1 Przestrzenie z iloczynem skalarnym

W definicji przestrzeni wektorowej nie ma mowy o długości wektora (chyba że jest to wektor zerowy). Wprowadzając pojęcie normy, mamy zdefiniowaną długość wektora. Nie umiemy jednak mierzyć *kątów* między wektorami (tzn. pojęcie kąta nie jest tam zdefiniowane). W przestrzeniach z iloczynem skalarnym jest to możliwe i w takich przestrzeniach możemy pełniej wykorzystywać intuicje geometryczne.

1.1 Iloczyn skalarny i jego własności

Niech V będzie przestrzenią wektorową nad ciałem $\mathbb{K}$ (w praktyce $\mathbb{K}$ będzie $\mathbb{R}$ lub $\mathbb{C}$). Odwzorowanie $(\cdot|\cdot)$: $V \times V \to \mathbb{K}$ nazywamy *iloczynem skalarnym*, jeśli dla dowolnych wektorów $x, y, z \in V$ oraz $\lambda, \mu \in \mathbb{K}$ zachodzą warunki:

1. $(z|\lambda x + \mu y) = \lambda(z|x) + \mu(z|y)$ (tzn. iloczyn skalarny w drugim argumencie jest liniowy);
2. (a) $(x|y) = (y|x)$ dla $\mathbb{K} = \mathbb{R}$ (symetria),
   (b) $(x|y) = \overline{(y|x)}$ dla $\mathbb{K} = \mathbb{C}$ (hermitowskość).
3. $(x|x) \geqslant 0$, oraz $(x|x) = 0 \Longrightarrow x = \mathbf{0}$.

Uwaga. Dla $\mathbb{K} = \mathbb{C}$, iloczyn skalarny na ogół przyjmuje wartości zespolone. Natomiast p. 2 powyżej mówi, że $(x|x) \in \mathbb{R}$ (a więc sensowny jest p. 3, i zgodnie z nim $(x|x) \geqslant 0$).

Def. Odwzorowanie $F : V \to W$ przestrzeni wektorowych nad $\mathbb{C}$ nazywa się *antyliniowe*, jeśli jest addytywne (tzn. $F(x+y) = F(x)+F(y)$) oraz $F(\lambda x) = \bar{\lambda}x$ (tu $x, y \in F$, $\lambda \in \mathbb{C}$).

Zauważmy, że ze względu na własność 2, w przypadku zespolonym iloczyn skalarny jest *antyliniowy* w pierwszym argumencie:

$$(\lambda x|y) = \overline{(y|\lambda x)} = \overline{\lambda(y|x)} = \bar{\lambda}\overline{(y|x)} = \bar{\lambda}(x|y).$$

Pamiętając definicję formy biliniowej, możemy powiedzieć, że:

- Dla $\mathbb{K} = \mathbb{R}$ iloczyn skalarny to forma biliniowa symetryczna, a stowarzyszona z nią forma kwadratowa jest dodatnio określona (p. 3).

  Def. Rzeczywista przestrzeń wektorowa wyposażona w iloczyn skalarny nazywa się *przestrzenią euklidesową.*

- Dla $\mathbb{K} = \mathbb{C}$ iloczyn skalarny to forma liniowa w drugim argumencie i antyliniowa w pierwszym (takie formy nazywa się *półtoraliniowymi*). Jest ona ponadto *hermitowska* (p. 2) a stowarzyszona z nią forma kwadratowa jest dodatnio określona (p.3).

  Def. Zespolona przestrzeń wektorowa wyposażona w iloczyn skalarny nazywa się *przestrzenią unitarną.*

Przykłady. (Własności iloczynu skalarnego wszędzie sprawdza się łatwo)

1. $V = \mathbb{R}^n$; $(x|y) = x^1y^1 + x^2y^2 + \cdots + x^ny^n$. Jest to *standardowy* (bądź *kanoniczny*) iloczyn skalarny w $\mathbb{R}^n$.

2. $V = \mathbb{C}^n$; $(x|y) = \overline{x^1}y^1 + \overline{x^2}y^2 + \cdots + \overline{x^n}y^n$. Jest to *standardowy* (bądź *kanoniczny*) iloczyn skalarny w $\mathbb{C}^n$.

3. $V = \mathbb{R}_n[\cdot]$; $(u|v) = \int_a^b u(x)v(x)\mathsf{d}x$ (tu zakładamy, że $b > a$)

4. $V = \mathbb{C}_n[\cdot]$; $(u|v) = \int_a^b \overline{u(x)}v(x)\mathsf{d}x$ (tu też zakładamy, że $b > a$)

5. $V = \mathbb{R}_n[\cdot]$; $(u|v) = \int_{-\infty}^{\infty} u(x)v(x)e^{-x^2}\mathsf{d}x$.

