Iloczyn skalarny

Klasyfikacja wszystkich endomorfizmów (nad ciałem algebraicznie domkniętym) w języku postaci Jordana mimo niewątpliwego uroku ma też swoje wady – wiemy wprawdzie coś o geometrii każdego endomorfizmu z osobna, ale nie rozumiemy zbyt dobrze „struktury algebraicznej" stojącej za tym opisem geometrycznym. Mówiąc prościej: mając dwa endomorfizmy ϕ, ψ przestrzeni V, na przykład nad ciałem $\mathbb{C}$, możemy wyznaczyć bazy (i postaci) Jordana $\mathcal{J}_1$, $\mathcal{J}_2$ każdego z tych endomorfizmów. Uzyskamy informację geometryczną o każdym z nich. Ale nie wiemy za bardzo czym jest z geometrycznego punktu widzenia $\phi \circ \psi$. Jego macierzą nie musi być przecież iloczyn macierzy w postaci Jordana $M(\psi)_{\mathcal{J}_2}^{\mathcal{J}_2} \cdot M(\phi)_{\mathcal{J}_1}^{\mathcal{J}_1}$, który nie musi mieć nawet postaci Jordana.

Wiemy, że dla przestrzeni V wymiaru n nad K struktura $End(V)$ to tak naprawdę struktura macierzy $M_{n\times n}(K)$, ale jest tak tylko wtedy, gdy każdemu przekształceniu przypiszemy macierz w tej samej bazie. Wtedy złożeniu przekształceń odpowiada iloczyn macierzy. A więc musielibyśmy napisać $M(\psi)_{\mathcal{J}_1}^{\mathcal{J}_1} \cdot M(\phi)_{\mathcal{J}_1}^{\mathcal{J}_1}$ lub $M(\psi)_{\mathcal{J}_2}^{\mathcal{J}_2} \cdot M(\phi)_{\mathcal{J}_2}^{\mathcal{J}_2}$, albo niestety $M(\psi)_{\mathcal{J}}^{\mathcal{J}} \cdot M(\phi)_{\mathcal{J}}^{\mathcal{J}}$, dla jakiejś bazy $\mathcal{J}$. Takie podejście niestety nie zadziała. Aby spełnić powyższe wymagania przyjrzymy się za jakiś czas możliwości „rozkładania" przekształceń na pewne „przekształcenia podstawowe". Musimy też określić nowy rodzaj baz, wymagający rozważania dodatkowej struktury na przestrzeni liniowej. Struktura ta pozwala na mówienie o prostopadłości wektorów, i to nie tylko w kontekście geometrycznym.

Przez kilka wykładów zajmować się będziemy jedynie przestrzeniami liniowymi nad ciałem liczb rzeczywistych. Dlaczego? Dla ilustracji rozmaitych zagadnień potrzebujemy odróżniać skalary dodatnie od ujemnych. Później przyjrzymy się analogicznym strukturom definiowanym nad dowolnymi ciałami. Poprowadzi nas to do pięknych i głębokich rezultatów. Sama zaś prostopadłość, jak się okaże w trakcie całych Państwa studiów, jest pojęciem absolutnie fundamentalnym dla rozumienia wielu struktur matematycznych.

Definicja 47. *Niech V będzie przestrzenią liniową nad ciałem $\mathbb{R}$. Funkcję:*

$$\langle\,,\rangle : V \times V \to \mathbb{R}$$

nazywamy ILOCZYNEM SKALARNYM NA PRZESTRZENI V*, jeśli dla każdych $\alpha, \beta, \gamma, \delta \in V$ i $a, b, c, d \in \mathbb{R}$ zachodzi:*

(1)	$\langle a\alpha + b\beta, \gamma\rangle = a\langle \alpha, \gamma\rangle + b\langle \beta, \gamma\rangle$	*liniowość względem pierwszej zmiennej,*
(2)	$\langle \alpha, c\gamma + d\delta\rangle = c\langle \alpha, \gamma\rangle + d\langle \alpha, \delta\rangle$	*liniowość względem drugiej zmiennej,*
(3)	$\langle \alpha, \beta\rangle = \langle \beta, \alpha\rangle$	*symetria,*
(4)	$\alpha \neq 0 \Rightarrow \langle \alpha, \alpha\rangle > 0$	*dodatnia określoność.*

Odnotujmy kilka trywialnych własności dowolnego iloczynu skalarnego $\langle\,,\rangle$ na przestrzeni rzeczywistej.

