Metody algebry liniowej w fizyce

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Kraków 2026

Literatura

[1] Z. Furdzik, J. Maj-Kluskowa, A. Kulczycka, M. Sękowska, *Nowoczesna matematyka dla inżynierów. Część I - Algebra*, Wydawnictwo AGH, Kraków, 1993

[2] B. Gleichgewicht, *Algebra*, PWN, Warszawa, 1983, wydanie III zmienione

[3] T. Jurlewicz, Z. Skoczylas, *Algebra i geometria analityczna. Definicje, twierdzenia, wzory*, Oficyna Wydawnicza GiS, Wrocław, 2020, wydanie XXII

[4] T. Jurlewicz, Z. Skoczylas, *Algebra liniowa. Definicje, twierdzenia, wzory*, Oficyna Wydawnicza GiS, Wrocław, 2015, wydanie VIII poprawione

[5] A. I. Kostrikin, *Wstęp do algebry. Tomy 1 - Postawy algebry*, Wydawnictwo Naukowe PWN, 2004.

[6] A. I. Kostrikin, *Wstęp do algebry. Tom 2 - Algebra liniowa*, Wydawnictwo Naukowe PWN, 2004.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Algebra
Wykład nr 1
Data: .....

TEMAT: *Działania i wybrane struktury algebraiczne*

1.1 Notacja

Przyjmujemy następująca notację.

$\mathbb{N} = \{1, 2, 3, \ldots\}$	zbiór liczb naturalnych
$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$	
$\mathbb{Z}$	zbiór liczb całkowitych
$\mathbb{Q}$	zbiór liczb wymiernych
$\mathbb{R}$	zbiór liczb rzeczywistych

Oznaczenia kwantyfikatorów

$\forall$	kwantyfikator ogólny (duży)	*dla każdego*
$\exists$	kwantyfikator szczegółowy (mały)	*istnieje*
$\exists!$		*istnieje dokładnie jeden*

Symbol sumy i iloczynu

Niech $n \in \mathbb{N}$ oraz niech $a_1, a_2, \ldots, a_m, a_{m+1}, \ldots, a_n$ to dowolny ciąg liczb.

Sumę $a_1 + a_2 + \ldots + a_n$ oznaczamy symbolem $\sum_{i=1}^{n} a_i$. Symbol i to wskaźnik sumowania lub indeks sumowania, m to dolna granica sumowania, zaś n to górna granica sumowania.

Zauważmy że dla dowolnego $m \in \mathbb{N}$, $m < n$ oraz $c \in \mathbb{R}$ zachodzą równości
$a_1 + a_2 + \ldots + a_n = (a_1 + a_2 + \ldots + a_m) + (a_{m+1} + a_{m+2} + \ldots + a_n)$ oraz
$c(a_1 + a_2 + \ldots + a_n) = ca_1 + ca_2 + \ldots + ca_n$, czyli

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{m} a_i + \sum_{i=m+1}^{n} a_i, \qquad c\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} ca_i.$$

Przykład 1.1.1. $\sum_{i=1}^{6} i = 1 + 2 + 3 + 4 + 5 + 6 = 21$
$\sum_{i=5}^{9} \frac{i}{i+2} = \frac{5}{7} + \frac{6}{8} + \frac{7}{9} + \frac{8}{10} + \frac{9}{11}$
$\sum_{i=0}^{n} \left(\frac{1}{2}\right)^i = 1 + \frac{1}{2} + \frac{1}{4} + \ldots + \left(\frac{1}{2}\right)^n = \frac{1-\left(\frac{1}{2}\right)^{n+1}}{1-\frac{1}{2}}$

Iloczyn $a_1 \cdot a_2 \cdot \ldots \cdot a_{n-1} \cdot a_n$ oznaczamy symbolem $\prod_{i=1}^{n} a_i$. Symbol i to wskaźnik iloczynu, m to dolny wskaźnik iloczynu, zaś n to górny wskaźnik iloczynu.

Przykład 1.1.2. $\prod_{i=0}^{6}(9-j) = 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3$
$\prod_{i=1}^{n} i = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (-1) \cdot n = n!$

$\underbrace{x \cdot x \cdot \ldots \cdot x}_{n \text{ razy}} = \prod_{i=1}^{n} x$

Definicja 1.1.3. *Iloczynem kartezjańskim* zbiorów A i B nazywamy zbiór

$$A \times B := \{(a,b) : \ a \in A \wedge b \in B\}.$$

Analogicznie definiujemy
$A_1 \times A_2 \times \ldots \times A_n := \{(a_1, a_2, \ldots a_n) : \ a_1 \in A_1 \wedge a_2 \in A_2 \wedge \ldots \wedge a_n \in A_n\}$.
Oznaczamy $A^2 = A \times A$ oraz $A^n = \underbrace{A \times A \times \ldots \times A}_{n-razy}$.

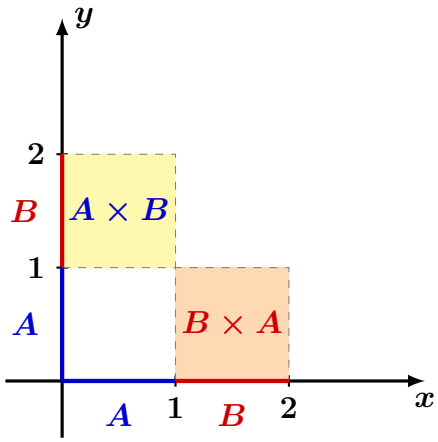
Przykład 1.1.4. Wyznacz $A \times B$, $B \times A$, B^2.

i) $A = \{3,4,5\}$, $B = \{5,7\}$

$A \times B = \{(3,5),(3,7),(4,5),(4,7),(5,5),(5,7)\}$
$B \times A = \{(5,3),(5,4),(5,5),(7,3),(7,4),(7,5)\}$ $\quad A \times B \neq B \times A$
$B^2 = \{(5,5),(5,7),(7,5),(77)\}$

ii) przedziały $A = (0,1) \subset \mathbb{R}$, $B = (1,2) \subset \mathbb{R}$

$A \times B = \{(x,y) : 0 < x < 1 \wedge 1 < y < 2\}$ $\quad B \times A = \{(x,y) : 1 < x < 2 \wedge 0 < y < 1\}$



iii) $A = B = \mathbb{R}$

$A \times B = B \times A = A^2 = B^2$
Oznaczamy $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x,y) : \ x,y \in \mathbb{R}\}$ - płaszczyzna rzeczywista

1.2 Działania wewnętrzne i ich własności

Niech A będzie zbiorem niepustym.

Definicja 1.2.1. Dowolną funkcję $h : A \times A \to A$ nazywamy *działaniem wewnętrznym* w zbiorze A. Wartość funkcji $h(a, b) \in A$ nazywamy *wynikiem* działania dla pary argumentów $a, b \in A$.

Przykład 1.2.2. i) Dodawanie jest działaniem wewnętrznym w $\mathbb{N}$.

ii) Odejmowanie nie jest działaniem wewnętrznym w $\mathbb{N}$.

iii) Dodawanie nie jest działaniem wewnętrznym w $\mathbb{R} \setminus \mathbb{Q}$.

Własności działań wewnętrznych

Definicja 1.2.3. Niech $* : A \times A \to A$ będzie działaniem wewnętrznym.

i) Działanie $*$ nazywamy *przemiennym*, jeśli $\forall\, a, b \in A \quad a * b = b * a$.

ii) Działanie $*$ nazywamy *łącznym*, jeśli $\forall\, a, b, c \in A \quad (a * b) * c = a * (b * c)$.

iii) Element $e \in A$ nazywamy *elementem neutralnym* działania $*$, jeśli $\forall\, a \in A \quad a * e = e * a = a$.

Przykład 1.2.4. i) Dodawanie jest działaniem wewnętrznym łącznym i przemiennym w $\mathbb{R}$. 0 jest elementem neutralnym.

ii) Mnożenie jest działaniem wewnętrznym łącznym i przemiennym w $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$. 1 jest elementem neutralnym.

Twierdzenie 1.2.5. Jeżeli element neutralny istnieje to jest jedyny.

Dowód. Przypuśćmy, że istnieją $e_1, e_2 \in A$ będące elementami neutralnymi w A. Wówczas $e_2 = e_1 * e_2 = e_1$. $\square$

Przykład 1.2.6. Niech X to dowolny zbiór.

i) Zbiór pusty $\varnothing$ jest elementem neutralnym w $\Big(P(X), \cup\Big)$.

ii) Zbiór X jest elementem neutralnym w $\Big(P(X), \cap\Big)$.

Definicja 1.2.7. Niech $* : A \times A \to A$ będzie działaniem wewnętrznym, posiadającym element neutralny $e \in A$. Dla dowolnego $a \in A$ każdy element $a' \in A$ taki, że $a * a' = a' * a = e$, nazywamy *elementem symetrycznym* do a względem działania $*$.

Przykład 1.2.8. i) W $(\mathbb{R}, +)$ elementem symetrycznym do 5 jest -5 (tzw. element przeciwny).

ii) W $(\mathbb{R}^*, \cdot)$ elementem symetrycznym do 5 jest $\frac{1}{5}$ (tzw. element odwrotny).

iii) W $(\mathbb{R}, \circ)$, gdzie $x \circ y = x + y + 1$, elementem neutralnym jest -1, zaś elementem symetrycznym do x jest $-2 - x$.
Istotnie, dla dowolnego $x \in \mathbb{R}$ mamy

$$x \circ e = x \Leftrightarrow x + e + 1 = x \Rightarrow e + 1 = 0 \Leftrightarrow e = -1 \in \mathbb{R},$$

$$x \circ x' = e \Leftrightarrow x + x' + 1 = -1 \Rightarrow x' = -2 - x \in \mathbb{R}.$$

Twierdzenie 1.2.9. Jeżeli działanie wewnętrzne jest łączne oraz posiada element neutralny, to każdy element posiada co najwyżej jeden element symetryczny.

1.3 Podstawowe struktury algebraiczne

1.3.1 Pojęcie grupy

Struktura algebraiczna - zbiór wraz z działaniami wewnętrznymi w tymże zbiorze

Niech G będzie zbiorem niepustym, zaś $* : G \times G \to G$ działaniem wewnętrznym w G.

Definicja 1.3.1. i) Parę $(G, *)$ nazywamy *półgrupą*, jeżeli działanie jest łączne.

ii) Parę $(G, *)$ nazywamy *monoidem*, jeżeli działanie jest łączne i posiada element neutralny.

iii) Parę $(G, *)$ nazywamy *grupą*, jeżeli działanie jest łączne, posiada element neutralny oraz każdy element G posiada element symetryczny względem $*$ w G.

Jeśli dodatkowo działanie $*$ jest przemienne, to mówimy o *półgrupie przemiennej, monoidzie przemiennym, grupie przemiennej (abelowej).*

Przykład 1.3.2. Oznaczmy $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$. Czy podane struktury to półgrupy/grupy (przemienne)?

	$(\mathbb{N}, +)$	$(\mathbb{Z}, \cdot)$	$(\mathbb{Q}, +)$	$(\mathbb{Q}, \cdot)$	$(\mathbb{R}^*, +)$	$(\mathbb{R}^*, \cdot)$
wewnętrzność	✓	✓	✓	✓	nie $-1+1=0$ $0 \notin \mathbb{R}^*$	✓
łączność	✓	✓	✓	✓		✓
przemienność	✓	✓	✓	✓		✓
el. neutralny	brak $0 \notin \mathbb{N}$	✓ $1 \in \mathbb{Z}$	✓ $0 \in \mathbb{Q}$	✓ $1 \in \mathbb{Q}$		✓ $1 \in \mathbb{R}^*$
el. symetryczny		brak $2 \cdot b = 1$ $b = \frac{1}{2} \notin \mathbb{Z}$	✓ $a + a' = 0$ $a' = -a \in \mathbb{Q}$	brak $a \cdot a' = 1$ $a' = \frac{1}{a}$ nie dla $a = 0$		✓ $a \cdot a' = 1$ $a' = \frac{1}{a}$ $\forall a \in \mathbb{R}^*$
	półgrupa przemienna bez el. neutr.	monoid przemienny	grupa abelowa	monoid przemienny		grupa abelowa

Przykład 1.3.3. $(\mathbb{Z}, +)$ jest grupą przemienną. Rozwiązując w $\mathbb{Z}$ równanie $x + 5 = 8$, korzystamy z własności działania w grupie.

$$
\begin{array}{rl}
x + 5 = 8 & \\
(x + 5) - 5 = 8 - 5 & \text{istnieje element przeciwny do 5, należący do } \mathbb{Z} \\
x + (5 - 5) = 3 & \text{dodawanie jest łączne w } \mathbb{Z} \\
x + 0 = 3 & \text{istnieje element neutralny} \\
x = 3 &
\end{array}
$$

Przykład 1.3.4. Niech $X = \{1, 2, 4, \ldots, 2^n, \ldots\}$, $n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, zaś działanie $\circ$ to mnożenie liczb. Czy $(X, \circ)$ jest półgrupą/grupą (abelową)?

Jeśli $a, b, c \in X$, to istnieją $k, m, n \in \mathbb{N}_0$ takie, ze $a = 2^n, b = 2^m, c = 2^k$.

wewnętrzne $\quad a \circ b = 2^n \cdot 2^m = 2^{n+m} \in A$

przemienne $\quad a \circ b = 2^n \cdot 2^m = 2^{n+m} = 2^{m+n} = 2^m \cdot 2^n = b \circ a$

łączne $\quad (a \circ b) \circ c = 2^{n+m} \cdot 2^k = 2^{n+m+k} = 2^n \cdot 2^{m+k} = a \circ (b \circ c)$

el. neutralny $\quad e = 2^s, s \in \mathbb{N}_0 \quad a \circ e = a \Leftrightarrow 2^{n+s} = 2^n \Rightarrow s = 0,\ e = 1 \in X$

brak el. odwrotnego $\quad a \circ b = 1 \Leftrightarrow 2^{n+m} = 2^0 \Leftrightarrow m = -n \Rightarrow m = -n \notin \mathbb{N}_0$

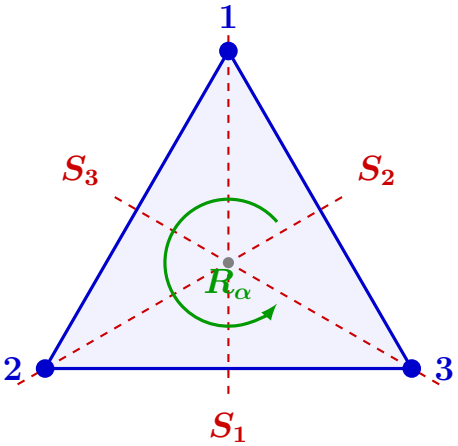
Wniosek: monoid przemienny (tj. półgrupa przemienna z jedynką)

Przykład 1.3.5.
Niech $T \subset \mathbb{R}^2$ będzie trójkątem równobocznym o wierzchołkach oznaczonych jako $1, 2, 3$.
Rozważmy zbiór D_3 wszystkich izometrii własnych trójkąta T, czyli przekształceń geometrycznych płaszczyzny, które nakładają ten trójkąt na samego siebie. Jest to tzw. *grupa dihedralna stopnia 3* lub *grupa symetrii trójkąta równobocznego*.

Oznaczamy

id	przekształcenie tożsamościowe,
$R_{\frac{2}{3}\pi}, R_{\frac{4}{3}\pi}$	obroty dokonywane wokół środka trójkąta o kąty odpowiednio $\frac{2}{3}\pi$ i $\frac{4}{3}\pi$ w kierunku przeciwnym do ruchu wskazówek zegara,
S_1, S_2, S_3	symetrie osiowe względem symetralnych boków przechodzących przez odpowiadające im wierzchołki.

Zauważmy, że struktura $(D_3, \circ)$, gdzie działaniem wewnętrznym $\circ$ jest składanie przekształceń, ma strukturę grupy. Jest to najmniejszy przykład grupy nieprzemiennej (nieabelowej).



$\boldsymbol{D_3 = \{id, R_{\frac{2}{3}\pi}, R_{\frac{4}{3}\pi}, S_1, S_2, S_3\}}$

- $\boldsymbol{id}$ – identyczność
- $\boldsymbol{R_\alpha}$ – obroty o kąt $\boldsymbol{\alpha \in \{\frac{2}{3}\pi, \frac{4}{3}\pi\}}$
- $\boldsymbol{S_i}$ – odbicia względem osi $\boldsymbol{i}$-tego wierzchołka

$\boldsymbol{S_1 \circ R_{\frac{2}{3}\pi} = S_2}$

$\boldsymbol{R_{\frac{2}{3}\pi} \circ S_1 = S_3 \implies}$ **grupa nieabelowa**

Tabela działań w grupie D_3. Wynikiem $g \circ f$ jest element w wierszu g i kolumnie f.

$\circ$	id	$R_{\frac{2}{3}\pi}$	$R_{\frac{4}{3}\pi}$	S_1	S_2	S_3
id	id	$R_{\frac{2}{3}\pi}$	$R_{\frac{4}{3}\pi}$	S_1	S_2	S_3
$R_{\frac{2}{3}\pi}$	$R_{\frac{2}{3}\pi}$	$R_{\frac{4}{3}\pi}$	id	S_3	S_1	S_2
$R_{\frac{4}{3}\pi}$	$R_{\frac{4}{3}\pi}$	id	$R_{\frac{2}{3}\pi}$	S_2	S_3	S_1
S_1	S_1	S_2	S_3	id	$R_{\frac{2}{3}\pi}$	$R_{\frac{4}{3}\pi}$
S_2	S_2	S_3	S_1	$R_{\frac{4}{3}\pi}$	id	$R_{\frac{2}{3}\pi}$
S_3	S_3	S_1	S_2	$R_{\frac{2}{3}\pi}$	$R_{\frac{4}{3}\pi}$	id

1.3.2 Grupa permutacji

W trakcie wykładu poznamy wiele przykładów grup, w tym grupy macierzy takie jak ogólna grupa liniowa, grupa ortogonalna i grupa unitarna. Zajmiemy się teraz grupami permutacji zbiorów skończonych.

Definicja 1.3.6. Niech X będzie zbiorem niepustym. Wzajemnie jednoznaczne (tj. bijektywne) odwzorowanie $\sigma : X \to X$ nazywamy *premutacją* zbioru X.

Zbiór wszystkich premutacji zbioru X będziemy oznaczać symbolem $S(X)$.

Twierdzenie 1.3.7. Niech X będzie zbiorem niepustym. Zbiór $S(X)$ wraz z działaniem składania odwzorowań ma strukturę grupy.

Dowód. Złożenie bijekcji jest bijekcją, zatem działanie jest wewnętrze w zbiorze $S(X)$. Składanie odwzorowań jest działaniem łącznym. Elementem neutralnym jest funkcja identycznościowa id_X, która jest bijekcją. Funkcja odwrotna do danej bijekcji również jest bijekcją. □

Grupę $(S(X), \circ)$ nazywamy *grupą permutacji* zbioru X lub *grupą symetryczną* zbioru X.

Zajmiemy się teraz grupami permutacji zbiorów skończonych.

Niech X będzie zbiorem n-elementowym, gdzie $n \in \mathbb{N}$. Natura elementów nie będzie miała dla nas znaczenia. Bez straty dla ogólności rozumowań możemy przyjąć, że $X = \{1, 2, \ldots, n\}$.

Grupę permutacji zbioru n-elemetowego będziemy oznaczać symbolem S_n.
Liczba wszystkich permutacji zbioru n-elementowego (to jest moc zbioru S_n) jest równa $n!$.

Jeśli $\sigma, \tau \in S_n$, to zamiast $\sigma \circ \tau$ będziemy pisać krótko $\sigma\tau$.
Podobnie zamiast $\sigma \circ \sigma$ będziemy pisać σ^2.
Często też zamiast o złożeniu permutacji mówi się o *iloczynie* permutacji.
Permutację $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, $j \mapsto \sigma(j)$ zapisujemy w postaci dwuwierszowej, podając wszystkie wartości przekształcenia. [1]

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & \mathrm{n} \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}$$

W szczególności $\mathrm{id} : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, $\mathrm{id}(j) = j$ zapisujemy w postaci

$$\mathrm{id} = \begin{pmatrix} 1 & 2 & \ldots & \mathrm{n} \\ 1 & 2 & \ldots & \mathrm{n} \end{pmatrix}$$

[1] Ciąg $\sigma(1)$,$\sigma(2)$,…,$\sigma(n)$, to liczby $1, 2, \ldots, n$ ustawione w innym porządku.

Przykład 1.3.8. Permutacje $\tau, \sigma \in S_8$ dane są za pomocą przyporządkowań

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 5 & 8 & 4 & 7 & 1 & 6 & 2 \end{pmatrix}, \qquad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 6 & 3 & 7 & 8 & 1 & 5 & 2 \end{pmatrix}.$$

Podamy postać permutacji σ^{-1}, $\sigma\tau$ oraz $\tau\sigma$.
Wyznaczamy σ^{-1}. Z definicji $\sigma^{-1}(j) = k \Leftrightarrow \sigma(k) = j$.

$$\begin{array}{cccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \sigma & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 4 & 6 & 3 & 7 & 8 & 1 & 5 & 2 \end{array} \qquad \begin{array}{cccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \sigma^{-1} & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 6 & 8 & 3 & 1 & 7 & 2 & 4 & 5 \end{array}$$

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 8 & 3 & 1 & 7 & 2 & 4 & 5 \end{pmatrix}$$

Wyznaczamy $\tau\sigma(j) = \tau\Big(\sigma(j)\Big)$, dla $j = 1, 2, \ldots, n$.

$$\begin{array}{cccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \sigma & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 4 & 6 & 3 & 7 & 8 & 1 & 5 & 2 \\ \tau & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 4 & 1 & 8 & 6 & 2 & 3 & 7 & 5 \end{array}$$

$$\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 5 & 8 & 4 & 7 & 1 & 6 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 6 & 3 & 7 & 8 & 1 & 5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 1 & 8 & 6 & 2 & 3 & 7 & 5 \end{pmatrix}$$

W podobny sposób wyznaczamy $\sigma\tau(j) = \sigma\Big(\tau(j)\Big)$, dla $j = 1, 2, \ldots, n$.

$$\begin{array}{cccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \tau & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 3 & 5 & 8 & 4 & 7 & 1 & 6 & 2 \\ \sigma & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ & 3 & 8 & 2 & 7 & 5 & 4 & 1 & 6 \end{array} \qquad \tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 8 & 2 & 7 & 5 & 4 & 1 & 6 \end{pmatrix}$$

Zauważmy, że $\tau\sigma \neq \sigma\tau$. [2]

Twierdzenie 1.3.9. i) Grupy S_1 oraz S_2 są grupami przemiennymi.

ii) Dla $n \geq 3$ grupy S_n są nieprzemienne.

Dowód. i) Jest oczywiste, że grupa jednoelementowa $S_1 = \left\{\sigma = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ jest przemienna. Dla $n = 2$ mamy

$$S_2 = \left\{\sigma = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \tau = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\right\}.$$

[2] Składanie odwzorowań nie jest działaniem przemiennym,

Możemy sprawdzić, że $\sigma^2 = \tau^2 = \sigma$ oraz $\sigma\tau = \tau\sigma = \tau$. $\square$

Definicja 1.3.10. Permutację $\sigma \in S_n$ nazywamy *permutacją cykliczną rzędu k* lub *cyklem k-wyrazowym* lub *cyklem długości k*, jeśli istnieje w zbiorze $\{1, 2, \ldots, n\}$ podzbiór k-elementowy $\{a_1, a_2, \ldots, a_k\}$ taki, że

$$\sigma(a_1) = a_2, \;\; \sigma(a_2) = a_3, \;\; \ldots \;\; \sigma(a_{k-1}) = a_k, \;\; \sigma(a_k) = a_1$$

oraz

$$\forall j \in \{1, 2, \ldots, n\} \setminus \{a_1, a_2, \ldots, a_k\} \quad \sigma(j) = j.$$

Cykl k-wyrazowy oznaczamy symbolem $\big(a_1, a_2, \ldots, a_k\big)$, pomijając w zapisie wyrazy, które przechodzą same na siebie.

Zgodnie z przyjętą definicją

$$\big(a_1, a_2, \ldots, a_k\big) = \big(a_2, a_3, \ldots, a_k, a_1\big) = \big(a_3, a_4, \ldots, a_k, a_1, a_2\big) \quad \text{itd.}$$

Przykład 1.3.11. i) Rozważmy permutację $\sigma \in S_6$ postaci $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 5 & 1 & 2 & 4 \end{pmatrix}$.

Zauważmy, że $1 \to 3 \to 5 \to 2 \to 6 \to 4 \to 1$. Zatem $\sigma = \big(1, 3, 5, 2, 6, 4\big)$

ii) Rozważmy permutację $\sigma \in S_7$ postaci $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 4 & 3 & 6 & 5 & 2 & 7 \end{pmatrix}$.

Permutacja ta jest cyklem 3-wyrazowym i możemy zapisać $\sigma = \big(2, 4, 6\big)$.

Zauważmy, że pisząc $\sigma = \big(2, 4, 6\big)$, musimy podać do jakiego zbioru S_n należy σ.

Dla przykładu, jeśli $\sigma \in S_6$, to $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 3 & 6 & 5 & 2 \end{pmatrix}$.

Jeśli zaś $\sigma \in S_7$, to $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 4 & 3 & 6 & 5 & 2 & 7 \end{pmatrix}$.

Definicja 1.3.12. Niech $k, l \in \mathbb{N}$. Dwa cykle $\big(a_1, a_2, \ldots, a_k\big)$, $\big(b_1, b_2, \ldots, b_l\big)$ nazywamy *rozłącznymi*, jeśli zbiory $\big\{a_1, a_2, \ldots, a_k\big\}$, $\big\{b_1, b_2, \ldots, b_l\big\}$ są rozłączne, to jest

$$\big\{a_1, a_2, \ldots, a_k\big\} \;\cap\; \big\{b_1, b_2, \ldots, b_l\big\} = \varnothing.$$

Twierdzenie 1.3.13. i) Składanie cykli rozłącznych jest przemienne.

ii) Jesli $\sigma = \Big(a_1, a_2, \ldots, a_k\Big)$ jest cyklem długości k to wówczas

$$\sigma^{-1} = \Big(a_k, a_{k-1}, \ldots, a_2, a_1\Big),$$

$$\sigma^k = \mathrm{id} \quad \text{oraz} \quad \sigma^r \neq \mathrm{id} \quad \text{dla} \quad 1 < r < k.$$

Ponadto dla dowolnego $m \in \mathbb{Z}$

$$\sigma^m = \sigma^{m \mod k}.$$

iii) Każda permutacja jest cyklem lub może być przedstawiona jako złożenie (iloczyn) rozłącznych cykli. Przedstawienie to jest jednoznaczne z dokładnością do porządku czynników.

Przykład 1.3.14. Niech $\sigma = (1,2,3,4,5,6) \in S_6$. Obliczymy kolejne potęgi tego cyklu.

$$\sigma^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 1 & 2 \end{pmatrix} = (1,3,5)(2,4,6)$$

$$\sigma^3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 1 & 2 & 3 \end{pmatrix} = (1,4)(2,5)(3,6)$$

$$\sigma^4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 1 & 2 & 3 & 4 \end{pmatrix} = (1,5,3)(2,6,4)$$

$$\sigma^5 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 1 & 2 & 3 & 4 & 5 \end{pmatrix} = (1,6,5,4,3,2)$$

$$\sigma^6 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} = \mathrm{id}$$

Zauważmy, że złożenie cykli nie musi być cyklem. [3]

[3] Można uzasadnić, że jeśli σ jest cyklem długości k, to wówczas σ^l będzie cyklem długości k wtedy i tylko wtedy, gdy k i l są względnie pierwsze.

Uwaga 1.3.15. Niech $\sigma \in S_n$. Zbiór [a]

$$\operatorname{supp}(\sigma) = \{a \in \{1, 2, \ldots, n\} : \ \sigma(a) \neq a\}$$

nazywamy *nośnikiem* permutacji. Mówimy, że permutacje $\sigma, \tau \in S_n$ są *rozłączne* lub *niezależne*, gdy [b]

$$\operatorname{supp}(\sigma) \cap \operatorname{supp}(\tau) = \varnothing.$$

Można uzasadnić, że składanie permutacji rozłącznych jest przemienne, czyli jeśli $\sigma, \tau \in S_n$ są rozłączne, to wówczas $\sigma\tau = \tau\sigma$.

Cykle rozłączne są permutacjami rozłącznymi. Można stąd wywnioskować, że jeśli $\sigma = \sigma_1\sigma_2 \ldots \sigma_r$ jest rozkładem permutacji σ na parami rozłączne cykle, to wówczas dla dowolnej liczby naturalnej n zachodzi

$$\sigma^n = \sigma_1^n \sigma_2^n \ldots \sigma_r^n.$$

[a] angielskie *support* - nośnik

[b] Dwie permutacje σ, τ nazywamy rozłącznymi, jeśli zbiór elementów przesuniętych przez σ jest rozłączny od zbioru elementów przesuniętych przez τ.

Przykład 1.3.16. Rozważmy permutację $\sigma \in S_{10}$ postaci $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 6 & 2 & 4 & 3 & 5 & 9 & 8 & 10 & 1 & 7 \end{pmatrix}$.
Dokonamy jej rozkładu na cykle.

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 6 & 2 & 4 & 3 & 5 & 9 & 8 & 10 & 1 & 7 \end{pmatrix} = (1,6,9) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 2 & 4 & 3 & 5 & 6 & 8 & 10 & 9 & 7 \end{pmatrix} =$$

$$= (1,6,9)(3,4) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 2 & 3 & 4 & 5 & 6 & 8 & 10 & 9 & 7 \end{pmatrix} =$$

$$= (1,6,9)(3,4)(7,8,10) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{pmatrix} = (1,6,9)(3,4)(7,8,10)$$

Możemy wykorzystać otrzymany rozkład na parami rozłączne cykle, aby wyznaczyć σ^{15} oraz σ^{-15}. Mnożenie cykli rozłącznych jest przemienne, zatem

$$\sigma^{15} = (1,6,9)^{15}(3,4)^{15}(7,8,10)^{15} = (3,4)^{15} = (3,4),$$

bowiem $(1,6,9)^3 = \mathrm{id}$, $(3,4)^2 = \mathrm{id}$, $(7,8,10)^3 = \mathrm{id}$.
Zatem

$$\sigma^{-15} = (3,4)^{-1} = (4,3).$$

Definicja 1.3.17. i) Cykl dwuwyrazowy nazywamy *transpozycją*.

ii) Transpozycję postaci $(k, k+1)$ nazywamy *transpozycją liczb sąsiednich*.

Twierdzenie 1.3.18. i) Każdą permutację $\sigma \in S_n$, dla $n > 1$ można przedstawić jako złożenie pewnej liczby transpozycji.

ii) Każda transpozycja jest złożeniem nieparzystej liczby transpozycji liczb sąsiednich.

Dowód. i) Każda permutacja rozkłada się na iloczyn rozłącznych cykli. Każdy cykl $\big(a_1, a_2, \ldots, a_k\big)$ rozkłada się na iloczyn transpozycji w sposób następujący [4]

$$\big(a_1, a_2, \ldots, a_k\big) = (a_1, a_k)(a_1, a_{k-1}) \ldots (a_1, a_3)(a_1, a_2).$$

ii) Niech $i < j$. Dla transpozycji (i, j) otrzymujemy następujący rozkład

$$(i, j) = (i, i+1)(i+1, i+2) \ldots (j-2, j-1)(j-1, j)(j-2, j-1) \ldots (i+1, i+2)(i, i+1).$$

W rozkładzie tym występuje $2(j-1) - 1$ czynników. □

Uwaga 1.3.19. Rozkład permutacji na iloczyn transpozycji nie jest jednoznaczny.

Przykład 1.3.20. Rozważmy permutację $\sigma \in S_5$ postaci $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$.
Dokonamy rozkładu σ na iloczyn transpozycji. Otrzymujemy

$$\sigma = \big(1, 2, 3, 4, 5\big) = (1, 5)(1, 4)(1, 3)(1, 2)$$

Dla dowolnych $j, k \in \{1, 2, 3, 4, 5\}$ mamy stąd

$$\sigma = \big(1, 2, 3, 4, 5\big) = (1, 5)(2, 3)(2, 3)(1, 4)(1, 3)(1, 2) = (1, 5)(1, 4)(1, 3)(3, 5)(2, 4)(2, 4)(3, 5)(1, 2).$$

Definicja 1.3.21. i) Niech $\sigma \in S_n$. Mówimy, że liczby $j, k \in \{1, 2, \ldots, n\}$ tworzą *inwersję* [a] w permutacji σ, jeśli $j < k$ oraz $\sigma(j) > \sigma(k)$.

ii) Permutację σ nazywamy *parzystą*, jeśli $\sigma \in S_1$ lub przy $n > 1$ liczba inwersji w permutacji σ jest parzysta.

iii) Permutację σ nazywamy *nieparzystą*, jeśli liczba inwersji w permutacji σ jest nieparzysta.

iv) Liczbę $(-1)^l$, gdzie l to liczba inwersji permutacji σ, nazywamy *znakiem permutacji* i oznaczamy symbolem $\operatorname{sgn}(\sigma)$.

[a] Niektórzy autorzy używają określenia *nieporządek*.

Liczba inwersji dla permutacji identycznościowej to 0.
Zatem $\boxed{\operatorname{sgn}(\sigma) \in \{-1, 1\}}$. Permutacja σ jest parzysta wtedy i tylko wtedy gdy $\operatorname{sgn}(\sigma) = 1$. Podobnie, permutacja σ jest nieparzysta wtedy i tylko wtedy gdy $\operatorname{sgn}(\sigma) = -1$.

[4] Wzór ten można uzasadnić indukcyjnie.

Przykład 1.3.22. Rozważmy permutację $\sigma \in S_6$ postaci $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 5 & 6 & 2 \end{pmatrix}$.
Wygodny jest następujący zapis, służący zliczaniu inwersji.

						liczba inwersji
3	4	1	5	6	2	2
	4	1	5	6	2	2
		1	5	6	2	0
			5	6	2	1
				6	2	1
						6

Mamy sześć inwersji, zatem $\operatorname{sgn}(\sigma) = (-1)^6 = 1$ i permutacja σ jest parzysta.

Związek parzystości permutacji z rozkładem na iloczyn transpozycji

Twierdzenie 1.3.23. i) Permutacje σ oraz $\sigma(j,\, k)$ różnią się znakiem. Podobnie permutacje σ oraz $(j,\, k)\sigma$ różnią się znakiem.

ii) Iloczyn k transpozycji jest permutacją, której znak jest równy $(-1)^k$.

iii) Przy dowolnym rozkładzie permutacji σ na iloczyn transpozycji parzystość liczby czynników jest zawsze taka sama i zgodna z parzystością permutacji σ.

Wniosek 1.3.24. i) Permutacja σ jest parzysta (nieparzysta) wtedy i tylko wtedy gdy jest iloczynem parzystej (nieparzystej) liczby transpozycji.

ii) Dla dowolnego $n \in \mathbb{N}$, dowolnych $\sigma, \tau \in S_n$

$$\operatorname{sgn}(\tau \circ \sigma) = \operatorname{sgn}(\tau) \cdot \operatorname{sgn}(\sigma).$$

Przykład 1.3.25. Rozważmy permutację $\sigma \in S_6$ postaci $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 5 & 6 & 2 \end{pmatrix}$.
Otrzymujemy następujący rozkład na cztery transpozycje.

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 5 & 6 & 2 \end{pmatrix} = (1,3) \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 3 & 5 & 6 & 2 \end{pmatrix} = (1,3)(2,4,5,6) = (1,3)(2,6)(2,5)(2,4)$$

Permutacja σ jest parzysta.

1.3.3 Pierścienie i ciała

Definicja 1.3.26. Zespół $(A, \circ, *)$ złożony z niepustego zbioru A i określonych w nim działań wewnętrznych $\circ : A \times A \to A$, $* : A \times A \to A$ nazywamy *pierścieniem*, jeśli $(A, \circ)$ jest grupą abelową, zaś działanie $*$ jest łączne oraz rozdzielne względem działania $\circ$, tzn.

$$\forall a, b, c \in A \quad (a \circ b) * c = (a * c) \circ (b * c) \ \wedge \ c * (a \circ b) = (c * a) \circ (c * b).$$

Pierścień, w którym działanie $*$ posiada element neutralny, nazywamy *pierścieniem z jedynką* lub *z jednością*. Pierścień, w którym działanie $*$ jest przemienne, nazywamy *pierścieniem przemiennym* lub *komutatywnym*.

	Notacja addytywna		Notacja multiplikatywna
$\circ \,/\, +$	dodawanie	$* \,/\, \cdot$	mnożenie
$a + b$	suma	$a \cdot b$	iloczyn
$e = 0$	zero	$e = 1$	jedynka
$a' = -a$	element przeciwny	$a' = a^{-1}$	element odwrotny

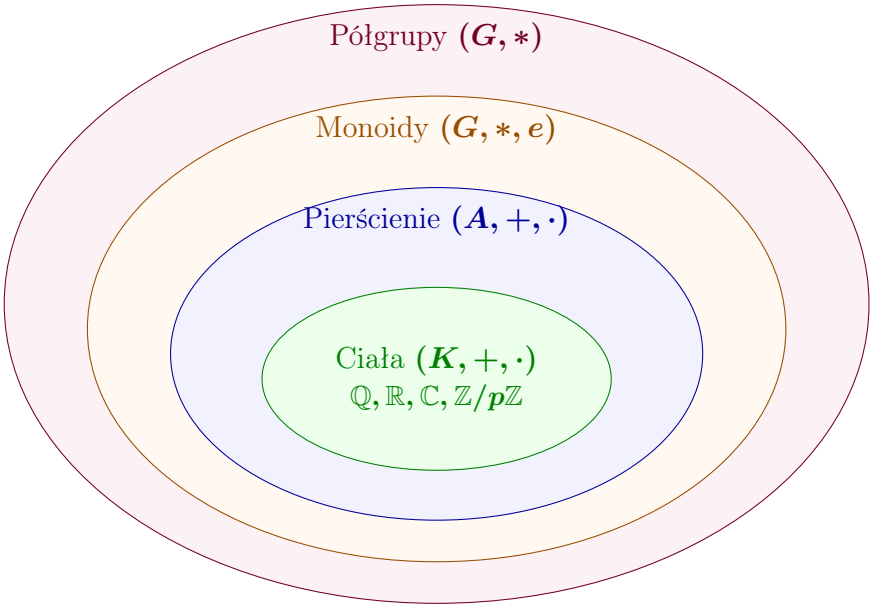
	$na = \underbrace{a + \ldots + a}_{n}, \ n \in \mathbb{N}$	$a^n = \underbrace{a \cdot \ldots \cdot a}_{n}$
	$0 \cdot a = 0$	$a^0 = e = e = 1$
$m \in \mathbb{Z}, m < 0$	$ma = \underbrace{(-a) + \ldots + (-a)}_{m}$	$a^m = \underbrace{a^{-1} \cdot \ldots \cdot a^{-1}}_{n}$

Używając powyższej notacji, piszemy po prostu $(A, +, \cdot)$.

Definicja 1.3.27. Niech K będzie zbiorem zawierającym co najmniej dwa elementy. Pierścień przemienny z jedynką $(K, \circ, *)$ nazywamy *ciałem*, jeśli [a]

$$\forall a \in K \setminus \{0\} \ \ \exists a' \in K \ \ a * a' = 1.$$

[a] Ciało to pierścień przemienny z jedynką (różną od zera, tj. $1 = e_* \neq e_\circ = 0$), w którym wszystkie niezerowe (tj. różne od elementu neutralnego $e_\circ$) elementy są odwracalne. Każde ciało zawiera zatem co najmniej dwa elementy, mianowicie 0 i 1. Zbiór elementów niezerowych jest niepusty i tworzy grupę z mnożeniem.



Hierarchia struktur algebraicznych.

Uwaga 1.3.28.

i) $(K, \circ, *)$ jest ciałem wtedy i tylko wtedy, gdy $(K, \circ)$ jest grupą abelową (z elementem neutralnym $e_\circ$ nazywanym *zerem* i oznaczanym jako 0) oraz $(K \setminus \{0\}, *)$ jest grupą abelową i ponadto działanie $*$ jest rozdzielne względem działania $\circ$. [a] [b]

ii) Jeśli $(K, \circ, *)$ jest ciałem, to wówczas grupę $(K, \circ)$ nazywamy *grupą addytywną* ciała K, zaś grupę $(K \setminus \{0\}, *)$ nazywamy *grupą multiplikatywną* ciała K.

iii) W ciele można zdefiniować operację *dzielenia* w sposób następujący

$$\frac{a}{b} := a * b^{-1}, \quad a, b \in K,\ b \neq 0.$$

[a] Warunek rozdzielności mnożenia względem dodawania pozwala rozróżniać działania mnożenia i dodawania. Stąd zamiast pisać $(a \cdot b) + c$ piszemy $a \cdot b + c$. Oznacza to, że mnożenie wiąże argumenty silniej niż dodawanie.

[b] Implikacja z lewa na prawo jest oczywista. By dowieść implikacji w przeciwnym kierunku wystarczy zaobserwować, że $a * 0 = 0 * a = 0$ dla dowolnego $a \in K$. Równość ta wynika z własności rozdzielności mnożenia względem dodawania, mianowicie $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$, skąd $a \cdot 0 = 0$. Analogicznie dowodzimy, że $0 \cdot a = 0$. Z przemienności i łączności działania $*$ w zbiorze $K \setminus \{0\}$ oraz z równości $a * 0 = 0 * a = 0$ wynika przemienność i łączność działania $*$ w zbiorze K.

Przykład 1.3.29. i) $(\mathbb{R}, +, \cdot)$ ciało liczb rzeczywistych
$(\mathbb{R}, +)$ grupa addytywna ciała
$(\mathbb{R}^*, \cdot)$ grupa multiplikatywna ciała

ii) $(\mathbb{Q}, +, \cdot)$ ciało liczb wymiernych

iii) $(\mathbb{Z}, +, \cdot)$ pierścień przemienny z jedynką, ale nie ciało (bowiem np. $2^{-1} \notin \mathbb{Z}$)

iv) $(\mathbb{R}[x], +, \cdot)$ zbiór wszystkich wielomianów jednej zmiennej x o współczynnikach rzeczywistych z działaniami dodawania wielomianów i mnożenia wielomianów ma strukturę pierścienia przemiennego jedynką.

v) $(\mathbb{R}(x), +, \cdot)$ zbiór wszystkich funkcji wymienych jednej zmiennej x o współczynnikach rzeczywistych z działaniami dodawania i mnożenia funkcji ma strukturę ciała.

vi) $(\mathbb{Z}/p\mathbb{Z}, +_p, \cdot_p)$, gdzie $p \in \mathbb{N}$, $p \geq 2$
$+_p, \cdot_p$ dodawanie i mnożenie modulo p

Jeśli p to liczba pierwsza, to $\mathbb{Z}/p\mathbb{Z}$ jest ciałem [5] (tzw. *ciało reszt modulo p*). Jeśli p nie jest liczbą pierwszą, to $\mathbb{Z}/p\mathbb{Z}$ jest pierścieniem, ale nie ciałem.

$\mathbb{Z}/5\mathbb{Z}$ to ciało

$+_5$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot_5$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Zauważmy, że w ciele $\mathbb{Z}/5\mathbb{Z}$ zachodzi równość $5 \cdot 1 = \underbrace{1 + \ldots + 1}_{5\,razy} = 0$.

$\mathbb{Z}/4\mathbb{Z}$ nie jest ciałem

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot_4$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Definicja 1.3.30. Niech $(A, +, \cdot)$ będzie pierścieniem. Elementy $a, b \in A, a \neq 0, b \neq 0$ nazywamy *dzielnikami zera*, jeśli $a \cdot b = 0$.

Uwaga 1.3.31. W ciele nie ma dzielników zera.

Dowód. Niech $(K, +, \cdot)$ będzie ciałem oraz niech $a, b \in K$. Załóżmy, że $a \cdot b = 0$ oraz $a \neq 0$. Jeśli $a \neq 0$, to istnieje $a^{-1} \in K$. Otrzymujemy

$$0 = a^{-1} \cdot 0 = a^{-1} \cdot (a \cdot b_{=}(a^{-1} \cdot a) \cdot b = 1 \cdot b = b.$$

Zatem $b = 0$. □

[5] Jest to przykład ciała o skończonej liczbie elementów. Tego typu ciała nazywamy *ciałami skończonymi* lub inaczej *ciałami Galois.*

Przykład 1.3.32. Sprawdź, że zbiór $\mathbb{R}^2$ wraz z działaniami dodawania i mnożenia zdefiniowanymi poniżej ma strukturę ciała.

$$\forall (x_1,y_1),(x_2,y_2)\in\mathbb{R}^2 \quad (x_1,y_1)+(x_2,y_2) := (x_1+x_2, y_1+y_2),\ (x_1,y_1)\cdot(x_2,y_2) := (x_1x_2-y_1y_2, x_1y_2+x_2y_1).$$

$(\mathbb{C},+)$ jest grupą abelową.

Działanie $+$ jest wewnętrzne, łączne, przemienne. Elementem neutralnym jest $(0,0)\in\mathbb{R}^2$, zaś elementem przeciwnym do $(x_1,y_1)\in\mathbb{R}^2$ jest $(-x_1,-y_1)\in\mathbb{R}^2$.

$(\mathbb{C}\setminus\{(0,0)\},\cdot)$ jest grupą abelową.

Działanie $\cdot$ jest wewnętrzne.

przemienność:

$$(x_2,y_2)\cdot(x_1,y_1) = (x_2x_1-y_2y_1, x_2y_1) = \ = (x_1x_2-y_1y_2, x_1y_2+x_2y_1) = (x_1,y_1)\cdot(x_2,y_2)$$

łączność: $\quad L = P$

$$L = (a,b)\cdot[(x_1,y_1)\cdot(x_2,y_2)] = (a,b)\cdot(x_1x_2-y_1y_2, x_1y_2+x_2y_1) =$$

$$= (ax_1x_2-ay_1y_2-bx_1y_2-bx_2y_1, ax_1y_2+ax_2y_1+bx_1x_2-by_1y_2)$$

$$P = [(a,b)\cdot(x_1,y_1)]\cdot(x_2,y_2) = (ax_1-by_1, ay_1+bx_1)\cdot(x_2,y_2) =$$

$$= (ax_1x_2-by_1x_2-ay_1y_2-bx_1y_2, ax_1y_1-by_1y_2+ay_1x_2+bx_1x_2)$$

el. neutralny: $\quad e = (1,0)$

$$\forall (x_1,y_1)\in\mathbb{R}^2 \quad (1,0)\cdot(x_1,y_1) = (1\cdot x_1-0\cdot y_1, 1\cdot y_1+0\cdot x_1) = (x_1,y_1)$$

el. odwrotny do $(x_1,y_1)\neq(0,0)$:

$$(x_1,y_1)\cdot(a,b) = (1,0) \Leftrightarrow (ax_1-by_1, ay_1+bx_1) = (1,0) \Leftrightarrow \begin{bmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{bmatrix}\cdot\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\forall (x_1,y_1)\neq(0,0) \quad a = \frac{x_1}{x_1^2+y_1^2},\ \ b = \frac{-y_1}{x_1^2+y_1^2}$$

Działanie $\cdot$ jest rozdzielne względem $+$, bowiem $L = P$.

$$L = (a,b)\cdot[(x_1,y_1)+(x_2,y_2)] = (a,b)\cdot(x_1+x_2, y_1+y_2) = (ax_1+ax_2-by_1-by_2, ay_1+ay_2+bx_1+bx_2)$$

$$P = [(a,b)\cdot(x_1,y_1)] + [(a,b)\cdot(x_2,y_2)] = (ax_1-by_1, ay_1+bx_1) + (ax_2-by_2, ay_2+bx_2)$$

Definicja 1.3.33. Zdefiniowane powyżej ciało nazywamy *ciałem liczb zespolonych* i oznaczamy symbolem $\mathbb{C}$. Elementy tego ciała nazywamy *liczbami zespolonymi.*

Q U I Z

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. Działanie $\circ$ takie, że $x \circ y = \frac{2x+y}{3}$ jest działaniem łącznym w $\mathbb{R}$.
2. Działanie $\circ$ określone w zbiorze $\mathbb{R}$ wzorem $\forall a, b \in \mathbb{R}\ \ a \circ b = a \cdot |b|$ posiada element neutralny.
3. Niech $*$ będzie dowolnym działaniem wewnętrznym w $\mathbb{R}$, posiadającym element neutralny. Wówczas, jeśli dla danego elementu $a \in \mathbb{R}$ istnieje element symetryczny, to jest on jedyny.
4. W zbiorze $\mathbb{R}$ określamy działania $x \circ y = x + y + 1$ oraz $x * y = xy + x + y$. Działanie $*$ jest rozdzielne względem działania $\circ$.
5. Struktura $(\mathbb{R}, \cdot)$, gdzie $\cdot$ jest mnożeniem liczb, to grupa abelowa.
6. Struktura $(\mathbb{R}[x], +, \cdot)$, gdzie $\cdot, +$ jest odpowiednio mnożeniem i dodawaniem wielomianów, jest ciałem.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Algebra
Wykład nr 2
Data: .....

TEMAT: *Liczby zespolone*

2.1 Ciało liczb zespolonych

Motywacja

$$\mathbb{N} \xhookrightarrow[n\mapsto n]{} \mathbb{Z} \xhookrightarrow[n\mapsto \frac{n}{1}]{} \mathbb{Q} \hookrightarrow \mathbb{R} \xhookrightarrow[x\mapsto (x,0)]{h} \mathbb{C}$$

$X^2 - 2 = 0$ równanie o współczynnikach z $\mathbb{Q}$, jego rozwiązania $\pm\sqrt{2} \notin \mathbb{Q}$
Ćwiczenie: $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ jest ciałem takim, że $\mathbb{Q} \hookrightarrow \mathbb{Q}(\sqrt{2}) \hookrightarrow \mathbb{R}$

$X^2 + 1 = 0$ równanie o współczynnikach z $\mathbb{R}$, jego rozwiązania $\pm i$ nie należą do $\mathbb{R}$
$\mathbb{R}(i) = \{a + bi : a, b \in \mathbb{R}\} = \mathbb{C}$
$\mathbb{C}$ ciało *algebraicznie domknięte* - tzn. rozwiązania równań algebraicznych (wielomianowych) o współczynnikach z $\mathbb{C}$ należą do $\mathbb{C}$

Zanurzenie $\mathbb{R}$ w $\mathbb{C}$

Niech $\Omega = \mathbb{R} \times \{0\} = \{(x, 0) : x \in \mathbb{R}\} \subset \mathbb{R}^2$. Wówczas $(\Omega, +, \cdot)$ jest ciałem.

wewnętrzność: $(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0) \in \Omega$, $(x_1, 0) \cdot (x_2, 0) = (x_1 x_2, 0) \in \Omega$

przemienność:
$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0) = (x_2 + x_1, 0) = (x_2, 0) + (x_1, 0)$$
$$(x_1, 0) \cdot (x_2, 0) = (x_1 x_2, 0) = (x_2 x_1, 0) = (x_2, 0) \cdot (x_1, 0)$$

łączność:
$$[(x_1, 0) + (x_2, 0)] + (x_3, 0) = (x_1 + x_2 + x_3, 0) = (x_1, 0) + [(x_2, 0) + (x_3, 0)]$$
$$[(x_1, 0) \cdot (x_2, 0)] \cdot x_3, 0 = (x_1 x_2 x_3, 0) = (x_1, 0) \cdot [(x_2, 0) \cdot (x_3, 0)]$$

el. neutralne: $(0, 0)$ dla dodawania oraz $(1, 0)$ dla mnożenia
el. symetryczne do $(x_1, 0)$: $(-x_1, 0)$ względem $+$, $(\frac{1}{x_1}, 0)$ względem $\cdot$

rozdzielność:
$$[(x_1, 0) + (x_2, 0)] \cdot (x_3, 0) = (x_1 + x_2, 0) \cdot (x_3, 0) = ((x_1 + x_2)x_3, 0) =$$
$$= (x_1 x + 3 + x_2 x_3, 0) = [(x_1, 0) \cdot (x_3, 0)] + [(x_2, 0) \cdot (x_3, 0)]$$

Niech $h : \mathbb{R} \to \Omega$, $h(x) = (x, 0)$. Jest to *zanurzenie*, czyli bijekcja taka, że
$h(x_1 + x_2) = h(x_1) + h(x_2)$ oraz $h(x_1 \cdot x_2) = h(x_1) \cdot h(x_2)$.
Utożsamiamy zbiory $\mathbb{R}$ oraz Ω i piszemy x zamiast $h(x)$.

Zdefiniujmy $i := (0,1)$ tzw. *jednostka urojona*. Wówczas

$$i^2 = i \cdot i = (0,1) \cdot (0,1) = (-1,0) = -1$$

$$\mathbb{C} \ni z = (x,y) = (x,0) + (0,y) = (x,0) + (0,1)(y,0) = x + iy$$

Postać $z = x + iy$, gdzie $x, y, \in \mathbb{R}$ to tzw. *postać kanoniczna (algebraiczna, Gaussa)* liczby zespolonej. Liczbę $x \in \mathbb{R}$ nazywamy *częścią rzeczywistą* liczby z i oznaczamy $\mathrm{Re}z$. Liczbę $y \in \mathbb{R}$ nazywamy *częścią urojoną* liczby z i oznaczamy $\mathrm{Im}z$. Liczby postaci iy, $y \in \mathbb{R}$ nazywamy *czysto urojonymi*.

$$z_1 = z_2 \Leftrightarrow \big(\mathrm{Re}z_1 = \mathrm{Re}z_1 \wedge \mathrm{Im}z_1 = \mathrm{Im}z_2\big)$$

Postać algebraiczna pozwala na dodawanie i mnożenie liczb zespolonych jak wielomianów zmiennej i, przy warunku $i^2 = -1$.

$$(x_1,y_1) + (x_2,y_2) = (x_1 + x_2, y_1 + y_2) \qquad (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$
$$(x_1,y_1) \cdot (x_2,y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1) \qquad (x_1 + iy_1) \cdot (x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

Przykład 2.1.1. $(2 + 7i) - (4 - 2i) = -2 + 9i$

$$(3 - i) \cdot (2 + 3i) = 6 + 9i - 2i - 3i^2 = 9 + 7i$$

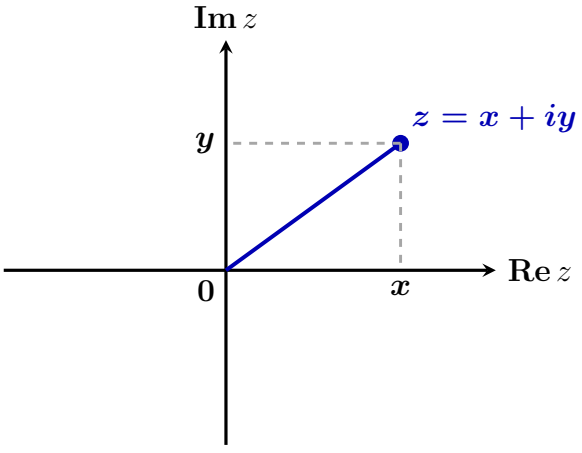
$$\frac{2+3i}{2-5i} = \frac{(2+3i)(2+5i)}{(2-5i)(2+5i)} = \frac{4+10i+6i+15i^2}{4-25i^2} = \frac{-11+16i}{29} = -\frac{11}{29} + \frac{16}{29}i$$

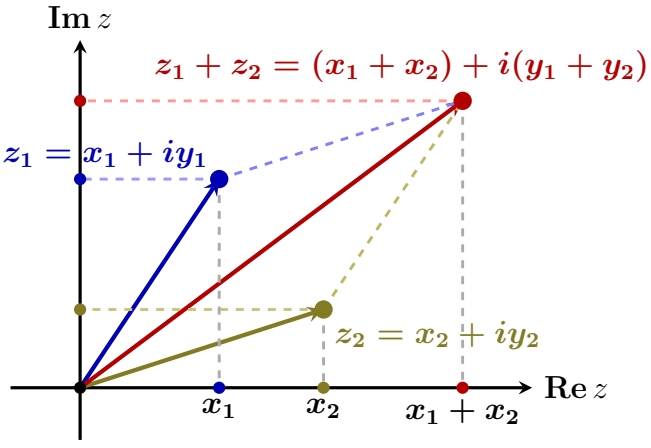
Uwaga 2.1.2. W ciele $\mathbb{C}$ nie można określić porządku liniowego.

$-1 = i \cdot i = (-i) \cdot (-i) \qquad 1 = 1 \cdot 1 = (-1) \cdot (-1)$

Utożsamiamy liczby zespolone z punktami na płaszczyźnie lub wektorami zaczepionymi w $(0,0)$.

Płaszczyzna zespolona - geometryczny model ciała liczb zespolonych $\mathbb{C}$

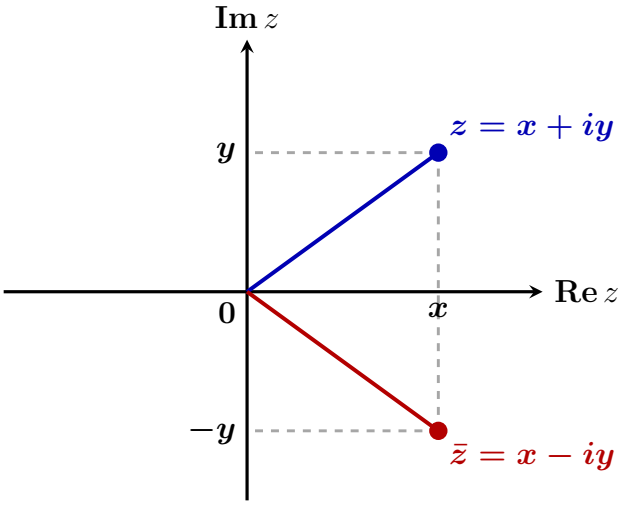




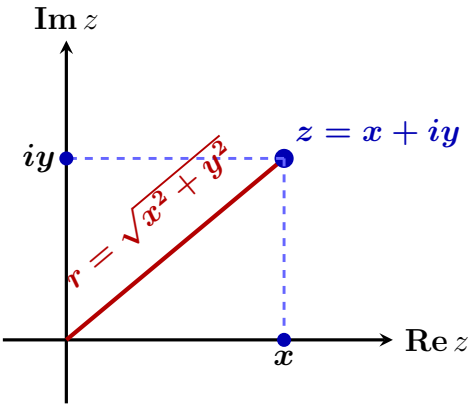
Definicja 2.1.3. Niech $z = x + iy \in \mathbb{C}$, $x, y \in \mathbb{R}$.

i) Liczbę zespoloną $w = x - iy$ nazywamy *liczbą sprzężoną* do liczby z. Oznaczamy ją $\bar{z}$.

i) Liczbę rzeczywistą $\sqrt{x^2 + y^2}$ nazywamy modułem liczby z. Oznaczamy ją $|z|$.



liczba sprzężona do z



moduł liczby zespolonej

Twierdzenie 2.1.4 (Własności liczb zespolonych)**.** Niech $z, z_1, z_2 \in \mathbb{C}$. Wówczas prawdziwe są następujące równości.

i) $\mathrm{Re}(z_1 + z_2) = \mathrm{Re}z_1 + \mathrm{Re}z_2, \quad \mathrm{Im}(z_1 + z_2) = \mathrm{Im}z_1 + \mathrm{Im}z_2$

ii) $\mathrm{Re}z = \frac{1}{2}(z + \bar{z}), \quad \mathrm{Im}z = \frac{1}{2i}(z - \bar{z})$

iii) $\overline{(\bar{z})} = z$

iv) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2, \quad \overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$

iv) $\overline{(\frac{z_1}{z_2})} = \frac{\bar{z}_1}{\bar{z}_2}$, dla $z_2 \neq 0$

vi) $|z| = |-z| = |\bar{z}|$

vii) $z \cdot \bar{z} = |z|^2$

viii) $\mathrm{Re}z \leq |\mathrm{Re}z| \leq |z|, \quad \mathrm{Im}z \leq |\mathrm{Im}z| \leq |z|$

ix) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$

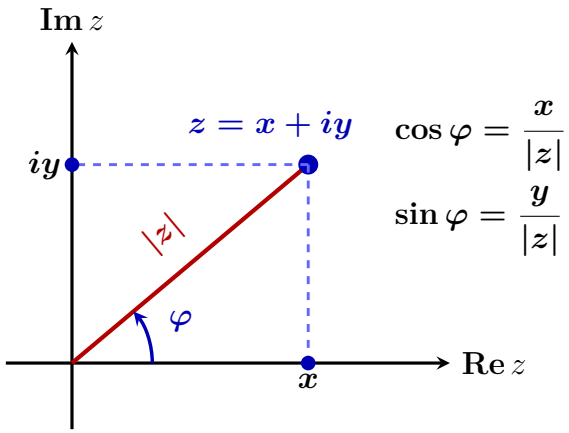
x) $|z_1 + z_2| \leq |z_1| + |z_2|$ (nierówność trójkąta) $\quad \big||z_1| - |z_2|\big| \leq |z_1 - z_2|$

2.2 Postać trygonometryczna i wykładnicza

Postać trygonometryczna liczby zespolonej

Niech $z = x + iy \neq 0$. Wówczas $z = |z|\left(\frac{x}{|z|} + i\frac{y}{|z|}\right)$. Ponieważ $\left(\frac{x}{|z|}\right)^2 + \left(\frac{y}{|z|}\right)^2 = 1$, więc istnieje kąt φ taki, że

$$z = |z|(\cos\varphi + i\sin\varphi).$$



argument liczby zespolonej

Dowolny taki kąt nazywamy *argumentem* liczby z.

KONWENCJA:
Ten z argumentów liczby zespolonej, **który leży w przedziale** $[\mathbf{0}, \mathbf{2\pi})$, nazywamy *argumentem głównym* liczby z i oznaczamy $\arg z$.

Zatem dowolny argument liczby z ma postać $\arg z + 2k\pi$, $k \in \mathbb{Z}$.

Przyjmujemy, że argument liczby $z = 0$ jest **nieokreślony**.

WNIOSEK: Dowolną liczbę $z \in \mathbb{C}, z \neq 0$ możemy zatem przedstawić w postaci

$$z = |z|(\cos\varphi + i\sin\varphi),$$

gdzie φ to jeden z jej argumentów. Powyższe przedstawienie nazywamy *postacią trygonometryczną* liczby zespolonej z.

Gdy $z_1 \neq 0, z_2 \neq 0$, $z_1 = |z_1|(\cos\alpha + i\sin\alpha)$, $z_2 = |z_2|(\cos\beta + i\sin\beta)$, to wówczas $z_1 = z_2$ wtedy i tylko wtedy gdy $|z_1| = |z_2|$ oraz $\beta = \alpha + 2k\pi$ dla pewnego $k \in \mathbb{Z}$.

Przykład 2.2.1. Przedstaw podane liczby w postaci trygonometrycznej.

$z = 7$	$z = -i$	$z = -\sqrt{27} - 3i$
$z = 7(1 + 0i)$	$\lvert z\rvert = \sqrt{0^2 + (-1)^2} = 1$	$\lvert z\rvert = \sqrt{27+9} = 6$
	$z = 1\,(0 + (-1)\cdot i)$	$z = 6\left(-\frac{\sqrt{3}}{2} - \frac{1}{2}i\right)$
	$\begin{cases} \cos\varphi = 0 \\ \sin\varphi = -1 \end{cases}$	$\begin{cases} \cos\varphi = -\frac{\sqrt{3}}{2} \\ \sin\varphi = -\frac{1}{2} \end{cases}$
$\varphi = 0 + 2k\pi$	$\varphi = -\frac{\pi}{2} + 2k\pi$	$\varphi = \frac{7}{6}\pi + 2k\pi, \qquad k \in \mathbb{Z}$
$\arg z = 0$	$\arg z = \frac{3}{2}\pi$	$\arg z = \frac{7}{6}\pi$
$z = 7(\cos 0 + i\sin 0)$	$-i = 1(\cos\frac{3}{2}\pi + i\sin\frac{3}{2}\pi)$	$-\sqrt{27} - 3i = 6(\cos\frac{7}{6}\pi + i\sin\frac{7}{6}\pi)$

Mnożenie i dzielenie liczb w postaci trygonometrycznej

Twierdzenie 2.2.2. Niech $z = |z|(\cos\varphi + i\sin\varphi)$,
$z_1 = |z_1|(\cos\alpha + i\sin\alpha)$, $z_2 = |z_2|(\cos\beta + i\sin\beta)$. Wówczas

i) $z_1 \cdot z_2 = |z_1| \cdot |z_2| \cdot \Big(\cos(\alpha+\beta) + i\sin(\alpha+\beta)\Big)$, [a]

ii) $\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|} \cdot \Big(\cos(\alpha-\beta) + i\sin(\alpha-\beta)\Big)$,

iii) $z^n = |z|^n \cdot \Big(\cos n\varphi + i\sin n\varphi\Big)$ tzw. *wzór de Moivre'a*

[a] Mnożeniu liczb zespolonych można przypisać następującą interpretacje geometryczną, prowadzi ono do zmiany długości wektora i jego obrotu. Zobacz także tutaj.

Dowód. i) $z_1 \cdot z_2 = |z_1| \cdot |z_2|(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta)$
$= |z_1| \cdot |z_2|\Big(\underbrace{(\cos\alpha\cos\beta - \sin\alpha\sin\beta)}_{\cos(\alpha+\beta)} + i\underbrace{(\sin\alpha\cos\beta + \cos\alpha\sin\beta)}_{\sin(\alpha+\beta)}\Big)$
ii) analogicznie
iii) Na mocy i) dla $n = 2$ mamy $z^2 = z \cdot z = |z|^2(\cos 2\varphi + i\sin 2\varphi)$.
Przeprowadzając dowód indukcyjny, otrzymujemy tezę. $\square$

Przykład 2.2.3. Przedstaw liczbę $z = -\sin\frac{\pi}{10} + i\cos\frac{\pi}{10}$ w postaci trygonometrycznej.

METODA I: Ponieważ $|z| = 1$, zatem $z = \cos\varphi + i\sin\varphi$ dla pewnego $\varphi \in [0, 2\pi)$.
$\begin{cases} \cos\varphi = -\sin\frac{\pi}{10} < 0 \\ \sin\varphi = \cos\frac{\pi}{10} > 0 \end{cases} \quad \Rightarrow \quad \varphi$ z II ćwiartki
Na mocy wzorów redukcyjnych $\varphi = \frac{\pi}{2} + \frac{\pi}{10}$ oraz $z = \cos\left(\frac{\pi}{2} + \frac{\pi}{10}\right) + i\sin\left(\frac{\pi}{2} + \frac{\pi}{10}\right)$.

METODA II: $z = -\sin\frac{\pi}{10} + i\cos\frac{\pi}{10} = i^2\sin\frac{\pi}{10} + i\cos\frac{\pi}{10} = i\left(\cos\frac{\pi}{10} + i\sin\frac{\pi}{10}\right) =$
$\left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right)\left(\cos\frac{\pi}{10} + i\sin\frac{\pi}{10}\right) = \cos\left(\frac{\pi}{2} + \frac{\pi}{10}\right) + i\sin\left(\frac{\pi}{2} + \frac{\pi}{10}\right)$.

Przykład 2.2.4. Oblicz $\left(\frac{1+i\sqrt{3}}{1-i}\right)^{20}$.

$1+i\sqrt{3} = 2\left(\frac{1}{2}+\frac{\sqrt{3}}{2}i\right) = 2(\cos\frac{\pi}{3}+i\sin\frac{\pi}{3}) \quad 1-i = \sqrt{2}\left(\frac{\sqrt{2}}{2}-\frac{\sqrt{2}}{2}i\right) = \sqrt{2}\left(\cos(-\frac{\pi}{4})+i\sin(-\frac{\pi}{4})\right)$

$\frac{1+i\sqrt{3}}{1-i} = \frac{2}{\sqrt{2}}\Big(\cos(\frac{\pi}{3}+\frac{\pi}{4})+i\sin(\frac{\pi}{3}+\frac{\pi}{4})\Big) = \sqrt{2}\left(\cos(\frac{7}{12}\pi)+i\sin(\frac{7}{12}\pi)\right)$

$\left(\frac{1+i\sqrt{3}}{1-i}\right)^{20} = 2^{10}\left(\cos(\frac{140}{12}\pi)+i\sin(\frac{140}{12}\pi)\right) = 2^{10}\left(\cos(\frac{35}{3}\pi)+i\sin(\frac{35}{3}\pi)\right) =$

$= 2^{10}\left(\cos(12\pi-\frac{\pi}{3})+i\sin(12\pi-\frac{\pi}{3})\right) = 2^{10}\left(\cos\frac{\pi}{3}-i\sin\frac{\pi}{3}\right) = 2^{10}\left(\frac{1}{2}-\frac{\sqrt{3}}{2}i\right) =$

$= 2^9(1-\sqrt{3}i)$

Twierdzenie 2.2.5 (Własności argumentu)**.** Niech $z, z_1, z_2 \in \mathbb{C}$.
Wówczas prawdziwe są następujące równości.

i) $\arg(z_1z_2) = \arg z_1 + \arg z_2 + 2k\pi,\ k\in\mathbb{Z} \qquad (k=0 \text{ lub } k=-1)$

ii) $\arg(\frac{z_1}{z_2}) = \arg z_1 - \arg z_2 + 2k\pi,\ k\in\mathbb{Z} \qquad (k=0 \text{ lub } k=-1)$

iii) $\arg(z^n) = n\cdot\arg z + 2k\pi,\ k\in\mathbb{Z}$

iv) $\arg\bar{z} = 2\pi - \arg z$, gdy $\arg z \neq 0$

Przykład 2.2.6. i) $\arg i = \frac{\pi}{2} \qquad \arg(-1)=\pi \qquad \arg(-i)=\frac{3}{2}\pi$
$i = (-1)\cdot(-i) \ \Rightarrow\ \arg i = \pi + \frac{3}{2}\pi + 2k\pi = \frac{5}{2}\pi + 2k\pi$ dla pewnego $k\in\mathbb{Z}$, tj. $k=-1$

ii) Niech $z = \frac{\sqrt{12}-2i}{-\sin\frac{\pi}{5}+i\cos\frac{\pi}{5}}$. Wyznaczymy argument główny $\arg z$ liczby z.

$$\sqrt{12}-2i = 4\left(\frac{\sqrt{3}}{2}-\frac{1}{2}i\right) = 4\left(\cos\frac{\pi}{6}-i\sin\frac{\pi}{6}\right) = 4\left[\cos\left(-\frac{\pi}{6}\right)+i\sin\left(-\frac{\pi}{6}\right)\right]$$

$$-\sin\frac{\pi}{5}+i\cos\frac{\pi}{5} = \cos\left(\frac{\pi}{2}+\frac{\pi}{5}\right)+i\sin\left(\frac{\pi}{2}+\frac{\pi}{5}\right) = 1\left(\cos\frac{7}{10}\pi+i\sin\frac{7}{10}\pi\right)$$

$$z = 4\left[\cos\left(-\frac{1}{6}-\frac{7}{10}\right)\pi+i\sin\left(-\frac{1}{6}-\frac{7}{10}\right)\pi\right] = 4\left[\cos\left(-\frac{13}{15}\pi\right)+i\sin\left(-\frac{13}{15}\pi\right)\right]$$

Jednym z argumentów liczby z jest $\varphi = -\frac{13}{15}\pi$. Stąd dla pewnego $k\in\mathbb{Z}$ mamy

$$\arg z + 2k\pi = -\frac{13}{15}\pi.$$

Biorąc $k=-1$, otrzymujemy $\arg z = \frac{17}{15}\pi \in [0,2\pi)$.

Postać wykładnicza liczby zespolonej

Dla $\varphi\in\mathbb{R}$ definiujemy $e^{i\varphi} = \cos\varphi + i\sin\varphi$. Zatem dowolną liczbę zespoloną $z\neq 0$ można zapisać w postaci $z = |z|e^{i\varphi}$, gdzie φ to pewien argument liczby z.

Przykład 2.2.7. a) $e^{\frac{\pi}{2}i} = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = 0 + 1\cdot i = i$

b) $e^{i\pi} = \cos\pi + i\sin\pi = -1 + 0\cdot i \;\Rightarrow\; \boxed{e^{i\pi} + 1 = 0}$

najpiękniejszy wzór w matematyce

Twierdzenie 2.2.8. Niech $\alpha, \beta \in \mathbb{R}$. Wówczas:

i) $e^{i(\alpha+\beta)} = e^{i\alpha}\cdot e^{i\beta}$, $e^{i(\alpha-\beta)} = \frac{e^{i\alpha}}{e^{i\beta}}$,

ii) $(e^{i\alpha})^k = e^{ik\alpha}$, dla $k \in \mathbb{Z}$,

iii) $e^{i(\alpha+2k\pi)} = e^{i\alpha}$, dla $k \in \mathbb{Z}$,

iv) $e^{i\alpha} \neq 0$, $|e^{i\alpha}| = 1$.

Dowód. iv) Mamy $|e^{i\alpha}| = |\cos\alpha + i\sin\alpha| = \sqrt{\cos^2\alpha + \sin^2\alpha} = 1$, zatem $e^{i\alpha} = 0 \Leftrightarrow \cos\alpha = \sin\alpha = 0$, co nie jest możliwe, gdyż gdy $\cos\alpha = 0$, to $\sin\alpha = \pm 1$. $\square$

Wniosek 2.2.9. Niech $z = re^{i\varphi}$, $z_1 = r_1e^{i\alpha}$, $z_2 = r_2e^{i\beta}$ będą liczbami zespolonymi w postaci wykładniczej. Wówczas:

i) $z_1z_2 = r_1r_2e^{i(\alpha+\beta)}$, $\frac{z_1}{z_2} = \frac{r_1}{r_2}e^{i(\alpha-\beta)}$,

ii) $z^k = r^ke^{ik\varphi}$, dla $k \in \mathbb{Z}$,

iii) $\bar{z} = re^{-i\varphi}$.

Dowód. iii) Jeśli $\arg z = \varphi$, to $\arg\bar{z} = 2\pi - \varphi$, skąd $e^{(2\pi i-\varphi)i} = e^{2\pi}e^{-\varphi i} = e^{-\varphi i}$. $\square$

Wzory Eulera

$e^{i\varphi} = \cos\varphi + i\sin\varphi$

$e^{-i\varphi} = \cos(-\varphi) + i\sin(-\varphi) = \cos\varphi - i\sin\varphi$

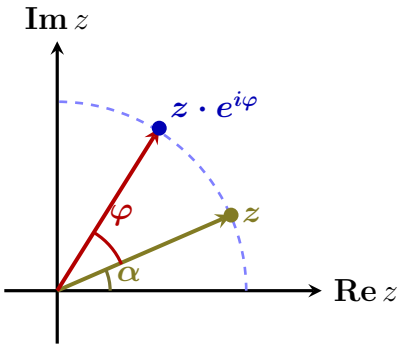
Dodając lub odejmując stronami, otrzymujemy :

$$\cos\varphi = \frac{e^{i\varphi} + e^{-i\varphi}}{2}, \quad \sin\varphi = \frac{e^{i\varphi} - e^{-i\varphi}}{2i}, \quad \varphi \in \mathbb{R}.$$

Obrót o kąt φ

$z = re^{i\alpha}$

$z\cdot e^{i\varphi} = re^{i(\alpha+\varphi)}$



obrót o kąt φ
w kierunku przeciwnym
do ruchu wskazówek zegara

Przykład 2.2.10. Rozwiążemy równanie $z^3 = (\bar{z})^3$.

Widzimy, że $z = 0$ spełnia równanie.

Załóżmy teraz, że $z \neq 0$. Wówczas $z = re^{i\varphi}$, dla pewnych $r \in \mathbb{R}, r > 0$

oraz $\varphi \in \mathbb{R}$. Zatem

$$\begin{aligned} z^3 &= (\overline{z})^3 \\ r^3 e^{3i\varphi} &= r^3 e^{-3i\varphi} \\ e^{6i\varphi} &= 1 = e^0 \\ 6\varphi &= 0 + 2k\pi, \quad k \in \mathbb{Z} \\ \varphi &= \tfrac{k\pi}{3}, \quad k \in \mathbb{Z} \end{aligned}$$

Rozwiązaniami są punkt $z = 0$ oraz półproste $\varphi_k = \frac{k\pi}{3}, \quad k \in \{0, 1, 2, 3, 4, 5\}$.
Podsumowując, rozwiązaniami są punkty należące do prostych $y = 0$, $y = \sqrt{3}x$ oraz $y = -\sqrt{3}x$.

2.3 Pierwiastkowanie, równania wielomianowe

Pierwiastkowanie
Niech $n \in \mathbb{N}$, $w \in \mathbb{C}$ będą ustalone.

Definicja 2.3.1. Każdą liczbę $z \in \mathbb{C}$ spełniającą równanie $z^n = w$, nazywamy *pierwiastkiem n-tego stopnia* z liczby w.

Przykład 2.3.2. Rozwiąż równanie $z^2 = 8 + 6i$.

I sposób:
postać algebraiczna

$z = x + iy$, $w = 8 + 6i$
$z^2 = w$
$(x + iy)^2 = 8 + 6i$
$x^2 + 2xyi - y^2 = 8 + 6i$
$\begin{cases} x^2 - y^2 = 8 \\ 2xy = 6 \end{cases}$
$y \neq 0$ (gdy $y = 0$, to $z = x$, $x^2 \neq 8 + 6i$)
$x = \frac{3}{y}$, $\frac{9}{y^2} - y^2 = 8$
$y^4 + 8y^2 - 9 = 0$, $\Delta = 100$
$y^2 = 1$, $y = \pm 1$

$z = 3 + i \ \vee \ z = -3 - i$

II sposób:
postać wykładnicza

$z = |z|e^{i\varphi}$, $w = 8 + 6i$
$|z|^2 e^{2i\varphi} = w = |w|e^{i\alpha}$

$|w| = 10$

$w = 8 + 6i = 10(\frac{4}{5} + i\frac{3}{5})$
$\alpha = \arcsin \frac{3}{5}$
$z^2 = |z|^2 e^{2i\varphi} = 10e^{i\alpha}$
$2\varphi = \alpha + 2k\pi$, $|z| = \sqrt{10}$
$\varphi_k = \frac{1}{2} \arcsin \frac{3}{5} + k\pi$

$z = \sqrt{10} e^{i\varphi_k}, \ k \in \{0, 1\}$

III sposób:
postać algebraiczna

$8 + 6i = 9 + 6i - 1 =$
$= 3^2 + 2 \cdot 3 \cdot i + i^2 =$
$= (3 + i)^2$

$z = 3 + i \ \vee \ z = -3 - i$

Twierdzenie 2.3.3. Jeśli $n \in \mathbb{N}$, $w \in \mathbb{C} \setminus \{0\}$, $w = r(\cos\varphi + i\sin\varphi)$, to wówczas równanie $z^n = w$ posiada n różnych rozwiązań. Rozwiązania te mają postać

$$z_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}\right), \quad k = 0, 1, \ldots, n-1.$$

Dowód. Niech $z = \rho(\cos\alpha + i\sin\alpha)$, wówczas

$z^n = w \;\Leftrightarrow\; \rho^n(\cos n\alpha + i\sin n\alpha) = r(\cos\varphi + i\sin\varphi) \;\Leftrightarrow\; \begin{cases} \rho^n = r \\ n\alpha = \varphi + 2k\pi,\ k \in \mathbb{Z} \end{cases}$. Otrzymujemy

$z_k = \sqrt[n]{r}(\cos\alpha_k + i\sin\alpha_k)$, gdzie $\alpha_k = \frac{\varphi+2k\pi}{n}$, $k = 0, 1, \ldots, n-1$. $\square$

Symbolem $\sqrt[n]{w}$ oznaczamy zbiór wszystkich rozwiązań równania. Zatem

$$\sqrt[n]{w} = \{z \in \mathbb{C} : z^n = w\} = \{z_0, z_1, \ldots, z_{n-1}\}.$$

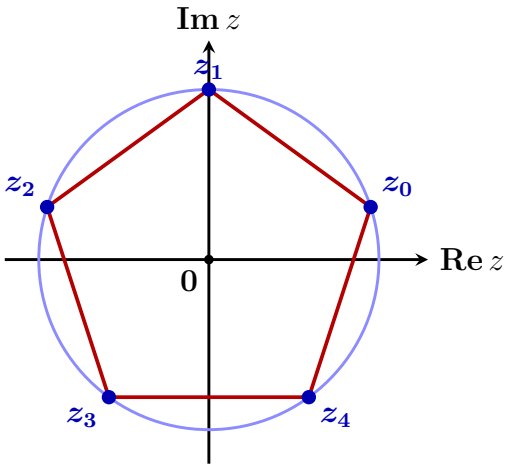
Przykład 2.3.4. Rozwiąż równanie $z^5 = \sqrt{8} - i\sqrt{24}$.
Niech $z = \rho(\cos\alpha + i\sin\alpha)$ oraz $w = \sqrt{8} - i\sqrt{24}$.
Obliczamy $|w| = \sqrt{32} = 4\sqrt{2}$, skąd $w = 4\sqrt{2}(\frac{1}{2} - \frac{\sqrt{3}}{2}i)$. Zatem

$z^5 = w \;\Leftrightarrow\; \rho^5(\cos 5\alpha + i\sin 5\alpha) = 4\sqrt{2}(\cos(-\frac{\pi}{3}) + i\sin(-\frac{\pi}{3})) \;\Leftrightarrow\; \begin{cases} \rho^5 = 4\sqrt{2} = (\sqrt{2})^5 \\ 5\alpha = -\frac{\pi}{3} + 2k\pi,\ k \in \{0,1,2,3,4\} \end{cases}$.