Łatwo zobaczyć, że jeśli V jest przestrzenią z iloczynem skalarnym, to każda jej podprzestrzeń również jest przestrzenią z iloczynem skalarnym.

1.2 Norma indukowana przez iloczyn skalarny. Nierówność Schwarza

Dla wektora $x \in V$ zdefiniujmy wielkość $||x||$:

$$||x|| = \sqrt{(x|x)}. \tag{1}$$

Z własności iloczynu skalarnego, mamy: $||x|| \geqslant 0$; $||\lambda x|| = |\lambda| \cdot ||x||$.

Stw. (*Nierówność Schwarza*). Dla dowolnych wektorów $x, y \in V$ zachodzi nierówność:

$$|(x|y)| \leqslant ||x|| \cdot ||y|| \tag{2}$$

przy czym równość ma miejsce wtedy i tylko wtedy, gdy x, y są proporcjonalne (tzn $x = \alpha y$ dla pewnego $\alpha \in \mathbb{K}$).

Dowody przeprowadzimy oddzielnie dla przypadków $\mathbb{R}$ i $\mathbb{C}$ (wersja zespolona jest nieco bardziej skomplikowana. Wersja dla $\mathbb{R}$ już – w nie najogólniejszym przypadku – była, ale dla kompletu zamieścimy i ją).

- $\mathbb{K} = \mathbb{R}$. Weźmy wektory $x, y \in V$ oraz $t \in \mathbb{R}$. Zdefiniujmy wielkość: $f(t) = (x - ty|x - ty)$. Ponieważ iloczyn skalarny jest nieujemny, mamy: $f(t) \geqslant 0$. Wypiszmy dokładniej:

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

  widzimy, że $f(t)$ jest *trójmianem kwadratowym* w zmiennej t. Trójmian ten jest nieujemny, zatem jego wyróżnik jest niedodatni:

  $$\Delta = 4(x|y)^2 - 4||x||^2||y||^2 \leqslant 0 \quad \Longrightarrow \quad |(x|y)| \leqslant ||x|| \cdot ||y||,$$

  czyli dostalimy nierówność (2). Kiedy ma miejsce równość? Wtedy, gdy $\Delta = 0$, a to oznacza, że $f(t)$ ma pierwiastek (podwójny): $\exists_{t_0} : f(t_0) = 0$, a to znaczy, że $||x - t_0 y|| = 0$, czyli $x = t_0 y$.

- $\mathbb{K} = \mathbb{C}$. Tu iloczyn skalarny nie jest symetryczny (przy zamianie argumentów pojawia się sprzężenie). Postąpimy trochę inaczej: Zapiszmy iloczyn skalarny $(x|y)$ w postaci (zespolonej) trygonometycznej:

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

  stąd możemy zapisać:

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

Rozpatrzmy teraz następującą funkcję $f(t)$ zmiennej rzeczywistej t:

$$f(t) = (e^{i\phi}x - ty|e^{i\phi}x - ty) = (x|x) - t(e^{i\phi}x|y) - t(y|e^{i\phi}x) + t^2(y|y)$$
$$= ||x||^2 - 2t|(x|y)| + t^2||y||^2.$$

Funkcja $f(t)$ jest nieujemnym trójmianej kwadratowym w zmiennej t. Widać, że od tego momentu można dalej postępować jak w przypadku rzeczywistym.

Wniosek. Dla dowolnych dwóch wektorów x, y, zarówno w przypadku rzeczywistym jak i zespolonym, zachodzi:

$$||x + y|| \leqslant ||x|| + ||y||. \tag{3}$$

Dow. Mamy:

$$||x + y||^2 = (x + y|x + y) = (x|x) + (x|y) + (y|x) + (y|y) = ||x||^2 + ||y||^2 + 2\mathrm{Re}(x|y)$$
$$\leqslant ||x||^2 + ||y||^2 + 2|(x|y)| \leqslant ||x||^2 + ||y||^2 + 2||x|| \cdot ||y|| = (||x|| + ||y||)^2$$

(w ostatniej nierówności korzystaliśmy z nierówności Schwarza). Obie strony otrzymanej nierówności są dodatnie, więc możemy wyciągnąć dodatni pierwiastek kwadratowy i otrzymujemy (3).

CBDO

Uwaga. W ten sposób stwierdziliśmy, że wielkość $||x||$ definiowana przez (1) spełnia trzy warunki *normy.* Normę definiowaną przez (1) nazywamy *normą indukowaną przez iloczyn skalarny*

1.3 Ortogonalność

Weźmy dwa niezerowe wektory $x, y \in V$ (V jest nad $\mathbb{R}$). Dzięki nierówności Schwarza mamy nierówność:

$$-1 \leqslant \frac{(x|y)}{||x|| \cdot ||y||} \leqslant 1. \tag{4}$$

Znak równości zachodzi, gdy x jest proporcjonalny do y.