- $\langle a \cdot \alpha, b \cdot \beta\rangle = ab\langle \alpha, \beta\rangle$, dla dowolnych $, a, b \in \mathbb{R}$ oraz $\alpha, \beta \in V$,
- $\langle \alpha, 0\rangle = \langle \alpha, \alpha - \alpha\rangle = \langle \alpha, \alpha\rangle - \langle \alpha, \alpha\rangle = 0$, dla dowolnego $\alpha \in V$.
- $\langle \alpha + \beta, \alpha + \beta\rangle = \langle \alpha, \alpha + \beta\rangle + \langle \beta, \alpha + \beta\rangle = \langle \alpha, \alpha\rangle + 2\langle \alpha, \beta\rangle + \langle \beta, \beta\rangle$ dla dowolnych $\alpha, \beta \in V$.

Możliwych iloczynów skalarnych na danej przestrzeni jest całkiem sporo. Zobaczmy kilka przykładów dotyczących przestrzeni $V = \mathbb{R}^3$. Zacznijmy od fundamentalnego.

- Definiujemy $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ wzorem:

  $$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

  zadaje iloczyn skalarny na przestrzeni $\mathbb{R}^n$, który nazywać będziemy STANDARDOWYM ILOCZYNEM SKALARNYM i oznaczamy jako $\langle\,,\rangle_{st}$. Dzięki (standardowemu) iloczynowi skalarnemu odczytamy w nowym języku wiele poznanych wcześniej konfiguracji. Jest to również podstawowy obiekt w elementarnej geometrii analitycznej, choć niestety już nie geometrii uczonej w szkole (patrz `https://www.mimuw.edu.pl/~krych/chemia/2016-2017/ch10-11_geoman.pdf`).

  Rozważmy układ równań o współczynnikach rzeczywistych postaci:

  $$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \tag{4}$$

  Przyjmijmy, że $v = (x_1, \ldots, x_n)$ oraz $\alpha_i = (\alpha_{i1}, \ldots, \alpha_{in})$ są elementami $\mathbb{R}^n$, dla $1 \leq i \leq m$. Wówczas powyższy układ równań liniowych zapisać możemy w postaci układu warunków:

  $$\langle \alpha_1, v\rangle_{st} = b_1, \quad \ldots \quad, \langle \alpha_n, v\rangle_{st} = b_n.$$

Zauważmy, że jeśli $\epsilon_1, \ldots, \epsilon_n$ jest bazą standardową przestrzeni $\mathbb{R}^n$, to mamy:

$$\langle \epsilon_i, \epsilon_j \rangle = \begin{cases} 1, & \text{dla } i = j, \\ 0, & \text{dla } i \neq j \end{cases}.$$

To powinno rodzić skojarzenia z bazą dualną. I rzeczywiście jest tu bardzo istotny związek z przestrzenią sprzężoną. Każdemu wektorowi $\alpha = (a_1, \ldots, a_n)$ przypisać możemy funkcjonał $f_\alpha \in (\mathbb{R}^n)^*$ dany wzorem:

$$f_\alpha(x_1, \ldots, x_n) = a_1x_1 + \ldots + a_nx_n = \langle \alpha, x \rangle,$$

gdzie $x = (x_1, \ldots, x_n)$. Powyższe przyporządkowanie jest bijekcją, ilustrującą izomorfizm pomiędzy V oraz V^*, gdy $\dim V < \infty$.

- Funkcja $\langle \, , \rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ dana wzorem

  $$\langle (x_1, x_2, x_3), (y_1, y_2, y_3) \rangle = x_1y_1 + 2x_1y_2 + 2x_2y_1 + 9x_2y_2 + 4x_3y_3 \quad (\dagger)$$

  jest iloczynem skalarnym. Sprawdzenie warunków (1)-(3) wykonamy poniżej w bardziej ogólnym kontekście. Warunek (4) jest spełniony, bo wyrażenie $\langle (x_1, x_2, x_3), (x_1, x_2, x_3) \rangle$ zapisuje się jako:

  $$x_1^2 + 2x_1x_2 + 2x_2x_1 + 9x_2^2 + 4x_3^2 = (x_1 + 2x_2)^2 + 5x_2^2 + 4x_3^2.$$