Stąd $\rho = \sqrt{2}$, $\alpha_k = -\frac{\pi}{15} + \frac{2}{5}k\pi$ oraz $z_k = \sqrt{2}(\cos\alpha_k + i\sin\alpha_k)$, gdzie $k \in \{0,1,2,3,4\}$.

Interpretacja geometryczna pierwiastka z liczby zespolonej

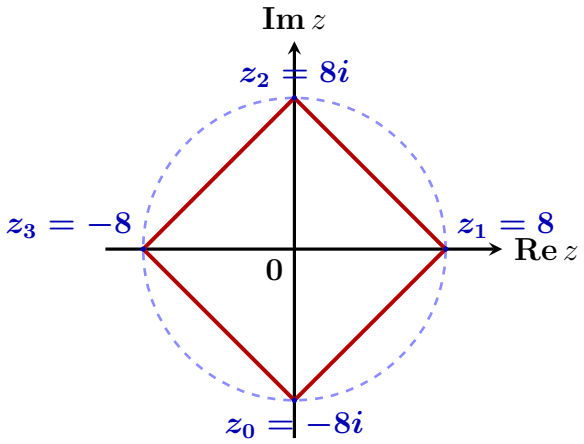
Liczby $z_0, z_1, \ldots, z_{n-1}$ będące rozwiązaniami równania $z^n = w$ stanowią wierzchołki n-kąta foremnego, wpisanego w koło o środku $z = 0$ i promieniu $\sqrt[n]{r}$.



Przykład 2.3.5. Rozwiąż równanie $z^4 = (\sqrt{3} - i)^{12}$.

Równanie ma 4 rozwiązania z_0, z_1, z_2, z_3. Będą one wierzchołkami kwadratu.

$z^4 = \left((\sqrt{3} - i)^3\right)^4$
Niech $z_0 = (\sqrt{3} - i)^3 = 3\sqrt{3} - 9i - 3\sqrt{3} + i = -8i$.
Kolejne wierzchołki kwadratu otrzymujemy przez obrót o kąt $\frac{\pi}{2}$, co odpowiada mnożeniu przez $i = e^{i\frac{\pi}{2}}$.
$z_1 = z_0 \cdot i = 8, \quad z_2 = z_1 \cdot i = 8i, \quad x_3 = z_2 \cdot i = -8$



Uwaga 2.3.6. Rozwiązywanie równań w $\mathbb{R}$ i w $\mathbb{C}$

w $\mathbb{R}$	w $\mathbb{C}$
$\sqrt{9} = 3$	$\sqrt{9} = \{-3, 3\}$
$\sqrt{-1}$ nie istnieje	$\sqrt{-1} = \{-i, i\}$
$\sqrt[4]{1} = 1$	$\sqrt[4]{1} = \{-1, 1, -i, i\}$
$\sqrt{x^2} = \lvert x\rvert$	$\sqrt{z^2} = \{-z, z\}$

Równania wielomianowe

Rozważmy wielomian zmiennej zespolonej z stopnia n.
$W(z) = c_n z^n + c_{n-1} z^{n-1} + \ldots + c_1 z + c_0, \quad c_0, \ldots, c_n \in \mathbb{C}, \quad c_n \neq 0$

Definicja 2.3.7. Liczbę $z_0 \in \mathbb{C}$ nazywamy *pierwiastkiem wielomianu* W, jeżeli $W(z_0) = 0$.

Twierdzenie 2.3.8 (Bézout)**.** Liczba $z_0 \in \mathbb{C}$ jest pierwiastkiem niezerowego wielomianu W wtedy i tylko wtedy gdy istnieje wielomian P taki, że $W(z) = (z - z_0)P(z)$.

Niech $k \in \mathbb{N}$.

Definicja 2.3.9. Liczbę $z_0 \in \mathbb{C}$ nazywamy *pierwiastkiem k-krotnym wielomianu* W, jeżeli istnieje wielomian P taki, że $W(z) = (z - z_0)^k P(z)$ oraz $P(z_0) \neq 0$.

Przykład 2.3.10. Niech $W(z) = z^3 - z^2 - z + 1$. Faktoryzując, otrzymujemy $W(z) = z^2(z-1) - (z-1) = (z-1)(z^2-1) = (z-1)^2(z+1)$.
Zatem $z = 1$ jest pierwiastkiem dwukrotnym.

Twierdzenie 2.3.11 (Zasadnicze twierdzenie algebry)**.** Każdy wielomian zespolony dodatniego stopnia ma pierwiastek w $\mathbb{C}$.

Wniosek 2.3.12. Każdy wielomian zespolony stopnia n ma dokładnie n pierwiastków w $\mathbb{C}$, licząc z krotnościami.

Dowód. Niech W będzie wielomianem stopnia n oraz niech $z_1, z_2 \ldots, z_m$ to jego wszystkie pierwiastki o krotnościach $k_1, k_2 \ldots, k_m$, odpowiednio. Na mocy zasadniczego twierdzenia algebry i twierdzenia Bézout otrzymujemy $k_1 + k_2 \ldots + k_m = n$ oraz

$$W(z) = (z - z_1)^{k_1} \cdot (z - z_2)^{k_2} \cdot \ldots \cdot (z - z_m)^{k_m}. \quad \square$$

Trójmian kwadratowy $az^2 + bz + c = 0, \quad a, b, c \in \mathbb{C}, \quad a \neq 0$

Obliczamy $\Delta = b^2 - 4ac \in \mathbb{C}$ oraz $\sqrt{\Delta} = \{-\delta, \delta\}$.

Gdy $\Delta \neq 0$, otrzymujemy dwa różne pierwiastki zespolone $z_1 = \frac{-b-\delta}{2a}$, $z_2 = \frac{-b+\delta}{2a}$.

Gdy $\Delta = 0$, otrzymujemy jeden pierwiastek podwójny $z_1 = z_2 = \frac{-b}{2a}$.

Przykład 2.3.13. Rozwiąż równanie $z^2 + 2iz + 3 = 0$.
Obliczamy $\Delta = 4i^2 - 12 = -16 = 16i^2$, $\sqrt{\Delta} = \{-4i, 4i\}$.
Niech $\delta = 4i$, wówczas $z_1 = -3i$ oraz $z_2 = i$.

Wielomiany zmiennej zespolonej o współczynnikach rzeczywistych

Niech $k \in \mathbb{N}$.

Twierdzenie 2.3.14. Niech W będzie wielomianem zmiennej zespolonej o współczynnikach rzeczywistych. Wówczas liczba $z_0 \in \mathbb{C}$ jest k-krotnym pierwiastkiem wielomianu W wtedy i tylko wtedy, gdy liczba $\bar{z}_0 \in \mathbb{C}$ jest k-krotnym pierwiastkiem wielomianu W.

Wniosek 2.3.15. Wielomian stopnia nieparzystego o współczynnikach rzeczywistych, ma przynajmniej jeden pierwiastek rzeczywisty.

Przykład 2.3.16. Rozwiąż równanie $z^3 - 3z^2 + 6z - 4 = 0$.

$0 = (z-1)(z^2 - 2z + 4) = (z-1)[(z-1)^2 + 3] = (z-1)[(z-1)^2 - (\sqrt{3}i)^2] =$
$= (z-1)(z-1-\sqrt{3}i)(z-1+\sqrt{3}i)$
rozwiązania $z_1 = 1, \quad z_2 = 1 + \sqrt{3}i, \quad z_3 = 1 - \sqrt{3}i, \qquad \bar{z}_2 = z_3$

Przykład 2.3.17. Wskaż wszystkie pierwiastki wielomianu $W(z) = z^4 + 2z^3 - 6z^2 - 22z - 15 = 0$, wiedząc, że $z_1 = -2 - i$ jest jednym z nich.

Wielomian ma współczynniki rzeczywiste, zatem $z_2 = -2 + i$.
Na mocy twierdzenia Bézout, $W(z)$ dzieli się bez reszty przez $(z - z_1)(z - z_2)$.

$$(z - z_1)(z - z_2) = (z + 2 + i)(z + 2 - i) = (z+2)^2 - i^2 = (z+2)^2 + 1 = z^2 + 4z + 5$$

Dzieląc wielomiany otrzymujemy

$$W(z) = (z^2 + 4z + 5)(z^2 - 2z - 3) = (z^2 + 4z + 5)(z - 3)(z + 1).$$

Pierwiastki wielomianu to $z_1 = -2 - i$, $z_2 = -2 + i$, $z_2 = 3$, $z_4 = 1$.

Przykład 2.3.18. Rozwiąż równanie $z^4 + (2+i)z^3 + (7+2i)z^2 + (12+i)z + 6 = 0$, wiedząc, że $z_1 = 2i$ jest jednym z jego rozwiązań.

$(z-2i)\Big(z^3 + (2+3i)z^2 + (1+6i)z + 3i\Big) = 0$

$(z-2i)(z+1)\Big(z^2 + (1+3i)z + 3i\Big) = (z-2i)(z+1)^2(z+3i) = 0$

rozwiązania $z_1 = 2i$, $z_2 = z_3 = -1$, $z_4 = -3i$

2.4 Interpretacja geometryczna

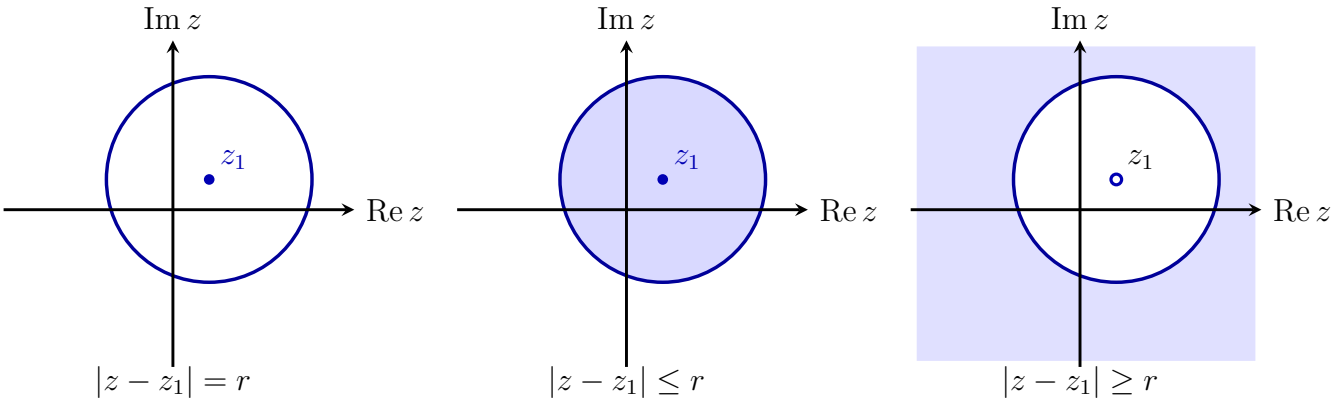
Niech $z = x + iy$, $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$ będą liczbami zespolonymi w postaci algebraicznej. Niech $r \in \mathbb{R}$, $r > 0$.

1) $|z_1 - z_2|$ odległość z_1 od z_2

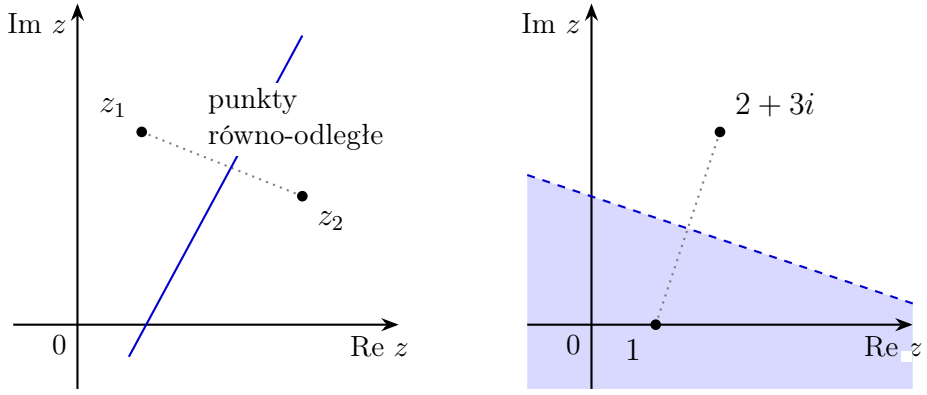
$$|(x_1 + iy_1) - (x_2 + iy_2)| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

2) $|z - z_1| = r$ równanie okręgu o środku z_1 i promieniu r

$r = |(x - x_1) + (y - y_1)i| = \sqrt{(x - x_1)^2 + (y - y_1)^2} \;\Rightarrow\; r^2 = (x - x_1)^2 + (y - y_1)^2$

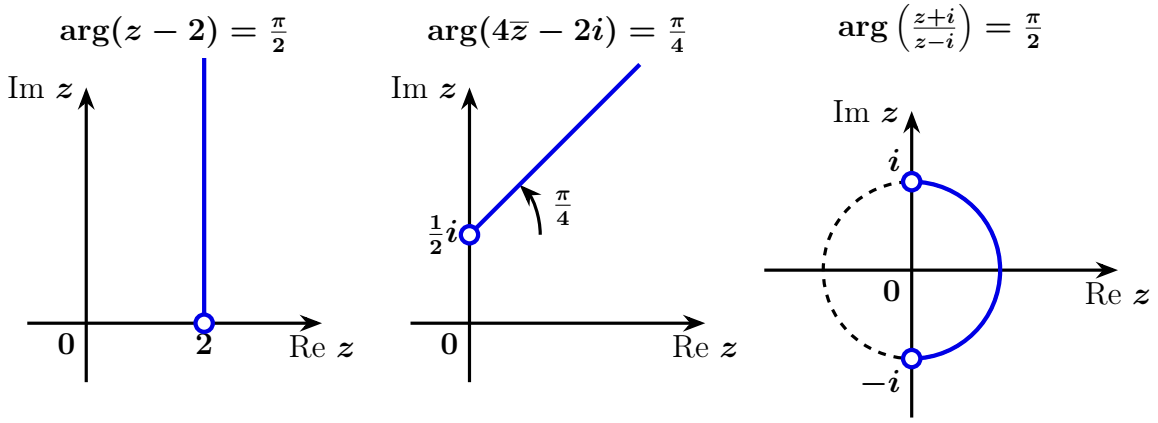


3) $|z - z_1| = |z - z_2|$ równanie symetralnej odcinka o końcach z_1 i z_2



np. $|z - 1| < |z - 2 - 3i|$

4) Zbiór $\{z \in \mathbb{C} : \ \arg z = \varphi\}$, gdzie $\varphi \in \mathbb{R}$ ustalony, to półprosta.



$$\frac{3}{4}\pi = \arg(4\overline{z} - 2i) = \arg[4(\overline{z} - \frac{1}{2}i)] = \arg(\overline{z} - \frac{1}{2}i) = 2\pi - \arg\left(\overline{\overline{z} - \frac{1}{2}i}\right) = 2\pi - \arg(z + \frac{1}{2}i)$$

$$\arg(z - (-\frac{1}{2}i)) = 2\pi - \frac{3}{4}\pi = \frac{\pi}{4}$$

5) Zbiór $\{z \in \mathbb{C} : \ \arg(\frac{z+i}{z-i}) = \frac{\pi}{2}\}$ to łuk na okręgu.

METODA I:

$$w = \frac{z+i}{z-i} = \frac{x+(y+1)i}{x+(y-1)i} = \frac{[x+(y+1)i]\cdot[x-(y-1)i]}{x^2+(y-1)^2} = \frac{x^2 - xyi + xi + xyi + xi + y^2 - 1}{x^2+(y-1)^2}$$

$$\mathrm{Re}w = \frac{x^2+y^2-1}{x^2+(y-1)^2}, \quad \mathrm{Im}w = \frac{2x}{x^2+(y-1)^2} \ \Rightarrow \ \begin{cases} \mathrm{Re}w = 0 \ \Leftrightarrow \ x^2+y^2 = 1 \\ \mathrm{Im}w > 0 \ \Leftrightarrow \ x > 0 \end{cases}$$

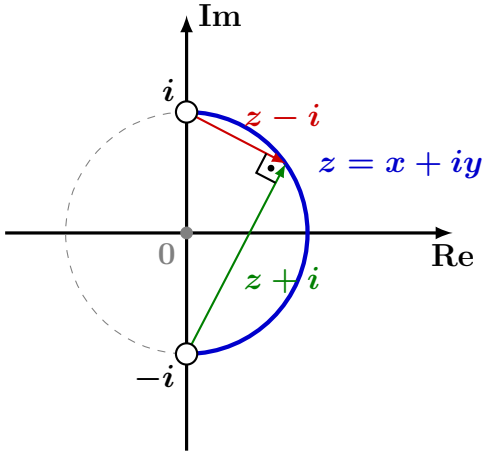
METODA II:

$$\frac{\pi}{2} = \arg\left(\frac{z+i}{z-i}\right) = \arg(z+i) - \arg(z-i).$$

Zatem $\arg(z+i) = \frac{\pi}{2} + \arg(z-i)$, co oznacza, że kąt skierowany między wektorami $\boldsymbol{z-i}$ oraz $\boldsymbol{z+i}$ wynosi dokładnie $\frac{\pi}{2}$.

Na mocy twierdzenia Talesa o kącie wpisanym opartym na średnicy okręgu, zbiorem punktów $\boldsymbol{z}$ spełniających ten warunek jest łuk okręgu, którego średnicą jest odcinek łączący punkty $\boldsymbol{-i}$ oraz $\boldsymbol{i}$.

Ponieważ różnica argumentów musi być dodatnia i równa $\frac{\pi}{2}$, rozwiązanie stanowi wyłącznie prawy półokrąg (dla $\mathbf{Re}\, \boldsymbol{z} > \mathbf{0}$), z wyłączeniem punktów osobliwych $\boldsymbol{i}$ oraz $\boldsymbol{-i}$.



QUIZ

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. $Re\Big[(\sin\frac{\pi}{6} - i\cos\frac{\pi}{3})^{12}\Big] = 0$
2. Niech $z = a + bi$, gdzie $a, b \in \mathbb{R}$, $a > 0$, $b > 0$. Wówczas $Re(z^2) > 0$ oraz $Im(z^2) > 0$.
3. $i^{4n} + i^{4n+1} + i^{4n+2} + i^{4n+3} = 0$ dla dowolnego $n \in \mathbb{N}$.
4. Argumentem głównym liczby $w = \frac{(1+i)^{17}}{i^9(\cos\frac{\pi}{3} - i\sin\frac{\pi}{3})}$ jest $\frac{\pi}{12}$.
5. Każdy wielomian zespolony dodatniego stopnia ma pierwiastek w ciele $\mathbb{C}$.
6. Nie jest możliwe, by wielomian zmiennej zespolonej o współczynnikach rzeczywistych stopnia 5 miał dwa pierwiastki rzeczywiste i trzy zespolone.
7. Nie jest możliwe, by wielomian $P(z) \in \mathbb{C}[z]$ stopnia 5 miał dwa pierwiastki rzeczywiste i trzy zespolone.
8. Zbiór $\{z \in \mathbb{C} : |z + 1 + i| = 1 \wedge |z + 2 + i| = |z + i|\}$ jest zbiorem dwuelementowym.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Algebra
Wykład nr 3
Data: .....

TEMAT: *Macierze i ich własności. Wyznacznik macierzy.*

3.1 Macierze i ich własności

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$.

Definicja 3.1.1. Funkcję $A : \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \to K$, $A(i, j) = a_{ij}$ nazywamy *macierzą* (rzeczywistą gdy $K = \mathbb{R}$, zespoloną gdy $K = \mathbb{C}$) o m wierszach i n kolumnach.

Wartości a_{ij} nazywamy *wyrazami* lub *elementami* macierzy. Odwzorowanie A oznaczamy symbolem $[a_{ij}]_{m\times n}$.

$$A = [a_{ij}]_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

Ciąg $a_{i1}, a_{i2}, \ldots, a_{in}$ nazywamy *i-tym wierszem* macierzy A, zaś ciąg $a_{1j}, a_{2j}, \ldots, a_{nj}$ nazywamy *j-tą kolumną* macierzy A.

Oznaczmy symbolem $M_{m\times n}(K)$ zbiór wszystkich macierzy o m wierszach i n kolumnach i elementach z K. Gdy $m = n$, piszemy krócej $M_n(K)$. Macierz $A \in M_n(K)$ nazywamy *macierzą kwadratową stopnia n*.

Dla $A, B \in M_{m\times n}(K)$, $A = [a_{ij}]$, $B = [b_{ij}]$ mamy
$A = B \Leftrightarrow \forall i \in \{1, 2, \ldots, m\}\ \forall j \in \{1, 2, \ldots, n\}\ a_{ij} = b_{ij}$.

Przykład 3.1.2. $A \in M_{3\times 4}(\mathbb{R})$, $\quad A = \begin{bmatrix} 1 & -1 & 2 & -2 \\ 3 & -3 & 4 & 4 \\ 5 & -5 & 6 & -6 \end{bmatrix}$

$$B \in M_3(\mathbb{R}), \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\mathbf{0}_{m\times n} = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ & \vdots & & \\ 0 & 0 & \ldots & 0 \end{bmatrix}_{m\times n} \quad \text{macierz zerowa wymiaru } m \times n$$

Typy macierzy

Definicja 3.1.3. Macierz kwadratową $A = [a_{ij}] \in M_n(K)$ nazywamy macierzą:

a) *diagonalną* lub *przekątniową*, gdy $a_{ij} = 0$ dla $i \neq j$

b) *trójkątną górną*, gdy $a_{ij} = 0$ dla $i > j$

c) *trójkątną dolną*, gdy $a_{ij} = 0$ dla $i < j$

Oznaczmy:
$D_n(K)$ zbiór macierzy diagonalnych stopnia n
$T_n^G(K)$ zbiór macierzy trójkątnych górnych stopnia n
$T_n^D(K)$ zbiór macierzy trójkątnych dolnych stopnia n

$$T_n^G(K) \cap T_n^D(K) = D_n(K)$$

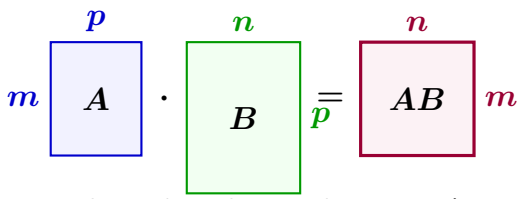
Przykład 3.1.4. I_n - macierz jednostkowa stopnia n

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \in D_3(K)$$

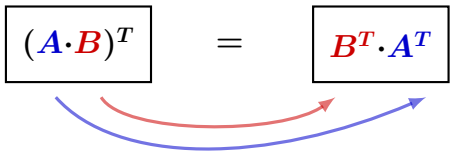
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \in T_3^G(K) \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 3 \end{bmatrix} \in T_3^D(K)$$

Działania na macierzach

- Dodawanie macierzy: Niech $A, B \in M_{m\times n}(K)$, $A = [a_{ij}]$, $B = [b_{ij}]$.
  $C = A + B$, $C = [c_{ij}] \in M_{m\times n}(K)$, $c_{ij} = a_{ij} + b_{ij}$
- Mnożenie macierzy przez skalar: Niech $\alpha \in K$, $A \in M_{m\times n}(K)$, $A = [a_{ij}]$.
  $C = \alpha \cdot A$, $C = [c_{ij}] \in M_{m\times n}(K)$, $c_{ij} = \alpha \cdot a_{ij}$
  Oznaczamy $-A = (-1) \cdot A$ oraz $A - B = A + (-B)$.
- Mnożenie macierzy: Niech $A \in M_{m\times \mathbf{p}}(K)$, $A = [a_{ik}]$, $B \in M_{\mathbf{p}\times n}(K)$, $B = [b_{kj}]$.
  $C = A \cdot B$, $C = [c_{ij}] \in M_{m\times n}(K)$, $c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}$
  Dla $r \in \mathbb{N}$ oznaczamy $A^r = \underbrace{A \cdot \ldots \cdot A}_{r-razy}$
- Transponowanie: Niech $A \in M_{m\times n}(K)$, $A = [a_{ij}]$
  $C = A^T$, $C = [c_{ij}] \in M_{n\times m}(K)$, $c_{ij} = a_{ji}$



Warunki wykonalności iloczynu $\boldsymbol{AB}$.



Transpozycja iloczynu. Odwrócenie kolejności mnożenia czynników.

Przykład 3.1.5.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 2 & 2 \\ -1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 1 \\ 2 & 5 \end{bmatrix}, \quad D = \begin{bmatrix} -1 & 0 \\ 3 & 4 \end{bmatrix}$$

Jakie mnożenia są wykonalne? Wyznacz $C - 2D$, $\frac{1}{2} \cdot A^T$, AB, BA, D^2.

$$C - 2D = \begin{bmatrix} 2 & 1 \\ 2 & 5 \end{bmatrix} - 2\begin{bmatrix} -1 & 0 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 5 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ -6 & -8 \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ -4 & -3 \end{bmatrix}$$

$$\frac{1}{2} \cdot A^T = \frac{1}{2}\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 2 \\ 1 & \frac{5}{2} \\ \frac{3}{2} & 3 \end{bmatrix}$$

$D = AB \in M_2(\mathbb{R})$ schemat Falka

			0	1
			2	2
			-1	1
1	2	3	1	8
4	5	6	4	20

$1 = 1 \cdot 0 + 2 \cdot 2 + 3 \cdot (-1)$

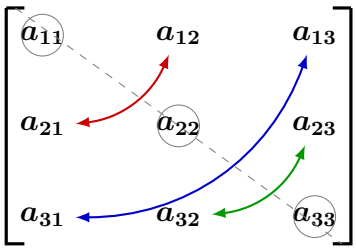
$G = BA \in M_3(\mathbb{R})$

		1	2	3
		4	5	6
0	1	4	5	6
2	2	10	14	18
-1	1	3	3	3

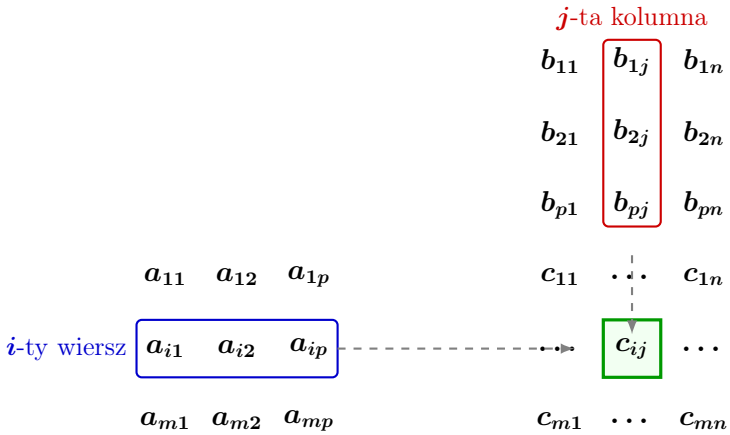
$AB \neq BA$

$D^2 = D \cdot D \in M_2(\mathbb{R})$, $CA \in M_{2\times 3}(\mathbb{R})$, zaś AC jest niewykonalne

$$CD \in M_2(\mathbb{R}),\ DC \in M_2(\mathbb{R}),\ CD = \begin{bmatrix} 1 & 4 \\ 13 & 20 \end{bmatrix},\ DC = \begin{bmatrix} -2 & -1 \\ 14 & 23 \end{bmatrix},\ CD \neq CD$$



Interpretacja geometryczna transponowania $\boldsymbol{A^T}$ macierzy kwadratowej $\boldsymbol{A}$ jako odbicia zwierciadlanego elementów względem głównej przekątnej.



Mnożenie macierzy za pomocą schematu Falka.

Własności działań na macierzach

Twierdzenie 3.1.6. Niech $A, B, C, \mathbf{0} \in M_{m\times n}(K)$, $\alpha, \beta \in K$. Wówczas prawdziwe są następujące równości.

i) $A + B = B + A$

ii) $(A + B) + C = A + (B + C)$

iii) $A + \mathbf{0} = \mathbf{0} + A$

iv) $(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$

v) $\alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$

vi) $(\alpha \cdot \beta) \cdot A = \alpha \cdot (\beta \cdot A)$

Wniosek 3.1.7. $(M_{m\times n}(K), +)$ jest grupą abelową.

Twierdzenie 3.1.8. Niech A, B, C będą macierzami o elementach z K oraz niech $\alpha \in K$. Jeśli działania są wykonalne, to wówczas prawdziwe są następujące równości.

i) $(AB)C = A(BC)$

ii) $(A + B)C = AC + BC$

iii) $A(B + C) = AB + AC$

iv) $\alpha(AB) = (\alpha A)B = A(\alpha B)$

v) $AI = A$

vi) $IA = A$

Wniosek 3.1.9. i) $(M_n(K), \cdot)$ jest półgrupą nieprzemienną z jedynką.

ii) $(M_n(K), +, \cdot)$ jest pierścieniem nieprzemiennym z jedynką.

Poniższy przykład pokazuje, że $(M_n(K), +, \cdot)$ nie jest ciałem, bowiem w zbiorze macierzy kwadratowych stopnia n o elementach z ciała K istnieją *dzielniki zera*.

Przykład 3.1.10. Rozważmy macierze $A, B \in M_2(\mathbb{R})$, $A = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 6 & 6 \\ 3 & 3 \end{bmatrix}$. Obliczamy $BA = \begin{bmatrix} -12 & 24 \\ -6 & 12 \end{bmatrix}$ oraz $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Twierdzenie 3.1.11. Niech A, B będą macierzami o elementach z K oraz niech $\alpha \in K$, $r \in \mathbb{N}$. Jeśli działania są wykonalne, to wówczas prawdziwe są następujące równości.

i) $(A^T)^T = A$

ii) $(A+B)^T = A^T + B^T$

iii) $(\alpha A)^T = \alpha \cdot A^T$

iv) $(AB)^T = B^T A^T$

v) $(A^r)^T = (A^T)^r$

Definicja 3.1.12. Niech $A = [a_{ij}] \in M_n(K)$. Sumę elementów na przekątnej $a_{11}+a_{22}+\ldots+a_{nn}$ nazywamy *śladem macierzy* A i oznaczamy symbolem $\mathrm{tr}(A)$.

Przykład 3.1.13. $\mathrm{tr}\begin{bmatrix} 1 & 0 & 0 \\ 2 & -7 & 0 \\ 0 & 1 & 3 \end{bmatrix} = 1 - 7 + 3 = -3$

Własności śladu macierzy

Twierdzenie 3.1.14. Niech $A, B \in M_n(K)$, $\alpha \in K$. Wówczas prawdziwe są następujące równości.

i) $\mathrm{tr}(A) = \mathrm{tr}(A^T)$ ii) $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$

iii) $\mathrm{tr}(\alpha A) = \alpha \mathrm{tr}(A)$ iv) $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

Wniosek 3.1.15. Jeśli $A, B \in M_{m\times n}(K)$, to wówczas $\mathrm{tr}(A^T B) = \mathrm{tr}(AB^T) = \mathrm{tr}(B^T A) = \mathrm{tr}(BA^T)$.

Definicja 3.1.16. Macierz kwadratową $A = [a_{ij}] \in M_n(K)$ nazywamy macierzą

a) *symetryczną*, gdy $A = A^T$,

b) *antysymetryczną*, gdy $A = -A^T$.

Twierdzenie 3.1.17. Każdą macierz kwadratową można przedstawić w postaci sumy macierzy symetrycznej i antysymetrycznej.

Dowód. $A = B + C$, gdzie $B = \frac{1}{2}(A + A^T)$, $C = \frac{1}{2}(A - A^T)$, $B = B^T$, $C = -C^T$. Ponadto $B^T = \left[\frac{1}{2}(A+A^T)\right]^T = \frac{1}{2}\left[(A+A^T)\right]^T = \frac{1}{2}(A^T+(A^T)^T) = \frac{1}{2}(A^T+A) = B$. Analogicznie sprawdzamy, że $C^T = -C$. $\square$

Przykład 3.1.18. $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 3 & 0 & 7 \end{bmatrix}$ macierz symetryczna

$B = \begin{bmatrix} 0 & -2 & 3 \\ 2 & 0 & 6 \\ -3 & -6 & 0 \end{bmatrix}$ macierz antysymetryczna

Definicja 3.1.19. Macierz utworzoną z macierzy A_{ij}, dla $1 \leq i \leq m, 1 \leq j \leq n$ postaci

$$\left[\begin{array}{c|c|c|c} A_{11} & A_{12} & \dots & A_{1n} \\ \hline A_{21} & A_{22} & \dots & A_{2n} \\ \hline \dots & \dots & \ddots & \dots \\ \hline A_{m1} & A_{m2} & \dots & A_{mn} \end{array}\right]$$

nazywamy *macierzą blokową*. Macierze $A_{i1}, A_{i2}, \dots, A_{in}$ stojące w i-tym wierszu macierzy blokowej muszą mieć te same liczby wierszy, zaś macierze $A_{1j}, A_{2j}, \dots, A_{mj}$ stojące w j-tej kolumnie muszą mieć te same liczby kolumn.

3.2 Wyznacznik macierzy

Definicja indukcyjna wyznacznika

Definicja 3.2.1. *Wyznacznikiem* macierzy kwadratowej $A = [a_{ij}] \in M_n(K)$ nazywamy liczbę $\det A \in K$ określoną następująco:

- gdy $n = 1$, to $\det A = a_{11}$
- gdy $n \geq 2$, to $\det A = \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det A_{1j} =$
  $= (-1)^{1+1}a_{11}\det A_{11} + (-1)^{1+2}a_{12}\det A_{12} + \dots (-1)^{1+n}a_{1n}\det A_{1n}$,

gdzie A_{1j} oznacza macierz stopnia $n-1$ otrzymaną z macierzy A przez skreślenie pierwszego wiersza i j-tej kolumny.

Oznaczenia:

$$\det A = \det \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

Przykład 3.2.2. i) $A = \begin{bmatrix} 1 & 5 \\ -3 & 2 \end{bmatrix}$

$$\det A = (-1)^{1+1} a_{11} \det A_{11} + (-1)^{1+2} a_{12} \det A_{12} = 1 \cdot 1 \cdot 2 - 5 \cdot (-3) = 17$$

ii) $B = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & -6 \\ -3 & -6 & 8 \end{bmatrix}$

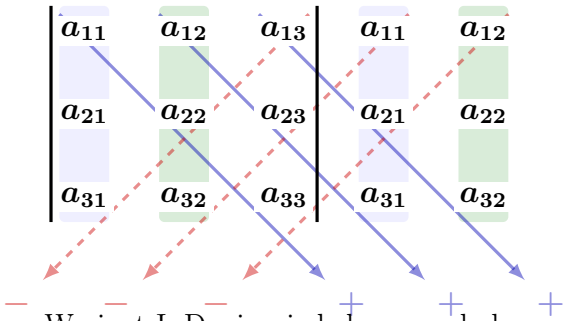
$$\det B = (-1)^{1+1} a_{11} \det B_{11} + (-1)^{1+2} a_{12} \det B_{12} + (-1)^{1+3} a_{13} \det B_{13} =$$

$$= 1 \cdot \begin{vmatrix} 1 & -6 \\ -6 & 8 \end{vmatrix} - (-2) \cdot \begin{vmatrix} 2 & -6 \\ -3 & 8 \end{vmatrix} + 3 \cdot \begin{vmatrix} 2 & 1 \\ -3 & -6 \end{vmatrix} = -59$$

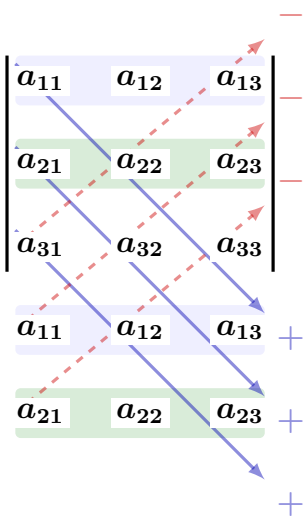
Metoda Sarrusa (**tylko dla wyznaczników stopnia 3!**)

$$\begin{vmatrix} 1 & -2 & 3 \\ 2 & 1 & -6 \\ -3 & -6 & 8 \end{vmatrix} = 1 \cdot 1 \cdot 8 + 2 \cdot (-6) \cdot 3 + (-3) \cdot (-2) \cdot (-6)$$

$$\begin{matrix} 1 & -2 & 3 \\ 2 & 1 & -6 \end{matrix} \qquad - \Big[3 \cdot 1 \cdot (-3) + (-6) \cdot (-6) \cdot 1 + 8 \cdot (-2) \cdot 2 \Big] = -59$$



Wariant I: Dopisanie kolumn na boku



Wariant II: Dopisanie wierszy na dole

Metoda Sarrusa

Definicja 3.2.3. Niech $A = [a_{ij}]$ będzie macierzą kwadratową stopnia n o elementach z K.

i) *Minorem elementu* a_{ij} nazywamy wyznacznik macierzy A_{ij} stopnia $n-1$ otrzymanej poprzez skreślenie i-tego wiersza oraz j-tej kolumny macierzy A. Oznaczamy go symbolem M_{ij}.

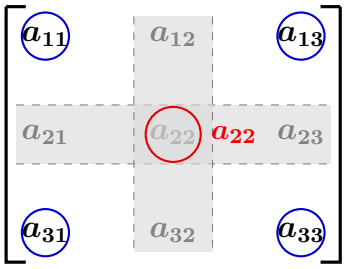
ii) *Dopełnieniem algebraicznym* elementu a_{ij} nazywamy liczbę $D_{ij} := (-1)^{i+j} \cdot M_{ij} \in K$.

Przykład 3.2.4. $B = \begin{bmatrix} 1 & -22 & 3 \\ 11 & 17 & 16 \\ -3 & -6 & 80 \end{bmatrix}$

$B_{32} = \begin{bmatrix} 1 & 3 \\ 11 & 16 \end{bmatrix}$, $\quad M_{32} = \det B_{32} = \begin{vmatrix} 1 & 3 \\ 11 & 16 \end{vmatrix} = 16 - 33 = -17$

$D_{32} = (-1)^{3+2} M_{23} = 17$

Minor $\boldsymbol{M_{ij}}$ wyznaczamy bezpośrednio jako wyznacznik macierzy stopnia $\boldsymbol{n-1}$, powstałej poprzez całkowite wykreślenie $\boldsymbol{i}$-tego wiersza oraz $\boldsymbol{j}$-tej kolumny wyjściowej macierzy $\boldsymbol{A}$.



$$\Longrightarrow \boldsymbol{M_{22}} = \begin{vmatrix} \boldsymbol{a_{11}} & \boldsymbol{a_{13}} \\ \boldsymbol{a_{31}} & \boldsymbol{a_{33}} \end{vmatrix}$$

$$\boldsymbol{D_{22} = (-1)^{2+2} \cdot M_{22}}$$

Twierdzenie 3.2.5 (Laplace'a)**.** Niech $A = [a_{ij}] \in M_n(K)$, gdzie $n \geq 2$. Wówczas:

i) $\forall i \in \{1, 2, \ldots, n\} \quad \det A = \sum_{k=1}^{n} a_{ik} D_{ik}$,

(rozwinięcie względem i-tego wiersza)

ii) $\forall j \in \{1, 2, \ldots, n\} \quad \det A = \sum_{k=1}^{n} a_{kj} D_{kj}$.

(rozwinięcie względem j-tej kolumny)

Przykład 3.2.6. $B = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 2 & -1 \\ 1 & 2 & 2 \end{bmatrix}$ Rozwijamy względem drugiego wiersza.

$$\det B = 0 \cdot D_{21} + 2 \cdot D_{22} + (-1) \cdot D_{23} = 2 \cdot (-1)^4 \cdot \begin{vmatrix} 1 & 4 \\ 1 & 2 \end{vmatrix} + (-1) \cdot (-1)^5 \cdot \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = -4$$

$C = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 2 & 0 & 2 & 0 & 0 \\ 2 & 1 & 3 & 0 & 1 \\ 3 & 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 4 & 1 \end{bmatrix}$ Rozwijamy względem czwartej kolumny,

a potem względem ostatniego wiersza.

$$\det C = 4 \cdot (-1)^9 M_{54} = -4 \cdot \begin{vmatrix} 1 & 0 & 1 & 1 \\ 2 & 0 & 2 & 0 \\ 2 & 1 & 3 & 1 \\ 3 & 2 & 0 & 0 \end{vmatrix} = -4(3D_{41} + 2D_{42}) =$$

$$-12 \cdot (-1)^5 M_{41} - 8 \cdot (-1)^6 M_{42} = 12 \cdot \begin{vmatrix} 0 & 1 & 1 \\ 0 & 2 & 0 \\ 1 & 3 & 1 \end{vmatrix} - 8 \cdot \begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 0 \\ 2 & 3 & 1 \end{vmatrix} = -40$$

Wniosek 3.2.7. Niech $A = [a_{ij}] \in T_n^G(K)$ lub $A = [a_{ij}] \in T_n^D(K)$. Wówczas

$$\det A = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}.$$

Przykład 3.2.8. $A = \begin{bmatrix} 5 & 2 & 0 & 0 \\ 0 & 3 & -11 & 22 \\ 0 & 0 & -1 & 7 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

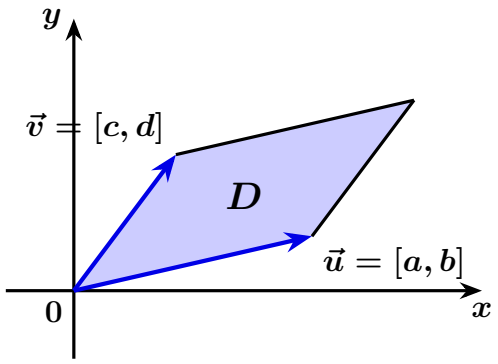
$$\det A = 5 \cdot (-1)^2 \begin{vmatrix} 3 & -11 & 22 \\ 0 & -1 & 7 \\ 0 & 0 & 2 \end{vmatrix} = 5 \cdot 3 \cdot (-1)^2 \begin{vmatrix} -1 & 7 \\ 0 & 2 \end{vmatrix} = 5 \cdot 3 \cdot (-1) \cdot 2$$

$$\det B = 10 \cdot 2 \cdot 3 \cdot (-1) = -60$$

Interpretacja geometryczna wyznacznika macierzy stopnia 2

Rozważmy wektory na płaszczyźnie $\vec{u} = [a, b]$, $\vec{v} = [c, d]$. Niech D będzie równoległobokiem rozpiętym na tych wektorach. Wówczas pole $|D|$ tego równoległoboku jest równe $|\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}|$.

Dowód można znaleźć w [3].



Własności wyznaczników

Niech $A = [a_{ij}] \in M_n(K)$. Oznaczmy przez A_k k-tą kolumnę macierzy A, czyli $A = [A_1, A_2, \ldots, A_n]$.

Twierdzenie 3.2.9. Niech $A \in M_n(K)$. Wówczas prawdziwe są następujące stwierdzenia.

i) $\det A = \det(A^T)$

ii) Jeśli pewna kolumna składa się z samych zer, to wówczas $\det A = 0$.

iii) Jeśli macierz B powstaje z macierzy A poprzez pomnożenie jednej kolumny przez liczbę $\lambda \neq 0$, to wówczas $\det B = \lambda \cdot \det A$.

iv) Jeśli $A_k = B_k + C_k$, to wówczas $\det A = \det[A_1, \ldots, B_k, \ldots, A_n] + \det[A_1, \ldots, C_k, \ldots, A_n]$

v) Jeśli macierz B powstaje z macierzy A poprzez przestawienie między sobą dwóch kolumn, to wówczas $\det B = -\det A$.

vi) Jeśli istnieją $k, l \in \{1, 2, \ldots, n\}$ takie, że $k \neq l$ oraz $A_k = \lambda A_l$, dla pewnego $\lambda \in K$, to wówczas $\det A = 0$.

vii) Jeśli jedna z kolumn jest kombinacją liniową pozostałych, tzn. $\exists \lambda_1, \ldots, \lambda_{k-1}, \lambda_{k+1}, \ldots, \lambda_n \in K : \; A_k = \sum_{j=1, j\neq k}^{n} \lambda_j \cdot A_j$, to wówczas $\det A = 0$.

viii) Jeśli do dowolnie wybranej kolumny dodamy kombinację liniową pozostałych kolumn macierzy A, to wyznacznik nie zmieni się.

ix) Jeśli $B \in M_n(K)$, to wówczas $\det(AB) = \det A \cdot \det B$.

x) Jeśli macierz A jest macierzą blokową postaci

$$\left[\begin{array}{c|c|c|c} B_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \hline ? & B_2 & \ldots & \mathbf{0} \\ \hline \ldots & \ldots & \ddots & \ldots \\ \hline ? & ? & \ldots & B_n \end{array}\right],$$

gdzie $B_1, B_2, \ldots, B_n$ są macierzami kwadratowymi (na ogół różnych stopni), $\mathbf{0}$ macierzami zerowymi, zaś **?** dowolnymi macierzami odpowiednich wymiarów, to wówczas $\det A = \det B_1 \cdot \det B_2 \cdot \ldots \cdot \det B_n$.

Przykład 3.2.10. Wiedząc, że $\begin{vmatrix} a & 2 & d \\ b & 2 & e \\ c & 2 & f \end{vmatrix} = 5$ oraz $\begin{vmatrix} a & 1 & d \\ b & 3 & e \\ c & 5 & f \end{vmatrix} = 1$, obliczymy $\begin{vmatrix} a & 2 & d \\ b & -2 & e \\ c & -6 & f \end{vmatrix}$.

$$\begin{vmatrix} a & 2 & d \\ b & -2 & e \\ c & -6 & f \end{vmatrix} = \begin{vmatrix} a & 3\cdot 2 - 4\cdot 1 & d \\ b & 3\cdot 2 - 4\cdot 3 & e \\ c & 3\cdot 2 - 4\cdot 5 & f \end{vmatrix} \underset{iv)}{=} \begin{vmatrix} a & 3\cdot 2 & d \\ b & 3\cdot 2 & e \\ c & 3\cdot 2 & f \end{vmatrix} + \begin{vmatrix} a & -4\cdot 1 & d \\ b & -4\cdot 3 & e \\ c & -4\cdot 5 & f \end{vmatrix} \underset{iii)}{=} 3\cdot 5 - 4\cdot 1 = 11$$

Przykład 3.2.11.

$$A = \left[\begin{array}{cc|cc} 2 & 1 & 0 & 0 \\ 7 & 1 & 0 & 0 \\ \hline 0 & 0 & 10 & 3 \\ 0 & 0 & 5 & 4 \end{array}\right] \quad macierz \quad blokowo-diagonalna$$

$$\det A = \det \begin{bmatrix} 2 & 1 \\ 7 & 1 \end{bmatrix} \cdot \det \begin{bmatrix} 10 & 3 \\ 5 & 4 \end{bmatrix} = (2-7)\cdot(40-15) = -125$$

Wniosek 3.2.12. i) Analogiczne własności zachodzą dla wierszy.

ii) $\det(\lambda A) = \lambda^n \cdot \det A$ dla dowolnego $0 \neq \lambda \in K$

iii) $\det(A^r) = (\det A)^r$ dla dowolnego $r \in \mathbb{N}$.

Metoda operacji elementarnych obliczania wyznacznika

Operacje elementarne:	$w_i \leftrightarrow w_j$	zamiana wierszy miejscami (między sobą)
	$\lambda \cdot w_i,\ \lambda \in K, \lambda \neq 0$	pomnożenie wiersza przez liczbę (różną od zera)
	$w_i + \lambda \cdot w_j,\ \lambda \in K$	dodanie do w_i wielokrotności w_j

Przykład 3.2.13.

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 1 & 10 & 1 \\ 3 & 2 & 10 & 3 \\ 1 & 1 & 5 & 5 \end{bmatrix}$$

$$\det A \overset{w_1 \leftrightarrow w_2}{=} - \begin{vmatrix} 1 & 1 & 10 & 1 \\ 2 & 1 & 3 & 4 \\ 3 & 2 & 10 & 3 \\ 1 & 1 & 5 & 5 \end{vmatrix} \overset{\substack{w_2-2w_1 \\ w_4-w_1 \\ w_3-3w_1}}{=} - \begin{vmatrix} 1 & 1 & 10 & 1 \\ 0 & -1 & -17 & 2 \\ 0 & -1 & -20 & 0 \\ 0 & 0 & -5 & 4 \end{vmatrix} \overset{w_3-w_2}{=} - \begin{vmatrix} 1 & 1 & 10 & 1 \\ 0 & -1 & -17 & 2 \\ 0 & 0 & -3 & -2 \\ 0 & 0 & -5 & 4 \end{vmatrix} \overset{-\frac{1}{3}\cdot w_3}{=} 3 \cdot$$

$$\begin{vmatrix} 1 & 1 & 10 & 1 \\ 0 & -1 & -17 & 2 \\ 0 & 0 & 1 & \frac{2}{3} \\ 0 & 0 & -5 & 4 \end{vmatrix} \overset{w_4+5w_3}{=} 3 \cdot \begin{vmatrix} 1 & 1 & 10 & 1 \\ 0 & -1 & -17 & 2 \\ 0 & 0 & 1 & \frac{2}{3} \\ 0 & 0 & 0 & \frac{22}{3} \end{vmatrix} = 3 \cdot 1 \cdot (-1) \cdot 1 \cdot \tfrac{22}{3} = -22$$

Przykład 3.2.14. Dane są macierze $A, B \in M_5(\mathbb{R})$ takie, że $A = [a_{ij}]$, gdzie $a_{ij} = \begin{cases} 1 & ;\ i \neq j \\ 2(i-1)+j & ;\ i = j \end{cases}$, zaś $\det B = 9$. Oblicz wyznacznik macierzy $C = 3B^3A^2B^T$.

Korzystając z własności wyznaczników, obliczamy

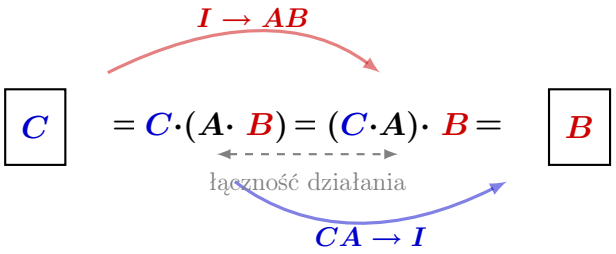
$$\det C = 3^5 \det(B^3)\det(A^2)\det(B^T) = 3^5(\det B)^4(\det A)^2.$$

Ponadto $\det A = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 4 & 1 & 1 & 1 \\ 1 & 1 & 7 & 1 & 1 \\ 1 & 1 & 1 & 10 & 1 \\ 1 & 1 & 1 & 1 & 13 \end{vmatrix} \overset{\substack{w_2-w_1 \\ w_3-w_1 \\ = \\ w_4-w_1 \\ w_5-w_1}}{} \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 0 & 12 \end{vmatrix} = 3 \cdot 6 \cdot 9 \cdot 12 = 3^5 \cdot 2^3.$

Zatem $\det C = 3^5 \cdot 9^4 \cdot (3^5 \cdot 2^3)^2 = 3^{23} \cdot 2^6$.

3.3 Macierz odwrotna

Jeśli dla danej macierzy kwadratowej $A \in M_n(K)$ istnieją macierze $B, C \in M_n(K)$ takie, że $AB = I$ oraz $CA = I$, to wówczas $B = C$. [6]



Definicja 3.3.1. Macierz $B \in M_n(K)$ nazywamy *macierzą odwrotną* do macierzy $A \in M_n(K)$, jeżeli $AB = BA = I_n$. Oznaczamy ją wówczas symbolem A^{-1}.

Przykład 3.3.2.

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ?$$

$$AA^{-1} = I \Leftrightarrow \begin{bmatrix} c & d \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow c = 0 \wedge \quad c = 1 \text{ sprzeczność}$$

Zatem nie istnieje A^{-1}.

Definicja 3.3.3. i) Macierz $A \in M_n(K)$, dla której istnieje macierz odwrotna, nazywamy macierzą *odwracalną*.

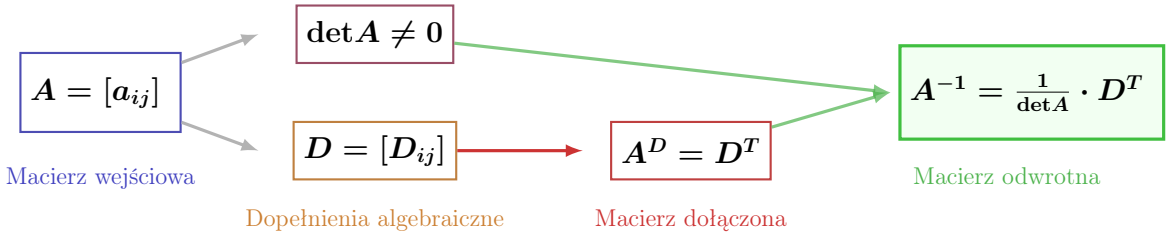
ii) Macierz $A \in M_n(K)$ taką, że $\det A = 0$ nazywamy macierzą *osobliwą*. W przeciwnym wypadku nazywamy ją macierzą *nieosobliwą*.

Twierdzenie 3.3.4. a) Macierz $A \in M_n(K)$ jest odwracalna wtedy i tylko wtedy, gdy jest nieosobliwa. Macierz odwrotna jest wówczas określona jednoznacznie.

b) Niech $A \in M_n(K)$ będzie macierzą nieosobliwą, zaś $D = [D_{ij}]$ macierzą jej dopełnień algebraicznych. Wówczas $A^{-1} = \frac{1}{\det A} \cdot D^T$.

[6] Można uzasadnić tenże fakt przy słabszych założeniach, mianowicie jeśli dla danej macierzy $A \in M_n(K)$ istnieje macierz $B \in M_n(K)$ taka, że $BA = I$, to wówczas $AB = I$ i vice versa.

Definicja 3.3.5. Niech $A \in M_n(K)$ będzie macierzą nieosobliwą, zaś $D = [D_{ij}]$ macierzą jej dopełnień algebraicznych. Macierz D^T nazywamy *macierzą dołączoną* do macierzy A i oznaczamy symbolem A^D.



Wniosek 3.3.6. Zbiór macierzy kwadratowych nieosobliwych stopnia n o elementach z ciała K wraz z działaniem mnożenia macierzy tworzy grupę nieprzemienną. Grupę tę oznaczamy symbolem $GL_n(K)$ i nazywamy *ogólną grupą liniową*.

Metody wyznaczania macierzy odwrotnej

1. **Za pomocą definicji**

Przykład 3.3.7. $A = \begin{bmatrix} 3 & 5 \\ -1 & 2 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ?$

$\det A = 11 \neq 0$, zatem A jest odwracalna.

$$AA^{-1} = I \Leftrightarrow \begin{bmatrix} 3a-b & 5a+2b \\ 3c-d & 5c+2d \end{bmatrix} = I \Leftrightarrow \begin{cases} 3a-b=1 \\ 5a+2b=0 \\ 3c-d=0 \\ 5c+2d=1 \end{cases} \Rightarrow A^{-1} = \begin{bmatrix} \frac{2}{11} & -\frac{5}{11} \\ \frac{1}{11} & \frac{3}{11} \end{bmatrix}$$

2. **Metoda dopełnień algebraicznych (metoda wyznacznikowa)**

Przykład 3.3.8.

$$A = \begin{bmatrix} 2 & 7 & 3 \\ 3 & 9 & 4 \\ 1 & 5 & 3 \end{bmatrix}, \quad A^{-1} = \frac{1}{\det A} \cdot A^D = ?$$

$\det A = -3 \neq 0$, zatem A jest odwracalna. Niech $M = [M_{ij}]$ oznacza macierz minorów elementów a_{ij}. Wówczas

$$M = \begin{bmatrix} \begin{vmatrix} 9 & 4 \\ 5 & 3 \end{vmatrix} & \begin{vmatrix} 3 & 4 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 3 & 9 \\ 1 & 5 \end{vmatrix} \\ \begin{vmatrix} 7 & 3 \\ 5 & 3 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 2 & 7 \\ 1 & 5 \end{vmatrix} \\ \begin{vmatrix} 7 & 3 \\ 9 & 4 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} & \begin{vmatrix} 2 & 7 \\ 3 & 9 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 7 & 5 & 6 \\ 6 & 3 & 3 \\ 1 & -1 & -3 \end{bmatrix}$$

$$D = [D_{ij}] = [(-1)^{j+j} M_{ij}] = \begin{bmatrix} 7 & -5 & 6 \\ -6 & 3 & -3 \\ 1 & 1 & -3 \end{bmatrix},\ A^{-1} = -\tfrac{1}{3} \begin{bmatrix} 7 & -6 & 1 \\ -5 & 3 & 1 \\ 6 & -3 & -3 \end{bmatrix}$$

3. **Metoda operacji elementarnych (metoda bezwyznacznikowa)**

Niech $A \in M_n(K)$ będzie macierzą nieosobliwą.

$$[A|I] \xrightarrow[\textbf{tylko na wierszach!}]{operacje\quad elementarne} [I|A^{-1}]$$

Algorytm Gaussa: macierz nieosobliwa $\longrightarrow$ macierz trójkątna górna $\longrightarrow I$

Metoda eliminacji Gaussa to praktyczny sposób używania metody operacji elementarnych.

Przykład 3.3.9.

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & -1 & 2 \\ 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad A^{-1} = ?$$

$$[A|I] = \left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{w_3 - w_1} \left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -2 & 2 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[w_2 \text{ na koniec}]{-\frac{1}{2} w_3}$$

$$\left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 2 & 0 & 1 & 0 & 0 \end{array}\right] \xrightarrow{w_4 + w_3} \left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow{\frac{1}{2} w_4}$$

$$\left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right] \xrightarrow{w_1 - w_4} \left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & 1 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & -1 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right] \xrightarrow{w_2 + w_3}$$

$$\left[\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & 1 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right] \xrightarrow{w_1 - 2w_2} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 1 & -\frac{5}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right]$$

Zatem $A^{-1} = \begin{bmatrix} 0 & -\frac{1}{2} & 1 & -\frac{5}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} & 1 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$.

Własności macierzy odwrotnej

Twierdzenie 3.3.10. Niech $A, B \in M_n(K)$, $\alpha \in K$, $\alpha \neq 0$, $r \in \mathbb{N}$. Jeśli macierze A i B są odwracalne, to wówczas macierze $A^{-1}, A^T, AB, \alpha A, A^r$ również są odwracalne i prawdziwe są następujące równości.

i) $\det(A^{-1}) = \frac{1}{\det A}$
ii) $\left(A^{-1}\right)^{-1} = A$
iii) $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$
iv) $(\alpha A)^{-1} = \frac{1}{\alpha} \cdot A^{-1}$
v) $(AB)^{-1} = B^{-1}A^{-1}$
vi) $\left(A^r\right)^{-1} = \left(A^{-1}\right)^r$

Przykład 3.3.11. Rozwiążemy równanie macierzowe $E^4(X - 4I)^T = \frac{1}{2}E^3F^3D^{-1}D^T$ wiedząc, że $D, E, F \in M_4(\mathbb{R})$ są nieosobliwe, ponadto macierz D jest macierzą symetryczną, zaś

$$E = \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 3 & 1 & 0 \\ 4 & 2 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad F = \begin{bmatrix} \sqrt[3]{2} & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & \sqrt[3]{2} \end{bmatrix}.$$

Ponieważ D jest symetryczna, spełnia warunek $D = D^T$. Stąd $D^{-1}D^T = D^{-1}D = I$.

Ponadto F jest diagonalna, zatem $F^3 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & -8 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$.

Macierz E jest nieosobliwa, a zatem odwracalna. Obliczamy

$$\begin{aligned} E^4(X - 4I)^T &= \tfrac{1}{2}E^3F^3D^{-1}D^T = \tfrac{1}{2}E^3F^3 \\ (X - 4I)^T &= E^{-4}\tfrac{1}{2}E^3F^3 = \tfrac{1}{2}E^{-1}F^3 \\ X - 4I &= (\tfrac{1}{2}E^{-1}F^3)^T = \tfrac{1}{2}(F^3)^T(E^{-1})^T \\ X &= 4I + \tfrac{1}{2}(F^3)^T(E^{-1})^T. \end{aligned}$$

Pozostaje obliczyć E^{-1} i wykonać odpowiednie mnożenia oraz dodawania.

Twierdzenie 3.3.12. Jeśli macierz kwadratowa A jest macierzą blokowo-diagonalną postaci $A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & A_k \end{bmatrix}$, to A jest odwracalna wtedy i tylko wtedy, gdy odwracalne są macierze $A_1, A_2, \dots, A_k$. Wówczas $A^{-1} = \begin{bmatrix} A_1^{-1} & 0 & \dots & 0 \\ 0 & A_2^{-1} & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & A_k^{-1} \end{bmatrix}$.

Przykład 3.3.13. Czy $A = \begin{bmatrix} 22 & 0 & 0 & 0 \\ 0 & 8 & -6 & 0 \\ 0 & 16 & -8 & 0 \\ 0 & 0 & 0 & \pi \end{bmatrix}$ jest odwracalna? Jesli tak, oblicz A^{-1}.

Jest to macierz blokowo-diagonalna $\left[\begin{array}{c|cc|c} 22 & 0 & 0 & 0 \\ \hline 0 & 8 & -6 & 0 \\ 0 & 16 & -8 & 0 \\ \hline 0 & 0 & 0 & \pi \end{array}\right]$.

Każdy z bloków jest macierzą nieosobliwą, zatem A jest odwracalna. Obliczamy

$$\begin{bmatrix} 22 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{22} \end{bmatrix}, \qquad \begin{bmatrix} 8 & -6 \\ 16 & -8 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{1}{4} & \frac{3}{16} \\ -\frac{1}{2} & \frac{1}{4} \end{bmatrix}, \qquad \begin{bmatrix} \pi \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{\pi} \end{bmatrix}.$$

Stąd

$$A^{-1} = \begin{bmatrix} \frac{1}{22} & 0 & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{3}{16} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{\pi} \end{bmatrix}.$$

Q U I Z

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. Niech $A, B \in M_n(\mathbb{R})$. Jeśli $AB = \mathbf{0}$, to wówczas $A = \mathbf{0}$ lub $B = \mathbf{0}$.
2. $\det(A - B) = \det A - \det B$ dla dowolnych $A, B \in M_n(\mathbb{R})$.
3. $\det(A^2 - I) = \det(A - I) \cdot \det(A + I)$ dla dowolnej $A \in M_n(\mathbb{R})$.
4. Dana jest macierz $A = \begin{bmatrix} 0 & 1 & -p \\ 0 & p & 1 \\ 1 & 0 & 0 \end{bmatrix} \in M_3(\mathbb{R})$, gdzie $p \in \mathbb{R}$. Można tak dobrać $p \in \mathbb{R}$, by A była osobliwa.
5. Jeśli $A \in M_5(\mathbb{R})$ oraz $A^T = -A$, to wówczas A jest osobliwa.
6. Jeśli $A = A^{-1}$, to wówczas $A = I$ lub $A = -I$.
7. Jeśli macierze $A, B \in M_9(\mathbb{R})$ są symetryczne, to wówczas macierz $C = BA^3B$ również jest symetryczna.
8. Zbiór macierzy $M_5(\mathbb{R})$ wraz z działaniem mnożenia macierzy tworzy półgrupę nieprzemienną z jedynką.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Algebra
Wykład nr 4
Data: .....

TEMAT: *Układy równań liniowych*

4.1 Układy równań liniowych - podstawowe pojęcia

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$. Rozważmy układ m równań liniowych z n niewiadomymi $x_1, \ldots, x_n$ o współczynnikach z K.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots & \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{cases}, \quad a_{ij}, b_i \in K,\ i \in \{1, 2, \ldots, m\}, j \in \{1, 2, \ldots, n\}$$

Liczby $a_{ij} \in K$ nazywamy *współczynnikami* układu, zaś liczby b_i *wyrazami wolnymi*.

Jeśli $b_1 = \ldots = b_m = 0$ to układ równań nazywamy *jednorodnym*.

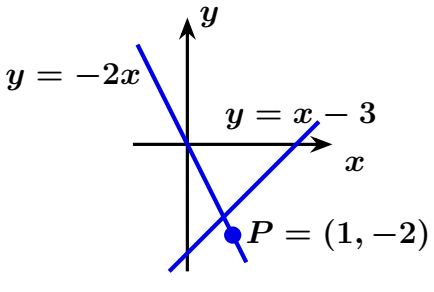
Jeśli $b_i \neq 0$ dla pewnego $i \in \{1, 2, \ldots, m\}$, to układ nazywamy *niejednorodnym*.

Rozwiązaniem układu nazywamy dowolny ciąg $(\gamma_1, \gamma_2, \ldots, \gamma_n) \in K^n$ spełniający ten układ.

Układ nie posiadający rozwiązania nazywamy *układem sprzecznym*, układ posiadający dokładnie jedno rozwiązanie - *układem oznaczonym*, zaś układ posiadający nieskończenie wiele rozwiązań - *układem nieoznaczonym*.

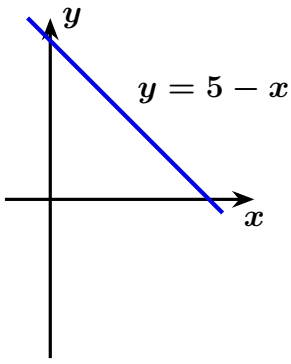
Przykład 4.1.1.

$$\begin{cases} x - y &= 3 \\ 2x + y &= 0 \end{cases}$$



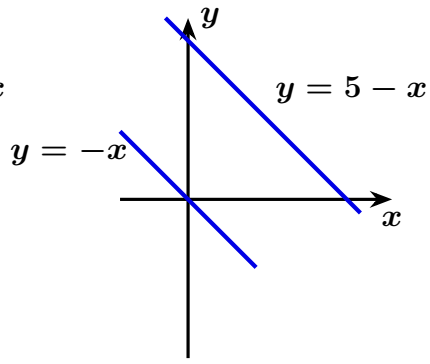
układ oznaczony

$$\begin{cases} x + y &= 5 \\ 2x + 2y &= 10 \end{cases}$$



układ nieoznaczony

$$\begin{cases} x + y &= 0 \\ 2x + 2y &= 10 \end{cases}$$



układ sprzeczny

Rozpatrywany układ równań liniowych jest równoważny z równaniem macierzowym $AX = B$, gdzie

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

macierz współczynników — kolumna niewiadomych — kolumna wyrazów wolnych

Przykład 4.1.2.

$$\begin{cases} x_1 + x_2 + 2x_4 = 1 \\ x_1 + x_2 - x_4 = 0 \\ x_1 - x_2 + x_3 - x_4 = 9 \end{cases} \Leftrightarrow \begin{bmatrix} 1 & 1 & 0 & 2 \\ 1 & 1 & 0 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 9 \end{bmatrix}$$

4.2 Układy Cramera

Definicja 4.2.1. Jeśli $m = n$ oraz macierz $A \in M_n(K)$ jest nieosobliwa, to układ równań $AX = B$ nazywamy *układem Cramera*.

Twierdzenie 4.2.2 (Cramera)**.** Układ Cramera jest układem oznaczonym.

Dowód. Ponieważ $\det A \neq 0 \Leftrightarrow \exists\ A^{-1}$, mamy $X = A^{-1}B$. $\square$

Wniosek 4.2.3. Rozwiązanie układu Cramera ma postać $X = A^{-1}B$. Można je również znaleźć za pomocą *wzorów Cramera*

$$\forall i \in \{1, 2, \dots, n\} \quad x_i = \frac{\det A_i}{\det A},$$

gdzie A_i jest macierzą powstałą z macierzy A poprzez zastąpienie i-tej kolumny kolumną wyrazów wolnych B.

Przykład 4.2.4.

Układ $\begin{cases} 3x - 2y = 5 \\ 2x + 4y = 1 \end{cases}$ jest równoważny równaniu $\begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

$W = \det \begin{bmatrix} 3 & -2 \\ 2 & 4 \end{bmatrix} = 16 \neq 0 \Rightarrow$ układ jest układem Cramera

Metoda eliminacji:

$y = \frac{3}{2}x - \frac{5}{2}$, zatem $2x + 4(\frac{3}{2}x - \frac{5}{2}) = 1$, skąd $x = \frac{11}{8}$ oraz $y = -\frac{7}{16}$

Rozwiązanie równania macierzowego:

$X = A^{-1}B$, $A^{-1} = ?$

$\det A = 16$, $D = \begin{bmatrix} 4 & -2 \\ 2 & 3 \end{bmatrix}$, $A^{-1} = \frac{1}{16}\begin{bmatrix} 4 & 2 \\ -2 & 3 \end{bmatrix}$, $X = \frac{1}{16}\begin{bmatrix} 4 & 2 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \end{bmatrix}$

$X = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{16}\begin{bmatrix} 22 \\ -7 \end{bmatrix}$, $\quad x = \frac{11}{8}$, $y = -\frac{7}{16}$

Wzory Cramera:

$W = \det A = 16 \neq 0$, $W_x = \det A_1 = \begin{vmatrix} 5 & -2 \\ 1 & 4 \end{vmatrix} = 22$, $W_y = \det A_2 = \begin{vmatrix} 3 & 5 \\ 2 & 1 \end{vmatrix} = -7$, $\quad x = \frac{W_x}{W} = \frac{22}{16}$, $y = \frac{W_y}{W} = -\frac{7}{16}$

Wniosek 4.2.5. i) Jedynym rozwiązaniem jednorodnego układu Cramera jest rozwiązanie zerowe $x_1 = \ldots = x_n = 0$.

ii) Jeśli $\det A = 0$, to układ $AX = B$ jest nieoznaczony lub sprzeczny.

Przykład 4.2.6.
Układ $\begin{cases} 0x + 0y = 0 \\ 0x + 0y = 5 \end{cases}$ jest układem sprzecznym. Ponadto $W = W_x = W_y = 0$.

Uwaga 4.2.7. i) Można wykazać, że jeśli $\det A = 0$ oraz $\det A_k \neq 0$ dla pewnego $k \in \{1, 2, \ldots, n\}$, to układ jest sprzeczny.

ii) Jeśli $\det A = 0$ oraz $\det A_i = 0$ dla każdego $i \in \{1, 2, \ldots, n\}$, to układ jest nieoznaczony lub sprzeczny.

4.3 Rząd macierzy i twierdzenie Kroneckera-Capellego

Definicja 4.3.1. Niech $A = [a_{ij}] \in M_{m\times n}(K)$. *Minorem stopnia* k macierzy A, gdzie $k \in \mathbb{N}$, $k \leq \min\{m, n\}$ nazywamy wyznacznik każdej macierzy kwadratowej otrzymanej z macierzy A poprzez skreślenie $m - k$ wierszy oraz $n - k$ kolumn.

Definicja 4.3.2. *Rzędem macierzy* $A \in M_{m\times n}(K)$, $A \neq \mathbf{0}$ nazywamy największy stopień jej niezerowego minora. Liczbę tę oznaczamy $r(A)$ lub $\operatorname{rank}(A)$. Ponadto przyjmujemy, że rząd dowolnej macierzy zerowej jest równy zero.

$$\mathbf{0 \leq r(A) \leq \min\{m, n\}}$$

Przykład 4.3.3.

$A = \begin{bmatrix} 5 & -2 & 1 & 3 \\ 0 & 0 & 3 & -1 \\ 10 & 2 & 3 & 9 \end{bmatrix}$, $r(A) \leq 3$, $\begin{vmatrix} 5 & 1 & 3 \\ 0 & 3 & -1 \\ 10 & 3 & 9 \end{vmatrix} = -50 \neq 0$, $r(A) = 3$

$B = \begin{bmatrix} 1 & 3 \\ 4 & 12 \\ 3 & 9 \\ 2 & 6 \end{bmatrix}$, $r(A) \leq 2$, $\begin{vmatrix} 1 & 3 \\ 4 & 12 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 3 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = \begin{vmatrix} 4 & 12 \\ 3 & 9 \end{vmatrix} = 0$

$\begin{vmatrix} 4 & 12 \\ 2 & 6 \end{vmatrix} = \begin{vmatrix} 3 & 9 \\ 2 & 6 \end{vmatrix} = 0$, $r(B) = 1$, Można zauważyć, że $k_2 = 3k_1$.

$C = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 7 \end{bmatrix}$, $r(A) \leq 2$, $\begin{vmatrix} 0 & 1 \\ 0 & 7 \end{vmatrix} = 0$, ale $\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} \neq 0$, $r(C) = 2$

$D = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$, $r(A) \leq 3$, $w_2 = 2w_1$, $w_3 = 3w_1$, $r(D) = 1$

Twierdzenie 4.3.4 (Własności rzędu macierzy)**.** Niech $A \in M_{m\times n}(K)$. Wówczas

i) $r(A) = r(A^T)$,

ii) operacje elementarne na wierszach (kolumnach) macierzy nie zmieniają jej rzędu.

Dowód. Teza wynika z własności wyznaczników. $\square$.

Definicja 4.3.5. Macierz $A \in M_{m\times n}(K)$ nazywamy *macierzą schodkową*, gdy pierwsze niezerowe elementy (tzw. schodki) w kolejnych niezerowych wierszach tej macierzy znajdują się w kolumnach o rosnących numerach. Ponadto przyjmujemy, że macierz o jednym niezerowym wierszu, a także dowolna macierz zerowa, są macierzami schodkowymi.

Przykład 4.3.6.

$A = \begin{bmatrix} 2 & 2 & 2 & 6 & 3 & 0 \\ 0 & 6 & 6 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 9 \end{bmatrix}$ 4 schodki

$B = \begin{bmatrix} 1 & 2 & 9 \\ 0 & 1 & -7 \\ 0 & 0 & 0 \\ 0 & 0 & 15 \end{bmatrix}$ 3 schodki

$C = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 2 & 2 & 1 \\ 0 & 0 & 6 & 1 \\ 0 & 0 & 7 & 0 \end{bmatrix}$ to nie postać schodkowa

Twierdzenie 4.3.7. Rząd macierzy schodkowej jest równy liczbie jej niezerowych wierszy (tj. schodków).

Dowód. Skreślając wiersze zerowe i kolumny nie wyznaczające schodków, otrzymamy macierz trójkątna górną nieosobliwą. □

Przykład 4.3.8.
$$D = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 5 & 7 & 10 \\ 4 & 6 & 8 & 9 & 3 \\ 4 & 7 & 7 & 8 & 3 \end{bmatrix} \xrightarrow[w_4-4w_1]{\substack{w_2-2w_1 \\ w_3-4w_1}} \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & -2 & -4 & -7 & -17 \\ 0 & -1 & -5 & -8 & -17 \end{bmatrix} \xrightarrow[w_4+w_2]{w_3+2w_2}$$
$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & -6 & -9 & -17 \\ 0 & 0 & -6 & -9 & -17 \end{bmatrix}, \quad r(D) = 3$$

Rozważmy układ m równań liniowych z n niewiadomymi $x_1, \ldots, x_n$ o współczynnikach z K postaci $AX = B$.

Macierz $U = [A|B] = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{bmatrix} \in M_{m\times(n+1)}(K)$ nazywamy *macierzą uzupełnioną* układu $AX = B$.

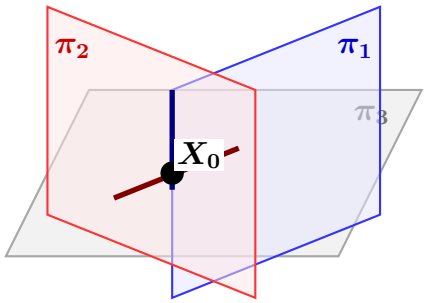
Twierdzenie 4.3.9 (Kroneckera-Capellego)**.** Układ równań liniowych $AX = B$ ma rozwiązanie wtedy i tylko wtedy, gdy $r(A) = r(U)$.

Wniosek 4.3.10.
i) Gdy $r(A) \neq r(U)$, układ jest sprzeczny.

ii) Gdy $r(A) = r(U) = n$, układ jest oznaczony.

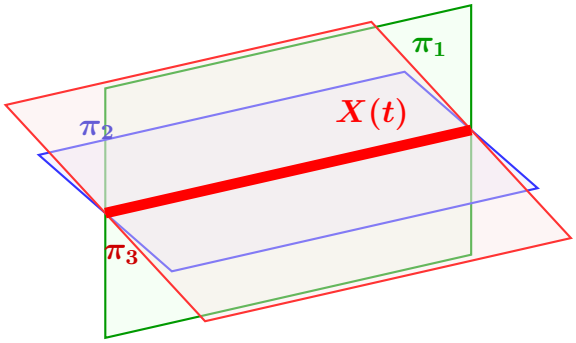
iii) Gdy $r(A) = r(U) = r < n$, układ ma nieskończenie wiele rozwiązań zależnych od $n - r$ parametrów.

$r(A) = r(U) = 3 = n$



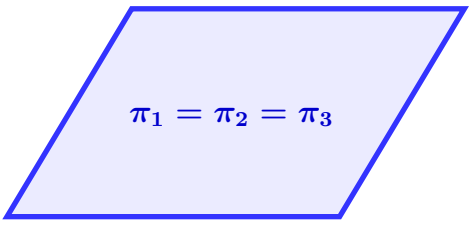
Rys. 1. Układ oznaczony:
Jedno rozwiązanie (punkt X_0).

$r(A) = r(U) = 2 < n$



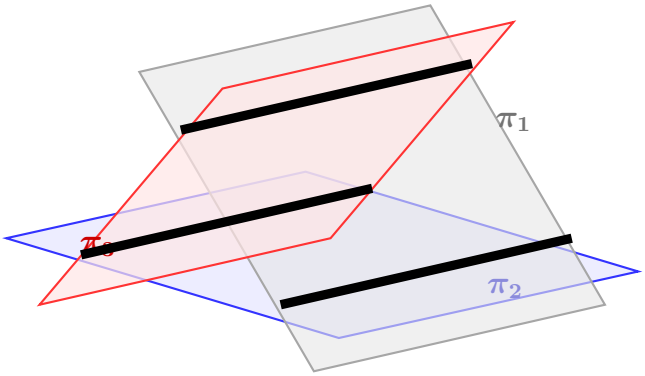
Rys. 2. Układ nieoznaczony (1 par.):
Rozwiązaniem jest wspólna prosta.

$r(A) = r(U) = 1 < n$



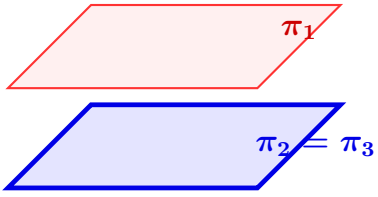
Rys. 3. Układ nieoznaczony (2 par.):
Płaszczyzny pokrywają się.

$r(A) = 2 \neq r(U) = 3$



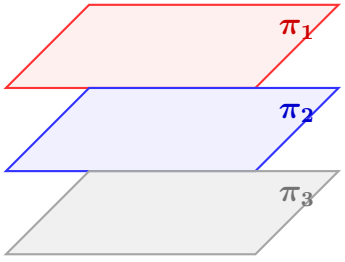
Rys. 4. Układ sprzeczny („Tunel”):
Przecięcia parami, brak punktu wspólnego.

