoś się z tym pojęciem nie zetknął, ma okazję uczynić to teraz: Metryka na jakimś zbiorze X jest to funkcja $d: X \times X \rightarrow \mathbf{R}$ o własnościach: $\forall x, y, z \in X$: \bullet) $d(x, y) = d(y, x)$; $\bullet\bullet$) $d(x, y) \geq 0$, $d(x, y) = 0 \iff x = y$; $\bullet\bullet\bullet$) $d(x, y) \leq d(x, z) + d(z, y)$. Sprawdzenie dwóch pierwszych własności jest natychmiastowe, natomiast własność trzecia: $d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$.

1.3 Ortogonalność

Dla $\mathbf{K} = \mathbf{R}$ i niezerowych wektorów x i y mamy, dzięki nierówności Schwarza,

$$-1 \leq \frac{(x|y)}{\|x\| \cdot \|y\|} \leq 1$$

(znak równości ma miejsce, gdy x jest proporcjonalny do y). Określimy kąt między wektorami x oraz y jako takie $0 \leq \alpha \leq \pi$, że

$$\alpha = \frac{(x|y)}{\|x\| \cdot \|y\|}.$$

Kąt dla wektorów proporcjonalnych wynosi 0 lub π (zależnie od tego, czy mają one ten sam zwrot, czy przeciwny). Kąt wynosi $\frac{\pi}{2}$, jeśli $(x|y) = 0$. To motywuje do przyjęcia następującej definicji:

Definicja. Mówimy (zarówno w rzeczywistym, jak i zespolonym przypadku), że wektory x i y są do siebie *ortogonalne (prostopadłe)*, jeśli $(x|y) = 0$.

Zwróćmy uwagę, że jeżeli x, y są ortogonalne, to mamy:

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

równość ta – przez analogię z faktem z elementarnej geometrii – nazywana jest *twierdzeniem Pitagorasa*.

Definicja. Bazę w przestrzeni V nazywamy *ortogonalną*, jeśli jej wektory są parami do siebie ortogonalne.

Bazy ortogonalne mają wiele pożytecznych własności i ułatwiają szereg rachunków. Dlatego warto umieć znaleźć bazę ortogonalną. Odpowiedź na pytanie, jak to zrobić, daje

1.3.1 Ortogonalizacja Grama – Schmidta.

(działa ona zarówno w przypadku rzeczywistym, jak i zespolonym). Jeśli $\{e_1, \dots, e_n\}$ jest bazą w przestrzeni V , to baza f_1, \dots, f_n , otrzymana przez *rekurencję Grama-Schmidta*:

$$f_1 := e_1, \quad f_i = e_i - \sum_{j=1}^{i-1} \frac{(f_j|e_i)}{(f_j|f_j)} f_j, \quad (3)$$

jest ortogonalna.

Dowód jest indukcyjny. Przypuśćmy, że f_1, \dots, f_{i-1} są parami ortogonalne. Wtedy f_i jest ortogonalny do f_k dla $k = 1, \dots, i-1$, ponieważ

$$(f_k|f_i) = (f_k|e_i) - \sum_{j=1}^{i-1} \frac{(f_j|e_i)}{(f_j|f_j)}(f_j|f_i) = (f_k|e_i) - \frac{(f_k|e_i)}{(f_k|f_k)}(f_k|f_k) = 0.$$

Uwaga: W przypadku zespolonym, we wzorze (3) istotna jest prawidłowa kolejność czynników w iloczynach skalarnych $(f_i|e_j)$.

Przechodząc od bazy ortogonalnej $\{f_i\}$ do bazy unormowanej $\{h_i\}$, gdzie $h_i := f_i/\|f_i\|$, otrzymamy bazę ortogonalną złożoną z wektorów o długości jednostkowej:

$$(h_i|h_j) = \delta_{ij} = \begin{cases} 0 & \text{dla } i \neq j \\ 1 & \text{dla } i = j \end{cases}.$$

Bazę o tej własności nazywamy *bazą ortonormalną* (w skrócie: o.n.). W każdej (skończenie-wymiarowej) przestrzeni z iloczynem skalarnym istnieje baza o.n. We współrzędnych w bazie o.n. iloczyn skalarny przyjmuje postać kanoniczną:

$$(x|y) = \left(\sum_{i=1}^n x^i h_i \mid \sum_{j=1}^n y^j h_j \right) = \sum_{i,j=1}^n \bar{x}^i y^j (h_i|h_j) = \sum_{i,j=1}^n \bar{x}^i y^j \delta_{ij} = \sum_{i=1}^n \bar{x}^i y^i.$$