- Funkcja $f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ dana wzorem

  $$f((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 + 4x_1y_2 + 4x_2y_1 + 5x_2y_2$$

  spełnia (1)-(3), ale

  $$f((1, -1, 0), (1, -1, 0)) = 1 \cdot 1 + 4 \cdot 1 \cdot (-1) + 4 \cdot (-1) \cdot 1 + 5 \cdot (-1) \cdot (-1) = -2,$$

  więc nie jest ona iloczynem skalarnym na $\mathbb{R}^3$.

Fakt 61. *Każdy iloczyn skalarny na $\mathbb{R}^n$ zadany jest wzorem:*

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \sum_{1 \leq i,j \leq n} a_{ij}x_iy_j,$$

dla pewnych $a_{ij} \in \mathbb{R}$ spełniających $a_{ij} = a_{ji}$. Co więcej, dla każdych liczb a_{ij} spełniających $a_{ij} = a_{ji}$ dla $i, j = 1, \ldots, n$ powyższy wzór zadaje funkcję $\langle \, , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ spełniającą warunki (1)-(3).

Dowód. Niech $\langle \, , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ będzie iloczynem skalarnym oraz niech $(\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$ będzie bazą standardową $\mathbb{R}^n$. Dla każdych $i, j = 1, 2, \ldots, n$ określamy $a_{ij} = \langle \epsilon_i, \epsilon_j \rangle$. Wówczas wartość $\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle$ to:

$$\begin{aligned}
&= \langle x_1\epsilon_1 + x_2\epsilon_2 + \ldots + x_n\epsilon_n, y_1\epsilon_1 + y_2\epsilon_2 + \ldots + y_n\epsilon_n \rangle = \\
&\overset{(1)}{=} x_1\langle \epsilon_1, y_1\epsilon_1 + y_2\epsilon_2 + \ldots + y_n\epsilon_n \rangle + \ldots + x_n\langle \epsilon_n, y_1\epsilon_1 + y_2\epsilon_2 + \ldots + y_n\epsilon_n \rangle = \\
&\overset{(2)}{=} x_1y_1\langle \epsilon_1, \epsilon_1 \rangle + \ldots + x_1y_n\langle \epsilon_1, \epsilon_n \rangle + \ldots + x_ny_1\langle \epsilon_n, \epsilon_1 \rangle + \ldots + x_ny_n\langle \epsilon_n, \epsilon_n \rangle = \\
&= a_{11}x_1y_1 + a_{12}x_1y_2 + \ldots + a_{1n}x_1y_n + \ldots + a_{n1}x_ny_1 + a_{n2}x_ny_2 + \ldots + a_{nn}x_ny_n.
\end{aligned}$$

Druga część dowodu wynika z następującej obserwacji, będącej przy okazji bardzo wygodnym narzędziem do obliczania wartości iloczynów skalarnych, zwłaszcza niestandardowych, na $\mathbb{R}^n$. Mianowie mamy:

$$\left[\sum_{1\le i,j\le n} a_{ij}x_iy_j\right] = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

A zatem to, że funkcja zadana wzorem wyżej spełnia (1) i (2) wynika łatwo z rozdzielności mnożenia macierzy względem dodawania. Natomiast łatwo widać, że jeśli $a_{ij} = a_{ji}$, to funkcja ta spełnia (3). □

Powyższa uwaga daje nam dwie informacje: po pierwsze iloczynów skalarnych w $\mathbb{R}^n$ jest całkiem sporo, a po drugie nie wiemy na starcie do końca jak je opisać, bo nie wiemy jak dobrać współczynniki a_{ij} tak, by spełniały warunek (4). Będzie to treścią kryterium Sylvestera, które poznamy za kilka wykładów.

Iloczyny skalarne pełnią fundamentalną rolę w badaniu ważnych klas przestrzeni nieskończonego wymiaru. Ograniczmy się w tym miejscu do podania dwóch przykładów ze skryptu.