$r(A) = 1 \neq r(U) = 2$



Rys. 5. Układ sprzeczny:
Dwie pokrywające się, trzecia równoległa.

$r(A) = 1 \neq r(U) = 2$



Rys. 6. Układ sprzeczny:
Trzy płaszczyzny równoległe rozłączne.

Uwaga 4.3.11. Dla n=3, przypadek $r(A) = 1$ oraz $r(U) = 3$ nie jest możliwy. Macierz uzupełniona $U = [A|B]$ powstaje poprzez dopisanie do macierzy A jednej kolumny (wyrazów wolnych), co może zwiększyć rząd macierzy *maksymalnie o 1.*
Rząd $r(A)$ określa liczbę niezależnych [a]

[a]Wielkość $r(A)$ określa wymiar przestrzeni rozpiętej przez wektory normalne płaszczyzn.

- $r(A) = 1$: Wszystkie wektory normalne są współliniowe. Oznacza to, że wszystkie trzy płaszczyzny są do siebie równoległe (mogą się rozłącznie mijać lub na siebie nakładać).
- $r(A) = 2$: Wektory normalne rozpinają przestrzeń dwuwymiarową (płaszczyznę). Trzy płaszczyzny przecinają się wzdłuż jednej prostej niczym skrzydła wiatraka (pęk płaszczyzn, jak na Rys. 2). W szczególności możliwe jest, że dwie płaszczyzny pokrywają się, a trzecia jest do nich nierównoległa i przecina je pod skosem.
- $r(A) = 3$: Wektory normalne są liniowo niezależne i rozpinają przestrzeń $\mathbb{R}^3$. Płaszczyzny przecinają się w jednym punkcie (jak narożnik pokoju, patrz Rys. 1).

Jeśli układ jest sprzeczny, to $r(U) = r(A) + 1$.

Przykład 4.3.12.

$$\begin{cases} x - y + 3z &= 5 \\ 4x + y + 7z &= 8 \\ 5x + 2y + 8z &= 4 \end{cases}$$

$$U = [A|B] = \left[\begin{array}{ccc|c} 1 & -1 & 3 & 5 \\ 4 & 1 & 7 & 8 \\ 5 & 2 & 8 & 4 \end{array}\right],\ r(A) \leq 3,\ r(U) \leq 3$$

$$U \xrightarrow[w_3-5w_1]{w_2-4w_1} \left[\begin{array}{ccc|c} 1 & -1 & 3 & 5 \\ 0 & 5 & -5 & -12 \\ 0 & 7 & -7 & -21 \end{array}\right] \xrightarrow[\frac{1}{7}w_3]{\frac{1}{5}w_2} \left[\begin{array}{ccc|c} 1 & -1 & 3 & 5 \\ 0 & 1 & -1 & -\frac{12}{5} \\ 0 & 1 & -1 & -3 \end{array}\right] \xrightarrow{w_3-w_2} \left[\begin{array}{ccc|c} 1 & -1 & 3 & 5 \\ 0 & 1 & -1 & -\frac{12}{5} \\ 0 & 0 & 0 & -\frac{3}{5} \end{array}\right]$$

$r(U) = 3 \neq r(A) = 2 \Rightarrow$ układ sprzeczny, brak rozwiązań

Algorytm rozwiązywania układów równań liniowych

1. Jeśli $r(A) < r(U)$, układ jest sprzeczny.

2. Niech $r(A) = r(U) = r$. Istnieje niezerowy minor M stopnia r macierzy A (będący również minorem macierzy U). Wówczas układ ma przynajmniej jedno rozwiązanie.

2a. Skreślamy $m - r$ wierszy macierzy U (równań układu), które nie tworzą M. Jeśli $r = n$, to otrzymujemy układ Cramera.

2b. Jeśli $r < n$, to $n-r$ niewiadomych, których współczynniki nie tworzą M, przenosimy na stronę wyrazów wolnych i traktujemy je dalej jako parametry (zmienne niezależne). Układ ma nieskończenie wiele rozwiązań zależnych od $n - r$ parametrów. Mówiąc dokładnej r spośród niewiadomych oznaczanych $x'_1, \ldots, x'_r$ zależy od pozostałych $n - r$ niewiadomych $x'_{r+1}, \ldots, x'_n$.

Przykład 4.3.13.

$$\begin{cases} 2x - y + 3z &= 7 \\ 3x + 2y - 5z &= 4 \\ 4x + 5y - 13z &= 1 \end{cases}$$

$U = [A|B] = \left[\begin{array}{ccc|c} 2 & -1 & 3 & 7 \\ 3 & 2 & -5 & 4 \\ 4 & 5 & -13 & 1 \end{array}\right]$, $r(A) \leq 3$, $r(U) \leq 3$

$$U \xrightarrow[\substack{zmienne\ \mathbf{y}\ \ \mathbf{x}\ \ \mathbf{z}}]{k_1 \leftrightarrow k_2} \left[\begin{array}{ccc|c} -1 & 2 & 3 & 7 \\ 2 & 3 & -5 & 4 \\ 5 & 4 & -13 & 1 \end{array}\right] \xrightarrow[w_3+5w_1]{w_2+2w_1} \left[\begin{array}{ccc|c} -1 & 2 & 3 & 7 \\ 0 & 7 & 1 & 18 \\ 0 & 14 & 2 & 36 \end{array}\right] \to \left[\begin{array}{ccc|c} -1 & 2 & 3 & 7 \\ 0 & 7 & 1 & 18 \end{array}\right]$$

$\left|\begin{array}{cc} -1 & 2 \\ 0 & 7 \end{array}\right| \neq 0$ minor niezerowy

$$\begin{cases} -y + 2x &= 7 - 3z \\ 7x &= 18 - z \end{cases} \text{ lub } \begin{bmatrix} -1 & 2 \\ 0 & 7 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} 7 - 3z \\ 18 - z \end{bmatrix}$$

układ nieoznaczony, rozwiązania $\begin{cases} x = \frac{18}{7} - \frac{1}{7}z \\ y = -\frac{13}{7} + \frac{20}{7}z \\ z \in \mathbb{R} \end{cases}$

4.4 Metoda eliminacji Gaussa

Dwa układy równań liniowych nazywamy *równoważnymi*, gdy posiadają ten sam zbiór rozwiązań.

Operacje elementarne na wierszach macierzy U, skreślenie wiersza złożonego z samych zer lub skreślenie jednego z dwu wierszy proporcjonalnych (równych) nie zmieniają rzędu macierzy U, zatem prowadzą one do równoważnego układu równań.

Ewentualne przestawienie kolumn macierzy A prowadzi do zmiany kolejności występowania niewiadomych.

Dokonując równoważnych przekształceń układu równań $AX = B$, sprowadzamy macierz $U = [A|B]$ do postaci

$$[A'|B'] = \left[\begin{array}{ccc|ccc|c} & & & s_{1,r+1} & \dots & s_{1,n} & z_1 \\ & I_r & & & \dots & & \dots \\ & & & s_{r,r+1} & \dots & s_{r,n} & z_r \\ \hline 0 & \dots & 0 & 0 & \dots & 0 & z_{r+1} \end{array}\right].$$

Jeśli $z_{r+1} \neq 0$, to układ jest sprzeczny.
Jeśli ostatni wiersz się nie pojawi oraz $r = n$, układ jest oznaczony i jego rozwiązaniem jest $x_i = z_i$, dla $i \in \{1, 2, \dots, r\}$.
Jeśli ostatni wiersz się nie pojawi oraz $r < n$, układ jest nieoznaczony. Rozwiązania mają postać

$$\begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_r \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_r \end{bmatrix} - \begin{bmatrix} s_{1,r+1} & \dots & s_{1,n} \\ & \dots & \\ s_{r,r+1} & \dots & s_{r,n} \end{bmatrix} \begin{bmatrix} x'_{r+1} \\ x'_{r+2} \\ \vdots \\ x'_n \end{bmatrix}.$$

Przykład 4.4.1.

$$\begin{cases} x_1 + 2x_2 - x_3 - x_4 &= 1 \\ x_1 + x_2 + x_3 + 3x_4 &= 2 \\ 3x_1 + 5x_2 - x_3 + x_4 &= 3 \end{cases}$$

$$U = \left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 1 \\ 1 & 1 & 1 & 3 & 2 \\ 3 & 5 & -1 & 1 & 3 \end{array}\right] \xrightarrow[w_3-3w_1]{w_2-w_1} \left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 1 \\ 0 & -1 & 2 & 4 & 1 \\ 0 & -1 & 2 & 4 & 0 \end{array}\right] \xrightarrow{w_3-w_2} \left[\begin{array}{cccc|c} 1 & 2 & -1 & -1 & 1 \\ 0 & -1 & 2 & 4 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right]$$

Równanie $0x_1 + 0x_2 + 0x_3 + 0x_4 = -1$ jest sprzeczne, zatem układ jest sprzeczny.

Przykład 4.4.2.

$$\begin{cases} 2x + 4y + 9z - 5t + 6u &= 13 \\ x + 2y + 4z - 4t + 2u &= 5 \\ 3x + 6y + 7z - 11t + 2u &= 18 \\ 4x + 8y + 19z - 15t + 11u &= 20 \\ -\frac{1}{2}x - y + z + 3t + 2u &= -\frac{5}{2} \end{cases}$$

$$U = \left[\begin{array}{ccccc|c} 2 & 4 & 9 & -5 & 6 & 13 \\ 1 & 2 & 4 & -4 & 2 & 5 \\ 3 & 6 & 7 & -11 & 2 & 18 \\ 4 & 8 & 19 & -15 & 11 & 20 \\ -\frac{1}{2} & -1 & 1 & 3 & 2 & -\frac{5}{2} \end{array}\right] \xrightarrow[-2\cdot w_5]{w_2 \leftrightarrow w_1} \left[\begin{array}{ccccc|c} 1 & 2 & 4 & -4 & 2 & 5 \\ 2 & 4 & 9 & -5 & 6 & 13 \\ 3 & 6 & 7 & -11 & 2 & 18 \\ 4 & 8 & 19 & -15 & 11 & 20 \\ 1 & 2 & -2 & -6 & -4 & 5 \end{array}\right] \xrightarrow[\substack{w_3-3w_1 \\ w_4-4w_1 \\ w_5-w_1}]{w_2-2w_1}$$

$$\left[\begin{array}{ccccc|c} 1 & 2 & 4 & -4 & 2 & 5 \\ 0 & 0 & 1 & 3 & 2 & 3 \\ 0 & 0 & -5 & 1 & -4 & 3 \\ 0 & 0 & 3 & 1 & 3 & 0 \\ 0 & 0 & -6 & -2 & -6 & 0 \end{array}\right] \xrightarrow[\substack{k_2 \text{ za } k_5 \\ \text{zmienne } \mathbf{x\,z\,t\,u\,y}}]{} \left[\begin{array}{ccccc|c} 1 & 4 & -4 & 2 & 2 & 5 \\ 0 & 1 & 3 & 2 & 0 & 3 \\ 0 & -5 & 1 & -4 & 0 & 3 \\ 0 & 3 & 1 & 3 & 0 & 0 \end{array}\right] \xrightarrow[w_4-3w_2]{w_3+5w_2}$$

$$\left[\begin{array}{ccccc|c} 1 & 4 & -4 & 2 & 2 & 5 \\ 0 & 1 & 3 & 2 & 0 & 3 \\ 0 & 0 & 16 & 6 & 0 & 18 \\ 0 & 0 & -8 & -3 & 0 & -9 \end{array}\right] \xrightarrow{(-1)\cdot w_4} \left[\begin{array}{ccccc|c} 1 & 4 & -4 & 2 & 2 & 5 \\ 0 & 1 & 3 & 2 & 0 & 3 \\ 0 & 0 & 8 & 3 & 0 & 9 \end{array}\right] \xrightarrow[\text{zmienne } \mathbf{x\ z\ u\ t\ y}]{k_3 \leftrightarrow k_4}$$

$$\left[\begin{array}{ccc|cc|c} 1 & 4 & 2 & -4 & 2 & 5 \\ 0 & 1 & 2 & 3 & 0 & 3 \\ 0 & 0 & 3 & 8 & 0 & 9 \end{array}\right] \xrightarrow{\frac{1}{3}w_3} \left[\begin{array}{ccc|cc|c} 1 & 4 & 2 & -4 & 2 & 5 \\ 0 & 1 & 2 & 3 & 0 & 3 \\ 0 & 0 & 1 & \frac{8}{3} & 0 & 3 \end{array}\right] \xrightarrow[w_2-2w_3]{w_1-2w_3} \left[\begin{array}{ccc|cc|c} 1 & 4 & 0 & -\frac{28}{3} & 2 & -1 \\ 0 & 1 & 0 & -\frac{7}{3} & 0 & -3 \\ 0 & 0 & 1 & \frac{8}{3} & 0 & 3 \end{array}\right]$$

$$\xrightarrow{w_1-4w_2} \left[\begin{array}{ccc|cc|c} 1 & 0 & 0 & 0 & 2 & 11 \\ 0 & 1 & 0 & -\frac{7}{3} & 0 & -3 \\ 0 & 0 & 1 & \frac{8}{3} & 0 & 3 \end{array}\right]$$

równoważny układ $\begin{cases} x + 2y = 11 \\ z - \frac{7}{3}t = -3 \\ u + \frac{8}{3}t = 3 \end{cases}$ rozwiązania $\begin{cases} x = 11 - 2y \\ z = -3 + \frac{7}{3}t \\ u = 3 - \frac{8}{3}t \\ y, t \in \mathbb{R} \end{cases}$

$r(A) = r(U) = 3 < n = 5$ układ nieoznaczony, nieskończenie wiele rozwiązań zależnych od $5 - 3 = 2$ parametrów

Uwaga 4.4.3. Podział niewiadomych na zmienne zależne i parametry nie jest jednoznaczny, lecz nie jest dowolny.

Przykład 4.4.4. Wskaż zbiory niewiadomych, które mogą być parametrami w rozwiązaniu układu równań.

$$\begin{cases} x - 3y + z - 2s + t = -5 \\ 2x - 6y - 4s + t = -10 \\ 2z + t = 0 \end{cases}$$

$$U = [A|B] = \left[\begin{array}{ccccc|c} 1 & -3 & 1 & -2 & 1 & -5 \\ 2 & -6 & 0 & -4 & 1 & -10 \\ 0 & 0 & 2 & 0 & 1 & 0 \end{array}\right] \xrightarrow{w_2-2w_1} \left[\begin{array}{ccccc|c} 1 & -3 & 1 & -2 & 1 & -5 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{ccccc|c} 1 & -3 & 1 & -2 & 1 & -5 \\ 0 & 0 & 2 & 0 & 1 & 0 \end{array}\right] \qquad r := r(A) = r(U) = 2,\ n = 5,\ n - r = 3 \text{ parametry}$$

Trzy zmienne spośród pięciu można wybrać na $\binom{5}{3} = 10$ sposobów.
Nie wszystkie wybory są dozwolone.

minory niezerowe	parametry
k_1, k_3	$\{y, s, t\}$
k_1, k_5	$\{y, z, s\}$
k_2, k_3	$\{x, s, t\}$
k_2, k_5	$\{x, z, s\}$
k_4, k_3	$\{x, y, t\}$
k_4, k_5	$\{x, y, z\}$
k_3, k_5	$\{x, y, s\}$

Uwaga 4.4.5. W przypadku gdy układ równań $AX = B$ jest układem Cramera, wykonując operacje elementarne **na wierszach** macierzy uzupełnionej $U = [A|B]$, sprowadzamy tę macierz do postaci $[I|X]$, gdzie X jest poszukiwanym rozwiązaniem układu.

$$[A|B] \xrightarrow[\textbf{na}\quad\textbf{wierszach}]{operacje\quad elementarne} [I|X]$$

Przykład 4.4.6.

$$\begin{cases} x+y+z &= 3 \\ 2x-y-z &= 0 \\ x+3y+z &= 5 \end{cases}$$

$$[A|B] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 2 & -1 & -1 & 0 \\ 1 & 3 & 1 & 5 \end{array}\right] \xrightarrow[w_3-w_1]{w_2-2w_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -3 & -3 & -6 \\ 0 & 2 & 0 & 2 \end{array}\right] \xrightarrow[\frac{1}{2}w_3]{-\frac{1}{3}w_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow{w_3-w_2}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & -1 & -1 \end{array}\right] \xrightarrow[w_2+w_3]{w_1+w_3} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & -1 \end{array}\right] \xrightarrow[(-1)\cdot w_3]{w_1-w_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

układ oznaczony, jedyne rozwiązanie $x = 1$, $y = 1$, $z = 1$

Przykład 4.4.7. Określ ilość rozwiązań układu w zależności od parametru $p \in \mathbb{R}$.

$$\begin{cases} x_1 + 2x_2 + (p+1)x_3 + (p+3)x_4 &= 4 \\ 2x_1 + 4x_2 + (2p+4)x_3 + (4p+6)x_4 &= p+15 \\ (p+1)x_1 + 2x_2 + (p+1)x_3 + (p+3)x_4 &= p+4 \\ x_1 + 2x_2 + (p+3)x_3 + (2p+2)x_4 &= 11 \end{cases}$$

$$U = \left[\begin{array}{cccc|c} 1 & 2 & p+1 & p+3 & 4 \\ 2 & 4 & 2p+4 & 4p+6 & p+15 \\ p+1 & 2 & p+1 & p+3 & p+4 \\ 1 & 2 & p+3 & 2p+2 & 11 \end{array}\right] \xrightarrow[\substack{w_3-w_2 \\ w_4-w_1}]{w_2-2w_1} \left[\begin{array}{cccc|c} 1 & 2 & p+1 & p+3 & 4 \\ 0 & 0 & 2 & 2p & p+7 \\ p & 0 & 0 & 0 & p \\ 0 & 0 & 2 & p+1 & 7 \end{array}\right]$$

$$\xrightarrow[\substack{zmienne\ \mathbf{x_2\ x_3\ x_1\ x_4}}]{k_1\ za\ k_3} \left[\begin{array}{cccc|c} 2 & p+1 & 1 & p+3 & 4 \\ 0 & 2 & 0 & 2p & p+7 \\ 0 & 0 & p & 0 & p \\ 0 & 2 & 0 & p+1 & 7 \end{array}\right] \xrightarrow{w_4-w_2} \left[\begin{array}{cccc|c} 2 & p+1 & 1 & p+3 & 4 \\ 0 & 2 & 0 & 2p & p+7 \\ 0 & 0 & p & 0 & p \\ 0 & 0 & 0 & 1-p & -p \end{array}\right]$$

Wnioski:

Dla $p \in \mathbb{R} \setminus \{0, 1\}$ mamy $r(A) = r(U) = n = 4$, zatem układ jest oznaczony.

Dla $p = 0$ ostatnia macierz przyjmuje postać $\left[\begin{array}{cccc|c} 2 & 1 & 1 & 3 & 4 \\ 0 & 2 & 0 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right]$,
skąd otrzymujemy, że $r(A) = r(U) = 3 < n = 4$.
Zatem układ jest nieoznaczony.
Posiada nieskończenie wiele rozwiązań zależnych od jednego parametru.

Dla $p = 1$ ostatnia macierz przyjmuje postać $\left[\begin{array}{cccc|c} 2 & 1 & 1 & 4 & 4 \\ 0 & 2 & 0 & 0 & 8 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right]$,
skąd otrzymujemy, że układ jest sprzeczny.

QUIZ

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. Układ równań $\begin{cases} 2x + 3y + 2z = 1 \\ 3x + 4y + 2z = 2 \\ 4x + 2y + 3z = 3 \end{cases}$ jest układem Cramera.
2. Dla dowolnego $p \in \mathbb{R}$ układ równań $\begin{bmatrix} p & 1 & 2 \\ 3 & 1 & 1 \\ 1 & 4 & p \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ jest układem oznaczonym.
3. Dla dowolnego $p \in \mathbb{R}$ układ równań $\begin{bmatrix} p & 1 \\ 3 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ p \end{bmatrix}$ jest układem sprzecznym.
4. Dana jest macierz $A = \begin{bmatrix} 0 & 1 & -p \\ 0 & p & 1 \\ 1 & 0 & 0 \end{bmatrix} \in M_3(\mathbb{R})$, gdzie $p \in \mathbb{R}$. Można tak dobrać $p \in \mathbb{R}$, by $r(A) = 3$.
5. Istnieje macierz $A \in M_5(\mathbb{R})$ taka, że $r(A) = 4$ oraz $\det A = 1$.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

TEMAT: *Geometria analityczna w* $\mathbb{R}^3$

5.1 Wektory w przestrzeni

$\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$ możemy interpretować jako:

- zbiór punktów $P = (x, y, z)$, gdzie x, y, z to współrzędne punktu
- zbiór wektorów zaczepionych w początku układu współrzędnych $\vec{a} = \overrightarrow{OP} = [x, y, z]$, gdzie x, y, z to współrzędne wektora
- zbiór wektorów swobodnych $\vec{a}$. Wektor swobodny to zbiór wszystkich wektorów zaczepionych (w różnych punktach) mających ten sam zwrot, kierunek i długość.

Oznaczamy przez $\vec{\mathbf{0}} = [0, 0, 0]$ wektor zerowy.

Działania na wektorach

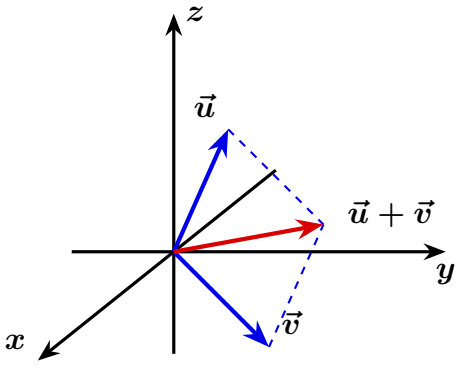
Niech $\vec{u} = [u_x, u_y, u_z]$, $\vec{v} = [v_x, v_y, v_z]$, $\lambda \in \mathbb{R}$.

$$\vec{u} + \vec{v} = [u_x + v_x, u_y + v_y, u_z + v_z] \quad \text{suma wektorów}$$

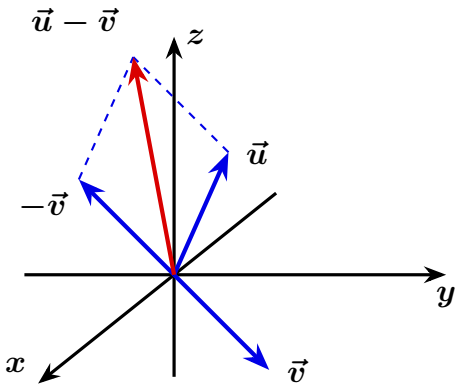
$$\lambda \cdot \vec{u} = [\lambda u_x, \lambda u_y, \lambda u_z] \quad \text{iloczyn wektora przez skalar}$$

$$-\vec{u} = [-u_x, -u_y, -u_z] \quad \text{wektor przeciwny do } \vec{u}$$

$$\vec{u} - \vec{v} = \vec{u} + (-\vec{v}) = \vec{u} + (-1) \cdot \vec{v} \quad \text{różnica wektorów}$$



dodawanie wektorów



odejmowanie wektorów

Długość wektora

Oznaczamy przez $|\vec{u}|$ długość wektora $\vec{u}$. Jeśli $P_1 = (x_1, y_1, z_1), P_2 = (x_2, y_2, z_2)$, to wówczas

$$\overrightarrow{P_1P_2} = [x_2 - x_1, y_2 - y_1, z_2 - z_1], \quad |\overrightarrow{P_1P_2}| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Wektor długości 1 nazywamy *wersorem*.

Oznaczamy przez $\hat{i} = [1, 0, 0], \hat{j} = [0, 1, 0], \hat{k} = [0, 0, 1]$ wersory osi układu współrzędnych.

$$\begin{aligned} \vec{\boldsymbol{u}} &= [\boldsymbol{u_x}, \boldsymbol{u_y}, \boldsymbol{u_z}] \\ &= [\boldsymbol{u_x}, \mathbf{0}, \mathbf{0}] + [\mathbf{0}, \boldsymbol{u_y}, \mathbf{0}] + [\mathbf{0}, \mathbf{0}, \boldsymbol{u_z}] \\ &= \boldsymbol{u_x} \cdot [\mathbf{1}, \mathbf{0}, \mathbf{0}] + \boldsymbol{u_y} \cdot [\mathbf{0}, \mathbf{1}, \mathbf{0}] + \boldsymbol{u_z} \cdot [\mathbf{0}, \mathbf{0}, \mathbf{1}] \\ &= \boldsymbol{u_x}\hat{\boldsymbol{i}} + \boldsymbol{u_y}\hat{\boldsymbol{j}} + \boldsymbol{u_z}\hat{\boldsymbol{k}} \end{aligned} \qquad |\vec{\boldsymbol{u}}| = \sqrt{\boldsymbol{u_x^2} + \boldsymbol{u_y^2} + \boldsymbol{u_z^2}}$$

Uwaga 5.1.1 (Własności długości)**.**

$|\vec{u}| = 0 \Leftrightarrow \vec{u} = \vec{\mathbf{0}}$ $\qquad |\lambda \cdot \vec{u}| = |\lambda| \cdot |\vec{u}|$ $\qquad |\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$

Wersorem niezerowego wektora $\vec{u}$ nazywamy wersor o tym samym kierunku i zwrocie co $\vec{u}$. Oznaczamy go $\hat{u}$.

$$\hat{\boldsymbol{u}} = \tfrac{\mathbf{1}}{|\vec{\boldsymbol{u}}|} \cdot \vec{\boldsymbol{u}}$$

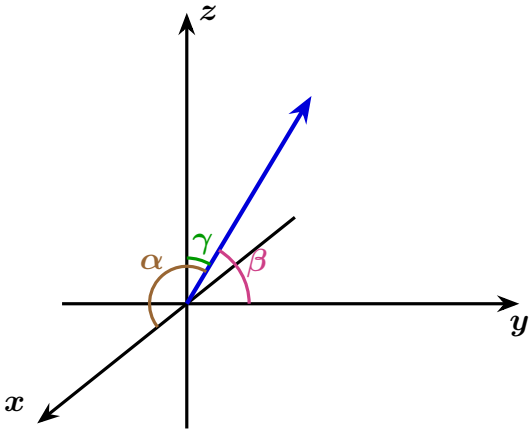
Istotnie, jeśli $\vec{u} = [u_x, u_y, u_z]$, to $\hat{u} = \left[\frac{u_x}{|\vec{u}|}, \frac{u_y}{|\vec{u}|}, \frac{u_z}{|\vec{u}|}\right]$ oraz

$$|\hat{u}| = \sqrt{\frac{u_x^2}{|\vec{u}|^2} + \frac{u_y^2}{|\vec{u}|^2} + \frac{u_y^2}{|\vec{u}|^2}} = \frac{1}{|\vec{u}|^2}\sqrt{u_x^2 + u_y^2 + u_z^2} = 1.$$

Jeśli wektor $\vec{u} = [u_x, u_y, u_z]$ tworzy z dodatnimi półosiami układu współrzędnych kąty α, β, γ, odpowiednio, to kąty te nazywamy *kątami kierunkowymi*, zaś współrzędne wersora $\hat{u}$, czyli liczby

$$\cos\boldsymbol{\alpha} = \frac{\boldsymbol{u_x}}{|\vec{\boldsymbol{u}}|}, \quad \cos\boldsymbol{\beta} = \frac{\boldsymbol{u_y}}{|\vec{\boldsymbol{u}}|}, \quad \cos\boldsymbol{\gamma} = \frac{\boldsymbol{u_z}}{|\vec{\boldsymbol{u}}|}$$

nazywamy *cosinusami kierunkowymi* wektora $\vec{u}$.



kąty kierunkowe

Iloczyn skalarny

Oznaczamy przez $\measuredangle(\vec{u},\vec{v})$ kąt między wektorami $\vec{u},\vec{v}$,.
Przyjmujemy, że jego miara należy do przedziału $[0,\pi]$.

Definicja 5.1.2. *Iloczynem skalarnym* dwóch niezerowych wektorów $\vec{u},\vec{v}$ nazywamy liczbę $\boldsymbol{|\vec{u}|\cdot|\vec{v}|\cdot\cos\measuredangle(\vec{u},\vec{v})}$. Oznaczamy ją symbolem $\boldsymbol{\vec{u}\circ\vec{v}}$. Gdy jeden z wektorów jest zerowy, przyjmujemy, że iloczyn jest równy 0.

Twierdzenie 5.1.3. Jeśli $\vec{u}=[u_x,u_y,u_z]$, $\vec{v}=[v_x,v_y,v_z]$, to wówczas $\vec{u}\circ\vec{v}=u_xv_x+u_yv_y+u_zv_z$.

Przykład 5.1.4. $[1,4,0]\circ[2,-1,1]=1\cdot2+4\cdot(-1)+0\cdot1=-2<0$
Cosinus kąta między wektorami jest ujemny, zatem kąt jest rozwarty.

Twierdzenie 5.1.5 (Własności iloczynu skalarnego)**.** Dla dowolnego $\lambda\in\mathbb{R}$ i dla dowolnych wektorów $\vec{u},\vec{v},\vec{w}$ prawdziwe są następujące równości.

i) $\vec{u}\circ\vec{v}=\vec{v}\circ\vec{u}$

ii) $\vec{u}\circ\vec{u}=|\vec{u}|^2$

iii) $(\lambda\vec{u})\circ\vec{v}=\lambda(\vec{u}\circ\vec{v})$

iv) $(\vec{u}+\vec{v})\circ\vec{w}=(\vec{u}\circ\vec{w})+(\vec{v}\circ\vec{w})$

v) $|\vec{u}\circ\vec{v}|\leq|\vec{u}|\cdot|\vec{v}|$

vi) $\vec{u}\neq\vec{\mathbf{0}},\vec{v}\neq\vec{\mathbf{0}}$, $\vec{u}\circ\vec{v}=0\ \Rightarrow\ \vec{u}\perp\vec{v}$

Przykład 5.1.6. Dane są wektory $\vec{a},\vec{b}$ w przestrzeni $\mathbb{R}^3$ takie, że $\vec{a}$ jest rzutem prostokątnym wektora $\vec{c}=[\frac{1}{\sqrt{2}},3,4]$ na oś Ox, zaś $\vec{b}$ wersorem takim, że $2\vec{a}-\vec{b}\perp\vec{b}$.
Wyznaczymy miarę kąta α pomiędzy wektorami $\vec{a}$ i $\vec{b}$.

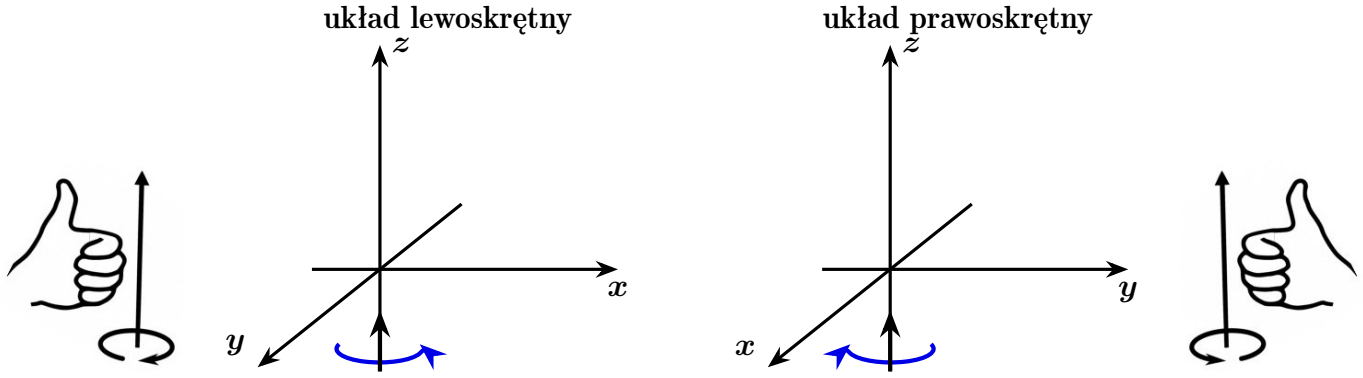
Zauważmy, że $|\vec{b}|=1$ oraz $\vec{a}=[\frac{1}{\sqrt{2}},0,0]$ i $|\vec{a}|=\frac{1}{\sqrt{2}}$. Ponadto

$$0=(2\vec{a}-\vec{b})\circ\vec{b}=2\vec{a}\circ\vec{b}-\vec{a}\circ\vec{b}=2|\vec{a}|\cdot|\vec{b}|\cdot\cos\alpha-|\vec{b}|^2=2\cdot\frac{1}{\sqrt{2}}\cos\alpha-1.$$

Stąd $\cos\alpha=\frac{\sqrt{2}}{2}$, zatem $\alpha=\frac{\pi}{4}$.

Układ współrzędnych

Układ współrzędnych w $\mathbb{R}^3$ - trójka wzajemnie prostopadłych prostych, przecinających się w jednym punkcie, zwanym początkiem układu współrzędnych.



Definicja 5.1.7. Uporządkowana trójka wektorów $\left(\vec{u}, \vec{v}, \vec{w}\right)$ ma *orientację zgodną* z orientacją układu współrzędnych, jeśli

$$\begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} > 0.$$

Iloczyn wektorowy

Dwa wektory $\vec{u}, \vec{v}$ nazywamy *współliniowymi* lub *kolinearnymi*, gdy istnieje prosta, w której zawate są te wektory. Piszemy wówczas $\boldsymbol{\vec{u} \parallel \vec{v}}$.

Definicja 5.1.8. *Iloczynem wektorowym* uporządkowanej pary niewspółliniowych wektorów $\vec{u}, \vec{v}$ nazywamy wektor $\vec{w}$ taki, że:

i) $\vec{w} \perp \vec{u}$, $\vec{w} \perp \vec{v}$,

ii) $|\vec{w}| = |\vec{u}| \cdot |\vec{v}| \cdot \sin \measuredangle(\vec{u}, \vec{v})$,

iii) orientacja trójki $\left(\vec{u}, \vec{v}, \vec{w}\right)$ jest zgodna z orientacją układu współrzędnych.

Wektor $\vec{w}$ oznaczamy symbolem $\boldsymbol{\vec{u} \times \vec{v}}$.
Jeśli $\vec{u} = \vec{0}$ lub $\vec{v} = \vec{0}$, to przyjmujemy $\vec{u} \times \vec{v} = \vec{0}$.

Przykład 5.1.9. $\hat{\imath} \times \hat{\imath} = \hat{\jmath} \times \hat{\jmath} = \hat{k} \times \hat{k} = 0$
$\hat{\imath} \times \hat{\jmath} = \hat{k}, \quad \hat{\jmath} \times \hat{k} = \hat{\imath}, \quad \hat{k} \times \hat{\imath} = \hat{\jmath}, \quad \hat{\jmath} \times \hat{\imath} = -\hat{k}, \quad \hat{k} \times \hat{\jmath} = -\hat{\imath}, \quad \hat{\imath} \times \hat{k} = -\hat{\jmath}$

Twierdzenie 5.1.10. Jeśli $\vec{u} = [u_x, u_y, u_z]$, $\vec{v} = [v_x, v_y, v_z]$, to wówczas

$$\vec{u} \times \vec{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix}.$$

Przykład 5.1.11. Niech $\vec{u} = [1, 2, -3], \vec{v} = [3, 4, 5]$. Korzystamy z twierdzenia Laplace'a

$$\vec{u} \times \vec{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 2 & -3 \\ 3 & 4 & 5 \end{vmatrix} = \hat{i} \begin{vmatrix} 2 & -3 \\ 4 & 5 \end{vmatrix} - \hat{j} \begin{vmatrix} 1 & -3 \\ 3 & 5 \end{vmatrix} + \hat{k} \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = [22, -14, -2]$$

lub metody Sarrusa

$$\vec{u} \times \vec{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 2 & -3 \\ 3 & 4 & 5 \end{vmatrix} = 10\hat{i} + 4\hat{k} - 9\hat{j} - 6\hat{k} + 12\hat{i} - 5\hat{j} = 22\hat{i} - 14\hat{j} - 2\hat{k}.$$

Twierdzenie 5.1.12 (Własności iloczynu wektorowego)**.** Dla dowolnego $\lambda \in \mathbb{R}$ i dla dowolnych wektorów $\vec{u}, \vec{v}, \vec{w}$ prawdziwe są następujące równości.

i) $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$

ii) $\lambda(\vec{u} \times \vec{v}) = (\lambda\vec{u}) \times \vec{v} = \vec{u} \times (\lambda\vec{v})$

iii) $(\vec{u} + \vec{v}) \times \vec{w} = (\vec{u} \times \vec{w}) + (\vec{v} \times \vec{w})$

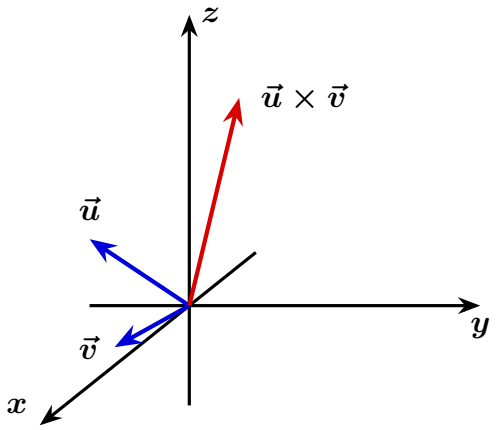
iv) $\vec{u} \times (\vec{v} + \vec{w}) = (\vec{u} \times \vec{v}) + (\vec{u} \times \vec{w})$

v) $|\vec{u} \times \vec{v}| \leq |\vec{u}| \cdot |\vec{v}|$

vi) $\vec{u} \times \vec{v} = \vec{0} \Leftrightarrow \left(\vec{u} \parallel \vec{v} \vee \vec{u} = \vec{0} \vee \vec{v} = \vec{0}\right)$

Reguła prawej dłoni





Uwaga 5.1.13. Pole równoległoboku rozpiętego na wektorach $\vec{u}, \vec{v}$ równe jest $|\vec{u} \times \vec{v}|$.

Iloczyn mieszany

Definicja 5.1.14. Niech $\vec{u} = [u_x, u_y, u_z]$, $\vec{v} = [v_x, v_y, v_z]$,$\vec{w} = [w_x, w_y, w_z]$. *Iloczynem mieszanym* uporządkowanej trójki wektorów $\vec{u}, \vec{v}, \vec{w}$ nazywamy liczbę $\boldsymbol{(\vec{u} \times \vec{v}) \circ \vec{w}}$.
Oznaczamy ją symbolem $\boldsymbol{\left(\vec{u}, \vec{v}, \vec{w}\right)}$.

Twierdzenie 5.1.15. Niech $\vec{u} = [u_x, u_y, u_z]$, $\vec{v} = [v_x, v_y, v_z]$,$\vec{w} = [w_x, w_y, w_z]$. Wówczas

$$\left(\vec{u}, \vec{v}, \vec{w}\right) = \begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix}.$$

Dowód. $\begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} = w_x \begin{vmatrix} u_y & u_z \\ v_y & v_z \end{vmatrix} - w_y \begin{vmatrix} u_x & u_z \\ v_x & v_z \end{vmatrix} + w_z \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} =$

$= [w_x, w_y, w_z] \circ \left[\begin{vmatrix} u_y & u_z \\ v_y & v_z \end{vmatrix}, \begin{vmatrix} u_x & u_z \\ v_x & v_z \end{vmatrix}, \begin{vmatrix} u_x & u_y \\ v_x & v_y \end{vmatrix} \right] = \vec{w} \circ (\vec{u} \times \vec{v}) \quad \square$

Uwaga 5.1.16. Objętość równoległościanu rozpiętego na wektorach $\vec{u}, \vec{v}, \vec{w}$ równa jest $|\left(\vec{u}, \vec{v}, \vec{w}\right)|$.

Dowód. Niech $\alpha = \measuredangle(\vec{u} \times \vec{v}, \vec{w})$. Ponieważ $V = P_p \cdot h = |\vec{u} \times \vec{v}| \cdot h$ oraz $\cos\alpha = \frac{h}{|\vec{w}|}$, zatem $V = |\vec{u} \times \vec{v}| \cdot |\vec{w}| \cos\alpha = \vec{w} \circ (\vec{u} \times \vec{v})$. $\square$

Przykład 5.1.17. Czy punkty $A = (1, 0, 2), B = (5, 1, 5), C = (3, -1, 2), D = (1, 3, 5)$ leżą w jednej płaszczyźnie?

Punkty leżą w jednej płaszczyźnie wtedy i tylko wtedy, gdy objętość czworościanu rozpiętego na wektorach $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}$ jest równa zero.

$\overrightarrow{AB} = [4, 1, 3]$, $\overrightarrow{AC} = [2, -1, 0]$, $\overrightarrow{AD} = [0, 3, 3]$

$$\left(\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}\right) = \begin{vmatrix} 4 & 1 & 3 \\ 2 & -1 & 0 \\ 0 & 3 & 3 \end{vmatrix} = -12 + 18 - 6 = 0$$

Punkty są współpłaszczyznowe (komplanarne).

Twierdzenie 5.1.18 (Własności iloczynu mieszanego)**.** Dla dowolnego $\lambda \in \mathbb{R}$ i dla dowolnych wektorów $\vec{u}, \vec{v}, \vec{w}, \vec{a}$ prawdziwe są następujące równości.

i) $\left(\vec{u}, \vec{v}, \vec{w}\right) = -\left(\vec{v}, \vec{u}, \vec{w}\right) = \left(\vec{v}, \vec{w}, \vec{u}\right)$

ii) $\lambda\left(\vec{u}, \vec{v}, \vec{w}\right) = \left(\lambda\vec{u}, \vec{v}, \vec{w}\right) = \left(\vec{u}, \lambda\vec{v}, \vec{w}\right) = \left(\vec{u}, \vec{v}, \lambda\vec{w}\right)$

iii) $\left(\vec{u} + \vec{v}, \vec{w}, \vec{a}\right) = \left(\vec{u}, \vec{w}, \vec{a}\right) + \left(\vec{v}, \vec{w}, \vec{a}\right)$

iv) $|\left(\vec{u}, \vec{v}, \vec{w}\right)| \leq |\vec{u}| \cdot |\vec{v}| \cdot |\vec{w}|$

5.2 Symbol Kroneckera i symbol całkowicie antysymetryczny

Podamy definicję dwóch symboli, które często pojawiają się w notacji dotyczącej rachunków na wektorach.

Rozważmy funkcję postaci

$$f : \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\} \to \{0, 1\}, \qquad f(i, j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Definiuje ona symbol $\boldsymbol{\delta_{ij} := f(i, j)}$ zwany *symbolem Kroneckera* lub *deltą Kroneckera*.

Przykład 5.2.1. i) Macierz jednostkowa $I_n \in M_n(\mathbb{R})$ ma elementy postaci δ_{ij}. Innymi słowy, $I_n = [\delta_{ij}]$, dla $i, j \in \{1, 2, \ldots, n\}$.

ii) Niech $\vec{u} = [u_1, u_2, u_3]$, oraz $\vec{v} = [v_1, v_2, v_3]$ będą dowolnymi wektorami. Ich iloczyn skalarny możemy zapisać w następującej postaci.

$$\vec{u} \circ \vec{v} = u_1v_1 + u_2v_2 + u_3v_3 = \sum_{i=1}^{3} u_iv_i = \sum_{i,j=1}^{3} \delta_{ij}u_iv_j$$

Definicja 5.2.2. Niech $n \in \mathbb{N}$ oraz $i_1, i_2, \ldots, i_n$, to liczby ze zbioru $\{1, 2, \ldots, n\}$ nie koniecznie różne. *Symbolem całkowicie antysymetrycznym* lub *tensorem Levi-Civita* w n wymiarach nazywamy symbol postaci [a]

$$\boldsymbol{\varepsilon_{i_1, i_2, \ldots, i_n}} := \begin{cases} 1 \;;\; \text{permutacja} \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix} \text{jest parzysta} \\ -1 \;;\; \text{permutacja} \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix} \text{jest nieparzysta} \\ 0 \;;\; \text{nie wszystkie liczby } i_1, i_2, \ldots, i_n \text{ są od siebie różne} \end{cases}.$$

[a] Gdy mówimy o permutacji, to oczywiście mamy na myśli przypadek, gdy liczby $i_1, i_2, \ldots, i_n$ są od siebie różne, to jest $\{i_1, i_2, \ldots, i_n\} = \{1, 2, \ldots, n\}$.

Innymi słowy

$$\varepsilon_{i_1,i_2,\ldots,i_n} = \begin{cases} \operatorname{sgn}(\sigma) & ; \text{ gdy } \{i_1,i_2,\ldots,i_n\} = \{1,2,\ldots,n\} \text{ oraz } \sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix} \\ 0 & ; \text{ nie wszystkie liczby } i_1,i_2,\ldots,i_n \text{ są od siebie różne} \end{cases}.$$

Uwaga 5.2.3. Permutacje σ oraz $\sigma(j,k)$ różnią się znakiem. [a] Wprost z definicji wynika zatem, że zamiana miejscami dowolnych dwóch indeksów w ciągu $i_1,i_2,\ldots,i_n$ powoduje zmianę symbolu z 1 na -1 i odwrotnie. Dla przykładu, gdy $n=3$ mamy

$$\varepsilon_{ijk} = -\varepsilon_{jik} = \varepsilon_{jki} = -\varepsilon_{kji} = \varepsilon_{kij} = -\varepsilon_{ikj}.$$

Stąd też nazwa symbol całkowicie antysymetryczny.

[a] Porównaj z twierdzeniem 1.3.23.

Przykład 5.2.4. i) Dla $n=2$ otrzymujemy

$$\varepsilon_{ij} = \begin{cases} 1 & ; \ (i,j) = (1,2) \\ -1 & ; \ (i,j) = (2,1) \\ 0 & ; \ i = j \end{cases}.$$

Jak wiemy, pole równoległoboku na płaszczyźnie rozpiętego na wektorach $\vec{u} = [u_1,u_2], \vec{v} = [v_1,v_2]$ jest równe

$$P = |\det \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix}| = |u_1v_2 - u_2v_1| = \left| \sum_{i,j=1}^{2} \varepsilon_{ij} u_i v_j \right|.$$

ii) Dla $n=3$ otrzymujemy

$$\varepsilon_{ijk} = \begin{cases} 1 & ; \ (i,j,k) \in \{(1,2,3),(2,3,1),(3,1,2)\} \\ -1 & ; \ (i,j,k) \in \{(3,2,1),(1,3,2),(2,1,3)\} \\ 0 & ; \ i=j \text{ lub } j=k \text{ lub } k=i \end{cases}$$

Iloczyn mieszany uporządkowanej trójki wektorów $(\vec{u},\vec{v},\vec{w})$, gdzie $\vec{u} = [u_1,u_2,u_3], \vec{v} = [v_1,v_2,v_3], \vec{w} = [w_1,w_2,w_3]$, jest równy

$$(\vec{u},\vec{v},\vec{w}) = \vec{u} \circ \left(\vec{v} \times \vec{w}\right) = \det \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{bmatrix}.$$

Można sprawdzić, że

$$(\vec{u},\vec{v},\vec{w}) = u_1(v_2w_3 - v_3w_2) - u_2(v_1w_3 - v_3w_1) + u_3(v_1w_2 - v_2w_1) = \sum_{i,j,k=1}^{3} \varepsilon_{ijk} u_i v_j w_k.$$

iii) Niech $\vec{u} = [u_1, u_2, u_3], \vec{v} = [v_1, v_2, v_3]$. Iloczyn wektorow $\vec{u} \times \vec{v}$, to wektor $\vec{w} = [w_1, w_2, w_3]$ taki, że

$$w_1 = u_2v_3 - u_3v_2 = \sum_{i,j=1}^{3} \varepsilon_{1ij} u_i v_j,$$

$$w_2 = -u_1v_3 + u_3v_1 = \sum_{i,j=1}^{3} \varepsilon_{2ij} u_i v_j,$$

$$w_3 = u_1v_2 - u_2v_1 = \sum_{i,j=1}^{3} \varepsilon_{3ij} u_i v_j.$$

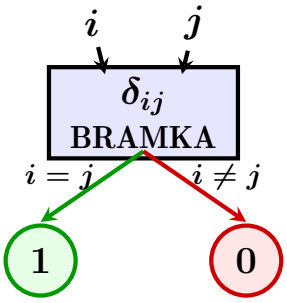
Istotnie,

$$\vec{u} \times \vec{v} = \hat{i} \cdot (\varepsilon_{123} u_2 v_3 + \varepsilon_{132} u_3 v_2) + \hat{j} \cdot (\varepsilon_{213} u_1 v_3 + \varepsilon_{231} u_3 v_1) + \hat{k} \cdot (\varepsilon_{312} u_1 v_2 + \varepsilon_{321} u_2 v_1) =$$

$$= \hat{i} \cdot (u_2 v_3 - u_3 v_2) + \hat{j} \cdot (-u_1 v_3 + u_3 v_1) + \hat{k} \cdot (u_1 v_2 - u_2 v_1)$$

iv) Niech $A \in M_n(\mathbb{R})$, $A = [a_{ij}]$. Wówczas

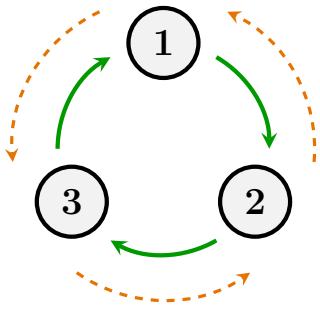
$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)} = \sum_{i_1, i_2, \dots, i_n = 1}^{n} \varepsilon_{i_1, i_2, \dots, i_n} a_{1 i_1} a_{2 i_2} \dots a_{n i_n}.$$



DELTA KRONECKERA (δ_{ij})

STAN $i = j$ (ZIELONE ŚWIATŁO):
Obwód zamknięty, wynik wynosi **1**.
Filtruje sumę indeksów: $\sum_j \delta_{ij} v_j = v_i$

STAN $i \neq j$ (BLOKADA):
Obwód przerwany, wynik wynosi **0**.
Wyraz zostaje całkowicie wykasowany z sumy.



SYMBOL LEVI-CIVITY (ε_{ijk})

RUCH ZGODNY (→):
Permutacja parzysta (zgodna z zegarem).
$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = +1$

RUCH PRZECIWNY (⇢):
Permutacja nieparzysta (lustrzane odbicie).
$\varepsilon_{321} = \varepsilon_{132} = \varepsilon_{213} = -1$

ZACIĘCIE MECHANIZMU:
Dowolne powtórzenie indeksu (np. $\varepsilon_{112} = \varepsilon_{323} = 0$).

Wizualizacja operacji na indeksach za pomocą analogii mechaniczno-blokowych.

Twierdzenie 5.2.5 (Własności symbolu całkowicie antysymetrycznego). [a]

i) $n = 2$

$$\varepsilon_{ik}\varepsilon_{jl} = \begin{vmatrix} \delta_{ik} & \delta_{il} \\ \delta_{jk} & \delta_{jl} \end{vmatrix} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}, \qquad \sum_{k=1}^{2} \varepsilon_{ik}\varepsilon_{jk} = \delta_{ij}, \qquad \sum_{i,j=1}^{2} \varepsilon_{ij}\varepsilon_{ij} = 2$$

ii) $n = 3$

$$\varepsilon_{ijk}\varepsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{jl} & \delta_{kl} \\ \delta_{im} & \delta_{jm} & \delta_{km} \\ \delta_{in} & \delta_{jn} & \delta_{kn} \end{vmatrix}, \qquad \sum_{k=1}^{3} \varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl},$$

$$\sum_{j,k=1}^{3} \varepsilon_{ijk}\varepsilon_{ljk} = 2\delta_{il}, \qquad \sum_{i,j,k=1}^{3} \varepsilon_{ijk}\varepsilon_{ijk} = 3!$$

[a] Aby udowodnić podane związki, w każdym wypadku należy porównać lewą i prawą stronę równania, wypisując wszystkie możliwości.

Przykład 5.2.6. Niech $\vec{u} = [u_1, u_2, u_3], \vec{v} = [v_1, v_2, v_3], \vec{w} = [w_1, w_2, w_3]$. Wyznaczymy wektor $\vec{a} = \vec{u} \times (\vec{v} \times \vec{w})$. Jeśli $\vec{a} = [a_1, a_2, a_3]$, to na mocy twierdzenia 5.2.5 otrzymujemy

$$a_1 = \sum_{j,k=1}^{3} \varepsilon_{1jk}u_j \sum_{l,m=1}^{3} \varepsilon_{klm}v_l w_m = \sum_{j,l,m=1}^{3} u_j v_l w_m \sum_{k=1}^{3} \varepsilon_{1jk}\varepsilon_{klm} = \sum_{j,l,m=1}^{3} u_j v_l w_m \sum_{k=1}^{3} \varepsilon_{1jk}\varepsilon_{lmk} =$$

$$= \sum_{j,l,m=1}^{3} u_j v_l w_m \Big(\delta_{1l}\delta_{jm} - \delta_{1m}\delta_{jl}\Big) = u_2 v_1 w_2 + u_3 v_1 w_3 - u_2 v_2 w_1 - u3 v_3 w_1 =$$

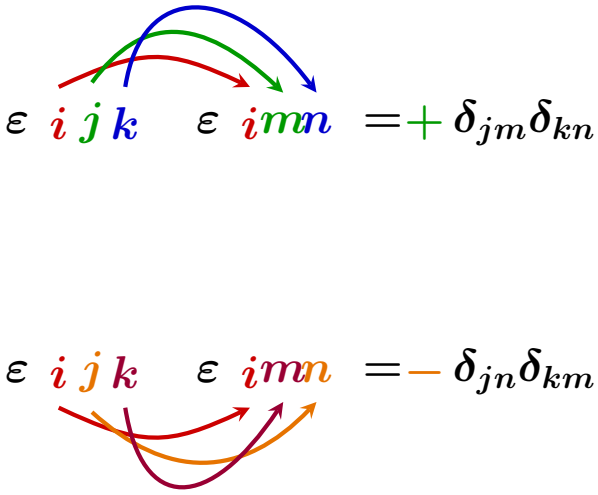
$$= (u_1w_1 + u_2w_2 + u_3w_3)v_1 - (u_1v_1 + u_2v_2 + u_3v_3)w_1 = \Big(\sum_{j=1}^{3} u_j w_j\Big)v_1 - \Big(\sum_{j=1}^{3} u_j v_j\Big)w_1 = (\vec{u}\circ\vec{w})v_1 - (\vec{u}\circ\vec{v})w_1$$

Analogicznie

$$a_2 = (\vec{u} \circ \vec{w})v_2 - (\vec{u} \circ \vec{v})w_2, \qquad a_3 = (\vec{u} \circ \vec{w})v_3 - (\vec{u} \circ \vec{v})w_3.$$

Stąd

$$\vec{a} = \vec{u} \times (\vec{v} \times \vec{w}) = \Big((\vec{u}\circ\vec{w})v_1 - (\vec{u}\circ\vec{v})w_1, (\vec{u}\circ\vec{w})v_2 - (\vec{u}\circ\vec{v})w_2, (\vec{u}\circ\vec{w})v_3 - (\vec{u}\circ\vec{v})w_3\Big) = (\vec{u} \circ \vec{w})\vec{v} - (\vec{u} \circ \vec{v})\vec{w}.$$



→ **Wskaźnik sumacyjny i** (występujący w obu symbolach) zawsze łączy się w parę i znika.

→ **Parowanie proste (+):**
Pozostałe wolne indeksy łączymy ze sobą w naturalnej kolejności: pierwszy z pierwszym ($j \to m$), drugi z drugim ($k \to n$). Daje to iloczyn $\delta_{jm}\delta_{kn}$.

→ **Parowanie krzyżowe (−):**
Wolne indeksy łączymy "na krzyż": pierwszy z drugim ($j \to n$), drugi z pierwszym ($k \to m$). Każde skrzyżowanie linii w fizyce zmienia znak, stąd **minus** przed $\delta_{jn}\delta_{km}$.

Schemat graficznego parowania indeksów w tożsamości $\sum_i \varepsilon_{ijk}\varepsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$.

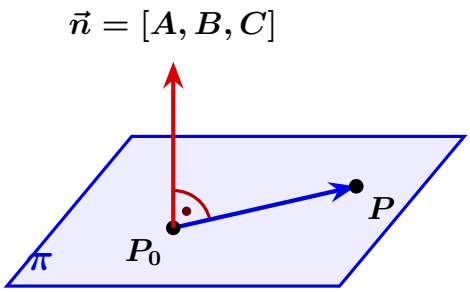
5.3 Płaszczyzna w przestrzeni $\mathbb{R}^3$

Równanie ogólne i normalne płaszczyzny

Niech $P_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ będzie ustalonym punktem, zaś $\vec{n} = [A, B, C] \neq \vec{0}$ ustalonym wektorem. Wówczas zbiór

$$\boldsymbol{\pi = \{P = (x, y, z) \in \mathbb{R}^3 : \overrightarrow{P_0P} \perp \vec{n}\}}$$

jest płaszczyzną przechodząca przez punkt P_0 i prostopadłą do wektora $\vec{n}$.
Wektor $\vec{n}$ nazywamy *wektorem normalnym* płaszczyzny π.



$P \in \pi \Leftrightarrow \overrightarrow{P_0P} \circ \vec{n} = 0 \Leftrightarrow A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$

Równanie normalne	$\boldsymbol{\pi : A(x - x_0) + B(y - y_0) + C(z - z_0) = 0}$
Równanie ogólne	$\boldsymbol{\pi : Ax + By + Cz + D = 0}, \quad D = -Ax_0 - By_0 - Cz_0$

Przykład 5.3.1. i) $P_0 = (1,2,5), \vec{n} = [1,-1,3]$, $P_0 \in \pi$, $\vec{n} \perp \pi$, $\pi = ?$

$\pi : 1 \cdot (x-1) + (-1) \cdot (y-2) + 3 \cdot (z-5) = 0$ równanie normalne

$\pi : x - y + 3z - 14 = 0$ równanie ogólne

ii) Płaszczyzna Oxz opisana jest równaniem $y = 0$.

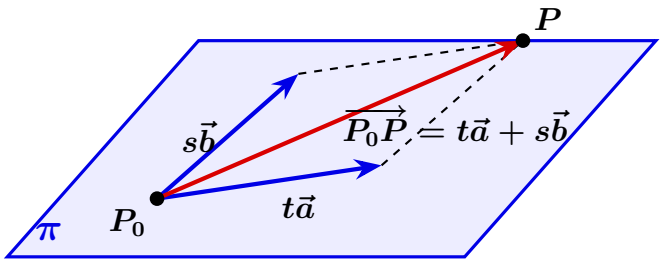
iii) Z równania plaszczyzny $\pi : x + z = 0$ możemy odczytać wektor normalny $\vec{n} = [1,0,1]$ oraz punkt należący do płaszczyzny $P_0 = (0,0,0)$.

Równanie parametryczne płaszczyzny

Niech $P_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ będzie ustalonym punktem, zaś $\vec{a} = [a_x, a_y, a_z] \neq \vec{0}$, $\vec{b} = [b_x, b_y, b_z] \neq \vec{0}$ ustalonymi wektorami niewspółliniowymi, tj. $\vec{a} \nparallel \vec{b}$. Wówczas zbiór

$$\boldsymbol{\pi = \{P = (x,y,z) \in \mathbb{R}^3 : \quad \exists t,s \in \mathbb{R} \quad \overrightarrow{P_0P} = t\vec{a} + s\vec{b}\}}$$

jest płaszczyzną przechodząca przez punkt P_0 i równoległą do wektorów $\vec{a}, \vec{b}$.
Mówimy, że wektory $\vec{a}, \vec{b}$ są *wektorami rozpinającymi* płaszczyznę π.



$$\overrightarrow{P_0P} = t\vec{a} + s\vec{b} \Leftrightarrow [x - x_0, y - y_0, z - z_0] = t[a_x, a_y, a_z] + s[b_x, b_y, b_z]$$

Równanie parametryczne $\quad \boldsymbol{\pi : \begin{cases} x = x_0 + t \cdot a_x + s \cdot b_x \\ y = y_0 + t \cdot a_y + s \cdot b_y \\ z = z_0 + t \cdot a_z + s \cdot b_z \end{cases}, \quad t,s \in \mathbb{R}}$

Przykład 5.3.2. i) Napiszemy równanie parametryczne i równanie ogólne płaszczyzny π przechodzącej przez punkty $A = (1,1,4)$, $B = (2,5,4)$ i równoległej do osi Oy.

$\overrightarrow{AB} = [1,4,0] \parallel \pi, \quad \hat{j} = [0,1,0] \parallel Oy \Rightarrow \hat{j} \parallel \pi, \quad A \in \pi$
Płaszczyznę π możemy opisać równaniem parametrycznym postaci

$$\pi : \begin{cases} x = 1 + 1 \cdot t + 0 \cdot s & = 1 + t \\ y = 1 + 4 \cdot t + 1 \cdot s & = 1 + 4t + s \\ z = 4 + 0 \cdot t + 0 \cdot s & = 4 \end{cases}, \; t,s \in \mathbb{R}.$$

Wektor $\vec{n} := \overrightarrow{AB} \times \hat{j}$ jest wektorem normalnym płaszczyzny π. Znajdziemy jego współrzędne.

$$\overrightarrow{AB} \times \hat{j} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 4 & 0 \\ 0 & 1 & 0 \end{vmatrix} = [0, 0, 1]$$

Równanie ogólne płaszczyzny π ma postać $\pi: \ 0 \cdot (x-1) + 0 \cdot (y-1) + 1 \cdot (z-4) = 0$, skąd $\pi: \ z = 4$.

ii) Zapiszemy równanie parametryczne płaszczyzny π danej równaniem ogólnym

$$\pi: \ x - y + 3z - 14 = 0.$$

Przyjmijmy $y = t$, $z = s$, wówczas

$$\pi: \begin{cases} x = t - 3s + 14 \\ y = t \\ z = s \end{cases}, \ t, s \in \mathbb{R}.$$

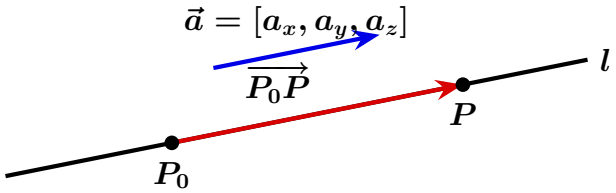
5.4 Prosta w przestrzeni $\mathbb{R}^3$

Równanie parametryczne i kierunkowe prostej

Niech $P_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ będzie ustalonym punktem, zaś $\vec{a} = [a_x, a_y, a_z] \neq \vec{0}$ ustalonym wektorem. Wówczas zbiór

$$\boldsymbol{l = \{P = (x, y, z) \in \mathbb{R}^3 : \ \overrightarrow{P_0P} \parallel \vec{a}\}}$$

jest prostą przechodząca przez punkt P_0 i równoległą do wektora $\vec{a}$.
Wektor $\vec{a}$ nazywamy *wektorem kierunkowym* prostej l.



$$P_0 \in l \quad \Leftrightarrow \quad \exists t \in \mathbb{R} \ : \ \overrightarrow{P_0P} = t\vec{a}$$

Równanie postaci $\boldsymbol{l}: \begin{cases} \boldsymbol{x = x_0 + t \cdot a_x} \\ \boldsymbol{y = y_0 + t \cdot a_y} \\ \boldsymbol{z = z_0 + t \cdot a_z} \end{cases}, \boldsymbol{t} \in \mathbb{R}$ nazywamy *równaniem parametrycznym* lub *parametryzacją* prostej $\boldsymbol{l}$.

Rugując z każdego z powyższych równań parametr $\boldsymbol{t}$, otrzymujemy równanie postaci

$$\boldsymbol{l}: \ \frac{\boldsymbol{x - x_0}}{\boldsymbol{a_x}} = \frac{\boldsymbol{y - y_0}}{\boldsymbol{a_y}} = \frac{\boldsymbol{z - z_0}}{\boldsymbol{a_z}},$$

które nazywamy *równaniem kierunkowym* prostej $\boldsymbol{l}$.

Przykład 5.4.1. Zapiszemy równanie prostej l przechodzącej przez punkt $P_0 = (2, 3, 1)$ i równoległej do płaszczyzn $\pi_1 : 6x - y + z - 2 = 0$, $\pi_2 : x + 3y - 2z + 1 = 0$.

Oznaczmy $\vec{n}_1 = [6, -1, 1] \perp \pi_1$, $\vec{n}_2 = [1, 3, -2] \perp \pi_2$ oraz $\vec{a} = \vec{n}_1 \times \vec{n}_2 \parallel l$.

Wówczas $\vec{a} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 6 & -1 & 1 \\ 1 & 3 & -2 \end{vmatrix} = [-1, 13, 19]$ oraz $l : \begin{cases} x = 2 - t \\ y = 3 + 13t \\ z = 1 + 19t \end{cases}, t \in \mathbb{R}$.

Przykład 5.4.2. Dana jest prosta $l : 3x = 2y + 1 = 5 - 5z$.
Odczytamy wektor kierunkowy prostej $\vec{a} \parallel l$ i podamy współrzędne punktu $P \in l$.

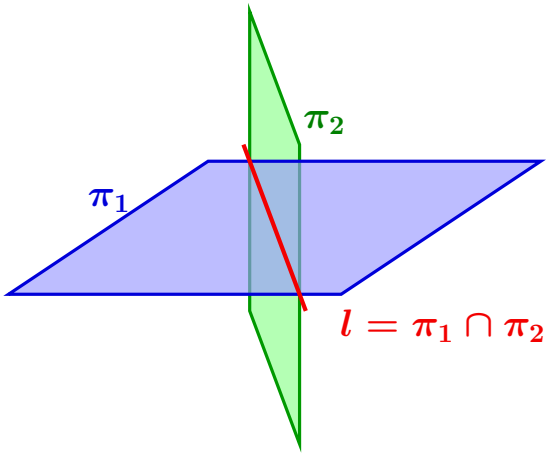
Prosta l opisana jest równaniem kierunkowym.

$$3x = 2y + 1 = 5 - 5z \;\Leftrightarrow\; \frac{x - 0}{\frac{1}{3}} = \frac{y - (-\frac{1}{2})}{\frac{1}{2}} = \frac{z - 1}{-\frac{1}{5}}$$

Zatem $\vec{a} = \left[\frac{1}{3}, \frac{1}{2}, -\frac{1}{5}\right]$ oraz $P = (0, -\frac{1}{2}, 1)$.

Równanie krawędziowe prostej

Niech $\pi_1 : A_1x + B_1y + C_1z + D_1 = 0$, $\pi_2 : A_2x + B_2y + C_2z + D_2 = 0$, gdzie $A_1^2 + B_1^2 + C_1^2 \neq 0$, $A_2^2 + B_2^2 + C_2^2 \neq 0$ będą dwiema nierównoległymi płaszczyznami.
Ich częścią wspólną jest prosta $\boldsymbol{l = \pi_1 \cap \pi_2}$.



$$P \in l \quad \Leftrightarrow \quad (P \in \pi_1 \wedge P \in \pi_2)$$

Równanie krawędziowe $\quad \boldsymbol{l :} \begin{cases} \boldsymbol{A_1x + B_1y + C_1z + D_1 = 0} \\ \boldsymbol{A_2x + B_2y + C_2z + D_2 = 0} \end{cases}$.

Przykład 5.4.3. Zapiszemy równanie parametryczne prostej $l: \begin{cases} x+2y-z+1=0 \\ x+y=0 \end{cases}$.

Odczytujemy wektory normalne podanych płaszczyzn $\vec{n}_1 = [1,2,-1]$, $\vec{n}_2 = [1,1,0]$.
Wektorem kierunkowym prostej l jest wektor $\vec{n}_1 \times \vec{n}_2 = [1,-1,-1]$.

Wybierzmy punkt należący do l. Gdy $x=0$, to wówczas $y=-x=0$ oraz $z=x+2y+1=1$, zatem jednym z takich punktów jest punkt $P=(0,0,1)$.

Prostą l możemy opisać równaniem parametrycznym postaci

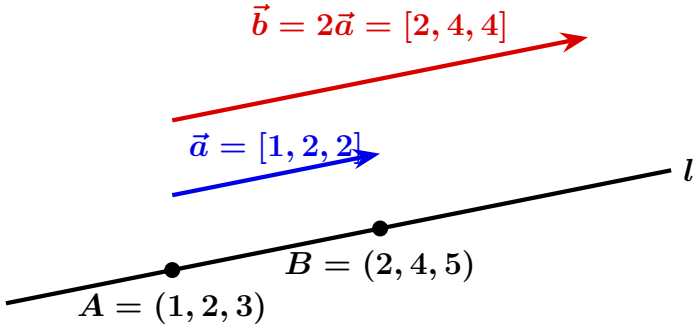
$$l: \begin{cases} x=t \\ y=-t \\ z=1-t \end{cases}, \quad t \in \mathbb{R}.$$

Uwaga 5.4.4. W powyższym zadaniu punkt P mogliśmy wybrać inaczej. Na tej podstawie możemy zauważyć, że każda zadana prosta posiada nieskończenie wiele różnych parametryzacji.

Przykład 5.4.5 (Dwie różne parametryzacje tej samej prostej)**.**
Rozważmy prostą l przechodzącą przez punkty $A=(1,2,3)$ oraz $B=(2,4,5)$. Wyznaczamy wektor kierunkowy $\vec{a} = \overrightarrow{AB} = [1,2,2]$.

- **Parametryzacja I** (punkt A, wektor $\vec{a}$) $\quad l: \begin{cases} x=1+t \\ y=2+2t \\ z=3+2t \end{cases}, \quad t \in \mathbb{R}$
- **Parametryzacja II** (punkt B, wektor $2\vec{a}=[2,4,4]$) $\quad l: \begin{cases} x=2+2s \\ y=4+4s \\ z=5+4s \end{cases}, \quad s \in \mathbb{R}$

Oba układy równań opisują geometrycznie tę samą prostą, mimo użycia innych stałych i różnych parametrów (t oraz s).



5.5 Wzajemne położenie płaszczyzn i prostych

Wzajemne położenie płaszczyzn

Niech $\pi_1: A_1x + B_1y + C_1z + D_1 = 0, \quad \vec{n_1} = [A_1, B_1, C_1] \neq \vec{\mathbf{0}}$,
$\pi_2: A_2x + B_2y + C_2z + D_2 = 0, \quad \vec{n_2} = [A_2, B_2, C_2] \neq \vec{\mathbf{0}}$.

Szukanie punktów wspólnych π_1 oraz π_2 polega na rozwiązaniu układu równań

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases} \Leftrightarrow \begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -D_1 \\ -D_2 \end{bmatrix}.$$

Niech $A = \begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{bmatrix}$, $U = \begin{bmatrix} A_1 & B_1 & C_1 & -D_1 \\ A_2 & B_2 & C_2 & -D_2 \end{bmatrix}$.

Płaszczyzny mogą być równoległe.

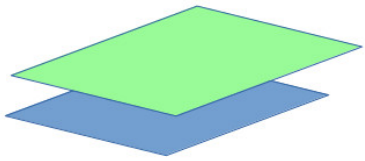
$$\boldsymbol{\pi_1 \| \pi_2 \Leftrightarrow n_1 \| n_2 \Leftrightarrow \vec{n_1} \times \vec{n_2} = \vec{\mathbf{0}}}$$

Wówczas albo $\pi_1 = \pi_2 \Leftrightarrow \frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} = \frac{D_1}{D_2}$, gdy $r(U) = r(A) = 1$

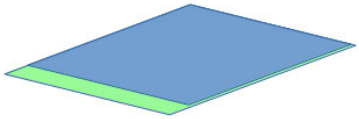
albo $\pi_1 \cap \pi_2 = \varnothing \Leftrightarrow \frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} \neq \frac{D_1}{D_2}$, gdy $r(U) = 2$, $r(A) = 1$.

Gdy $r(U) = r(A) = 2$, płaszczyzny $\pi_1 \nparallel \pi_2$ przecinają się wzdłuż prostej.
W szczególności mogą być prostopadłe.

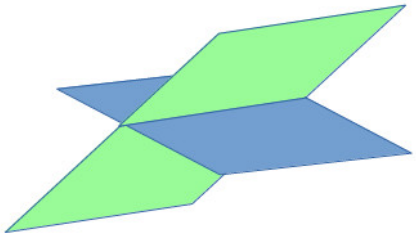
$$\pi_1 \perp \pi_2 \Leftrightarrow n_1 \perp n_2 \Leftrightarrow \vec{n_1} \circ \vec{n_2} = 0$$



płaszczyzny równoległe
$\pi_1 \cap \pi_2 = \varnothing$



$\pi_1 = \pi_2$



płaszczyzny nierównoległe
przecinają się wzdłuż prostej

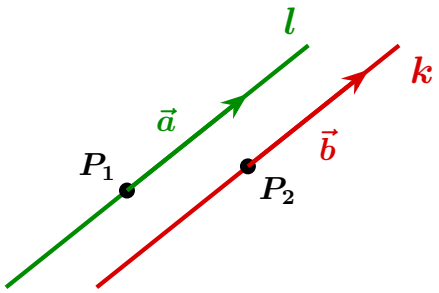
Wzajemne położenie prostych

Niech $\vec{a}$ będzie wektorem kierunkowym prostej l, przechodzącej przez punkt P_1, zaś $\vec{b}$ wektorem kierunkowym prostej k, przechodzącej przez punkt P_2.

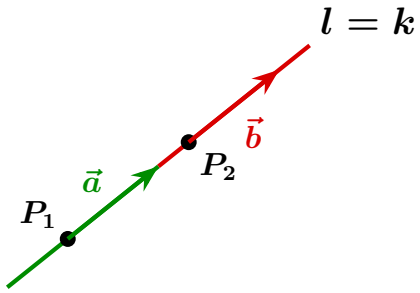
I. Proste mogą być równoległe.

$$l \parallel k \Leftrightarrow \vec{a} \parallel \vec{b} \Leftrightarrow \vec{a} \times \vec{b} = \vec{0}$$

Wówczas albo $l = k$, albo $l \cap k = \varnothing$.



proste równoległe
$\boldsymbol{l \cap k = \emptyset}$



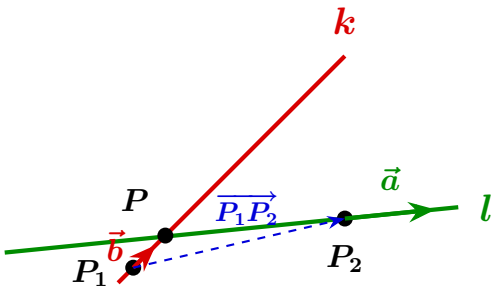
proste równoległe
$\boldsymbol{l = k}$

II. Gdy $l \nparallel k$, możliwe są dwie sytuacje.

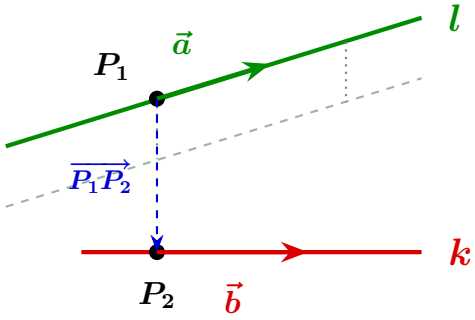
1) Proste l i k leżą w jednej płaszczyźnie, co jest równoważne stwierdzeniu, że wektory $\overrightarrow{P_1P_2}, \vec{a}, \vec{b}$ leżą w jednej płaszczyźnie. Wówczas l i k mają jeden punkt wspólny tj. $l \cap k = \{P\}$,

2) Proste l i k nie leżą w jednej płaszczyźnie (tzw. *proste skośne*), co jest równoważne stwierdzeniu, że wektory $\overrightarrow{P_1P_2}, \vec{a}, \vec{b}$ nie leżą w jednej płaszczyźnie. Wówczas $l \cap k = \varnothing$.

Zatem proste l i k są skośne wtedy i tylko wtedy gdy $\left(\overrightarrow{P_1P_2}, \vec{a}, \vec{b}\right) \neq 0$.



proste nierównoległe leżące w jednej płaszczyźnie mają punkt wspólny

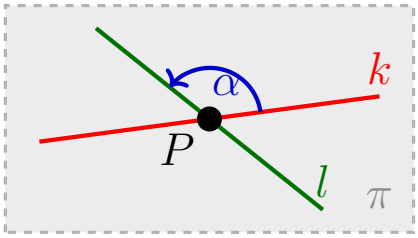


proste nierównoległe nie leżące w jednej płaszczyźnie – brak punktów wspólnych, tzw. proste skośne

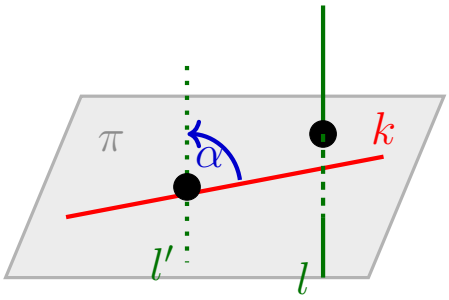
Kąty

Definicja 5.5.1. i) *Kątem między dwiema prostymi* nazywamy kąt ostry (lub prosty, gdy proste są prostopadłe) między odpowiednio zwróconymi wektorami kierunkowymi tychże prostych.

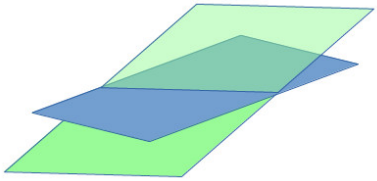
ii) *Kątem między dwiema płaszczyznami* nazywamy kąt ostry (lub prosty, gdy płaszczyzny są prostopadłe) między odpowiednio zwróconymi wektorami normalnymi tychże płaszczyzn.



Kąt między prostymi przecinającymi się



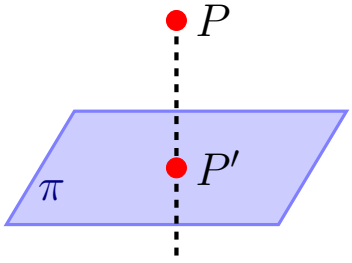
Kąt α między prostymi skośnymi



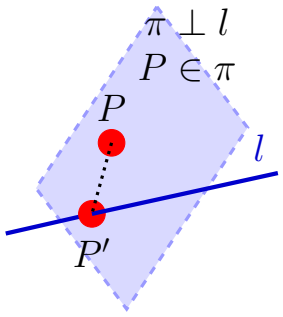
Kąt między nierównoległymi płaszczyznami

Definicja 5.5.2. i) *Rzutem prostokątnym punktu P na płaszczyznę π* nazywamy punkt $P' \in \pi$ taki, że $PP' \perp \pi$.

ii) *Rzutem prostokątnym punktu P na prostą l* nazywamy punkt $P' \in l$ taki, że $PP' \perp l$.

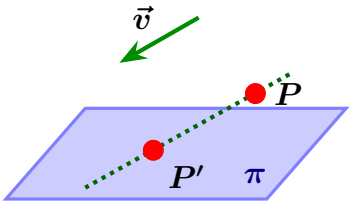


rzut prostokątny punktu P na płaszczyznę π

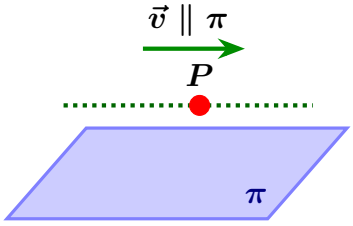


rzut prostokątny punktu P na prostą l

Można zdefiniować *rzut ukośny* w kierunku zadanego wektora.

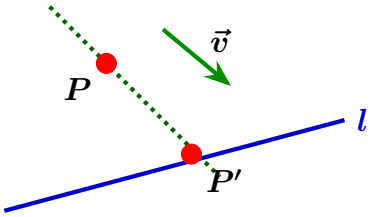


Rzut punktu P na płaszczyznę π w kierunku wektora $\vec{v} \nparallel \pi$.

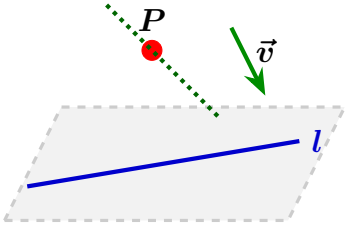


Brak punktu wspólnego

Jeśli $\vec{v} \parallel \pi$, rzut ten nie istnieje.