1.3.2 Dopełnienie ortogonalne podprzestrzeni.

Dla podzbioru $E \subset V$ weźmy

$$E^\perp := \{y \in V : (x|y) = 0 \text{ dla } x \in E\}$$

(innymi słowy, E^\perp jest zbiorem wektorów prostopadłych do E). Łatwo sprawdzić, że zbiór E^\perp jest podprzestrzenią V .

Stwierdzenie. Jeśli V jest skończenie wymiarową przestrzenią z iloczynem skalarnym, a E – podprzestrzenią, to $V = E \oplus E^\perp$.

Dowód: Jeśli $x \in E \cap E^\perp$, to $(x|x) = 0$, więc $x = 0$. Stąd $E \cap E^\perp = \{0\}$. Pokażemy, że $E + E^\perp = V$. Niech e_1, \dots, e_k będzie bazą w E i $\dim V = n$. Mamy: $E^\perp = \{y \in V : (e_j|y) = 0 \text{ dla } j = 1, \dots, k\}$, skąd $\dim E^\perp \geq n - k$, więc $\dim(E + E^\perp) = \dim E + \dim E^\perp \geq k + (n - k) = n$.

Wniosek: $(E^\perp)^\perp = E$, bo $E \subset (E^\perp)^\perp$ oraz $\dim E = n - (n - \dim E) = \dim(E^\perp)^\perp$.

1.3.3 Rzut ortogonalny (prostopadły) na podprzestrzeń

Z każdą podprzestrzenią E w V związany jest rzut P_E na E wzdłuż E^\perp (*rzut ortogonalny na E*).

Pokażemy, że jeśli (e_1, \dots, e_k) jest bazą o.n. w E , to

$$P_E x = \sum_{j=1}^k (e_j | x) e_j.$$

Istotnie, ponieważ

$$(e_l | x - \sum_{j=1}^k (e_j | x) e_j) = - \sum_{j=1}^k (e_j | x) (e_l | e_j) = (e_l | x) - (e_l | x) = 0$$

dla $l = 1, \dots, k$, to wektor $x - \sum_{j=1}^k (e_j | x) e_j$ należy do E^\perp , zatem

$$x = \sum_{j=1}^k (e_j | x) e_j + (x - \sum_{j=1}^k (e_j | x) e_j)$$

jest rozkładem x na składowe w sumie prostej $E \oplus E^\perp$.

Uwaga: Wzór (3) w rekurencji Grama-Schmidta ma postać

$$f_i = e_i - \sum_{j=1}^{i-1} (h_j | e_i) h_j = e_i - P_E e_i,$$

gdzie $h_j := f_j / \|f_j\|$, $E = \langle h_1, \dots, h_{i-1} \rangle$, zatem f_i jest rzutem prostopadłym e_i na $\langle h_1, \dots, h_{i-1} \rangle^\perp$.

1.4 Kilka pożytecznych faktów z geometrii metrycznej

1. Odległość punktu x od podprzestrzeni E definiujemy następująco:

$$d(x, E) := \inf_{y \in E} d(x, y).$$

Ponieważ zachodzi:

$$\begin{aligned} d(x, E)^2 &= \inf_{y \in E} \|x - y\|^2 = \inf_{y \in E} \|(x - P_E x) + (P_E x - y)\|^2 = \\ &= \inf_{y \in E} (\|(x - P_E x)\|^2 + \|(P_E x - y)\|^2) = \|x - P_E x\|^2 \end{aligned}$$

(w przedostatniej równości skorzystaliśmy z faktu, iż $x - P_E x \in E^\perp$, zaś $P_E x - y \in E$ oraz z twierdzenia Pitagorasa). Mamy więc wzór

$$d(x, E) = \|x - P_E x\| = \|P_{E^\perp} x\|. \quad (4)$$

WTRĘT: Podstawowe pojęcia geometrii afinicznej

Ozn.: Dla podzbioru Z w przestrzeni wektorowej V i elementu $x \in V$ wprowadzimy oznaczenie:

$$x + Z = Z + x := x + z : z \in Z.$$

Mamy własności:

- (a) $x + (y + Z) = (x + y) + Z$
- (b) Z – podprzestrzeń, $x \in Z \implies x + Z = Z$.

