2 Wektory, przestrzenie wektorowe

Matematycy zamiast pytań *"czym jest dany obiekt"* wolą zadawać pytanie:*"jakie ma on własności"*. Bo okazuje się, że w wielu analogicznych sytuacjach mamy do czynienia z obiektami równoważnymi w tym sensie, że każda własność obiektu A jest również własnością obiektu B i jeden z tych obiektów można wybrać jako model dla szerokiej klasy jemu podobnych. Dość dobrym przykładem jest pojęcie wektora i przestrzeni wektorowej wymiaru d.

Gdy przesuwamy obiekt z punktu P_1 do punktu P_2 wzdłuż odcinka łączącego te punkty wystarczy znać 3 wielkości: kierunek (prostą, wzdłuż której będziemy się poruszali), zwrot (czy zaczynamy w punkcie P_1 i lądujemy w P_2, czy też na odwrót) oraz długość drogi (oznaczmy ją $\|P_2 - P_1\|$). Te 3 dane możemy wykorzystać przemieszczając się z innego punktu P_3 do punktu P_4, wówczas ten sam kierunek oznacza, że proste (trajektorie ruchu) będą równoległe, zwrot będzie taki sam, jak i długość drogi -identyczna. Można powiedzieć, że wektory ruchu będą jednakowe. Traktując pary punktów jako "wektory zaczepione" - wyznaczone przez ich punkt początkowy i koniec (na kórym rysujemy strzałkę) zauważamy, że dla opisu ruchu wystarczy znać wektor swobodny - to klasa równoważności wektorów zaczepionych, gdzie za równoważne uznajemy wektory o takich samych 3 cechach: kierunek, zwrot, długość.

Będziemy więc rozważali jedynie wektory swobodne, które można traktować jak wektory zaczepione w punkcie 0 -początku układu. Początkiem wektora jest więc 0, koniec, to pewien punkt przestrzeni, który jednoznacznie wyznacza dany wektor. W tym sensie, każdy element P ze zbioru $\mathbb{R}^d$ możemy traktować jako wektor (wektor o początku w zerze i końcu P.

Najważniejszym działaniem na wektorach jest dodawanie. Jeśli wektor $\vec{u}$ ma początek w P_1 i koniec w P_2, zaś $\vec{w}$ ma początek w P_2 i koniec w P_3, to $\vec{u} + \vec{w}$ jest wektorem o początku P_1 i końcu P_3. Dodawanie wektorów swobodnych sprowadzamy do tego przypadku zaczepiając wektor pierwszy w zerze, koniec będzie w $P_2 - P_1$, następnie składamy go z wektorem zaczepionym w $P_2 - P_1$ równoważnym wektorowi $\vec{w}$. Czyli przesuwamy równolegle ten drugi wektor tak, by jego początek znalazł się na końcu wektora pierwszego: $\vec{v}$. Trochę to skomplikowane, ale najlepiej zrobić rysunek.

Najprościej jednak mówić o wektorach jako o punktach przestrzeni (w przypadku przestrzeni $\mathbb{R}^d$ wymiaru d złożonej z układów d liczb rzeczywistych). Gdy $d = 2$, to parom liczb (a,b) przyporządkujwemy punkty na płaszczyźnie o współrzędnych a, b -lub liczbę zespoloną $a + ib$. W przestrzeni trójwymiarowej mamy współrzędne zazwyczaj oznaczane x, y, z, gdzie np. z oznacza wysokość punktu nad płaszczyzną OXY. Współrzędne punktu, to wartości jego rzutu prostopadłego na daną oś układu współrzędnych. Punkty $P \in \mathbb{R}^3$ o współrzędnych (x_P, y_P, z_P) wyznaczają wektory $\vec{v_P}$ zaczepione w zerze, czyli o początku 0 i o końcu P. Gdy mamy drugi wektor $\vec{v_Q}$ wyznaczony przez inny punkt $Q \in \mathbb{R}^3$, to

$$\texttt{współrzędnymi wektora } \vec{v_P} + \vec{v_Q} \texttt{ są } (x_P + x_Q, y_P + y_Q, z_P + z_Q).$$

W przypadku dwuwymiarowym jeśli traktować wektory jako liczby zespolone, sumą wektorów będzie to, co zdefiniowaliśmy jako suma liczb zespolonych. Zawsze dodawanie wektorów jest przemienne i łączne, dodanie wektora zerowego $\vec{0}$ nie zmienia wyniku. Wydłużanie wektora, to mnożenie wszystkich jego współrzędnych przez tę samą liczbę $k > 0$. Na przykład gdy $k = 3$, to $k\vec{v}$ jest wektorem 3-krotnie dłuższym. Każdą z jego współrzędnych mnożymy przez to samo k. Gdy $k = -1$, to $(-1)\vec{v}$, czyli wektor $\vec{-v}$ nazywamy wektorem przeciwnym do $\vec{v}$. Jest to jedyny wektor o tej własności,że $\vec{v} + \vec{-v} = \vec{0}$. Ogólnie, definiujemy iloczyn wektora $\vec{v}$ o współrzędnych (x, y, z) przez liczbę $k \in \mathbb{R}$ jako wektor $k\vec{v}$ o współrzędnych (kx, ky, kz).