Rzut ukośny punktu P na prostą l
w kierunku wektora $\vec{v}$,o którym zakładamy,
że należy do płaszczyzny zawierającej P oraz l.



Linie są skośne

Jeśli l i prosta o kierunku $\vec{v}$
przechodząca przez P są skośne,
rzut ten nie istnieje.

Przykład 5.5.3. Wyznacz rzut prostokątny punktu $P = (4, 5, -3)$ na płaszczyznę

$$\pi : \begin{cases} x = 2 + 2t + s \\ y = 1 + 3s \\ z = 3 + t + s \end{cases}, t, s \in \mathbb{R}.$$

Niech k będzie prostą taką, że $k \perp \pi$, $P \in k$.

$\vec{u} = [2, 0, 1] \parallel \pi, \quad \vec{v} = [1, 3, 1] \parallel \pi, \quad A = (2, 1, 3) \in \pi$

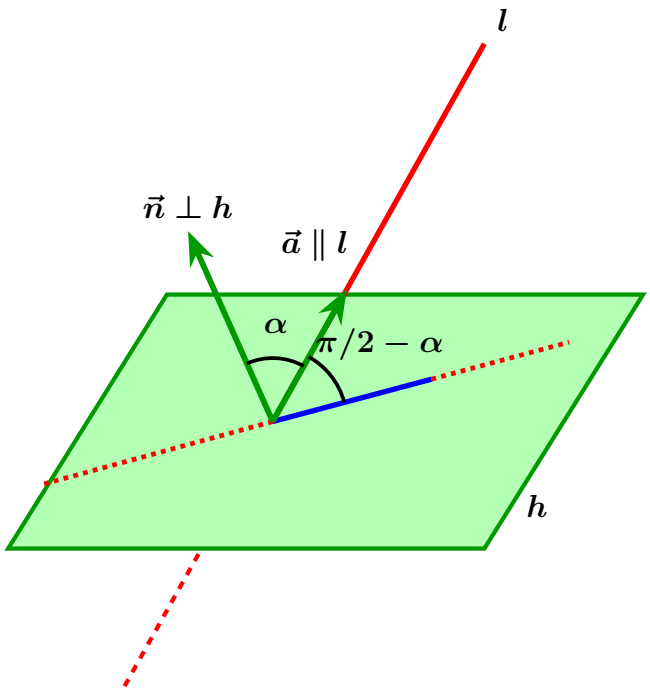
$$\vec{n} = \vec{u} \times \vec{v} \perp \pi, \quad \vec{n} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 2 & 0 & 1 \\ 1 & 3 & 1 \end{vmatrix} = [-3, -1, 6], \quad k \perp \pi \Rightarrow k \parallel \vec{n}$$

$$k : \begin{cases} x = 4 - 3t \\ y = 5 - t \\ z = -3 + 6t \end{cases}, t \in \mathbb{R}, \quad \pi : -3(x - 2) - (y - 1) + 6(z - 3) = 0$$

$\pi : 3x + y - 6z + 11 = 0 \quad \{P'\} = k \cap \pi = ?$

$3(4 - 3t) + 5 - t - 6(-3 + 6t) = 0 \Rightarrow t = 1, \quad P' = (1, 4, 3)$

Definicja 5.5.4. *Kątem między płaszczyzną a prostą* nazywamy kąt o mierze $\frac{\pi}{2} - \alpha$, gdzie α to miara kąta ostrego (lub prostego, gdy prosta i płaszczyzna są równoległe) między odpowiednio zwróconym wektorem kierunkowym prostej a wektorem normalnym płaszczyzny.



Kąt między płaszczyzną $\boldsymbol{h}$ a prostą $\boldsymbol{l}$.

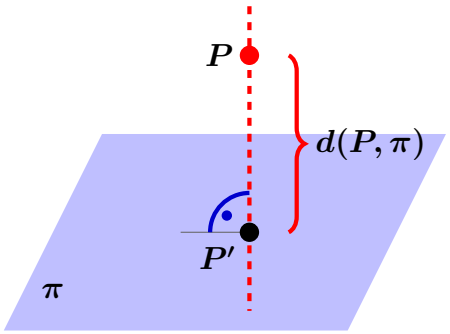
Przykład 5.5.5. Wyznacz kąt między prostą $l: \begin{cases} x = 2 - t \\ y = 1 \\ z = 2 + 2t \end{cases}, t \in \mathbb{R}$ a płaszczyzną $\pi:\ 3x + y + z + 1 = 0$.

Mamy $\vec{a} = [-1, 0, 2] \parallel l$, $\vec{n} = [3, 1, 1] \perp \pi$. Oznaczmy $\beta = \measuredangle(\vec{n}, \vec{a})$, $\alpha = \frac{\pi}{2} - \beta$.
Obliczamy $\cos\beta = \frac{|\vec{n} \circ \vec{a}|}{|\vec{n}| \cdot |\vec{a}|} = \frac{|-3+0+2|}{\sqrt{9+1+1} \cdot \sqrt{1+4}} = \frac{1}{\sqrt{55}}$, $\alpha = \frac{\pi}{2} - \arccos\frac{1}{\sqrt{55}}$,
albo $\sin\alpha = \cos\beta \Rightarrow \alpha = \arcsin\frac{1}{\sqrt{55}}$

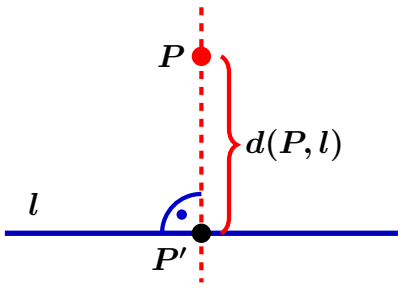
Odległości

Definicja 5.5.6. i) *Odległością punktu P od płaszczyzny π*, nazywamy długość odcinka PP', gdzie P' jest rzutem prostokątnym P na π. Oznaczamy ją $d(P, \pi)$.

ii) *Odległością punktu P od prostej l*, nazywamy długość odcinka PP', gdzie P' jest rzutem prostokątnym P na l. Oznaczamy ją $d(P, l)$.



Odległość punktu $\boldsymbol{P}$ od płaszczyzny $\boldsymbol{\pi}$.



Odległość punktu $\boldsymbol{P}$ od prostej $\boldsymbol{l}$.

Wzór na odległość punktu od płaszczyzny

Niech $P_0 = (x_0, y_0, z_0)$ oraz $\pi: \ Ax + By + Cz + D = 0$, $\vec{n} = [A, B, C] \neq \vec{0}$. Wówczas

$$d(P_0, \pi) = \frac{|Ax_0 + By_0 + Cz_0 + D|}{|\vec{n}|}.$$

Istotnie, niech k będzie prostą taką, że $P_0 \in k$, $k \perp \pi$. Wówczas $k: \begin{cases} x = x_0 + At \\ y = y_0 + Bt \\ z = z_0 + Ct \end{cases}$

$\{P_0'\} = k \cap \pi = ?$

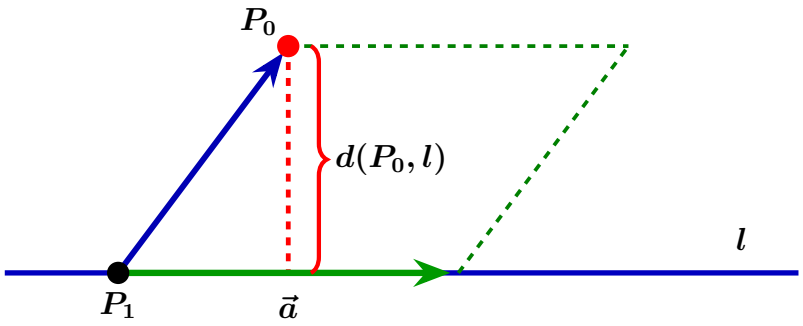
$A(x_0 + At) + B(y_0 + Bt) + C(z_0 + Ct) + D = 0, \quad t = -\frac{Ax_0+By_0+Cz_0+D}{A^2+B^2+C^2}$

$d(P, \pi) = |P_0P_0'| = \sqrt{(At)^2 + (Bt)^2 + (Ct)^2} = |t|\sqrt{A^2 + B^2 + C^2} = \frac{|Ax_0+By_0+Cz_0+D|}{\sqrt{A^2+B^2+C^2}}$

Wzór na odległość punktu od prostej

Niech $P_0 = (x_0, y_0, z_0)$ oraz niech l będzie prostą przechodząca przez punkt P_1 o wektorze kierunkowym $\vec{a}$. Wówczas

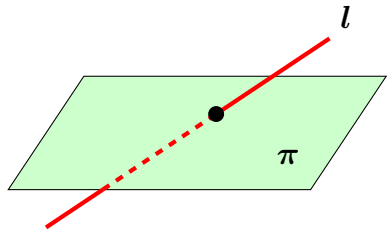
$$d(P_0, l) = \frac{|\overrightarrow{P_1P_0} \times \vec{a}|}{|\vec{a}|}.$$



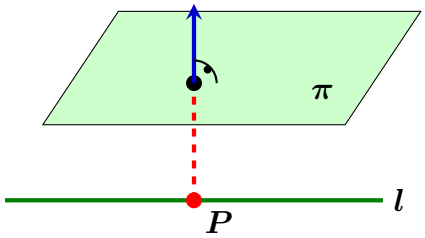
Uzasadnienie wzoru na odległość punktu $\boldsymbol{P_0}$ od prostej $\boldsymbol{l}$.

Odległość prostej od płaszczyzny

Jeśli prosta l nie przecina płaszczyzny π, to wówczas odległością prostej l od płaszczyzny π nazywamy odległość dowolnego punkty prostej od płaszczyzny.



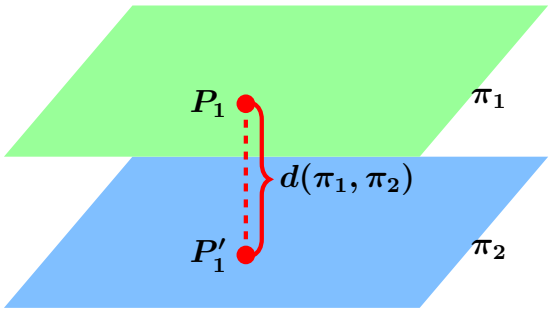
$l \cap \pi \neq \varnothing \Rightarrow d(l, \pi) = 0$



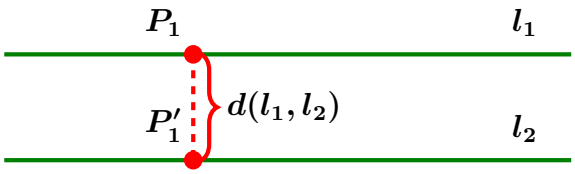
$l \cap \pi = \varnothing \Rightarrow d(l, \pi) = d(P, \pi)$

Definicja 5.5.7. *Odległością dwóch płaszczyzn (prostych) równoległych* nazywamy odległość dowolnego punktu jednej z nich od drugiej.

Można wykazać, że dla płaszczyzn równoległych $\pi_1 : \ Ax + By + Cz + D_1 = 0$, $\pi_2 : \ Ax + By + Cz + D_2 = 0$, $\vec{n_2} = [A, B, C] \neq \vec{0}$ zachodzi wzór $d(\pi_1, \pi_2) = \frac{|D_2 - D_1|}{|\vec{n}|}$.



odległość płaszczyzn równoległych



odległość prostych równoległych

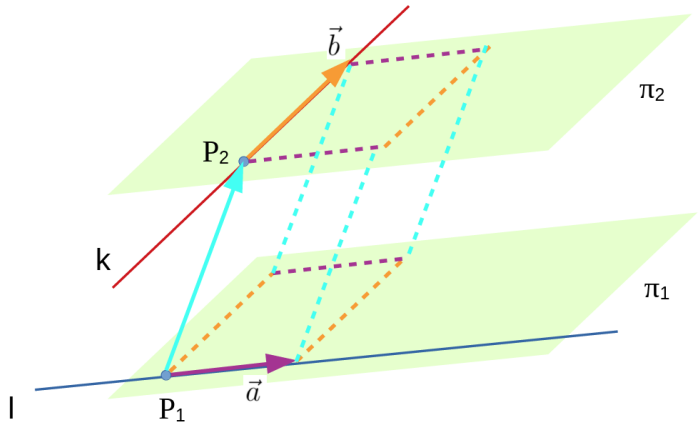
Odległość prostych skośnych

Przypomnijmy, że jeśli dwie proste leżą w jednej płaszczyźnie, to przecinają się lub są równoległe. Proste, które nie są do siebie równolegle i nie mają punktów wspólnych nazywamy skośnymi. Zatem proste są skośne, gdy nie leżą w jednej płaszczyźnie.

Niech $\vec{a}$ będzie wektorem kierunkowym prostej l, przechodzącej przez punkt P_1, zaś $\vec{b}$ wektorem kierunkowym prostej k, przechodzącej przez punkt P_2.

Załóżmy że proste te są skośne. Wówczas istnieją płaszczyzna π_1 zawierającą prostą l oraz płaszczyzna π_2 zawierająca prostą k takie, że $\pi_1 \parallel \pi_2$ oraz $\pi_1 \cap \pi_2 = \varnothing$. Ponadto płaszczyzny równoległe, na których leżą dane dwie proste skośne są wyznaczone jednoznacznie. Istotnie, π_1 jest płaszczyzną

zawierającą l i równoległą do wektora $\vec{b}$. Jej wektorem normalnym jest wektor $\vec{a} \times \vec{b}$. Podobnie, π_2 jest płaszczyzną zawierającą k i równoległą do wektora $\vec{a}$. Jej wektorem normalnym również jest wektor $\vec{a} \times \vec{b}$.



odległość prostych skośnych

Definicja 5.5.8. *Odległością dwóch prostych skośnych* nazywamy odległość dwóch płaszczyzn równoległych zawierających te proste.

Uwaga 5.5.9. Niech l i k będą jak wyżej. Załóżmy że proste te są skośne, co jest równoważne stwierdzeniu, że wektory $\overrightarrow{P_1P_2}, \vec{a}, \vec{b}$ nie leżą w jednej płaszczyźnie. Wówczas

$$d(k,l) = \frac{|\left(\overrightarrow{P_1P_2}, \vec{a}, \vec{b}\right)|}{|\vec{a} \times \vec{b}|}.$$

Przykładowe zadanie polegające na wyznaczaniu odległości dwóch prostych skośnych można znaleźć tutaj.

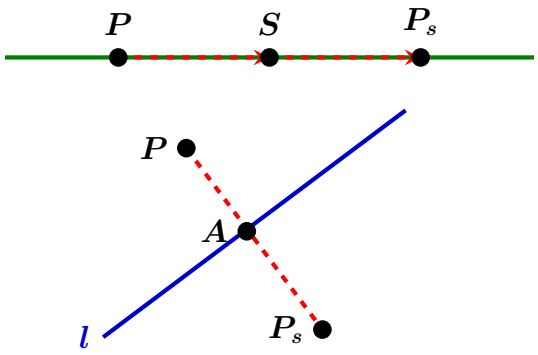
Symetrie

Definicja 5.5.10. Niech S będzie ustalonym punktem, l ustaloną prostą oraz π ustaloną płaszczyzną.

i) Punkt P_s jest *punktem symetrycznym* do punktu P względem punktu S, jeżeli $\overrightarrow{PS} = \overrightarrow{SP_s}$.

ii) Punkt P_s jest punktem symetrycznym do punktu P względem prostej l, jeżeli istnieje $A \in l$ taki, że $\overrightarrow{PA} = \overrightarrow{AP_s}$ oraz $\overrightarrow{PA} \perp l$.

iii) Punkt P_s jest punktem symetrycznym do punktu P względem płaszczyzny π, jeżeli istnieje $A \in \pi$ taki, że $\overrightarrow{PA} = \overrightarrow{AP_s}$ oraz $\overrightarrow{PA} \perp \pi$.



punkty symetryczny względem punktu $\boldsymbol{S}$

punkty symetryczny względem prostej $\boldsymbol{l}$

punkty symetryczny względem płaszczyzny $\boldsymbol{\pi}$

QUIZ

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. Kąt pomiędzy wektorami $\vec{a} = [1, 4, 0]$ oraz $\vec{b} = [2, -1, 100]$ jest ostry.
2. Iloczyn wektorowy jest działaniem przemiennym.
3. Niech $\vec{u}, \vec{v}, \vec{w}$ będą dowolnymi wektorami w przestrzeni $\mathbb{R}^3$. Wówczas $(\vec{u} \times \vec{v}) \circ \vec{w} = \vec{u} \circ (\vec{v} \times \vec{w})$.
4. Dowolne trzy punkty w przestrzeni $\mathbb{R}^3$ jednoznacznie wyznaczają płaszczyznę.
5. Niech π_1 oraz π_2 to dwie dowolne płaszczyzny. Niech A będzie punktem należącym do π_1, zaś A' rzutem prostokątnym punktu A na płaszczyznę π_2. Wówczas $|\overrightarrow{AA'}|$ jest odległością płaszczyzny π_1 od płaszczyzny π_2.
6. Niech l będzie prostą przechodząca przez punkty A i B. Jeśli dla punktu P spełniony jest warunek $\overrightarrow{AP} = t \cdot \overrightarrow{AB}$, dla pewnego $t > 1$, to wówczas punkt P leży na prostej l i punkt B leży pomiędzy punktami A i P.
7. Odległość punktu $P = (3, 5, -1)$ od osi $0y$ wynosi $\sqrt{10}$.
8. Proste skośne mogą leżeć w jednej płaszczyźnie.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Algebra
Wykład nr 6
Data: .....

TEMAT: *Przestrzenie liniowe*

6.1 Przestrzenie liniowe i ich podprzestrzenie

Niech $K = (K, +, \cdot)$ będzie ciałem, gdzie $K = \mathbb{R}$ lub $K = \mathbb{C}$, zaś $V \neq \varnothing$ zbiorem. Niech dane będzie działanie wewnętrzne $\oplus : V \times V \ni (u, v) \mapsto u \oplus v \in V$ oraz działanie zewnętrzne $\odot : K \times V \ni (\alpha, v) \mapsto \alpha \odot v \in V$.

Definicja 6.1.1. Zespół $V = (V, \oplus, K, \odot)$ taki, że

i) $(V, \oplus)$ jest grupą abelową,

ii) $\forall u, v \in V \quad \forall \alpha \in K \quad \alpha \odot (u \oplus v) = (\alpha \odot u) \oplus (\alpha \odot v)$

iii) $\forall v \in V \quad \forall \alpha, \beta \in K \quad (\alpha + \beta) \odot v = (\alpha \odot v) \oplus (\beta \odot v)$

iv) $\forall v \in V \quad \forall \alpha, \beta \in K \quad (\alpha \cdot \beta) \odot v = \alpha \odot (\beta \odot v)$

v) $\forall v \in V \quad 1 \odot v = v$

nazywamy *przestrzenią wektorową* bądź *przestrzenią liniową* nad ciałem K (albo przestrzenią K-liniową). Elementy zbioru V nazywamy *wektorami*, zaś elementy ciała K *skalarami*.

Przykład 6.1.2. Poniższe struktury są przestrzeniami wektorowymi.

i) $\mathbb{R}^n = (\mathbb{R}^n, \oplus, \mathbb{R}, \odot)$
Dla $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ oraz $\alpha \in \mathbb{R}$ definiujemy
$u \oplus v = (u_1 + v_1, \ldots, u_n + v_n)$, $\alpha \odot u =)\alpha u_1, \ldots, \alpha u_n)$.

ii) $(K^n, \oplus, K, \odot)$, gdzie $K = (K, +, \cdot)$ to dowolne ciało
Dla $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n) \in K^n$ oraz $\alpha \in K$ definiujemy
$u \oplus v = (u_1 + v_1, \ldots, u_n + v_n)$, $\alpha \odot u = (\alpha \cdot u_1, \ldots, \alpha \cdot u_n)$.

iii) $(\mathbb{R}^{\mathbb{R}}, +, \mathbb{R}, \cdot)$, gdzie $\mathbb{R}^{\mathbb{R}} = \{f \mid f : \mathbb{R} \to \mathbb{R}\}$
Dla $f_1 : \mathbb{R} \to \mathbb{R}$, $f_2 : \mathbb{R} \to \mathbb{R}$ oraz $\alpha \in \mathbb{R}$ definiujemy
$f_3 = f_1 \oplus f_2$ takie, że $\forall x \in \mathbb{R}\ f_3(x) = (f_1 \oplus f_2)(x) := f_1(x) + f_2(x)$
oraz $f_4 = \alpha \odot f_1$ takie, że $\forall x \in \mathbb{R}\ f_4(x) = (\alpha \odot f_1)(x) := \alpha \cdot f_1(x)$
Elementem neutralnym działania $\oplus$ jest funkcja stale równa zero.

iv) $M_{m\times n}(\mathbb{R}) = (M_{m\times n}(\mathbb{R}), +, \mathbb{R}, \cdot)$, gdzie działania $+, \cdot$ to działania dodawania macierzy i mnożenia macierzy przez liczbę.

v) $\mathbb{R}[x] = (\mathbb{R}[x], +, \mathbb{R}, \cdot)$, gdzie $\mathbb{R}[x]$ to zbiór wszystkich wielomianów jednej zmiennej x o współczynnikach rzeczywistych z działaniami dodawania wielomianów i mnożenia wielomianu przez liczbę.

Uwaga 6.1.3. Często tym samym symbolem oznaczamy działania w ciele $K = (K, +, \cdot)$ i działania w przestrzeni wektorowej $V = (V, +, K, \cdot)$.

Wówczas dla $u, v \in V$, $\alpha, \beta \in K$

$u + v$	suma wektorów
$\alpha + \beta$	suma skalarów
$\alpha \cdot u$	iloczyn wektora przez skalar
$\alpha \cdot \beta$	iloczyn skalarów

Twierdzenie 6.1.4. Niech $V = (V, +, K, \cdot)$ będzie przestrzenią wektorową. Niech $\mathbf{0}$ oznacza element neutralny dodawania w V, zaś $0, 1$ elementy neutralne działań w ciele K. Wówczas:

i) $\forall v \in V \quad \forall \alpha \in K \quad 0 \cdot v = \alpha \cdot \mathbf{0} = \mathbf{0}$

ii) $\forall v \in V \quad \forall \alpha \in K \quad \alpha \cdot v = \mathbf{0} \Leftrightarrow (\alpha = 0 \vee v = \mathbf{0})$

iii) $\forall v \in V \quad \forall \alpha \in K \quad (-\alpha) \cdot v = \alpha \cdot (-v) = -(\alpha \cdot v)$

iv) $\forall v \in V \quad -1 \cdot v = -v$

v) $\forall v \in V \quad \forall \alpha, \beta \in K \quad (\alpha - \beta) \cdot v = (\alpha \cdot v) - (\beta \cdot v)$

vi) $\forall u, v \in V \quad \forall \alpha \in K \quad \alpha \cdot (u - v) = (\alpha \cdot u) - (\alpha \cdot v)$

Niech $V = (V, \oplus, K, \odot)$ będzie przestrzenią wektorową nad ciałem $K = (K, +, \cdot)$ i niech $U \subset V$ będzie niepustym podzbiorem zbioru V.

Definicja 6.1.5. Jeśli zbiór U wraz z działaniami $\oplus|_{U\times U} : U \times U \ni (u, v) \mapsto u \oplus v \in U$, $\odot|_{K\times U} : K \times U \ni (\alpha, v) \mapsto \alpha \odot v \in U$ jest przestrzenią liniową nad ciałem K, to $U = (U, \oplus|_{U\times U}, K, \odot|_{K\times U})$ nazywamy *podprzestrzenią wektorową* lub *podprzestrzenią liniową* przestrzeni V.

Twierdzenie 6.1.6. Jeśli $V = (V, \oplus, K, \odot)$ jest przestrzenią wektorową oraz $\varnothing \neq U \subset V$, to wówczas U jest podprzestrzenią wektorową przestrzeni V wtedy i tylko wtedy, gdy

$$\forall u_1, u_2 \in U \quad \forall \alpha \in K \quad u_1 \oplus u_2 \in U \wedge \alpha \odot u_1 \in U$$

lub równoważnie

$$\forall u_1, u_2 \in U \quad \forall \alpha, \beta \in K \quad (\alpha \odot u_1) \oplus (\beta \odot u_2) \in U.$$

Uwaga 6.1.7. Niech $V = (V, \oplus, K, \odot)$ będzie przestrzenią liniową.

i) $U = \{\mathbf{0}\}$ jest podprzestrzenią liniową. Nazywamy ją *podprzestrzenią trywialną.*

ii) $U = V$ jest podprzestrzenią liniową V. [a] Nazywamy ją *podprzestrzenią niewłaściwą.*

[a] Innymi słowy, każda przestrzeń liniowa jest swoją własną podprzestrzenią.

Uwaga 6.1.8. Każda podprzestrzeń liniowa zawiera wektor zerowy.

Dowód. Jeśli U jest podprzestrzenią liniową, to $\forall u_1, u_2 \in U \quad \forall \alpha \in K \quad u_1 \oplus u_2 \in U \wedge \alpha \odot u_1 \in U$. W szczególności $-1 \odot u_1 = -u_1 \in U$ oraz $u_1 \oplus (-u_1) = \mathbf{0} \in U$. $\square$

Wniosek 6.1.9. Niech $V = (V, \oplus, K, \odot)$ będzie przestrzenią liniową, zaś $U \subset V$ podzbiorem V. Jeśli $\mathbf{0} \notin U$, to U nie jest podprzestrzenią liniową przestrzeni V.

Przykład 6.1.10.

i) Niech $\mathbb{R}^{\mathbb{R}} := \{f : \mathbb{R} \to \mathbb{R} : \ f - funkcja\}$ oraz $V = (\mathbb{R}^{[0,1]}, +, \mathbb{R}, \cdot)$.

$U = \{f : \mathbb{R} \to \mathbb{R} : \lim_{x \to 0} f(x) = 1\}$ nie jest podprzestrzenią liniową V, bowiem funkcja stale równa zero nie należy do U.

ii) Niech $\mathbb{R}^{[0,1]} := \{f : [0,1] \to \mathbb{R} : \ f - funkcja\}$ oraz $V = (\mathbb{R}^{[0,1]}, +, \mathbb{R}, \cdot)$, $U = (\mathcal{C}([0,1], \mathbb{R}), +, \mathbb{R}, \cdot)$

U jest podprzestrzenią liniową V, bowiem suma funkcji ciągłych jest funkcją ciągłą oraz iloczyn funkcji ciągłej przez liczbę jest funkcja ciągłą.

iii) $V = (\mathbb{R}[x], +, \mathbb{R}, \cdot)$, $U = \{f \in \mathbb{R}[x] : \deg f \ - \ parzysty\}$

U **nie** jest podprzestrzenią liniową V. Niech $f(x) = x^4 + x^3$ oraz $g(x) = -x^4$. Wówczas $(f+g)(x) = x^3$. Zatem $f, g \in U$, ale $f + g \notin U$.

iv) $V = (\mathbb{R}^4, +, \mathbb{R}, \cdot)$, $U = \{(x, y, z, t) \in \mathbb{R}^4 : \ 2x + z - 3t = 0 \wedge y = 0\}$

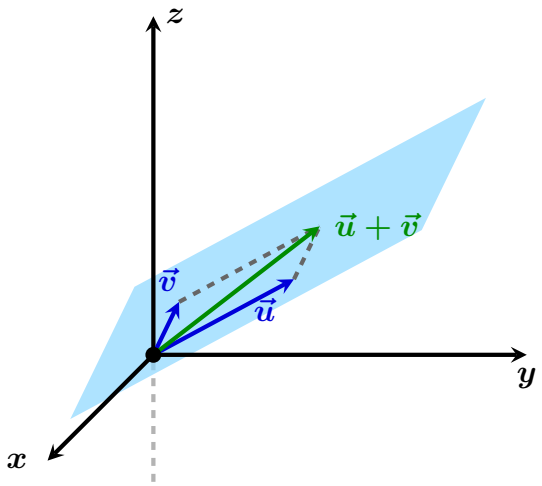
U jest podprzestrzenią liniową V. Skoro $z = 3t - 2x$ oraz $y = 0$, zatem dowolny element $u \in U$ jest postaci $u = (x, 0, 3t - 2x, t)$. Weźmy $u_1 = (x_1, 0, 3t_1 - 2x_1, t_1) \in U$, $u_2 = (x_2, 0, 3t_2 - 2x_2, t_2) \in U$ oraz $\alpha \in \mathbb{R}$, wówczas
$u_1 + u_2 = (x_1 + x_2, 0, 3(t_1 + t_2) - 2(x_1 + x_2), t_1 + t_2) \in U$
oraz $\alpha u_1 = (\alpha x_1, 0, \alpha(3t_1 - 2x_1), \alpha t_1) = (\alpha x_1, 0, 3\alpha t_1 - 2\alpha x_1, \alpha t_1) \in U$

v) $V = (\mathbb{R}[x], +, \mathbb{R}, \cdot)$, $U = \mathbb{R}_n[x] := \{p \in \mathbb{R}[x] : \ \deg p \leq n\}$.
Przyjmujemy, że $\deg \mathbf{0} = -\infty$. Wówczas U jest podprzestrzenią liniową V.

Zatem $\mathbb{R}_n[x]$ to przestrzeń liniowa wielomianów stopnia co najwyżej n-tego. [7]

Podprzestrzenie wektorowe przestrzeni $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$ **i** $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$

Jedynymi właściwymi podprzestrzeniami liniowymi $\mathbb{R}^2$ są proste przechodzące przez $(0, 0)$.
Jedynymi właściwymi podprzestrzeniami liniowymi $\mathbb{R}^3$ są płaszczyzny i proste przechodzące przez $(0, 0, 0)$.



Płaszczyzna przechodząca przez $(0, 0, 0)$ jest podprzestrzenią liniową $\mathbb{R}^3$.
Jest zbiorem zamkniętym na dodawanie wektorów i mnożenie wektorów przez skalar.

NOTACJA:
Jeśli nie wyszczególniono inaczej, symbolem $\mathbb{R}^n$ oznaczamy przestrzeń liniową $(\mathbb{R}^n, +, \mathbb{R}, \cdot)$.

Podobnie $\mathbb{R}[x]$, $\mathbb{R}_n[x]$, $M_{m\times n}(\mathbb{R})$, $\mathcal{C}([a, b], \mathbb{R})$ będą oznaczać
$(\mathbb{R}[x], +, \mathbb{R}, \cdot)$, $(\mathbb{R}_n[x], +, \mathbb{R}, \cdot)$, $(M_{m\times n}(\mathbb{R}), +, \mathbb{R}, \cdot)$, $(\mathcal{C}([a, b], \mathbb{R}), +, \mathbb{R}, \cdot)$ odpowiednio.

6.2 Liniowa niezależność wektorów, baza i wymiar przestrzeni liniowej

Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową nad ciałem K. Niech $\alpha_1, \dots \alpha_m \in K$, $v_1, \dots, v_m \in V$. Niech $W \neq \varnothing$ będzie podzbiorem zbioru V.

[7] Niektórzy autorzy stosują również oznaczenie $\mathcal{P}_n(\mathbb{R})$.

Definicja 6.2.1. i) Wektor $\alpha_1 v_1 + \ldots + \alpha_m v_m \in V$ nazywamy *kombinacją liniową* wektorów $v_1, \ldots, v_m \in V$ o współczynnikach $\alpha_1, \ldots \alpha_m \in K$.

ii) Jeśli $\alpha_1 v_1 + \ldots + \alpha_m v_m = \mathbf{0}$, mówimy, że jest to *kombinacja zerowa*.

iii) Kombinację liniową $\alpha_1 v_1 + \ldots + \alpha_m v_m$ nazywamy *kombinacją trywialną* wtedy i tylko wtedy gdy $\alpha_1 = 0, \ldots, \alpha_m = 0$.

iv) Zbiór $\Big\{v = \alpha_1 w_1 + \ldots + \alpha_k w_k \,:\, \alpha_1, \ldots \alpha_k \in K;\, w_1, \ldots, w_k \in W;\, k \in \mathbb{N}\Big\}$, będący zbiorem wszystkich kombinacji liniowych wszystkich skończonych układów wektorów w zbiorze W, nazywamy *powłoką liniową* zbioru W i oznaczamy symbolem $\mathrm{lin}_K W$ lub krótko $\mathrm{lin} W$.

Gdy W jest zbiorem skończonym $W = \{w_1, \ldots, w_m\}$ piszemy też $\mathrm{lin}\{w_1, \ldots, w_m\}$.
Czyli $\mathrm{lin}\{w_1, \ldots, w_m\} = \{\alpha_1 w_1 + \ldots + \alpha_m w_m : \alpha_1, \ldots \alpha_m \in K\}$.

Przykład 6.2.2. Niech $u = (1, 2, 3), v = (4, 5, 6) \in \mathbb{R}^3$.

$$\mathrm{lin}\{u, v\} = \{au + bv : a, b \in \mathbb{R}\} = \{a(1, 2, 3) + b(4, 5, 6) : a, b \in \mathbb{R}\} = \{(a + 4b, 2a + 5b, 3a + 6b) : a, b \in \mathbb{R}\}$$

Z równości

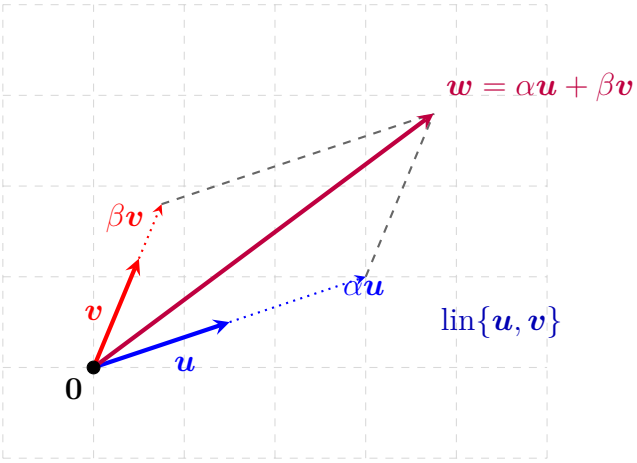
$$2u - v = 2(1, 2, 3) - (4, 5, 6) = (2, 4, 6) - (4, 5, 6) = (-2, -1, 0)$$

wynika, że wektor $w = (-2, -1, 0)$ jest kombinacją liniową wektorów u i v o współczynnikach $2, -1$. Możemy zapisać, że $w \in \mathrm{lin}\{u, v\}$.

Twierdzenie 6.2.3. Niech $V = (V, +, K, \cdot)$ oraz $\varnothing \neq W \subset V$. Wówczas zbiór $\mathrm{lin} W$ jest podprzestrzenią liniową przestrzeni V. Jest to najmniejsza (w sensie relacji inkluzji [a]) podprzestrzeń V zawierająca zbiór W.

[a] Chodzi o relację zawierania się jednego zbioru w drugim.

Definicja 6.2.4. Elementy zbioru W nazywamy *generatorami* przestrzeni $\mathrm{lin} W$, zaś podprzestrzeń $\mathrm{lin} W$ nazywamy podprzestrzenią *generowaną* przez zbiór W.



Wizualizacja kombinacji liniowej dwóch wektorów $\boldsymbol{u}, \boldsymbol{v}$ generujących płaszczyznę $\mathrm{lin}\{\boldsymbol{u}, \boldsymbol{v}\}$.

Przykład 6.2.5. Wersory $\hat{\imath} = (1, 0)$ oraz $\hat{\jmath} = (0, 1)$ generują przestrzeń $\mathbb{R}^2$, bowiem dla dowolnego $\vec{u} = (u_x, u_y) \in \mathbb{R}^2$ mamy $\vec{u} = u_x(1, 0) + u_y(0, 1) = u_x\hat{\imath} + u_y\hat{\jmath}$.

Przykład 6.2.6. Czy wektory $u = (1, 1, -1)$, $v = (2, 1, 0)$, $w = (5, 2, 2)$ generują $\mathbb{R}^3$?

Sprawdzamy, czy dla dowolnego $b = (x, y, z) \in \mathbb{R}^3$ istnieją $\alpha, \beta, \gamma \in \mathbb{R}$ takie, że $b = \alpha u + \beta v + \gamma w$.

$(x, y, z) = \alpha(1, 1, -1) + \beta(2, 1, 0) + \gamma(5, 2, 2) = (\alpha + 2\beta + 5\gamma, \alpha + \beta + 2\gamma, -\alpha + 2\gamma)$

$$\begin{bmatrix} 1 & 2 & 5 \\ 1 & 1 & 2 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Wektory generują $\mathbb{R}^3$, jeśli powyższy układ jest oznaczony.

$$\left[\begin{array}{ccc|c} 1 & 2 & 5 & x \\ 1 & 1 & 2 & y \\ -1 & 0 & 2 & z \end{array}\right] \xrightarrow[w_3+w_1]{w_2-w_1} \left[\begin{array}{ccc|c} 1 & 2 & 5 & x \\ 0 & -1 & -3 & y-x \\ 0 & 2 & 7 & z+x \end{array}\right] \xrightarrow[(-1)\cdot w_2]{w_3+2w_2} \left[\begin{array}{ccc|c} 1 & 2 & 5 & x \\ 0 & 1 & 3 & x-y \\ 0 & 0 & 1 & -x+2y+z \end{array}\right]$$

Układ oznaczony, posiada rozwiązanie $\gamma = -x + 2y + z$, $\beta = x - y - 3\gamma = 4x - 7y - 3z$, $\alpha = x - 2\beta - 5\gamma = -2x + 4y + z$. Zatem układ wektorów u, v, w generuje $\mathbb{R}^3$.

Twierdzenie 6.2.7. Wektory $v_1, \ldots, v_k \in \mathbb{R}^n$ generują $\mathbb{R}^n$ wtedy i tylko wtedy, gdy rząd macierzy A, której kolejne wiersze to współrzędne wektorów $v_1, \ldots, v_k$, jest równy n.

Dowód. Wektory $v_1, \ldots, v_k \in \mathbb{R}^n$ generują $\mathbb{R}^n$, gdy dla dowolnego $x \in \mathbb{R}^n$ układ $x = \alpha_1 v_1 + \ldots + \alpha_k v_k$, w którym niewiadomymi są $\alpha_1, \ldots \alpha_k \in \mathbb{R}$, ma jedyne rozwiązanie. □

Wniosek 6.2.8. Jeśli wektory $v_1, \ldots, v_k \in \mathbb{R}^n$ **generują $\mathbb{R}^n$, to $k \geq n$.**

Przykład 6.2.9. Czy wektory $u = (1, 2, 3)$, $v = (0, 4, 5)$ generują $\mathbb{R}^3$?
Na pewno nie, gdyż $n = 3$, a wektory są tylko dwa.

Wniosek 6.2.10. Jeśli $u = \lambda_1 v_1 + \ldots + \lambda_m v_m$, dla pewnych $v_1, \ldots, v_m \in V$, $\lambda_1, \ldots, \lambda_m \in K$ oraz $\lambda_1 \neq 0$, to wówczas $\text{lin}\{v_1, v_2, \ldots, v_m\} = \text{lin}\{u, v_2, \ldots, v_m\}$

Przykład 6.2.11.

$$\mathbb{R}^2 = \text{lin}\{(1,0),(0,1)\} = \text{lin}\{(1,0),(1,1)\} = \text{lin}\{(2,1),(0,1) = \text{lin}\{(1,0),(0,1),(1,1)\}$$

Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową nad ciałem K.

Definicja 6.2.12. Wektory $v_1, \ldots, v_m \in V$ nazywamy *liniowo niezależnymi* lub mówimy, że tworzą *układ liniowo niezależny*, gdy każda kombinacja zerowa jest trywialna, to znaczy jeśli dla dowolnych skalarów $\beta_1, \ldots, \beta_m \in K$ zachodzi

$$\beta_1 v_1 + \ldots + \beta_m v_m = \mathbf{0} \Rightarrow \beta_1 = \ldots = \beta_m = 0.$$

Wektory, które nie są liniowo niezależne, nazywamy *liniowo zależnymi*.

Przykład 6.2.13. Czy układ $\{A, B, C\}$ jest układem liniowo niezależnym?

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \in M_2(\mathbb{R})$$

Sprawdzamy, czy dowolna kombinacja zerowa jest trywialna.
Niech $\alpha, \beta, \gamma \in \mathbb{R}$ będą dowolne takie, że $\alpha A + \beta B + \gamma C = \mathbf{0}_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Stąd $\begin{bmatrix} \alpha - \beta & -\alpha \\ 0 & \alpha + \beta + \gamma \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow \begin{cases} \alpha = 0 \\ \beta = \alpha = 0 \\ \gamma = -\alpha - \beta = 0 \end{cases}$

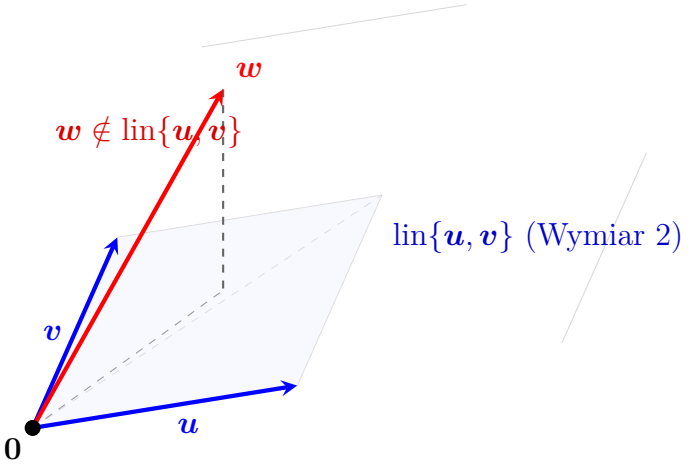
Zatem macierze A, B, C tworzą układ liniowo niezależny.

Przykład 6.2.14. Czy wektory $u = (1,2,3,4)$, $v = (1,2,0,-1)$, $w = (0,1,3,1) \in \mathbb{R}^4$ są liniowo niezależne?

Niech $\alpha, \beta, \gamma \in \mathbb{R}$ będą dowolne takie, że $\alpha u + \beta v + \gamma w = \mathbf{0} = (0,0,0,0)$.
Stąd $(\alpha + \beta, 2\alpha + 2\beta + \gamma, 3\alpha + 3\gamma, 4\alpha - \beta + \gamma) = (0,0,0,0)$.
Wektory u, v, w będą liniowo niezależne, gdy powyższy układ jednorodny ma jedyne rozwiązanie $\alpha = \beta = \gamma = 0$.

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 \\ 3 & 0 & 3 & 0 \\ 4 & -1 & 1 & 0 \end{array}\right] \xrightarrow[\substack{w_3 - 3w_1 \\ w_4 - 4w_1}]{w_2 - 2w_1} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & -5 & 0 & 0 \end{array}\right] \xrightarrow[-\frac{1}{5}w_4]{-\frac{1}{3}w_3} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$$

$\Rightarrow \; r(U) = r(A) = n = 3$
Na mocy twierdzenia Kroneckera-Capellego układ jest oznaczony. Zatem wektory u, v, w są liniowo niezależne.



Geometryczny sens rzędu macierzy $\mathrm{r}(A) = 3$: trzeci wektor posiada niezerową składową prostopadłą do przestrzeni stanów rozpiętej przez pierwsze dwa wektory.

Przykład 6.2.15. Czy funkcje f_1, f_2, f_3 takie, że $f_1(x) = 1$, $f_2(x) = \sin x$, $f_3(x) = \cos x$, tworzą układ liniowo niezależny w $\mathcal{C}(\mathbb{R})$?

Skorzystamy z definicji liniowej niezależności. Niech $C_1, C_2, C_3 \in \mathbb{R}$ będa takie, że $C_1 f_1 + C_2 f_2 + C_3 f_3 = \mathbf{0}$, gdzie $\mathbf{0}$ oznacza funkcję tożsamościowo równą zero.

Skoro $C_1 f_1(x) + C_2 f_2(x) + C_3 f_3(x) = C_1 + C_2 \sin x + C_3 \cos x = 0$ dla dowolnego $x \in \mathbb{R}$,
zatem w szczególności równość jest prawdziwa dla $x = 0$, $x = \frac{\pi}{2}$ oraz $x = \pi$.
Podstawiając te wartości, otrzymujemy

$$C_1 + C_3 = 0 \quad \wedge \quad C_1 + C_2 = 0 \quad \wedge \quad C_1 - C_3 = 0.$$

Stąd $C_1 = C_2 = C_3 = 0$, a zatem układ funkcji f_1, f_2, f_3 jest liniowo niezależny.

Obserwacja: Badanie liniowej niezależności wektorów w $\mathbb{R}^n$ polega na wyliczaniu rzędu macierzy, której kolumnami są podane wektory. Wektory są liniowo niezależne, gdy rząd macierzy równy jest liczbie wektorów.

Wniosek 6.2.16.
Niech $v_1, \ldots, v_k \in \mathbb{R}^n$. Jeśli $k > n$, to wektory $v_1, \ldots, v_k$ są liniowo zależne.
Równoważnie, jeśli wektory $v_1, \ldots, v_k$ są **liniowo niezależne, to $k \leq n$**.

Przykład 6.2.17.
Czy wektory $(1, 2, 3)$, $(1, 0, 2)$, $(-1, -1, 5)$, $(6, 7, 8) \in \mathbb{R}^3$ są liniowo niezależne?
Na pewno nie, gdyż $n = 3$, a wektory są cztery.

Twierdzenie 6.2.18. Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową oraz niech $m \in \mathbb{N}, m \geq 2$. Wektory $v_1, \ldots, v_m \in V$ są liniowo zależne wtedy i tylko wtedy, gdy jeden z nich jest kombinacją liniową pozostałych.

Dowód. Załóżmy, że wektory $v_1, \ldots, v_m$ są liniowo zależne, czyli istnieją $\alpha_1, \ldots \alpha_m \in K$ nie wszystkie równe zeru, takie że $\alpha_1 v_1 + \ldots + \alpha_m v_m = \mathbf{0}$. Bez straty dla ogólności możemy założyć, że $\alpha_1 \neq 0$. Wówczas $v_1 = -\frac{\alpha_2}{\alpha_1} v_2 - \ldots - \frac{\alpha_m}{\alpha_1} v_m$.
Załóżmy teraz, że v_1 jest kombinacja liniową $v_2, \ldots, v_m$, czyli istnieją $\beta_2, \ldots \beta_m \in K$ takie, że $v_1 = \beta_2 v_2 + \ldots + \beta_m v_m$. Wówczas $\beta_1 v_1 + \beta_2 v_2 + \ldots + \beta_m v_m = 0$, gdzie $\beta_1 = -1 \neq 0$. $\square$

Przykład 6.2.19. Czy funkcje f_1, f_2, f_3 takie, że $f_1(x) = 1$, $f_2(x) = \sin^2 x$, $f_3(x) = \cos^2 x$, tworzą układ liniowo niezależny w $\mathcal{C}(\mathbb{R})$?

Nie, ponieważ $\sin^2 x + \cos^2 x = 1$, zatem $f_1 = f_2 + f_3$. Jest to układ liniowo zależny.

Twierdzenie 6.2.20. Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową.

i) Układ $\{v\}$, $v \in V$ jest liniowo zależny wtedy i tylko wtedy, gdy $v = \mathbf{0}$.

ii) Układ wektorów zawierający podukład liniowo zależny jest liniowo zależny.

iii) Jeśli układ wektorów jest liniowo niezależny, to każdy jego podukład jest liniowo niezależny.

iv) Układ wektorów zawierający wektor zerowy jest liniowo zależny.

Przykład 6.2.21. Wektory $u = (1, 2, 3)$, $v = (4, 5, 6)$, $w = (0, 0, 0)$ tworzą układ liniowo zależny, bowiem istnieje nietrywialna kombinacja zerowa postaci $0 \cdot u + 0 \cdot v + 1 \cdot w = 0$.

Przykład 6.2.22. Układ wektorów $\hat{i} = (1, 0), \hat{j} = (0, 1), \hat{i} + \hat{j} = (1, 1)$ generuje przestrzeń $\mathbb{R}^2$, to jest $\mathbb{R}^2 = \text{lin}\{\hat{i}, \hat{j}, \hat{i} + \hat{j}\}$.

Rozważmy wekotr $v = (1, 2) \in \mathbb{R}^2$. Zauważmy, że

$$v = 1 \cdot \hat{i} + 2 \cdot \hat{j}, \qquad v = 1 \cdot \hat{j} + 1 \cdot (\hat{i} + \hat{j}).$$

Wektor v jest kombinacją liniową generatorów przestrzeni $\mathbb{R}^2$, lecz przedstawienie to nie jest jednoznaczne. [8]

Wyróżnimy układy wektorów generujących daną przestrzeń liniową o tej własności, że dowolny wektor z tej przestrzeni przedstawia się jednoznacznie jako ich kombinacja liniowa.

[8] Istnieje więcej niż jeden sposób przedstawienia wektora v jako kombinacji liniowej wektorów $\hat{i}, \hat{j}, \hat{i} + \hat{j}$.

Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową nad ciałem K. Niech $b_1, \ldots, b_n \in V$.

Definicja 6.2.23. Układ wektorów $\mathcal{B} = \{b_1, \ldots, b_n\}$ nazywamy *bazą* przestrzeni V jeśli jest on liniowo niezależny oraz $V = \text{lin}\{b_1, \ldots, b_n\}$.

Wniosek 6.2.24. Jeśli wektory $v_1, \ldots, v_k \in \mathbb{R}^n$ tworzą bazę przestrzeni $\mathbb{R}^n$, to wówczas $k = n$.

Dowód. Tezę otrzymujemy na mocy wniosków 6.2.8 oraz 6.2.16. □

Przykład 6.2.25. Baza przestrzeni $\mathbb{R}^n$

Układ wektorów $\{e_1, \ldots, e_n\}$ stanowi bazę przestrzeni $\mathbb{R}^n$.

$$e_1 = (1, 0, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$$

Przykład 6.2.26. Wskaż bazę podprzestrzeni U przestrzeni $\mathbb{R}^4$, jeśli
$U = \text{lin}\{(1, 3, 2, 1), (1, 2, 1, 1), (1, 1, 0, 1), (1, 2, 2, 1), (3, 4, 1, 3)\}$.

Podane generatory na pewno nie tworzą bazy U, gdyż układ 5 wektorów w przestrzeni $\mathbb{R}^4$ jest liniowo zależny.

$$r\begin{bmatrix} 1 & 1 & 1 & 1 & 3 \\ 3 & 2 & 1 & 2 & 4 \\ 2 & 1 & 0 & 2 & 1 \\ 1 & 1 & 1 & 1 & 3 \end{bmatrix} \leq 4 \neq 5$$

$$r\begin{bmatrix} 1 & 1 & 1 & 1 & 3 \\ 3 & 2 & 1 & 2 & 4 \\ 2 & 1 & 0 & 2 & 1 \\ 1 & 1 & 1 & 1 & 3 \end{bmatrix} = r\begin{bmatrix} 1 & 3 & 2 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 \\ 3 & 4 & 1 & 3 \end{bmatrix} = r\begin{bmatrix} 1 & 3 & 2 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -2 & -2 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -5 & -5 & 0 \end{bmatrix} = r\begin{bmatrix} 1 & 3 & 2 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} = 3 \Rightarrow U =$$

$\text{lin}\{(1, 3, 2, 1), (0, -1, -1, 0), (0, -1, 0, 0)\}$
Układ wektorów $(1, 3, 2, 1), (0, -1, -1, 0), (0, -1, 0, 0)$ jest bazą przestrzeni U
(porównaj wniosek 6.2.10).

Twierdzenie 6.2.27. i) Każda przestrzeń wektorowa różna od $\{\mathbf{0}\}$ posiada bazę.

ii) Wszystkie bazy danej przestrzeni wektorowej są równoliczne. Jeśli baza danej przestrzeni liniowej jest nieskończona, to każda inna jej baza także jest nieskończona.

iii) Każdy układ wektorów liniowo niezależnych przestrzeni wektorowej może być uzupełniony do jej bazy.

Twierdzenie 6.2.28. (Warunki równoważne do bycia bazą przestrzeni wektorowej)

Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową nad ciałem K. Niech $b_1, \ldots, b_n \in V$. Następujące warunki są równoważne.

i) Układ wektorów $b_1, \ldots, b_n$ jest bazą przestrzeni V.

ii) Układ wektorów $b_1, \ldots, b_n$ jest minimalnym (w sensie relacji inkluzji [a]) układem generującym V.

iii) Układ wektorów $b_1, \ldots, b_n$ jest maksymalnym (w sensie relacji inkluzji) układem liniowo niezależnym w tej przestrzeni.

iv) Dowolny wektor $v \in V$ ma jednoznaczne przedstawienie jako kombinacja liniowa elementów $b_1, \ldots, b_n$.

[a] Chodzi o relację zawierania się jednego zbioru w drugim.

Definicja 6.2.29. Liczbę elementów bazy przestrzeni wektorowej $V = (V, +, K, \cdot)$ nazywamy *wymiarem przestrzeni wektorowej* V i oznaczamy $\dim_K V$ lub krótko $\dim V$. Mówimy wówczas, że przestrzeń V jest *n-wymiarowa*. Jeśli żaden skończony układ wektorów nie tworzy bazy przestrzeni V, to przyjmujemy $\dim V = \infty$. Ponadto przyjmujemy $\dim\{\mathbf{0}\} = 0$.

Wniosek 6.2.30.
i) W przestrzeni liniowej n-wymiarowej każdy układ m wektorów, gdzie $m > n$ jest liniowo zależny.

ii) W przestrzeni liniowej n-wymiarowej każdy układ n wektorów liniowo niezależnych stanowi bazę tej przestrzeni.

iii) W przestrzeni liniowej n-wymiarowej każde n wektorów generujących tę przestrzeń stanowi jej bazę.

Jeśli zatem znamy wymiar przestrzeni linowej i wektorów jest dokładnie tyle, ile wynosi wymiar, to wystarczy sprawdzić tylko jeden spośród dwóch warunków z definicji bazy.

Przykład 6.2.6 - raz jeszcze
Wiemy, że wektory $u = (1, 1, -1)$, $v = (2, 1, 0)$, $w = (5, 2, 2)$ generują przestrzeń $\mathbb{R}^3$. Ponadto $\dim \mathbb{R}^3 = 3$, zatem układ $\{u, v, w\}$ jest bazą $\mathbb{R}^3$.

Przykład 6.2.31. Niech $u = (1, 2, 3)$, $v = (4, 5, 6) \in \mathbb{R}^3$. Czy układ $\{u, v\}$ może być uzupełniony do bazy przestrzeni $\mathbb{R}^3$?

Wektory u, v są liniowo niezależne, bowiem $r\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \underset{w_2-4w_1}{=} r\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \end{bmatrix} = 2.$

Jedną z baz przestrzeni $\mathbb{R}^3$ jest $\{\hat{i} = (1,0,0), \hat{j} = (0,1,0), \hat{k} = (0,0,1)\}$.
Baza ta składa się z trzech wektorów.
Na mocy twierdzenia 6.2.27, ii) każda inna baza przestrzeni $\mathbb{R}^3$ składa się z trzech wektorów.

Na mocy twierdzenia 6.2.27, iii) możliwe jest wybranie jeszcze jednego wektora $w \in \mathbb{R}^3$ w taki sposób, by układ $\{u, v, w\}$ stanowił bazę przestrzeni $\mathbb{R}^3$.

Jak to zrobić? Ponieważ $\dim \mathbb{R}^3 = 3$, na mocy twierdzenia 6.2.30 wystarczy dobrać $w \in \mathbb{R}^3$ w taki sposób, by układ $\{u, v, w\}$ był liniowo niezależny.

$$r\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 1 \end{bmatrix} \underset{w_2-4w_1}{=} r\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{bmatrix} = 3$$

Zatem możemy wybrać $w = (0,0,1)$.

Wniosek 6.2.32. Wektory $v_1 = (v_{11}, v_{12}, \ldots, v_{1n}), v_2 = (v_{21}, v_{22}, \ldots, v_{2n}), \ldots, v_n = (v_{n1}, v_{n2}, \ldots, v_{nn})$ tworzą bazę przestrzeni $\mathbb{R}^n$, wtedy i tylko wtedy gdy $\det[v_{ij}] \neq 0$.

Dowód. Tezę otrzymujemy na mocy twierdzeń 6.2.7 oraz 6.2.30 iii). □

Przykład 6.2.6 - po raz trzeci
Wektory $u = (1,1,-1)$, $v = (2,1,0)$, $w = (5,2,2)$ tworzą bazę przestrzeni $\mathbb{R}^3$, bowiem

$$\begin{vmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 5 & 2 & 2 \end{vmatrix} = 2 - 4 + 5 - 4 = -1 \neq 0.$$

Wniosek 6.2.33.

i) Na płaszczyźnie $\mathbb{R}^2$ dwa dowolne wektory niewspółliniowe tworzą jej bazę.

ii) W przestrzeni $\mathbb{R}^3$ trzy dowolne wektory niewspółpłaszczyznowe tworzą jej bazę.

Niech V będzie przestrzenią liniową, zaś $\mathcal{B} = \{b_1, \ldots, b_n\}$ jej bazą.

Definicja 6.2.34. Uporządkowany ciąg wektorów bazowych $(b_1, \ldots, b_n)$ nazywamy *reperem bazowym* lub *bazą uporządkowaną*.

Często mówimy po prostu o *bazie*, zaznaczając w zapisie uporządkowanie wektorów bazowych, np. poprzez ich ponumerowanie.

Bazy standardowe (kanoniczne) wybranych przestrzeni liniowych

1) $(\mathbb{R}^n, +, \mathbb{R}, \cdot)$

$\mathcal{B}_k^n = \Big(e_1, \ldots, e_n\Big)$, gdzie $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$

2) $(\mathbb{R}[x], +, \mathbb{R}, \cdot)$

$\mathcal{B} = \Big(1, x, x^2, x^3, \ldots\Big), \quad \dim \mathbb{R}[x] = \infty, \quad p(x) = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 + \ldots + a_s \cdot x^s$

3) $(\mathbb{R}_n[x], +, \mathbb{R}, \cdot), \quad \mathbb{R}_n[x] = \{f \in \mathbb{R}[x] : \deg f \leq n\}$

$\mathcal{B} = \Big(1, x, x^2, \ldots, x^n\Big), \quad \dim \mathbb{R}_n[x] = n + 1$
dla dowolnego $p \in \mathbb{R}_n[x]$ mamy $p(x) = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 + \ldots + a_n \cdot x^n$.

4) $(M_{m\times n}(\mathbb{R}), +, \mathbb{R}, \cdot)$
$\mathcal{B} = \Big(E_{11}, E_{12}, \ldots, E_{1n}, E_{21}, \ldots, E_{mn}\Big)$, gdzie $E_{kl} = [e_{ij}^{kl}]$, zaś $e_{ij}^{kl} = \begin{cases} 1 & (i,j) = (k,l) \\ 0 & (i,j) \neq (k,l) \end{cases}$
$\dim M_{m\times n}(K) = m \cdot n$

Przykład 6.2.35. Bazą $M_2(\mathbb{R})$ jest układ $(E_{11}, E_{12}, E_{21}, E_{22})$, gdzie

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Dla dowolnej macierzy $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ mamy $A = aE_{11} + bE_{12} + cE_{21} + dE_{22}$.

Twierdzenie 6.2.36. Niech V będzie przestrzenią liniową, zaś W jej podprzestrzenią. Wówczas

i) $W \neq V \Rightarrow \dim W < \dim V$,

ii) $\dim W = \dim V < \infty \Rightarrow W = V$.

Przykład 6.2.37. Niech U, V będą podprzestrzeniami liniowymi przestrzeni $\mathbb{R}^{100}$. Jeśli $U \subseteq W$, $\dim U = 99$ oraz $W \neq \mathbb{R}^{100}$, to wówczas $\dim W = 99$.

Przykład 6.2.38. Podprzestrzenie liniowe przestrzeni $\mathbb{R}, \mathbb{R}^2$ oraz $\mathbb{R}^3$.

	podprzestrzenie $\mathbb{R}$	podprzestrzenie $\mathbb{R}^2$	podprzestrzenie $\mathbb{R}^3$
wymiaru 0	$\{0\}$	$\{(0,0)\}$	$\{(0,0,0)\}$
wymiaru 1	$\mathbb{R}$	proste przechodzące przez $(0,0)$	proste przechodzące przez $(0,0,0)$
wymiaru 2		$\mathbb{R}^2$	płaszczyzny przechodzące przez $(0,0,0)$
wymiaru 3			$\mathbb{R}^3$

Niech V będzie przestrzenią liniową, zaś $\mathcal{B} = (b_1, \ldots, b_n)$ jej bazą uporządkowaną. Wówczas dla każdego $v \in V$ istnieją skalary $\alpha_1, \ldots \alpha_n \in K$ takie, że $v = \alpha_1 b_1 + \ldots + \alpha_n b_n$. Można uzasadnić, że przy ustalonym reperze bazowym skalary $\alpha_1, \ldots \alpha_n$ są wyznaczone jednoznacznie.

Definicja 6.2.39. Skalary $\alpha_1, \ldots \alpha_n \in K$ takie, że $v = \alpha_1 b_1 + \ldots + \alpha_n b_n$ nazywamy *współrzędnymi* wektora v w bazie $\mathcal{B}$. Piszemy wówczas $v = [\alpha_1, \ldots \alpha_n]_{\mathcal{B}}$.

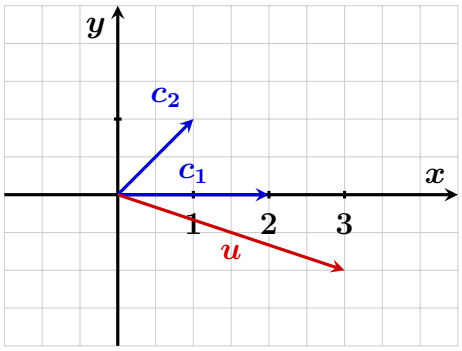
NOTACJA:
Dla $\boldsymbol{v \in \mathbb{R}^n}$ będziemy pisać $\boldsymbol{v = (v_1, v_2, \ldots, v_n)}$ zamiast $\boldsymbol{u = [v_1, v_2, \ldots, v_n]_{\mathcal{B}_k^n}}$.

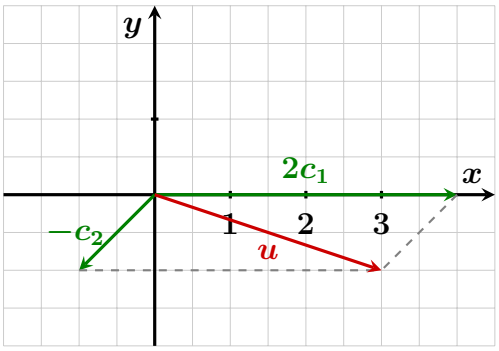
Przykład 6.2.40. i) Niech $p \in \mathbb{R}_2[x]$, $p(x) = ax^2 + bc + c$.
W bazie standardowej $\mathcal{B} = \left(1, x, x^2\right)$ wektor ten ma współrzędne $p = [c, b, a]_{\mathcal{B}}$.

ii) Niech $E_{12} = \begin{bmatrix} -1 & 1 \\ 0 & 5 \end{bmatrix} \in M_2(\mathbb{R})$. W bazie standardowej $\mathcal{B} = \left(E_{11}, E_{12}, E_{21}, E_{22}\right)$ wektor ten ma współrzędne $A = [-1, 1, 0, 5]_{\mathcal{B}}$.

Przykład 6.2.41. $\mathbb{R}^2 = (\mathbb{R}^2, +, \mathbb{R}, \cdot)$
$\mathcal{B}_k^2 = (e_1, e_2))$, gdzie $e_1 = (1,0), e_2 = (0,1)$ to baza kanoniczna.
$\mathcal{C} = (c_1, c_2))$, gdzie $c_1 = (2,0), c_2 = (1,1)$ to inna baza $\mathbb{R}^2$.

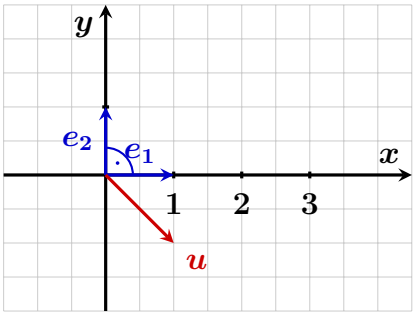
$u = [2, -1]_{\mathcal{C}} = 2c_1 - c_2 = 2 \cdot (2,0) - (1,1) = (3, -1)$

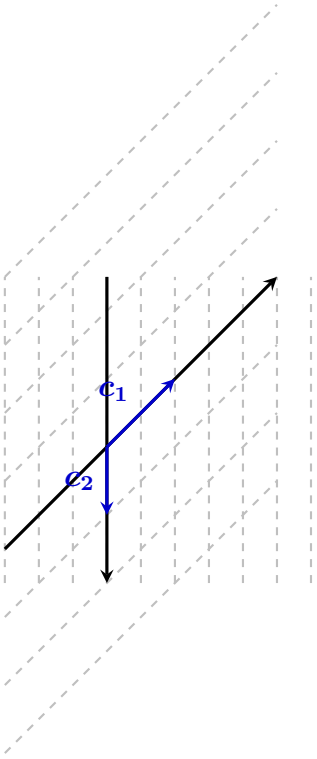


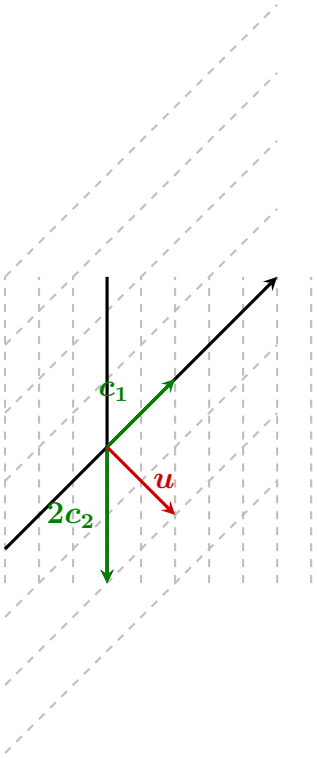


Przykład 6.2.42. $\mathbb{R}^2 = (\mathbb{R}^2, +, \mathbb{R}, \cdot)$
$\mathcal{B}_k^2 = (e_1, e_2))$, gdzie $e_1 = (1, 0), e_2 = (0, 1)$ to baza kanoniczna.
$\mathcal{C} = (c_1, c_2))$, gdzie $c_1 = (1, 1), c_2 = (0, -1)$ to inna baza $\mathbb{R}^2$.

$u = (1, -1)$ oraz $u = [\alpha, \beta]_{\mathcal{C}}$, $\alpha = ?$, $\beta = ?$
$(1, -1) = \alpha c_1 + \beta c_2 = \alpha(1, 1) + \beta(0, -1) = (\alpha, \alpha - \beta)$
Stąd $\alpha = 1$, $\beta = 2$ oraz $u = [1, 2]_{\mathcal{C}}$.







Przykład 6.2.43. $\mathbb{R}^3 = (\mathbb{R}^3, +, \mathbb{R}, \cdot)$, $\mathcal{B} = \mathcal{B}_k^3$ - baza kanoniczna,
$\mathcal{B}' = (b_1', b_2', b_3')$ inna baza, gdzie $b_1' = (4, 2, 1), b_2' = (-5, 2, 3), b_3' = (1, 3, 0)$

Skalary $4, 2, 1$ to współrzędne b_1' w bazie kanonicznej.
Piszemy $b_1' = (4, 2, 1)$ zamiast $b_1' = [4, 2, 1]_{\mathcal{B}_k^3}$.

Ponadto $b'_1 = [1,0,0]_{\mathcal{B}'}$.

Niech $v = (-3,15,7) \in \mathbb{R}^3$. Wyznaczymy współrzędne wektora v w bazie $\mathcal{B}'$.
Niech $v = [\alpha,\beta,\gamma]_{\mathcal{B}'}$, tzn. $v = \alpha b'_1 + \beta b'_2 + \gamma b'_3$. Otrzymujemy

$$(-3,15,7) = \alpha(4,2,1) + \beta(-5,2,3) + \gamma(1,3,0) = (4\alpha - 5\beta + \gamma, 2\alpha + 2\beta + 3\gamma, \alpha + 3\beta).$$

Aby wyznaczyć współrzędne α,β,γ, należy rozwiązać układ równań.

$$\begin{bmatrix} 4 & -5 & 1 \\ 2 & 2 & 3 \\ 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} -3 \\ 15 \\ 7 \end{bmatrix}.$$

$$\left[\begin{array}{ccc|c} 4 & -5 & 1 & -3 \\ 2 & 2 & 3 & 15 \\ 1 & 3 & 0 & 7 \end{array}\right] \xrightarrow{w_3 \leftrightarrow w_1} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 2 & 2 & 3 & 15 \\ 4 & -5 & 1 & -3 \end{array}\right] \xrightarrow[w_3 - 4w_1]{w_2 - 2w_1} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & -4 & 3 & 1 \\ 0 & -17 & 1 & -31 \end{array}\right] \xrightarrow{-4 \cdot w_2} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 16 & -12 & -4 \\ 0 & -17 & 1 & -31 \end{array}\right]$$

$$\xrightarrow{w_3 + w_2} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 16 & -12 & -4 \\ 0 & -1 & -11 & -35 \end{array}\right] \xrightarrow[w_3 \leftrightarrow w_2]{-1 \cdot w_2} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 1 & 11 & 35 \\ 0 & 16 & -12 & -4 \end{array}\right] \xrightarrow{w_3 - 16w_2} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 1 & 11 & 35 \\ 0 & 0 & -188 & -564 \end{array}\right] \xrightarrow{-\frac{1}{188} \cdot w_3}$$

$$\left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 1 & 11 & 35 \\ 0 & 0 & 1 & 3 \end{array}\right] \xrightarrow{w_2 - 11w_3} \left[\begin{array}{ccc|c} 1 & 3 & 0 & 7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \xrightarrow{w_1 - 3w_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \Rightarrow v = [1,2,3]_{\mathcal{B}'}$$

Niech V będzie przestrzenią liniową, zaś $\mathcal{B} = (b_1,\ldots,b_n)$ oraz $\mathcal{B}' = (b'_1,\ldots,b'_n)$ jej dwiema bazami. Wówczas istnieją skalary $\alpha_{ij} \in K$, $i,j \in \{1,2,\ldots,n\}$ takie, że

$$\begin{aligned} b'_1 &= \alpha_{11}b_1 + \alpha_{21}b_2 + \ldots + \alpha_{n1}b_n \\ b'_2 &= \alpha_{12}b_1 + \alpha_{22}b_2 + \ldots + \alpha_{n2}b_n \\ &\ldots \\ b'_n &= \alpha_{1n}b_1 + \alpha_{2n}b_2 + \ldots + \alpha_{nn}b_n \end{aligned}.$$

Definicja 6.2.44. Macierz $P \in M_n(K)$ postaci

$$P = [\alpha_{ij}] = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ & \ldots & \ldots & \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{bmatrix}$$

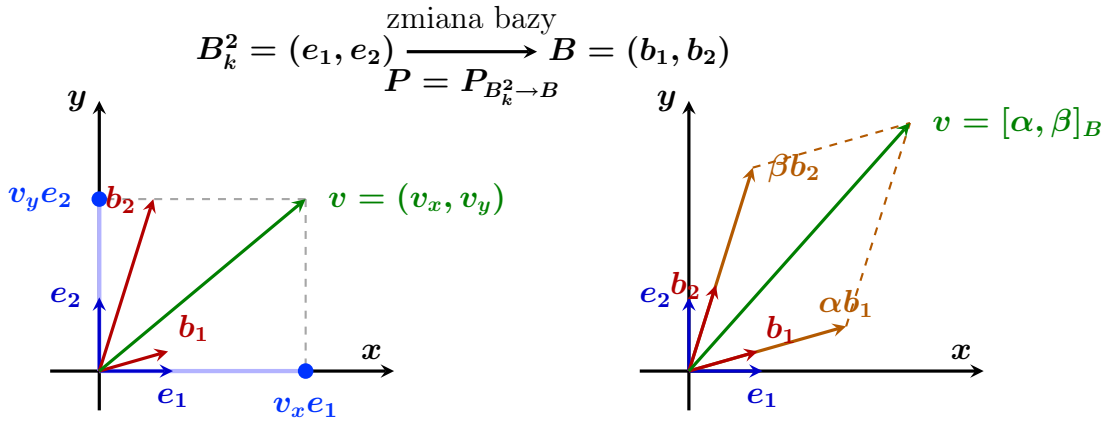
nazywamy *macierzą przejścia* lub *macierzą zmiany bazy* od bazy $\mathcal{B}$ do bazy $\mathcal{B}'$. Oznaczamy ją symbolem $P_{\mathcal{B}\to\mathcal{B}'}$.

Macierz przejścia $P_{\mathcal{B}\to\mathcal{B}'}$ zawsze jest nieosobliwa. Wynika to z faktu, że wektory bazy $\mathcal{B}'$ są liniowo niezależne.

Zmiana współrzędnych wektora przy zmianie bazy

Niech V będzie przestrzenią liniową, zaś $\mathcal{B} = (b_1, \ldots, b_n)$ oraz $\mathcal{B}' = (b'_1, \ldots, b'_n)$ jej dwiema bazami. Niech $P = P_{\mathcal{B} \to \mathcal{B}'}$.

Twierdzenie 6.2.45. Niech $v \in V$, $v = [x_1, \ldots, x_n]_{\mathcal{B}} = [x'_1, \ldots, x'_n]_{\mathcal{B}'}$. Wówczas $X = PX'$, gdzie $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $X' = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix}$.



Przykład 6.2.43 - ciąg dalszy
$\mathbb{R}^3 = (\mathbb{R}^3, +, \mathbb{R}, \cdot)$, $\mathcal{B} = \mathcal{B}_k^3$ - baza kanoniczna,
$\mathcal{B}' = (b'_1, b'_2, b'_3)$ inna baza, gdzie $b'_1 = (4, 2, 1), b'_2 = (-5, 2, 3), b'_3 = (1, 3, 0)$

Wyznaczymy współrzędne wektora $v = (-3, 15, 7)$ w bazie $\mathcal{B}'$.

$P = P_{\mathcal{B}_k^3 \to \mathcal{B}'} = \begin{bmatrix} 4 & -5 & 1 \\ 2 & 2 & 3 \\ 1 & 3 & 0 \end{bmatrix}$, $X' = P^{-1}X = ?$.

Wyznaczamy macierz $P^{-1} = \frac{1}{47} \begin{bmatrix} 9 & -3 & 17 \\ -3 & 1 & 10 \\ -4 & 17 & -18 \end{bmatrix}$ i obliczamy

$X' = \frac{1}{47} \begin{bmatrix} 9 & -3 & 17 \\ -3 & 1 & 10 \\ -4 & 17 & -18 \end{bmatrix} \begin{bmatrix} -3 \\ 15 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Przykład 6.2.46. Rozważmy przestrzeń $\mathbb{R}_2[x]$ oraz jej dwie bazy
$\mathcal{B} = (1+x, x+x^2, 1+x^2)$, $\mathcal{B}' = (1, 1+x, 1+x+x^2)$. Wyznaczmy macierz $P_{\mathcal{B}\to\mathcal{B}'}$.

$1 = [\alpha_1, \beta_1, \gamma_1]_{\mathcal{B}} = \alpha_1(1+x) + \beta_1(x+x^2) + \gamma_1(1+x^2) = (\alpha_1+\gamma_1) + (\alpha_1+\beta_1)x + (\beta_1+\gamma_1)x^2$

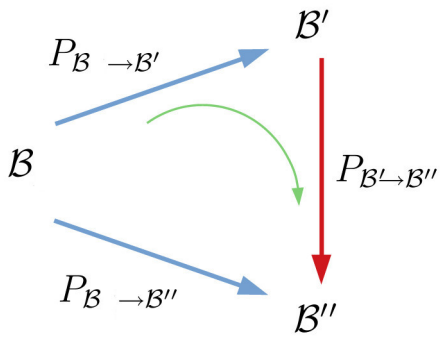
Stąd $\begin{cases} \alpha_1 + \gamma_1 = 1 \\ \alpha_1 + \beta_1 = 0 \\ \beta_1 + \gamma_1 = 0 \end{cases}$ i ostatecznie $\alpha_1 = \frac{1}{2}$, $\beta_1 = -\frac{1}{2}$, $\gamma_1 = \frac{1}{2}$.

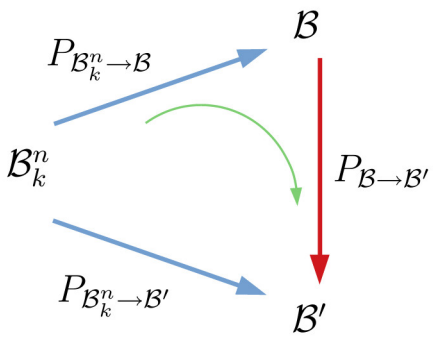
Łatwo zauważyć, że $1 + x = [\alpha_2, \beta_2, \gamma_2]_{\mathcal{B}} = [1, 0, 0]_{\mathcal{B}}$.
Analogicznie obliczenia przeprowadzamy dla $1 + x + x^2 = [\alpha_3, \beta_3, \gamma_3]_{\mathcal{B}}$
i otrzymujemy $P_{\mathcal{B}\to\mathcal{B}'} = \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$.

Twierdzenie 6.2.47. Niech V będzie przestrzenią liniową skończenie wymiarową, zaś $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ jej bazami. Wówczas

i) $P_{\mathcal{B}'\to\mathcal{B}} = \left(P_{\mathcal{B}\to\mathcal{B}'}\right)^{-1}$,

ii) $P_{\mathcal{B}\to\mathcal{B}'} \cdot P_{\mathcal{B}'\to\mathcal{B}''} = P_{\mathcal{B}\to\mathcal{B}''}$.





Uwaga 6.2.48. W przypadku przestrzeni $\mathbb{R}^n$, gdy baza początkowa to baza kanoniczna $\mathcal{B}_k^n$, łatwo wypisać macierze przejścia do nowych baz $P_{\mathcal{B}_k^n\to\mathcal{B}}$ oraz $P_{\mathcal{B}_k^n\to\mathcal{B}'}$. Wypisanie macierzy $P_{\mathcal{B}\to\mathcal{B}'}$ wymaga rachunków. Na mocy twierdzenia 6.2.47 otrzymujemy

$$P_{\mathcal{B}_k^n\to\mathcal{B}'} = P_{\mathcal{B}_k^n\to\mathcal{B}} \cdot P_{\mathcal{B}\to\mathcal{B}'} \qquad \Rightarrow \qquad P_{\mathcal{B}\to\mathcal{B}'} = P^{-1}_{\mathcal{B}_k^n\to\mathcal{B}} \cdot P_{\mathcal{B}_k^n\to\mathcal{B}'}.$$

Q U I Z

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. Zbiór $U = \{(x, y) : y = 3x^2\}$ jest podprzestrzenią liniową przestrzeni $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$.
2. Płaszczyzna $\pi : 7x + 20y + 27z = 0$ jest podprzestrzenią liniową przestrzeni $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$.
3. Przestrzeń liniowa $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$ ma nieskończenie wiele dwuwymiarowych podprzestrzeni liniowych.
4. Zbiór wielomianów rzeczywistych podzielnych przez $x^4 + 1$ jest podprzestrzenią liniową przestrzeni $(\mathbb{R}[x], +, \mathbb{R}, \cdot)$.
5. Zbiór $Z = \{A \in M_3(\mathbb{R}) : r(A) = 1\}$ jest podprzestrzenią liniową przestrzeni $(\dim M_3(\mathbb{R}), +, \mathbb{R}, \cdot)$.
6. Jeśli U jest podprzestrzenią liniową przestrzeni $(\mathbb{R}^7, +, \mathbb{R}, \cdot)$ taką, że $\dim U = 7$, to wówczas $U = \mathbb{R}^7$.
7. Istnieje podprzestrzeń liniowa przestrzeni $(\mathbb{R}[x], +, \mathbb{R}, \cdot)$ zawierająca dokładnie 10 różnych wektorów.
8. Niech S będzie dowolnym zbiorem wektorów z przestrzeni $(\mathbb{R}^6, +, \mathbb{R}, \cdot)$.
   Jeśli $(0, 0, 0, 0, 0, 0) \in S$, to wówczas układ S jest liniowo zależny.
9. Układ $\{f = \sin^2 x,\ g = \cos^2 x,\ h = \cos 2x\}$ jest liniowo niezależny w przestrzeni $(\mathcal{C}^\infty(\mathbb{R}), +, \mathbb{R}, \cdot)$ funkcji klasy $\mathcal{C}^\infty$ w $\mathbb{R}$.
10. Układ $\{p = 2,\ q = x + 2,\ r = x^2 + 2,\ s = x^2 + x + 100\}$ jest liniowo niezależny w przestrzeni $(\mathbb{R}_2[x], +, \mathbb{R}, \cdot)$.
11. Jeśli wektory u, v, w tworzą układ liniowo zależny, to każdy z nich jest kombinacją liniową dwóch pozostałych.
12. Wektory $u = (1, 1, 1),\ v = (0, 2, 2),\ w = (0, 0, 3)$ generują $(\mathbb{R}^3, +, \mathbb{R}, \cdot)$.
13. $\dim \mathbb{R}_7[x] = 8$
14. Dowolna baza przestrzeni $(M_{3\times 4}(\mathbb{R}), +, \mathbb{R}, \cdot)$ jest siedmioelementowa.
15. Macierz przejścia od jednej bazy przestrzeni linowej do drugiej bazy zawsze jest nieosobliwa.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Algebra
Wykład nr 7
Data: .....