Zdefiniujmy kąt α między wektorami x, y jako takie $\alpha \in [0, \pi]$, że

$$\cos\alpha = \frac{(x|y)}{||x|| \cdot ||y||}. \tag{5}$$

Gdy $x = ty$, $t > 0$, to $\cos\alpha = 1$, więc $\alpha = 0$ – zgodnie z intuicją z $\mathbb{R}^3$ ('wektory x oraz y mają zwroty zgodne'). Podobnie gdy $x = ty$, $t < 0$, to $\cos\alpha = -1$, więc $\alpha = \pi$ ('wektory x oraz y mają zwroty przeciwne'). Gdy zachodzi $(x|y) = 0$, to $\alpha = \frac{\pi}{2}$ – mówimy wtedy, że wektory x, y są *prostopadłe* (lub *ortogonalne*). Zauważmy, że definicja ortogonalności jest też sensowna w przestrzeniach *zespolonych* (nie ma tam pojęcia kąta, ale gdy $(x|y) = 0$, to mówimy, że x, y są ortogonalne).

Zauwaṁy, że jeżeli x, y są ortogonalne, to

$$||x + y||^2 = (x + y|x + y) = (x|x) + (x|y) + (y|x) + (y|y) = ||x||^2 + ||y||^2.$$

Przez analogię z elementarną geometrią, równość tę nazywamy *twierdzeniem Pitagorasa.*

Def. Niech zbiór wektorów $(e = e_1, e_2, \ldots, e_n)$ będzie bazą w V. Mówimy, że baza e jest *ortogonalna*, jeśli dla dowolnych $i \neq j$ zachodzi:

$$(e_i|e_j) = 0.$$

Jeśli ponadto dla $i = 1, 2, \ldots, n$ zachodzi:

$$(e_i|e_i) = 1,$$

to bazę e nazywamy *ortonormalną*.

Powyższe dwa warunki można – przy użyciu symbolu *delty Kroneckera*, zapisać jako jeden; mamy wtedy definicję:

Bazę $e = (e_1, e_2, \ldots, e_n)$ nazywamy *ortonormalną*, jeżeli

$$(e_i|e_j) = \delta_{ij}.$$

Przy użyciu baz ortogonalnych, a jeszcze lepiej ortonormalnych, wiele rachunków się ułatwia i wzory w tych bazach na ogół wyglądają prościej, niż w dowolnej bazie. Dlatego – mając jakąś bazę – warto umieć znaleźć bazę ortogonalną. Metodą znajdowania bazy ortogonalnej jest *ortogonalizacja Grama-Schmidta.*

Ale wprzód jeszcze przykład, jak użycie bazy ortogonalnej ułatwia życie. Zobaczmy, jak wygląda *iloczyn skalarny* dwu wektorów rozłożonych w bazie ortogonalnej.

Niech $x, y \in V$, $h = (h_1, \ldots, h_n)$ – baza ortnormalna w V. Mamy:

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

zatem, *w bazie ortonormalnej iloczyn skalarny ma postać standardową.*

1.4 Ortogonalizacja Grama-Schmidta

Metoda działa zarówno w przypadku rzeczywistym, jak i zespolonym (w tym ostatnim trzeba bardziej uważać – o tym za chwilę). Załóżmy, że mamy bazę, na ogół nieortogonalną $e = (e_1, e_2, \ldots, e_n)$. Zdefiniujmy bazę $f = (f_1, f_2, \ldots, f_n)$ rekurencyjnym wzorem (*rekurencja Grama-Schmidta*)

$$f_1 = e_1; \quad f_2 = e_2 - \frac{(e_2|f_1)}{(f_1|f_1)} f_1; \quad \ldots, \quad f_j = e_j - \sum_{k=1}^{j-1} \frac{(e_j|f_k)}{(f_k|f_k)} f_k \tag{6}$$

Stw. Baza $f = (f_1, f_2, \ldots, f_n)$ jest ortogonalna.