W dalszym ciągu będziemy utożsamiali wektory z układami ich współrzędnych, np. pisząc $(x,y)+(\hat{x},\hat{y})$ na oznaczenie wektora $(x+\hat{x},y+\hat{y})$, czyli sumy tych wektorów. Podobnie, elementy przestrzeni $\mathbb{R}^d$, czyli układy d liczb $(x_1,x_2,\dots,x_d)$ traktujemy jako wektory, sumą jest wektor, którego j-ta współrzędna jest sumą j-tych współrzędnych dodawanych wektorów. Analogicznie, jak dla $d=3$ definiujmy iloczyn takiego wektora przez liczbę $k\in\mathbb{R}$ jako wektor $(kx_1,kx_2,\dots,kx_n)$. Długość euklidesową (czyli normę) wektora definiujemy jako liczbę $\|\vec{v}\|$ daną (w przypadku wektora o d współrzędnych) wzorem

$$\|(x_1,x_2,\dots,x_d)\| := \sqrt{x_1^2+x_2^2+\dots+x_d^2}.$$

W przypadku ciała liczb zespolonych traktowanego jako zbiór $\mathbb{R}^2$ taka norma, to moduł liczby: $|a+ib|=\sqrt{a^2+b^2}=\|(a,b)\|$.

Zdefiniujmy jeszcze iloczyn skalarny wektorów $v=(x_1,x_2,\dots,x_d)$ oraz $w=(y_1,y_2,\dots,y_d)\in\mathbb{R}^d$ jako liczbę

$$\langle v,w\rangle := x_1y_1+x_2y_2+\dots+x_dy_d.$$

Iloczyn skalarny często oznaczany jest również przez kropkę w zapisie $v\cdot w$, można też spotkać oznaczenie $(v|w)$ lub $u\circ w$. Użycie nawiasów zamiast kropki wiąże się z tym, że w przyszłości będziemy rozważali też przestrzenie, w których wektorami będą funkcje, a dla nich kropka oznacza zwykły iloczyn, zaś "$\circ$" oznacza złożenie. Iloczyn skalarny dla $f,g\in R[a,b]$, to $\langle f,g\rangle := \int_a^b f(t)g(t)\,dt$.

Pora na ogólną definicję:

Definicja. Przestrzenią wektorową (rzeczywistą, czyli nad ciałem $\mathbb{R}$) nazywamy zbiór X z działaniem wewnętrznym "dodawania wektorów":

$$X\times X\ni(v,w)\mapsto v+w\in X$$

które tworzy strukturę grupy przemiennej, czyli jest przemienne, łączne, ma element neutralny $0\in X$ taki, że $\forall_{v\in X}\, v+0=v$ i dla każdego $v\in X$ istnieje element przeciwny $-v\in X$ taki, że $v+(-v)=0$. Ponadto określone jest działanie zewnętrzne "mnożenia wektorów przez skalary":

$$\mathbb{R}\times X\ni(t,v)\mapsto tv\in X,$$

które jest rozdzielne względem dodawania wektorów: $t(v+w)=tv+tw$, dodawania skalarów: $(t+s)v=tv+sv$, spełniające jeszcze 2 warunki: $(ts)v=t(sv), 1v=v$ dla wszystkich $t,s\in\mathbb{R}, v,w\in X$.

Przykładem przestrzeni wektorowej jest $\mathbb{R}^d$, ale również mamy przestrzenie, w których wektorami są funkcje, np. $C[a,b]$, czy $R[a,b]$. W przestrzeniach funkcyjnych złożonych z funkcji na zbiorze D wynikiem dodawania funkcji f,g jest funkcja $f+g: D\ni z\mapsto f(z)+g(z)$. Mnożenie przez skalar $c\in\mathbb{R}$, to mnożenie przez stałą c, czyli funkcja $cf: D\ni z\mapsto cf(z)$.

Nazwa "iloczyn skalarny wektorów" oznacza, że wynikiem takiego mnożenia dwu wektorów jest liczba (czyli "wielkość skalarna").

Własności definiujące iloczyn skalarny (w przypadku przestrzeni rzeczywistych)[1], to:

- *symetria*: $\langle v,w\rangle=\langle w,v\rangle$,
- *dodatnia określoność:* $\forall_{v\neq0}\,\langle v,v\rangle>0$ oraz
- *liniowość wzgl. 1. zmiennej:* $\langle v+w,u\rangle=\langle v,u\rangle+\langle w,u\rangle, \quad \langle\alpha v,u\rangle=\alpha\langle v,u\rangle$.