Definicja. Podprzestrzenią afiniczną w przestrzeni wektorowej V nazywamy podzbiór postaci

$$P = p_0 + L, \tag{5}$$

gdzie p_0 jest pewnym punktem V , zaś L – podprzestrzenią (wektorową) V .

Oczywiście $p_0 \in P$. Ponadto, punkt p_0 w (5) można zastąpić dowolnym innym punktem z P : jeśli $p'_0 \in P$, to $p'_0 = p_0 + l$ dla pewnego $l \in L$, zatem

$$p'_0 + L = (p_0 + l) + L = p_0 + (l + L) = p_0 + L = P.$$

Okazuje się (proste ćwiczenie), że podprzestrzeń L jest dla danej podprzestrzeni afinicznej P jednoznacznie wyznaczona. L będziemy nazywać *przestrzenią wektorów stycznych* do P (lub: *przestrzenią styczną* do P).

Kla $\dim L = 0, 1, 2$ podprzestrzeń afiniczna P nazywa się odpowiednio *punktem*, *prostą*, *płaszczyzną*.

Przykład: Zbiór rozwiązań równania

$$Tx = y,$$

gdzie $T \in L(V, W)$, $y \in W$, jest pusty albo jest podprzestrzenią afiniczną w V .

Mówimy, że podprzestrzenie afiniczne P_1 i P_2 są *równoległe*, jeśli $L_1 \subset L_2$ lub $L_2 \subset L_1$ (L_1, L_2 są tu przestrzeniami stycznymi odpowiednio do P_1 i P_2).

Przecięcie dowolnej rodziny podprzestrzeni afinicznych jest puste albo jest podprzestrzenią afiniczną.

Odwzorowanie liniowe $T \in L(V, W)$ ma następujące własności "zachowywania afiniczności":

- (a) jeśli P jest podprzestrzenią afiniczną (p.a.) w V , to $T(P)$ jest p.a. w W .
- (b) jeśli P jest podprzestrzenią afiniczną (p.a.) w W , to $T^{-1}(P)$ jest puste lub jest p.a. w V .

Uwaga: Powyższe własności przysługują również odwzorowaniom postaci:

$$F(v) = T(v) + w_0,$$

gdzie $T \in L(V, W)$, a $w_0 \in W$ (takie odwzorowania nazywają się *afinicznymi*).

2. **Odległość podprzestrzeni afinicznych** P i P' definiujemy następująco:

$$d(P, P') := \inf_{\substack{p \in P \\ p' \in P'}} d(p, p').$$

Używając przedstawienia

$$P = p_0 + L, \quad P' = p'_0 + L'$$

gdzie L i L' są podprzestrzeniami wektorowymi, otrzymamy

$$d(P, P') = \inf_{\substack{y \in L \\ y' \in L'}} \|p_0 + y - (p'_0 + y')\| = \inf_{\substack{y \in L \\ y' \in L'}} \|p_0 - p'_0 - (y + y')\| =$$

$$d(p_0 - p'_0, L + L') = \|P_{(L+L')^\perp}(p_0 - p'_0)\|.$$

W ostatniej równości skorzystaliśmy z wyprowadzonego wyżej wzoru (4).

Przykład. Odległość prostych skośnych w \mathbf{R}^3 .

3. **Objętość równoległościanu rozpiętego na k wektorach.**

Równoległościan rozpięty na wektorach v_1, \dots, v_k to zbiór $\{\lambda_1 v_1 + \dots + \lambda_k v_k : 0 \leq \lambda_j \leq 1\}$. Definiujemy jego *objętość* następującym wzorem indukcyjnym ("pole podstawy razy wysokość"):

$$\text{Vol}(v_1, \dots, v_k) := \text{Vol}(v_1, \dots, v_{k-1}) \cdot d(v_k, \langle v_1, \dots, v_{k-1} \rangle), \quad \text{Vol}(v_1) := \|v_1\|.$$

Macierz Grama układu wektorów v_1, \dots, v_k nazywamy macierz $G(v_1, \dots, v_k) \in \mathbf{R}^k_k$, której elementami są iloczyny skalarne $(v_i | v_j)$ dla $i, j = 1, \dots, k$.