Dow. jest indukcyjny. Sprawdźmy, że $(f_1|f_2) = 0$:

$$(f_2|f_1) = (e_2|f_1) - \frac{(e_2|f_1)}{(f_1|f_1)} (f_1|f_1) = 0.$$

Załóżmy teraz (założenie indukcyjne), że $(f_1, f_2, \ldots, f_{j-1})$ są prostopadłe, i sprawdźmy, że f_j jest prostopadły do wszystkich poprzednich (tzn. $(f_1, f_2, \ldots, f_{j-1})$): Weźmy jakiś wskaźnik $m < j$, i mamy:

$$(f_j|f_m) = (e_j|f_m) - \sum_{k=1}^{j-1} \frac{(e_j|f_k)}{(f_k|f_k)} (f_k|f_m);$$

w powyższej sumie, wszystkie (z wyjątkiem jednego) iloczyny $(f_k|f_m)$ są równe zeru (z założenia indukcyjnego). Wyraz nie równy zeru to ten, dla którego $k = m$. Mamy więc

$$(f_j|f_m) = (e_j|f_m) - \frac{(e_j|f_m)}{(f_m|f_m)}(f_m|f_m) = 0.$$

CBDO

Przykł. Zortogonalizujmy układ trzech wektorów w $\mathbb{R}^3$:

$$e_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Zgodnie z receptą (6), mamy:

$$f_1 = e_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad f_2 = e_2 - \frac{(e_2|f_1)}{(f_1|f_1)}f_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix},$$

... policzmy na boku: $(f_2|f_2) = \frac{1}{9}(1+1+4) = \frac{2}{3}$, $(f_2|e_3) = \frac{1}{3}(1-2) = \frac{-1}{3}$... i ostatni wektor:

$$f_3 = e_3 - \frac{(e_3|f_1)}{(f_1|f_1)}f_1 - \frac{(e_3|f_2)}{(f_2|f_2)}f_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{-1}{3}\cdot\frac{3}{2}\cdot\frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

$$= \frac{1}{6}\begin{bmatrix} 6-4+1 \\ 0-4+1 \\ 6-4-2 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Rachunkami do zrobienia już w pamięci widać, że f_1, f_2, f_3 są ortogonalne.

1.5 Dopełnienie ortogonalne podprzestrzeni

Niech $E \subset V$ (E może, ale nie musi, być podprzestrzenią). Zdefiniujmy zbiór $E^\perp$:

$$E^\perp = \{x \in V : \ (x|y) = 0 \ \forall y \in E\}; \tag{7}$$

innymi słowy, $E^\perp$ jest zbiorem wektor ów prostopadłych do E (tzn. do dowolnego wektora z E). Łatwo zobaczyć, że $E^\perp$ jest podprzestrzenią V.

Def. Gdy $E \subset V$ jest podprzestrzenią, to $E^\perp$ nazywamy *dopełnieniem ortogonalnym* podprzestrzeni E.

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

Dow. Jeśli $x \in E \cap E^\perp$, to $(x|x) = 0$, a więc $x = \mathbf{0}$. Stąd $E \cap E^\perp = \{\mathbf{0}\}$. Pokażemy, że $E + E^\perp = V$.

Niech $e_1, e_2, \ldots e_k$ będzie bazą w E i niech $\dim V = n$. Mamy:

$$E^\perp = \{y \in V : \ (y|e_j) = 0 \text{ dla } j = 1, \ldots, k\};$$

jest to k warunków w $n-$ wymiarowej przestrzeni; nie testujemy na razie czy wszystkie są niezależne, więc wymiar $E^\perp$ jest przynajmniej $n-k$: $\dim E^\perp \geqslant n-k$. Wymiar sumy $E+E^\perp$ jest więc:

$$\dim(E+E^\perp) \geqslant k+(n-k) \geqslant n.$$

Ostra nierówność nie może mieć miejsca, więc $\dim(E+E^\perp)$, czyli $E+E^\perp = V$.

Tak więc, skoro $E \cap E^\perp = \{\mathbf{0}\}$ i $E+E^\perp = V$, to $V = E \oplus E^\perp$.

CBDO

Wniosek: Jeśli E – podprzestrzenń V, to $(E^\perp)^\perp = E$. Wystarczy sprawdzić rozkład wektorów w bazach: Niech $(e_1,\ldots,e_k)$ – – baza w E, $(e_{k+1},\ldots,e_n)$ – baza w $E^\perp$; wtedy: $x \in E \Longrightarrow x = \sum_{i=1}^k a^i e_i \Longrightarrow x^\perp = \sum_{j=n-k}^n b^j e_j \Longrightarrow (x^\perp)^\perp = \sum_{i=1}^k c^i e_i$, zatem $E \subset (E^\perp)^\perp$. Mamy ponadto: $\dim E = k = n-(n-\dim E)) = \dim(E^\perp)^\perp$.