TEMAT: *Przekształcenia liniowe*

7.1 Definicja przekształcenia liniowego i podstawowe własności

Niech $V = (V, +, K, \cdot)$ oraz $W = (W, \oplus, K, \odot)$ będą przestrzeniami wektorowymi nad tym samym ciałem K.

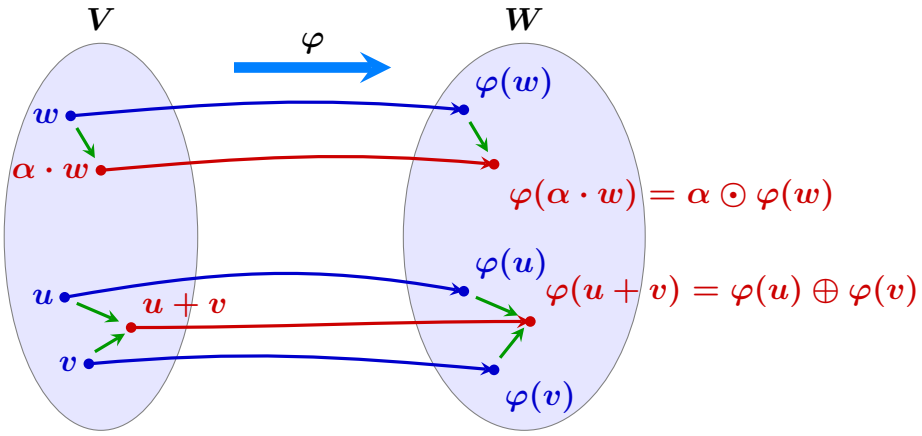
Definicja 7.1.1. Odwzorowanie $\varphi : V \to W$ spełniające warunki

i) własność addytywności $\forall u, v \in V \quad \varphi(u + v) = \varphi(u) \oplus \varphi(v)$

ii) własność jednorodności $\forall v \in V \quad \forall \alpha \in K \quad \varphi(\alpha \cdot v) = \alpha \odot \varphi(v)$,

nazywamy *odwzorowaniem liniowym* lub *przekształceniem liniowym* lub *homomorfizmem przestrzeni liniowych.* [a]

[a] Odwzorowanie to wektorowi przypisuje wektor.



Odwzorowanie liniowe

Twierdzenie 7.1.2. Niech $\varphi : V \to W$. Wówczas następujące warunki są równoważne.

i) φ jest liniowe

ii) $\forall u, v \in V \quad \forall \alpha, \beta \in K \quad \varphi(\alpha \cdot u + \beta \cdot v) = \alpha \odot \varphi(u) \oplus \beta \odot \varphi(v)$

iii) $\forall v_1, \ldots, v_n \in V \quad \forall \alpha_1, \ldots, \alpha_n \in K$
$\varphi(\alpha_1 \cdot v_1 + \ldots + \alpha_n \cdot v_n) = \alpha_1 \odot \varphi(v_1) \oplus \ldots \oplus \alpha_n \odot \varphi(v_n)$

Zbiór wszystkich odwzorowań liniowych odwzorowujących V w W oznaczamy $\mathcal{L}_K(V,W)$ lub $Hom_K(V,W)$ (lub krótko $\mathcal{L}(V,W)$ lub $Hom(V,W)$, gdy wiemy, z jakim ciałem mamy do czynienia).

Uwaga 7.1.3. Często notujemy $V = (V,+,K,\cdot)$ oraz $W = (W,+,K,\cdot)$, to znaczy używamy tych samych symboli dla działań w przestrzeniach V i W, mimo że są to różne działania.

Przykład 7.1.4. Czy φ jest liniowe?

1) $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(x) = ax$, gdzie $a \in \mathbb{R}$ ustalone

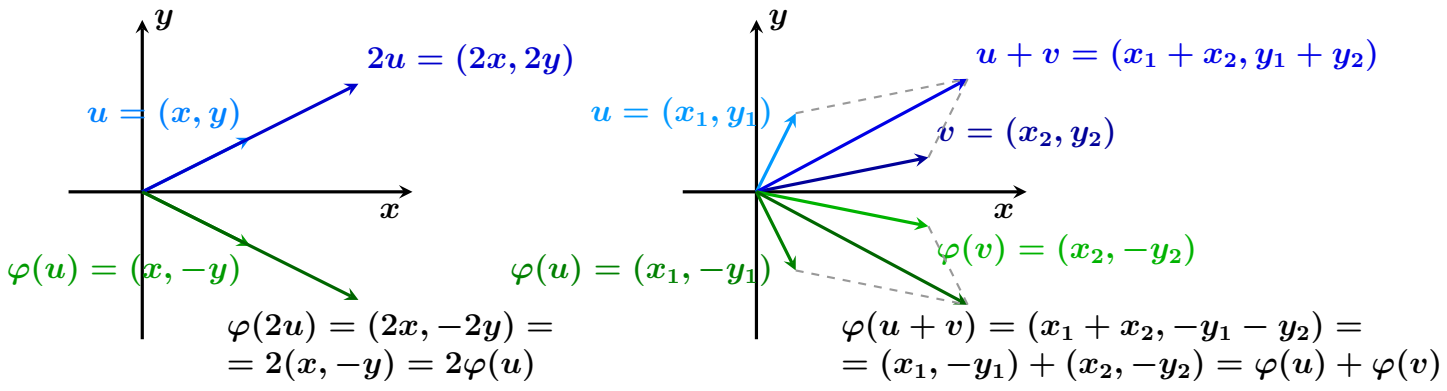
Odwzorowanie jest liniowe, bowiem dla dowolnych $\alpha \in \mathbb{R}$, $x_1, x_2 \in R$ mamy
$\varphi(\alpha x) = a(\alpha x) = \alpha(ax) = \alpha\varphi(x)$ oraz
$\varphi(x_1 + x_2) = a(x_1 + x_2) = ax_1 + ax_2 = \varphi(x_1) + \varphi(x_2)$.

2) $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(x) = ax + b$, gdzie $a, b \in \mathbb{R}$ ustalone

Jeśli $b \neq 0$, to odwzorowanie nie jest liniowe, bowiem
$\varphi(1+1) = \varphi(2) = 2a + b \neq \varphi(1) + \varphi(1) = (a+b) + (a+b) = 2a + 2b$.

3) $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$, symetria względem osi Ox

Ponieważ $\varphi(x,y) = (x,-y)$, zatem dla dowolnych $\alpha \in \mathbb{R}$, $(x,y),(x_1,y_1),(x_2,y_2) \in \mathbb{R}^2$
$\varphi\left(\alpha(x,y)\right) = \varphi(\alpha x, \alpha y) = (\alpha x, -\alpha y) = \alpha(x,-y) = \alpha\varphi(x,y)$ oraz
$\varphi\Big((x_1,y_1)+(x_2,y_2)\Big) = \varphi(x_1+x_2, y_1+y_2) = (x_1+x_2, -(y_1+y_2)) = (x_1+x_2, -y_1-y_2) = (x_1,-y_1) + (x_2,-y_2) = \varphi(x_1,y_1) + \varphi(x_2,y_2)$.
Odwzorowanie jest liniowe.



symetria względem osi $\boldsymbol{Ox}$ jako odwzorowanie liniowe

4) $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$, $\varphi(x, y, z) = (x - y + z, 2y + z + 1)$

Odwzorowanie nie jest liniowe.
$L = \varphi\Big(\alpha(x, y, z)\Big) = \varphi(\alpha x, \alpha y, \alpha z) = (\alpha x - \alpha y + \alpha z, 2\alpha y + \alpha z + 1)$
$P = \alpha\varphi(x, y, z) = \alpha(x - z + z, 2y + z + 1) = (\alpha x - \alpha y + \alpha z, 2\alpha y + \alpha z + \alpha)$
Na ogół $L \neq P$. Możemy podać kontrprzykład
$\varphi\Big(5 \cdot (1, 0, 0)\Big) = \varphi(5, 0, 0) = (5, 1) \neq 5 \cdot \varphi(1, 0, 0) = 5 \cdot (1, 1) = (5, 5)$

Uwaga 7.1.5. i) Odwzorowanie $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ jest liniowe wtedy i tylko wtedy, gdy istnieją $a, b, c, d \in \mathbb{R}$ takie, że $\boxed{\varphi(x, y) = (ax + by, cx + dy).}$

ii) Odwzorowanie $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ jest liniowe wtedy i tylko wtedy, gdy istnieją $a, b, c, d, e, f, g, h, j \in \mathbb{R}$ takie, że $\boxed{\varphi(x, y, z) = (ax+by+cz, dx+ey+fz, gx+hy+jz).}$

Przykład 7.1.6. Czy odwzorowanie $\varphi : \mathbb{R}_2[x] \to \mathbb{R}_1[x]$, dane wzorem $\varphi(p)(x) = (3 - x)p''(x) + 4p'(x)$, dla dowolnego $p \in \mathbb{R}_2[x]$, jest liniowe?

Sprawdzimy, że φ jest liniowe. Wynika to z liniowości różniczkowania.
Dla dowolnych $p, q \in \mathbb{R}_2[x]$ mamy
$\varphi(p + q)(x) = (3 - x)(p + q)''(x) + 4(p + q)'(x) = (3 - x)\Big(p''(x) + q''(x)\Big) + 4\Big(p'(x) + q'(x)\Big) =$
$\Big((3 - x)p''(x) + 4p'(x)\Big) + \Big((3 - x)q''(x) + 4q'(x)\Big) = \varphi(p)(x) + \varphi(q)(x),$
dla dowolnego $x \in \mathbb{R}$.
Zatem $\varphi(p + q) = \varphi(p) + \varphi(q)$.

Dla dowolnych $p \in \mathbb{R}_2[x], \alpha \in \mathbb{R}$ mamy
$\varphi(\alpha p)(x) = (3-x)(\alpha p)''(x)+4(\alpha p)'(x) = (3-x)\alpha \cdot p''(x)+4\alpha \cdot p'(x) = \alpha \cdot \Big((3-x)p''(x)+4p'(x)\Big) =$
$\alpha \cdot \varphi(p)(x)$. dla dowolnego $x \in \mathbb{R}$.
Zatem $\varphi(\alpha p) = \alpha\varphi(p)$.

Niech $V = (V, +, K, \cdot)$ oraz $W = (W, \oplus, K, \odot)$.

Twierdzenie 7.1.7. Jeśli odwzorowanie $\varphi : V \to W$ jest liniowe, to wówczas

i) $\varphi(\mathbf{0}_V) = \mathbf{0}_W$,

ii) $\forall v \in V\ \varphi(-v) = -\varphi(v)$.

Wniosek 7.1.8. Niech $\varphi : V \to W$. Jeśli $\varphi(\mathbf{0}_V) \neq \mathbf{0}_W$, to φ nie jest liniowe.

Dowód. Teza wynika z twierdzenia 7.1.7 i) na mocy prawa kontrapozycji. $\square$

Przykład 7.1.9.

i) Odwzorowanie $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2x + 5$ nie jest liniowe, bowiem $f(0) = 5 \neq 0$.

ii) Odwzorowanie $g : \mathbb{R}[x] \to \mathbb{R}[x]$, $g(p)(x) = x + p'(x)$ nie jest liniowe, bowiem $f(0 + 0x + 0x^2 + \ldots) = x \neq 0 + 0x + 0x^2 + \ldots$.

iii) Odwzorowanie $h : \mathbb{R}^2 \to \mathbb{R}^2$, $h(x, y) = (x + y, xy)$ spełnia warunek $h(0, 0) = (0, 0)$. Nie jest jednak liniowe, bowiem $2 \cdot h(1, 1) = 2 \cdot (2, 1) = (4, 2) \neq h(2, 2) = (4, 4)$.

Definicja 7.1.10. Odwzorowanie liniowe $\varphi : V \to W$ nazywamy:

i) *monomorfizmem*, jeśli φ jest injekcją,

ii) *epimorfizmem*, jeśli φ jest surjekcją,

iii) *izomorfizmem*, jeśli φ jest bijekcją,

iv) *endomorfizmem*, jeśli $V = W$,

v) *automorfizmem*, jeśli $V = W$ i φ jest bijekcją,

vi) *formą liniową*, jeśli $W = K$.

Twierdzenie 7.1.11. Niech V oraz W będą przestrzeniami wektorowymi nad tym samym ciałem K. Niech $(b_1, \ldots, b_n)$ będzie bazą przestrzeni V oraz niech $w_1, \ldots, w_n \in W$ będzie dowolnym układem wektorów. Wówczas istnieje dokładnie jedno odwzorowanie liniowe $\varphi \in \mathcal{L}(V, W)$ takie, że $\varphi(b_i) = w_i$, dla $i \in \{1, 2, \ldots, n\}$.

Uwaga 7.1.12. Z powyższego twierdzenia wynika, że aby w pełni określić odwzorowanie liniowe na przestrzeni liniowej V wystarczy określić obrazy wektorów bazowych $b_1, \ldots, b_n$ przestrzeni V.

$$\varphi(v) = \varphi(\alpha_1 b_1 + \ldots + \alpha_n b_n) = \alpha_1 \varphi(b_1) + \ldots + \alpha_n \varphi(b_n) \quad \text{dla} \quad v = \alpha_1 b_1 + \ldots + \alpha_n b_n \in V.$$

Przykład 7.1.13.

i) Podamy wzór odwzorowania liniowego $\varphi : \mathbb{R} \to \mathbb{R}$, o którym wiadomo, że $\varphi(1) = 3$.

$$\varphi(x) = \varphi\Big(x \cdot 1\Big) = x \cdot \varphi(1) = 3x.$$

Jeśli pomyślimy o φ jako o funkcji, to jej wykresem jest prosta przechodząca przez punkty $(0, 0)$ oraz $(1, 3)$. Informacje te jednoznacznie określają, jaka to prosta. [9]

[9] Gdybyśmy rozważyli funkcję $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3x^2$, to informacja o tym, że $f(0) = 0$ i $f(1) = 3$ w żaden sposób nie pozwala określić, że krzywą przechodzącą przez te dwa punkty jest parabola. Wynika to z faktu, że f nie jest odwzorowaniem liniowym.

ii) Podamy wzór odwzorowania liniowego $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$, o którym wiadomo, że $\varphi(1,0) = (2,0)$, $\varphi(0,1) = (0,1)$.

$$\varphi(x,y) = \varphi\Big(x(1,0) + y(0,1)\Big) = x\varphi(1,0) + y\varphi(0,1) = x(2,0) + y(0,1) = (2x, y)$$

iii) Podamy wzór odwzorowania liniowego $\varphi : \mathbb{R}_2[x] \to \mathbb{R}_1[x]$ takiego, że

$$\varphi(x^2 + x) = 6x + 10, \quad \varphi(x - 1) = 4, \quad \varphi(2x) = 8.$$

W przestrzeni wektorowej $\mathbb{R}_2[x]$ bazą standardową jest $\mathcal{B} = (1, x, x^2)$. Mamy
$$\begin{cases} \varphi(x^2 + x) = \varphi(x^2) + \varphi(x) = 6x + 10 \\ \varphi(x - 1) = \varphi(x) - \varphi(1) = 4 \\ \varphi(2x) = 2\varphi(x) = 8 \end{cases}.$$
Stąd $\varphi(x) = 4$, $\varphi(1) = \varphi(x) - 4 = 0$, $\varphi(x^2) = 6x + 10 - \varphi(x) = 6x + 6$.

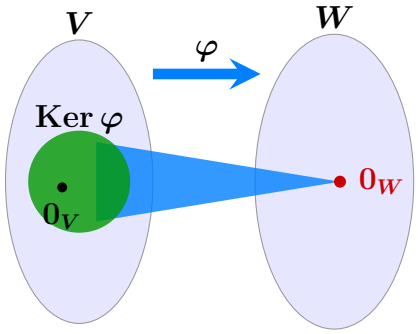
Dowolny $p \in \mathbb{R}_2[x]$ jest postaci $p(x) = ax^2 + bx + c$, dla pewnych $a, b, c \in \mathbb{R}$. Zatem $\varphi(ax^2 + bx + c) = a\varphi(x^2) + b\varphi(x) + c\varphi(1) = a \cdot (6x + 6) + b \cdot 4 + c \cdot 0 = 6ax + 6a + 4b$.

7.2 Jądro, obraz i rząd odwzorowania liniowego

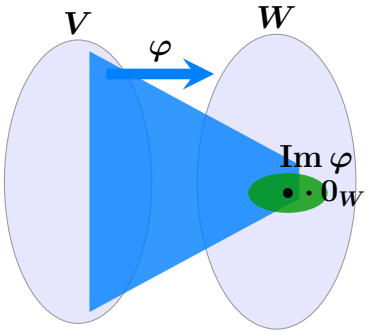
Niech V oraz W będą przestrzeniami liniowymi nad ciałem K. Niech $\varphi \in \mathcal{L}(V, W)$.

Definicja 7.2.1. i) Zbiór $\{v \in V : \varphi(v) = \mathbf{0}_W\} \subseteq V$ nazywamy *jądrem* odwzorowania liniowego φ i oznaczamy $\mathrm{Ker}\varphi$.

ii) Zbiór $\{w \in W : \exists v \in V\, \varphi(v) = w\} = \{\varphi(v) : v \in V\} \subseteq W$ nazywamy *obrazem* odwzorowania liniowego φ i oznaczamy $\mathrm{Im}\varphi$ lub $\varphi(V)$.



jądro odwzorowania liniowego



obraz odwzorowania liniowego

Uwaga 7.2.2. Dla dowolnego odwzorowania liniowego $\varphi : V \to W$ zbiory $\mathrm{Ker}\varphi$ oraz $\mathrm{Im}\varphi$ są niepuste.

Dowód. Ponieważ $\varphi(\mathbf{0}_V) = \mathbf{0}_W$, zatem $\mathbf{0}_V \in \mathrm{Ker}\varphi$ oraz $\mathbf{0}_W \in \mathrm{Im}\varphi$. $\square$

Twierdzenie 7.2.3. Dla dowolnego $\varphi \in \mathcal{L}(V, W)$ zbiór $\mathrm{Ker}\varphi$ jest podprzestrzenią liniową przestrzeni V, zaś zbiór $\mathrm{Im}\varphi$ jest podprzestrzenią liniową przestrzeni W.

Niech V oraz W będą przestrzeniami wektorowymi nad ciałem K.

Twierdzenie 7.2.4. Dla dowolnego $\varphi \in \mathcal{L}(V, W)$

i) φ jest injekcją $\Leftrightarrow \mathrm{Ker}\varphi = \{\mathbf{0}_V\}$,

ii) φ jest surjekcją $\Leftrightarrow \mathrm{Im}\varphi = W$.

Definicja 7.2.5. Jeśli $\dim \mathrm{Im}\varphi < \infty$, to liczbę tę nazywamy *rzędem* odwzorowania liniowego $\varphi : V \to W$ i oznaczamy $r(\varphi)$ lub $\mathrm{rank}(\varphi)$.

Twierdzenie 7.2.6. (Twierdzenie o rzędzie, Rank-nullity theorem) Niech V oraz W będą przestrzeniami wektorowymi skończenie wymiarowymi nad ciałem K. Niech $\varphi \in \mathcal{L}(V, W)$. Wówczas

$$\boxed{r(\varphi) + \dim \mathrm{Ker}\varphi = \dim V.}$$

Wniosek 7.2.7. Dla dowolnego $\varphi \in \mathcal{L}(V, W)$ mamy $\dim \mathrm{Im}\varphi \leq \dim V$.

Przykład 7.2.8. Niech $f : \mathbb{R}^5 \to \mathbb{R}^5$ będzie dowolnym odwzorowaniem liniowym.
Na mocy twierdzenia o rzędzie $\dim Ker f + \dim Im f = 5$.
Ponieważ 5 jest liczbą nieparzystą, zatem $\dim Ker f \neq \dim Im f$ i w konsekwencji $Ker f \neq Im f$.

Przykład 7.2.9. Czy odwzorowanie liniowe $f : \mathbb{R}^3 \to \mathbb{R}^5$ może być epimorfizmem?
Aby f było epimorfizmem, spełniony musiałby być warunek $Im f = \mathbb{R}^5$.
Na mocy twierdzenia o rzędzie, $\dim Im f = \dim \mathbb{R}^3 - \dim Ker f \leq 3$.
Zatem f nie może być epimorfizmem.

Przykład 7.2.10.

$$\varphi : \mathbb{R}^4 \to \mathbb{R}^3, \quad \varphi(x, y, z, t) = (x + y + z + 2t, x - y + z + 6t, x + y - z - 4t)$$

Wyznacz jądro oraz obraz φ, ich bazy i wymiary. Podaj własności φ.

$$\varphi(x,y,z,t)=(0,0,0) \Leftrightarrow \begin{cases} x+y+z+2t=0 \\ x-y+z+6t=0 \\ x+y-z-4t=0 \end{cases}$$

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 2 & 0 \\ 1 & -1 & 1 & 6 & 0 \\ 1 & 1 & -1 & -4 & 0 \end{array}\right] \xrightarrow[w_3-w_1]{w_2-w_1} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 2 & 0 \\ 0 & -2 & 0 & 4 & 0 \\ 0 & 0 & -2 & -6 & 0 \end{array}\right] \xrightarrow[-\frac{1}{2}w_3]{-\frac{1}{2}w_2} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 2 & 0 \\ 0 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & 0 \end{array}\right] \xrightarrow{w_1-w_3}$$

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & 0 \end{array}\right] \xrightarrow{w_1-w_2} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 3 & 0 \end{array}\right] \Rightarrow \begin{cases} x=-t \\ y=2t \\ z=-3t \\ t\in\mathbb{R} \end{cases}$$

$\mathrm{Ker}\varphi=\{(-t,2t,-3t,t),\, t\in\mathbb{R}\}=\mathrm{lin}\{(-1,2,-3,1)\}$

Układ $\{(-1,2,-3,1)\}$ jest bazą $\mathrm{Ker}\varphi$ oraz $\dim\mathrm{Ker}\varphi=1$.
Zatem φ nie jest monomorfizmem.

Ponadto $r(\varphi)=\dim\mathbb{R}^4-\dim\mathrm{Ker}\varphi=4-1=3=\dim\mathbb{R}^3$, zatem φ jest epimorfizmem.
Stąd na mocy twierdzenia 6.2.36 otrzymujemy $\mathrm{Im}\varphi=\mathbb{R}^3$.

Można to również sprawdzić bezpośrednim rachunkiem.
$\varphi(x,y,z,t)=x(1,1,1)+y(1,-1,1)+z(1,1,-1)+t(2,6,-4)$
$\mathrm{Im}\varphi=\mathrm{lin}\{(1,1,1),(1,-1,1),(1,1,-1),(2,6,-4)\}$

$$r\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 2 & 6 & -4 \end{bmatrix} = r\begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \\ 0 & 4 & 6 \end{bmatrix} = 3 \Rightarrow \mathrm{Im}\varphi=\mathrm{lin}\{(1,1,1),(0,-2,0),(0,0,-2)\}$$

Układ $\{(1,1,1),(0,-2,0),(0,0,-2)\}$ jest bazą przestrzeni $\mathrm{Im}\varphi$.
$\mathrm{Im}\varphi\subseteq\mathbb{R}^3 \wedge \dim\mathrm{Im}\varphi=3=\dim\mathbb{R}^3 \Rightarrow \mathrm{Im}\varphi=\mathbb{R}^3$.

Twierdzenie 7.2.11. Niech V oraz W będą przestrzeniami wektorowymi skończenie wymiarowymi nad ciałem K. Niech $\varphi\in\mathcal{L}(V,W)$ będzie injekcją, zaś wektory $v_1,\ldots,v_n\in V$ tworzą układ liniowo niezależny. Wówczas wektory $\varphi(v_1),\ldots,\varphi(v_n)\in W$ również tworzą układ liniowo niezależny.

Wniosek 7.2.12. Niech V oraz W będą przestrzeniami wektorowymi skończenie wymiarowymi nad ciałem K takimi, że $\dim V=\dim W=n$. Niech $\varphi\in\mathcal{L}(V,W)$ będzie injekcją, zaś wektory $v_1,\ldots,v_n\in V$ tworzą bazę przestrzeni V. Wówczas wektory $\varphi(v_1),\ldots,\varphi(v_n)\in W$ tworzą bazę przestrzeni W.

Przykład 7.2.13. Dane jest odwzorowanie liniowe $f:\mathbb{R}_2[x]\to\mathbb{R}_2[x]$, $f(p)(x)=p(x)+p'(x)$.
Sprawdzimy, czy f jest monomorfizmem.
Jeśli $p\in\mathbb{R}_2[x]$, to wówczas $p(x)=ax^2+bx+c$, dla pewnych $a,b,c\in\mathbb{R}$.

Zatem
$$f(ax^2+bx+c) = ax^2+bx+c+2ax+b = ax^2+(2a+b)x+b+c.$$
Zatem $f(p)$ jest wielomianem zerowym wtedy i tylko wtedy gdy $\begin{cases} a=0 \\ 2a+b=0 \\ b+c=0 \end{cases}$, a zatem gdy $a=b=c=0$, czyli gdy p jest wielomianem zerowym.

Stąd $Kerf = \{\mathbf{0} = 0x^2+0x+0\}$. Na mocy twierdzenia 7.2.4, f jest monomorfizmem.
Układ wektorów $\{p=1, q=x, r=x^2\}$ jest liniowo niezależny w $\mathbb{R}_2[x]$.
Ponieważ f jest monomorfizmem, zatem układ

$$\{f(p)(x)=1, \quad f(q)(x)=x+1, \quad f(r)(x)=x^2+2x\}$$

również jest liniowo niezależny w $\mathbb{R}_2[x]$.

Przykład 7.2.14.
Jeśli odwzorowanie liniowe nie jest injektywne, teza twierdzenia 7.2.12 nie zachodzi.
Rozważmy $f:\mathbb{R}^2\to\mathbb{R}^2$ będące rzutowaniem prostokątnym na oś Ox.
Odwzorowanie to dane jest wzorem $f(x,y)=(x,0)$.
Odwzorowanie f nie jest injekcją, bowiem $f(1,1)=f(1,2)=(1,0)$.
Obrazami wektorów bazy kanonicznej przestrzeni $\mathbb{R}^2$ poprzez f są wektory

$$f(1,0)=(1,0), \quad f(0,1)=(0,0).$$

Układ $\{(1,0),(0,0)\}$ jest liniowo zależny, jako układ zawierający wektor zerowy (patrz twierdzenie 6.2.20).

7.3 Działania na odwzorowaniach liniowych

Niech U,V, W będą przestrzeniami wektorowymi nad ciałem K.

Twierdzenie 7.3.1. Zbiór $\mathcal{L}(V,W)$ wraz z działaniami

$$+ : \mathcal{L}(V,W)\times\mathcal{L}(V,W)\to\mathcal{L}(V,W), \quad \cdot : K\times\mathcal{L}(V,W)\to\mathcal{L}(V,W)$$

określonymi wzorami $(f+g)(x)=f(x)+g(x)$, $(\alpha\cdot f)(x)=\alpha\cdot f(x)$,
jest przestrzenią liniową nad ciałem K.

Niech $f\in\mathcal{L}(V,W)$ będzie bijekcją. [10] *Odwzorowaniem odwrotnym* do f nazywamy odwzorowanie liniowe $g\in\mathcal{L}(W,V)$ takie, że

$$\forall v\in V \ \ \forall w\in W \ \ f(v)=w \Leftrightarrow g(w)=v.$$

[10] W szczególności wynika stąd, że $\dim V = \dim W$.

Jesli takie odwzorowanie istnieje, to jest jedyne. Oznaczamy je symbolem f^{-1}.

Twierdzenie 7.3.2.

i) Jeśli $f \in \mathcal{L}(U,V)$ oraz $g \in \mathcal{L}(V,W)$, to wówczas $g \circ f \in \mathcal{L}(U,W)$. [a]

ii) Jeśli $f \in \mathcal{L}(V,W)$ jest bijekcją, to $f^{-1} \in \mathcal{L}(W,V)$. [b]

[a] Innymi słowy, złożenie odwzorowań liniowych jest odwzorowaniem liniowym.
[b] Innymi słowy, jesli f jest izomorfizmem, to f^{-1} również jest izomorfizmem.

Oznaczmy przez $Aut_K(V) = \{f \in \mathcal{L}_K(V,V) : \ f - bijekcja\}$ zbiór wszystkich automorfizmów przestrzeni liniowej V.

Wniosek 7.3.3. Zbiór $Aut_K(V)$ wraz z działaniem składania odwzorowań jest grupą nieprzemienną.

Dowód. Wewnętrzność działania składania wynika na mocy twierdzenia 7.3.2 i). Składania odwzorowań jest łączne. Elementem neutralnym jest odwzorowanie identycznościowe id_V. Elementem symetrycznym do $f \in Aut_K(V)$ jest f^{-1}, bowiem na mocy twierdzenia 7.3.2 ii) $f^{-1} \in Aut_K(V)$. $\square$

Grupa $Aut_K(V)$ bywa też oznaczana symbolem $GL(V)$ i nazywana *pełną* lub *ogólną grupą liniową przestrzeni liniowej* V.

Przykład 7.3.4.
Czy odwzorowanie liniowe $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$, $\varphi(x_1,x_2,x_3) = (x_1+x_2+x_3, x_2+x_3, x_1+x_2)$ jest odwracalne?
Wyznaczymy jądro φ.

$$\varphi(x_1,x_2,x_3) = (0,0,0) \Leftrightarrow (x_1+x_2+x_3, x_2+x_3, x_1+x_2) = (0,0,0) \Rightarrow \begin{cases} x_1+x_2+x_3 = 0 \\ x_2+x_3 = 0 \\ x_1+x_2 = 0 \end{cases}$$

Jedynym rozwiązaniem powyższego układu jest $x_1 = x_2 = x_3 = 0$.
Zatem $Ker\varphi = \{(0,0,0)\}$. Na mocy twierdzenia 7.2.4, φ jest monomorfizmem.
Na mocy twierdzenia 7.2.6 o rzędzie

$$\dim Im\varphi = \dim \mathbb{R}^3 - \dim Ker\varphi = 3 - 0 = 3.$$

Ponieważ $Im\varphi \subseteq \mathbb{R}^3$ oraz $\dim Im\varphi = 3$, zatem $Im\varphi = \mathbb{R}^3$, czyli φ jest epimorfizmem.
W konsekwencji φ jest automorfizmem przestrzeni $\mathbb{R}^3$.
W szczególności odwzorowanie odwrotne $\varphi^{-1} : \mathbb{R}^3 \to \mathbb{R}^3$ również jest automorfizmem.

7.4 Reprezentacja macierzowa odwzorowania liniowego

Niech V i W będą przestrzeniami liniowymi skończenie wymiarowymi nad tym samym ciałem K. Niech $\mathcal{B}_V = \big(b_1, \ldots, b_n\big)$ oraz $\mathcal{B}_W = \big(c_1, \ldots, c_m\big)$ będą ustalonymi bazami przestrzeni V i W odpowiednio. Rozważmy odwzorowanie liniowe $\varphi : V \to W$.

Definicja 7.4.1. *Macierzą* (lub *reprezentacją macierzową*) *odwzorowania liniowego* φ w bazach $\mathcal{B}_V, \mathcal{B}_W$ nazywamy macierz $A \in M_{m\times n}(K)$, której kolejne kolumny to współrzędne wektorów $\varphi(b_1), \ldots, \varphi(b_n)$ w bazie $\mathcal{B}_W$. Oznaczamy ją symbolem $M_\varphi(\mathcal{B}_V, \mathcal{B}_W)$.

Przykład 7.4.2. Wyznacz macierz odwzorowania liniowego $\varphi : \mathbb{R}^3 \to \mathbb{R}^2, \quad \varphi(x, y, z) = (3x, 2y+z)$, gdy rozważamy

a) w $\mathbb{R}^3$ i $\mathbb{R}^2$ bazy kanoniczne

$$\varphi(1,0,0) = (3,0),\ \varphi(0,1,0) = (0,2),\ \varphi(0,0,1) = (0,1), \qquad M_\varphi(\mathcal{B}^3_k, \mathcal{B}^2_k) = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x \\ 2y+z \end{bmatrix}$$

b) bazy $\mathcal{B} = \big(b_1 = (1,2,0), b_2 = (1,1,1), b_3 = (0,0,1)\big)$, $\mathcal{C} = \big(c_1 = (1,2), c_2 = (0,1)\big)$

$$\varphi(b_1) = \varphi(1,2,0) = (3,4) = [\alpha_1, \beta_1]_\mathcal{C}, \qquad (3,4) = \alpha_1 c_1 + \beta_1 c_2 = (\alpha_1, 2\alpha_1 + \beta_1),$$
$$\Rightarrow\ \alpha_1 = 3, \beta_1 = -2,\ \varphi(b_1) = [3,-2]_\mathcal{C}$$
$$\varphi(b_2) = \varphi(1,1,1) = (3,3) = [\alpha_2, \beta_2]_\mathcal{C}, \qquad (3,3) = \alpha_2 c_1 + \beta_2 c_2 = (\alpha_2, 2\alpha_2 + \beta_2)$$
$$\Rightarrow\ \alpha_2 = 3, \beta_2 = -3,\ \varphi(b_2) = [3,-3]_\mathcal{C}$$
$$\varphi(b_3) = \varphi(0,0,1) = (0,1) = [\alpha_3, \beta_3]_\mathcal{C}, \qquad (0,1) = \alpha_3 c_1 + \beta_3 c_2 = (\alpha_3, 2\alpha_3 + \beta_3)$$
$$\Rightarrow\ \alpha_3 = 0, \beta_3 = 1,\ \varphi(b_3) = [0,1]_\mathcal{C}$$

$$M_\varphi(\mathcal{B}, \mathcal{C}) = \begin{bmatrix} 3 & 3 & 0 \\ -2 & -3 & 1 \end{bmatrix}$$

Obserwacja: Postać macierzy reprezentującej dane odwzorowanie liniowe zależy od wyboru baz.

Uwaga 7.4.3. Odwzorowanie $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ jest liniowe wtedy i tylko wtedy, gdy istnieje macierz $A \in M_{m\times n}(\mathbb{R})$ taka, że

$$\forall (x_1, \ldots, x_n) \in \mathbb{R}^n \quad \varphi(x_1, \ldots, x_n) = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Wówczas $A = M_\varphi(\mathcal{B}_k^n, \mathcal{B}_k^m)$.

Przykład 7.4.4. Niech $A \in M_{4\times 7}(\mathbb{R})$ reprezentuje odwzorowanie liniowe f w bazach kanonicznych. Wówczas $f : \mathbb{R}^7 \to \mathbb{R}^4$.
Ponieważ $dim Imf \subseteq \mathbb{R}^4$, zatem $\dim Imf \leq 4$.
Na mocy twierdzenia 7.2.6 o rzędzie $\dim Kerf >= \dim \mathbb{R}^7 - \dim Imf \geq 7 - 4 = 3$.

Uwaga 7.4.5. Rozważmy dwie przestrzenie liniowe V oraz W nad $\mathbb{R}$ takie, że $\dim V = n$ oraz $\dim W = m$. Niech $\mathcal{B}$ będzie bazą przestrzeni V, zaś, $\mathcal{C}$ bazą przestrzeni W.

$$\begin{array}{ccc} & \text{odwzorowanie liniowe} & \\ & \boldsymbol{\varphi} & \\ \boldsymbol{V} & \longrightarrow & \boldsymbol{W} \\ \boldsymbol{v} & \longmapsto & \boldsymbol{w = \varphi(v)} \end{array}$$

$$\begin{array}{ccc} & \boldsymbol{A = M_\varphi(\mathcal{B}_k^n, \mathcal{B}_k^m)} & \\ \mathbb{R}^{\boldsymbol{n}} & \longrightarrow & \mathbb{R}^{\boldsymbol{m}} \\ \boldsymbol{[v_1, v_2, \ldots, v_n]_B} & \longmapsto & \boldsymbol{[w_1, w_2, \ldots, w_m]_C} \end{array}$$

Przykład 7.4.6. Wyznacz macierz odwzorowania liniowego

$$f : \mathbb{C}_1[z] \to M_2(\mathbb{C}), \quad f(\alpha z + \beta) = \alpha A + \beta I_2, \quad \text{gdzie} \quad A = \begin{bmatrix} 2+i & 1 \\ 3 & 4-i \end{bmatrix},$$

względem baz standardowych danych przestrzeni $(\mathbb{C}_1[z], +, \mathbb{C}, \cdot)$ oraz $(M_2(\mathbb{C}), +, \mathbb{C}, \cdot)$.

Weźmy dowolny $p \in \mathbb{C}_1[z]$. Jest on postaci $p(z) = \alpha z + \beta$.
Rozważamy bazę $\mathcal{B} = \big(1, z\big)$ przestrzeni $\mathbb{C}_1[z]$ oraz bazę $\mathcal{C} = \big(E_{11}, E_{12}, E_{21}, E_{22}\big)$ przestrzeni $M_2(\mathbb{C})$.

$$f(1) = 0 \cdot A + 1 \cdot I_2 = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [1,0,0,1]_{\mathcal{C}}$$
$$f(z) = 1 \cdot A + 0 \cdot I_2 = A = [2+i, 1, 3, 4-i]_{\mathcal{C}}$$

$$\dim_{\mathbb{C}} \mathbb{C}_1[z] = 2,\ \dim_{\mathbb{C}} M_2(\mathbb{C}) = 4,\ \Rightarrow\ M_f(\mathcal{B},\mathcal{C}) \in M_{4\times 2}(\mathbb{C}),\ M_f(\mathcal{B},\mathcal{C}) = \begin{bmatrix} 1 & 2+i \\ 0 & 1 \\ 0 & 3 \\ 1 & 4-i \end{bmatrix}$$

Twierdzenie 7.4.7. Niech V oraz W będą przestrzeniami liniowymi nad ciałem K. Niech $\mathcal{B}_V = \left(b_1, \ldots, b_n\right)$ oraz $\mathcal{B}_W = \left(c_1, \ldots, c_m\right)$ będą ustalonymi bazami przestrzeni V i W. Niech $\varphi \in \mathcal{L}(V,W)$ oraz $A = M_\varphi(\mathcal{B}_V, \mathcal{B}_W)$. Oznaczmy

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix},$$

gdzie $v = [x_1, \ldots, x_n]_{\mathcal{B}_V}, w = [y_1, \ldots, y_m]_{\mathcal{B}_W}$. Wówczas

$$\varphi(v) = w \quad \Leftrightarrow \quad AX = Y.$$

Uwaga 7.4.8. Istnieje wzajemnie jednoznaczna odpowiedniość [a] między odwzorowaniami liniowymi a macierzami (przy ustalonych bazach). Macierz odwzorowania liniowego φ w pełni opisuje to odwzorowanie, można zatem badać macierz, zamiast odwzorowania.

[a] czyli bijekcja

Wniosek 7.4.9. Rząd macierzy A przekształcenia liniowego φ nie zależy od wyboru baz $\mathcal{B}_V, \mathcal{B}_W$. Ponadto $\operatorname{rank}(\varphi) = \operatorname{rank} A$. [a]

[a] Innymi słowy rząd odwzorowania liniowego równy jest rzędowi reprezentującej go macierzy (w dowolnych bazach).

Wniosek 7.4.10. Przyjmijmy oznaczenia jak w twierdzeniu 7.4.7. Wówczas

i) φ jest epimorfizmem $\Leftrightarrow$ $r(A) = m$,

ii) φ jest monomorfizmem $\Leftrightarrow$ $r(A) = n$.

Przykład 7.4.11. Niech $D = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} \in M_2(\mathbb{R})$. Czy odwzorowanie liniowe $f : \mathbb{R}_2[x] \to M_2(\mathbb{R})$ dane wzorem $f(ax^2 + bx + c) = aD^2 + bD + cI$ jest epimorfizmem?

Zapiszemy macierz $A = M_f(\mathcal{B}, \mathcal{C})$ tego odwzorowania w bazach $\mathcal{B} = (b_1 = 1, b_2 = x, b_3 = x^2)$, $\mathcal{C} = (E_{11}, E_{12}, E_{21}, E_{22})$. Obliczamy

$$f(b_1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad f(b_2) = D = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}, \quad f(b_3) = D^2 = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}.$$

Zatem

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 3 & 9 \end{bmatrix} \quad \text{oraz} \quad r(A) = 2 \neq \dim M_2(\mathbb{R}) = 4.$$

Zatem $Imf \neq M_2(\mathbb{R})$, czyli f nie jest epimorfizmem.

Przykład 7.4.2 - ciąg dalszy
Oblicz $\varphi(1, 2, 3)$ dwoma sposobami, za pomocą macierzy $M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^2)$ oraz za pomocą $M_\varphi(\mathcal{B}, \mathcal{C})$.
Czy φ jest monomorfizmem /epimorfizmem?

Oznaczmy $A = M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^2) = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix}$, $A' = M_\varphi(\mathcal{B}, \mathcal{C}) = \begin{bmatrix} 3 & 3 & 0 \\ -2 & -3 & 1 \end{bmatrix}$.

$$A \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

Sprawdzenie $\varphi(1, 2, 3) = (3 \cdot 1, 2 \cdot 2 + 2) = (3, 7)$
$(1, 2, 3) = 1(1, 2, 0) + 3(0, 0, 1) = 1 \cdot b_1 + 3 \cdot b_3 = [1, 0, 3]_{\mathcal{B}}$

$$A' \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 0 \\ -2 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Stąd $\varphi(1, 2, 3) = [3, 1]_{\mathcal{C}} = 3c_1 + c_2 = 3(1, 2) + (0, 1) = (3, 7)$.

Dodatkowo zauważmy, że $r(A) = r(A') = 2 = \dim \mathbb{R}^2$, zatem φ jest epimorfizmem.
Ponadto $\dim \mathrm{Ker}\varphi = 3 - 2 = 1$, więc φ nie jest monomorfizmem.

Niech V oraz W będą przestrzeniami liniowymi nad ciałem K. Niech $\mathcal{B}_V$ oraz $\mathcal{B}_W$ będą ustalonymi bazami. Ponadto niech $\alpha \in K$, $f, g \in \mathcal{L}(V, W)$ oraz $A = M_f(\mathcal{B}_V, \mathcal{B}_W)$, $B = M_g(\mathcal{B}_V, \mathcal{B}_W)$.

Twierdzenie 7.4.12. Przy powyższych założeniach

$$A + B = M_{f+g}(\mathcal{B}_V, \mathcal{B}_W) \quad \text{oraz} \quad \alpha A = M_{\alpha f}(\mathcal{B}_V, \mathcal{B}_W).$$

Twierdzenie 7.4.13. Jeśli $\boxed{\dim V = \dim W,}$ to wówczas następujące warunki są równoważne.

i) f jest izomorfizmem.

ii) $\mathrm{Ker} f = \{\mathbf{0}_V\}$

iii) $\mathrm{Im} f = W$

iv) $r(A) = \dim V$

v) $\det A \neq 0$

Wniosek 7.4.14. Niech $f \in \mathcal{L}(V, W)$ będzie izomorfizmem oraz niech $A = M_f(\mathcal{B}_V, \mathcal{B}_W)$. Wówczas $A^{-1} = M_{f^{-1}}(\mathcal{B}_W, \mathcal{B}_V)$.

Przykład 7.4.15. Czy odwzorowanie liniowe $f : \mathbb{R}_2[x] \to \mathbb{R}^3$ dane wzorem $f(p)(x) = \Big(p(-1), p(0), p(1)\Big)$ jest izomorfizmem?
Zapiszemy macierz $A = M_f(\mathcal{B}, \mathcal{B}_k^3)$, gdzie $\mathcal{B} = \Big(1, x, x^2\Big)$.

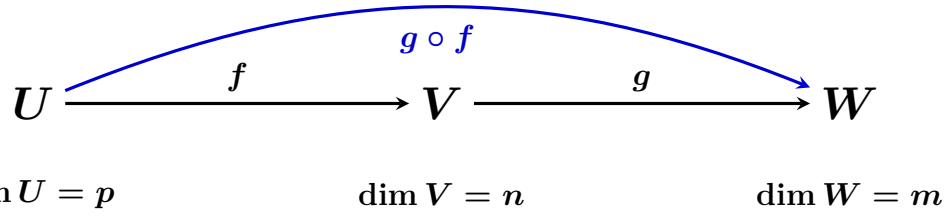
$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

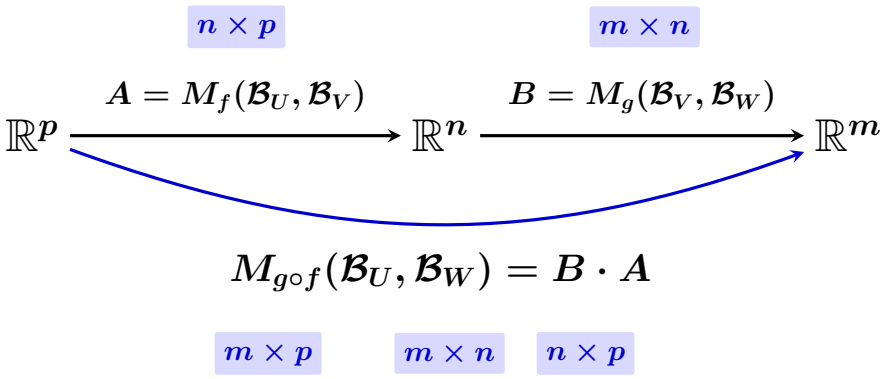
Obliczamy $\det A = 2 \neq 0$. Na mocy twierdzenia 7.4.13, f jest izomorfizmem.

Twierdzenie 7.4.16. Niech U, V, W będą przestrzeniami liniowymi skończenie wymiarowymi nad ciałem K. Niech $\mathcal{B}_U, \mathcal{B}_V, \mathcal{B}_W$ będą ustalonymi bazami. Ponadto niech $f \in \mathcal{L}(U, V)$, $g \in \mathcal{L}(V, W)$ oraz $A = M_f(\mathcal{B}_U, \mathcal{B}_V)$, $B = M_g(\mathcal{B}_V, \mathcal{B}_W)$. Wówczas [a]

$$B \cdot A = M_g(\mathcal{B}_V, \mathcal{B}_W) \cdot M_f(\mathcal{B}_U, \mathcal{B}_V) = M_{g \circ f}(\mathcal{B}_U, \mathcal{B}_W).$$

[a] Pamiętajmy, że działanie mnożenia macierzy nie jest przemienne. W podanym wzorze istotna jest kolejność czynników.





macierz złożenia odwzorowań liniowych

Przykład 7.4.17. Dane są odwzorowania liniowe
$f : \mathbb{R}^3 \to \mathbb{R}^2$, $f(x, y, z) = (x - y + z, 2y + z)$,
$g : \mathbb{R}^3 \to \mathbb{R}^2$, $g(x, y, z) = (x - 3z, x + y)$,
$h : \mathbb{R}^2 \to \mathbb{R}^2$, $h(x, y) = (2x + y, x - y)$.
Za pomocą rachunku macierzowego uzasadnimy, że h jest izomorfizmem, a następnie wyznaczymy wzory odwzorowań h^{-1} oraz $\varphi = 2h^{-1} \circ h^{-1} \circ (f + g)$ i obliczymy $\varphi(1, 2, 3)$.

Czy h jest odwracalne?

$h : \mathbb{R}^2 \to \mathbb{R}^2 \Rightarrow M_f := M_f(\mathcal{B}_k^2, \mathcal{B}_k^2) \in M_2(\mathbb{R}),\ M_h = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$

$\det M_h = -3 \neq 0 \Rightarrow h$ jest odwracalne

$M_{h^{-1}} := M_{h^{-1}}(\mathcal{B}_k^2, \mathcal{B}_k^2) \in M_2(\mathbb{R}), \quad M_{h^{-1}} = \left(M_h\right)^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$

$$f:\mathbb{R}^3\to\mathbb{R}^2 \;\Rightarrow\; M_f:=M_f(\mathcal{B}_k^3,\mathcal{B}_k^2)\in M_{2\times 3}(\mathbb{R}),\; M_f=\begin{bmatrix}1&-1&1\\0&2&1\end{bmatrix}$$
$$g:\mathbb{R}^3\to\mathbb{R}^2 \;\Rightarrow\; M_g:=M_g(\mathcal{B}_k^3,\mathcal{B}_k^2)\in M_{2\times 3}(\mathbb{R}),\; M_g=\begin{bmatrix}1&0&-3\\1&1&0\end{bmatrix}$$
$$M_{f+g}:=M_{f+g}(\mathcal{B}_k^3,\mathcal{B}_k^2)\in M_{2\times 3}(\mathbb{R}),\quad M_{f+g}=M_f+M_g=\begin{bmatrix}2&-1&-2\\1&3&1\end{bmatrix}$$

$$\varphi:\mathbb{R}^3\to\mathbb{R}^2 \;\Rightarrow\; M_\varphi:=M_\varphi(\mathcal{B}_k^3,\mathcal{B}_k^2)\in M_{2\times 3}(\mathbb{R}),$$
$$M_\varphi=2M_{h^{-1}}^2M_{f+g}=\tfrac{2}{9}\begin{bmatrix}1&1\\1&-2\end{bmatrix}\begin{bmatrix}1&1\\1&-2\end{bmatrix}\begin{bmatrix}2&-1&-2\\1&3&1\end{bmatrix}=\tfrac{2}{9}\begin{bmatrix}3&-5&-5\\3&16&7\end{bmatrix}$$

$\varphi(x,y,z)=?$
$$M_\varphi\begin{bmatrix}x\\y\\z\end{bmatrix}=\tfrac{2}{9}\begin{bmatrix}3x-5y-5z\\3x+16y-7z\end{bmatrix}\;\Rightarrow\;\varphi(x,y,z)=(\tfrac{2}{3}x-\tfrac{10}{9}y-\tfrac{10}{9}z,\tfrac{2}{3}x+\tfrac{32}{9}y-\tfrac{14}{9}z)$$
$\varphi(1,2,3)=?$
$$M_\varphi\begin{bmatrix}1\\2\\3\end{bmatrix}=\tfrac{2}{9}\begin{bmatrix}3&-5&-5\\3&16&7\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix}=\tfrac{2}{9}\begin{bmatrix}-22\\56\end{bmatrix}\;\Rightarrow\;\varphi(1,2,3)=(-\tfrac{44}{9},\tfrac{112}{9})$$

Uwaga 7.4.18. Niech V będzie przestrzenią liniową skończenie wymiarową, zaś $\mathcal{B}$ oraz $\mathcal{B}'$ jej dwiema bazami. Wówczas macierz przejścia $P_{\mathcal{B}\to\mathcal{B}'}$ jest reprezentacją macierzową odwzorowania identycznościowego $\mathrm{id}:V\to V$ przestrzeni V z bazą $\mathcal{B}'$ w przestrzeń V z bazą $\mathcal{B}$.

Dowód. Niech $\mathcal{B}=(b_1,\ldots,b_n)$, $\mathcal{B}'=(b'_1,\ldots,b'_n)$. Wówczas
$$\begin{cases}\mathrm{id}(b'_1)=b'_1=a_{11}b_1+\ldots+a_{n1}b_1\\ \ldots\\ \mathrm{id}(b'_n)=b'_n=a_{1n}b_1+\ldots+a_{nn}b_n\end{cases}\text{, skąd } P_{\mathcal{B}\to\mathcal{B}'}=M_{\mathrm{id}}(\mathcal{B}',\mathcal{B}).\quad\square$$

odwzorowanie identycznościowe
$$\boldsymbol{(V,B')}\xrightarrow{\quad\mathbf{id}\quad}\boldsymbol{(V,B)}$$
$$\boldsymbol{v}\longmapsto\mathbf{id}(\boldsymbol{v})=\boldsymbol{v}$$
$\boldsymbol{\dim V=n}$
$$\mathbb{R}^{\boldsymbol{n}}\xrightarrow{\quad\boldsymbol{P_{B\to B'}=M_{\mathrm{id}}(B',B)}\quad}\mathbb{R}^{\boldsymbol{n}}$$
$$[\boldsymbol{v'_1,v'_2,\ldots,v'_n}]_{B'}\longmapsto[\boldsymbol{v_1,v_2,\ldots,v_n}]_B$$

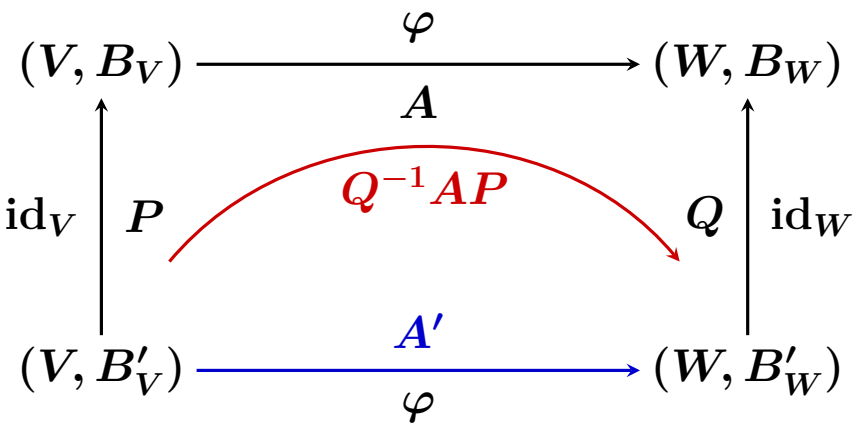
macierz przejścia jako macierz odwzorowania identycznościowego

7.5 Zmiana baz

Twierdzenie 7.5.1. Niech V oraz W będą przestrzeniami liniowymi skończenie wymiarowymi nad ciałem K. Niech $\mathcal{B}_V$ oraz $\mathcal{B}_W$ będą bazami przestrzeni V i W. Rozważmy nowe bazy $\mathcal{B}'_V, \mathcal{B}'_W$ oraz odwzorowanie liniowe $\varphi : V \to W$. Niech $P = P_{\mathcal{B}_V \to \mathcal{B}'_V}$, $Q = P_{\mathcal{B}_W \to \mathcal{B}'_W}$ oraz $A = M_\varphi(\mathcal{B}_V, \mathcal{B}_W)$, $A' = M_\varphi(\mathcal{B}'_V, \mathcal{B}'_W)$. Wówczas

$$\boxed{A' = Q^{-1}AP.}$$

Dowód. Ponieważ $X = PX'$, $Y = QY'$, $AX = Y$, $A'X' = Y'$, zatem $QY' = Y = AX = APX' \Rightarrow Y' = (Q^{-1}AP)X'$. $\square$



zmiana macierzy odwzorowania liniowego przy zmianie baz

Przykład 7.4.2 raz jeszcze
Korzystając ze wzoru ma zmianę macierzy odwzorowania liniowego przy zmianie baz, wyznaczymy macierz $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$, $\varphi(x, y, z) = (3x, 2y + z)$ w bazach $\mathcal{B} = \Big(b_1 = (1, 2, 0), b_2 = (1, 1, 1), b_3 = (0, 0, 1)\Big)$, $\mathcal{C} = \Big(c_1 = (1, 2), c_2 = (0, 1)\Big)$.

$$A = M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^2) = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix}, \quad A' = M_\varphi(\mathcal{B}, \mathcal{C}) = ?$$

$$P = P_{\mathcal{B}_k^3 \to \mathcal{B}} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, Q = P_{\mathcal{B}_k^2 \to \mathcal{C}} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad Q^{-1} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$$

$$A' = Q^{-1}AP = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 0 \\ -2 & -3 & 1 \end{bmatrix}$$

Uwaga 7.5.2. Zmiana bazy przestrzeni wektorowej prowadzi do *liniowej zmiany zmiennych* lub *podstawienia liniowego*.

Równość $X = PX'$ oznacza, że każda ze zmiennych $x_1, \ldots, x_n$ wyraża się jako kombinacja liniowa $x'_1, \ldots, x'_n$.

Ponieważ $\det P \neq 0$, możemy wywnioskować, że $X' = P^{-1}X$ i wyrazić każdą ze zmiennych $x'_1, \ldots, x'_n$ jako kombinację liniową $x_1, \ldots, x_n$.

Wektory pozostają te same, ale ich opis za pomocą współrzędnych zmienia się w zależności od przyjętego układu współrzędnych.

W powyższym przykładzie zmiana baz w przestrzeniach $\mathbb{R}^3$ i $\mathbb{R}^2$ geometrycznie oznacza zmianę układów współrzędnych (a zatem zmianę zmiennych). Powoduje to zmianę przepisu odwzorowania φ, ale w żaden sposób nie wpływa na to jaki wektor $\varphi(v) \in \mathbb{R}^2$ jest przypisywany danemu wektorowi $v \in \mathbb{R}^3$.

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \; X' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}, \; PX' = X,$$

Korzystamy z równości $X' = P^{-1}X$. Obliczamy $P^{-1} = \begin{bmatrix} -1 & 1 & 0 \\ 2 & -1 & 0 \\ -2 & 1 & 1 \end{bmatrix}$ oraz

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 2 & -1 & 0 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} -x + y \\ 2x - y \\ -2x + y + z \end{bmatrix}.$$

Ponadto $Y = \begin{bmatrix} a \\ b \end{bmatrix}, \; Y' = \begin{bmatrix} a' \\ b' \end{bmatrix}, \; Y = QY' \Leftrightarrow \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} a' \\ 2a' + b' \end{bmatrix}$.

Korzystając z macierzy A oraz A' otrzymujemy

$$(a, b) = \varphi(x, y, z) = (3x, 2y + z) \quad \text{oraz} \quad (a', b') = \varphi(x', y', z') = (3x' + 3y', -2x' - 3y' + z').$$

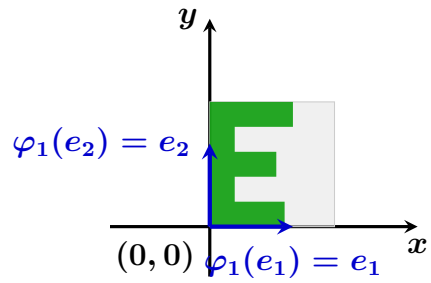
Możemy sprawdzić, że

$$(a', b') = \varphi(x', y', z') = (3x' + 3y', -2x' - 3y' + z') =$$

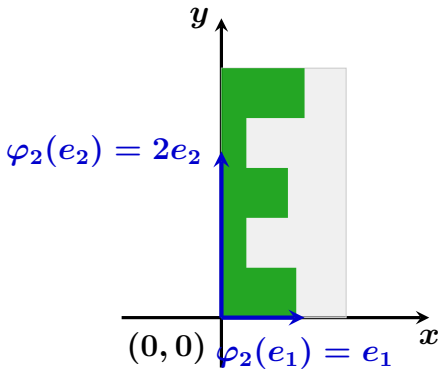
$$= \Big(3(-x + y) + 3(2x - y), -2(-x + y) - 3(2x - y) + (-2x + y + z)\Big) = (3x, -6x + 2y + z).$$

Stąd $a = a' = 3x$ oraz $b = 2a' + b' = 6x - 6x + 2y + z = 2y + z$.

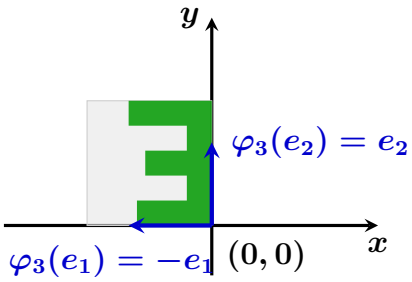
Przykład 7.5.3. Endomorfizmy przestrzeni $\mathbb{R}^2$



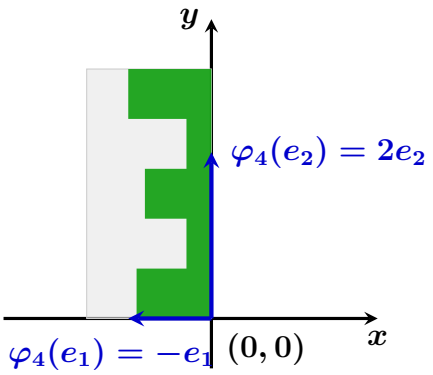
identyczność $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$



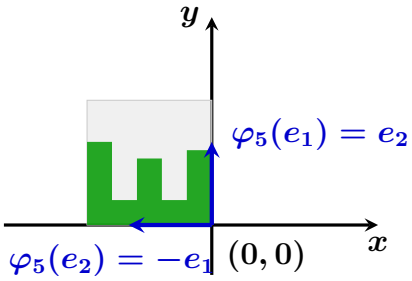
rozciąganie $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$



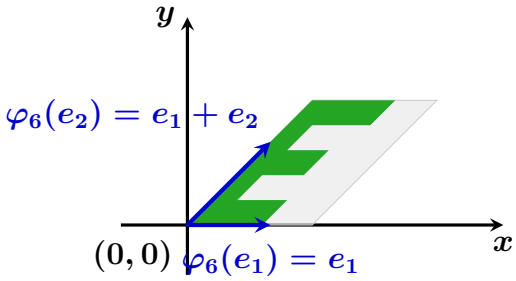
odbicie (symetria osiowa) $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$



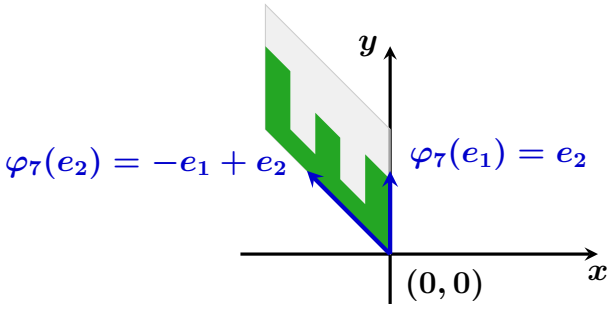
rozciąganie i odbicie $\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$



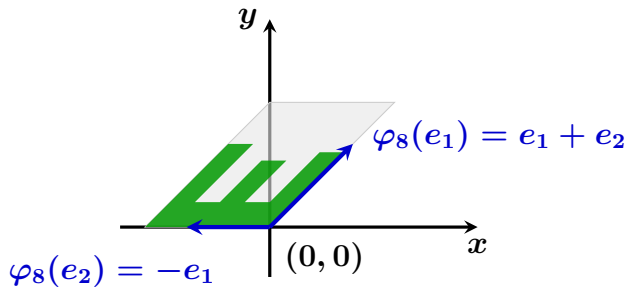
obrót (rotacja) o kąt $\frac{\pi}{2}$ $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$



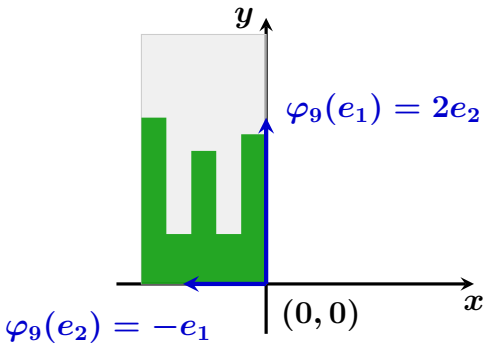
powinowactwo ścinające (ang. shear) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$



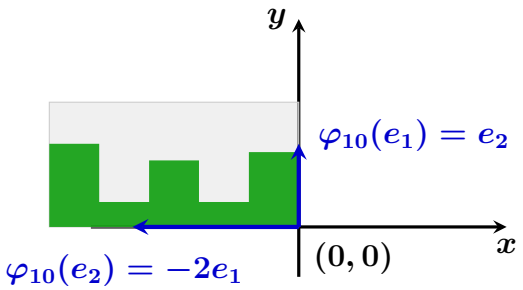
powinowactwo ścinające i rotacja $\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$



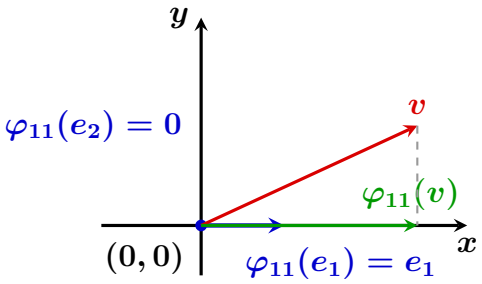
rotacja i powinowactwo ścinające $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$



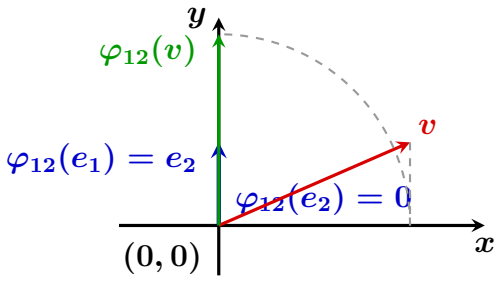
rotacja i rozciąganie $\begin{bmatrix} 0 & -1 \\ 2 & 0 \end{bmatrix}$



rozciąganie i rotacja $\begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix}$



rzutowanie (projekcja) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$



projekcja i rotacja $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

Czym jest powinowactwo ścinające, dowiesz się tutaj.

7.6 Uwaga na temat rozwiązywania układów równan liniowych

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$. Rozważmy układ m równań liniowych z n niewiadomymi $x_1, \ldots, x_n$ o współczynnikach z K.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots & \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{cases}, \quad a_{ij}, b_i \in K,\ i \in \{1, 2, \ldots, m\}, j \in \{1, 2, \ldots, n\}$$

Rozpatrywany układ równań liniowych jest równoważny z równaniem macierzowym $AX = B$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

macierz współczynników — kolumna niewiadomych — kolumna wyrazów wolnych

Przy ustalonych bazach przestrzeni K^n, K^m macierz A jednoznacznie wyznacza odwzorowanie liniowe $\varphi_A : K^n \to K^m$. Rozwiązywanie układu jednorodnego $AX = \mathbf{0}$ polega na wyznaczeniu jądra $\mathrm{Ker}\varphi_A$ odwzorowania φ_A.

Wniosek 7.6.1. Zbiór rozwiązań jednorodnego układu równań liniowych $AX = \mathbf{0}$ ma strukturę przestrzeni liniowej nad ciałem K. Jej wymiar jest równy $n - r(A)$.

W szczególności wektor zerowy zawsze należy do zbioru rozwiązań układu jednorodnego. Ponadto suma rozwiązań układu jednorodnego jest rozwiązaniem tegoż układu jednorodnego.

Twierdzenie 7.6.2. Niech $(x_1^*, x_2^*, \ldots, x_n^*)$ będzie jednym z rozwiązań niesprzecznego niejednorodnego układu równań liniowych $AX = B$. Wówczas zbiór wszystkich rozwiązań układu $AX = B$ ma postać

$$\Big\{(x_1^*, x_2^*, \ldots, x_n^*) + (\gamma_1, \gamma_2, \ldots, \gamma_n) :\ (\gamma_1, \gamma_2, \ldots, \gamma_n) \text{ dowolne rozw. } AX = \mathbf{0}\Big\}.$$

Przykład 7.6.3. Rozważmy raz jeszcze układ równań z przykładu 4.4.2.

$$\begin{cases} 2x + 4y + 9z - 5t + 6u &= 13 \\ x + 2y + 4z - 4t + 2u &= 5 \\ 3x + 6y + 7z - 11t + 2u &= 18 \\ 4x + 8y + 19z - 15t + 11u &= 20 \\ -\frac{1}{2}x - y + z + 3t + 2u &= -\frac{5}{2} \end{cases}$$

Jak już wiemy, $r(A) = r(U) = 3 < n = 5$ i jest to układ nieoznaczony, posiadający nieskończenie wiele rozwiązań zależnych od 2 parametrów.

Rozwiązania są postaci $\begin{cases} x = 11 - 2y \\ z = -3 + \frac{7}{3}t \\ u = 3 - \frac{8}{3}t \\ y, t \in \mathbb{R} \end{cases}$. Równoważnie

$(x, y, z, t, u) = (11 - 2y, y, -3 + \frac{7}{3}t, t, 3 - \frac{8}{3}t) = (11, 0, -3, 0, 3) + (-2, 1, 0, 0, 0)y + (0, 0, \frac{7}{3}, 1, -\frac{8}{3})t.$

Dla $y = 0$ i $t = 0$ otrzymujemy $(11, 0, -3, 0, 3)$ jako jedno z rozwiązań rozważanego układu.

Zbiór rozwiązań korespondującego układu jednorodnego jest przestrzenią liniową wymiaru 2, zaś układ wektorów $\{(-2, 1, 0, 0, 0), (0, 0, \frac{7}{3}, 1, -\frac{8}{3})\}$ jest bazą tej przestrzeni.

Przykład 7.6.4. Czy istnieje macierz $A \in M_2(\mathbb{R})$ taka, że rozwiązania układu $AX = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$ są postaci $y = 5x + 1$, $x \in \mathbb{R}$, zaś rozwiązania układu $AX = \begin{bmatrix} -4 \\ 10 \end{bmatrix}$ są postaci $y = 4x - 1$, $x \in \mathbb{R}$?

Taka macierz nie istnieje. Na mocy twierdzenia 4.4.2 rozwiązanie układu postaci $AX = B$ jest sumą rozwiązania układu jednorodnego $AX = \mathbf{0}$ i pewnego rozwiązania układu niejednorodnego $AX = B$.

Gdyby taka macierz istniała, podane proste miałyby takie same współczynniki kierunkowe (byłyby równoległe).

Q U I Z

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. Odwzorowanie $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (y, |x|, -2z)$ jest liniowe.
2. Odwzorowanie $\varphi : \mathbb{R}[x] \to \mathbb{R}[x]$, $\varphi(p)(x) = p'(x) \cdot p''(x)$ jest liniowe.
3. Jądro dowolnego odwzorowania liniowego $f : \mathbb{R}_3[x] \to M_3(\mathbb{R})$ jest podprzestrzenią liniową przestrzeni $M_3(\mathbb{R})$.
4. Odwzorowanie liniowe $f : \mathbb{R}[x] \to \mathbb{R}[x]$, $f(p)(x) = x \cdot p(x)$ jest monomorfizmem.
5. Niech $(V, +, \mathbb{R}, \cdot)$ będzie przestrzenią liniową wszystkich nieskończonych rzeczywistych ciągów liczbowych $(a_n)_{n\in\mathbb{N}}$ (z naturalnymi działaniami dodawania ciągów i mnożenia ciągu przez liczbę). Odwzorowanie $f : V \to V$ dane wzorem $f(a_1, a_2, a_3, \ldots) = (0, a_1, a_2, a_3, \ldots)$ jest monomorfizmem.
6. Dany jest endomorfizm $\varphi : M_3(\mathbb{R}) \to M_3(\mathbb{R})$, $\varphi(A) = AB - BA$, gdzie $B \in M_n(\mathbb{R})$ to ustalona macierz. Można tak wybrać B, by φ był izomorfizmem.

7. Żadne odwzorowanie liniowe $\mathbb{R}^5 \to \mathbb{R}^3$ nie jest monomorfizmem.

8. Żadne odwzorowanie liniowe $\mathbb{R}_2[x] \to M_2(\mathbb{R})$ nie jest epimorfizmem.

9. Niech U, V będą przestrzeniami liniowymi nad tym samym ciałem. Jeśli $f : U \to V$ jest izomorfizmem, to wówczas $\dim U = \dim V$.

10. Niech $\varphi : \mathbb{R}^7 \to \mathbb{R}^5$ będzie odwzorowaniem liniowym. Możliwe jest, by $\dim \mathrm{Ker}\varphi = 1$.

11. Dana jest macierz $A \in M_{6\times 10}(\mathbb{R})$ będąca reprezentacją macierzową pewnego odwzorowania liniowego φ. Jeśli $\dim Ker\varphi = 4$, to wówczas dla dowolnej macierzy B (odpowiedniego wymiaru), układ równań $AX = B$ posiada co najmniej jedno rozwiązanie.

12. Niech $f : \mathbb{R}^3 \to \mathbb{R}^3$ będzie dowolnym odwzorowaniem liniowym. Jeśli wektory $u, v, w \in \mathbb{R}^3$ tworzą układ liniowo zależny, to wówczas wektory $f(u), f(v), f(w)$ również tworzą układ liniowo zależny.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Algebra
Wykład nr 8
Data: .....

TEMAT: *Zagadnienie własne operatora liniowego*

8.1 Wartości własne i wektory własne endomorfizmu

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$.

Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową wymiaru n. Oznaczmy $End(V) = \mathcal{L}(V, V)$. Endomorfizmy przestrzeni V nazywamy również *operatorami liniowymi*.

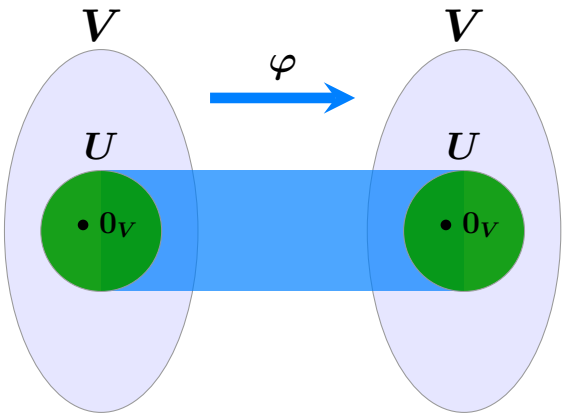
Twierdzenie 8.1.1. i) Zbiór $End(V) = (End(V), +, K\cdot)$ wraz z działaniami dodawania odwzorowań i mnożenia odwzorowania przez liczbę jest przestrzenią wektorową wymiaru n^2.

ii) Zbiór $End(V) = (End(V), +, \circ)$ wraz z działaniami dodawania i składania odwzorowań ma strukturę pierścienia nieprzemiennego.

iii) $\forall \alpha \in K \quad \forall f, g \in End(V) \quad \alpha \cdot (f \circ g) = (\alpha \cdot f) \circ g = f \circ (\alpha \cdot g)$

Definicja 8.1.2. Podprzestrzeń liniową U przestrzeni V nazywamy *niezmienniczą względem endomorfizmu* $\varphi \in End(V)$ lub krótko *φ-niezmienniczą*, jeżeli

$$\varphi(U) \subset U, \quad tzn. \quad \forall u \in U \, \varphi(u) \in U.$$



podprzestrzeń φ-niezmiennicza U

Przykład 8.1.3. 1) Niech $V = \mathbb{R}^3$, $U = \{(0,0,t) \in \mathbb{R}^3 : t \in \mathbb{R}\}$, $\varphi(x,y,z) = (-y,x,z)$

Zauważmy, że φ jest obrotem o kąt $\frac{\pi}{2}$ wokół osi Oz.
Zatem dla $(0,0,t) \in U$ mamy $\varphi(0,0,t) = (0,0,t) \in U$ i U jest φ-niezmiennicza.

2) Niech V będzie dowolna przestrzenią liniową, zaś $\varphi \in End(V)$ dowolnym endomorfizmem.
Podprzestrzeniami φ-niezmienniczymi są $\{\mathbf{0}_V\}$, V, Kerφ, Imφ.

$\{\mathbf{0}_V\}$ jest φ-niezmiennicza, bowiem $\varphi(\mathbf{0}_V) = \mathbf{0}_V \in \{\mathbf{0}_V\}$.

Niech $u \in$ Kerφ, wówczas $\varphi(u) = \mathbf{0}_V$. Oczywiście $\mathbf{0}_V \in$ Kerφ, bowiem $\varphi(\mathbf{0}_V) = \mathbf{0}_V$.
Zatem Kerφ jest φ-niezmiennicza.

Niech $u \in$ Imφ, wówczas $\varphi(u) \in$ Imφ. Zatem Imφ jest φ-niezmiennicza.

3) Rozważmy $f : \mathbb{R}[x] \to \mathbb{R}[x]$ takie, że $f = \frac{d}{dx}$. Innymi słowy

$$\forall p \in \mathbb{R}[x] \quad f(p)(x) = \frac{dp}{dx} = p'(x).$$

Podprzestrzeń $\mathbb{R}_n[x]$ dla dowolnego $n \in \mathbb{N}$ jest podprzestrzenią f-niezmienniczą przestrzeni $\mathbb{R}[x]$, bowiem

$$\forall p \in \mathbb{R}_n[x] \quad f(p) = p' \in \mathbb{R}_{n-1}[x] \subset \mathbb{R}_n[x].$$

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$. Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową oraz $\varphi \in End(V)$.

Definicja 8.1.4.
i) Liczbę $\lambda \in K$ nazywamy *wartością własną* endomorfizmu φ, jeżeli istnieje niezerowy wektor $v \in V$ taki, że $\varphi(v) = \lambda v$. Każdy taki wektor nazywamy *wektorem własnym* endomorfizmu φ, odpowiadającym wartości własnej λ.

ii) Problem polegający na wyznaczeniu dla danego endomorfizmu φ wartości własnych i odpowiadających im wektorów własnych nazywamy *zagadnieniem własnym* dla endomorfizmu φ.

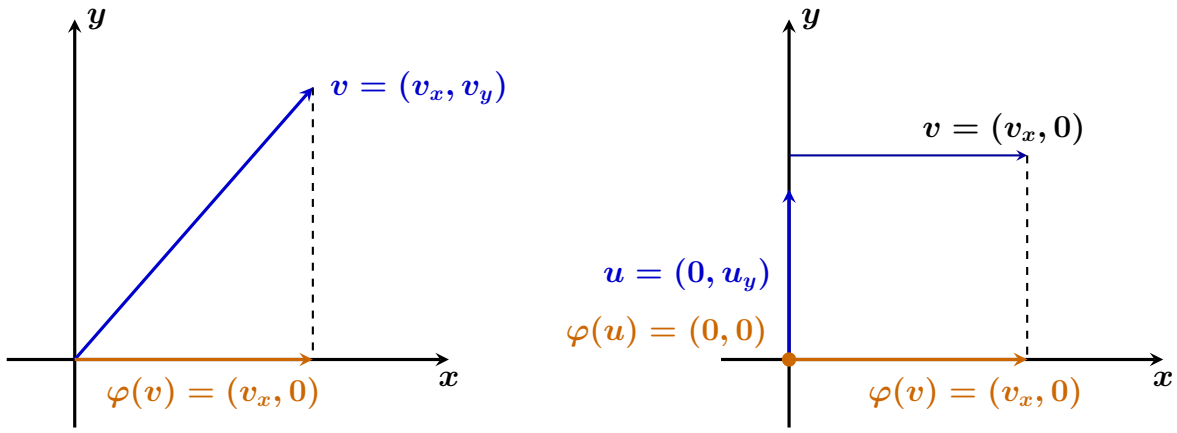
iii) Zbiór wszystkich wartości własnych operatora liniowego φ oznaczamy symbolem $Spec(\varphi)$ i nazywamy *widmem* bądź *spektrum* tego operatora.

Przykład 8.1.5. Niech $\varphi \in End(\mathbb{R}^2)$, będzie rzutem prostokątnym na oś Ox. Rozważ zagadnienie własne dla operatora φ.