Zauważmy, że $\|v\|=\sqrt{\langle v,v\rangle}$. Ponadto następujące odwzorowanie zmiennej $t\in\mathbb{R}$: $\|v+tw\|^2=\langle v+tw,v+tw\rangle=\|v\|^2+2t\langle v,w\rangle+t^2\|w\|^2$ jest trójmianem kwadratowym stale nieujemnym. Jego wyróżnik Δ nie może być dodatni (wówczas trójmian ten powinien zmieniać znak). Tak więc $0\geqslant\Delta=(2\langle v,w\rangle)^2-4\|v\|^2\|w\|^2$, co (po spierwiastkowaniu i podzieleniu przez 2) daje

$$\texttt{NIERÓWNOŚĆ Schwarza:}\quad |\langle v,w\rangle|\leqslant\|v\|\|w\|.$$

[1] Dla przestrzeni zespolonych zamiast symetrii jest "skośna symetria": $\langle w,u\rangle=\overline{\langle u,w\rangle}$. Na przykład, dla funkcji całkowalnych $f,g:[a,b]\to\mathbb{C}$ definiujemy $\langle f,g\rangle=\int_a^b f(t)\overline{g(t)}\,dt$.

Wynika stąd

$$\texttt{nierówność trójkąta dla normy:}\ \|v+w\| \leqslant \|v\| + \|w\|.$$

Faktycznie, ponieważ składniki są tu nieujemne, wystarczy sprawdzić ją po podniesieniu stronami do kwadratu. Ale z własności iloczynu skalarnego,

$$\|v+w\|^2 = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 \leqslant \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2.$$

Gdy $d = 2$, wynika stąd w szczególności nierówność trójkąta dla modułów liczb zespolonych. Innym wzorem, w którym występuje kąt $\angle(v, w)$ między dwoma wektorami (np. w $\mathbb{R}^3$) jest następujące twierdzenie:

$$\texttt{Wzór kosinusów} \qquad \langle v, w\rangle = \|v\|\|w\| \cos\angle(v, w).$$

W szczególnosci, mówimy, że wektory $v, w \in \mathbb{R}^d$ są prostopadłe, co zapisujemy $v \perp w$, gdy $\langle v, w\rangle = 0$. W takim przypadku z poprzednio używanego wzoru na $\|v + w\|^2$ wynika, że gdy $v \perp w$, to zachodzi

$$\texttt{TWIERDZENIE Pitagorasa:} \qquad \|v+w\|^2 = \|v\|^2 + \|w\|^2.$$

Współrzędne wektora $v = (x, y, z)$ są iloczynami z wektorami jednostkowymi poszczególnych osi. Te wektory jednostkowe oznaczane są często symbolami i, j, k, przy czym i nie ma tu związku z jednostką urojoną -oznacza wektor $(1, 0, 0)$. W przestrzeni $\mathbb{R}^d$ używamy raczej symboli $\varepsilon_j,\ j = 1, 2, \dots, k$, gdzie np. $\varepsilon_2 = (0, 1, 0, \dots, 0)$, a wektor ε_j ma zame zera z wyjątkiem jedynki na miejscu o numerze j. Oczywiście, $\langle (x_1, x_2, \dots, x_d), \varepsilon_j\rangle = x_j$. Ze wzoru kosinusów i z faktu, że $\|\varepsilon_j\| = 1$ wynika, że tak zwane kosinusy keirunkowe (odpowiadające kątom między wektorem v i wektorem ε_j dane są wzorami

$$\cos(\angle(v, \varepsilon_j)) = \frac{x_j}{\|v\|} \quad \texttt{gdy} \quad v = (x_1, x_2, \dots, x_d).$$

Oczywiście,

$$\sum_{j=1}^{d} \cos^2(\angle(v, \varepsilon_j)) = 1.$$

Kombinacją liniową wektorów $v_1, \dots, v_k$ (kombinacją o współczynnikach skalarnych $c_1, \dots, c_k$) nazywamy wektor

$$c_1 v_1 + c_2 v_2 + \dots + c_k v_k = \sum_{m=1}^{k} c_m v_m. \tag{1}$$

Układ liniowo niezależny, to układ, dla którego współczynniki skalarne kombinacji liniowych są wyznaczone jednoznacznie. Liniowa niezależność jest równoważna implikaacji $(c_1 v_1 + c_2 v_2 + \dots + c_k v_k = 0 \Rightarrow c_1 = c_2 = \dots = c_k = 0)$.

Twierdzenie. Układy ortogonalne (czyli takie układy wektorów v_n, dla których $n \neq m \Rightarrow v_n \perp v_m$) są liniowo niezależne.