1.6 Rzut ortogonalny (prostopadły) na podprzestrzeń

Pamiętamy (było to w rozdz. o wart. i wekt. własnych), że gdy mamy rozkład przestrzeni na sumę prostą: $V = V_1 \oplus V_2$, to możemy zdefiniować *rzut* P_1 na podprzestrzeń V_1. (mówimy wtedy: 'Rzut na podprzestrzeń V_1 wzdłuż V_2') Rzutowanie to polega na zapominaniu składowej, należącej do drugiej podprzestrzeni. Tzn. gdy $v = v_1+v_2$, gdzie $v_1 \in V_1, v_2 \in V_2$, to $P_1 v = v_1$. Dla przestrzeni z iloczynem skalarnym też oczywiście można takie rzuty zdefiniować. Szczególnie ważne są rzuty w sytuacji, gdy przestrzeń V rozkładamy na podprzestrzeń E i jej dopełnienie ortogonalne $E^\perp$. Tak więc:

Def. *Rzutem ortogonalnym* na podprzestrzeń $E \subset V$ nazywamy rzut na podprzestrzeń E wzdłuż $E^\perp$. Oznaczamy go P_E.

Pokażemy, że gdy $(e_1,\ldots,e_k)$ jest bazą ortonormalną w E, to rzut ortogonalny wektora $v \in V$ na E jest dany wzorem:

$$P_E x = \sum_{i=1}^k (e_i|x) e_i. \tag{8}$$

Zauważmy najsampierw, że wzór ma szansę być prawdziwy, gdyż $P_E x \in E$ (jako liniowa kombinacja wektorów bazy E). Mamy ponadto, dla $j = 1,\ldots,k$:

$$(e_j|x-P_E x) = (e_j|x - \sum_{i=1}^k (e_i|x)e_i) = (e_j|x) - \sum_{i=1}^k (e_i|x)(e_i|e_j)$$

$$= (e_j|x) - \sum_{i=1}^k (e_i|x)\delta_{ij} = (e_j|x)-(e_j|x) = 0.$$

Skoro tak, to wektor: $(x-P_E x) \in E^\perp$. Otrzymaliśmy zatem rozkład wektora x na składowe w sumie prostej $E \oplus E^\perp$:

$$x = \underbrace{\sum_{i=1}^k (e_i|x)e_i}_{\in E} + \underbrace{x - \sum_{i=1}^k (e_i|x)e_i}_{\in E^\perp} \tag{9}$$

Uwagi.

1. Gdy baza w E jest *nieortonormalna*, to wyrażenie na rzut P_E też można napisać, ale jest ono znacznie bardziej skomplikowane.

2. Powróćmy do wzoru (6) w rekurencji Grama-Schmidta. Zapiszmy go w postaci:

$$f_k = e_k - \sum_{j=1}^{k-1} (e_k|h_j)h_j,$$

gdzie h_j są unormowane: $h_j = \frac{f_j}{||f_j||}$, tworzą więc bazę *ortonormalną*. Porównując powyższe wyrażenie z wzorem (8), możemy napisać, że

$$f_k = e_k - P_{E_{k-1}}e_k, \quad \text{gdzie} \quad E_{k-1} = \langle f_1, f_2, \ldots, f_{k-1} \rangle.$$

2 Kilka faktów z geometrii metrycznej, ciekawych [1] i pożytecznych [2]

2.1 Odległość punktu od podprzestrzeni

Def. *Odległość* punktu x od podprzestrzeni E definiujemy jako:

$$d(x, E) = \inf_{y \in E} d(x, y). \tag{10}$$

RYS. Uwaga. W ten sposób też definiujemy odległość punktu od *dowolnego* zbioru, nie tylko podprzestrzeni.

Ale dla podprzestrzeni, definicję tę można przekształcić, używając odpowiedniości: punkt przestrzeni $\longleftrightarrow$ wektor, w wyliczalny wzór:

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

... zauważamy, że $x - P_E x \in E^\perp$, zaś $P_E x - y \in E$, i korzystamy z tw. Pitagorasa, dostając...

$$\inf_{y \in E}(||x - P_E x||^2 + ||P_E x - y||^2) = \inf_{y \in E} ||x - P_E x||^2 = ||x - P_E x||^2. \tag{11}$$

Przedostatnia równość wynika z faktu, że wektorem y przebiegamy *całą* podprzestrzen E, więc trafiamy na $y = P_E x$, a w tym przypadku $||P_E x - y|| = 0$. **RYS.** Zaś ostatnia równość wynika z faktu, że $||x - P_E x||$ nie zależy od y, jest więc taka sama dla wszystkich $y \in E$.

Przykł. Znaleźć odległość pomiędzy prostymi $y = 2x$ i $y = 2x + 2$.