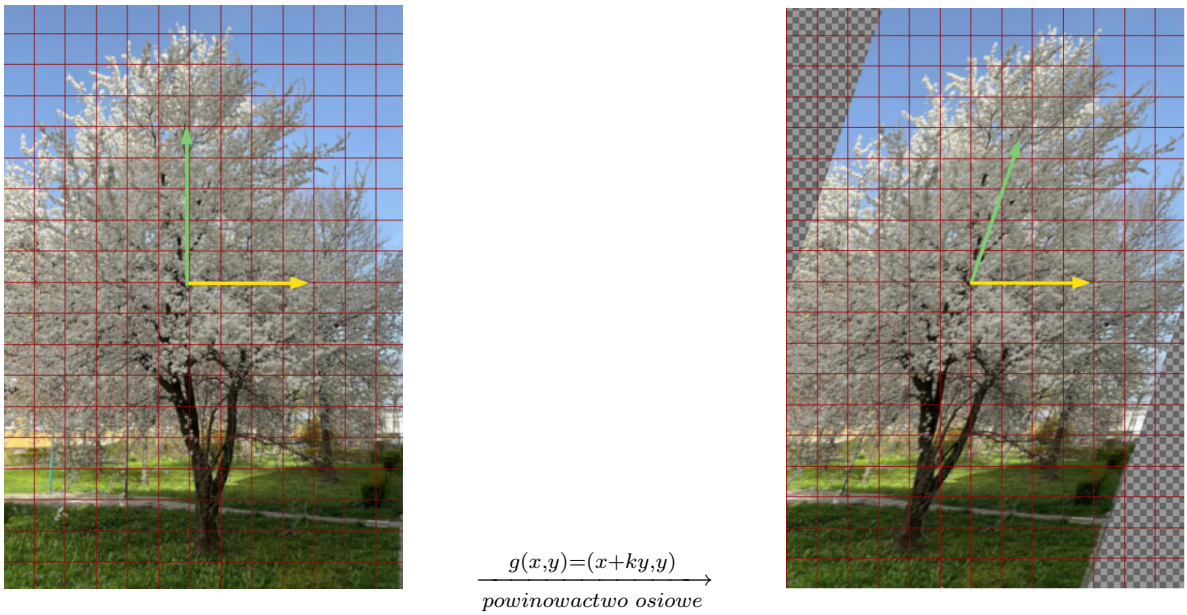
Zauważmy, że $\varphi(v) = \lambda v$, gdy $\varphi(v)$ ma ten sam kierunek co v, tj. $\varphi(v) \in \text{lin}(v)$.
Zatem możliwe są dwie sytuacje.

1) $\lambda_1 = 1$, $\varphi(v) = v$, gdy $v \parallel Ox$, czyli $v = (v_x, 0)$

2) $\lambda_2 = 0$, $\varphi(v) = \mathbf{0}$, gdy $v \perp Ox$, czyli $v = (0, v_y)$



Przykład 8.1.6. Niech $g \in End(\mathbb{R}^2)$ to powinowactwo osiowe dane wzorem $g(x, y) = (x + ky, y)$, gdzie $k \in \mathbb{R}$ to ustalona stała.



W wyniku działania przekształcenia g zielony wektor zmienia kierunek, zaś żłóty wektor nie zmienia kierunku. Żółty wektor jest zatem wektorem własnym $g \in End(\mathbb{R}^2)$. Ponieważ długość żółtego wektora nie ulega zmianie, odpowiada on wartości własnej $\lambda = 1$.

Zobacz także tutaj.

Przykład 8.1.7.
1) Niech φ będzie endomorfizmem płaszczyzny danym wzorem $\varphi(v) = 3v$ dla dowolnego wektora v. Dowolny niezerowy wektor jest wektorem własnym φ odpowiadającym wartości własnej $\lambda = 3$. $Spec(\varphi) = \{3\}$.

2) Niech φ będzie rotacją na płaszczyźnie (w kierunku przeciwnym do ruchu wskazówek zegara) o kąt $\frac{\pi}{2}$. Zatem $\varphi(x, y) = (-y, x)$.
Nie istnieją niezerowe wektory, które w wyniku rotacji zachowują swój kierunek.
Brak zatem wektorów własnych dla φ. A zatem $Spec(\varphi) = \varnothing$.

3) Niech $V = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ oraz $\varphi = \frac{d}{dx}$, tzn. dla dowolnego $f \in V$ mamy $\varphi(f) = \frac{df}{dx} = f'$.
Przekonamy się, że dowolna liczba rzeczywista jest wartością własną operatora $\frac{d}{dx}$.

Niech $\lambda \in \mathbb{R}$ będzie dowolne. Zdefiniujmy

$$g_\lambda : \mathbb{R} \to \mathbb{R}, \quad g_\lambda(x) = a \cdot e^{\lambda x}, \quad \text{gdzie} \quad a \in \mathbb{R} \neq \{0\}.$$

Oczywiście $g_\lambda \in \mathbb{C}^\infty(\mathbb{R}, \mathbb{R})$. Ponadto

$$\varphi(g_\lambda)(x) = a \cdot (e^{\lambda x})' = a\lambda e^{\lambda x} = \lambda(a \cdot e^{\lambda x}) = \lambda g_\lambda(x).$$

Zatem g_λ jest wektorem własnym odpowiadającym wartości własnej λ. Stąd $Spec(\varphi) = \mathbb{R}$.

Wektory własne znajdują zastosowania na przykład w systemach rozpoznawania twarzy. Więcej informacji można znaleźć tutaj lub tutaj.

Uwaga 8.1.8. Każdy wektor własny odpowiada dokładnie jednej wartości własnej.

Uwaga 8.1.9. Cel: diagonalizacja macierzy endomorfizmu
Przy spełnieniu odpowiednich warunków, wyznaczymy taką bazę przestrzeni liniowej V, że macierz endomorfizmu $\varphi \in End(V)$ w tejże bazie będzie diagonalna.
Taka baza nie zawsze istnieje. Będzie to baza złożona z wektorów własnych φ.
Na macierzach diagonalnych łatwo wykonywać działania mnożenia i potęgowania, co odpowiada składaniu i iterowaniu endomorfizmów.

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$. Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową oraz niech $\varphi \in End(V)$, $\lambda \in Spec(\varphi)$. Oznaczmy przez E_λ zbiór wszystkich wektorów własnych odpowiadających wartości własnej λ, uzupełniony o wektor zerowy. Zatem

$$E_\lambda = \{v \in V : \varphi(v) = \lambda v\}.$$

Twierdzenie 8.1.10. i) Zbiór E_λ jest podprzestrzenią liniową przestrzeni V.

ii) Zbiór E_λ jest podprzestrzenią φ-niezmienniczą.

iii) $E_\lambda = \mathrm{Ker}\psi$, gdzie $\psi = \varphi - \lambda \cdot \mathrm{id}_V$.

Definicja 8.1.11. Przestrzeń wektorową E_λ nazywamy *podprzestrzenią własną* endomorfizmu φ, odpowiadającą wartości własnej λ.

Uwaga 8.1.12. Na mocy uwagi 8.1.8 otrzymujemy $\lambda_1 \neq \lambda_2 \Rightarrow E_{\lambda_1} \cap E_{\lambda_2} = \{\mathbf{0}_V\}$.

Przykład 8.1.13. Rozważmy ponownie $\varphi \in End(\mathbb{R}^2)$, $\varphi(v) = 3v$ dla dowolnego wektora $v \in \mathbb{R}^2$. Czyli $\varphi(x, y) = 3(x, y) = (3x, 3y)$ dla dowolnego $(x, y) \in \mathbb{R}^2$.

$Spec(\varphi) = \{3\}$

$\mathbb{R}^2 \setminus \{(0,0)\} = \{(x, y) \in \mathbb{R}^2 : (x, y) \neq (0, 0)\}$ to zbiór wszystkich wektorów własnych

$E_3 = \mathbb{R}^2$

Przykład 8.1.14. Dany jest endomorfizm $f : \mathbb{R}^3 \to \mathbb{R}^3$ oraz wektory $u, v, w \in \mathbb{R}^3$ takie, że

$$f(u) = u, \quad f(v) = 2v, \quad f(w) = 2w.$$

Wówczas $v + w, v - w, 5v + 7w$ to również wektory własne endomorfizmu f odpowiadające wartości własnej $\lambda = 2$, a zatem należące do przestrzeni własnej E_2.

Wynika to z faktu, że E_2 ma strukturę przestrzeni liniowej.
Można to również sprawdzić bezpośrednim rachunkiem. Przykładowo

$$f(5v + 7w) = 5f(v) + 7f(w) = 5 \cdot 2v + 7 \cdot 2w = 2(5v + 7w)$$

Wektor $u + v$ nie jest wektorem własnym endomorfizmu f, bowiem $E_1 \cap E_2 = \{(0, 0, 0)\}$.
Można to również sprawdzić bezpośrednim rachunkiem.

$$f(u + v) = f(u) + f(v) = u + 2v \neq \lambda(u + v).$$

Twierdzenie 8.1.15. Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$. Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową oraz niech $\varphi \in End(V)$, $\lambda \in K$. Niech A będzie macierzą odwzorowania φ w dowolnej ustalonej bazie przestrzeni V. Wówczas następujące warunki są równoważne.

i) $\lambda \in Spec(\varphi)$

ii) $\mathrm{Ker}(\varphi - \lambda \cdot \mathrm{id}_V) \neq \{\mathbf{0}_V\}$

iii) $\det(A - \lambda I) = 0$

Wniosek 8.1.16. Niech $\varphi \in End(V)$, $\lambda \in Spec(\varphi)$. Niech $\mathcal{B}$ będzie bazą przestrzeni V oraz $A = M_\varphi(\mathcal{B}, \mathcal{B})$. Wówczas wektor $\mathbf{0}_V \neq v \in V$, $v = [x_1, x_2, \ldots, x_n]_\mathcal{B}$ jest wektorem własnym endomorfizmu φ, odpowiadającym wartości własnej λ wtedy i tylko wtedy, gdy

$$(A - \lambda I)X = \mathbf{0}, \quad \text{gdzie} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Przykład 8.1.17. Rozważmy zagadnienie własne dla operatora liniowego

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad f(x, y) = (x + 2y, 2x + y).$$

I. Wyznaczymy widmo operatora f. Macierzą f jest $A = M_f(\mathcal{B}_k, \mathcal{B}_k) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \in M_2(\mathbb{R})$.

$$f(x, y) = \lambda(x, y) \Leftrightarrow \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} = \lambda I \begin{bmatrix} x \\ y \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Układ jednorodny posiada niezerowe rozwiązania, gdy jest nieoznaczony.

$$\chi_f(\lambda) = \det(A - \lambda I) = (1 - \lambda)^2 - 4 = \lambda^2 - 2\lambda - 3 = (\lambda + 1)(\lambda - 3) = 0$$

Zatem $Spec(f) = \{\lambda_1 = -1, \lambda_2 = 3\}$.

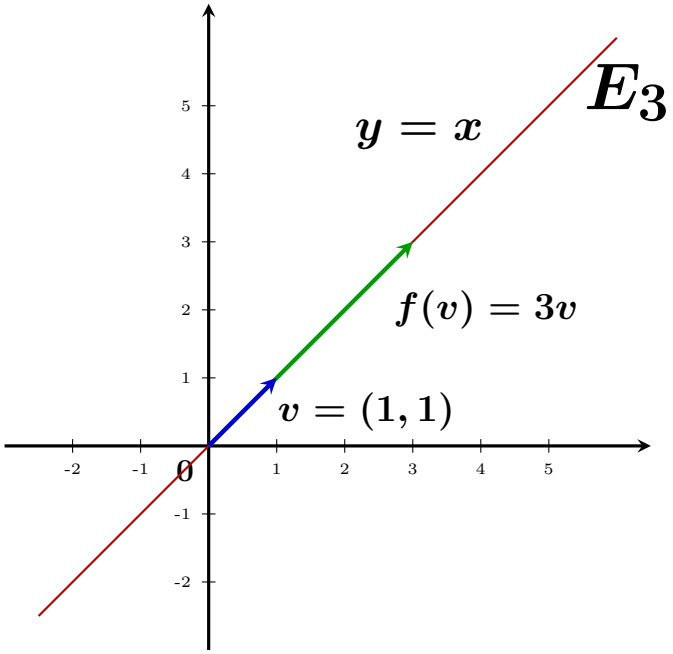
II. Wyznaczymy wektory własne odpowiadające wartości własnej $\lambda_2 = 3$.

$$f(x, y) = 3(x, y) \Leftrightarrow \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 3 \begin{bmatrix} x \\ y \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1-3 & 2 \\ 2 & 1-3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Stąd $-2x + 2y = 0$, czyli $x = y$.
Zbiór rozwiązań jednorodnego układu równań liniowych ma strukturę przestrzeni liniowej.
W tym wypadku jest to podprzestrzeń własna

$$E_{\lambda_2} = E_3 = \{(x, y) \in \mathbb{R}^2 : x = y\} = \{(x, x) : x \in \mathbb{R}\} = \text{lin}\{(1, 1)\}.$$

Każdy niezerowy wektor z przestrzeni E_3 jest wektorem własnym f odowiadającym wartości własnej $\lambda_2 = 3$. Jednym z nich jest wektor $v = (1, 1)$.



Podprzestrzeń własna E_3 jest prostą przechodzącą przez $(0, 0)$ o równaniu $y = x$.

Twierdzenie 8.1.18. Wartości własne endomorfizmu $\varphi \in End(V)$ nie zależą od wyboru bazy przestrzeni V.

Definicja 8.1.19. *Wielomianem charakterystycznym* endomorfizmu $\varphi \in End(V)$ nazywamy wielomian $\chi_\varphi \in K[\lambda]$ postaci $\chi_\varphi(\lambda) = \det(A - \lambda I)$, gdzie A jest reprezentacją macierzową odwzorowania φ w pewnej bazie przestrzeni V. [a] Równanie $\chi_\varphi(\lambda) = 0$ nazywamy *równaniem charakterystycznym* tego endomorfizmu. Pierwiastki wielomianu χ_φ nazywamy *pierwiastkami charakterystycznymi* odwzorowania φ.

[a] Na mocy twierdzenia 8.1.18 baza ta może być dowolnie wybrana.

Uwaga 8.1.20.

i) Pierwiastki charakterystyczne wielomianu χ_φ należące do ciała K to wartości własne endomorfizmu φ.

ii) Na mocy twierdzenia 8.1.18 wielomian χ_φ nie zależy od wyboru bazy przestrzeni V.

iii) Jeśli $\dim V = n$, to wówczas $\deg \chi_\varphi = n$.

Definicja 8.1.21. Niech $A \in M_n(K)$, gdzie $K = \mathbb{R}$ lub $K = \mathbb{C}$.

i) *Wielomianem charakterystycznym* macierzy A nazywamy wielomian $\chi_A \in K[\lambda]$ postaci $\chi_A(\lambda) = \det(A - \lambda I)$. Równanie $\chi_A(\lambda) = 0$ nazywamy *równaniem charakterystycznym* macierzy A.

ii) Każdy pierwiastek wielomianu χ_A należący do ciała K nazywamy *wartością własną* macierzy A. Zbiór wszystkich wartości własnych macierzy A oznaczamy symbolem $Spec(A)$ i nazywamy *widmem* bądź *spektrum* tej macierzy.

iii) Każdy niezerowy wektor $x = (x_1, x_2, \ldots, x_n) \in K^n$ spełniający równanie

$$AX = \lambda X, \quad \text{gdzie} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

nazywamy *wektorem własnym* macierzy A, odpowiadającym wartości własnej λ.

Uwaga 8.1.22. Wartości i wektory własne endomorfizmu $\varphi \in End(\mathbb{R}^n)$ są identyczne z wartościami i wektorami własnymi macierzy $A \in M_n(\mathbb{R})$, będącej reprezentacją macierzową odwzorowania φ w bazie kanonicznej przestrzeni $\mathbb{R}^n$.

Przykład 8.1.23. Czy wektor $v = (1, 1, 1)$ jest wektorem własnym macierzy $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$?

Jeśli tak, jakieś wartości własnej odpowiada?

$$A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Tak, jest to wektor własny macierzy A, który odpowiada wartości własnej $\lambda = 3$.

Uwaga 8.1.24. Niech $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ będzie dowolna macierzą kwadratową stopnia 2. Wówczas jej wielomian charakterystyczny ma postać

$$\chi_A(\lambda) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A).$$

Istotnie

$$\chi_A(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 - \underbrace{(a+d)}_{\mathrm{tr}A}\lambda + \underbrace{(ad - bc)}_{\det A}.$$

Przykład 8.1.25. Zapiszemy wielomian charakterystyczny macierzy $A \in M_2(\mathbb{R})$ takiej, że $tr(A) = -5$ oraz $\det A = -6$.

$$\chi_A(\lambda) = \lambda^2 + 5\lambda - 6$$

Niech V będzie przestrzenią liniową n-wymiarową. Niech $\varphi \in End(V)$ oraz $\lambda \in Spec(\varphi)$.

Definicja 8.1.26.
i) Krotność k_λ liczby λ jako pierwiastka wielomianu charakterystycznego χ_φ nazywamy *krotnością algebraiczną* wartości własnej λ.

ii) Wymiar $\dim E_\lambda$ podprzestrzeni własnej E_λ nazywamy *krotnością geometryczną* wartości własnej λ.

iii) Wartości własne o krotności geometrycznej 1 nazywamy *prostymi*. Widmo składające się z n różnych (a zatem prostych) wartości własnych nazywamy *widmem prostym*.

Twierdzenie 8.1.27. Niech $\varphi \in End(V)$, $\lambda \in Spec(\varphi)$ oraz niech A będzie reprezentacją macierzową φ w pewnej bazie przestrzeni V. Niech k_λ oznacza krotność algebraiczną $\lambda \in Spec(\varphi)$. Wówczas

i) $1 \leq \dim E_\lambda \leq k_\lambda$,

ii) $\dim E_\lambda = \dim V - \operatorname{rank}(A - \lambda I)$.

Szkic dowodu. i) Jeśli $\lambda \in Spec(\varphi)$, to istnieje $v \in V$, $v \neq \mathbf{0}_V$ taki, że $v \in E_\lambda$, zatem $1 \leq \dim E_\lambda$. Rozumowanie uzasadniające, że $\dim E_\lambda \leq k_\lambda$ można znaleźć w [4].
ii) Ponieważ $E_\lambda = \operatorname{Ker}(\varphi - \lambda \cdot \mathrm{id}_V)$, zatem
$\dim E_\lambda = \dim V - \dim \operatorname{Im}(\varphi - \lambda \cdot \mathrm{id}_V) = \dim V - \operatorname{rank}(A - \lambda I)$. $\square$

Przykład 8.1.28. Wyznacz wszystkie wartości własne endomorfizmu

$$\varphi \in End(\mathbb{R}^3), \quad \varphi(x, y, z) = (x + 2y, 2y, -2x - 2y - z).$$

Określ ich krotności algebraiczne. Wyznacz odpowiadające im wektory własne, podprzestrzenie własne i wymiary tychże podprzestrzeni.

$$A = M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^3) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ -2 & -2 & -1 \end{bmatrix}$$

$$\chi_A(\lambda) = \chi_\varphi(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 0 & 2-\lambda & 0 \\ -2 & -2 & -1-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)(-1-\lambda)$$

$Spec(\varphi) = \{\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -1\}$, $k_1 = k_2 = k_3 = 1$ widmo proste

$E_{\lambda_3} = E_{-1} = ?$

$$(A - \lambda_3 I)X = \mathbf{0}_{\mathbb{R}^3} \Leftrightarrow (A + I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 2 & 2 & 0 \\ 0 & 3 & 0 \\ -2 & -2 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} 2x + 2y &= 0 \\ 3y &= 0 \\ -2x - 2y &= 0 \end{cases} \Rightarrow x = y = 0,\ z \in \mathbb{R}$$

METODA I:
Wektory własne odpowiadające $\lambda_3 = -1$ są postaci $v = (0, 0, z)$, gdzie $z \in \mathbb{R} \setminus \{0\}$.
$E_{-1} = \{(0, 0, z) : z \in \mathbb{R}\} = \text{lin}\{(0, 0, 1)\}$ oraz $\dim E_{-1} = 1$

METODA II:
Alternatywnie, korzystając z twierdzenia 8.1.27 ii), możemy obliczyć

$$\dim E_{-1} = \dim \mathbb{R}^3 - r(A + I) = 3 - r\begin{bmatrix} 2 & 2 & 0 \\ 0 & 3 & 0 \\ -2 & -2 & 0 \end{bmatrix} = 3 - 2 = 1.$$

METODA III:
Tak naprawdę, wiemy to z góry, bowiem na mocy twierdzenia 8.1.27 i) mamy
$1 \leq \dim E_{\lambda_3} \leq k_3 = 1$, zatem $\dim E_{\lambda_3} = 1$.

Analogicznie wyznaczamy E_{λ_1} oraz E_{λ_2}. Z góry wiemy, że $\dim E_{\lambda_1} = \dim E_{\lambda_2} = 1$

8.2 Diagonalizacja

Twierdzenie 8.2.1. Wektory własne operatora liniowego $\varphi \in End(V)$ odpowiadające różnym wartościom własnym są liniowo niezależne.

Przykład 8.2.2. Dany jest endomorfizm $f : \mathbb{R}^3 \to \mathbb{R}^3$ oraz wektory $u, v, w \in \mathbb{R}^3$ takie, że

$$f(u) = \pi u, \quad f(v) = \sqrt{2}v, \quad f(w) = \sqrt{2}w.$$

Zatem $u \in E_\pi$, $v, w \in E_{\sqrt{2}}$.
Wektory u, v są liniowo niezależne. Podobnie wektory u, w są liniowo niezależne

Wektory v, w nie muszą być liniowo niezależne.
Jeśli $\dim E_{\sqrt{2}} = 1$, to wektory v, w są liniowo zależne.
Jeśli $\dim E_{\sqrt{2}} = 2$, to wektory v, w mogą być liniowo zależne lub liniowo niezależne.

Twierdzenie 8.2.3. Niech V będzie rzeczywistą (zespoloną) przestrzenią liniową n-wymiarową, zaś $\varphi \in End(V)$.

i) Jeśli φ ma widmo proste $Spec(\varphi) = \{\lambda_1, \ldots, \lambda_n\}$, to wektory własne $v_1, \ldots, v_n$, gdzie v_i odpowiada λ_i, dla $i \in \{1, 2, \ldots, n\}$, tworzą bazę przestrzeni V.

ii) Jeśli wektory własne $v_1, \ldots, v_n \in V$ endomorfizmu φ (nie koniecznie odpowiadające różnym wartościom własnym) tworzą bazę przestrzeni V oraz $\varphi(v_i) = \lambda_i v_i$, dla $i \in \{1, 2, \ldots, n\}$, to macierz przekształcenia φ w tejże bazie ma postać diagonalną

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & & \ldots & \lambda_n \end{bmatrix}.$$

iii) (**Twierdzenie spektralne**) Jeśli wielomian charakterystyczny χ_φ rozkłada się na czynniki liniowe

$$\chi_\varphi = (t - \lambda_1)^{k_1}(t - \lambda_2)^{k_2} \ldots (t - \lambda_p)^{k_p},$$

(tzn. $\lambda_i \neq \lambda_j$, gdy $i \neq j$ oraz $k_1 + k_2 + \ldots + k_p = n$) i ponadto $k_i = \dim E_{\lambda_i}$, dla $i \in \{1, 2 \ldots, p\}$, to wówczas istnieje baza przestrzeni V złożona z wektorów własnych endomorfizmu φ.

Definicja 8.2.4. Operator liniowy $\varphi \in End(V)$ nazywamy *diagonalizowalnym*, gdy istnieje taka baza przestrzeni V, że macierz operatora φ w tejże bazie jest macierzą diagonalną.

Wniosek 8.2.5. Operator liniowy $\varphi \in End(V)$ jest diagonalizowalny wtedy i tylko wtedy, gdy istnieje baza przestrzeni V złożona z wektorów własnych φ. Dokładniej mówiąc, $\varphi \in End(V)$, gdzie $\dim V = n$, jest diagonalizowalny wtedy i tylko wtedy, gdy wielomian charakterystyczny ma n pierwiastków w ciele $\mathbb{R}$ (licząc z krotnościami) oraz dla każdej wartości własnej można wybrać tyle liniowo niezależnych wektorów własnych, ile wynosi krotność tej wartości własnej jako pierwiastka wielomianu charakterystycznego.

Uwaga 8.2.6. Z zasadniczego twierdzenia algebry wynika, że każdy operator liniowy na zespolonej przestrzeni liniowej ma wektor własny.

Przykład 8.2.7. Niech $\varphi \in End(\mathbb{R}^2)$ oraz $A = M_\varphi(\mathcal{B}_k^2, \mathcal{B}_k^2) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in M_2(\mathbb{R})$.

$\chi_A(\lambda) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 \neq 0.$
Wielomian charakterystyczny nie posiada pierwiastków rzeczywistych, zatem $Spec(A) = \varnothing$ i nie istnieje baza $\mathbb{R}^2$ złożona z wektorów własnych φ.

Przykład 8.2.8. Dana jest macierz $A \in M_2(\mathbb{R})$ taka, że $\lambda = \pi$ jest jej wartością własną. Macierz A nie jest diagonalizowalna. Jak wygląda jej wielomian charakterystyczny?

Wielomian charakterystyczny jest wielomianem stopnia drugiego i posiada pierwiastek rzeczywisty, zatem drugi pierwiastek również jest rzeczywisty.

Jedyna możliwość, to taka, że wielomian charakterystyczny ma jeden pierwiastek podwójny $\lambda = \pi$. Gdyby posiadał dwa różne pierwiastki o krotnościach algebraicznych 1, to ich krotności geometryczne również wynosiłyby 1 (widmo proste) i A byłaby diagonalizowalna.

Zatem $\chi_A(\lambda) = (\lambda - \pi)^2$ lub jest wielokrotnością tego wielomianu.

Macierz diagonalizująca

Definicja 8.2.9. Mówimy, że macierze $A, B \in M_n(K)$ *są podobne*, jeżeli istnieje macierz nieosobliwa $C \in GL_n(K)$ taka, że $A = C^{-1}BC$.

Twierdzenie 8.2.10 (o niezmiennikach macierzy podobnych)**.** Jeżeli macierze A i B są podobne, to wówczas

i) $r(A) = r(B)$,

ii) $\det A = \det B$

iii) $\mathrm{tr}(A) = \mathrm{tr}(B)$

iv) $Spec(A) = Spec(B)$

Uwaga 8.2.11. Jeżeli $\varphi \in End(V)$, to dla dowolnych baz $\mathcal{B}, \mathcal{B}'$ przestrzeni V macierze $M_\varphi(\mathcal{B}, \mathcal{B})$, $M_\varphi(\mathcal{B}', \mathcal{B}')$ są podobne. Macierz ustanawiająca relację podobieństwa jest macierzą $P_{\mathcal{B}\to\mathcal{B}'}$ zmiany bazy przestrzeni V.

Definicja 8.2.12. Macierz $A \in M_n(K)$ nazywamy *diagonalizowalną*, gdy jest podobna do macierzy diagonalnej, tzn. istnieje macierz nieosobliwa $P \in GL_n(K)$ taka, że macierz $P^{-1}AP$ jest diagonalna. Mówimy wówczas, że macierz P *diagonalizuje* macierz A.

Wniosek 8.2.13. Macierz $A \in M_n(K)$ jest diagonalizowalna wtedy i tylko wtedy, gdy istnieje baza przestrzeni K^n złożona z wektorów własnych A.

Uwaga 8.2.14. Z uwagi na związek między widmem endomorfizmu a widmem macierzy wszystkie twierdzenia udowodnione dla endomorfizmu są prawdziwe dla macierzy.

Uwaga 8.2.15. Niech V będzie przestrzenią liniową z bazą kanoniczną $\mathcal{B}$. Niech $\varphi \in End(V)$ będzie diagonalizowalny oraz niech $\mathcal{B}'$ będzie bazą przestrzeni V złożoną z wektorów własnych φ. Oznaczmy $A = M_\varphi(\mathcal{B}_k, \mathcal{B}_k)$, $D = M_\varphi(\mathcal{B}', \mathcal{B}')$ oraz $P = P_{\mathcal{B}_k \to \mathcal{B}'}$. Wówczas $A = PDP^{-1}$.

$$\boldsymbol{AX = P\,D\,P^{-1}X}$$

przejście od bazy wektorów własnych do bazy kanonicznej

macierz diagonalna

przejście od bazy kanonicznej do bazy wektorów własnych

Każda współrzędna wektora $\boldsymbol{P^{-1}X}$ jest mnożona przez stałą (będącą wartością własną).

Zmiana każdej współrzędnej zależy tylko od niej samej, a nie od innych współrzędnych.

związek pomiędzy macierzami $\boldsymbol{A}$ i $\boldsymbol{D}$ reprezentującymi ten sam endomorfizm w różnych bazach

Przykład 8.2.16. Czy $A = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \in M_2(\mathbb{R})$ jest diagonalizowalna?

$$\chi_A(\lambda) = \begin{vmatrix} 2-\lambda & -3 \\ 1 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda) + 3 = \lambda^2 - 3\lambda + 5 = \left(\lambda - \tfrac{3}{2}\right)^2 + \tfrac{11}{4} \neq 0.$$

Wielomian charakterystyczny nie posiada pierwiastków rzeczywistych, zatem $Spec(A) = \varnothing$ i macierz rzeczywista A nie jest diagonalizowalna.

Innymi słowy, nie istnieje baza przestrzeni $(\mathbb{R}^2, +, \mathbb{R}, \cdot)$ złożona z wekorów własny macierzy A.

Przykład 8.2.17. Czy $A = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \in M_2(\mathbb{C})$ jest diagonalizowalna?

$\chi_A(\lambda) = \begin{vmatrix} 2-\lambda & -3 \\ 1 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)+3 = \lambda^2 - 3\lambda + 5.$

Wielomian charakterystyczny to wielomian zespolony o współczynnikach rzeczywistych.
Posiada zatem dwa sprzężone pierwiastki zespolone.

Mamy $\chi_A(\lambda) = (\lambda - \frac{3-\sqrt{11}i}{2})(\lambda - \frac{3+\sqrt{11}i}{2})$.
Macierz zespolona A ma widmo proste, zatem krotności geometryczne równają się krotnościom algebraicznym (i są równe 1), jest więc diagonalizowalna.

Innymi słowy, istnieje baza przestrzeni $(\mathbb{C}^2, +, \mathbb{C}, \cdot)$ złożona z wekorów własny macierzy A.

Przykład 8.2.18. Czy $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \in M_2(\mathbb{C})$ jest diagonalizowalna?

$\chi_A(\lambda) = \begin{vmatrix} 1-\lambda & 2 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$ oraz $Spec(A) = \{\lambda_1 = 1\}$.
Ponadto $\lambda_1 = 1$ ma krotność algebraiczną $k_1 = 2$.
Wyznaczamy podprzestrzeń własną $E_{\lambda_1} = \{v = (x,y) \in \mathbb{C}^2 : (A-I)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\}$.

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = (A-I)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2y \\ 0 \end{bmatrix}.$$

Stąd $E_{\lambda_1} = \{(x,0) : x \in \mathbb{C}\}$ i krotność geometryczna $\dim E_{\lambda_1} = 1 \neq k_1 = 2$.
Macierz A nie jest diagonalizowalna.

Przykład 8.2.19. Czy $\varphi \in End(\mathbb{R}^3)$ taki, że $\varphi(1,1,1) = (-1,-1,-1)$, $\varphi(0,1,1) = (0,1,1)$, $\varphi(0,0,1) = (0,0,2)$ jest diagonalizowalny?

Odczytujemy wartości własne i wektory własne

$$\lambda_1 = -1, v_1 = (1,1,1), \quad \lambda_2 = 1, v_1 = (0,1,1), \quad \lambda_3 = 2, v_1 = (0,0,1).$$

Widmo jest proste, a zatem na mocy twierdzenia 8.2.3 i), operator φ jest diagonalizowalny.

Przykład 8.2.20. Czy $\varphi \in End(\mathbb{R}^3)$ dany wzorem
$\varphi(x,y,z) = (3x+8z, 3x-y+6z, -2x-5z)$ jest diagonalizowalny?

$$A = M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^3) = \begin{bmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{bmatrix}$$

$$\chi_A(\lambda) = \begin{vmatrix} 3-\lambda & 0 & 8 \\ 3 & -1-\lambda & 6 \\ -2 & 0 & -5-\lambda \end{vmatrix} = (-1-\lambda)(-1)^4 \begin{vmatrix} 3-\lambda & 8 \\ -2 & -5-\lambda \end{vmatrix} =$$

$= -(1+\lambda)(1+2\lambda+\lambda^2) = -(1+\lambda)^3$
$Spec(\varphi) = \{\lambda_1 = -1\}$, $k_1 = 3$

$$E_{\lambda_1} = E_{-1} = \{v = (x,y,z) : (A+I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\}$$

$$\dim E_{-1} = 3 - r(A+I) = 3 - r\begin{bmatrix} 4 & 0 & 8 \\ 3 & 0 & 6 \\ -2 & 0 & -4 \end{bmatrix} = 3 - 1 = 2 \neq k_1 = 3.$$

Endomorfizm φ nie jest diagonalizowalny. Nie istnieje baza $\mathbb{R}^3$ złożona z wektorów własnych φ.

Przykład 8.2.21.
Wyznacz wszystkie wartości własne endomorfizmu f i określ ich krotności algebraiczne.
Wyznacz odpowiadające im podprzestrzenie własne i wymiar tychże podprzestrzeni.
Czy f jest diagonalizowalny? Odpowiedź uzasadnij.
Jeśli tak, podaj bazę wektorów własnych, macierz D endomorfizmu f w tejże bazie oraz macierz diagonalizującą P.

$$f \in End(M_2(\mathbb{R})), \quad f(A) = A + A^T$$

I. Wyznaczymy macierz odwzorowania f.

$$\mathcal{B} = \left(E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) \text{ baza } M_2(\mathbb{R})$$

$$f(E_{11}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$f(E_{12}) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$f(E_{21}) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$f(E_{22}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

$$A = M_f(\mathcal{B},\mathcal{B}) = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

II. Wyznaczymy widmo macierzy A.

$$\chi_f(\lambda) = \det(A - \lambda I) = (2-\lambda)\begin{vmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)^2\begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} =$$

$$= (2-\lambda)^2\Big((1-\lambda)^2 - 1\Big) = (2-\lambda)^2(\lambda^2 - 2\lambda) = -\lambda(2-\lambda)^3$$

$Spec(f) = \{\lambda_1 = 0, \lambda_2 = 2\}$, $k_1 = 1, k_2 = 3$, $\dim E_0 = 1$, $1 \leq \dim E_2 \leq 3$

III. Wyznaczymy podprzestrzenie własne.

$E_{\lambda_1} = E_0 = \mathrm{Ker}(f) = \{B \in M_2(\mathbb{R}) : B + B^T = \mathbf{O}\}$
Niech $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, gdzie $a, b, c, d \in \mathbb{R}$.

$$B + B^T = \mathbf{O} \Leftrightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow \begin{cases} 2a &= 0 \\ b+c &= 0 \\ 2d &= 0 \end{cases}$$

Zatem $B = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$ oraz $E_0 = \left\{ \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} : b \in \mathbb{R} \right\} = \mathrm{lin}\left\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$

$E_{\lambda_2} = E_2 = \mathrm{Ker}(f - 2 \cdot \mathrm{id}_{M_2(\mathbb{R})}) = \{B \in M_2(\mathbb{R}) : B + B^T = 2B\}$
$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = [a, b, c, d]_{\mathcal{B}}$, $\mathbf{O}_{2\times 2} = [0, 0, 0, 0]_{\mathcal{B}}$

$$(A - 2I) \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \mathbf{O} \Leftrightarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow b = c, \quad a, b, d \in \mathbb{R}$$

$$E_2 = \left\{ \begin{bmatrix} a & b \\ b & d \end{bmatrix} : a, b, d \in \mathbb{R} \right\} = \mathrm{lin}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ są liniowo niezależne $\Rightarrow \dim E_2 = 3$

IV. Podsumowanie.

$$\text{Baza wektorów własnych } \mathcal{C} = \left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$$

$$D = M_f(\mathcal{C}, \mathcal{C}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \; P_{\mathcal{B}\to\mathcal{C}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Przykład 8.2.22. Rozważmy $\varphi \in End(\mathbb{R}^2)$ dany wzorem $\varphi(x, y) = (4x + 2y, y - x)$.

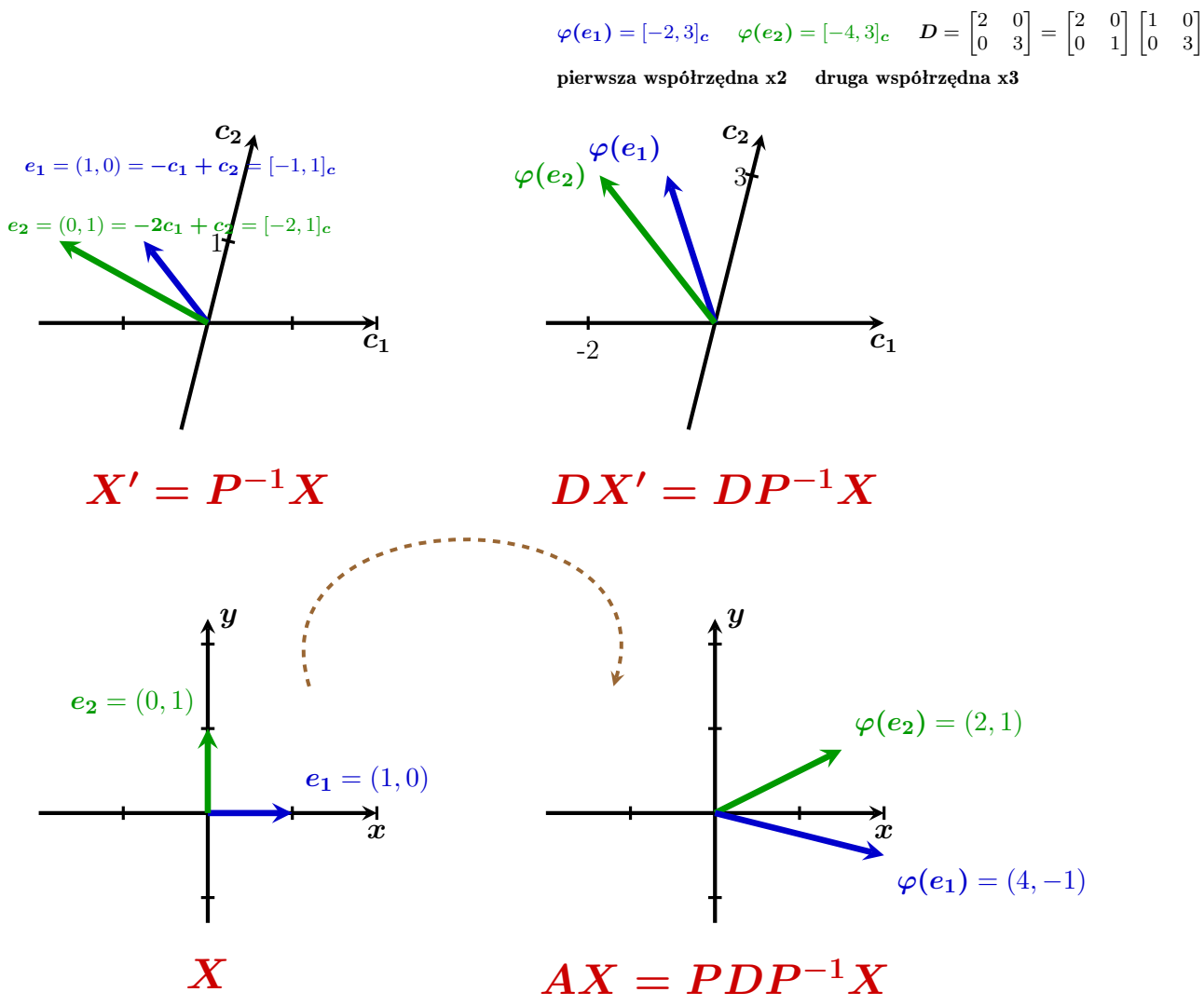
Możemy wyliczyć $Spec(\varphi) = \{\lambda_1 = 2, \lambda_2 = 3\}$.
Widmo jest proste, zatem φ jest diagonalizowalny.
Układ $\mathcal{C} = \big(v_1 = (1, -1), v_2 = (2, -1)\big)$ jest bazą wektorów własnych.

$$A = M_\varphi(\mathcal{B}_k^2, \mathcal{B}_k^2) = \begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix}, \; P = P_{\mathcal{B}_k^2\to\mathcal{C}} = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}, \; D = M_\varphi(\mathcal{C}, \mathcal{C}) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

Obliczamy $P^{-1} = P_{\mathcal{C}\to\mathcal{B}_k^2} = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}$ oraz

$$DP^{-1} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ 3 & 3 \end{bmatrix},$$

$$PDP^{-1} = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -2 & -4 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix} = A.$$



interpretacja geometryczna diagonalizacji

Uwaga 8.2.23. W dalszej części wykładu przekonamy się, że każdej wartości własnej o krotności algebraicznej k zeczywistej macierzy symetrycznej odpowiada k liniowo niezależnych wektorów własnych. A zatem rzeczywiste macierze symetryczne są zawsze diagonalizowalne. Co więcej, wśród wszystkich możliwych macierzy diagonalizujących daną rzeczywistą macierz symetryczną, znajdziemy zawsze taką macierz diagonalizująca P, dla której $P^{-1} = P^T$.

Twierdzenie 8.2.24. Niech V będzie skończenie wymiarową przestrzenią liniową nad ciałem $\mathbb{C}$. Niech $A \in M_n(\mathbb{C})$, $Spec(A) = \{\lambda_1, \ldots, \lambda_n\}$. Wówczas

i) $\det(A) = \chi_A(0) = \lambda_1 \cdot \ldots \cdot \lambda_n$,

ii) Jeśli $\lambda \in Spec(A)$, ale wielomian charakterystyczny χ_A macierzy A ma współczynniki rzeczywiste, to $\overline{\lambda} \in Spec(A)$. Ponadto wektor $w = (w_1, \ldots, w_n) \in V$ taki, że $\overline{w_i} = v_i$ jest wektorem własnym odpowiadającym $\overline{\lambda}$.

iii) $\mathrm{tr}(A) = \lambda_1 + \ldots + \lambda_n$

iv) $r(A)$ jest sumą krotności niezerowych wartości własnych.

Wniosek 8.2.25. Niech V będzie skończenie wymiarową przestrzenią liniową nad ciałem K. Macierz $A \in M_n(K)$ jest odwracalna $\Leftrightarrow 0 \notin Spec(A)$.

Przykład 8.2.26. Dana jest macierz diagonalizowalna $A \in M_3(\mathbb{R})$ taka, że $\lambda_1 = 5$ jest jej wartością własną o krotności algebraicznej 2, zaś $\lambda_2 = -2$ jest jej wartością własną o krotności algebraicznej 1. Obliczymy ślad, wyznacznik i rząd macierzy A.

$$\mathrm{tr}A = 5 + 5 - 2 = 8, \qquad \det A = 5 \cdot 5 \cdot (-2) = -50, \qquad r(A) = 3.$$

Przykład 8.2.27. Dana jest macierz diagonalizowalna $A \in M_4(\mathbb{R})$ taka, że $\mathrm{tr}A = 6$.
Ponadto $v_1 = (1, 0, 3, 0), v_2 = (0, 1, 2, 0)$ to wektory własne A odpowiadające wartości własnej $\lambda_1 = 2$, zaś $w_1 = (0, 0, 1, 1), w_2 = (0, 0, 5, 5)$ to wektory własne A odpowiadające wartości własnej $\lambda_2 = 3$. Obliczymy $\det A$.

Wektory v_1, v_2 są liniowo niezależne, zatem $\lambda_1 = 2$ ma krotność algebraiczną ≥ 2.
Wektory v_3, v_4 są liniowo zależne, zatem $\lambda_2 = 3$ ma krotność algebraiczną ≥ 1.

Wielomian charakterystyczny jest wielomianem stopnia 4.
Macierz A jest diagonalizowalna, więc posiada cztery wartości własne (licząc z krotnościami).

$$6 = \mathrm{tr}A = 2 + 2 + 3 + \lambda \quad \Rightarrow \quad \lambda = -1 \quad \Rightarrow \quad \det A = 2 \cdot 2 \cdot 3 \cdot (-1) = -12$$

Przykład 8.2.28. Niech $B \in M_3(\mathbb{C})$ będzie diagonalizowalna i taka, że wszystkie jej elementy mają zerową część urojoną.
Ponadto $\mathrm{tr}B = 10$ oraz $v_1 = (1, i, 0)$ jest wektorem własnym odpowiadającym wartości własnej $2 + i$.

Wielomian charakterystyczny macierzy B jest wielomianem zespolonym stopnia 3 o współczynnikach rzeczywistych. Jeśli $2 + i$ jest jednym z jego pierwiastków, to jest nim również $2 - i$.

Wektor $v_2 = (1, -i, 0)$ jest wektorem własnym odpowiadającym wartości własnej $2 - i$.

$$10 = \mathrm{tr}A = 2 + i + 2 - i + \lambda \quad \Rightarrow \quad \lambda = 6 \quad \Rightarrow \quad \det A = (2 + i) \cdot (2 - i) \cdot 6 = (4 - i^2) \cdot 6 = 30$$

8.3 Zastosowania diagonalizacji

Znajdowanie wartości złożenia endomorfizmu

Wniosek 8.3.1. Niech V będzie rzeczywistą przestrzenią liniową n-wymiarową oraz niech $\varphi \in End(V)$. Niech $\mathcal{B}$ będzie ustaloną bazą przestrzeni V. Jeśli wektory własne $v_1 \ldots, v_n$ odpowiadające wartościom własnym $\lambda_1, \ldots, \lambda_n$ (niekoniecznie różnym), tworzą bazę $\mathcal{C} = \left(v_1 \ldots, v_n\right)$ przestrzeni V, to wówczas dla dowolnego $r \in \mathbb{N}$

$$M_{\varphi^r}(\mathcal{B},\mathcal{B}) = PD^rP^{-1}, \quad \text{gdzie} \quad P = P_{\mathcal{B}\to\mathcal{C}}, \quad D = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & & \ldots & \lambda_n \end{bmatrix}.$$

Dowód. Niech $A = M_\varphi(\mathcal{B},\mathcal{B})$, $D = M_\varphi(\mathcal{C},\mathcal{C})$ oraz $P = P_{\mathcal{B}\to\mathcal{C}}$. Wówczas $A^r = M_{\varphi^r}(\mathcal{B},\mathcal{B})$ oraz $D^r = M_{\varphi^r}(\mathcal{C},\mathcal{C})$. Ponadto $D = P^{-1}AP$, skąd $A = PDP^{-1}$ oraz

$$A^r = \left(PDP^{-1}\right)^r = \underbrace{(PDP^{-1})(PDP^{-1})\ldots(PDP^{-1})}_{r-razy} = PD(P^{-1}P)D\ldots(PP^{-1})DP^{-1} = PD^rP^{-1}.$$

□

Przykład 8.3.2. Wyznacz wszystkie wartości własne endomorfizmu φ i określ ich krotności algebraiczne. Wyznacz odpowiadające im wektory własne, podprzestrzenie własne i wymiar tychże podprzestrzeni. Czy φ jest diagonalizowalny? Jeśli tak, podaj bazę wektorów własnych, macierz D endomorfizmu φ w tejże bazie oraz macierz diagonalizującą P. Oblicz $\varphi^{101}(1,2,3)$.

$$\varphi : \mathbb{R}^3 \to \mathbb{R}^3, \quad \varphi(x,y,z) = (4x - 2y + 2z, 2x + 2z, y + z - x)$$

I. Wyznaczymy widmo endomorfizmu φ.

$$A = M_\varphi(\mathcal{B}_k^3, \mathcal{B}_k^3) = \begin{bmatrix} 4 & -2 & 2 \\ 2 & 0 & 2 \\ -1 & 1 & 1 \end{bmatrix}$$

$$\chi_A(\lambda) = \begin{vmatrix} 4-\lambda & -2 & 2 \\ 2 & -\lambda & 2 \\ -1 & 1 & 1-\lambda \end{vmatrix} \overset{w_2+2w_3}{=} \begin{vmatrix} 4-\lambda & -2 & 2 \\ 0 & 2-\lambda & 4-2\lambda \\ -1 & 1 & 1-\lambda \end{vmatrix} \overset{k_3+k_2}{=} \begin{vmatrix} 4-\lambda & -2 & 0 \\ 0 & 2-\lambda & 6-3\lambda \\ -1 & 1 & 2-\lambda \end{vmatrix} \overset{k_2+k_1}{=}$$

$$= \begin{vmatrix} 4-\lambda & 2-\lambda & 0 \\ 0 & 2-\lambda & 6-3\lambda \\ -1 & 0 & 2-\lambda \end{vmatrix} = (4-\lambda)(2-\lambda)^2 - 3(2-\lambda)^2 = (2-\lambda)^2(1-\lambda)$$

$Spec\varphi = \{\lambda_1 = 1, \lambda_2 = 2\}$, $k_1 = 1$, $k_2 = 2$, $\dim E_1 = 1$, $1 \le \dim E_2 \le 2$

II. Uzasadnimy, że φ jest diagonalizowalny.

Wielomian charakterystyczny χ_A rozkłada się na iloczyn czynników liniowych.
Ponadto $\dim E_1 = 1 = k_1$ oraz

$$\dim E_2 = 3 - r(A - 2I) = 3 - r \begin{bmatrix} 2 & -2 & 2 \\ 2 & -2 & 2 \\ -1 & 1 & 1 \end{bmatrix} = 3 - 1 = 2 = k_2.$$

Na mocy twierdzenia spektralnego, φ jest diagonalizowalny.

III. Wyznaczymy bazę złożoną z wektorów własnych.

Rozważmy $\lambda_1 = 1$.

$$(A - I) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 3 & -2 & 2 \\ 2 & -1 & 2 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 3 & -2 & 2 & 0 \\ 2 & -1 & 2 & 0 \\ -1 & 1 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 3 & -2 & 2 & 0 \\ 2 & -1 & 2 & 0 \end{array}\right] \xrightarrow[w_3 - 2w_1]{w_2 - 3w_1} \left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right]$$

$$\begin{cases} x - y &= 0 \\ y + 2z &= 0 \end{cases} \Rightarrow x = y = -2z, z \in \mathbb{R}$$

$E_1 = \{(-2z, -2z, z) : z \in \mathbb{R}\} = \text{lin}\{(-2, -2, 1)\}$

Rozważmy $\lambda_2 = 2$.

$$(A - 2I) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 2 & -2 & 2 \\ 2 & -2 & 2 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$-x + y - z = 0 \Rightarrow z = y - x,\ x, y \in \mathbb{R}$

$E_2 = \{(x, y, y - x) : x, y \in \mathbb{R}\} = \text{lin}\{(1, 0, -1), (0, 1, 1)\}$

Bazą wektorów własnych jest $\mathcal{C} = \Big(c_1 = (-2, -2, 1), c_2 = (1, 0, -1), c_3 = (0, 1, 1)\Big)$

$D = M_\varphi(\mathcal{C}, \mathcal{C}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, macierz diagonalizująca $P := P_{\mathcal{B}_k^3 \to \mathcal{C}} = \begin{bmatrix} -2 & 1 & 0 \\ -2 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$

IV. Obliczymy $\varphi^{101}(1, 2, 3)$ na dwa różne sposoby.

METODA I: Wykorzystamy macierz D.

Niech $v := (1, 2, 3) = [a, b, c]_\mathcal{C}$. Wówczas $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = P \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, skąd $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = P^{-1} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Obliczamy $P^{-1} = \begin{bmatrix} 1 & -1 & 1 \\ 3 & -2 & 2 \\ 2 & -1 & 2 \end{bmatrix}$ oraz $v = [2, 5, 6]_\mathcal{C}$.

$$D^{101} \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^{101} & 0 \\ 0 & 0 & 2^{101} \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \cdot 2^{101} \\ 6 \cdot 2^{101} \end{bmatrix}$$

$$\varphi^{101}(v) = [2, 5\cdot 2^{101}, 6\cdot 2^{101}]_{\mathcal{C}} = 2c_1 + 5\cdot 2^{101}c_2 + 6\cdot 2^{101}c_3 = (-4+5\cdot 2^{101}, -4-6\cdot 2^{101}, 2+2^{101})$$

METODA II: Wykorzystamy macierz A.

$$A^{101}\begin{bmatrix}1\\2\\3\end{bmatrix} = PD^{101}P^{-1}\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}-2 & 1 & 0\\ -2 & 0 & 1\\ 1 & -1 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 2^{101} & 0\\ 0 & 0 & 2^{101}\end{bmatrix}\begin{bmatrix}1 & -1 & 1\\ 3 & -2 & 2\\ 2 & -1 & 2\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix} =$$

$$= \begin{bmatrix}3\cdot 2^{101}-2 & 2-2^{102} & 2^{102}-2\\ 2^{102}-2 & 2-2^{101} & 2^{102}-2\\ 1-2^{101} & -1+2^{101} & 1\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}-4+5\cdot 2^{101}\\ -4-6\cdot 2^{101}\\ 2+2^{101}\end{bmatrix}$$

Relacje rekurencyjne

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$.

Definicja 8.3.3. Niech $(a_n)_{n\in\mathbb{N}}$ będzie ciągiem zdefiniowanym następująco
i) $a_1 = r_1, a_2 = r_2, \ldots, a_k = r_k$, gdzie $r_1, r_2, \ldots, r_k \in K$, $k \geq 1$.
ii) $a_n = p_1a_{n-1} + p_2a_{n-2} + \ldots + p_ka_{n-k}$, gdzie $p_1, p_2, \ldots, p_k \in K$, $p_k \neq 0$, $n \geq k+1$.
Równanie ii) nazywamy *jednorodną liniową relacją rekurencyjną rzędu k*, zaś równania i) nazywamy *warunkami początkowymi rekurencji*.

Niech $X_n = [a_n, a_{n-1}, \ldots, a_{n-k+1}]^T$. Wówczas powyższy układ równań można zapisać w postaci

$$\begin{bmatrix}a_n\\a_{n-1}\\a_{n-2}\\\vdots\\a_{n-k+1}\end{bmatrix} = \underbrace{\begin{bmatrix}p_1 & p_2 & p_3 & \ldots & p_k\\ 1 & 0 & 0 & \ldots & 0\\ 0 & 1 & 0 & \ldots & 0\\ 0 & 0 & 1 & \ldots & 0\\ & \ldots & & \ldots & \\ 0 & 0 & & 1 & 0\end{bmatrix}}_{A_k}\begin{bmatrix}a_{n-1}\\a_{n-2}\\a_{n-3}\\\vdots\\a_{n-k}\end{bmatrix},$$

czyli $X_n = A_kX_{n-1}$, dla $n \geq k+1$.

Zauważmy, że $X_n = A_kX_{n-1} = A_k^2X_{n-2} = A_k^3X_{n-3} = \ldots = A_k^{n-k}X_k$, czyli

$$\begin{bmatrix}a_n\\a_{n-1}\\a_{n-2}\\\vdots\\a_{n-k+1}\end{bmatrix} = \begin{bmatrix}p_1 & p_2 & p_3 & \ldots & p_k\\ 1 & 0 & 0 & \ldots & 0\\ 0 & 1 & 0 & \ldots & 0\\ 0 & 0 & 1 & \ldots & 0\\ & \ldots & & \ldots & \\ 0 & 0 & & 1 & 0\end{bmatrix}^{n-k}\begin{bmatrix}r_k\\r_{k-1}\\r_{k-2}\\\vdots\\r_1\end{bmatrix}.$$

OBSERWACJA: Aby znaleźć wzór ogólny a_n, należy wyznaczyć potęgę macierzy A_k.
Jeśli A_k jest diagonalizowalna, możemy wykorzystać metodę z poprzedniego przykładu.

Przykład 8.3.4. Znajdziemy n-ty wyraz ciągu zadanego rekurencyjnie

$$a_1 = 0, \quad a_2 = 8, \quad a_n = 2a_{n-1} + 3a_{n-2}.$$

Otrzymujemy układ $\begin{bmatrix} a_n \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_{n-2} \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}^{n-2} \begin{bmatrix} 8 \\ 0 \end{bmatrix}$.

Diagonalizujemy macierz $A_2 = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}$.

Obliczamy $\det(A_2 - \lambda I) = \begin{vmatrix} 2-\lambda & 3 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1)$.

$Spec(A_2) = \{-1, 3\}$ widmo proste, macierz diagonalizowalna

Wybieramy bazę wektorów własnych. Rozważmy $\lambda_1 = -1$.

$(A_2 + I)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $v_1 = (-1, 1)$

Rozważmy $\lambda_2 = 3$.

$(A_2 - 3I)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 1 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $v_2 = (3, 1)$

Macierz diagonalizująca to $P = \begin{bmatrix} -1 & 3 \\ 1 & 1 \end{bmatrix}$. Stąd

$$\begin{bmatrix} a_n \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}^{n-2} \begin{bmatrix} 8 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}^{n-2} \begin{bmatrix} -1 & 3 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 8 \\ 0 \end{bmatrix} =$$

$$= \begin{bmatrix} -1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (-1)^{n-2} & 0 \\ 0 & 3^{n-2} \end{bmatrix} \frac{1}{4} \begin{bmatrix} -1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \cdot (-1)^n + 2 \cdot 3^{n-1} \\ 2 \cdot (-1)^{n-1} + 2 \cdot 3^{n-2} \end{bmatrix}.$$

Zatem $a_n = 2 \cdot (-1)^n + 2 \cdot 3^{n-1}$.

8.4 Twierdzenie Cayleya-Hamiltona

Twierdzenie 8.4.1 (Cayleya-Hamiltona)**.** Każda macierz kwadratowa nad ciałem liczb rzeczywistych lub zespolonych spełnia swoje równanie charakterystyczne, czyli

$$\boxed{\chi_A(A) = \mathbf{0}.}$$

Przykład 8.4.2. Niech $A = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}$. Wiemy, że $\chi_A(\lambda) = \lambda^2 - 2\lambda - 3$.

Na mocy powyższego twierdzenia $A^2 - 2A - 3I = \mathbf{0}$. Sprawdzimy to rachunkiem.

$$\begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}^2 - 2\begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix} - 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 6 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} 4 & 6 \\ 2 & 0 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Wniosek 8.4.3. Niech $A \in M_n(K)$, gdzie $K = \mathbb{R}$ lub $K = \mathbb{C}$, $n \in \mathbb{N}$, $n \geq 2$. Niech $\chi_A(\lambda) = c_n\lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0$ będzie wielomianem charakterystycznym macierzy A.

i) Jeśli A jest odwracalna, to wówczas
$A^{-1} = -\frac{1}{c_0}(c_nA^{n-1} + c_{n-1}A^{n-2} + \ldots + c_2A + c_1I)$

ii) $A^n = -\frac{1}{c_n}(c_{n-1}A^{n-1} + \ldots + c_1A + c_0I)$

iii) Dowolną całkowitą potęgę macierzy stopnia n można zapisać w postaci wielomianu macierzy stopnia co najwyżej $n-1$.

Przykład 8.4.4. Dana jest macierz $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \in M_2(\mathbb{R})$. Korzystając z twierdzenia Cayleya-Hamiltona, wyznaczymy macierz odwrotną A^{-1}.
$\det A = -5 \neq 0$ macierz jest odwracalna

$$\chi_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{vmatrix} = \lambda^2 - 4\lambda - 5 \quad \Rightarrow \quad A^2 - 4A - 5I = 0$$

$$A^{-1}(A^2 - 4A - 5I) = A - 4I - 5A^{-1} = 0 \quad \Rightarrow \quad A^{-1} = \tfrac{1}{5}(A - 4I) = \tfrac{1}{5}\left(\begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}\right)$$

$$A^{-1} = \tfrac{1}{5}\begin{bmatrix} -3 & 4 \\ 2 & -1 \end{bmatrix}$$

Przykład 8.4.5. Dana jest macierz $B = \begin{bmatrix} 0 & 2 & 6 \\ 2 & -8 & -26 \\ -2 & 2 & 8 \end{bmatrix} \in M_3(\mathbb{R})$. Korzystając z twierdzenia Cayleya-Hamiltona, wyznaczymy B^9.

Obliczamy $\chi_B(\lambda) = \det(B - \lambda I) = \begin{vmatrix} -\lambda & 2 & 6 \\ 2 & -8-\lambda & -26 \\ -2 & 2 & 8-\lambda \end{vmatrix} \overset{w_1-w_3,\, w_2+w_3}{=} \begin{vmatrix} -\lambda+2 & 0 & -2+\lambda \\ 0 & -6-\lambda & -18-\lambda \\ -2 & 2 & 8-\lambda \end{vmatrix}$

$$\overset{k_3-3k_2}{=} \begin{vmatrix} -\lambda+2 & 0 & -2+\lambda \\ 0 & -6-\lambda & 2\lambda \\ -2 & 2 & 2-\lambda \end{vmatrix} \overset{w_3+w_1}{=} \begin{vmatrix} -\lambda+2 & 0 & -2+\lambda \\ 0 & -6-\lambda & 2\lambda \\ -\lambda & 2 & 0 \end{vmatrix} =$$

$= (2-\lambda)(6+\lambda)\lambda - 4\lambda(2-\lambda) = -\lambda^3 + 4\lambda$.
Na mocy twierdzenia Cayleya-Hamiltona $B^3 = 4B$, skąd

$$B^9 = 4^3B^3 = 4^4B = 256\begin{bmatrix} 0 & 2 & 6 \\ 2 & -8 & -26 \\ -2 & 2 & 8 \end{bmatrix}.$$

Przykład 8.4.6. Dana jest macierz $C = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in M_2(\mathbb{R})$. Korzystając z twierdzenia Cayleya-Hamiltona, wyznaczymy $C^5 + C^3$ oraz C^{50}.

I. Oznaczmy $p(\lambda) = \lambda^5 + \lambda^3$. Obliczamy $\chi_C(\lambda) = (1-\lambda)^2$.

Dzielimy wielomian p przez wielomian χ_C (na ogół z resztą).

Zatem istnieją wielomiany q_1, r_1 takie że

$$p(\lambda) = q_1(\lambda)\chi_C(\lambda) + r_1(\lambda) \quad \text{oraz} \quad \deg(r_1) < \deg(\chi_C).$$

Wykonując dzielenie, otrzymujemy $q_1(\lambda) = \lambda^3 + 2\lambda^2 + 4\lambda + 6,\ r_1(\lambda) = 8\lambda - 6.$

Na mocy twierdzenia Cayleya-Hamiltona
$p(C) = \chi_C(C)q_1(C) + r_1(C) = r_1(C) = 8C - 6I = 8\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - 6\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ 0 & 2 \end{bmatrix}.$

II. Oznaczmy $w(\lambda) = \lambda^{50}$. Istnieją wielomiany q_2, r_2 takie że

$$w(\lambda) = q_2(\lambda)\chi_C(\lambda) + r_2(\lambda) \quad \text{oraz} \quad \deg(r_2) < \deg(\chi_C) = 2.$$

Oznaczmy $r_2(\lambda) = a\lambda + b$, $a, b \in \mathbb{R}$.
Różniczkując obustronnie otrzymujemy $w'(\lambda) = q_2'(\lambda)\chi_C(\lambda) + q_2(\lambda)\chi_C'(\lambda) + r_2'(\lambda)$.
Podstawiamy $\lambda = 1$ do układu równań

$$\begin{cases} \lambda^{50} &= q_2(\lambda)(1-\lambda)^2 + a\lambda + b \\ 50\lambda^{49} &= q_2'(\lambda)(1-\lambda)^2 - 2q_2(\lambda)(1-\lambda) + a \end{cases}.$$

Stąd $\begin{cases} 1 &= a + b \\ 50 &= a \end{cases}$. Zatem $r_2(\lambda) = 50\lambda - 49$ oraz $C^{50} = 50C - 49I = \begin{bmatrix} 1 & 50 \\ 0 & 1 \end{bmatrix}.$

QUIZ

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. Jeśli $f : \mathbb{R}^2 \to \mathbb{R}^2$ jest odbiciem względem prostej $y = 678x$, to $Spec(f) = \{-1, 1\}$.
2. Jeśli u, v to dowolne wektory własne macierzy A, to wówczas $2u + 3v$ również jest wektorem własnym macierzy A.
3. Jeśli v jest wektorem własnym macierzy A, to jest on również wektorem własnym macierzy $2A$.
4. Jeśli $A \in M_2(\mathbb{R})$ oraz $A\begin{bmatrix} 2 \\ 10 \end{bmatrix} = \begin{bmatrix} 4 \\ 20 \end{bmatrix}$, to macierz $A - 2I$ jest odwracalna.
5. Wektory własne danego endomorfizmu odpowiadające różnym wartościom własnym są liniowo niezależne.
6. Macierz $A = \begin{bmatrix} \pi & 1 & 1 & 1 \\ 0 & \pi & 0 & 1 \\ 0 & 0 & \pi & 1 \\ 0 & 0 & 0 & \pi + 1 \end{bmatrix} \in M_4(\mathbb{R})$ jest diagonalizowalna.
7. Dla dowolnego $p \in \mathbb{R}$ macierz $A = \begin{bmatrix} 2 & p & 0 \\ 0 & 3 & p \\ 0 & 0 & 2 \end{bmatrix}$ jest diagonalizowalna.

8. Niech $\varphi \in End(\mathbb{R}^5)$ będzie taki, że $Spec(\varphi) = \{\lambda_1 = -10, \lambda_2 = 5\}$. Jeśli wymiar podprzestrzeni własnej odpowiadającej $\lambda_2 = 5$ równa się 4, to wówczas φ jest diagonalizowalny.

9. Niech A będzie macierzą endomorfizmu $f : \mathbb{R}^3 \to \mathbb{R}^3$. Jeśli $\det(A - \lambda I) = (1 + \lambda)(\lambda - 7)^2$ oraz $f(0, 1, 0) = (0, 7, 0)$, $f(1, 2, 0) = (7, 14, 0)$, to wówczas f jest diagonalizowalny.

10. Jeśli macierz A jest odwracalna i diagonalizowalna, to macierz A^{-1} również jest diagonalizowalna.

11. Dana jest macierz $A \in M_2(\mathbb{R})$. Jeśli $Spec(A) = \{-3, \frac{1}{3}\}$, to wówczas $\det(A^{100}) = 1$.

12. Dowolna macierz $A \in M_3(\mathbb{R})$ taka, że $\det(A - \lambda I) = \lambda(\lambda - 5)(\lambda + 10)$ jest odwracalna.

13. Jeśli $Spec(f) = \{-2, -1, 0\}$, to f jest izomorfizmem.

14. Niech $A \in M_3(\mathbb{R})$. Jeśli $A \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, to wówczas $\det A = 0$.

15. Jeśli $A \in M_3(\mathbb{R})$ oraz $\det(A - \lambda I) = -\lambda^3 + 100$, to układ równań $AX = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ jest oznaczony.

16. Jeśli $A = \begin{bmatrix} 1 & 0 & 0 \\ 5 & -1 & 3 \\ 10 & 0 & 0 \end{bmatrix} \in M_3(\mathbb{R})$, to wówczas $A^2 = A^4$.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

TEMAT: *Przestrzenie euklidesowe*

9.1 Iloczyn skalarny i norma

Niech $V = (V, +, \mathbb{R}, \cdot)$ będzie rzeczywistą przestrzenią liniową.

Definicja 9.1.1. Funkcję $s : V \times V \to \mathbb{R}$ nazywamy *iloczynem skalarnym* lub *iloczynem wewnętrznym*, jeżeli spełnia ona następujące warunki.

i) $\forall u, v, w \in V \quad \forall \alpha, \beta \in \mathbb{R} \quad s(\alpha u + \beta v, w) = \alpha s(u, w) + \beta s(v, w)$, [a]

ii) $\forall u, v \in V \quad s(u, v) = s(v, u)$, [b]

iii) $\forall v \in V \quad s(v, v) \geq 0 \quad \wedge \quad s(v, v) = 0 \Leftrightarrow v = \mathbf{0}_V$. [c]

[a] Jest to warunek *liniowości ze względu na pierwszą zmienną.*
[b] warunek symetrii
[c] warunki *dodatniej określoności* $\forall v \in \setminus\{\mathbf{0}\} \quad s(v, v) > 0$ i niezdegenerowania

Parę (V, s) nazywamy wówczas *przestrzenią euklidesową*. Bywa ona oznaczana symbolem E. Zamiast $s(u, v)$ będziemy również pisać $u \circ v$ lub $\langle u, v \rangle$. Zamiast s będziemy też pisać $\langle ., . \rangle$.



Uwaga 9.1.2. Jeśli $\mathcal{B}$ jest bazą przestrzeni V, oraz $u, v \in V$, $u = [x_1, x_2, \ldots, x_n]_{\mathcal{B}}$, $v = [y_1, y_2, \ldots, y_n]_{\mathcal{B}}$, to każdy iloczyn skalarny jest postaci $\boxed{s(u,v) = X^T A Y}$, gdzie $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ oraz $A = [a_{ij}] \in M_n(K)$, gdzie $a_{ij} = s(b_i, b_j)$.
Warunki i) oraz ii) równoważne są stwierdzeniu, że $s(u,v)$ jest wielomianem jednorodnym stopnia 2 postaci $\sum_{i,j=1}^{n} a_{ij} x_i y_j$, zaś macierz A jest macierzą symetryczną.

Warunek iii) w powyższej definicji równoważny jest stwierdzeniu, że macierz A jest dodatnio określona, co zachodzi wtedy i tylko wtedy, gdy wszystkie minory wiodące główne macierzy A są dodatnie, lub równoważnie, gdy wszystkie wartości własne macierzy A są dodatnie.
Minorem wiodącym głównym stopnia k macierzy $A = [a_{ij}]$ nazywamy wyznacznik macierzy powstałej przez skreślenie $n-k$ ostatnich wierszy i kolumn. Oznaczamy go symbolem Δ_k.

Przykład 9.1.3. Poniższe funkcje są iloczynami skalarnymi.

1) *Standardowym iloczynem skalarnym w* $\mathbb{R}^n$ nazywamy $\circ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ taką, że

$\forall u = (x_1, \ldots, x_n), v = (y_1, \ldots, y_n) \in \mathbb{R}^n \quad \boxed{u \circ v = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \ldots + x_n y_n.}$

Przestrzeń euklidesową $(\mathbb{R}^n, \circ)$ oznaczamy symbolem $\mathbb{E}^n$.

2) $V = \mathcal{C}([a,b], \mathbb{R})$, $\quad \langle ., . \rangle : V \times V \to \mathbb{R}$, $\quad \forall f, g \in V \quad \boxed{\langle f, g \rangle = \int_a^b f(x) g(x) dx.}$

Całka oznaczona ma własność liniowości.
Ponadto $\int_a^b f^2(x) dx \geq 0$, bowiem $f^2(x) \geq 0$ i całka oznaczona zachowuje nierówność słabą.

3) $V = \mathbb{R}_n[x]$, $\quad \langle ., . \rangle : \mathbb{R}_n[x] \times \mathbb{R}_n[x] \to \mathbb{R}$,

$\forall p, q \in V \quad \boxed{\langle p, q \rangle = \sum_{i=0}^{n} p(x_i) q(x_i)}$, gdzie $x_0 < x_1 < \ldots < x_n$ liczby ustalone

Rozważmy $\mathbb{R}_2[x]$ z iloczynem skalarnym $\langle p, q \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$.
Wówczas $\langle p, p \rangle = [p(-1)]^2 + [p(0)]^2 + [p(1)]^2 \geq 0$.
Jeśli $\langle p, p \rangle = 0$, to oczywiście $p(-1) = p(0) = p(1) = 0$. Wielomian stopnia co najwyższej drugiego, posiadający trzy różne pierwiastki jest wielomianem zerowym.

4) $V = M_{m \times n}(\mathbb{R})$, $\quad \langle ., . \rangle : M_{m \times n}(\mathbb{R}) \times M_{m \times n}(\mathbb{R}) \to \mathbb{R}$, $\forall A, B \in M_{m \times n}(\mathbb{R}) \quad \boxed{\langle A, B \rangle = \operatorname{tr}(AB^T).}$

Na mocy twierdzenia 3.1.14, mówiącego o własnościach śladu macierzy, można wywnioskować, że jest to iloczyn skalarny.

Przykład 9.1.4. Uzasadnimy, że odwzorowanie $s : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ dane wzorem

$$s(x,y) = 13x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2 \quad \text{dla} \quad x = (x_2, x_2), y = (y_1, y_2) \in \mathbb{R}^2$$

jest iloczynem skalarnym w przestrzeni $\mathbb{R}^2$.

METODA I : Sprawdzimy warunki z definicji iloczynu skalarnego.

Dla dowolnych $x = (x_2, x_2), y = (y_1, y_2), z = (z_1, z_2) \in \mathbb{R}^2$, $\alpha, \beta \in \mathbb{R}$ mamy

$$s(\alpha x + \beta z, y) = 13(\alpha x_1 + \beta z_1)y_1 - 2(\alpha x_1 + \beta z_1)y_2 - 2(\alpha x_2 + \beta z_2)y_1 + (\alpha x_2 + \beta z_2)y_2 =$$

$$= \alpha\Big(13x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2\Big) + \beta\Big(13z_1y_1 - 2z_1y_2 - 2z_2y_1 + z_2y_2\Big) = \alpha s(x,y) + \beta s(z,y)$$

Zatem odwzorowanie s jest liniowe ze względu na pierwszą zmienną. Ponadto $s(x,y) = s(y,x)$ oraz

$$s(x,x) = 13x_1^2 - 4x_1x_2 + x_2^2 = 9x_1^2 + (4x_1^2 - 4x_1x_2 + x_2^2) = 9x_1^2 + (2x_1 - x_2)^2 \geq 0$$

Stąd $s(x,x) = 0 \Leftrightarrow x_1 = 0 \wedge 2x_1 - x_2 = 0 \Leftrightarrow x = (x_1, x_2) = (0,0)$.

METODA II : Skorzystamy z uwagi 9.1.2.

$s(x,y) = 13x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2$ jest wielomianem jednorodnym stopnia 2 postaci $\sum_{i,j=1}^{n} a_{ij}x_iy_j$.

Rozważmy bazę $\mathcal{B}_k^2 = (e_1 = (1,0), e_2 = (0,1))$.
Macierz $A = [a_{ij}]$, gdzie $a_{ij} = s(e_i, e_j)$ dla $i, i = 1, 2$ ma postać $A = \begin{bmatrix} 13 & -2 \\ -2 & 1 \end{bmatrix}$.
Łatwo zauważyć, że $A^T = A$. Ponadto $\Delta_1 = 13 > 0$ oraz $\Delta_2 = \det A = 9 > 0$.
Zatem A jest symetryczna i dodatnio określona. Stąd wnioskujemy, że s jest iloczynem skalarnym.

Przykład 9.1.5. W przestrzeni euklidesowej $(M_{2\times 3}(\mathbb{R}), \langle ., . \rangle)$, gdzie $\langle A, B \rangle = \mathrm{tr}(AB^T)$ obliczymy iloczyn skalarny wektorów $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ i $B = \begin{bmatrix} -1 & 0 & 3 \\ 2 & -5 & 1 \end{bmatrix}$.

$$\begin{aligned} \langle A, B \rangle &= \mathrm{tr}\left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 3 \\ 2 & -5 & 1 \end{bmatrix}^T \right) \\ &= \mathrm{tr}\left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} -1 & 2 \\ 0 & -5 \\ 3 & 1 \end{bmatrix} \right) \\ &= \mathrm{tr} \begin{bmatrix} 8 & 0 \\ 14 & -11 \end{bmatrix} = 8 - 11 = -3 \end{aligned}$$

Twierdzenie 9.1.6. Niech (V, s) będzie przestrzenią euklidesową. Wówczas

i) $\forall u, v, w \in V \quad \forall \alpha, \beta \in \mathbb{R} \quad s(u, \alpha v + \beta w) = \alpha s(u, v) + \beta s(u, w)$, [a]

ii) $\forall v \in V \quad s(v, \mathbf{0}_V) = 0$,

iii) $\forall u, v \in V \quad \Big(s(u, v)\Big)^2 \leq s(u, u) \cdot s(v, v)$ *nierówność Schwarza*

[a] Jest to warunek *liniowości ze względu na drugą zmienną.*

W przestrzeni $\mathbb{R}^2$ ze standardowym iloczynem skalarnym $u \circ v$ łatwo dowieść nierówności Schwarza.

$$(u \circ v)^2 = \Big(|u| \cdot |v| \cdot \cos \measuredangle(u, v)\Big)^2 \leq |u|^2 \cdot |v|^2 = (u \circ u) \cdot (v \circ v)$$

Niech $V = (V, +, \mathbb{R}, \cdot)$ będzie przestrzenią liniową.

Definicja 9.1.7. Funkcję $||.|| : V \to \mathbb{R}$ nazywamy *normą*, jeżeli spełnia ona następujące warunki:

i) $\forall v \in V \quad ||v|| \geq 0 \quad \wedge \quad ||v|| = 0 \Leftrightarrow v = \mathbf{0}_V$,

ii) $\forall v \in V \quad \forall \alpha \in \mathbb{R} \quad ||\alpha v|| = |\alpha| \cdot ||v||$,

iii) $\forall u, v \in V \quad ||u + v|| \leq ||u|| + ||v||$. tzw. warunek trójkąta

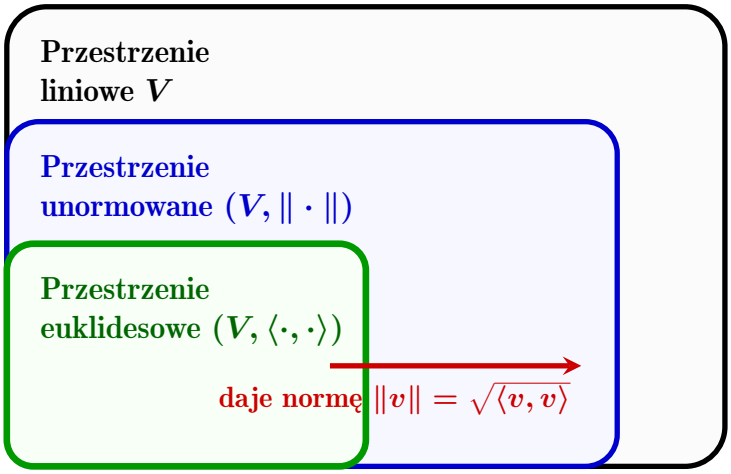
Liczbę $||v|| \geq 0$ nazywamy *normą (lub długością) wektora* v.
Parę $(V, ||.||)$ nazywamy *przestrzenią unormowaną.*

Twierdzenie 9.1.8. Jeśli (V, s) jest przestrzenią euklidesową, to odwzorowanie $||.||_s : V \to \mathbb{R}$ dane wzorem

$$\forall v \in V \quad ||v||_s = \sqrt{s(v, v)}$$

jest normą w przestrzeni V. Mówimy, że jest to *norma określona przez iloczyn skalarny.*

Wniosek 9.1.9. Każda przestrzeń euklidesowa jest przestrzenią unormowaną.



• **Możemy**
dodawać wektory i mnożyć przez liczby.
Nie znamy długości ani kątów.

• **Zyskujemy**
pojęcie **długości wektora** $(\|\cdot\|)$
oraz odległości między punktami.

• **Dodatkowo zyskujemy**
pojęcie **kąta** między wektorami
oraz **prostopadłości** $(\perp)$.

Jak iloczyn skalarny wprowadza geometrię (długości i kąty) do przestrzeni liniowej.

Przykład 9.1.10. Rozważmy przestrzeń $\mathbb{E}^n$, tj. $\mathbb{R}^n$ ze standardowym iloczynem skalarnym. Dla $x = (x_1, x_2, \ldots, x_n)$ mamy $||x|| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$. Jest to tzw. *norma euklidesowa* w $\mathbb{R}^n$.