Stwierdzenie. $\text{Vol}(v_1, \dots, v_k) = \sqrt{\det G(v_1, \dots, v_k)}$.

Dowód będzie indukcyjny. Dla $k = 1$ teza jest prawdziwa: $\text{Vol}(v_1) = \|v_1\| = \sqrt{(v_1 | v_1)} = \sqrt{\det G(v_1)}$. Pokażemy, że z założenia o prawdziwości tezy dla $k-1$ wynika jej prawdziwość dla k . Mamy:

$$\begin{aligned} \text{Vol}(v_1, \dots, v_k)^2 &= \text{Vol}(v_1, \dots, v_{k-1})^2 \cdot d(v_k, \langle v_1, \dots, v_{k-1} \rangle)^2 = \\ &= \det G(v_1, \dots, v_{k-1}) \cdot (v_k - P v_k, v_k - P v_k) = \det G(v_1, \dots, v_{k-1}, v_k - P v_k), \end{aligned}$$

gdzie P - rzut ortogonalny na $\langle v_1, \dots, v_{k-1} \rangle$ (skorzystaliśmy tu z (4) oraz z prostopadłości $v_k - P v_k$ i v_j dla $j = 1, \dots, k-1$). Ostatni wyznacznik jest równy $\det G(v_1, \dots, v_k)$ na mocy następującego lematu.

Lemat. Jeśli w_1, \dots, w_k są kombinacjami liniowymi wektorów v_1, \dots, v_k :

$$w_r = \sum_{i=1}^k a^i_r v_i, \quad r = 1, \dots, k$$

to

$$G(w_1, \dots, w_k) = A^T G(v_1, \dots, v_k) A,$$

gdzie A jest macierzą współczynników (a^i_r) , $i, r = 1, \dots, k$. W szczególności,

$$\det G(w_1, \dots, w_k) = (\det A)^2 G(v_1, \dots, v_k). \quad (6)$$

Dowód (lematu):

$$(w_r | w_s) = \left(\sum_i a^i_r v_i \mid \sum_j a^j_s v_j \right) = \sum_{ij} a^i_r a^j_s (v_i | v_j) = \sum_{ij} (a^T)_i^r (v_i | v_j) a^j_s.$$

Dokończenie dowodu stwierdzenia: Wystarczy wziąć $w_1 = v_1, \dots, w_{k-1} = v_{k-1}, w_k = v_k - P v_k$ i skorzystać z (6), biorąc pod uwagę, że w naszym przypadku $\det A = 1$.

Uwagi:

1. Ze wzoru (6) wynika, że permutowanie wektorów v_1, \dots, v_k nie zmienia $\det G(v_1, \dots, v_k)$, a więc też (p. ostatnie Stwierdzenie) $\text{Vol}(v_1, \dots, v_k)$.
2. Z definicji $\text{Vol}(v_1, \dots, v_k)$ i powyższego Stwierdzenia wynika jeszcze inna postać wyrażenia na odległość punktu x od podprzestrzeni $\langle v_1, \dots, v_k \rangle$:

$$d(x, \langle v_1, \dots, v_k \rangle) = \left(\frac{\det G(v_1, \dots, v_k, x)}{\det G(v_1, \dots, v_k)} \right)^{\frac{1}{2}}$$

Dla liniowo niezależnych v_1, \dots, v_k mianownik jest różny od zera (jako wyznacznik macierzy dodatniej formy kwadratowej).

3. Jeśli w (6) rozważać sytuację, gdy $k = n = \dim V$ i $v = (v_1, \dots, v_n)$ – baza o.n., to otrzymamy

$$\text{Vol}(w_1, \dots, w_n) = \sqrt{G(w_1, \dots, w_n)} = |\det A| = |\det(a_1, \dots, a_n)|,$$

gdzie $a_j = [w_j]^v$ – wektor kolumnowy złożony ze współrzędnych wektora w_j w bazie v . Widać stąd, że wartość bezwzględna formy objętości (por. rozdział o wyznacznikach) jest równa, z dokładnością do czynnika, objętości w sensie metrycznym (dlatego też noszą one wspólną nazwę 'objętość').