Rozw. Widać, że są to proste równoległe. Niby nie podpada to pod otrzymany przez nas wzór, ale jest inaczej – wystarczy wybrać *dowolny* punkt na jednej prostej i znaleźć odległość tego punktu od drugiej prostej. Tu, wybierzmy punkt $(x = 0, y = 2)$ na drugiej prostej. Przeformułujmy teraz zadanie w języku punktów i podprzestrzeni:

Znaleźć odległość punktu $v = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ od podprzestrzeni E rozpiętej przez wektor $f = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. **RYS.**

[1] kwestia gustu

[2] przynajmniej na egzaminie

Odległość znajdujemy korzystając z wzoru (11). Musimy najsampierw unormowa'ḅazę rozpinającą E; tu baza składa się z jednego wektora f. Wektor unormowany to $e = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Mamy:

$$P_E v = (v|e)e = \frac{1}{\sqrt{5}}(0 \cdot 1 + 2 \cdot 2)\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{4}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

więc

$$v - P_E v = \begin{bmatrix} 0 \\ 2 \end{bmatrix} - \frac{4}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -4 \\ -2 \end{bmatrix},$$

a szukana odległość to długość tego wektora:

$$d(v, E) = ||v - P_E v|| = \sqrt{\frac{20}{25}} = \frac{2}{\sqrt{5}}.$$

Kto nie wierzy, niech na papierze milimetrowym narysuje i zmierzy, albo prostą prostopadła narysuje, punkty przecięcia poznajduje, i odległość między nimi wykalkuluje

Powyższy przykład byłjedynie ilustracją działania wzoru (11). W większej liczbie wymiarów liczy się analogicznie – jedynie trudności rachunkowe mogą być większe.

2.2 Podstawowe pojęcia geometrii afinicznej

Dotąd zajmowaliśmy się przestrzeniami, podprzestrzeniami i odwzorowaniami liniowymi. Są jednak naturalne sytuacje, gdy warto te pojęcia w określonym kierunku rozszerzyć, otrzymując obiekty tzw. *afiniczne.*

Przed ich zdefiniowaniem, wprowadźmy najsampierw oznaczenie:

Ozn. Niech Z – podzbiór przestrzeni wektorowej V, zaś x – jakiś wektor z V. Wprowadźmy oznaczenie (możemy je roboczo nazwać 'zbiór Z przesunięty o wektor x')

$$x + Z = Z + x = \{y = z + x, z \in Z\}.$$

RYS. Jest oczywiste, że zachodzi:

1. $x + (y + Z) = (x + y) + Z$,
2. Jeśli Z – podprzestrzeń, $x \in Z$ to $x + Z = Z$.

Def. *Podprzestrzenią afiniczną* P przestrzeni V nazywamy zbiór postaci

$$P = p_0 + L, \tag{12}$$

gdzie p_0 – pewien wektor należący do V, zaś L – podprzestrzeń *wektorowa* przestrzeni V.

Oczywiście, $p_0 \in P$. Ponadto p_0 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 dla zadanej podprzestrzeni afinicznej P, podprzestrzeń L jest wyznaczona jednoznacznie. L nazywa się *przestrzenią styczną* do P.

Dla $\dim L = 0,1,2$ podprzestrzenie afiniczne nazywają się odpowiednio *punktem, prostą, płaszczyzną* afiniczną.

Przykład. Niech $T : V \to W$ – odwzorowanie liniowe z przestrzeni V do W. Wiemy, że zbiór rozwiązań równania $Tx = \mathbf{0}$ jest podprzestrzenią liniową w V. Natomiast zbiór rozwiązań równania *niejednorodnego* $Tx = y$, $y \in W$ – dany wektor, *nie* jest podprzestrzenią liniową w V. Jest natomiast podprzestrzenią *afiniczną*. (może też być pusty).

Def. Mówimy, że podprzestrzenie afiniczne P_1, P_2 są *równoległe*, jeśli $L_1 \subset L_2$ lub $L_2 \subset L_1$ (mogą się też pokrywać). (Tu L_1, L_2 są przestrzeniami stycznymi odpowiednio do P_1 i P_2).

Łatwo sprawdzić, że przecięcie dwóch podprzestrzeni afinicznych jest puste albo też podprzestrzenią afiniczną. To samo dotyczy dowolnej ilości przestrzeni afinicznych.

2.3 Odległość podprzestrzeni afinicznych

Def. *Odległość* dwóch podprzestrzeni afinicznych Π oraz Π' definiujemy jako

$$d(\Pi, \Pi') = \inf_{p \in \Pi,\ p' \in \Pi'} d(p, p'). \tag{13}$$

Jak i uprzednio, przekształćmy to wyrażenie we wzór bardziej nadający się do wyliczeń. W tym celu zacznijmy od przedstawienia:

$$\Pi = p_0 + L, \quad \Pi' = p_0' + L'$$

gdzie L, L' są podprzestrzeniami wektorowymi (stycznymi do Π oraz Π' odpowiednio). Wynika stąd przedstawienie dla dowolnych punktów $p \in \Pi$, $p' \in \Pi'$:

$$p = p_0 + y, \quad p' = p_0' + y' \quad \text{gdzie} \quad y \in L, \quad y' \in L'$$

Wykorzystując to, mamy:

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

$$= \inf_{y \in L,\ y' \in L'} ||(p_0 - p_0') + (y - y')|| = \inf_{z \in (L + L')} ||(p_0 - p_0') + z|| = \dots$$

... w czym rozpoznajemy *odległość punktu* $(p_0 - p_0')$ *od podprzestrzeni* $(L + L')$, na którą to odległość wyprowadziliśmy niedawno wzór (11), który tu doprowadzi do kalkulowalnego wyrażenia na odległość:

$$d(\Pi, \Pi') = d((p_0 - p_0'), (L + L')) = ||(p_0 - p_0') - P_{(L+L')}(p_0 - p_0')||. \tag{14}$$

Przykład. Dla naprzykładu[3] obliczmy odległość pomiędzy prostymi skośnymi, zadanymi parametrycznie (parametr $t \in \mathbb{R}$), w $\mathbb{R}^3$:

$$\Pi = p_0 + L = p_0 + tl = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \Pi' = p_0' + L' = p_0' + tl' = \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

[3]jak mawiał kolega kierownik w audycjach Jacka Fedorowicza

Rozw. Liczymy co trzeba:

$$\pi = p_0' - p_0 = \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix}, \quad L + L' = \langle l, l' \rangle \,.$$

Będziemy liczyć rzuty na $L+L'$, zatem musimy mieć tam bazę *ortonormalną*. Baza (l_1, l_2) ortonormalna nie jest. Skoro tak, to musimy ją zortonormalizować. Najpierw zortogonalizujmy, oznaczając bazę ortogonalną jako (f_1, f_2):

$$f_1 = l_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad f_2 = l_2 - \frac{(f_2|l_1)}{l_1|l_1)} l_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \frac{6}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},$$

które – jak widać gołym okiem – są ortogonalne. Normalizujemy (dzieląc przez długośći) i otrzymujemy bazę ortonormalną:

$$e_1 = \frac{1}{\sqrt{3}} f_1, \qquad e_2 = \frac{1}{\sqrt{2}} f_2.$$

Dalej:

$$P_{L+L'}\pi = (\pi|e_1)e_1 + (\pi|e_2)e_2 = \frac{1}{3}(3\cdot 1 + 5\cdot 1 + 8\cdot 1) f_1 + \frac{1}{2}(3\cdot(-1) + 5\cdot 0 + 8\cdot 1) = \frac{16}{3} f_1 + \frac{5}{2} f_2$$

$$= \frac{1}{6}\begin{bmatrix} 32 - 15 \\ 32 + 0 \\ 32 + 15 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 17 \\ 32 \\ 47 \end{bmatrix},$$

$$\pi - P_{L+L'}\pi = \frac{1}{6}\begin{bmatrix} 18 - 17 \\ 30 - 32 \\ 48 - 47 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix},$$

i mamy odległość:

$$d(\Pi, \Pi') = \sqrt{\frac{1^2 + (-2)^2 + 1^2}{6^2}} = \frac{1}{\sqrt{6}}.$$

Przykład 2. Obliczyć odległość pomiędzy prostymi skośnymi, zadanymi parametrycznie (parametr $t \in \mathbb{R}$), w $\mathbb{R}^3$:

$$\Pi = p_0 + L = p_0 + t\, l = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \Pi' = p_0' + L' = p_0' + t\, l' = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Odp. Tu proste, rzekomo skośne, przecinają się! Oblicz, Czytelniku (Czytelniczko), w jakim punkcie.

2.4 Objętość równoległościanu rozpiętego na k wektorach w $\mathbb{R}^n$

Weźmy k wektorów $v_1, v_2, \ldots, v_k$ należących do $\mathbb{R}^n$. Definiujemy *równoległościan* $\mathcal{P}$ rozpięty na tych wektorach jako zbiór **RYS.**:

$$\mathcal{P} = \{\lambda^1 v_1 + \lambda^2 v_2 + \cdots + \lambda^k v_k\}, \quad 0 \leqslant \lambda^i \leqslant 1\}. \tag{15}$$

Definiujemy jego $k-$wymiarową objętość indukcyjnie wzorem 'pole podstawy razy wysokość':

$$Vol(v_1, \ldots, v_k) = Vol(v_1, \ldots, v_{k-1}) \cdot d(v_k, \langle v_1, \ldots, v_{k-1} \rangle) \tag{16}$$

zaś punkt startowy indukcji ('objętość jednowymiarowa') to *długość* wektora:

$$Vol(v_1) = ||v_1||.$$

Def. *Macierzą Grama* układu wektorów $v_1, v_2, \ldots, v_k$ (ozn. $G(v_1, v_2, \ldots, v_k)$) nazywamy macierz $k \times k$, której elementami są iloczyny skalarne wektorów v_i, tzn.