Przykład 9.1.11. i) W przestrzeni euklidesowej $\Big(\mathcal{C}([0,1],\mathbb{R}), \langle .,. \rangle\Big)$, gdzie $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$ obliczymy normę wektora $h = x^3$.

$$||h|| = \sqrt{\langle h, h\rangle} = \sqrt{\int_0^1 x^6 dx} = \sqrt{\left[\frac{x^7}{7}\right]_0^1} = \sqrt{\frac{1}{7}} = \frac{\sqrt{7}}{7}$$

ii) W przestrzeni euklidesowej $(\mathbb{R}_2[x], \langle .,. \rangle)$, gdzie $\langle p, q\rangle = 2p(0)q(0) + 3p(1)q(1) + 4p(2)q(2)$ obliczymy normę wektora $r = x^2 - 1$.

$$||r|| = \sqrt{\langle r, r\rangle} = \sqrt{2[r(0)]^2 + 3[r(1)]^2 + 4[r(2)]^2} = \sqrt{2 \cdot 2 + 3 \cdot 0 + 4 \cdot 9} = \sqrt{40} = 2\sqrt{10}$$

iii) W przestrzeni euklidesowej $(\mathbb{R}^2, s)$, gdzie $s(x, y) = 13x_1y_1 - 2x_1y_2 - 2x_2y_1 + x_2y_2$ dla $x = (x_2, x_2), y = (y_1, y_2) \in \mathbb{R}^2$, obliczymy normę wektora $v = (-1, 3)$.

$$||v|| = \sqrt{s(v, v)} = \sqrt{13(-1)^2 - 4 \cdot (-1) \cdot 3 + 3^2} = \sqrt{34}$$

9.2 Układy ortogonalne

UMOWA: Jeśli nie wyszczególniono inaczej, zawsze w danej przestrzeni euklidesowej rozpatrujemy normę pochodzącą od ustalonego iloczynu skalarnego.

Niech $(V, \langle .,. \rangle)$ będzie przestrzenią euklidesową.
Wektor $v \in V$, którego długość jest równa 1 nazywamy *unormowanym* lub *wersorem*.

Każdy wektor $v \in V, v \neq \mathbf{0}_V$ można *unormować*, tj. znaleźć wersor $\hat{v}$ o tym samym zwrocie i kierunku co v. Istotnie, $\hat{v} := \frac{v}{||v||}$ jest poszukiwanym wersorem, bowiem

$$||\hat{v}|| = \left|\left|\frac{v}{||v||}\right|\right| = \left|\frac{1}{||v||}\right| \cdot ||v|| = \frac{1}{||v||} \cdot ||v|| = 1.$$

Miara kąta między wektorami
W przestrzeni euklidesowej można wprowadzić pojęcie kąta między niezerowymi wektorami.
Na mocy nierówności Schwarza dla dowolnych $u, v \in V$ mamy

$$|\langle u, v\rangle| \leq \sqrt{\langle u, u\rangle}\sqrt{\langle v, v\rangle} = ||u|| \cdot ||v|| \;\Rightarrow\; \frac{|\langle u, v\rangle|}{||u|| \cdot ||v||} \leq 1 \;\Leftrightarrow\; -1 \leq \frac{\langle u, v\rangle}{||u|| \cdot ||v||} \leq 1.$$

Jeśli $v \neq \mathbf{0}_V, u \neq \mathbf{0}_V$, możemy widzieć $\frac{\langle u,v\rangle}{||u||\cdot||v||}$ jako cosinus jednoznacznie określonego kąta $\alpha \in [0, \pi]$.

Definiujemy kąt między wektorami u i v jako α.
Utożsamiamy tutaj kąt z jego miarą. Zatem

$$\cos \measuredangle(u, v) = \frac{\langle u, v\rangle}{||u|| \cdot ||v||}, \quad \measuredangle(u, v) = \arccos \frac{\langle u, v\rangle}{||u|| \cdot ||v||}.$$

Przyjmujemy, że kąt pomiędzy wektorem zerowym $\mathbf{0}_V$ a innym wektorem jest nieokreślony.

Przykład 9.2.1. Rozważmy przestrzeń euklidesową $(\mathbb{R}_1[x], \langle ., .\rangle)$, gdzie $\langle p, q\rangle = p(0)q(0) + p(1)q(1)$. Wyznaczymy miarę kąta pomiędzy wektorami $u(x) = 2, v(x) = 5 - x$.

$\langle u, v\rangle = u(0)v(0) + u(1)v(1) = 2 \cdot 5 + 2 \cdot 4 = 18$

$||u|| = \sqrt{\langle u, u\rangle} = \sqrt{[u(0)]^2 + [u(1)]^2} = \sqrt{2^2 + 2^2} = 2\sqrt{2}$

$||v|| = \sqrt{\langle v, v\rangle} = \sqrt{[v(0)]^2 + [v(1)]^2} = \sqrt{5^2 + 4^2} = \sqrt{41}$

$\measuredangle(u, v) = \arccos \frac{18}{2\sqrt{2}\cdot\sqrt{41}} = \arccos \frac{9}{\sqrt{82}}$

Definicja 9.2.2. i) Dwa wektory u, v nazywamy *ortogonalnymi*, jeśli ich iloczyn skalarny jest równy zero. Piszemy wówczas $u \perp v$.

ii) Układ wektorów $\{v_1, \ldots v_n\} \subset V$ nazywamy *układem ortogonalnym*, jeżeli

$$\forall i, j \in \{1, 2, \ldots, n\} \quad i \neq j \;\Rightarrow \langle v_i, v_j\rangle = 0.$$

ii) Układ wektorów nazywamy *układem ortonormalnym*, jeśli jest układem ortogonalnym i każdy wektor jest unormowany, czyli

$$\forall i, j \in \{1, 2, \ldots, n\} \quad \langle v_i, v_j\rangle = \delta_{ij} = \begin{cases} 0 & ; \quad i \neq j \\ 1 & ; \quad i = j \end{cases}.$$

Uwaga 9.2.3. Wektor zerowy $\mathbf{0}_V$ jest ortogonalny do każdego wektora.

Przykład 9.2.4. Rozważmy przestrzeń euklidesową $(\mathcal{C}([-\pi,\pi],\mathbb{R}),\langle .,.\rangle)$, gdzie $\langle f,g\rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$. Sprawdzimy, czy wektory $f(x)=\sin x$, $g(x)=\cos x$ są ortogonalne/ortonormalne.

$\langle f,g\rangle = \int_{-\pi}^{\pi} \sin x\cos x dx = \frac{1}{2}\int_{-\pi}^{\pi}\sin 2x dx = \left[-\frac{1}{4}\cos 2x\right]_{-\pi}^{\pi} = \frac{1}{4}\Big(-1-(-1)\Big) = 0$

$\langle f,f\rangle = ||f||^2 = \int_{-\pi}^{\pi}\sin^2 x dx = \int_{-\pi}^{\pi}\frac{1-\cos 2x}{2}dx = \left[\frac{1}{2}x - \frac{1}{4}\sin 2x\right]_{-\pi}^{\pi} = \pi$

Wektory są ortogonalne, ale nie ortonormalne.

Przykład 9.2.5. Rozważmy przestrzeń euklidesową $(\mathbb{R}_2[x],\langle .,.\rangle)$, gdzie $\langle f,g\rangle = \int_0^1 f(x)g(x)dx$. Podamy wartości parametru $a\in\mathbb{R}$, dla których wektory $f=ax^2+1$ i $g=x+3$ są ortogonalne.

$$\langle f,g\rangle = \int_0^1 (ax^2+1)(x+3)dx = \int_0^1 (ax^3+3ax^2+x+3)dx = \left[a\frac{x^4}{4}+ax^3+\frac{1}{2}x^2+3x\right]_0^1 = \frac{a}{4}+a+\frac{1}{2}+3$$

$$\langle f,g\rangle = 0 \;\Leftrightarrow\; \frac{5}{4}a+\frac{7}{2}=0 \;\Leftrightarrow\; a=-\frac{14}{5}$$

Twierdzenie 9.2.6 (Pitagorasa)**.** Niech V będzie przestrzenią euklidesową. Wówczas

$$\forall u,v\in V \quad u\perp v \;\Leftrightarrow\; ||u+v||^2 = ||u||^2+||v||^2.$$

Dowód. $||u+v||^2 = \langle u+v,u+v\rangle = \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle \overset{\langle u,v\rangle=0}{=} ||u||^2+||v||^2$ □

Twierdzenie 9.2.7. Układ ortogonalny nie zawierający wektora zerowego jest liniowo niezależny.

Wniosek 9.2.8. i) Układ ortonormalny jest liniowo niezależny.

ii) W n-wymiarowej przestrzeni euklidesowej układ ortonormalny (lub układ ortogonalny nie zawierający wektora zerowego) nie może zawierać więcej niż n wektorów.

Definicja 9.2.9. Bazę przestrzeni euklidesowej, która jest układem ortogonalnym (ortonormalnym), nazywamy *bazą ortogonalną (ortonormalną)* tej przestrzeni.

Przykład 9.2.10. W przestrzeni $\mathbb{R}^n$ ze standardowym iloczynem skalarnym bazą ortogonalną jest baza kanoniczna $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)$.

Przykład 9.2.11. W przestrzeni $\mathbb{R}^3$ rozważamy standardowy iloczyn skalarny.
Dana jest podprzestrzeń liniowa $U = \text{lin}\{b_1 = (1, -3, -1),\ b_2 = (2, 1, -1)\}$ przestrzeni $\mathbb{R}^3$.
Podamy bazę przestrzeni U i sprawdzimy, czy jest ona ortogonalna.

Wektory b_1, b_2 są liniowo niezależne, bowiem $b_1 \neq \alpha b_2$, dla dowolnego $\alpha \in \mathbb{R}$.
Zatem układ (b_1, b_2) jest bazą przestrzeni U. Ponadto

$$b_1 \circ b_2 = 1 \cdot 2 - 3 \cdot 1 + (-1)^2 = 2 - 3 + 1 = 0,$$

zatem jest to baza ortogonalna.

Uwaga 9.2.12. Pojęcie ortogonalności jest zdefiniowane w danej przestrzeni euklidesowej, to jest za pomocą ustalonego iloczynu skalarnego.

Przykład 9.2.13. i) Rozważmy przestrzeń euklidesową $(\mathbb{R}_1[x], \langle ., . \rangle)$, gdzie $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$.
Sprawdzimy, czy wektory $p = x, q = x - \frac{2}{3}$ są wzajemnie ortogonalne i wyznaczymy kąt między nimi.

$$\langle p, q \rangle = \int_0^1 x\left(x - \frac{2}{3}\right) dx = \int_0^1 \left(x^2 - \frac{2}{3}x\right) dx = \left[\frac{x^3}{3} - \frac{x^2}{3}\right]_0^1 = 0$$

Zatem $p \perp q$. W konsekwencji $\measuredangle(p, q) = \frac{\pi}{2}$.

ii) Rozważmy przestrzeń euklidesową $(\mathbb{R}_1[x], \langle ., . \rangle)$, gdzie $\langle f, g \rangle = f(0)g(0) + f(1)g(1)$.
Sprawdzimy, czy wektory $p = x, q = x - \frac{2}{3}$ są wzajemnie ortogonalne i wyznaczymy kąt między nimi.

$$\langle p, q \rangle = p(0)q(0) + p(1)q(1) = 0 \cdot \left(-\frac{2}{3}\right) + 1 \cdot \frac{1}{3} = \frac{1}{3} \neq 0$$

Zatem $p \not\perp q$. Wyznaczymy kąt $\measuredangle(p, q)$.

$$||p|| = \sqrt{\langle p, p \rangle} = \sqrt{[p(0)]^2 + [p(1)]^2} = \sqrt{0^2 + 1^2} = 1$$

$$||q|| = \sqrt{\langle q, q \rangle} = \sqrt{[q(0)]^2 + [q(1)]^2} = \sqrt{\left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2} = \sqrt{\frac{5}{9}} = \frac{\sqrt{5}}{3}$$

$$\cos \measuredangle(p, q) = \frac{\langle p, q \rangle}{||p|| \cdot ||q||} = \frac{\frac{1}{3}}{1 \cdot \frac{\sqrt{5}}{3}} = \frac{1}{\sqrt{5}} = \frac{\sqrt{5}}{5} \quad \Rightarrow \quad \measuredangle(p, q) = \arccos \frac{\sqrt{5}}{5}$$

Przykład 9.2.14. W przestrzeni wektorowej $\mathbb{R}^2$ zadamy iloczyn skalarny w taki sposób by wektory $u = (2, -1)$ i $v = (0, 1)$ były ortogonalne oraz $||u|| = 1$, $||v|| = 4$.

Na mocy uwagi 9.1.2 wiemy, że $s\Big((x_1, x_2), (y_1, y_2)\Big) = ax_1y_1 + bx_1y_2 + bx_2y_1 + cy_1y_2$, dla pewnych $a, b, c \in \mathbb{R}$. Otrzymujemy zatem układ równań

$$\begin{cases} 0 = s(u,v) = 2b - c \\ 1 = s(u,u) = 4a - 4b + c \\ 4 = c \end{cases} \Rightarrow \begin{cases} b = \frac{c}{2} = 2 \\ a = \frac{1}{4}(1 + 4b - c) = \frac{5}{4} \\ c = 4 \end{cases}$$

Zatem

$$s\Big((x_1, x_2), (y_1, y_2)\Big) = \frac{5}{4}x_1y_1 + 2x_1y_2 + 2x_2y_1 + 4y_1y_2.$$

Współrzędne wektora w bazie ortogonalnej

MOTYWACJA: Przestrzeń $\mathbb{E}^3$
baza ortogonalna $\mathcal{B}_k^3 = \Big(\hat{i} = (1,0,0), \hat{j} = (0,1,0), \hat{k} = (0,0,1)\Big)$
Dla dowolnego wektora $v = (v_x, v_y, v_z)$ mamy $\boxed{v_x = v \circ \hat{i},\ v_y = v \circ \hat{j},\ v_z = v \circ \hat{k}.}$

Z dokładnością do znaku skalary v_x, v_y, v_z to długości rzutów ortogonalnych wektora v na osie Ox, Oy, Oz odpowiednio.
Wektory $v_x \cdot \hat{i}, v_y \cdot \hat{j}, v_z \cdot \hat{k}$, to rzuty ortogonalne wektora v na osie Ox, Oy, Oz odpowiednio.

Przykład 9.2.15. Rozważmy przestrzeń $\mathbb{R}^3$ ze standardowym iloczynem skalarnym oraz wektor $v = (1, -2, 3)$. Wówczas

$$v \circ \hat{i} = (1,-2,3) \circ (1,0,0) = 1, \quad v \circ \hat{j} = (1,-2,3) \circ (0,1,0) = -2, \quad v \circ \hat{k} = (1,-2,3) \circ (0,0,1) = 3.$$

$(v \circ \hat{i}) \cdot \hat{i} = \quad 1 \cdot \hat{i} = (1,0,0)$ rzut prostokątny wektora v na oś $0x$, tj. na prostą $x = t, y = z = 0,\ t \in \mathbb{R}$

$(v \circ \hat{j}) \cdot \hat{j} = \quad -2 \cdot \hat{j} = (0,-2,0)$ rzut prostokątny wektora v na oś $0y$, tj. na prostą $x = z = 0, y = t,\ t \in \mathbb{R}$

$(v \circ \hat{k}) \cdot \hat{k} = \quad 3 \cdot \hat{k} = (0,0,3)$ rzut prostokątny wektora v na oś $0z$, tj. na prostą $x = y = z = 0,\ z = t,\ t \in \mathbb{R}$

Twierdzenie 9.2.16. Niech $\mathcal{B} = \Big(b_1, b_2, \ldots, b_n\Big)$ będzie bazą ortogonalną przestrzeni euklidesowej $(V, \langle ., . \rangle)$. Niech $v \in V$, $v = [\alpha_1, \alpha_2, \ldots, \alpha_n]_{\mathcal{B}}$. Wówczas

$$\forall i \in \{1, 2, \ldots, n\} \quad \alpha_i = \frac{\langle v, b_i \rangle}{||b_i||^2}.$$

Ponadto

$$\forall u, w \in V \quad \langle u, w \rangle = \frac{\langle u, b_1 \rangle \langle w, b_1 \rangle}{||b_1||^2} + \frac{\langle u, b_2 \rangle \langle w, b_2 \rangle}{||b_2||^2} + \ldots \frac{\langle u, b_n \rangle \langle w, b_n \rangle}{||b_n||^2}.$$

Wniosek 9.2.17. Jeśli $\mathcal{B} = \left(b_1, b_2, \ldots, b_n\right)$ jest bazą ortonormalną przestrzeni euklidesowej $(V, \langle ., . \rangle)$ oraz $V \ni v = [\alpha_1, \alpha_2, \ldots, \alpha_n]_{\mathcal{B}}$, to wówczas

$$\forall i \in \{1, 2, \ldots, n\} \quad \alpha_i = \langle v, b_i \rangle$$

oraz

$$\forall u, w \in V \quad \langle u, w \rangle = \langle u, b_1 \rangle \langle w, b_1 \rangle + \langle u, b_2 \rangle \langle w, b_2 \rangle + \ldots + \langle u, b_n \rangle \langle w, b_n \rangle.$$

Dowód. Wystarczy w twierdzeniu 9.2.16 przyjąć $||b_i|| = 1$. $\square$

Wniosek 9.2.18. Niech $\mathcal{B} = \left(b_1, b_2, \ldots, b_n\right)$ będzie bazą przestrzeni euklidesowej $(V, \langle ., . \rangle)$ oraz niech $v, w \in V$, $v = [\alpha_1, \alpha_2, \ldots, \alpha_n]_{\mathcal{B}}$, $w = [\beta_1, \beta_2, \ldots, \beta_n]_{\mathcal{B}}$. Wówczas baza $\mathcal{B}$ jest ortonormalna wtedy i tylko wtedy gdy $\boxed{\langle u, w \rangle = \alpha_1\beta_1 + \alpha_2\beta_2 \ldots + \alpha_n\beta_n.}$

Dowód. Zauważmy, że $\langle v, w \rangle = \langle \sum_{i=1}^n \alpha_i b_i, \sum_{i=1}^n \beta_j b_j \rangle = \sum_{i,j=1}^n \alpha_i\beta_j \langle b_i, b_j \rangle$ równa się $\sum_i^n \alpha_i\beta_i$ wtedy i tylko wtedy, gdy $\langle b_i, b_j \rangle = \begin{cases} 0 & ; \quad i \neq j \\ 1 & ; \quad i = j \end{cases}$. $\square$

Przykład 9.2.19. Niech $(V, \langle ., . \rangle)$ będzie przestrzenią euklidesową, zaś $\mathcal{B} = (b_1, b_2, b_3, b_4, b_5)$ jej bazą ortonormalną. Obliczymy iloczyn skalarny wektorów $u = 3b_1 + 2b_2 - b_3 + 2b_5$ i $v = 2b_1 + b_3 + 6b_4 + b_5$.

Ponieważ $u = [3, 2, -1, 0, 2]_{\mathcal{B}}$, $v = [2, 0, 1, 6, 1]_{\mathcal{B}}$ oraz baza $\mathcal{B}$ jest ortonormalna, na mocy wniosku 9.2.18 otrzymujemy

$$\langle u, v \rangle = 3 \cdot 2 + 2 \cdot 0 - 1 \cdot 1 + 0 \cdot 6 + 2 \cdot 1 = 6 - 1 + 2 = 7.$$

Dopełnienie ortogonalne

Niech $(V, \langle ., . \rangle)$ będzie przestrzenią euklidesową, zaś $S \subset V$ dowolnym podzbiorem.

Definicja 9.2.20. Zbiór $S^\perp := \{v \in V : \forall x \in S \ \langle v, x \rangle = 0\}$ nazywamy *dopełnieniem ortogonalnym zbioru* S w przestrzeni euklidesowej V. Jeżeli zbiór S jest jednoelementowy tj. $S = \{x\}$, to zbiór $S^\perp$ nazywamy *dopełnieniem ortogonalnym wektora* $x \in V$.

Twierdzenie 9.2.21. Niech V będzie przestrzenią euklidesową.

i) Jeśli $U \subset V$ jest podprzestrzenią liniową przestrzeni V, to wówczas $U^\perp$ również jest podprzestrzenią liniową.

ii) Dla dowolnego wektora $u \in V$ zbiór $\{u\}^\perp$ jest podprzestrzenią liniową przestrzeni V.

Przykład 9.2.22. W przestrzeni $\mathbb{E}^4$ wyznaczymy wszystkie wektory v ortogonalne do $u = (1, 0, 1, 0)$.

Niech $v = (x, y, z, t) \in \mathbb{R}^4$. Wówczas $u \perp v \Leftrightarrow u \circ v = 0 \Leftrightarrow x + z = 0$. Zatem $v = (x, y, -x, t)$ oraz

$$\{u\}^\perp = \{(x, y, -x, t) : \ x, y, t \in \mathbb{R}\} = \text{lin}\{(1, 0, -1, 0), (0, 1, 0, 0), (0, 0, 0, 1)\}.$$

Definicja 9.2.23. Niech U będzie podprzestrzenią liniową przestrzeni V. Jeśli $w \in U^\perp$, to mówimy, że wektor w jest *ortogonalny do podprzestrzeni* U i piszemy $w \perp U$.

Uwaga 9.2.24. Niech U będzie podprzestrzenią liniową przestrzeni V, zaś $\mathcal{B} = \{b_1, \ldots, b_k\}$ bazą tej podprzestrzeni. Wówczas dla dowolnego wektora $w \in V$

$$w \perp U \Leftrightarrow \forall i \in \{1, 2, \ldots, k\} \quad w \perp b_i$$

Dowód. Implikacja z lewa na prawo jest oczywista. Ponadto jeśli dla dowolnego $i \in \{1, 2, \ldots, n\}$ mamy $\langle w, b_i\rangle = 0$, to wówczas dla dowolnego $u \in U, u = [\alpha_1, \ldots, \alpha_n]_\mathcal{B}$ mamy $\langle w, u\rangle = \alpha_1\langle w, b_1\rangle + \ldots + \alpha_n\langle w, b_n\rangle = 0$. Zatem $w \perp U$. $\square$

Przykład 9.2.25. Rozważmy $V = \mathbb{R}_2[x]$ z iloczynem skalarnym $\langle p, q\rangle = \int_0^1 p(x)q(x)dx$, podprzestrzeń $U = \mathbb{R}_1[x]$ oraz $w = 6x^2 - 6x + 1$. Czy $w \perp U$?

$w \perp U = \mathbb{R}_1[x] = \text{lin}\{1, x\} \Leftrightarrow w \perp 1 \wedge w \perp x$

$$\langle w, 1\rangle = \int_0^1 (6x^2 - 6x + 1)dx = 2x^3 - 3x^2 + x\Big|_0^1 = 0$$

$$\langle w, x\rangle = \int_0^1 (6x^3 - 6x^2 + x)dx = \frac{3}{2}x^4 - 2x^3 + \frac{1}{2}x^2\Big|_0^1 = 0$$

Na mocy uwagi 9.2.24 wnioskujemy, że $w \perp U$.

Własności dopełnienia ortogonalnego

Twierdzenie 9.2.26. Niech V będzie przestrzenią euklidesową, zaś U, U_1, U_2 jej dowolnymi podprzestrzeniami liniowymi. Wówczas

i) $U_1 \subset U_2 \Rightarrow U_2^\perp \subset U_1^\perp$,

ii) $U^\perp = (\mathrm{lin}U)^\perp$,

iii) $U \subset (U^\perp)^\perp$.

iv) Jeśli $\dim U < \infty$, to wówczas $U = (U^\perp)^\perp$ oraz $\dim U + \dim U^\perp = \dim V$. [a]

[a] Jeśli U jest przestrzenią nieskończenie wymiarową, to równość na ogół nie zachodzi. Przykładem może być podprzestrzeń U przestrzeni $V = \mathcal{C}([0,1], \mathbb{R})$ z iloczynem skalarnym $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$, postaci $U = \{f \in V : f(0) = 0\}$. Można uzasadnić, że $U^\perp = \{0\}$ i w konsekwencji $U \neq (U^\perp)^\perp = V$.

Przykład 9.2.27. i) Rozważmy $\mathbb{R}^n$ ze standardowym iloczynem skalarnym.
Każdy wektor jest ortogonalny do wektor zerowego (porównaj twierdzenie 9.1.6), zatem

$$\{\mathbf{0}\}^\perp = \{(0, 0, \ldots, 0)\}^\perp = \mathbb{R}^n.$$

Przestrzeń $\{\mathbf{0}\}$ jest skończenie wymiarowa, zatem na mocy twierdzenia 9.2.26

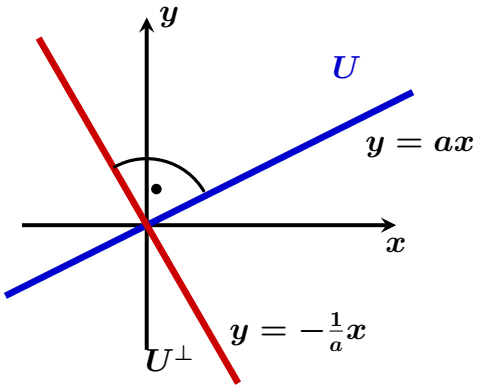
$$\left(\mathbb{R}^n\right)^\perp = \left(\left(\{\mathbf{0}\}\right)^\perp\right)^\perp = \{\mathbf{0}\}.$$

ii) Rozważmy $\mathbb{R}^2$ ze standardowym iloczynem skalarnym oraz jej podprzestrzeń

$$U = \{(x, ax) : x \in \mathbb{R}\} = \mathrm{lin}\{(1, a)\}, \quad \text{gdzie} \quad a \in \mathbb{R} \setminus \{0\}.$$

Podprzestrzeń U to prosta $y = ax$.
Jej dopełnienie ortogonalne $U^\perp$ to prosta przechodząca przez $(0, 0)$ i prostopadła do $y = ax$, czyli prosta o równaniu $y = -\frac{1}{a}x$. A zatem

$$U^\perp = \left\{\left(x, -\frac{1}{a}x\right) : x \in \mathbb{R}\right\} = \mathrm{lin}\left\{\left(1, -\frac{1}{a}\right)\right\}.$$

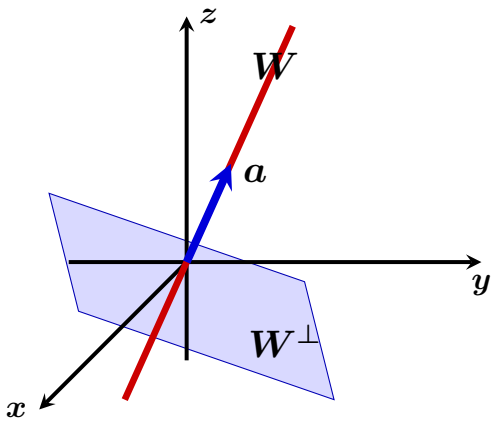


iii) Rozważmy $\mathbb{R}^3$ ze standardowym iloczynem skalarnym oraz jej podprzestrzeń

$$W = \{(t, 2t, 3t) : t \in \mathbb{R}\} = \text{lin}\{(1, 2, 3)\}.$$

Podprzestrzeń W to prosta w przestrzeni, przechodząca przez punkt $(0, 0, 0)$, której wektorem kierunkowym jest $a = (1, 2, 3)$.
Jej dopełnienie ortogonalne $W^\perp$ to płaszczyzna przechodząca przez $(0, 0, 0)$ i prostopadła do W, czyli płaszczyzna przechodząca przez $(0, 0, 0)$, której wektorem normalnym jest $a = (1, 2, 3)$. A zatem

$$\begin{aligned} W^\perp &= \{(x, y, z) : x + 2y + 3z = 0\} = \{(x, y, z) : x = -2y - 3z\} \\ &= \{(-2y - 3z, y, z) : y, z \in \mathbb{R}\} = \text{lin}\{(-2, 1, 0), (-3, 0, 1)\} \end{aligned}$$



Ponieważ $\dim \mathbb{R}^3 < \infty$, to zgodnie z tezą twierdzenia 9.2.26 otrzymujemy

$$\dim W^\perp = \dim \mathbb{R}^3 - \dim W = 3 - 1 = 2.$$

9.3 Metody ortogonalizacji

Twierdzenie 9.3.1. Niech $(V, \langle ., . \rangle)$ będzie przestrzenią euklidesową. Niech $\{b_1, \ldots b_m\} \subset V$ będzie układem wektorów liniowo niezależnych. Wówczas istnieje układ ortogonalny $\{c_1, \ldots c_m\} \subset V$ taki, że $\boxed{\text{lin}\{b_1, \ldots b_m\} = \text{lin}\{c_1, \ldots c_m\}.}$

Wniosek 9.3.2. i) Każda skończenie wymiarowa przestrzeń euklidesowa różna od $\{\mathbf{0}\}$ ma bazę ortogonalną i ortonormalną.

ii) W przestrzeni euklidesowej skończenie wymiarowej każdy układ ortonormalny można uzupełnić do bazy ortonormalnej.

Metoda ortogonalizacji Grama-Schmidta
Jest to sposób efektywnej konstrukcji bazy ortogonalnej.

Wniosek 9.3.3. Jeśli $\mathcal{B} = \Big(b_1, b_2, \ldots, b_n\Big)$ jest dowolną bazą przestrzeni euklidesowej $(V, \langle ., . \rangle)$, to wówczas ciąg wektorów $\mathcal{C} = \Big(c_1, c_2, \ldots, c_n\Big)$, zdefiniowany poniżej, jest bazą ortogonalną tej przestrzeni.

$$c_1 := b_1 \qquad c_2 := b_2 - \frac{\langle b_2, c_1\rangle}{||c_1||^2} c_1 \qquad c_3 := b_3 - \frac{\langle b_3, c_1\rangle}{||c_1||^2} c_1 - \frac{\langle b_3, c_2\rangle}{||c_2||^2} c_2 \qquad \ldots \qquad c_n := b_n - \sum_{i=1}^{n-1} \frac{\langle b_n, c_i\rangle}{||c_i||^2} c_i$$

Przykład 9.3.4. Rozważmy przestrzeń $\mathbb{R}_2[x]$ z iloczynem skalarnym

$$\langle p, q\rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1).$$

Dokonamy ortogonalizacji bazy $\mathcal{B} = \Big(1, x, x^2\Big)$,
a następnie wskażemy bazę ortonormalną przestrzeni euklidesowej $(\mathbb{R}_2[x], \langle ., . \rangle)$.

Niech $b_1 = 1, b_2 = x, b_3 = x^2$.
Baza $\mathcal{B}$ nie jest ortogonalna, bowiem $b_1 \circ b_3 = 1 \cdot (-1)^2 + 1 \cdot 0^2 + 1 \cdot 1^2 = 2 \neq 0$.
Niech $\mathcal{C} = (c_1, c_2, c_3)$ oznacza poszukiwaną bazę ortogonalną.

I KROK: Niech $\boxed{c_1 := b_1} = 1$. Oczywiście $\lin\{b_1\} = \lin\{c_1\}$.

II KROK: Poszukujemy $c_2 \in \mathbb{R}_2[x]$ takiego, że $\lin\{b_1, b_2\} = \lin\{c_1, c_2\}$ oraz $c_2 \perp c_1$.

$c_2 \in \lin\{b_1, b_2\} = \lin\{c_1, b_2\}$, zatem $\boxed{c_2 = b_2 + \alpha c_1}$, dla pewnego $\alpha \in \mathbb{R}$. [11]
Dobierzemy α w taki sposób, by $\boxed{\langle c_2, c_1\rangle = 0}$.
Iloczyn skalarny jest liniowy ze względu na pierwszą zmienną, zatem

$$\mathbf{0} = \langle c_2, c_1\rangle = \langle b_2 + \alpha c_1, c_1\rangle = \boldsymbol{\langle b_2, c_1\rangle + \alpha \cdot \langle c_1, c_1\rangle} = (-1 \cdot 1 + 0 \cdot 1 + 1 \cdot 1) + \alpha(1 + 1 + 1) = 3\alpha.$$

Skąd $\alpha = 0$. [12] Zatem $c_2 = b_2 = x$. (Mogliśmy to zauważyć wcześniej.)

III KROK: Poszukujemy $c_3 \in \mathbb{R}_2[x]$ takiego, że $\lin\{b_1, b_2, b_3\} = \lin\{c_1, c_2, c_3\}$ oraz $c_3 \perp c_1$, $c_3 \perp c_2$.

$c_3 \in \lin\{b_1, b_2, b_3\} = \lin\{c_1, c_2, b_3\}$, zatem $\boxed{c_3 = b_3 + \beta_1 c_1 + \beta_2 c_2}$, dla pewnych $\beta_1, \beta_2 \in \mathbb{R}$.
Dobierzemy β_1, β_2 w taki sposób, by $\boxed{\langle c_3, c_1\rangle = 0,\ \langle c_3, c_1\rangle = 0}$.

[11] Warunek $c_2 \in \lin\{c_1, b_2\}$ oznacza, że c_2 jest kombinacja liniową wektorów c_1 i b_2, czyli $c_2 = \gamma b_2 + \alpha c_1$ dla pewnych $\gamma, \alpha \in \mathbb{R}$. W tym kroku chcemy zapewnić, by kąt między c_1 i c_2 miał miarę $\frac{\pi}{2}$. Nie dbamy na razie o długość wektora c_2. Z tego powodu bez straty dla ogólności rozumowania możemy założyć, że $\gamma = 1$.

[12] Z warunku $\langle b_2, c_1\rangle + \alpha \cdot \langle c_1, c_1\rangle = 0$ wynika, że $\alpha = -\frac{\langle b_2, c_1\rangle}{||c_1||^2}$ oraz $c_2 = b_2 + \alpha c_1 = b_2 - \frac{\langle b_2, c_1\rangle}{||c_1||^2} c_1$, jak we wniosku 9.3.3.

Iloczyn skalarny jest liniowy ze względu na pierwszą zmienną, zatem

$$\mathbf{0} = \langle c_3, c_1\rangle = \langle b_3 + \beta_1 c_1 + \beta_2 c_2, c_1\rangle = \langle b_3, c_1\rangle + \beta_1\langle c_1, c_1\rangle + \beta_2 \underbrace{\langle c_2, c_1\rangle}_{=0} =$$

$$= \mathbf{\langle b_3, c_1\rangle + \beta_1 ||c_1||^2} = (1\cdot 1 + 0\cdot 1 + 1\cdot 1) + \beta_1 \cdot 3,$$

skąd $\beta_1 = -\frac{2}{3}$. [13] Analogicznie

$$\mathbf{0} = \langle c_3, c_2\rangle = \langle b_3 + \beta_1 c_1 + \beta_2 c_2, c_2\rangle = \langle b_3, c_2\rangle + \beta_1 \underbrace{\langle c_1, c_2\rangle}_{=0} + \beta_2\langle c_2, c_2\rangle =$$

$$= \mathbf{\langle b_3, c_2\rangle + \beta_2 ||c_2||^2} = (1\cdot(-1) + 0\cdot 0 + 1\cdot 1) + \beta_2\cdot((-1)^2 + 0^2 + 1^2) = 2\beta_2,$$

skąd $\beta_2 = 0$. [14]

Zatem $c_3 = b_3 - \frac{2}{3}c_1 = x^2 - \frac{2}{3}$ i $\mathcal{C} = \left(c_1 = 1, c_2 = x, c_3 = x^2 - \frac{2}{3}\right)$ jest bazą ortogonalną.

IV KROK: Unormujemy wektory, by wskazać bazę ortonormalną. Obliczamy

$$||c_1|| = \sqrt{\langle c_1, c_1\rangle} = \sqrt{1+1+1} = \sqrt{3},$$
$$||c_2|| = \sqrt{\langle c_2, c_2\rangle} = \sqrt{1+0+1} = \sqrt{2},$$
$$||c_3|| = \sqrt{\langle c_3, c_3\rangle} = \sqrt{\frac{1}{9} + \frac{4}{9} + \frac{1}{9}} = \frac{\sqrt{6}}{3}.$$

Bazą ortonormalną przestrzeni euklidesowej $(\mathbb{R}_2[x], \langle .,.\rangle)$ jest

$$\mathcal{C}' = \left(c_1' = \frac{c_1}{||c_1||} = \frac{\sqrt{3}}{3}, \quad c_2' = \frac{c_2}{||c_2||} = \frac{\sqrt{2}}{2}x, \quad c_3' = \frac{c_3}{||c_3||} = \frac{\sqrt{6}}{2}\left(x^2 - \frac{2}{3}\right) = \frac{\sqrt{6}}{2}x^2 - \frac{\sqrt{6}}{3}\right).$$

Przykład 9.3.5. Rozważmy przestrzeń $\mathbb{R}_2[x]$ z iloczynem skalarnym $\langle p, q\rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$ oraz dwie bazy tej przestrzeni $\mathcal{B} = \left(1, x, x^2\right)$ oraz $\mathcal{C}' = \left(c_1' = \frac{\sqrt{3}}{3}, c_2' = \frac{\sqrt{2}}{2}x, c_3' = \frac{\sqrt{6}}{2}(x^2 - \frac{2}{3}) = \frac{\sqrt{6}}{2}x^2 - \frac{\sqrt{6}}{3}\right)$.

Dany jest wektor $p = \sqrt{6}x^2 + 3\sqrt{2}x$. Obliczymy jego normę.

$$||p||^2 = \langle p, p\rangle = [p(-1)]^2 + [p(0)]^2 + [p(1)]^2 = (\sqrt{6} - 3\sqrt{2})^2 + (\sqrt{6} + 3\sqrt{2})^2 = 12 + 36 = 48.$$

Zatem $||p|| = \sqrt{48} = 4\sqrt{3}$. Zauważmy, że

$$p = [2\sqrt{2}, 6, 2]_{\mathcal{C}'} \quad \text{oraz} \quad p = [0, 3\sqrt{2}, \sqrt{6}]_{\mathcal{B}}.$$

Baza $\mathcal{C}'$ jest bazą ortonormalną, zatem na mocy wniosku 9.2.18 otrzymujemy równość

$$||p||^2 = (2\sqrt{2})^2 + 6^2 + 2^2 = 48.$$

Baza $\mathcal{B}$ nie jest bazą ortonormalną, dlatego też $||p||^2 \neq (3\sqrt{2})^2 + (\sqrt{6})^2 = 24$.

[13] Z warunku $\langle b_3, c_1\rangle + \beta_1||c_1||^2 = 0$ wynika, że $\beta_1 = -\frac{\langle b_3, c_1\rangle}{||c_1||^2}$, jak we wniosku 9.3.3.

[14] Z warunku $\langle b_3, c_2\rangle + \beta_2||c_2||^2 = 0$ wynika, że $\beta_2 = -\frac{\langle b_3, c_2\rangle}{||c_2||^2}$ oraz $c_3 = b_3 + \beta_1 c_1 + \beta_2 c_2 = b_3 - \frac{\langle b_3, c_1\rangle}{||c_1||^2}c_1 - \frac{\langle b_3, c_2\rangle}{||c_2||^2}c_2$, jak we wniosku 9.3.3.

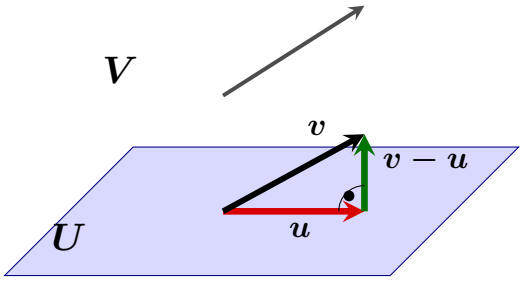
9.4 Rzut ortogonalny na podprzestrzeń

Niech V będzie przestrzenią euklidesową, zaś U jej podprzestrzenią liniową.

Definicja 9.4.1. Operator liniowy $\pi \in End(V)$ dany wzorem

$$\forall v \in V \ \pi(v) = u, \quad \text{gdzie} \quad v - u \perp U,$$

nazywamy *rzutowaniem ortogonalnym* lub *projekcją ortogonalną* na podprzestrzeń U.
Obraz wektora v poprzez π nazywamy *rzutem ortogonalnym* wektora $v \in V$ na podprzestrzeń U.



projekcja ortogonalna na podprzestrzeń U

Twierdzenie 9.4.2 (Jednoznaczność rzutu ortogonalnego)**.** Jeśli $\dim U < \infty$, to wówczas dla dowolnego $v \in V$ istnieje jednoznacznie wyznaczony rzut ortogonalny $u \in U$ tego wektora na podprzestrzeń U.

i) Jeśli $\mathcal{B} = \left(b_1, b_2, \ldots, b_k\right)$ jest dowolną bazą podprzestrzeni U, wówczas $u = \pi(v) = [\alpha_1, \alpha_2, \ldots, \alpha_k]_{\mathcal{B}}$, gdzie

$$\begin{bmatrix} \langle b_1, b_1\rangle & \langle b_1, b_2\rangle & \ldots & \langle b_1, b_k\rangle \\ \langle b_2, b_1\rangle & \langle b_2, b_2\rangle & \ldots & \langle b_2, b_k\rangle \\ \ldots & & \ldots & \\ \langle b_k, b_1\rangle & \langle b_k, b_2\rangle & \ldots & \langle b_k, b_k\rangle \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \begin{bmatrix} \langle v, b_1\rangle \\ \langle v, b_2\rangle \\ \vdots \\ \langle v, b_k\rangle \end{bmatrix}. \tag{1}$$

ii) Jeśli $\mathcal{B} = \left(b_1, b_2, \ldots, b_k\right)$ jest bazą ortogonalną podprzestrzeni U, wówczas

$$u = \pi(v) = \frac{\langle v, b_1\rangle}{||b_1||^2} b_1 + \frac{\langle v, b_2\rangle}{||b_2||^2} b_2 + \ldots + \frac{\langle v, b_k\rangle}{||b_k||^2} b_k.$$

iii) Jeśli $\mathcal{B} = \left(b_1, b_2, \ldots, b_k\right)$ jest bazą ortonormalną podprzestrzeni U, wówczas

$$u = \pi(v) = \langle v, b_1\rangle b_1 + \langle v, b_2\rangle b_2 + \ldots + \langle v, b_k\rangle b_k.$$

Macierz $\begin{bmatrix} \langle b_1, b_1\rangle & \langle b_1, b_2\rangle & \dots & \langle b_1, b_k\rangle \\ \langle b_2, b_1\rangle & \langle b_2, b_2\rangle & \dots & \langle b_2, b_k\rangle \\ \dots & & \dots & \\ \langle b_k, b_1\rangle & \langle b_k, b_2\rangle & \dots & \langle b_k, b_k\rangle \end{bmatrix}$ występującą w powyższym twierdzeniu nazywamy *macierzą Grama* układu wektorów $\big(b_1, b_2, \dots, b_k\big)$.

Wniosek 9.4.3. Niech V będzie przestrzenią euklidesową, zaś U jej podprzestrzenią liniową. Niech $\mathcal{B} = \big(b_1, b_2, \dots, b_k\big)$ będzie bazą ortonormalną podprzestrzeni U, zaś $A \in M_{n\times k}(\mathbb{R})$ macierzą, której kolumnami są kolejne wektory bazy $\mathcal{B}$. Wówczas AA^T jest reprezentacją macierzową projekcji ortogonalnej na podprzestrzeń U.

Przykład 9.4.4. W przestrzeni euklidesowej $\mathbb{E}^4$ wyznaczymy rzut ortogonalny wektora $v = (1, 1, 1, 0)$ na podprzestrzeń $U = \text{lin}\{(2, 1, 1, 2), (1, 1, -3, 0)\}$.

Zauważmy, że $\mathcal{B} := (b_1, b_2)$ jest bazą U, bowiem $r \begin{bmatrix} 2 & 1 & 1 & 2 \\ 1 & 1 & -3 & 0 \end{bmatrix} = 2$.

METODA I
$\boldsymbol{u := \pi_U(v) = ?}$
$\boldsymbol{u \in U}$, zatem istnieją $\alpha, \beta \in \mathbb{R}$ takie, że $\boldsymbol{u = \alpha b_1 + \beta b_2}$.
$\boldsymbol{w := v - u \perp U \Leftrightarrow w \perp b_1 := (2, 1, 1, 2) \wedge w \perp b_2 := (1, 1, -3, 0)}$
Wówczas
$w = (1, 1, 1, 0) - \alpha(2, 1, 1, 2) - \beta(1, 1, -3, 0) = (1 - 2\alpha - \beta, 1 - \alpha - \beta, 1 - \alpha + 3\beta, -2\alpha)$

Otrzymujemy układ dwóch równań
$0 = \langle w, b_1\rangle = (1 - 2\alpha - \beta, 1 - \alpha - \beta, 1 - \alpha + 3\beta, -2\alpha) \circ (2, 1, 1, 2) = 4 - 10\alpha,$
$0 = \langle w, b_2\rangle = (1 - 2\alpha - \beta, 1 - \alpha - \beta, 1 - \alpha + 3\beta, -2\alpha) \circ (1, 1, -3, 0) = -1 - 11\beta.$

Stąd $\alpha = \frac{2}{5}, \beta = -\frac{1}{11}$ oraz

$u = [\alpha, \beta]_{\mathcal{B}} = [\frac{2}{5}, -\frac{1}{11}]_{\mathcal{B}} = \frac{2}{5}b_1 - \frac{1}{11}b_2 = \left(\frac{4}{5}, \frac{2}{5}, \frac{2}{5}, \frac{4}{5}\right) - \left(\frac{1}{11}, \frac{1}{11}, \frac{-3}{11}, 0\right) = \left(\frac{39}{55}, \frac{17}{55}, \frac{37}{55}, \frac{4}{5}\right)$

METODA II
Zauważmy, że baza $\mathcal{B} := (b_1, b_2)$ jest **ortogonalna**.
Istotnie $b_1 \circ b_2 = (2, 1, 1, 2) \circ (1, 1, -3, 0) = 0$.

Na mocy twierdzenia 9.2.16 mamy $\boldsymbol{u = \left[\frac{\langle u, b_1\rangle}{||b_1||^2}, \frac{\langle u, b_2\rangle}{||b_2||^2}\right]}$.

Ale $\boldsymbol{\langle v, b_i\rangle = \langle u, b_i\rangle}$, zatem $\boldsymbol{u = \left[\frac{\langle v, b_1\rangle}{||b_1||^2}, \frac{\langle v, b_2\rangle}{||b_2||^2}\right]}$.

Obliczamy $\langle v, b_1\rangle = (1, 1, 1, 0) \circ (2, 1, 1, 2) = 4$, $\langle v, b_2\rangle = (1, 1, 1, 0) \circ (1, 1, -3, 0) = -1$
oraz $||b_1||^2 = 10$, $||b_2||^2 = 11$. Stąd $u = [\alpha, \beta]_{\mathcal{B}}$.

UWAGA: Gdyby baza $\mathcal{B} := (b_1, b_2)$ nie była ortogonalna, zawsze możemy metodą Grama-Schmidta ją zortogonalizować i w dalszych rachunkach wykorzystać znalezioną bazę ortogonalną $\mathcal{C} := (c_1, c_2)$.

METODA III

Znajdziemy **macierz rzutowania** na U. **Normalizujemy bazę** $\mathcal{B}$.

Niech $\mathcal{B}' = \{b_1', b_2'\}$, gdzie $b_1' = \frac{b_1}{||b_1||} = \frac{1}{\sqrt{10}}(2,1,1,2), b_2' = \frac{b_2}{||b_2||} = \frac{1}{\sqrt{11}}(1,1,-3,0)$.

Niech $A = \begin{bmatrix} \frac{2}{\sqrt{10}} & \frac{1}{\sqrt{11}} \\ \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{11}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{11}} \\ \frac{2}{\sqrt{10}} & 0 \end{bmatrix}$. Macierzą projekcji jest $AA^T = \begin{bmatrix} \frac{27}{55} & \frac{16}{55} & \frac{-4}{55} & \frac{2}{5} \\ \frac{16}{55} & \frac{21}{110} & \frac{-19}{110} & \frac{1}{5} \\ \frac{-4}{55} & \frac{-19}{110} & \frac{101}{110} & \frac{1}{5} \\ \frac{2}{5} & \frac{1}{5} & \frac{1}{5} & \frac{2}{5} \end{bmatrix}$.

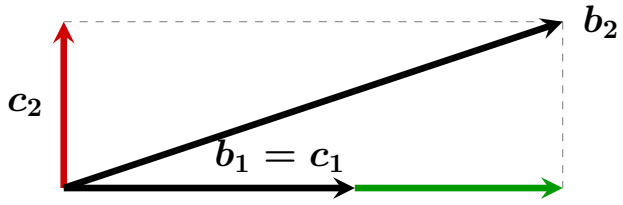
Stąd $AA^T \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{39}{55} \\ \frac{17}{55} \\ \frac{37}{55} \\ \frac{4}{5} \end{bmatrix}$.

Interperatacja geometryczna metody Grama-Schmidta

Rozważmy przestrzeń $\mathbb{E}^3$ i jej bazę $\mathcal{B} = (b_1, b_2, b_3)$. Niech $\mathcal{C} = (c_1, c_2, c_3)$ będzie szukaną bazą ortogonalną.

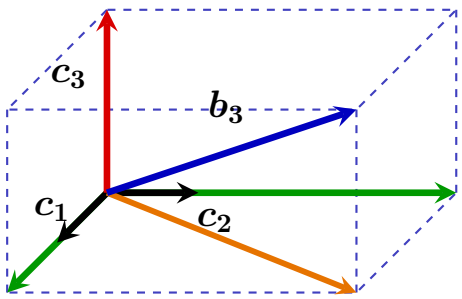
KROK I: $c_1 := b_1$

KROK II:



$$c_2 := b_2 - \frac{\langle b_2, c_1 \rangle}{||c_1||^2} c_1 = b_2 - \pi_{\text{lin}\{c_1\}}(b_1)$$

KROK III:



$$\begin{aligned} c_3 &:= b_3 - \frac{\langle b_3, c_1 \rangle}{||c_1||^2} c_1 - \frac{\langle b_3, c_2 \rangle}{||c_2||^2} c_2 \\ &= b_3 - \pi_{\text{lin}\{c_1\}}(b_3) - \pi_{\text{lin}\{c_2\}}(b_3) \\ &= b_3 - \pi_{\text{lin}\{c_1, c_2\}}(b_3) \end{aligned}$$

9.5 Macierze ortogonalne i izometrie liniowe

Przykład 9.5.1. $\begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \\ 0 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$.
Wektor $(1,0,0)$ ma długość 1, zaś wektor $(1,6,0)$ ma długość $\sqrt{37}$.

Przekształcając dowolne wektory w $\mathbb{R}^n$ za pomocą macierzy przejścia możemy zmienić ich długość oraz kąt między nimi. Wyróżnimy teraz te macierze, które pozostawią powyższe wielkości niezmienione.

Zauważmy, że dla dowolnej macierzy $A \in M_n(\mathbb{R})$, jeśli $A^TA = I_n$, to wówczas $A^TA = AA^T$. Istotnie, z równości $A^TA = I_n$ wynika, że $\det A \neq 0$ oraz $A^T = A^{-1}$. Zatem $I = AA^{-1} = AA^T$.

Definicja 9.5.2. Niech $n \in \mathbb{N}$. Macierz $A \in M_n(\mathbb{R})$ nazywamy *macierzą ortogonalną*, jeśli $A^TA = I_n$ (lub równoważnie $AA^T = I_n$).

Przykład 9.5.3. Macierz $A = \frac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$ jest ortogonalna.

Twierdzenie 9.5.4. Zbiór wszystkich macierzy ortogonalnych stopnia n wraz z działaniem mnożenia macierzy jest grupą. Nazywamy ją *grupą ortogonalną* i oznaczamy symbolem $O(n)$.

Wniosek 9.5.5. i) Wyznacznik macierzy ortogonalnej jest równy 1 lub -1.

ii) Wiersze macierzy ortogonalnej są wzajemnie ortogonalnymi wersorami (jako wektory w $\mathbb{R}^n$ względem standardowego iloczynu skalarnego). Analogiczne stwierdzenie jest prawdziwe dla kolumn.

iii) Niech $(V, \langle .,. \rangle)$ będzie przestrzenią euklidesową. Niech $\mathcal{B}, \mathcal{B}'$ to dwie bazy ortonormalne tej przestrzeni. Wówczas macierz przejścia $P_{\mathcal{B}\to\mathcal{B}'}$ jest macierzą ortogonalną. I odwrotnie, dowolna macierz ortogonalna jest macierzą przejścia między dwiema bazami ortonormalnymi.

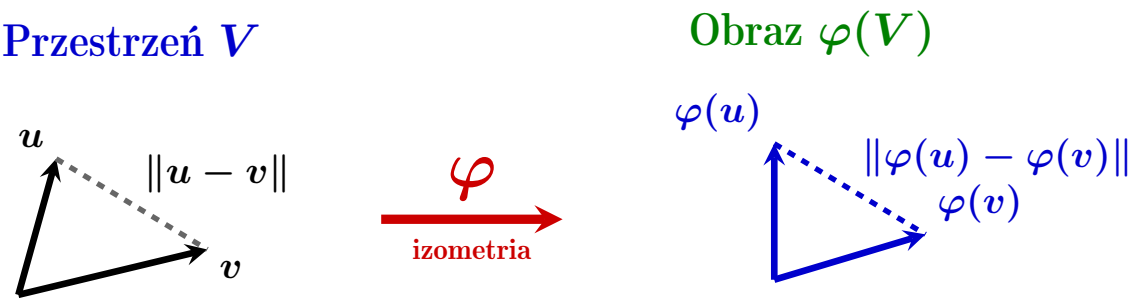
Dowód. i) Wynika to z równości $1 = \det I_n = \det(AA^T) = (\det A)^2$. $\square$

Oczywiście nie każda macierz kwadratowa o wyznaczniku równym 1 lub -1 jest ortogonalna. Przykładami mogą być macierze $A = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$ oraz $B = \begin{bmatrix} 3 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$.
Łatwo sprawdzić, że $\det A = 1$, zaś $AA^T \neq I$. Podobnie $\det B = 1$ i mimo tego, że kolumny B są wektorami ortogonalnymi, to jednak $BB^T = \begin{bmatrix} 9 & 0 \\ 0 & \frac{1}{9} \end{bmatrix} \neq I$.

Niech $(V, \langle ., . \rangle)$ będzie przestrzenią euklidesową. Rozpatrujemy normę zadaną przez iloczyn skalarny.

Definicja 9.5.6. Operator liniowy $\varphi \in End(V)$ nazywamy *izometrią liniową*, jeśli zachowuje on odległość, tzn. spełniony jest warunek

$$\forall u, v \in V \quad ||\varphi(u) - \varphi(v)|| = ||u - v||.$$



Izometria to przekształcenie, które nie zmienia odległości między punktami.

Definicja 9.5.7. Załóżmy, że $\dim V < \infty$. Operator liniowy $\varphi \in End(V)$ nazywamy *ortogonalnym*, jeśli jego macierz w pewnej bazie ortonormalnej przestrzeni V jest macierzą ortogonalną.

Stwierdzenie, że macierz φ w pewnej bazie ortonormalnej przestrzeni V jest macierzą ortogonalną, jest równoważne stwierdzeniu, że macierz φ w dowolnej bazie ortonormalnej przestrzeni V jest macierzą ortogonalną.

Wynika to z faktu, że macierz przejścia między dwiema bazami ortonormalnymi jest macierzą ortogonalną oraz z faktu, że macierze ortogonalne tworzą grupę.
Oznacza to, że odwzorowanie ortogonalne to takie odwzorowanie liniowe, które przeprowadza pewną (każdą) bazę ortonormalną przestrzeni V na bazę ortonormalną.

Uwaga 9.5.8. Niech $\varphi \in End(V)$ będzie izometrią liniową. Wówczas

i) φ jest monomorfizmem, a zatem jest izomorfizmem.

ii) $Spec(\varphi) \subset \{-1, 1\}$

Dowód. i) Jeśli $v \in \text{Ker}\varphi$, to $\varphi(v) = \mathbf{0}_V$, zatem $||\varphi(v)|| = ||v|| = 0$. Z definicji normy, $||v|| = 0$ wtedy i tylko wte $v = \mathbf{0}_V$. Ponadto $\dim V = r(\varphi)$, więc φ jest surjekcją.
ii) Dla dowolnego $v \in V$ mamy $||\varphi(v)|| = ||v|| = 1 \cdot ||v|| = |\pm 1| \cdot ||v||$. $\square$

Twierdzenie 9.5.9. Niech $(V, \langle ., . \rangle)$ będzie skończnie wymiarową[a] przestrzenią euklidesową oraz niech $\varphi \in End(V)$. Następujące warunki są równoważne.

i) φ jest izometrią liniową.

ii) φ *zachowuje normę* tzn. $\forall v \in V \ ||\varphi(v)|| = ||v||$.

iii) φ *zachowuje iloczyn skalarny* tzn. $\forall u, v \in V \ \langle u, v \rangle = \langle \varphi(u), \varphi(v) \rangle$.

iv) φ jest operatorem ortogonalnym

[a]Jeśli V jest *nieskończenie wymiarowa*, wówczas każdy operator ortogonalny jest izometrią liniową, jednak nie każda izometria liniowa musi być operatorem ortogonalnym. Przykładem może być przestrzeń l_2 ciągów zbieżnych z kwadratem, tj. takich rzeczywistych ciągów liczbowych $(a_n)_{n\in\mathbb{N}}$ dla których $\sum_{n=1}^{\infty} a_n^2 < \infty$, wyposażona w iloczyn skalarny $\langle (a_n), (b_n) \rangle = \sum_{n=1}^{\infty} a_n b_n$. Operator liniowy $\varphi(a_1, a_2, a_3, \ldots) = (0, a_1, a_2, \ldots)$ jest izometrią liniową, jednak nie jest ortogonalny, bowiem $\varphi \circ \varphi^* \neq \mathrm{id}$, gdzie φ^* jest operatorem sprzężonym do φ.

Wniosek 9.5.10. Izometrie liniowe *zachowują kąty*, tzn. jeżeli kąt pomiędzy wektorami u i v wynosi α, to kąt pomiędzy wektorami $\varphi(u)$ i $\varphi(v)$ również wynosi α.

Uwaga 9.5.11. Jeśli rozpatrujemy w $\mathbb{R}^n$ niestandardowy iloczyn skalarny, baza kanoniczna nie musi być ortonormalna. Nie zawsze zatem wystarczy badać macierz w bazie kanonicznej.

Przykład 9.5.12. Rozważmy przestrzeń euklidesową $\mathbb{R}^2$ ze standardowym iloczynem skalarnym. Rozważmy $\varphi \in End(\mathbb{R}^2)$ dany wzorem $\varphi(x, y) = (-y, x)$ oraz dwie bazy przestrzeni $\mathbb{R}^2$, bazę kanoniczną $\mathcal{B}_k^2 = \Big(e_1 = (1, 0), e_2 = (0, 1)\Big)$ oraz bazę $\mathcal{C} = \Big(c_1 = (1, 1), c_2 = (0, 1)\Big)$.

Możemy wypisać reprezentacje macierzowe odwzorowania φ.

$$A = M_\varphi(\mathcal{B}_k^2, \mathcal{B}_k^2) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \text{ oraz } B = M_\varphi(\mathcal{C}, \mathcal{C}) = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}.$$

Sprawdzenie warunku $BB^T = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \neq I$ nic nam nie mówi o tym, czy φ jest przekształceniem ortogonalnym, bowiem baza $\mathcal{C}$ nie jest bazą ortonormalną rozważanej przestrzeni euklidesowej.

Baza standardowa jest bazą ortonormalną, gdy rozważamy standardowy iloczyn skalarny.
Zatem z równości $AA^T = I$ wynika, że φ jest przekształceniem ortogonalnym przestrzeni euklidesowej $\mathbb{R}^2$ ze standardowym iloczynem skalarnym.

Jest to rotacja o kąt prosty w kierunku przeciwnym do ruchu wskazówek zegara (a zatem jasne jest, że zachowuje odległości).

Przykład 9.5.13. Rozważmy $\mathbb{R}^2$ ze standardowym iloczynem skalarnym. Czy endomorfizm $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ dany wzorem $\varphi(x, y) = (-x + y, y)$ jest izometrią liniową?

Baza standardowa jest bazą ortonormalną, gdy rozważamy standardowy iloczyn skalarny. Wyznaczmy macierz A endomorfizmu φ w tejże bazie.
Otrzymujemy $A = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$. Sprawdźmy, czy A jest ortogonalna.

$$A^T A = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \neq I$$

Zatem φ nie jest izometrią liniową.

Przykład 9.5.14. Rozważmy przestrzeń euklidesową $(\mathbb{R}^2, g)$ z iloczynem skalarnym

$$g : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \quad g(x, y) = x_1 y_1 - \frac{1}{2} x_1 y_2 - \frac{1}{2} x_2 y_1 + \frac{1}{2} x_2 y_2, \quad \text{dla} \quad x = (x_1, x_2), y = (y_1, y_2)$$

oraz odwzorowanie $\varphi(x_1, x_2) = (-x_1 + x_2, x_2)$.

METODA I
Niech $A = M_\varphi(\mathcal{B}_k^2, \mathcal{B}_k^2) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.
Baza kanoniczna $\mathcal{B}_k^2 = \Big(e_1 = (1, 0), e_2 = (0, 1)\Big)$ nie jest ortonormalna, bowiem $g(e_1, e_2) = -\frac{1}{2} \neq 0$.
Zatem z warunku $AA^T = I$ <u>nie można</u> wnioskować, że φ jest izometrią liniową.

Rozważmy bazę $\mathcal{C} = \Big(c_1 = (1, 0), c_2 = (1, 2)\Big)$ oraz $B = M_\varphi(\mathcal{C}, \mathcal{C}) = \begin{bmatrix} 0 & -2 \\ 1 & 1 \end{bmatrix}$.

Baza $\mathcal{C}$ jest ortonormalna, bowiem
$g(c_1, c_2) = 1 - 1 = 0$, $||c_1||^2 = g(c_1, c_1) = 1$, $||c_2||^2 = g(c_2, c_2) = 1 - 1 - 1 + 2 = 1$.

Z warunku $BB^T = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} \neq I$ <u>można</u> wnioskować, że φ nie jest izometrią liniową w przestrzeni euklidesowej $(\mathbb{R}^2, g)$.

METODA II
Alternatywnie można zauważyć, że $||c_2|| = 1$, zaś $||\varphi(c_2)|| = ||(-2, 1)|| = \sqrt{4 + 1 + 1 + \frac{1}{2}} \neq 1$.
Zatem φ nie zachowuje normy.

Izometrie liniowe przestrzeni $\mathbb{R}^n$ ze standardowym iloczynem skalarnym

Jeśli $v = (v_x, v_y, v_z)$ jest wektorem w $\mathbb{R}^3$, to symbolem v^T oznaczamy odpowiadający mu wektor kolumnowy $\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$.

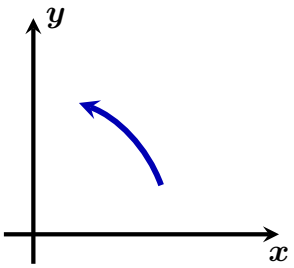
n = 1

$\varphi : \mathbb{R} \to \mathbb{R}$ jest linowe, zatem $\varphi(x) = ax$, $a \in \mathbb{R}$
φ jest izometrią, zatem $|x| = |ax| = |a| \cdot |x|$ oraz $a = \pm 1$.
Otrzymujemy dwie izometrie liniowe $\varphi_1 = \mathrm{id}_{\mathbb{R}}$, czyli $\varphi_1(x) = x$ oraz φ_2 takie, że $\varphi_2(x) = -x$.

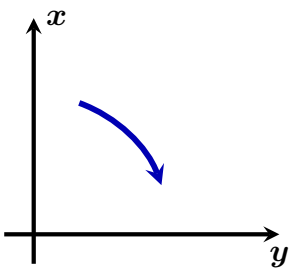
n = 2

Orientacja płaszczyzny

Jeśli obrót osi Ox dookoła punktu O o kąt $\frac{\pi}{2}$ doprowadzający do pokrycia się jej z osią Oy odbywa się w kierunku przeciwnym do ruchu wskazówek zegara, to mówimy, że układ jest *prawoskrętny (zorientowany dodatnio)*, a jeśli zgodnie z ruchem wskazówek zegara, to *lewoskrętny*. O płaszczyźnie, na której ustalono kierunek dodatni obrotu, mówimy, że została *zorientowana*.



układ prawoskrętny



układ lewoskrętny

Poniżej rozważamy płaszczyznę zorientowaną dodatnio.

I. PODEJŚCIE ALGEBRAICZNE

Z założenia $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ jest linowe, zatem możemy rozważyć jego macierz w bazie standardowej $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \in M_2(\mathbb{R})$. Wiemy, że macierz A jest ortogonalna oraz $\det A = \pm 1$.

Z równości $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A^T A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{bmatrix}$

otrzymujemy układ równań $\begin{cases} a^2+b^2 = ||[a,b]||^2 = 1 \\ c^2+d^2 = ||[c,d]||^2 = 1 \\ ac+bd = [a,b] \circ [c,d] = 0 \end{cases}$.

Punkty $(a,b),(c,d)$ leżą na okręgu jednostkowym, zatem istnieje taki kąt α, że $\boldsymbol{a = \cos\alpha}$ oraz $\boldsymbol{b = \sin\alpha}$.

Wówczas na mocy warunku ortogonalności $ac + bd = 0$ mamy, że

$$c = \cos(\frac{\pi}{2} + \alpha) = -\sin\alpha, \qquad d = \sin(\frac{\pi}{2} + \alpha) = \cos\alpha$$

lub też

$$c = \cos(\frac{3}{2}\pi + \alpha) = \sin\alpha, \qquad d = \sin(\frac{3}{2}\pi + \alpha) = -\cos\alpha.$$

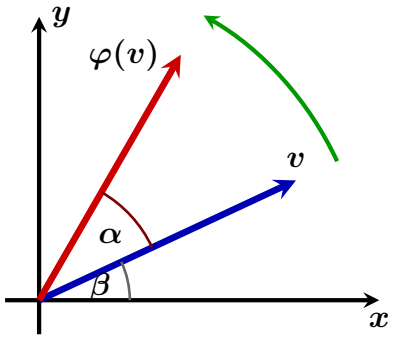
Podsumowując,

$$A = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad lub \quad A = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix} = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}.$$

I. 1. W pierwszym przypadku $\boldsymbol{\det A = a^2 + b^2 = 1}$.
Niech $v = (x,y) = (r\cos\beta, r\sin\beta)$. Obliczamy $\varphi(v)$.

$$A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} r\cos\beta \\ r\sin\beta \end{bmatrix} = r \begin{bmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta \end{bmatrix} = r \begin{bmatrix} \cos(\alpha+\beta) \\ \sin(\alpha+\beta) \end{bmatrix}$$

Mamy do czynienia z *rotacją (obrotem) o kąt* α wokół punktu $(0,0)$ w kierunku przeciwnym do ruchu wskazówek zegara (jeśli kąt α jest dodatni). Odwzorowanie oznaczamy symbolem $\boldsymbol{R_\alpha}$ lub $\boldsymbol{R(\alpha)}$.



Macierzą obrotu w kierunku zgodnym z ruchem wskazówek zegara (obrotu o kąt $-\alpha$) będzie

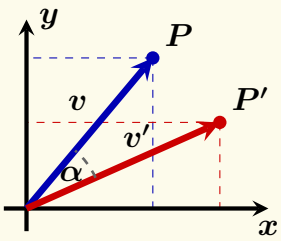
$$A^{-1} = A^T = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}.$$

Zauważmy jeszcze, że $\det(A - tI) = (a-t)^2 + b^2$ i $Spec(A) = \varnothing$.
Macierz A nie jest diagonalizowalna.
Ponadto *rotacja nie zmienia orientacji* układu współrzędnych.

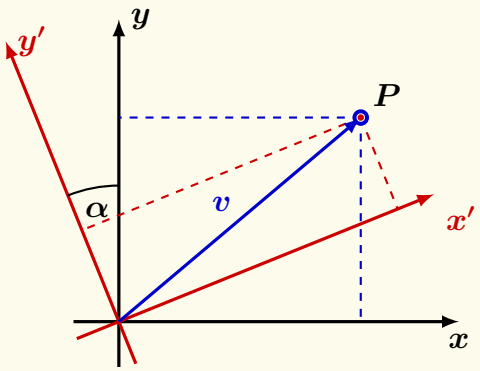
Uwaga 9.5.15. W naszych rozważaniach przyjęliśmy prawoskrętny układ współrzędnych i obracaliśmy wektor (przeciwnie do ruchu wskazówek zegara).

Gdybyśmy zmienili układ na lewoskrętny, ruch odbywałby się w przeciwnym kierunku.

Alternatywne podejście używa obrotu osi układu współrzędnych.
Wówczas wyliczona macierz obrotu reprezentuje obrót osi o ten sam kąt ale w przeciwnym kierunku (zgodnie ze wskazówkami zegara).



Punkt (wektor) zmienia swoje położenie. Zmieniają się jego współrzędne w układzie współrzędnych.

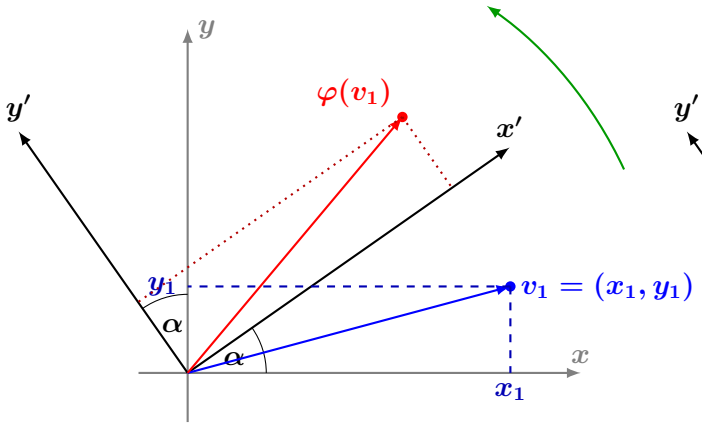


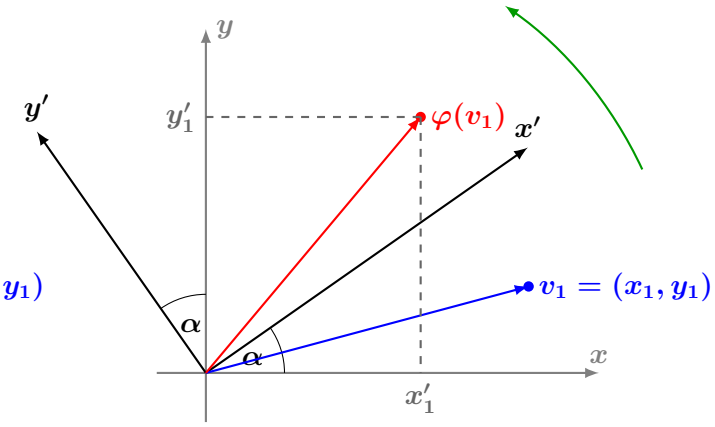
Punkt (wektor) są nieruchome. Obracamy układ współrzędnych, tzn. dokonujemy zmiany bazy. W nowej bazie punkt ma nowe współrzędne.

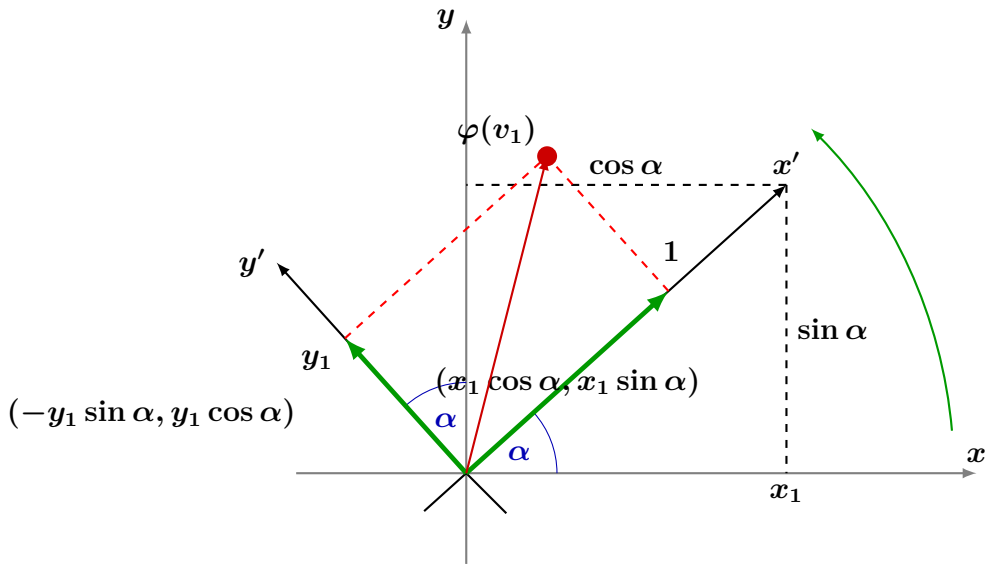
Współrzędne punktu P' w "starej bazie" (rysunek po lewej) są takie same, jak współrzędne punktu P w "nowej bazie" (rysunek po prawej).

PODEJŚCIE GEOMETRYCZNE

Obrotem układu współrzędnych o kąt α nazywamy obrót bez zmiany początku układu oraz bez zmiany jednostek miary wzdłuż wszystkich osi.







$$\varphi(v_1) = (x_1 \cos\alpha - y_1 \sin\alpha, x_1 \sin\alpha + y_1 \cos\alpha)$$

Przykład 9.5.16.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{identyczność, tj. rotacja o kąt miary } 0$$

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix} \quad \text{rotacja o kąt miary } \tfrac{\pi}{2}$$

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix} \quad \text{rotacja o kąt miary } \pi$$

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ -x \end{bmatrix} \quad \text{rotacja o kąt miary } \tfrac{3}{2}\pi$$

Przykład 9.5.17. Dana jest macierz $A = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix} \in M_2(\mathbb{R})$.
Sprawdzając warunek $AA^T = I$, wiemy, że mamy do czynienia z izometrią liniową.

Obliczając wyznacznik $\det A = 1$, wnioskujemy, że macierz ta reprezentuje rotację.
Znajdziemy kąt obrotu.

Ponieważ $\cos\alpha = -\frac{1}{2}$, zatem α ma miarę $\frac{2}{3}\pi$ lub $\frac{4}{3}\pi$.

Ponadto $\sin\frac{2}{3}\pi = \frac{\sqrt{3}}{2}$, zaś $\sin\frac{4}{3}\pi = -\frac{\sqrt{3}}{2}$.

Zatem jest to obrót w kierunku zgodnym z ruchem wskazówek zegara o kąt $\frac{4}{3}\pi$ lub (równoważnie) obrót w kierunku przeciwnym do ruchu wskazówek zegara o kąt $\frac{2}{3}\pi$.

Przykład 9.5.18. Wyznaczymy równanie prostej l' powstałej przez obrót prostej l o równaniu $y = mx$, $m \in \mathbb{R}$ o kąt $\frac{\pi}{4}$.

Obliczamy $x' = x\cos\frac{\pi}{4} - y\sin\frac{\pi}{4} = \frac{\sqrt{2}}{2}(x-y)$, $y' = x\sin\frac{\pi}{4} + y\cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}(x+y)$.

Dodając i odejmując równania stronami, obliczamy $x = \frac{\sqrt{2}}{2}(x'+y')$, $y = \frac{\sqrt{2}}{2}(-x'+y')$.

Wstawiając do równania $y = mx$, otrzymujemy $-x'+y' = m(x'+y')$.

Zatem $(1-m)y' = (1+m)x'$. Równaniem prostej obróconej jest $(m+1)x + (m-1)y = 0$.

I. 2. W drugim przypadku $\det A = -1$. Obliczamy

$$\det(A - tI) = \begin{vmatrix} a-t & b \\ b & -a-t \end{vmatrix} = -(a-t)(a+t) - b^2 = t^2 - a^2 - b^2 = t^2 - 1.$$

Zatem $Spec(A) = \{-1, 1\}$ i macierz A jest diagonalizowalna.
W bazie wektorów własnych macierz operatora φ jest postaci $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.
Niech $v = (x, y)$. Obliczamy współrzędne $\varphi(v)$ w bazie wektorów własnych.

$$D\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}.$$

Mamy do czynienia z *symetrią ortogonalną (odbiciem) względem prostej* $y = 0$.
$A = P^{-1}DP$, gdzie P to macierz przejścia do bazy wektorów własnych.

Zmiana bazy oznacza zmianę układu współrzędnych z zachowaniem początku układu współrzędnych (bo przekształcenie jest liniowe).

Zatem $A\begin{bmatrix} x \\ y \end{bmatrix}$ jest symetrią ortogonalną względem pewnej prostej przechodzącej przez punkt $(0, 0)$.

Jaka to prosta?

PODEJŚCIE ALGEBRAICZNE
Przypuśćmy, że l jest prostą o równaniu $y = \mathrm{tg}\beta \cdot x$.

Aby znaleźć odbicie wektora v względem l posłużymy się inną bazą ortonormalną.
Pierwszy wektor bazowy b_1 będzie wersorem leżącym na prostej l, a drugi b_2 będzie wersorem do niego prostopadłym.

Zatem $b_1 = (\cos\beta, \sin\beta), b_2 = (-\sin\beta, \cos\beta)$. Istnieją skalary $c_1, c_2 \in \mathbb{R}$ takie, że $v = c_1b_1 + c_2b_2$. Obliczamy

$$v^T = c_1\begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix} + c_2\begin{bmatrix} -\sin\beta \\ \cos\beta \end{bmatrix} = \begin{bmatrix} c_1\cos\beta - c_2\sin\beta \\ c_1\sin\beta + c_2\cos\beta \end{bmatrix} = \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

Podobnie obliczamy $\varphi(v) = c_1 b_1 - c_2 b_2$.

$$\varphi(v)^T = c_1 \begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix} - c_2 \begin{bmatrix} -\sin\beta \\ \cos\beta \end{bmatrix} = \begin{bmatrix} c_1\cos\beta + c_2\sin\beta \\ c_1\sin\beta - c_2\cos\beta \end{bmatrix} = \begin{bmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

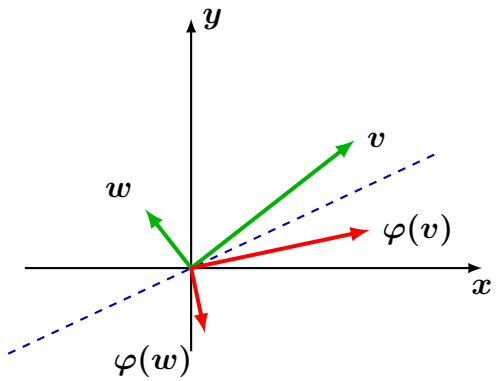
Stąd

$$\varphi(v)^T = \begin{bmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{bmatrix} \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}^{-1} v^T =$$

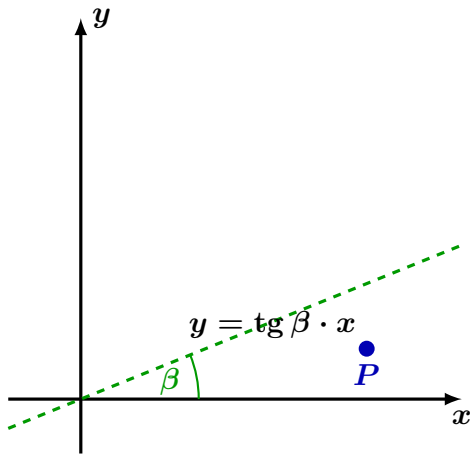
$$= \begin{bmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} v^T = \begin{bmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{bmatrix} v^T$$

Czyli $\alpha = 2\beta$ i macierz $\begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$ reprezentuje odbicie względem prostej o równaniu

$$\boldsymbol{y = \operatorname{tg}\frac{\alpha}{2} \cdot x}.$$

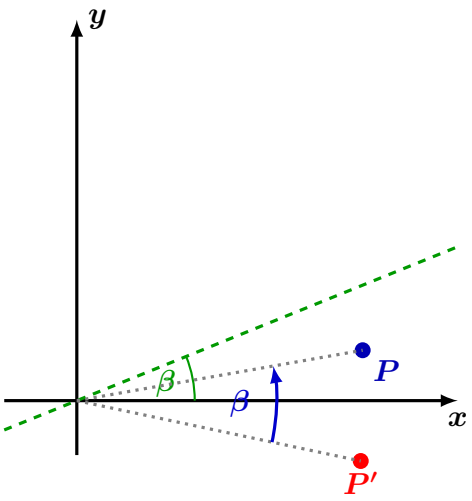


PODEJŚCIE GEOMETRYCZNE



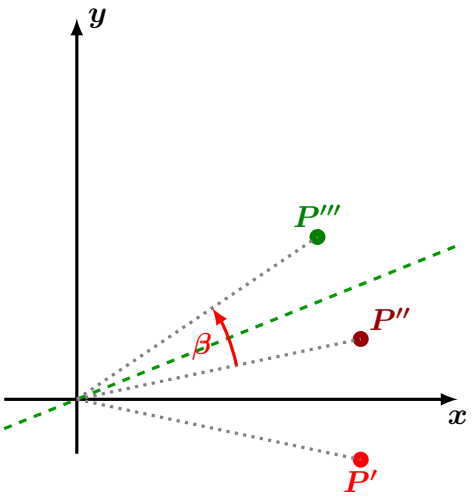
1) Obracamy punkt o kąt $-\beta$.

$$\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}^{-1} = \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix}$$



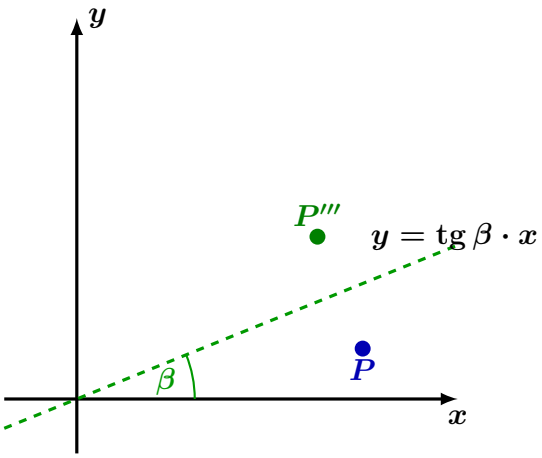
2) Dokonujemy odbicia względem prostej $y = 0$.