- W przestrzeni $\mathbb{R}^\infty$ rozpatrujemy podprzestrzeń l^2 złożoną ze wszystkich ciągów (x_i) takich, że

  $$\sum_{i=1}^{\infty} x_i^2 < \infty.$$

  Wówczas (i to trzeba udowodnić) funkcja $\langle\, , \rangle : l^2 \times l^2 \to \mathbb{R}$ dana wzorem

  $$\langle\, (x_i), (y_i) \rangle = \sum_{i=1}^{\infty} x_i y_i$$

  jest iloczynem skalarnym na l^2.

- Niech $C[a,b]$ oznacza przestrzeń wszystkich funkcji ciągłych $[a,b] \to \mathbb{R}$. Funkcja $\langle\, , \rangle : C[a,b] \times C[a,b] \to \mathbb{R}$ dana wzorem

  $$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

  jest iloczynem skalarnym na przestrzeni $C[a,b]$.

Widzimy, że w obydwu przypadkach wprowadzenie iloczynu skalarnego wymaga nietrywialnych definicji i faktów analitycznych (i nie tylko). Stąd tez rozważania na temat nieskończenie wymiarowych przestrzeni liniowych (m.in) z iloczynem skalarnym stanowią przedmiot semestralnego wykładu na wyższych latach.

Definicja 48. *Parę $(V, \langle\,,\rangle)$ gdzie V jest skończenie wymiarową przestrzenią liniową nad $\mathbb{R}$ a $\langle\,,\rangle$ jest iloczynem skalarnym na V nazywamy* PRZESTRZENIĄ EUKLIDESOWĄ LINIOWĄ.

Definicja 49. *Niech $\langle\,,\rangle$ będzie iloczynem skalarnym na przestrzeni V.* DŁUGOŚCIĄ *(albo* NORMĄ*) wektora $\alpha \in V$ nazywamy liczbę $\|a\| = \sqrt{\langle \alpha, \alpha\rangle}$.*

Norma zależy od wyboru iloczynu skalarnego. Wektor $(3, 4)$ ma:

- normę $\sqrt{3^2 + 4^2} = 5$, jeśli $\langle\,,\rangle$ to standardowy iloczyn skalarny,
- normę $\sqrt{3^2 + (2 \cdot 4)^2} = \sqrt{73}$, jeśli $\langle (x_1, x_2), (y_1, y_2)\rangle = x_1y_1 + 2x_2y_2$.

Nowo zdefiniowana funkcja „zachowuje" się jak znana ze szkoły długość, co wyrażają następujące „znane" fakty.

Fakt 62. *Niech $\langle\,,\rangle$ będzie iloczynem skalarnym na przestrzeni V. Wówczas dla każdych $\alpha, \beta \in V$ zachodzi:*

- *nierówność Schwarza: $|\langle \alpha, \beta\rangle| \leq \|\alpha\| \cdot \|\beta\|$,*
- *nierówność trójkąta: $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$,*
- *twierdzenie Pitagorasa:*

$$\langle \alpha, \beta\rangle = 0 \quad \Longrightarrow \quad \|\alpha\|^2 + \|\beta\|^2 = \|\alpha + \beta\|^2.$$

Dowód. Dowód (1). Jeśli $\alpha = 0$, to obie strony nierówności są zerami. W przypadku, gdy $\alpha \neq 0$ rozważamy funkcję $f : \mathbb{R} \to \mathbb{R}$ zadaną wzorem $f(t) = \|t\alpha + \beta\|^2$. Zatem

$$\begin{aligned} f(t) = \langle t\alpha + \beta, t\alpha + \beta\rangle &= \\ = t\langle \alpha, t\alpha + \beta\rangle + \langle \beta, t\alpha + \beta\rangle &= \\ = t^2\langle \alpha, \alpha\rangle + t\langle \alpha, \beta\rangle + t\langle \beta, \alpha\rangle + \langle \beta, \beta\rangle &= \\ = t^2\|\alpha\|^2 + 2t\langle \alpha, \beta\rangle + \|\beta\|^2&. \end{aligned}$$

Na mocy warunku (4) z definicji iloczynu skalarnego dla każdego $t \in \mathbb{R}$ mamy $f(t) \geq 0$, a zatem wyróżnik trójmianu kwadratowego $t^2\|\alpha\|^2 + 2\langle \alpha, \beta\rangle t + \|\beta\|^2$ musi być niedodatni, tj. $\Delta \leq 0$, gdzie

$$\Delta = (2\langle \alpha, \beta\rangle)^2 - 4\|\alpha\|^2\|\beta\|^2 = 4\langle \alpha, \beta\rangle^2 - 4\|\alpha\|^2\|\beta\|^2.$$

Stąd wynika (1). Do dowodu (2) i (3) wykorzystujemy przydatną w rachunkach tożsamość, wynikającą z naszych wcześniejszych obserwacji.