$$G_{ij} = (v_i|v_j).$$

Stw.

$$Vol(v_1, \ldots, v_k) = \sqrt{\det G(v_1, v_2, \ldots, v_k)}.$$

Dow. będzie indukcyjny. Dla $k = 1$ teza jest prawdziwa:

$$Vol(v_1) = ||v_1|| = \sqrt{(v_1|v_1)} = \sqrt{\det G(v_1)}.$$

Pokażemy, że z założenia o prawdziwości tezy dla $k - 1$ wynika jej prawdziwość dla k. Mamy:

$$Vol(v_1, \ldots, v_k)^2 = Vol(v_1, \ldots, v_{k-1})^2 \cdot d(v_k, \langle v_1, \ldots, v_{k-1} \rangle)^2$$
$$= Vol(v_1, \ldots, v_{k-1})^2 \cdot (v_k - Pv_k|v_k - Pv_k) = \det G(v_1, v_2, \ldots, v_{k-1}) \cdot (v_k - Pv_k|v_k - Pv_k)$$
$$= \det G(v_1, v_2, \ldots, v_{k-1}, v_k - Pv_k)$$

gdzie P jest rzutem na podprzestrzeń $\langle v_1, \ldots, v_{k-1} \rangle$. W ostatniej równości wykorzystaliśmy fakt, iż $v_k - Pv_k$ jest ortogonalny do wszystkich wektorów $v_1, \ldots, v_{k-1}$, zatem macierz Grama : $G(v_1, v_2, \ldots, v_{k-1}, v_k - Pv_k)$ ma, w ostatnim wierszu (i kolumnie) niezerowy tylko wyraz diagonalny G_{kk}, a pozostałe $G_{kj}, j < k$ są równe zeru.

Pozostaje pokazać, że

$$\det G(v_1, v_2, \ldots, v_{k-1}, v_k - Pv_k) = \det G(v_1, v_2, \ldots, v_{k-1}, v_k).$$

Wynika to z następującego lematu.

Lemat. Jeśli wektory $w_1, \ldots, w_k$ są kombinacjami liniowymi wektorów $v_1, v_2, \ldots, v_k$):

$$w_r = \sum_{i=1}^{k} a^i{}_r v_i, \qquad r = 1, \ldots, k$$

to

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

gdzie A jest macierzą współczynników $(a^i{}_r)$. Stąd wynika, że

$$\det G(w_1, \ldots, w_k) = (\det A)^2 \det G(v_1, \ldots, v_k). \tag{17}$$

Dow. (lematu): Liczymy iloczyn skalarny:

$$(w_r|w_s) = \left(\sum_{i=1}^{k} a^i{}_r v_i \middle| \sum_{j=1}^{k} a^j{}_s v_j \right) = \sum_{i=1}^{k} \sum_{j=1}^{k} a^i{}_r a^j{}_s (v_i|v_j) = \sum_{i,j} (A^T)^r{}_i G_{ij} (A)^j{}_s$$

CBDO

(lematu)

Koniec dowodu Stwierdzenia: W macierzy A wybieramy: $w_1 = v_1, \ldots, w_{k-1} = v_{k-1}$, $w_k = v_k - Pv_k$. Ponieważ Pv_k jest liniową kombinacją wektorów $v_1, \ldots, v_{k-1}$, to macierz A ma strukturę *górnotrójkątną*, a na diagonali ma same jedynki, więc $\det A = 1$. Korzystamy z wzoru (17) i mamy co trzeba.

CBDO

Przykł. Obliczyć pole powierzchni równoległoboku, wyciętego przez graniastosłup o podstawie będącej jednostkowym kwadratem o wierzchołkach w punktach: $(0,0), (0,1), (1,0), (1,1)$ z płaszczyzny zadanej równaniem $z = ax + by + c$.

Tzn. jest to pole powierzchni (objętość dwuwymiarowa) równoległościanu rozpiętego na wektorach:

$$v_x = \begin{bmatrix} 1 \\ 0 \\ a \end{bmatrix}, \qquad v_y = \begin{bmatrix} 0 \\ 1 \\ b \end{bmatrix}$$

Liczymy macierz Grama:

$$G = \begin{bmatrix} (v_x|v_x) & (v_x|v_y) \\ (v_y|v_x) & (v_y|v_y) \end{bmatrix} = \begin{bmatrix} 1+a^2 & ab \\ ab & 1+b^2 \end{bmatrix},$$

a szukane pole powierzchni $\mathcal{V}_2$ to pierwiastek z wyznacznika:

$$\mathcal{V}_2 = \sqrt{\det G} = \sqrt{(1+a^2)(1+b^2) - a^2b^2} = \sqrt{1+a^2+b^2}.$$

Wzór ten przyda się nam zaniedługo...