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$



3) Obracamy punkt o kąt β.

$$\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$$



4) Otrzymujemy

$$\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} =$$

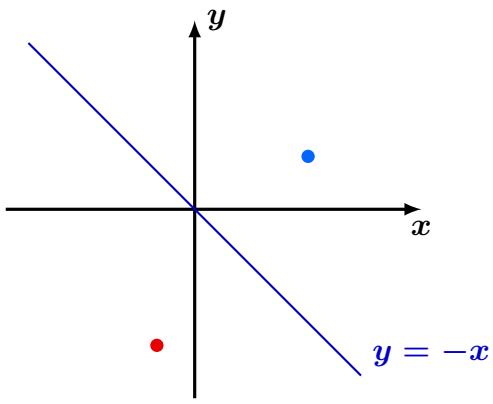
$$\begin{bmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} =$$

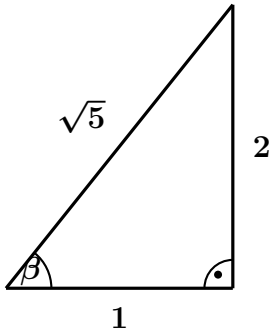
$$\begin{bmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{bmatrix}.$$

Wniosek 9.5.19. Izometria liniowa płaszczyzny jest rotacją (obrotem) wokół punktu $(0,0)$ o pewien kąt lub symetrią ortogonalną (odbiciem) względem pewnej prostej przechodzącej przez początek układu współrzędnych.

Przykład 9.5.20. Macierzą symetrii ortogonalnej względem prostej $y=-x$ jest macierz $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$.

Otrzymujemy $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ -x \end{bmatrix}$.





Przykład 9.5.21. Wyznaczymy macierz izometrii liniowej płaszczyzny będącej symetrią ortogonalną względem prostej $y=2x$.

METODA I: Jeśli $\operatorname{tg}\beta = 2$, to wówczas $\sin\beta = \frac{2}{\sqrt{5}}$, $\cos\beta = \frac{1}{\sqrt{5}}$.

Obliczamy $\sin 2\beta = 2\sin\beta\cos\beta = \frac{4}{5}$ oraz $\cos 2\beta = \cos^2\beta - \sin^2\beta = -\frac{3}{5}$. Otrzymujemy

$$\begin{bmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{bmatrix} = \begin{bmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}.$$

METODA II: Niech F oznacza poszukiwaną macierz, zaś F_0 macierz odbicia względem prostej $y=0$.

Wówczas $F = R(\beta)F_0R(-\beta)$. Ponieważ $\sin\beta = \frac{2}{\sqrt{5}}$, $\cos\beta = \frac{1}{\sqrt{5}}$, zatem

$$F = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}.$$

Uwaga 9.5.22. Składaniu odwzorowań odpowiada mnożenie reprezentujących ich macierzy w pewnej ustalonej bazie. Geometrycznie oznacza to, że układ współrzędnych jest nieruchomy, a obrotowi/odbiciu podlegają wektory.

Przykład 9.5.23. Dana jest macierz $A = \begin{bmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix} \in M_2(\mathbb{R})$.
Sprawdzając warunek $AA^T = I$, wiemy, że A reprezentuje izometrię liniową.

Obliczając wyznacznik, $\det A = -1$, wnioskujemy, że macierz ta reprezentuje odbicie.
Znajdziemy oś symetrii.

Punkty na osi symetrii nie zmieniają swego położenia w wyniku obicia.
Są to punkty (x, y) spełniające równanie $\begin{bmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$.
Otrzymujemy układ równań $\begin{cases} 3x + 4y &= -5x \\ -4x + 3y &= 5y \end{cases}$.

Zatem prosta $y = -2x$ jest osią symetrii.

Uwaga 9.5.24. Symetria ortogonalna φ jest *inwolucją*, tzn. spełnia warunek $\varphi \circ \varphi = \mathrm{id}$. Równoważnie, jeśli A jest macierzą odbicia φ, to wówczas $A^2 = I$ lub inaczej $A^{-1} = A$.

Q U I Z

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. Funkcja $s : M_2(\mathbb{R}) \times M_2(\mathbb{R}) \to \mathbb{R}$, $s(A, B) = \mathrm{tr}(A + B)$ jest iloczynem skalarnym w przestrzeni $M_2(\mathbb{R})$.

2. W przestrzeni euklidesowej dowolny układ ortonormalny jest liniowo niezależny.

3. Niech $(V, \langle ., . \rangle)$ będzie przestrzenią euklidesową, zaś $||.||$ normą w V zadaną przez iloczyn skalarny. Niech $u \in V$ będzie wersorem, zaś $v \in V$ wektorem o długości $\sqrt{2}$. Jeśli $||u + v||^2 = 4$, to wówczas $\langle u, v \rangle = \frac{1}{2}$.

4. W przestrzeni euklidesowej wektor zerowy jest ortogonalny do każdego wektora.

5. Niech $(V, \langle ., . \rangle)$ będzie przestrzenią euklidesową. Jeśli dla danych wektorów $x, y, z \in V$ zachodzi równość $\langle x, y \rangle = \langle x, z \rangle$, to wówczas wektory x oraz $y - z$ są ortogonalne.

6. Dana jest przestrzeń euklidesowa $(\mathbb{R}^n, \langle ., . \rangle)$, jej podprzestrzeń liniowa $U \subset \mathbb{R}^n$ oraz wektor $v \in \mathbb{R}^n$. Niech u będzie rzutem ortogonalnym wektora v na podprzestrzeń U. Wówczas $\langle u, v \rangle = 0$.

7. W przestrzeni $\mathbb{R}^3$ rozpatrujemy standardowy iloczyn skalarny. Niech $f : \mathbb{R}^3 \to \mathbb{R}^3$ będzie odwzorowaniem rzutowania ortogonalnego na podprzestrzeń $U = \{(x, y, z) : x - y + 2z = 0\}$. Wówczas $\dim Imf = 2$.

8. Zbiór macierzy ortogonalnych stopnia n jest podprzestrzenią liniową przestrzeni $M_n(\mathbb{R})$.

9. Dane są macierze $A \in M_5(\mathbb{R})$ oraz $B \in M_{5\times 1}(\mathbb{R})$. Jeśli A jest ortogonalna, to układ równań $AX = B$ jest oznaczony.

10. Jeśli kolumny macierzy A są ortogonalne, to A^TA jest macierzą diagonalną.

11. Macierz rotacji na płaszczyźnie o kąt $\frac{\pi}{3}$ w kierunku przeciwnym do ruchu wskazówek zegara jest odwracalna.

12. W $\mathbb{R}^2$ rozpatrujemy standardowy iloczyn skalarny. Dane są $f, g \in End(\mathbb{R}^2)$ takie, że f jest symetrią ortogonalną względem prostej o równaniu $3x = 4y$, zaś g jest rotacją wokół punktu $(0, 0)$ o kąt prosty w kierunku przeciwnym do ruchu wskazówek zegara. Niech $\varphi = g \circ f$. Wówczas $M_\varphi(\mathcal{B}_k^2, \mathcal{B}_k^2) = \frac{1}{25}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 7 & 24 \\ 24 & -7 \end{bmatrix}$.

13. Odwzorowanie $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ to symetria ortogonalna względem prostej $y = -3x$. Wówczas $\text{lin}\{(1, -3)\}$ jest podprzestrzenią własną φ.

Elżbieta Adamus
Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza w Krakowie

Algebra
Wykład nr 10
Data: .....

TEMAT: *Przestrzenie unitarne*

10.1 Definicja przestrzeni unitarnej i podstawowe własności

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$. Niech $V = (V, +, K, \cdot)$ będzie przestrzenią liniową. Dla dowolnego $\alpha \in \mathbb{C}$, jeśi $z = x + iy$, $x, y, \in \mathbb{R}$, to wówczas $\overline{z} = x - iy$.

Definicja 10.1.1. Funkcję $s : V \times V \to \mathbb{C}$ nazywamy *iloczynem skalarnym (iloczynem hermitowskim)*, jeżeli spełnia ona następujące warunki:

i) $\forall u, v, w \in V \quad \forall \alpha, \beta \in \mathbb{C} \quad s(\alpha u + \beta v, w) = \alpha s(u, w) + \beta s(v, w)$,

ii) $\forall u, v \in V \quad s(u, v) = \overline{s(v, u)}$,

ii) $\forall v \in V \quad s(v, v) \geq 0 \quad \wedge \quad s(v, v) = 0 \Leftrightarrow v = \mathbf{0}_V$.

Parę (V, s) nazywamy wówczas *przestrzenią unitarną*. Zamiast $s(u, v)$ będziemy również pisać $\langle u, v \rangle$.

Uwaga 10.1.2. Każda przestrzeń euklidesowa jest przestrzenią unitarną.

Twierdzenie 10.1.3. Niech (V, s) będzie przestrzenią unitarną. Wówczas

i) $\forall u, v, w \in V \quad s(u, v + w) = s(u, v) + s(u, w)$,

ii) $\forall u, v \in V \quad \forall \alpha \in \mathbb{C} \quad s(u, \alpha v) = \overline{\alpha} s(u, v)$,

ii) $\forall v \in V \quad s(v, \mathbf{0}_V) = 0 = s(\mathbf{0}_V, v)$,

Przykład 10.1.4. Poniższe funkcje są iloczynami skalarnymi.

1) *Standardowym iloczynem skalarnym w* $\mathbb{C}^n$ nazywamy $s : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ taką, że

$$\forall u = (x_1, \ldots, x_n), v = (y_1, \ldots, y_n) \in \mathbb{C}^n \quad s(u, v) = \sum_{i=1}^{n} x_i \overline{y_i} = x_1\overline{y_1} + \ldots + x_n\overline{y_n}.$$

Przykład liczbowy:
W przestrzeni $\mathbb{C}^2$ rozważmy dwa wektory $u = (1 + i, 2 - 3i)$ oraz $v = (2i, 1 + 2i)$. Obliczamy ich iloczyn hermitowski

$$s(u, v) = x_1\overline{y_1} + x_2\overline{y_2} = (1 + i)\overline{(2i)} + (2 - 3i)\overline{(1 + 2i)}$$

Ponieważ $\overline{2i} = -2i$ oraz $\overline{1 + 2i} = 1 - 2i$, otrzymujemy

$$s(u, v) = (1 + i)(-2i) + (2 - 3i)(1 - 2i) = (-2i - 2i^2) + (2 - 4i - 3i + 6i^2)$$

Wiedząc, że $i^2 = -1$

$$s(u, v) = (-2i + 2) + (2 - 7i - 6) = (2 + 2 - 6) + (-2i - 7i) = -2 - 9i$$

2) Niech l^2 będzie zbiorem ciągów $(z_i)_{i\in\mathbb{N}}$ liczb zespolonych takich, że ciąg $a_n = \sum_{i=1}^{n} z_i^2$ jest zbieżny. Granicę ciągu $(a_n)_{n\in\mathbb{N}}$ oznaczamy symbolem $\sum_{i=1}^{\infty} z_i^2$.

l^2 jest przestrzenią wektorową ze zwykłymi działaniami dodawania ciągów po współrzędnych i mnożenia przez liczbę. Definiujemy iloczyn hermitowski

$$\langle (z_i)_{i\in\mathbb{N}}, (w_i)_{i\in\mathbb{N}} \rangle = \sum_{n=1}^{\infty} z_i\overline{w_i}.$$

Przykład liczbowy:
Weźmy dwa ciągi geometryczne zbieżne w l^2 postaci $(z_n)_{n\in\mathbb{N}} = \left(\frac{i^n}{2^n}\right)$ oraz $(w_n)_{n\in\mathbb{N}} = \left(\frac{1}{3^n}\right)$.
Obliczamy ich iloczyn hermitowski.

$$\langle (z_n), (w_n) \rangle = \sum_{n=1}^{\infty} z_n\overline{w_n} = \sum_{n=1}^{\infty} \frac{i^n}{2^n} \cdot \frac{1}{3^n} = \sum_{n=1}^{\infty} \left(\frac{i}{6}\right)^n$$

Korzystając ze wzoru na sumę szeregu geometrycznego $\sum_{n=1}^{\infty} q^n = \frac{q}{1-q}$ dla $q = \frac{i}{6}$, mamy

$$\langle (z_n), (w_n) \rangle = \frac{\frac{i}{6}}{1 - \frac{i}{6}} = \frac{i}{6 - i} = \frac{i(6 + i)}{(6 - i)(6 + i)} = \frac{6i + i^2}{36 - i^2} = \frac{-1 + 6i}{37} = -\frac{1}{37} + \frac{6}{37}i$$

3) Niech $\mathcal{C}([a, b], \mathbb{C}) = \{f : [a, b] \to \mathbb{C} \mid f\text{−ciągła}\}$.
Jest to przestrzeń wektorowa nad ciałem $\mathbb{C}$, tj. *zespolona* przestrzeń wektorowa.
Jej elementami są funkcje postaci $f(t) = u(t) + iv(t)$, gdzie $u, v \in \mathcal{C}([a, b], \mathbb{R})$.
Na przykład $f(t) = t + it^3$, $t \in [-1, 1]$. Definiujemy iloczyn hermitowski

$$\langle f, g \rangle = \int_a^b f(t)\overline{g(t)}dt.$$

Przykład liczbowy:
Rozważmy funkcje $f(t) = t+it^2$ oraz $g(t) = i\cdot t$ na przedziale $[0,1]$. Obliczamy ich iloczyn hermitowski

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,dt = \int_0^1 (t+it^2)\overline{(it)}\,dt$$

Ponieważ $\overline{it} = -it$ (dla rzeczywistego t), funkcja podcałkowa przyjmuje postać

$$(t+it^2)(-it) = -it^2 - i^2t^3 = t^3 - it^2$$

Obliczamy całkę oznaczoną.

$$\langle f, g\rangle = \int_0^1 (t^3 - it^2)\,dt = \left[\frac{t^4}{4} - i\frac{t^3}{3}\right]_0^1 = \frac{1}{4} - \frac{1}{3}i$$

4) Zdefiniujemy iloczyn skalarny w przestrzeni macierzy $M_n(\mathbb{C})$.
Niech $A, B \in M_n(\mathbb{C})$, $A = [a_{ij}], B = [b_{ij}]$.

$$\langle A, B\rangle = \sum_{i,j=1}^{n} a_{ij}\overline{b_{ij}} = \mathrm{tr}(A^T\,\overline{B})$$

Przykład liczbowy:
Dane są dwie macierze w przestrzeni $M_2(C)$ postaci $A = \begin{bmatrix} i & 1 \\ 0 & 2 \end{bmatrix}$, $\quad B = \begin{bmatrix} 1 & -i \\ 2 & i \end{bmatrix}$.
Wówczas

$$\overline{B} = \begin{bmatrix} 1 & i \\ 2 & -i \end{bmatrix} \implies B^* = \overline{B}^T = \begin{bmatrix} 1 & 2 \\ i & -i \end{bmatrix}.$$

Iloczyn skalarny macierzy obliczamy jako ślad macierzy AB^*.

$$AB^* = \begin{bmatrix} i & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ i & -i \end{bmatrix} = \begin{bmatrix} i\cdot 1 + 1\cdot i & i\cdot 2 + 1\cdot(-i) \\ 0\cdot 1 + 2\cdot i & 0\cdot 2 + 2\cdot(-i) \end{bmatrix} = \begin{bmatrix} 2i & i \\ 2i & -2i \end{bmatrix}$$

Z definicji śladu (suma elementów na głównej przekątnej) mamy

$$\langle A, B\rangle = \mathrm{tr}(AB^*) = 2i + (-2i) = 0.$$

Definicja 10.1.5. Macierzą sprzężoną do macierzy $A = [a_{ij}] \in M_{m\times n}(\mathbb{C})$ nazywamy macierz $B = [b_{ij}] \in M_{m\times n}(\mathbb{C})$, której każdy element jest liczbą sprzężoną do odpowiadającego mu elementu macierzy A, tj. $b_{ij} = \overline{a_{ij}}$. Oznaczamy ją symbolem $\overline{A}$.

Przykład 10.1.6. Niech $A = \begin{bmatrix} 1+i & 2 & 0 \\ -3 & 7-5i & i \\ -1-i & 1+i & 1 \end{bmatrix}$. Wówczas $\overline{A} = \begin{bmatrix} 1-i & 2 & 0 \\ -3 & 7+5i & -i \\ -1+i & 1-i & 1 \end{bmatrix}$.

Twierdzenie 10.1.7 (Własności maceirzy sprzężonej)**.**

i) $A, B \in M_{m\times n}(\mathbb{C}) \implies \overline{A+B} = \overline{A} + \overline{B}$

ii) $A \in M_{m\times n}(\mathbb{C}), a \in \mathbb{C} \implies \overline{a \cdot A} = \overline{a} \cdot \overline{A}$

iii) $A \in M_{m\times n}(\mathbb{C}) \wedge B \in M_{n\times p}(\mathbb{C}) \implies \overline{A \cdot B} = \overline{A} \cdot \overline{B}$

iv) $\overline{A}^{-1} = \overline{A^{-1}}$ dla macierzy odwracalnej A

v) $A \in M_{m\times n}(\mathbb{R}) \implies \overline{A} = A$

vi) $\det\overline{A} = \overline{\det A}$ dla $A \in M_n(\mathbb{C})$

Definicja 10.1.8. *Sprzężeniem hermitowskim* macierzy $A = [a_{ij}] \in M_{m\times n}(\mathbb{C})$ nazywamy macierz $B = [b_{ij}] \in M_{n\times m}(\mathbb{C})$ taką, że $B = \overline{A}^T$, $b_{ij} = \overline{a_{ji}}$. Oznaczamy ją symbolem A^*. Stosuje się również oznaczenia A^H, $A^\dagger$.

Przykład 10.1.9. Niech $A = \begin{bmatrix} 1+i & 2 & 0 \\ -3 & 7-5i & i \end{bmatrix}$. Wówczas $A^* = \begin{bmatrix} 1-i & -3 \\ 2 & 7+5i \\ 0 & -i \end{bmatrix}$.

Twierdzenie 10.1.10 (Własności sprzężenia hermitowskiego)**.** Niech $A, B \in M_{m\times n}(\mathbb{C}), C \in M_{n\times k}(\mathbb{C})$, $\alpha \in \mathbb{C}$. Wówczas

i) $(A+B)^* = A^* + B^*$

ii) $(AC)^* = C^*A^*$

iii) $(\alpha A)^* = \overline{\alpha}A^*$

iv) $(A^*)^* = A$

v) Jeśli $n = m$, to $\det A^* = \overline{\det A}$ oraz $\operatorname{tr}A^* = \overline{\operatorname{tr}A}$.

vi) Jeśli $n = m$ oraz $\alpha \in Spec(A)$, to $\overline{\alpha} \in Spec(A^*)$.

Definicja 10.1.11. Macierz kwadratową A nazywamy macierzą

i) *hermitowską*, jeśli $A = A^*$,

ii) *antyhermitowską*, jeśli $A = -A^*$,

Jeśli $A = [a_{ij}]$ jest macierzą hermitowską, to oznacza to, że $a_{ij} = \overline{a_{ji}}$. W szczególności dla $i = j$ mamy $a_{ii} = \overline{a_{ii}}$. Innymi słowy elementy leżące na diagonali macierzy hermitowskiej są rzeczywiste.

Przykładem macierzy hermitowskiej jest macierz $A = \begin{bmatrix} 1 & 1-i \\ 1+i & 2 \end{bmatrix}$.

Wartości własne macierzy hermitowskiej są rzeczywiste, co pozwala sformułować następująca definicję.

Definicja 10.1.12. Macierz hermitowską $A_n(\mathbb{C})$ nazywamy *dodatnio (ujemnie) określoną*, jeśli wszystkie jej wartości własne są dodatnie (ujemne).

Uwaga 10.1.13. Jeśli $\mathcal{B}$ jest bazą przestrzeni V, oraz $u, v \in V$, $u = [x_1, x_2, \ldots, x_n]_{\mathcal{B}}$, $v = [y_1, y_2, \ldots, y_n]_{\mathcal{B}}$, to każdy iloczyn skalarny jest postaci $s(u, v) = X^T A \overline{Y}$, gdzie
$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ oraz $A = [a_{ij}] \in M_n(K)$, gdzie $a_{ij} = s(b_i, b_j)$.

Warunki i) oraz ii) w definicji iloczynu hermitowskiego równoważne są stwierdzeniu, że s jest postaci $s(u, v) = \sum_{i,j=1}^{n} a_{ij} x_i \overline{y_j}$ oraz macierz A jest macierzą hermitowską.

Jeśli $s : V \times V \to \mathbb{C}$ jest iloczynem hermitowskim, to wówczas $s(v, v) = \overline{s(v, v)}$, dla dowolnego $v \in V$. Innymi słowy $s(v, v)$ jest liczbą rzeczywistą. Zatem warunek iii) w definicji iloczynu hermitowskiego ma sens. Jest on równoważny stwierdzeniu, że macierz A jest dodatnio określona.

Przykład 10.1.14. Niech $f : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ dane będzie wzorem $f(z, w) = z_1\overline{w_1} + iz_2\overline{w_1} - iz_1\overline{w_2}$, gdzie $z = (z_1, z_2), w = (w_1, w_2) \in \mathbb{C}^2$.

Odwzorowanie f jest liniowe ze względu na pierwszą zmienną.
Dla dowolnych $\alpha, \beta \in \mathbb{C}$, $z = (z_1, z_2), w = (w_1, w_2), v = (v_1, v_2) \in \mathbb{C}^2$ mamy

$$f(\alpha z+\beta v, w) = f\Big((\alpha z_1+\beta v_1, \alpha z_2+\beta v_2), (w_1, w_2)\Big) = (\alpha z_1+\beta v_1)\overline{w_1}+i(\alpha z_2+\beta v_2)\overline{w_1}-i(\alpha z_1+\beta v_1)\overline{w_2} =$$

$$= \alpha f\Big((z_1, z_2), (w_1, w_2)\Big) + \beta f\Big((v_1, v_2), (w_1, w_2)\Big) = \alpha f(z, w) + \beta f(v, w),$$

Macierz $A = \begin{bmatrix} 1 & -i \\ i & 0 \end{bmatrix} \in M_2(\mathbb{C})$ spełnia warunek $A = A^*$.

Wartości własne macierzy hermitowskiej sa rzeczywiste (omówione w dalszej części wykładu).
Sprawdzimy czy A jest dodatnio określona, co równoważne jest stwierdzeniu, że wszystkie jej wartości

własne są dodatnie.
Obliczamy $Spec(A) = \{\lambda_1 = \frac{1-\sqrt{5}}{2}, \lambda_2 = \frac{1+\sqrt{5}}{2}\}$.

Ponieważ $\lambda_1 < 0$, macierz A nie jest dodatnio określona. Zatem f nie jest iloczynem hermitowskim.

Wiele stwierdzeń jest analogicznych jak w przypadku przestrzeni euklidesowych.

Można udowodnić, że jeśli $(V, \langle ., . \rangle)$ jest przestrzenią unitarną, to odwzorowanie $||.|| : V \to \mathbb{R}$ dane wzorem $\forall v \in V \quad ||v|| = \sqrt{\langle v, v \rangle}$, jest normą w przestrzeni V.

Zachodzi również *nierówność Schwarza* $\forall u, v \in V \quad \langle u, v \rangle \leq ||u|| \cdot ||v||$.

Przykład 10.1.15. Rozważamy w $\mathbb{C}^3$ standardowy iloczyn skalarny.
Dla danego wektora $(p, q, r) \in \mathbb{C}^3$ mamy

$$||(p, q, r)|| = \sqrt{p\bar{p} + q\bar{q} + r\bar{r}} = \sqrt{|p|^2 + |q|^2 + |r|^2}.$$

Obliczymy normę wektora $u = (3, 4i, 0) \in \mathbb{C}^3$.

$$||(3 - i, 4i, 0)|| = \sqrt{(3 - i)(3 + i) - 16i^2} = \sqrt{9 + 1 + 16} = \sqrt{26}$$

Wektory u, v nazywamy *ortogonalnymi*, jeśli ich iloczyn skalarny jest równy zero. Piszemy wówczas $u \perp v$. Układ wektorów $\{v_1, \ldots v_n\} \subset V$ nazywamy *układem ortogonalnym*, jeżeli $\forall i, j \in \{1, 2, \ldots, n\} \quad i \neq j \Rightarrow \langle v_i, v_j \rangle = 0$. Układ wektorów nazywamy *układem ortonormalnym*, jeśli jest układem ortogonalnym i każdy wektor jest unormowany, czyli

$$\forall i, j \in \{1, 2, \ldots, n\} \quad \langle v_i, v_j \rangle = \begin{cases} 0 & ; \quad i \neq j \\ 1 & ; \quad i = j \end{cases}.$$

Każdy ortonormalny układ wektorów jest liniowo niezależny. Jeśli przestrzeń wektorowa jest skończenie wymiarowa, to każdy ortonormalny układ wektorów można uzupełnić do bazy ortonormalnej.

Jeśli $\mathcal{B} = \left(b_1, b_2, \ldots, b_n\right)$ jest bazą ortonormalną przestrzeni unitarnej $(V, \langle ., . \rangle)$, zaś $v, w \in V$, to wówczas $v = \sum_{i=1}^{n} \langle v, b_i \rangle b_i$ oraz

$$\langle v, w \rangle = \sum_{i=1}^{n} \langle v, b_i \rangle \langle b_i, w \rangle.$$

Zatem jeśli $\mathcal{B} = \left(b_1, b_2, \ldots, b_n\right)$ jest bazą przestrzeni unitarnej $(V, \langle ., . \rangle)$, zaś $v, w \in V$, $v = [\alpha_1, \alpha_2, \ldots, \alpha_n]_{\mathcal{B}}$, $w = [\beta_1, \beta_2, \ldots, \beta_n]_{\mathcal{B}}$, to wówczas baza $\mathcal{B}$ jest ortonormalna wtedy i tylko wtedy gdy

$$\langle u, w \rangle = \alpha_1 \overline{\beta_1} + \alpha_2 \overline{\beta_2} \ldots + \alpha_n \overline{\beta_n}.$$

10.2 Macierze unitarne i normalne

Definicja 10.2.1. Macierz kwadratową A nazywamy macierzą

i) *normalną*, jeśli $AA^* = A^*A$,

ii) *unitarną*, jeśli $AA^* = A^*A = I$ lub równoważnie $A^{-1} = A^*$.

Uwaga 10.2.2. i) Każda macierz rzeczywista symetryczna (traktowana jako macierz nad ciałem $\mathbb{C}$) jest hermitowska, a każda macierz rzeczywista antysymetryczna (traktowana jako macierz nad ciałem $\mathbb{C}$) jest antyhermitowska.

ii) Macierz rzeczywista ortogonalna (traktowana jako macierz nad ciałem $\mathbb{C}$) jest macierzą unitarną.

iii) Macierze hermitowskie, antyhermitowskie, unitarne są macierzami normalnymi.

iv) Każdą macierz zespoloną kwadratową można zapisać jako sumę macierzy hermitowskiej i macierzy antyhermitowskiej.

Przykład 10.2.3.
$A = \begin{bmatrix} 1 & 1-i \\ 1+i & 2 \end{bmatrix}$, $A^* = A$, $AA^* = A^2 = \begin{bmatrix} 3 & 3-3i \\ 3+3i & 6 \end{bmatrix} = 3A$
Macierz jest hermitowska (a zatem normalna), ale nie unitarna.

$B = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$, $B^* = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}$, $BB^* = I$
Macierz nie jest hermitowska, ani antyhermitowska, ale jest unitarna (a zatem normalna).

$C = \begin{bmatrix} \cos\alpha & i\sin\alpha \\ -i\sin\alpha & -\cos\alpha \end{bmatrix}$, $\alpha \in \mathbb{R}$, $C = C^*$, $CC^* = I$
Macierz jest hermitowska i unitarna (a zatem normalna).

$D = \begin{bmatrix} 1 & i \\ 2i & 3 \end{bmatrix}$, $D^* = \begin{bmatrix} 1 & -2i \\ -i & 3 \end{bmatrix}$, $DD^* = \begin{bmatrix} 2 & i \\ -i & 13 \end{bmatrix}$
Macierz nie jest hermitowska, ani antyhermitowska, ani normalna, ani unitarna.

$E = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$, $E^* = E^T = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, $EE^* = E^*E = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$
Macierz nie jest hermitowska, ani antyhermitowska, ani unitarna, ale jest normalna.

Twierdzenie 10.2.4. Jeśli macierz $A \in M_n(\mathbb{C})$ jest unitarna, to wówczas $|\det A| = 1$.

Wyznacznik macierzy unitarnej to liczba zespolona o module równym 1, tj. leżąca na okręgu jednostkowym. W szczególności oznacza to, że macierze unitarne są *nieosobliwe*.

Twierdzenie 10.2.5.

i) Zbiór macierzy unitarnych stopnia n wraz z działaniem mnożenia macierzy jest grupą. Nazywamy ją *grupą unitarną* i oznaczamy symbolem $U(n)$. Macierze te reprezentują izometrie liniowe przestrzeni V.

ii) Zbiór macierzy unitarnych stopnia n o wyznaczniku równym 1 wraz z działaniem mnożenia macierzy tworzy grupę. Grupę tę nazywamy *specjalną grupą unitarną* i oznaczamy symbolem $SU(n)$.

iii) Niech $\mathcal{B}$, $\mathcal{B}'$ będą dwiema bazami ortonormalnymi przestrzeni unitarnej $(V, \langle ., . \rangle)$. Macierz przejścia $P = P_{\mathcal{B} \to \mathcal{B}'}$ jest macierzą unitarną. I odwrotnie, dowolna macierz unitarna jest macierzą przejścia między dwiema bazami ortonormalnymi.

10.3 Operatory liniowe w przestrzeniach unitarnych

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$. Niech $(V, \langle ., . \rangle_V)$ oraz $(W, \langle ., . \rangle_W)$ będą przestrzeniami unitarymi nad tym samym ciałem K, zaś $\varphi : V \to W$ odwzorowaniem liniowym.

Definicja 10.3.1. Odwzorowanie $\varphi^* : W \to V$ spełniające warunek

$$\forall v \in V, w \in W \quad \langle \varphi(v), w \rangle_W = \langle v, \varphi^*(w) \rangle_V$$

nazywamy *odwzorowaniem sprzężonym do przekształcenia liniowego* φ.

Twierdzenie 10.3.2.

i) Jeśli dla danego odwzorowania liniowego $\varphi : V \to W$ istnieje φ^*, to wówczas φ^* również jest odwzorowaniem liniowym i jest jedyne.

ii) Jeśli V oraz W są przestrzeniami unitarnymi skończnie wymiarowaymi, zaś $A = M_\varphi(\mathcal{B}, \mathcal{C})$ jest reprezentacja macierzową odwzorowania liniowego φ w bazach ortonormalnych $\mathcal{B}$, $\mathcal{C}$ przestrzeni V oraz W odpowiednio, to wówczas $M_{\varphi^*}(\mathcal{C}, \mathcal{B}) = A^*$.

iii) Jeśli V jest przestrzenią unitarną skończenie wymiarową, to dla dowolnego odwzorowania liniowego $\varphi : V \to W$ istnieje φ^*.[a]

[a] W przypadku, gdy V jest nieskończenie wymiarowa, dla danego φ może nie istnieć φ^*.

Dowód. ii) Niech $\mathcal{B} = (b_1, \ldots, b_n)$, $\mathcal{C} = (c_1, \ldots, c_m)$ to bazy ortonormalne przestrzeni V oraz W odpowiednio. Niech $v \in V, w \in W$ będą takie, że $v = [x_1, \ldots, x_n]_\mathcal{B}$, $w = [y_1, \ldots, y_m]_\mathcal{C}$. Oznaczmy

$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$. Wówczas
$\langle v, \varphi^*(w)\rangle_V = \langle \varphi(v), w\rangle_W = (AX)^T\overline{Y} = X^T A^T \overline{Y} = X^T\overline{A^*Y}$. $\square$

UWAGA: Operator sprzężony jest jedyny w danej przestrzeni unitarnej (zależy on od iloczynu skalarnego!).

Twierdzenie 10.3.3. Niech $(V, \langle .,. \rangle_V)$ oraz $(W, \langle .,. \rangle_W)$ będą przestrzeniami unitarymi nad tym samym ciałem K, zaś $\varphi, \psi : V \to W$ odwzorowaniami liniowym, dla których istnieją φ^*, ψ^*. Wówczas prawdziwe są następujące równości.

i) $(\varphi + \psi)^* = \varphi^* + \psi^*$

ii) $(\alpha \cdot \varphi)^* = \overline{\alpha} \cdot \varphi^*$

iii) $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$

iv) $(\varphi^*)^* = \varphi$

Przykład 10.3.4. Rozważmy przestrzenie $\mathbb{C}^2$ oraz $\mathbb{C}^3$ ze standardowymi iloczynami skalarnymi i standardowymi bazami $\mathcal{B}_k^2, \mathcal{B}_k^3$, które to są bazami ortonormalnymi.
Niech odwzorowanie liniowe $\varphi : \mathbb{C}^2 \to \mathbb{C}^3$ dane będzie wzorem

$$\varphi(z_1, z_2) = \Big(z_1 + iz_2, (2+i)z_1, -3iz_1 + \sqrt{2}z_2\Big).$$

Na mocy twierdzenia 10.3.2 istnieje jedyne odwzorowanie liniowe $\varphi^* : \mathbb{C}^3 \to \mathbb{C}^2$.
Podamy jego wzór. Niech $A = M_\varphi(\mathcal{B}_k^2, \mathcal{B}_k^3) \in M_{3\times 2}(\mathbb{C})$.

Wyznaczamy $A = \begin{bmatrix} 1 & i \\ 2+i & 0 \\ -3i & \sqrt{2} \end{bmatrix}$ oraz obliczamy $A^* = \begin{bmatrix} 1 & 2-i & 3i \\ -i & 0 & \sqrt{2} \end{bmatrix} \in M_{2\times 3}(\mathbb{C})$.

Ponownie na mocy twierdzenia 10.3.2 $A^* = M_{\varphi^*}(\mathcal{B}_k^3, \mathcal{B}_k^2)$. Zatem φ^* dane jest wzorem

$$\varphi^*(w_1, w_2, w_3) = \Big(w_1 + (2-i)w_2 + 3iw_3, -iw_1 + \sqrt{2}w_3\Big).$$

Weźmy $v = (1+i, -i) \in \mathbb{C}^2$ oraz $w = (0, 2+3i, \pi i) \in \mathbb{C}^3$. Wówczas

$$\langle \varphi(v), w\rangle = \langle \varphi(1+i, -i), (0, 2+3i, \pi i)\rangle = \langle (2+i, 1+3i, 3-(3+\sqrt{2})i), (0, 2+3i, \pi i)\rangle =$$
$$= (1+3i)(2-3i) + (3-(3+\sqrt{2})i)(-\pi i) = 11 - (3+\sqrt{2})\pi + 3(1-\pi)i$$

oraz

$$\langle v, \varphi^*(w)\rangle = \langle (1+i, -i), \varphi^*(0, 2+3i, \pi i)\rangle = \langle (1+i, -i), (7-3\pi+4i, \pi\sqrt{2}i)\rangle =$$
$$= (1+i)(7-3\pi-4i) + (-i)(-\pi\sqrt{2}i) = 11 - (3+\sqrt{2})\pi + 3(1-\pi)i.$$

W dalszej części rozważamy już jedynie endomorfizmy przestrzeni unitarnej $(V, \langle ., . \rangle)$.

Definicja 10.3.5. Operator liniowy $\varphi \in End(V)$ nazywamy

i) *samosprzężonym* lub *hermitowskim*, jeśli $\varphi^* = \varphi$,

ii) *antyhermitowskim*, jeśli $\varphi^* = -\varphi$,

iii) *unitarnym*, jeśli $\varphi^* \circ \varphi = \varphi \circ \varphi^* = \mathrm{id}_V$,

iii) *normalnym*, jeśli $\varphi^* \circ \varphi = \varphi \circ \varphi^*$.

W przypadku rzeczywistym operator hermitowski (antyhermitowski) nazywamy *symetrycznym (antysymetrycznym)*, zaś operator unitarny to operator *ortogonalny*.

Uwaga 10.3.6. i) Każdy operator hermitowski, antyhermitowski bądź unitarny jest operatorem normalnym.

ii) Jeśli V jest przesrzenią skończenie wymiarową, to operator liniowy $\varphi \in End(V)$ jest hermitowski / antyhermitowski / unitarny / normalny
wtedy i tylko wtedy, gdy jego reprezentacja macierzowa w pewnej bazie ortonormalnej przestrzeni V jest macierzą
hermitowską / antyhermitowską / unitarną / normalną.

10.4 Unitarna diagonalizacja macierzy normalnych

Niech $(V, \langle ., . \rangle)$ będzie skończenie wymiarową przestrzenią unitarną.

Jeśli dla danego operatora liniowego $\varphi \in End(V)$ (reprezentowanego w bazie standardowej przestrzeni V przez macierz A) istnieje baza ortonormalna $\mathcal{B}$ przestrzeni $(V, \langle ., . \rangle)$, w której macierz operatora φ jest diagonalna, to mówimy wówczas, że operator φ jest *unitarnie diagonalizowalny* oraz macierz A jest *unitarnie diagonalizowalna*.

Istnieje wówczas macierz unitarna $P \in U(n)$, będącą macierzą przejścia od bazy standardowej do bazy $\mathcal{B}$, taka, że $P^{-1}AP = P^*AP$ jest macierzą diagonlaną.

Przykład 10.4.1. Dla macierzy $A = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$ istnieje unitarna macierz $P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}$ taka, że

$$P^{-1}AP = P^*AP = \frac{1}{2} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & i \\ -1 & i \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Twierdzenie 10.4.2. Przy powyższych założeniach, prawdziwe są poniższe stwierdzenia.

i) Operator liniowy $\varphi \in End(V)$ jest unitarnie diagonalizowalny wtedy i tylko wtedy, gdy jest on operatorem normalnym.

ii) Macierz jest unitarnie diagonalizowalna wtedy i tylko wtedy, gdy jest macierzą normalną. [a]

[a] Macierze, które nie są macierzami normalnymi, nie są unitarnie diagonalizowalne. Nie oznacza to jednak, że nie są diagonalizowalne. Macierz $A = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix} \in M_2(\mathbb{C})$ nie jest macierzą normalną, bowiem $AA^T = \begin{bmatrix} 5 & -7 \\ -7 & 17 \end{bmatrix} \neq A^T A = \begin{bmatrix} 2 & 2 \\ 2 & 20 \end{bmatrix}$. Widmo macierzy A jest proste ($\lambda_1 = 2, \lambda_2 = 3$), zatem jest ona diagonalizowalna.

Twierdzenie 10.4.3. i) Wartości własne operatora hermitowskiego są rzeczywiste. Wartości własne macierzy hermitowskiej (lub rzeczywistej symetrycznej) są rzeczywiste.

ii) Wektory własne odpowiadające różnym wartościom własnym operatora hermitowskiego (macierzy hermitowskiej lub rzeczywistej symetrycznej) są wzajemnie ortogonalne.

ii) Każdej wartości własnej o krotności algebraicznej k operatora hermitowskiego (macierzy hermitowskiej lub rzeczywistej symetrycznej) odpowiada k liniowo niezależnych wektorów własnych.

Twierdzenie 10.4.4 (Twierdzenie spektralne dla operatorów i macierzy hermitowskich)**.**

i) Niech $(V, \langle ., . \rangle)$ będzie przestrzenią unitarną, zaś $\varphi \in End(V)$ operatorem hermitowskim. Wówczas istnieje taka baza ortonormalna przestrzeni V, że macierz operatora liniowego φ w tejże bazie jest macierzą diagonalną. Co więcej, elementy tej macierzy są liczbami rzeczywistymi.

ii) Dla każdej macierzy hermitowskiej $A \in M_n(\mathbb{C})$ istnieją macierz diagonalna $D \in M_n(\mathbb{R})$ oraz macierz unitarna $P \in U(n)$ takie, że

$$P^{-1}AP = P^*AP = D.$$

Mówimy wówczas, że macierz A jest *unitarnie diagonalizowalna* [a].

[a] Innymi słowy, jest unitarnie podobna do macierzy diagonalnej.

Wniosek 10.4.5 (Twierdzenie spektralne dla rzeczywistej macierzy symetrycznej)**.** Dla każdej macierzy symetrycznej $A \in M_n(\mathbb{R})$ istnieje macierz diagonalna $D \in M_n(\mathbb{R})$ oraz macierz ortogonalna $P \in O(n)$ takie, że

$$P^{-1}AP = P^TAP = D.$$

Mówimy wówczas, że macierz A jest *ortogonalnie diagonalizowalna* [a]. [b]

[a] Innymi słowy, jest ortogonalnie podobna do macierzy diagonalnej.

[b] Zauważmy, że z powyższych twierdzeń nie wynika, że symetryczne macierze zespolone są (unitarnie) diagonalizowlane. Macierz $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix} \in M_2(\mathbb{C})$ jest symetryczna, jednakże nie jest diagonalizowalna. Jej jedyną wartością własną jest $\lambda = 0$, o krotności algebraicznej równej 2 i krotności geometrycznej równej 1.

Przypomnijmy, że macierze $A, B \in M_n(K)$ *są podobne*, jeżeli istnieje macierz nieosobliwa $C \in GL_n(K)$ taka, że $A = C^{-1}BC$. Jeśli dodatkowo C jest macierzą unitarną (ortogonalną), to mówimy, że A jest *unitarnie (ortogonalnie) podobna* do B.

Diagonalizacja macierzy za pomocą macierzy unitarnych:

1) Znajdujemy wartości własne i odpowiadające im wektory własne.

2) Normalizujemy wektory własne odpowiadające różnym wartościom własnym o krotności algebraicznej równej 1.

3) Wektory własne odpowiadające wartości własnej o krotności algebraicznej większej niż 1 dobieramy w taki sposób, by były ortogonalne i następnie normalizujemy.

Przykład 10.4.6. Macierz $A = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix} \in M_3(\mathbb{R})$ jest symetryczna, a zatem ortogonalnie diagonalizowalna. Obliczamy $\chi_A(t) = \det(A - tI) = -t(t-6)^2$.

$Spec(A) = \{\lambda_1 = 0, \lambda_2 = 6\}$, przy czym $k_1 = 1, k_2 = 2$.

$$(A - \lambda_1 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 3 & 0 & -3 \\ 2 & 2 & 2 \\ -3 & 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Stąd $\begin{cases} x = z \\ x + y + z = 0 \end{cases}$. Zatem $E_0 = \{(x, -2x, x), x \in \mathbb{R}\}$.

Wybieramy wektor własny $u = (1, -2, 1)$ i normalizujemy go $\hat{u} = \frac{1}{\sqrt{6}}(1, -2, 1)$.

$$(A - \lambda_2 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow x - 2y + z = 0.$$

Zatem $E_6 = \{(2y - z, y, z), y, z \in \mathbb{R}\}$. Wybieramy wektory własne $v = (2, 1, 0), w = (-1, 0, 1)$.

W $\mathbb{R}^3$ rozpatrujemy standardowy iloczyn skalarny. Oczywiście $\langle u, v\rangle = 0$ oraz $\langle u, w\rangle = 0$.

Zauważmy, że $\langle v, w\rangle = -2 \neq 0$, zatem v i w nie są do siebie ortogonalne.

Zortogonalizujemy układ $\{v, w\}$ metodą Grama-Schmidta.

Niech $c_1 := v$.

Poszukujemy $c_2 = w + \alpha c_1$, $\alpha \in \mathbb{R}$. Dobierzmy α tak, by $\langle c_2, c_1\rangle = 0$.

Obliczamy $\langle c_2, c_1\rangle = \langle w + \alpha c_1, c_1\rangle = \langle w, c_1\rangle + \alpha\langle c_1, c_1\rangle = (-2 + 0 + 0) + \alpha(4 + 1 + 0)$.

Skąd $5\alpha - 2 = 0$, czyli $\alpha = \frac{2}{5}$ oraz $c_2 = w + \frac{2}{5}v = (-\frac{1}{5}, \frac{2}{5}, 1)$. Mamy $\langle c_1, c_2\rangle = 0$.

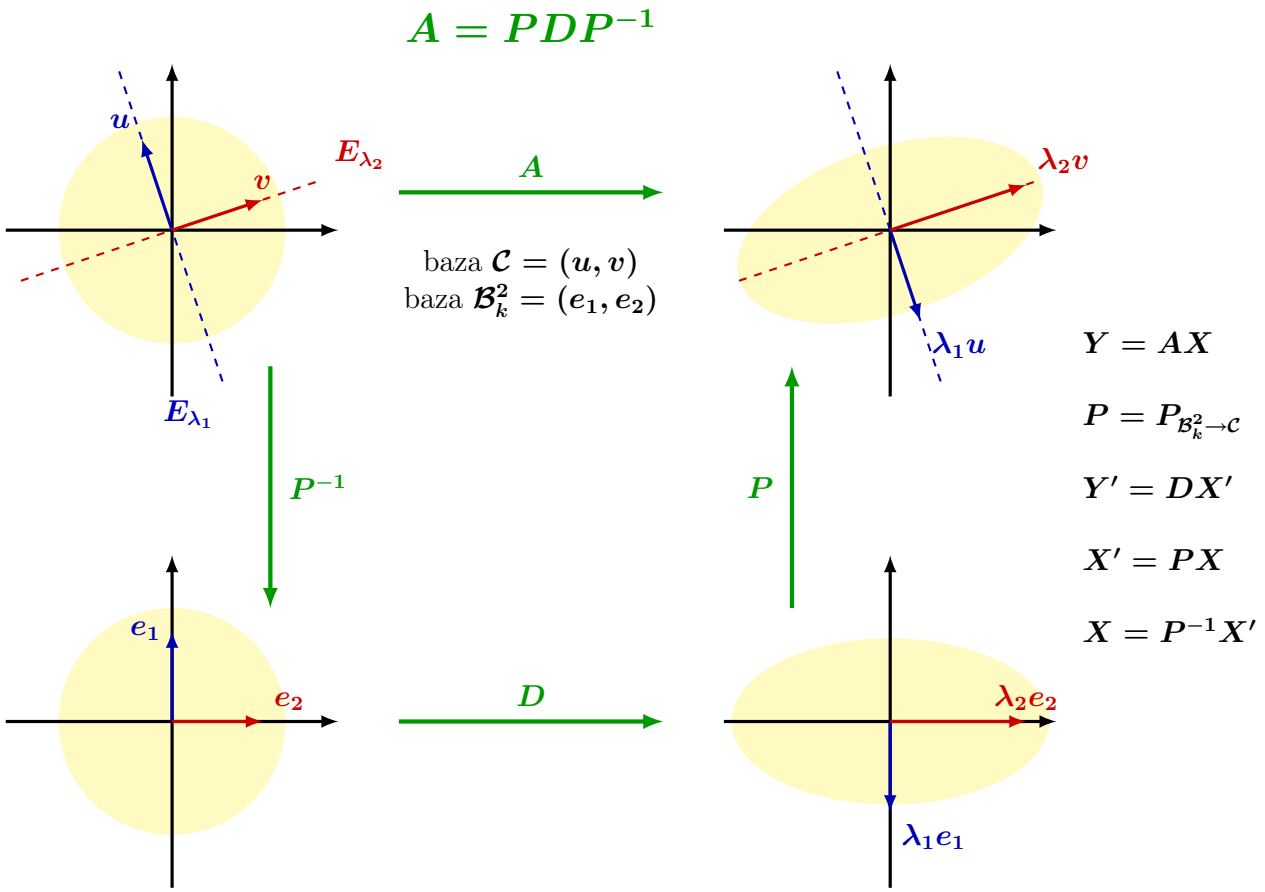
Normujemy wektory

$$\hat{c_1} = \frac{c_1}{|c_1|} = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0), \quad \hat{c_2} = \frac{c_2}{|c_2|} = \sqrt{\frac{5}{6}}(-\frac{1}{5}, \frac{2}{5}, 1).$$

Rozważamy bazę $(\hat{u}, \hat{c_1}, \hat{c_2})$. Macierz diagonalizująca ma postać $P = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & -\frac{1}{5\sqrt{30}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{\sqrt{5}}{\sqrt{6}} \end{bmatrix}$.

Można sprawdzić, że $PP^T = I$, zatem $P \in O(3)$ oraz $P^{-1}AP = P^TAP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}$.

Uwaga 10.4.7. Ortogonalna diagonalizacja macierzy symetrycznej może być widziana jako *rotacja osi układu współrzędnych* w taki sposób, *aby były one równoległe do wektorów własnych.* Zobacz tutaj.



Uwaga 10.4.8.

i) Wektory własne macierzy hermitowskiej na ogół są zespolone, zatem unitarna macierz diagonalizująca jest na ogół macierzą zespoloną.

ii) Rzeczywista macierz symetryczna jest macierzą hermitowską, zatem jej wartości własne są rzeczywiste. Ponieważ wartości własne i sama macierz są rzeczywiste, więc wektory własne można wybrać jako rzeczywiste, a w konsekwencji macierz diagonalizująca jest rzeczywistą macierzą ortogonalną.

iii) Macierze unitarne oraz rzeczywiste macierze ortogonalne można zdiagonalizować przy pomocy macierzy unitarnej. Ponieważ wartości własne i wektory własne są na ogół zespolone, macierz diagonalizująca jest macierzą unitarną, nie zaś ortogonalną.

Przykład 10.4.9. Macierz obrotu jest rzeczywistą macierzą ortogonalną.
W ogólności może być zdiagonalizowana przez zespoloną macierz unitarną.

Niech $A = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \in M_3(\mathbb{C})$, gdzie $\alpha \in \mathbb{R}$ ustalone.

$$\det(A - tI) = \begin{vmatrix} \cos\alpha - t & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha - t & 0 \\ 0 & 0 & 1-t \end{vmatrix} = (1-t)\Big((\cos\alpha - t)^2 + \sin\alpha^2\Big) =$$

$(1-t)(1 - 2t\cos\alpha + t^2) = \mathbf{0} \quad \Rightarrow \quad t = 0 \quad \vee \quad 1 - 2t\cos\alpha + t^2 = 0$
$\Delta = 4\cos^2\alpha - 4 = 4i^2\sin^2\alpha, \quad t = \cos\alpha \pm i\sin\alpha$

$Spec(A) = \{\lambda_1 = 1, \lambda_2 = \cos\alpha + i\sin\alpha, \lambda_3 = \cos\alpha - i\sin\alpha\}$

$\boxed{\lambda_1 = 1}$

$$(A - \lambda_1 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos\alpha - 1 & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha - 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} x(\cos\alpha - 1) + y\sin\alpha = 0 \\ -x\sin\alpha + (\cos\alpha - 1)y = 0 \end{cases}$$

Jeśli $\cos\alpha = 1$, wówczas $\sin\alpha = 0$ oraz $A = I$ (postać diagonalna).

Załóżmy, że $\cos\alpha \neq 1$. Wówczas $\sin\alpha \neq 0$ oraz $y = -\frac{\cos\alpha - 1}{\sin\alpha}x$. Stąd

$$x\sin\alpha + \frac{(\cos\alpha - 1)^2}{\sin\alpha}x = \frac{\sin^2\alpha + (\cos\alpha - 1)^2}{\sin\alpha}x = \frac{2\cos\alpha - 2}{\sin\alpha}x = 0 \Rightarrow x = 0.$$

W konsekwencji $y = 0$, zaś $z \in \mathbb{C}$ dowolne. Wybieramy wektor własny $u = (0, 0, 1)$.

$\boxed{\lambda_2 = \cos\alpha + i\sin\alpha}$

$$(A - \lambda_2 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -i\sin\alpha & \sin\alpha & 0 \\ -\sin\alpha & -i\sin\alpha & 0 \\ 0 & 0 & 1 - \cos\alpha - i\sin\alpha \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Stąd $x\sin\alpha + yi\sin\alpha = 0$. Jeśli $\sin\alpha = 0$, wówczas $\cos\alpha = \pm 1$ oraz A ma postać diagonalną.

Załóżmy, że $\sin\alpha \neq 0$. Wówczas $y = ix$ oraz $z = 0$. Wybieramy wektor własny $v = (1, i, 0)$.

$\boxed{\lambda_3 = \cos\alpha - i\sin\alpha}$

$$(A - \lambda_3 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} i\sin\alpha & \sin\alpha & 0 \\ -\sin\alpha & i\sin\alpha & 0 \\ 0 & 0 & 1 - \cos\alpha + i\sin\alpha \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Stąd $x\sin\alpha - yi\sin\alpha = 0$. Jeśli $\sin\alpha = 0$, wówczas $\cos\alpha = \pm 1$ oraz A ma postać diagonalną.

Załóżmy, że $\sin\alpha \neq 0$. Wówczas $y = -ix$ oraz $z = 0$. Wybieramy wektor własny $w = (-1, i, 0)$.

Układ (u, v, w) jest bazą wektorów własnych. Normujemy wektory bazowe

$$\hat{u} = u, \hat{v} = \frac{v}{|v|} = \frac{1}{\sqrt{2}}v, \hat{w} = \frac{w}{|w|} = \frac{1}{\sqrt{2}}w.$$

Zauważmy, że układ $\{\hat{u}, \hat{v}, \hat{w}\}$ jest układem ortonormalnym.
Istotnie $\langle \hat{u}, \hat{v} \rangle = \langle \hat{u}, \hat{w} \rangle = 0$ oraz $\langle \hat{v}, \hat{w} \rangle = 1 \cdot (-1) + i \cdot \bar{i} = -1 + 1 = 0$.

Macierz $P = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & -1 \\ 0 & i & i \\ \sqrt{2} & 0 & 0 \end{bmatrix} \in U(3)$, zatem $P^{-1} = P^* = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \\ -1 & -i & 0 \end{bmatrix}$ oraz

$$D = P^{-1}AP = \frac{1}{2} \begin{bmatrix} 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \\ -1 & -i & 0 \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \\ 0 & i & i \\ \sqrt{2} & 0 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha + i\sin\alpha & 0 \\ 0 & 0 & \cos\alpha - i\sin\alpha \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{i\alpha} & 0 \\ 0 & 0 & e^{-i\alpha} \end{bmatrix}.$$

PODSUMOWANIE

Niech $K = \mathbb{R}$ lub $K = \mathbb{C}$. Niech $V = (V, +, K, \cdot)$, $u = (u_1, \ldots, u_n) \in V, v = (v_1, \ldots, v_n) \in V$ oraz $A = [a_{ij}], U \in M_n(K)$

$K = \mathbb{R}$	$K = \mathbb{C}$
iloczyn skalarny $\langle u, v \rangle = u_1v_1 + \ldots + u_nv_n$	iloczyn hermitowski $\langle u, v \rangle = u_1\overline{v_1} + \ldots + u_n\overline{v_n}$
transponowanie $A^T : [a_{ij}]^T = [a_{ji}]$	sprzężenie hermitowskie $A^* : [a_{ij}]^* = [\overline{a_{ji}}]$
$\det A^T = \det A$	$\det A^* = \overline{\det A}$
macierz ortogonalna $UU^T = U^TU = I$	macierz unitarna $UU^* = U^*U = I$
$\det U = \pm 1$	$\lvert\det U\rvert = 1$
macierz symetryczna $A = A^T$	macierz hermitowska $A = A^*$
Macierz może nie mieć wartości własnych.	Każda macierz ma n wartości własnych (licząc z krotnościami).
Macierz symetryczna ma n rzeczywistych wartości własnych.	Macierz hermitowska ma n rzeczywistych wartości własnych.

10.5 Równoczesna diagonalizacja pary macierzy

W tym paragrafie zakładamy, że V jest skończenie wymiarową zespoloną przestrzenią liniową.

Definicja 10.5.1. i) Mówimy, że operatory liniowe $\varphi, \psi \in End(V)$ na n-wymiarowej przestrzeni liniowej V są *równocześnie diagonalizowalne*, jeśli istnieje taka baza przestrzeni V, w której macierze tych operatorów są równocześnie diagonalne.

ii) Mówimy, że macierze $A, B \in M_n(\mathbb{C})$ są *równocześnie diagonalizowalne*, jeśli istnieje $P \in GL_n(\mathbb{C})$ taka, że macierze $P^{-1}AP$ oraz $P^{-1}BP$ są diagonalne.

Przykład 10.5.2. Rozważmy dwie macierze $A, B \in M_2(\mathbb{C})$ postaci

$$A = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & \frac{1}{2} \end{bmatrix}.$$

A jest diagonalizowalna, bowiem jej widmo jest proste $Spec(A) = \{\lambda_1 = -2i, \lambda_2 = 2i\}$.

B jest macierzą symetryczną o elementach rzeczywistych, jest więc diagonalizowalna.

Możemy wyznaczyć $Spec(B) = \{\mu_1 = -\frac{1}{2}, \mu_2 = \frac{3}{2}\}$.

Ponadto $E_{\lambda_1} = \text{lin}_{\mathbb{C}}\{(i, 1)\}$, $E_{\lambda_2} = \text{lin}_{\mathbb{C}}\{(-i, 1)\}$ oraz $E_{\mu_1} = \text{lin}_{\mathbb{C}}\{(-1, 1)\}$, $E_{\mu_2} = \text{lin}_{\mathbb{C}}\{(1, 1)\}$.

Zatem A i B nie są równocześnie diagonalizowalne, gdyż nie istnieje taka baza przestrzeni $\mathbb{C}^2$, która składałaby się z wektorów będących równocześnie wektorami własnymi A i B.

Zauważmy, że jeśli dwa operatory liniowe $\varphi, \psi \in End(V)$ są równocześnie diagonalizowalne w pewnej wspólnej bazie przestrzeni V, nie koniecznie ortonormalnej, to wówczas operatory te są ze sobą przemienne lub inaczej mówiąc *komutują*, tj. $\varphi \circ \psi = \psi \circ \varphi$.

Istotnie, niech $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ będzie wspomnianą bazą (zbudowaną z wektorów własnych) oraz niech $v \in V, v = \sum_{i=1}^{n} \alpha_i b_i$. Wówczas istnieją $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ takie, że $\varphi(b_i) = \lambda_i b_i$ oraz $\mu_1, \ldots, \mu_n \in \mathbb{C}$ takie, że $\psi(b_i) = \mu_i b_i$, dla $i \in \{1, 2, \ldots, n\}$. Stąd

$$\psi \circ \varphi(v) = \sum_{i=1}^{n} \alpha_i \cdot \psi(\varphi(b_i)) = \sum_{i=1}^{n} \alpha_i \cdot \mu_i \lambda_i b_i = \sum_{i=1}^{n} \alpha_i \cdot \lambda_i \mu_i b_i = \sum_{i=1}^{n} \alpha_i \cdot \varphi(\psi(b_i)) = \varphi \circ \psi(v).$$

Podobnie jeśli macierze $A, B \in M_n(\mathbb{C})$ są równocześnie diagonalizowalne, to wówczas macierze A i B *komutują*, tj. $AB = BA$. Istotnie, ponieważ macierze diagonalne komutują, otrzymujemy

$$AB = (PP^{-1})A(PP^{-1})B(PP^{-1}) = P(P^{-1}AP)(P^{-1}BP)P^{-1} = P(P^{-1}BP)(P^{-1}AP)P^{-1} = BA.$$

Twierdzenie 10.5.3. Niech V będzie skończenie wymiarową zespoloną przestrzenią liniową.

i) Jeśli dwa operatory liniowe $\varphi, \psi \in End(V)$ są ze sobą przemienne, to wówczas dla dowolnego $\lambda \in Spec(\varphi)$ podprzestrzeń własna E_λ jest ψ-niezmiennicza i operatory φ, ψ mają wspólny wektor własny.

ii) Jeśli dwa diagonalizowalne operatory liniowe $\varphi, \psi \in End(V)$ są ze sobą przemienne, to są równocześnie diagonalizowalne.[a]

[a] Analogiczne stwierdzenie jest prawdziwe dla dowolnej skończonej liczby komutujących diagonalizowalnych operatorów liniowych. Zauważmy, że twierdzenie 10.5.3 nie mówi, że dwa operatory liniowe, które są ze sobą przemienne są równocześnie diagonalizowalne. Mówi ono, że dwa diagonalizowalne operatory liniowe, które są ze sobą przemienne są równocześnie diagonalizowalne. Macierze $A, B \in M_2(\mathbb{C})$, $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$ komutują, gdyż $AB = BA = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}$. Jednakże ani A, ani B nie jest diagonalizowalna.

Dowód. i) Niech $\lambda \in Spec(\varphi)$ oraz $E_\lambda = \{v \in V : \varphi(v) = \lambda v\}$. Dla dowolnego $v \in E_\lambda$ mamy

$$\varphi(\psi(v)) = \psi(\varphi(v)) = \psi(\lambda v) = \lambda\psi(v).$$

Zatem jeśli $v \in E_\lambda$, to również $\psi(v) \in E_\lambda$, czyli E_λ jest ψ-niezmiennicza.
Na mocy zasadniczego twierdzenia algebry, każdy operator liniowy na zespolonej przestrzeni liniowej ma wektory własne. W szczególności $\psi|_{E_\lambda} \in End(E_\lambda)$ ma wektor własny $w \in E_\lambda$, tzn. istnieje $\mu \in Spec(\psi)$ takie, że $\psi(w) = \mu w$. Równocześnie $\varphi(w) = \lambda w$, zatem φ, ψ mają wspólny wektor własny. □

Wniosek 10.5.4. Jeżeli dwie macierze hermitowskie (lub dwie macierze unitarne) $A, B \in M_n(\mathbb{C})$ komutują, to istnieje macierz $P \in U(n)$ oraz macierze diagonalne $D_A, D_B \in M_n(\mathbb{C})$ takie, że $P^{-1}AP = D_A$ oraz $P^{-1}BP = D_B$.

Dowód. Macierze hermitowskie (unitarne) sa macierzami normalnymi, a zatem są (unitarnie) diagonalizowalne. □

Zauważmy, że w przypadku, gdy A, B są hermitowskie, macierze D_A, D_B są rzeczywiste.

Przykład 10.5.5. Niech $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix}$.
Są to rzeczywiste macierze symetryczne, zatem każda z nich jest ortogonalnie diagonalizowalna.

Ponieważ $AB = \begin{bmatrix} 4 & 0 & 4 \\ 0 & 0 & 0 \\ 4 & 0 & 4 \end{bmatrix} = BA$, zatem istnieje (można tak wybrać) $P \in O(3)$ taka, że macierze $P^{-1}AP$, $P^{-1}BP$ są diagonalne.

Znajdujemy wartości własne i wektory własne macierzy A.
$\chi_A(t) = \det(A - tI) = \begin{vmatrix} 1-t & 0 & 1 \\ 0 & -t & 0 \\ 1 & 0 & 1-t \end{vmatrix} = -t(1-t)^2 + t = t^2(2-t)$
$Spec(A) = \{\lambda_1 = 2, \lambda_2 = 0\}$, krotności algebraiczne $k_1 = 1, k_2 = 2$

$\boxed{\lambda_1 = 2}$
$(A - \lambda_1 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (A - 2I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow y = 0,\ -x + z = 0$
Podprzestrzeń własna to $E^A_{\lambda_1} = \{(x, 0, x) \in \mathbb{R}^3 : x \in \mathbb{R}\} = \text{lin}\{(1, 0, 1)\}$

Wybieramy wektor własny $c_1 = (1, 0, 1)$ i normujemy go $c_1' = \frac{c_1}{||c_1||} = \frac{1}{\sqrt{2}}(1, 0, 1)$.

$\boxed{\lambda_2 = 0}$
$(A - \lambda_2 I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow x + z = 0,\ y\ -\ \text{dowolne}$
Podprzestrzeń własna to $E^A_{\lambda_2} = \{(x, y, -x) \in \mathbb{R}^3 : x, y \in \mathbb{R}\} = \text{lin}\{(1, 0, -1), (0, 1, 0)\}$

Wybieramy wektory własne $c_2 = (1, 0, -1), c_3 = (0, 1, 0)$ i normujemy je $c_2' = \frac{c_2}{||c_2||} = \frac{1}{\sqrt{2}}(1, 0, -1)$, $c_3' = c_3$.

Macierz, której kolumnami sa wersory z bazy ortonormalnej $\mathcal{C}' = (c_1', c_2', c_3')$ przestrzeni $\mathbb{R}^3$ diagonalizuje macierz A. Nie koniecznie jednak diagonalizuje macierz B.

Znajdujemy wartości własne i wektory własne macierzy B.
$\chi_B(t) = \det(B - tI) = \begin{vmatrix} 3-t & 1 & 1 \\ 1 & -t & -1 \\ 1 & -1 & 3-t \end{vmatrix} = (4-t)(t^2 - 2t - 2)$
$Spec(B) = \{\mu_1 = 4, \mu_2 = 1 - \sqrt{3}, \mu_3 = 1 + \sqrt{3}\}$, krotności algebraiczne $k_1 = k_2 = k_3 = 1$ (widmo proste)

Wybierzemy teraz wektory własne macierzy B spośród wektorów własnych macierzy A.
Wiemy że dla dowolnego $v \in E^A_{\lambda_1}$ istnieje $\mu \in \mathbb{R}$ takie, że $Bv^T = \mu v^T$.

W szczególności dla c_1' mamy

$$B(c_1')^T = B\begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{4}{\sqrt{2}} \\ 0 \\ \frac{4}{\sqrt{2}} \end{bmatrix} = 4(c_1')^T.$$

Zatem c_1' jest wektorem własnym ψ odpowiadającym wartości własnej $\mu_1 = 4$.

Niech u, w to dwa wektory własne ψ odpowiadające μ_2,μ_3.
Wiemy, że $u, w \in E^A_{\lambda_2}$.
Są zatem kombinacjami liniowymi postaci $ac_2' + bc_3' = (\frac{a}{\sqrt{2}}, b, -\frac{a}{\sqrt{2}})$, dla pewnych $a, b \in \mathbb{R}$.

Dla $\mu_2 = 1 - \sqrt{3}$ otrzymujemy $B\begin{bmatrix} x \\ y \\ -x \end{bmatrix} = \begin{bmatrix} 2x - y \\ -2x \\ x - 2y \end{bmatrix} = (1 - \sqrt{3})\begin{bmatrix} x \\ y \\ -x \end{bmatrix}$, skąd $y = (1 + \sqrt{3})x$.

Zatem $E^B_{\mu_2} = \{(x, (1 + \sqrt{3})x, -x) : x \in \mathbb{R}\} = \text{lin}\{(1, 1 + \sqrt{3}, -1)\}$.

Wybieramy wektor własny $d_2 = (1, 1 + \sqrt{3}, -1)$ i przyjmujemy $d_2' = \frac{d_2}{||d_2||} = \frac{1}{\sqrt{6+2\sqrt{3}}}(1, 1 + \sqrt{3}, -1)$.

Dla $\mu_3 = 1 + \sqrt{3}$ otrzymujemy $B\begin{bmatrix} x \\ y \\ -x \end{bmatrix} = \begin{bmatrix} 2x - y \\ -2x \\ x - 2y \end{bmatrix} = (1 + \sqrt{3})\begin{bmatrix} x \\ y \\ -x \end{bmatrix}$, skąd $y = (1 - \sqrt{3})x$.

Zatem $E^B_{\mu_3} = \{(x, (1 - \sqrt{3})x, -x) : x \in \mathbb{R}\} = \text{lin}\{(1, 1 - \sqrt{3}, -1)\}$.

Wybieramy wektor własny $d_2 = (1, 1 - \sqrt{3}, -1)$ i przyjmujemy $d_3' = \frac{d_3}{||d_3||} = \frac{1}{\sqrt{6-2\sqrt{3}}}(1, 1 - \sqrt{3}, -1)$.

Stąd

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6+2\sqrt{3}}} & \frac{1}{\sqrt{6-2\sqrt{3}}} \\ 0 & \frac{1+\sqrt{3}}{\sqrt{6+2\sqrt{3}}} & \frac{1-\sqrt{3}}{\sqrt{6-2\sqrt{3}}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6+2\sqrt{3}}} & \frac{-1}{\sqrt{6-2\sqrt{3}}} \end{bmatrix}.$$

Otrzymujemy

$$P^{-1}BP = P^TBP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 - \sqrt{3} & 0 \\ 0 & 0 & 1 + \sqrt{3} \end{bmatrix} \quad \text{oraz} \quad P^{-1}AP = P^TAP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

10.6 Rozkład Schura

Twierdzenie 10.6.1 (Lemat Schura o triangularyzacji)**.** [a]
Niech (V, s) będzie przestrzenią unitarną nad ciałem $\mathbb{C}$. Niech $\varphi \in End(V)$. Wówczas istnieje taka baza ortonormalna przestrzeni V, że macierz operatora φ w tejże bazie jest macierzą trójkątną górną.

[a] Macierz A podobną do macierzy trójkątnej nazywamy *triangularyzowalną*. Oznacza to, że istnieje taka macierz nieosobliwa P, że $P^{-1}AP$ jest macierzą trójkątną.

Wniosek 10.6.2. Dowolna macierz $A \in M_n(\mathbb{C})$ może być przedstawiona w postaci $A = UTU^*$, gdzie $T \in M_n(\mathbb{C})$ jest macierzą trójkątną górną, zaś $U \in M_n(\mathbb{C})$ jest macierzą unitarną.

Powyższe przedstawienie A nazywamy *rozkładem Schura* dla macierzy A.

Uwaga 10.6.3. Rozkład Schura nie jest jedyny. Możliwe są różne wybory bazy. Nawet jeśli elementy macierzy A będą rzeczywiste, to na ogół elementy macierzy T oraz U są zespolone.

Twierdzenie 10.6.4. Niech $(V, \langle ., . \rangle)$ będzie przestrzenią euklidesową. Niech $\varphi \in End(V)$. Jeśli φ ma n rzeczywistych wartości własnych (licząc z krotnościami), to wówczas istnieje taka baza ortonormalna przestrzeni V, że macierz operatora φ w tejże bazie jest macierzą trójkątną górną.

Wniosek 10.6.5. Macierz $A \in M_n(\mathbb{R})$, która posiada n rzeczywistych wartości własnych (licząc z krotnościami), może być przedstawiona w postaci $A = PTP^T$, gdzie $T \in M_n(\mathbb{R})$ jest macierzą trójkątną górną, zaś $P \in M_n(\mathbb{R})$ jest macierzą ortogonalną.

Przykład 10.6.6. W przestrzeni $\mathbb{C}^2$ rozpatrujemy standardowy iloczyn hermitowski.
Dana jest macierz $A = \begin{bmatrix} 2 & 8 \\ -2 & 2 \end{bmatrix} \in M_2(\mathbb{C})$.

Obliczamy $\chi_A(t) = (2-t)^2 + 16$. Rozwiązaniami $\chi_A(t) = 0$ są $\lambda_1 = 2 - 4i$, $\lambda_2 = 2 + 4i$.

Oczywiście $\lambda_2 = \overline{\lambda_1}$, gdyż współczynniki wielomianu są rzeczywiste.

Widmo jest proste, zatem A jedt diagonalizowalna.

Jednakże $AA^T \neq A^TA$, zatem macierz nie jest macierzą normalną, więc nie jest unitarnie diagonali-

zowalna.

Wyznaczymy bazę ortonormalną $\mathcal{C} = (c_1, c_2)$ przestrzeni $\mathbb{R}^2$ taką, że $U := P^{-1}AP = P^*AP$, gdzie $P = P_{\mathcal{B}_k^2 \to \mathcal{C}}$ jest macierzą trójkątną górną.

Jako pierwszy wektor bazowy wybierzemy wersor własny macierzy A.

Dla $\lambda_1 = 2 - 4i$ otrzymujemy $(A - \lambda_1 I) \begin{bmatrix} z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 4i & 8 \\ -2 & 4i \end{bmatrix} \begin{bmatrix} z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Stąd $-2z + 4iw = 0$. Wybieramy wektor własny $v_1 = (2i, 1)$ i przyjmujemy $c_1 = \frac{v_1}{||v_1||} = \frac{1}{\sqrt{5}}(2i, 1)$

Wyznaczymy teraz wektor $v_2 \in (\text{lin}\{v_1\})^\perp$.
Przypomnijmy, że dwa niezerowe wektory ortogonalne są liniowo niezależne.
Zatem $v_2 = (a, b) \in \mathbb{C}^2$ jest taki, że

$$v_2 \circ v_1 = (a, b) \circ (2i, 1) = a \cdot (-2i) + b \cdot 1 = 0.$$

Stąd $b = 2ia$. Wybieramy $v_2 = (1, 2i)$ i przyjmujemy $c_1 = \frac{v_2}{||v_2||} = \frac{1}{\sqrt{5}}(1, 2i)$.
Zatem $P = \frac{1}{\sqrt{5}} \begin{bmatrix} 2i & 1 \\ 1 & 2i \end{bmatrix}$. Obliczamy

$$U = P^{-1}AP = P^*AP = \frac{1}{\sqrt{5}} \begin{bmatrix} -2i & 1 \\ 1 & -2i \end{bmatrix} \begin{bmatrix} 2 & 8 \\ -2 & 2 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 2i & 1 \\ 1 & 2i \end{bmatrix} = \begin{bmatrix} 2 - \frac{2}{5}i & 6 \\ 0 & 2 + \frac{2}{5}i \end{bmatrix}.$$

Macierz A ma rozkład Schura postaci

$$A = PUP^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2i & 1 \\ 1 & 2i \end{bmatrix} \begin{bmatrix} 2 - \frac{2}{5}i & 6 \\ 0 & 2 + \frac{2}{5}i \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} -2i & 1 \\ 1 & -2i \end{bmatrix}.$$

Pomimo tego, że elementy macierzy A są rzeczywiste, macierz ta nie jest unitarnie podobna do żadnej rzeczywistej macierzy trójkątnej

Uwaga 10.6.7. Na mocy twierdzenia 8.2.10 macierze podobne mają ten sam wielomian charakterystyczny. Stąd możemy wywnioskować, że macierz A w powyższym przykładzie nie jest unitarnie podobna (ani nawet podobna) do żadnej rzeczywistej macierzy trójkątnej górnej.

Istotnie, wartości własne macierzy A są zespolone, zaś wartości własne macierzy trójkątnej górnej U (będące dokladnie wartościami własnymi A) leżą na jej przekątnej.
A zatem U nie może być macierzą rzeczywistą.

QUIZ

Czy podane stwierdzenia są prawdziwe, czy fałszywe?

1. Macierz jest unitarnie diagonalizowalna wtedy i tylko wtedy, gdy jest macierzą normalną.
2. Każda macierz hermitowska jest unitarnie diagonalizowalna.
3. Każda zespolona macierz symetryczna jest unitarnie diagonalizowalna.
4. Każda rzeczywista macierz symetryczna jest ortogonalnie diagonalizowalna.
5. Wektory własne odpowiadające różnym wartościom własnym macierzy hermitowskiej są wzajemnie ortogonalne.
6. Wektory własne macierzy hermitowskiej zawsze mają współrzędne rzeczywiste.
7. Każda rzeczywista macierz ortogonalna jest unitarnie diagonalizowalna.
8. Każda macierz unitarna jest unitarnie diagonalizowalna.

Spis treści

1

1.1 Notacja 1
1.2 Działania wewnętrzne i ich własności 3
1.3 Podstawowe struktury algebraiczne 4
1.3.1 Pojęcie grupy 4
1.3.2 Grupa permutacji 7
1.3.3 Pierścienie i ciała 14

19

2.1 Ciało liczb zespolonych 19
2.2 Postać trygonometryczna i wykładnicza 22
2.3 Pierwiastkowanie, równania wielomianowe 26
2.4 Interpretacja geometryczna 30

33

3.1 Macierze i ich własności 33
3.2 Wyznacznik macierzy 38
3.3 Macierz odwrotna 44

49

4.1 Układy równań liniowych - podstawowe pojęcia 49
4.2 Układy Cramera 50
4.3 Rząd macierzy i twierdzenie Kroneckera-Capellego 51
4.4 Metoda eliminacji Gaussa 57

61

5.1 Wektory w przestrzeni 61
5.2 Symbol Kroneckera i symbol całkowicie antysymetryczny 67
5.3 Płaszczyzna w przestrzeni $\mathbb{R}^3$ 71
5.4 Prosta w przestrzeni $\mathbb{R}^3$ 73
5.5 Wzajemne położenie płaszczyzn i prostych 76

85

6.1 Przestrzenie liniowe i ich podprzestrzenie 85
6.2 Liniowa niezależność wektorów, baza i wymiar przestrzeni liniowej 88

104

7.1 Definicja przekształcenia liniowego i podstawowe własności 104
7.2 Jądro, obraz i rząd odwzorowania liniowego 108
7.3 Działania na odwzorowaniach liniowych 111
7.4 Reprezentacja macierzowa odwzorowania liniowego 113
7.5 Zmiana baz 120
7.6 Uwaga na temat rozwiązywania układów równan liniowych 124

127
8.1 Wartości własne i wektory własne endomorfizmu 127
8.2 Diagonalizacja 136
8.3 Zastosowania diagonalizacji 145
8.4 Twierdzenie Cayleya-Hamiltona 148

152
9.1 Iloczyn skalarny i norma 152
9.2 Układy ortogonalne 156
9.3 Metody ortogonalizacji 164
9.4 Rzut ortogonalny na podprzestrzeń 167
9.5 Macierze ortogonalne i izometrie liniowe 170

184
10.1 Definicja przestrzeni unitarnej i podstawowe własności 184
10.2 Macierze unitarne i normalne 190
10.3 Operatory liniowe w przestrzeniach unitarnych 191
10.4 Unitarna diagonalizacja macierzy normalnych 193
10.5 Równoczesna diagonalizacja pary macierzy 200
10.6 Rozkład Schura 204

209

Linki do stron wykorzystane w dokumencie

Mnożenie liczb zespolonych
https://en.wikipedia.org/wiki/Complex_number#/media/File:Complex_number_multiplication_visualisatio

Przykładowe zadanie polegające na wyznaczaniu odległości dwóch prostych skośnych
https://enauczanie.pg.edu.pl/lab/geo3d/index.html?lang=pl&program=e.

Powinowactwo ścinające
https://en.wikipedia.org/wiki/Shear_mapping
https://pl.wikipedia.org/wiki/Powinowactwo_osiowe
https://pl.wikipedia.org/wiki/Powinowactwo_ścinające

Wektory własne
https://en.wikipedia.org/wiki/File:Eigenvectors-extended.gif

Rozpoznawanie twarzy
https://km.pcz.pl/konferencja/dokumenty/MMFT2017/Caban L.pdf
https://en.wikipedia.org/wiki/Eigenface

Kąty Eulera
https://compsci290-s2016.github.io/CoursePage/Materials/EulerAnglesViz/index.html
http://www.geometrictools.com/Documentation/EulerAngles.pdf

Rotacja w przestrzeni
https://www.andre-gaschler.com/rotationconverter/.

Ortogonalna diagonalizacja jako rotacja
https://en.wikipedia.org/wiki/Diagonalizable_matrix#/media/File:Diagonalization_as_rotation.gif.