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + 2\langle \alpha, \beta\rangle + \|\beta\|^2,$$

którą wyprowadza się jak powyżej. Wówczas na mocy (1) mamy:

$$\|\alpha + \beta\|^2 \leq \|\alpha\|^2 + 2\|\alpha\| \cdot \|\beta\| + \|\beta\|^2 = (\|\alpha\| + \|\beta\|)^2.$$

$\square$

Z nierówności Schwarza wynika natychmiast, że dla każdej pary niezerowych wektorów $\alpha, \beta \in V$ mamy:

$$-1 \leq \frac{\langle \alpha, \beta \rangle}{\|\alpha\| \cdot \|\beta\|} \leq 1.$$

Dla każdej liczby rzeczywistej $r \in [-1, 1]$ istnieje dokładnie jedna liczba $\theta \in [0, \pi]$, że $\cos\theta = r$.

Definicja 50. *Niech $\langle\, , \rangle$ będzie iloczynem skalarnym na przestrzeni V. Niech α, β będą niezerowymi wektorami przestrzeni V. Liczbę $\theta \in [0, \pi]$ taką, że*

$$\frac{\langle \alpha, \beta \rangle}{\|\alpha\| \cdot \|\beta\|} = \cos\theta$$

nazywamy (niezorientowanym) KĄTEM MIĘDZY WEKTORAMI *α i β.*

Fakt 63. *W przestrzeni liniowej V z iloczynem skalarnym $\langle\, , \rangle$ dla każdych niezerowych α, β mamy: $\langle \alpha, \beta \rangle = \|\alpha\| \cdot \|\beta\| \cos\theta$, gdzie θ jest kątem pomiędzy wektorami α i β.*

Zobaczmy przykład rachunków związanych z wprowadzonymi pojęciami rozwiązując następujące zadanie.

Fakt 64. *Niech α, β będą liniowo niezależnymi wektorami w przestrzeni liniowej V z iloczynem skalarnym takimi, że $\|\alpha\| = \|\beta\| = 1$. Pokazać, że dla każdego $t \in (0, 1)$ zachodzi nierówność*

Jaka jest geometryczna interpretacja tego ćwiczenia?

$$\|(1-t)\alpha + t\beta\| < 1.$$

ROZWIĄZANIE. Dowód 1. Korzystamy z nierówności trójkąta, która jest ostra dla wektorów liniowo niezależnych (dlaczego?):

$$\|(1-t)\alpha + t\beta\| < \|(1-t)\alpha\| + \|t\beta\| = 1 - t + t = 1.$$

Ostatni krok jest OK, bo $1 - t > 0$ i $t > 0$.

Dowód 2. Skorzystamy z wniosku z nierówności Schwarza:

$$\langle v, w \rangle < \|v\|\|w\|,$$

dla wektorów liniowo niezależnych (dlaczego?) $u, v \in V$. Rozpisujemy $\|(1-t)\alpha + t\beta\|^2 = \langle (1-t)\alpha + t\beta, (1-t)\alpha + t\beta \rangle$ i dalej:

$$\begin{aligned}
\langle (1-t)\alpha + t\beta, (1-t)\alpha + t\beta \rangle &= \langle (1-t)^2 \langle \alpha, \alpha \rangle + t^2 \langle \beta, \beta \rangle + (1-t)t\langle \alpha, \beta \rangle + t(1-t)\langle \beta, \alpha \rangle = \\
&= \langle (1-t)^2 \|\alpha\|^2 + t^2 \|\beta\|^2 + 2(1-t)t\langle \alpha, \beta \rangle \\
&< (1-t)^2 \|\alpha\|^2 + t^2 \|\beta\|^2 + 2(1-t)t\|\alpha\|\|\beta\| \\
&= (1-t)^2 + t^2 + 2t(1-t) = (1-t+t)^2 = 1
\end{aligned}$$

■