Obwiednią liniową układu wektorów $v_1, v_2, \dots, v_k$, oznaczaną span $\{v_1, \dots, v_k\}$ nazywamy zbiór wszystkich kombinacji liniowych postaci (1), gdzie c_m są dowolnymi układami liczb rzeczywistych.

Układ wektorów $v_1, v_2, \dots, v_k$ nazywany bazą przestrzeni wektorowej X, jeśli jest on liniowo niezależny oraz span $\{v_1, \dots, v_k\} = X$. Najprostszym przykładem bazy jest układ wektorów jednostkowych w kierunkach osi, czyli $\varepsilon_n,\ n = 1, \dots, d$. Jest to nawet przykład bazy ortogonalnej, czyli takiej bazy, w której każde dwa różne wektory są wzajemnie prostopadłe. Można wykazać, że każde dwie bazy danej przestrzeni wektorowej X mają tyle samo elementów, ta ilość elementów baz nazywa się wymiarem przestrzeni, oznaczamy ją symbolem $\dim(X)$. Tak więc np. $\dim(\mathbb{R}^d) = d$.

Teraz jeśli X, Y są dwiema przestrzeniami wektorowymi, to odwzorowanie $T : X \to Y$ nazywamy odwzorowaniem (lub operatorem) liniowym, gdy

$$\forall_{u,w\in X}\forall_{\alpha,\beta\in\mathbb{R}}\ T(\alpha u + \beta w) = \alpha T(u) + \beta T(w).$$

Zbiór wszystkich odwzorowań liniowych z przestrzeni wektorowej X do Y oznaczamy symbolem $\mathcal{L}(X, Y)$. Odwzorowania liniowe z X do X nazywamy też endomorfizmami liniowymi X i ich zbiór oznaczamy symbolem $\mathcal{L}(X)$. Operatory liniowe o wartościach skalarnych nazywamy funkcjonałami liniowymi, ich zbiór, czyli $\mathcal{L}(X, \mathbb{R})$ oznaczamy symbolem X^*. Jak można łatwo wykazać, każdy funkcjonał liniowy $\phi : \mathbb{R}^d \to \mathbb{R}$ jest mnożeniem skalarnym przez pewien wektor $z \in \mathbb{R}^d$, czyli $\phi(x) = \langle x, z\rangle$. Takim wektorem jest $z = (z_1, \dots, z_d)$, gdzie $z_j = \phi(\varepsilon_j)$.

Wartości dowolnego operatora liniowego T na przestrzeni X są jednoznacznie określone poprzez wartości T na wektorach dowolnie ustalonej bazy. Gdy $X = \mathbb{R}^d$, $Y = \mathbb{R}^k$, to mamy takich d wektorów $T(\varepsilon_1), T(\varepsilon_2), \dots, T(\varepsilon_d)$. Każdy z tych wektorów zapisujemy w postaci kolumny k liczb, otrzymując macierz złożoną z k wierszy i d kolumn.

Jeżeli $e_1, \dots, e_k$ oznaczają wektory kanonicznej bazy 0-1 -kowej w $\mathbb{R}^k$, to współczynnikami wektora $T(\varepsilon_j)$ są iloczyny skalarne tego wektora przez kolejne wektory $e_1, \dots, e_k$. Zapisując przez a_{mj} wyrazy macierzy będące w j-tej kolumnie i w m-tym wierszu widzimy, że $a_{mj} = \langle T(\varepsilon_j), e_m\rangle$.

Przykładem przestrzeni nieskończenie wymiarowej jest przestrzeń $C[a, b]$ złożona z funkcji ciągłych ma odcinku $[a, b]$. Aby się o tym przekonać wystarczy dla dowolnego $n \in \mathbb{N}$ znaleźć n liniowo niezależnych funkcji ciągłych. Takimi są np. funkcje f_j różne od zera w pewnych punktach $t_j \in [a, b]$, dla których zbiory $\{s \in [a, b] : f_j(s) \neq 0\}$ są parami rozłączne. Liniowo niezależnym -a nawet ortogonalnym układem funkcji na odcinku $[-\pi, \pi]$ jest tak zwany

$$\texttt{układ trygonometryczny:}\quad \cos(nx),\ \sin(nx),\ \texttt{gdzie}\ \ n \in \mathbb{N}.$$

Funkcjonałem ciągłym na $C[a, b]$ jest np. funkcjonał ϕ_g, gdzie

$$\phi_g(f) = \int_a^b f(t)g(t)\,dt,$$

zaś g jest pewną funkcją całkowalną. Ale są też funkcjonały nieco innej postaci, np. gdy $c \in [a, b]$, takim funkcjonałem jest tzw. "*delta Diraca* w punkcie c", to funkcjonał określony wzorem $\delta_c(f) := f(c)$.